# Applied Intelligence


## Exploring Emotional Stability: From Conventional Approaches to Machine Learning Insights

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | APIN-D-24-03623R2 |
| **Full Title:** | Exploring Emotional Stability: From Conventional Approaches to Machine Learning Insights |
| **Article Type:** | Original Submission |
| **Keywords:** | Machine Learning;  data mining, emotional stability, psychology |
| **Corresponding Author:** | Luis Gonzaga Baca Ruiz<br>University of Granada: Universidad de Granada<br>Granada, Spain / Andalucía / Granada SPAIN |
| **Corresponding Author Secondary Information:** | |
| **Corresponding Author's Institution:** | University of Granada: Universidad de Granada |
| **Corresponding Author's Secondary Institution:** | |
| **First Author:** | Marcos Romero Madroñal |
| **First Author Secondary Information:** | |
| **Order of Authors:** | Marcos Romero Madroñal |
| | Eduar S. Ramirez |
| | Luis Gonzaga Baca Ruiz |
| | María José Serrano Fernández |
| | Elena Pérez Moreiras |
| | María del Carmen Pegalajar Jiménez |
| **Order of Authors Secondary Information:** | |
| **Funding Information:** | |
| **Abstract:** | In contemporary psychological assessments, diverse traits are often evaluated using extensive questionnaires. This study focuses on the trait of emotional stability, and acknowledges the inherent limitations and issues associated with prolonged survey instruments. To address these challenges, we propose a Machine Learning (ML) approach to directly predict emotional stability, offering a more efficient alternative to bulky questionnaires. The study carefully selected variables with previously established relationships to emotional stability, utilizing a dataset of 2203 individuals who responded to a series of psychometric questionnaires. The proposed method yields promising results, achieving an R2 score of approximately 0.71 on the test set, indicating robust predictive performance. These models highlighted the significance of variables such as emotional stress and self-esteem, emphasizing their substantial role in predicting emotional stability. It is noteworthy that even with a reduced set of variables, the models remained statistically equivalent. The results provide valuable insights for predicting stability with smaller sets of variables and contribute knowledge that complements the understanding of emotional stability. |

We thank the reviewers for their work during this process.

All typos commented by the reviewers have been corrected.
No more modifications were mentioned.

# Exploring Emotional Stability: From Conventional Approaches to Machine Learning Insights

Marcos Romero Madroñal[1], Eduar S. Ramírez[2], Luis Gonzaga Baca Ruiz[3], María José Serrano-Fernández[4], Elena Pérez-Moreiras[5], María del Carmen Pegalajar Jiménez[1,*]

[1]Department of Computer Science and Artificial Intelligence, University of Granada, Granada, 18014, Andalucía, Spain; e.maromad@go.ugr.es, mcarmen@decsai.ugr.es
[2]Faculty of Health Sciences, Universidad Villanueva, Calle Costa Brava 2, 28034 Madrid, Spain; eduar.ramirez@villanueva.edu
[3]Department of Software Engineering, University of Granada, Granada, 18014, Andalucía, Spain; bacaruiz@ugr.es
[4]Department of Psychology, Rovira I Virgili University, 43002 Tarragona, Spain; mariajose.serrano@urv.cat
[5]RH Asesores Improving S.L., Instituto para el Desarrollo del Talento Natural y el Nuevo Liderazgo, 28221 Madrid, Spain; eperez-moreiras@gruporh.com

**Abstract**. In contemporary psychological assessments, diverse traits are often evaluated using extensive questionnaires. This study focuses on the trait of emotional stability, and acknowledges the inherent limitations and issues associated with prolonged survey instruments. To address these challenges, we propose a Machine Learning (ML) approach to directly predict emotional stability, offering a more efficient alternative to bulky questionnaires. The study carefully selected variables with previously established relationships to emotional stability, utilizing a dataset of 2203 individuals who responded to a series of psychometric questionnaires. The proposed method yields promising results, achieving an R2 score of approximately 0.71 on the test set, indicating robust predictive performance. These models highlighted the significance of variables such as emotional stress and self-esteem, emphasizing their substantial role in predicting emotional stability. It is noteworthy that even with a reduced set of variables, the models remained statistically equivalent. The results provide valuable insights for predicting stability with smaller sets of variables and contribute knowledge that complements the understanding of emotional stability.

Keywords: machine learning; data mining, emotional stability, psychology

## 1 Introduction

Research on emotional stability or neuroticism continues to arouse profound interest because of its enduring influence on various psychology aspects. This construct stands as one of the most extensively studied and frequently used features on academic platforms like PsycINFO. From a classical perspective, it is defined as a personality dimension that reflects an individual's propensity to experience negative emotions, such as anxiety, sadness, anger, and guilt, with greater frequency and intensity than their peers (Costa Jr & McCrae, 1995).

In the workplace, research suggests that emotional stability plays a mediating role in the relationship between performance and work engagement, implying that emotionally stable employees may experience fewer adverse consequences in their performance, even when facing variability in their work engagement (Tewfik et al., 2023). These findings are also supported by certain previous meta-analyses (Frye, 2000).

Conversely, contemporary research, such as the study conducted by Kobylińska et al. (2020), has established connections between this personality dimension and various outcomes, including life satisfaction, the experience of positive affect, and emotional regulation strategies. Since the 1980s, personality studies have underscored the profound influence of emotional stability on mental health. An illustrative example of this is the study by Costa and McCrae (1980), which demonstrated a positive correlation between emotional stability and subjective well-being. Likewise, they noted that neurotic traits such as anxiety and worry were linked to subjective distress or unhappiness.

These studies share a common approach in their investigation of emotional stability and its relationship with other significant variables. They employ various methodologies, including correlation studies, regression analysis, analysis of variance, and structural equation modeling. For instance, a notable study led by Hills and Argyle (2001), found that emotional stability was more strongly correlated with life satisfaction than extraversion, with coefficients of 0.49 and 0.53, respectively. And recently, it has been found that emotional stability is positively associated with life satisfaction (Kobylińska et al., 2020). Furthermore, a stronger correlation was observed between emotional stability and self-esteem than extraversion and self-esteem, with coefficients of 0.34 and 0.63, respectively, all statistically significant at a $p < 0.001$ level. Equally important is the finding that emotional stability explains a significantly larger portion of the total variability in multiple regression analyses. These studies have contributed to a shift in the previous paradigm, which suggested a prominent relationship between extraversion and well-being, demonstrating that emotional stability more strongly predicts the latter.

Despite the significance of these methodologies, there is a growing interest in approaches based on machine learning (ML) and artificial intelligence (AI). In this paradigm, the focus is on predicting data patterns rather than conducting traditional inferential tests (Orrù et al., 2020). This approach offers the advantage of not requiring the same assumptions as inferential methodologies. A prominent example is the study conducted by Wardenaar et al. (2021), which investigated the factors influencing the severity of depression and anxiety symptoms over 9 years. Data from individuals diagnosed with depression and/or anxiety, considering 152 variables, were analyzed. Machine learning techniques were employed to identify key factors. The results indicated that age at early onset age, respiratory rate, participation disability, somatic diseases, and low income were related to overall psychological distress and symptoms of depression and anxiety.

A similar example is the study by Bai et al. (2021). They monitored and tracked patients with major depression, allowing the prediction of variables that explain symptom variability. Moreover, studies have compared traditional methods for assessing personality traits, such as emotional stability (Koffi, 2020; Yuan et al., 2018), with approaches based on ML, highlighting the significant utility of ML in the evaluation of psychological phenomena.

These two examples of research in psychology that utilise machine learning methods offer promising perspectives in this field. As noted previously, standard statistical models used in psychological research assume linear relationships between predictors and dependent variables. However, these studies indicate that many psychological processes, such as depression, may not be adequately described by a linear model and result from complex and unknown interaction effects based on numerous influential predictors (Henninger et al., 2023). In this regard, as the prediction complexity increases, more flexible models must be considered (Breiman, 2001; Hastie et al., 2009).

In this context, conducting a study on emotional stability using a training dataset with a machine learning method would be a valuable contribution. This approach involves extracting information from the training set, where the model learns patterns and relationships in the data. Subsequently, the model is applied to the test set to predict the emotional stability of previously unseen individuals. An exemplary method, such as Random Forest (Breiman, 2001), offers advantages such as handling large and complex datasets, capturing non-linear patterns and interactions between variables, resistance to overfitting, and providing information on the relative importance of different predictor variables (Philipp et al., 2018). This approach could be an effective tool for anticipating emotional stability based on multiple factors, thereby enhancing the understanding and prediction of this phenomenon.

Given the information presented in the following sections, we justify the selected variables for the training set's prediction of emotional stability and provide reasons supporting their inclusion.

## 2 Selected variables

In our research, we employ an innovative approach supported by artificial intelligence (AI) to predict emotional stability based on various psychological variables. This method allows the prospective assessment of emotional stability, highlighting elements that could influence levels of emotional stability or neuroticism without direct measurements.

