Applied Computer Science, vol. 20, no. 2, pp. 175–193
doi: 10.35784/acs-2024-23

Submitted: 2024-03-15 | Revised: 2024-05-30 | Accepted: 2024-05-31

Keywords: Machine Learning, mental health prediction, treatment willingness, healthcare, predictive analysis

Mohammed Chachan YOUNIS [0000-0002-9035-0738]*

PREDICTION OF PATIENT'S WILLINGNESS FOR TREATMENT OF MENTAL ILLNESS USING MACHINE LEARNING APPROACHES

Abstract

Mental illness is a physical condition that significantly changes a person's thoughts, emotions, and capacity to interact with others. The purpose of this study was to explore the application of Artificial Intelligence (AI) and Machine Learning (ML) algorithms in predicting behaviour regarding seeking treatment for mental illnesses, to support healthcare providers in reaching out to and supporting individuals more likely to seek treatment, leading to early detection, enhanced outcomes. The Open Sourcing Mental Illness (OSMI) dataset contains 1259 samples used for research and experiment. The study uses several classifiers (Random Forest, Gradient Boosting, SVM, KNN, and Logistic Regression) to take advantage on their unique capabilities and applicability for various parts of the prediction task. Experiments performed in Jupiter notebook and the major findings revealed varying levels of accuracy among the classifiers, with the Random Forest and 0.81 and Gradient Boosting classifiers 0.83 achieving highest accuracy, while the accuracy for SVM 0.82 and KNN 0.83 also give good result but Logistic Regression classifier had a lower accuracy 0.8. In conclusion, this research demonstrates the potential of AI and machine learning in predicting individual behaviour and offers valuable insights into mental health treatment-seeking behaviour.

1. INTRODUCTION

The human brain is made up of millions of neurons which play an important part in both the body's internal and external communication as well as in interpersonal contact with other people. Nevertheless, disturbances in neuron transmission can have an effect not only on the interior state of the individual, but also on their overall well-being. A person's thoughts, feelings, and the ways in which they interact with other people are all severely impacted when that person has a mental illness. Mental diseases provide a huge worldwide health concern, affecting millions and having far-reaching social and economic consequences (Tan et al., 2024). Regrettably, stigmas from society frequently encourage individuals to keep their mental health concerns a secret, which contributes to the widespread misunderstanding that doing so indicates a lack of personal integrity on their behalf with the result $p < 0.05$. The workforce in the scientific community is subjected to high stress due to long working

* University of Mosul, Department of Computer Science, College of Computer Science and Mathematics, Iraq

hours, and the strain of achieving deadlines, which contribute to the possibility of difficulties with mental health (Bijl et at., 1998). When the far-reaching effects of mental illness on a society are considered, it becomes clear that innovative approaches to provide treatment and prevention are required, in order to detect the problem and act faster. An integral part of these efforts is timely monitoring an individual's psychological well-being (Sun et al., 2010), which includes intellectual, psychological, and policy-related components and influences an individual's thoughts, feelings, and responses in a wide range of situations. Anxiety, social phobia, sadness, obsessive compulsive disorder, addictions, and borderline personality disorder are only a few of the mental disorders that can be caused by a wide variety of psychological difficulties. It is essential to address mental health and encourage a desire to seek therapy across a wide range of demographic groupings, as anxiety and discontent are experiences that are common to all people regardless of their background. Machine learning models can help improve mental health treatment (Nova, 2023). ML approaches identify hidden characteristics in high-dimensional imaging data that may be undetected by traditional statistical methods (Yeung et al., 2023). The rapid development of machine learning (ML) methods is happening in every field, including medicine. Data preprocessing is an extremely important step in the biomedical industry, as it helps transform raw data into a form that can be understood by machine learning models and extracts information that may be used to make predictions (Chen et al., 2018). Considering the arise of these problems, there's also an immediate need to evaluate how AI applications in mental health care might supplement or improve psychologists' existing practices. To fill in the gaps, proposed study will assess the effect of AI applications in mental health care, investigate demographic variations in the prevalence of mental illness, examine the link between mental illness and individual behaviour, and develop a system to predict individual behaviour using ML algorithms. This paper primarily aims to aid medical professionals, such as mental health care facilities and clinics, in anticipating individual decisions to seek treatment for mental illnesses. In particular, this study focuses on people, who live and work on an IT context, because it is unclear how big the impact these new technologies have on the brain and thus the mental stability of the people. To accomplish this goal, ML algorithms and other feature selection strategies that are supported by different data analysis approaches should be used. By analysing the quantitative insights gained from data cleaning and preparation, this study aims to identify differences in the prevalence of mental illness in different geographical regions, genders and ages, as well as in people's attitudes towards and treatment of mental health issues, and in people's behaviour on these issues. In addition, statistical analysis is used to explore the relationship between mental illness and the behaviour of those affected. With the rapid development of technology, AI applications such as wearable technologies and tools for monitoring mental health may supplement or enhance the work of psychologists. The utilization of classification strategies is essential to the process of developing a prediction model for determining a person's current state of mental health (Goodman et al., 2000). To reach that goal, the application of ML models on mental illness must start with fixed dataset taken from patients, which this study focuses on.

Random Forest is well-known for being able to handle big datasets with multiple variables while maintaining high accuracy through ensemble learning. Gradient Boosting, on the other hand, excels by constructing models consecutively to correct flaws in preceding ones, hence enhancing prediction accuracy. To ensure model reliability and generalizability, the data was separated into training and testing sets. The data used to assess model performance is split between training and testing, with 70% of the data used for training and 30% for testing. The result shows that among all machine learning models, two models give the best results: Random Forest at 0.81 and Gradient Boosting classifiers at 0.83. The result for the other two models like SVM obtain 0.82 accuracies and KNN obtain 0.83 accuracy, but logistic gives us the lowest accuracy than other machine learning algorithm used in this study. This study has some limitations. One of them is bias, as the dataset used in this study does not use a diverse population of patients with mental disorders. Another limitation is model overfit due to the generalizability of another dataset. In future studies, the study expands the research with the diverse dataset for mental illness and treatment for patient willingness.

The purpose of this study is to apply Artificial Intelligence and Machine Learning algorithms to predict treatment-seeking behaviors in patients with mental diseases. Using a range of classifiers such as Random Forest and Gradient Boosting, the study reveals profound insights into the factors that influence patients' decisions to seek treatment. This strategy not only helps healthcare providers identify and support patients who are more likely to seek help, but it also improves intervention tactics.

The paper is structured as follows: the coming section discusses the related studies that have already been conducted on machine learning and mental health. The next section describes the proposed machine learning techniques that are used for predicting mental health-related individual behavior. Then it presents the findings of exploratory data analysis and the accuracy metrics i.e., classification reports and confusion matrices, along with the comparison of testing accuracy of machine learning classifiers that are used to predict the willingness of the individuals to seek mental treatment. The study was then completed with conclusions and findings.

