UNIVERSITAT ROVIRA I VIRGILI

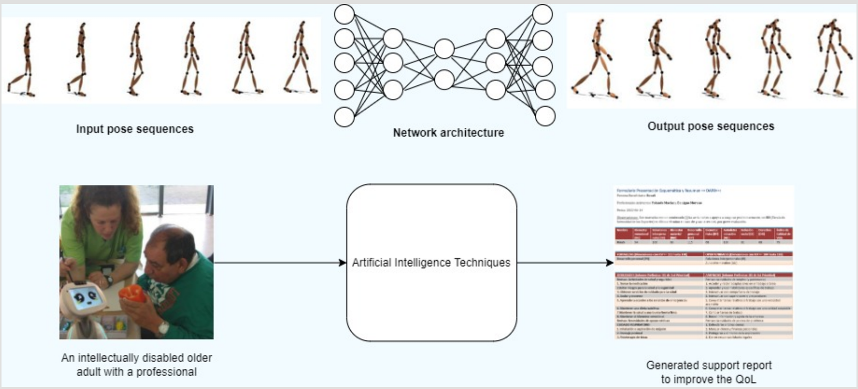
IMPROVING THE QUALITY OF LIFE FOR INTELLECTUALLY DISABLED ELDERLY PEOPLE USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Gaurav Kumar Yadav

UNIVERSITAT ROVIRA i VIRGILI

Improving the Quality of Life for Intellectually Disabled Elderly People using Artificial Intelligence Techniques

Author:
Gaurav Kumar YADAV



DOCTORAL THESIS
2023

Gaurav Kumar YADAV

Improving the Quality of Life for Intellectually Disabled Elderly People using Artificial Intelligence Techniques

DOCTORAL THESIS

Supervisors:
Domènec PUIG
G. C. NANDI

Department of Computer Engineering and Mathematics

UNIVERSITAT ROVIRA i VIRGILI

July, 2023

UNIVERSITAT
ROVIRA i VIRGILI

Departament d'Enginyeria Informàtica
i Matemàtiques
Av. Països Catalans, 26
43007 Tarragona
Tel. (+34) 977 559 703
Fax. (+34) 977 559 710

CERTIFICATE FROM SUPERVISORS

We do hereby forward the Thesis Report titled **"Improving the Quality of Life for Intellectually Disabled Elderly People using Artificial Intelligence Techniques"** prepared by **Gaurav Kumar Yadav** under our supervision in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy at **Universitat Rovira I Virgili, Tarragona** and **Indian Institute of Information Technology Allahabad**. We certify that the thesis is an authentic research work he carried out under our supervision.

Prof. Domenec Puig
URV Tarragona

Prof. G. C. Nandi
IIIT Allahabad

Abstract

As much independence as possible should be given to those with disabilities, according to the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) Guide, 2010, which a lot of people who have intellectual disabilities (IDs) want. Quality of life (QoL) research on reliant people, especially one with an ID, is gaining significant attention due to the rise in cases. Intelligence quotients (IQ) are commonly used to identify ID, typically described as intellectual and adaptive functioning deficiencies (e.g., an IQ score of less than 70). It illustrates a person's incapacity to carry out the duties, obligations, and socially required tasks. The disabilities emerge during the age of development and result in daily restrictions that demand ongoing support. It lasts a lifetime. We can not eliminate it, but we can lessen its effect using artificial intelligence (AI). The advancement of AI may assist ID elder people in improving their QoL with the help of collaborative robots (cobots) and many intelligent technologies. There is a significant need for cobots to help improve the QoL of ID older adults. This research work includes two parts. The research's initial phase is dedicated to developing the technology for the ID older adults once they are diagnosed with the level of QoL and referred for robot-assisted care. However, the present level of robot-assisted technology must be more robust to provide adequate care. One of the critical challenges in this area is developing cobots that can predict human motion accurately. This ability is essential because it allows the cobot to anticipate the person's movements and provide assistance. For instance, if an older adult is walking and suddenly loses balance, the cobot can use its motion prediction ability to anticipate the fall and take action to prevent it. This ability could involve providing physical support or alerting a caregiver. In this part, we try to develop human motion prediction techniques where cobots can judge the kind of services required by the IDs emphasizing correct human motion predictions using deep learning techniques. It has been

observed that the mechanics of human motion (HM) have never been consistent or predictable, and as such, they have constantly been varied and complex Lyu et al., 2022. Now that deep learning has advanced, it is possible to predict the HM by looking at the sequence of recent motions and guessing in what directions the future motion will most likely go. Existing research addresses this problem using discriminative models and displays the results for cases where the data has a homogeneous distribution (in distribution) but does not address the domain shift problem, which arises when the training and testing data have a heterogeneous distribution (out of distribution) like it does when these models are applied in real-world situations. A current study, however, suggested tackling domain shift difficulties by adding a generative model to the discriminative model Bourached et al., 2022, Shukla et al., 2022, and it outperforms other techniques in terms of results. In the current investigation, for training the complete end-to-end network and lowering the rank of the latent space, we recommend regularising the extended network by adding linear layers. We could enhance the model's efficiency for both the in-distribution(ID) and out-of-distribution (OoD) cases. Improving human motion prediction ability is an important task. On the other hand, caring for older ID people by improving their deficit QoL dimensions using AI techniques is also challenging. The second part of the proposed research analyzes every aspect of the QoL of ID elderly and provides adequate support to improve their deficit aspects. ID older adults face many more challenges than ordinary older adults, and their QoL and community involvement may increase by strengthening their capacity to manage their affairs independently. Despite the multiple challenges people with ID may face daily, technology interventions and support can help attain a higher QoL. Supporting a person with ID typically causes family members and professionals to feel overwhelmed. Over the years, efforts have been made to balance the various elements that directly influence a person's behaviour to achieve the desired outcome. The present investigation is dedicated to developing

an assistive tool with three models, details of which are presented in the thesis, which can significantly improve the quality of lives of IDs to a level where they can live their daily lives with their families. The proposed method is used to judge whether a person needs support. If the answer is affirmative, then the proposed method calculates the care priority for each aspect individually. It generates a SWOT (Strength, Weakness, Opportunity and Threat) report. By using the report, professional physicians can direct the patient to implement the necessary actions to improve the deficit dimension of the QoL. Our research studies involve developing two deep learning-based models for human motion predictions- one based on the discriminative model, which produces better results for in-distribution scenarios. Another one is based on the augmented model, which produces better results even for out-of-distribution scenarios, and three models for improving the QoL validated with rigorous experiments with data collected from real subjects from Ave Maria Foundation.

Resum

S'ha de donar la màxima independència possible a les persones amb discapacitat, segons la Convenció de les Nacions Unides sobre els Drets de les Persones amb Discapacitat (UNCRPD) Guide, 2010, que volen moltes persones amb discapacitat intel·lectual (DI). La investigació sobre la qualitat de vida (QoL) sobre persones dependents, especialment una amb DNI, està rebent una atenció important a causa de l'augment dels casos. Els quocients d'intel·ligència (QI) s'utilitzen habitualment per identificar l'ID, normalment descrits com a deficiències de funcionament intel·lectual i adaptatiu (per exemple, una puntuació de coeficient intel·lectual inferior a 70). Il·lustra la incapacitat d'una persona per dur a terme els deures, les obligacions i les tasques socialment requerides. Les discapacitats sorgeixen durant l'edat del desenvolupament i donen lloc a restriccions diàries que requereixen un suport continu. Dura tota la vida. No podem eliminar-lo, però podem disminuir el seu efecte mitjançant la intel·ligència artificial (IA). L'avenç de la IA pot ajudar les persones grans d'ID a millorar la seva qualitat de vida amb l'ajuda de robots col·laboratius (cobots) i moltes tecnologies intel·ligents. Hi ha una necessitat important de cobots per ajudar a millorar la QoL de les persones grans ID. Aquest treball de recerca consta de dues parts. La fase inicial de la investigació es dedica a desenvolupar la tecnologia per a les persones grans amb ID un cop se'ls diagnostica el nivell de QoL i es deriva a l'atenció assistida per robot. Tanmateix, el nivell actual de tecnologia assistida per robots ha de ser més robust per oferir una atenció adequada.

Un dels reptes crítics en aquesta àrea és desenvolupar cobots que puguin predir el moviment humà amb precisió. Aquesta habilitat és essencial perquè permet al cobot anticipar-se als moviments de la persona i donar assistència. Per exemple, si un adult gran camina i de sobte perd l'equilibri, el cobot pot utilitzar la seva capacitat de predicció de moviment per anticipar la caiguda i prendre mesures per prevenir-la.

Aquesta capacitat podria implicar proporcionar suport físic o alertar a un cuidador. En aquesta part, intentem desenvolupar tècniques de predicció del moviment humà on els cobots puguin jutjar el tipus de serveis que requereixen els ID fent èmfasi en les prediccions correctes del moviment humà mitjançant tècniques d'aprenentatge profund. S'ha observat que la mecànica del moviment humà (HM) mai ha estat coherent ni previsible i, com a tal, ha estat constantment variada i complexa Lyu et al., 2022. Ara que l'aprenentatge profund ha avançat, és possible predir l'HM mirant la seqüència de moviments recents i endevinant en quines direccions anirà probablement el moviment futur. La investigació existent aborda aquest problema mitjançant models discriminatius i mostra els resultats per als casos en què les dades tenen una distribució homogènia (en distribució) però no aborda el problema del canvi de domini, que sorgeix quan les dades d'entrenament i proves tenen una distribució heterogènia (fora de distribució).) com passa quan aquests models s'apliquen en situacions del món real. Un estudi actual, però, va suggerir abordar les dificultats de canvi de domini afegint un model generatiu al model discriminatiu Bourached et al., 2022, Shukla et al., 2022, i supera altres tècniques en termes de resultats. En la investigació actual, per entrenar la xarxa completa d'extrem a extrem i reduir el rang de l'espai latent, recomanem regularitzar la xarxa estesa afegint capes lineals. Podríem millorar l'eficiència del model tant en els casos en distribució (ID) com en els casos fora de distribució (OoD).

Millorar la capacitat de predicció del moviment humà és una tasca important. D'altra banda, tenir cura de persones grans amb identificació millorant les dimensions de la seva qualitat de vida deficitària mitjançant tècniques d'IA també és un repte. La segona part de la recerca proposada analitza tots els aspectes de la QoL de la gent gran amb DI i ofereix un suport adequat per millorar els seus aspectes de dèficit. Els adults grans amb ID s'enfronten a molts més reptes que els adults majors normals, i la seva qualitat de vida i la seva implicació en la comunitat poden augmentar en reforçar la seva capacitat per gestionar els seus assumptes de manera independent. Malgrat els múltiples reptes que les persones amb ID poden afrontar diàriament, les intervencions i el suport tecnològics poden ajudar a aconseguir una qualitat de vida més alta. Donar suport a una persona amb DNI normalment fa que els familiars i els professionals se sentin aclaparats. Al llarg dels anys, s'han fet esforços per equilibrar els diferents elements que influeixen directament en el comportament d'una persona per aconseguir el resultat desitjat.

La present investigació es dedica a desenvolupar una eina d'assistència amb tres models, els detalls dels quals es presenten a la tesi, que poden millorar significativament la qualitat de vida dels DNI fins a un nivell on puguin viure la seva vida diària amb les seves famílies. El mètode proposat s'utilitza per jutjar si una persona necessita suport. Si la resposta és afirmativa, el mètode proposat calcula la prioritat assistencial per a cada aspecte individualment. Genera un informe DAFO (Força, Debilitat, Oportunitat i Amenaça). Mitjançant l'ús de l'informe, els metges professionals poden indicar al pacient que implementi les accions necessàries per millorar la dimensió de dèficit de la QoL. Els nostres estudis de recerca impliquen desenvolupar dos models basats en l'aprenentatge profund per a les prediccions del moviment humà, un basat en el model discriminatiu, que produeix millors resultats per a escenaris de distribució. Un altre es basa en el model augmentat, que produeix millors resultats fins i tot en escenaris fora de distribució, i tres models per millorar la QoL validats amb experiments rigorosos amb dades recollides de subjectes reals de la Fundació Ave Maria.

Resumen

Según la Convención de las Naciones Unidas sobre los Derechos de las Personas con Discapacidad (CNUDPD)Guide, 2010, las personas con discapacidad intelectual (DI) deben gozar de la mayor independencia posible, algo que desean muchas de ellas. La investigación sobre la calidad de vida (CdV) de las personas dependientes, especialmente las que padecen una DI, está recibiendo una gran atención debido al aumento de casos. Los cocientes intelectuales (CI) se utilizan habitualmente para identificar la DI, que suele describirse como deficiencias intelectuales y de funcionamiento adaptativo (por ejemplo, una puntuación de CI inferior a 70). Ilustra la incapacidad de una persona para llevar a cabo los deberes, obligaciones y tareas socialmente requeridas. Las discapacidades surgen durante la edad de desarrollo y dan lugar a restricciones cotidianas que exigen un apoyo continuo. Dura toda la vida. No podemos eliminarla, pero podemos atenuar su efecto utilizando la inteligencia artificial (IA). El avance de la IA puede ayudar a las personas mayores con DI a mejorar su calidad de vida con la ayuda de robots colaborativos (cobots) y muchas tecnologías inteligentes. Existe una gran necesidad de cobots que ayuden a mejorar la CdV de las personas mayores con DI. Este trabajo de investigación consta de dos partes. La fase inicial de la investigación está dedicada al desarrollo de la tecnología para los adultos mayores con DI una vez que se les diagnostica el nivel de QoL y se les deriva para recibir cuidados asistidos por robots. Sin embargo, el nivel actual de tecnología asistida por robots debe ser más robusto para proporcionar una atención adecuada. Uno de los retos fundamentales en este campo es el desarrollo de cobots capaces de predecir con precisión el movimiento humano. Esta capacidad es esencial porque permite al cobot anticiparse a los movimientos de la persona y prestarle asistencia. Por ejemplo, si un adulto mayor está caminando y de repente pierde el equilibrio, el cobot puede utilizar su capacidad de predicción

del movimiento para anticiparse a la caída y tomar medidas para evitarla. Esta capacidad podría consistir en proporcionar apoyo físico o alertar a un cuidador. En esta parte, intentamos desarrollar técnicas de predicción de movimiento humano en las que los cobots puedan juzgar el tipo de servicios requeridos por las ID haciendo hincapié en las predicciones correctas de movimiento humano utilizando técnicas de aprendizaje profundo. Se ha observado que la mecánica del movimiento humano (HM) nunca ha sido consistente o predecible, y como tal, ha sido constantemente variada y compleja Lyu et al., 2022. Ahora que el aprendizaje profundo ha avanzado, es posible predecir la MH observando la secuencia de movimientos recientes y adivinando en qué direcciones es más probable que vaya el movimiento futuro. Las investigaciones existentes abordan este problema utilizando modelos discriminativos y muestran los resultados para casos en los que los datos tienen una distribución homogénea (dentro de la distribución), pero no abordan el problema del cambio de dominio, que surge cuando los datos de entrenamiento y prueba tienen una distribución heterogénea (fuera de la distribución) como ocurre cuando estos modelos se aplican en situaciones del mundo real. Un estudio actual, sin embargo, sugirió abordar las dificultades de cambio de dominio añadiendo un modelo generativo al modelo discriminativo Bourached et al., 2022, Shukla et al., 2022, y supera a otras técnicas en términos de resultados. En la presente investigación, para entrenar la red completa de extremo a extremo y reducir el rango del espacio latente, recomendamos regularizar la red extendida añadiendo capas lineales. Podríamos mejorar la eficacia del modelo tanto en el caso de la distribución interna (ID) como en el de la distribución externa (OoD).

Mejorar la capacidad de predicción del movimiento humano es una tarea importante. Por otro lado, atender a las personas mayores con DI mejorando sus dimensiones deficitarias de calidad de vida mediante técnicas de IA también supone un reto. La segunda parte de la investigación propuesta analiza todos los aspectos de la CdV de las personas mayores con DI y proporciona el apoyo adecuado para mejorar sus aspectos deficitarios. Los mayores con DI se enfrentan a muchos más retos que los mayores normales, y su CdV y su participación en la comunidad pueden aumentar si se refuerza su capacidad para gestionar sus asuntos de forma independiente. A pesar de los múltiples retos a los que las personas con DI pueden enfrentarse a diario, las intervenciones tecnológicas y el apoyo pueden ayudar a conseguir una mayor CdV. Normalmente, el apoyo a una persona con DI hace que los familiares y los profesionales se sientan desbordados. A lo largo de los años, se ha intentado equilibrar los distintos elementos que influyen directamente en el comportamiento de una persona para lograr el resultado deseado. La presente investigación está dedicada a desarrollar una herramienta de ayuda con tres modelos, cuyos detalles se presentan en la tesis, que puede mejorar significativamente la calidad de vida de las personas con DI hasta un nivel en el que puedan vivir su día a día con sus familias. El método propuesto se utiliza para juzgar si una persona necesita apoyo. Si la respuesta es afirmativa, el método propuesto calcula la prioridad asistencial de cada aspecto individualmente. Genera un informe DAFO (Debilidades, Amenazas, Fortalezas y Oportunidades). Utilizando el informe, los médicos profesionales pueden dirigir al paciente para implementar las acciones necesarias para mejorar la dimensión deficitaria de la CdV. Nuestros estudios de investigación implican el desarrollo de dos modelos basados en aprendizaje profundo para predicciones de movimiento humano: uno basado en el modelo discriminativo, que produce mejores resultados para escenarios en distribución. Otro se basa en el modelo aumentado, que produce mejores resultados incluso para escenarios fuera de distribución, y tres modelos para mejorar la QoL validados con experimentos rigurosos con datos recogidos de sujetos reales de la Fundación Ave María.

Acknowledgements

I want to express my deep gratitude to my supervisors, Prof. Domènec Puig, and Prof. G. C. Nandi, for their unwavering support and guidance throughout my PhD. Their insights, feedback, and encouragement have been invaluable in shaping my research. I am also grateful to my thesis doctoral committee members, Dr Rahul Kala from IIIT Allahabad and Prof. Ashish Dutta from IIT Kanpur, for their constructive comments and suggestions that helped refine my work. I thank Dr Hatem A Rashwan from URV Spain and Dr Mohamed Abdel-Nasser from Aswan University, Egypt, for their insightful feedback and valuable suggestions. Their support immensely improved the quality of my thesis. I am grateful to Mr Benigno Moreno, Sara Duenas, and the entire IRD team for their support in completing the most crucial part of the thesis. My thanks also go to my seniors Shruti Jaiswal, Priya Shukla, Aditya Singh, Vandana Kushwaha, and Saddam Abdulwahab and friends Padmini, Eddardaa, Ritesh, Shwetang, and Surendra for their camaraderie and support during the ups and downs of the PhD journey. I am indebted to my family for their love, support, and understanding. Their constant encouragement and belief in me have been the driving force behind my success. I am indebted to all the Center of Intelligent Robotics members of IIIT-A, headed by Prof. G.C.Nandi, Robotics and Intelligent Vision Laboratory of the URV, headed by Prof. Domènec Puig, and Instituto de Robotica para la Dependencia (IRD) you all made this research possible. Finally, I offer my humble gratitude and heartfelt thanks to lord Krishna for his infinite blessings, guidance, and love that he bestows upon me.

Thank you all for your contributions, guidance, and support, which have been instrumental in helping me to achieve this milestone.

Contents

List of Figures

List of Tables

List of Abbreviations

WHO: World Health Organisation

UNCRPD: United Nations Convention on the Rights of Persons with Disabilities

ID: Intellectual Disability

IDs: Intellectual Disabilities

QoL: Quality of Life

IQ: Intelligence Quotients

AI: Artificial Intelligence

Cobots: Collaborative Robots

HM: Human Motion

HMP: Human Motion Prediction

ID: In-Distribution

OoD: Out-of-Distribution

SWOT: Strength, Weakness, Opportunity and Threat

SIS: Support Intensity Scale

POC: Priority of Care

GCN: Graph Convolutional Network

CNNs: Convolutional Neural Networks

GNNs: Graph neural networks

GC-LSTM: Graph Convolutional LSTM

VAE: Variational Autoencoder

GAN: Generative Adversarial Network

RNN: Recurrent Neural Network

CNN: Convolutional Neural Network

ANOVA: Analysis of Variance

H_0: Null Hypothesis

H_a: Alternative Hypothesis

IRB: Inception Residual Block

DCT: Discrete Cosine Transform

ERD: Encoder-Recurrent-Decoder

S-RNN: Structural Recurrent Neural Networks

CVAE: Conditional Variational Autoencoder

SVRNN: Semi-supervised Recurrent Neural Network

STMN: Spatio-temporal Manifold Network

ADMM-BP: Alternating Direction Method of Multipliers and Backward Propagation

TRAM: Temporal Attention Re-calibration Module

STCM: Spatio-Temporal Convolution Module

3DHoTs: 3D Histograms of Texture

MBC: multi-class boosting classifier

MPJPE: Mean Per Joint Position Error

CMU-MoCap: CMU motion capture

GP-GPU: General Purpose Graphical Processing Unit

NIF: Number of Input Frames

NOF: Number of Output Frames

IDCT: Inverse Discrete Cosine Transform

GCL: Graph Convolutional Layer

ResGCN: Residual Graph Convolutional Network

GCBs: Graph Convolution Blocks

EW: Emotional Well-being

PD: Personal Development

PW: Physical Well-being

SD: Self-determination

IR: Interpersonal Relation

SI: Social Inclusion

MW: Material Well-being

RI: Rights

S1E: Health and Healthcare

S2: Protection and Defense

S3B: Behavioral Support Need

S1A: Homelife Activities

S1C: Lifelong Learning

S3A: Exceptional Medical Need

S1F: Social Activities

S1D: Employment Activities

S1B: Community Life Activities

RT: Regression Tree

RF: Random Forest

GB: Gradient Boosting

MLP: Multiple Linear Regression

MLPRegressor: Multilayer Perceptron Regressor

ANFIS: Adaptive Neuro-fuzzy Inference System

SD: Standard Deviation

SMOTE: Synthetic Minority Over-Sampling Technique

SMOTE-R: Synthetic Minority Over-Sampling Technique for Regression

SMOGN: Synthetic Minority Over-Sampling Technique for Regression with random Gaussian Noise

ANNs: Artificial Neural Networks

MAE: Mean Absolute Error

MSE: Mean Squared Error

RMSE: root Mean Square Error

IDD: Intellectual and Developmental Disabilities

NDDs: Neurodevelopmental Disorders

ASD: Autism Spectrum Disorder

ADHD: Attention Deficit Hyperactivity Disorder

ETR: Extra Trees Regressor

CART: Classification and Regression Tree

GBR: Gradient Boosting Regressor

KNN: K-Nearest Neighbors

Lasso: Least Absolute Shrinkage and Selection Operator

OMP: Orthogonal Matching Pursuit

Ada: Adaptive Boosting Regressor

DT: Decision Tree

LAR: Least Angle Regression

PAR: Passive Aggressive Regressor

BR: Bayesian Ridge

LR: Linear Regression

Huber: Huber Regressor

EN: Elastic Net

OMP: Orthogonal Matching Pursuit

ML: Machine Learning

I would like to dedicate this thesis to
my parents, and my brothers
for their endless love, support and
encouragement.

Part I

Introduction

Chapter 1

Introduction

1.1 Motivation

The most significant medical and social demographic issue facing the globe now is the ageing of the population. The world celebrates Senior Citizens Day on August 21 to raise awareness of the importance of providing for the elderly and enabling them to live with dignity. Although we cannot stop ageing, we may learn how to manage its effects so our loved ones remain healthy. According to historical standards, an ageing population is a relatively recent issue. No nation had more than 11% of its 65 years or older population in 1950 Cohen et al., 2001. The highest percentage, as compared to 2000, was 18%. However, the issue will significantly worsen by 2050, when it might reach 38% Cohen et al., 2001. According to projections, there will be more seniors (those 60 or older) than adolescents (adults that are 10 to 24 years old) in 2050 (2.1 billion older adults versus 2.0 billion adolescents) (WHO), 2021. Figure 1.1 shows the percentage population of nine countries from 1970-2021, provided by OECD (the Organization for Economic Co-operation and Development) OECD, 2014. The highest percentage is 28.86% in Japan currently. In Spain, it is 19.95% whereas, in European Union, including

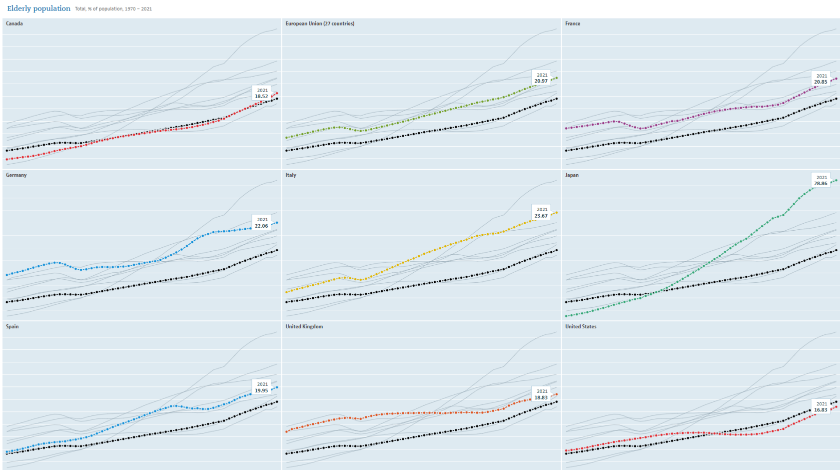


Figure 1.1: Current aging population of nine countries OECD, 2014

27 countries, it is 20.97%. In the diagram, various lines show the population of all nine countries for comparison. The black line shows the population of total OECD countries, which includes 56 countries across the world. The average elder people population for the total OECD countries in 2021 is 17.64 %, which shows the alarming call for these nations to work in this area because older age can not be mitigated. However, with the intervention of technology, life can be made more accessible. The ageing population is the most significant medical and social demographic issue globally.

As people age, they may be more susceptible to developing disabilities, such as mobility limitations, visual impairment, or cognitive decline. It may be more challenging for older persons with disabilities to maintain independence and actively engage in society (WHO), 2021. 15% of the global population, or one billion people, are estimated to have one or more disabilities Krahn, 2011. Most nations and regions have significantly increased the number of older people in recent years.

This trend is anticipated to accelerate in the following decades, according to statistics from World Population Prospects Economic and Social Affairs, 2022. As a result, there is also a rise in the number of elderly people with disabilities. Governments are compelled to reevaluate and further explore the intersections between the discourses on ageing and disability in light of the higher-than-average rates of disability among older people, which are brought on by the accumulation of a lifetime's worth of health hazards from illness, injury, and chronic disease Disease Control and (CDC), 2022. According to estimates from the World Health Organisation (WHO), 1% to 3% of people worldwide have an ID Schepens, Van Puyenbroeck, and Maes, 2019a, and the prevalence is higher in developing countries due to the higher incidence of factors such as poverty, malnutrition, and lack of access to healthcare. Studies on the senior population have indicated that ID prevalence rises with age. In the United States, a survey was carried out, and it found 2.5% of persons 60 and older have ID on average Conrad, 2020. According to different research done in Spain, 2.2% of adults 65 and older have ID González-Valero et al., 2021. It is crucial to remember that the prevalence of ID in the elderly is influenced by various factors, including genetic conditions, ageing-related cognitive decline, and acquired brain injuries. Early detection and diagnosis of ID can help improve outcomes for elderly individuals with this condition Innes, McCabe, and Watchman, 2012. ID, formerly known as mental retardation, is a common neurodevelopmental disease characterised by severe disability in intelligent and adaptive behaviour Schalock, Luckasson, and Tassé, 2021a. In contrast to adaptive functioning, described as routine tasks like independent living and communication, intellectual functioning includes problem-solving, learning, and judgement. Deficits in intelligence and adaptability can be seen as early as birth. The IQ score assesses ID Schepens, Van Puyenbroeck, and Maes, 2019a. ID symptoms can range from mild to profound. If a child's IQ is below 70, they are categorised as ID children McKenzie et al., 2016. ID is differentiated based on IQ scores, so mild

ID would be considered when a child's IQ is between 50 and 70. This category includes almost 85 % of the whole ID population. These kids can graduate from school, become independent, receive training, and possibly look for employment. A moderate IQ ranges from 35 to 50. They typically require care and attention even if they can function independently. The ID population as a whole comprises 10% of those in this category. A severe intellectual ID occurs when an IQ falls between 20 and 35. These people have lower competence levels and need help with reading and math. They must be watched over frequently. This group comprises 4% of the entire ID population. An IQ below 20 is considered to be a profound ID occurrence. This category includes 1% of the entire population of ID individuals Wieczorek, 2018. While some children are genetically predisposed to ID, others may develop it due to drugs taken as young children to treat the illness. However, in some instances, the reason for ID is not evident Oliveira et al., 2020. People with IDs are living into old age for the first time in the history of mankind Putnam and Bigby, 2021. Caregivers are challenged by their evolving requirements as they strive to provide the highest level of care. Elderly individuals with IDs have extremely distinct life course trajectories from old individuals without chronic IDs, and many have been socially isolated Bigby, 2004. Few people with IDs have children to take care of them, and many still rely on their elderly parents for help. Several of these ID individuals have been in residential facilities since they were young, depending on frequently shifting teams of professional support staff for their daily care Foster and Boxall, 2015.

ID people feel uncomfortable making a good relationship with new caregivers. In the same way, new caregivers and other professionals who care for them feel bored and tired by fulfilling the requirement of the ID people. Caregivers face several challenges related to communication, behaviour, physical care, emotional support, medical care, and caregiver stress Innes, McCabe, and Watchman, 2012; due to these issues, Most senior citizens with ID prefer to remain in their homes instead of in residential aged care centres. Whether they live in a home or residential services, they require support frequently and fulfilling their need is overhead for the caregivers. Therefore, Assistive technology intervention can assist caregivers in providing adequate care to ID elderly people Perry, Beyer, and Holm, 2009. It may be possible to feel more secure and comfortable at home with the help of assistive technologies, sensors, and/or body-worn devices that track people's movements Yusif, Soar, and Hafeez-Baig, 2016. Collaborative/Assistive robots (Cobots) can play an essential part in improving the QoL of ID older adults. These cobots can help them with tasks that may be challenging due to their cognitive or physical limitations, thereby promoting independence and reducing dependence on caregivers Pu et al., 2019. Cobots can potentially improve the QoL of ID older people by promoting independence, enhancing social interaction, improving safety, promoting physical activity, and providing medication reminders. However, it is essential to note that cobots should be used with human support and care. They cannot replace the essential human connection and compassion important for overall well-being Bemelmans et al., 2012. For Cobot to care for elderly people with IDs, the ability to predict human motion is crucial. It is because individuals with IDs may have difficulty with movement coordination, balance, and spatial awareness, which can increase their risk of falls and injuries March, 2017. Cobots that predict human motion can help prevent accidents by anticipating a person's movements and providing necessary support or assistance. For instance, a cobot can predict when an individual is about to lose balance and offer a stabilizing arm or handrail to prevent falls. A cobot that can predict human motion can also help with mobility and navigation. In order to avoid collisions or impediments, the cobot can predict where the person intends to go and adjust its movement accordingly Fang Wang, 2019. Furthermore, a cobot that can predict human motion

Figure 1.2: Various robots are assisting the elderly people in providing them ease in their daily life

can provide personalized care and assistance. The cobot can adapt its care and support to better meet its needs and preferences by analyzing a person's movement patterns. Hence, the ability to predict human motion is essential for cobots that care for ID older adults. Being able to predict human movement by observing their past movement is one of the challenging tasks for cobots. In addition, cobots must be capable of predicting accurate human motion by observing their past motion to avoid a possible collision. This leads us towards the primary objective of this research work.

Assisting the ID elderly people with the help of assistive technologies such as cobots, online games, and others can directly impact the condition of the ID elderly. On the other hand, monitoring their cognitive and adaptive behaviour is essential to improve their lifestyle. Based on the monitoring, professionals or doctors can suggest required actions by implementing them, and the present condition of the ID elderly may improve. Therefore, many researchers worked to build a support system to improve the QoL of ID elderly Schepens, Van Puyenbroeck, and Maes, 2019a; Verdugo et al., 2012. Providing support to elderly people contains many challenges, like analysing the cognitive ability of the ID elderly people, finding the area in which ID people face problems, and calculating the QoL's deficit dimensions Parmenter, 2021, etc. The second objective of this research work is focused on this field to help the ID elderly live a better life by implementing the provided actions in their life. The following section discusses our approach to dealing with both tasks.

1.2 Approach

In the first part of the thesis, we try to develop collision-free human motion prediction techniques where collaborative robots can judge the kind of services required by the IDs emphasizing correct human motion predictions using deep learning techniques. Caring for older people using AI techniques, especially a collaborative robot as an assistant, is also challenging. A machine must be capable of predicting human motion by observing past motion to avoid the chance of collision. The machine can learn about human behaviour by recognizing human motion. Apart from the human-machine interaction for healthcare, this field has numerous applications. It includes motion generation Liu and Mu, 2021, human tracking Gong et al., 2011 Gupta et al., 2014, action identification and prediction Kong and Fu, 2022, and internet gaming Espinoza et al., 2022. Existing research addresses the problem when the data distribution is homogeneous (in distribution) (ID). Still, it does not address the domain shift problem, which arises when there is a heterogeneous distribution (out-of-distribution) (OoD), as happens when these models are applied in real-world situations. Although, we address both situations separately by proposing two models in this work. First, in the

ID example, we aim to predict human motion more accurately. To forecast future human motion, we have provided past human motion. The essential thing is to find temporal features to get higher accuracy. To find temporal features, researchers Fragkiadaki et al., 2015 Martinez, Black, and Romero, 2017 Yadav and Nandi, 2020 employ RNNs. With the RNN-based models, there are two issues. The first issue is that errors in RNN accumulate at every stage of sequencing Fragkiadaki et al., 2015 Martinez, Black, and Romero, 2017. At the time of the test, it results in nonsensical predictions. Second, there is a discrepancy between the most recent observed and the first anticipated frames Martinez, Black, and Romero, 2017. As a result of the frame-by-frame regression's inability to promote global smoothness, this inconsistency exists. Utilizing the residual block with the inception module can improve the consistency between an earlier observed stance and the upcoming anticipated pose. Since 1D CNN is essential for identifying temporal features, we recommend employing it in the inception module. In the second OoD scenario, current research suggested tackling domain shift difficulties by enhancing the discriminative model with a generative model Bourached et al., 2022 and producing improved results. In the current study, we suggest regularising the extended network by adding linear layers to train the complete end-to-end network and reduce the rank of the latent space Jing, Zbontar, et al., 2020. We enhance the model by regularising the network to handle domain shift circumstances better. We strengthen our network by including additional linear layers to the network encoder to account for the distribution sets for the training and testing sets that differ. We tested our model using Human3.6M Ionescu et al., 2013 and CMU Motion Capture Mao et al., 2019 benchmark datasets. We showed that it performed better regarding the Euclidean distance between anticipated and actual joint angle values than 14 OoD actions from H3.6M and 7 OoD actions from CMU MoCap.

Improving the human motion prediction ability using deep learning for the in-distribution and out-of-distribution cases to provide the prediction ability to collaborative robots is an important task. On the other hand, improving the QoL of ID elderly people through the intervention of the support system is challenging. ID is a lifelong disease; therefore, ID people face age-specific challenges at each stage of life. Specifically, the situation becomes more severe in their older age because of the other diseases that come with older age. Studies focusing on psychiatric disorders in older individuals with ID suggest that age increases the risk of general psychiatric morbidity, dementia, anxiety disorder, and depression Axmon et al., 2018. These problems make the life of ID people in older age very complex. Generally, people of older age live in care houses where they depend on the caregiver. Caregivers frequently change; therefore, they face challenges maintaining good relationships with their caregivers. Professionals, researchers, and government organizations are working to simplify ID people's lives. It can not be cured, but the effect of ID can be minimized by regular inspection and providing support. This work proposed using AI techniques to analyze an individual's various aspects of QoL. The idea of QoL has several facets and includes both etic (universal) and emic (culture-bound) elements. It has both subjective as well as objective traits, and social and environmental factors have an impact on it. The researcher suggested eight dimensions to account for all facets of ID persons Gómez Sánchez, Schalock, Verdugo Alonso, et al., 2021. These eight aspects determine the QoL. An improvement in these demonstrates a rise in QoL. Verdugo et el. Verdugo et al., 2010 suggested utilizing a GENCAT scale tool to gauge the values of these eight aspects from respondents' questionnaire responses. The GENCAT scale uses predetermined criteria and specific tables to calculate the eight dimension values. After that, the GENCAT tool calculates the QoL index value using the values for the eight dimensions. The QoL index value says either beneficiary needs support or not Gómez Sánchez et al., 2022. Improving the QoL of an ID person depends on the QoL index value. Utilizing the GENCAT tool to calculate the index value is laborious and technical. Usually, a specialist must use the GENCAT tool to determine the index value. Although the GENCAT tool effectively and precisely determines the index value, it requires time and expertise because it is an entirely statistical process. As a result, in this study, we suggest using an ML-based model to derive the index value directly from the answers to the questionnaire. Previously, experts had to use the 69 questionnaire responses to calculate the eight dimensions value and then use the GENCAT tool to convert the eight dimensions value to the index value. ML-based model is an efficient and faster approach to calculating these dimension's values. Our proposed method uses an ML-based model to calculate the index value and decide the need for support. Based on the decision model, proceed to the following steps, where the first model calculates the priority of care (POC) for every eight aspects of QoL based on the PoC value. The model decides the set of actions to help the deficit aspects. After that model generate a report, which consists of the possible actions required to improve an individual's QoL. This report helps professionals to track the progress of the individual health condition.

1.3 Thesis objectives

The primary goals of this thesis include:

1. To predict human motion more accurately in in-distribution scenarios. We proposed an inception residual block and GCN-based deep learning model to learn the inconsistency between the first predicted pose and the last observed pose.

2. To measure and enhance the effectiveness of a deep augmented network for the out-of-distribution scenario to predict human motion. We proposed a regularization technique by minimizing the rank of the latent variable of the augmented network.

3. To improve the QoL of ID, elderly people. We proposed a support system that analysed the QoL dimensions of the ID elderly people and provided a support report based on the analysis. The report contains the necessary actions related to life's personal, social and judicial areas. Implementing these actions improves the QoL of the ID elderly.

4. To decide the deficit dimension of the QoL. We suggested a way to calculate each dimension's priority of care value. Based on the priority of care value model, decide whether dimensions need immediate or optional support or do not need any support.

1.4 Contributions

The following list summarises this thesis' significant contributions:

1. Development of a residual connection between the input and inception blocks that generates temporal encoding and aids in efficiently learning salient features. To determine an ideal number of features that are neither overfitted nor under fitted, parametric studies of the Human3.6M dataset for the number of temporal features are conducted. End-to-end learning is accomplished with the Inception Residual Block and Graph Convolutional Network (GCN).