We carefully selected a specific set of variables, starting with the expectation of self-efficacy as defined in (Bandura, 1977, 1992). The inclusion of this variable is grounded in its well-established connection with the ability to manage emotions and stress. Longitudinal studies, such as the one conducted by Caprara et al. (2013) to investigate personality development, reveal a strong correlation between initial levels of emotional stability and self-efficacy beliefs. Although emotional stability is perceived as a constant trait in most stages of life, self-efficacy beliefs play an essential role in modifying this trait. Furthermore, findings from this study indicate that the rate of increase in perceived self-efficacy in managing negative emotions can predict the rate of growth in emotional stability.

Other studies support the importance of self-efficacy and emotional stability as protective factors for mental health, mitigating adverse experiences in childhood (Cohrdes & Mauz, 2020). Additionally, self-efficacy has proven to be an essential mechanism linking emotional stability to the reduction of symptoms of occupational burnout, as evidenced by the study of Alessandri et al. (2018).

Another key variable in our study is the concept of "flourishing" or psychosocial prosperity, defined as the perception of success in significant areas of life, such as personal relationships, self-esteem, sense of purpose, and optimism (Diener et al., 2010). This variable has gained notable importance, especially from the perspective of well-being. Evidence reveals a positive correlation with emotional stability, indicating that individuals with higher emotional stability experience higher levels of "flourishing" (Omoregie & Carson, 2023). It demonstrated that elevated levels of emotional stability and "flourishing" are associated with a lower incidence of negative intrusive thoughts, a variable of great interest in the last decade because of its relationship with a wide variety of psychological conditions. In this line Pérez-Moreiras (2020) and Serrano-Fernández et al. (in press) also found a high correlation between emotional stability and flourishing.

Additionally, we have included self-esteem, according to Rosenberg's definition (Rosenberg, 2015), representing an individual's overall attitude towards themselves, whether positive or negative. The relevance of self-esteem in an individual's life has been widely recognized from this perspective. Evidence regarding the connection between self-esteem and emotional stability indicates that individuals with stable high self-esteem reported high levels of emotional stability (Crowe et al., 2016; Pérez-Moreiras, 2020). These studies have also identified connections between high self-esteem and traits such as kindness, conscientiousness, and openness. Furthermore, some research, such as that conducted by Bajaj et al. (2019) has demonstrated a relationship between emotional stability and self-esteem, highlighting the mediating role of both variables in the relationship between mindfulness and happiness.

Considering the results of the aforementioned studies, we have adopted a comprehensive approach that encompasses various dimensions. This involves the incorporation of life satisfaction, from the perspective of Diener et al. (1985) and Kjell and Diener (2021), as well as the experience of flow, conceptualized by Getzels and Csikszentmihalyi (1976) and Csikszentmihalyi and Larson (2014), emphasizing the importance of immersion and intrinsic gratification in work activities. This choice is based on more than two decades of insistence on the relevant role of emotional stability in predicting happiness and life satisfaction (Hills & Argyle, 2001). Recent evidence presented by Olaru et al. (2023), reinforces the strong relationship between these variables. Regarding flow, it is noteworthy that a recent study using machine learning has demonstrated its predictive ability and linked it to emotional understanding, self-efficacy, and emotional stability (Pegalajar et al., 2023).

In our research, we have also integrated emotional intelligence, following the framework proposed by Fernández-Berrocal et al. (2004), recognizing its impact on managing emotions, a crucial factor identified in the cited studies that contribute to emotional stability.

Additionally, we have incorporated variables of personal and organizational quality, according to (Barrios-Choplin & Atkinson, 2004; Buchanan & Reilly, 2019), with the purpose of exploring their predictive capacity for emotional stabilitys and observing their relationship with other study variables.

Our research incorporates the "Energetic Intelligence Inventory (ENII-33)" (Pérez-Moreiras, 2020) as a key predictor. This inventory encompasses various subdimensions, such as Body and Movement Intelligence (awareness and constructive use of bodily information), Linguistic Intelligence (use of language to create positive contexts), Transitive-Spiritual Intelligence (seeking a life purpose and transcending individual self), Emotional Intelligence (recognition and utilization of emotions), and Energetic Awareness (being aware of the energetic dimension of the human being). This innovative model interprets energy capacity as a continuously modifiable and improvable variable, aligning with the life purpose and definitions of energy cited by Schippers and Hogenes (2011).

By integrating these variables, we aim to offer a comprehensive and updated perspective on the determinants of emotional stability, thus enriching the understanding of how these factors interact and affect mental health and well-being. This vision is supported by machine learning methods that facilitate the prediction of emotional stability. These AI models allow us to extract hidden information and that traditional methods fail to provide.

The subsequent sections of this paper are organized as follows. Section 3 provides an in-depth exploration of the proposed methodology, offering detailed insights into the instruments, participants, and machine learning techniques, as well as the measures and metrics employed in the study. Subsequently, Section 4 outlines the experiments, shedding light on the experimental setup and procedures. Section 5 presents the experimental results. The paper concludes with Section 6, presenting the conclusions drawn from the research findings and outlining future work.

## 3 Methodology

Figure 1 illustrates the methodology used in this study. Our research begins by gathering a comprehensive dataset derived from the responses of participants to questionnaires designed to measure the psychological variable of emotional stability. After collecting the data, we analyse each variable in detail, including personal, business/work, and psychological factors. The dataset then experiences a thorough preprocessing phase involving normalisation and codification to ensure data quality and uniformity.

During the exploratory analysis stage, we use techniques such as correlation analysis to find insights into the relationships within the dataset. Next, we employ a set of ML techniques, including Linear Regression, Bayesian Ridge, k-NN, SVR, MLP, Decision

Tree, Random Forest, and Gradient Boosting, to model and predict emotional stability. At the same time, we examine variable importance through dominance analysis, which provides information about the most significant predictors in our models. Then, we aim to identify the best model for predicting emotional stability using two distinct approaches. The first approach utilises all available psychological variables, while the second approach employs a reduced set of the most informative variables. To improve the accuracy of our findings, we used confidence intervals to estimate the precision of our model's performance. Finally, the study concludes with a comprehensive analysis of the results and conclusions, providing practical knowledge of the variables that influence emotional stability.

To offer a detailed explanation of the proposed methodology, subsequent subsections focus on key aspects of each phase.

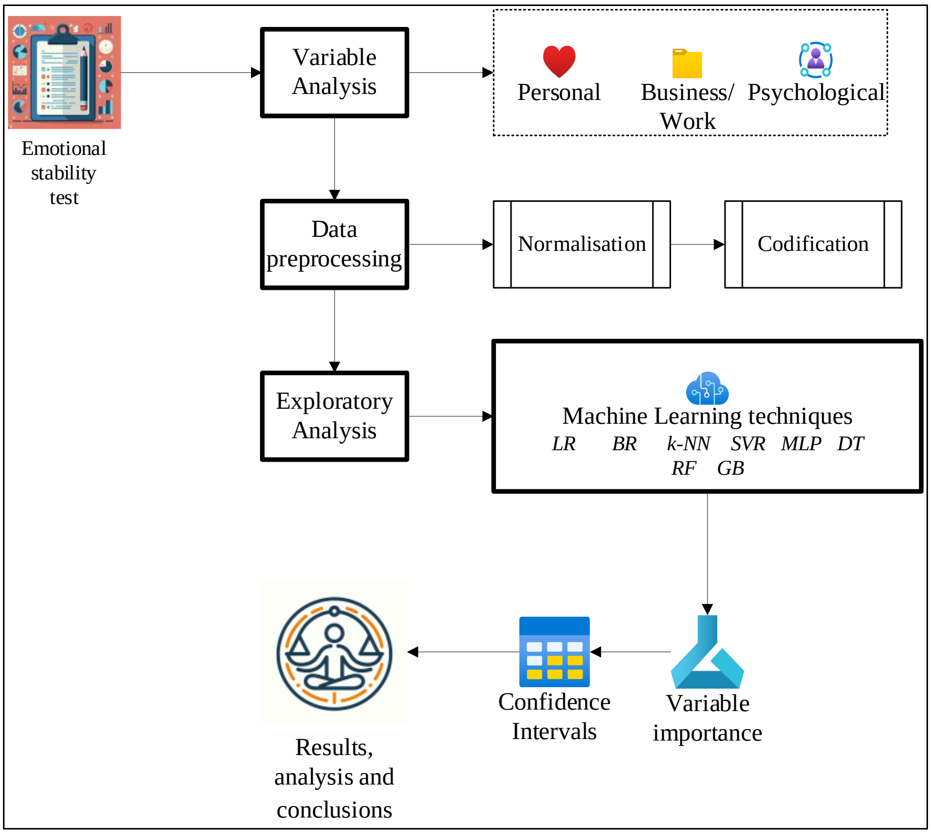


Figure 1.Flow of the proposed methodology.

