PLOS ONE

Predicting depression severity using machine learning models: Insights from mitochondrial peptides and clinical factors

Toheeb Salahudeen[1], **Maher Maalouf**[1]*,
Ibrahim (Abe) M. Elfadel[2,3], **Herbert F. Jelinek**[4,5]

1 Department of Management Science and Engineering, Khalifa University, Abu Dhabi, UAE,
2 Department of Computer and Communication Engineering, Khalifa University, Abu Dhabi, UAE,
3 Center for Secure Cyber Physical Systems, Khalifa University, Abu Dhabi, UAE, 4 Department of
Medical Sciences and Health Engineering Innovation Group, Khalifa University, Abu Dhabi, UAE,
5 Biotechnology Center, Khalifa University, Abu Dhabi, UAE

* maher.maalouf@ku.ac.ae

Abstract

Depression presents a significant challenge to global mental health, often intertwined with factors including oxidative stress. Although the precise relationship with mitochondrial pathways remains elusive, recent advances in machine learning present an avenue for further investigation. This study employed advanced machine learning techniques to classify major depressive disorders based on clinical indicators and mitochondrial oxidative stress markers. Six machine learning algorithms, including Random Forest, were applied and their performance was investigated in balanced and unbalanced data sets with respect to binary and multiclass classification scenarios. Results indicate promising accuracy and precision, particularly with Random Forest on balanced data. RF achieved an average accuracy of 92.7% and an F1 score of 83.95% for binary classification, 90.36% and 90.1%, respectively, for the classification of three classes of severity of depression and 89.76% and 88.26%, respectively, for the classification of five classes. Including only oxidative stress markers resulted in accuracy and an F1 score of 79.52% and 80.56%, respectively. Notably, including mitochondrial peptides alongside clinical factors significantly enhances predictive capability, shedding light on the interplay between depression severity and mitochondrial oxidative stress pathways. These findings underscore the potential for machine learning models to aid clinical assessment, particularly in individuals with comorbid conditions such as hypertension, diabetes mellitus, and cardiovascular disease.

1 Introduction

Recent research indicates an increase in neuropsychiatric disorders [1]. This makes it crucial to better understand these conditions and their complex interactions with chronic diseases, as well as associated biochemical markers for improved personalized treatment and outcomes

[2]. Depression is the most common chronic stress-related disorder, affecting 20% of people at some point in their lives [3]. According to the World Health Organization (WHO), this disorder is responsible for approximately 800,000 deaths annually and is one of the leading causes of suicide [4].

Several studies have indicated a possible role of mitochondrial oxidative stress in the development of depression severity—the intensity and number of depressive symptoms and functional impairments [5]—and other mental health problems. The severity of depression is commonly assessed using standardized tools such as the Hamilton Depression Rating Scale (HAM-D) or the Patient Health Questionnaire (PHQ-9) [5]. Understanding and measuring the severity of depression is crucial to determine the appropriate treatment approach, [6] and assessing the severity of depression [7]. Incorporating mitochondrial peptides in the assessment of the severity of depression holds significant promise [8]. Biomarkers can provide objective measures of depression, assist in identifying disease in a timely manner, predict treatment responses, and identify subtypes of depression [3]. Furthermore, mitochondrial biomarkers may provide information on the underlying biological mechanisms of depression, which can lead to more personalized and effective treatment strategies [9]. Integrating biomarkers in the evaluation of the severity of depression increases the number of factors taken into account for diagnosis and avoids possible misdiagnosis or under-/over-prescribing [10].

However, information on the role of specific mitochondrial-associated peptides is still lacking [11–19]. This study aims to fill this critical gap by: (1) examining the impact of Mitochondria-derived peptides (MDPs), such as Humanin, MOTS-c, and p66Shc, together with ACE genotypes in the context of cardiovascular disease (CVD), hypertension (HT), and diabetes mellitus (DM); (2) clarifying the intricate interplay between these mitochondrial factors and the progression of depressive disorders, and advance understanding of the role of mitochondrial-derived oxidative stress peptides in the progression of depressive disorders.

1.1 Mitochondrial oxidative stress peptides

Mitochondria-derived peptides (MDPs), include humanin (HN) and mitochondrial ORF (Open Reading Frame) of 12S rRNA Type C (MOTS-c), are encoded by mitochondrial DNA, and play crucial roles in maintaining mitochondrial function [20,21]. MDPs have been associated with neuroprotection and cardioprotection, affecting the development of cardiovascular disease [13]. Humanin (HN) has also been shown to be involved in major depressive disorder [22]. HN affects mitochondrial oxidative phosphorylation, adenosine triphosphate (ATP) production, and inhibition of apoptosis, by influencing glucose metabolism [23]. Fig 1 describes and displays the HN pathways. HN binds to two types of receptors and activates signaling pathways: protein kinase B (AKT), extracellular signal-regulated kinase 1/2 (ERK1/2) and signal transducer and transcription activator 3 (STAT3)—a transcription factor that plays a critical role in mediating cellular responses to cytokines and growth factors. It is involved in various biological processes, including cell growth, differentiation, and inflammation, which in turn trigger several downstream signaling pathways, such as Janus kinase 2 activation (JAK2), a non-receptor tyrosine kinase involved in signaling by cytokines, which subsequently activates STAT3 [24,25].

MOTS-c also has various physiological functions (as shown in Fig 2 such as reducing insulin resistance, preventing obesity, and regulating bone remodeling, primarily through the Folate-AICAR-AMPK pathway. The Folate-AICAR-AMPK pathway is a series of interconnected chemical reactions in the body that involve the molecules folate, AICAR
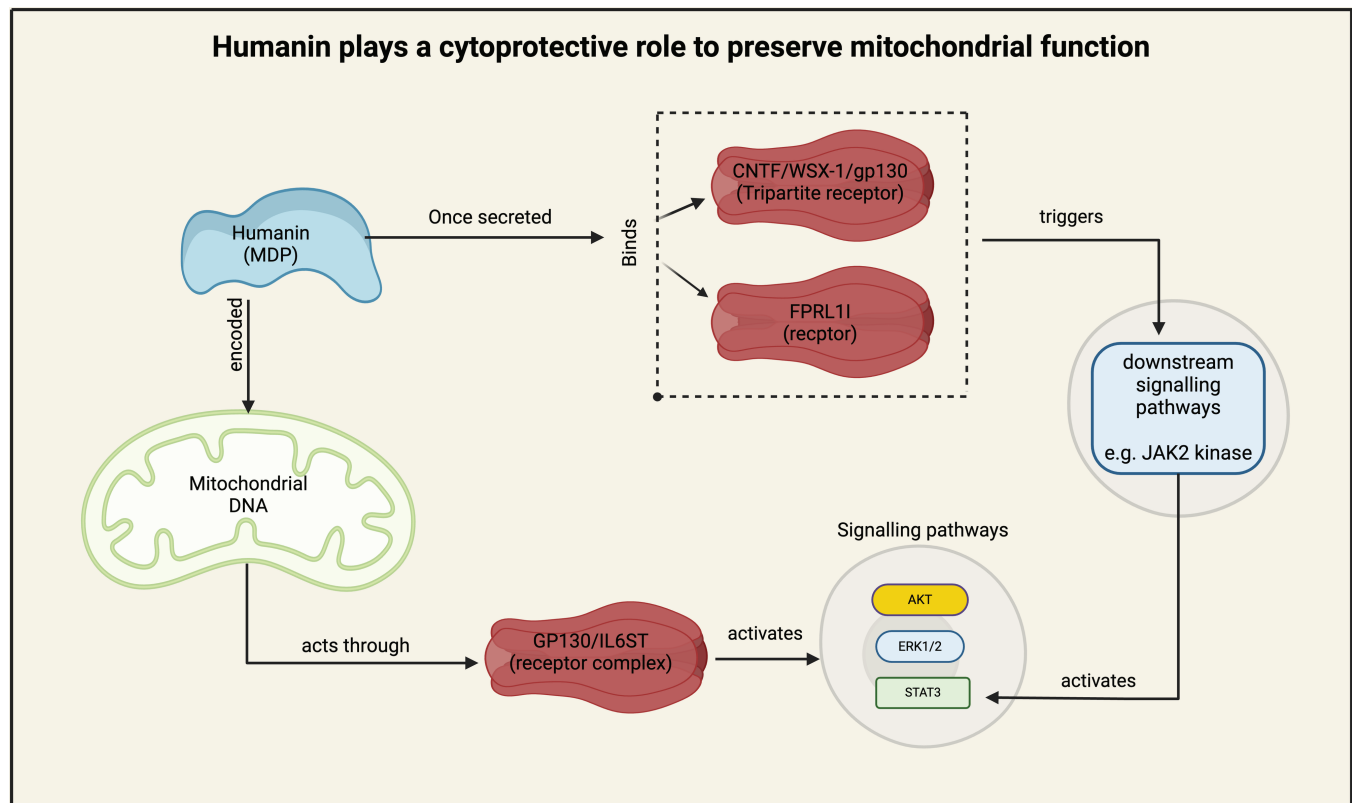
Fig 1. Pathways of humanin.

(5-aminoimidazole-4-carboxamide ribonucleotide), and AMPK (Adenosine Monophosphate-activated protein kinase). Folate is a type of B vitamin that is important for various cellular functions, including the production of DNA and other genetic material. AICAR is a compound that can activate the AMPK protein, which is involved in the regulation of energy levels in cells. When AICAR activates AMPK, it can influence various processes in the body, such as cell growth, metabolism, and responses to stress [26]. This pathway is important for understanding how the body regulates energy and other essential functions at the cellular level. The expression and function of MOTS-c have been implicated in the pathogenesis of various diseases [23,27,28]. Similarly, Mots-c has been associated with psychiatric disease, but further research is required to determine any connection to major depressive disorder [29].

p66shc is a nucleus-derived peptide but acts on the mitochondrial electron transport chain [30,31]. It serves as a redox enzyme that regulates reactive oxygen species (ROS)[1] metabolism and apoptosis, influencing cell injury pathways [32–34]. Its role in cardiovascular pathophysiology is well established [35,36]. Fig 3 shows the pathways of p66Shc, where Protein Kinase C (PKC), a family of enzymes that play a crucial role in cellular signaling pathways, phosphorylates p66shc and translocates to mitochondria, thus promoting ROS production. Dysregulation of p66shc expression and function has been implicated in the pathogenesis of various diseases and psychopathology [30,37–39].

---

[1]   ROS is a group of highly reactive molecules that contain oxygen, such as superoxide, hydrogen peroxide, and hydroxyl radicals. These molecules can cause damage to cellular components, including DNA, proteins, and lipids.
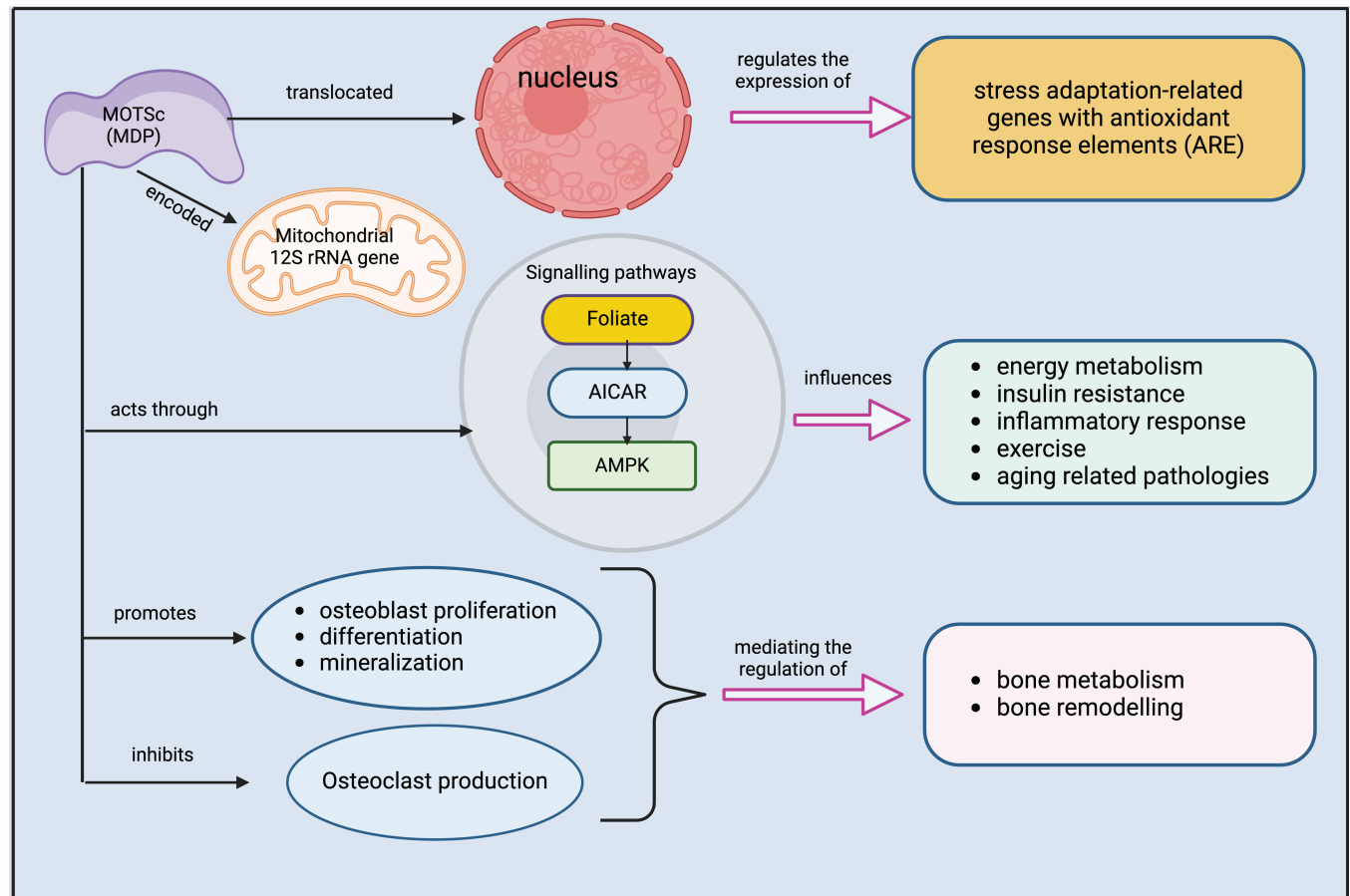
Fig 2. Pathways of MOTS-c.

