Journal of Universal Computer Science, vol. 30, no. 10 (2024), 1316-1342
submitted: 25/11/2023, accepted: 5/3/2024, appeared: 28/9/2024

Smart healthcare: developing a pattern to predict the stress and anxiety among university students using machine learning technology

Farhad Lotfi*
(University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
https://orcid.org/0000-0002-3308-6977, farhadlotfi1990@gmail.com)

Branka Rodić
(College of Health Sciences, Academy for applied studies Belgrade, Belgrade, Serbia
https://orcid.org/0000-0003-1965-3899, brodic@gmail.com)

Aleksandra Labus
(University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
https://orcid.org//0000-0002-7716-5845, aleksandra@elab.rs)

Zorica Bogdanović
(University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
https://orcid.org/0000-0003-4799-1588, zorica@elab.rs)

Abstract: Background: Anxiety among students has become a fairly major problem. In the current era, Machine Learning (ML) can be used as a quick technology to predict students' anxiety with the high-level accuracy.
Objectives: This research aims to predict university students' anxiety by using supervised learning algorithms with providing pertinent feedback.
Methods: A total of 231 students from the University of Belgrade filled out the standard questionnaire called the State-Trait Anxiety Inventory (STAI). In addition, deeper information related to students' anxiety like physical activity, Grade Point Average (GPA), and smoking cigarettes were collected. The Linear Regression algorithm was chosen to examine STAI using Python.
Results: Linear regression as an appropriate algorithm was exploited for this purpose. The accuracy metric obtained by using the Mean Absolute Error (MAE), was 7.86 for state anxiety and 5.68 for trait anxiety. In addition, the Mean Squared Error (MSE) has also been calculated with state anxiety at 7.80, and trait anxiety at 9.66. Moreover, to find the factor with the highest impact after training, a regression analysis method (LASSO) was used. K-Nearest Neighbour (KNN) algorithm also checked the accuracy of training by overfitting and underfitting.
Conclusion: The purpose of this study was the analysis of anxiety factors with the highest impact as well as the analysis of the STAI by linear regression to improve a smart healthcare model by discovering an acceptable output with the highest accuracy.

Keywords: Smart Healthcare, Anxiety, Machine Learning, Supervised Learning, Linear Regression, Quantitative Analysis

Categories: H.3.1, H.3.2, H.3.3, H.3.7, H.5.1
DOI:10.3897/jucs.116174

1 Introduction

Smart healthcare is a good choice for the new lifestyle and is known as one of the most important parts of the smart city [Pramanik et al., 2017, Souza and Soares, 2023]. One of the applications of smart healthcare is monitoring people's behaviours as well as mental illnesses (e.g., stress and anxiety) [Caragliu et al., 2009]. Stress can be a temporary turmoil, while anxiety can bother a person for a while and cause irreparable damage [Barton et al., 2014; Stojanović et al., 2020]. The measurement of anxiety examines physiological, behavioural, and psycho-cognitive aspects among students [Spielberger et al., 1971]. Many factors can exacerbate stress and anxiety among students, such as exam time, student life, living expenses, and similar issues [Huang and Wan, 2013]. Studies have also shown that stress and anxiety can exacerbate physical and mental well-being conditions among students because of vulnerable times in students' lives [Sau and Bhakta, 2017, van Vuuren et al., 2021]. According to the social viewpoint, and in most developed countries, the detection of stress or anxiety in the context of electronic systems is being implemented in real-time. The level of anxiety and stress of patients are extracted through sensors and wearables so that the accuracy in these systems has reached over 90% and this accuracy could not be seen in traditional systems [Labus et al., 2021]. As mentioned, youth of different ages are vulnerable to anxiety, while on the other hand, the emergence of new technologies in their lives has positive feedback and increases their motivation [Valle et al., 2021]. Thus, using machine learning to diagnose many physical and mental illnesses like stress and anxiety is done with an acceptable accuracy. The devastating significance of COVID-19 disease on students' behaviours in recent years has been undeniable. Therefore, using ML to detect students' anxiety seems to be an effective solution, and this methodology was able to identify students those who were prone to anxiety [Alharthi, 2020]. The main purpose of this research is to predict university students' anxiety by using ML technology with high accuracy. The approach of analysing the anxiety data using machine learning algorithms has been successful.
The scientific reports also argued that approximately 35% of individuals had faced anxiety during COVID-19 all around the world [Delpino et al., 2022].

So, such reasons increased our motivation to participate in this research as fully as possible. Finding the factors with the highest impact by using a regression analysis method (LASSO) on both state and trait anxiety after training was one of the vital gaps in this research.

This paper is organised as follows: section two provides the literature review. Section three provides the research approach, followed by the research methodology. In section four we present the results of the research, and in section five we discuss the research outcomes. The conclusion section is presented in the sixth section.

2 Literature review

2.1 Measuring the anxiety

Anxiety has been diagnosed as an emotional reaction that may be pathological among individuals [Stojanović et al., 2020]. Fortunately, anxiety is known as one of the diseases that can be treated with self-management [Pullyblank et al., 2022]. Besides, measuring, managing, and controlling anxiety requires a multi-year process [Bokma et al., 2020]. Machine learning technology has been categorised into intelligent systems because this allows computers to train from datasets to improve models' performance [Das et al., 2024]. However, using intelligent systems to detect stress and anxiety has more advantages. Nowadays, researchers have been able to detect the users' perceived anxiety scale by evaluating millions of tweets [Gruda and Hasan, 2019]. Previous research confirms that users have expressed their desire to use social media, which helps to identify their personality [Youyou et al., 2015]. Anxiety detection by micro-blogs has become very accurate. For instance, researchers with 500 million tweets in 2017 more than 328 million active users, and 100 million daily users have found that analysing this amount of big data can be beneficial in predicting anxiety [Aslam, 2018, Gruda and Hasan, 2019]. In addition to smart technologies, standard questionnaires, measuring skin temperature, heart rate, and blood pressure could be useful in detecting anxiety and stress [Hermens et al., 2014].

2.2 Predicting anxiety in smart healthcare systems

Today, Mobile Health (M-Health) applications with smart mobile phones have taken positive steps to reduce anxiety, and this technology has also provided real-time responsiveness for young people's diseases [Pramana et al., 2014]. For instance, to improve a new lifestyle without anxiety, it has been argued that the use of Electronic Health (E-Health) and m-health have a direct impact on improving life conditions [Hwang and Nam, 2022]. The results of a survey with more than 1000 samples regarding lifestyle improvements showed that 47.2% of applicants used e-health technology, 23.2% used m-health technology, and 10.7% used wearable devices [Leung and Chen, 2019]. However, monitoring smart healthcare is running in the context of the Internet of Things (IoT), and in addition to speeding up diagnosis and treatment, it has also reduced treatment costs [Sujith et al., 2022]. Besides, smart healthcare interventions extend beyond this point, and due to the increasing development of smart devices like smartwatches, smart bands as well as mobile health programmes, the younger generation has enthusiastically adopted such technologies [Du et al., 2020, Fan and Zhao, 2022, Khanjari et al., 2021, Pai et al., 2021]. Research findings confirm that the existence of e-health as well as m-health has increased the level of physical activities among young people and motivated them to perform physical activities [McIntosh et al., 2017]. In the United States in 2014, almost half of students felt frustrated, according to a report by the American College Health Association. Moreover, by using e-health, young people's anxiety was detected better at 4.5% compared to previous studies [Jacobson et al., 2015].

2.3 Machine learning and algorithms in predicting the stress and anxiety

Predicting anxiety and stress by using ML has been increased because of its accuracy and high sensitivity in real-time [Sau and Bhakta, 2017]. As mentioned in previous studies, the analysis of anxiety data in the context of ML can be very economical [Balogh et al., 2015, Jacobson et al., 2021]. Also, ML technology has been able to analyse quite complex data on stress and anxiety as well [Shin et al., 2013]. Today, using ML to predict anxiety with high accuracy has been able to improve the performance of experts in this regard. As a result of the research, multivariate and Logistic Regression have provided acceptable results. Data were analysed with six cognitive-behavioural tasks. ML-based prediction of anxiety disorders showed 76.81% specificity with 69.66% sensitivity among 25 individuals with no psychiatric diagnosis [Richter et al., 2021]. Thus, the most important point of this research has been enhancing the confidence between patient and clinician to achieve an acceptable precision and individually tailored therapy [Richter et al., 2021].