   - S. Gupta, Gaurav Kumar Yadav, G. C. Nandi, "Development of Human Motion Prediction Strategy using Inception Residual Block", is published in Multimedia tools and applications, SCI indexed, IF: 2.396).

2. Development of an enhanced discriminate and generative integrated model, which contains graph convolution networks with residual connection between the branches' respective encoder and

decoder parts of the model with additional linear layers integrated with the output of the encoder to regularize the extended network for predicting the improved future human motion in an out-of-distribution scenario.

- Gaurav Kumar Yadav, M.A.Nasser, H.A.Rashwan, Domenec Puig, G. C. Nandi, "Implicit Regularization of a Deep Augmented Neural Network Model for Human Motion Prediction", is published in Applied Intelligence 2022, (SCI indexed, IF: 5.086).

3. Development of an implicit regularization method for predicting human motion by reducing the rank solutions of the covariance matrix. The redundant information is removed by adding more linear layers to the encoder model, reducing the latent space size of the variational autoencoder network.

   - Gaurav Kumar Yadav, M.A.Nasser, H.A.Rashwan, Domenec Puig, G. C. Nandi, "Implicit Regularization of a Deep Augmented Neural Network Model for Human Motion Prediction", is published in Applied Intelligence 2022, (SCI indexed, IF: 5.086).

4. Developing an efficient model for predicting QoL index value to decide the need for support and calculating the priority of care value for each dimension to identify the deficit dimension using artificial intelligence techniques.

   - Gaurav Kumar Yadav, B.M.Vidales, H.A.Rashwan, J.Oliver, Domenec Puig, G. C. Nandi, M.A.Nasser, "Effective ML-based quality of life prediction approach for dependent people in guardianship entities", is published in Alexandria Engineering Journal 2022, (SCI indexed, IF: 6.626).

5. Development of a model which takes the values of eight dimensions as input, analyses each dimension and generates a SWOT (strength, weakness, opportunity, and threat), and produces a report containing the necessary actions to improve the deficit dimensions of quality of life and helps professionals to track the condition of the patient frequently.

    - Gaurav Kumar Yadav, B.M.Vidales, H.A.Rashwan, J.Oliver, Domenec Puig, G. C. Nandi, M.A.Nasser, "Effective ML-based quality of life prediction approach for dependent people in guardianship entities", is published in Alexandria Engineering Journal 2022, (SCI indexed, IF: 6.626).
    - Gaurav Kumar Yadav, B.M.Vidales, Sara Dueñas, M.A.Nasser, H.A.Rashwan, Domenec Puig, G. C. Nandi, "Predicting Personalized Quality of Life of an Intellectually Disabled Person Utilizing Machine Learning", is published in 24th International Conference of the Catalan Association for Artificial Intelligence (CCIA 2022) (Frontiers in Artificial Intelligence and Applications 2022).

6. Development of a private dataset named NewtonOne containing the information of forty-six people, twenty-six collected in the first stage of the Never Alone project and twenty more people during the Siempre Contigo project. This dataset has the age group of people above fifty-five who have some symptoms of intellectual disability. Using this dataset, we trained a learning-based model instead of a statistical model to predict the quality of life index value, which helps to automate the decision of whether a patient needs support or not.

    - Gaurav Kumar Yadav, B.M.Vidales, H.A.Rashwan, J.Oliver, Domenec Puig, G. C. Nandi, M.A.Nasser, "Effective ML-based quality of life prediction approach for dependent people in guardianship entities", is published in Alexandria Engineering Journal 2022, (SCI indexed, IF: 6.626).

- Gaurav Kumar Yadav, B.M.Vidales, J.Oliver, H.A.Rashwan, M.A.Nasser, G. C. Nandi, Domenec Puig,"Un enfoque eficaz basado en la IA para mejorar la calidad de vida de las personas con discapacidad intelectual", is published in Brains, the business, research, ageing, Innovation, neuroscience and social journal, 2022.
- Gaurav Kumar Yadav, "Quality of Life Analysis of Dependent People using Multiple Linear Regression Model", is published in the preceding 7th URV Doctoral workshop in computer science and mathematics.

7. Develop a machine learning-based model to predict the values of eight dimensions by taking the responses of sixty-nine questionnaires. This model replaces the role of the traditional GENCAT tool, a statistical tool, with a set of rules and some lookup table to calculate the values of the eight dimensions from the responses of the questionnaires. Our model attempts to replace the tedious and laborious job requiring a professional to calculate the dimension values correctly.

    - Gaurav Kumar Yadav, H.A.Rashwan, B.M.Vidales, M.A.Nasser, J.Oliver, G. C. Nandi, Domenec Puig,"A Data-driven Model to Predict Quality of Life Dimensions of Intellectually Disabled People based on the GENCAT Scale", is under review in Social Indicators Research 2022, (SSCI indexed, IF: 3.024).

8. Development of a machine learning-based model to predict the support intensity scale (SIS) value by inputting the values of eight dimensions. SIS value integrates the individual aspect (questionnaire response) with the environmental factors (sensory information) to enhance the dependent person's quality of life.

- Gaurav Kumar Yadav, B.M.Vidales, H.A.Rashwan, Joan Oliver, M.A.Nasser, Domenec Puig, G. C. Nandi, "Predicting the Quality of Life Index Value for determining the Requirement of Support Needs for Intellectually Disabled Individual using Machine Learning Methods", is published in IEEE Xplore, IEEE Bangalore Humanitarian Technology Conference 2023.

Other publications

- Gaurav Kumar Yadav, Domenec Puig, G. C. Nandi, "Designing an Adaptive Cost Function for Dynamic Human Pose Predictions", under review in Multimedia tools and applications, SCI indexed, IF: 2.396).

- Gaurav Kumar Yadav, G.C.Nandi. "Development of Adaptive Sampling Based strategy for Human ActivityPredictions Using Sequential Networks.", is published in 4th Conference on Information and Communication Technology (CICT 2020) (IEEE Xplore), DOI: 10.1109/CICT51604.2020.9312097.

- Gaurav Kumar Yadav, Shruti Jaiswal, G.C.Nandi. "Trajectory Learning for Stable Bipedal Walking Robots using Sequential Networks", is published in 7th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON 2020) (IEEE Xplore), DOI: 10.1109/UPCON50219.2020.9376558.

- Gaurav Kumar Yadav, Shruti Jaiswal, G.C.Nandi "Generic Walking Trajectory Generation of Biped using Sinusoidal Function and Cubic Spline", is published in 7th International Conference on Signal Processing and Integrated Networks (SPIN 2020), (IEEE Xplore), DOI: 10.1109/SPIN48934.2020.9071083.

1.5 Thesis organization

The proposed draft for the thesis is distributed into eight chapters and organised as follows:

Chapter 1: This chapter is presented as the introduction to the thesis. It holds the importance of elderly care by improving their quality of life and providing a model which may be used in care robots to advance them with a human motion prediction ability. In this chapter, we have explained ageing, intellectual disability, the causes of intellectual disability, possible techniques to improve intellectual disability, human motion prediction, its real-time application, and the motivation behind our work. Finally, this chapter presents the thesis's outline and the primary contribution of our research.

Chapter 2: This chapter discusses the analysis of previous research and subsequent problem formulation. First, we present the role of AI in ID older people care. Then we discussed the importance of assistive robots' human motion prediction ability to improve the quality of life of ID older people. Further, we discussed the recent research on human motion prediction. After that, we discussed the recent research in the quality of life support paradigm. After that, we discussed the GENCAT tool. Finally, we have discussed the statistical analysis technique named the t-test.

Chapter 3: This chapter discusses our proposed inception residual block-based model to learn the salient temporal features to predict more accurately future human motion. Here we use IRB to learn the temporal features and graph convolutional network to learn the spatial features. It also shows the parametric analyses of the Human3.6M dataset Ionescu et al., 2014 for the number of temporal features performed to establish an ideal number of features that are neither overfitted nor under fitted. We compare our model results with other existing results, and our model performance is better than theirs.

Chapter 4: In this chapter, we have discussed the development of a technique to regularise augmented discriminative and generative models to overcome the limitation of domain shift scenarios and predict more accurate human motion. It discusses the role of the linear layers to regularise by minimizing the rank of the covariance matrix. This extended network helps to achieve more accurate prediction results, especially for the out-of-distribution scenario. We also test the model for the in-distribution case, and the results are better than other methods. This chapter explains it in detail.

Chapter 5: This chapter deals with developing a practical model to assist a physician professional by generating a support report of his/her intellectually disabled elderly patient. This model contains three ML-based algorithms to predict the QoL index value, standard score value, and support intensity scale value, and finally, generates a support report. This chapter contains it in detail.

Chapter 6: In this chapter, we have discussed developing a deep neural network-based model to predict the values of eight QoL dimensions from the responses of sixty-nine questionnaires. Earlier, professionals used the GENCAT tool to do this work. The GENCAT tool contains a set of rules and relatable tables to perform this task; therefore, it requires much time and skills to complete the task. This chapter explains the deep neural network and machine learning models instead of the GENCAT tool to do this task in detail.

Chapter 7: This chapter discusses our proposed ML-based model to predict the QoL index value. The QoL index value decides the need for support for a person, whether required or not. Earlier, the GENCAT tool calculates it from the values of eight QoL dimensions. Using the GENCAT tool, calculating the index value is a two-step process. First, the GENCAT tool converts the response of sixty-nine questionnaires to values of eight dimensions and, afterwards, from the values of eight dimensions to the QoL index value. In this chapter, we explain the ML-based model to calculate the QoL index directly from the responses of sixty-nine questionnaires.

Chapter 8: The chapter concludes the research and experiments and summarises the research's key conclusions and contributions to the relevant research field. Finally, this chapter describes potential future work based on the experiment's and research's conclusions.

Chapter 2

Background

2.1 AI for ID elderly

We are currently in a significant demographic transition, going from a society where most people are reasonably young to one where a sizable portion of the population is over 65 Pollack, 2005. This transformation presents a challenge and a chance for the development of intelligent technologies. Although many older persons will still be in good health and able to work, this group is generally more susceptible to physical and cognitive decline than younger ones Kharbat, Alshawabkeh, and Woolsey, 2021. AI could be essential in caring for elderly persons with intellectual disabilities by assisting with daily living activities, enhancing their QoL, and promoting independence. The following are a few ways that AI can be beneficial in providing care for the elderly and those who are intellectually disabled. AI can develop personalized care plans for elderly persons with IDs based on their medical history, lifestyle, and preferences. It can help caregivers provide tailored care that meets the individual needs of the person Taimoor and Rehman, 2021. AI-powered sensors can remotely monitor elderly persons with IDs and alert caregivers in case of any emergency or unusual activity. These sensors can monitor the person's movement, behaviour, and sleep patterns and can even detect falls Al-Khafajiy et al., 2019. AI-powered virtual assistants can provide companionship and emotional support and engage elderly persons with IDs in activities such as games or storytelling. These virtual assistants can also provide personalized reminders, offer medication management support, or answer general questions Tian et al., 2019. AI algorithms can analyze data to predict potential health issues, medication side effects, or changes in behaviour that may require attention. It can help caregivers take proactive measures to prevent adverse events from occurring Rubeis, 2020. AI-powered virtual assistants can provide social interaction and engagement for elderly persons with intellectual disabilities, which can be especially helpful for those isolated or living alone Chattaraman et al., 2019. AI can be programmed to send reminders and prompts to the person to take their medications, attend appointments, or engage in other activities necessary for their health and well-being Mihailidis et al., 2007.

2.2 Addressing the need for human motion Prediction in collaborative robots for ID elderly

Collaborative robots can be highly beneficial in caring for elderly persons with IDs, as they can support various daily life activities, such as mobility, communication, and personal care. Human motion prediction ability is an essential feature that cobots should possess to effectively assist elderly persons with IDs Ansari et al., 2017. The ability to predict human motion can help cobots to anticipate the needs and movements of elderly persons with IDs, which can improve their safety and comfort. For instance, a cobot with HMP ability can detect when an elderly person is about to fall and provide support to prevent the fall. It can also anticipate the person's movements during activities such as walking or transferring from one surface to another and assist accordingly. Additionally, HMP ability can help cobots to adapt to the changing needs of elderly persons with IDs. For instance, if an ID elderly person develops a new movement pattern due to an injury or illness, the cobot can use its motion prediction ability to learn and adjust its assistance accordingly Fiorini et al., 2021. Furthermore, predicting human motion can help cobots provide more personalized care to ID elderly people. The cobot can use its motion prediction ability to understand the specific needs and preferences of the person and tailor its assistance accordingly.

2.3 Recent studies in the field of predicting human motion

Human motion prediction (HMP) has recently gained interest in computer vision and robotics because it enables machines to comprehend human behaviour, plan intended actions, and enhance interaction tactics. Deep neural networks are effective in predicting human motion in in-distribution scenarios. In general, these models use large datasets of motion capture data to learn complex patterns of human motion and use this knowledge to predict future movements. One successful approach is to use recurrent neural networks (RNNs) Martinez, Black, and Romero, 2017; Fragkiadaki et al., 2015; Jain et al., 2016; Cui et al., 2019 to model the temporal dependencies in human motion. RNNs are well-suited to this task because they can consider the body's previous movements when predicting future movements. In addition, models based on RNNs can be trained using backpropagation through time, allowing them to learn from data sequences Shu et al., 2021. Another approach is to employ a combination of convolutional neural networks (CNNs), and RNNs Pavllo et al., 2020. CNNs are good at extracting spatial features from input data, while RNNs can capture temporal dependencies. By combining these two types of networks, models can learn to predict human motion based on spatial and temporal cues. For example, some models use a hybrid approach that combines a neural network with a physics engine Zhang et al., 2022, which can help to ensure that the predicted motions are physically plausible. Generally, the ability of deep neural networks to forecast human movement in in-distribution scenarios has shown considerable potential. The capacity to generalize to out-of-distribution scenarios is an issue that still needs to be resolved.

2.4 Graph neural networks based models to predict human motion

HMP is challenging, as it involves capturing complex temporal dependencies and interactions between different body parts. Graph neural networks (GNNs) Zhou et al., 2020 have emerged as a powerful tool for modelling such dependencies in various applications, including HMP. In a GNN-based model for HMP, the human body is shown as a graph, with every node corresponding to a body part, and the edges depict the connections between them. The graph can be constructed using different techniques, such as skeleton-based models, which use a predefined set of joints to represent the human body, or point cloud-based models, which use 3D points to represent the human body Li et al., 2020c. The GNN model then processes this graph by iteratively updating each node's concealed state based on the information from its neighbours, using message-passing algorithms. It allows the model to capture the spatiotemporal dependencies between different body parts and determine how the human body will move in the future Dang et al., 2021. One popular approach to GNN-based HMP is the convolutional graph LSTM (GC-LSTM) model, which extends the traditional LSTM architecture to operate on graphs Shi B, 2019. In this model, each node has an associated LSTM cell, which receives input from its neighbours and alters its hidden state following this information. The model then outputs the predicted future motion of each body part. Another approach uses the graph transformer network (GTN) model, which applies the transformer architecture to graphs Tang J, 2021. In this model, each node is represented by an embedding, updated using attention mechanisms that capture the relationships between different body parts. Furthermore, a model proposed by li et al. Li et al., 2020c takes multiscale as input to the graph neural network-based encoder; the multiscale phenomenon helps to model the learning enter-dependencies between joints, which helps to predict the more accurate future motion.

2.5 Augmented model to predict the human motion in out-of-distribution scenario

In the past few years, there have been important developments in generative and discriminative models for human motion prediction for in-distribution scenarios. Generative models such as Variational Autoencoder (VAE) Bourached et al., 2022, Generative Adversarial Networks (GAN) Cui et al., 2021, and Autoregressive models Li et al., 2020a can generate new human motion sequences by getting to know the training data's underlying distribution. As opposed to that, discriminative models such as Recurrent Neural Networks (RNN) Fragkiadaki et al., 2015; Jain et al., 2016; Cui et al., 2019, Convolutional Neural Networks (CNN) Li et al., 2018, and Graph Neural Networks (GNN) Li et al., 2020c learn directly map the output sequence from the input sequence.

However, most of these models are trained on a specific dataset. Their performance may suffer when applied to out-of-distribution scenarios where the test data's distribution is substantially different from the training data. To address this issue, researchers have proposed augmented models that combine generative and discriminative models to enhance human motion prediction efficiency in out-of-distribution scenarios Bourached et al., 2022. Augmented generative and discriminative models can effectively predict human motion in out-of-distribution scenarios. By leveraging the strengths of generative and discriminative models, these approaches can learn to grasp the intricacy and variability of human motion and generalize well to new and unseen scenarios.

2.6 Various methods to improve the QoL of ID elderly

A multimodal strategy that considers the needs of the body, the society, and the emotions of older individuals with IDs is needed to enhance their standard of living Schepens, Van Puyenbroeck, and Maes, 2019a. The following are some helpful tactics. For the maintenance of both bodily and psychological well-being, exercise is crucial. Encourage senior citizens with IDs to exercise in ways they prefer, including swimming, dancing, or walking Schepens, Van Puyenbroeck, and Maes, 2019a. To preserve health and prevent disease, proper nutrition is crucial. Provide healthy meals and snacks tailored to their specific needs, taking into account any dietary restrictions or food preferences Riquelme et al., 2016. Social isolation is a common problem among elderly people with IDs. Encourage social interactions through group outings, volunteer work, or participation in social clubs Schepens, Van Puyenbroeck, and Maes, 2019a. Elderly individuals with ID may experience depression, anxiety, or other mental health conditions. Provide access to mental health services, such as counselling or therapy, to help them manage their emotions and cope with stress Garcia-Domnguez et al., 2020. Elderly individuals with ID may have complex medical needs that require specialized care. Ensure they receive regular medical check-ups and appropriate treatment for any health conditions Garcia-Domnguez et al., 2020. Ensure safe and comfortable living conditions: Elderly individuals with intellectual disabilities may require modifications to their living arrangements to ensure their safety and comfort. Ensure their living environment is safe, accessible, and adapted to their needs Miskimmin et al., 2019. Engage elderly individuals with ID in meaningful and fulfilling activities like art, music, or hobbies. This can help promote a sense of purpose and well-being Schepens, Van Puyenbroeck, and Maes, 2019a. Support their independence: Encourage elderly individuals with ID to be as independent as possible while providing support and assistance as needed. This can help them maintain a sense of autonomy and self-worth Schepens, Van Puyenbroeck, and Maes, 2019a.

2.7 The GENCAT tool

GENCAT is a tool developed by Verdugo et al. Verdugo et al., 2010 under the sponsorship of the government of Catalonia, Spain, to assess its citizens' living level. The acronym stands for "Generalitat de Catalunya," the name of the regional government, and "Indicators de Qualitat de Vida de Catalunya," which translates to "Indicators of Quality of Life in Catalonia." The GENCAT tool measures various aspects of life, including health, education, environment, social participation, safety, and economic well-being. It includes a set of indicators for each area, which are used to collect data and evaluate the region's general standard of living Gómez et al., 2013. The GENCAT tool is essential because it provides policymakers and researchers with a comprehensive system for comprehending the factors contributing to the QoL of citizens in Catalonia. By collecting and analyzing data on these indicators, the government can identify areas where improvements are needed and develop policies to address them. Moreover, the GENCAT tool calculates the quality of life dimensions value and corresponding index value. Quality of life index value indicates whether a dependent person needs support or not Yadav et al., 2023. The GENCAT tool is essential for measuring and improving Catalonia's living standards. It is a model for other regions and countries seeking to develop similar frameworks.

2.8 Statistical analysis

F-test, Z-test, Chi-square test and t-test are statistical tests used to make inferences about population parameters based on sample data. Each test has its specific use case and assumptions.

- F-test: Two populations' variances are compared using an F-test. In ANOVA (analysis of variance), it is frequently utilised to determine whether the average of three or more groups are equivalent Griffiths and Hill, 2022.
- Z-test: A z-test is performed once the population variance has been determined and the sample size is big (usually $n > 30$). When there is a large sample size, it is utilised to evaluate population mean hypotheses Mishra et al., 2019.
- Chi-square test: The significance of a connection between two category variables is assessed using the chi-square test. It is utilised to assess how closely observed and anticipated frequency distributions of various categories match up in a sample Goual and Yousof, 2020.
- t-test: When there is uncertainty regarding a population's variation or with a tiny sample size (usually $n < 30$), a t-test is utilised. In situations where the limited sample size, it is employed to test population mean hypotheses Mishra et al., 2019.

2.8.1 Student's t-test

The Student's t-test, a statistical hypothesis test, determines whether the two data groups' mean differ significantly. The test has the name of William Sealy Gosset, who wrote under the pen name "Student" and released his work in 1908 Mishra et al., 2019. The t-test contrasts the sample means, standard deviations, and sample sizes to determine a test statistic called t. It compares the mean of two samples Abebe, 2019. In the context of the t-test, a hypothesis is a statement about the population traits from which the samples are drawn. The two main hypotheses in the t-test are the null hypothesis, as well as the alternative hypothesis.

The mean of the two groups under comparison does not differ, according to the null hypothesis (H0) in a t-test. Put another way, any observable difference in the sample results from randomness or sampling variability. Usually, the null hypothesis is expressed as follows:

$$H0 : \mu 1 - \mu 2 = 0, \tag{2.1}$$

Where $\mu 1$ and $\mu 2$ are the comparison population mean of both groups.

The alternative hypothesis (Ha), the research hypothesis, is the opposite of the null hypothesis. According to this statement, there is a clear distinction between the means of the two groups under comparison.

$$Ha : \mu 1 - \mu 2 \neq 0, \tag{2.2}$$

The t-test contrasts the alternative hypothesis with the null hypothesis using a test statistic (t-value) and a critical value calculated from the t-distribution Abebe, 2019. Rejecting the null hypothesis whenever the t value exceeds the critical value, it is determined that the two groups' mean differ significantly. Consider the scenario where the t-value is less than the crucial value. If that is the case, we cannot rule out the null hypothesis and draw the inference that there is insufficient evidence to demonstrate that the means of the two groups do not differ significantly.

2.9 Problem description

The advancement of technology leads towards intelligent homes, which may be used to care for older people. These homes contain various cobots and are equipped with various sensors, which observe the activities performed by the ID older people. They may assist them by performing the required tasks and analysing the quality of life based on the inputs. One of the significant challenges currently involved is to provide the human-like with the capacity to anticipate motion based on an observer's motion. This ability of cobots makes them advanced to understand the requirement of the disabled person by observing them. Therefore, firstly, we tackled the issue of human motion prediction. After that, the focus shifted towards developing a support system to improve the QoL of the ID elderly. A sophisticated, intelligent system that can analyse the dimensions of QoL and help ID individuals is necessary to enhance ID senior citizens' quality of life. The following discussed tasks contain many challenges. Below is a summary of the problems we have resolved:

1. Human motion (HM) mechanics are constantly changing and complex since they have never been consistent or predictable. The model must simultaneously detect the temporal and spatial features to predict HM. Various approaches use different techniques to detect temporal features. Still, after applying the methods, there needs to be more consistency between the last observed and the first predicted pose. We employ the residual inception module to learn the temporal feature because it can learn the salient features using multiple different-size kernels. Residual connection helps to cope with the inconsistency between the last observed and first predicted pose. And then, the proposed model learns the spatial features using graph convolutional blocks.

2. Current research uses discriminative models to handle the HMP problem and presents the results when the data has a homogeneous distribution. It does not deal with the domain shift issue, which occurs when the training and testing data are heterogeneous, as is when such models are deployed in practice. We use the linear

matrices-based augmented deep learning model to predict human motion in out-of-distribution scenarios.

3. ID older people must be analysed frequently to improve their QoL. It is very challenging for the caregiver and the doctors to analyse the patient's situation frequently; therefore, there is a need for an intelligent system that can analyse the patient by taking inputs from the patient and suggest possible actions to improve their QoL. We develop an intelligent system using machine learning algorithms to provide the necessary action to improve their QoL and track the patient's situation.

4. Analysing QoL requires analysing its dimensions. Earlier, professionals used the GENCAT tool to calculate the QoL's dimension values using sixty-nine questionnaires, which is tedious and requires experts to calculate dimensions value. In this study, we propose to forecast the value of the QoL's dimension using a trained machine-learning model.

5. A critical term to deciding whether an ID older person needs support is quality of life index value. Calculating index value from the eight QoL dimensions value requires time and expertise to calculate the value. Therefore, we use a trained machine learning model to predict the index value directly from the sixty-nine questionnaires. It removes the two-step process of calculating the eight dimensions value first and then from the eight dimensions value to the index value.

Part II

Improving Human Motion Prediction through Deep Neural Network Architectures

Chapter 3

Development of Human Motion Prediction Strategy using Inception Residual Block

Summary

In robotics and computer vision, predicting human motion is a critical task. It has versatile application potentials, such as human-robot interactions, human action tracking for airport security systems, autonomous car navigation, and computer gaming, to name a few. However, predicting human motion based on past actions is challenging due to the difficulties in correctly detecting spatial and temporal features. We propose an Inception Residual Block(IRB) detect temporal features in human poses due to its inherent capability of processing multiple kernels to capture salient features. Here, we propose using multiple 1-D Convolution Neural Networks (CNN) with different kernel sizes and input

sequence lengths to concatenate them to get proper embedding. As kernels stride over different receptive fields, they detect smaller and bigger salient features at multiple temporal scales. Our main contribution is to propose a residual connection between the input and the output of the inception block to have a continuity between the previously observed pose and the next predicted pose. With this proposed architecture, it learns prior knowledge much better about human poses, and we achieve much higher prediction accuracy, as detailed in the paper. Subsequently, we further propose to feed the output of the IRB as an input to the Graph Convolution Neural Network (GCN) due to its better spatial feature learning capability. We perform a parametric analysis for a better design of our model. Subsequently, we assess our methodology on the Human 3.6M dataset and the CMU MoCap dataset. We compare our short-term and long-term predictions with the leading papers, where our model outperforms most of the pose results, the detailed reasons for which have been elaborated in the paper.

3.1 Introduction

Predicting human movements is essential for the machine to predict human movement. Recognizing the human pose tells human behaviour to the machine. It involves human tracking Gong et al., 2011 Gupta et al., 2014, motion generation Liu and Mu, 2021, human action recognition and prediction Kong and Fu, 2022, online gaming Espinoza et al., 2022. Generally, it is used where machines need to interact with humans and is useful in computer vision and robotics. In human-robot interaction, Koppula and Saxena, 2013, if a human is walking and a robot has to approach a human, the robot must predict human motion. In Autonomous car driving Paden et al., 2016, a car must predict pedestrian motion Kalatian and Farooq, 2022 to prevent accidents.

Our task in this paper is to predict human motion and achieve higher accuracy. We have given past human poses and must forecast the next human pose. To achieve higher accuracy key point is to detect temporal features. Authors in Fragkiadaki et al., 2015 Martinez, Black, and Romero, 2017 Yadav and Nandi, 2020 use RNNs to detect temporal features. There are two problems with the RNN-based models. The first problem Fragkiadaki et al., 2015 Martinez, Black, and Romero, 2017 is, in RNN, errors add up to every step of sequences. It leads to illogical predictions at testing time. Second, as it is found in Fragkiadaki et al., 2015 and Martinez, Black, and Romero, 2017, there is an inconsistency between the last observed and first forecasted frames. This inconsistency exists because global smoothness is not encouraged by the frame-by-frame regression.

Mao et al. Mao et al., 2019 consider this task a forward feed network. They have used the Discrete Cosine Transform (DCT) to encrypt temporal features and graph convolution network Velickovic et al., 2017 as a feed-forward network for prediction. Lebailly et al. Lebailly et al., 2020 follow Mao et al., 2019 research and, instead of DCT, use inception module Szegedy et al., 2015. In line with Lebailly et al., 2020, for our specific problem of predicting human poses based on past actions, we hypothesize that using residual block with inception module can provide better continuity between previously observed pose and the next predicted pose. In the inception module, we propose using 1D CNN since 1D CNN plays a crucial role in detecting temporal features.

CNN considers three architectural concepts: shared weights, spatial sub-sampling, and local receptive field Cruttwell et al., 2022. The receptive field slides over data by the predefined value called stride. While sliding, they overlap, cover the whole data series, and detect input data features. CNN shares weight and bias for a layer, making it more computational and cost-efficient than artificial neural networks. Weights and bias learn by training. By doing one convolution operation, one feature is detected. When the kernel strides in 1-dimensional time series data, it detects temporal features of input data.

We applied the inception module and the residual block in our work to detect temporal features. Now the question is how to detect spatial features of a human pose. In our work, spatial features mean learning dependencies between human joints of the pose. Our approach should not depend on fixed convolution filter size, like detecting temporal features. We are using GCN Welling and Kipf, 2016 to learn graph connectivity and detect spatial features. GCNs are a robust feed-forward neural network that learns human pose joint dependencies Mao et al., 2019. As an input, they take temporal encoding and predict human motion.

Our proposed architecture has two parts, the first is on IRB, and the second is on the Graph convolution network. Our main contribution is IRB. In the inception module Szegedy et al., 2015, we apply the different kernel sizes to get different receptive fields to detect salient features. By considering multiple sizes of kernels, we have two advantages. First, we are relieved of the concern of selecting the ideal kernel size for input. Second, the model gets a better look at the input data. It detects smaller and more prominent vital features. Adding residual blocks makes it easier to learn residual mapping than original mapping He et al., 2016. It adds continuity between a previously observed pose and the temporal encoding of the pose. Using the Inception residual module, we encode human poses for temporal encoding. IRB takes input joint trajectory and prepares output that is input for GCN by generating temporal features.

The major contributions of the paper:

- To generate temporal encoding, we proposed a residual connection between the input and inception blocks, which helps to learn salient features effectively.

- We have done parametric analyses on the Human3.6M dataset for the number of temporal features to identify an optimal number of features that are neither overfit nor underfit.

- With the help of the Inception Residual Block and GCN, we carry out end-to-end learning.

This paper has divided into five parts. In the First Part, we discussed human motion prediction, its application, and problem analysis and gave a little overview of 1D CNN. In the second part, we have discussed related work. In the third part, we have shared our methodology, where we formulate a problem statement, discuss preliminaries, and explain our architecture and training details. In the Fourth part, we have shown our implementation details, discussed the parametric analysis, and compared our result with baseline methods. In the final part, we have written our conclusion and future possibilities.

3.2 Related Work

3.2.1 Prediction of human motion using RNN-based algorithms

RNNs are the default choice to detect time-dependent features. Liu and Liu, 2020 considers human motion prediction as sequence-to-sequence prediction. Fragkiadaki et al., 2015 have introduced 2 approaches: LSTM-3LR and Encoder-Recurrent-Decoder(ERD). In both architectures, they used a layer of LSTM cells. However, they have added a non-linear encoder for data pre-processing in ERD architecture. To prevent error accumulation, they added noise to the input. However, adding noise made long-term predictions tough and inconsistent. Sang et al. Sang, Chen, and He, 2020 also proposed two models: At-seq2seq and seq2seq. The At-seq2seq model with the attention mechanism used GRU cells in the encoder and decoder. In seq2seq, they did not use the attention mechanism and got better results with the attention mechanism. Jain et al. Jain et al., 2016 have suggested S-RNN(structural Recurrent Neural Network). Using S-RNN, they made a Spatio-temporal graph and detected spatial and temporal features of the human pose. However, they manually designed graphs, which were not flexible and lacked long-term dependencies. Martinez et al. Martinez, Black, and Romero,

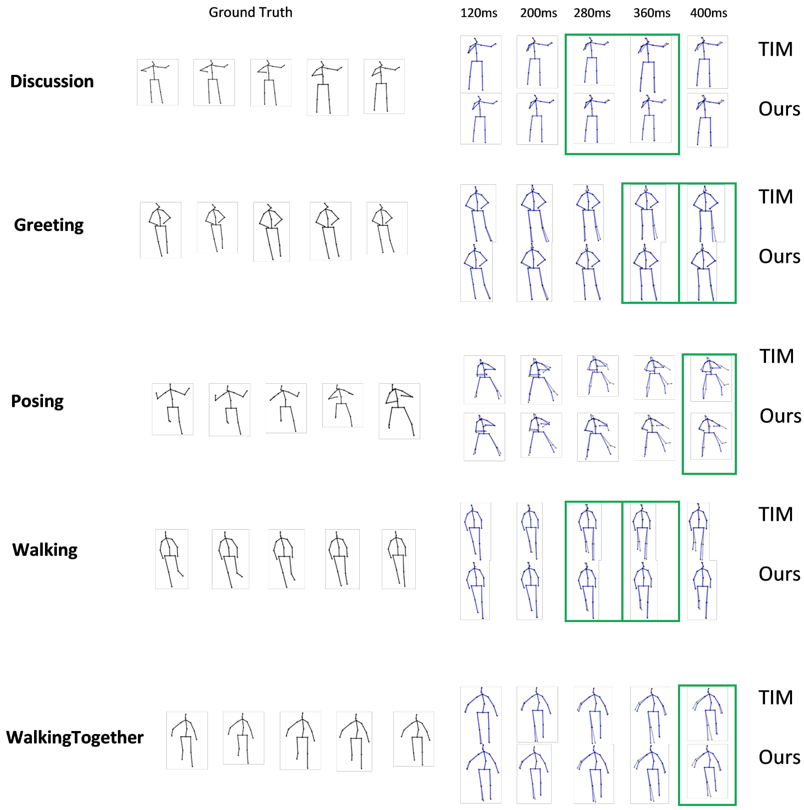


Figure 3.1: Pose comparison between TIM model and ours: For an activity On the left side, we have shown ground truth(in black colour). On the right side, we have shown prediction. TIM model prediction is shown in 1st row, and our model is in 2nd row(written at the end). We have superimposed ground truth with both methods where the blue pose is predicted, and the black pose is the ground truth. Some of our best prediction comparisons with the TIM model are highlighted.

2017 have proposed a simple baseline method. They have also used the sequence-to-sequence architecture and added residual connections between the input and output of RNN cells. Surprisingly their approach outperforms most of the previous solutions. Yadav et al. Yadav and Nandi, 2020 have applied an adaptive sampling-based cost function instead of a sampling-based loss function and got a better result for many poses than Martinez, Black, and Romero, 2017.

3.2.2 Feed-forward approaches forecasts human motion

To encounter discontinuity in RNNs, fully-connected network Butepage et al., 2017 and convolutions Li et al., 2018 approaches were studied. Butepage et al. Butepage et al., 2017 used various techniques to detect temporal characteristics while modelling human poses as input to fully connected layers. To detect spatial features, they used kinematic trees. They have built a tree on a specific dataset which made it inflexible. So kinematic trees have not synced with different body parts such as limbs. Li et al. Li et al., 2018 have introduced a convolutional sequence to sequence. Convolution kernels have captured spatial and temporal features. They have fixed kernel sizes, so they were inflexible and could not detect features properly.

Kjellstróm et al. Bütepage, Kjellström, and Kragic, 2018 have suggested Conditional Variational Autoencoder(CVAE), a probabilistic approach to predict and generate human motion. Based on CVAE, Kragic et al. Butepage, Kjellstrom, and Kragic, 2019 have proposed a semi-supervised recurrent neural network(SVRNN) to detect, classify, and predict human pose. Aliakbarian et al. Aliakbarian et al., 2020 have used a recurrent-encoder-decoder network with CVAE for prediction. Between the encoder and decoder, they used a CVAE-equipped future pose autoencoder.

Human pose prediction leads to recognizing action and helps to understand the action performed by a human. Zhang et. alLi et al., 2019 proposed a new architecture, Spatio-temporal manifold network(STMN), for Action recognition. They modelled input videos using the embedding method and later applied the Alternating Direction Method of Multipliers and Backward Propagation (ADMM-BP) algorithm. They outperformed peer publications in terms of results. Li et al. Li et al., 2021 have developed a memory attention network that incorporates a Temporal Attention Re-calibration Module (TARM) and a Spatio-Temporal Convolution Module (STCM). TARM has incorporated a residual connection to detect temporal features, while STCM detects both temporal and spatial features. Shum et al. Zhang et al., 2018 demonstrated a brand-new, inexpensive descriptor called 3D Histograms of Texture (3DHoTs). It is used to extract distinguishing characteristics from a collection of depth maps. They divided frames into x,y, and z planes and found salient features in each plane. They also presented a multi-class boosting classifier(MBC) to utilize plane features for action classification. Yang et al. in Zhang et al., 2017 use 3D texture histograms and a multi-class boosting classifier to get the best results for the action recognition task.

3.2.3 Graph convolution networks

GCN captures information passing through the node in the graph. It involves mainly node classification Welling and Kipf, 2016 Hamilton, Ying, and Leskovec, 2017 , image classification Monti et al., 2017, machine translation problem Marcheggiani, Bastings, and Titov, 2018, recommendation system Ying et al., 2018. GCN does convolution operations on graph data structure in two ways defined in Welling and Kipf, 2016, and Yuan et al., 2022. Kipf et al. Welling and Kipf, 2016 have used a layer-wise propagation rule for nodes inspired by first-order approximation of spectral convolutions on graphs. They have applied convolution operation based on the graph structure. It is limited to the characteristics of the graph. While Yuan et al. Yuan et al., 2022, have learned

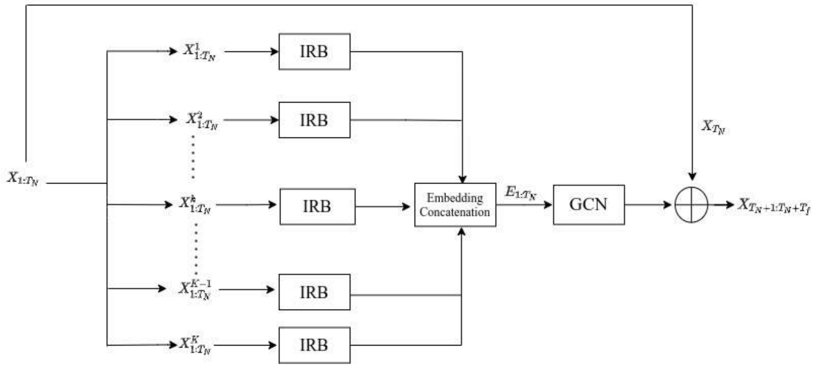


Figure 3.2: The complete architecture of our suggested method: We break human poses $X_{1:T_N}$ into K human joint $X^1_{1:T_N}...X^K_{1:N}$ trajectory. Each $X^k_{1:T_N}$ pose fed to IRB(Fig. 3) and generate $E^k_{1:T_N}$ embedding. All embedding are concatenated and made node features $E_{1:T_N}$ of the graph, which is fed to GCN and added with X_N most recent pose.

node connectivity using the node neighbourhood. So it is more flexible for GCN. So we are using GCN architecture motivated by Yuan et al., 2022 to learn node embedding adaptively.

3.2.4 Inception block

The inception Module was first introduced in the ILSVRC14 image classification competition by Szegedy et al. Szegedy et al., 2015. They were used in GooleLeNet, and their result proved that one could achieve higher accuracy using the inception block. It makes the network broader rather than more profound, which helps to reduce computation costs. Ioffe et al. Szegedy et al., 2016 propose four design principles to build CNN architecture. Based on these principles, they introduced inception v2 and v3 architecture and got a better result. Szeged et. al Szegedy et al., 2017 suggested inception-v4. It was a bit more complex than the previous ones but gave a good result. We are using simple inception architecture Szegedy et al., 2015 without a max-pooling layer.

