RESEARCH ARTICLE

Major depressive disorder prediction using data science

Vincent Peter C. Magboo*, Ma. Sheila A. Magboo

*Corresponding author's email address: vcmagboo@up.edu.ph

Department of Physical Sciences and Mathematics, College of Arts and Sciences, University of the Philippines Manila, Manila, Philippines

ABSTRACT

Background: Major depressive disorder is a mood disorder that has affected many people worldwide. It is characterized by persistently low or depressed mood, anhedonia or decreased interest in pleasurable activities, feelings of guilt or worthlessness, lack of energy, poor concentration, appetite changes, psychomotor retardation or agitation, sleep disturbances, or suicidal thoughts.

Objective: The objective of the study was to predict the presence of major depressive disorder using a variety of machine learning classification algorithms (logistic regression, Naive Bayes, support vector machine, random forest, adaptive boosting, and extreme gradient boosting) on a publicly available depression dataset.

Methodology: After data pre-processing, several experiments were performed to assess the recursive feature elimination with cross validation as a feature selection method and synthetic minority over-sampling technique to address dataset imbalance. Several machine learning algorithms were applied on an anonymized publicly available depression dataset. Feature importance of the top performing models were also generated. All simulation experiments were implemented via Python 3.8 and its machine learning libraries (Scikit-learn, Keras, Tensorflow, Pandas, Matplotlib, Seaborn, NumPy).

Results: The top performing model was obtained by logistic regression with excellent performance metrics (91% accuracy, 93% sensitivity, 85% specificity, 93% recall, 93% F1-score, and 0.78 Matthews correlation coefficient). Feature importance scores of the most relevant attribute were also generated for the best model.

Conclusion: The findings suggest the utility of data science techniques powered by machine learning models to make a diagnosis of major depressive disorders with acceptable results. The potential deployment of these machine learning models in clinical practice can further enhance the diagnostic acumen of health professionals. Using data analytics and machine learning, data scientists can have a better understanding of mental health illness contributing to prompt and improved diagnosis thereby leading to the institution of early intervention and medical treatments ensuring the best quality of care for our patients.

Keywords: *Major Depressive Disorder (MDD) prediction, machine learning, recursive feature elimination with cross validation (RFE-CV), synthetic minority over-sampling technique (SMOTE), feature importance*

Introduction

Major depressive disorder (MDD) is a mood disorder that has affected many people worldwide. It is characterized by persistently low or depressed mood, anhedonia or decreased interest in pleasurable activities, feelings of guilt or worthlessness, lack of energy, poor concentration, appetite changes, psychomotor retardation or agitation, sleep disturbances, or suicidal thoughts [1]. Like any other mental health problem, the overall functioning at home, school, and workplace is significantly sidelined by depression and often leads to suicide. One study reported a close association of MDD with chronic diseases such as diabetes and heart disease [2]. Patients with depression are also likely to drop out of school, be unemployed, prone to substance abuse, or worse, suffer incarceration. MDD is usually diagnosed by psychologists or psychiatrists using a mental screening tool. Unfortunately, many patients do not seek professional help and, thus, are unable to receive early intervention efforts. It is imperative that recognition of depression be made promptly such that early professional intervention can be made to mitigate the devastating effects on the overall well-being [3,4].

It is in this area of early identification where data science methods particularly using machine learning (ML) algorithms can be utilized, thus enhancing the whole diagnostic process [5,6]. The use of these ML tools developed by data scientists is seen as a supplement to mental health professionals as these are cost-effective and can shorten the time between diagnosis and initiation of early intervention efforts and medical therapy.

In the literature, there were research studies applying ML classification algorithms to diagnose depression on varied types of depression datasets. Sabab Zulfiker et al. applied six ML classifiers with three feature selection methods and synthetic minority oversampling technique (SMOTE) for a diagnosis of depression. The results indicated AdaBoost with SelectKBest feature selection technique to be the best performing model with a 92.56% accuracy rate [2]. In another study by Grzenda *et al.*, several ML classifiers – Support Vector Machine (SVM), random forest (RF), and logistic regression (LR) – were compared to predict treatment outcomes in late-life depression using sociodemographic characteristics, baseline clinical self-reports, cognitive tests, and structural magnetic resonance imaging features as attributes [7]. RF generated an area under receiver operating characteristic curve (AUROC) of 0.83 while SVM and LR recorded an AUROC of 0.80 and 0.79, respectively. A 2022 study [8] compared regression-based models (LR, lasso, ridge) and RF in depression forecasting among home-based elderly Chinese. The results showed these models having good diagnostic performance in differentiating depression versus no depression. In the study by Nemesure et al. [9], an ensemble of ML models [SVM, k-Nearest-Neighbors (kNN), LR, RF, and Extreme Gradient Boosting (XGBoost)] was applied to predict depression and Generalized Anxiety Disorder (GAD) with moderate predictive performance (AUROC of 0.73 for GAD and 0.60 for depression, respectively). Zhou et al. [10] developed ML-based models (DT, LR with least absolute shrinkage and selection operator (LASSO), RF, and gradient-boosting tree) for predicting depression during COVID-19 outbreak in China. Results showed ML models to be suitable for making predictions about mentally at-risk healthcare workers in a public health emergency setting which could support stakeholders' decision-making on possible psychological interventions and proper mental health management.