Recently, research findings have demonstrated that the prediction of stress and anxiety among college students by using Logistic Regression and Support Vector Machine (SVM) showed better performance in classifying such mental illnesses than other ML algorithms [Daza, et al., 2023]. For this purpose and for determining stress and anxiety, they used 29 research articles in their systematic review from 7 databases so they identified 43 psychological features, 15 demographic features, and some socio-economic, environmental, academic, and other features. Finally, for example, Saudi Arabian students' anxiety (N= 3017) using SVM achieved 100.00% accuracy [Daza, et al., 2023]. Physical activities like swimming, cycling, and similar activities are among vital factors in reducing stress and anxiety, while concurrently fostering overall health and enhancing physical strength [Abbas et al., 2019]. In the United States, Deep Learning (DL) by using mobile wearables has been able to predict the escalation of anxiety among young people with a degree of accuracy deemed acceptable (sensitivity= 84.6%, accuracy = 68.7%, and AUC = 0.696, CI [0.598, 0.793]) [Jacobson et al., 2021]. The deep learning models predicted a long-term generalised anxiety disease between 17 and 18 ages from data on sleeping habits and physical activity during the day. Furthermore, the data extracted from a phone-based interview showed increasing anxiety disorders among 265 participants [Jacobson et al., 2021]. Thus, the feedback obtained from the output of the analysis of machine learning techniques in the field of anxiety was the discovery of new algorithms, models, and methods for anxiety screening. Also, some psychological topics such as anxiety have been diagnosed through image processing and deep learning algorithms, in which case the disease can be controlled more accurately [Sau and Bhakta, 2019]. For instance, in the automatic diagnosis of anxiety, a health-related predictive model, with 510 patients, was validated by the ML algorithms and achieved a prediction accuracy of 89% by Random Forest (RF) [Sau and Bhakta, 2017]. Women between 18 and 35 years old in Bangladesh have been suffering from anxiety disorders more than men, according to a study by the World Health Organisation (WHO). Anxiety data were collected through a standard questionnaire with 35 questions among patients from Dhaka Medical Hospital and Rangpur Medical College. Convolutional Neural Network (CNN) was used to diagnose anxiety so that results were based on different measurement criteria, the CNN algorithm had the most favourable effect with 96% accuracy [Ahmed et al., 2020]. On the one hand, the detection of anxiety among university students via ML algorithms has been expanded. For example, the research done among 127 engineering students in India resulted in an accuracy of 78.9% by using the random forest. This research used 40 questions via a questionnaire to analyse the Indian students' anxiety with categories of demographic, Probable Causes of Anxiety, and Probable Effects of Anxiety. One of the goals of this research, recognising the risk of a severe mental breakdown [Bhatnagar et al., 2023]. Insufficient attention is accorded to a noteworthy issue in anxiety management, particularly among younger ages, thereby leading to potentially unsafe consequences. ML algorithms with 80% accuracy and 93% sensitivity, using audio data, have been able to predict childhood anxiety [McGinnis et al., 2019a]. It was found that children affected by anxiety speak in low voices and with deviation. Also, the system's results showed children and teenagers experiencing more stress in clinical environments [McGinnis et al., 2019a]. In addition, according to [Malik and Khan, 2023], most college students who are suffering from stress and anxiety cannot figure out such mental illnesses. These authors collected data on students' mental illnesses like stress and anxiety via a standard questionnaire including 21 questions, and analysed data by using 5 ML algorithms like RF, Decision Tree (DT), logistic regression, KNN, and Naïve Bayes. The K-nearest neighbour algorithm showed the best performance with 0.972 precision, F1-Score= 0.932, and the 0.935% of accuracy. KNN also showed the best performance of stress so that precision= 0.984, F1-Score= 0.988, and 0.992% of accuracy were achieved [Malik and Khan, 2023]. Another research reported anxiety data extracted from the STAI questionnaire and analysed this by linear regression [Gruda and Hasan, 2019].

Therefore, the linear regression algorithm has been used for our research.

2.4 Challenges of Predicting the Anxiety Based on ML & E-Health

Machine learning technology has made a significant change in predicting students' anxiety [Alharthi, 2020, Soysal et al., 2022]. A summary of relevant research in the field of predicting anxiety and stress based on e-health or the use of ML technology, including its findings and benefits, is presented in Table 1.

No	Author's Name	Challenges or Limitations	Focus → Goal
1	[Gruda and Hasan, 2019]	Tweets were generally short text and had misspelled words.	The goal of predicting future outcomes, such as popularity and social engagement.
2	[Richter *et al.*, 2021]	Most of the participants were female, that is why there was no gender balance.	Analysis focusing on anxiety to differentiate between clinical anxiety and depression disorders.
3	[Garcia-Chimeno et al., 2017]	There were not a large number of samples for the dataset.	Focusing on migraine pathologies and magnetic resonance images to automate the diagnosis.
4	[Eeden et al., 2021]	Some differences between DSM-IV- TR (A Kind of mental disorders) and DSM-	Using the ML algorithms to predict the onset of anxiety.

		5 (A tool for mental health clinicians to diagnose) for anxiety, and some diagnostic algorithms that were used in this study were slightly outdated.	
5	[Sharma and Verbeke, 2021]	Challenges: because of the small sample of data, predicting people's anxiety was quite difficult. Limitations: research funds were limited, and the lifeline database had some limitations to the choice number of biomarkers.	This research article has been focused on Dutch citizens (11081 sample size) to extract types of anxiety disorder.
6	[Demiris et al., 2020]	Challenges of self-reported anxiety. The limitation was about the training data set.	The goal of this research article was to extract features of transcribed audio communication data extracted using ML.
7	[McGinnis et al., 2019b]	The challenges and limitations focused on a huge dataset that examined internalising disorders.	This study used ML techniques to improve childhood anxiety that had just been beginning.
8	[Ihmig et al., 2020]	This research article has some challenges during data collection, and that is why three records were disregarded.	The important focus of this study was to survey the HRV – Heart Rate Variability – biofeedback intervention in the therapeutic. The goal of this research was the usage of physiological responses to predict anxiety.
9	[McGinnis et al., 2018]	There were some limitations while clinics were collecting reports from parents, children, and teachers.	Reducing the time of diagnosis, the mental well-being.
10	[Masino et al., 2019]	Physiological measures.	Prevent aggravation of behaviour.
11	[Labus et al., 2021]	Data collection from dentistry.	Managing anxiety and stress to relaxation of the mind and reduces physical tension in dentistry.
12	[Bokma et al., 2020]	Predicting the anxiety disorders were not available.	Providing individual disease period predictions.
13	[Priya et al., 2020]	Collecting data from employed and unemployed	Improving the prediction of stress and anxiety problems.

		individuals across different cultures.	
14	[Nemesure et al., 2021]	There were several limitations regarding the dataset of French college-aged students, sample size, and similar factors.	Help to identify predictors of Major Depressive Disorder (MDD) & Generalised Anxiety Disorder (GAD) to the early detection of MDD & GAD.
15	[Sau and Bhakta, 2019]	Studying a difficult case.	Periodic screening for anxiety of seafarers' using machine learning technology as a quick and automated screening procedure to identify at-risk seafarers.
16	[Daza, et al., 2023]	Confounding factors, specifically scan and gender. Detection of such illnesses during COVID-19 was also one of the limitations and challenges.	A better understanding of the features, and analysing such features of college students' stress and anxiety by using ML algorithms and functions.

Table 1: Outstanding research in the area of predicting anxiety via machine learning

Real-time responsiveness can occur using wearable technology and has been able to detect and prevent some disorders with impending stress and anxiety. The important challenges of predicting anxiety by using new methods like machine learning include data collection, the limited number of participants, and sometimes the study on rather a difficult case. One of the current study goals was to solve such challenges.

Moreover, studies in Canada have shown accruing more physical activity of a human being can reduce anxiety [Duncan et al., 2022, Omura et al., 2020]. In addition to these reasons, research findings in Germany showed that students who had anxiety were more likely to smoke [Lavallee et al., 2021]. Participants were asked to answer these questions (Table 7).

3 Research methodology

3.1 The research approach

This research article has analysed and assessed students' anxiety at the University of Belgrade. Due to the COVID-19 pandemic, most of the courses at the university were held in an online context. For this research, we were not able to measure anxiety via sensors and e-health, which is why we decided to assess students' anxiety via the standard questionnaire using ML algorithms. As such, this research completes the previous study that argued about monitoring anxiety among young people by using ML algorithms [Lotfi et al., 2022]. This methodology allows for fast assessment and analysis of anxiety. Figure 1 shows our research approach to the prediction of the students' anxiety using ML [Bordoloi et al., 2022, Sau and Bhakta, 2019].