3.2.5 Baseline methods

We are comparing our results in mean per joint position error (MPJPE) Error Ionescu et al., 2013 because all of the baseline methods we are considering also use the MPJPE metric. We use the following baselines to compare our results: 2 famous sequences to sequence-based methods. First, Martinez et al. Martinez, Black, and Romero, 2017 used the well-known RNN method, and Li et al. Li et al., 2018 used Convolution to encode and decode data. Mao et al. Mao et al., 2019, and Tim et al. Lebailly et al., 2020 encode human poses. Mao et al. Mao et al., 2019 used DCT + GCN, and Tim et al. Lebailly et al., 2020 used TIM + GCN. We didn't compare our result with Li et al., 2020c Yadav and Nandi, 2020 papers because they used different metric mean angle errors.

3.3 Methodology

3.3.1 Problem statement

We are using human pose 3D joint positions. For each joint, we have its x,y, and z coordinate. We have given time-series T_N 3D-joint coordinates as shown in equation 1.

$$X_{1:T_N} = [X_1, X_2, ...X_{T_N}], \tag{3.1}$$

where $X_{1:T_N} \in \mathbb{R}^{T_N * K * D}$, T_N is the number of input time steps, K is several joints in human pose and $D = 3$ is feature dimension x,y,z. Our task is to predict $X_{T_N+1:T_N+T_f}$, continuous human poses, where T_f number of future time steps.

3.3.2 Proposed approach

We are making a graph represented by an adjacency matrix where the human joints of the pose represent the nodes. To calculate node features,

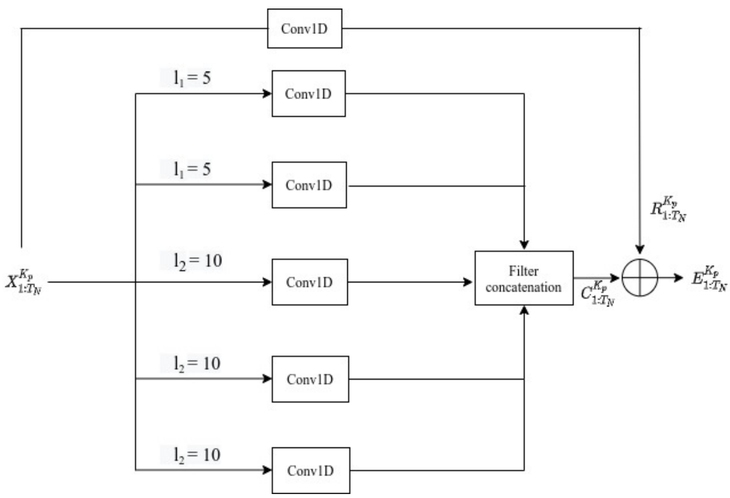


Figure 3.3: Full architecture of Inception Residual Block: For *k*th joint, feature dimension $p \in \{x, y, z\}$, $X^{k_p}_{1:T_N}$ is fed to 5 Conv1d filters with different input length and their output is concatenated and generates $C^{k_p}_{1:T_N}$. To match $X^{k_p}_{1:T_N}$ same shape whole feature dimension trajectory is passed to kernel size 1 Conv1d added with $C^{k_p}_{1:T_N}$ and produces feature dimension embedding $E^{k_p}_{1:T_N}$

we propose an IRB. These node features are fed to GCN. GCN learns node features of the graph and produces the future expected poses $X_{T_N+1:T_N+T_f}$. Details of network architecture we have discussed in the following sections.

3.3.3 Inception module

When we increase the depth of CNN, the model becomes more susceptible to overfitting Szegedy et al., 2015 as parameters increase model becomes harder to train. Determining the proper kernel size is another challenge because salient features are great in some training examples while some are smaller. Szegedy et al. Szegedy et al., 2015 apply multiple kernels parallel in the model to encounter all these problems simultaneously. In the inception module Szegedy et al., 2015, kernel size used 1, 3, 5 and max-pooling layer. Using multiple kernel models can capture both bigger and smaller salient features. By concatenation, the model becomes wider, computationally inexpensive, and overcomes overfitting.

3.3.4 Temporal encoding using inception residual block

We have shown the IRB module in Figure 3. The main purpose of this block is to detect temporal features of the human joint's feature dimension trajectory (x, y, z). It takes feature dimension trajectory $X^{k_p}_{1:T_N}$ and generates embedding $E^{k_p}_{1:T_N}$.

IRB takes two different size $l_1 = 5$ and $l_2 = 10$ input length feature dimension trajectory. In inception block, For l_1 length, we apply 2 1D convolution operation and for l_2 length 3 convolution operation. On each input, different size kernels are applied. For more details on kernel size and the number of kernels, refer to Table 1. As Lebailly et al. Lebailly et al., 2020 suggested, a larger kernel size is applied for shorter input length, shorter kernel size, and larger input length. Small kernel

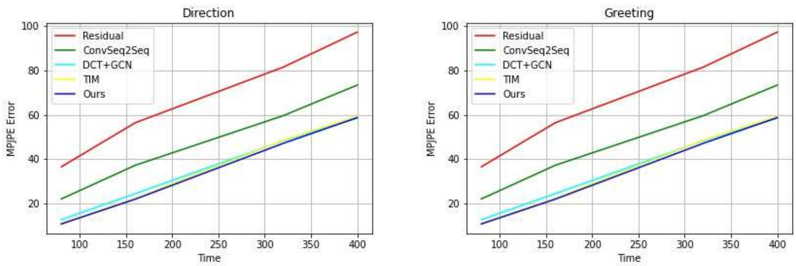


Figure 3.4: Comparison of MPJPE error for Direction and Greeting pose with "Our" approach and Baseline methods

detects small salient, and large kernel detects big salient features. All 1D CNN output are concatenated into one embedding $C^{k_p}_{1:T_N}$.

To match same shape $X^{k_p}_{1:T_N}$ with $C^{k_p}_{1:T_N}$ we apply 1D convolution with kernel size 1 and getting $R^{k_p}_{1:T_N}$. We have added residual connection between $R^{k_p}_{1:T_N}$ and $C^{k_p}_{1:T_N}$. In other words, between the input and the output of the inception block. There are two main advantages. First, removing discontinuity between the previously observed pose and the next predicted pose addition helps to learn prior knowledge much better about human poses. Second, during the backpropagation algorithm, the gradient can flow through the model directly, which prevents gradients from vanishing or exploding. So, residual connection He et al., 2016 improves performance. Finally residual connection generates embedding $E^{k_p}_{1:T_N}$.

3.3.5 Graph convolution networks

We use the same GCN suggested by Mao et al., 2019 to predict human motion. A human 3D skeleton can interpret as a graph with K fully connected nodes, where K is several body joints. The graph is denoted by $A \in \mathbb{R}^{K*K}$ adjacency matrix where the nodes are human joints. Graph

Row No.	Sequence input length	Number of kernels	Kernel size	Number of features
1	5	17	2	17*4
2	5	16	3	16*3
3	10	14	3	14*8
4	10	13	5	13*6
5	10	11	7	11*4
6	10	-	-	10
7	-	-	-	360

Table 3.1: Details of Inception Block in IRB

convolution layer takes hidden feature matrix $H \in \mathbb{R}^{K*F}$, where F is features of the preceding layer. Inside the top layer, the hidden matrix is the output of IRB output. In the Graph Convolution layer, each node accumulates node neighbour features. We use multiple stacked graph convolution layers, and each layer performs a given operation shown in Equation 2.

$$H^{(p+1)} = \sigma(A^{(p)}H^{(p)}W^{(p)}), \tag{3.2}$$

here $\sigma(.)$ is the activation function and $A^{(p)}$ is adjacency matrix for layer p. $H^{(p)}$ is hidden feature matrix for layer p. $W^{(p)}$ is the weight matrix of layer p where $W^{(p)} \in \mathbb{R}^{F*F}$. $H^{(p+1)}$ is output hidden feature matrix for layer (p).

Both A and W are learnable matrices. The learnable matrix leads to better results, proven by Mao et al. Mao et al., 2019. A and W are trained using a standard backpropagation algorithm.

GCN takes IRB generated embedding $E_{1:T_N}$. GCN learns the node features of the graph. As it learns node features, it detects spatial features of the graph. The output of GCN will be added with residual block as the most recent human pose X_N is shown in Fig 2. It gives our predicted output $X_{T_N+1:T_N+T_f}$.

Motion	Walking	Eating	Smoking	Discussion
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
ResidualMartinez, Black, and Romero, 2017	23.8 40.4 62.9 70.9	17.6 34.7 71.9 87.7	19.7 36.6 61.8 73.9	31.7 61.3 96.0 103.5
ConSeq2Seq Jain et al., 2016	17.1 31.2 53.8 61.5	13.7 25.9 52.5 63.3	11.1 21.0 33.4 38.3	18.9 39.3 67.7 75.7
DCT + GCN Mao et al., 2019	**8.9** 15.7 29.2 33.4	8.8 18.9 39.4 47.2	7.8 14.9 25.3 28.7	9.8 22.1 39.6 44.1
TIMLebailly et al., 2020	9.3 15.9 30.1 34.1	8.4 18.5 **38.1 46.6**	**6.9** 13.8 24.6 29.1	8.8 21.3 40.2 45.5
Ours	**8.9 15.2 29.1 32.7**	**8.1 18.3** 39.1 47.0	**6.9 13.6 23.9 28.2**	**8.4 19.7 36.1 40.7**

Table 3.2: Detail results of short-term prediction, measure in MPJPE error of different poses, e.g., Walking, Eating, Smoking, and Discussing.

Motion	Directions	Greeting	Phoning	Posing
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
ResidualMartinez, Black, and Romero, 2017	36.5 56.4 81.5 97.3	37.9 74.1 139.0 158.8	25.6 44.4 74.0 84.2	27.9 54.7 131.3 160.8
ConvSeq2Seq Jain et al., 2016	22.0 37.2 59.6 73.4	24.5 46.2 90.0 103.1	17.2 29.7 53.4 61.3	16.1 35.6 86.2 105.6
DCT + GCN Mao et al., 2019	12.6 24.4 48.2 **58.4**	14.5 30.5 74.2 89.0	11.5 20.2 37.9 **43.2**	9.4 23.9 66.2 82.9
TIMLebailly et al., 2020	11.0 22.3 48.4 59.3	**13.7** 29.1 72.6 **88.9**	11.5 19.8 38.5 44.4	7.5 22.3 64.8 80.8
Ours	**10.7 21.9 47.1** 58.7	**13.7 28.7 71.6** 89.7	**11.2 19.4 37.5** 43.9	**7.2 21.1 57.9 72.1**

Table 3.3: Detail results of short-term prediction, measure in MPJPE error of different poses, e.g., Direction, Greeting, Phoning, and Posing.

Motion	Purchases	Sitting	Sitting Down	Taking Photo
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
ResidualMartinez, Black, and Romero, 2017	40.8 71.8 104.2 109.8	34.5 69.9 126.3 141.6	28.6 55.3 101.6 118.9	23.6 47.4 94.0 112.7
ConvSeq2Seq Jain et al., 2016	29.4 54.9 82.2 93.0	19.8 42.4 77.0 88.4	17.1 34.9 66.3 77.7	14.0 27.2 53.8 66.2
DCT + GCN Mao et al., 2019	19.6 38.5 64.4 **72.2**	10.7 24.6 50.6 62.0	11.4 27.6 56.4 67.6	6.8 15.2 **38.2** 49.6
TIMLebailly et al., 2020	19.0 39.2 65.9 74.6	**9.3 22.3 45.3 56.0**	11.3 28.0 54.8 64.8	**6.4** 15.6 41.4 53.5
Ours	**18.4 37.2 64.3** 74.5	**9.3** 22.8 47.0 58.5	**10.5 25.5 49.6 60.0**	**6.4 15.2** 39.9 **51.6**

Table 3.4: Detail results of short-term prediction, measure in MPJPE error of different poses, e.g., Purchases, Sitting, Sitting down, and Taking a photo.

Motion	Waiting	Walking Dog	Walking Together	Average
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
ResidualMartinez, Black, and Romero, 2017	29.5 60.5 119.9 140.6	60.5 101.9 160.8 188.3	23.5 45.0 71.3 82.8	30.8 57.0 99.8 115.5
ConvSeq2Seq Jain et al., 2016	17.9 36.5 74.9 90.7	40.6 74.7 116.6 138.7	15.0 29.9 54.3 65.8	19.6 37.8 68.1 80.2
DCT + GCN Mao et al., 2019	9.5 22.0 57.5 73.9	32.2 58.0 102.2 122.7	**8.9 18.4** 35.3 44.3	12.1 25.0 51.0 61.3
TIMLebailly et al., 2020	9.2 **21.7 55.9 72.1**	29.3 **56.4** 99.6 119.4	**8.9** 18.6 35.5 44.3	11.4 24.3 50.4 60.9
Ours	**9.0 21.7** 56.8 72.5	**29.2** 58.3 **98.3 114.9**	9.1 18.7 **34.2 43.0**	**11.1 23.8 48.8 59.2**

Table 3.5: Detail results of short-term prediction, measure in MPJPE error of different poses, e.g., Waiting, Walking with the dog, Walking together, and Average.

3.4 Experiments and evaluation

To learn the numerous parameters that the deep architecture creates, a sizable dataset is needed. We use the two most popular datasets (H3.6M and CMU MoCap) for the human motion prediction task. Details about the datasets are given in the following section.

3.4.1 Human 3.6M dataset

We are utilising the Human 3.6M dataset Ionescu et al., 2013 to predict human movements. In this dataset, seven subjects are doing 15 actions walking, eating, smoking, walking together, etc. 32 joints represent a human pose. We used the same data pre-processing Martinez, Black, and Romero, 2017 Mao et al., 2019 and removed global rotations and translations. Similar to previous work Mao et al., 2019 Li et al., 2018 Martinez, Black, and Romero, 2017 for training we used 5 subjects. One subject is used for checking the validation, and one for testing. We are considering 25 frames per second. So, for 80ms, we are considering one frame. For testing, we use subject number 5, and similar to previous work same sequences are considered for testing.

3.4.2 CMU MoCap dataset

Additionally, we present findings from the CMU motion capture dataset, as described in Mao et al., 2019 (CMU-Mocap). We replicate the previous study's data format, train/test splits, and analysis methods Mao et al., 2019 to make a fair comparison. We utilize two representations: a 64-dimensional vector representing the joint angle as an exponential map and a 75-dimensional vector representing the 25 joints in 3D Cartesian coordinates. The dataset includes eight distinct professional actions. Basketball, directing basketball signals, traffic, soccer, running, walking, jumping, and window-washing are some of these actions.

Motion	Walking	Eating	Smoking	Discussion	Average
milliseconds	560 1000	560 1000	560 1000	560 1000	560 1000
ResidualMartinez, Black, and Romero, 2017	73.8 86.7	101.3 119.7	85.0 118.5	120.7 147.6	95.2 118.1
ConvSeq2Seq Jain et al., 2016	59.2 71.3	66.5 85.4	42.0 67.9	84.1 116.9	62.9 85.4
DCT + GCN Mao et al., 2019	42.3 51.3	56.5 68.6	32.3 60.5	70.5 103.5	50.4 71.0
TIMLebailly et al., 2020	39.6 46.9	56.9 68.6	33.5 61.7	**68.5** 97.0	50.4 71.0
Ours	**35.6 44.3**	**55.4 68.2**	**31.1 56.0**	68.8 **75.5**	**47.7 61.0**

Table 3.6: For long-term prediction, MPJPE error of different poses of four actions and their average for H3.6M dataset.

3.4.3 Implementation details and model configuration

We train our model to 50 epochs. For short-term prediction used, ten input and output frames were used. We followed the same work as Mao et al., 2019 and set the learning rate to 0.0005, decaying by 0.96 with every two epochs. For training, our batch size is set to 16. For GCN, we use 12 stacked convolution layers. For the implementation, we use PyTorch Paszke et al., 2017 and use ADAM optimizer Zhang, 2018 and Tanh activation function. We run our code on NVIDIA Tesla V-100(16GB) GP-GPU(General Purpose GPU). We conducted multiple experiments and chose the best result.

3.4.4 Training details

For K joints, $K * 3$ IRB is used. In each IRB block, multiple different-size 1D kernels are used. Smaller kernels like 2 and 3 detect more minor features. E.g., in a smoking pose, only one hand is moving while another body part is still. A smaller kernel size can easily detect smoking features. A larger kernel size like 5, or 7 considers a sizeable receptive field. In the walking pose the whole body is moving, both hands and legs. To detect continuously moving legs and hands, larger kernel sizes are required.

We have used 1D convolution as we are feeding feature dimension trajectory at multiple temporal scales, which is taking multiple length input $l_1 = 5$ and $l_2 = 10$. In Table 1, we have shown details of IRB. In row 1, we take input sequence length five and apply kernel size 2, stride 1, and zero-padding with 17 Conv1D kernels. We are concatenating Conv1D output, getting 17 * (5 - 2 + 1) features and row 7 concatenate input sequence length 10. After performing all concatenation, we get temporal embedding 360. This 360-size embedding is added with input as a residual connection. Every IRB block embedding $E^{k_p}_{1:T_N}$ concatenated where $k \in K$ and $p \in x, y, z$ (p = 3) & $E_{1:T_N}$ fed to GCN. IRB and GCN are training for an end to end learning.

3.4.5 Evaluation metric

For the loss function, we use MPJPE Ionescu et al., 2013 instead of mean square error to calculate the error between the predicted pose and the actual pose. We also use this evaluation metric for comparing our results with other states of the art results. MPJPE is calculated by the following equation

$$Error = \frac{1}{K(T_f + T_N)} \sum_{n=1}^{T_N+T_f} \sum_{j=1}^{K*3} ||\hat{p}_{j,n} - p_{j,n}||^2, \tag{3.3}$$

where $\hat{p}_{j,n} \in R^3$ is predicted jth joint position at nth time frame and $p_{j,n}$ is corresponding ground-truth and K is several joints in human pose(multiplied by three because x,y,z feature dimension).

3.4.6 Optimization of features with parametric analysis

We train the inception module with a different number of kernels; as we are varying kernels, we have considered temporal embedding features from 223 to 460. However, on 460 features model becomes more prone to overfitting. In Fig 5. for 460 and 420 features, training error is significantly less while validation error is very high, indicating overfitting. The same behaviour can be seen in Table 7. For walkingdog3d400, the prediction error is very high.

For 223 and 300 features model was not overfit but can train more with features. Their validation error and training error is to the same

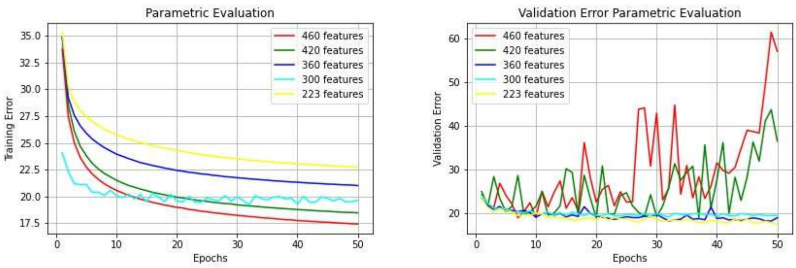


Figure 3.5: Analysis of Training and Validation error based on number of temporal features

No. of Features	Avg Training Loss	Avg Validation loss	walking3d400	walkingdog3d400	walkingtogether3d400
223	22.8	18.0	35.4	118.8	45.0
300	19.4	19.6	35.3	114.8	45.2
360	**21.1**	**18.6**	**32.7**	**114.8**	**43.0**
420	18.5	38.4	33.5	218.4	41.6
460	17.5	51.6	34.7	515.2	42.0

Table 3.7: Parametric Analysis of Number of Temporal features

degree. This behaviour is shown in Fig 5 and Table 7. We have trained our model with 360 features which we found best. We have chosen 360 features on some poses because we are getting better results, e.g., in table 7, walking3d400, walkingdog3d400, walkingtogether3d400.

We have modified residual connection Conv1D kernels based on the number of features, and we have changed only the number of kernels while other parameters are kept constant. In Table 7, we have taken the average over the last five epochs.

3.5 Results

We have shown our results of H3.6M in table no. 3.2,3.3,3.4,3.5 for all poses. We report our results over short-term prediction for 80, 160, 320, and 400ms and use the same input length frame ten and output length frame 10 to capture short-term prediction. We have taken the average of our results over the last five epochs. We get a better result for most of

Motion	**Basketball**	**Basketball Signal**	**Directing Traffic**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Residual sup. Martinez, Black, and Romero, 2017	18.4 33.8 59.5 70.5 106.7	12.7 23.8 40.3 46.7 77.5	15.2 29.6 55.1 66.1 127.1
convSeq2Seq Li et al., 2018	16.7 30.5 53.8 64.3 **91.5**	8.4 16.2 30.8 37.8 76.5	10.6 20.3 38.7 48.4 **115.5**
DCT + GCN Mao et al., 2019	14.0 25.4 49.6 61.4 106.1	3.5 6.1 11.7 15.2 **53.9**	7.4 15.1 31.7 42.2 152.4
TIM Lebailly et al., 2020	12.7 22.6 44.6 55.6 102.0	3.0 **5.6** 11.6 15.5 57.0	**7.1** 14.1 31.1 **41.4** 138.3
Ours	**11.9 20.7 41.3 54.6** 98.1	**2.7** 5.8 **10.8 14.9** 55.8	7.4 **13.8 30.9** 41.7 120.4
Motion	**Jumping**	**Running**	**Soccer**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Residual sup Martinez, Black, and Romero, 2017	36.0 68.7 125.0 145.5 195.5	15.6 19.4 31.2 36.2 43.3	20.3 39.5 71.3 84 129.6
convSeq2Seq Li et al., 2018	22.4 44.0 87.5 106.3 162.6	**14.3 16.3 18.0 20.2 27.5**	12.1 21.8 41.9 **52.9 94.6**
DCT + GCN Mao et al., 2019	16.9 34.4 76.3 96.8 164.6	25.5 36.7 39.3 39.9 58.2	11.3 21.5 44.2 55.8 117.5
TIM Lebailly et al., 2020	14.8 31.1 71.2 **91.3** 163.5	24.5 37.0 39.9 41.9 62.6	11.2 22.1 45.1 58.1 122.1
Ours	**13.6 30.3 70.8** 93.5 **159.8**	18.2 35.8 32.7 35.3 58.8	**10.4 20.1 41.7** 54.7 115.9
Motion	**Walking**	**Washing Window**	**Average**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Residual sup Martinez, Black, and Romero, 2017	8.2 13.7 21.9 24.5 32.2	8.4 15.8 29.3 35.4 61.1	16.8 30.5 54.2 63.6 96.6
convSeq2Seq Li et al., 2018	7.6 12.5 23.0 27.5 49.8	8.2 15.9 32.1 39.9 **58.9**	12.5 22.2 40.7 49.7 **84.6**
DCT + GCN Mao et al., 2019	7.7 11.8 19.4 23.1 40.2	5.9 11.9 30.3 40.0 79.3	11.5 20.4 37.8 46.8 96.5
TIM Lebailly et al., 2020	7.1 11.1 19.9 22.8 39.3	5.9 12.3 32.1 42.6 80.4	10.8 19.5 36.9 46.2 95.7
Ours	**7.1 10.8 19.3 21.6 35.8**	**5.3 11.6 30.6 38.7** 78.4	**9.6 18.6 34.8 44.4** 90.4

Table 3.8: For all activities in the CMU motion capture dataset for both short-term and long-term prediction in MPJPE.

the actions than baseline methods. In fig., 3.4 comparison of MPJPE error with the short-term prediction time frame is further elaborated. We conduct the parametric analysis and find that kernel size and the number of filters we use to give the best result. More than 360 node features are more prone to overfitting. We use the Tanh activation function. We also experiment with different input lengths to get the best result with $l_1 = 5$ and $l_2 = 10$.

We have shown our long-term prediction result for 560ms and 1000ms on the Human 3.6M dataset in table 3.6. We have also outperformed almost all baseline methods except for discussion pose 560ms. We have used the same input length frame ten and output length frame 25 to capture long-term prediction. Note that, on average, we get a better result for 560 and 1000ms.

Table 3.8 compares the MPJPE value with the other research works for the CMU MoCap dataset. It shows short-term (80ms, 160ms, 320ms, 400ms) and long-term (1000ms) results. Our results outperform the maximum scenario of others' work results for both short-term and long-term conditions. For the short-term scenario, our results outperform the other results, whereas, in the long-term case, our result is second best in this

	Residual sup Martinez, Black, and Romero, 2017	TIMLebailly et al., 2020	Our
Train Batch Size	16	16	16
Time/Pass	75ms	75ms	70ms

Table 3.9: The table displays the speed-based performance of the various models.

comparison.

We also contrasted our findings regarding the speed with findings from other studies. The training times for each pass for various algorithms are displayed below in table 3.9. Every pass includes the timing for both a forward and a backward pass. The batch size we specified during the training is 16, and our algorithm takes 70ms per pass.

3.6 Conclusions and recommendations for future works

Human Motion Prediction plays an essential role in action recognition and autonomous driving. It is addressed as a sequence-to-sequence problem using RNNs and GRUs extensively. We have discussed problems with RNNs architecture. Adding a residual connection between the inception block and input removes discontinuity between the previous and next predicted frame. It outperforms baseline models in both short and long-term prediction. We use IRB to detect temporal features and feed them to Graph Convolution Network in our work, and GCN learns spatial features of the graph and predicts human motion. Multiple IRBs can be used and tested with other datasets in future work.

Chapter 4

Implicit Regularization of a Deep Augmented Neural Network Model for Human Motion Prediction

Summary

Predicting human motion based on past observed motion is one of the challenging issues in computer vision and graphics. Existing research works are dealing with this issue by using discriminative models and showing the results for cases that follow a homogeneous distribution (in distribution) and not discussing the issues of the domain shift problem, where training and testing data follow a heterogeneous (out of distribution) problem, which is the reality when such models are used in practice. However, recent research proposed addressing domain shift issues by augmenting the discriminative model with a generative model and obtained better results. In the present investigation, we propose regularizing the extended network by inserting linear layers to minimize the

rank of the latent space and train the entire end-to-end network. We regularize the network to strengthen the model to deal effectively with domain shift scenarios. Both training and testing data come from different distribution sets; to deal with this, we toughen our network by adding extra linear layers to the network encoder. We tested our model with the benchmark datasets, CMU Motion Capture and Human3.6M, and proved that our model outperforms 14 OoD actions of H3.6M and 7 OoD actions of CMU MoCap in terms of the Euclidean distance calculated between predicted and ground truth joint angle values. Our average results of 14 OoD actions for short-term (80, 160, 320, 400) are 0.34, 0.6, 0.96, 1.07, and for CMU MoCap of 7 OoD actions for short-term and long-term (80, 160, 320, 400, 1000) are 0.28, 0.45, 0.77, 0.89, 1.46. All these results are much better than the other state-of-the-art results.

4.1 Introduction

Human motion (HM) has always been complex and diverse, and movement mechanics have never been distinct or predictable. The growth of deep learning results the way for predicting the HM by observing the sequence of immediate past motion and predicting the most likely future motion. Predicting HM plays a critical role in numerous real-world applications. In human-robot interactions, where certain robots work with humans, they need to predict the motion of the humans around them for safe navigation Gui et al., 2018. Geertsema et al., 2018 portrays how people with convulsive seizures can be saved from sudden death if the movement can be predicted. Studies Shirai, Geslin, and Richir, 2007; Rofougaran et al., 2018 have demonstrated how human motion prediction can be applied to entertainment systems involving video game controllers, which is helpful for nonprofessional game development. Zhang, Zhong, and Cai, 2022; Ma et al., 2022 showed the applicability of prediction for pedestrian intention in prominent public places. These are vast, diverse applications where human motion prediction can be used.

Considering the two categories of actions, both might scarcely vary, while the mechanics might vary within one category. That establishes a wide range of human motion possibilities, many not specified as specific human actions. These actions, which the model has not seen or ever been trained upon, are known as out-of-distribution (OoD) Hsu et al., 2020. Comparing OoD with In-distribution (ID) actions, upon which the model can be trained, demonstrates that robustness is necessary and critical for the model to adjust and predict OoD actions that can occur in the wild. Thus, urgent importance will be gained since real-world applications depend significantly on predicting human motion, not ID. Suppose such actions deviating from the standard-specific set of actions that the model is not trained on are not anticipated. In that case, it could even prove a life-or-death situation in particular applications. One instance of such applications is autonomous driving, where forecasting OoD actions is necessary to avoid accidents, as discussed by Singh and Srivastava, 2022; Kalatian and Farooq, 2022; Dafrallah et al., 2021. The conception of ID becomes questionable since the model will not be acquainted with the entire domain of possible human actions.

The kinematics and mechanics of joints vary significantly among the OoD actions. Since most OoD actions are not part of the domain of actions when performing on ID, the accuracy and generalizability of the predicted actions become debatable Bourached et al., 2020. In real-life situations, there are kinematic anomalies; thus, the OoD method should efficiently predict them in the short window. Thus, testing needs to be done for OoD actions rather than ID for the sake of the robustness of the prediction results. As a result, two frameworks are needed: one to assess OoD performance and the other to harden a model against anomalous cases.

Existing discriminative models of human kinematics do not discuss the OoD performance; they only show their results for ID performance.

Bourached et al., 2020 has an augmented discriminative model Mao et al., 2019 with a variational autoencoder (VAE) and shows their model performance primarily for the OoD scenario together with ID. VAE aims to strengthen the latent variable's representation capabilities. Bourached et al., 2020 learned the representation of human kinematics using the concept of the latent variable. This study motivates us to improve the latent variable's capacity by minimizing its implicit rank. In addition, Jing, Zbontar, et al., 2020 demonstrated that adding linear layers between the encoder and the decoder could penalize the rank of the latent variable and has an implicit regularizing effect. Reducing the rank of the latent variable means removing the redundant rows or columns, which leads to obtaining independent features in the latent space. These independent features help the decoder network of the autoencoder to generate the most expected future pose during training.

This paper addresses the domain shift issues by augmenting the discriminative model with a generative model to obtain better OoD prediction results. Specifically, we propose regularizing an extended network by inserting linear layers to minimize the rank of the latent space and train the entire end-to-end network. We regularize the network to strengthen the model to deal effectively with domain shift scenarios. In this study, we demonstrate that adding linear matrices with a generative model, which combines low in-distribution prediction error with maximum generalizability, can harden modern discriminative designs against significant distributional shifts. The added linear layers are 'absorbed' during the Inference process and thus did not cause any unwanted variation in the results. The test dataset comes from a distinct distribution compared to the training dataset. Despite this domain shift, the proposed model achieves better results on OoD actions than the current techniques. The following are the paper's main contributions:

- An out-of-distribution learning scenario is approached for a time-series problem to perform human pose prediction using an augmented graph convolution network. Using a residual graph-based encoder-decoder network, the network efficiently predicts future poses.

- The latent space size of the variational autoencoder network is minimized by removing the redundant information by introducing extra linear layers in the encoder model to minimize the rank solutions of the covariance matrix representation.

- We have shown superior results of our approach with other cutting-edge techniques for OoD actions on two standard datasets: Human3.6 M Ionescu et al., 2014 and CMU MoCap *CMU Graphics Lab Motion Capture Database* n.d.

This paper is organised as follows for the remainder. Section 2 discusses the related work. Section 3 presents our methodology and training details and discusses our architecture. Section 4 provides the experimental details, datasets, and results. The study is concluded in Section 5 along with suggestions for additional research.

4.2 Related work

Recent propositions have been using graph-based neural networks to encode the set of joints, as shown by Li et al., 2020c. This method used a multiscale graph neural network, which extracted features on different scales based on the graphs and then used these features across the different scales. It has an encoder-decoder architecture. Mao et al., 2019 showed the effectiveness of encoding spatial dependencies between the joints, which relieved them from manually specifying the range of temporal frequencies based on a deep feedforward network. Butepage et al., 2017 proposed another deep feedforward network whose representation could be used for classification and prediction problems. These feedforward networks are innately faster than recurrent neural networks (RNNs) Fragkiadaki et al., 2015; Li et al., 2020c while also being easier to train. Bourached et al., 2020 showed the effectiveness of augmenting the previous state-of-the-art graph-based approaches with feedforward networks as proposed by Mao et al., 2019, and the motion attention span extension as demonstrated by Mao, Liu, and Salzmann, 2020, in facilitating the robustness of the model when faced with the OoD samples. This can be particularly applicable to safety-specific situations where some human motion seen by the model immediately may not have been seen in the training data. Earlier methods Li et al., 2020c; Mao et al., 2019; Yu et al., 2022 have been discussed only in distribution scenarios, whereas real-world applications involve domain shift scenarios. Our focus is mainly on the OoD scenario, whereas we are also concerned about the ID scenario.

4.2.1 Autoencoder-based models

Autoencoders are a fundamental aspect of representation learning Zhang et al., 2020; Aldhubri et al., 2021. Variational autoencoders Lopez et al., 2020; Zietlow, Rolinek, and Martius, 2021 have an intrinsic probabilistic nature and thus generate blurry images. Since the data are contained in a lower-dimensional space, it is critical to explicitly control the latent capacity Chen et al., 2020a. Between the encoder and decoder, the technique Jing, Zbontar, et al., 2020 suggested introducing linear layers, which led to minimum rank solutions, thus achieving better, and in some cases, on par performance than other autoencoders such as naïve autoencoders and VAEs. The proposed approach also employs linear transformations to the incoming data from the encoder and passes these linear transformations' output to the decoder. This method proved simple, deterministic, and learned a compact latent space. It should be noted that the more linear layers added, the better the regularization effect is. An added benefit is the random initialization of the layers instead of manually initializing the square matrices.

4.2.2 Generative models

Some studies such as Liang, Li, and Srikant, 2018; Hendrycks, Mazeika, and Dietterich, 2018 demonstrated how deep generative networks could be used comprehensively for detecting OoD inputs. They are also shown to generalize competently. In Gustafsson, Danelljan, and Schon, 2020, Gustafsson et al. explained how recent propositions of deep generative models have been several issues with using density estimations to discover essential OoD, but are still good enough when the application requires the detection of anomalies. The study in Lee et al., 2018 listed the major provocations to AI safety, and one of them is robustness to distributional shift. Bourached et al. Bourached et al., 2020 proposed the evaluation of OoD results of the deep generative model on only one action, derived from one action distribution, which was accessible for hyperparameter search and training. The other classes are supposed to be used for carrying out testing.

4.2.3 Variational autoencoder parameters

A specific type of neural network called a variational autoencoder (VAE) can be trained to understand how to represent complex input in a lower-dimensional space. It has an encoder, a decoder, and a loss function that help the latent representation be useful and effective at precisely capturing the data's basic structure Kingma and Welling, 2013. In a VAE, the encoder maps the input data to a lower-dimensional latent space, where each point represents a possible variation of the input data. The decoder then reconstructs the input by mapping the latent representation back to the initial data space Doersch, 2016. To make the training of the VAE tractable, we introduce two parameters, namely the mean μ and the covariance σ, which represent the Gaussian distribution parameters that we believe the latent variables to follow. In order to create samples from the latent space during the decoding phase, the mean and covariance are acquired through the training process Liu et al., 2020. The

μ indicates the distribution's predicted value, which is the centre of the distribution, while the σ represents the spread of the distribution around the mean. Together, they define the shape of the distribution. During the training of the VAE, we introduce a noise parameter ϵ, which is an insignificant random number drawn from a typical Gaussian distribution. Kingma, Welling, et al., 2019. This noise is added to the mean and covariance parameters to ensure that the VAE is not overfitting to the data and can generate new samples that are not identical to the training data. The latent variables' Gaussian distribution is defined by the mean and covariance parameters learned during training. The epsilon parameter is introduced during the training and adds noise to the mean and covariance parameters to prevent overfitting.

4.2.4 Generalization with gradient descent

Implicit regularization, a beneficial facet of gradient descent optimization, is vital when deep neural networks require generalization. Saxe, McClelland, and Ganguli, 2019 demonstrated how continuous gradient descent could lead to a low-rank solution. Gunasekar et al., 2018 successfully showed how gradient descent optimization leads to a minimal nuclear norm solution, which was broadened to deep linear networks. Soudry et al., 2018 explained how gradient descent-based implicit bias is used to minimize exponential tail-based monotone loss functions, and this type of loss is the same as deep learning loss functions. Gidel, Bach, and Lacoste-Julien, 2019 explained how to optimize parameterized models in linear neural networks with the help of discrete gradient dynamics. These approaches are employed to enhance the generalization ability of deep neural networks. The proposed approach also employs implicit regularization techniques to improve the generalization capability of the network while learning robust representations.

4.2.5 Baseline methods

In this study, we evaluated the proposed approach and other compared methods regarding mean angle error and mean per joint position error (MPJPE) Ionescu et al., 2013. We compared our method with two well-known baselines that achieved good results with the domain shift scenario. The first baseline is Bourached et al., 2020, in which the well-known GCN-based encoder-decoder was used with a variational autoencoder. The second baseline Mao et al., 2019, in which GCN was used to encode and decode the discrete cosine transform (DCT) based input data. We did not compare our result with Li et al., 2020c; Yadav and Nandi, 2020 because they presented their results only for the ID scenario.

4.3 Methodology

Learning interdependently between various human skeleton joints is a tedious task. Human motion depends on the joint's relation, and human motion incorporates the motion of the joints. This work aims to improve the networks' prediction ability to understand the interdependence between joints of the human skeleton in the OoD scenario. We improve the capability of our deep network by adding multiple linear matrices, which helps to understand the interdependency between joints by improving the capability of the latent variable. Adding extra linear matrices between the encoder and the decoder penalizes the rank of the latent variable by reducing the number of dimensions used by the latent variables. The more linear matrices added, the better the regularization effect produced. The output of the encoder is fed to the linear matrices, and its output is provided to the decoder. This can result in an implicit regularization effect.

4.3.1 Problem description

In the literature on human motion, three-dimensional joint coordinates and angles are the two most common ways to represent a human position. These two methods are completely static. In this work, we use joint angle representation. As other methods discuss, The input of our model is an observed continuous motion sequence $X_{1:N} = [X_1, X_2, ...X_i, .., X_N]$ consisting N continuous observed pose frames, and the output of the model is T continuous frames containing the predicted poses. Input and output frames have a single scale. In other words, motion sequence $X_{1:N}$ is a matrix consisting of N rows and K columns. Each pose $X_i \in R^K$ has K joints to represent the pose. Now the task is to build a model $f_{predict}(.)$ to predict the unknown future pose sequence $\hat{X}_{N+1:N+T}$ given $X_{1:N}$ that approximates the ground truth $X_{N+1:N+T}$. The column of $X_{1:N}$ represents the human pose each time step, whereas each row of $X_{1:N}$ shows the motion of each joint. Here $X_N^k = [X_1^k, X_2^k, ...X_N^k]$ shows the trajectory of the k^{th} joint across the N frames. In our study, N is set to 10, and T varies from 2 to 25.