2. LITERATURE REVIEW

The arising and the treatment of mental health related problems require developing custom-made strategic planning that considers also cultural and social factors. The three proposed studies aim to contribute to this objective by proposing a system that collects data from relevant research publications, collating them, and using the results to build ML and statistical tools that can be used in the field of mental healthcare. Using feature selection and machine learning algorithms, the study (Bijl et al., 1998), attempt to predict the prevalence of mental illness in the technology industry. Meanwhile, the study (Bijl et al., 1998) looked at the prevalence of psychiatric disorders in the Dutch population and quantified the extent of mental illness in terms of statistics. When it comes to one’s mental health, it’s crucial to catch any problems early on, because they get bigger and thus hardly resolvable. In another study (Hewner et al., 2023), according to the author, coordinating transitional care for people with multiple chronic or complicated chronic diseases, functional limitations, and/or social requirements necessitates collaboration with service providers outside of the health-care system. To ease the computational load of high-dimensional datasets in multiclass problems, for instance, a new feature selection approach was proposed in Sun et al. (2010). In Chen et al. (2018), the authors updated feature selection in the Binary Particle Swarm Optimization (BPSO) technique with the pbest and gbest methods, assessing the efficacy of each partial and partial swarm with a Pearson correlation test. The evaluation of various facets of mental health using machine learning algorithms was the subject of another study (Goodman et al., 2000). The paper Goodman et al., (2000), investigates how the Internet of Things (IoT) could be used to evaluate people's mental health by tracking how much time they spent on various devices during different states of mind. Additional help in overcoming mental health issues have been proposed in Strauss et al. (2013), Jung & Yoon (2017) and Soomro et al. (2022). Serious mental health individuals must be monitored closely, with reliable alarm systems to alert medical staff as soon as possible. Health care sensor network security and risk assessment is discussed in studies (Iyer et al., 2022; Bh et al., 2022; Dao, 2011). Another study (Strauss et al., 2013), investigates the evaluating mental health clinical forms using machine learning techniques like cluster analysis, K-nearest neighbours (KNN), decision trees, and support vector machines (SVM). In (Jung & Yoon, 2017) the authors used biosensors to detect the individual's mental state, based on aspects such as tension, normality and relaxation, and these states are then classified using fuzzy logic and support vector machines. By combining data from multiple types of magnetic resonance imaging (MRI) scans with demographic information from the UK Bio-bank (the world's largest biological cohort), the study (Soomro et al., 2022) aims to develop proxy metrics. The authors Hassan et al. (2023) have invented numerous image processing-based ways to diagnosis, including Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, but they are extremely expensive and require significant training. In Balaji et al. (2023) the study uses hybrid deep learning approaches. The effectiveness of these proxy metrics in capturing real-world health behaviours (e.g., sleep, exercise, and substance consumption) was demonstrated by comparing them with actual measurements. Mental health information was more easily extracted from population models using machine learning when brain signals and questionnaire data were combined in study (Iyer et al., 2022), in which the authors aimed to explaining how mental illness could be diagnosed by looking at the underlying methodologies of existing chatbot applications. While the number of consumer-facing healthcare apps has grown in the years, the mental healthcare sector slowly adopted them, according to analysis exposed in (Iyer et al., 2022). According to that survey (Iyer et al., 2022), over 70% of customers are unhappy with the assistance they received from chatbot applications, while different types of classification models (e.g., Logistic Regression, Random Forest classifier, Decision Tree classifier, stacking, and KNN) were used to improve prediction accuracy. The purpose of the research (Iyer et al., 2022), was to study and understand more about the nature, frequency, and origins of mental health problems. Also, the authors aimed to develop an efficient method of detecting mental illness by utilizing data sources, machine learning methodologies, and feature 4 extraction tools, such as thought processes, decision-making capacities, personality disorders, phobias, psychotic disorders, major depression. The author (Muehlensiepen et al., 2024) uses various algorithms like Logistic regression, Lasso regression, ridge regression, support vector machine (SVM) with linear classifier, SVM with polynomial basis kernel, SVM with radial basis kernel, random forest, neural network, AdaBoost, k-nearest neighbours, naive Bayes, and extreme gradient boosting (XGBoost). The goal of the study presented in (Bh et al., 2022), was to investigate the causes of stress and anxiety among IT workers. The authors used ML algorithms to examine the connection between being self-employed, having a family history of mental illness, having access to mental health benefits at work, and experiencing mental illness at work. The aforementioned (Bh et al., 2022), educate workers about mental illness and help reduce the occurrence of mental health issues in the workplace.

Several stochastic algorithms were used in (Dao, 2011), based on natural phenomena have been explored in the field of computational intelligence that include genetic algorithms, evolutionary programming, evolution methods, and genetic programming. Among the five classification algorithms tested in (Dao, 2011), the KNN classifier had the highest accuracy of 82%, while the reference study attempted to predict the prevalence of mental illness, the proposed model instead attempts to predict individuals' openness to treatment. Another study (Hingorani, 2021), focused on implementing ML tools to use auditory features for the diagnosis of mental disorders. Regarding the used methodology (Hingorani, 2021), supervised ML models were chosen to prevent any misunderstandings from occurring between various degrees of enthusiasm. The author of (Hingorani, 2021) shown that Mental health prediction models can benefit from such kind of data. More information from mobile sensors and internet use can improve the accuracy of predictions.The author (Chung & Teo, 2023) suggests that one viable technique to obtaining fully automated computer-based systems for predicting mental health concerns is using machine learning.The urgency of mental health research was highlighted in study (Singh et al., 2022), which examined the development of various mental health technologies. Because of the intrinsic relationship between the brain and the body, caring for one's mental health is essential to achieving complete wellness of individuals. Although there have been great strides in the diagnosis and treatment of many diseases, more studies are needed to catch mental disorders at an early stage according to the latest study presented in Singh et al., (2022). People's perspectives of trust in organizations like the media, businesses, and government have been significantly influenced by social media according to the study (Ameer et al., 2022). However, there are still difficulties in determining how to categorize feelings based on natural language whereas. In Ameer et al. (2022), the authors investigated if that the problems can arise when data classification methods are adopted inappropriately. To address these difficulties, the research (Sumathy et al., 2022), presented an approach that categorizes Indian tweets sentiments. The tone of a tweet can be used to determine an individual's emotional state, and the adjectives, negations, slang, and acronyms used in Indian tweets all play significant roles in conveying meaning. Depression, anxiety, bipolar disorder, attention deficit hyperactivity disorder, and post-traumatic stress disorder are just some of the mental health conditions that are investigated in Ameer et al. (2022), in which multiclass models for mental disorders detection are used. In order to have a better understanding about what motivates people with depression symptoms to seek treatment, a Seeking Mental Health Care Model was designed and tested in McLaren et al. (2023). The model proposed by the authors of (McLaren et al., 2023), presented a high fit and confirmed most predicted outcomes in online research with 1368 participants. Intention to seek assistance was a predictor of actual assistance seeking. In Wu et al. (2023) Five-fold cross-validation was used for evaluation. The purpose of another related research (Saito et al., 2022), was to create a model that could assist in avoiding the emergence of mental illness by using biometric data from wearable devices and medical examination data. A total of 4,612 people were enrolled in the observational research proposed in Saito et al., (2022), mostly from Japanese medical insurance database. Using the XGBoost method, the authors were able to get an AUC of 0.712 for the receiver-operator-characteristic curve. The research Saito et al., (2022), provides preliminary evidence that wearable technologies may one day be used for early diagnosis and treatment of mental health problems. Using the same survey dataset, related research (Chung & Teo, 2023), compared different machine learning algorithms for mental health prediction.