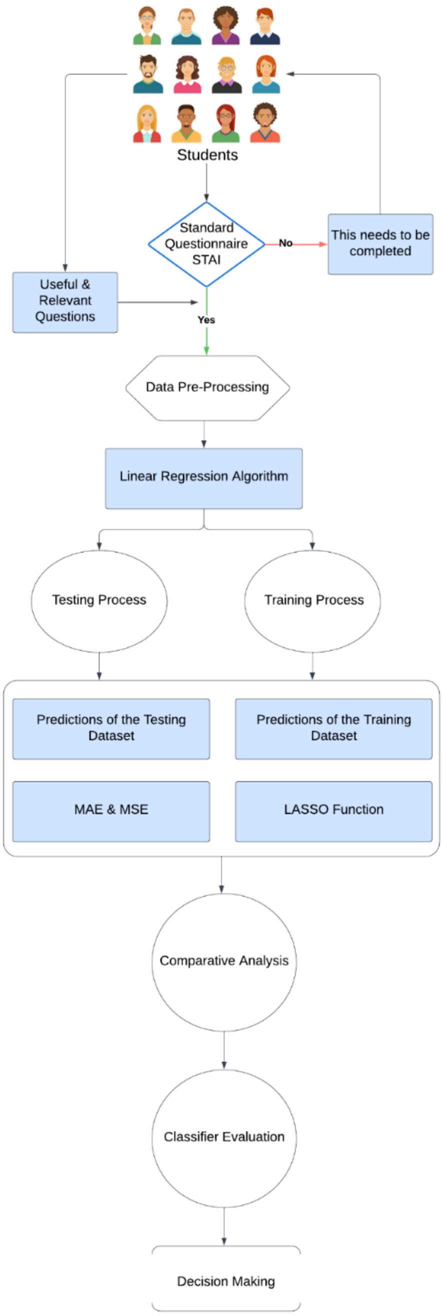

Figure 1: The research approach of the prediction of the students' anxiety using machine learning

At first, the research approach included different parts of the prediction of the students' anxiety using machine learning. The standard questionnaire of the STAI, and two additional questions were asked to assess students' use of modern technologies in this regard (Table 6).

1. Do you use mobile health apps like the anxiety management app on your smartphone?

2. Would you like your anxiety to be predicted using ML technology?

Moreover, data pre-processing was the second section of this approach, and the dataset then was prepared to be put into the linear regression ML algorithm. In this study, the linear regression approach is selected based on the results from the literature [Gruda and Hasan, 2019]. In addition, MAE and MSE were used to separate state and trait anxiety unobserved quantity [Yang, 2022]. After the training step, the LASSO function shows the features with a high impact on the model [Centofanti et al., 2022]. The training process had an impact on LASSO function, and the MAE and MSE factors were achieved from both the training process and testing process. Finally, after extracting the accuracy of the model from the test and train, comparative analysis, classifier evaluation, and decision-making are mentioned in the research approach.

3.2 Participants

The standard questionnaire was distributed via Google Forms among 78 students at the University of Belgrade. After extracting and solving the problems, it was distributed among various groups of students on social media and e-mail. Most participants were from the Faculty of Organisational Sciences, the Faculty of Medicine, and the Faculty of Dentistry. Data were collected from April to July 2022. Finally, 265 students participated in this study, 231 responses were correct and used for analysis. Most of the participants (89.2%) were between 18 and 27 years old (Table 2).

Characters		**Frequency**	**Percentage (%)**
Gender	Female	173	74.9%
	Male	56	24.2%
	Androgynous	2	0.9%
Age	18 – 22	169	73.2%
	23 – 27	37	16%
	28 – 32	11	4.8%
	> 32	14	6%
Marital Status	Single	214	92.6%
	Married	15	6.6%
	Divorced	1	0.4%
	Widowed	1	0.4%
Educational Level	Bachelor	196	84.8%
	MSc	22	9.5%
	MPhil	3	1.3%
	PhD	10	4.4%

Table 2: The demographic profile of respondents, and audiences' characters

In addition, 92.6% of participants were single, and finally, attributes like gender, age, marital status, and educational levels were collected.

3.3 Instruments

One of the most common methods for measuring anxiety is named the State-Trait Anxiety Inventory [Spielberger et al., 1983]. It consists of 20 items for state anxiety and 20 items for assessing trait anxiety. Scores are defined in the range of 20 to 80, where higher scores indicate high anxiety. All items are rated on a four-point Likert type scale of state anxiety (1- "not at all", 2- "somewhat", 3- "moderately so", and 4- "very much so"), and similarly for trait anxiety (1- "almost never", 2- "sometimes", 3- "often", 4- "almost always") [Spielberger et al., 1971, Spielberger, 1989, Stojanović et al., 2020]. The Spielberger STAI scores were classified as low (20 - 40), moderate (41 - 50), and high (51 - 80) [Spielberger et al., 1983, Tze Ping et al., 2008].

3.4 Data Analysis

According to data analysis and anxiety scoring, a final dataset (N=231) was prepared and scored using ML. The data were collected via an online questionnaire and imported in CSV – an Excel datasheet – to analyse them. The CSV was used because in this case the database can be displayed in the form of a data frame in Pandas Library. Hence, features of anxiety as extracted in the questionnaire were measured, and the system was able to predict reasons for individuals' anxiety as a predictor. The machine learning algorithms were coded and implemented using Python (JupyterLab 3.2.1), 70% of the data were used for the training and 30% (0.3) for the test. Features of this research are extracted from the anxiety questionnaire content to learn which factors had a negative impact on users' anxiety.

3.4.1 Data Pre-processing

After collecting data, a number of scales were reversed, and then scores between 20 and 80 were sorted [Spielberger et al., 1983]. To ensure the reliability of the questionnaire and the obtained results, a total of 231 samples were transferred from EXCEL to "IBM SPSS Statistics 26" software. Since the obtained Cronbach's Alpha should be ≥ 0.7 ($\alpha \geq 0.7$), the reliability of the questionnaire was established, hence the results were prepared to be extracted and converted into anxiety data [de Vet et al., 2017].

Reliability	
N of Items	**Cronbach's Alpha**
State – Anxiety (1 – 20)	0.89
Trait – Anxiety (21 – 40)	0.88
Total (1 – 40)	0.93

Table 3: Cronbach's Alpha of the anxiety dataset

In addition, the Pearson correlation between state and trait anxiety was 0.83.

Generally, data pre-processing is done for the accuracy and efficiency of testing and training in machine learning to increase reliability [Alharthi, 2020]. In this dataset, samples were placed on rows, and attributes were placed on columns. Data pre-processing including data cleaning, duplicates, instance selection, normalisation, and feature extraction had quite an impact on performance in this research. In the data cleaning section, some columns were removed from the machine learning process regarding lack of need. Duplicate data were also removed or modified to increase the accuracy of calculations. Moreover, data normalisation implies the removal of very large and very small data that had a significant impact on the final result (dividing each level by the norm equals 1). To standardise the data, each sample was subtracted from the mean value and divided by the standard deviation. Finally, noises were diagnosed with outliers, and noises were removed from the data set.

3.5 ML algorithms and functions

This section provides brief information about one of the machine learning algorithms which has shown successful results in this research. Furthermore, as has been noted in the literature review in previous studies, supervised learning algorithms were appropriate for this study. Both state anxiety and trait anxiety were examined (train & test) by linear regression and ridge regression so that the obtained scores were the same, and that is why linear regression was chosen to use and explain. Also, linear regression is a method employed for predictive analysis across datasets [Ali et al., 2020]. To solve problems in regression, linear regression is one of the common algorithms in supervised learning that can develop optimised solutions [Ilic et al., 2021].

Moreover, the reliability of the pattern in this study depends on the validity of linear regression [Luu et al., 2021]. This algorithm is known as a flexible and powerful prediction technique [Tunc and Genç, 2021]. Training data from students' anxiety scores were defined at 70% and 30% of the data were selected as the test section. The implementation of linear regression was coded by using the Scikit-learn library. Furthermore, both state anxiety and trait anxiety were prepared to fit and then predict.

After linear regression processes, it is necessary to calculate the difference between the predicted values and the actual values. The MSE and MAE were also calculated for this dataset to estimate an unobserved quantity as well.

These functions have been able to find the risk rate between the actual values and the estimated values. As much as the error model reaches zero, the model will be optimal.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{\text{n}} = \frac{\sum_{i=1}^{n} |e_i|}{\text{n}}$$

MAE as the unbiased estimator has been able to measure a wide diversity of disciplines of ML. Greek symbol (Σ) is used for summation, actual values for observation are calculated with "y_i", and "x_i" calculates the value for the observation, and the total number of observations is shown with "n". The arithmetical average of absolute errors $|e_i| = |y_i - x_i|$ where "x_i" is chosen as the true value, and "$\hat{Y}_i$" is the prediction [Karunasingha, 2022].

$$MSE = \frac{1}{n} \sum_{i=0}^{n} \left(Y_i - \hat{Y}_i\right)^2$$

Further, "n" denotes the number of data points so that "Y_i" observed values, and "$\hat{Y}_i$" can process predicted values [Yang, 2022]. Finally, LASSO function as an efficient estimator is proposed for this model. The factors with the highest impact on both state anxiety and trait anxiety were extracted after training using the LASSO function of the Scikit-learn library [Centofanti et al., 2022].