As stated in the literature, given a time horizon of 40ms, the short-term predictions correspond to 40, 80, 160, 320, and 400ms (2, 4, 6, 8, and 10 frames), while long-term predictions correspond to 1000ms (25 frames). In our implementation, we set the number of input frames (NIF) and the number of output frames (NOF) that determines the type of prediction (short or long). With the CMU MoCap dataset, our model has been trained to predict long-term poses (25 frames). For short-term predictions, we set the NOF of the trained model to 2, 4, 6, 8, or 10 frames to make the short-term predictions 40, 80, 160, 320, or 400ms.

We use temporal encoding, bypassing the input frame to a discrete cosine transform function. The goal of temporal encoding is to capture each joint's motion pattern. It provides compact representation by discarding the high frequencies that help capture the smoothness of human motion. In the proposed model, we use a linear layer-enabled discriminative model incorporated with a generative model for training the network by minimizing the combined loss calculated loss functions for both network branches. In the testing phase, we use only the discriminator network is shown in Figure 4.4 to predict the expected future pose $\hat{X}_{N+1:N+T}$.

We apply discrete cosine transformations (DCT) on each input to the network, along with the freedom to eliminate low or high frequencies. Every single coefficient contains information at a distinct temporal frequency of the whole sequence. We define a vector $x_k = [x_{k,1}, x_{k,2}, .., .., x_{k,N}]$, which is a vector of a specific joint k, taken over N time steps. This is converted to a DCT vector defined as $C_k = [C_{k,1}, C_{k,2}, .., .., C_{k,N}]$, where $C_{k,l}$ denotes the l^{th} DCT coefficient, which can be computed as (4.1):

$$C_{k,l} = \sqrt{\frac{2}{N}} \sum_{n=1}^{N} x_{k,n} \frac{1}{\sqrt{1+\delta_{l1}}} \cos(\frac{\pi}{2N}(2n-1)(l-1)). \tag{4.1}$$

This is invertible on the condition that none of the frequencies are cropped. Invertibility is achieved by employing inverse discrete cosine transform (IDCT), which can be expressed as (4.2):

$$x_{k,l} = \sqrt{\frac{2}{N}} \sum_{l=1}^{N} C_{k,l} \frac{1}{\sqrt{1+\delta_{l1}}} \cos(\frac{\pi}{2N}(2n-1)(l-1)). \tag{4.2}$$

The IDCT is applied to the outputs for conversion to the original coordinate system. This effectively allows comparison with the ground truths.

4.3.2 Graph convolutional networks

The early version of neural networks works effectively with Euclidean data. However, many datasets, such as social networks, and biological, chemical, etc., are irregular, non-Euclidean data. The graph uses a node to represent these inconsistent data, and mathematical operations can be implemented on these data types. All the mathematical operations are applied to the GCN following the previous CNN equation with some modification Lian et al., 2022. The forward propagation of GCNs at each layer is represented by (4.3). The term GCL($A^{[l-1]}$) represents the function of feature representation. σ is the activation function. We presume that $C \in R^{K\times(N+T)}$ is elucidated on a graph with N+T dimensions and k nodes. $A^{[0]} = C \in R^{K\times(N+T)}$, where C is the feature matrix at the first layer, which can also be written as $A^{[0]}$, it is also known as input features or node features. $W \in R^{n^{[l-1]}\times n^{[l]}}$ are the weights and $b \in R^{n^{[l]}}$ are the bias parameters. The count of hidden units in each layer l is represented by $n^{[l]}$. The normalized adjacency matrix S, the weights W, and the biases b are all learnable parameters. $A^{[l-1]}$ shows the feature representation matrix at layer l-1.

$$GCL(A^{[l-1]}) = \sigma(SA^{[l-1]}W + b). \tag{4.3}$$

To ensure that the model converges, we must normalize the features to avoid numerical instability and vanishing/exploding gradients. S is the normalized adjacency matrix calculated using (4.4).

$$S = D^{-\frac{1}{2}} * s * D^{-\frac{1}{2}}. \tag{4.4}$$

Moreover, $S \in R^{K\times K}$ is a normalized graph Laplacian that constitutes the connection between the different joints and is layer specific. The adjacency matrix s, in contrast, represents the connections or edges between the nodes in forward propagation. Moreover, D is the diagonal node degree matrix, which normalizes the adjacency matrix s.

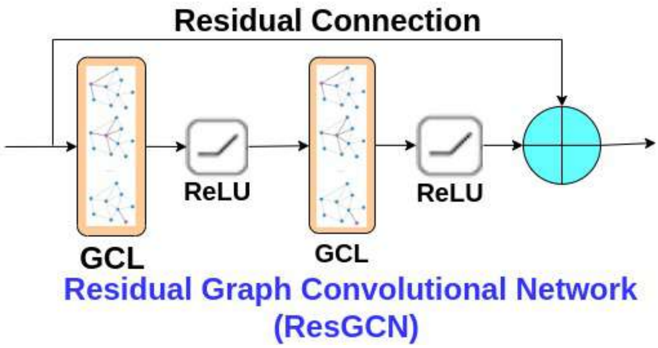


Figure 4.1: Complete figure of residual graph convolutional network block

4.3.3 Residual graph convolutional network block

Convolutional graph layers make up the residual graph network (ResGCN). with residual connections. By doing this, we are trying to enhance the capability of simple GCN. The graphical depiction of the ResGCN is in Figure 4.1. First, we use this ResGCN in the encoder of our network, and one after another, we use six ResGCNs in the encoder network. Similarly, we use six ResGCNs for both branches of the decoder network, followed by a single graph convolutional layer. Each ResGCN works as follows:

$$ResGCN(A) = A + S(SAW)W, \tag{4.5}$$

where $A \in R^{K \times N}$ is the previous layer's activation passing to the ResGCN block. $S \in R^{K \times K}$ is a normalized adjacency matrix, and $W \in R^{N \times N}$ is a weight matrix. The output of this block passes through a nonlinear activation function. In this instance, we use the ReLU activation function. ReLU function production is the result of the ResGCN block. We use six such blocks for both the encoder and decoder.

4.3.4 Implicit regularization using linear matrices

We use generative and discriminative models to learn the OoD scenario. Earlier research works Li et al., 2020c; Yadav and Nandi, 2020; Fragkiadaki et al., 2015 used a discriminative model for the human motion prediction task, and they discussed only the ID scenario. For the first time, Bourached et al., 2020 proposed augmentation of the generative model with a discriminative model to deal with the OoD scenario. Bourached et al., 2020 uses the variational autoencoder as a generative model to deal with the OoD cases. Decreasing or restricting the latent representation's information capacity is crucial to autoencoder approaches. Jing, Zbontar, et al., 2020 proposed an approach to minimize the rank of latent variables by adding linear matrices between the encoder and decoder of the autoencoder. To obtain motivation, we use this concept in our case, where our task to predict the future poses an OoD scenario. In our model, we use an autoencoder and discriminative branches. By adding the linear matrices, we improve the learning representation capability of the latent variable of the generative branch.

Implicit regularization is a phenomenon whereby gradient descent optimization of a linear multilayer neural network results in a low-rank solution. Our goal is to create an architecture that benefits from this phenomenon to enhance the quality of learned representations in predicting human motion. Learning good representations between various human joints remains a core issue in human motion prediction tasks. To address this scenario in the proposed deep network architecture, we add extra linear matrices W_1, W_2,.,W_j..W_l between the encoder and decoder networks. We used eight linear matrices. The dimensions of each matrix are $W_j \in R^{d \times d}$, and they are initialized randomly. We train all W_j matrices jointly with both branches. Figure 4.2 shows the encoder part of our network. Linear matrices are added after residual GCN blocks. These matrices successfully minimize the rank of the covariance matrix of the latent space while training latent variables by encouraging

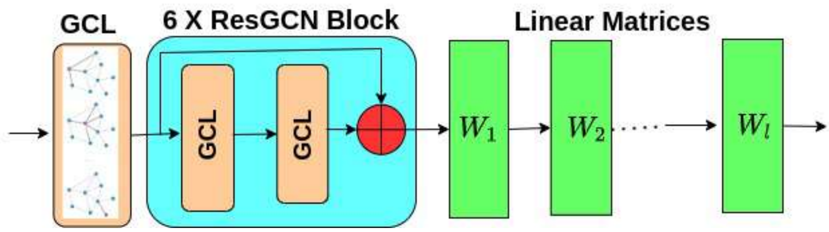


Figure 4.2: Figure shows the encoder part of our deep augmented network

them to use lower dimensions. Moreover, by accomplishing this, they intensify the regularization effect of our deep network. The regularized network learns interdependently between the various joints of the 3D skeleton of a human, and this capability of the network helps to predict more accurate motion during inference.

The inference network used for testing purpose only is shown in Figure 4.4. The predicted pose is the output of the decoder network with a residual connection and is represented by the following equation. Equation (4.6) mathematically expresses the Figure 4.4.

$$\hat{C}_{k,n} = Dec(W_l....W_2W_1Enc(C_{k,n})) + C_{k,n}. \tag{4.6}$$

Because linear matrix multiplication is associative, all W_j matrices can be absorbed into the encoder during inference. As a result, this linearly adjusted decoder can be used directly to generate the future pose.

4.3.5 Network architecture

Figure 4.3 shows the details of the train network architecture of our network. There exists one GCL at the start of the network. This GCL is followed by the first 6 Graph Convolution Blocks (GCBs). Each GCB consists of two graph convolution layers with residual connections. For each layer l, the number of hidden units, $n^{[l]} = 256$. The first 6 GCBs and linear layers constitute the encoder part of our network architecture. These blocks are defined with a latent variable. The latent variable can be represented as (4.7):

$$z \in \mathbb{R}^{k \times n_z} = N(\mu_z, \sigma_z), \tag{4.7}$$

where $\mu_z \in \mathbb{R}^{k \times n_z}$, $\sigma_z \in \mathbb{R}^{k \times n_z}$. n_z is 8 or 32 according to the stability of the training process. Equation (4.8) defines the KL divergence between a latent space distribution and the spherical Gaussian $N(0, I)$ distribution.

$$l_l = KL(q(z|c) \| q(z)) = \frac{1}{2}\sum_{1}^{n_z}(\mu_z{}^2 + \sigma_z{}^2 - 1 - \log(\sigma_z{}^2). \tag{4.8}$$

The output of these six GCBs is passed through the linear layers. The number of linear layers is set as 8. The linear layers perform the task of applying a linear transformation of whatever input goes into them. The learnable weights of these layers are of the shape $(out_{features}, in_{features})$ and are taken from a uniform distribution $\mu(-\sqrt{k}, \sqrt{k})$, where

$$k = \frac{1}{in_{features}^2},$$

where $in_{features}$ is the size of each input sample, and $out_{features}$ is the size of each output sample. The linear transformation applied to the incoming data is $y = w^T * x$. Our network's encoder design is depicted in Figure 4.2. The output of the linear transformations is fed to the decoder part of the VAE that has the same composition as the discriminative branch. It has 6 GCBs. The output neurons are parameterized as $\mu_z \in \mathbb{R}^{K \times (N+T)}$, and $log(\sigma_2) \in \mathbb{R}^{K \times (N+T)}$. The second component of VAE has a negative variational lower bound (VLB) as shown in Equation (4.12). It reconstructs the inputs as samples from a maximum likelihood Gaussian distribution.

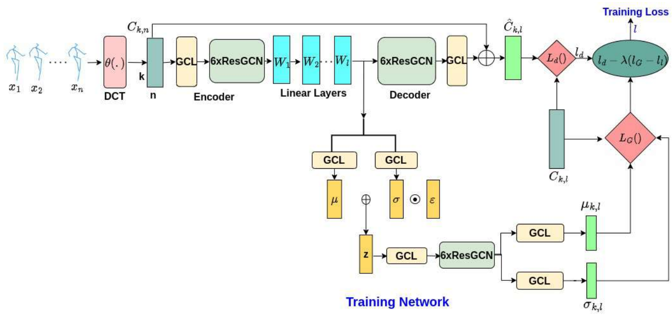


Figure 4.3: The proposed deep network architecture consists of GCN, linear matrices, and a variational autoencoder branch.$C_{k,n}$ is the input DCT coefficient matrix. $\hat{C}_{k,l}$ is the DCT coefficient predicted by the discriminative decoder branch. $C_{k,l}$ is the DCT coefficient of the ground truth corresponding to the output. $L_d()$ is the loss function for discriminative branch. $L_G()$ is the loss function calculated by the VAE branch. This architecture encoder consists of one graph convolutional layer (GCL), six residual graph convolutional network (ResGCN) blocks, and eight linear matrices. For training, the network consists of two branches. The first discriminative branch consists of six ResGCN blocks followed by GCL and a residual connection. Second, the variational autoencoder (VAE) branch uses two GCL layers to make a code variable. The code variable passes through a GCL followed by six ResGCN blocks and two GCL blocks separately to construct $\mu_{k,l}$ and σ_{kl}. Then, we calculate the loss for training using the loss function discussed in the methodology section. We use the combined loss l for the training process.

4.3.6 Training

The complete network is trained in conjunction with the negative VLB. Our loss function is defined by (4.9).

$$l = \frac{1}{K(N+T)} \sum_{n=1}^{N+T} \sum_{k=1}^{K} |\hat{x}_{k,n} - x_{k,n}| - \lambda(l_c - l_l). \tag{4.9}$$

In each layer, the graph size shows that the count of the parameters differs with K, which is the number of joints. The values of K are 48 for H3.6 M, 64 for CMU joint angles, and K = 65 for CMU Cartesian coordinates. λ is a hyperparameter, and our entire model has approximately 4.2 M parameters.

The average $l1$ distance between the ground-truth joint angles and the projected joint angles is given the joint angle loss, which can be defined as (4.10):

$$l_a = \frac{1}{K(N+T)} \sum_{n=1}^{N+T} \sum_{k=1}^{K} |\hat{x}_{k,n} - x_{k,n}|, \tag{4.10}$$

where $\hat{x}_{k,n}$ is the predicted k^{th} joint, and $x_{k,n}$ is the analogous ground truth, both at time-step n. As has already been proposed by Ionescu et al., 2013 and implemented in Mao et al., 2019; Mao, Liu, and Salzmann, 2020, the MPJPE has been used for the training of joint angle loss on 3D joint coordinate prediction. For each training sample, the MPJPE has been established as (4.11):

$$l_m = \frac{1}{K(N+T)} \sum_{n=1}^{N+T} \sum_{k=1}^{K} |\hat{P}_{k,n} - P_{k,n}|^2, \tag{4.11}$$

where K is the count of the skeletal joints, $\hat{P}_{k,n} \in \mathbb{R}^3$ is the predicted k^{th} joint position, and $P_{k,n}$ is the analogous ground truth, both at time-step n.

Out of these different loss functions, our final loss function to train

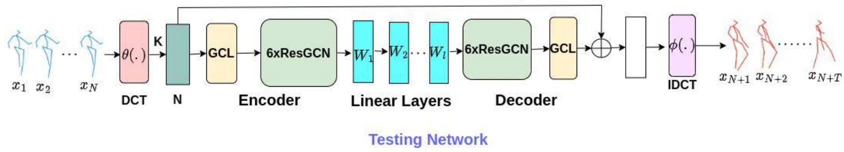


Figure 4.4: Overview of the linear transformation incorporated discriminator network to predict frame x_{N+1} to x_{N+T} by taking x_1 to x_N. Each frame passes through the DCT function. Here, K represents the number of joints in the 3D human skeleton. N shows the number of input frames. Graph convolutional layer in the encoder section used for upsampling. The graph convolution block is used for convoluting the features. Linear layers are used to transform the features. GCL in the decoder network is used for downsampling.

our network architecture is represented as l. l integrates the loss of both branches of the network. l_a shows the loss of the discriminator branch, where l_c and l_l are losses of the VAE branch. We use l_m (MPJPE) for test cases to compare our results with the other approach results. We also compare our results in the Euclidean distance between anticipated and ground truth joint angles.

$$l_c = log(p(c|z)) = -\frac{1}{2}\sum_{n=1}^{N+T}\sum_{l=1}^{l}(log(\sigma_{kl}^2) + log(2\pi) + \frac{|c_{k,l} - \mu_{k,l}|^2}{e^{log(\sigma_{k,l}^2)}}), \tag{4.12}$$

where $c_{k,l}$ are the ground truth DCT coefficients.

4.4 Experimental results and discussion

4.4.1 Datasets

The Human3.6 M and CMU motion capture datasets are used to evaluate the proposed method.

4.4.1.1 Human3.6 M

This dataset includes 3.6 million 3D human poses and their comparable photos, all performed by seven different subjects. Every subject features 15 actions, including 'walking', 'smoking', 'discussion', 'eating', 'greeting', 'directions', 'phoning', 'posing', 'purchases', 'sitting down', 'sitting', 'waiting', 'walking with dog', 'taking photo', and 'walking together'. For training and cross-validation, we employ the simple act of 'walking' and test on all other sets of actions. The model was trained for six actors, with five being used for training and one for cross-validation. As with the previous literature, Martinez, Black, and Romero, 2017; Li et al., 2020c; Yadav and Nandi, 2020 all have followed this same training method, with the metrics being evaluated on Subject 5, the remaining one. This technique ensures the selection of all the parameters based on ID performance.

4.4.1.2 CMU motion capture

This dataset contains a collection of complex motions enacted by 144 subjects. As reported in Mao et al., 2019; Li et al., 2020b, the rest of the previous literature, we make use of 8 detailed actions, namely 'directing traffic,' 'basketball,' 'jumping,' 'basketball signal,' 'walking,' 'soccer,' 'running,' and 'washing the window.' No hyperparameters are tuned for this dataset, and as with the previous literature, the identical data split into training and testing sets is used. One representation bestows an exponential map representation of the joint angle, a vector with 64 dimensions. The other representation gives the 3D Cartesian coordinates of 25 joints and is a vector with 75 dimensions.

4.4.2 Experimental settings

Myronenko et al. Myronenko, 2018 arbitrarily set the hyperparameter λ to 0.1. Selecting the optimum value for the weighting parameter between the discriminative and the generative models is of utmost importance since this weighting relates to other regularization parameters. For impartial comparison, the hyperparameter is conducted on the GCN model for dropout probability drops ranging from 0.1 to 0.9. and the generative GCN model. Although the search for GCN is done only for short-term (up to 400 ms) predictions for the H3.6 M dataset, this is used for future evaluations, thus depicting the model's generalizability. To search, a random hyperparameter search for our model, along with the generative GCN model, is carried out for drop, which is the dropout probability, with λ in the ranges drop [10, 0.00001] on a logarithmic scale, and λ = [0, 0.5] on a linear scale.

Congruous with the preceding literature, we show the resulting short-term predictions (up to 400 ms) and long-term (up to 1000 ms). To compare our model with the GCN and the generative GCN, we consider a short history, ideally ten frames (up to 400 ms) for predicting short-term and 25 frames (up to 1000 ms) for long-term human motion. We use the Euler angle representation for the training of the network. We calculate the loss using the Euclidean distance between the predicted and ground truth joint angles. The second loss function, MPJPE is used to calculate the loss between the predicted and ground-truth 3D joint positions. Both baseline methods Mao et al., 2019, Bourached et al., 2020 use graph convolutional network (GCN) with DCT inputs. In particular, Mao et al., 2019 used only GCN in both encoder-decoder sections of the network, whereas Bourached et al., 2020 used a generative branch along with GCN to train the network. In the OoD situation, we compare our results for both baseline techniques.

The suggested model is implemented in PyTorch Paszke et al., 2017 and trained using a training batch size of 16 for 50 iterations. We clipped off the gradients to a maximum l2-norm of 1 with a restriction between -20 and 3. The learning rate was set to 0.0005 without any decay since the dynamic correlation between the generative and discriminative loss renders a decaying redundant. We used ADAM Zhang, 2018 as an optimizer.

Table 4.1: Analysis of the performance of the suggested approach with and without DCT of all activities using the H3.6 M dataset for short-term prediction in terms of the Euclidean distance.

Motion	**Walking (ID)**	**Eating (OoD)**	**Smoking (OoD)**	**Discussion (OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
Without DCT	0.23 0.40 0.57 0.68	0.20 0.36 0.59 0.72	0.25 0.46 **0.87 0.85**	0.29 **0.60 0.91** 1.00
With DCT	**0.22 0.39 0.55 0.62**	**0.19 0.33 0.54 0.67**	**0.24 0.45** 0.89 **0.85**	**0.28** 0.61 **0.91 0.98**
Motion	**Directions (OoD)**	**Greeting (OoD)**	**Phoning (OoD)**	**Posing (OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
Without DCT	**0.33** 0.52 0.76 0.87	0.44 0.75 **1.13** 1.31	0.57 1.09 1.43 1.57	0.33 0.67 1.38 1.49
With DCT	**0.33 0.51 0.72 0.80**	**0.43 0.73 1.13 1.30**	**0.56 1.08 1.42 1.55**	**0.31 0.66 1.35 1.49**
Motion	**Purchases(OoD)**	**Sitting(OoD)**	**Sitting Down(OoD)**	**Taking Photo(OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
Without DCT	0.55 0.80 1.16 1.22	0.36 0.53 0.90 **1.05**	**0.35 0.69** 1.00 1.11	0.21 0.47 0.83 0.87
With DCT	**0.53 0.77 1.08 1.15**	**0.35 0.52 0.89** 1.07	**0.35 0.69 0.98 1.09**	**0.20 0.45 0.72 0.85**
Motion	**Waiting(OoD)**	**Walking Dog(OoD)**	**Walking Together(OoD)**	**Average (of 14 for OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
Without DCT	0.28 0.57 1.06 1.30	0.51 0.89 1.19 1.36	**0.21 0.44** 0.62 0.68	0.36 0.63 0.98 1.10
With DCT	**0.27 0.56 1.02 1.25**	**0.51 0.88 1.15 1.30**	**0.21 0.44 0.61 0.67**	**0.34 0.62 0.96 1.07**

Table 4.2: Analysis of the performance of the suggested approach with and without DCT for all activities using the CMU motion capture dataset for short-term and long-term predictions in terms of the Euclidean distance.

Motion	**Basketball (ID)**	**Basketball Signal (OoD)**	**Directing Traffic (OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Without DCT	0.41 0.67 1.19 1.39 1.85	0.17 0.28 0.49 0.60 1.07	0.23 **0.42 0.68** 0.85 2.06
With DCT	**0.37 0.63 1.14 1.34 1.83**	**0.15 0.26 0.45 0.57 1.05**	**0.21 0.42** 0.70 **0.83 2.04**
Motion	**Jumping (OoD)**	**Running (OoD)**	**Soccer (OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Without DCT	0.36 0.63 1.50 1.71 1.91	0.44 0.72 0.84 0.88 1.46	0.22 0.37 0.74 0.87 1.65
With DCT	**0.34 0.56 1.39 1.57 1.79**	**0.42 0.70 0.83 0.84 1.42**	**0.21 0.35 0.73 0.86 1.62**
Motion	**Walking (OoD)**	**Washing Window (OoD)**	**Average (of 7 for OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
Without DCT	0.38 **0.49 0.60** 0.71 1.03	0.27 0.41 0.69 **0.87** 1.38	0.30 0.47 0.79 0.93 1.51
With DCT	**0.37** 0.51 0.62 **0.69 0.99**	**0.24 0.35 0.67** 0.88 **1.31**	**0.28 0.45 0.77 0.89 1.46**

4.4.3 Ablation study

Each subject of dataset H3.6 M has 32 joints, and we only use the joints with nonzero values after transforming the joint positions into exponential maps (20 joints remain). The dataset initially consists of sequential numeric data of approximately a thousand rows and ninety-nine columns. After preprocessing, the input to the proposed model is a matrix of 10×48 for the H3.6M dataset. Moreover, input fed to the proposed model is a matrix of 10×64 in the case of the CMU MoCap dataset, where 10 represents ten-time steps (10 frames) and 48 and 64 rows consisting of the number of the joint angle values of the two datasets, respectively. Moreover, we followed the standard preprocessing method, followed by every literature work Bourached et al., 2020, Mao et al., 2019, Li et al., 2020c. The output of our network is 10, 48 for H3.6 M and 10, 64 for CMU Mocap. Between the ten anticipated postures and the associated ground truth poses, we computed the findings as a Euclidean distance. We also calculated the mean per joint position error.

In this ablation study, we investigate the effect of DCT on our model. We show the results with and without DCT for both datasets in Tables 4.1 and 4.2. Table 4.1 shows the results for the H3.6 M dataset, and Table 4.2 shows the result for the CMU MoCap dataset; the DCT-based data show minor error for each time step compared to the data without DCT. However, one can directly use the input-output trajectory without DCT, as suggested by other works. The results for both datasets with DCT show superior or equal results in most cases. In our representation and work in trajectory space, DCT encodes the temporal character of human motion. The goal of temporal encoding is to record each joint's motion pattern. The fundamental aim for this is because the DCT may offer a more compact representation by dropping the high frequencies, which nicely portrays the smoothness of human motion, especially in terms of 3D coordinates. Therefore, the DCT-based approach is recommended in this scenario of human motion.

4.4.4 Comparisons with existing methods

Tables 4.3, and 4.4 and Figure 4.5 present comparisons between the proposed method and two existing methods: Mao et al., 2019 and Bourached et al., 2020. The method of Mao et al., 2019 proposed a simple encoder-decoder-based discriminative model that contains GCN in the encode and decoder section to deal with the OoD scenario. The method of Bourached et al., 2020 proposed an augmented network that combines the discriminative network with the generative network to deal with the OoD scenario. It should be noted that Mao et al., 2019 and Bourached et al., 2020 are the only recent works that presented results for the OoD scenario as per our best knowledge.

Table 4.3 presents the short-term predictions in Euclidean distance between the predicted and ground truth joint angles. It is performed on the H3.6M dataset. In Table 4.3, the first component shows the results calculated in the Euclidean distance for walking activities, which is ID activities. In this case, the results demonstrate a significant improvement for the short-term period of 80 ms, 320 ms, and 400 ms; but in the case of 160 ms, the baseline method result is better than ours. Similarly, in the case of some OoD actions Posing and walking with the dog, the baseline methods loss is slightly less than our methods loss in some cases. Restoring all the other OoD actions for the short-term period (80 ms, 160 ms, 320 ms, and 400 ms), we report significantly improved loss compared with the other baseline methods. It depicts the role of implicit regularization of the latent variable. The last component of Table 4.3 shows the average results for each period (80 ms, 160 ms, 320 ms, and 400 ms) for fourteen out of the distribution activities. Our overall average results outperformed state-of-the-art results.

Table 4.4 presents the short and long-term predictions of the Euclidean distance between the expected joint angle and the ground truth

Table 4.3: Comparing the proposed method with two existing methods (Mao et al., 2019, and Bourached et al., 2020) for all activities in the H3.6 M dataset for short-term prediction in terms of the Euclidean distance.

Motion	**Walking(ID)**	**Eating(OoD)**	**Smoking(OoD)**	**Discussion(OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
GCN Mao et al., 2019	0.22 0.37 0.60 0.65	0.22 0.38 0.65 0.79	0.28 0.55 1.08 1.10	0.29 0.65 0.98 1.08
Generative GCN Bourached et al., 2020	0.23 **0.37** 0.59 0.64	0.21 0.37 0.59 0.72	0.28 0.54 1.01 0.99	0.31 0.65 0.97 1.07
Ours	**0.22** 0.39 **0.55 0.62**	**0.19 0.33 0.54 0.67**	**0.24 0.45 0.89 0.85**	**0.28 0.61 0.91 0.98**
Motion	**Directions (OoD)**	**Greeting (OoD)**	**Phoning (OoD)**	**Posing (OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
GCN Mao et al., 2019	0.38 0.59 0.82 0.92	0.48 0.81 1.25 1.44	0.58 1.12 1.52 1.61	**0.27 0.59** 1.26 1.53
Generative GCN Bourached et al., 2020	0.38 0.58 0.79 0.90	0.49 0.81 1.24 1.43	0.57 1.10 1.52 1.61	0.33 0.68 **1.25** 1.51
Ours	**0.33 0.51 0.72 0.80**	**0.43 0.73 1.13 1.30**	**0.56 1.08 1.42 1.55**	0.31 0.66 1.35 **1.49**
Motion	**Purchases(OoD)**	**Sitting(OoD)**	**Sitting Down(OoD)**	**Taking Photo(OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
GCN Mao et al., 2019	0.62 0.90 1.34 1.42	0.40 0.66 1.15 1.33	0.46 0.94 1.52 1.69	0.26 0.53 0.82 0.93
Generative GCN Bourached et al., 2020	0.62 0.89 1.23 1.31	0.39 0.63 1.05 1.20	0.40 0.79 1.19 1.33	0.26 0.52 0.81 0.95
Ours	**0.53 0.77 1.08 1.15**	**0.35 0.52 0.89 1.07**	**0.35 0.69 0.98 1.09**	**0.20 0.45 0.72 0.85**
Motion	**Waiting(OoD)**	**Walking Dog(OoD)**	**Walking Together(OoD)**	**Average (of 14 for OoD)**
milliseconds	80 160 320 400	80 160 320 400	80 160 320 400	80 160 320 400
GCN Mao et al., 2019	0.29 0.59 1.06 1.30	0.52 **0.86** 1.18 1.33	0.21 0.44 0.67 0.72	0.38 0.69 1.09 1.27
Generative GCN Bourached et al., 2020	0.29 0.58 1.06 1.29	0.52 0.88 1.17 1.34	0.21 0.44 0.66 0.74	0.38 0.68 1.07 1.21
Ours	**0.27 0.56 1.02 1.25**	**0.51** 0.88 **1.15 1.30**	**0.21 0.44 0.61 0.67**	**0.34 0.62 0.96 1.07**

Table 4.4: Comparing the proposed method with two existing methods (Mao et al., 2019, and Bourached et al., 2020) for all activities in the CMU motion capture dataset for short-term and long term prediction in terms of the Euclidean distance.

Motion	**Basketball (ID)**	**Basketball Signal (OoD)**	**Directing Traffic (OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
GCN Mao et al., 2019	0.40 0.67 **1.11 1.25 1.63**	0.27 0.55 1.14 1.42 2.18	0.31 0.62 1.05 1.24 2.49
Generative GCN Bourached et al., 2020	0.40 0.66 1.12 1.29 1.76	0.28 0.57 1.15 1.43 2.07	0.28 0.56 0.96 1.10 2.33
Ours	**0.37 0.63** 1.14 1.34 1.83	**0.15 0.26 0.45 0.57 1.05**	**0.21 0.42 0.70 0.83 2.04**
Motion	**Jumping (OoD)**	**Running (OoD)**	**Soccer (OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
GCN Mao et al., 2019	0.42 0.73 1.72 1.98 2.66	0.46 0.84 1.50 1.72 1.57	0.29 0.54 1.15 1.41 2.14
Generative GCN Bourached et al., 2020	0.38 0.72 1.74 2.03 2.70	0.46 0.81 1.36 1.53 2.09	0.28 0.53 1.07 1.27 1.99
Ours	**0.34 0.56 1.39 1.57 1.79**	**0.42 0.70 0.83 0.84 1.42**	**0.21 0.35 0.73 0.86 1.62**
Motion	**Walking (OoD)**	**Washing Window (OoD)**	**Average (of 7 for OoD)**
milliseconds	80 160 320 400 1000	80 160 320 400 1000	80 160 320 400 1000
GCN Mao et al., 2019	0.40 0.61 0.97 1.18 1.85	0.36 0.65 1.23 1.51 2.31	0.36 0.65 1.41 1.49 2.17
Generative GCN Bourached et al., 2020	0.38 0.54 0.82 0.99 1.27	0.35 0.63 1.20 1.51 2.26	0.34 0.62 1.35 1.41 2.10
Ours	**0.37 0.51 0.62 0.69 0.99**	**0.24 0.35 0.67 0.88 1.31**	**0.28 0.45 0.77 0.89 1.46**

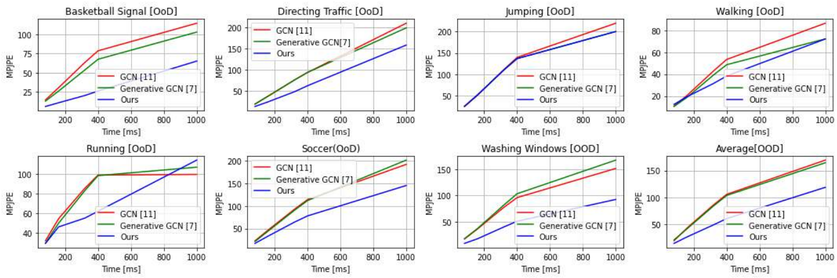


Figure 4.5: Comparing the proposed method with two existing methods (Mao et al., 2019, and Bourached et al., 2020) for all seven OoD activities of the CMU MoCap dataset for short (up to 400 ms) and long (up to 1000 ms) by considering MPJPE.

joint angles. It is performed on the CMU Mocap dataset. The first component of the table also shows the result for Basketball ID activity. For the short-term span of 80 ms and 160 ms, our loss report is less than both baseline approaches; later, for another period, 320 ms, 400 ms, 1000 ms short-term and extended the baseline approaches report lower loss than ours. In the OoD scenario, our method results are better than those of the baseline methods for all seven actions for a short time (80 ms, 160 ms, 320 ms, and 400 ms) and long term (1000 ms). The results show the importance of adding linear layers between the encoder and latent variable of the network. The last component shows the average results for each step, and we also offer promising results here.

Figure 4.5 depicts the short (up to 400 ms) and long (up to 1000 ms) term MPJPE for all seven OoD activities of the CMU MoCap dataset. We calculated MPJPE error just for comparing our result through baseline approaches, whereas we do not use this loss function in the network's training. In this, all components are for the OoD scenario. Blue lines represent our results, and different colours represent the other two baseline approaches. The figure reports that our loss in the case of jumping action is quite similar up to 400 ms for the short-term, but later in the long term (560 ms, 720 ms, and 1000 ms), we improved the loss in terms of

MPJPE error compared to the baseline. Similarly, our loss is similar to the baseline for walking MPJPE error for the initial short-term period of 80 ms and 160 ms. However, in the case of the long term, including 320 ms, and 400 ms, our results are significantly improved compared with both baseline approaches, and at 1000 ms, one of the baseline results is similar to our. In the case of running action, our results for 80 ms, 160 ms, 320 ms, 400 ms, 560 ms, and 720 ms are better than the baseline approach, but for the long run, 1000 ms is one of the baselines that shows better results than our result. Whereas for the rest of the actions (Basketball, Directing traffic, soccer, and washing windows), our MPJPE error is improved compared to the short-term and long-term baseline approach, including the average scenario. During this work, we focused primarily on short-term prediction for both datasets. We have shown long-term results in the case of CMU MoCap, but in the case of H3.6 M, only short-term results in the Euclidean distance. Therefore, this work has not demonstrated the long-term effects and MPJPE error results for H3.6 M data.

4.4.5 Complexity analysis

The additional linear layers reduce the training's time complexity. Training the proposed model without a linear layer takes approximately five hours using an NVIDIA GeForce RTX 3060. It should be noted that adding linear layers to the model reduces the training time to approximately four hours. Table 4.5 compares our model's training time per pass (forward and backward pass) with state-of-the-art methods (Martinez, Black, and Romero, 2017, Lebailly et al., 2020 and Mao et al., 2019) for 16 batch sizes. All methods are trained on a single GPU, an NVIDIA GeForce RTX 3060, an 8 GB GPU RAM system, CUDA 11.3, and PyTorch 1.81. During the training process, our network's forward and backward pass takes 30ms, which is shorter than Martinez, Black, and Romero, 2017 and Lebailly et al., 2020 and equal to Mao et al., 2019. It should be mentioned that the proposed model has 4.22 million parameters.

Table 4.5: Contrasting the suggested method's complexity with that of competing methods.

Method	Residual sup Martinez, Black, and Romero, 2017	TIM Lebailly et al., 2020	GCN Mao et al., 2019	Proposed
Train Batch Size	16	16	16	16
Time/Pass	75ms	75ms	30ms	30ms

4.5 Conclusion and future work

This paper focused on the robustness of distributional shifts in human motion prediction. We presented a hybrid framework and evaluated its performance for the most extensive publicly available datasets, H3.6 M and CMU MoCap. The previous baseline models have shown promising results with graph convolutional networks, attention-based GCNs, and even multiscale graph neural networks. Augmenting these with generative models improves the robustness of the network. However, since these frameworks follow an encoder-decoder architecture, we delved deeper into state-of-the-art autoencoder architectures and methods. We have taken inspiration from the autoencoder literature and inserted linear layers that take the encoder's output and feed its output to the decoder as input. This causes the implicit regularization effect. Combined with the augmentation of previous state-of-the-art human motion prediction architectures, these layers can provide better performance and enhance robustness to out-of-distribution actions without sacrificing in-distribution performance. Future work will focus on adjusting the proposed approach to handle the uncertainty while making long-term predictions in OoD and ID scenarios with different datasets.

Part III

Improving the Quality of Life of Intellectually Disabled Elderly People

Chapter 5

Effective ML-Based Quality of Life Prediction Approach for Dependent People in Guardianship Entities

Summary

This paper proposes an effective approach for predicting quality of life (QoL) for dependent individuals in guardianship entities. In addition, it aims to improve the QoL of people with intellectual disabilities. The proposed QoL prediction approach employs machine learning (ML) techniques to model the relationship between eight aspects of QoL and the corresponding QoL index. It determines whether or not a person needs assistance based on the index value. The proposed approach determines the priority of care (PoC) value for each aspect of a person. Based on PoC, the deficit aspect is determined, followed by the type of assistance a person requires, based on the decision priorities. It also generates a

support report with suggested actions to highlight the level in that aspect. In addition, we train multiple ML models to predict the standard score (SS), which represents the support value related to the eight aspects of QoL. Finally, we use SS values to train an ML model to predict the support intensity scale (SIS). On a dataset compiled from guardianship entities, the proposed approach is validated. The QoL index, SS, and SIS prediction models achieve average R^2 values of 0.9897, 0.9998, and 0.9977 with a standard deviation of 0.0051, 0.0002, and 0.0007, respectively.

5.1 Introduction

Due to the increase in cases, QoL assessment of a dependent individual, particularly one with an intellectual disability (ID), attracts significant research interest. ID is typically defined as intellectual and adaptive functioning deficits determined by standardized testing (e.g., an IQ score of less than 70). It demonstrates a person's inability to perform socially expected functions, responsibilities, and tasks. The disabilities appear during the developing period and cause daily limits that require continuous support. These deficiencies have an impact not only on autonomous functioning at home but also on involvement in the community, social, and academic activities Sandjojo et al., 2019. Individuals with intellectual impairments should be allowed to live as independently as possible, as stated in the United Nations Convention on the Rights of Persons with disabilities, which is something that many individuals with intellectual disabilities desire Haigh et al., 2013; Kuijken et al., 2016; Bond and Hurst, 2010. Improving their ability to manage their affairs independently could improve their QoL and engagement in the community Dollar et al., 2012. Individuals with ID can encounter numerous difficulties in their daily lives, but interventions and support can alleviate them Schalock et al., 2018. Typically, family members and professionals feel overburdened when supporting an individual with

IDs Dawson et al., 2016; Hermsen et al., 2014; Vilaseca et al., 2017. Consequently, interventions encouraging people with IDs to handle their affairs are necessary.

