Predictive Analysis of Mental Health Conditions Using AdaBoost Algorithm

5 authors, including:

Elizabeth Oluyemisi Ogunseye
University of Ibadan
1 PUBLICATION   **21** CITATIONS

SEE PROFILE

Adenusi Ajowho Ajowho
University of Agriculture, Abeokuta
10 PUBLICATIONS   **32** CITATIONS

SEE PROFILE

Andrew Chinonso Nwanakwaugwu
University of Salford
15 PUBLICATIONS   **61** CITATIONS

SEE PROFILE

Sunday Adeola Ajagbe
First Technical University
95 PUBLICATIONS   **654** CITATIONS

SEE PROFILE

PARADIGM *Plus*

Predictive Analysis of Mental Health Conditions Using AdaBoost Algorithm

Elizabeth Oluyemisi Ogunseye 🅾 1, Cecilia Ajowho Adenusi 🅾 2, Andrew C. Nwanakwaugwu 🅾 3,✉, Sunday Adeola Ajagbe 🅾 4,✉, and Solomon O. Akinola 🅾 1

1*University of Ibadan, Ibadan, Nigeria*
elizabetholuyemisio@yahoo.com, solom202@yahoo.co.uk
2*Linux Professional Institute, Ibadan, Nigeria*
ceceresearch@gmail.com
3*University of Salford, Greater Manchester, United Kingdom*
a.c.nwanakwaugwu@edu.salford.ac.uk
4*Ladoke Akintola University of Technology, Ogbomoso, Nigeria*
saajagbe@pgschool.lautech.edu.ng

Abstract

The presented research responds to increased mental illness conditions worldwide and the need for efficient mental health care (MHC) through machine learning (ML) implementations. The datasets employed in this investigation belong to a Kaggle repository named "Mental Health Tech Survey." The surveys for the years 2014 and 2016 were downloaded and aggregated. The prediction results for bagging, stacking, LR, KNN, tree class, NN, RF, and AdaBoost yielded 75.93%, 75.93%, 79.89%, 90.42%, 80.69%, 89.95%, 81.22%, and 81.75% respectively. The AdaBoost ML model performed data cleaning and prediction on the datasets, reaching an accuracy of 81.75%, which is good enough for decision-making. The results were further used with other ML models such as Random Forest (RF), K-Nearest Neighbor (KNN), bagging, and a few others, with reported accuracy ranging from 81.22% to 75.93% which is good enough for decision making. Out of all the models used for predicting mental health treatment outcomes, AdaBoost has the highest accuracy.

Keywords: Artificial Intelligence · Predictive Analysis · Mental Health Conditions · Machine Learning · AdaBoost Algorithm

1 Introduction

Mental Health Disorder (MHD) is also known as a mental illness disorder with diverse symptoms ranging from feeling down and sad, excessive worries or fears, sleeping issues, low energy, and significant tiredness. There are diverse kinds of mental health, which include bipolar disorder, anxiety disorder, suicidal ideation, and many more. According to Rehm et al., [1], millions of people are suffering from mental disorders and if care is not taken early enough, it can lead to the progression of all kinds of disease and discomfort, which will then lead to severe treatments. And it is being treated and evaluated using subjective, self-reported, or questionnaire-based methods, which takes time and

can be difficult at times. MH issues have gotten to the systematic research stage. The World Health Organization (WHO) listed many mental conditions that cause depression, psychological problems, and schizophrenia [2]. The United Nations stated that, according to their statistics, almost 20 million deaths are recorded every year from mental health disorders. In response to this, WHO stated that MH is now given an upper priority and therefore leads to more study and research on MH. Universities and most citizens' understanding and warning remain at the traditional level of psychological warning, which is limited and theoretical. In this new age of Information Technology (IT), data mining, and big data, it has started to play a vital role in the MH field, most especially in the aspect of rising alarm of the early warning sigma or symptoms, and most institutions have realized the use of Informa ionization of student management, which only entails the collection stage, storage means, student information management, and combining such with the literature [3, 4].

The vast increase in the usage of the Internet has led to fast growth in data generation, which has led to the utilization of data mining technology for mining values from big data. Such operations commence with classification, which apparently saves a lot of time and economic cost as in data analysis, tuning & fitting, predicting & forecasting processes, followed by regression operations. The methods that are most ideal for classification and regression operations are known as ensemble methods. This combines several models to form one good model via reducing variance and bias in order to enhance the prediction accuracy percentage [5]. Common methods used for these are boosting [6] and bagging, of which boosting produces better results than bagging [7]. A boosting algorithm is generally a learning algorithm in the field of machine learning, which makes use of a "probably approximately correct" model designed by Valiant. The algorithm was further improved by [8] and the improved version was named the AdaBoost algorithm. This algorithm has now been improved to a more vital feature classification algorithm in ML and is widely used for diverse areas like intrusion detection [9], feature extraction [10], image retrieval, object recognition [11], and other applications. The error rate increases during training as weights assigned to samples via AdaBoost rise and likewise fall as new weights fall [12].

ML tools have become useful tools for psychiatry in order to disentangle parameters associated with patient outcomes. The majority of ML research in psychiatry has focused on classification and diagnosis, and as a result, is prone to poor performance due to a lack of understanding of the application of data science and ML techniques, as well as corresponding diagnostic processes or psychiatric disorders. But advancement has been made in some areas via decision tree (DT) models in predicting suicide in U.S. military personnel and adolescents, and DT has been able to show how variables can be used to give insight into the models being used. This research looks into the ML algorithms in analyzing and predicting MH conditions and treatments via the use of the Mental Health Tech Survey dataset gotten from the Kaggle repository.

The remaining sections of this work are organized as follows. Section 2 and 3 present related work and methodology respectively. Section 4 contains the results and discussion, while sections 5 and 6 conclude the study and highlight future scopes.

2   Related Work

In transforming mental healthcare, Graham et al., [13] opined that AI technology holds amazing and promising roles in achieving solutions to many problems. The researchers reviewed 28 papers on AI and mental health that made use of brain imaging data, electronic health records (ERs), novel smart monitoring systems, mood rating scales, and platforms on social media in classifying and predicting schizophrenia, health illness (depression), suicide attempts, and ideation. The outcome of the studies showed high accuracy and provided good examples of the potential of AI in healthcare for mental illness.

According to Madan et al., [14], there are vast real-world databases of health care records available worldwide that can be used to improve personalized medicine and predictive analysis. Of these, the

major source of information collection in mental health has been text-based records. Therefore, the main approach to detecting psychiatric disease is via text mining, according to the researchers. In this paper, 150 text files were manually used for psychiatric symptoms and attributes, and they were divided into two (training and test set). A system was designed to automatically detect the attributes, and the workflow made use of a neural network model. The overall result shows 86% of the F1 score with 91% of the independent test set.