### 3.1 Instruments

The instruments utilized in the study are outlined below. The predictor variables encompass the scales and subscales presented subsequently.

The General Self-Efficacy Scale (Baessler & Scharzer, 1993) was adapted into Spanish by Sanjuán et al. (2000) and consists of items answered on a four-point scale (from 1 = no agreement, not at all true to 4 = totally agree, completely true). The level of General Self-Efficacy is measured, which is understood as the stable belief of people regarding their ability to adequately handle several stressors in daily life. The scale showed internal consistency (α = .87).

The Psychological Well-Being Scale (Flourishing Scale) (Biswas-Diener, 2008; Diener et al., 2010) was adapted into Spanish by Serrano-Fernández et al. (in press) for Spanish workers. It consists of 5 items (α =.87; for example, 1.- I lead a meaningful and purposeful life) on a Likert-type response scale of 7 alternatives (1 = Strongly disagree, 7 = Strongly agree). It has a good level of internal consistency (Cronbach's α .88). The scale provides a single rating of psychological Well-being related to Flourishing or Personal Growth and it concerns the feeling of happiness and well-being that a person experiences. It is strongly correlated with the results on other scales of psychological well-being and feelings (Diener et al., 2010).

The Short Dispositional Flow Scale (Jackson et al., 2012) was adapted to Spanish by Godoy-Izquierdo et al. (2009) in a sample of Spanish athletes. The scale consists of nine items (α = .80; e.g. “2. I can act spontaneously and automatically, without having to think”). This scale has a Likert-type response scale of five alternatives (from 1 = I never experience these sensations to 5 = I always experience these sensations).

The Rosenberg Self-Esteem Scale (Rosenberg, 1965), in the Spanish version adapted by Martín-Albo et al. (2007), contains 10 items that are assessed on a four-point Likert scale (1 = strongly disagree to 4 = strongly agree). Five items were positively and five were negatively written. The scale showed internal consistency (α = .86; e.g. 4. I can do things like most people).

The Satisfaction with Life Scale (Diener et al., 2010), in its Spanish version, as adapted byAtienza et al. (2000). This is a single factor scale, (α = .84) made up of 5 items (e.g., 2.- Until now I have got the things out of life that I consider important). The response format is a 5-point Likert type scale. (1 = Totally disagree to 5=Totally agree).

The Energetic Intelligence Inventory (IIEN-33) measures the ability of people to identify the energy they feel within and outside themselves, distinguish one from the other, and use this information to achieve individual and collective goals (Pérez-Moreiras, 2020). It has 33 items with a Likert-type response scale of five alternatives (1 = strongly disagree to 5 = strongly agree).  Cronbach's α values for each factor are (f1) Body & Movement Intelligence (BMI = .84); (f2) Emotional Intelligence (EI = .86); (f3) Linguistic Intelligence (LI =.86); (f4) Transitive-Spiritual Intelligence (TSI =. 90); and (f5) Energetic Awareness (EA =.91).

The Overall Personality Assessment Scale (Vigil-Colet et al., 2013) is based on the Big Five model. The test comprises 42 items that were answered on a five-point scale (from 1 = completely disagree to 5 = completely agree) with a five-factor structure. The first of these is “F1. Extraversion”, made up of seven items (α = .86); the second is “F2. Emotional Stability” made up of seven items (α = .86); the third is “F3. Conscientiousness”, made up of seven items (α = .77); the fourth is “F4. Agreeableness”, made up of eight items (α = .71); and the fifth is “F5. Openness to Experience”, made up of eight items (α = .81). It is important to note that the variable of emotional stability serves as the predictor variable in the data training of this study. The score is calculated through a simple weighting of the 7 items that constitute the subscale. For data training, scores above 28 points are considered indicative of high emotional stability, whereas scores below 23 points are indicative of low emotional stability or neuroticism. Direct scores between 24 and 27 are considered normal.

The Personal and Organizational Quality Assessment Revised 4 Scale (POQA-R4) is an instrument designed by Barrios-Choplin and Atkinson (2004) and Buchanan and Reilly (2019) to assess the quality of the work environment. The POQA-R4 comprises four specific subscales: Emotional Vitality, Emotional Stress, Organizational Stress, and Physical Stress. These subscales quantify the perception of improvement or detriment to efficient performance in the work environment using a Likert scale ranging from 1 to 7, where 1 represents “never” and 7 represents “always”. Various measures of internal consistency were recorded, with Cronbach's α values ranging from 0.76 to 0.92 for the four subscales (Barrios-Choplin & Atkinson, 2004; Buchanan & Reilly, 2019).

The Spanish version of the Trait Meta-Mood Scale (TTMS) (Extremera & Fernández-Berrocal, 2004) is the most widely used Emotional Intelligence scale in psychological and educational research. The TTMS-24 contains three dimensions of the Emotional Intelligence (emotional Attention, Clarity and Repair) with 8 items in each dimension. Reliability analysis of the three dimensions of TMMS-24 revealed a Cronbach’s α higher than 0.80. The dimension of emotional Attention has an α=0.89, the dimension of emotional Clarity shows an α=0.84, and the dimension of emotional Repair has an α=0.83.

Regarding the description of the dataset at hand, these surveys encompass a variety of questions, including both personal inquiries and those designed to assess various psychological variables. Initially, without undergoing any filtering processes, the dataset comprises 12 personal variables, 16 variables related to individuals' workplaces, and 26 psychological variables. The raw survey data obtained from the respondents comprised various personal and psychological variables. Initially, 54 potential variables were identified from a comprehensive set of psychological instruments, including scales and subscales. For our experimental focus on emotional stability, we refined our analysis to specifically consider 18 key psychological variables. These variables were selected based on their relevance to emotional stability and were derived through a consolidation process that integrated scales and subscales from the original instruments.

Our study focuses on predicting emotional stability scores from various psychological variables, utilizing the Overall Personality Assessment Scales (OPERAS) (Vigil-Colet et al., 2013). OPERAS is a concise version of the Big Five personality traits and has demonstrated psychometric robustness. It assesses five major dimensions: Extraversion, Emotional Stability, Conscientiousness, Agreeableness, and Openness to Experience.

Emotional Stability, in particular, is a distinct subscale within the OPERAS questionnaire and has shown reliability indices of 0.86. The score for emotional stability is calculated through a simple weighting of the seven items that comprise this subscale. For the purposes of this study, we categorized scores above 28 points as indicative of high emotional stability, scores below 23 points as indicative of low emotional stability or neuroticism, and scores between 24 and 27 as normal. To enhance the predictive model, we excluded factors directly related to the emotional stability construct and composite variables such as total personal and organizational quality, energetic intelligence, and emotional intelligence, resulting in a total of 18 predictors.

### 3.2 Participants

The sample consisted of 2203 Spanish-speaking workers, comprising 22.8% men and 77.2% women. The average age of the participants was 46.09 years (SD = 10.54). Regarding marital status, 50% were married, 34% were single, 15% were divorced or separated, and 1% were widowed. In terms of academic qualifications, 2% had primary education, 22% had secondary education, 46% had university education, and 28% had master's or doctoral studies.

In the context of psychological and social research, working with large sample sizes is often challenging and not always feasible. Our study acknowledges this limitation and, as discussed, the sample size used is consistent with similar research in this field.

### 3.3 Machine Learning Algorithms

In this section, we describe the machine learning methods used to predict the psychological emotional stability variable. Our aim is to discern patterns and relationships within our dataset to reveal the multifaceted factors that influence emotional stability. We will be using a range of advanced computational techniques, including Linear Regression, Bayesian Ridge, k-Nearest Neighbour, Multi-Layer Perceptron, Support Vector Machine, Decision Tree, Random Forest, and Gradient Boosting. Each algorithm brings a particular perspective and computational power to our problem.

In addressing the task of predicting emotional stability, our focus shifts towards regression techniques. Unlike classification, where the goal is to categorize data into distinct groups, regression methods enable forecasting and understanding of the magnitude of specific outcomes. In our study, emotional stability was measured as a continuous variable, and the implementation of regression algorithms becomes pivotal. These techniques allowed us to model the relationships between various input features and continuous emotional stability scores.

**Linear Regression**

Linear Regression (LR) is a straightforward yet powerful algorithm used to estimate numerical outcomes (Maydeu-Olivares et al., 2020). In our study, this is the first tool to model the relationship between input features and the continuous variable of emotional stability. By fitting a linear equation to the data, LR provides a clear understanding of how changes in independent variables relate to changes in emotional stability. This simplicity makes LR a valuable starting point for predictive analysis.