Over the past decades, efforts have been made to reconcile the various aspects that directly influence a person's behaviour to achieve the desired outcome Gómez Sánchez et al., 2022. Living life based on one's preference leads to autonomy, which also seeks to provide for those dependent on others Verdugo et al., 2020. Recently, many support instruments have been developed to aid dependent individuals in living everyday life with their families. The recent development in intellectual development disorder (IDD) reveals methods to enhance the QoL using support paradigms. Various countries have recognized a different number of aspects to represent the QoL. For instance, specialists have adopted eight aspects in Spain to evaluate the QoL of people with ID Verdugo et al., 2010. These eight aspects are emotional well-being (EW), material well-being (MW), personal development (PD), self-determination (SD), physical well-being (PW), social inclusion (SI), interpersonal relations (IR), and rights (RI) Gómez Sánchez, Schalock, Verdugo Alonso, et al., 2021. The assessment of the QoL in the eight aspects covers the overall domain. Enhancing QoL means improving these eight areas of an individual with ID. These eight aspects encompass all the necessities of a dependent individual to live life equally as an average person. The support paradigm helps make a required support plan to improve the QoL aspects for an individual. Integrating the support with QoL generates a new paradigm QoL support model (QoLSM) Gómez, Schalock, and Verdugo, 2021, whereas integrating the support paradigm with QoL indexes enhances the lifestyle of the dependent individuals. For instance, Gómez et al. Gómez, Schalock, and Verdugo, 2021 demonstrate the way to evaluate the model effect on individuals and organizations and improve its performance. Therefore, Verdugo et al. Verdugo et al., 2020 reviewed the recent works in this field based on measurement tools, descriptive correlation studies, predictive studies, and interventions. These four criteria cover the current investigation of intellectual and developmental disabilities. Recent research focuses on various QoL-related factors, such as the use of technology Geppert et al., 2022; Desideri et al., 2021, prompting Walters et al., 2021, employment Garrels and Sigstad, 2021; Gjertsen, Hardonk, and Ineland, 2021, and health behaviour Sanchez, 2021, Morán et al., 2022. Due to the studies' narrow focus and methodological limitations, it is difficult to draw definitive conclusions about the efficacy of self-management interventions. Nevertheless, the majority of previous research has revealed positive results.

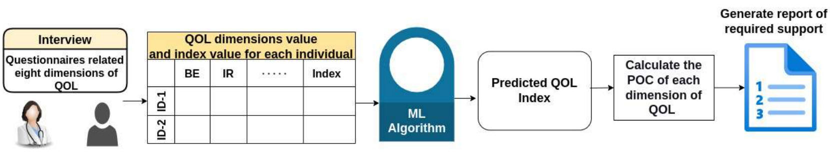


Figure 5.1: Schematic diagram of the proposed method for predicting dependent people QoL in guardianship entities.

Indeed, ML techniques can assist in analyzing the patient's report to predict the patient's potential need for support. This paper proposes a practical approach for predicting QoL for dependent people in guardianship entities. Our approach has numerous stages, from accepting input from eight aspects of QoL to automatically generating a report on necessary assistance without the involvement of psychologists, specialists, physicians, or other professionals. Specifically, we employ ML techniques to build QoL index, SS, and SIS prediction models. We also determine the PoC value for each aspect of a person. This study is based on the Newton-One private dataset of 26 subjects. Each subject contains eight aspects of QoL and the QoL index value, with eight to nine questions. A total of eight aspects represent the QoL. The proposed approach can contribute to understanding different aspects of QoL and fulfil the requirement of self-management to reduce reliance on family members and professionals. These are this paper's primary contributions, in order:

- Proposing an efficient ML approach for predicting dependent people QoL in guardianship entities.
- Presenting a new ML-based method for determining the PoC for each quality aspect to decide the QoL aspect that needs support.
- Proposing an ML-based method for predicting the SIS value to integrate this value with the other sensory information inputs to strengthen the QoL of dependent people.
- Analyzing the efficacy of different ML techniques to model the QoL to help professionals in decision-making and patients improve their QoL.

The following describes the paper's structure: Section 2 contains an overview of the literature on intellectual disability. The proposed technique and dataset specifics are illustrated in Section 3. The outcomes are displayed in section 4. The conclusion and future efforts are presented in section 5 lastly.

5.2 Related works

5.2.1 QoL's aspects and support paradigm

QoL refers to an individual's general well-being while meeting fundamental needs. Independently enjoying a high quality of life is difficult for individuals with IDs. People with ID who have IQs below 70 have difficulty performing personal and social, and behavioural activities effectively Sandjojo et al., 2019. Therefore, improving the QoL of these individuals is a fundamental challenge for the researchers. QoL encompasses three aspects of life— the personal, the social, and the judicial.

Improving QoL entails focusing on these three aspects of an individual's identity. These three areas are further subdivided into eight aspects of QoL González Fernández, Laborda Molla, and Jariot García, 2021; Verdugo et al., 2005, as shown in Table 5.1. These eight aspects satisfy all necessary substitutes for a person's QoL.

The type and intensity of support necessary to complete a specified task determine the support need Gómez Sánchez, Schalock, Verdugo Alonso, et al., 2021. The support paradigm's primary goal is comprehending people based on their basic assistance needs. Using this paradigm of support, organizations assist families of individuals with intellectual disabilities with inclusive education, independent living, supported employment, and other rights using this paradigm of support Thompson, Shogren, and Wehmeyer, 2016; Schalock, Luckasson, and Tassé, 2021b. The capability of the individual with ID to operate in their environment will be improved so that the assistance provided meets their needs, goals, and preferences. An individual's support plan calculates support needs, which are then aggregated as data. Utilization of combined data to improve organizational efficiency and research allocation Schalock, Luckasson, and Tassé, 2021b; Lombardi et al., 2016. Support needs assessment aims to develop an individualized and generalized support system to help people with ID and their families improve their QoL over time. Therefore, the best practice for intellectual impairment is to assess support needs. There are numerous ways to assist with these support requirements. Each QoL aspect is associated with support requirements in the SIS column of Table 5.1. The area indicates the required area where support is required, corresponding to each aspect. These support areas are protection and defence, health and healthcare, social activities, Behavioral support need, employment activities, home life activities, lifelong learning, exceptional medical needs, and community life activities.

Part –B Life activities in the community	frequency					Daily support time					Type of support					Direct Score
1. Moving from one place to another throughout the community (Transportation)	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
2. Participate in recreational and leisure activities in community settings	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
3. Access buildings and public environments	0	1	2	3		0	1	2	3	4	0	1	2	3	4	7
4. Go to visit friends and family	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
5. Participate in preferred community activities (parish, volunteering..)	0	1	2	3		0	1	2	3	4	0	1	2	3	4	9
6. Go shopping and buy groceries and services.	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	8
7. Interact with members of community	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	6
8. Access public buildings and environments	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	7

Score direct total Life activities in the community	64

Figure 5.2: A fragment of questionnaires used to collect the dataset. Each question has three parameters(daily support time, frequency, and type of support). The beneficiary must select the appropriate response from the scale provided. The brown colour number indicates the responses that have been selected, while the black colour alternatives indicate that there is no option to choose this number. The direct score is the sum of the question's three parameters, and the sum of all direct scores for all questions is the direct score total.

5.2.2 GENCAT scale

The GENCAT scale is a tool created by INICO. It provides an impartial assessment of the QoL of an individual Verdugo et al., 2013; Carrión-Nessi et al., 2022. In this assessment, the user must reply to the 69 objective questions using a frequency scale based on observation. Figure 5.2 presents a fragment of questionnaires used to collect the dataset. The GENCAT scale has been developed and validated using Schalock and Verdugo's multidimensional model. The scale generates valid and

Table 5.1: Details of aspects of QoL and support intensity scale metric with the maximum value of support, and also with several activities

QoL aspects	Support intensity scale metric			
	Area of SIS	Number of support activities	Maximum value of support	Area of support
Emotional well-being (EW)	Health and healthcare (S1E)	8	94	Personal area
	Protection and defense (S2)	8	94	
	Behavioral support need (S3B)	13	26	
Personal development (PD)	Home life activities (S1A)	8	92	
	Lifelong learning (S1C)	9	104	
Physical well-being (PW)	Health and healthcare (S1E)	8	94	
	Exceptional medical need (S3A)	16	32	
Self-determination (SD)	Protection and defense (S2)	8	94	
Interpersonal relation (IR)	Social activities (S1F)	8	93	Social area
Social inclusion (SI)	Community life activities (S1B)	8	91	
	Social activities (S1F)	8	93	
Material well-being (MW)	Employment activities (S1D)	8	87	Judicial area
Rights (RI)	Protection and defense($S2$)	8	94	
	Health and healthcare (S1E)	8	94	

trustworthy scores for each of the eight aspects of QoL and a global QoL index based on the multidimensional model.

5.3 Methodology

The relevance of this domain in society is demonstrated in the literature. Learning algorithms aid the psychologist in analyzing the patient's situation during professional interviews. Professionals administer a set of 69 questions covering all possible aspects of an intellectually disabled person's life. Professionals record the answers to the questions based on four points frequency scale. Professionals use a tool made by INICO known as the GENCAT scale Verdugo et al., 2013, which transforms 69 questions into eight aspect score values and corresponding index values. We interviewed 26 subjects and prepared our dataset using the GENCAT scale tool. Hence, the initial dataset contains 26 rows and 9 columns.

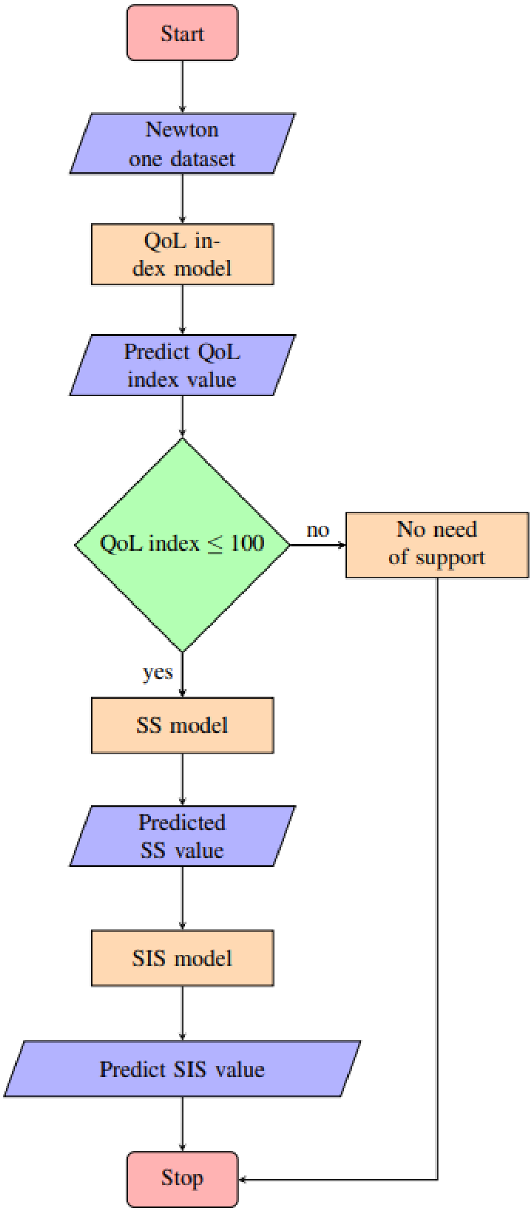


Figure 5.3: Flowchart of constructing the QoL index, SS, and SIS models.

In this study, we train three ML-based models: the QoL index, the SS, and the SIS. The flowchart for constructing the three models is depicted in Figure 5.3. First, we divide the dataset three-fold. Each fold has the train and test sets. After that, the SMOTE-R algorithm is used to augment the training dataset of each fold. We train the index model using the augmented dataset. The index model predicts the QoL index value, which decides whether either beneficiary needs support. If the QoL index value is more significant than 100, the beneficiary does not need support, and the process stops. The process moves forward if the index value is less than or equal to 100.

Furthermore, we calculate the PoC value for each aspect and beneficiary and incorporate it into the augmented dataset. We also determine the SS value by summing the PoC value of each aspect of the beneficiary's QoL (detailed in Section 3.4). We train the SS model using this modified dataset. The SS model primarily calculates the PoC score value for each aspect and predicts the SS value of support for the beneficiary. Finally, we train the SIS model using another dataset containing the SS and corresponding SIS values. The SIS model predicts the SIS value and finally takes the SS value of support. The process ends once the SIS value has been predicted.

The flowchart of the process for generating the required support report is demonstrated in Figure 5.4. After calculating the PoC value in the SS model, a support report is generated based on the PoC value. If the PoC value is above 25, the beneficiary suffers in this aspect of life and needs immediate support to improve it. After that, if the PoC value lies between 15 to 25, the beneficiary does not need direct support but requires support to perform better in this aspect of life. Finally, if the PoC value is between 10 to 15, the beneficiary only needs optional support as a precaution. In contrast, the beneficiary does not require any support in this aspect.

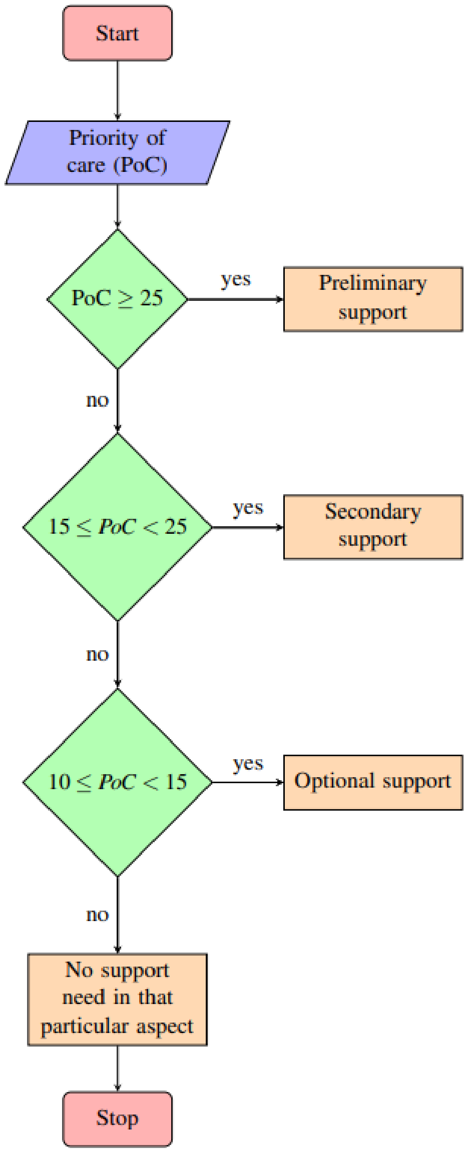


Figure 5.4: Flowchart of generating required support reports

5.3.1 Dataset preparation

We own the newton-One dataset, as we collected the information by interviewing various individuals. Based on questionnaires interviewer asked questions to each individual. The dataset includes eight QoL aspects, each with eight to nine questions. Professionals pose these questions to individuals and collect the answers based on a four-point frequency scale Verdugo et al., 2012. Following the interview, professionals use the GENCAT scale to transform the answers to each aspect question into the aspect's QoL value. There are eight aspect values in the dataset. These eight aspect values are used to generate the QoL index value based on the GENCAT Scale method Verdugo et al., 2008. The beneficiary's QoL index value indicates whether or not the beneficiary needs assistance.

The value of the QoL index varies according to the Gaussian distribution. The maximum populations are found on the Gaussian curve. The minimum value starts at 68, and the maximum reaches 130. The average of this distribution is 100. Therefore, if a person's QoL index value is above 100, she/he does not need any support. Whereas if the person's QoL index value is equal to or less than 100. Consequently, this person needs support in relation to each aspect. We have a dataset containing the information of 26 individuals affected with ID. The original dataset includes 26 subjects, eight QoL aspects, and a QoL index value. The mean and standard deviation of our dataset for each aspect of QoL and index value are depicted in Table 5.2. It also contains the minimum and maximum values of the dataset.

Table 5.2: Mean, standard deviation, minimum, and maximum values of each aspect of the Newton-One dataset.

	EW	IR	MW	PD	PW	SD	SI	RI	QoL Index
Mean	101.192308	102.076923	92.038462	104.461538	81.346154	114.730769	97.961538	94.384615	97.461538
std	14.759456	15.478819	17.340083	15.862486	15.184050	8.951235	20.588309	17.024869	16.322330
min	81.000000	81.000000	68.000000	68.000000	68.000000	94.000000	68.000000	68.000000	72.000000
max	126.000000	130.000000	115.000000	130.000000	119.000000	130.000000	126.000000	115.000000	125.000000

5.3.2 ML techniques

We employ various ML techniques in this study, including regression tree (RT) Mahmoud and Abdel-Nasser, 2018, random forest (RF) Liang et al., 2022, gradient boosting (GB) Bentéjac, Csörgő, and Martínez-Muñoz, 2021, multiple linear regression (MLR) Montgomery, Peck, and Vining, 2021, multilayer perceptron regressor (MLPRegressor) Dutt and Saadeh, 2022, and adaptive neuro-fuzzy inference system (ANFIS) Panahi et al., 2020. These algorithms are famous for dealing with linear as well as nonlinear datasets. One dependent variable and two or more independent variables are analysed using MLR to identify their relationship. MLR follows some assumptions such as independence of observations, there should be no hidden relation among variables, and data follow normality and linearity.

$$Y^i = W_0^i + W_1^i.X_1^i + W_2^i.X_2^i + ... + W_n^i.X_n^i, \tag{5.1}$$

where W_0 is the bias term, and W_1 to W_n is the weight coefficient. X_1 to X_n are the input features, and Y is the output. i represents the number of the dataset. Here, we primarily have eight aspects of QoL as input to the model, in which X_1 to X_8 and the QoL index is the output, representing Y.

The random forest regression algorithm employs the ensemble learning technique as part of its supervised learning approach to regression. The ensemble learning approach combines predictions from various ML algorithms to create a forecast that is more accurate than one from a single model Biau, 2012. MLPRegressor Dutt and Saadeh, 2022 is a neural network that contains an input layer, hidden layers, and an output layer. The MLPRegressor algorithm is adjustable and can commonly be used to become proficient in an input-output mapping, meaning it can solve complex linear and nonlinear regression problems. The ANFIS integrates the benefits of artificial neural networks and fuzzy inference systems Panahi et al., 2020. ANFIS is a reliable technique for constructing complex and nonlinear relationships between input and output data sets.

5.3.3 Detail of experiments

The algorithm must first examine the QoL index value after receiving the dataset. The QoL index ranges from 68 to 130. The beneficiary does not need assistance if the index value exceeds 100, indicating the typical distribution. Only data with a QoL index larger than 100 is eliminated. Once we have obtained data with QoL index values less than or equal to 100, the dataset is divided into a train and validation fold. This study's dataset is small enough to train an ML system. As a result, we use the synthetic minority oversampling strategy for regression to augment the train data. The number of train data points used to train our algorithm is increased. After augmentation, we acquire a large number of distinct train datasets. Nothing is added to the validation data set to validate our trained model on the original dataset. The results reveal that the trained model is accurate. After receiving the augmented train and original validation dataset, the algorithm calculates the SS value of support in the case of the SS model for both augmented train and validation data. We divided the original data by three and performed the aforementioned steps for each fold. We train our data using MLR, RT, RF, GB, MLPRegressor, and ANFIS models for all threefold. For every three models, we train separately based on the performance in each case, and we select the model accordingly. In the case of the MLR model, we evaluated the trained model and obtained an excellent R^2 score for the training testing case.

Algorithm 1 Priority of care for each QoL aspect algorithm

Require: Data, Dict = {}, List = []

Ensure: (Dict contains the maximum value of each support intensity scale, and the list contains no values initially)

function LOOP(Data)
    for $j \leftarrow 0$ to $D-1$ **do** ▷ D = Number of aspects of QoL
        if $j = 0$ **then**
            Calculate PoC_{EW} score value by averaging the three support metric values.
        else if $j = 1$ **then**
            Calculate PoC_{IR} score value by averaging the corresponding support metric values.
        else if $j = 2$ **then**
            Calculate PoC_{MW} score value by averaging the corresponding support metric values
        else if $j = 3$ **then**
            Calculate PoC_{PD} score value by averaging the corresponding support metric values.
        else if $j = 4$ **then**
            Calculate PoC_{PW}score value by averaging the corresponding support metric values.
        else if $j = 5$ **then**
            Calculate PoC_{SD} score value by averaging the corresponding support metric values.
        else if $j = 6$ **then**
            Calculate PoC_{SI} score value by averaging the corresponding support metric values.
        else if $j = 7$ **then**
            Calculate PoC_{RI} score value by averaging the corresponding support metric values
        end if
    end for
    $SS \leftarrow Sum(PoC_{EW}, PoC_{IR}, PoC_{MW}, PoC_{PD}, PoC_{PW}, PoC_{SD}, PoC_{SI}, PoC_{RI})$
    $List \leftarrow SS$
end function

Figure 5.5: Algorithm to calculate the POC value and standard score value

5.3.4 Standard score of support corresponding QoL

We calculate the PoC value corresponding to each aspect of QoL for every beneficiary. Initially, we calculate PoC for an area of action separately, and each aspect has corresponding support group actions as depicted in Table 5.1 in the second column. Table 5.1 illustrates that the first aspect of QoL, EW has three support areas known as S1E, S2, and S3B. The PoC value for every three areas is calculated separately using (5.2).

$$PoC = \frac{A - E}{A} \times I, \tag{5.2}$$

where A is the average QoL index value, E is the aspect value, and I is the maximum support need value. The values of I are listed in Table 5.1, A value is 100, and the beneficiary gives the E value.

We calculate the final PoC_{EW} value for the EW aspect by averaging the PoC values for each of the three areas for the EW aspect. Likewise, we calculate the PoC value of other aspects. After obtaining all the PoC values for each aspect, we compute the SS by summing them as shown in equation (5.3).

$$SS = PoC_{EW} + PoC_{PD} + PoC_{PW} + PoC_{SD} + PoC_{IR} + PoC_{SI} + PoC_{MW} + PoC_{RI}, \tag{5.3}$$

Algorithm 5.5 presents the steps to calculate the PoC value for each aspect by following equations (5.2) and (5.3). Once the PoC value has been calculated, they are summed to obtain the SS value. SS value represents the support score value corresponding to eight aspects of QoL. Then, for the SS model's training, the new column is added to the original data. We calculated the SS value of each augmented data threefold. The training of the SS model takes eight aspects as input and SS value as a target. Once training was completed, we validated our trained model using the original validation dataset.

5.3.5 Data augmentation using SMOTE-R

The synthetic minority over-sampling technique for regression SMOTE-R Torgo et al., 2013 is an oversampling technique to increase the imbalanced dataset. The SMOTE concept is proposed to reduce the imbalanced distribution of a dataset for classification tasks Chawla et al., 2002. It utilizes the Gaussian noise concept to maximize the amount of data. In regression, entering the data and the corresponding column helps determine the minimum value of the column. It is a sampling strategy that undersamples the standard classes and oversamples the rare categories to address an unbalanced distribution of classes. It generates synthetic data corresponding to minority target values. Algorithm 2 presents the detailed working process of SMOTE-R Torgo et al., 2013.

Algorithm 1 SMOTE for regression algorithm

function SMOTE-R(D, thr, o, u, k) ▷ D - dataset, thr- The value of the target variable's relevance threshold, o and u represent the percentage of over and under-sampling, k -In case of generation, the number of neighbours used.
  $rearL \leftarrow \{(x,y) \in D : \phi(y) > thr \bigwedge y < \hat{y}\}$ ▷ $\hat{y}$ is the median of the target Y
  $newCasesL \leftarrow GenSynthCases(rareL, \%o, k)$ ▷ generate synthetic cases for rareL
  $rareH \leftarrow \{(x,y) \in D : \phi(y) > thr \bigwedge y > \hat{y}\}$
  $newCasesH \leftarrow GenSynthCases(rearH, \%o, k)$ ▷ generate synthetic cases for rareH
  $newCases \leftarrow newCasesL \cup newCasesH$
  $nrNorm \leftarrow \%uof|newCases|$
  $normCases \leftarrow sampleofnrNormcases \in D\{rareL \cup rareH\}$ ▷ under-sampling **return** $newCases \cup normCases$
end function

5.3.6 Developing the prediction models

The three prediction models are the index, SS, and SIS models. The index model is the first to determine if a person requires assistance, and it predicts the QoL index value using eight aspects of the QoL value as input. The SS model is trained to predict the SS value. It estimates the SS value based on eight aspects of QoL. As a result, it is the most common model. The SIS model uses the output of the SS model to predict the SIS value. The SS model's output is used to predict the SIS value by the SIS model. These three models accomplish their tasks sequentially. A multivariate linear regression approach is used to train all three models. The second branch is followed to provide a beneficiary with a support report corresponding to the requirement in a particular aspect. The second branch uses the index and SS models, which estimate the PoC for each beneficiary's eight aspects. The approach to computing the PoC for each aspect is illustrated in algorithm 5.5. The value of PoC determines the level of support.

As demonstrated in Figure 5.4, it contains three conditions. If an aspect's PoC value is greater than or equal to twenty-five, the beneficiary needs an immediate action plan in relation to that QoL aspect. Furthermore, if the PoC of an aspect is more significant than fifteen but less than twenty-five, a person requires secondary-level assistance. When a person's PoC of an aspect is more significant than ten or equal to ten but less than fifteen, the recipient will only require partial assistance, which is not mandatory. Furthermore, no assistance is required if a person's PoC in any aspect is less than ten. Our model generates all three sorts of support for each aspect of each person. Beneficiaries can use these tips without the assistance of a professional, and they will improve their lives.

5.3.7 Validation and evaluation metrics

To cross-validate the model performance of the random train and validation case data in the K fold, the k-fold cross-validation technique is utilised in this work Montgomery, Peck, and Vining, 2021. The new dataset displays how well the model performed. It is generally used in the case of a limited data sample. This procedure has a single parameter, K, which depicts how many folds we need to split the original dataset. The data are randomly divided into the fold. In this work, we split our dataset three-fold. We get to train and validate data in each fold differently. Now we use the SMOTE-R algorithm to increase the train data in each fold and keep validation data to verify the model accuracy.

In this work, we assess the effectiveness of the trained model using three assessment metrics: MAE, RMSE, and R^2. MAE determines the absolute difference between the expected and actual values. MAE can be formulated as follows:

$$MAE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}, \tag{5.4}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the total number of data points.

RMSE measures the dispersion of the residuals, not the distance between the data points and the regression line. RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{n}}, \tag{5.5}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of data points.

Finally, we calculate the R^2 score value, known as the coefficient of determination. This statistic demonstrates a model's ability to fit the data. R^2 can be defined as follows:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}, \tag{5.6}$$

where $\sum(y_i - \bar{y}_i)^2$ is the total sum of the square, and $\sum(y_i - \hat{y}_i)^2$ is the sum of squares of residuals.

5.4 Results and discussion

Tables 5.3, 5.4, and 5.5 display the results of the index, SS, and SIS models. Table 5.3 presents the results for the first case, where the input is the eight aspects of QoL, and the target is the QoL index. Results are shown here in the form of various metrics. Six regression models we trained for three different folds. Table 5.3 contains the best results threefold for all six algorithms. The fundamental metrics representing the regression performance are MAE, RMSE, and R^2 scores. It should be noted that we use a three-fold cross-validation technique in our experiments. Each cross-validation fold produces MAE, RMSE, and R^2 values. In all Tables, we report the mean and standard deviation (SD) of the evaluation metrics values of the three folds. R^2 score demonstrates the closeness of the predicted value with the actual value, varying between zero and one. One demonstrated the overlap between the predicted value with the regression line. Therefore, a higher value of R^2 shows the model's goodness. Both MAE and RMSE measure the accuracy of the predictions of the ML regression models and demonstrate the amount of deviation from the actual values. Table 5.3 presents the index model's mean and SD of MAE, RMSE and R^2 values. MLPRegressor and ANFIS produce MAE of 4.1344 and 2.6936, respectively. RT, RF, and GB produce MAE and RMSE values higher than 5. MLR achieves the smallest MAE and RMSE values (1.2247 and 1.5236) and the highest R^2 score value than other regression techniques results for the index model.

Table 5.3: The MAE, RMSE, and R^2 values of the six ML techniques used for building the index model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD
MAE	1.2247 $\pm$ 0.0904	9.2030 $\pm$ 3.0083	5.5271 $\pm$ 2.1788	5.9723 $\pm$ 1.2817	4.1344 $\pm$ .4683	2.6936 $\pm$ 0.1325
RMSE	1.5236 $\pm$ 0.1644	11.6664 $\pm$ 4.2920	6.5353 $\pm$ 2.4031	7.2981 $\pm$ 1.6082	5.1202 $\pm$ 0.3524	2.8361 $\pm$ 0.1722
R^2 Score	0.9897 $\pm$ 0.0051	0.4213 $\pm$ 0.3698	0.8227 $\pm$ 0.0935	0.7650 $\pm$ 0.1257	0.8698 $\pm$ 0.0095	0.9259 $\pm$ 0.0619

Table 5.4: The MAE, RMSE, and R^2 values of the six ML techniques used for building the SS model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD
MAE	0.2260 $\pm$ 0.2655	39.9563 $\pm$ 2.3875	32.8962 $\pm$ 2.4674	34.7427 $\pm$ 1.6301	7.8524 $\pm$ 0.3678	0.7362 $\pm$ 0.0785
RMSE	0.2765 $\pm$ 0.3173	47.3084 $\pm$ 3.9119	39.2744 $\pm$ 1.4503	40.0628 $\pm$ 0.6041	9.5437 $\pm$ 0.2537	0.9821 $\pm$ 0.1080
R^2 Score	0.9998 $\pm$ 0.0002	-0.7837 $\pm$ 0.4098	-0.2575 $\pm$ 0.4463	-0.3001 $\pm$ 0.4129	0.9341 $\pm$ 0.7849	0.9632 $\pm$ 0.0445

Table 5.4 presents the results for the SS model. Here the input parameters are the eight aspects of QoL, and the corresponding target is the PoC value. RT, RF, GB, MLPRegressor, and ANFIS produce an average MAE higher than 0.6. MLR achieves the highest R^2 score value (0.9977) and the lowest MAE value (0.7702). Therefore, we use MLR trained model for future integration with a SS model to predict the SS value.

Table 5.5 presents the results of the SIS model. It takes input as SS value and has a corresponding SIS value, representing the final support metric for a person's fundamental aspects of QoL. MLR achieves MAE, RMSE, and R^2 of 0.7702, 1.3403, and 0.9977, respectively, which are much better than the results of RF, GB, MLPRegressor, and ANFIS. However, RT produces RMSE values lower than MLR, and the MAE and R^2 values of MLR are better than RT. Consequently, we selected MLR to build the SIS model.

Table 5.5: The MAE, RMSE, and R^2 values of the six ML techniques used for building the SIS model.

Algorithms	MLR	RT	RF	GB	MLPRegressor	ANFIS
metrics	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD	Avg $\pm$ STD
MAE	0.7702 $\pm$ 0.1025	0.9091 $\pm$ 0.1285	0.6657 $\pm$ 0.03816	0.9364 $\pm$ 0.1030	1.3324 $\pm$ 0.0968	1.7536 $\pm$ 0.5361
RMSE	1.3403 $\pm$ 0.4674	1.1692 $\pm$ 0.0253	1.02734 $\pm$ 0.4342	1.1821 $\pm$ 0.0381	1.6998 $\pm$ 0.1068	1.9642 $\pm$ 0.7385
R^2 Score	0.9977 $\pm$ 0.0007	0.9976 $\pm$ 0.0002	0.9981 $\pm$ 0.0014	0.9976 $\pm$ 0.0005	0.9950 $\pm$ 0.0008	0.9825 $\pm$ 0.0125

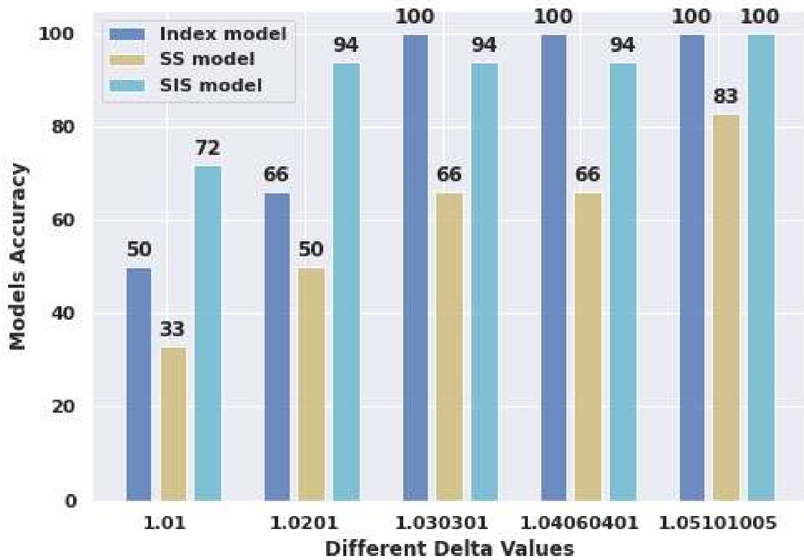


Figure 5.6: The threshold accuracy measure of the index, SS and SIS models with five threshold values (five z values). The blue bars represent the index model, the yellow bars represent the SS model, and the green bars represent the SIS model

.

Our entire model has two branches. First, it contains a trained MLR model that considers eight aspects of QoL and predicts the QoL index value. The algorithm determines whether or not the patient needs support based on the QoL index value. The algorithm is trained to pass only those patient information to further for an analysis whose QoL index value is less or equal to 100. After detecting the QoL index value, it then goes through the second trained model after detecting the QoL index value. During testing, the SS model receives input from the beneficiary and predicts the SS value. The third trained SIS model receives the projected value as input and predicts the SIS value after the SS model predicts the SS value, which shows the support value corresponding to a beneficiary's eight aspects of QoL. In the second branch, after predicting the QoL index value by the index model, the model decides whether the person needs support. Once the decision is made model calculates the PoC value and, based on that, generates the support report as shown in flowchart 2 in Figure 5.4.

Assessing the accuracy of a regression model is a tedious task. Here, we use the λ accuracy measure and threshold accuracy measure Abdulwahab et al., 2020. As presented in algorithm 3, λ is an error threshold value between the predicted and actual values. If the error calculated between the actual and predicted value is less than or equal to the λ value, we consider the prediction accurate; otherwise, inaccurate. In the current research, the accuracy is directly dependent on the λ value. If the value of λ is high, accuracy is directly increased. If the λ value is small, the error acceptance is compassionate, and the accuracy may be less. We set the λ value to 0.75, yielding 100, 83.33, and 66.67% accuracy with the index, SS and SIS models, respectively.

We also utilise the threshold accuracy metric to assess the ML regression models' accuracy in estimating errors under various thresholds, demonstrating how frequently our estimate is accurate. Essentially, the threshold accuracy measure is the expectation E that the predicted value error of the input is less than a threshold thr^z Abdulwahab et al., 2020:

$$\delta_z = E[F(max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y}) < thr^z)], \tag{5.7}$$

where $F(.)$ is an indicator function that outputs a binary value (0 or 1), y and $\hat{y}$ are the actual and predicted values, respectively, and thr is a threshold value set as 1.01. z is the parameter that controls the acceptance span. Here, the values of z range from 1 to 5 that correspond to $thr^z = \{1.01, 1.0201, 1.030301, 1.04060401, 1.0510100501\}$. As the value of z increases, the span of acceptance increases.

Figure 5.6 presents the accuracy of the index, SS and SIS models with five threshold values (five z values). In the case of $z = 1$, the threshold

$thr^1 = 1.01$, resulting in accuracy values of 50, 33, and 72% with the index, SS and SIS models, respectively. The accuracy of the SIS model is identical for the threshold values 1.0201, 1.030301, and 1.04060401. When we set a high acceptance span z to 5 (thr^5 = 1.05101005), it achieves the highest accuracy values of the index, SS and SIS models. For the index and SIS models when $z = 5$, the term $max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y})$ of all test samples is lower than 1.05101005, and thus the accuracy is 100 %. While in the case of the SS model, the term $max(\frac{y}{\hat{y}}, \frac{\hat{y}}{y})$ of all test samples is not lower than 1.05101005.

5.4.1 Statistical analysis of index model

For the statistical analysis of our trained index model, we collected additional twenty data by conducting the rigorous interview of the twenty ID elderly people. Hence, the sample size for the statistical analysis is twenty. The sample size is less than thirty; therefore, we perform a t-test. We predicted the index value using the trained index model for the new twenty data. Now, we have predicted index value and actual index value for the twenty sample data. We apply the t-test to these two samples using the following equation 5.8. Here, s is the sample's standard deviation as determined by equation 5.9.

$$t = \frac{(\bar{x} - \mu)}{\frac{s}{\sqrt{n}}}, \tag{5.8}$$

where, μ is the theoretical mean of the sample, $\bar{x}$ is the observed mean of the sample, s is the standard deviation of the sample, n is the sample size.

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}, \tag{5.9}$$

here $n - 1$ is the degree of freedom. For the testing, we imported the t-test function from the Scipy library. After the testing, we got a t-test value lower than the critical value. As a result, the null hypothesis is accepted because the t-test value is less than the critical value. Which states the two samples' means are the same. As a result, our trained model is trusted to forecast the index value for the new case data.

5.5 Conclusion and future work

This paper proposed an efficient ML-based approach for predicting the QoL of dependent people with intellectual and developmental disabilities in guardianship entities by analyzing the various aspects of QoL. In addition, we proposed three ML models for predicting QoL index value, SS value, and SIS value, along with implementing and testing different ML techniques: MLR, RT, RF, GB, MLPRegressor, and ANFIS. Based on the QoL index, the proposed approach determines the PoC for each aspect of QoL. We validated our approach on a dataset collected from guardianship entities. We found that MLR yields the best prediction results for the QoL index, SS, and SIS. The QoL index, SS, and SIS ML models achieved MAE values of 1.2247, 0.2260, and 0.7702, respectively, and also obtained average R^2 scores of 0.9897, 0.9998, and 0.9977, respectively. The proposed ML approach can assist professionals in analyzing the QoL of a beneficiary to determine which measures are required to improve their QoL.

Future research will focus on using sensory data (e.g., data of sensors that monitor the activity of dependent individuals) collected from guardianship entities to improve the efficiency of the suggested strategy.