https://doi.org/10.1371/journal.pone.0320955.g002

1.2 Angiotensin-converting enzyme (ACE) genotype and depression

The ACE genotype has been shown to be associated with depression in several studies, including variants of the ACE gene linked to cortisol secretion and late life depression [40,41]. Furthermore, the ACE genotype has been associated with cardiovascular disease, diabetes [42], cardiac autonomic neuropathy as part of diabetes progression [43], COVID-19, and recurrent aphthous stomatitis [40,41]. In addition, mice lacking ACE have been shown to have a lower systolic blood pressure [44]. In particular, the ACE DD genotype is associated with increased plasma ACE concentrations, which affects the conversion of angiotensin I to angiotensin II [41,45]. However, the precise nature of the relationship between the ACE genotype and depression remains a subject of ongoing debate, necessitating further investigation [41,46].

1.3 Clinical factors in depression

Research has confirmed the coexistence of depression and cardiovascular disease (CVD). Patients recovering from a myocardial infarction often experience depressive symptoms, which complicates the prognosis of CVD, including heart failure and [47–50]. This association is influenced by shared risk factors, unhealthy lifestyles, and chronic diseases [49]. Mental stress can negatively affect heart health, increasing the risk of adverse cardiac events
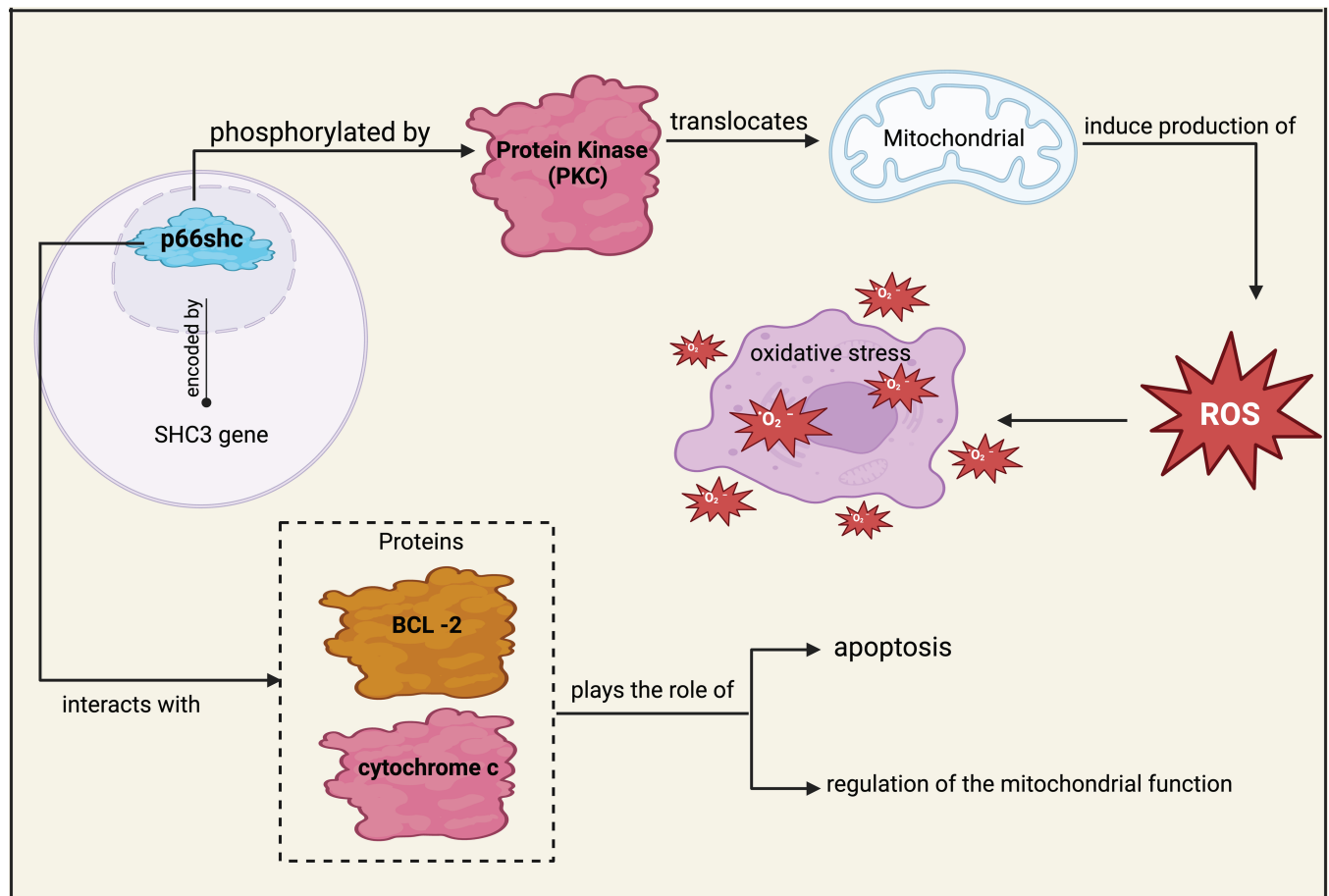
Fig 3. Pathways of p66shc.

https://doi.org/10.1371/journal.pone.0320955.g003

[51,52]. Depression in patients with heart failure leads to higher hospital readmission and mortality risks due to its association with heart attacks and blood clots [53]. Hypertension and diabetes, common chronic diseases, are also linked to CVD and depression [48, 54]. Despite the well-established connection between CVD and depression, the underlying mechanism remains largely unknown [55,56]. The complex relationship between depression, CVD, hypertension, and diabetes requires further research for effective management strategies.

The potential roles played by Humanin, MOTS-c, p66shc, and the ACE genotype have remained largely unexplored, representing a significant gap in understanding depression and disease progression. This study aims to fill this critical gap by examining the impact of Humanin, MOTS-c, p66Shc, and ACE genotypes in the context of cardiovascular disease (CVD), hypertension (HT) and diabetes mellitus (DM). This research sought to clarify the intricate interplay between these mitochondrial factors and the progression of depressive disorders, shed light on potential avenues for therapeutic interventions, and advance understanding of the role of mitochondrial-derived oxidative stress peptides in the progression of depressive disorders.

2 Methodology

2.1 Data description

The data set used in this study was obtained from a public screening program (DiabHealth) between 2002 and 2015, which was approved by the Human Research Ethics Committee (ethics approval number: 2006/042) and followed the Helsinki protocol. All participants gave their informed consent in writing to participate in this study. The data set consists of 261 screening reports with patient ages ranging from 25 to 92 years. To limit the data points reported from the available clinical report, the last reported entry was used for each patient when there were multiple attendances. Therefore, the final data set consisted of 830 individuals. Humanin, MOTS-c, p66Shc, ACE genotype, gender, age, medication use, CVD, hypertension (HT) and type 2 diabetes mellitus characteristics were extracted. Depression scores were determined using the Patient Health Questionnaire (PHQ-9). The PHQ-9 score of 5 categories or classes of depression is shown in Table 1. The higher the score, the more symptoms of depression there are and the more severe the depression.

2.2 Data preprocessing

The data set consists of 13 variables, including the target variable (the PHQ-9 score). Predictor variables include four biomarkers (Humanin, MOTS-c, p66shc, and ACE levels), gender, birth age, use of other medications, including antidepressants, CVD, hypertension (HT), use of HT medication, diabetes mellitus (DM), and use of DM medication. Feature selection was carried out using the Pearson correlation test.

However, there are more advanced and robust system of feature selection that could be used. One of such is the Boruta selection technique. The Boruta feature selection technique is an all-relevant feature selection method designed to identify all the important features in a dataset. It works by creating shadow features, which are duplicates of the original features with randomly shuffled values. A random forest classifier is then trained on the dataset with both original and shadow features. The importance of each feature is measured, and features that have a higher importance score than the best-performing shadow feature are considered important. This process is iteratively repeated to ensure robust feature selection, effectively distinguishing between relevant and irrelevant features [57]. However, Boruta can sometimes eliminate features that might be relevant in a specific research context, making it less suitable for studies requiring the inclusion of specific types of variables. Specifically, in this study, upon using Boruta, only four features (Humanin, MOTSc, P66, and BirthAge) were reported as important. This limited selection of features does not align with the research focus, which aims to investigate the role of mitochondrial peptides in the presence of various clinical factors.

Additionally, the Boruta feature selection technique, while robust, may not be the best applicable technique in the context of this study. By its nature of comparing feature

Table 1. Depression severity level with corresponding PHQ-9 scores

Depression severity	PHQ-9 score
Normal or no depression	0–4
Mild depression	5–9
Moderate depression	10–14
Moderately severe depression	15–19
Severe depression	≥ 20

https://doi.org/10.1371/journal.pone.0320955.t001

importance against randomized features, Boruta can sometimes lead to overfitting, especially if the dataset is not sufficiently large. Also, Boruta's performance is highly dependent on the quality of the data. Noisy or irrelevant features can skew the results, leading to less reliable feature selection. In contrast, the Pearson correlation test offers some advantages for feature selection, in the context of this study. The Pearson correlation test is straightforward to implement and computationally less demanding compared to Boruta. It is effective at identifying linear relationships between features and the target variable. This can be particularly useful for initial feature selection in datasets where linear relationships are expected. Furthermore, by selecting features based on their linear correlation with the target variable, the risk of overfitting is reduced, making the model potentially more generalizable to new data.

Therefore, while Boruta might overly simplify the feature set by focusing on only a few variables, Pearson correlation allows for a more inclusive selection process. This inclusivity is crucial for the research as it focuses on the peptides within the broader context of clinical factors, ensuring that the complexity of depression severity is adequately captured and analyzed.

Using the Pearson correlation test a threshold of 0.7 was used and, as evident in the correlation map in Fig 4, the DM status and HT status are highly correlated with DM-MedUse and HT-MedUse with values above the threshold of 0.7: 0.80 and 0.77, respectively. Consequently, DM-MedUse and HT-MedUse were removed to ensure that there is no high correlation between predictors, which reduced their number to ten as shown in Fig 5.

Depression classes were divided into binary classes (no depression and depression), three classes (no depression, mild depression, and severe depression), and five classes (from no depression to severe depression). The study adhered to the conventional multiclass division of five categories, as stipulated by the recognized standards of the Patient Health Questionnaire-9 (PHQ-9) score, which serves as a benchmark for assessing depression levels. However, the investigation conducted supplementary experiments to explore the feasibility of refining the classification scheme by reducing the severity levels of depression into binary and three-class divisions. Tables 2–4 show that there is an expected class imbalance for multiclass classification (5-class and 3-class) and binary classification.

Due to imbalanced data sets, a synthetic minority oversampling algorithm (SMOTE) [58] was included in the analysis when dealing with imbalanced data, with precision, recall, and F1 scores reported.

2.3 Machine learning algorithms

To classify the progression of depression, six different Machine Learning (ML) classifiers were used, including random forest (RF), support vector machine (SVM), logistic regression (LR), K-nearest neighbor (KNN), gradient boosting (GB), and artificial neural network, also known as multilayer perception (MLP).

After feature selection, grid search was used with k-fold cross-validation (k = 10) to adjust the hyperparameters of the classifiers. The data set was then divided into training and testing sets to avoid overfitting. This study used 80% of the sample data as training data and the remaining 20% as test data. The performance of the best-tuned models was validated using accuracy, precision, recall, and F1 score. Furthermore, the permutation importance was used to determine the importance of predictor variables by calculating their statistical contribution to the performance of the models.