Multiple machine learning algorithms including Logistic Regression, Gradient Boosting, Neural Networks, K-Nearest Neighbours, and Support Vector Machine were evaluated both independently and collectively.The maximum accuracy i.e., 88.80% was obtained by Gradient Boosting, followed by Neural Networks i.e., 88.00%. The accuracy of deep neural networks was 86.40%, whereas that of extreme gradient boosting was 87.20%. An accuracy of 85.60% was reached by the ensemble classifier. The findings in (Chung & Teo, 2023), suggest that machine learning algorithms can predict mental health issues with an accuracy of 80% or higher, indicating great prospects for automated clinical diagnosis in the field of mental health. The proposed research focuses on the prevalence of mental illness within the tech industry. However, the literature that studies this aspect of mental illness is limited. The main that aim to answer is: what factors influence the tech industry's treatment-seeking willingness, and how can targeted interventions promote mental well-being. The study aims to identify barriers, propose strategies, and enhance mental health support within this unique subset of population.

3. MATERIALS AND METHODS

The purpose of the proposed method is to predict whether an individual will seek treatment for a mental health issue by involving ML algorithms. In this study, random forest was used because it can handle large data sets without overfitting the model. Instead of searching for the most essential feature when dividing a node, it selects the best feature from a random group of features and gradient Boosting continually takes a function that leads in the direction of a weak hypothesis or a negative gradient in order to minimize a loss function. The analysis includes feature selection, performing advanced correlation tests, and fine-tuning the parameters of the model that are relevant. The data is also analysed statistically to draw conclusions about the prevalence of mental illness among the selected tech employees. Six Data collections, analysis, and interpretation procedures, as well as details about the study's sample, sample size, measures, number of variables, and the operationalization of the dependent variable ”willingness to seek treatment,” are provided below.

3.1. Data description and data acquisition

The Open Sourcing Mental Illness (OSMI) in Tech Survey "Mental Health in Tech Survey" dataset is a 2014 non-profit organization survey available on the Kaggle website (Open Sourcing Mental Illness, 2016) that explores attitudes towards mental health and the prevalence of mental health disorders in the tech workplace. The initial raw and unstructured form of the data contains some abnormal values (e.g., missing, and null values) and is composed of 1259 rows (samples) and 27 columns (features) with target variable comment. It includes various demographic and work-related variables, as well as information on treatment seeking attitudes. The dataset allows for analysing geographic variations and identifying predictors of mental health illness and attitudes in the workplace.

The dataset includes the following attributes (features):

– Timestamp: Date and time of survey, not directly linked to the target variable.

- Age: Age of participants provides insights into age-related patterns in mental health attitudes.
- Gender: Gender of participants helps understand the influence of gender on mental health attitudes.
- Country: Country of residence allows exploration of regional variations in mental health attitudes.
- State: State or territory of residence (US only), enables analysis of state-level differences in attitudes.
- Self_employed: Indicates if participants are self-employed, useful for comparing attitudes between self-employed and non-self-employed individuals.
- Family_history: Captures family history of mental illness, can uncover attitudes and treatment-seeking behaviour based on family history.
- Treatment: Indicates if participants sought treatment for mental health conditions, essential for understanding treatment-seeking behaviour.
- Work_interfere: Measures how much mental health conditions interfere with work, reveals the relationship between work interference and treatment-seeking behaviour.
- No_employees: Number of employees in the respondent's organization, helps assess the impact of organization size on attitudes and treatment-seeking behaviour.
- Remote_work: Indicates if participants work remotely at least 50% of the time, enables analysis of the impact of remote work on mental health attitudes.
- Tech_company: Indicates if the employer is primarily a tech company, explores the influence of the tech industry on attitudes and treatment-seeking behaviour.
- Benefits: Thus, reveals whether the employer offers mental health benefits, permits exploration of the relationship between the extent of providing benefits and evidence of treatment.
- Advantages: Identifies whether or not the provider supplies mental health benefits which enables an examination of how the presence of the benefits influences the likelihood of the individual to seek treatment.
- Care_options: Describes the type of mental healthcare that can be accessed through work and presents the impact on acknowledgement on perspectives as well as utilization.
- Wellness_program: All the three objectives determine if the business addresses mental health in employee wellness programmes and also investigates the changes in attitudes and treatment-seeking behaviours.
- Seek_help: Essentially, it checks if the employer provides info on issues to do with mental health and where to seek assistance in cases of such issues exist. Evaluates how a resource enhances or is a hindrance to attitudes and treatment-seeking behaviours.
- Anonymity protection: Evaluates the effects of anonymity on perceived attitudes and willingness, the treatment one receives while seeking help for mental health issues or drug addiction.
- The Leave study evaluates mental health care workers' beliefs regarding the accessibility of medical leave for mental health conditions and investigates effects of the guidelines on attitudes as well as self- reported behaviors related to leave.

- Mental_health_consequence: Captures beliefs about negative consequences of discussing mental health with employers, examines the influence of perceived consequences on attitudes and treatment-seeking behaviour.
- Phys_health_consequence: Captures beliefs about negative consequences of discussing physical health with employers, compares mental and physical health attitudes and treatment-seeking behaviour.
- Coworkers: Measures willingness to discuss mental health with coworkers assesses the impact of social support on attitudes and treatment-seeking behaviour.
- Supervisor: Measures willingness to discuss mental health with direct supervisors, explores the relationship between supervisor support and attitudes/treatment-seeking behaviour.
- Mental_health_interview: Indicates willingness to discuss mental health in a job interview, analyses the influence on employment decisions and attitudes.
- Phys_health_interview: Indicates willingness to discuss physical health in a job interview, compares mental and physical health discussions during job interviews and their impact on employment decisions.
- Mental_vs_physical: Captures perception of employer attitudes towards mental health compared to physical health, examines the influence on attitudes and treatment-seeking behaviour.
- Obs_consequence: Indicates awareness of negative consequences for coworkers with mental health conditions, and explores the prevalence and impact of negative consequences in the workplace.

Analyzing these attributes of the target variable provides insights into factors influencing mental health attitudes and treatment-seeking behavior in the tech workplace. To perform statistical and predictive analyses, the data must undergo cleaning and preprocessing techniques to handle missing or null values and ensure it is structured and suitable for machine learning models.

3.2. Data preprocessing and cleaning

The raw data is processed using Python Pandas. As first step, samples containing missing values are removed, and undesirable columns (Time, Comments, State and no employees) are removed. This work used a variety of data preparation techniques to reduce dataset from 1259 rows and 27 variables to 977 rows and 23 variables. This study's analysis was carried out using Jupyter Notebook, an interactive computing environment that allows for data purification, transformation, visualization, and modeling. This preprocessing was essential to ensure that the data was accurate and relevant to investigation. Categorical data categories are converted into numerical 8 values, beginning with 0. Also, the preprocessed and cleaned data are all integers, as seen in Fig. 1.