The objectives of this research were as follows: (1) to benchmark the performance of ML algorithms using AutoGluon in predicting depression in a publicly available depression dataset; (2) to apply hyperparameter tuning on a variety of ML classification algorithms, namely, Logistic Regression (LR), Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), Adaptive Boosting (AdaBoost), and Extreme Gradient Boosting (XGBoost) to determine if performance can be improved; (3) to assess the effect of recursive feature elimination with cross validation (RFECV) as a feature selection technique on the classification performance of each hyperparameter-tuned ML algorithms; (4) to assess the effects of synthetic minority oversampling technique (SMOTE) to address data imbalance on the classification performance of each hyperparameter-tuned ML algorithms, and; (5) to determine the feature importance of various independent attributes relevant to depression prediction in the top-performing hyperparameter-tuned ML models.

Methodology

An anonymized publicly available depression dataset was obtained from a machine learning (ML) repository. To benchmark the performance of ML algorithms on this dataset, AutoGluon was used. AutoGluon is an open source framework developed by Amazon Web Service to quickly prototype potential ML solutions on a given raw data with a few lines of code. The goal is to improve the performance of the ML algorithms through hyperparameter tuning. As a prerequisite to this goal, pre-processing steps such as data cleaning and dataset normalization were applied. Feature selection techniques to select important predictors and address data imbalance were also applied to the dataset. After pre-processing, the dataset was split into training and testing. Various ML algorithms were applied to the training dataset after which hyperparameter tuning was performed for each algorithm to improve individual model performance. These tuned models were then applied to the testing dataset after which an assessment of performance was done using the following metrics: (1) accuracy, (2) precision, (3) recall/sensitivity, (4) specificity, (5) F1-scores, and (6) Matthews correlation coefficient. Feature importance scores of the attributes of the best model were also generated. These are highlighted in the ML pipeline shown in Figure 1.

Dataset Description

A publicly available anonymized depression dataset from an ML data repository is utilized in this study [11]. This dataset contains 604 records, majority of which are males numbering 455 (75%). In terms of age, a significant number of samples belong to the 21-25 age group at 395 (65.4%) followed by the 26-30 age bracket at 146 (24.2%), and the young age group (16-20 years old) at 48 (7.9%) while the rest are over 31 years old at 15 (2.48%). In the dataset, most were unmarried at 514 (85%) while in terms of profession, most were classified as students at 441 (73%). There were 30 predictor variables and 1 target variable (depressed or