All experiments in this study were performed with Python using the *scikit-learn* machine learning library version 1.4.0 and the *Tensorflow* 2.0 neural network library on a 16GiB RAM macOS Monterey machine with the Apple M1 Pro Chip. Fig 6 shows the flowchart of the steps
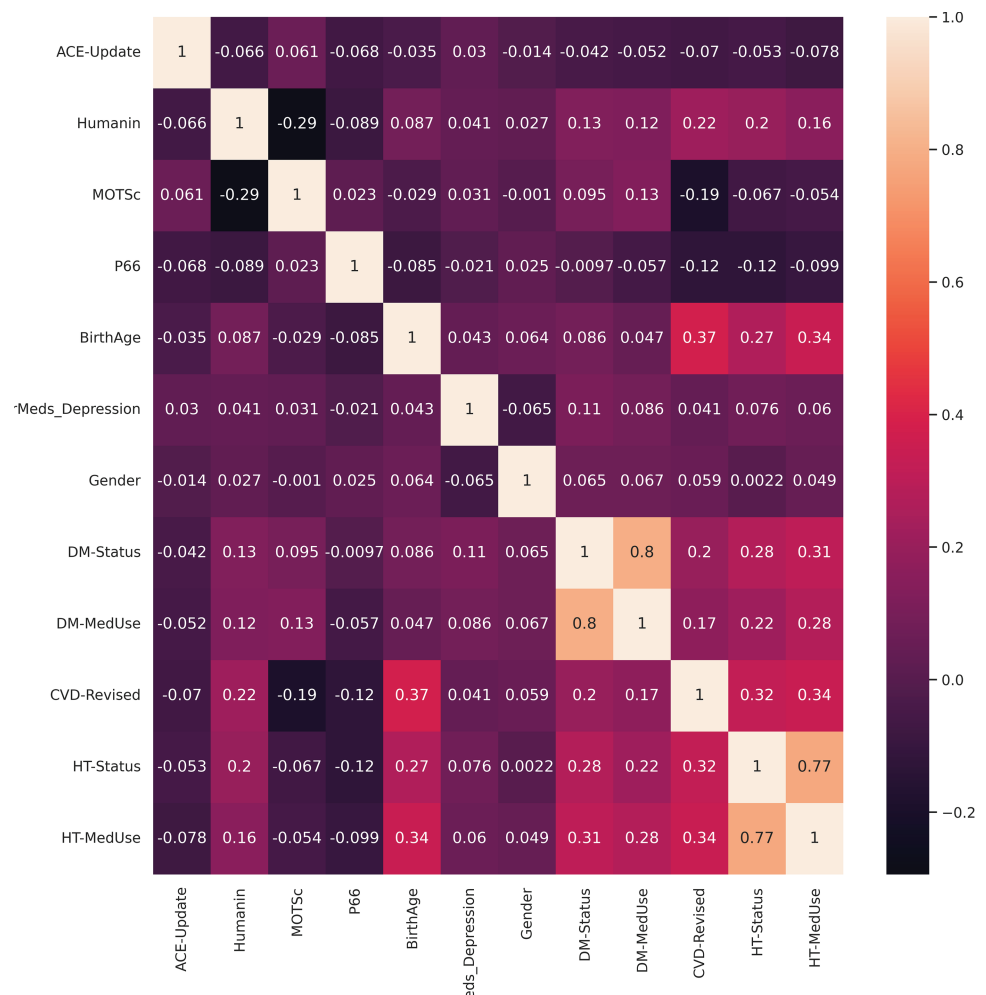
Fig 4. Correlation matrix of all twelve features.

used in this study to predict the severity of depression. Below is a brief description of the ML algorithms used.

2.3.1 Random forest Random forest (RF) classifiers fall under the broad umbrella of ensemble-based learning methods. They are simple to implement, easy to operate and have proven to be extremely successful in a variety of domains [59]. The key principle underlying the RF approach comprises the construction of many "simple" decision trees in the training stage and the majority vote (mode) across them in the classification stage. Among other benefits, this voting strategy has the effect of correcting for the undesirable tendency of decision trees to overfit training data. In the training stage, RF applies bagging to individual trees in the ensemble. Bagging repeatedly selects a random sample with replacement from the training set and fits trees to these samples. Each tree is grown without any pruning. The number of trees in the ensemble is a free parameter that is easily learned automatically using the so-called out-of-bag error.

RF is chosen for this study for its simplicity, ease of operation, and success across various domains. It is robust to overfitting due to the ensemble nature of combining multiple decision
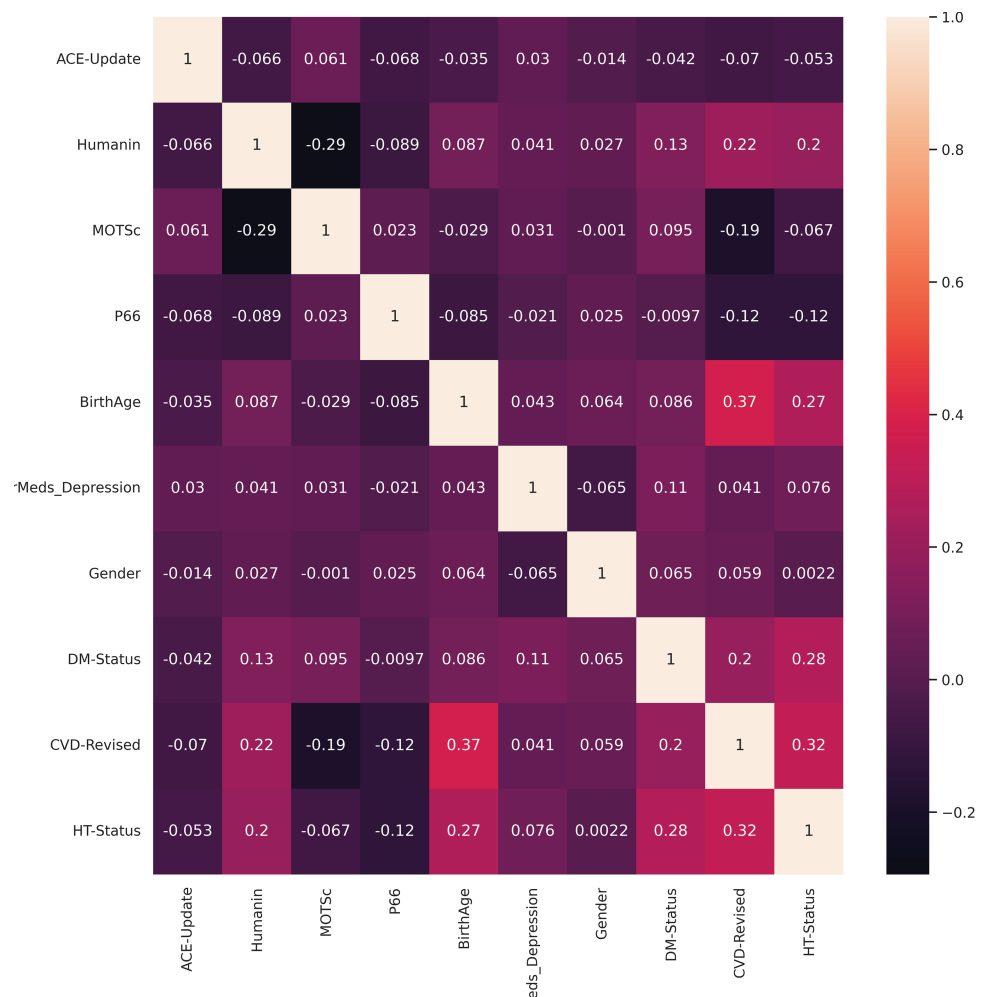
Fig 5. Correlation matrix of features after the dropping HT-MedUse and DM-MedUse.

trees, and it performs well with imbalanced data as it can handle large datasets with higher dimensions and capture complex interactions between features.

2.3.2 Support vector machine (SVM) Support Vector Machine (SVM) is a powerful supervised learning algorithm that is used for classification, regression and outlier detection tasks. SVMs are particularly effective for solving binary classification problems, where the algorithm aims to find the best possible line or decision boundary (a hyperplane) that separates the data points of different classes [60]. This decision boundary is established in such a way that it maximizes the margin, which is the distance between the hyperplane and the closest data points in each category, making it easier to distinguish data classes. SVM can also be adapted for multiclass classification and regression tasks. SVM has different types and variants that provide specific functionalities, such as linear SVM and nonlinear SVMs. In this study, the RBF kernel SVM was used because it is capable of capturing complex relationships in the data, making it suitable for nonlinear classification tasks, especially in low-dimensional spaces.

Support Vector Machine (SVM) is selected for its effectiveness in binary classification problems and its versatility in handling linear and non-linear data through kernel functions,

Table 2. Class distribution for 5-class classification.

Depression severity	PHQ-9 score	Distribution
Normal or No depression	0–4	611
Mild depression	5–9	170
Moderate depression	10–14	29
Moderately severe depression	15–19	11
Severe depression	≥20	9

https://doi.org/10.1371/journal.pone.0320955.t002

Table 3. Class distribution for 3-class classification.

Depression severity	PHQ-9 score	Distribution
Normal or no depression	0–4	611
Mild depression	5–14	199
Moderate depression	≥15	20

https://doi.org/10.1371/journal.pone.0320955.t003

Table 4. Class distribution for binary classes.

Depression severity	PHQ-9 score	Distribution
Normal or no depression	0–4	611
With depression	≥20	219

https://doi.org/10.1371/journal.pone.0320955.t004

such as the RBF kernel. SVM is particularly effective for high-dimensional spaces, making it a robust choice for various classification tasks.

2.3.3 Logistic regression Logistic regression is a supervised learning classification algorithm that uses a logistic function to predict the probability of a target variable, typically binary in nature [61]. The algorithm works by computing the sum of the input features (including a bias term) and calculating the logistic of the result. The output of logistic regression is always between 0 and 1, which is suitable for binary classification tasks. The higher the value, the higher the probability that the current sample is classified as Class 1. Logistic regression can be adapted for multinomial classification by creating multiple binary classification problems. It is used in this study for both binary and multiclass classifications.

Logistic Regression (LR) is utilized for its suitability in predicting binary outcomes and its ability to adapt for multinomial classification. It is straightforward and interpretable, making it a good baseline model that works well with binary and multiclass classification problems.

2.3.4 Gradient boosting Gradient boosting (GB) is a form of machine learning boosting. It builds a sequential ensemble of predictors, each of which corrects the previous one. This strategy aims to adapt the new predictor to the residual errors of the old one. When the best future potential model is integrated with the past models, the total prediction error is reduced. The essential concept is to set the target results for the subsequent model to minimize the error. In gradient boosting, the target outcomes are dependent on the error gradient relative to the forecast. Each new model minimizes the prediction error for each training instance [62]. The goal for each case depends on how much the forecast changes the total error. If a small adjustment in the forecast reduces the error, the following target result will be high. The new model's near predictions will decrease the error. However, if a small adjustment in the forecast does not affect the error, the target result is zero, and this forecast cannot be improved on[63].
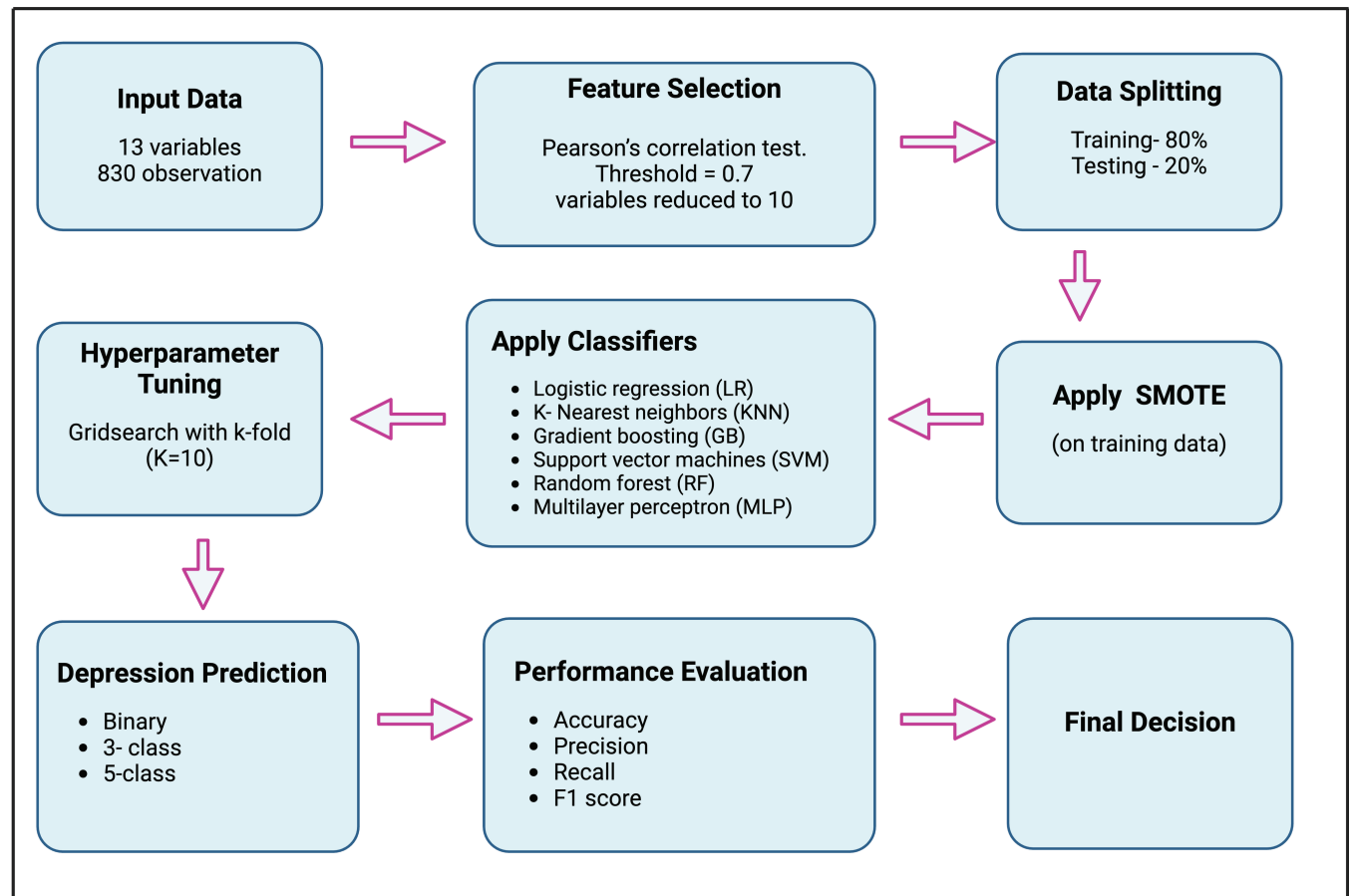
Fig 6. Flowchart for predicting depression severity.