Chapter 6

An Data-driven Model to Predict Quality of Life Dimensions of Intellectually Disabled People based on GENCAT Scale

Summary

Recently, it has been observed that the life of intellectually disabled (ID) people in their older age is getting more complex. Many initiatives launched by organizations and government bodies are rigorously working to improve ID people's quality of life . QoL is a multifaceted notion having elements that are both etic (universal) and emic (culture-bound). It possesses objective and subjective characteristics and is influenced by

environmental and societal influences. The researcher proposed eight dimensions to cover every aspect of ID people. In the last decades, professionals made the GENCAT scale to predict these eight dimensions' values through a set of questionnaires containing 69 questions. Professionals respond to 69 questions based on four point frequency scale by interviewing the beneficiary. The answers to these 69 questions were transformed into values for eight dimensions by the GENCAT scale programme. The GENCAT tool uses a set of rules and some correlatable tables to evaluate the eight dimensions values. In this work, we propose using a machine learning-based model instead of the GENCAT tool. We train various ML models using the NewtonOne dataset to predict the eight dimensions value. A trained ML model predicts the eight dimension values by inputting 69 questions and responses. The root means square error, mean absolute error, and R^2 score values are used to evaluate how well different models perform.

6.1 Introduction

According to the World Health Organisation, there are over a billion disabled individuals globally Krahn, 2011. It represents almost 15 % of the world's population. According to government statistics, in Spain, 8.7% of people with disabilities have an ID González-Valero et al., 2021. ID is a common neurodevelopmental condition, formerly known as mental retardation, and is defined by significant impairment in intellectual and adaptive functioning Schalock, Luckasson, and Tassé, 2021a. Problem-solving, learning, and judgment are examples of intellectual functioning, whereas adaptive functioning is defined as regular activities like independent living and communication. Intelligence and adaptability deficiencies appear as early as birth. If someone under 18 has an intelligence quotient (IQ) below 70 and shows signs of reduced adaptive functioning, they are deemed to have an ID Patel et al., 2020. It could show itself in milder or more severe ways. An average IQ is between 50 and 70 in about 85% of cases. These children can finish their education, develop independence, get training, and even pursue employment. An average IQ is between 35 and 50. Even though they can function independently, they frequently need care and attention. IQs between 20 and 35 usually indicate a severe intellectual disability. These people are less competent and struggle to understand numbers and reading. They need to be attended over frequently. ID is a severe intelligence deficit when the IQ is less than 20 Wieczorek, 2018; McKenzie et al., 2016. There are numerous causes of ID. Some are brought on by issues that develop after childbirth when a child receives medical treatment for an illness. Moreover, others inherit. Numerous children have IDs that continue throughout their lives for inexplicable reasons. Nevertheless, ongoing care and prompt diagnosis can improve adaptive functioning throughout childhood and adulthood.

In recent decades, the lifespan of those with ID has grown. The increased life expectancy of older people with ID is challenging for older adults, and their requirement to live a healthy life needs specific support. The demand for support tailored to each age group is raised by people with ID living longer. According to the World Health Organization, ID policies in industrialized nations should emphasize "productive or prosperous ageing" Alftberg, Johansson, and Ahlström, 2021; Organization, 2001. Active ageing refers to older adults (with ID) continuing to lead fulfilling, active lives in their communities after retirement Schepens, Van Puyenbroeck, and Maes, 2019a. They want to be active in their communities, maintain significant connections and networks, and age comfortably in their chosen places. They like to live longer, have a higher quality of life, and feel more satisfied while functionally capable and essentially free from disease or emotional stress. The nature and frequency of (lifelong) care for ID people must change with ageing to achieve optimal QoL Simões and Santos, 2017.

The recent literature works Yadav et al., 2022a; González-Valero et al., 2021 used to analyze the dimensions of the QoL using the GENCAT tool. The GENCAT tool is an essential factor in every work to transform the responses of the ID people on questions related to their QoL dimensions. The GENCAT tool was proposed by Verdugo et al. Verdugo et al., 2010. GENCAT tool follows the rules and tables to calculate the eight dimensions value. The steps recommended for creating multidimensional QOL questionnaires were used in creating the GENCAT Scale Maestro-Gonzalez et al., 2018. The GENCAT scale contains sixty-nine questions related to QoL dimensions Verdugo et al., 2010. The GENCAT tool is used in many areas nowadays and primarily for intellectual and developmental disabilities (IDD). In this area, the GENCAT tool plays a role in evaluating the 69 questions responses. The GENCAT has a set of rules and tables, with the help of which it calculates the eight dimensions values. This method of using all the steps of the GENCAT tool is completely rule-based. This process is laborious and needs expert professionals to conduct this experiment. These manual calculations will be complex with many patients, which may cause delays in taking corrective actions to improve their QoL. To cope with the issues mentioned above, for the first time, as per our knowledge, we propose an ML-based approach to predict the eight dimensions' values. The massive success of ML-based models in various fields motivates us to use these learning-based algorithms to predict the QoL dimensions' values. Thus, we train various ML models in this work using the NewtonOne dataset, a private dataset collected by some ID professionals. We use the trained ML models in place of the GENCAT tool to calculate the eight QoL dimensions' values. We trained the ML models in different scenarios of the dataset. Such as, we use an augmented dataset and data normalization in order to train the various ML-based models. The trained model enables us to effectively learn the relationship between the 69 questionnaires to 8 dimensions of QoL. Earlier GENCAT tools used various rules and tables to map the relationship between questionnaires to QoL dimensions. Once the neural network model is trained, we can use it to generate QoL dimensions directly without any interference from the professional to use

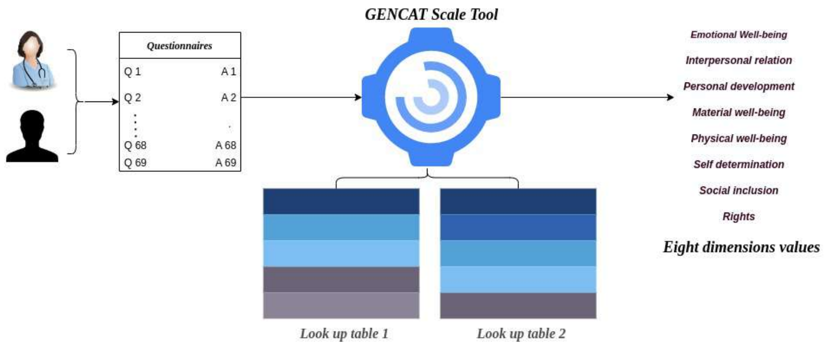


Figure 6.1: GENCAT tool approach to calculate the eight quality of life dimensions value using various look-up tables, which makes this process tedious. For every beneficiary professional need to calculate the eight dimensions value using look tables which time-consuming and need attention to calculate the eight dimensions values accurately

the GENCAT scale to transform the response of the questionnaires.

- Developing an AI-based approach to learning the relationship between the set of questionnaires and estimating the corresponding QoL dimensions.
- Using a learning-based model instead of a statistical tool to generate the eight dimensions of QoL.
- Training ML models using Newton One dataset. Data contains 102 beneficiaries' information having 69 questions and corresponding eight dimensions values.

This paper consists of five sections. The second section depicts the related research work. The third section shows the details of the methodology and experimental details, and the fourth section shows the result of this work. Finally, section five concludes this work and gives an idea about future work.

6.2 Background

6.2.1 Intellectual disability

According to estimates, between 1% and 3% of the general population suffers from some form of ID Patel et al., 2020. A person with an ID has severe limits in their capacity to think conceptually, interact socially, use practical skills, and adapt to their environment Schalock, Luckasson, and Tassé, 2021b. This disability first manifests during the developing phase, which is defined as occurring before the person reaches the age of 18. IDs were previously referred to as mental retardation and are no longer used. An underlying biologic cause is more likely in severe to profound intellectual deficits, affecting a small percentage of people with intellectual disabilities. Since moderate intellectual disorders are more common than severe ones, a biological cause is less likely to be identified in this population Patel et al., 2018. The health-related issue is prevalent in ID people. Compared to the general population, ID people face significant health disparities, obstacles to accessing quality healthcare, and a higher risk of dying young and from preventable causes O'Leary, Cooper, and Hughes-McCormack, 2018; Trollor et al., 2017. People with ID have a higher prevalence of physical diseases than the general population, particularly neurological problems, constipation, obesity, sensory impairments, and congenital malformations Kinnear et al., 2018; Liao et al., 2021.

6.2.2 ID people during ageing

Individuals with ID who live into old age have various obstacles to leading a typical life. Their evolving needs challenge caregivers who want to provide the highest level of care. Ageing persons with ID have had highly different life path trajectories than those without ID, and many have been marginalized in society Schepens, Van Puyenbroeck, and Maes, 2019a. Children support few people with ID, and many

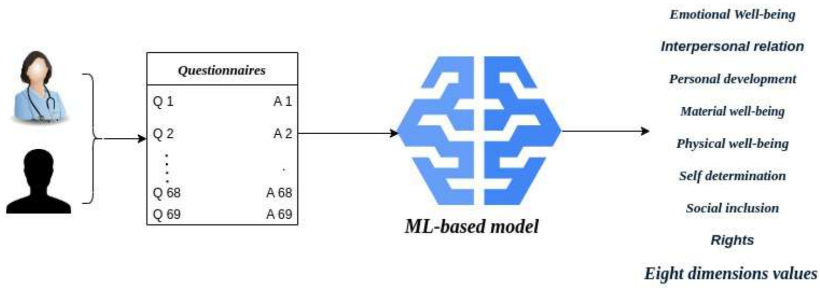


Figure 6.2: The proposed work replaces the tedious and time-consuming steps using an ML-based model. The proposed model takes 69 questions responses and gives eight dimensions value more accurately and fast than the GENCAT tool approach

still depend on their elderly parents. Some individuals with ID have resided in residential facilities since they were young. They rely on frequently-changing professional support staff for daily care. Due to frequent changes, ID people face difficulties forming good relationships with their caregivers. It is not very pleasant to the ID people. Numerous common disorders, such as memory loss, depression, and loneliness, are brought on by ageing. These disorders have a severe impact on people with ID. As they age, they require particular attention. Professionals are working to provide age-specific support plans to help them live optimally.

6.2.3 Quality of life

For a long time, there has been discussion surrounding the QoL. Over 200 definitions of QoL can be found in the earliest quality of life literature Van Loon, 2013; Urizarna-Varona et al., 2018. However, as it was introduced into the intellectual disability (ID) field, the emphasis moved from defining the QoL to figuring out its essential components and related measure. This change was made possible by the realization that QoL is more of a multifaceted phenomenon than a personal characteristic or a particular stage of life. We now understand that personal and environmental influences influence the quality of people's lives; therefore, it needs to be evaluated using subjective and objective metrics Grabowska et al., 2021.

6.2.4 QoL's dimensions

The elements that make up the QoL are known as its dimensions. The range represented by this set encompasses the QoL and encapsulates its multidimensionality. These dimensions of QoL cover every aspect of ID people's life. The three fundamental areas, the personal, the social and the judicial, need support to improve the QoL of a person. The eight dimensions of the QoL cover these three areas broadly. These eight fundamental aspects are interpersonal relationship (IR), material well-being (MW), personal development (PD), physical well-being (PW), self-determination (SD), social inclusion (SI), and right (RI). The personal area of life covers EW, PD, PW, and SD. The social area of life covers IR, and SI Yadav et al., 2022b. Finally, the judicial area covers MW and RI. Support paradigms in intellectual and developmental disease (IDD) help beneficiaries to improve these eight dimensions Van Loon, 2013. Table 6.1 shows the eight dimensions of QoL, the indicator of QoL, examples related to each dimension of the QoL and the area of QoL. Indicators depict the specific characteristics of each dimension. A dimension has various indicators to represent the dimension. The indicator tells us which dimensions need help. There could be many possible indicators of a dimension, a few of which we showed here in Table 6.1. The third column of Table 6.1 shows an example of an indicator related to its dimension. In turn, the fourth column shows the area of the QoL related to each QoL's dimensions.

6.2.5 GENCAT Scale

With the help of a self-administered questionnaire called the GENCAT Scale, professionals can assess the user's quality of life by providing objective, observable answers to questions based on personal experience. It comprises 69 items that are third-person declarative statements, which are distributed randomly across the eight scales that correlate to the eight dimensions of quality of life. The format for the response is a frequency scale containing four choices (always, frequently, sometimes, never). The GENCAT Scale was created utilizing a methodical and exacting process. Other multidimensional QoL measures that prioritise context have been developed using it as a paradigm on a global scale Verdugo et al., 2010. All types of consumers of human and social services, including the elderly and persons with disabilities, as well as those who have mental health issues or are infected with AIDS or HIV, are eligible to utilize the GENCAT Scale Gómez, Schalock, and Verdugo, 2021. The GENCAT tool contains a set of rules and lookup tables that help professionals map the relationship between the 69 questions and the eight dimensions of QoL. Finding a relation between the questions and QoL dimensions is a tedious task requiring an ID professional to complete. The complete GENCAT scale contains many steps; however, they are statistical-based. In our approach, we propose using learning-based algorithms to calculate the eight-dimension values from 69 questionnaire responses.

6.2.6 Likert scale

To quantify "attitude" in a scientifically recognized and proven way, the Likert scale was developed in 1932 Joshi et al., 2015. An attitude is a preferred plan of action or response to a certain situation based on a structured system of beliefs and ideas formed through time through social interaction (around a particular object, subject, or concept). The description above clarifies that attitude delivery in a particular scenario

Table 6.1: The indicator of the QoL dimensions and related examples for each dimension.

QoL aspects	Indicators of QoL	Example	Area of support
Emotional well-being (EW)	Positivity, self-esteem, contentment, and stress relief, Mental stability, negative feelings, happiness with life, experience, lack of stress, satisfaction,	They exhibit depressive symptoms.	Personal area
Personal development (PD)	education, work abilities, and day-to-day chores, Education Job, learning opportunities, activities of daily living functional abilities (personal competency; adaptive behaviour, etc.)	She/he is participating in creating their own personal planning.	Personal area
Physical well-being (PW)	basic health, sleep, physical exercise, Health status, sanitary care, Health care, health consequences, technical assistance, mobility	He or she struggles with sleeping.	Personal area
Self-determination (SD)	The ability to make decisions, goals and personal preferences, choices personal preferences, Autonomy, decisions, Goals	How to spend his/her money are decided by others.	Personal area
Interpersonal relations (IR)	relations in society, Social relationships, to have stable and identifies friends Relation within the family, family relationships	She/he bemoans how they get along with their buddies.	Social area
Social inclusion (SI)	Inclusion, integration, support involvement, assistance, Participation	Support comes from his or her family.	Social area
Material well-being (MW)	wages, Services conditions, Workplace conditions possessions housing conditions, belongings workplace conditions Housing conditions, employment, incomes	His/her earnings are insufficient to support desires.	Judicial area
Rights (RI)	law human, the exercise of rights, Knowledge of rights, respect privacy, defence of rights	He or she experiences exploitation, assault, or abuse.	Judicial area

consists of thinking, feeling, and acting, all in various combinations and permutations Joshi et al., 2015. Several assertions are offered for a hypothetical or real-world event under consideration using the Likert scale. The presented statement must be rated on a metric scale by participants as to their level of agreement or disagreement. Participants show their responses by choosing one of these answers. This work uses a four-point scale for every 69 questions. The four points shown to the beneficiaries for each questionnaire are "always, frequently, sometimes, and never". These four points scale from 1 to 4 based on the requirement of the questions.

6.3 Methodology

6.3.1 NewtonOne dataset

This dataset is collected during the Never Alone project by professionals. During the interaction between professionals and ID people, professionals need to answer the 69 questions response based on the ID people's conditions and replies. The 69 questions are one part of the GENCAT scale. The dataset contains 102 beneficiary data, of which 26 were collected in the first stage of the project Never Alone, and 20 more were

Table 6.2: Representation of the GENCAT tool that employs 69 questions for estimating eight aspects of the QoL. The first column contains the eight aspects' names. The second column contains two questions sample corresponding to each aspect and followed by the response to the questions based on a four-point Likert scale.

QoL aspects	Questionnaires and corresponding response				
	Questionnaires	Always	Frequently	Sometimes	Never
Emotional well-being (EW)	He is satisfied with his present life.	4	3	2	1
	You have symptoms of depression.	1	2	3	4
Interpersonal relation (IR)	Do activities you like with other people.	4	3	2	1
	He complains about the lack of stable friends.	1	2	3	4
Material well-being (MW)	The place where you work complies with safety regulations.	4	3	2	1
	He is dissatisfied with where he lives.	1	2	3	4
Personal development (PD)	Has access to new technologies (Internet, mobile phone, etc.).	4	3	2	1
	Has difficulty solving problems effectively that are put to him.	1	2	3	4
Physical well-being (PW)	You have sleep problems.	1	2	3	4
	Technical aids are available if you need them.	4	3	2	1
Self-determination (SD)	Has personal goals, objectives and interests.	4	3	2	1
	Other people decide about your personal life.	1	2	3	4
Social inclusion (SI)	His family supports him when he needs it.	4	3	2	1
	There are physical, cultural or social barriers that hinder their social inclusion.	1	2	3	4
Rights (RI)	In his environment he is treated with respect.	4	3	2	1
	Shows difficulties in defending their rights when they they are raped.	1	2	3	4

collected in the second stage. Finally, 56 data were taken from the repository made during the project. Table 6.2 shows the two questions for each eight dimension. The response to the questions is collected based on 4 points Likert scale. These four points are always, frequently, sometimes, and never. The response's value varies from 1 to 4. Each dimension has eight to nine questions; the professionals need to provide the question's answers by analysing the beneficiary response. Professionals must respond to this self-administered questionnaire's observable and objective questions about ID people QOL. It consists of 69 questions from every eight dimensions, and the number of questions from each dimension is as follows: IR=10, EW = 8, PD=8, MW = 8, PW = 8, SI=8, SD = 9, and RI = 10. Each question is a third-person declarative statement randomly ordered and grouped by domain. Questions have a positive valence in half (n = 35) and a negative valence in half (n = 34). The Newton One dataset contains 65 % female beneficiaries and 35 % male beneficiaries over 55 years.

6.3.2 Data augmentation

Our dataset has data from 102 beneficiaries. The data shape is $[102, 77]$, where 102 rows show the number of beneficiaries. Out of 77 columns, 69 contain the 69 questionnaire responses collected during the interview by professionals, and eight columns have corresponding eight dimension values. We need more data to train the machine learning models, so we augmented our training data. Primarily we split our data into train and test cases. We kept 30% (31) data for the test case and the remaining 70%(71) data for the train case. We augmented the train case data from 71 to 135 using the SMOGN algorithm Branco, Torgo, and Ribeiro, 2017.

6.3.3 Synthetic minority over-sampling technique for regression with random Gaussian noise (SMOGN)

SMOGN Branco, Torgo, and Ribeiro, 2017 combines two oversampling methods-SMOTE-R (Synthetic minority over-sampling technique for regression) and the addition of Gaussian Noise with random under-sampling. SMOTE-R Torgo et al., 2013is a sampling technique that solves regression issues with of minority of data. It originated from SMOTE (Synthetic minority over-sampling technique)Fernández et al., 2018 algorithm, which is very popular for classification issues with unequal distribution. The main advantage of SMOTE is how it balances undersampling the majority classes while oversampling the minority classes. The original SMOTE method used an over-sampling technique that creates "synthetic" examples with an uncommon target value. SMOTE builds these synthetic instances using an interpolation approach. The approach is to randomly choose one of each case's k-nearest neighbours from the collection of observations with unusual values for each observation. The attribute values of the new example are interpolated from the values of the two previous cases using these two observations Camacho, Douzas, and Bacao, 2022.

Table 6.3: MAE, RMSE, R^2 score value of MLP, for all four case scenarios discussed in ablation study

Metrics	Original data	Original and normalized data	Augmented and normalized	Augmented data
	Train|*Test*	*Train*|*Test*	*Train*|*Test*	*Train*|*Test*
MAE	0.3230|5.1130	0.0021|0.0176	0.0007|0.0067	0.2153|2.1091
RMSE	0.4056|6.5046	0.0023|0.0227	0.0009|0.0137	0.2827|4.3343
$R^2 Score$	0.9988|0.7388	0.9929|0.3742	0.9990|0.7848	0.9996|0.8419

The fundamental concept behind the SMOGN algorithm is to combine both methods for creating synthetic examples to simultaneously reduce the risks that SMOTE-R may face by employing the more cautious strategy of introducing Gaussian Noise and permitting an increase in the diversity of examples generation, which is not possible by utilizing only the introduction of Gaussian Noise. When the seed example and the chosen k-nearest neighbour are "close enough," SMOGN will employ SMOTE-R to create new synthetic examples; however, when the two examples are "further distant," SMOGN will use the insertion of Gaussian Noise Branco, Torgo, and Ribeiro, 2017. Two factors drive SMOGN: 1) a reduction in the risks associated with using SMOTE-R because it does not use the farthest examples within the interpolation process; and 2) the ability to expand the decision boundaries for such rare cases to increase generalization capability, which is more challenging to accomplish with the incorporation of Gaussian Noise because it is a more excellent conservative approach Branco, Torgo, and Ribeiro, 2017.

6.3.4 Data preprocessing

The data contains 69 input features and eight output features. The range of input features varies between 1 to 4, whereas the output varies from approximately 68 to 130. The data distribution follows the Gaussian distribution, 90% of the population of eight dimensions follow between 68 to 130. The variation in the range of input and output features is relatively high. Therefore, normalization techniques can improve the model's performance in the data preprocessing process. In the columns of the dataset, normalisation converts the numerical values to a standard scale without removing any information or changing the value ranges. Some algorithms need normalization to model the data correctly. There are many techniques to normalize the data. The two most prevalent techniques are min-max normalization and z score.

Min-max normalisation is one of the most popular methods for reshaping data. All values are transformed to decimals between 0 and 1, except each feature's minimum and maximum values, which are individually converted to 0 and 1, respectively.

$$z = \frac{x - min(x)}{max(x) - min(x)}, \tag{6.1}$$

6.3.4.1 Z Score normalization

Z-score normalization transforms each item in a dataset so that the μ is 0 and the σ is 1.

$$z = \frac{x - \bar{x}}{\sigma}, \tag{6.2}$$

where z is the new normalized value. σ is the standard deviation, and $\bar{x}$ is the mean of the data.

6.3.5 Multi-layer perceptron regressor (MLP)

Multilayer Perceptron is a fully connected multi-layer neural network. It has three layers: an input, a concealed, and an output layer. Deep artificial neural networks (ANNs) are referred to as having more than one hidden layer. Biological neural networks mainly inspire the idea behind ANN. In addition to "ANN," other names for them include "artificial neural systems," "parallel distributed processing systems," and "connectionist systems". ANN acquires a huge collection of units that are linked together in some way to permit communication between the units. These elements are crucial parallel processors, also known as nodes or neurons. Through a connecting link, every neuron is joined to every other neuron. A weight with knowledge of the input signal is connected to each connection link. The weight typically excites or suppresses the signal that is being conveyed. Therefore this information is the most helpful for neurons to resolve a specific issue. An activation signal is a name given to each neuron's internal state. Combining the input and activation rules results in output signals that can be delivered to other components.

6.3.6 Experimental details

We use the Sci-kit learn Pedregosa et al., 2011 in the experiment to import the regression models. We import MLP directly, where it uses Adam as an optimizer and ReLU as an activation function. The original dataset contains 102 beneficiary information, where we use 70% for training and 30% for testing. After that, we augmented the training dataset. We also use the Sci-kit learn machine learning library to import the various evaluation metrics. There are many ml models used in this work. Because of the nature of the problem, we use ml regression-based models discussed in the following sections. The limitation of the data is a significant constraint to not using the deep neural network architecture because it will be prone towards the overfitting scenario.

6.3.6.1 ML-based regression models

The nature of the dataset is regressive and has output labels corresponding to input features. Various machine learning algorithms are prone to deal with regressive data. Based on the nature of the data, we use multiple linear regression (MLR) Ottaviani and De Marco, 2022, regression tree (RT) Achenbach et al., 2022, random forest (RF) Wang et al., 2021 and finally, multi-layer perceptron regressor model (MLPRegressor) AL-Rousan et al., 2021. These algorithms can deal with the linearity and non-linearity of the data. In the MLR regression model, the relationship between two or more independent variables and a quantitative dependent variable is depicted by a straight line. A regression tree is a decision tree used for regression, predicting continuous-valued outputs rather than discrete outputs. The RF creates many randomly chosen decision trees from the input using ensemble learning techniques and the decision tree architecture. The algorithm then takes an average of the results to give a fresh prediction that is frequently accurate. MLP is an artificial neural network in which there are no cycles in the connections between the neurons. The typical MLP architecture has one hidden layer.

6.3.7 Evaluation metricise

In the regression task, generally, the three most common metricise are used to evaluate the performance of a model. These metrics are mean absolute error; root means square error and R^2 score value. The details of these three are given as follows.

Regression analysis issues are assessed using Mean Absolute Error (MAE). The total absolute value of each prediction error across all test set occurrences is used to establish a model's mean absolute error with regard to the test set. The difference between the instance's actual and expected values represents each prediction error.

$$MAE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}, \tag{6.3}$$

where $\hat{y}_i$ is the predicted value and y_i is the actual value, and n is the number of data.

The root mean square error, commonly referred to as the root mean square deviation, is one of the techniques most frequently used to evaluate the accuracy of forecasts. The difference between forecasts and actual measurements is shown to be Euclidean. It obtains the residual (difference between the forecast and reality), the residual's norm, and the residuals' mean for each data point. Then, calculate the root mean square error by taking the square root of that mean (RMSE).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}, \tag{6.4}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of data.

The amount of a dependent variable's variance that an independent variable or set of independent variables in a regression model can explain is expressed statistically as R-squared. R-squared measures how effectively the variance of one variable accounts for the variance of the other, as opposed to correlation, which represents the strength of the relationship between independent and dependent variables.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}, \tag{6.5}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the mean, and i varies from 1 to n.

6.4 Results

6.4.1 Ablation study

We conducted ablation research to assess the training procedure and the potency of various modules employed in our model. We experimented with many different scenarios, first training and testing our models with original and augmented data, normalized and without normalized data.

Performance of models on original data We used original data without preprocessing, such as augmentation and normalization. The MLPRegressor model training was done using the original 102 datasets. The Performance of the models is shown in Table 6.3. The Performance in

Table 6.4: Train and test case results augmented data without normalization for different ML approaches

	MLR	RT	RF	MLP
	Train|*Test*	*Train*|*Test*	*Train*|*Test*	*Train*|*Test*
MAE	0.9632|6.4981	0.4826|5.8723	2.5815|7.5517	0.8561|1.3729
RMSE	1.3870|9.3084	0.5342|7.4372	3.2979|9.3197	1.3227|3.0928
R^2	0.9894|0.4017	0.9983|0.5284	0.9309|0.2573	0.9917|0.9281

the form evaluation metrics score value shows it is lower than the Performance of the augmented data. The R^2 score value is lower in the case of original data for both the train and test case scenarios. In the case of MAE and RMSE errors, the value is higher than in the augmented case.

Performance of models on original with normalization Normalization of the data brings the value of input features and corresponding output values between 0 and 1. Table 6.3 shows the model's Performance; the MAE and RMSE values are lower for the training and test cases, but the R^2 score value is poor.

Performance of model with augmented and normalized data The Performance of the models on normalized and augmented data is better than the earlier two scenarios. In this case, the MAE and RMSE are much better than in the earlier case, and the R^2 score value is higher than in the earlier scenario. However, the Performance is not adequate.

Performance of model with augmented data The Performance of the augmented data is better than every three cases. We use augmented data without normalization to train the models in this case. After that, we use non-augmented original test case data to evaluate the model performance. The Performance in the form of R^2 score value is better than all other scenarios. Table 6.4 shows the results for this case for all the ML approaches and DNN models for the train and test case.

Based on the ablation study, we observed the effectiveness of the models with augmented data is better than the other three scenarios. Therefore we present detailed results for all ML models for this case scenario. We trained every ml model using the augmented data and tested the Performance using the original test case data. Table 6.4 depicts the evaluation metrics results for ml-based algorithms. The MAE, RMSE, and R^2 score value for the MLR algorithm during the test case is 6.4981, 9.3084, and 0.4017. The metrics score for the regression tree is 5.8723, 7.4372, and 0.5284. For the random forest algorithm, the scores are 7.5517, 9.3197, and 0.2573. The Performance of the all ml model is shown in Table 6.4. Out of these ml model, the results of MLPRegressor is better than other ml models. The results for the MLPRegressor are as follows. The RMSE, MAE, and R^2 score values are 1.2729, 3.0928, and 0.9281. The model's performance for the best case is 0.9281, which needs further improvement. The absence of enough data is the primary cause of this Performance. Therefore, we further collect more data to increase the efficiency of our model.

6.5 Conclusion and future works

This paper proposed a novel ML-based approach to predicting QoL dimensions value. This work details the creation and validation of an ML-based model that enables the unbiased evaluation of the QoL of ID people. A set of questionnaires, which cover the eight dimensions of the QoL, have been asked to ID people a professional during an interview between professional and ID people. Earlier professionals transformed the questionnaire's response into the eight dimensions value using the GENCAT scale. However, this work proposes to use a trained ml model to generate the eight dimensions value by getting 69 dimensions as input. However, in future work, we will approach predicting the QoL index value using the trained AI model instead of the GENCAT tool.

Chapter 7

Predicting the Quality of Life Index Value for determining the Requirement of Support Needs for Intellectually Disabled Individual using Machine Learning Methods

7.1 Summary

Recently, the number of cases of intellectually disabled people has been increasing. People with intellectual disability (ID) in their older age face many challenges that ordinary people do not. Each elder individual

needs a different kind to support to improve their lifestyle. The concept of quality of life (QoL) in intellectual disability provides a trait to improve the individual QoL by providing support. There is a set of 69 questionnaires related to every aspect of the individual life of an ID person. Based on the questionnaire response, professionals calculate the QoL index value, which decides Who needs support and who doesn't. So far, professionals have used the GENCAT tool to calculate the QoL index value by initially converting the 69 items' responses into the eight dimensions value and, after that, from the eight dimensions value to the QoL index. This process is lengthy and tedious and needs an expert to calculate the index value. This work proposes using a trained state-of-the-art machine learning (ML) based model to directly predict the QoL index value by inputting the response of the 69 items. This paper proposes different ML-based models using the original and augmented NewtonOne dataset. We choose the Ridge regression algorithm to predict the QoL index value based on evaluation metrics such as mean squared error, mean absolute error, root mean squared error, and R^2 score value. The test case value of various evaluation metrics is $1.6676, 3.8585, 1.9643$, and 0.9745 for the MAE, MSE, RMSE, and R^2 score, respectively, for the ridge regression model.

7.2 Introduction

In recent decades the population of intellectually disabled (ID) people has increased worldwide. ID people cover around 1 to 3 % of the global population Krahn, 2011. This increasing figure motivates researchers and other professionals to work in this area to improve the lifestyle of ID people. ID is a severe, widespread neurodevelopmental condition that has lifelong effects. It appears early stage of life till 18 years of life Patel et al., 2020. ID people show inefficiency in their intellectual and adaptive behaviour. ID is measured by using an intelligence quotient (IQ) level score. It could show itself in mild to profound ways. If a kid has an IQ score of less than 70, considered an ID kid McKenzie et al., 2016. ID differentiate based on the IQ score, such as if a kid's IQ is less than 70 and higher than 50, categorised as mild ID. Around 85% of the total ID population belongs to this category. These children can finish their education, develop independence, get training, and possibly pursue employment. A moderate IQ falls between 35 and 50. Despite having the potential to function independently, they frequently need care and attention. A total of 10% of the ID population belongs to this category. A severe intellectual disability exists in cases with IQ scores between 20 and 35. These people are less competent and struggle to understand numbers and reading. They require regular supervision. 4% of the total ID population belongs to this category. A profound ID instance is defined as having an IQ below 20. 1% of the total ID people population belongs to this category Wieczorek, 2018. The cause of ID is genetically inherited in some kids, whereas in some cases, it appears due to the effect of medication consumed to cure the disease at an early age. However, the cause of ID is unclear in some cases Oliveira et al., 2020.

Difficulties with social interaction, constrained interests, and repetitive behaviours characterize the neurological disease known as an autism spectrum disorder (ASD) Hodges, Fealko, and Soares, 2020. In comparison, ID shows limitations in intellectual functioning and adaptive behaviour. Around 35% of the ID population has ASD Bishop-Fitzpatrick and Rubenstein, 2019. Both diseases belong to the same neurodevelopmental disorder category but have different symptoms and cures. People in either category face many differences in their older age. The lifestyle of ID older people is challenging and complex due to many other age-related diseases arising in people Schepens, Van Puyenbroeck, and Maes, 2019b. Old age increases the chance of general psychiatric morbidity, anxiety disorder, dementia, and depression, according to studies concentrating on psychiatric conditions among older persons with ID. People with ID may be more susceptible to psychiatric disorders than the general population in younger and/or broader age groups Axmon et al., 2018. The most common diseases among ID people in their older age are dementia and psychiatric disorderTakenoshita et al., 2020. Dementia is more common among adults with ID in the age groups of 45 to 54, 55 to 64, and 65 to 74, with prevalence rates of 0.8%, 3.5%, and 13.9%, respectively Takenoshita et al., 2020. Due to these diseases, ID people suffer a lot more than the general population in their older age.

To support ID people to improve their lifestyle, they must analyse their quality of life (QoL). QoL is multidimensional, having etic (universal) and emic (cultural bound) components. It is affected by both personal and environmental circumstances and has objective and subjective components Grabowska et al., 2021. Improving QoL improves the three (personal, social and judicial) areas of ID people's life. These three areas are broad categories into the eight dimensions of QoL. Improving the QoL of a person requires improving these eight dimensions Yadav et al., 2022b. These eight dimensions are interpersonal relation (IR), emotional well-being (EW), physical well-being (PW), material well-being (MW), personal development (PD), social inclusion (IS), self-determination (SD), and rights (RI). These eight dimensions define the quality of life. Improvement in these shows the improvement in QoL. To measure the values of these eight dimensions using the response of questionnaires, Verdugo el proposed a tool known as the GENCAT scale Verdugo et al., 2010. The GENCAT scale calculates the eight dimension values using specified rules and following some tables. After that, using the eight dimensions value, the GENCAT tool uses to calculate the QoL index value. The QoL index value defines the need for support. QoL index value shows if a person needs support or not Yadav et al., 2022a, Gómez Sánchez et al., 2022. The index value is essential in improving an ID person's QoL. Calculating the index value using the GENCAT tool is tedious and requires knowledge. Generally, a professional needs to calculate the index value using the GENCAT tool. However, the GENCAT tool is efficient and accurate in calculating the index value, but it is an entirely statistical method and requires time and professionalism to calculate the index value.

Therefore, in this work, we propose to use an ML-based model to calculate the index value directly from questionnaire responses. Whereas earlier, professionals need to calculate the eight dimensions value from the response of the 69 questionnaires and then from eight dimensions value to the index value using the GENCAT tool. ML models learn the weights and biases parameters attached to the input features during training. After training, the trained ML-based models predict the QoL index value accurately. In this study, we trained ML-based models using the Newton One dataset. Various regression-based ML models are assessed and compared during training and testing. Some ML-based models' performances are close to each other.

- Unlike existing manual approaches for predicting QoL, we propose an accurate ML-based model for predicting QoL automatically in this paper.
- Providing comparative study between 15 ML-based models for QoL index prediction.

This paper is organised into five sections. Section two speaks about the background of the work. Whereas in section three, we show the methodology and experimental details of the paper. After that, section four depicts the results of the work. And finally, section five concludes our work and proposes future aspects related to this work.

7.3 Background

7.3.1 Quality of life and its dimensions

A person's QoL is a multifaceted phenomenon of fundamental domains influenced by personal and environmental factors. Although they may differ in relative worth and significance, these fundamental areas apply to everyone Gomez et al., 2012; Maestro-Gonzalez et al., 2018. The present method of measuring QoL is distinguished by: (1) its multidimensional nature, which includes fundamental areas and metrics; (2) the use of many methodologies, including both subjective and objective metrics; (3) the application of a systems viewpoint that takes into account the various contexts that have an impact on people at the meso, micro, and macrosystem levels; (4) the greater participation of those at risk of being socially excluded in the planning and execution processes Gomez et al., 2012. QoL has eight dimensions, interpersonal relation (IR), emotional well-being (EW), physical well-being(PW), social inclusion (SI), personal development (PD), material well-being (MW), rights (RI), and self-determination (SD). These eight dimensions cover three essential areas, the personal, the social, and the judicial area of life.

7.3.2 Neurodevelopmental disorder

A relatively broad range of disabilities that interfere in some way with brain development is called "neurodevelopmental" disabilities. People with neurodevelopmental disorders (NDDs) cannot meet developmental milestones in their cognitive, emotional, or physical growth Parenti et al., 2020. NDDs are frequently connected to interference with the well-coordinated events that result in brain development. In current society, NDDs are a severe health issue that affects $> 3\%$ of children worldwide Gilissen et al., 2014. They have a variety of etiologies and are associated with cognitive, communicative, behavioural, and psychomotor impairments. NDDs include intellectual disability (ID), epilepsy, autism spectrum disorder (ASD), and attention deficit hyperactivity disorder (ADHD)Tărlungeanu and Novarino, 2018. Various circumstances can bring early-onset NDDs, including extreme social deprivation, metabolic diseases, immunological disorders, genetic risk, infectious diseases, nutritional issues, physical trauma, and toxic and environmental variables

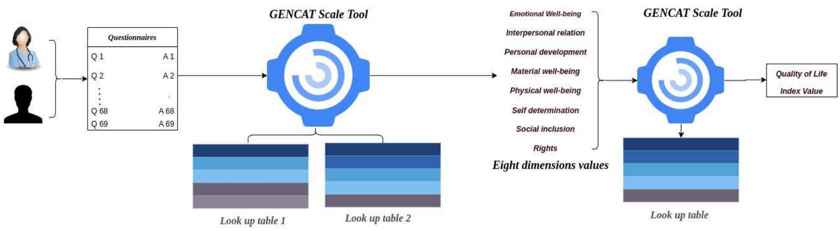


Figure 7.1: Figure shows the complete architecture to predict the quality of life index value using the Gencat tool. It starts with an interview conducted between the professional and the beneficiary. Professionals collect the response to the 69 questionnaires and, using the Gencat tool; he/she converts the response into the eight dimensions value. After that, using the eight-dimension value, professionals calculate the quality of life index value using the Gencat tool.

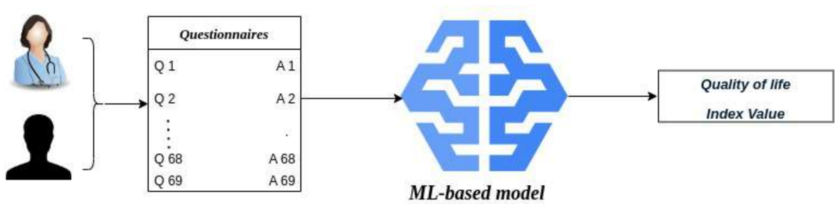


Figure 7.2: Figure shows the complete architecture of the proposed method. After the interview, a trained ML-based model takes the response of 69 items as input and predicts the quality of life index value.