4 Results

Before visualisation or pre-processing, the reliability of each part of the anxiety questions of the questionnaire was determined by Cronbach's Alpha, and the value of state anxiety with 0.89, and the value of trait anxiety was achieved at 0.88 (Table 3). Then, the scores for each part of anxiety were calculated (Table 4). In state anxiety, 89 samples were at the lowest level of anxiety category, 90 samples were moderate, and only 52 students had high anxiety. In trait anxiety, 81 samples included a low level of anxiety, 100 were moderate, and 50 students had a high level of anxiety.

	Anxiety Grade			
Test	**Low**	**Moderate**	**High**	**Total Average**
State Anxiety	89	90	52	43.75 (From 20 To 80)
Trait Anxiety	81	100	50	43.74 (From 20 To 80)

Table 4: Participants in State - Trait Anxiety Inventory test results

Average students' anxiety was reported as moderate so the average state of anxiety showed 43.75 out of 80 points, and the average trait of anxiety showed 43.74 out of 80. In addition, some of the details that may have adverse effects on students' anxiety were collected (Table 5). Among the students, 130 individuals (56.3%) were in their first year of study at the university, 38 (16.5%) second-year students, 23 (10%) third-year students, 28 (12.1%) fourth-year students, and 12 (5.1%) of them were as the advanced university students. The GPA is also one of the main factors that may have quite the effect on students' anxiety. In addition to these results, 76 students (32.9%) had GPA from 6.00 to 7.99 out of 10.00, and 155 of them (67.1%) had GPA between 8.00 and 10.00.

Questions	**Options**	**Frequency**	**Percentage (%)**
1.Year of study	First	130	56.3
	Second	38	16.5
	Third	23	10
	Fourth	28	12.1
	Advanced university	12	5.1
2. Grade Point Average	6 – 6.99	7	3
	7 – 7.99	69	29.9
	8 – 8.99	99	42.9
	9 – 10.00	56	24.2
3. Has it ever happened that it took you once or more times to take the exam in order to pass?	Yes – once	71	30.7
	Yes – many times	36	15.6
	Never	124	53.7
4. Are you employed?	Yes	171	74
	No	60	26

5. Are you a beneficiary of a scholarship/ student loan?	Yes	91	39.4
	No	140	60.6
6. Who supports your scholarships or tuition fees for your study?	State	76	32.9
	Company	4	1.7
	Your parents	114	49.4
	Yourself	37	16

Table 5: More information about students

It was found that 124 of the students (53.7%) had passed their courses successfully the first time, 71 of them (30.7%) sometimes failed to pass their exams successfully the first time and repeated their exams once, and finally, 36 students (15.6%) passed their exams over several times. They were also asked about the status of their occupation, 171 (74%) of students were employed and 60 of them (26%) did not have any occupations. Meanwhile, 91 students (39.4%) were benefiting from a scholarship or a student loan, and 140 students (60.6%) never used any scholarships or student loans.

They were asked to answer who supported their tuition fees or scholarships during their study. Moreover, 37 students (16%) supported the fees themselves, 114 (49.4) of them were supported by their parents, and finally, 76 students (32.9%) used state fully funded scholarships for their studies. It was decided to get the opinion of the participants about their use of smart healthcare programmes, and their interest in predicting anxiety by machine learning technology (Table 6). The first question (Q1) was about the use of mobile health apps like anxiety management by the smartphone, and the second question (Q2) was about their interest in predicting the students' anxiety by machine learning technology (Table 6).

Questions	Options	Frequency	Percentage
Q1	A. I quite often use these technologies.	27	11.7
	B. Almost less than 4 times a month.	15	6.5
	C. I occasionally use this technology.	37	16
	D. I do not use these technologies.	152	65.8
Q2	A. Yes, I would like to.	75	32.5
	B. I might use it later.	102	44.2
	C. No, I wouldn't like to use it.	54	23.3

Table 6: Advantageous questions related to interest in using smart healthcare among university students

In the continuation of the results, students' physical activities, and their smoking cigarettes rate were asked (Table 7). Results showed that 179 participants (77.5%) were not willing to smoke, and 52 (22.5) of them were used to smoking cigarettes.

No.	Questions	Options	Frequency	Percentage
1	Smoking cigarettes	No Yes	179 52	77.5 22.5
2	Physical activities (the approximate 150	Yes No	140 91	60.6 39.4

Table 7: Smoking and physical activities among the students

Although useful information related to stress and anxiety such as smoking cigarettes, physical activity, and GPA collected in this research article, we have never used them for any classifications or in ML algorithms and functions. In addition to these contents, 140 of the students (60.6) had an adequate amount of physical activities (approximately 150 minutes per week) in their lives, but 91 students (39.4%) hadn't done enough physical activities during the week. Figure 2 shows the distribution of state anxiety and trait anxiety so that extracted from the Seaborn Kernel Distribution Estimation plot by using Python programming language. The anxiety level of the participants, listed from 20 to 80, and the mean distribution of both modes have been accurately displayed. The Kernel Distribution Estimation plot represents the distribution of the STAI questionnaire using a continuous probability density curve.

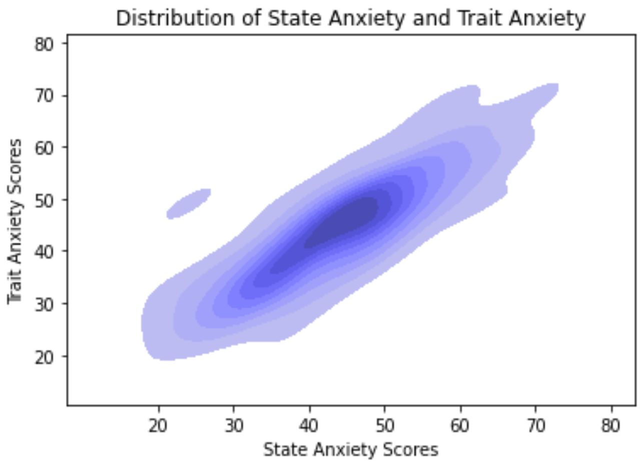

Figure 2: Distribution of samples in both state anxiety and trait anxiety

Moreover, Figure 3 shows a pair-plot of the dataset (the combination of two datasets as a matrix), and this plot creates a function so that each variable of this dataset is divided between the y-axis and the x-axis. Very few outliers have also been shown clearly in this plot, so it's obvious there hadn't been strong significance in the results and accuracy of the model.

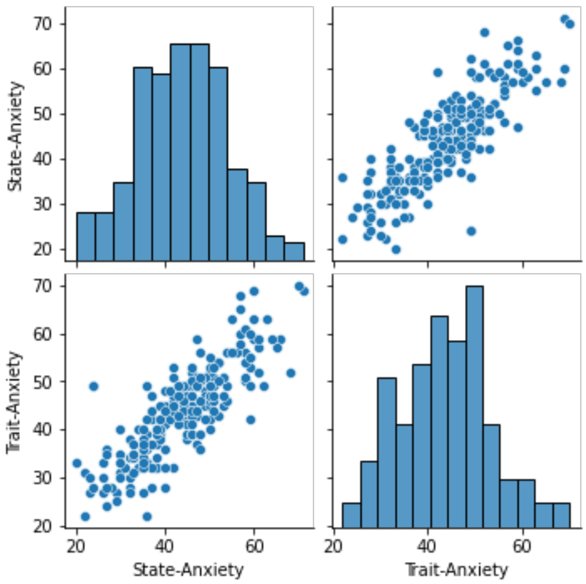

Figure 3: A pair-plot relationship in the dataset

Then, Figure 4 called the box-plot, shows the 231-number summary of the set of data, in the first quartile of the state anxiety, the minimum response showed 20 out of 80, and 71 out of 80 was the maximum range in the third quartile. For the trait anxiety that has been shown with orange colour in Figure 4, the minimum number was 22 out of 80, and 70 out of 80 was the third quartile and maximum.