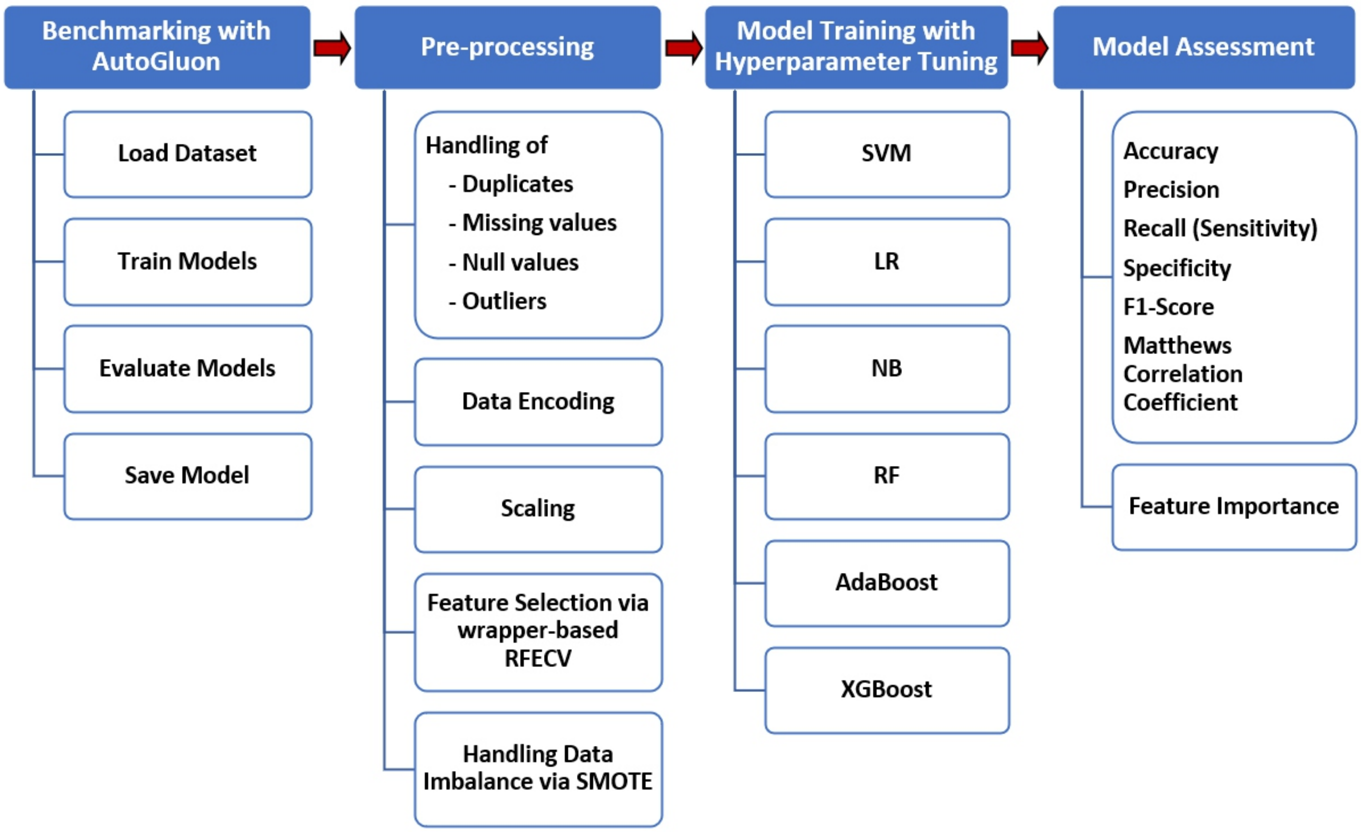

Figure 1. *Machine Learning Pipeline for Major Depressive Disorder Prediction*

Table 1. *List of Variables for Major Depressive Disorder*

Attribute	Attribute	Attribute
Age in years (AGERNG)	Physical Exercise (PHYEX)	Felt Cheated (CHEAT)
Gender	Smoker (SMOKE)	Faced threat (THREAT)
Educational Attainment (EDU)	Alcohol Drinker (DRINK)	Felt Abused (ABUSED)
Profession (PROF)	With Illness (ILLNESS)	Lost someone (LOST)
Marital Status (MARSTS)	Has Insomnia (INSOM)	Has Work/Study Pressure (WRKPRE)
Type of Residence (RESDPL)	Has Eating disorder (EATDIS)	Inferiority Complex (INFER)
Lives with Family or not (LIVWTH)	Average sleep hours (AVGSLP)	Suicidal thoughts (SUICIDE)
Satisfied with Environment or not (ENVSAT)	Taking Prescribed Med (PREMED)	In Conflict with Family or Friends (CONFLICT)
Satisfied with current position or achievements or not (POSSAT)	Ave hours in social network (TSSN)	Depressed
Financial stress (FINSTR)	Feels anxiety (ANXI)	
Had Debt (DEBT)	Feels deprived (DEPRI)	

not depressed) based on Burns Depression Checklist. The variables are listed in Table 1.

Benchmarking Using AutoGluon

AutoGluon, an open source toolkit, is an easy-to-use machine learning toolkit which facilitates a variety of AutoML

(Automated Machine Learning) tasks. It provides automatic data processing, model selection, and model architecture search [12]. After installing AutoGluon library, TabularDataset and TabularPredictor classes were imported. TabularDataset was used to load the dataset while TabularPredictor was used to train pre-defined ML models to make the predictions on the test dataset, and evaluate the performance of the pre-defined ML models.

Pre-processing Steps

The goal is to improve further the performance of the ML algorithms through hyperparameter tuning. As a prerequisite to this goal, pre-processing methods were applied to the dataset. On close inspection, there were no missing values in the dataset but there were 10 duplicate records which were promptly removed. No outliers were likewise detected in the dataset. The dataset has mild data imbalance with 391 (65.82%) with depression and 203 (34.18%) without depression. Data encoding for the attributes and feature scaling with normalization using the StandardScaler function of scikit-learn library were performed. All categorical predictors were dummified resulting in an increase in the number of columns. To assess the feature selection, a wrapper method RFE-CV was used. SMOTE was utilized to address the mild data imbalance.

Machine Learning Models

A split of 30% testing involving 179 records and 70% training involving 415 records with a 10-fold cross validation was made. Python 3.8 and its various ML libraries (scikit-learn, keras, tensorflow, pandas, Matplotlib, seaborn, and NumPy) were utilized in the simulation experiments. Four model configurations were built corresponding to the type of feature selection and application of SMOTE. These were : (1) no feature selection, no SMOTE, (2) with RFE-CV, no SMOTE, (3) no feature selection, with SMOTE, and (4) with RFE-CV, with SMOTE. The ML models tested in these four configurations were LR, NB, SVM, RF, AdaBoost, and XGBoost. This resulted in 24 simulations (4 model configurations x 6 ML models). To further improve the performance of each ML model, hyperparameter tuning was also performed resulting in another round of 24 simulations. In totality, 48 simulation experiments were performed. Matthews correlation coefficient (MCC) was used to ascertain the top performing ML model as MCC takes into consideration all the entries in the confusion matrix for each model. MCC appeared to be a more informative and reliable statistical measure that produces a high score only when prediction obtained good results in all four confusion matrix categories (true positives, false negatives, true negatives, and false positives) [13]. As such, MCC is more useful in evaluating binary classifications than accuracy and F1 scores.

Feature Importance

The feature importance scores to determine the most important attributes relevant to MDD prediction for the best ML models were generated. This method is useful in eliminating irrelevant attributes in predicting the presence of MDD.

Results and Discussion

The benchmarking results of ML algorithms to predict depression using AutoGluon are shown in Table 2. There was a wide variation in the accuracy rate of the different pre-defined ML models ranging from 52% to 84%. Likewise, the

Table 2. *Performance Metrics of the Pre-defined ML Models Using AutoGluon*

	Accuracy	Recall	Precision	F1-score	MCC
RandomForestEntr	0.85	0.91	0.90	0.90	0.71
LightGBM	0.84	0.92	0.89	0.90	0.69
WeightedEnsemble_L2	0.84	0.91	0.89	0.90	0.69
LightGBMXT	0.84	0.91	0.89	0.90	0.68
ExtraTreesEntr	0.82	0.91	0.87	0.89	0.65
XGBoost	0.83	0.88	0.89	0.88	0.65
CatBoost	0.78	0.89	0.84	0.87	0.57
LightGBMLarge	0.68	0.96	0.77	0.85	0.47
KNeighborsDist	0.52	0.76	0.68	0.72	0.05
KNeighborsUnif	0.52	0.76	0.68	0.72	0.05

MCC ranged from a poor 0.05 to 0.71. The best pre-defined ML model generated by AutoGluon was RandomForestEntr. As earlier described, AutoGluon can quickly prototype ML solutions with a few lines of code and generating acceptable performance results.