Zhang et al., [15] were concerned about depression, which is one of the mental disorders that keeps increasing in its effect on a person's physical health. The technology approach of AI has currently been implemented to aid mental health practitioners, most especially clinicians and psychiatrists, in decision-making via their patients' medical records. The researchers explored children's mental health and how to use AI to cope with MH issues in children. In order for it to be treated at an early stage, deep learning tools were proposed to be implemented in diagnosing and forecasting children's mental health. A convolutional neural network (CNN) was first used to examine the deep parts of children's behavior data features, followed by insights into the disruption, classification, and then finally, forecasting. The results showed that this current model had a 97.9% sensitivity ratio, 96.7% of the rate specified, 95.6% of ratio recall with 90.1% of precision ratio, and 95.6% of F rate measured with a 9.2% error rate, which is better than the existing system.

One of the complex illnesses associated with varieties of clinical associations and socioeconomic status is mental illness and it's being associated with diverse textual clinical data, coupled with those from social media and interviews [16]. Natural language processing (NLP) is one of the methods used by researchers as one means to aid in early diagnosis and empower the medical healthcare system. Overall, 399 papers from about 10,467 records were reviewed and the review results showed that there was an upward trend detected via NLP of mental illness. Therefore, for further studies, a recommendation on the use of deep learning was made.

Sharma et al., [17] studied diagnosis of mental illnesses based on direct interviews designed with some set of questions, which mostly consume more time and human energy, because men's health is beginning to gain more ground and the need for the use of Machine Learning (ML) is also beginning to arise, with the healthcare system not left behind. The researchers here then made use of ML in developing a model that can enhance the diagnosis of mental depression, and such a model was evaluated. A sample of about 11,081 records of Dutch citizens was used and Extreme Gradient Boosting (XGBoost), which is one of the types of ML algorithm, was applied to each sample, and classification was carried out with an accurate result of 0.90.

This research looks into mental health conditions by analyzing the survey of mental health tech to extract meaningful information, and this was achieved via the means of aggregation. The survey years 2014 and 2016 were merged together in order to extract more meaningful information, which they then used to predict the number of people responding to treatment.

3   Methodology

This section entails the method used in collating the information used for this research work and the flow of how the data collected will be utilized properly. The datasets used were downloaded from the Kaggle repository and were titled *Mental Health in Tech Survey* for the years 2014 and 2016. The 2014 dataset has 1259 instances with 27 parameters. The 2016 dataset has 1433 instances with 62 parameters. Both were used to carry out the analysis and prediction processes.

The structure in Figure 1 denotes the architecture of the mental health predictive analysis in this study hence, it shows how the whole system was implemented, commencing from when the dataset is able to be cleaned of such data in order to avoid missing values that may affect the accuracy of any models implemented, to the real analysis of all the parameters used in the survey, to the covariance matrix, correlation matrices, and probability of mental health occurrence, to how people respond to treatment, and predicting the respondents' treatment with ROC curve visualized results.
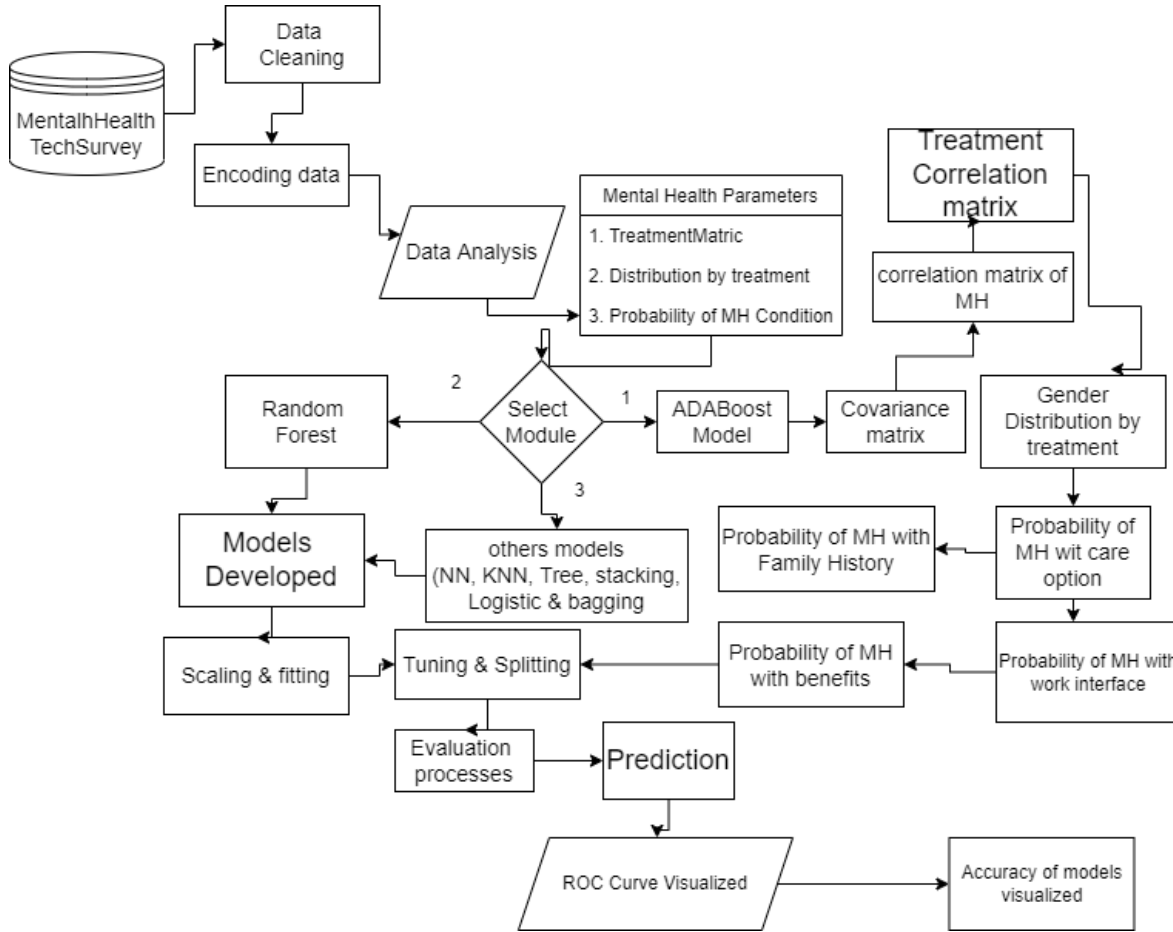
Figure 1: Architecture of Mental health Predictive analysis

3.1   Algorithm of MHD with AdaBoost

1° Input: Training dataset $= \{(MH_1, TR_1), (MH_2, TR_2), \ldots (MH_n, TR_n)\}$; Weak Learn algorithm; Ensemble size T.