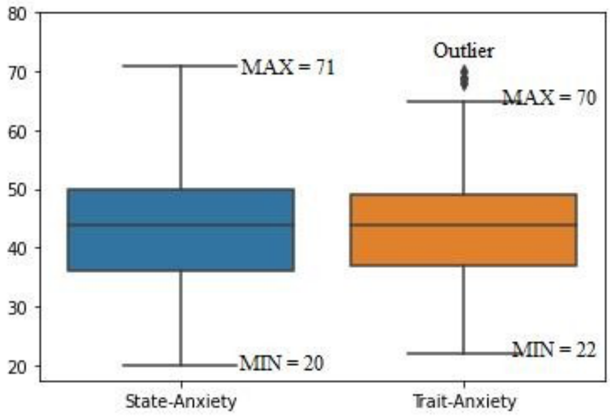

Figure 4: Distribution of differences in measured average across the dataset between state anxiety and trait anxiety

Although noises were removed from the dataset, the outlier was calculated. Figure 4 displays the outlier for the trait anxiety dataset, and this factor caused few impacts on the output. Eventually, after observation of 70 data samples, and as mentioned above, the standardisation operation was performed, and the data then were defined in the machine learning programme for testing and training. Machine learning results with 70 samples did not have an acceptable accuracy. Therefore, after the desired number of samples reached 231 students, the tests were performed again and a very high accuracy was obtained from the linear regression algorithm (Table 8). It should also be noted that, as argued above, all measures were examined separately for both state and trait anxiety.

Algorithm Type	Type of Anxiety	MAE	MSE
Linear Regression	State-Anxiety	7.86	7.80
Linear Regression	Trait-Anxiety	5.68	9.66

Table 8: Error/accuracy metrics of both state and trait anxiety in linear regression by using MAE and MSE

Mean squared error assessed the value of error in statistical models, and the value of MSE showed the system error close to zero and optimal (Table 8). Thus, Figure 5 and Figure 6 show the "Lasso Regression Lost Function" that found the factors with the highest impact in each state and trait anxiety after training.

Figure 5 showed the 12th question in state anxiety had the highest impact after training, and Figure 6 showed that the 13th question was the most important factor in trait anxiety after the training section. Finally, the k-nearest neighbour algorithm is used in this regard to find and show the way of training as well overfitting and underfitting problems.

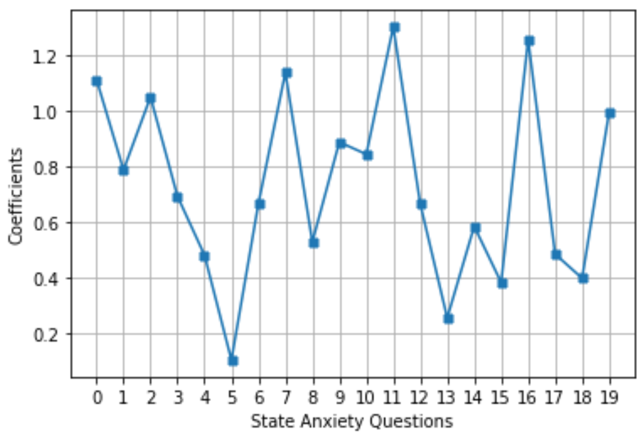

Figure 5: The option with the highest impact after training in state anxiety

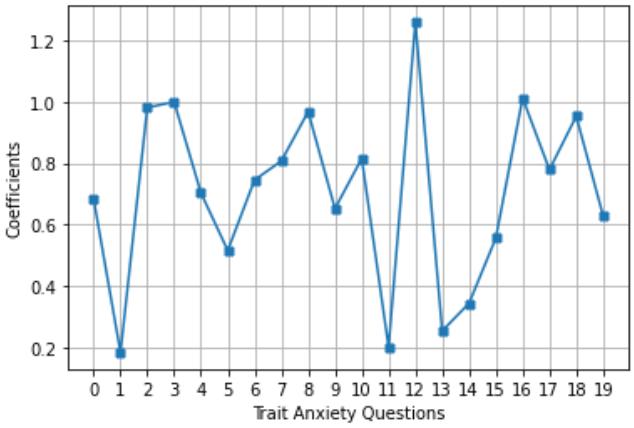

Figure 6: The option with the highest impact after training in the trait -anxiety

Figure 7 showed that approximately 140 samples that were randomly extracted from the dataset, were coded, and up to the 20th sample proximity, overfitting problems were seen. Figure 8 also showed the accuracy of the model starting from sample 85 onwards, and overfitting and underfitting problems were not seen in trait anxiety.

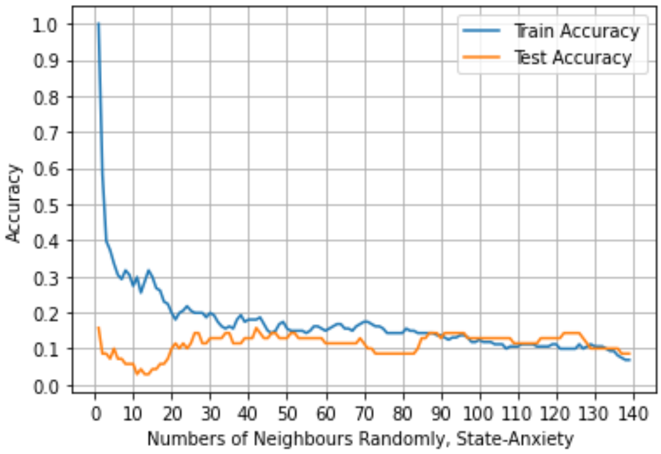

Figure 7: Checking the Overfitting and Underfitting the State Anxiety by K-Nearest Neighbour

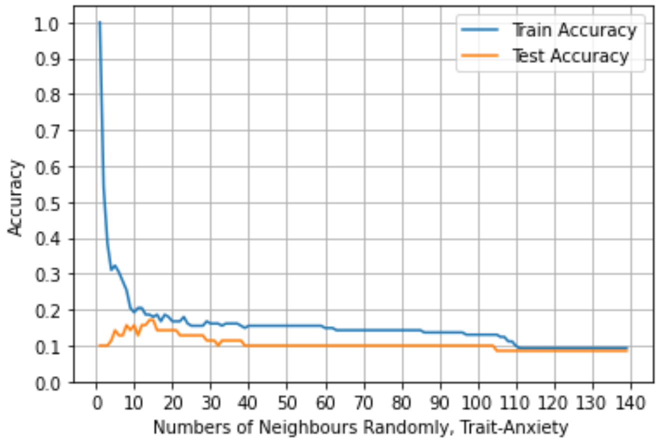

Figure 8: Checking the Overfitting and Underfitting the Trait Anxiety by K-Nearest Neighbour

5 Discussion

Nowadays, COVID-19 is known as one of the major reasons for anxiety [Metin et al., 2022]. As mentioned above, the education period is a sensitive time and students are vulnerable, so, this study focused on the prediction of students' stress and anxiety by using machine learning to develop smart healthcare services. It may be difficult for the youngest students to recognise anxiety, therefore using machine learning algorithms to detect and predict anxiety can be an effective solution [Gruda and Hasan, 2019]. However, smart healthcare is still in its infancy [Pramanik et al., 2017], and most of the students in this research (65.8%) have never used such technologies (Table 6).

Besides, only 23.3% of students have stated that they would not like to use ML technology to predict their anxiety. Some of the details in daily life that may cause anxiety among young students have been asked. For example, although using public bikes is useful, busy parking can cause anxiety [Tsai et al., 2019]. This research examined some cases that might affect students' anxiety such as students' year of study, grade point average, taking the exam more than once, employment status, and the situation with scholarships or student loans for tuition fees for the study. As argued before, most of the classes at the university were held in an online context, and as Table 7 shows, due to COVID-19, 39.4 % of students had not done enough physical activities per week [Abbas et al., 2019, McIntosh et al., 2017]. Moreover, the previous research argued that students who had anxiety were more likely to smoke [Lavallee et al., 2021].

Table 7 also shows that 52 students were used to smoking, this is one of the vital reasons for students' moderate anxiety. So, according to the material obtained and discussed, the anxiety rate (Table 4) seems a reasonable response. State anxiety points out the physiological temporary reactions directly related to harmful conditions at a specific point in time, on the other hand, trait anxiety argues for a trait of personality and the general condition of anxiety [Wiglusz et al., 2019]. In this study, both state and trait anxiety were examined separately [Spielberger et al., 1971].

The experiments showed unacceptable results after examining 70 samples, and when the data set reached 231 favourable cases, the error or accuracy metrics of the state anxiety achieved acceptable results, which shows the acceptable result compared to the conducted studies in this regard [Gruda and Hasan, 2019, Richter et al., 2020, Richter et al., 2021].

One of the studies has worked on users' state and trait anxiety by the STAI, and social engagement of the number of users on Twitter using linear regression. In addition, our MSE outputs (Table 8) have shown few more acceptable results than previous research [Gruda and Hasan, 2019]. Besides, in another study that focused on a clinical sample to predict 101 participants' anxiety based on STAI questionnaire, decision tree classification algorithm was used with 80.50% accuracy [Richter et al., 2021]. Also, another research article that collected 125 participants' anxiety data from the STAI and analysed them using the random forest algorithm reached 74.18% prediction accuracy [Richter et al., 2020]. From Figure 5 we can perceive users with these conditions (Table 4 and Table 7) may face nervousness in their life.