The performance metrics of the 6 ML models highlighting the effects of the feature selection method are seen in Table 3. When there was no feature selection method applied to the dataset, the best-performing model was obtained by LR with a 91% accuracy. On the other hand, when RFE-CV feature selection was applied to the dataset, XGBoost generated the highest accuracy at 85%. Generally, there was a decrease in the accuracy rates ranging from 3%-7% for most of the models while the NB model did not show any significant change. Hence, LR with no feature selection, appeared to be the top performing model with the highest MCC score at 0.78, 91% accuracy, 93% sensitivity, 85% specificity, 93% precision, and 93% F1-score. It is apparent that all attributes in this dataset appeared to be important in MDD prediction and that no attributes needed to be eliminated. It is also interesting to note the improvement of the metrics when compared to the earlier benchmarking results using AutoGluon.

SMOTE was applied to the dataset to address the mild data imbalance. The comparative performance metrics to assess the effect of SMOTE are shown in Table 4. There was a decrease in the accuracy rate (3%-18%) for most models when SMOTE was applied to the dataset with no feature selection while the rest of the models did not show any significant change in the accuracy rates. Nevertheless, LR obtained the highest MCC and accuracy at 0.74 and 84%, respectively. When SMOTE was applied for the dataset with RFE-CV feature selection, the highest accuracy rates were obtained by RF and XGBoost. Generally, there were no significant changes in the accuracy rates for most models except for NB with a decrement of 4% and AdaBoost with an increment of 3% in the accuracy rates. Overall, LR obtained the highest MCC at 0.74, 88% accuracy, 91% sensitivity, 83% specificity, 92% precision, and 91% F1-score for this experiment assessing the effects of SMOTE.

The best performing ML models in predicting MDD based on MCC were (a) LR (no feature selection, no SMOTE), (b) XGBoost (no feature selection, no SMOTE) and (c) LR (no

Table 3. *Performance Metrics for Predicting MDD – Assessment of Feature Selection*

	No Feature Selection						RFE-CV					
	LR	RF	NB	SVM	Ada Boost	XG Boost	LR	RF	NB	SVM	Ada Boost	XG Boost
Accuracy	0.91	0.87	0.84	0.87	0.87	0.89	0.84	0.84	0.84	0.83	0.80	0.85
Recall /Sensitivity	0.93	0.92	0.82	0.93	0.91	0.93	0.87	0.87	0.86	0.82	0.84	0.90
Specificity	0.85	0.78	0.92	0.75	0.78	0.81	0.80	0.80	0.80	0.83	0.71	0.76
Precision	0.93	0.89	0.95	0.88	0.89	0.91	0.90	0.90	0.90	0.91	0.86	0.89
F1-score	0.93	0.91	0.88	0.91	0.90	0.92	0.88	0.88	0.88	0.86	0.85	0.89
MCC	0.78	0.71	0.60	0.70	0.69	0.76	0.65	0.65	0.64	0.63	0.55	0.67

Table 4. *Comparative Performance Metrics of ML Models with SMOTE in MDD Prediction*

	No Feature Selection Plus SMOTE						RFE-CV Plus SMOTE					
	LR	RF	NB	SVM	Ada Boost	XG Boost	LR	RF	NB	SVM	Ada Boost	XG Boost
Accuracy	0.88	0.87	0.86	0.83	0.87	0.85	0.84	0.84	0.80	0.84	0.83	0.84
Recall /Sensitivity	0.91	0.91	0.86	0.83	0.74	0.89	0.86	0.87	0.78	0.86	0.82	0.87
Specificity	0.83	0.80	0.86	0.83	0.81	0.78	0.80	0.80	0.85	0.80	0.83	0.80
Precision	0.92	0.90	0.93	0.91	0.89	0.89	0.90	0.90	0.91	0.90	0.91	0.90
F1-score	0.91	0.90	0.89	0.87	0.81	0.89	0.88	0.88	0.84	0.88	0.86	0.88
MCC	0.74	0.71	0.70	0.64	0.53	0.67	0.64	0.65	0.60	0.64	0.63	0.65

feature selection, with SMOTE). The comparative performance of these best models is illustrated in Figure 2. The performance of the top three models seems to be similar or comparable to each other across all metrics suggesting that for this dataset, feature selection method may or may not be done. Likewise, SMOTE may or may not be performed as the data imbalance is only mild. Nonetheless, the overall top-performing model was obtained by LR without any feature selection method and with no SMOTE application as it yielded the highest values in all metrics. This model has an excellent recall (sensitivity) at 93% which indicates superior ability to predict persons with depression. It also has good specificity at 85% suggesting the ability to predict persons without depression. Likewise, the model has an excellent precision (or positive predictive value) at 93%. Precision measures the percentage of persons that the model correctly identified as being depressed out of all persons having depression. Finally, the F1-score of the model is also excellent at 93% as it represents the harmonic mean of recall and precision. F1-score is a more suitable form of metric than accuracy when there is a huge class imbalance or when the costs of false positives and false negatives are significantly different from each other [13].

Figures 3-5 show the radar plot highlighting the feature importance generated by LR (No Feature Selection, No SMOTE), XGBoost (No Feature Selection, No SMOTE), and LR (No Feature Selection, with SMOTE) as the top performing models in the simulation experiments. The most important attributes relevant to depression prediction by these three top models were ANXI (feels anxiety), DEPRI (feels deprived), POSSAT (satisfaction with current position/achievement), INFER (inferiority complex), and ENVSAT (satisfaction with the environment). These are in consonance with the clinical assessment of depression. These attributes have been shown in several studies as important precursors of stress which directly contribute to the development of depression [14-17].