Diester et al., 2015.

7.3.3 Intellectual disability in older people

The longer life expectancy of those with ID makes the requirement for age-specific support more urgent. Older ID individuals typically reside in residential services. Where caregivers give the best possible support to them to make their life easier. However, sometimes caregivers are challenged by their evolving requirements as they strive to provide the highest level of care Schepens, Van Puyenbroeck, and Maes, 2019a. On the other hand, ID people face many challenges in their personal and professional lives. Throughout their progressively longer lives, people with intellectual disabilities experience many unfavourable life events. According to reports, people with intellectual disabilities endure more traumatic, potentially upsetting life situations than average, not just when they're young but also throughout their much longer lifespans Berg et al., 2019. The prevalence of violent crime, abuse, neglect, and other forms of discrimination ranges from 30% to approximately 90% Baladerian, Coleman, and Stream, 2013. Such events may result in toxic stress, which then causes trauma. Trauma happens when a person's resources cannot handle stress Masten and Barnes, 2018.

7.3.4 The GENCAT tool

The GENCAT tool is proposed by Verdugo et al. Verdugo et al., 2010. The GENCAT scale is used to gauge beneficiaries' overall quality of life. This scale is created as a measurement tool to support social services' ongoing improvement. The multidimensional model on which GENCAT has based displays individual wants regarding eight key dimensions (EW, IR, PW, PD, MW, SD, SI, RI) Yadav et al., 2022a. The scale measures the quality of life using 69 questionnaires that all responded on a 4-point Likert scale, all declaratively presented in the third person. There are various scales in the GENCAT Scale depending on the type of population being studied: a general population scale, an elderly population scale (those who are 50 years of age or over), a scale for intellectually disabled persons, and a scale for other groups (e.g., physical disability, drug addicts, and mental health problems). We applied the scale for intellectually disabled issues Maestro-Gonzalez et al., 2018.

7.3.5 ML-based models

ML models are prone to learn from structured data. ML models are categorised into supervised, unsupervised, and reinforcement learning. The nature of our dataset is structured and labelled dataset. Data has 69 input features and corresponding index values as output labels. Therefore, supervised learning algorithms are the best fit for this dataset. In supervised learning, there are two categories: regression and classification. These two categories rely on the nature of output features. If the output feature is definite, then classification algorithms are helpful. In the case of regression, the nature of the output variable will be some numeric value that relies on the input features value. We use regression algorithms based on the nature of our data, where the output feature is an index value.

7.4 Methodology

This work proposes to use an ML-based model to predict the QoL index value directly from the response of the 69 items. Earlier, people used the GENCAT tool to calculate the index value, a traditional and statistical-based method with many rules and relatable tables involved in calculating the first eight dimensions value using 69 questionnaires response. After that, using the eight dimensions values, the professional again calculates the QoL index value using the GENCAT scale as shown in Fig. 7.1. This process of calculating the QoL index value from the 69 items response is very skills-required and needs a professional to calculate these values. We suggest a method based on learning to directly predict the QoL index value by taking 69 questionnaires response as shown in Fig. 7.2. We trained various regression-based algorithms using the NewtonOne dataset. We assessed the models' performance using error metrics, including mean absolute error, mean squared error, root mean squared error and R2 score value.

7.4.1 NewtonOne dataset

The NewtonOne dataset is a private dataset collected by professionals during the Never Alone project. This NeverAlone project is ongoing. Therefore, the number of data is increasing time after time. Currently, this data contains the QoL data of 113 persons. The category of people who participated in this data collection has a mild ID and an age group above 55. The demography of the data has 65% female participants and 35% male participants. One data contains 69 input features and corresponding QoL index values. The response of each 69 questionnaires is collected using four points Likert scale. Professionals during the interview asked questions to the participants and, based on the answers, gave the value as a response.

7.4.2 Data augmentation

Data is the backbone of machine learning models. For training an ML-based model, the amount of data is essential to the model's performance. Therefore, a substantial amount of data is required to train an ML-based model for optimal performance. To enhance the amount of data, the data augmentation approach is performed. Therefore, this technique is commonly used in case of less data. This work uses the NewtonOne data, which contains 113 original data points. We trained various ML-based; therefore, increasing the data as much as possible is demanded.

The model is trained through augmented data to improve performance. At the same time, the model evaluation requires the original data to justify the model's performance. Therefore, we split our dataset into train and test sets before augmentation. The ratio of the train and test is set as 70% and 30%. After that, we augmented the training dataset. Using the augmented train dataset, we train the ML-based and DNN models and test the model's performance using an original test set data.

In this experiment, we imported the ML algorithms using the Scikit-learn library Pedregosa et al., 2011. Implementation of the experiment done on Google Colaboratory Carneiro et al., 2018.

SMOGN Branco, Torgo, and Ribeiro, 2017 (Synthetic minority over-sampling technique for regression with random Gaussian noise) is proposed for the imbalanced distribution regression task. SMOTE-R Torgo et al., 2013 and adding Gaussian Noise are two oversampling techniques used with random under-sampling in SMOGN. When a high risk is associated with using SMOTE-R to generate a new example, a more cautious alternative is to generate the new example by adding Gaussian noise. The distance between the base scenario and the neighbour case, to be more precise, is what determines which approach to utilize. SMOTE-R is employed if the neighbour case is close enough; otherwise, Gaussian Noise generates the new case. Additionally, SMOGN resamples the dataset using both over and under-sampling Song, Dao, and Branco, 2022.

7.4.3 Quality of life index value

The quality of life index value is essential in providing support to intellectually disabled people Verdugo et al., 2012. This value decides whether a person needs support or not. Therefore the role of the index value is very important in providing support. The index value follows the Gaussian nature, and the 80% of the population lies between the Gaussian bell-shaped curve, whose values vary from a minimum of 68 to 130. Most of the population will come under these values. The mean of the curve lies at 100. therefore if the index value of a beneficiary is 100 or above 100, then he/she does not need any support. If the index value is below 100, she/he needs support. Professionals use the GENCAT tool to calculate the index value in steps. They calculate the first eight dimension values from the 69 questions and then the index value using the eight dimensions value. In this work, we propose to use an ML-based model to predict the index value directly using 69 items responses.

7.4.4 Developing ML-based QoL prediction models

Even though classic machine learning techniques have made tremendous progress in knowledge discovery, they sometimes struggle to perform well when faced with complex data, such as unbalanced, highly dimensional, noisy data, etc. The cause of this is that various qualities and the underlying structure of the data are complex for such methods to represent Dong et al., 2020. A set of features are first extracted through ensemble learning using various transformations. Various learning algorithms are used to get mediocre predicted outcomes based on these learned properties. Compared to a single model, this strategy enables more excellent predictive performance.

A variation of the random forest method, very randomised trees, also known as the extra trees regressor (ETR) Geurts, Ernst, and Wehenkel, 2006, is a relatively new machine learning technique less prone to overfit a dataset. Each base estimator is trained using a random selection of characteristics using the same process as a random forest in ETR. To split the node, it instead arbitrarily chooses the best characteristic and its matching value. Each regression tree is trained by ET using the entire training dataset Ahmad, Reynolds, and Rezgui, 2018.

To overcome the drawbacks of the conventional Classification and Regression Tree (CART) method, a random forest (RF) Speiser et al., 2019 was developed. This ensemble method is based on trees. To reduce the model's bias and variance concurrently, RF comprises numerous weak decision tree learners built in tandem Ahmad, Reynolds, and Rezgui, 2018. To train the model, RF employs a bootstrap replica. N bootstrapped sample sets are used from the source dataset to train an RF.

Gradient boosting regressor (GBR) Friedman, Hastie, and Tibshirani, 2000 is an error function-based optimization algorithm. By Using weak prediction models, like the decision tree, creates a robust prediction model. The fundamental notion is that a basic model performs each calculation and that with each subsequent calculation, the residual of the previous model is reduced, resulting in the creation of a new basic model that is directed in the direction of the gradient and has reduced residuals Wei et al., 2019. Therefore, the weight of a weak learner can be continuously improved and modified to make it a strong learner, lowering and optimizing the loss function.

Instance-based learning is used to implement K-Nearest Neighbors (KNN) Ahmed, Seraj, and Islam, 2020 with a value of k. This approach is used to classify test data based on test patterns and estimate the density function of the distribution of test data. The first step in using this technique is measuring the difference between the training and test data sets. The distance between training and test data sets is frequently calculated using the Euclidean distance. After finding the distance between the data (minimum similarity), the database samples were organised ascendingly from the smallest distance (maximal similarity) to the most significant distance (Euclidean distance). The next stage in this paradigm is finding the experiment's sample size (k) to estimate the intended database's characteristics. The model's peculiar single points have an impact on the outcomes. It may be conceivable to incorporate points from other classes inside the targeted range if k is considered significant even though k is supposed to be tiny.Shabani et al., 2020. Cross-validation is often used to find the ideal value for k.

Bayesian regression enables a natural process to persist in sparse or uneven data. Instead of using point estimates, it formulates linear regression utilizing probability distribution. The result is anticipated to be chosen from a probability distribution instead of evaluated as a single value M. Mostafa et al., 2020. Regularized linear regression is the basis of the Bayesian algorithm. The approach uses a hyperparameter to adjust regularisation strength to fully integrate over the hyperparameter inside the posterior distribution and apply a roughly noninformative hyperprior. Because Bayesian ridge regression is a subset of Bayesian linear regression and is a member of the ridge regression family, it possesses all the traits of ridge regression and Bayesian linear regression. Yang and Yang, 2020.

A "linear regression" model assumes that the input and output variables have a linear relationship. The relationship between the input and output variables can be seen as a line with a single input variable and a hyperplane linking them in higher dimensions. Using an optimisation process, the model's coefficients are chosen to reduce the overall squared error between predicted and actual values. A statistical method known as multiple linear regression (MLR) employs several input variables to forecast the result of an output variable Uyanık and Güler, 2013.

Regularization and feature selection are two critical tasks that can be accomplished with the help of the Least Absolute Shrinkage and Selection Operator (Lasso) Shafiee et al., 2021. The LASSO method constrains some of the model's absolute parameter values. The total must fall below a certain threshold (upper bound). The method employs a shrinkage (or regularization) mechanism that penalizes the regression variables' coefficients, decreasing some to zero. During the feature selection stage, the variables with non-zero coefficients that the shrinking technique left are picked to be a component of the model. The method's objective is to reduce prediction error Shafiee et al., 2021.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2 + \lambda\sum_{j=1}^{n}|\beta_j|, \tag{7.1}$$

where $\frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2$ is the mean squared loss function and $\lambda\sum_{j=1}^{n}|\beta_j|$ is the penalty $L1$ regularization.

Ridge regression is a refinement of conventional least squares that extends the optimization issue by including a regularization term Tour et al., 2022. The regularization hyperparameter in ridge regression needs to be determined from the data. If the predicted value is too low, the model will tend to overfit the noise in the data, and if it is too high, the model will not forecast as well as it could. Usually, the regularization parameter is chosen using the grid search with a cross-validation approach.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2 + \lambda\sum_{j=1}^{p}{\beta_j}^2, \tag{7.2}$$

where $\frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2$ is the mean squared loss function and $\lambda\sum_{j=1}^{p}{\beta_j}^2$ is the penalty $L2$ regularization.

Huber regression is a robust regression that gives outliers in a dataset less weight than other data and considers the potential of outliers Sun, Zhou, and Fan, 2020. The critical component in this regression is ϵ. Smaller values of the ϵ argument, which determine how much of the data is deemed an outlier, strengthen the model's resistance to outliers. We imported this regression from sci-kit learn, and the default value of the ϵ is 1.35.

Elastic net is a linear regression model. It uses the lasso and ridge penalties to penalize regression models to make them more consistent Liu et al., 2018. The approach combines ridge and lasso regression techniques, taking into account their limitations in improving the regularization of statistical models.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2 + r\lambda\sum_{j=1}^{n}|\beta_j| + \frac{1-r}{2}\lambda\sum_{j=1}^{n}{\beta_j}^2, \tag{7.3}$$

where $\frac{1}{m}\sum_{i=1}^{m}(Y_i - \hat{Y}_i)^2$ is mean squared loss function, $r\lambda\sum_{j=1}^{n}|\beta_j|$ is the penalty ($L1$ regularisation), and $\frac{1-r}{2}\lambda\sum_{j=1}^{n}{\beta_j}^2$ is the penalty ($L2$ regularization).

Orthogonal matching pursuit (OMP) is a greedy algorithm. It is a straightforward, iterative strategy for solving the sparse recovery of high-dimensionality vector problems. A priori knowledge of either the noise statistics or the sparsity of the regression vector is necessary for OMP to operate at its best. In sparse regression, the purpose is to identify the best sparse vector for minimizing a specified objective function. Since sparse models provide higher generalization assurances when the feature dimension is large or the data is slight, sparse regression is a crucial issue in statistical machine learning Somani et al., 2018.

Adaptive boosting regressor (Ada), a sequential ensemble machine learning technique, combines several weak learners randomly to develop a strong learner using the dataset. Weak learners are produced by using machine learning algorithms. Each hypothesis is taught using the weights assigned to each sample observation in a training dataset. The incorrect predictions are logged, and the subsequent base learner is given them with a focus on this inaccuracy. The exact steps are repeated until the algorithm can adequately classify the output. When using regression, an output instant's absolute value error, which might be any constant, determines whether it is accurate or incorrect, Shanmugasundar et al., 2021.

A decision tree (DT) is the most popular ML technique for handling classification and regression problems Chen et al., 2020b. As suggested by its name, the algorithm predicts the desired value using a model of decisions that resembles a tree. The splitting process starts at the root node and moves through a branching tree to reach the leaf node containing the prediction or algorithmic result. Typically, a decision tree is constructed from the top down, choosing the variable that best separates the set of objects at each stage. Regression trees are decision trees with either a continuous target variable or a continuous terminal node.

Least Angle Regression (LAR) is a technique used in regression for high dimensional data. LAR is similar to forward stepwise regression. Because LAR employs data comprising multiple qualities, it may identify the property most strongly correlated with the goal value at each

Table 7.1: Test case results of original data for different ML algorithms

No	ML Models	MAE	RMSE	R^2 score	Variable parameters
1	Extra trees regressor (ETR)	5.4073	48.7362	0.6785	num of estimators= 20
2	Random forest regressor (RFR)	5.1991	6.9963	0.6771	$maxdepth = 15$
3	Gradient boosting regressor (GBR)	5.7362	8.2507	0.5509	
4	K neighbors regressor (KNN)	7.4926	9.1553	0.4471	
5	Bayesian ridge (BR)	1.7648	2.1664	0.9690	
6	Linear regression (LR)	2.3722	3.1927	0.9327	
7	Lasso regression (Lasso)	1.9504	2.4140	0.9615	$alpha = 0.01$
8	Ridge regression (Ridge)	1.6667	1.9643	0.9745	
9	Huber regressor (Huber)	4.3128	5.4135	0.8066	
10	Elastic net (EN)	3.1787	3.9597	0.8965	
11	Orthogonal matching pursuit (OMP)	7.0742	8.4859	0.5250	
12	AdaBoost regressor (Ada)	5.4395	6.7094	0.7030	num estimators=300
13	Decision tree regressor (DT)	10.1764	14.2663	-0.3425	
14	Least angle regression (LAR)	5.2143	5.9376	0.7674	num nonzero coefs = 30
15	Passive aggressive regressor (PAR)	5.7502	6.7163	0.7024	

stage. The correlation may exist across multiple attributes. LAR considers the qualities in this situation and proceeds perpendicularly to the attributes. That is why this method is referred to as least angle regression. Rubio et al., 2020.

Passive aggressive regressor (PAR) is used for large-scale learning Crammer et al., 2006. It is a particular algorithm used in online learning. Online machine learning techniques update the model using input data in sequential order instead of batch learning, which simultaneously uses the entire training dataset. This is especially helpful when the data volume is significant, and training the entire dataset would not be easy. An algorithm for online learning will gather a training example, update the classifier, and then discard the example.

7.5 Results

7.5.1 Evaluation metrics

Error metrics are used to assess the model's performance in the regression process. These error metrics are root mean square error (RMSE), mean squared error (MSE), mean absolute error (MAE), and R^2 score.

Table 7.2: Test case results augmented data for different ML algorithms

No	ML Models	MAE	RMSE	R^2 score	Variable parameters
1	Extra trees regressor (ETR)	5.8897	7.2597	0.6865	num estimators=20
2	Random forest regressor (RFR)	5.9216	7.1776	0.6936	max depth=15
3	Gradient boosting regressor (GBR)	5.1671	6.1371	0.7760	
4	K neighbors regressor (KNN)	7.2669	8.7813	0.5414	
5	Bayesian ridge (BR)	2.0681	2.7359	0.9554	
6	Linear regression (LR)	3.8803	5.3497	0.8298	
7	Lasso regression (Lasso)	2.0397	2.5516	0.9612	$alpha = 0.01$
8	Ridge regression (Ridge)	1.5103	1.9956	0.9763	
9	Huber regressor (Huber)	4.1738	5.5782	0.8149	
10	Elastic net (EN)	2.6404	3.7010	0.9185	
11	Orthogonal matching pursuit (OMP)	11.0497	12.8025	0.0253	
12	AdaBoost regressor (Ada)	5.7854	7.5685	0.6593	num estimators=300
13	Decision tree regressor (DT)	10.7058	13.1507	−0.0284	
14	Least angle regression (LAR)	4.6589	5.8079	0.7994	num nonzero coefs=30
15	Passive aggressive regressor (PAR)	3.4341	4.7778	0.8642	

MAE measures the mean of the absolute difference between the ground truth value and the model's predicted value.

$$MAE = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n}, \tag{7.4}$$

here, $\hat{Y}_i$ is the predicted value and Y_i is the ground truth value. i varies from 1 to n. MSE measures the degree of inaccuracy in regression models. It calculates the average squared difference between the observed and projected values.

$$MSE = \frac{\sum (Y_i - \hat{Y}_i)^2}{n}, \tag{7.5}$$

where $\hat{Y}_i$ predicted value, Y_i is the ground truth value, and n is the number observation.

The standard deviation of the residuals is RMSE. Residuals and RMSE measure the spread of these residuals, which is the distance between the data points and the regression line. Put another way, it indicates how closely the data is centred on the line of best fit. RMSE is commonly used in regression analysis, forecasting, and climatology to verify the accuracy of experimental results.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}, \tag{7.6}$$

where, Y_i is the actual value and $\hat{Y}_i$ is the predicted value.

R-squared measures a linear regression model's ability to fit data. This statistic demonstrates the extent to which the independent variables may collectively explain the variation of the dependent variable. R-squared uses a proper 0–100 % scale to concisely quantify the strength of the link between the model and the dependent variable.

$$R^2 Score = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y}_i)^2}, \tag{7.7}$$

where, Y_i is the actual value, $\hat{Y}_i$ is the predicted value, $\bar{Y}$ is the mean value.

7.5.2 Ablation study

We experimented with two ways based on the availability of the data. We trained the model using the augmented and original data and reported the results for both in the tables 7.1 and 7.2

Table 7.1 shows the various ML-based model performance for the test case result of the original dataset. The original dataset contains 113 beneficiary's responses to 69 items and corresponding index values. We divided the original data into 30% (34) for testing and 70% (79) for training. The results are shown in the form of four evaluation metrics. The variable parameters show a change in variable values to the default value. In other models' cases, we directly imported the model. We did not change their variable because we achieved the best results from the default value, whereas in other cases, we achieved better results when we used the defined variable by us.

Table 7.2 depicts the various ML-based model performance for the augmented data. Initially, the data contained a total of 113 beneficiary

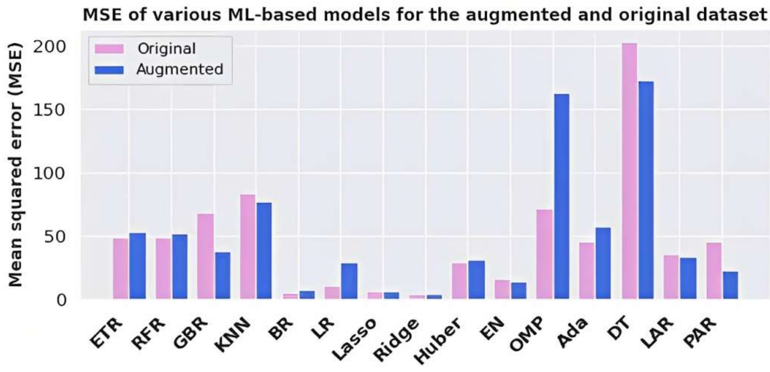


Figure 7.3: Test case mean squared error of various machine learning models for the augmented and original dataset

information. Before augmentation of the original data, We divided the data into the proportions of 70% train and 30% test. The train data contains 79 data, and we use the SMOGN algorithm to augment the train data. At the same time, we do not augment the test data. After the augmentation, the total data point in the data increases to 150. So, in this case, we trained various ML-based models using 150 data and tested the model's performance using 34 data.

Tables 7.1, 7.2 show the results of the two case data of the various ML-based model in the form of MAE, RMSE, and R^2 score value. We imported various ensemble based ML-models to learn the correlation between the 69 questionnaire responses and the corresponding index value. For both cases, the performance of the three models is higher than the others. Bayesian ridge, Lasso regression and Ridge regression have the minimum error and the highest R^2 score value. The evaluation metrics score for Bayesian ridge for the test case is 1.7648, 2.1664, and 0.9690 as MAE, RMSE, and R^2 scores, respectively, in the case of original data. The results of Lasso regression for the test case are $1.9504, 2.4140, and 0.9615$ as MAE, RMSE, and R^2 scores, respectively, in the case of original data.

And the result of the Ridge regression for the test case scenario is $1.6667, 1.9643$, and 0.9745 as MAE, RMSE, and R^2 scores, respectively, in the case of original data. The performance of these three models on original data is as good as or even better than augmented case results. The performance of the other models, like random forest regressor, gradient boosting, passive-aggressive regressor and others, is better in the augmented data case. In both instances, the Ridge regression model outperforms the other 14 models. The ridge regression results in augmented data cases are $1.5103, 1.9956$, and 0.9763 as MAE, RMSE, and R^2 scores, respectively.

The quantity of datasets affects how well each model performs. In our case, we have limitations in the original dataset. And one can not augment the data more because the augmented data may lose the originality of the distribution. Therefore, increasing the original data can improve the model's performance.

7.6 Conclusion

This paper proposed a novel approach for predicting the QoL index, and it solves the limitations of the existing manual approach. We use the NewtonOne dataset to train the various state-of-the-art machine learning models. So far, professionals have used the GENCAT scale to calculate the QoL index by taking first the response of the 69 questionnaires and, after that calculating the eight dimensions value and from eight dimensions value to the QoL index value, which is a lengthy and tedious task and requires a skilful professional to calculate the index value. We propose using a trained ML-based model which accurately predicts the QoL index value by inputting the response of the 69 questionnaires. In the traditional approach calculating the index value is a two-step process in which first calculates, the eight dimension value from the 69 questionnaires response and after that from eight dimension value to the QoL index value. In this approach, we directly calculate the index value.

Out of the various regression model, the performance of the Ridge regression model is superior in terms of the MAE, RMSE, and R^2 score value. Therefore we finalize using this algorithm to predict the QoL index value.

We will incorporate more data in the NewtonOne dataset in future work to improve our model performance and add sensor-based information affecting the QoL of the individual together with questionnaires response.

Part IV

Concluding Remarks and Future works

Chapter 8

Concluding remarks

8.1 Summary of contributions

After extensive research and analysis, this PhD work has contributed valuable insights and knowledge to the ID older people care by developing more precise models to predict accurate future human motion and developing a support system to improve the QoL of ID older people by providing a support report containing required actions to improve the deficit dimensions of QoL. There are two sections to this thesis. The first part of this study has provided a deeper understanding of human motion prediction for the in-distribution and out-of-distribution scenarios. They have implications for advancing the prediction ability of collaborative robots. The second part discusses the support systems to analyze the QoL's dimensions to improve the QoL of ID older people and other dependent people.

Chapters 1 and 2 discuss the introduction and the analysis of related research work in both areas. Chapter 3,4 contains the first part of the work, where we discussed our proposed models to predict more precise future human motion. In Chapter 3, we propose to predict future human motion by observing past motion. We use the inception residual blocks to learn the temporal feature and the Convolutional Graph Network to learn the spatial features. Using this method, we can produce future poses that are more precise than those produced by other cutting-edge methods. In Chapter 4, we empower our human motion prediction model to predict using even out-of-distribution(OoD) data by augmenting discriminative and generative models with regularization using linear matrices. This proposed network has shown much-improved results for both out-of-distribution and in-distribution scenarios.

In chapters 5, 6, and 7, we discussed the second part of the thesis. Where, Through rigorous data collection and further analysis, we move towards improving the QoL of dependent individuals with IDs in guardianship entities with the help of artificial intelligence techniques. More specifically, in Chapter 5, we proposed our ML-based models are trained and tested to predict the index value, standard scale value, and support intensity scale value related to the QoL. We evaluate each of the three models using various evaluation metrics of different ML algorithms calculating each task's performance. Subsequently, the proposed method generates a support report containing the required actions to improve the deficit dimensions and QoL. It helps professionals to track the patient's progress. In Chapter 6, models based on machine learning are used to forecast the values of eight dimensions from the results of sixty-nine questionnaires. Previously, professionals used the GENCAT tool to calculate these eight dimension values, a statistical model that uses a set of rules and look-up tables. This process is cumbersome, requiring time and expertise to calculate the eight-dimensional values. In Chapter 7, we use various machine learning algorithms to predict the quality of life index value directly from the sixty-nine items response instead of using the GENCAT tool to calculate it in two steps-calculating the values of eight dimensions and then from the values of eight dimensions to the index value.

8.2 Future Research lines

These findings have important implications for industry, which are developing an intelligent robot for the caring field, and researcher working in human motion prediction, older people care, and intellectual and developmental disorder. We think that these are intriguing and significant areas of study. Throughout this investigation, several future research directions have been discovered. Some of the future research lines are given as follows:

1. Future research may be directed towards effective robotization of the caregiving process; the human motion prediction model may be modified towards perfection using the suggested method to address the uncertainties while producing long-term forecasts for both OoD and ID scenarios.

2. Implementing the prediction models practically on real-world robots to provide them with the prediction ability to work with humans in working place.

3. Towards using sensory data obtained from the guardianship entities, such as data from sensors that monitor the dependent individuals' activities to increase the suggested approach's efficacy.

4. Building an augmented support system which contains both the environmental and individual aspects of QoL to enhance the lifestyle of the elderly ID people.

5. Exploring the proposed idea for other neurodevelopmental diseases like Autism, Dementia, Attention-Deficit/Hyperactivity Disorder etc, to help people to improve their quality of life.

This PhD work has significantly contributed to the literature on human motion prediction, intellectual disability, and older people care. Moreover, it has advanced our understanding of machine learning, deep learning, and neurodevelopmental disorder. It is hoped that the findings of this study will be used to inform policy decisions, improve motion prediction ability, provide proper support to ID elderly and inspire further research and inquiry into this critical area of study.

Bibliography

Abdulwahab, Saddam, Hatem A Rashwan, Miguel Angel Garcia, Mohammed Jabreel, Sylvie Chambon, and Domenec Puig (2020). "Adversarial Learning for Depth and Viewpoint Estimation from a Single Image". In: *IEEE Transactions on Circuits and Systems for Video Technology*.

Abebe, Teshome Hailemeskel (2019). "The Derivation and choice of appropriate test statistic (z, t, f and chi-square test) in research methodology". In: *Journal of Mathematics Letters* 5.3, pp. 33–40.

Achenbach, Peter, Markus Hippich, Jose Zapardiel-Gonzalo, Beate Karges, Reinhard W Holl, Agnese Petrera, Ezio Bonifacio, Anette-G Ziegler, DiMelli Study group, DPV Study group, et al. (2022). "A classification and regression tree analysis identifies subgroups of childhood type 1 diabetes". In: *EBioMedicine* 82, p. 104118.

Ahmad, Muhammad Waseem, Jonathan Reynolds, and Yacine Rezgui (2018). "Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees". In: *Journal of cleaner production* 203, pp. 810–821.

Ahmed, Mohiuddin, Raihan Seraj, and Syed Mohammed Shamsul Islam (2020). "The k-means algorithm: A comprehensive survey and performance evaluation". In: *Electronics* 9.8, p. 1295.

Al-Khafajiy, Mohammed, Thar Baker, Carl Chalmers, Muhammad Asim, Hoshang Kolivand, Muhammad Fahim, and Atif Waraich (2019). "Remote health monitoring of elderly through wearable sensors". In: *Multimedia Tools and Applications* 78.17, pp. 24681–24706.

AL-Rousan, Nadia, Nor Ashidi Mat Isa, Mohammad Khairunaz Mat Desa, and Hazem AL-Najjar (2021). "Integration of logistic regression and multilayer perceptron for intelligent single and dual axis solar tracking systems". In: *International Journal of Intelligent Systems* 36.10, pp. 5605–5669.

Aldhubri, Ali, Yu Lasheng, Farida Mohsen, and Majjed Al-Qatf (2021). "Variational autoencoder Bayesian matrix factorization (VABMF) for collaborative filtering". In: *Applied Intelligence* 51.7, pp. 5132–5145.

Alftberg, Åsa, Maria Johansson, and Gerd Ahlström (2021). "Ambivalence among staff regarding ageing with intellectual disabilities: Experiences and reflections". In: *Journal of Intellectual Disabilities* 25.2, pp. 192–209.

Aliakbarian, Sadegh, Fatemeh Sadat Saleh, Mathieu Salzmann, Lars Petersson, and Stephen Gould (2020). "A stochastic conditioning scheme for diverse human motion prediction". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5223–5232.

Ansari, Yasmin, Mariangela Manti, Egidio Falotico, Yoan Mollard, Matteo Cianchetti, and Cecilia Laschi (2017). "Towards the development of a soft manipulator as an assistive robot for personal care of elderly people". In: *International Journal of Advanced Robotic Systems* 14.2, p. 1729881416687132.

Axmon, A, P Björne, L Nylander, and G Ahlström (2018). "Psychiatric diagnoses in older people with intellectual disability in comparison with the general population: a register study". In: *Epidemiology and psychiatric sciences* 27.5, pp. 479–491.

Baladerian, Nora J, Thomas F Coleman, and Jim Stream (2013). "Abuse of people with disabilities". In: *Spectrum Institute Disability and Abuse Project*.

Bemelmans, Roger, Gert Jan Gelderblom, Pieter Jonker, and Luc De Witte (2012). "Socially assistive robots in elderly care: a systematic review into effects and effectiveness". In: *Journal of the American Medical Directors Association* 13.2, pp. 114–120.

Bentéjac, Candice, Anna Csörgő, and Gonzalo Martínez-Muñoz (2021). "A comparative analysis of gradient boosting algorithms". In: *Artificial Intelligence Review* 54.3, pp. 1937–1967.

Berg, Kristin L, Cheng-Shi Shiu, Rebecca T Feinstein, Kruti Acharya, Jonathan MeDrano, and Michael E Msall (2019). "Children with developmental disabilities experience higher levels of adversity". In: *Research in Developmental Disabilities* 89, pp. 105–113.

Biau, Gérard (2012). "Analysis of a random forests model". In: *The Journal of Machine Learning Research* 13.1, pp. 1063–1095.

Bigby, Christine (2004). *Ageing with a lifelong disability: A guide to practice, program, and policy issues for human services professionals*. Jessica Kingsley Publishers.

Bishop-Fitzpatrick, Lauren and Eric Rubenstein (2019). "The physical and mental health of middle aged and older adults on the autism spectrum and the impact of intellectual disability". In: *Research in Autism Spectrum Disorders* 63, pp. 34–41.

Bond, Rebecca J and Jenni Hurst (2010). "How adults with learning disabilities view living independently". In: *British Journal of Learning Disabilities* 38.4, pp. 286–292.

Bourached, Anthony, Ryan-Rhys Griffiths, Robert Gray, Ashwani Jha, and Parashkev Nachev (2020). "Generative Model-Enhanced Human Motion Prediction". In: *Applied AI Letters*.

Bourached, Anthony, Ryan-Rhys Griffiths, Robert Gray, Ashwani Jha, and Parashkev Nachev (2022). "Generative model-enhanced human motion prediction". In: *Applied AI Letters* 3.2, e63.

Branco, Paula, Luís Torgo, and Rita P Ribeiro (2017). "SMOGN: a preprocessing approach for imbalanced regression". In: *First international* *workshop on learning with imbalanced domains: Theory and applications*. PMLR, pp. 36–50.

Butepage, Judith, Michael J Black, Danica Kragic, and Hedvig Kjellstrom (2017). "Deep representation learning for human motion prediction and classification". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6158–6166.

Bütepage, Judith, Hedvig Kjellström, and Danica Kragic (2018). "Anticipating many futures: Online human motion prediction and generation for human-robot interaction". In: *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, pp. 4563–4570.

Butepage, Judith, Hedvig Kjellstrom, and Danica Kragic (2019). "Predicting the what and how-a probabilistic semi-supervised approach to multi-task human activity modelling". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 0–0.

Camacho, Luís, Georgios Douzas, and Fernando Bacao (2022). "Geometric SMOTE for regression". In: *Expert Systems with Applications*, p. 116387.

Carneiro, Tiago, Raul Victor Medeiros Da Nóbrega, Thiago Nepomuceno, Gui-Bin Bian, Victor Hugo C De Albuquerque, and Pedro Pedrosa Reboucas Filho (2018). "Performance analysis of google colaboratory as a tool for accelerating deep learning applications". In: *IEEE Access* 6, pp. 61677–61685.

Carrión-Nessi, Fhabián S, María V Marcano-Rojas, Diana C Freitas-DeNobrega, Sinibaldo R Romero Arocha, Allen W Antuarez-Magallanes, and Yurilís J Fuentes-Silva (2022). "Validación del LupusQoL en Venezuela: una medida específica de la calidad de vida en pacientes con lupus eritematoso sistémico". In: *Reumatología Clínica* 18.6, pp. 355–360.

Chattaraman, Veena, Wi-Suk Kwon, Juan E Gilbert, and Kassandra Ross (2019). "Should AI-Based, conversational digital assistants employ social-or task-oriented interaction style? A task-competency and reciprocity perspective for older adults". In: *Computers in Human Behavior* 90, pp. 315–330.

Chawla, Nitesh V, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer (2002). "SMOTE: synthetic minority over-sampling technique". In: *Journal of artificial intelligence research* 16, pp. 321–357.

Chen, Ting, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton (2020a). "A simple framework for contrastive learning of visual representations". In: *International conference on machine learning*. PMLR, pp. 1597–1607.

Chen, Wei, Yang Li, Paraskevas Tsangaratos, Himan Shahabi, Ioanna Ilia, Weifeng Xue, and Huiyuan Bian (2020b). "Groundwater spring potential mapping using artificial intelligence approach based on kernel logistic regression, random forest, and alternating decision tree models". In: *Applied Sciences* 10.2, p. 425.

CMU Graphics Lab Motion Capture Database (n.d.). URL: http://mocap.cs.cmu.edu/.

Cohen, Joel E et al. (2001). "World population in 2050: assessing the projections". In: *Conference Series-Federal Reserve Bank of Boston*. Vol. 46. Federal Reserve Bank of Boston; 1998, pp. 83–113.

Conrad, Jordan A (2020). "On intellectual and developmental disabilities in the United States: A historical perspective". In: *Journal of Intellectual Disabilities* 24.1, pp. 85–101.

Crammer, Koby, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer (2006). "Online passive aggressive algorithms". In.

Cruttwell, Geoffrey SH, Bruno Gavranović, Neil Ghani, Paul Wilson, and Fabio Zanasi (2022). "Categorical foundations of gradient-based learning". In: *European Symposium on Programming*. Springer, Cham, pp. 1–28.

Cui, Qiongjie, Huaijiang Sun, Yue Kong, Xiaoqian Zhang, and Yanmeng Li (2021). “Efficient human motion prediction using temporal convolutional generative adversarial network”. In: *Information Sciences* 545, pp. 427–447.

Cui, Qiongjie, Huaijiang Sun, Yupeng Li, and Yue Kong (2019). “A Deep Bi-directional Attention Network for Human Motion Recovery.” In: *IJCAI*, pp. 701–707.

Dafrallah, Safaa, Aouatif Amine, Stéphane Mousset, and Abdelaziz Bensrhair (2021). “Monocular pedestrian orientation recognition based on Capsule Network for a Novel Collision Warning System”. In: *IEEE Access* 9, pp. 141635–141650.

Dang, Lingwei, Yongwei Nie, Chengjiang Long, Qing Zhang, and Guiqing Li (2021). “Msr-gcn: Multi-scale residual graph convolution networks for human motion prediction”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 11467–11476.

Dawson, F, S Shanahan, E Fitzsimons, G O’Malley, N Mac Giollabhui, and J Bramham (2016). “The impact of caring for an adult with intellectual disability and psychiatric comorbidity on carer stress and psychological distress”. In: *Journal of Intellectual Disability Research* 60.6, pp. 553–563.

Desideri, Lorenzo, Giulio Lancioni, Massimiliano Malavasi, Arianna Gherardini, and Lisa Cesario (2021). “Step-instruction technology to help people with intellectual and other disabilities perform multistep tasks: a literature review”. In: *Journal of Developmental and Physical Disabilities* 33.6, pp. 857–886.

Diester, I, F Hefti, IM Mansuy, A Pascual-Leone, TW Robbins, LL Rubin, A Sawa, M Wernig, G Dölen, SE Hyman, et al. (2015). “Translational Neuroscience: Toward New Therapies”. In.

Disease Control, Centers for and Prevention (CDC) (2022). “Aging with a Disability”. In.

Doersch, Carl (2016). “Tutorial on variational autoencoders”. In: *arXiv preprint arXiv:1606.05908*.

Dollar, Chad A, Laura D Fredrick, Paul A Alberto, and Jaye K Luke (2012). "Using simultaneous prompting to teach independent living and leisure skills to adults with severe intellectual disabilities". In: *Research in developmental disabilities* 33.1, pp. 189–195.

Dong, Xibin, Zhiwen Yu, Wenming Cao, Yifan Shi, and Qianli Ma (2020). "A survey on ensemble learning". In: *Frontiers of Computer Science* 14.2, pp. 241–258.

Dutt, Muhammad Ibrahim and Wala Saadeh (2022). "A Multilayer Perceptron (MLP) Regressor Network for Monitoring the Depth of Anesthesia". In: *2022 20th IEEE Interregional NEWCAS Conference (NEWCAS)*. IEEE, pp. 251–255.

Economic, United Nations Department of and Population Division Social Affairs (2022). *World population prospects 2022: Summary of results*.

Espinoza, Jose Luis Vazquez, Alexander Liniger, Wilko Schwarting, Daniela Rus, and Luc Van Gool (2022). "Deep Interactive Motion Prediction and Planning: Playing Games with Motion Prediction Models". In: *Learning for Dynamics and Control Conference*. PMLR, pp. 1006–1019.