Figure 5 shows the factor with the highest impact after training in state anxiety, is called "I feel nervous". In addition, Figure 6 by using the LASSO function shows the option (I feel secure) with the highest impact after training in the trait anxiety. These results, which are mapped from ML algorithms and functions, show that because of the COVID-19 pandemic and several problems in the education period, some of the students have anxiety, but generally, they feel secure.

Furthermore, the current study introduces a pattern of prediction of students' anxiety by using machine learning technology so this pattern can develop smart healthcare systems. The performance of the linear regression algorithm is assessed using the KNN algorithm in checking the Overfitting and Underfitting for both the state and trait anxiety. In Figure 7, approximately between the 90th and 125th factors, it is clear that the training accuracy of our system has increased and is stable. Although Figure 8 hasn't touched the point of learning, this Figure shows a range of progress and is near the point of learning, specifically and approximately from the 110th factor onward. In addition to these features, Figure 8 shows monitoring the accuracy of training without the overfitting and underfitting the trait anxiety by K-NN so that up to the 20th factor in state anxiety, few overfitting problems were seen.

Finally, the implementation of this research allows researchers to learn about serious issues that affect anxiety and also know how ML can predict and analyse the participants' anxiety with few errors.

6 Conclusion

The current research explains how the findings of this study can expand experts' knowledge in predicting students' anxiety using machine learning technology. According to the challenges of predicting anxiety among students by using ML in Table 1, some of them were chosen as the gap in this study and have been solved. In addition, this research aimed to extract deep information and details about students, smoking, and physical activity rates by checking their final anxiety scores. Then, using linear regression, proved such data can be predicted with the highest accuracy metrics by this supervised learning algorithm. Data analysis reveals both state and trait anxiety of 231 university students with error metrics. Data analysis also showed the MAE and MSE of these factors the relationship between them, and the LASSO function, which found the factors with the highest impact after training. Moreover, overfitting and underfitting using K-NN show the lowest problems in state anxiety, especially in trait anxiety. In this research, some concerns have also been adequately addressed in the context of smart healthcare in this regard. Finally, the presented research approach can have several practical concepts in academia and industrial areas.

There were few limitations in this study. The examination period was one of this study's limitations because they did not have enough time and focus to fill out the online questionnaire. In addition, the most important limitation of this research was the students' English language knowledge. Some of the bachelor's students hadn't quite the knowledge of English language skills.

Future research can examine the amount of physical activity, smoking cigarettes, and GPA in the main model and ML algorithms and functions as well as classification algorithms (In two sections, first without examining such features with the main standard questionnaire of STAI and second, analysing the stress and anxiety dataset with these three features and more factors in this field). Future studies can also use previous anxiety records, and some issues like insomnia, and analyse them by using deep learning methods. Researchers can focus on air pollution and its effects on students' anxiety using DL as well.

Funding

The research was funded by institutional funding of University of Belgrade – Faculty of Organizational Sciences.

Conflict of interest

The authors do not report any conflicts of interest.

Ethical approval

All the students participated voluntarily. Data was fully anonymised. The ethical clearance for this study was obtained from the Department of e-business, University of Belgrade Faculty of Organizational Sciences, in accordance with the local rules valid at the time when the study was conducted.

Corresponding Author: FL

Contributions

FL: Data curation, Formal Analysis, Investigation, Software, Visualisation, Writing – Original Draft.
BR: Methodology, Validation, Supervision, Writing – Review & Editing.
AL: Conceptualisation, Methodology, Validation, Resources, Supervision, Writing – Review & Editing.
ZB: Conceptualisation, Methodology, Resources, Supervision, Funding Acquisition, Project Administration, Resources, Writing – Review & Editing.

Acknowledgments

Finally, we would like to thank our colleagues and friends for their technical assistance, continuous encouragement, and support throughout the research process.

References

[Abbas et al., 2019] Abbas, J., Aqeel, M., Abbas, J., Shaher, B., A. , J., Sundas, J. and Zhang, W. (2019). The moderating role of social support for marital adjustment, depression, anxiety, and stress: Evidence from Pakistani working and nonworking women. Journal of Affective Disorders, 244, 231–238. https://doi.org/10.1016/j.jad.2018.07.071.

[Ahmed et al., 2020] Ahmed, A., Sultana, R., Ullas, M. T. R., Begom, M., Rahi, M. M. I. and Alam, M. A. (2020). A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning. In 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CSDE50874.2020.9411642.

[Alharthi, 2020] Alharthi, H. (2020). Predicting the level of generalized anxiety disorder of the coronavirus pandemic among college age students using artificial intelligence technology. In Proceedings - 2020 19th Distributed Computing and Applications for Business Engineering and Science, DCABES 2020 (pp. 218–221). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/DCABES50732.2020.00064.

[Ali et al., 2020] Ali, M., Prasad, R., Xiang, Y. and Deo, R. C. (2020). Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms. Renewable and Sustainable Energy Reviews, 132, 110003. https://doi.org/10.1016/J.RSER.2020.110003.

[Aslam, 2018] Aslam, S. (2018). Twitter by the numbers: Stats, demographics & fun facts. Omnicoreagency.com.

[Balogh et al., 2015] Balogh, E. P., Miller, B. T. and Ball, J. R. (2015). Improving diagnosis in health care. Washington (DC): National Academies Press (US). https://pubmed.ncbi.nlm.nih.gov/26803862/.

[Barton et al., 2014] Barton, S., Karner, C., Salih, F., Baldwin, D. S. and Edwards, S. J. (2014). Clinical effectiveness of interventions for treatment- resistant anxiety in older people; a systematic review. Health Technology Assessment, 18(50), 1–60.

[Bhatnagar et al., 2023] Bhatnagar, S., Agarwal, J., & Sharma, O. R. (2023). Detection and classification of anxiety in university students through the application of machine learning. Procedia Computer Science, 218, 1542-1550. https://doi.org/10.1016/j.procs.2023.01.132.

[Bokma et al., 2020] Bokma, W. A., Bokma, W. A., Zhutovsky, P., Giltay, E. J., Schoevers, R. A., Penninx, B. W. J. H., Penninx, B. W. J. H., Van Balkom, A. L. J. M., Van Balkom, A. L. J. M., Batelaan, N. M., Batelaan, N. M. and Van Wingen, G. A. (2020). Predicting the naturalistic course in anxiety disorders using clinical and biological markers: A machine learning approach. Psychological Medicine. https://doi.org/10.1017/S0033291720001658.

[Bordoloi et al., 2022] Bordoloi, D., Singh, V., Sanober, S., Buhari, S. M., Ujjan, J. A. and Boddu, R. (2022). Deep Learning in Healthcare System for Quality of Service. https://doi.org/10.1155/2022/8169203.

[Caragliu et al., 2009] Caragliu, A., Del Bo, C. and Nijkamp, P. (2009). Smart cities in Europe. In In 3rd Central European Conference in Regional Science. https://air.unimi.it/handle/2434/72594.

[Centofanti et al., 2022] Centofanti, F., Fontana, M., Lepore, A. and Vantini, S. (2022). Smooth LASSO estimator for the Function-on-Function linear regression model. Computational Statistics and Data Analysis, 176. https://doi.org/10.1016/j.csda.2022.107556.

[Das et al., 2024] Das, A. K., Biswas, S. K., Mandal, A., Bhattacharya, A., & Sanyal, S. (2024). Machine Learning based Intelligent System for Breast Cancer Prediction (MLISBCP). Expert Systems with Applications, 242, 122673. https://doi.org/10.1016/j.eswa.2023.122673.

[Daza et al., 2023] Daza, A., Saboya, N., Necochea-Chamorro, J. I., Ramos, K. Z., & Valencia, Y. D. R. V. (2023). Systematic review of machine learning techniques to predict anxiety and stress in college students. Informatics in medicine unlocked, 101391. https://doi.org/10.1016/j.imu.2023.101391.

[de Vet et al., 2017] de Vet, H. C. W., Mokkink, L. B., Mosmuller, D. G. and Terwee, C. B. (2017). Spearman–Brown prophecy formula and Cronbach's alpha: different faces of reliability and opportunities for new applications. Journal of Clinical Epidemiology, 85, 45–49. https://doi.org/10.1016/J.JCLINEPI.2017.01.013.

[Delpino et al., 2022] Delpino, F. M., da Silva, C. N., Jerônimo, J. S., Mulling, E. S., da Cunha, L. L., Weymar, M. K., Alt, R., Caputo, E. L. and Feter, N. (2022). Prevalence of anxiety during the COVID-19 pandemic: A systematic review and meta-analysis of over 2 million people. Journal of Affective Disorders, 318, 272–282. https://doi.org/10.1016/J.JAD.2022.09.003.