**Bayesian Ridge**

Bayesian Ridge (BR) regression is an extension of traditional linear regression that incorporates probabilistic principles (Stuke et al., 2023). In our study, this method accounts for uncertainties in the data and optimizes the model by considering a range of possible solutions. By blending statistical techniques with Bayesian probability, BR enhances our ability to make accurate predictions of the target variable. This approach is particularly valuable when dealing with complex relationships in a dataset.

**k-Nearest Neighbour**

k-Nearest Neighbour (k-NN) is a non-parametric algorithm that classifies or predicts data points based on their proximity to other data points in the feature space (Chowdhury & Das, 2022). In the context of this study, k-NN considers the similarity between instances to estimate the emotional well-being of an individual. By relying on the closeness of neighbouring data points, k-NN provides a flexible approach to capture patterns in our dataset, making it suitable for scenarios with varying levels of data complexity.

**Multi-Layer Perceptron**

The fourth implemented method was a well-known Artificial Neural Networks, the Multi-Layer Perceptron (MLP) (Farahani et al., 2020). MLP belongs to the family of neural network models, and it can capture hidden patterns in the data. In our study, MLP offers a robust framework for learning the complex relationships between diverse input features and emotional stability. This model comprises multiple layers of interconnected nodes, MLP is capable of recognising non-linear patterns, making it a suitable choice for revealing the subtle relationships within our data.

**Support Vector Regression**

Support Vector Regression (SVR) is an extension of Support Vector Machines adapted for regression tasks (Peiqing, 2022). Similar to its classification counterpart, SVR seeks to find the optimal hyperplane, but in this context, it focuses on minimizing the deviation between predicted and actual values. Using a kernel function and support vectors, SVR efficiently captures nonlinear patterns in the dataset. SVR's adaptability and capacity to manage intricate relationships make SVR a valuable tool for modelling and predicting nuanced variations in the context at hand. In comparison with MLP, SVR provides a simpler but effective approach to handling non-linearities in the data, making it a practical choice for our regression task.

**Decision Tree**

A Decision Tree (DT) is a straightforward and efficient model designed to aid decision-making processes (Chen & Liu, 2021; Namazkhan et al., 2020). A DT acts as a graphical representation that illustrates potential outcomes determined by different input features. The algorithm divides the dataset into subsets through recursive splitting, allowing it to make decisions at each node and ultimately predict emotional stability. DT offers interpretability and are well-matched for comprehending the hierarchical structure of factors that influence emotional stability. This provides a clear and intuitive framework for understanding the relationships between input variables and predicting outcomes in our analysis.

**Random Forest**

Random Forest (RF) is an ensemble learning method that leverages the strengths of multiple DTs (Şevgin & Eranıl, 2023). This model creates a forest of trees, each trained on different subsets of the data. By combining the predictions of individual trees, RF mitigates overfitting and enhances the overall predictive accuracy. This approach is particularly beneficial relative to capturing diverse relationships in the dataset. Thus, it makes RF a robust choice with improved generalizability.

**Gradient Boosting**

The final model used in this study is Gradient Boosting (GB), which is an ensemble learning technique that builds a predictive model in a step-by-step manner (Chen & Liu, 2021). In our study, this model uses the abilities of so-called «weak learners», such as Decision Trees, to create a reliable predictive model. By minimising a specified loss function and correcting errors from the previous stage, GB learns to adapt to the subtleties of the data, progressively improving its predictive accuracy. The algorithm has a great ability to handle complicated relationships and identify subtle patterns. This makes it a valuable asset in our predictive modelling efforts, where we aim to minimise the loss function associated with the differences between the predicted and actual emotional stability scores.

The decision of the models employed in this study is closely linked to our dataset size and the results obtained as can be seen in the following sections. More complex models typically require larger datasets to perform effectively. Given our dataset's constraints, it is possible that these advanced models might not yield significantly better results and could potentially perform worse.

### 3.4 Measures

In order to quantify the performance of our model, the utilization of metrics is paramount. In this study, we specifically employ key metrics such as Mean Squared Error (MSE) and R-squared ($R^2$).

Firstly, MSE measures the average squared difference between the predicted values and the actual values. Importantly, MSE depends on the scale of the output variable and takes only positive values. Due to its crucial role in assessing prediction precision, minimizing MSE becomes a common optimization goal in many machine learning models during training. Let $n$ represents the total number of data points, and $y_i$ and $\hat{y}_i$ denote the actual and predicted values, respectively, for a specific data point. The MSE is mathematically defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (1)$$

Additionally, $R^2$ assesses the proportion of variance in the dependent variable explained by the independent variables, providing a comprehensive evaluation of the model's overall explanatory capability. The $R^2$ metric can range from $-\infty$ to 1, where 1 indicates a perfect fit, and values less than 0 suggest that the model performs worse than a basic mean-based model. The $R^2$ metric is calculated using the formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (2)$$

where $\bar{y}$ represents the mean of the actual values:

$$\bar{y} = \frac{1}{n} - \sum_{i=1}^{n} y_i \tag{3}$$

Lastly, we included a modification of $R^2$ metric that accounts for the number of predictors in the model ($p$). This adjusted-$R^2$ is computed using the expression:

$$R^2_{adjusted} = 1 - \frac{n-1}{n-p-1} (1 - R^2) \tag{4}$$

**Variable importance**

To gain a deeper understanding of the relative importance of each predictor in our model, we employed dominance analysis (Azen & Budescu, 2006), a statistical technique that allows us to assess the unique and shared contributions of individual predictors to the variance in the dependent variable.

The dominance analysis process involves systematically evaluating all possible combinations of predictors and determining their impact on the dependent variable. Unlike traditional variable importance measures, dominance analysis accounts for both independent and joint effects, providing a more nuanced perspective on predictor importance.

Let $\{X_1, X_2, \ldots, X_p\}$ represent the set of predictors in our model. The dominance analysis calculates the dominance weight for each predictor, denoted as $D_k$. This represents the proportion of times that the predictor dominates (i.e., contributes more to the explained variance) across all possible predictor combinations. Thus, the dominance weight $D_k$ for each predictor $X_k$ by considering all possible model combinations without that predictor $X_k$ is mathematically expressed as follows:

$$D_k = \frac{1}{p} \sum_{j=0}^{p-1} \left( \sum_{\substack{T \subseteq \{X_1, \ldots, X_p\} \setminus \{X_k\} \\ n(T)=j}} \frac{R^2(T \cup \{X_k\}) - R^2(T)}{\binom{p-1}{j}} \right) \tag{5}$$

where $T$ represents a subset of predictors, $n(T)$ denotes its cardinality, $R^2(T)$ is the R-squared value for the subset $T$, and $p$ is the total number of predictors in the model. Therefore, it compares the $R^2$ value of each subset $T$ with the $R^2$ value when the excluded predictor $X_k$ is added back $T \cup \{X_k\}$. The calculation considers the proportion of times where adding $X_k$ improves the model fit across all these combinations.

In addition to providing dominance weights, we explored hierarchical levels of dominance, distinguishing between complete, conditional, and general dominance. Complete dominance occurs when one predictor consistently outperforms others across all subsets, whereas conditional dominance and general dominance consider variations in performance across specific subsets and all subsets, respectively.

The dominance weights and hierarchical dominance levels offer a nuanced perspective on predictor importance, considering both individual and interactive effects. Higher dominance weights indicate stronger contributions to the explained variance in the dependent variable. Interpreting the dominance analysis results helps identify key predictors, guiding further investigations or interventions in the psychological context under study.

The application of dominance analysis enhances our model interpretation beyond traditional variable importance metrics, providing valuable insights into the interplay among predictors and their varying levels of importance in explaining the variability in the dependent variable.

**Confidence intervals**

To estimate the precision of our model performance metrics, we calculate confidence intervals around the mean values of the metrics obtained from several measurements. The utilization of confidence intervals allows us to quantify the uncertainty associated with our metric estimates and provides a range within which the true population parameter is likely to fall.

The confidence intervals are approximated using the Central Limit Theorem. Let $\bar{X}$ be the mean of the metric obtained from the different measurements, and $S$ be the sample standard deviation. The confidence interval for the mean can be calculated as defined in equation (6):

$$\left(\bar{X} + t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}\right) \tag{6}$$

Here, $S$ is the sample standard deviation, $n$ is the number of measurements (in our case, $n = 10$), and $\alpha$ represents the significance level (chosen as $\alpha = 0.05$). The values $t_{n-1,\frac{\alpha}{2}}$ and $t_{n-1,1-\frac{\alpha}{2}}$ are the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of the t-distribution with $n - 1$ degrees of freedom, respectively. We selected $n = 10$ to balance the need for statistical reliability with practical constrains such as time and is commonly used in many experimental studies. This number ensures that our results are statistically significant while remaining feasible.

This approach approximates a 95% confidence interval, giving insight into the precision of our metric estimates and the associated uncertainty.