Gradient Boosting (GB) is employed for its boosting technique, which builds a sequential ensemble of predictors that correct previous errors, thus minimizing prediction errors effectively. Known for its accuracy and ability to handle complex datasets, GB constructs models sequentially, correcting errors from previous models.

2.3.5 K-nearest neighbor The K-Nearest Neighbors (KNN) algorithm is a simple non-parametric supervised learning classifier that uses proximity to make classifications or predictions about the grouping of individual data points. It is typically used as a classification algorithm, working from the assumption that similar points can be found near each other [64]. It is easy to implement and understand, but has the major drawback of becoming significantly slower as the size of data grows. KNN works by finding the distances between a query and all the examples in the data, selecting the specified number of examples (K) closest to the query, and then voting for the most frequent label. Choosing the right K involves trying several K's and picking the one that works best. In this study, the best K was 4.

K-Nearest Neighbors (KNN) is chosen for its simplicity and effectiveness in classification tasks, based on the proximity of data points. It works well with smaller datasets and is straightforward to implement. As a non-parametric method, KNN makes predictions based on the closest data points in the feature space, which can be beneficial for certain types of imbalanced data.

2.3.6 Artificial neural network Artificial Neural Network (ANN) is a machine learning algorithm inspired by the structure and function of the human brain. It is also known as a multilayer perceptron (MLP) connectionist model or parallel distributed processing system [65]. ANN is widely used for classification tasks such as image recognition, speech recognition, and natural language processing. ANN consists of multiple layers of interconnected nodes, or neurons, that process and transmit information. Each neuron receives input from other neurons, processes the information, and then passes the output to other neurons in the next layer. The output of the final layer represents the classification or prediction of the input data. ANN is known for its ability to learn from data and generalize to new data, making it a powerful tool for classification tasks. ANN has several types, including feedforward neural networks, convolutional neural networks, and recurrent neural networks, each with specific architectures and applications [66]. In this study, feedforward neural networks were used.

Artificial Neural Network (ANN) is selected for its powerful classification capabilities, particularly in tasks like image and speech recognition, due to its multilayered structure that mimics the human brain. ANNs can model complex patterns and interactions in the data, making them highly flexible and adaptable for various tasks.

The choice of the ML methods used in this study was guided by their proven effectiveness and suitability for the classification tasks at hand. They were chosen from different categories of algorithms, including linear (LR), nonlinear (SVM, KNN), ensemble (RF, GB), and neural network (ANN). This diversity in the application of machine learning methods helps in the prediction process by leveraging the strengths of various types of algorithms. Linear models like LR are simple and interpretable, making them good baseline models. Nonlinear models such as SVM and KNN handle complex relationships and interactions in the data. Ensemble methods like RF and GB improve prediction accuracy by combining multiple models. ANNs are powerful in modeling complex patterns due to their multilayered structure. This level of diversity ensures a comprehensive approach to predicting depression severity, accommodating the complexities and nuances within the dataset.

Table 5 outlines the criteria and reasons for selecting these machine learning methods, emphasizing their strengths.

3 Results

This study presented comprehensive results that shed light on the performance of the ML models. Specifically, detailed tabulated data for each class were provided in three distinct models, namely Model 1, Model 2, and Model 3. Model 1 consisted of all predictors and Model 2 consisted of biomarkers of oxidative stress (p66shc, Humanin and MOTS-c), ACE, age, and gender. Model 3 consisted only of biomarkers of oxidative stress. Model 3 was included to investigate the impact of oxidative stress biomarkers on the classification of severity of depression. To ensure accessibility to the performance of the three models, all result tables have been included in the supplementary section of this study. S9–S11 Tables in S1 Text show the results for all classes for Model 1, S12–S14 Tables in S1 Text for Model 2, and S15–S17 Tables in S1 Text for Model 3. Each table shows the result values for the six classifiers that were used in this study. The most pertinent results are summarized in Section 3.2.

3.1 Understanding the data using explainable AI

In recent years, the application of explainable AI (XAI) techniques has become increasingly vital for interpreting complex machine learning models, particularly in high-stakes domains

Table 5. ML algorithms used in this study.

Type of Algorithm	Algorithm/method	Strengths/ reasons for selection
Linear	Logistic regression (LR)	● good baseline model that works well with binary
		● ability to adapt for multinomial classification
		● straightforward
		● Interpretable
Nonlinear	K-nearest neighbor (KNN)	● Simplicity and effectiveness in classification tasks, based on the proximity of data points
		● It works well with smaller datasets Straightforward to implement
	Support vector machine (SVM)	● Effectiveness in binary classification problems
		● Particularly effective for high-dimensional spaces
		● Versatility in handling linear and non-linear data through kernel functions
Ensemble	Randomforest (RF)	● Simplicity
		● Ease of operation
		● It is robust to overfitting due to the ensemble nature
		● It captures complex interactions between features It performs well with imbalanced data
	Gradientboosting (GB)	● Known for its accuracy and ability to handle complex datasets
		● It constructs models sequentially, correcting errors from previous models.
Neural Network	Artificial neural network (ANN)	● Ability to model complex patterns and interactions in the data
		● Highly flexible and adaptable for various tasks

such as healthcare, finance, and social sciences [67]. This study leverages several XAI methods, including permutation importance [68], Local Interpretable Model-agnostic Explanations (LIME) [69] and SHapley Additive exPlanations (SHAP) [70], to gain deeper insights into the data and the models trained on it. By utilizing these tools, we aim to elucidate the intricate relationships between features and model predictions, providing a transparent and interpretable framework for understanding the factors that drive model decisions.

Permutation feature importance as a technique can be leveraged to quantify the contribution of each feature to the predictive performance of the machine learning models. It can also be visualized using plots. In this study, calculating the importance of the variables provided additional information on how each predictor variable contributed to the performance of the models. This analysis helped identify the key biomarkers and clinical factors that were most predictive of depression severity.

Figs 7–9 depict the permutation importance of features for Model 1 for 2, 3, and 5 classes, respectively using the RF classifier. This analysis shows the similarities and differences in the importance of certain biomarkers (peptides) across these models, as well as in the consistency of the topmost and bottommost variables. The following are noticeable observations: Across all three models, the peptides (P66, Humanin, and MOTSc) consistently rank highly, with P66 often being the top variable. This indicates that these peptides are crucial biomarkers or features across the different classes of classification tasks. The top variable shifts from DM-Status in the 2-class model to P66 in the 3-class and 5-class models, suggesting that as the classification task becomes more complex, P66 becomes increasingly important. Similarly, the least important variable is consistently a feature related to other medications for depression or ACE updates, which may indicate that these factors do not contribute significantly to distinguishing between the classes in the dataset.
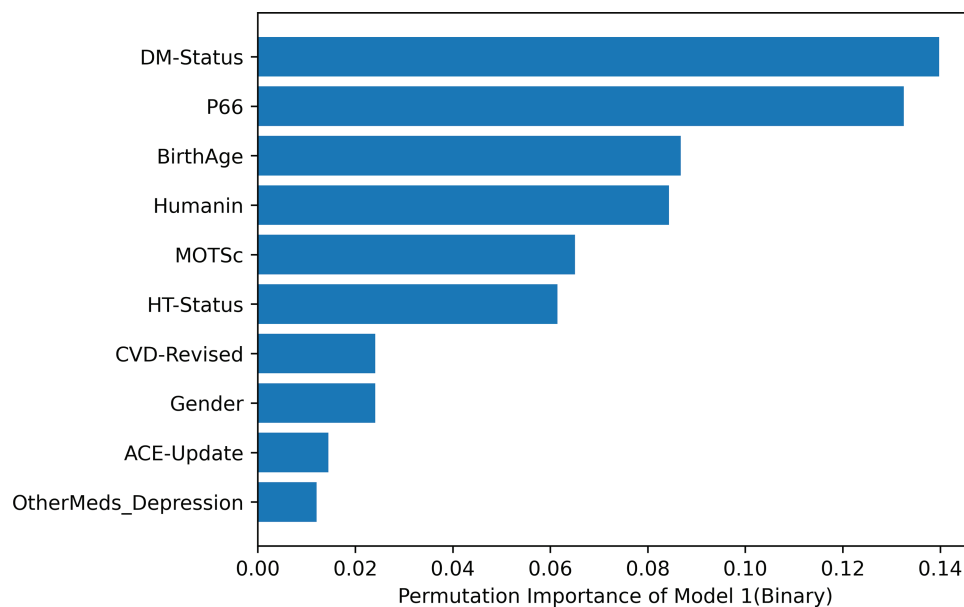
Fig 7. Model 1: All Variables for binary classification.
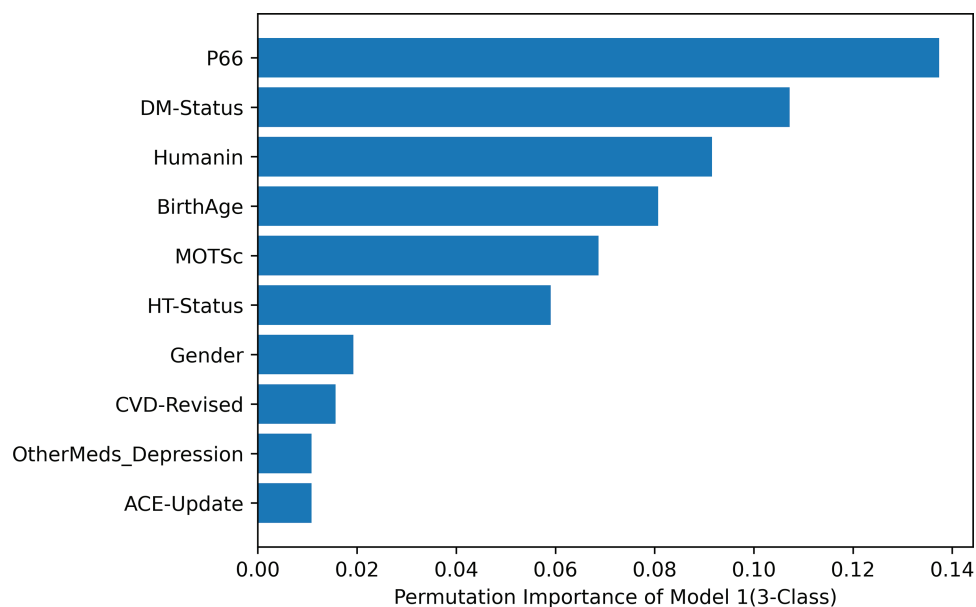
Fig 8. Model 1: All Variables for 3-class classification.

Furthermore, Figs 10 and 11 show the permutation importance of features for Model 1 and Model 2 respectively. In Model 2, biomarkers ranked highest on the order of p66shc, MOTS-c, and humanin while ACE ranks least showing that it is the least important variable in predicting the severity of depression in this particular Model. Model 3 indicated that of the three biomarkers of oxidative stress, p66shc has the highest rank, followed by MOTS-c and Humanin.