Age	Gender	Country	self_employed	family_history	treatment	work_interfere	remote_work	tech_company	benefits	...
46	35	38	1	1	0	3	1	1	2	...
29	19	38	0	1	1	3	0	1	2	...
31	35	38	1	0	0	0	1	1	1	...
46	19	38	0	0	1	1	1	1	2	...
41	19	38	0	0	1	0	0	0	0	...

Fig. 1. An example of preprocessed data in an organized manner

The original dataset contained a total of 1259 samples (rows) and 27 features after performing preprocessing and cleaned dataset was reduced to 977 samples with 23 columns.

3.3. Exploratory data analysis approaches to analyse the data

The proposed research on individuals seeking mental health care must include exploratory data analysis (EDA) in order to gain understanding of the dataset's properties and gain insights. Statistical analysis is needed to examine the links, patterns, and trends found in the study's data. Descriptive and inferential statistics are used to examine and evaluate the sampled data. Descriptive analysis will be used to examine variables such as mean, median, mode, standard deviation, variance, range, and percentiles of each column in order to give an overview of the dataset.

3.4. Exploratory data analysis and feature engineering

Feature selection is a method for discovering highly correlated features with the target, as described in a study (Bh et al., 2022). Correlation analysis allows for examining linear relationships between variables and quickly identifying potential associations. It aids in Exploratory Data Analysis (EDA) and highlights influential variables. Before moving on to additional modelling or other types of more systematic analyses, exploratory data analysis is a crucial step in the data analysis process. It involves examining and determining what the most significant features, characteristics, patterns, or relationships within any given dataset are. Correlations complement ML techniques and serve as an initial step to identify relevant features before exploring complex associations using more advanced methods. In addition, by combining a feature engineering step and before using a ML model the results of the latter can be improved, especially when operating on tabular data as in this case. The correlation is calculated using the Spearman’s correlation, which measures the magnitude and direction of a monotonic relationship between two variables. The maximum correlation is indicated by a score of 1 and the worst correlation is indicated by 0. The heat map color saturation technique represents the higher correlation with light color whereas the lowest 9 correlation is illustrated by high saturated color. The highly correlated independent features extracted from the correlation test are used to train ML algorithms.To obtain a fundamental comprehension of the distribution and central tendencies of the dataset, EDA can compute descriptive statistics. For continuous variables, statistics like mean, median, mode, range, variance, and standard deviation were calculated; for categorical variables, frequency counts and modes were employed. This stage assisted in locating any anomalies, including outliers or erroneous records. The correlation score bar plot shown in Fig. 2, indicates the relationship between various features and the target variable ’Treatment’. Positive correlation values imply that as the value of a feature increases, the likelihood of seeking treatment also increases. Whereas negative correlation values suggest that the feature value decreases the likelihood of seeking treatment decreases. From the positive correlations, can conclude that factors such as ’work interfere’, ’family history’, ’benefits’, ’care options’, and ’anonymity’ have a stronger positive influence on treatment-seeking behaviour.On the other hand, the negative correlations between ’tech_company’, ’Gender’, ’phys health consequence’, ’self_employed’, and ’mental_health_consequence’ and treatment-seeking behaviour are weaker. Although these factors have a negative influence, the magnitude of the correlations is relatively small.

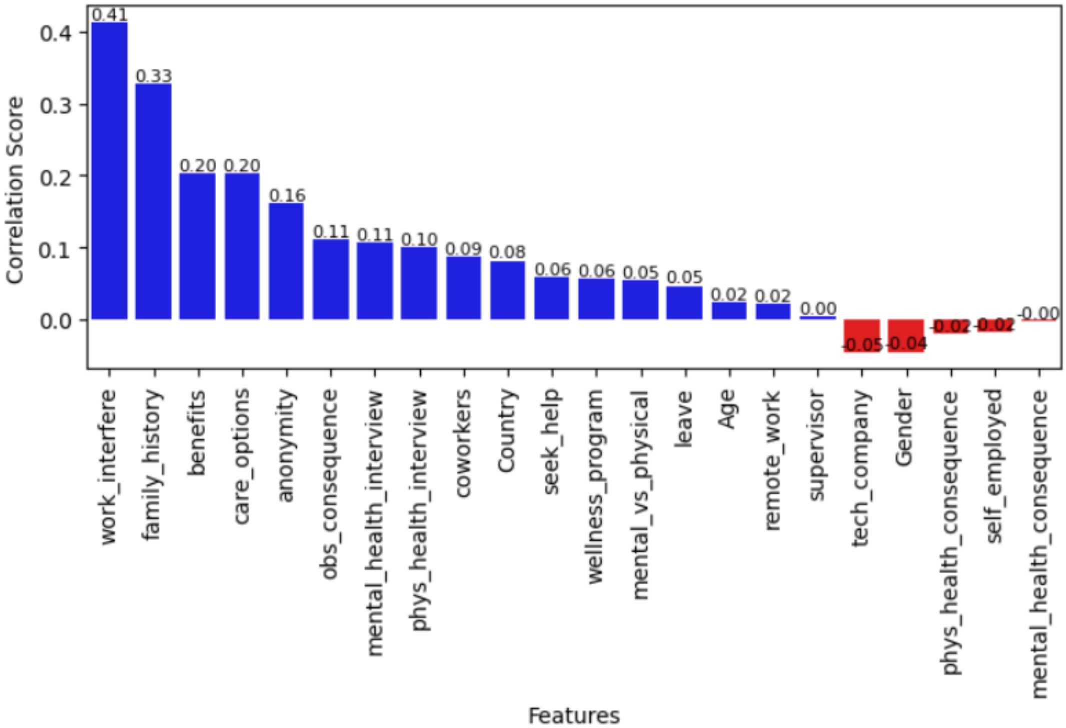

Fig. 2. Correlation of features with the target treatment

Age distribution counts shown in Fig. 3, concluded that there is a wide range of ages represented in the age column, from 18 to 49. The dataset covers a large range of ages, from 18 to 49, among the valid age values that can be used to drive the conclusion that this age range is working in the tech population. Numbers for each age bracket represent the total number of people in the dataset. The highest number can be used to extrapolate the median age of the people in the tech community who filled out the form. In this case, 32 is the oldest age group, with 66 totals. Therefore, 32 is the most common age group among the respondents from the tech community.