2° Initialization: Starting training set using distribution weight (see Equation 1).

$$w_c^1 : w_c^1 = \frac{1}{n}, c = 1, 2, \ldots, n \tag{1}$$

3° Do for $e = 1, 2, \ldots, T$

– Compute a weak classifier based on the classifier algorithm with present weight w^e

– Compute weighted error during training (see Equation 2).

$$\epsilon_e : \epsilon_e = \sum_{c=1}^{n} w_c^e; \ TR_c \neq h_e(MH_c) \tag{2}$$

– Update weight of training instances (see Equation 3).

$$w_c^{e+1} = w_c^e.\exp(-\propto_e.TR_c.h_e(MH_c)\frac{MH}{Z_e}, where Z_e = \sum_{c=1}^{n}.w_c^e.\exp(-\propto_e.TR_c.h_e(MH_c)) \tag{3}$$

. . . .

4° Output: ensemble classifier (see Equation 4).

$$f\left(MH\right) = sign\left(\sum_{e=1}^{T} .\propto_e .h_t\left(MH\right)\right) \tag{4}$$

3.2 Models used for Mental Health Conditions

3.2.1 AdaBoost

Adaptive boosting is a classification algorithm that forms a committee of weak classifiers and enhances ML [18] performance algorithms by combining them together into a strong classifier. The main goal of this classifier is to enhance the weight of uncategorized points and decrease the weight of categorized points. Some of the advantages of AdaBoost are its ease of use and the accuracy of the weak classifier, and some of the disadvantages are the need for quality data and optimization of hyperparameters. Another name for the tree generated is called Decision Stump [19], and the model will keep training till a lower error is achieved [20].

3.2.2 Random Forest (RF) Classifier

This classifier made use of training data to model many tree structures, and each tree suggested a class as an output, with vast numbers of results selected as the overall outcome [21]. The building structure requires specific numbers of trees and, as such can be used as a means of bagging and aggregating bootstrap data. This can be used to reduce variance in the final result.

3.2.3 Neural Network (NN)

An artificial neural network (ANN) refers to a biologically inspired sub-field of artificial intelligence modeled after the brain. Artificial neural networks (ANNs) are computer networks that are based on biological principles. Multilayer perceptrons (MLPs) with back propagation learning algorithms are the focus of this chapter among the numerous forms of ANNs [22]. MLPs are supervised ANNs with three layers: input, hidden, and output. They are the most commonly used ANNs for a wide range of tasks. We go over the structure, method, data pretreatment, overfitting, and sensitivity analysis of MLPs, among other things. advantages: the ability to work with insufficient knowledge and good fault tolerance. Disadvantages: the difficulty of showing the problem to the network and hardware dependence [23].

3.2.4 Decision Tree (DT)

This is a supporting tool that has a tree-shaped structure and models the cost of resources, possible outcomes, consequences, and utilities. It also provides a means of presenting the algorithm via a control statement. DT involves the use of a flowchart, which has nodes that denote each stage, and each branch denotes the attribute result, while the rule of classification is the path from the leaf to root. The tree structure denotes classification and If-then rules, which are quite easy to understand. The computation methods involved are quite short and efficient. It can be used in computing symptoms related to mental health. Mining can be achieved with the use of knowledge and law via algorithms and the provision of student mental health for efficient decision-making for student education [24].

3.2.5 K-Nearest Neighbor (KNN)

One of the simpler machine learning methods is this algorithm. Many in the industry favor it because of how simple it is to use and how quickly calculations can be completed. Using the points that are

most similar to it, the KNN model categorizes data points. It "informed guesses" what an unclassified point should be classed as using test data. Simple recommendation systems, image recognition software, and decision-making models frequently use KNN [25, 26].

4   Result and Discussion

This section analyzed, visualized, and predicted mental health using ML models, and the AdaBoost model ultimately performed better than other ML models in this study. The result was further validated with other models and the Ada result had an accuracy of 81.75% while others were lower than that. Figure 2 is the correlation matrix of MH parameters and Figure 3 is the respondents by countries with tech flag and company size orange.

Most respondents in the survey are in Tech. Most of them are from the US and work in medium or large companies. Even though the description of the research appeared to be global, which would have been a great way to amplify the problem in various communities, the research ended up focusing on the USA and a few European countries. As expected, most of the respondents are male, this happened because the tech industry is occupied by a big majority of men. The proportion of females
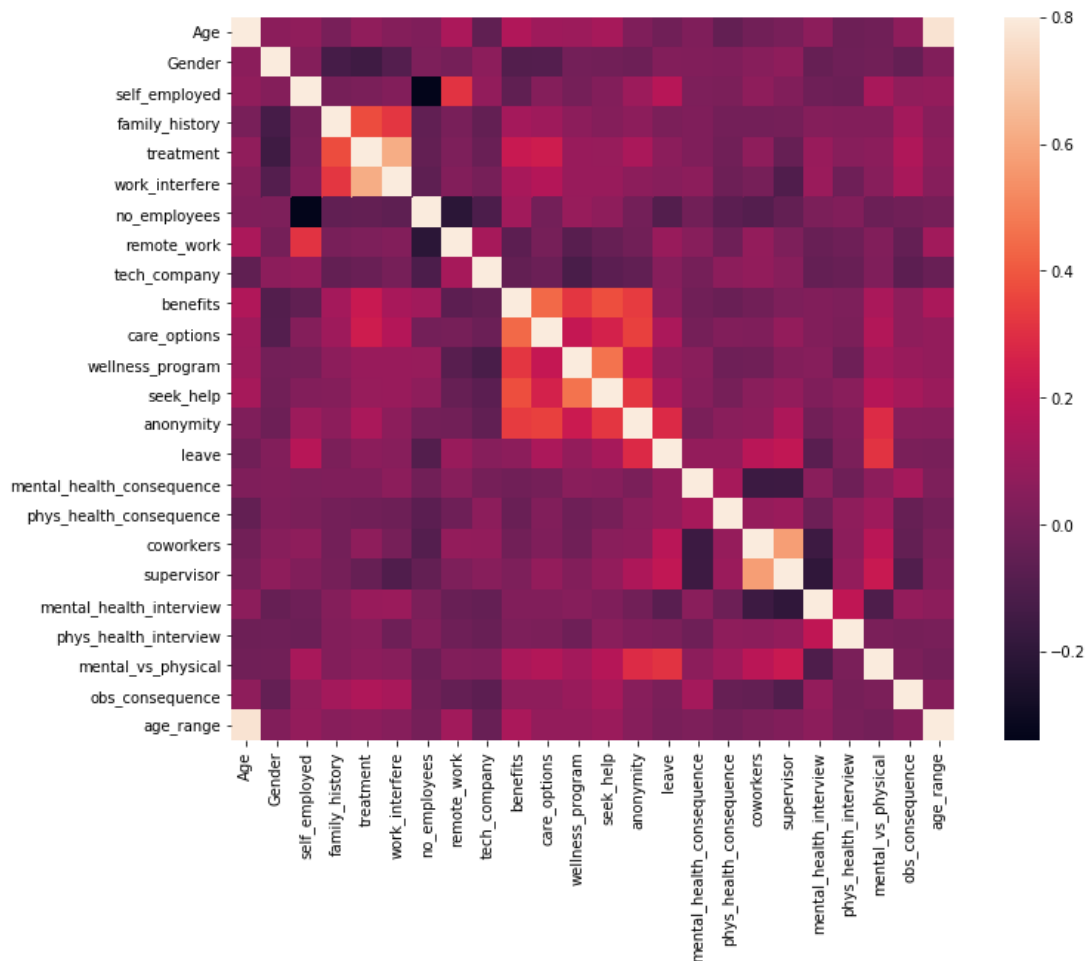