Fig 9. Model 1: All Variables for 5-class classification.
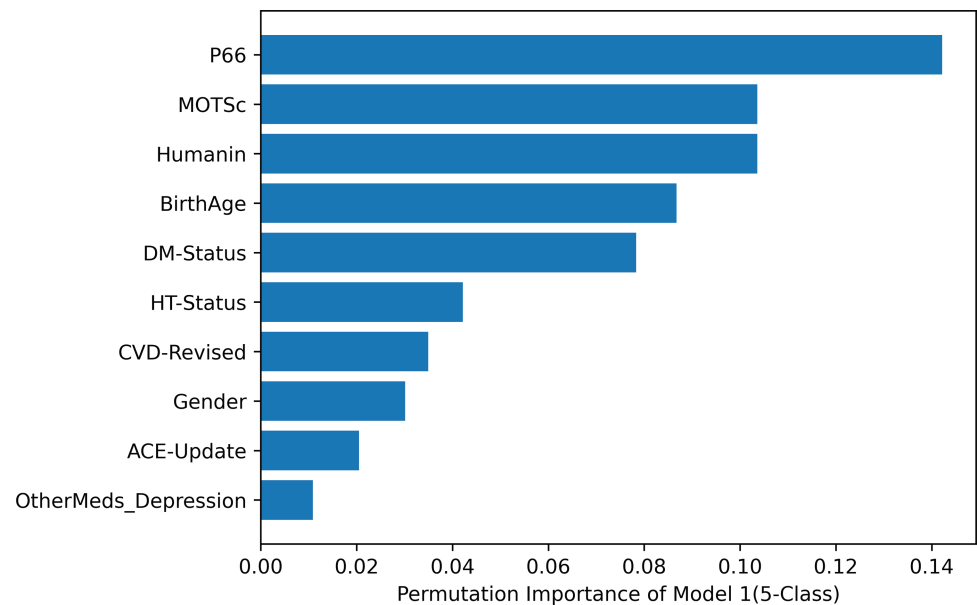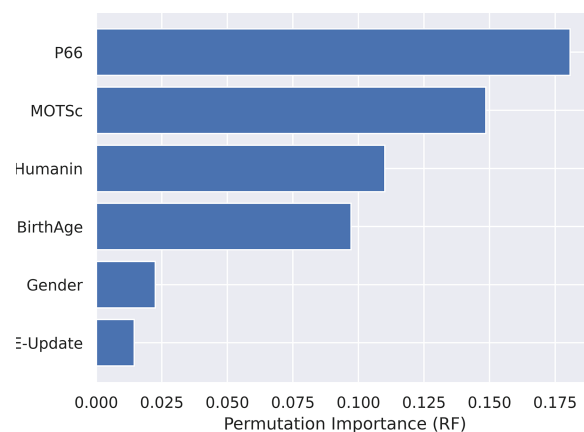
Fig 10. Model 2: Biomarkers + ACE + Age + Gender.

Beyond the ranking of features or variables using permutation importance, much more can done to understand and interpret data and models. A powerful XAI tool that can be used for this purpose is SHAP [70]. SHAP is a unified framework for interpreting predictions by assigning each feature an importance value (called SHAP values) that reflects its contribution to the model's output. Derived from cooperative game theory, SHAP values offer a consistent and theoretically sound approach to feature attribution, ensuring that the contributions of all features are fairly represented. In this study, SHAP is utilized to generate both global and local explanations of model 1's behavior, including feature importance rankings and detailed instance-level breakdowns of prediction contributions.

Fig 12 depicts the SHAP summary plot providing an overview of feature importance across all instances in the dataset and ranking them accordingly. Compared with Fig 7 which is the
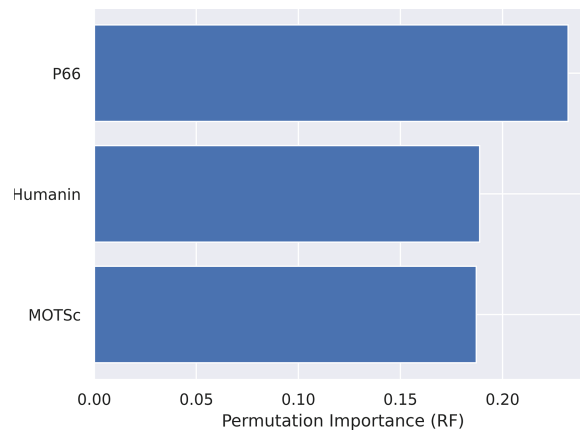
Fig 11. Model 3: Biomarkers only.

permutation feature importance generated by using the RF classifier, it is very similar. Both plots show that biomarkers (p66shc, MOTS-c, and humanin) are amongst the top-ranked features, while ACE and other medications for depression consistently rank lowest.

Fig 13 displays the SHAP summary plot that combines feature importance with feature effects across all instances in the training set. The plot sorts features by the sum of SHAP value magnitudes over all samples and uses SHAP values to show the distribution of the impacts each feature has on the model output. The y-axis position is determined by the feature, while the x-axis reflects the corresponding SHAP value. Each point on the plot represents a SHAP value for a specific feature and instance. The plot indicates that "DM-Status," "P66," and

Fig 12. SHAP feature importance for Model 1 (binary classification).

Fig 13. The SHAP Summary plot of Model 1's Feature effects.

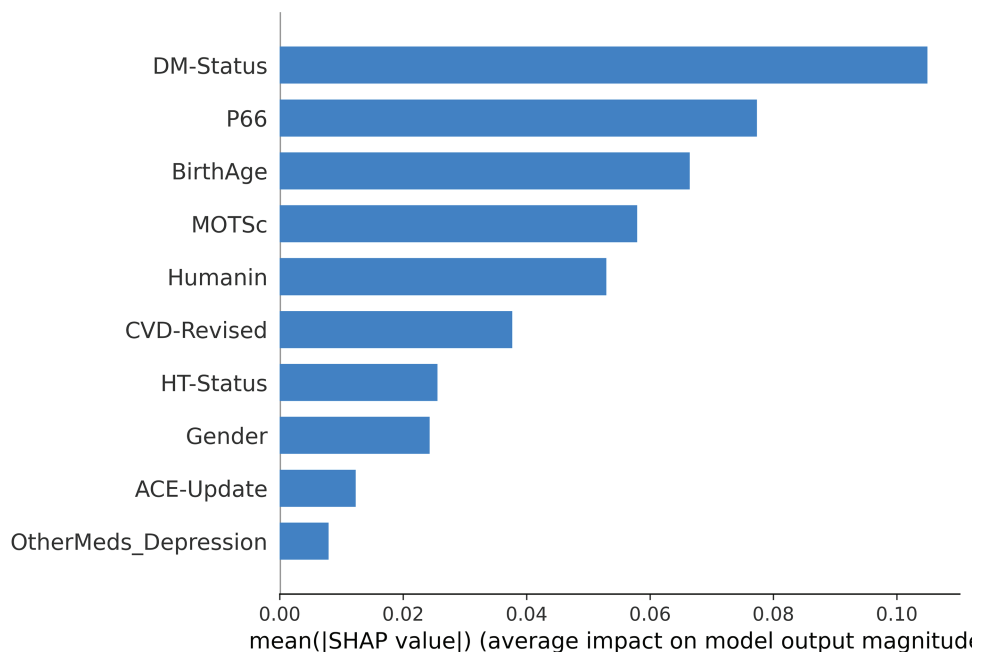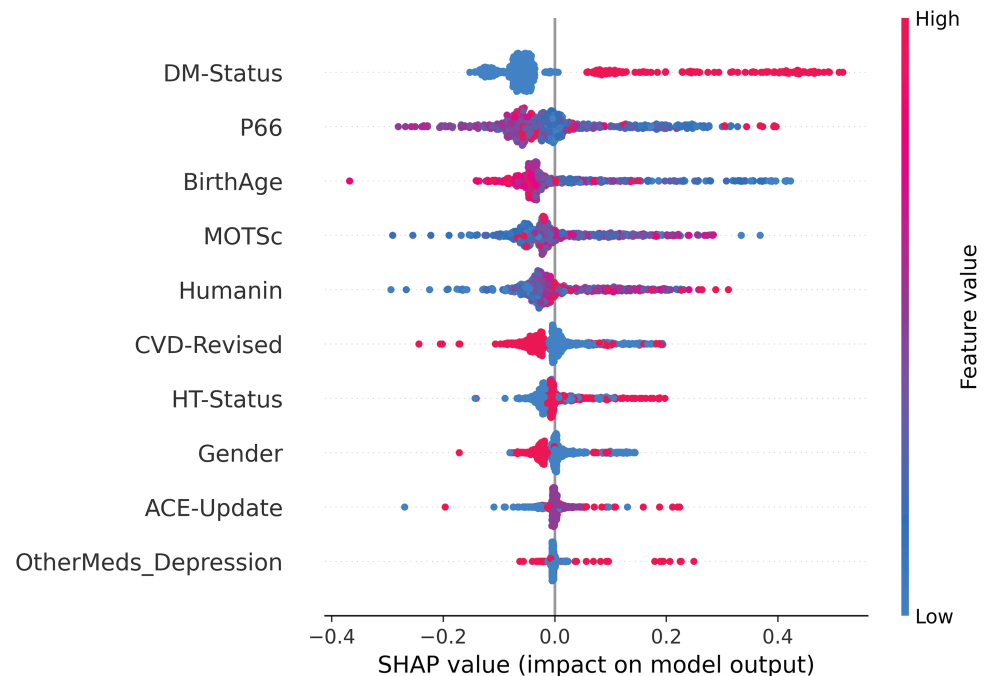"BirthAge" are among the most influential features in determining the model's output, with higher SHAP values indicating a stronger impact on the prediction. The color gradient from blue to red represents the feature values from low to high. For instance, a high "DM-Status" value is associated with a lower SHAP value, indicating a tendency towards predicting Class 0. This plot provides a comprehensive view of how each feature interacts with the model across the dataset, allowing for a deeper understanding of the model's behavior.

Fig 14 is the SHAP waterfall plot showing the breakdown of contributions from individual features to the model's prediction for Instance 6 (the sixth patient in the training dataset). Each bar represents the impact of a feature, with positive contributions (in red) pushing the prediction towards Class 1 (depressed) and negative contributions (in blue) pulling it towards Class 0 (not depressed) from the base value prediction$(E(f(x)) = 0.256)$, which is the average prediction the model would make if no features were known or if we considered all instances in the dataset without any specific information about an individual instance. "BirthAge" and "P66" are the most significant contributors, pushing the prediction closer to Class 1, while "DM-Status" and "CVD-Revised" slightly counterbalance this effect by contributing towards Class 0. The final prediction probability ($f(x) = 1$ - indicating that the patient is depressed) is represented at the rightmost point of the plot, summing up the cumulative effects of all features. The values in between (written on either the red or blue bars) indicate the SHAP value of the features. For example, while "BirthAge" contributes to the final prediction by pushing towards it with a SHAP value of 0.37 from the based value prediction, the "DM-Status" pushes away from the final prediction with a SHAP value of 0.06.

Another popular powerful XAI method for explaining individual predictions by approximating the local behavior of complex models with simple, interpretable models, such as linear models is LIME [71]. By perturbing the input data and observing the changes in predictions,
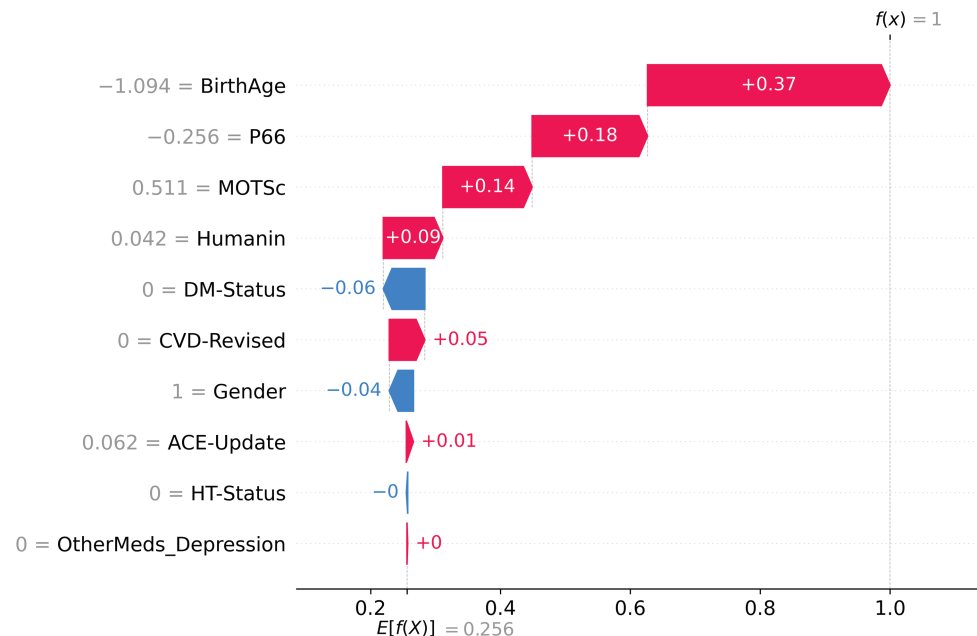
Fig 14. The SHAP Waterfall Plot for Model 1's Prediction for Instance 6.

LIME generates a set of explanations that highlight which features most significantly contribute to a particular prediction. In this study, LIME is applied to some instances to unravel the reasoning behind specific predictions made by the model.

Fig 15 shows the LIME explanation for Instance 20 illustrating the contribution of the top features to the prediction outcome of the binary classification model. The model predominantly associates "DM-Status" with a high likelihood of belonging to Class 0 (not depressed), as indicated by the large negative contribution (red bar). In contrast, features like "P66" and "Humamin" have substantial positive contributions (green bars) toward predicting Class 1 (depressed). This visualization highlights how the model's prediction for this specific instance is largely driven by the diabetic status (DM-Status), with other features such as P66 and Humanin influencing the probability distribution between Class 0 and Class 1. The class probabilities are presented at the bottom, showing a 0.82 probability for Class 0 and 0.16 for Class 1.

Figure 16 shows the LIME explanation for another instance (Instance 70), detailing the relative contributions of various features to the model's prediction. Similar to Instance 20, "DM-Status" plays a significant role, but here it is again linked with Class 0. Other features such as "BirthAge" and "P66" show notable negative contributions, pulling the prediction towards Class 0. On the other hand, "Humamin" and "Gender" contribute positively towards Class 1. The outcome probabilities reveal that this instance has a more balanced prediction, with a probability of 0.37 for Class 0 and 0.63 for Class 1. This suggests that while "DM-Status" heavily influences the prediction, other features also play critical roles, leading to a more nuanced outcome.