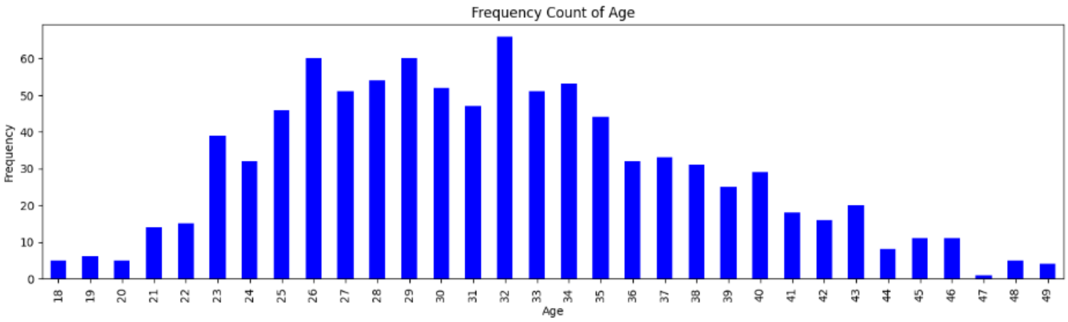

Fig. 3. Age distribution in the dataset

The provided cross-tabulation shown in Fig. 4, reveals the counts of individuals classified by their age and treatment-seeking willingness. The first row indicates that among individuals aged 32, there are 26 individuals who did not seek treatment and 40 individuals who sought treatment. These choices were heavily impacted by a number of important characteristics, such as symptom severity, workplace support, family history, and understanding of mental health issues. These observations emphasize how crucial it is to have accessible mental health resources and supportive settings in order encourage treatment involvement. Analyzing the entire table allows us to gain insights into the distribution of treatment-seeking behaviours across different age groups. Age 32 has the highest count of individuals seeking treatment compared to other age groups, with a count of 40. Age 33 follows closely with 38 individuals seeking treatment. Additional age groups such as 26, 28, and 25 also exhibit relatively high counts of individuals seeking treatment. It’s important to note that the counts vary across different age groups, with some showing a higher proportion of individuals seeking treatment compared to those not seeking treatment. Analysing cross-tabulations provides valuable information about the relationship between variables and helps identify patterns or trends.

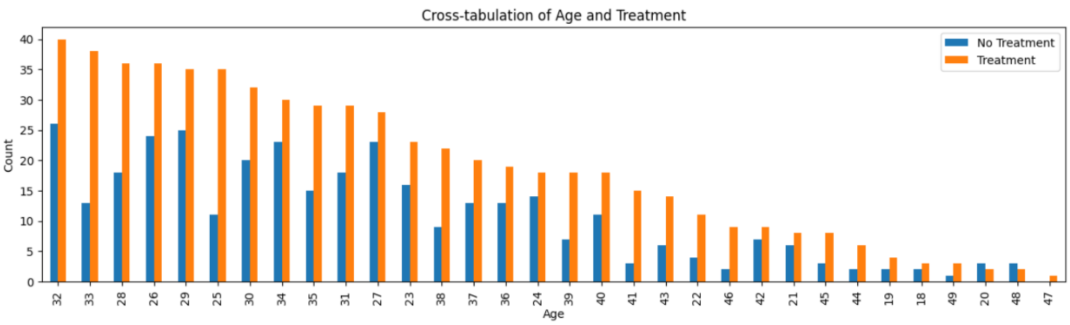

Fig. 4. An overview of the distribution of treatment-seeking among different age groups

3.5. Machine learning approaches

To predict whether an individual will seek treatment for their mental illness, the proposed system employs machine learning algorithms such as Support vector machines (SVM), Decision trees (DT), Naive Bayes (NB), Gradient boosting (GB), K-nearest neighbor (K-NN), logistic regression (LR), and Random Forest (RF). To predict whether an individual will choose to seek treatment for their mental disease or not, the system uses classifiers. The performance of the classification models is based on the accuracy of predictions, f1-score, recalls, precisions, and parameter-tuning strategies to improve the classifiers' accuracy. The proposed system diagram is shown in Fig. 5 which shows the overall workflow from dataset selection followed by various algorithms as well as their predicted results.

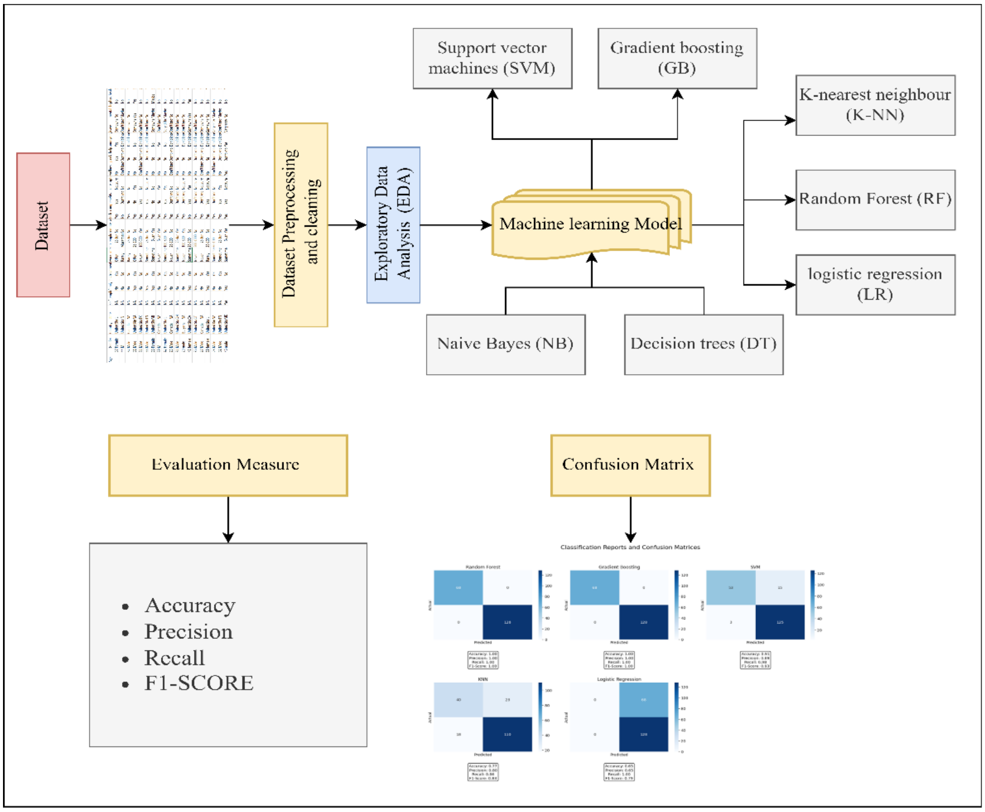

Fig. 5. Overall proposed diagram

The RF classifier is deployed using 500 (five hundred) decision trees, the GB classifier computes the results using 400 weak algorithms, and the SVM classifier uses the probability method, auto kernel coefficient gamma for RBF kernel as represented in Tab. 1. Moreover, the K-NN classifier considers the values of the 12 nearest neighbors for the prediction with the uniform weight and the algorithm used by K-NN is kd-tree. The LR classifier used balanced class weight with the multiclass solver 'sag'.

Tab. 1. Tuned parameters for achieving better accuracy of classifiers

Classifier	Tuned Parameters
RF	n_estimators=500
GB	n_estimators= 400
SVM	probability=True, gamma = 'auto', kernel = 'rbf'
K-NN	n_neighbours= 12, weights = 'uniform', algorithm= 'kd_tree'
LR	penalty = 'l2', class_weight = 'balanced', solver = 'sag', multi_class = 'auto'

3.6. Hyper-parameters of proposed study

Learning rate: The learning rate is an important hyperparameter in machine learning since it influences the step size throughout the optimization process. It controls how much the model's weights are altered relative to the loss gradient during each iteration. A high learning rate might result in faster convergence but risks overshooting the ideal solution.