The use of AutoGluon provides a quick prototype of ML solutions with a few lines of code. This distinct advantage of AutoGluon can be very useful for both beginners and experts of ML. AutoGluon automatically optimizes many procedures in ML including data pre-processing and model-architecture search. The data pre-processing techniques include classifying and formatting the input sampling data while model-architecture search involves automating architecture engineering to find the optimal design [12]. As different ML

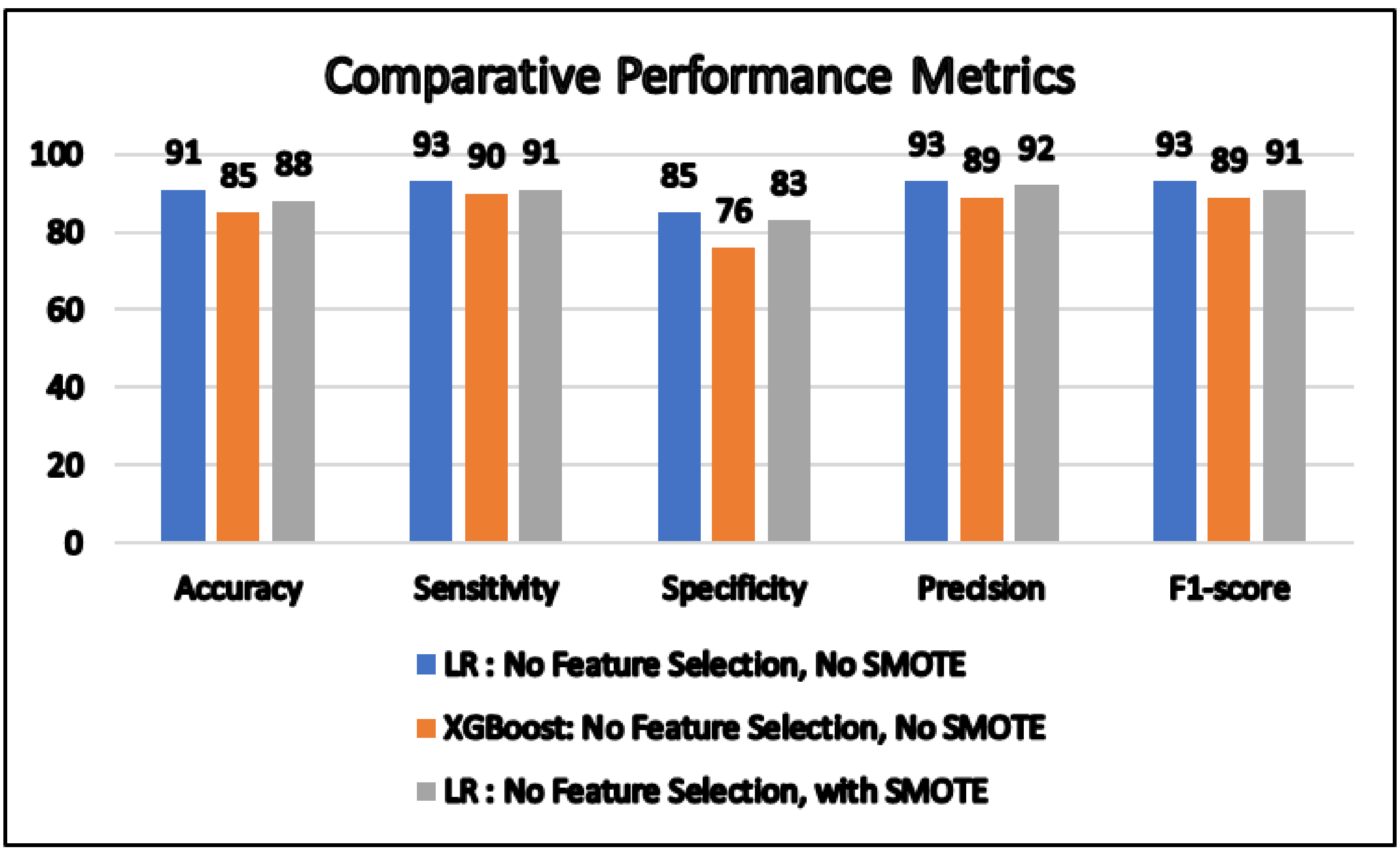

Figure 2. *Performance Metrics of Best Models for MDD Prediction*

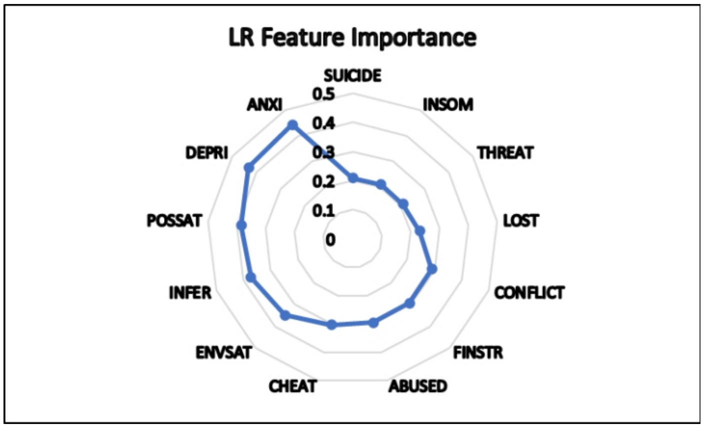

Figure 3. *Feature Importance for MDD Prediction, LR (No Feature Selection, No SMOTE)*