3.2 Performance evaluation of random forest results on the three models

Tables 6–8 describe our results with a particular focus on the best performing classifier, namely Random Forest (RF). Specifically, Model 3, which exclusively incorporates
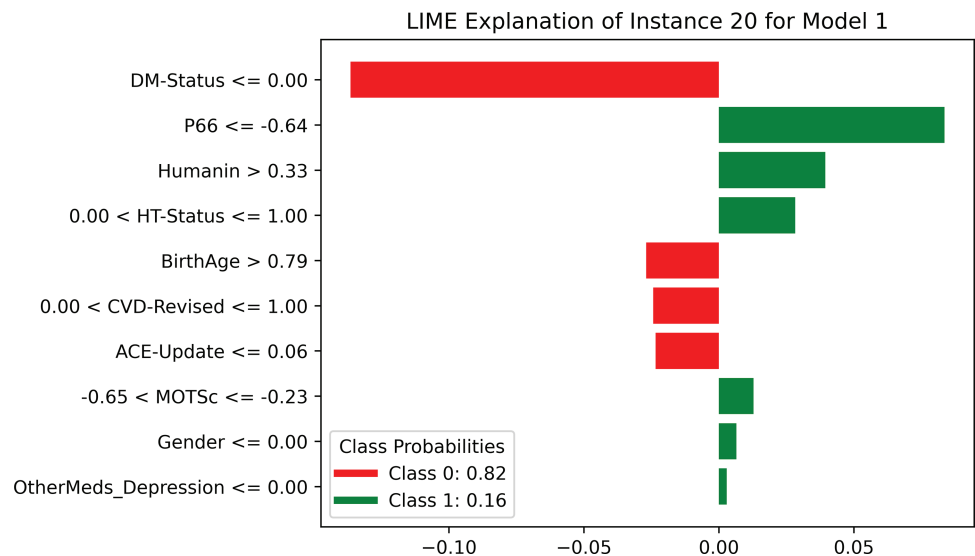
Fig 15. The LIME explanation for Instance 20.
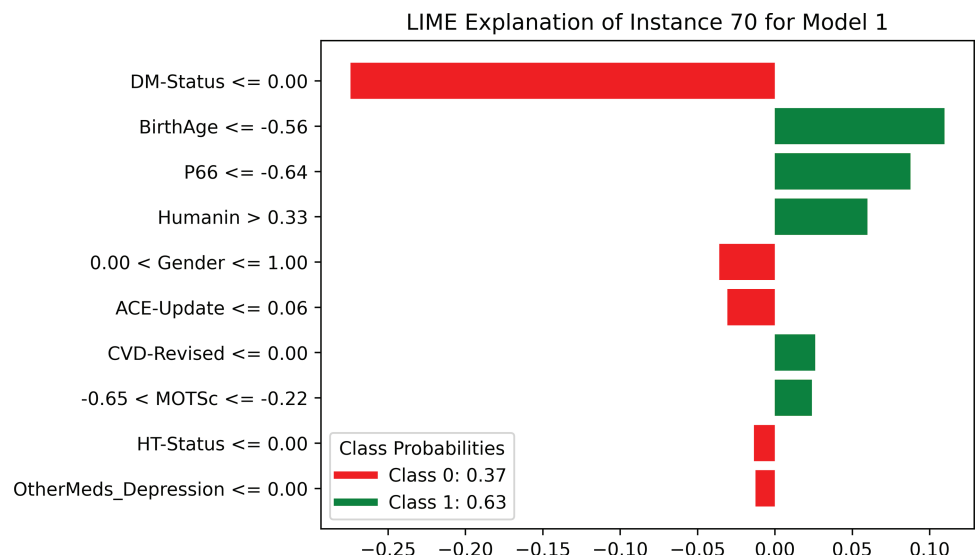
Fig 16. The LIME explanation for Instance 70.

biomarkers, is compared to Models 1 and 2. Model 1, which covers all variables considered in our experimental setup, consistently exhibits superior precision and F1 scores in all classification categories. This finding underscores the crucial role that a comprehensive set of variables plays in improving predictive accuracy. Model 3, which only includes biomarkers, produces results that are similar to those of Model 1, and even exceeds the performance of Model 1, especially when the data are unbalanced. This similarity in the results between Model 1 and Model 3 demonstrates the influence of oxidative stress biomarkers in the larger model (Model 1) and their association with depression.

Table 6. Comparison of the performance evaluation of Random Forest results for binary classification on both balanced and unbalanced data set.

Models	Unbalanced data set				Balanced data set			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
Model 1	90.96	96.77	68.14	80.00	92.17	91.89	77.27	80.86
Model 2	85.54	85.56	85.54	84.38	86.14	85.78	86.14	83.95
Model 3	89.76	90.12	89.76	89.13	87.35	87.13	87.35	87.2

Table 7. Comparison of the performance evaluation of Random Forest results for 3-class classification on both balanced and unbalanced data set.

Models	Unbalanced data set				Balanced data set			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
Model 1	89.67	89.17	89.76	89.1	90.36	90.5	90.36	90.1
Model 2	88.55	88.8	88.55	87.83	80.12	80.29	80.12	80.12
Model 3	89.16	89.05	89.16	88.56	84.94	86.74	84.94	85.62

Table 8. Comparison of the performance evaluation of Random Forest results for 5-class classification on both balanced and unbalanced data set.

Models	Unbalanced data set				Balanced data set			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
Model 1	87.35	85.64	87.35	85.35	89.76	87.76	89.76	88.26
Model 2	86.14	84.63	86.14	83.63	80.72	79.0	80.72	79.73
Model 3	87.35	87.02	87.35	86.35	79.52	82.76	79.52	80.86

The significance of presenting varying classification results, such as binary, three-class, and five-class classifications, lies in several key aspects that includes:

- Understanding Model Robustness: Evaluating the performance across different classification levels (binary, three-class, and five-class) allows researchers to understand the robustness and generalizability of their models. It helps in determining whether a model that performs well for simpler classifications (e.g., binary) also maintains its performance for more complex classifications (e.g., five-class).
- Benchmarking and Comparisons: The evaluation of the models' performance across different levels of classification complexity enables benchmarking against other studies or models that might have used different classification schemes. This provides valuable insights into how the models scale as the problem becomes more challenging, transitioning from a simple binary distinction to a more nuanced multi-class categorization.
- Understanding Depression Severity Thresholds and Informing Future Research Directions: This study adhered to the conventional PHQ-9 scoring system, but the supplementary experiments on binary and three-class models indicate that a more refined categorization may be feasible and beneficial. These findings can guide future research by highlighting the need to further investigate the most appropriate level of depression severity classification. It may be worthwhile to explore alternative categorization schemes or feature engineering approaches to improve the performance of the more granular multi-class models.
- Guiding Practical Application: Different real-world applications might require different levels of classification granularity. For instance, a healthcare application might need to distinguish between mild, moderate, and severe depression (three-class) rather than just depression vs. no depression (binary). By showing how models perform across these varying needs, researchers can provide guidance on the most appropriate use cases for their models.

3.3 Comparison for balanced and unbalanced data set

The imbalanced data set, comprising binary, 3-class and 5-class classification tasks, obtained accuracy rates of 90. 96%, 89. 67% and 87. 35%, respectively. These promising results established RF as the best classifier algorithm for these classification tasks. When data set balancing techniques were applied, a slight improvement in precision was observed and the F1 score for the RF classifier was recorded as 92. 17%, 90. 36%, and 89. 76% for binary, three-class and five-class classifications, respectively. However, it is important to note that these improvements were observed primarily in Model 1. In contrast, Models 2 and 3 exhibited lower performance for balanced data sets compared to their unbalanced counterparts. One possible explanation for this result is the selective removal of certain variables from Model 1. Consequently, this loss of information can lead to a decrease in overall performance when applied to the balanced data set, as observed in Models 2 and 3. Moreover, it is worth emphasizing that model performance is intricately linked to class balance within the tested data set. In the case of class imbalance, the effectiveness of a model in classifying minority classes may be compromised. This inherent dependence on data balance implies that classification performance is inherently influenced by the balance of the *test* data set.

3.4 Implication of the results

The results of this study have several significant implications for the field of mental health, particularly in the prediction and management of depression severity. The implications of these findings extend to clinical practice, preventive healthcare, and ongoing research, ultimately contributing to better management and understanding of depression.

The implications are discussed below:

Clinical implications:

- Biomarker-Based Assessment: The identification of specific mitochondrial-associated oxidative stress biomarkers, such as Humanin, MOTS-c, and p66shc, provides a new avenue for assessing depression severity. Clinicians can use these biomarkers to monitor the progression of depression, enabling more accurate and timely interventions. Also, the high-performing machine learning models, particularly the Random Forest classifier, demonstrate the potential to enhance clinical assessment and diagnosis of depression. By integrating mitochondrial biomarkers and clinical factors, these models can provide more accurate and objective measures of depression severity.
- Treatment Personalization: Understanding the role of these biomarkers allows for personalized treatment plans. Patients exhibiting certain biomarker profiles may respond better to specific treatments, which can lead to more effective management of depression and reduce the trial-and-error approach often associated with prescribing antidepressants. Therefore, the insights into the key predictors of depression severity, including the mitochondrial peptides, can inform the development of more personalized treatment approaches. Targeting specific biological pathways associated with depression may lead to more effective interventions.
- The robust and interpretable machine learning models developed in this study could be translated into practical clinical decision support tools. Such tools could assist healthcare providers in making more informed and personalized assessments and treatment decisions for patients with depression.

Early detection and intervention:

- Predictive Accuracy: With the capability to predict depression severity more accurately, there is potential for early identification of individuals at risk of developing severe depression. Early intervention can prevent the worsening of symptoms and improve long-term outcomes.
- Preventive Strategies: By identifying individuals with elevated levels of specific biomarkers, preventive strategies can be implemented to mitigate the onset or progression of severe depression. This proactive approach can reduce the burden on healthcare systems and improve patient quality of life.

Research implications:

- Expanded Biomarker Exploration: The study highlights the importance of exploring a broader range of biomarkers, including those associated with oxidative stress, inflammation, and neurobiological processes. This can uncover new insights into the underlying mechanisms of depression.
- Mechanistic Understanding: The importance of mitochondrial peptides in predicting depression severity sheds light on the underlying biological mechanisms linking oxidative stress and mitochondrial dysfunction to the development and progression of depressive disorders. This knowledge can guide future research into the pathophysiology of depression.
- Longitudinal Studies: Future research should focus on longitudinal studies with larger and more diverse sample sizes to validate these findings and explore the dynamic nature of depression and its biomarkers over time. This will help in understanding how changes in biomarker levels correlate with disease progression and treatment response.

4 Discussion

Mental health research has witnessed a surge in demand for tools that can effectively stratify mental health disorders, such as depression, and provide accurate outcome forecasts. The current study uniquely focused on the impact of mitochondrial-associated oxidative stress biomarkers in this context. In particular, the incorporation of Humanin, MOTS-c, and p66shc within the framework of machine learning techniques has enabled the prediction of the severity of depression based on a binary or multiclass model. Adding information on hypertension, diabetes mellitus, and cardiovascular disease alongside oxidative stress biomarkers added a comprehensive dimension to our predictive models. The combination of these factors reflects a holistic approach to understanding the multifaceted nature of the severity of depression and contributes to the robustness of predictive models. Further, the inclusion of mitochondrial biomarkers has led to a highly accurate assessment of the severity of depression, indicating the importance of mitochondrial oxidative stress in the progression of depressive disorders.

Identification of p66shc, MOTSc, and Humanin as primary biomarkers to predict the severity of depression highlights their biological and clinical relevance [72–75]. Despite the well-established link between elevated oxidative stress and neurological and psychiatric disorders, the inclusion of p66shc, MOTSc, and Humanin as key mitochondrial biomarkers provides new insight into the progression and severity of mental health diseases and strengthens the findings that mitochondrial peptide dynamics play an important role in the pathophysiology of depression, as well as providing an opportunity for a more comprehensive clinical intervention [72].

Machine learning and classification have previously been applied to the classification of depression severity. In a 5-class classification model where the severity of depression was

based on the PHQ-9 questionnaire and using social behavior in two different data sets (one from the questionnaire and the other from Twitter social networks), a comparative analysis of different machine learning algorithms showed that the XGBoost classifier gave a precision of 83.87% for the questionnaire-generated data set, while logistic regression resulted in the highest precision of 86.45% for the Twitter tweet data set [76]. Another investigation included length of stay at home, duration of sleep, and vigorous activity among 85 features to predict depression with a Random Forest and a set of SVM and KNN and resulted in an equivalent precision of 90% [77]. Similarly, a study using the Depression, Anxiety, and Stress Scale Questionnaire (DASS-21) on 348 participants and five depression categories applied Naive Bayes and achieved a precision of 85.5% for the detection of depression [78]. Furthermore, a comprehensive investigation was carried out that involved six different machine learning classifiers using various sociodemographic and psychosocial information from 604 participants to detect depression. AdaBoost classifier with the SelectKBest feature selection technique outperformed all other classifiers with an accuracy of 92.56% [79]. These collective findings, along with the current study, underscore the growing consensus on the efficacy of machine learning in discerning intricate patterns and predicting depression outcomes.