Epoch: An epoch is a learning interval in which the model evaluates the complete training dataset. It specifies the maximum number of iterations the algorithm can make across the entire training dataset, which can vary depending on the dataset used. An epoch is defined as the number of times all images are processed forward and backward across the network.

Batch size: In machine learning, the batch size refers to the number of training samples handled in a single iteration. As a result, the mini-batch size is the total number of sub-samples submitted to the network prior to parameter modifications. Batch sizes can be 32, 64, 128, 256, 512, and similar possibilities.

3.7. Interpretation of feature importance

This section goes deeper into the interpretation of the variables that are most influential in predicting treatment-seeking behavior. The Random Forest and Gradient Boosting classifiers identified critical factors that strongly influence treatment-seeking behavior. Work interference identified as a top predictor, showing that people whose mental health difficulties interfere with their employment are more likely to seek help. Family history was also important, implying that those who have a family history of mental illness are more conscious and aggressive in seeking help. Employer-provided benefits and knowledge of care options had a substantial impact on the likelihood of seeking treatment, highlighting the importance of workplace mental health support and effective communication about resources.

4. RESULTS AND DISCUSSION

4.1. Results

4.1.1. Results on training data

As shown in Fig. 6, the confusion matrix for all the machine algorithms used in the study RF and GB classifiers demonstrated exceptional accuracy by correctly predicting instances from both classes without any occurrences of false positives or false negatives. The support vector machine (SVM) classifier exhibited a relatively low number of misclassifications, with 10 instances being falsely identified as positive and only 1 instance being falsely identified as negative. Overall, the classifier demonstrated a high accuracy in correctly predicting the majority of instances. K-NN classifier exhibited a relatively higher rate of misclassifications. Specifically, it accurately predicted 181 instances belonging to one class and 455 instances belonging to the other class. However, it misclassified 104 instances from the former class and 41 instances from the latter class. The Logistic Regression classifier exhibited a lack of accurate predictions for one class, while successfully predicting 494 instances of the other class. It produced 285 false positives and 2 false negatives. The obtained results offer valuable insights regarding the performance and accuracy of each classifier in effectively classifying the provided training data.

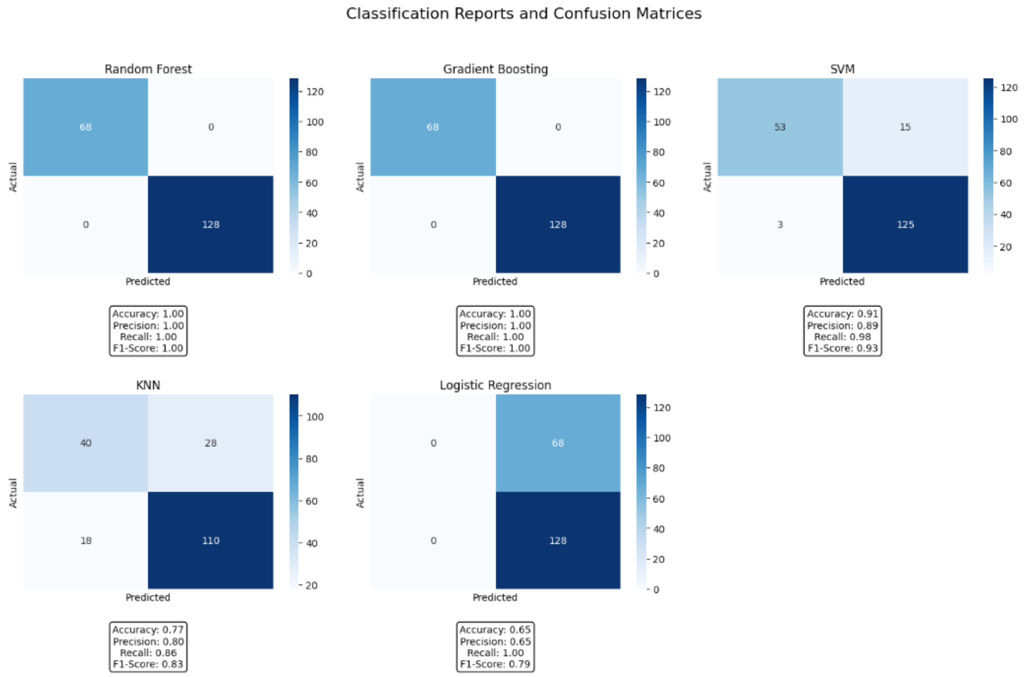

Fig. 6. Accuracy metrics i.e., confusion matrices and classification reports of classifiers on the Training dataset

The Random Forest and Gradient Boosting classifiers exhibited exceptional performance, achieving perfect accuracy. In contrast, the SVM classifier demonstrated high accuracy with only a limited number of misclassifications. Nevertheless, the K-NN classifier exhibited elevated rates of misclassification, while the Logistic Regression classifier encountered difficulties in accurately classifying one particular class but demonstrated satisfactory performance for the other class. The results shed light on the diverse performance exhibited by the classifiers, thereby facilitating their appropriate selection for particular classification endeavors, taking into account factors such as accuracy and rates of misclassification.

4.1.2. Results on testing data

The models are trained using 70% of the training data and tested using 30% of the testing data. Here, the performance of different classifiers on unseen test data was assessed using accuracy metrics. The results are shown in Tab. 2 and in Fig. 7. They accurately predicted the willingness of individuals to seek treatment, as evidenced by perfect precision, recall, and F1-scores for both classes.

Tab. 2. Scores obtained on the test dataset

Classifier	Accuracy	Precision	Recall	F1-SCORE
Random Forest	0.81	0.82	0.81	0.81
Gradient Boosting	0.83	0.85	0.83	0.83
SVM	0.82	0.84	0.82	0.82
K-NN	0.83	0.84	0.83	0.83
Logistic Regression	0.8	0.8	0.8	0.8

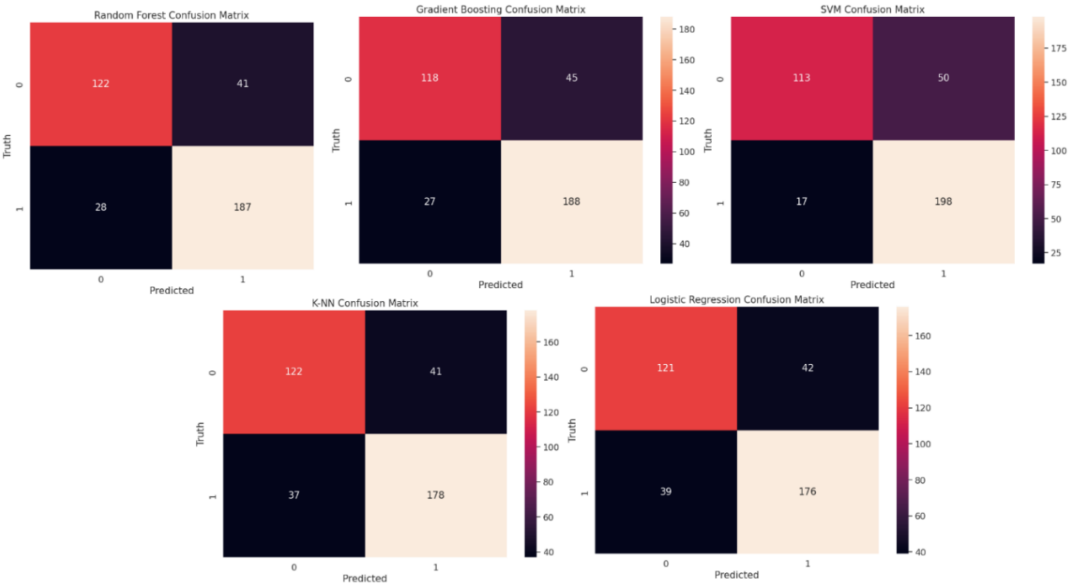

Fig. 7. Accuracy metrics i.e., confusion matrices and classification reports of classifiers on the testing dataset