[Demiris et al., 2020] Demiris, G., Corey Magan, K. L., Parker Oliver, D., Washington, K. T., Chadwick, C., Voigt, J. D., Brotherton, S. and Naylor, M. D. (2020). Spoken words as biomarkers: Using machine learning to gain insight into communication as a predictor of anxiety. Journal of the American Medical Informatics Association, 27(6), 929–933. https://doi.org/10.1093/jamia/ocaa049.

[Du et al., 2020] Du, S., Liu, W., Cai, S., Hu, Y. and Dong, J. (2020). The efficacy of e-health in the self-management of chronic low back pain: A meta analysis. International Journal of Nursing Studies, 106, 103507. https://doi.org/10.1016/j.ijnurstu.2019.103507.

[Duncan et al., 2022] Duncan, M. J., Riazi, N. A., Faulkner, G., Gilchrist, J. D., Leatherdale, S. T. and Patte, K. A. (2022). The association of physical activity, sleep, and screen time with mental health in Canadian adolescents during the COVID-19 pandemic: A longitudinal isotemporal substitution analysis. Mental Health and Physical Activity, 23, 100473. https://doi.org/10.1016/J.MHPA.2022.100473.

[Eeden et al., 2021] Eeden, W. A. Van, Luo, C., Hemert, A. M. Van, Carlier, I. V. E., Penninx, B. W., Wardenaar, K. J., Hoos, H. and Giltay, E. J. (2021). Predicting the 9-year course of mood and anxiety disorders with automated machine learning : A comparison between auto-sklearn , naïve Bayes classifier , and traditional logistic regression. Psychiatry Research, 299(February), 113823. https://doi.org/10.1016/j.psychres.2021.113823.

[Fan and Zhao, 2022] Fan, K. and Zhao, Y. (2022). Mobile health technology: a novel tool in chronic disease management. Intelligent Medicine, 2(1), 41–47. https://doi.org/10.1016/j.imed.2021.06.003.

[Garcia- Chimeno et al., 2017] Garcia-Chimeno, Y., Garcia-Zapirain, B., Gomez-Beldarrain, M., Fernandez-Ruanova, B. and Garcia-Monco, J. C. (2017). Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. BMC Medical Informatics and Decision Making, 17(1). https://doi.org/10.1186/s12911-017-0434-4.

[Gruda and Hasan, 2019] Gruda, D. and Hasan, S. (2019). Feeling anxious? Perceiving anxiety in tweets using machine learning. Computers in Human Behavior, 98, 245–255. https://doi.org/10.1016/j.chb.2019.04.020.

[Hermens et al., 2014] Hermens, H., op den Akker, H., Tabak, M., Wijsman, J. and Vollenbroek, M. (2014, 1st December). Personalized Coaching Systems to support healthy behavior in people with chronic conditions. Journal of Electromyography and Kinesiology, 24(6), 815–826. Elsevier Ltd. https://doi.org/10.1016/j.jelekin.2014.10.003.

[Huang and Wan, 2013] Huang, H. and Wan, K. (2013). Heart rate variability in junior high school students with depression and anxiety in Taiwan. Acta neuropsychiatrica, 25(3), 175–178.

[Hwang and Nam, 2022] Hwang, S. and Nam, J. Y. (2022). Do lifestyle risk behaviors affect anxiety, depressive symptoms, and suicidal ideation among adolescents? Psychiatry Research, 318, 114942. https://doi.org/10.1016/J.PSYCHRES.2022.114942.

[Ihmig et al., 2020] Ihmig, F. R., Antonio Gogeascoechea, H., Neurohr-Parakenings, F., Schäfer, S. K., Lass-Hennemann, J. and Michael, T. (2020). On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider- fearful individuals. PLoS ONE, 15(6), 1–20. https://doi.org/10.1371/journal.pone.0231517.

[Ilic et al., 2021] Ilic, I., Görgülü, B., Cevik, M. and Baydoğan, M. G. (2021). Explainable boosted linear regression for time series forecasting. Pattern Recognition, 120, 108144. https://doi.org/10.1016/J.PATCOG.2021.108144.

[Jacobson et al., 2021] Jacobson, N. C., Lekkas, D., Huang, R. and Thomas, N. (2021). Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17–18 years. Journal of Affective Disorders, 282, 104–111. https://doi.org/10.1016/j.jad.2020.12.086.

[Jacobson et al., 2015] Jacobson, N. C., Lekkas, D., Huang, R., Thomas, N., McIntosh, J. R. D., Jay, S., Hadden, N., Whittaker, P. J., Zhang, J., Xiong, H., Huang, Y., Wu, H., Leach, K. and Barnes, L. E. (2015). M-SEQ: Early detection of anxiety and depression via temporal orders of diagnoses in electronic health data. Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015, 148(October), 140–148. https://doi.org/10.1016/j.puhe.2017.04.001.

[Karunasingha, 2022] Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. Information Sciences, 585, 609–629. https://doi.org/10.1016/J.INS.2021.11.036.

[Khanjari et al., 2021] Khanjari, S., Bell, E. F., Sadeghi, L. A., Sabzehei, M. kazem and Haghani, S. (2021). The impact of a mobile health intervention on the sense of coherence and quality of life of mothers with premature infants. Journal of Neonatal Nursing, 27(6), 444– 450. https://doi.org/10.1016/j.jnn.2021.06.008.

[Labus et al., 2021] Labus, A., Radenković, B., Rodić, B., Barać, D. and Malešević, A. (2021). Enhancing smart healthcare in dentistry: an approach to managing patients' stress. Informatics for Health and Social Care, 46(3), 306–319. https://doi.org/10.1080/17538157.2021.1893322.

[Lavallee et al., 2021] Lavallee, K. L., Zhang, X. C., Schneider, S. and Margraf, J. (2021). A longitudinal examination of the relationship between smoking and panic, anxiety, and depression in Chinese and German students. Addictive Behaviors Reports, 14, 100347. https://doi.org/10.1016/j.abrep.2021.100347.

[Leung and Chen, 2019] Leung, L. and Chen, C. (2019). E-health/m-health adoption and lifestyle improvements: Exploring the roles of technology readiness, the expectation-confirmation model, and health-related information activities. Telecommunications Policy, 43(6), 563–575. https://doi.org/10.1016/j.telpol.2019.01.005.

[Lotfi et al., 2022] Lotfi, F., Rodić, B. and Bogdanović, Z. (2022). A System For Monitoring And Managing The Anxiety Among The Young People Using Machine Learning. In E-business technologies conference proceedings (vol. 2, pp. 91–94).

[Luu et al., 2021] Luu, Q. H., Lau, M. F., Ng, S. P. H. and Chen, T. Y. (2021). Testing multiple linear regression systems with metamorphic testing. Journal of Systems and Software, 182, 111062. https://doi.org/10.1016/J.JSS.2021.111062.

[Malik and Khan, 2023] Malik, S. S., & Khan, A. (2023, April). Anxiety, Depression and Stress prediction among College Students using Machine Learning Algorithms. In 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT) (pp. 1-5). IEEE. https://doi.org/10.1109/ICEEICT56924.2023.10157693.

[Masino et al., 2019] Masino, A. J., Forsyth, D., Nuske, H., Herrington, J., Pennington, J., Kushleyeva, Y. and Bonafide, C. P. (2019). M-Health and autism: Recognizing stress and anxiety with machine learning and wearables data. In Proceedings - IEEE Symposium on Computer-Based Medical Systems (vols 2019-June, pp. 714–719). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CBMS.2019.00144.

[McGinnis et al., 2019a] McGinnis, E. W., Anderau, S. P., Hruschak, J., Gurchiek, R. D., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L., Muzik, M. and McGinnis, R. S. (2019a). Giving Voice to Vulnerable Children: Machine Learning Analysis of Speech Detects Anxiety and Depression in Early Childhood. IEEE Journal of Biomedical and Health Informatics, 23(6), 2294–2301. https://doi.org/10.1109/JBHI.2019.2913590.

[McGinnis et al., 2018] McGinnis, R. S., McGinnis, E. W., Hruschak, J., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L. and Muzik, M. (2018). Wearable sensors and machine learning diagnose anxiety and depression in young children. 2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018, 2018-Janua(March), 410–413. https://doi.org/10.1109/BHI.2018.8333455.

[McGinnis et al., 2019b] McGinnis, R. S., McGinnis, E. W., Hruschak, J., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L. and Muzik, M. (2019b). Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning. PLoS ONE, 14(1), 1–16. https://doi.org/10.1371/journal.pone.0210267.

[McIntosh et al., 2017]] McIntosh, J. R. D., Jay, S., Hadden, N. and Whittaker, P. J. (2017). Do E-health interventions improve physical activity in young people: a systematic review. Public Health, 148, 140–148. https://doi.org/10.1016/j.puhe.2017.04.001.