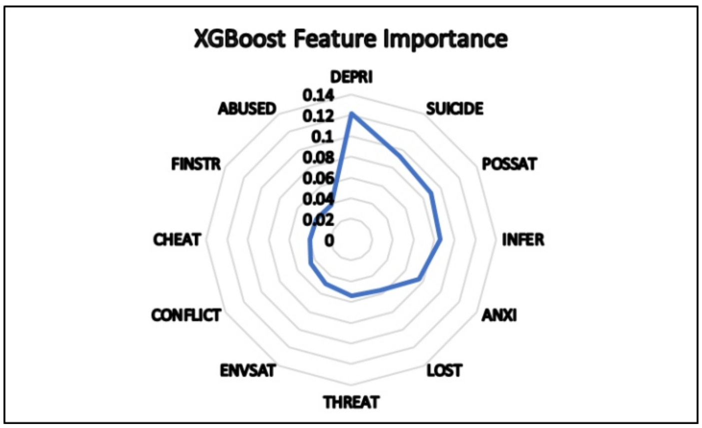

Figure 4. *Feature Importance for MDD Prediction, XGBoost (No Feature Selection, No SMOTE)*

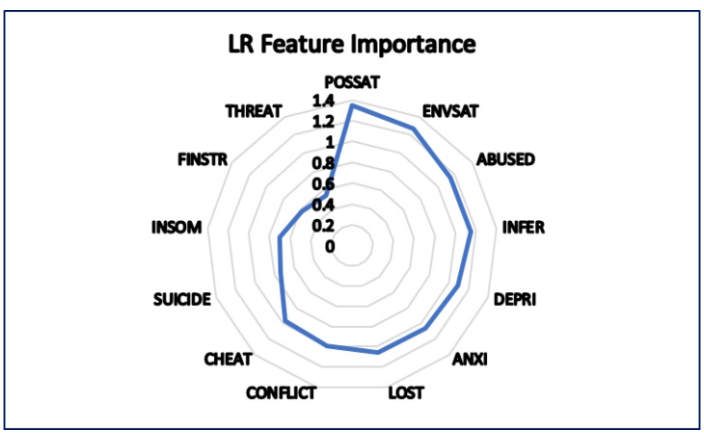

Figure 5. *Feature Importance for MDD Prediction, LR (No Feature Selection, with SMOTE)*

algorithms perform differently for various healthcare datasets, there is also a need to apply other algorithms to the same dataset in a reasonable time with optimal lines of codes [18]. Because of AutoML's high level of automation, non-experts can use ML approaches without the need for considerable computer science skills and experience [19]. The main drawback of the AutoML is the limitation of the available ML models. The results of the benchmarking process in this study suggest the feasibility of AutoML with a fairly good performance in the binary classification task. Nonetheless, the classification performance can still be improved with hyperparameter tuning as shown in the results.

In this study, a wrapper method using RFE-CV was applied to the dataset as a feature selection technique to select only the necessary attributes. Feature selection aims to remove redundant features which can be expressed by other attributes and irrelevant features which do not contribute to the performance of the model in predicting depression [20]. RFE-CV reduces model complexity by removing attributes one at a time until it automatically finds an optimal number of features based on the cross-validation score of the model [21,22]. It is commonly used due to its ease of use. Using the associated feature weights, those attributes with small feature weights close to zero contribute very little to predicting depression. But it should be noted that removing a single attribute would also lead to a change in the feature weights, suggesting that elimination of the features should be done in a stepwise fashion. Additionally, the use of feature selection technique can help boost the training speed of the classifiers [2].

SMOTE is an oversampling method that creates artificial minority data points within the cluster of minority class samples in a balanced way which render it to be an effective method in reducing negative effects of imbalance leading to increased performance [2,23-27]. It works by utilizing a k-Nearest Neighbors (kNN) algorithm to create synthetic data by first selecting a random data from the minority (no depression) class and then kNN from the data are set. That is, synthetic data is created between the random data and the randomly selected kNN. As such, there is not only an increase in the number of datapoints but an increase in its variety. However, SMOTE has its disadvantages such as sample overlapping, noise interference, and blindness of neighbor selection as well as the suitability for clinical datasets [25,28,29].

The use of feature importance makes it possible to detect which attributes in the depression dataset have predictive power. This is done by assigning a score to each attribute based on its ability to improve predictions and allows ranking these features. The feature importance is determined by the increase in the model prediction error after permuting the values of that feature. Likewise, an increase in the model error also increases the importance of that feature for predicting depression. On the other hand, if the accuracy of the model remains the same or slightly decreases, then the feature is deemed unimportant for depression prediction [30-32]. However, the accompanying computational cost of implementing this procedure as well as not being a substitute for statistical inference remains to be its undesirable disadvantage [33].

In comparison with other studies in the literature utilizing ML in predicting MDD, the results of the study are comparable [2,7-9,34,35]. While most of the previous works are focused on predicting depression among a select group of people, this research focused on predicting depression across age groups, occupations, and among people of different age ranges, occupation or profession, marital status, etc. The study is also very simple as only a few demographic and psychosocial information have been incorporated into the ML models. As such, a person suspected of being depressed shall not be intimidated to provide other personal information as required in various depression screening scales. Classification of mental disorders through ML can be via sensors (such as mobile phones and audio signals), text through social media platforms and text messages, structured data (such as those extracted from questionnaires, standard screening scales, medical health record), and multimodal technology interactions such as those obtained from human interactions with everyday technological equipment, robot, and virtual agents [36]. This research made use of structured data extracted from a questionnaire.

