Heliyon 9 (2023) e19422

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Single-label machine learning classification revealed some hidden but inter-related causes of five psychotic disorder diseases

Hilary I. Okagbue [a,*], Ogochukwu A. Ijezie [b], Paulinus O. Ugwoke [c,d], Temitope M. Adeyemi-Kayode [e], Oluranti Jonathan [f]

[a] *Department of Mathematics, Covenant University, Ota, Nigeria*
[b] *Faculty of Science and Technology, Bournemouth University, Poole, BH12 5BB, UK*
[c] *Department of Computer Science, University of Nigeria, Nsukka, Nigeria*
[d] *Digital Bridge Institute, International Centre for Information & Communications Technology Studies, Abuja, Nigeria*
[e] *Department of Electrical & Information Engineering, Covenant University, Ota, Nigeria*
[f] *Department of Computer & Information Sciences, Covenant University, Ota, Nigeria*

A R T I C L E I N F O

Keywords:
Classification
Diagnosis
Feature importance
Machine learning
Psychotic disorder
Single label approach

A B S T R A C T

Psychotic disorder diseases (PDD) or mental illnesses are group of illnesses that affect the minds and impair the cognitive ability, retard emotional ability and obstruct the process of communication and relationship with others and are characterized by delusions, hallucinations and disoriented or disordered pattern of thinking. Prognosis of PDD is not sufficient because of the nature of the diseases and as such adequate form of diagnosis is required to detect, manage and treat the illness. This paper applied the single-label classification (SLC) machine learning approach in mining of electronic health records of people with PDD in Nigeria using eleven independent (demographic) variables and five PDD as target variables. The five PDDs are Insomnia, Schizophrenia, Minimal Brain dysfunction (MBD), which is also known as Attention-Deficit/Hyperactivity Disorder (ADHD), Vascular Dementia (VD) and Bipolar Disorder (BD). The aim of using SLC is that it would be easier to detect some PDDs that are related to each other without the loss of information, which is a plus over multi-label classification (MLC). RelieffF algorithm was used at each experiment to precipitate the order of importance of the independent variables and redundant variables were excluded from the analysis. The order of the variables in feature selection was matched with feature importance after the classifications and quantified using the Spearman rank correlation coefficient. The data was divided into: 70% for training and 30% for testing. Four new performance metrics adapted from the root mean square (RMSE) were proposed and used to measure the differences between the performance results of the 10 Machine learning models in terms of the training and testing and secondly, feature and without feature selection. The new metrics are close to zero which is an indication that the use of feature selection and cross validation may not greatly affects the accuracy of the SLC. When the PDDs are included as predictors for classifying others, there was a tremendous improvement as revealed by the four new metrics for classification accuracy (CA), precision and recall. Analysis of variance showed the four different metrics differs significantly for classification accuracy (CA) and precision. However, there were no significant difference between the CA and precision when the duo are compared together across the four evaluation metrics at p value less than 0.05.

* Corresponding author.
E-mail address: hilary.okagbue@covenantuniversity.edu.ng (H.I. Okagbue).

https://doi.org/10.1016/j.heliyon.2023.e19422
Received 21 November 2022; Received in revised form 4 August 2023; Accepted 22 August 2023
Available online 23 August 2023

1. Introduction

Psychotic Disorder Disease (PDD) is a state of impaired mental or cognitive functioning which results in the inability to coordinate physical activities and emotions [1]. Difficulty in sustaining interaction and maintaining relationships are also some of its manifestations [2]. PDD has been described as the leading cause of disability [3] and the detailed epidemiology and prevalence can be found in Ref. [4]. According to the World Health Organization (WHO), PDD, suicide and neurological disorders are among the emerging causes of morbidity [5]. Low-and middle-income countries are home to a significant percentage of people living with PDD [6]. The burden of mental health compounds the challenges in the public health profiles of developing countries [7]. Avenues to diagnose, treat or manage them helps to reduce the social, psychosocial and economic losses caused by PDD [8,9], and generally improve the overall quality of life.