## 4 Experiments

This section details the experiments conducted to address the relevant psychological inquiries identified by our expert psychologist. These carefully designed experiments aim to develop predictive tools with the goal of estimating emotional stability without the conventional reliance on direct measurement through psychological questionnaires. Our objective is to develop a path towards predictive models capable of discerning emotional stability, thereby minimizing the need for extensive sets of psychological surveys. We identify the best machine learning model for predicting emotional stability based on the scores of various psychological variables measured through different tests.

Initially, we evaluated the information each variable provides for predicting emotional stability. Based on the importance of these variables (the amount of information each variable contributes to predicting emotional stability), we developed new machine learning models using only the five most informative variables. These models aim to achieve the same predictive accuracy as the models using all eighteen variables. This approach would significantly reduce the number of tests required, from the original set to a maximum of three, thereby restructuring the evaluation process while maintaining predictive accuracy. These three tests would be 1) The Personal and Organizational Quality Assessment Revised 4 Scale (POQA-R4) (comprises four specific subscales: Emotional Vitality, Emotional Stress, Organizational Stress, and Physical Stress), 2) the Rosenberg Self-Esteem Scale and 3) the Short Dispositional Flow Scale.

With the aim of ensuring robust and reliable results, we employed a nested cross-validation procedure with 10 folds at the outer level and 5 folds at the inner level. This methodology, unlike non-nested cross-validation, allows us to obtain an unbiased estimate of model performance. In each outer iteration, the dataset was divided into partitions for training and testing. Within each outer iteration, we perform an inner cross-validation to obtain the best hyperparameters for the model. Subsequently, we evaluated the model's various metrics on the test set, providing an assessment of its performance on unseen data during training. The entire process is depicted in **Figure 2**.

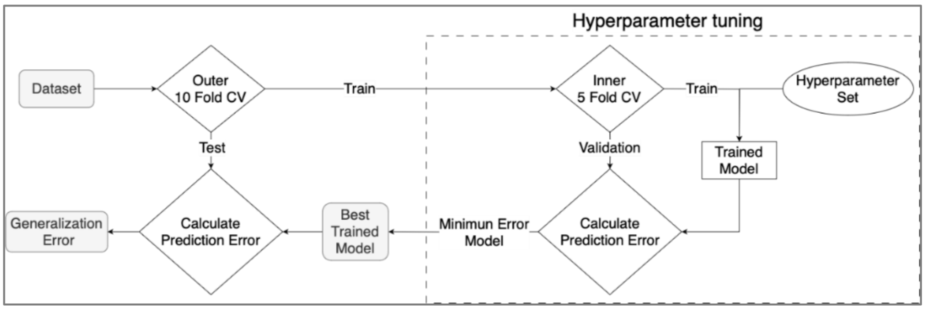


**Figure 2.** Flow chart of the experimental process of nested cross-validation with 10 outer folds and 5 inner folds for model evaluation and hyperparameter tuning.

Next, we present each of the models' subsets of hyperparameters that were adjusted, as well as the range of values explored. Note that LR does not require adjustments to its hyperparameters.

We calibrated the BR parameters to control the shape of the distribution ($\alpha_1, \alpha_2$) and the precision parameters associated with the distribution of the noise ($\lambda_1, \lambda_2$). Specifically, $\alpha_1, \alpha_2, \lambda_1, \lambda_2 \in \{10^{-6}, 10^{-5}, 10^{-4}\}$.

MLP experienced testing with various hidden layer sizes, including 2, 5, 10, (2,2), (5,5), and (10,10), along with testing up to 100 neurons per layer. However, it was observed that models with higher neuron counts per layer tended to overfit significantly. Consequently, the evaluation focused on the aforementioned layers and neurons to strike a balance between model complexity and generalizability. Additionally, the activation functions relu and tang were employed to observe their influence on the MLP's learning capabilities. Adam, sgb, and lbfgs solvers were tested to determine the most suitable optimization technique. The $\alpha$ parameter, representing regularization strength, was explored across a spectrum including $\{10^{-4}, 10^{-3}, 10^{-2}\}$ values.

In the evaluation of the k-NN algorithm, the parameter $k$, denoting the number of nearest neighbours considered during classification, was tested across a range from 3 to 30 in order to understand the impact of varying $k$ on the model's predictive accuracy and sensitivity to local patterns. Besides, the weight hyperparameter was examined using two distinct approaches: uniform and distance. The uniform weight assigns equal importance to each neighbour, while the distance weight considers the inverse of the distance, giving more weight to closer neighbours.

In the evaluation of SVR, three distinct kernels were tested, namely linear, radial basis function (rbf), and polynomial (poly). Moreover, another two hyperparameters were examined: the C parameter, which regulates the trade-off between achieving a smooth decision boundary and accurately fitting the training data, with values set at $\{0.1, 1, 10\}$; and the $\epsilon$ parameter, which represents the margin of tolerance for errors in the regression model, with values $\{0.01, 0.1, 1\}$.

DT was evaluated for different splitting criteria, including squared error, Friedman MSE, absolute error, and Poisson. These criteria dictate the measures used to evaluate the quality of a split and influence the tree's overall structure. The max depth parameter, which controls the depth of the individual trees and affects their complexity and potential overfitting, was varied across $\{None, 10, 20\}$.

RF was examined by mirroring the splitting criteria (squared error, Friedman MSE, absolute error, and Poisson) and max depth ($\{None, 10, 20\}$) for DT. The number of estimators was explored across $\{100, 200, 300\}$, determining the number of individual trees in the random forest.

Finally, GB was tested across $\{100, 200, 300\}$ estimators. The maximum depth of each decision tree in the ensemble was varied within $\{3, 4, 5\}$. Furthermore, the learning rate, which regulates the contribution of each tree to the ensemble, was explored across $\{0.01, 0.1, 0.2\}$. The subsample parameter, which determines the fraction of training data used for fitting individual trees, was tested with values of $\{0.8, 1\}$.

## 5 Results

This section presents and analyses the results derived from the models implemented to predict emotional stability. As discussed in the previous sections, numerous experiments were conducted; however, due to space constraints, we will focus on the most significant ones.

The 10-fold cross-validation results are shown in Table 1. It exhibits the $R^2$ scores for each of the 10 folds across all the models implemented. By analysing this table, we can see how well each model generalises to different subsets and its ability to handle variations in the data. Overall, the observed behaviours of all models remains notably consistent across different folds, indicating a robust performance across various aspects of our dataset. However, a discernible anomaly appears in the fourth fold, where a slight deviation from the overall pattern is observed. This divergence can be attributed to specific characteristics or anomalies present in the data subset associated with the fourth fold. Hence, this fact justifies our first proposal to perform this experiment because we achieve a more complete perspective of the models' performance, and the resulting mean offers a generalised assessment of our models' predictive capabilities.

Table 1. $R^2$ scores in the 10-fold procedure for emotional stability using several ML models.

| Fold | Model | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | LR | BR | MLP | kNN | SVR | DT | RF | GB |
| 1 | 0.7783 | 0.7786 | 0.7670 | 0.6952 | 0.7750 | 0.5005 | 0.7323 | 0.7349 |
| 2 | 0.7210 | 0.7217 | 0.7122 | 0.6634 | 0.7181 | 0.4359 | 0.7034 | 0.7222 |
| 3 | 0.7356 | 0.7356 | 0.7385 | 0.6583 | 0.7311 | 0.3911 | 0.7056 | 0.7134 |
| 4 | 0.6441 | 0.6441 | 0.6476 | 0.5880 | 0.6382 | 0.4205 | 0.6283 | 0.6382 |
| 5 | 0.7394 | 0.7388 | 0.7397 | 0.6742 | 0.7382 | 0.5369 | 0.7172 | 0.7206 |
| 6 | 0.7323 | 0.7320 | 0.7300 | 0.6754 | 0.7312 | 0.5508 | 0.7213 | 0.7232 |
| 7 | 0.6630 | 0.6634 | 0.6642 | 0.6064 | 0.6646 | 0.4887 | 0.6723 | 0.6750 |
| 8 | 0.6837 | 0.6839 | 0.6846 | 0.6190 | 0.6799 | 0.3956 | 0.6655 | 0.6761 |
| 9 | 0.7594 | 0.7595 | 0.7487 | 0.6685 | 0.7543 | 0.4899 | 0.7477 | 0.7511 |
| 10 | 0.7174 | 0.7179 | 0.7156 | 0.6828 | 0.7127 | 0.4865 | 0.7292 | 0.7108 |
| Mean | 0.717±0.030 | 0.718±0.030 | 0.714±0.028 | 0.653±0.026 | 0.714±0.030 | 0.470±0.040 | 0.702±0.026 | 0.707±0.024 |

After the preceding table, the following Table 2 summarizes the outcomes of the cross-validation procedure. This table presents the prediction results of Mean Squared Error (MSE) and R2, utilizing all the predictors during both the training and test phases. The most favourable R2 scores hover around 0.71 on the test, which is a noteworthy achievement. In addition, by considering the confidence interval, it becomes evident that both DT and kNN exhibit notably inferior performance compared to other models. In addition to this, there is a discernible issue of overfitting, as indicated by the training scores being significantly superior to those in the testing set. This phenomenon is particularly pronounced in the kNN model, which indicates the model's tendency to overfit itself to the training data. Similarly, the tree-based models also displayed signs of overfitting because their training scores are significantly better than the testing set.