The SVM classifier demonstrated good performance with an accuracy of 0.82. It classified most instances correctly, particularly for both classes, achieving a precision of 0.84 and recall of 0.9828. The F1-score was 0.9382 indicating a good balance between precision and recall. The K-NN classifier achieved an accuracy of 0.83 on the testing data, which was relatively lower compared to the previous classifiers. It performed reasonably well in predicting instances of both classes, with a precision of 0.84 and recall of 0.83. The F1-score of K-NN was 0.83, indicating a reasonable balance between precision and recall. Logistic regression produces the following results, which are inferior to those obtained by other algorithms used in this study, with precision 0.8, recall 0.8, and F1 score 0.8.

The above-mentioned results obtained from several classifiers provide important understanding on patient willingness for individual treatment. Random forest and gradient boosting achieve very good results, as the compared to the similar study (Chung & Teo, 2023). The SVM classifier performed well overall, but the KNN classifier provided reasonable precision and recall for class 1, while having much lower accuracy. The Logistic Regression classifier produced the most accurate predictions for class 1 cases, although having the lowest accuracy. These results facilitate an understanding of the advantages and disadvantages of each classifier within the study's parameters.

4.2. Discussion

The patient's willingness to discuss mental health issues is assessed by the study using a variety of algorithms. The Random Forest, Gradient Boosting, SVM, KNN, and Logistic Regression algorithms are utilized by the method for prediction. The results of this study advance our knowledge of how patient behaviour analysis using machine learning algorithms can be utilized to treat mental illnesses. Although the Strengths and Difficulties Questionnaire (SDQ) and traditional statistical models were used by Goodman et al., (2000) to predict psychiatric disorders in clinical settings, this study use latest models to predict mental health treatment-seeking behaviour in a tech workplace setting and also this study makes use of the "Mental Health in Tech Survey" dataset, which covers a broader range of characteristics including personal background, work-related issues, and mental health status across a sizable and diverse community of professionals in the tech industry. This makes it possible to analyse the variables impacting people's decisions to seek mental health therapy in a broader and more complex way. Although gradient boosting and random forest perform better than the others, all methods produce the best results. The significant positive correlations seen between these variables suggest that people who are facing obstacles at work, have a family history of mental health problems, and have access to mental health benefits and care options are more likely to pursue treatment. Healthcare professionals can provide timely interventions and individual care by accurately predicting how likely a patient is to seek treatment. This could result in early diagnosis and improved mental health results. Tab. 3 provide details about the compression of the current research with the stat-of the art models.

Tab. 3. Previous Study

References	Method	Dataset Details	Results
(Muehlensiepen et al., 2024)	Machine learning-powered analysis	German survey data	Prediction of telemedicine acceptance (different for all models)
(Tan et al., 2024)	Machine learning model	1,000 participants	$p < 0.05$
(McLaren et al., 2023)	Seeking Mental Health Care Model	Online research with 1368 participants	$M_{age} = 42.38$, $SD_{age} = 15.22$, 65.6% female
(Chung & Teo, 2023)	Single and ensemble machine learning	Various machine-learning datasets	Accuracy = 85.60%
(Ameer et al., 2022)	Deep learning and transfer learning	Social media texts	Accuracy = 80%
The proposed model	Random Forest	Mental Health in Tech Survey	Accuracy = 81%
	Gradient Boosting	Mental Health in Tech Survey	Accuracy = 83%
	SVM	Mental Health in Tech Survey	Accuracy = 0.82%
	K-NN	Mental Health in Tech Survey	Accuracy =83%
	Logistic Regression	Mental Health in Tech Survey	Accuracy = 80%

5. CONCLUSIONS

The proposed study examined how artificial intelligence (AI) and machine learning algorithms can be used to forecast an individual's behaviour when it comes to seeking mental health treatment. This study makes a contribution by using a machine learning algorithm to predict the patient's willingness to undergo mental disorder therapy. This contributes to the healthcare department and assists medical practitioners. And, by using real-time capabilities, this study aims to improve access to mental health services, thereby lowering the impact of this severe condition on individuals, families, and communities. By evaluating AI's influence on mental health treatment, looking at differences in the frequency of mental illness among different demographics, researching the relationship between mental disease and individual behaviour, and creating a machine learning-based prediction system, the study effectively addressed the research topics. People experience fear of confessing that they are suffering from a mental illness, and in this society promotes the misconception that clinical depression reflects a moral failure on the part of the individual who has it. This study found that multiple machine learning classifiers, including Random Forest, Gradient Boosting, SVM, K-NN, and Logistic Regression, had varying levels of success in forecasting an individual's behaviour while seeking mental health therapy. The Random Forest and Gradient Boosting classifiers were 100% accurate on the training data, but the SVM classifier had a small number of misclassifications. Whereas the Logistic Regression classifier did well in one class but struggled in another, the KNN classifier showed greater misclassification rates on the training data. The testing subset of the dataset yielded 100% accuracy for the Random Forest and Gradient Boosting classifiers, indicating their effectiveness in identifying patients who are likely to seek treatment. As compared to KNN algorithm, SVM give us better results. In conclusion, the testing dataset showed decreased accuracy for the Logistic Regression classifier. It makes it easier for healthcare providers to identify and assist people who are more willing to seek treatment. By initially determining which patients are more likely to seek assistance, healthcare practitioners can more effectively focus their services to individuals in need of mental health therapy. This could improve access to care, reduce wait times for specific patients, and fortify the mental health care system. A number of factors, including high levels of stress, long work hours, and demanding workplaces, can negatively impact an employee's mental health in the technology sector. The willingness of technology workers to ask for help can help companies tailor wellness initiatives, insurance plans and workplaces to the specific needs of their employees. Among tech workers, the suggested system may boost output, retention, and job satisfaction.

Future research could benefit from integrating larger and broader datasets to improve the model's generalizability and robustness, as well as novel deep learning methods such as transformers or attention mechanisms to provide more sophisticated data representations and improve prediction accuracy.

The study underlines the potential of artificial intelligence and machine learning in improving the prediction and comprehension of individual behavior in connection to mental health treatment. Machine learning algorithms can help improve mental health care practices by accurately predicting treatment-seeking behavior and guiding the development of focused interventions. However, further research is necessary to advance this field and enhance mental health care practices.