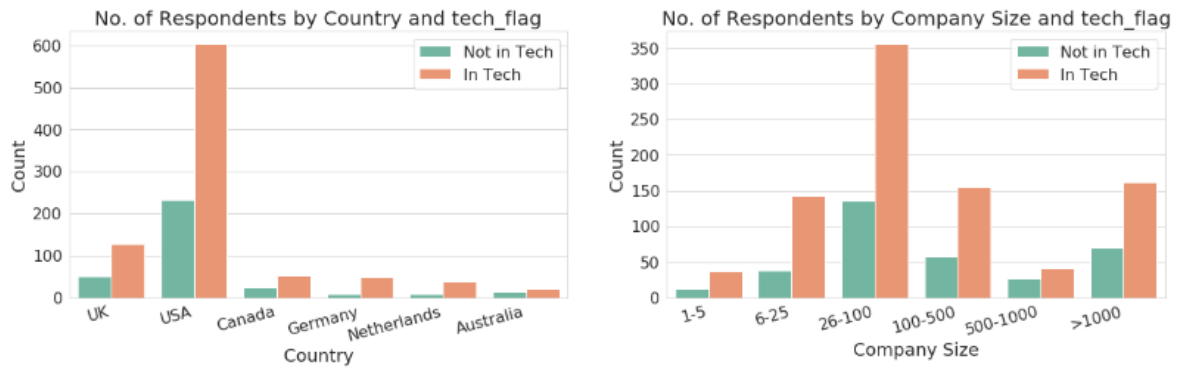Figure 2: Correlation matrix of MH parameters

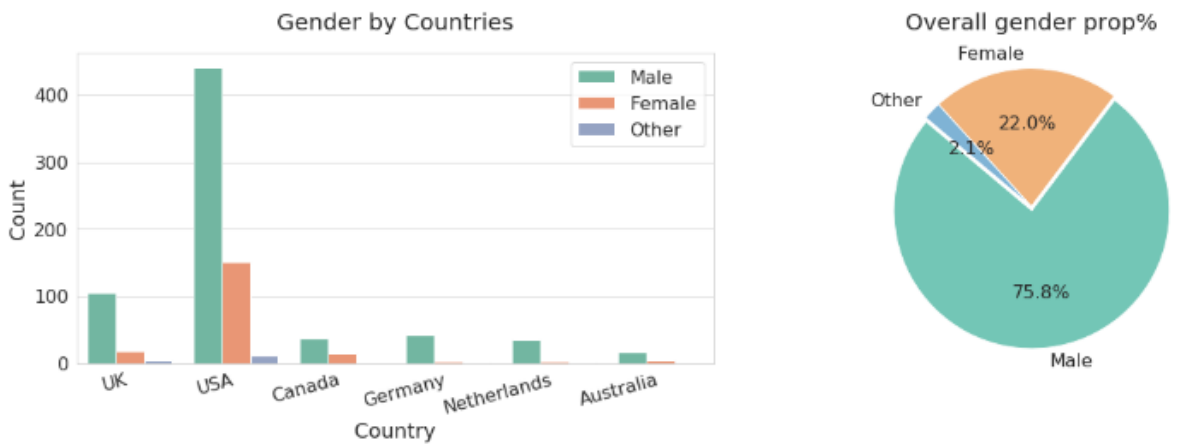Figure 3: Respondents by countries with tech flag and company size orange

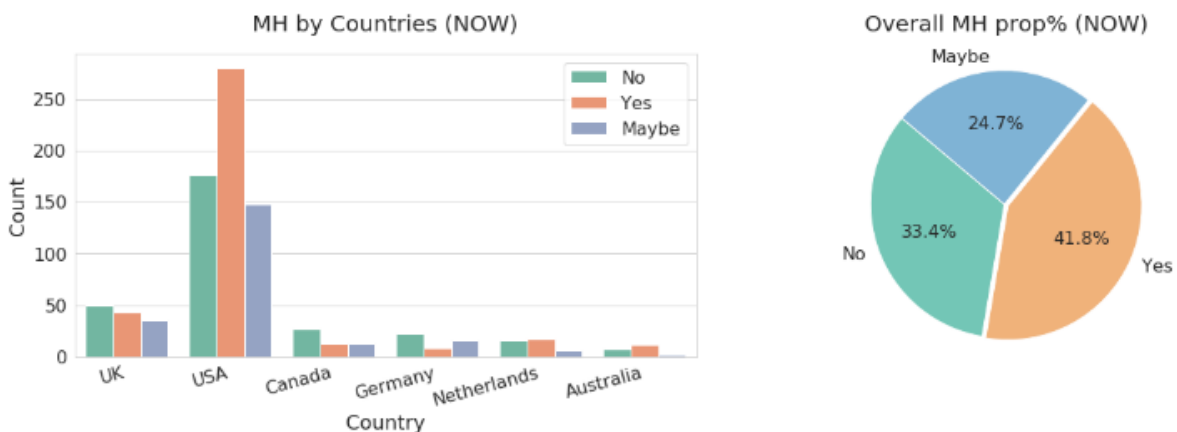Figure 4: Gender by countries with overall proportional (only techs)

Figure 5: Mental Health Disorder in Tech companies with overall percentage proportion (currently)

in tech in this survey in the United States (US) is the biggest one out of all countries. Figure 4 shows the gender of the participants per country, and Figure 5 is the mental health disorder in Tech. companies with overall percentage proportion. The results in Figure 6 show that there are tech workers with
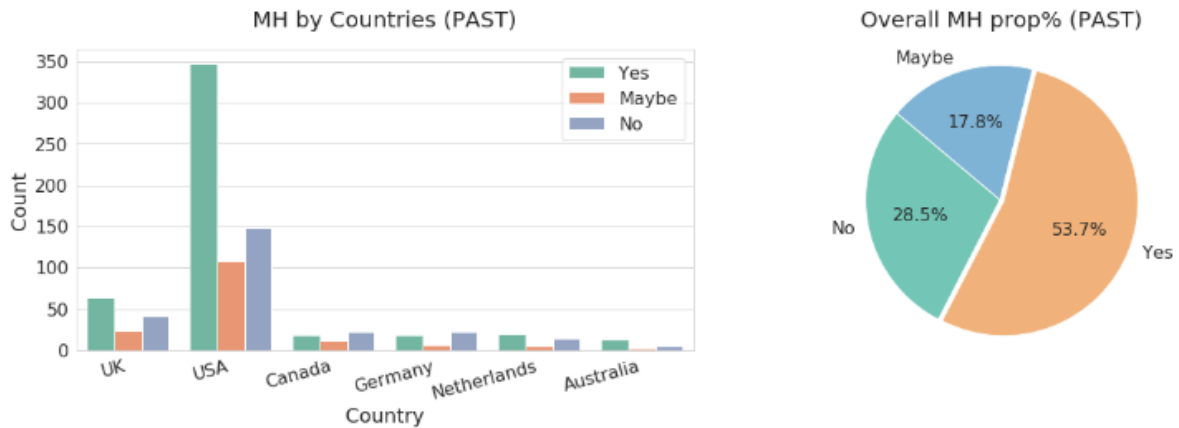
Figure 6: Mental Health Disorder in Tech with past overall percentage proportion