Table 2. Prediction results, MSE and R2, for Emotional Stability using all variables in the training and test sets across different models.

| Model | Train | | Test | |
|---|---|---|---|---|
| | MSE ($10^{-2}$) | $R^2$ | MSE ($10^{-2}$) | $R^2$ |
| LR | 0.913±0.010 | 0.728±0.003 | 0.937±0.087 | 0.717±0.030 |
| BR | 0.913±0.10 | 0.728±0.003 | 0.936±0.086 | 0.718±0.030 |
| MLP | 0.899±0.010 | 0.732±0.003 | 0.948±0.078 | 0.714±0.028 |

| | | | | |
|---|---|---|---|---|
| kNN | 0.000±0.000 | 1.000±0.000 | 1.151±0.074 | 0.653±0.026 |
| SVR | 0.920±0.009 | 0.726±0.003 | 0.947±0.085 | 0.714±0.030 |
| DT | 0.366±0.041 | 0.891±0.012 | 1.764±0.149 | 0.470±0.040 |
| RF | 0.285±0.068 | 0.915±0.020 | 0.988±0.074 | 0.702±0.026 |
| GB | 0.630±0.52 | 0.812±0.015 | 0.974±0.067 | 0.707±0.024 |

To evaluate the significance of the input variables, the next step excludes kNN and DT models due to their unsatisfactory performance in generalization. These models demonstrated inadequate predictive abilities when applied to unseen data, prompting their exclusion from this evaluation. This decision aims to ensure a more focused and reliable analysis of the remaining models. Given that both approaches—one using all variables and the other with a reduced set—yielded similar results, it is evident that emotional stress is a key variable. Emotional stress consistently appeared among the top five most important variables, reinforcing its significance. Had this not been the case, the performance of the models with fewer variables would have been notably poorer. Following closely in second place, self-esteem emerges as a significantly influential factor. Subsequently, emotional vitality occupies the third position in terms of importance. When considering these three variables together with physical stress and Flow, they collectively contribute to 50% of the overall importance.

Table 3. Prediction results, MSE and R2, for Emotional Stability using a reduced set of variables in the training and test sets across different models.

| Model | Train | | Test | |
|---|---|---|---|---|
| | MSE ($10^{-2}$) | $R^2$ | MSE ($10^{-2}$) | $R^2$ |
| LR | 0.990±0.011 | 0.705±0.003 | 0.999±0.100 | 0.699±0.033 |
| BR | 0.990±0.011 | 0.705±0.003 | 0.998±0.100 | 0.699±0.033 |
| MLP | 0.982±0.011 | 0.707±0.004 | 0.997±0.100 | 0.700±0.033 |
| SVR | 0.992±0.011 | 0.704±0.003 | 1.002±0.101 | 0.698±0.033 |
| RF | 0.473±0.005 | 0.859±0.002 | 1.069±0.070 | 0.677±0.027 |
| GB | 0.856±0.037 | 0.745±0.011 | 1.037±0.081 | 0.689±0.028 |

In the dominance analysis, the results showed that emotional stress and self-esteem demonstrated complete dominance over all other factors included in the study (see Figure 3). This outcome is evident because these two variables consistently exhibit better performance across all subsets. In other words, Emotional Stress emerges as the most informative variable, followed closely by Self-esteem, which also holds substantial importance. Additionally, Emotional Vitality stands out prominently among the variables. Including Physical Stress and Flow alongside these variables aggregates more than 50% of the total importance.

These findings highlight the critical roles of Emotional Stress and Self-esteem in our emotional stability model, where their contributions exceed those of other variables across various linear models. Furthermore, it is important to note that this analysis is grounded in LR, a model chosen for its simplicity and commendable performance in training data without displaying overfitting issues. For our target variable, we included the dominance analysis by computing $2^{18} - 1$ (262.143) models. We employed LR because of its interpretability, it yielded good results, and the training times were feasible. For the other models, this approach would not be viable due to increased complexity and significantly longer training times.

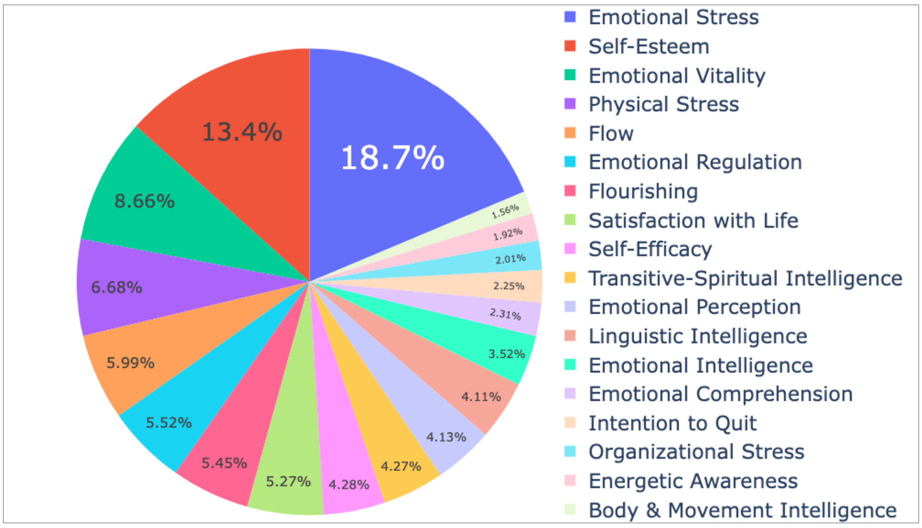


Figure 3. Feature importance to predict emotional stability.

Finally, for the purpose of comparing both approaches (with all variables and the simplified one), we computed the adjusted R2 for each. This analysis revealed that the models with the complete set of variables and the reduced set were equivalent in terms of their predictive performance. We refer to them as equivalent because we calculated the $R^2_{adjusted}$ for both approaches separately. Due to the different number of variables, direct comparison using $R^2$ is not appropriate. Once done, both approaches yield the same $R^2_{adjusted}$, indicating similar predictive accuracy.

In Figure 4, it is evident that the adjusted R2 improves as more variables are included. In particular, the model with three variables exhibits a noteworthy level of fit, with the highest point on the graph obtained when using 12 variables. Interestingly, the difference in predictive performance between a model employing three predictors and a model that incorporates the full set of 18 predictors is minimal. Even when reducing the number of predictors from 18 to just 5, the model's ability to predict emotional stability remains practically equivalent. This further reinforces the potential for achieving robust prediction with a significantly smaller set of variables, restructuring the evaluation process and potentially reducing participant burden. This suggests that a model with only three predictors provides a satisfactory level of performance, as well as the efficiency of a more concise variable selection.

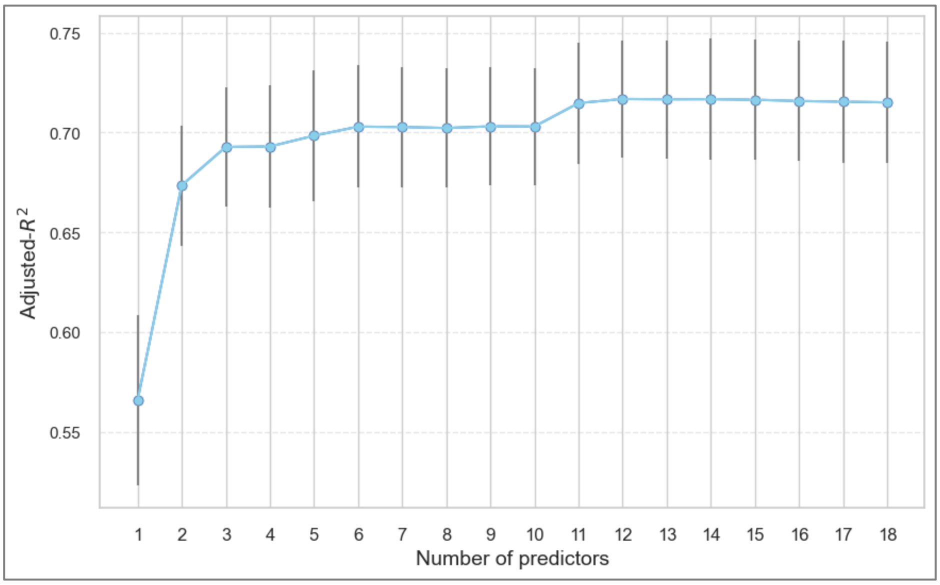


Figure 4. R2 adjusted according to the number of predictors.