The findings suggest the utility of data science techniques particularly using ML models to make a diagnosis of MDD with acceptable results. The clinical relevance of the experiments is even more highlighted with these ML models that can provide faster and with high reliability in assisting mental health professionals in the screening of patients for major depressive disorders. ML classification algorithms are being utilized in mental health for the prediction of mental disorders and can help in the institution of effective treatment outcomes [36]. An early accurate diagnosis leading to prompt intervention efforts is very crucial as it warrants prompt intervention methods to improve the patient's quality of life, diminished risk of developing chronic diseases, improve productivity, and prevention of suicide cases [2,5,7,37]. This research, thus, provided useful insights in the development of automated models that can assist healthcare workers in the assessment of major depressive disorders. Data science is indeed a rapidly evolving field that offers many valuable applications to mental health research, particularly that of major depressive disorders. Using data analytics and ML, data scientists can have a better understanding of mental health illness contributing to prompt and improved diagnosis thereby leading to the institution of early treatments for patients. Techniques in data science powered by artificial intelligence (AI) algorithms provide an option for mental health professionals as it is cost-effective and shortens the time between diagnosis and treatment. The deployment of AI-based applications to diagnose mental health conditions in clinical practice can provide the level of anonymity to the patients enabling them to share their thoughts and feelings without the feeling of being judged. Nevertheless, these apps are highly dependent on the clinical data and the algorithmic methods used in their development by data scientists.

Conclusion

Several ML models were applied to a publicly available anonymized depression dataset. A series of experiments to assess the effects of the use of feature selection methods and the technique to address dataset imbalance were performed. The best performing model was logistic regression (LR) with a 91% accuracy, 93% sensitivity, 85% specificity, 93% recall, 93% F1-score, and 0.78 Matthews correlation coefficient. The most important attributes relevant to the prediction of depression were determined using feature importance which are also in agreement with clinical assessment of depression. The potential deployment of these ML models in clinical practice can further enhance the diagnostic acumen of health professionals. The primary limitation of this research is the use of small datasets due unavailability of large and open-source depression datasets. Because of the sensitivity of depression data, it is difficult to gain access to mental health datasets. The availability of a labeled dataset in mental health for any machine learning study is a major issue as availability is scarce up to the present day.

Future enhancement of this study should focus on the inclusion of other tools for feature importance of attributes as well as techniques in explainable AI for better understanding of these ML models by mental health professionals. In the future, mixed types of datasets combining symptoms with neuroimaging features seen in functional magnetic resonance imaging can also be explored to generate more superior diagnostic accuracy. The findings of this study are promising and have generated useful insights in the development of automated models that are faster and with high reliability which can be of use to mental health professionals in predicting depression. Nonetheless, early intervention efforts and treatment for depression ensure the best quality of care for the patients.

References

1. Bains N, Abdijadid S. (2022) Major Depressive Disorder. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK559078/
2. Sabab Zulfiker M, Kabir N, Biswas AA, Nazneen T, Shorif Uddin M. (2021) An in-depth analysis of machine learning approaches to predict depression. Current Research in Behavioral Sciences, 2 (100044). https://doi.org/10.1016/j.crbeha.2021.100044
3. Uddin MZ, Dysthe KK, Følstad A, Brandtzaeg PB. (2022) Deep learning for prediction of depressive symptoms in a large textual dataset. Neural Computing and Applications 34 : 721–744. https://doi.org/10.1007/s00521-021-06426-4
4. Tao X, Chi O, Delaney PJ, Li L, Huang J. (2021) Detecting depression using an ensemble classifier based on Quality of Life scales. Brain Informatics, 8 (2). https://doi.org/10.1186/s40708-021-00125-5
5. Magboo VPC, Magboo MSA. (2021) Machine Learning Classifiers on Breast Cancer Recurrences. Procedia Computer Science, 192 : 2742-2752. https://doi.org/10.1016/j.procs.2021.09.044
6. Magboo VPC, Magboo MSA. (2022) Prediction Models for COVID-19 in Children. In: Chen, YW., Tanaka, S., Howlett, R.J., Jain, L.C. (eds) Innovation in Medicine and Healthcare. Smart Innovation, Systems and Technologies, 308. Springer, Singapore. https://doi.org/10.1007/978-981-19-3440-7_2
7. Grzenda A, Speier W, Siddarth P, *et al.* (2021) Machine Learning Prediction of Treatment Outcome in Late-Life Depression. Frontiers in Psychiatry, 12. https://www.frontiersin.org/article/10.3389/fpsyt.2021.738494.
8. Lin S, Wu Y, Fang Y. (2022) Comparison of Regression and Machine Learning Methods in Depression Forecasting Among Home-Based Elderly Chinese: A Community Based Study. Frontiers in Psychiatry, 12:764806. https://doi.org/10.3389/fpsyt.2021.764806
9. Nemesure MD, Heinz MV, Huang R, Jacobson N. (2021) Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. Scientific Reports, 11 : 1980. https://doi.org/10.1038/s41598-021-81368-4.
10. Zhou Z, Luo D, Yang BX, Liu Z. (2022) Machine Learning-Based Prediction Models for Depression Symptoms Among Chinese Healthcare Workers During the Early COVID-19 Outbreak in 2020: A Cross-Sectional Study. Frontiers in Psychiatry, 13. https://www.frontiersin.org/articles/10.3389/fpsyt.2022.876995
11. Sabab31/Depression-Repository, https://github.com/Sabab31/Depression-Repository.git.
12. Qi W, Xu C, Xu S. (2021) AutoGluon: A revolutionary framework for landslide hazard analysis. Natural Hazards Research 1(3):103–108. https://doi.org/10.1016/j.nhres.2021.07.002
13. Chicco D, Jurman G. (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21 (6). https://doi.org/10.1186/s12864-019-6413-7
14. Rusli BN, Edimansyah BA, Naing L. (2008) Working conditions, self-perceived stress, anxiety, depression, and quality of life: A structural equation modelling approach. BMC Public Health 8 (48). https://doi.org/10.1186/1471-2458-8-48
15. Keshabyan A, Day MV. (2020) Concerned Whether You'll Make It in Life? Status Anxiety Uniquely Explains Job Satisfaction. Frontiers in Psychology 11. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01523
16. Williams DR, Lawrence JA, David BA, Vu C. (2019) Understanding how discrimination can affect health. Health Serv Res. 2019 (54):1374–1388. https://doi.org/10.1111/1475-6773.13222
17. Seabrook EM, Kern ML, Rickard NS. (2016) Social Networking Sites, Depression, and Anxiety: A Systematic Review. JMIR mental health, 3(4) : e50. https://doi.org/10.2196/mental.5842
18. Poudel S. (2022).A Study of Disease Diagnosis Using Machine Learning. In Proceedings of the 2nd International Electronic Conference on Healthcare, 17 February–3 March 2022, MDPI: Basel, Switzerland, doi:10.3390/IECH2022-12311
19. Naga D, Muster W, Musvasva E, Ecker GF. (2022) Off-targetP ML: an open-source machine learning framework for off-target panel safety assessment of small molecules. Journal of cheminformatics, 14(1), 27. https://doi.org/10.1186/s13321-022-00603-w
20. Demircioğlu A. (2021) Measuring the bias of incorrect application of feature selection when using cross-validation in radiomics. Insights Imaging 12, 172. https://doi.org/10.1186/s13244-021-01115-1
21. Chang W, Ji X, Wang L, *et al.* (2021) A Machine-Learning Method of Predicting Vital Capacity Plateau Value for Ventilatory Pump Failure Based on Data Mining. Healthcare, 9;1306. https://doi.org/10.3390/