MHD more in the US followed by (the United Kingdom), from the experiment, the result from the tech dataset also reveals that the UK has more MHD than any other countries that participated in the survey with over 41.8% acknowledged having MHD. The response was either 'Yes', 'No', or 'Maybe'. There were tech workers with MHD more in the US followed by (the United Kingdom) UK in the past than any other countries that participated in the survey with over 53.7% acknowledged having MHD. So, the rate dropped from 53.7% to 41.8% from the present analysis.

Figure 7 shows the result of the extraction of important features from the tech dataset. These include age, gender, family history, benefits, care options, anonymity, leave, and work interference. This highlighted the importance of features extracted for further analysis of the mental health disordered,
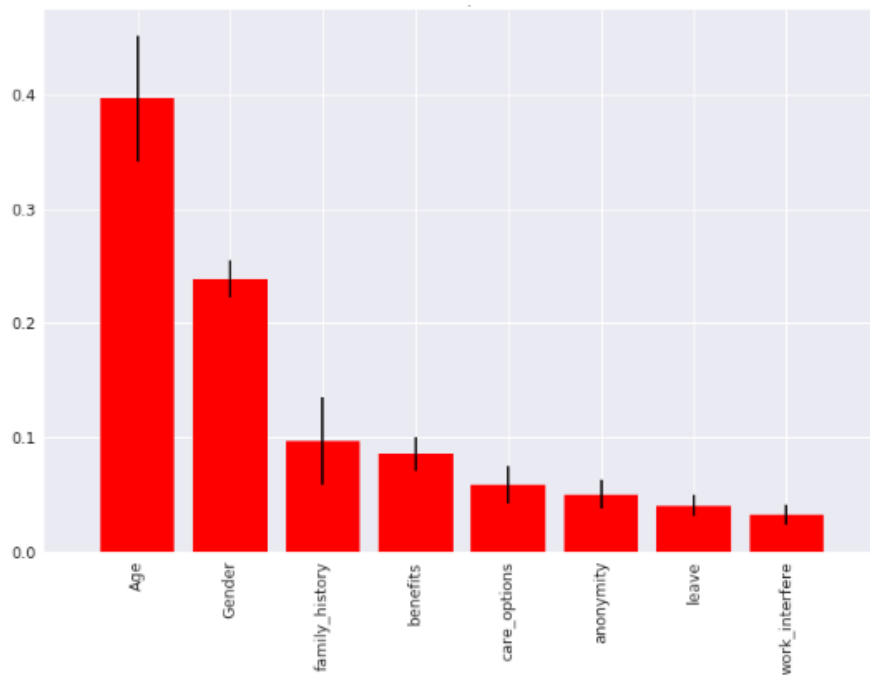

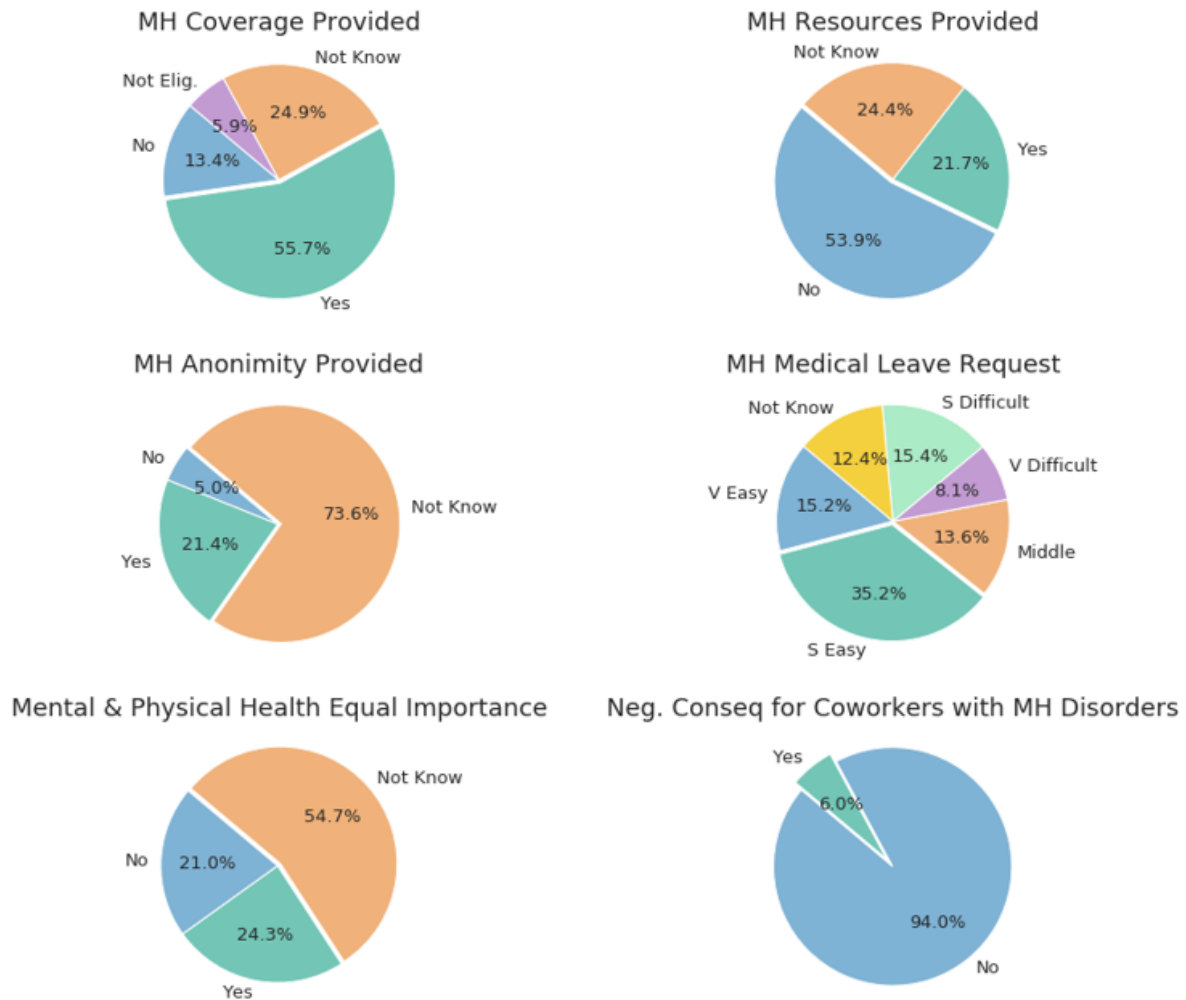Figure 7: Importance features extraction

Figure 8: MH considerations by Companies

which ranges from age down to work interface to arrive at a reasonable conclusion. The result in Figure 8 shows the results of MH considerations by Companies. that about 55.7% companies have MH coverage provided for their employees, 21.7% provided resources, 21.4% provided anonymity, 35.2%
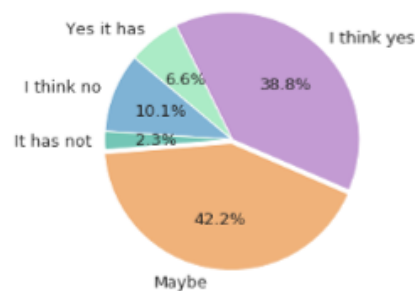