It should be noted that this study represents a novel approach to the use of specific mitochondrial-associated oxidative stress biomarkers within the framework of machine learning to predict the severity of depression. The inclusion of Humanin, MOTS-c, and p66Shc in the analysis represents a novel avenue that extends the application of machine learning techniques in mental health research.

Although ACE has previously been associated with DM, its inclusion in our study aims to explore its potential relevance to the severity of depression. The rationale lies in the intricate interplay between the renin-angiotensin system (RAS), of which ACE is a key component, and various physiological processes, including those involved in mental health. Emerging evidence suggests that RAS, beyond its traditional role in cardiovascular regulation, may have a modulatory effect on neurotransmission and neuroendocrine pathways [3,80]. Dysregulation of RAS has also been implicated in neuroinflammation, oxidative stress, and altered cerebral blood flow, all factors associated with depression. Furthermore, ACE inhibitors, commonly used to regulate blood pressure by targeting RAS, have shown neuroprotective effects and potential benefits in neurodegenerative disorders [3,80]. Although the specific link between ACE and depression is still an area of active investigation, the potential influence of ACE on neurobiological processes makes it a relevant candidate for exploration in the context of predicting the severity of depression [81,82]. However, in the current study, ACE genotype appears to be of relatively low importance in predicting the severity of depression. Several reasons can explain this observation. The ACE genotype may have a stronger association with hypertension, diabetes, and cardiovascular disease, which is not reflected when these diseases are included as comorbidities in the models [83,84]. Mitochondrial peptides and clinical factors played a more significant role in predicting disease severity compared to the ACE genotype in the current study. This suggests that the ACE genotype may have a limited impact on disease progression compared to other factors. Furthermore, the ACE genotype may interact with other genetic or environmental factors to influence the severity of the disease. By including ACE in this study, the objective was to contribute to an evolving understanding of the multifaceted connections between cardiovascular health, neurobiology, and mental health outcomes. More research is needed to elucidate the specific mechanisms through which ACE can impact depression severity and establish its potential as a biomarker to predict mental health disorders.

The difference between this study and previous research can be attributed to the use of new biomarkers, the specific focus of the study, and the limitations of previous research. Additionally, variability in study designs, sample populations, and methodologies in different investigations introduces a layer of complexity when attempting to compare results. Participant diversity, diagnostic criteria, and data analysis techniques can contribute to discrepancies between studies. Furthermore, the retrospective nature of some studies, including potential biases and confounding factors, may impact the generalizability of the findings.

Future studies with larger and more diverse sample sizes, as well as longitudinal designs, will provide a more comprehensive understanding of the intricate relationships between biomarkers and the severity of depression. The current study was limited by its relatively small sample size and cross-sectional nature. Larger studies with more participants from diverse backgrounds would help validate the findings and determine if the associations hold true across different populations. Longitudinal studies that follow participants over time would shed light on how changes in biomarker levels relate to changes in depression severity. This would help clarify the directionality of the relationships and whether biomarkers can predict future depression risk or track responses to treatment.

An expanded focus on biological markers associated with oxidative stress, inflammation, and neurobiological processes, including peripheral biomarkers such as cytokines, growth factors, hormones, and protein markers, could further enhance understanding of the underlying pathophysiology of depression. Incorporating behavioral data from smartphones and wearable technologies, including activity levels, sleep patterns, dietary habits, and social interactions, also holds promise for capturing the real-time dynamics of depressive symptoms and their interaction with biological processes. Furthermore, delving into genetic, molecular and neuroimaging factors, as well as exploring the interactions between the hypothalamic-pituitary-adrenal (HPA) axis and mitochondrial metabolism, offers the potential to uncover novel predictors and improve the overall robustness of predictive models for the severity of depression. While this study provides promising preliminary evidence, further research is needed to fully elucidate the complex interplay between mitochondrial peptides, clinical factors, and depression.

5 Conclusion

This study highlights the potential of machine learning (ML) techniques in conjunction with biomarkers to diagnose and understand the progression of depression. The findings highlight the considerable impact of Humanin, MOTS-c, and p66shc on predicting the severity of depression. In the context of clinical applications, the study has implications for practitioners who want to use ML and biomarkers to assess and manage the severity of depression.

Supporting information

S1 Text: Supplementary material. This file contains Tables S1 to S9.

- S1 to S3 are Tables 9–11 showing the results for all classes for Model 1.
- S4 to S6 are Tables 12–14 showing the results for all classes for Model 2.
- S7 to S9 are Tables 15–17 showing the results for all classes for Model 3.

(TEX)

Author contributions

Conceptualization: Maher Maalouf, Herbert F. Jelinek.

Data curation: Herbert F. Jelinek.

Formal analysis: Toheeb Salahudeen, Ibrahim (Abe) M. Elfadel.

Funding acquisition: Herbert F. Jelinek.

Investigation: Toheeb Salahudeen.

Methodology: Toheeb Salahudeen, Ibrahim (Abe) M. Elfadel.

Resources: Maher Maalouf.

Software: Toheeb Salahudeen.

Supervision: Maher Maalouf, Ibrahim (Abe) M. Elfadel, Herbert F. Jelinek.

Validation: Toheeb Salahudeen.

Visualization: Toheeb Salahudeen.

Writing – original draft: Toheeb Salahudeen.

Writing – review & editing: Maher Maalouf, Ibrahim (Abe) M. Elfadel, Herbert F. Jelinek.

References

1. Harrison PJ, Taquet M. Neuropsychiatric disorders following SARS-CoV-2 infection. Brain. 2023;146(6):2241–7. https://doi.org/10.1093/brain/awad008 PMID: 36729559

2. Skogstrand K, Borbye-Lorenzen N, Bækvad-Hansen M, Lausten-Thomsen U. Biomarkers to predict, prevent and find the appropriate treatments of disorders in childhood. Frontiers in Pediatrics. 2022;10:1093198. https://doi.org/10.3389/fped.2022.1093198 PMID: 36507127

3. Li Z, Ruan M, Chen J, Fang Y. Major depressive disorder: advances in neuroscience research and translational applications. Neurosci Bull. 2021;37(6):863–80. https://doi.org/10.1007/s12264-021-00638-3 PMID: 33582959

4. World Health Organization. Suicide worldwide in 2019: global health estimates. Geneva: World Health Organization; 2021

5. Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. Br J Gen Pract. 2008;58(546):32–6. https://doi.org/10.3399/bjgp08X263794 PMID: 18186994

6. Tolentino JC, Schmidt SL. DSM-5 criteria and depression severity: implications for clinical practice. Front Psychiatry. 2018;9:450. https://doi.org/10.3389/fpsyt.2018.00450 PMID: 30333763

7. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4(6):561–71. https://doi.org/10.1001/archpsyc.1961.01710120031004 PMID: 13688369

8. Fries GR, Saldana VA, Finnstein J, Rein T. Molecular pathways of major depressive disorder converge on the synapse. Mol Psychiatry. 2023;28(1):284–97. https://doi.org/10.1038/s41380-022-01806-1 PMID: 36203007

9. Bharadwaj P, Bates K, Porter T, Teimouri E, Perry G, Steele J, et al. Latrepirdine: molecular mechanisms underlying potential therapeutic roles in Alzheimer's and other neurodegenerative diseases. Transl Psychiatry. 2013;3(12):e332. https://doi.org/10.1038/tp.2013.97 PMID: 24301650

10. Targum SD, Schappi J, Koutsouris A, Bhaumik R, Rapaport MH, Rasgon N, et al. A novel peripheral biomarker for depression and antidepressant response. Mol Psychiatry. 2022;27(3):1640–6. https://doi.org/10.1038/s41380-021-01399-1 PMID: 34969978

11. Wallace DC. A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. Annu Rev Genet. 2005;39:359-407. https://doi.org/10.1146/annurev.genet.39.110304.095751 PMID: 16285865

12. Merry TL, Chan A, Woodhead JS, Reynolds JC, Kumagai H, Kim SJ, et al. Mitochondrial-derived peptides in energy metabolism. Am J Physiol Endocrinol Metab. 2020;319(4):E659–66. https://doi.org/10.1152/ajpendo.00249.2020 PMID: 32776825

13. Dabravolski SA, Nikiforov NG, Starodubova AV, Popkova TV, Orekhov AN. The role of mitochondria-derived peptides in cardiovascular diseases and their potential as therapeutic targets. Int J Mol Sci. 2021;22(16):8770. https://doi.org/10.3390/ijms22168770 PMID: 34445477

14. Allen J, Romay-Tallon R, Brymer KJ, Caruncho HJ, Kalynchuk LE. Mitochondria and mood: mitochondrial dysfunction as a key player in the manifestation of depression. Front Neurosci. 2018;12:386. https://doi.org/10.3389/fnins.2018.00386 PMID: 29928190

15. Ji N, Lei M, Chen Y, Tian S, Li C, Zhang B. How Oxidative Stress Induces Depression? ASN Neuro. 2023;15:17590914231181037. https://doi.org/10.1177/17590914231181037 PMID: 37331994

16. Black CN, Bot M, Scheffer PG, Cuijpers P, Penninx BW. Is depression associated with increased oxidative stress? A systematic review and meta-analysis. Psychoneuroendocrinology. 2015;51:164–75. https://doi.org/10.1016/j.psyneuen.2014.09.025 PMID: 25462890

17. Manji H, Kato T, Di Prospero NA, Ness S, Beal MF, Krams M, et al. Impaired mitochondrial function in psychiatric disorders. Nat Rev Neurosci. 2012;13(5):293–307. https://doi.org/10.1038/nrn3229 PMID: 22510887

18. Liu T, Zhong S, Liao X, Chen J, He T, Lai S, et al. A meta-analysis of oxidative stress markers in depression. PLoS One. 2015;10(10):e0138904. https://doi.org/10.1371/journal.pone.0138904 PMID: 26445247

19. Detka J, Głombik K. Insights into a possible role of glucagon-like peptide-1 receptor agonists in the treatment of depression. Pharmacol Rep. 2021;73(4):1020–32. https://doi.org/10.1007/s43440-021-00274-8 PMID: 34003475

20. Wan W, Zhang L, Lin Y, Rao X, Wang X, Hua F, et al. Mitochondria-derived peptide MOTS-c: effects and mechanisms related to stress, metabolism and aging. J Transl Med. 2023;21(1):36. https://doi.org/10.1186/s12967-023-03885-2 PMID: 36670507

21. Cobb LJ, Lee C, Xiao J, Yen K, Wong RG, Nakamura HK, et al. Naturally occurring mitochondrial-derived peptides are age-dependent regulators of apoptosis, insulin sensitivity, and inflammatory markers. Aging (Albany NY). 2016;8(4):796–809. https://doi.org/10.18632/aging.100943 PMID: 27070352

22. Goetzl EJ, Wolkowitz OM, Srihari VH, Reus VI, Goetzl L, Kapogiannis D, et al. Abnormal levels of mitochondrial proteins in plasma neuronal extracellular vesicles in major depressive disorder. Mol Psychiatry. 2021;26(12):7355–62. https://doi.org/10.1038/s41380-021-01268-x PMID: 34471251

23. Voigt A, Jelinek HF. Humanin: a mitochondrial signaling peptide as a biomarker for impaired fasting glucose-related oxidative stress. Physiol Rep. 2016;4(9):e12796. https://doi.org/10.14814/phy2.12796 PMID: 27173674

24. Kim SJ, Guerrero N, Wassef G, Xiao J, Mehta HH, Cohen P, et al. The mitochondrial-derived peptide humanin activates the ERK1/2, AKT, and STAT3 signaling pathways and has age-dependent signaling differences in the hippocampus. Oncotarget. 2016;7(30):46899–912. https://doi.org/10.18632/oncotarget.10380 PMID: 27384491

25. Zuccato CF, Asad AS, Nicola Candia AJ, Gottardo MF, Moreno Ayala MA, Theas MS, et al. Mitochondrial-derived peptide humanin as therapeutic target in cancer and degenerative diseases. Expert Opin Ther Targets. 2019;23(2):117–26. https://doi.org/10.1080/14728222.2019.1559300 PMID: 30582721

26. Steinberg GR, Carling D. AMP-activated protein kinase: the current landscape for drug development. Nat Rev Drug Discov. 2019;18(7):527–51. https://doi.org/10.1038/s41573-019-0019-2 PMID: 30867601

27. Yi X, Hu G, Yang Y, Li J, Jin J, Chang B. Role of MOTS-c in the regulation of bone metabolism. Front Physiol. 2023;14:1149120. https://doi.org/10.3389/fphys.2023.1149120 PMID: 37200834

28. Wan W, Zhang L, Lin Y, Rao X, Wang X, Hua F, et al. Mitochondria-derived peptide MOTS-c: Effects and mechanisms related to stress, metabolism and aging. J Transl Med. 2023;21(1):36. https://doi.org/10.1186/s12967-023-03885-2 PMID: 36670507

29. Kim Y, Vadodaria KC, Lenkei Z, Kato T, Gage FH, Marchetto MC, et al. Mitochondria, metabolism, and redox mechanisms in psychiatric disorders. Antioxid Redox Signal. 2019;31(4):275–317. https://doi.org/10.1089/ars.2018.7606 PMID: 30585734