Following this, our next and last step involves training models based on a reduced set of variables that are carefully selected based on their predictive dominance, as illustrated in Figure 3. Specifically, these selected variables were derived from the dominance experiment, encompassing those collectively contributing to over 50% dominance. The results obtained from this selection process are consolidated and presented in Figure 4. Upon scrutinizing the table results, a discernible trend emerges where the exclusion of certain psychological variables corresponds to a decrease in R2. This observation substantiates the notion that these variables genuinely play a crucial role in prediction. Although there is a reduction in precision, it is noteworthy that the impact is not highly significant.

## 6 Conclusions

The purpose of this research was to predict emotional stability using a training dataset and a machine learning method. This approach involves extracting information from the training set, allowing the model to identify patterns and relationships in the data. The key contribution of this study lies in the capability to apply these models to a test set to predict the emotional stability of individuals who had not been previously observed or assessed. In other words, by training our data, we can uncover information that remained hidden without the need for direct measurements.

Our motivation to employ machine learning models is based on the premise that highly complex variables, which result from the interaction of multiple factors, require more generalized and flexible approaches (Breiman, 2001). In this context, we assert that various factors influence emotional stability. As highlighted in the introduction, there is ample evidence emphasizing the relevance and connection of emotional stability, not only with mental well-being but also with fundamental aspects such as occupational performance (Frye, 2000; Kobylińska et al., 2020). These studies have identified numerous variables related to emotional stability. Although classical approaches in psychology studies are based on correlations (Costa & McCrae, 1980; Hills & Argyle, 2001) or have explored prospective relationships between variables using multivariate methods, such as structural equations, they may have limitations inherent to the nature and constraints of these statistical models. Nevertheless, we acknowledge in this work that these studies have been crucial to advance the understanding of emotional stability. However, we believe that these classical approaches can be significantly enriched using machine learning and artificial intelligence methods, representing a promise in our era, thanks to their differential potential in prediction based on learning from the primary dataset.

For these reasons, we carefully selected a set of variables that had previously been shown to be significantly related to emotional stability for training. The training measures or data for these variables were obtained from the responses of a group of 2203 individuals to a series of psychometric questionnaires that exhibited good reliability and validity. We opted for this approach to maintain some proximity to conventional studies, and, among other things, in addition to predicting emotional stability, we attempted to quantify the weight or contribution of these variables to emotional stability.

In this study, we introduced an ML approach to forecast the psychological variable "emotional stability". The findings revealed that DT exhibited the poorest predictive performance, with kNN often trailing behind in unfavourable outcomes. The remaining models demonstrated comparable performance metrics in the test. Notably, kNN and tree-based models demonstrated substantial overfitting. The variable selection process led to an enhancement in the performance of the overfit models. Models using both reduced and complete sets of variables were found to be statistically equivalent.

The results obtained from the machine learning models provide an essential contribution to understanding emotional stability. It was observed that 12 out of the 18 variables evaluated in the training generated a model that predicted emotional stability with an adjusted R2 of 0.72, and seven of our variables already produced a model with an adjusted R2 of 0.70.

Firstly, these models highlighted the significance of the emotional stress variable, a subscale of POQA-R4 (Barrios-Choplin & Atkinson, 2004). Defined as emotional discord that diminishes perceived quality of life and jeopardizes health and well-being, this subscale measures the extent to which individuals report negative emotions, experience difficulty controlling them, and feel that they negatively impact the quality and effectiveness of their life experiences. It is noteworthy that across all the test sets, this variable has consistently proven more relevant compared to others. Additionally, it exhibited a relatively higher average weight (18.7%), suggesting that it is one of the most robust predictors of emotional stability. Although it might be theoretically argued that emotional stability or neuroticism overlap with the definition of emotional stress, the measures' content differs, and there is evidence indicating separate effects for perceived psychological stress not shared with neuroticism (Kaiser & Ozer, 1997; Rietschel et al., 2014).

Another significant predictor was self-esteem, confirming previous findings such as those in (Bajaj et al., 2019), highlighting a positive relationship between emotional stability and self-esteem. Our study shows that self-esteem contributes with a relatively higher weight of 13.4% in the training sets. Therefore, it is one of the variables with greater dominance in predicting emotional stability. Self-esteem, along with emotional vitality, physical stress, flow, emotional regulation, and flourishing, constitute the set of most dominant variables.

Furthermore, upon analysing the results, we observed that the three initial variables already form a model that allows us to understand which variables could predict emotional stability. However, we believe that models that incorporate up to 14 variables can provide a more detailed explanation of the studied phenomenon. Variables such as transitive-spiritual intelligence or linguistic intelligence, the ENII-33 subscale, and emotional understanding, a subscale of the TMMS-24, exhibit modest but significant contributions according to the results obtained.

In conclusion, the training set of these data has provided a more detailed understanding of the predictive variables related to emotional stability. The strength of this work lies in aspects such as the breadth and heterogeneity of the sample, the meticulous selection of variables, and, especially, in machine learning-based strategies to explore relationships that might go unnoticed in conventional statistical analysis. Although we acknowledge the potential limitations associated with psychometric scales, we believe that these analyses can be enriched by integrating data from other sources.

In future research, it will be interesting to explore the implementation of a multi-model predictor to analyse systematic patterns in errors among various models and investigate potential improvements in predictive performance. Additionally, integrating association rules into the predictive framework could reveal complex relationships within the dataset, offering a better understanding of factors influencing emotional stability prediction. Furthermore, exploring more complex models, such as convolutional artificial neural networks could be a promising approach to capture patterns and relationships, which could potentially enhance the accuracy of emotional stability predictions.

# 7 Abbreviations

| | |
|---|---|
| BR | Bayesian Ridge |
| DT | Decision Tree |

| | |
|---|---|
| GB | Gradient Boosting |
| k-NN | k-Nearest Neighbour |
| LR | Linear Regression |
| MLP | Multi-Layer Perceptron |
| MSE | Mean Squared Error |
| $R^2$ | R-squared |
| RF | Random Forest |
| SVR | Support Vector Regression |

## 8 Data availability statement

Data available upon request to the corresponding author.

## 9 Competing interests

The authors declare that there is no conflict of interest.

## 10 Authors contribution statement

All authors have contributed equally to this work.

## 11 Ethical and informed consent for data used

Ethical approval and informed consent were obtained for all data used in this study.

## 12 Acknowledgements

The authors extend their heartfelt gratitude to RH Asesores Improving, Talent Development Institute for its invaluable contributions and support throughout the research process. Furthermore, we wish to emphasize that this research was undertaken without specific funding, and the authors conducted and concluded the study independently.

## 13 Bibliography


Alessandri, G., Perinelli, E., De Longis, E., Schaufeli, W. B., Theodorou, A., Borgogni, L., . . . Cinque, L. (2018). Job burnout: The contribution of emotional stability and emotional self-efficacy beliefs. *Journal of occupational and organizational psychology*, *91*(4), 823-851.

Atienza, F. L., Pons, D., Balaguer, I., & García-Merita, M. (2000). Propiedades psicométricas de la Escala de Satisfacción con la Vida en adolescentes. *Psicothema*, 314-319.

Azen, R., & Budescu, D. V. (2006). Comparing predictors in multivariate regression models: An extension of dominance analysis. *Journal of Educational and Behavioral Statistics*, *31*(2), 157-180.

Baessler, J., & Scharzer, R. (1993). Spanish Adaptation of the General Self-Efficacy Scale. *Ansiedad y Estrés*, *2(1)*, 1-8.

Bai, R., Xiao, L., Guo, Y., Zhu, X., Li, N., Wang, Y., . . . Yu, X. (2021). Tracking and monitoring mood stability of patients with major depressive disorder by machine learning models using passive digital data: prospective naturalistic multicenter study. *JMIR mHealth and uHealth*, *9*(3), e24365.

Bajaj, B., Gupta, R., & Sengupta, S. (2019). Emotional stability and self-esteem as mediators between mindfulness and happiness. *Journal of Happiness Studies*, *20*, 2211-2226.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1992). Social cognitive theory of social referencing. In *Social referencing and the social construction of reality in infancy* (pp. 175-208). Springer.

Barrios-Choplin, R., & Atkinson, M. (2004). Personal and Organizational Quality Assessment–Revised 4 Scale (52 questions). Boulder Creek, CA: Institute of HeartMath. Retrieved February 16, 2017. In.

Biswas-Diener, R. (2008). The new (smiley) face of economics.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

Buchanan, T. M., & Reilly, P. M. (2019). The impact of HeartMath resiliency training on health care providers. *Dimensions of Critical Care Nursing*, *38*(6), 328-336.

Caprara, G., Vecchione, M., Barbaranelli, C., & Alessandri, G. (2013). Emotional stability and affective self–regulatory efficacy beliefs: Proofs of integration between trait theory and social cognitive theory. *European Journal of Personality*, *27*(2), 145-154.