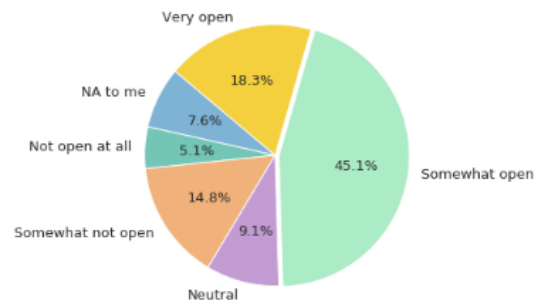Figure 9: Consequences of MHD on career observed

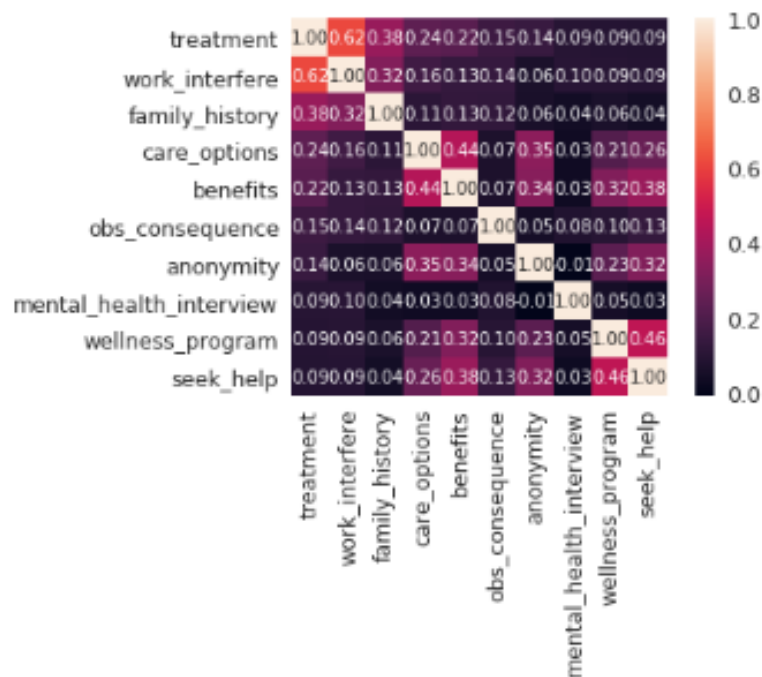Figure 10: Being open about Mental Health with friends and family

Figure 11: Treatment correlation matrix parameters

requested for medical leave, with 24.3 having both mental and physical health of equal importance and only 6% had negative consequences of MHD over their work and 94% were able to manage it effectively. In the main time, there are ripple effects of these considerations on the overall output of these companies.

Figure 9 shows that only 38.8% think MHD may hurt their careers while 42.2% were not so sure. 2.3% has not actually experienced MHD, 10.1% think no, 6.6% yes it has. Therefore, MHD will continue to affect the career of work since consciousness is not there to address it.

When considering openness about MH with friends and family, Figure 10 shows that only 18.3% were very open to their friends and family, 45.1% are sometimes open 9.1% neutral, 14.8% sometimes not open, while 5.1% were not opened at all and 7.6% NA to me. Therefore, MHD will continue to affect the career of workers since open to addressing the issue. Figure 11 shows the relationship between the MHD parameters used for the analysis and prediction while using ML models

Figure 12 visualized the histogram chart showing those with MHD and are undergoing treatment and those that have not appeared, those that are on treatment are more than those that have not.
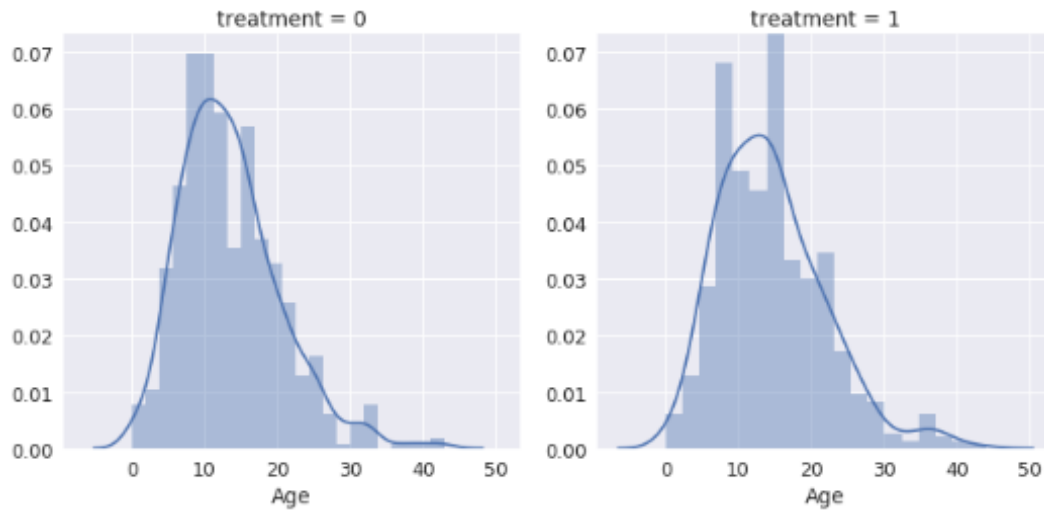
Figure 12: Separate by treatment or not
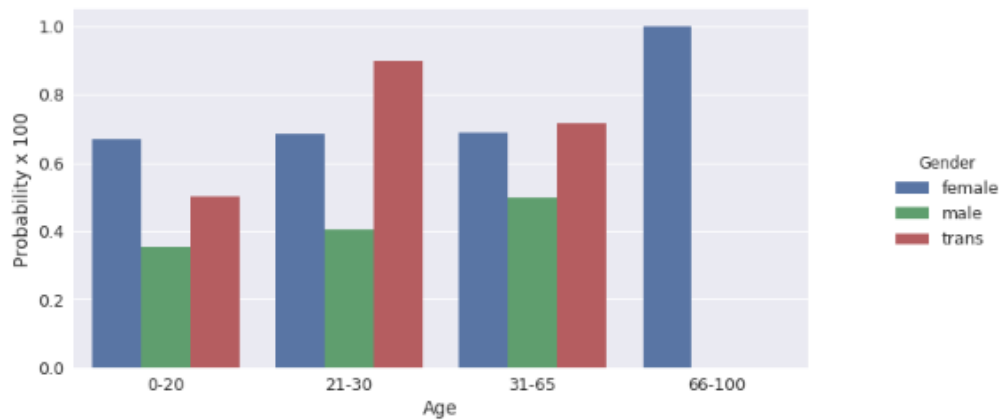
Figure 13: Probability of age-range of mental health condition with care option based on gender

Figure 13 results shows 100% (1.0) probability of female between the age-range of 66-100 having MHD 31-65 age-range shows about 0.72 probability of female having MHD, 21-30 age range with 0.7 probability of MHD and 07. Probability of females having MHD of 0-20 age-range.