healthcare9101306.
22. Li D, Wang Y, Hu W, *et al.* (2021) Application of Machine Learning Classifier to Candida auris Drug Resistance Analysis. Frontiers in Cellular and Infection Microbiology, 11. https://www.frontiersin.org/article/10.3389/fcimb.2021.742062.
23. Sousa S, Paúl C, Teixeira L. (2021) Predictors of Major Depressive Disorder in Older People. International Journal of Environmental Research and Public Health, 18;11894. https://doi.org/10.3390/ijerph182211894.
24. Huang CY, Dai HL. (2021) Learning from class-imbalanced data: review of data driven methods and algorithm driven methods. Data Science in Finance and Economics, 1(1):21-36. https://doi.org/10.3934/DSFE.2021002.
25. Wang S, Dai Y, Shen J, Xuan J. (2021) Research on expansion and classification of imbalanced data based on SMOTE algorithm. Scientific Reports, 11; 24039. https://doi.org/10.1038/s41598-021-03430-5.
26. Risi M, Wang JB, Zou CA, Fu GH. (2021) AWSMOTE: An SVM-Based Adaptive Weighted SMOTE for Class-Imbalance Learning. Scientific Programming, 2021; 9947621. https://doi.org/10.1155/2021/9947621.
27. Magboo VPC, Magboo, MSA. (2021) Imputation Techniques and Recursive Feature Elimination in Machine Learning Applied to Type II Diabetes Classification. In 2021 4th Artificial Intelligence and Cloud Computing Conference (AICCC '21). Association for Computing Machinery, New York, NY, USA, 201-207. https://doi.org/10.1145/3508259.3508288.
28. Jiang Z, Pan T, Zhang C, Yang J. (2021) A New Oversampling Method Based on the Classification Contribution Degree. Symmetry 13,194. https://doi.org/10.3390/sym13020194.
29. Beinecke J, Heider D. (2021) Gaussian noise up-sampling is better suited than SMOTE and ADASYN for clinical decision making. BioData Mining 14, 49. https://doi.org/10.1186/s13040-021-00283-6
30. Ljubobratovic D, Vukovic M, Brkic Bakaric M, Jemric T, Matetic M. (2021) Utilization of Explainable Machine Learning Algorithms for Determination of Important Features in 'Suncrest' Peach Maturity Prediction. Electronics 10,3115. https://doi.org/10.3390/electronics10243115.
31. Mi X, Zou B, Zou F, Hu J. (2021) Permutation-based identification of important biomarkers for complex diseases via machine learning models. Nature Communications 12;3008. https://doi.org/10.1038/s41467-021-22756-2.
32. Inglis A, Parnell A, Hurley C. (2021) Visualizing Variable Importance and Variable Interaction Effects in Machine Learning Models. Journal of Computational and Graphical Statistics, https://doi.org/10.1080/10618600.2021.2007935.
33. Sejong Oh. (2022) Predictive case-based feature importance and interaction. Information Sciences, 593:155-176. https://doi.org/10.1016/j.ins.2022.02.003.
34. Lin S, Wu Y, Fang Y. (2022) Comparison of Regression and Machine Learning Methods in Depression Forecasting Among Home-Based Elderly Chinese: A Community Based Study. Frontiers in psychiatry, 12,764806. https://doi.org/10.3389/fpsyt.2021.764806
35. Jan Z, AI-Ansari N, Mousa O, *et al.* (2021) The Role of Machine Learning in Diagnosing Bipolar Disorder: Scoping Review. Journal of Medical Internet Research, 23(11):e29749. https://doi.org/10.2196/29749
36. Aleem S, Huda NU, Amin R, Khalid S, Alshamrani SS, Alshehri A. (2022) Machine Learning Algorithms for Depression: Diagnosis, Insights, and Research Directions. Electronics 11(7):1111. https://doi.org/10.3390/electronics11071111
37. Magboo VPC, Abu PAR. (2022) Deep Neural Network for Diagnosis of Bone Metastasis. In 2022 The 5th International Conference on Software Engineering and Information Management (ICSIM) (ICSIM 2022). Association for Computing Machinery, New York, NY, USA, 144-151. https://doi.org/10.1145/3520084.3520107