Fang Wang Zhibin Li, Chun-Yi Su (2019). "Human Motion Prediction for Assistive Robots in Elder Care". In: *Journal of Healthcare Engineering*.

Fernández, Alberto, Salvador Garcia, Francisco Herrera, and Nitesh V Chawla (2018). "SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary". In: *Journal of artificial intelligence research* 61, pp. 863–905.

Fiorini, Laura, Marleen De Mul, Isabelle Fabbricotti, Raffaele Limosani, Alessandra Vitanza, Grazia D'Onofrio, Michael Tsui, Daniele Sancarlo, Francesco Giuliani, Antonio Greco, et al. (2021). "Assistive robots to improve the independent living of older persons: results from a needs study". In: *Disability and Rehabilitation: Assistive Technology* 16.1, pp. 92–102.

Foster, Liam and Kathy Boxall (2015). "People with learning disabilities and 'active ageing'". In: *British Journal of Learning Disabilities* 43.4, pp. 270–276.

Fragkiadaki, Katerina, Sergey Levine, Panna Felsen, and Jitendra Malik (2015). "Recurrent network models for human dynamics". In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4346–4354.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2000). "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)". In: *The annals of statistics* 28.2, pp. 337–407.

Garcia-Domnguez, Laura, Patricia Navas, Miguel Angel Verdugo, and Victor B Arias (2020). "Chronic health conditions in aging individuals with intellectual disabilities". In: *International Journal of Environmental Research and Public Health* 17.9, p. 3126.

Garrels, Veerle and Hanne Marie Høybråten Sigstad (2021). "Employment for persons with intellectual disability in the Nordic countries: A scoping review". In: *Journal of Applied Research in Intellectual Disabilities* 34.4, pp. 993–1007.

Geertsema, Evelien E, Roland D Thijs, Therese Gutter, Ben Vledder, Johan B Arends, Frans S Leijten, Gerhard H Visser, and Stiliyan N Kalitzin (2018). "Automated video-based detection of nocturnal convulsive seizures in a residential care setting". In: *Epilepsia* 59, pp. 53–60.

Geppert, Ana, Emma M Smith, Damian Haslett, Jennie Wong, Ikenna D Ebuenyi, Jessica Noske-Turner, and Malcom MacLachlan (2022). "Assistive Technology to Promote Participation in Sport for People with Disabilities". In.

Geurts, Pierre, Damien Ernst, and Louis Wehenkel (2006). "Extremely randomized trees". In: *Machine learning* 63.1, pp. 3–42.

Gidel, Gauthier, Francis Bach, and Simon Lacoste-Julien (2019). "Implicit regularization of discrete gradient dynamics in linear neural networks". In: *Advances in Neural Information Processing Systems* 32.

Gilissen, Christian, Jayne Y Hehir-Kwa, Djie Tjwan Thung, Maartje van de Vorst, Bregje WM van Bon, Marjolein H Willemsen, Michael Kwint, Irene M Janssen, Alexander Hoischen, Annette Schenck, et al. (2014). "Genome sequencing identifies major causes of severe intellectual disability". In: *Nature* 511.7509, pp. 344–347.

Gjertsen, Hege, Stefan Hardonk, and Jens Ineland (2021). "Work Inclusion for People with Intellectual Disabilities in Three Nordic Countries: The Current Policy and Challenges". In.

Gómez, Laura E, Robert L Schalock, and Miguel Ángel Verdugo (2021). "A quality of life supports model: Six research-focused steps to evaluate the model and enhance research practices in the field of IDD". In: *Research in developmental disabilities* 119, p. 104112.

Gómez, Laura E, Miguel Ángel Verdugo, Benito Arias, Patricia Navas, and Robert L Schalock (2013). "The development and use of Provider Profiles at the organizational and systems level". In: *Evaluation and Program Planning* 40, pp. 17–26.

Gomez, Laura Elisabet, Benito Arias, Miguel Angel Verdugo, and Patricia Navas (2012). "An outcomes-based assessment of quality of life in social services". In: *Social Indicators Research* 106.1, pp. 81–93.

Gómez Sánchez, Laura Elisabet, María Lucía Morán Suárez, Susana Al-Halabí Díaz, Chris Swerts, Miguel Ángel Verdugo Alonso, Robert L Schalock, et al. (2022). "Quality of life and the International Convention on the Rights of Persons with Disabilities: consensus indicators for assessment". In: *Psicothema*.

Gómez Sánchez, Laura Elisabet, Robert L Schalock, Miguel Ángel Verdugo Alonso, et al. (2021). "A new paradigm in the field of intellectual and developmental disabilities: Characteristics and evaluation". In: *Psicothema*.

Gong, Haifeng, Jack Sim, Maxim Likhachev, and Jianbo Shi (2011). "Multi-hypothesis motion planning for visual object tracking". In: *2011 International Conference on Computer Vision*. IEEE, pp. 619–626.

González Fernández, Henar, Cristina Laborda Molla, and Mercè Jariot García (2021). “Employability skills and quality of life among employees with mild and moderate intellectual disability who attend occupation centres in Catalonia (Spain)”. In.

González-Valero, Gabriel, Josep Vidal-Conti, Félix Zurita-Ortega, and Pere Palou-Sampol (2021). “Active Time in Cooperative Activities, Quality of Life and Body Mass Index in Individuals with Intellectual Disabilities. A Model of Structural Equations”. In: *Sustainability* 13.4, p. 2341.

Goual, Hafida and Haitham M Yousof (2020). “Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test”. In: *Journal of Applied Statistics* 47.3, pp. 393–423.

Grabowska, Izabela, Radosław Antczak, Jan Zwierzchowski, and Tomasz Panek (2021). “Individual quality of life and the environment–towards a concept of livable areas for persons with disabilities in Poland”. In: *BMC public health* 21.1, pp. 1–15.

Griffiths, William E and R Carter Hill (2022). “On the Power of the F-test for Hypotheses in a Linear Model”. In: *The American Statistician* 76.1, pp. 78–84.

Gui, Liang-Yan, Kevin Zhang, Yu-Xiong Wang, Xiaodan Liang, José MF Moura, and Manuela Veloso (2018). “Teaching robots to predict human motion”. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 562–567.

Guide, Training (2010). “The convention on the rights of persons with disabilities”. In.

Gunasekar, Suriya, Blake Woodworth, Srinadh Bhojanapalli, Behnam Neyshabur, and Nathan Srebro (2018). “Implicit regularization in matrix factorization”. In: *2018 Information Theory and Applications Workshop (ITA)*. IEEE, pp. 1–10.

Gupta, Ankur, Julieta Martinez, James J Little, and Robert J Woodham (2014). “3D pose from motion for cross-view action recognition via non-linear circulant temporal encoding". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2601–2608.

Gustafsson, Fredrik K, Martin Danelljan, and Thomas B Schon (2020). "Evaluating scalable bayesian deep learning methods for robust computer vision". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 318–319.

Haigh, Anna, Darren Lee, Carl Shaw, Michelle Hawthorne, Stephen Chamberlain, David W Newman, Zara Clarke, and Nigel Beail (2013). "What things make people with a learning disability happy and satisfied with their lives: An inclusive research project". In: *Journal of Applied Research in Intellectual Disabilities* 26.1, pp. 26–33.

Hamilton, William L, Rex Ying, and Jure Leskovec (2017). "Inductive representation learning on large graphs". In: *arXiv preprint arXiv:1706.02216*.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016). "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

Hendrycks, Dan, Mantas Mazeika, and Thomas Dietterich (2018). "Deep Anomaly Detection with Outlier Exposure". In: *International Conference on Learning Representations*.

Hermsen, MA, PJCM Embregts, AHC Hendriks, and N Frielink (2014). "The human degree of care. Professional loving care for people with a mild intellectual disability: an explorative study". In: *Journal of Intellectual Disability Research* 58.3, pp. 221–232.

Hodges, Holly, Casey Fealko, and Neelkamal Soares (2020). "Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation". In: *Translational pediatrics* 9.Suppl 1, S55.

Hsu, Yen-Chang, Yilin Shen, Hongxia Jin, and Zsolt Kira (2020). "Generalized odin: Detecting out-of-distribution image without learning from out-of-distribution data". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10951–10960.

Innes, Anthea, Louise McCabe, and Karen Watchman (2012). "Caring for older people with an intellectual disability: a systematic review". In: *Maturitas* 72.4, pp. 286–295.

Ionescu, Catalin, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu (2013). "Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments". In: *IEEE transactions on pattern analysis and machine intelligence* 36.7, pp. 1325–1339.

Ionescu, Catalin, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu (2014). "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.7, pp. 1325–1339.

Jain, Ashesh, Amir R Zamir, Silvio Savarese, and Ashutosh Saxena (2016). "Structural-rnn: Deep learning on spatio-temporal graphs". In: *Proceedings of the ieee conference on computer vision and pattern recognition*, pp. 5308–5317.

Jing, Li, Jure Zbontar, et al. (2020). "Implicit rank-minimizing autoencoder". In: *Advances in Neural Information Processing Systems* 33, pp. 14736–14746.

Joshi, Ankur, Saket Kale, Satish Chandel, and D Kumar Pal (2015). "Likert scale: Explored and explained". In: *British journal of applied science & technology* 7.4, p. 396.

Kalatian, Arash and Bilal Farooq (2022). "A context-aware pedestrian trajectory prediction framework for automated vehicles". In: *Transportation research part C: emerging technologies* 134, p. 103453.

Kharbat, Faten F, Abdallah Alshawabkeh, and M Lynn Woolsey (2021). "Identifying gaps in using artificial intelligence to support students with intellectual disabilities from education and health perspectives". In: *Aslib Journal of Information Management* 73.1, pp. 101–128.

Kingma, Diederik P and Max Welling (2013). "Auto-encoding variational bayes". In: *arXiv preprint arXiv:1312.6114*.

Kingma, Diederik P, Max Welling, et al. (2019). "An introduction to variational autoencoders". In: *Foundations and Trends® in Machine Learning* 12.4, pp. 307–392.

Kinnear, Deborah, Jill Morrison, Linda Allan, Angela Henderson, Elita Smiley, and Sally-Ann Cooper (2018). "Prevalence of physical conditions and multimorbidity in a cohort of adults with intellectual disabilities with and without Down syndrome: cross-sectional study". In: *BMJ open* 8.2, e018292.

Kong, Yu and Yun Fu (2022). "Human action recognition and prediction: A survey". In: *International Journal of Computer Vision* 130.5, pp. 1366–1401.

Koppula, Hema Swetha and Ashutosh Saxena (2013). "Anticipating human activities for reactive robotic response." In: *IROS*. Tokyo, p. 2071.

Krahn, Gloria L (2011). "WHO World Report on Disability: a review". In: *Disability and health journal* 4.3, pp. 141–142.

Kuijken, NMJ, J Naaldenberg, MW Nijhuis-Van der Sanden, and HMJ Van Schrojenstein-Lantman de Valk (2016). "Healthy living according to adults with intellectual disabilities: towards tailoring health promotion initiatives". In: *Journal of Intellectual Disability Research* 60.3, pp. 228–241.

Lebailly, Tim, Sena Kiciroglu, Mathieu Salzmann, Pascal Fua, and Wei Wang (2020). "Motion prediction using temporal inception module". In: *Proceedings of the Asian Conference on Computer Vision*.

Lee, Kimin, Kibok Lee, Honglak Lee, and Jinwoo Shin (2018). "A simple unified framework for detecting out-of-distribution samples and adversarial attacks". In: *Advances in neural information processing systems* 31.

Li, Bin, Jian Tian, Zhongfei Zhang, Hailin Feng, and Xi Li (2020a). "Multitask non-autoregressive model for human motion prediction". In: *IEEE Transactions on Image Processing* 30, pp. 2562–2574.

Li, Ce, Chunyu Xie, Baochang Zhang, Jungong Han, Xiantong Zhen, and Jie Chen (2021). “Memory attention networks for skeleton-based action recognition”. In: *IEEE Transactions on Neural Networks and Learning Systems*.

Li, Ce, Baochang Zhang, Chen Chen, Qixiang Ye, Jungong Han, Guodong Guo, and Rongrong Ji (2019). “Deep manifold structure transfer for action recognition”. In: *IEEE transactions on image processing* 28.9, pp. 4646–4658.

Li, Chen, Zhen Zhang, Wee Sun Lee, and Gim Hee Lee (2018). “Convolutional sequence to sequence model for human dynamics”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5226–5234.

Li, Dongxu, Cristian Rodriguez, Xin Yu, and Hongdong Li (2020b). “Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison”. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1459–1469.

Li, Maosen, Siheng Chen, Yangheng Zhao, Ya Zhang, Yanfeng Wang, and Qi Tian (2020c). “Dynamic multiscale graph neural networks for 3d skeleton based human motion prediction”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 214–223.

Lian, Jing, Weiwei Ren, Linhui Li, Yafu Zhou, and Bin Zhou (2022). “PTP-STGCN: Pedestrian Trajectory Prediction Based on a Spatio-temporal Graph Convolutional Neural Network”. In: *Applied Intelligence*, pp. 1–17.

Liang, Dehua, David A Frederick, Elia E Lledo, Natalia Rosenfield, Vincent Berardi, Erik Linstead, and Uri Maoz (2022). “Examining the utility of nonlinear machine learning approaches versus linear regression for predicting body image outcomes: The US Body Project I”. In: *Body Image* 41, pp. 32–45.

Liang, Shiyu, Yixuan Li, and R Srikant (2018). “Enhancing The Reliability of Out-of-distribution Image Detection in Neural Networks”. In: *International Conference on Learning Representations*.

Liao, Peiwen, Claire Vajdic, Julian Trollor, and Simone Reppermund (2021). “Prevalence and incidence of physical health conditions in people with intellectual disability–a systematic review”. In: *PloS one* 16.8, e0256294.

Liu, Chenchen and Yadong Mu (2021). “Searching Motion Graphs for Human Motion Synthesis”. In: *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 871–879.

Liu, Jie, Gangning Liang, Kimberly D Siegmund, and Juan Pablo Lewinger (2018). “Data integration by multi-tuning parameter elastic net regression”. In: *BMC bioinformatics* 19.1, pp. 1–9.

Liu, Ruixuan and Changliu Liu (2020). “Human motion prediction using adaptable recurrent neural networks and inverse kinematics”. In: *IEEE Control Systems Letters* 5.5, pp. 1651–1656.

Liu, Wenqian, Runze Li, Meng Zheng, Srikrishna Karanam, Ziyan Wu, Bir Bhanu, Richard J Radke, and Octavia Camps (2020). “Towards visually explaining variational autoencoders”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8642–8651.

Lombardi, Marco, Luigi Croce, Claudia Claes, Stijn Vandevelde, and Robert L Schalock (2016). “Factors predicting quality of life for people with intellectual disability: Results from the ANFFAS study in Italy”. In: *Journal of Intellectual & Developmental Disability* 41.4, pp. 338–347.

Lopez, Romain, Pierre Boyeau, Nir Yosef, Michael Jordan, and Jeffrey Regier (2020). “Decision-making with auto-encoding variational Bayes”. In: *Advances in Neural Information Processing Systems* 33, pp. 5081–5092.

Lyu, Kedi, Haipeng Chen, Zhenguang Liu, Beiqi Zhang, and Ruili Wang (2022). "3D human motion prediction: A survey". In: *Neurocomputing* 489, pp. 345–365.

M. Mostafa, Samih, Abdelrahman S. Eladimy, Safwat Hamad, and Hirofumi Amano (2020). "CBRL and CBRC: Novel algorithms for improving missing value imputation accuracy based on Bayesian ridge regression". In: *Symmetry* 12.10, p. 1594.

Ma, Qiulin, Qi Zou, Yaping Huang, and Nan Wang (2022). "Dynamic pedestrian trajectory forecasting with LSTM-based Delaunay triangulation". In: *Applied Intelligence* 52.3, pp. 3018–3028.

Maestro-Gonzalez, Alba, M Cruz Bilbao-Leon, David Zuazua-Rico, Jose M Fernandez-Carreira, Ricardo F Baldonedo-Cernuda, and M Pilar Mosteiro-Diaz (2018). "Quality of life as assessed by adults with cerebral palsy". In: *PLoS One* 13.2, e0191960.

Mahmoud, Karar and Mohamed Abdel-Nasser (2018). "Fast yet accurate energy-loss-assessment approach for analyzing/sizing PV in distribution systems using machine learning". In: *IEEE Transactions on Sustainable Energy* 10.3, pp. 1025–1033.

Mao, Wei, Miaomiao Liu, and Mathieu Salzmann (2020). "History repeats itself: Human motion prediction via motion attention". In: *European Conference on Computer Vision*. Springer, pp. 474–489.

Mao, Wei, Miaomiao Liu, Mathieu Salzmann, and Hongdong Li (2019). "Learning trajectory dependencies for human motion prediction". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9489–9497.

March, Salvatore L. (2017). "Human Motion Prediction for Intelligent Assistance to Elderly People with Disabilities". In: *IEEE Journal of Biomedical and Health Informatics*.

Marcheggiani, D, J Bastings, and I Titov (2018). "Exploiting Semantics in Neural Machine Translation with Graph Convolutional Networks". In: The Association for Computational Linguistics.

Martinez, Julieta, Michael J Black, and Javier Romero (2017). “On human motion prediction using recurrent neural networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2891–2900.

Masten, Ann S and Andrew J Barnes (2018). “Resilience in children: Developmental perspectives”. In: *Children* 5.7, p. 98.

McKenzie, Katherine, Meagan Milton, Glenys Smith, and Hélène Ouellette-Kuntz (2016). “Systematic review of the prevalence and incidence of intellectual disabilities: current trends and issues”. In: *Current Developmental Disorders Reports* 3.2, pp. 104–115.

Mihailidis, Alex, Jennifer Boger, Marcelle Canido, and Jesse Hoey (2007). “The use of an intelligent prompting system for people with dementia”. In: *interactions* 14.4, pp. 34–37.

Mishra, Prabhaker, Uttam Singh, Chandra M Pandey, Priyadarshni Mishra, and Gaurav Pandey (2019). “Application of student’s t-test, analysis of variance, and covariance”. In: *Annals of cardiac anaesthesia* 22.4, p. 407.

Miskimmin, Caley, Shahin Shooshtari, Verena Menec, Karen A Duncan, Toby Martin, and Brenda M Stoesz (2019). “Age-friendly communities for older persons with intellectual disabilities”. In: *Quality in Ageing and Older Adults* 20.4, pp. 206–218.

Montgomery, Douglas C, Elizabeth A Peck, and G Geoffrey Vining (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

Monti, Federico, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M Bronstein (2017). “Geometric deep learning on graphs and manifolds using mixture model cnns”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5115–5124.

Morán, Lucía, Laura E Gómez, Giulia Balboni, Asunción Monsalve, Miguel Ángel Verdugo, and Mar Rodríguez (2022). “Predictors of individual quality of life in young people with Down syndrome.” In: *Rehabilitation Psychology*.

Myronenko, Andriy (2018). “3D MRI brain tumor segmentation using autoencoder regularization”. In: *International MICCAI Brainlesion Workshop*. Springer, pp. 311–320.

OECD (2014). “Elderly population”. In: DOI: https://doi.org/https://doi.org/10.1787/8d805ea1-en. URL: https://www.oecd-ilibrary.org/content/data/8d805ea1-en.

O’Leary, Lisa, Sally-Ann Cooper, and Laura Hughes-McCormack (2018). “Early death and causes of death of people with intellectual disabilities: a systematic review”. In: *Journal of Applied Research in Intellectual Disabilities* 31.3, pp. 325–342.

Oliveira, Luan Freitas, Tiago Fernando Chaves, Nathacha Baretto, Gisele Rozone de Luca, Ingrid Tremel Barbato, Jorge Humberto Barbato Filho, Maristela Ocampos, and Angelica Francesca Maris (2020). “Etiology of intellectual disability in individuals from special education schools in the south of Brazil”. In: *BMC pediatrics* 20.1, pp. 1–12.

Organization, World Health (2001). “Healthy ageing–adults with intellectual disabilities: summative report”. In: *Journal of Applied Research in Intellectual Disabilities* 14.3, pp. 256–275.

Ottaviani, Filippo Maria and Alberto De Marco (2022). “Multiple Linear Regression Model for Improved Project Cost Forecasting”. In: *Procedia Computer Science* 196, pp. 808–815.

Paden, Brian, Michal Čáp, Sze Zheng Yong, Dmitry Yershov, and Emilio Frazzoli (2016). “A survey of motion planning and control techniques for self-driving urban vehicles”. In: *IEEE Transactions on intelligent vehicles* 1.1, pp. 33–55.

Panahi, Mahdi, Amiya Gayen, Hamid Reza Pourghasemi, Fatemeh Rezaie, and Saro Lee (2020). “Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms”. In: *Science of the Total Environment* 741, p. 139937.

Parenti, Ilaria, Luis G Rabaneda, Hanna Schoen, and Gaia Novarino (2020). "Neurodevelopmental disorders: from genetics to functional pathways". In: *Trends in Neurosciences* 43.8, pp. 608–621.

Parmenter, Trevor R (2021). "An analysis of the dimensions of quality of life for people with physical disabilities". In: *Quality of life for handicapped people*. Routledge, pp. 7–36.

Paszke, Adam, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer (2017). "Automatic differentiation in pytorch". In.

Patel, Dilip R, Roger Apple, Shibani Kanungo, and Ashley Akkal (2018). "Intellectual disability: definitions, evaluation and principles of treatment". In: *Pediatric Medicine* 1.11, pp. 10–21037.

Patel, Dilip R, Maria Demma Cabral, Arlene Ho, and Joav Merrick (2020). "A clinical primer on intellectual disability". In: *Translational pediatrics* 9.Suppl 1, S23.

Pavllo, Dario, Christoph Feichtenhofer, Michael Auli, and David Grangier (2020). "Modeling human motion with quaternion-based neural networks". In: *International Journal of Computer Vision* 128, pp. 855–872.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. (2011). "Scikit-learn: Machine learning in Python". In: *the Journal of machine Learning research* 12, pp. 2825–2830.

Perry, Jonathan, S Beyer, and Soren Holm (2009). "Assistive technology, telecare and people with intellectual disabilities: ethical considerations". In: *Journal of Medical Ethics* 35.2, pp. 81–86.

Pollack, Martha E (2005). "Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairment". In: *AI magazine* 26.2, pp. 9–9.

Pu, Lihui, Wendy Moyle, Cindy Jones, and Michael Todorovic (2019). “The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies”. In: *The gerontologist* 59.1, e37–e51.

Putnam, Michelle and Christine Bigby (2021). *Handbook on ageing with disability*. Routledge.

Riquelme, Luis, Rebecca Benjamin, Hannah Tahhan, Gracelynn Sandoval, Naga Alomari, and Alexandra Soyfer (2016). “Feeding/swallowing disorders: Maintaining quality of life in persons with intellectual disability”. In.

Rofougaran, Ahmadreza Reza, Maryam Rofougaran, Nambirajan Seshadri, Brima B Ibrahim, John Walley, and Jeyhan Karaoguz (2018). *Game console and gaming object with motion prediction modeling and methods for use therewith*. US Patent 9,943,760.

Rubeis, Giovanni (2020). “The disruptive power of Artificial Intelligence. Ethical aspects of gerontechnology in elderly care”. In: *Archives of gerontology and geriatrics* 91, p. 104186.

Rubio, José de Jesús, Enrique Garcia, Gustavo Aquino, Carlos Aguilar-Ibañez, Jaime Pacheco, and Alejandro Zacarias (2020). “Learning of operator hand movements via least angle regression to be teached in a manipulator”. In: *Evolving Systems* 11.2, pp. 317–332.

Sanchez, Laura Elisabet Gomez (2021). *The Environmental, Public Health, and Human Rights Impacts on Enhancing the Quality of Life of People with Intellectual Disability*. MDPI.

Sandjojo, Janice, Aglaia MEE Zedlitz, Winifred A Gebhardt, Joop Hoekman, Jeanet A den Haan, and Andrea WM Evers (2019). “Effects of a self-management training for people with intellectual disabilities”. In: *Journal of Applied Research in Intellectual Disabilities* 32.2, pp. 390–400.

Sang, Hai-Feng, Zi-Zhen Chen, and Da-Kuo He (2020). “Human Motion prediction based on attention mechanism”. In: *Multimedia Tools and Applications* 79.9, pp. 5529–5544.

Saxe, Andrew M, James L McClelland, and Surya Ganguli (2019). "A mathematical theory of semantic development in deep neural networks". In: *Proceedings of the National Academy of Sciences* 116.23, pp. 11537–11546.

Schalock, RL, R Luckasson, and MJ Tassé (2021a). "Twenty questions and answers regarding the 12th edition of the AAIDD manual: intellectual disability: definition, diagnosis, classification, and systems of supports". In: *American Association on Intellectual and Developmental Disabilities*.

Schalock, Robert L, Ruth Luckasson, and Marc J Tassé (2021b). "An Overview of Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Supports". In: *American Journal on Intellectual and Developmental Disabilities* 126.6, pp. 439–442.

Schalock, Robert L, Ruth Luckasson, Marc J Tassé, and Miguel Angel Verdugo (2018). "A holistic theoretical approach to intellectual disability: Going beyond the four current perspectives". In: *Intellectual and developmental disabilities* 56.2, pp. 79–89.

Schepens, Hadewych RMM, Joris Van Puyenbroeck, and Bea Maes (2019a). "How to improve the quality of life of elderly people with intellectual disability: A systematic literature review of support strategies". In: *Journal of applied research in intellectual disabilities* 32.3, pp. 483–521.

Schepens, Hadewych RMM, Joris Van Puyenbroeck, and Bea Maes (2019b). ""One does not forget, it all comes back": elderly people with intellectual disability review adversities and stress-protection in their lives". In: *Quality in Ageing and Older Adults*.

Shabani, Sevda, Saeed Samadianfard, Mohammad Taghi Sattari, Amir Mosavi, Shahaboddin Shamshirband, Tibor Kmet, and Annamária R Várkonyi-Kóczy (2020). "Modeling pan evaporation using Gaussian process regression K-nearest neighbors random forest and support vector machines; comparative analysis". In: *Atmosphere* 11.1, p. 66.

Shafiee, Sahameh, Lars Martin Lied, Ingunn Burud, Jon Arne Dieseth, Muath Alsheikh, and Morten Lillemo (2021). "Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery". In: *Computers and Electronics in Agriculture* 183, p. 106036.

Shanmugasundar, G, M Vanitha, Robert Čep, Vikas Kumar, Kanak Kalita, and M Ramachandran (2021). "A Comparative Study of Linear, Random Forest and AdaBoost Regressions for Modeling Non-Traditional Machining". In: *Processes* 9.11, p. 2015.

Shi B, Liu T (2019). "Graph Convolutional LSTM for Human Motion Prediction." In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 12538–12547.

Shirai, Akihiko, Erik Geslin, and Simon Richir (2007). "WiiMedia: motion analysis methods and applications using a consumer video game controller". In: *Proceedings of the 2007 ACM SIGGRAPH symposium on Video games*, pp. 133–140.

Shu, Xiangbo, Liyan Zhang, Guo-Jun Qi, Wei Liu, and Jinhui Tang (2021). "Spatiotemporal co-attention recurrent neural networks for human-skeleton motion prediction". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.6, pp. 3300–3315.

Shukla, Priya, Nilotpal Pramanik, Deepesh Mehta, and GC Nandi (2022). "Generative model based robotic grasp pose prediction with limited dataset". In: *Applied Intelligence*, pp. 1–15.

Simões, Cristina and Sofia Santos (2017). "The impact of personal and environmental characteristics on quality of life of people with intellectual disability". In: *Applied research in quality of life* 12.2, pp. 389–408.

Singh, Divya and Rajeev Srivastava (2022). "Graph Neural Network with RNNs based trajectory prediction of dynamic agents for autonomous vehicle". In: *Applied Intelligence*, pp. 1–16.

Somani, Raghav, Chirag Gupta, Prateek Jain, and Praneeth Netrapalli (2018). "Support recovery for orthogonal matching pursuit: upper and lower bounds". In: *Advances in Neural Information Processing Systems* 31.

Song, Xin Yue, Nam Dao, and Paula Branco (2022). "DistSMOGN: Distributed SMOGN for Imbalanced Regression Problems". In: *Fourth International Workshop on Learning with Imbalanced Domains: Theory and Applications*. PMLR, pp. 38–52.

Soudry, Daniel, Elad Hoffer, Mor Shpigel Nacson, Suriya Gunasekar, and Nathan Srebro (2018). "The implicit bias of gradient descent on separable data". In: *The Journal of Machine Learning Research* 19.1, pp. 2822–2878.

Speiser, Jaime Lynn, Michael E Miller, Janet Tooze, and Edward Ip (2019). "A comparison of random forest variable selection methods for classification prediction modeling". In: *Expert systems with applications* 134, pp. 93–101.

Sun, Qiang, Wen-Xin Zhou, and Jianqing Fan (2020). "Adaptive huber regression". In: *Journal of the American Statistical Association* 115.529, pp. 254–265.

Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi (2017). "Inception-v4, inception-resnet and the impact of residual connections on learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. 1.

Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich (2015). "Going deeper with convolutions". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9.

Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna (2016). "Rethinking the inception architecture for computer vision". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826.

Taimoor, Najma and Semeen Rehman (2021). "Reliable and resilient AI and IoT-based personalised healthcare services: A survey". In: *IEEE Access* 10, pp. 535–563.

Takenoshita, Shintaro, Seishi Terada, Ryozo Kuwano, Tomokazu Inoue, Atsushi Cyoju, Shigeru Suemitsu, and Norihito Yamada (2020). "Prevalence of dementia in people with intellectual disabilities: Cross-sectional study". In: *International journal of geriatric psychiatry* 35.4, pp. 414–422.

Tang J Liu Y, Zhao H (2021). "Graph transformer network for human motion prediction." In: *Multimedia Tools and Applications* 80.15, pp. 23339–23352.

Tărlungeanu, Dora C and Gaia Novarino (2018). "Genomics in neurodevelopmental disorders: an avenue to personalized medicine". In: *Experimental & molecular medicine* 50.8, pp. 1–7.

Thompson, James R, Karrie A Shogren, and Michael L Wehmeyer (2016). "Supports and support needs in strengths-based models of intellectual disability". In: *Handbook of research-based practices for educating students with intellectual disability*. Routledge, pp. 39–57.

Tian, Shuo, Wenbo Yang, Jehane Michael Le Grange, Peng Wang, Wei Huang, and Zhewei Ye (2019). "Smart healthcare: making medical care more intelligent". In: *Global Health Journal* 3.3, pp. 62–65.

Torgo, Luís, Rita P Ribeiro, Bernhard Pfahringer, and Paula Branco (2013). "Smote for regression". In: *Portuguese conference on artificial intelligence*. Springer, pp. 378–389.

Tour, Tom Dupré la, Michael Eickenberg, Anwar O Nunez-Elizalde, and Jack L Gallant (2022). "Feature-space selection with banded ridge regression". In: *NeuroImage* 264, p. 119728.

Trollor, Julian, Preeyaporn Srasuebkul, Han Xu, and Sophie Howlett (2017). "Cause of death and potentially avoidable deaths in Australian adults with intellectual disability using retrospective linked data". In: *BMJ open* 7.2, e013489.

Urizarna-Varona, Santiago, Pilar Ruiz, Lorena Naharro, Ignacio Martínez, Irene Sáenz, and Juan-Carlos Carriquiri (2018). "Planificación Centrada en la Persona: metodología de apoyo para que cada persona con discapacidad dirija su vida. La experiencia de ASPRODEMA." In: *Trabajo Social Global-Global Social Work* 8, pp. 164–187.

Uyanık, Gülden Kaya and Neşe Güler (2013). "A study on multiple linear regression analysis". In: *Procedia-Social and Behavioral Sciences* 106, pp. 234–240.

Van Loon, Jos (2013). "Un sistema de apoyos centrado en la persona. Mejoras en la calidad de vida a través de los apoyos". In.

Velickovic, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio (2017). "Graph attention networks". In: *stat* 1050, p. 20.

Verdugo, MA, B Arias, LE Gómez, and RL Schalock (2008). "Manual d'aplicació de l'Escala Gencat de Qualitat de vida". In: *Barcelona: Departament d'Acció Social i Ciutadania, Generalitat de Catalunya*.

Verdugo, MA, Patricia Navas, LE Gómez, and Robert L Schalock (2012). "The concept of quality of life and its role in enhancing human rights in the field of intellectual disability". In: *Journal of Intellectual Disability Research* 56.11, pp. 1036–1045.

Verdugo, MA, Robert L Schalock, Benito Arias, L Gómez, and B Jordán de Urríes (2013). "Calidad de vida". In: *MA Verdugo & RL Schalock (Coords.), Discapacidad e inclusión manual para la docencia*, pp. 443–461.

Verdugo, Miguel A, Virginia Aguayo, Victor B Arias, and Laura García-Domínguez (2020). "A systematic review of the assessment of support needs in people with intellectual and developmental disabilities". In: *International Journal of Environmental Research and Public Health* 17.24, p. 9494.

Verdugo, Miguel A, Robert L Schalock, Kenneth D Keith, and Roger J Stancliffe (2005). "Quality of life and its measurement: Important principles and guidelines". In: *Journal of intellectual disability research* 49.10, pp. 707–717.

Verdugo, Miguel Ángel, Benito Arias, Laura E Gómez, and Robert L Schalock (2010). "Development of an objective instrument to assess quality of life in social services: Reliability and validity in Spain". In: *International Journal of clinical and health psychology* 10.1, pp. 105–123.

Vilaseca, Rosa, Marta Gràcia, Francesc S Beltran, Mariona Dalmau, Elisabeth Alomar, Ana Luisa Adam-Alcocer, and David Simó-Pinatella (2017). "Needs and supports of people with intellectual disability and their families in Catalonia". In: *Journal of Applied Research in Intellectual Disabilities* 30.1, pp. 33–46.

Walters, Sharon, Jordan Stierle, Daniel Stockwell, Alex Carlson, and Joseph Ryan (2021). "Efficacy of video prompting using mobile technology to teach employment tasks to individuals with intellectual disability". In: *Journal of Inclusive Postsecondary Education* 3.1.

Wang, Feier, Yixu Wang, Kai Zhang, Ming Hu, Qin Weng, and Huichun Zhang (2021). "Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation". In: *Environmental Research* 202, p. 111660.

Wei, Lifei, Ziran Yuan, Yanfei Zhong, Lanfang Yang, Xin Hu, and Yangxi Zhang (2019). "An improved gradient boosting regression tree estimation model for soil heavy metal (Arsenic) pollution monitoring using hyperspectral remote sensing". In: *Applied Sciences* 9.9, p. 1943.

Welling, Max and Thomas N Kipf (2016). "Semi-supervised classification with graph convolutional networks". In: *J. International Conference on Learning Representations (ICLR 2017)*.

(WHO), World Health Organization (2021). "Disability and Health". In.

Wieczorek, Dagmar (2018). "Autosomal dominant intellectual disability". In: *medizinische genetik* 30.3, pp. 318–322.

Yadav, Gaurav Kumar and G C Nandi (2020). "Development of Adaptive Sampling Based strategy for Human Activity Predictions Using Sequential Networks". In: *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*. IEEE, pp. 1–6.

Yadav, Gaurav Kumar, Benigno Moreno Vidales, Sara Dueñas, Mohamed Abdel-Nasser, Hatem A Rashwan, Domenec Puig, and GC Nandi (2022a). "Predicting Personalized Quality of Life of an Intellectually Disabled Person Utilizing Machine Learning". In: *Artificial Intelligence Research and Development*. IOS Press, pp. 139–142.

Yadav, Gaurav Kumar, Benigno Moreno Vidales, Hatem A Rashwan, Joan Oliver, Domenec Puig, GC Nandi, and Mohamed Abdel-Nasser (2022b). "Effective ML-based quality of life prediction approach for dependent people in guardianship entities". In: *Alexandria Engineering Journal*.

Yadav, Gaurav Kumar, Benigno Moreno Vidales, Hatem A Rashwan, Joan Oliver, Domenec Puig, GC Nandi, and Mohamed Abdel-Nasser (2023). "Effective ML-based quality of life prediction approach for dependent people in guardianship entities". In: *Alexandria Engineering Journal* 65, pp. 909–919.

Yang, Yujun and Yimei Yang (2020). "Hybrid prediction method for wind speed combining ensemble empirical mode decomposition and Bayesian ridge regression". In: *IEEE Access* 8, pp. 71206–71218.

Ying, Rex, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec (2018). "Graph convolutional neural networks for web-scale recommender systems". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 974–983.

Yu, Yue, Niehao Tian, XiangYu Hao, Tao Ma, and Chunguang Yang (2022). "Human motion prediction with gated recurrent unit model of multi-dimensional input". In: *Applied Intelligence* 52.6, pp. 6769–6781.

Yuan, Jinliang, Meng Cao, Hao Cheng, Hualei Yu, Junyuan Xie, and Chongjun Wang (2022). "A unified structure learning framework for graph attention networks". In: *Neurocomputing* 495, pp. 194–204.

Yusif, Salifu, Jeffrey Soar, and Abdul Hafeez-Baig (2016). "Older people, assistive technologies, and the barriers to adoption: A systematic review". In: *International journal of medical informatics* 94, pp. 112–116.

Zhang, Baochang, Yun Yang, Chen Chen, Linlin Yang, Jungong Han, and Ling Shao (2017). "Action recognition using 3D histograms of texture and a multi-class boosting classifier". In: *IEEE Transactions on Image processing* 26.10, pp. 4648–4660.

Zhang, Bo, Jinghui Zhong, and Wentong Cai (2022). "A Data-Driven Approach for Pedestrian Intention Prediction in Large Public Places". In: *SIGSIM Conference on Principles of Advanced Discrete Simulation*, pp. 33–36.

Zhang, Chao, Zichao Yang, Xiaodong He, and Li Deng (2020). "Multimodal intelligence: Representation learning, information fusion, and applications". In: *IEEE Journal of Selected Topics in Signal Processing* 14.3, pp. 478–493.

Zhang, Jingtian, Hubert PH Shum, Jungong Han, and Ling Shao (2018). "Action recognition from arbitrary views using transferable dictionary learning". In: *IEEE transactions on image processing* 27.10, pp. 4709–4723.

Zhang, Zhibo, Yanjun Zhu, Rahul Rai, and David Doermann (2022). "PIMNet: Physics-Infused Neural Network for Human Motion Prediction". In: *IEEE Robotics and Automation Letters* 7.4, pp. 8949–8955.

Zhang, Zijun (2018). "Improved adam optimizer for deep neural networks". In: *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. Ieee, pp. 1–2.

Zhou, Jie, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun (2020). "Graph neural networks: A review of methods and applications". In: *AI open* 1, pp. 57–81.

Zietlow, Dominik, Michal Rolinek, and Georg Martius (2021). "Demystifying Inductive Biases for (Beta-) VAE Based Architectures". In: *International Conference on Machine Learning*. PMLR, pp. 12945–12954.

UNIVERSITAT ROVIRA i VIRGILI