The actual result of MHD with treatment is presented as follows, while Figure 14 shows AdaBoost Predicted result of MHD with treatment.

```
########## AdaBoosting ###############
Accuracy: 0.81746031746
Null accuracy:
 0    191
1    187
Name: treatment, dtype: int64
Percentage of ones: 0.4947089947089947
Percentage of zeros: 0.5052910052910053
True: [0 0 0 0 0 0 0 1 1 0 1 1 0 1 1 0 1 0 1 0 0 0 1 1 0 0]
Pred: [1 0 0 0 0 1 0 1 1 1 0 1 1 0 1 1 1 0 0 0 0 1 0 0]
```

Figure 14: AdaBoost Predicted result

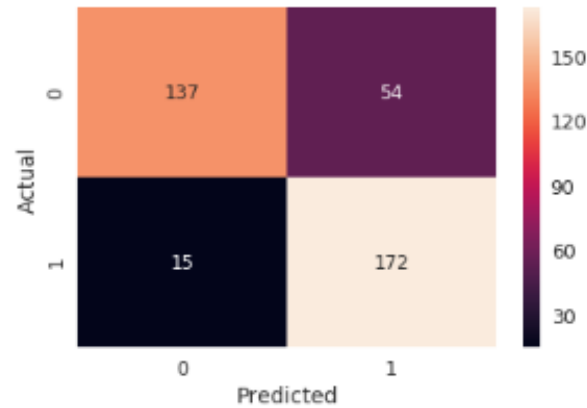The performance of eight ML models used in this study is presented in Table 1, and it shows how AdaBoosting outperformed other ML models in this study.

Table 1: Table Style

S/N	ML models	Accuracy (%)
1	Bagging	75.93
2	Stacking	75.93
3	LR	79.89
4	KNN	80.42
5	Tree class	80.69
6	NN	80.95
7	RF	81.22
8	**AdaBoost**	**81.75**

5   Conclusion

The study presents an ML implementation that impacts how mental disorders are being diagnosed. To this end, the Kaggle repository has been used as a training/testing dataset. We attempted to classify the patients with different supervised algorithms and concluded that AdaBoost was the most successful tool for the task. In other words, this research work has been able to explore the mental health (MH) tech survey dataset, and diverse insights have been extracted and visualized from the dataset. It shows the application of ML will aid in the faster collection, sorting, and understanding of the data, as well as prediction. This visualization includes the gender, age, where the survey took place, the reaction of people and workers to their employees having MH issues, how their work has been affected, treatment or not issues down to prediction using ML models. The accuracy of prediction results was compared among the eight ML models used, the results of the prediction for bagging, stacking, LR, KNN, tree class, NN, RF, and AdaBoost yielded 75.93%, 75.93%, 79.89%, 90.42%, 80.69%, 89.95%, 81.22%, and 81.75% respectively. The results also show that more males responded than females and from all the countries that the survey was done, respondents from the United States responded well. Work and family history were also reflected in the analysis, with the majority of the respondents having a history in their family, and most of them were under one care or the other, es-

pecially the female respondents. The results further show that MH often interfaces with their work performance. And the predictions made with ML show the probability of respondents responding to treatments keeps increasing.

6 Future Work

Kaggle is a bonafide source of data. However, the data set seems very limited when it comes to MH dataset as it contains MH datasets of a few of the developed countries. Hence, there is a need to populate MH dataset for an informed decision-making process. This research can also be further improved upon by making use of the mental health tech survey 2020 dataset if one exists or conducting another survey for 2020 and then applying another model using an ensemble model with Natural Language Processing (NLP) and all the missing values can be further removed for more accurate results. The survey in 2014 and 2016 was great. Replica of such survey is suggested, now that during the COVID-19 restrictions and social distancing is prevailing and compare the results

Authors' Information

– **Elizabeth Oluyemisi Ogunseye** obtained B.Sc in computer science from Ambrose Alli University, Ekpoma in 2007, a master's degree in computer systems in 2009, another master's degree in computer science in 2016, and currently, she is doing a Ph.D. at the Department of Computer Science at the University of Ibadan, Nigeria. Her research interests include Health Informatics, Artificial Intelligence, Data Science, and Data Mining

– **Cecilia Ajowho Adenusi** is a Postgraduate student of the Department of computer science at the Federal University of Agriculture, Abeokuta, Ogun State, Nigeria. She currently works at Linux professional institute, Nigeria master affiliate as the head of training at the institute. She is a member of the Computer Professional Council of Nigeria(CPN), member of Nigeria Computer Society (NCS), member of Academia in Information Technology (AITP), and member of Nigerian Women in Information Technology(NIWIIT).

– **Andrew C. Nwanakwaugwu** is a scholarly Data Science enthusiast with years of IT (Information Technology) related working experience. He is currently working as a PT Operations Advisor at HMRC United Kingdom (UK). He worked as a Data Scientist at Greater Manchester AI Foundry where he used Artificial Neural Networks (ANN) in deep learning and the development of a customized speech synthesis system with an interactive user interface. He also worked at Jamasoft Concept as a System Analyst/Data Scientist, Media/Musical Instrumentalist at RCCG Manchester. He studied Computer Engineering in his first degree and Data Science in his Second degree in Nigeria and the University of Salford, UK respectively.

– **Sunday Adeola Ajagbe** is a Ph.D. candidate in Computer Engineering at the Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria. He obtained an M.Sc. and B.Sc. in Information Technology and Communication Technology respectively at the National Open University of Nigeria (NOUN), and his Postgraduate Diploma in Electronics and Electrical Engineering at LAUTECH, and Higher National Diploma in Electrical and Electronics Engineering, The Polytechnic, Ibadan. His specialization includes Artificial Intelligence (AI), Natural language processing (NLP), Information Security, Quantum computing, Data Science, Internet of Things (IoT), and Smart solution. He is also licensed by The Council Regulating Engineering in Nigeria (COREN) as a professional Electrical Engineer, student member of the Institute of Electrical and Electronics Engineers (IEEE), International Association of Engineers (IAENG), and corporate member of the Nigeria Computer Society (NCS).