Cohrdes, C., & Mauz, E. (2020). Self-efficacy and emotional stability buffer negative effects of adverse childhood experiences on young adult health-related quality of life. *Journal of Adolescent Health*, *67*(1), 93-100.

Costa Jr, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment*, *64*(1), 21-50.

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. *Journal of personality and social psychology*, *38*(4), 668.

Crowe, M. L., LoPilato, A. C., Campbell, W. K., & Miller, J. D. (2016). Identifying two groups of entitled individuals: Cluster analysis reveals emotional stability and self-esteem distinction. *Journal of personality disorders*, *30*(6), 762-775.

Csikszentmihalyi, M., & Larson, R. (2014). *Flow and the foundations of positive psychology* (Vol. 10). Springer.

Chen, Y., & Liu, Y. (2021). Which risk factors matter more for psychological distress during the covid-19 pandemic? An application approach of gradient boosting decision trees. *International Journal of Environmental Research and Public Health*, *18*(11), 5879.

Chowdhury, B., & Das, N. (2022). Predictive Analysis of Child's Mental Health/Psychology During the COVID-19 Pandemic. *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2022*, *480*, 183.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of personality assessment*, *49*(1), 71-75.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D.-w., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social indicators research*, *97*, 143-156.

Extremera, N., & Fernández-Berrocal, P. (2004). La importancia de desarrollar la inteligencia emocional en el profesorado. *Revista iberoamericana de educación*, *34*(3), 1-9.

Farahani, H., Azadfallah, P., & Rahiminezhad, A. (2020). Comparison of Logistic Regression (LR), Fuzzy Rule-Based Classification (FRBC), and Multilayer Perceptron (MLP) Classifier using Anxiety Data.

Fernández-Berrocal, P., Extremera, N., & Ramos, N. (2004). Validity and reliability of the Spanish modified version of the Trait Meta-Mood Scale. *Psychological reports*, *94*(3), 751-755.

Frye, C. M. (2000). *The effect of emotional stability on job satisfaction: A meta-analysis*. The University of Iowa.

Getzels, J. W., & Csikszentmihalyi, M. (1976). The creative vision: A longitudinal study of problem finding in art. *(No Title)*.

Godoy-Izquierdo, D., Vélez, M., Rodríguez, Z., & Jiménez, M. (2009). Flow in sport: Concept, evaluation and empirical findings. *XII Andalusian Congress of Psychology of Physical Activity and Sport*.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.

Henninger, M., Debelak, R., Rothacher, Y., & Strobl, C. (2023). Interpretable machine learning for psychological research: Opportunities and pitfalls. *Psychological methods*.

Hills, P., & Argyle, M. (2001). Emotional stability as a major dimension of happiness. *Personality and individual differences*, *31*(8), 1357-1364.

Jackson, S., Eklund, B., & Martin, A. (2012). The Flow Scale Manual. *Mind Garden, Inc.*

Kaiser, R. T., & Ozer, D. J. (1997). Emotional stability and goal-related stress. *Personality and individual differences*, *22*(3), 371-379.

Kjell, O. N., & Diener, E. (2021). Abbreviated three-item versions of the satisfaction with life scale and the harmony in life scale yield as strong psychometric properties as the original scales. *Journal of personality assessment*, *103*(2), 183-194.

Kobylińska, D., Zajenkowski, M., Lewczuk, K., Jankowski, K. S., & Marchlewska, M. (2020). The mediational role of emotion regulation in the relationship between personality and subjective well-being. *Current Psychology*, 1-14.

Koffi, B. J. B. (2020). *Inferring Big Five personality factors using text analysis its assessment and impact on prosocial behavior and IS security compliance*. The University of Texas at Arlington.

Martín-Albo, J., Núñez, J. L., Navarro, J. G., & Grijalvo, F. (2007). The Rosenberg Self-Esteem Scale: translation and validation in university students. *The Spanish journal of psychology*, *10*(2), 458-467.

Maydeu-Olivares, A., Shi, D., & Fairchild, A. J. (2020). Estimating causal effects in linear regression models with observational data: The instrumental variables regression model. *Psychological methods*, *25*(2), 243.

Namazkhan, M., Albers, C., & Steg, L. (2020). A decision tree method for explaining household gas consumption: The role of building characteristics, socio-demographic variables, psychological factors and household behaviour. *Renewable and Sustainable Energy Reviews*, *119*, 109542.

Olaru, G., van Scheppingen, M. A., Bleidorn, W., & Denissen, J. J. (2023). The link between personality, global, and domain-specific satisfaction across the adult lifespan. *Journal of personality and social psychology*.

Omoregie, J., & Carson, J. (2023). Reductive mechanisms for unwanted intrusive thoughts: exploring affectivity in clinical and non-clinical samples. *Mental Health and Social Inclusion*, *27*(1), 51-65.

Orrù, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2020). Machine learning in psychometrics and psychological research. *Frontiers in psychology*, *10*, 2970.

Pegalajar, M., Ruiz, L., Pérez-Moreiras, E., Boada-Grau, J., & Serrano-Fernandez, M. (2023). An Intelligent Approach Using Machine Learning Techniques to Predict Flow in People. *Big Data and Cognitive Computing*, *7*(2), 67.

Peiqing, H. (2022). Multidimensional state data reduction and evaluation of college students' mental health based on SVM. *Journal of Mathematics*, *2022*, 1-11.

Pérez-Moreiras, E. (2020). LA INTELIGENCIA Y EL COACHING ENERGÉTICOS: Una aproximación basada en la evidencia desde la Psicología para el Desarrollo de la Inteligencia Energética (Energetic Intelligence), el Fluir (Flow) y el Florecer (Flourishing) | Energetic Intelligence and Coaching: An evidence-Based Approach from Psychology for the Development of Energetic Intelligence, Flow, and Flourishing.

Philipp, M., Rusch, T., Hornik, K., & Strobl, C. (2018). Measuring the stability of results from supervised statistical learning. *Journal of Computational and Graphical Statistics*, *27*(4), 685-700.

Rietschel, L., Zhu, G., Kirschbaum, C., Strohmaier, J., Wüst, S., Rietschel, M., & Martin, N. G. (2014). Perceived stress has genetic influences distinct from neuroticism and depression. *Behavior genetics*, *44*, 639-645.

Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). *Acceptance and commitment therapy. Measures package*, *61*(52), 18.

Rosenberg, M. (2015). *Society and the adolescent self-image*. Princeton university press.

Sanjuán, P., Pérez, A., & Bermúdez, J. (2000). Escala de Autoeficacia General: Datos psicométricos de la adaptación para población Española [The General Self-efficacy Scale: Psychometric data from the Spanish adaptation]. Psicothema, 12 (supl. 2), 509-513. In.

Schippers, M. C., & Hogenes, R. (2011). Energy Management of People in Organizations: A Review and Research Agenda. *Journal of Business and Psychology*, *26*(2), 193-203.

Serrano-Fernández, M. J., Pérez-Moreiras, E., Boada-Cuerva, M., Assens-Serra, J., & Boada-Grau, J. (in press). Personality, self-efficacy and self-esteem as predictors of psychological well-being of workers: the flourishing scale (5-FS). *International Journal of Business Environment*. https://doi.org/10.1504/IJBE.2023.10058504

Şevgin, H., & Eranıl, A. K. (2023). Investigation of Turkish Students' School Engagement through Random Forest Methods Applied to TIMSS 2019: A Problem of School Psychology. *International Journal of Psychology and Educational Studies*, *10*(4), 896-909.

Stuke, H., Priebe, K., Weilnhammer, V. A., Stuke, H., & Schoofs, N. (2023). Sparse models for predicting psychosocial impairments in patients with PTSD: An empirical Bayes approach. *Psychological trauma: theory, research, practice, and policy*, *15*(1), 80.

Tewfik, B. A., Kim, D., & Patil, S. V. (2023). The ebb and flow of job engagement: Engagement variability and emotional stability as interactive predictors of job performance. *Journal of Applied Psychology*.

Vigil-Colet, A., Morales-Vives, F., Camps, E., Tous, J., & Lorenzo-Seva, U. (2013). Desarrollo y validacion de las Escalas de Evaluacion Global de la Personalidad (OPERAS). *Psicothema*, *25*(1), 100-107.

Wardenaar, K. J., Riese, H., Giltay, E. J., Eikelenboom, M., van Hemert, A. J., Beekman, A. F., . . . Schoevers, R. A. (2021). Common and specific determinants of 9-year depression and anxiety course-trajectories: A machine-learning investigation in the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Affective Disorders*, *293*, 295-304.

Yuan, C., Wu, J., Li, H., & Wang, L. (2018). Personality recognition based on user generated content. 2018 15th International Conference on Service Systems and Service Management (ICSSSM),