30. Lone A, Harris RA, Singh O, Betts DH, Cumming RC. p66Shc activation promotes increased oxidative phosphorylation and renders CNS cells more vulnerable to amyloid beta toxicity. Sci Rep. 2018;8(1):17081. https://doi.org/10.1038/s41598-018-35114-y PMID: 30459314

31. Galimov E. The role of p66shc in oxidative stress and apoptosis. Acta Naturae. 2010;2(4):44–51. PMID: 22649663

32. Villalpando-Rodriguez GE, Gibson SB. Reactive oxygen species (ROS) regulates different types of cell death by acting as a rheostat. Oxid Med Cell Longev. 2021;2021:9912436. https://doi.org/10.1155/2021/9912436 PMID: 34426760

33. Arfin S, Jha NK, Jha SK, Kesari KK, Ruokolainen J, Roychoudhury S, et al. Oxidative stress in cancer cell metabolism. Antioxidants (Basel). 2021;10(5):642. https://doi.org/10.3390/antiox10050642 PMID: 33922139

34. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. Curr Biol. 2014;24(10):R453–62. https://doi.org/10.1016/j.cub.2014.03.034 PMID: 24845678

35. Di Lisa F, Kaludercic N, Carpi A, Menabò R, Giorgio M. Mitochondrial pathways for ROS formation and myocardial injury: the relevance of p66Shc and monoamine oxidase. Basic Res Cardiol. 2009;104(2):131–9. https://doi.org/10.1007/s00395-009-0008-4 PMID: 19242637

36. Cosentino F, Francia P, Camici GG, Pelicci PG, Volpe M, Lüscher TF. Final common molecular pathways of aging and cardiovascular disease: role of the p66Shc protein. Arterioscler Thromb Vasc Biol. 2008;28(4):622–8. https://doi.org/10.1161/ATVBAHA.107.156059 PMID: 18162611

37. Giorgio M, Migliaccio E, Orsini F, Paolucci D, Moroni M, Contursi C, et al. Electron transfer between cytochrome c and p66Shc generates reactive oxygen species that trigger mitochondrial apoptosis. Cell. 2005;122(2):221–33. https://doi.org/10.1016/j.cell.2005.05.011 PMID: 16051147

38. Berry A, Amrein I, Nötzli S, Lazic SE, Bellisario V, Giorgio M, et al. Sustained hippocampal neurogenesis in females is amplified in P66Shc-/- mice: An animal model of healthy aging. Hippocampus. 2012;22(12):2249–59. https://doi.org/10.1002/hipo.22042 PMID: 22707391

39. Xu Y, Wang C, J Klabnik J, M O'Donnell J. Novel therapeutic targets in depression and anxiety: antioxidants as a candidate treatment. Curr Neuropharmacol. 2014;12(2):108–19. https://doi.org/10.2174/1570159X11666131120231448 PMID: 24669206

40. Ancelin ML, Carrière I, Scali J, Ritchie K, Chaudieu I, Ryan J. Angiotensin-converting enzyme gene variants are associated with both cortisol secretion and late-life depression. Transl Psychiatry. 2013;3(11):e322. https://doi.org/10.1038/tp.2013.95 PMID: 24193727

41. Zettergren A, Kern S, Gustafson D, Gudmundsson P, Sigström R, Östling S, et al. The ACE gene is associated with late-life major depression and age at dementia onset in a population-based cohort. Am J Geriatr Psychiatry. 2017;25(2):170–7. https://doi.org/10.1016/j.jagp.2016.06.009 PMID: 27639288

42. Ng E, Lu Y, Hambly B, Jelinek HF, Yu B, Matthews S, et al. Angiotensin-converting enzyme gene DD genotype is associated with increased systolic blood pressure in an Australian rural type 2 diabetic cohort. Hypertens Res. 2013;36(4):381–2. https://doi.org/10.1038/hr.2012.195 PMID: 23235711

43. Witzel II, Jelinek HF, Khalaf K, Lee S, Khandoker AH, Alsafar H. Identifying common genetic risk factors of diabetic neuropathies. Front Endocrinol (Lausanne). 2015;6:88. https://doi.org/10.3389/fendo.2015.00088 PMID: 26074879

44. Esther Jr CR, Marino EM, Bernstein KE. The role of angiotensin-converting enzyme in blood pressure control, renal function, and male fertility. Trends Endocrinol Metab. 1997;8(5):181–6. https://doi.org/10.1016/s1043-2760(97)00039-8 PMID: 18406804

45. Alexopoulos GS. Mechanisms and treatment of late-life depression. Transl Psychiatry. 2019;9(1):188. https://doi.org/10.1038/s41398-019-0514-6 PMID: 31383842

46. Akbari M, Eghtedarian R, Hussen BM, Eslami S, Taheri M, Ghafouri-Fard S. Angiotensin I converting enzyme gene polymorphisms and risk of psychiatric disorders. BMC Psychiatry. 2022;22(1):351. https://doi.org/10.1186/s12888-022-04007-w PMID: 35606706

47. Wynn A. Unwarranted emotional distress in men with ischaemic heart disease (IHD). Med J Aust. 1967;2(19):847–51. https://doi.org/10.5694/j.1326-5377.1967.tb74314.x PMID: 6081071

48. Bot M, Pouwer F, Zuidersma M, Van Melle JP, De Jonge P. Association of coexisting diabetes and depression with mortality after myocardial infarction. Diabetes Care. 2012;35(3):503–9. https://doi.org/10.2337/dc11-1749 PMID: 22301118

49. Jia Z, Li S. Risk of cardiovascular disease mortality in relation to depression and 14 common risk factors. Int J Gen Med. 2021;14:441–9. https://doi.org/10.2147/IJGM.S292140 PMID: 33603451

50. Musselman DL, Evans DL, Nemeroff CB. The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. Arch Gen Psychiatry. 1998;55(7):580–92. https://doi.org/10.1001/archpsyc.55.7.580 PMID: 9672048

51. Sara JDS, Toya T, Ahmad A, Clark MM, Gilliam WP, Lerman LO, et al. Mental stress and its effects on vascular health. In: Mayo clinic proceedings. vol. 97. Elsevier; 2022. pp. 951–90.

52. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J. 2017;16:1057. https://doi.org/10.17179/excli2017-480 PMCID: PMC5579396

53. De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. Dialogues Clin Neurosci. 2018;20(1):31–40. https://doi.org/10.31887/DCNS.2018.20.1/mdehert PMID: 29946209

54. Jung I, Rhee EJ, Lee WY. Increased risk of cardiovascular disease and mortality in patients with diabetes and coexisting depression: a nationwide population-based cohort study (Diabetes Metab J 2021; 45: 379-89). Diabetes Metab J. 2021;45(5):793–4. https://doi.org/10.4093/dmj.2021.0222 PMID: 34610724

55. Shao M, Lin X, Jiang D, Tian H, Xu Y, Wang L, et al. Depression and cardiovascular disease: shared molecular mechanisms and clinical implications. Psychiatry Res. 2020;285:112802. https://doi.org/10.1016/j.psychres.2020.112802 PMID: 32036152

56. Almas A, Moller J, Iqbal R, Lundin A, Forsell Y. Does depressed persons with non-cardiovascular morbidity have a higher risk of CVD? A population-based cohort study in Sweden. BMC Cardiovasc Disord. 2019;19(1):260. https://doi.org/10.1186/s12872-019-1252-7 PMID: 31752710

57. Kursa MB, Jankowski A, Rudnicki WR. Boruta—a system for feature selection. Fundamenta Informaticae. 2010;101(4):271–85. http://dx.doi.org/10.3233/FI-2010-288

58. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57. http://dx.doi.org/10.1613/jair.953

59. Qi Y. Random forest for bioinformatics. In: Zhang C, Ma Y, editors. Ensemble machine learning: methods and applications. Springer; 2012, pp. 307–23.

60. Rueping S. SVM classifier estimation from group probabilities. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10); 2010. pp. 911–8.

61. Wright RE. Logistic regression.; 1995. Wright RE (1995). Logistic regression. In: Grimm LG, Yarnold PR, editors. Reading and understanding multivariate statistics. American Psychological Association; 1995, pp. 217–-44.

62. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorob. 2013;7:21. https://doi.org/10.3389/fnbot.2013.00021

63. Géron A. Hands-on machine learning with scikit-learn, Keras & TensorFlow, vol. 2007. O'Reilly; 2019.

64. Taunk K, De S, Verma S, Swetapadma A. A brief review of nearest neighbor algorithm for learning and classification. In: 2019 international conference on intelligent computing and control systems (ICCS). IEEE; 2019. pp. 1255–60.

65. Heo S, Lee JH. Fault detection and classification using artificial neural networks. IFAC-PapersOnLine. 2018;51(18):470–5. https://doi.org/10.1016/j.ifacol.2018.09.380

66. Karrach L, Pivarčiová E. Using Different Types of Artificial Neural Networks to Classify 2D Matrix Codes and Their Rotations—A Comparative Study. J Imaging. 2023;9(9):188. https://doi.org/10.3390/jimaging9090188 PMID: 37754952

67. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on data science and advanced analytics (DSAA). IEEE; 2018. pp. 80–89.

68. Breiman L. Random forests. Machine Learn. 2001;45:5–32. http://dx.doi.org/10.1023/A:1010950718922

69. Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. pp. 1135–44.

70. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, (NIPS'17). Red Hook, NY, USA: Curran Associates Inc., pp. 4768-–77.

71. Chen J, Song L, Wainwright M, Jordan M. Learning to explain: An information-theoretic perspective on model interpretation. In: International conference on machine learning. PMLR; 2018. pp. 883–892. http://dx.doi.org/10.48550/arXiv.1802.07814

72. Pant T, Uche N, Juric M, Bosnjak ZJ. Clinical Relevance of lncRNA and Mitochondrial Targeted Antioxidants as Therapeutic Options in Regulating Oxidative Stress and Mitochondrial Function in Vascular Complications of Diabetes. Antioxidants (Basel). 2023;12(4):898. http://dx.doi.org/10.3390/antiox12040898 PMID: 37107272

73. Büttiker P, Weissenberger S, Esch T, Anders M, Raboch J, Ptacek R, et al. Dysfunctional mitochondrial processes contribute to energy perturbations in the brain and neuropsychiatric symptoms. Front Pharmacol. 2023;13:1095923. http://dx.doi.org/10.3389/fphar.2022.1095923 PMID: 36686690

74. Cojocaru KA, Luchian I, Goriuc A, Antoci LM, Ciobanu CG, Popescu R, et al. Mitochondrial dysfunction, oxidative stress, and therapeutic strategies in diabetes, obesity, and cardiovascular disease. Antioxidants (Basel). 2023;12(3):658. http://dx.doi.org/10.3390/antiox12030658 PMID: 36978905

75. Negah SS, Forouzanfar F. Oxidative Stress is a New Avenue for Treatment of Neuropsychiatric Disorders: Hype of Hope? Curr Mol Med. 2024;24(12):1494–505. http://dx.doi.org/10.2174/1566524023666230904150907 PMID: 37670697

76. Kilaskar M, Saindane N, Ansari N, Doshi D, Kulkarni M. Machine learning algorithms for analysis and prediction of depression. SN Comput Sci. 2022;3(2):103. https://doi.org/10.1007/s42979-021-00967-0

77. Narayanrao PV, Kumari PLS. Analysis of machine learning algorithms for predicting depression. In: 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE; 2020. pp. 1–4.

78. Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. Procedia Comput Sci. 2020;167:1258–67.

79. Zulfiker MS, Kabir N, Biswas AA, Nazneen T, Uddin MS. An in-depth analysis of machine learning approaches to predict depression. Curr Res Behav Sci. 2021;2:100044. https://doi.org/10.1016/j.crbeha.2021.100044

80. Bondy B, Baghai TC, Zill P, Schule C, Eser D, Deiml T, et al. Genetic variants in the angiotensin I-converting-enzyme (ACE) and angiotensin II receptor (AT1) gene and clinical outcome in depression. Prog Neuropsychopharmacol Biol Psychiatry. 2005;29(6):1094–9. https://doi.org/10.1016/j.pnpbp.2005.03.015 PMID: 15949885

81. Kim Y, Lee H, Park A. Patterns of adverse childhood experiences and depressive symptoms: self-esteem as a mediating mechanism. Soc Psychiatry Psychiatr Epidemiol. 2022;57(2):331–41. https://doi.org/10.1007/s00127-021-02129-2 PMID: 34191037

82. Meyer T, Rothe I, Staab J, Deter HC, Fangauf SV, Hamacher S, et al. Length polymorphisms in the angiotensin I-converting enzyme gene and the serotonin-transporter-linked polymorphic region constitute a risk haplotype for depression in patients with coronary artery disease. Biochem Genet. 2020;58(4):631–48. https://doi.org/10.1007/s10528-020-09967-w PMID: 32367400

83. O'Donnell CJ, Lindpaintner K, Larson MG, Rao VS, Ordovas JM, Schaefer EJ, et al. Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. Circulation. 1998;97(18):1766–72. https://doi.org/10.1161/01.cir.97.18.1766 PMID: 9603529

84. Yang M, Qiu CC, Xu Q, Xiang HD. Association of angiotensin converting enzyme gene I/D polymorphism with type 2 diabetes mellitus. Biomed Environ Sci. 2006;19(4):323–7. PMID: 17044652