– **Solomon O. Akinola** is a Professor of Computer Science in the Department of Computer Science at Nigeria's premier university, the University of Ibadan, Nigeria. He has a Ph.D. in Computer Science at the University of Ibadan, Nigeria. He researches in the areas of Software Engineering and Data Mining.

Authors' Contributions

– **Elizabeth Oluyemisi Ogunseye** participated in conceptualization, review, validation, and Supervision,

– **Cecilia Ajowho Adenusi** participated in conceptualization, writhing of the original draft, methodology, and software development.

– **Andrew C. Nwanakwaugwu** participated in software design, project administration, and software development.

– **Sunday Adeola Ajagbe** participated in the methodology and writing of the original draft.

– **Solomon O. Akinola** participated in editing, review, and supervision.

Competing Interests

The authors declare that they have no competing interests.

Funding

No funding was received for this project.

References

[1] J. Rehm and K. D. Shield, "Global burden of disease and the impact of mental and addictive disorders," *Current psychiatry reports*, vol. 21, no. 2, pp. 1–7, 2019.

[2] G. Singer and M. Golan, "Applying data mining algorithms to encourage mental health disclosure in the workplace," *International Journal of Business Information Systems*, vol. 36, no. 4, pp. 553–571, 2021.

[3] M. Fokkema, J. Edbrooke-Childs, and M. Wolpert, "Generalized linear mixed-model (glmm) trees: A flexible decision-tree method for multilevel and longitudinal data," *Psychotherapy Research*, vol. 31, no. 3, pp. 329–341, 2021.

[4] S. A. Ajagbe, I. R. Idowu, J. Oladosu, and A. Adesina, "Accuracy of machine learning models for mortality rate prediction in a crime dataset," *International Journal of Information Processing and Communication (IJIPC)*, vol. 10, no. 1, pp. 150–160, 2020.

[5] M. Zounemat-Kermani, O. Batelaan, M. Fadaee, and R. Hinkelmann, "Ensemble machine learning paradigms in hydrology: A review," *Journal of Hydrology*, vol. 598, p. 126266, 2021.

[6] P. Karimi and H. Jazayeri-Rad, "Comparing the fault diagnosis performances of single neural networks and two ensemble neural networks based on the boosting methods," *Journal of Automation and Control*, vol. 2, no. 1, pp. 21–32, 2014.

[7] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *Journal of artificial intelligence research*, vol. 11, pp. 169–198, 1999.

[8] M. Payal, A. Kumar, and S. A. Ajagbe, "Support vector machines (svms) based advanced healthcare system using machine learning techniques," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 10, no. 5, pp. 3007–3014, 2022.

[9] Y. Zhou, T. A. Mazzuchi, and S. Sarkani, "M-adaboost-a based ensemble system for network intrusion detection," *Expert Systems with Applications*, vol. 162, p. 113864, 2020.

[10] B. Liu, C. Liu, Y. Xiao, L. Liu, W. Li, and X. Chen, "Adaboost-based transfer learning method for positive and unlabelled learning problem," *Knowledge-Based Systems*, vol. 241, p. 108162, 2022.

[11] G. Hu, C. Yin, M. Wan, Y. Zhang, and Y. Fang, "Recognition of diseased pinus trees in uav images using deep learning and adaboost classifier," *Biosystems Engineering*, vol. 194, pp. 138–151, 2020.

[12] N. Gupta and S. Nigam, "Crude oil price prediction using artificial neural network," *Procedia Computer Science*, vol. 170, pp. 642–647, 2020.

[13] S. Graham, C. Depp, E. E. Lee, C. Nebeker, X. Tu, H.-C. Kim, and D. V. Jeste, "Artificial intelligence for mental health and mental illnesses: an overview," *Current psychiatry reports*, vol. 21, no. 11, pp. 1–18, 2019.

[14] S. Madan, F. J. Zimmer, H. Balabin, S. Schaaf, H. Fröhlich, J. Fluck, I. Neuner, K. Mathiak, M. Hofmann-Apitius, and P. Sarkheil, "Deep learning-based detection of psychiatric attributes from german mental health records," *International Journal of Medical Informatics*, vol. 161, p. 104724, 2022.

[15] X. Zhang, R. Wang, A. Sharma, and G. G. Deverajan, "Artificial intelligence in cognitive psychology—influence of literature based on artificial intelligence on children's mental disorders," *Aggression and Violent Behavior*, p. 101590, 2021.

[16] T. Zhang, A. M. Schoene, S. Ji, and S. Ananiadou, "Natural language processing applied to mental illness detection: a narrative review," *NPJ digital medicine*, vol. 5, no. 1, pp. 1–13, 2022.

[17] A. Sharma and W. J. Verbeke, "Improving diagnosis of depression with xgboost machine learning model and a large biomarkers dutch dataset (n= 11,081)," *Frontiers in big Data*, p. 15, 2020.

[18] C. A. Adenusi, O. R. Vincent, J. Bolarinwa, and O. Oluwakemi, "Predicting the pandemic effect of covid 19 on the nigeria economic, crude oil as a measure parameter using machine learning," in *Decision Sciences for COVID-19*, pp. 79–93, Springer, 2022.

[19] A. S. Hussein, T. Li, C. W. Yohannese, and K. Bashir, "A-smote: A new preprocessing approach for highly imbalanced datasets by improving smote," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1412–1422, 2019.

[20] P. C. Reddy, R. M. S. Chandra, P. Vadiraj, M. A. Reddy, T. Mahesh, and G. S. Madhuri, "Detection of plant leaf-based diseases using machine learning approach," in *2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, pp. 1–4, IEEE, 2021.

[21] M. Sarveshvar, A. Gogoi, A. K. Chaubey, S. Rohit, and T. Mahesh, "Performance of different machine learning techniques for the prediction of heart diseases," in *2021 International Conference on Forensics, Analytics, Big Data, Security (FABS)*, vol. 1, pp. 1–4, IEEE, 2021.

[22] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. e00938, 2018.

[23] J. Zhang, Z. Gu, J. Jang, H. Wu, M. P. Stoecklin, H. Huang, and I. Molloy, "Protecting intellectual property of deep neural networks with watermarking," in *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, pp. 159–172, 2018.

[24] M. Sheikhhassan, M. Sadeghzadeh, and M. Faghihi, "Analysis of health tourist's behaviors using data mining method (international tourists)," *Razavi International Journal of Medicine*, vol. 9, no. 2, pp. 48–55, 2021.

[25] D. Prasad, S. K. Goyal, A. Sharma, A. Bindal, and V. S. Kushwah, "System model for prediction analytics using k-nearest neighbors algorithm," *Journal of Computational and Theoretical Nanoscience*, vol. 16, no. 10, pp. 4425–4430, 2019.

[26] Y. Wang and Z.-O. Wang, "A fast knn algorithm for text categorization," in *2007 international conference on machine learning and cybernetics*, vol. 6, pp. 3436–3441, IEEE, 2007.