[Metin et al., 2022] Metin, A., Erbiçer, E. S., Şen, S. and Çetinkaya, A. (2022). Gender and COVID-19 related fear and anxiety: A meta-analysis. Journal of Affective Disorders, 310, 384–395. https://doi.org/10.1016/J.JAD.2022.05.036.

[Nemesure et al., 2021] Nemesure, M. D., Heinz, M. V., Huang, R. and Jacobson, N. C. (2021). Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-81368-4.

[Omura et al., 2020] Omura, J. D., Brown, D. R., McGuire, L. C., Taylor, C. A., Fulton, J. E. and Carlson, S. A. (2020). Cross-sectional association between physical activity level and subjective cognitive decline among US adults aged ≥45 years, 2015. Preventive Medicine, 141, 106279. https://doi.org/10.1016/J.YPMED.2020.106279.

[Pai et al., 2021] Pai, R. R., Alathur, S., Khanjari, S., Bell, E. F., Sadeghi, L. A., Sabzehei, M. kazem, Haghani, S., Fan, K. and Zhao, Y. (2021). Bibliometric Analysis and Methodological Review of Mobile Health Services and Applications in India. Journal of Neonatal Nursing, 2(6), 41–47. https://doi.org/10.1016/j.ijmedinf.2020.104330.

[Pramana et al., 2014] Pramana, G., Parmanto, B., Kendall, P. C. and Silk, J. S. (2014). The SmartCAT: An m-health platform for ecological momentary intervention in child anxiety treatment. Telemedicine and e-Health, 20(5), 419–427. https://doi.org/10.1089/tmj.2013.0214.

[Pramanik et al., 2017] Pramanik, M. I., Lau, R. Y. K., Demirkan, H. and Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. Expert Systems with Applications, 87, 370–383. https://doi.org/10.1016/j.eswa.2017.06.027.

[Priya et al., 2020] Priya, A., Garg, S. and Tigga, N. P. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. In Procedia Computer Science (vol. 167, pp. 1258–1267). Elsevier B.V. https://doi.org/10.1016/j.procs.2020.03.442.

[Pullyblank et al., 2022] Pullyblank, K., Brunner, W., Scribani, M., Krupa, N., Ory, M. G. and Smith, M. L. (2022). Recruitment and engagement in disease self-management programs: Special concerns for rural residents reporting depression and/or anxiety. Preventive Medicine Reports, 26, 101761. https://doi.org/10.1016/J.PMEDR.2022.101761.

[Richter et al., 2021] Richter, T., Fishbain, B., Fruchter, E., Richter-Levin, G. and Okon-Singer, H. (2021). Machine learning-based diagnosis support system for differentiating between clinical anxiety and depression disorders. Journal of Psychiatric Research, 141, 199– 205. https://doi.org/10.1016/j.jpsychires.2021.06.044.

[Richter et al., 2020] Richter, T., Fishbain, B., Markus, A., Richter-Levin, G. and Okon-Singer, H. (2020). Using machine learning-based analysis for behavioral differentiation between anxiety and depression. Scientific Reports, 10(1), 1–12. https://doi.org/10.1038/s41598-020-72289-9.

[Sau and Bhakta, 2017] Sau, A. and Bhakta, I. (2017). Predicting anxiety and depression in elderly patients using machine learning technology. Healthcare Technology Letters, 4(6), 238–243. https://doi.org/10.1049/htl.2016.0096

[Sau and Bhakta, 2019] Sau, A. and Bhakta, I. (2019). Screening of anxiety and depression among the seafarers using machine learning technology. Informatics in Medicine Unlocked, 16. https://doi.org/10.1016/j.imu.2018.12.004.

[Sharma and Verbeke, 2021] Sharma, A. and Verbeke, W. J. M. I. (2021). Understanding importance of clinical biomarkers for diagnosis of anxiety disorders using machine learning models. PLoS ONE, 16(5 May 2021), 1–16. https://doi.org/10.1371/journal.pone.0251365.

[Shin et al., 2013] Shin, L. M., Davis, F., VanElzakker, M. B., Dahlgren, M. K. and Dubois, S. J. (2013). Neuroimaging predictors of treatment response in anxiety disorders. Biology of Mood & Anxiety Disorders, 3(1), 15. https://doi.org/10.1186/2045-5380-3-15.

[Souza and Soares, 2023] Souza, L. S. and Soares, M. S. (2023). Combining SysML and Timed Coloured Petri Nets for Designing Smart City Applications. JUCS - Journal of Universal Computer Science, 29(10), 1217–1250. https://doi.org/10.3897/jucs.97170.

[Soysal et al., 2022] Soysal, D., Bani-Yaghoub, M. and Riggers-Piehl, T. A. (2022). A Machine Learning Approach to Evaluate Variables of Math Anxiety in STEM Students. Pedagogical Research, 7(2), em0125. https://doi.org/10.29333/pr/11978.

[Spielberger, 1989] Spielberger, C. D. (1989). State-Trait Anxiety Inventory: Bibliography . Palo Alto, CA: Consulting PsychologistsPress.

[Spielberger et al., 1971] Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. S. and Natalicio, D. S. (1971). The state-trait anxiety inventory. Revista Interamericana de Psicologia/Interamerican Journal of Psychology, 5(3 & 4).

[Spielberger et al., 1983] Spielberger, C. D., Gorsuch, L. R., Lushene, R., Vagg, P. R. and Jacobs, G. A. (1983). Consulting Psychologists Press; Palo Alto, CA: 1983.

[Stojanović et al., 2020] Stojanović, N. M., Randjelović, P. J., Nikolić, G., Stojiljković, N., Ilić, S., Stoiljković, B. and Radulović, N. S. (2020). Reliability and validity of the spielberger's state-trait anxiety inventory (STAI) in serbian university student and psychiatric non- psychotic outpatient populations. Acta Facultatis Medicae Naissensis, 327(2), 149–159. https://doi.org/10.5937/afmnai37-25011.

[Sujith et al., 2021] Sujith, A. V. L. N., Sajja, G. S., Mahalakshmi, V., Nuhmani, S. and Prasanalakshmi, B. (2022). Systematic review of smart health monitoring using deep learning and Artificial intelligence. Neuroscience Informatics, 2(3), 100028. https://doi.org/10.1016/j.neuri.2021.100028.

[Tsai et al., 2019] Tsai, M. F., Chen, P. and Hong, Y. J. (2019). Enhancing the utilization of public bike sharing systems using return anxiety information. Future Generation Computer Systems, 92, 961–971. https://doi.org/10.1016/J.FUTURE.2017.12.063.

[Tunc and Genç, 2021] Tunc, H. and Genç, B. (2021). A column generation based heuristic algorithm for piecewise linear regression. Expert Systems with Applications, 171, 114539. https://doi.org/10.1016/J.ESWA.2020.114539.

[Tze Ping et al., 2008] Tze Ping, L., Subramaniam, K. and Krishnaswamy, S. (2008). Test anxiety: state, trait and relationship with exam satisfaction, 15(2), 18–23.

[Valle et al., 2021] Valle, N., Antonenko, P., Valle, D., Dawson, K., Huggins-Manley, A. C. and Baiser, B. (2021). The influence of task-value scaffolding in a predictive learning analytics dashboard on learners' statistics anxiety, motivation, and performance. Computers and Education, 173. https://doi.org/10.1016/j.compedu.2021.104288.

[van Vuuren et al., 2021] van Vuuren, C. L., van Mens, K., de Beurs, D., Lokkerbol, J., van der Wal, M. F., Cuijpers, P. and Chinapaw, M. J. M. (2021). Comparing machine learning to a rule-based approach for predicting suicidal behavior among adolescents: Results from a longitudinal population-based survey. Journal of Affective Disorders, 295, 1415–1420. https://doi.org/10.1016/j.jad.2021.09.018.

[Wiglusz et al., 2019] Wiglusz, M. S., Landowski, J. and Cubała, W. J. (2019). Psychometric properties and diagnostic utility of the State–Trait Anxiety Inventory in epilepsy with and without comorbid anxiety disorder. Epilepsy & Behavior, 92, 221–225. https://doi.org/10.1016/J.YEBEH.2019.01.005.

[Yang, 2022] Yang, D. (2022). Correlogram, predictability error growth, and bounds of mean square error of solar irradiance forecasts. Renewable and Sustainable Energy Reviews, 167(April), 112736. https://doi.org/10.1016/j.rser.2022.112736.

[Youyou et al., 2015] Youyou, W., Kosinski, M. and Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. Proceedings of the National Academy of Sciences of the United States of America, 112(4), 1036–1040. https://doi.org/10.1073/pnas.1418680112.