Author Contributions

"Mohammed Chachan Younis" is the sole author for this investigation.

Funding

Not Applicable.

Acknowledgments

The author would like to thank Kaggle data science website for the availability of this valuable dataset to conduct this research. Also, the author would like to acknowledge the support provided by MHESR and Univ. of Mosul.

Conflicts of Interest

The author declares no competing interests.

REFERENCES

Ameer, I., Arif, M., Sidorov, G., Gòmez-Adorno, H., & Gelbukh, A. (2022). Mental illness classification on social media texts using deep learning and transfer learning. *ArXiv, abs/2207.01012*. https://doi.org/10.48550/arXiv.2207.01012

Balaji, P., Chaurasia, M. A., Bilfaqih, S. M., Muniasamy, A., & Alsid, L. E. G. (2023). Hybridized Deep Learning approach for detecting Alzheimer's disease. *Biomedicines, 11*(1), 149. https://doi.org/10.3390/biomedicines11010149

Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social psychiatry and psychiatric epidemiology*, *33*, 587-595. https://doi.org/10.1007/s001270050098

Chen, Y., Wang, Y., Cao, L., & Jin, Q. (2018). An effective feature selection scheme for healthcare data classification using binary particle swarm optimization. *2018 9th international conference on information technology in medicine and education (ITME)* (pp. 703-707). IEEE. https://doi.org/10.1109/ITME.2018.0016017

Chung, J., & Teo, J. (2023). Single classifier vs. ensemble Machine Learning approaches for mental health prediction. *Brain informatics*, *10*, 1. https://doi.org/10.1186/s40708-022-00180-6

Dao, T. T. (2011). Investigation on evolutionary computation techniques of a nonlinear system. *Modelling and Simulation in Engineering*, *2011*(1), 496732. https://doi.org/10.1155/2011/496732

Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European child & adolescent psychiatry*, *9*, 129-134. https://doi.org/10.1007/s007870050008

Hassan, F., Hussain, S. F., & Qaisar, S. M. (2023). Fusion of multivariate EEG signals for schizophrenia detection using CNN and machine learning techniques. *Information Fusion*, *92*, 466-478. https://doi.org/10.1016/j.inffus.2022.12.019

Hewner, S., Smith, E., & Sullivan, S. S. (2023). Identifying high-need primary care patients using nursing knowledge and Machine Learning methods. *Applied Clinical Informatics*, *14*(03), 408-417. https://doi.org/10.1055/a-2048-7343

Hingorani, M. (2021). Detection of mental illness using machine learning and deep learning. *6th North American International Conference on Industrial Engineering and Operations Management*.

Iyer, M. V., Kadlag, M. T., Patil, M. M., Pillai, M. A., & Moholkar, K. P. (2022). Virtual self care companion - Detection of mental illness using machine learning and deep learning. *Specialusis Ugdymas*, *1*(43), 5955-5964.

Jung, Y., & Yoon, Y. I. (2017). Multi-level assessment model for wellness service based on human mental stress level. *Multimedia Tools and Applications, 76*, 11305-11317. https: //doi.org/10.1007/s11042-016-3444-9

McLaren, T., Peter, L. J., Tomczyk, S., Muehlan, H., Schomerus, G., & Schmidt, S. (2023). The seeking mental health care model: prediction of help-seeking for depressive symptoms by stigma and mental illness representations. *BMC Public Health*, *23*, 69. https://doi.org/10.1186/s12889-022-14937-5

Open Sourcing Mental Illness. (2016). *Mental Health in Tech Survey: Survey on Mental Health in the Tech Workplace in 2014*. Kaggle. https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey

Saito, T., Suzuki, H., & Kishi, A. (2022). Predictive modeling of mental illness onset using wearable devices and medical examination data: Machine Learning approach. *Frontiers in Digital Health, 4*, 861808. https://doi.org/10.3389/fdgth.2022.861808

Singh, P., Srinivas, K. K., Peddi, A., Shabarinath, B., Neelima, I., & Bhagavathi, K. A. (2022). Artificial Intelligence based early detection and timely diagnosis of mental illness - A review. *2022 International Mobile and Embedded Technology Conference (MECON)* (pp. 282-286). IEEE. https://doi.org/10.1109/MECON53876.2022.9752219

Soomro, T. A., Zheng, L., Afifi, A. J., Ali, A., Soomro, S., Yin, M., & Gao, J. (2022). Image segmentation for MR brain tumor detection using machine learning: A Review. *IEEE Reviews in Biomedical Engineering*, *16*, 70-90. https://doi.org/10.1109/RBME.2022.3185292

Strauss, J., Peguero, A. M., & Hirst, G. (2013). Machine learning methods for clinical forms analysis in mental health. *IOS Press, 192*, 1024. https://doi.org/10.3233/978-1-61499-289-9-1024

Sumathy, B., Kumar, A., Sungeetha, D., Hashmi, A., Saxena, A., Kumar Shukla, P., & Nuagah, S. J. (2022). Machine Learning technique to detect and classify mental illness on social media using lexicon-based recommender system. *Computational Intelligence and Neuroscience, 2022*(1), 9790823. https://doi.org/10.1155/2022/5906797

Sun, Y., Todorovic, S., & Goodison, S. (2009). Local-learning-based feature selection for high-dimensional data analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1610-1626. https://doi.org/10.1109/TPAMI.2009.190

Muehlensiepen, F., Petit, P., Knitza, J., Welcker, M., & Vuillerme, N. (2024). Prediction of the acceptance of telemedicine among rheumatic patients: a machine learning-powered secondary analysis of German survey data. *Rheumatology International*, *44*, 523-534. https://doi.org/10.1007/s00296-023-05518-9

Nova, K. (2023). Machine Learning approaches for automated mental disorder classification based on social media textual data. *Contemporary Issues in Behavioral and Social Sciences*, *7*(1), 70-83.

Tan, M., Xiao, Y., Jing, F., Xie, Y., Lu, S., Xiang, M., & Ren, H. (2024). Evaluating machine learning-enabled and multimodal data-driven exercise prescriptions for mental health: a randomized controlled trial protocol. *Frontiers in psychiatry, 15*, 1352420. https://doi.org/10.3389/fpsyt.2024.1352420

Wu, C.-H., Hsu, J.-H., Liou, C.-R., Su, H.-Y., Lin, E. C.-L., & Chen, P.-S. (2023). Automatic bipolar disorder assessment using Machine Learning with smartphone-based digital phenotyping. *IEEE Access*, *11*, 121845-121858. https://doi.org/10.1109/ACCESS.2023.3328342

Yeung, H. W., Stolicyn, A., Buchanan, C. R., Tucker-Drob, E. M., Bastin, M. E., Luz, S., McIntosh, A. M., Whalley, H. C., Cox, S. R., & Smith, K. (2023). Predicting sex, age, general cognition and mental health with machine learning on brain structural connectomes. *Human Brain Mapping, 44*(5), 1913-1933. https://doi.org/10.1002/hbm.26182