The traditional methods of diagnosis of PDD are gradually being replaced with modern methods because of the heterogeneity and comorbidity of some of the diseases. The traditional methods of diagnosis are failing to capture the necessary features that can provide insights into the nature of PDD because the data generated are becoming larger and complex to analyze [10]. The heterogeneity underlying the nature of the human brain limits the ability of statistical methods to interpret or integrate different datasets generated from the analysis of brain function [11]. Furthermore, the multidimensional data obtained in the analysis of brain function with the aim of detecting PDD could better be managed using advanced methods that can create patterns by clustering, reduce dimensionality without loss of information, classify disease instances or predict the likelihood of disease occurrence with permissible error [12]. The accurate prediction could create room for a more focused intervention targeted at reducing the occurrence of psychotic episodes [13].

Advances in medical sciences have provided better alternatives such as applying machine learning (ML), evolutionary and optimization algorithms and neuroimaging. ML techniques utilize the computational strength of algorithms in creating patterns that explain latent relationships within a given data [14]. The field of psychiatry has benefited from using ML in assistance with diagnosis, prediction of psychosis [15], prognosis and treatment of PDD [16]. An example of the predictive capacity of ML is predicting the cognitive brain for patients with Alzheimer's disease [17]. ML has been able to unravel hidden patterns, hence deepening the understanding of the aetiology and pathogenesis of PDD [18], thereby opening new treatment options and better management of the diseases. ML can link disease symptoms to the part of the brain from a given data, thereby precipitating from a large quantum of data to hidden associations that maps symptoms to PDD with minimal human input [19]. Such interconnections and associations among different regions of the human brain could be modelled using different data mining models and network analysis [20]. Apart from symptoms, ML is routinely used to determine the nature, magnitude and order of the risks factors that predict a given mental illness [21]. Unsupervised (clustering), supervised, semi-supervised, and reinforcement learning are applied in different aspects of PDD. Deep learning (DL) which is a subset of ML models are increasingly used because their computational strengths are matched with the complex nature of the brain. For instance, predicting the onset of schizophrenia has been an unmet diagnostic challenge until recently [22,23]. The knowledge gaps in extant research paradigms are filled with the application of DL with evidence of a better understanding of the nature, prevalence and orientation of PDD [24,25]. In addition, ML and DL are used in detecting PDD in magnetic resonance imaging datasets obtained via medical examination [26,27]. Hence, making it possible for psychiatric prognosis [28]. ML has proved in identifying key risk factors in the management, treatment of PDD, and consequently, reducing hospital re-admission [29] and optimum allocation of mental healthcare resources [30]. ML can be used in training datasets generated from different sources in different formats such as datasets of emotions fluctuation or digital expressions of human behaviour obtained via smartphones [31], wearable devices and digital phenotyping techniques [32]. The usefulness of ML in the analysis of PDD data has some unsettled issues. Concerns about ethics [33], reliability and interpretability of the results [34], methodology, data integration [35], quality [36] and privacy of data obtained from different electronic health records. Tackling these challenges will help advance the course of using ML and DL in analysing PDD data.

Recent studies have applied ML learning techniques for multi-label classification of five PDDs using 11 demographic variables [37,38]. The goal of Multi-Label Classification (MLC) is to allot an instance to a set of different labels [37,38]. In this case, the five PDDs are Insomnia, Schizophrenia, Minimal Brain dysfunction (MBD), which is also known as Attention-Deficit/Hyperactivity Disorder (ADHD), Vascular Dementia (VD) and Bipolar Disorder (BD). The use of MLC allows for a simultaneous diagnosis of the 5 PDD. Three major problems are associated with the use of MLC. First, information may be lost during the transformation of the variables. For instance, binary reduction may not be feasible for some variables. Second, MLC could lead to class imbalance problems, especially where few instances are observed for a given variable. Third, the loss of information and huge technicalities involved in MLC may lead to misleading information such as comorbidity where an individual can have more than one disorder. MLC allows for simultaneous diagnosis of more than one disease. Lastly, combining the PDD into one instance reduces the efficiency of the machine learning classifiers as seen in the performance metrics reported in Refs. [37,38] because it is easier to understand the risk factors of a disease individually [39].

This paper applied the Single Label Approach (SLA), which considers the possibility of comorbidity by using some of the PDD as independent variables. Hence, it would be easier to detect some PDDs that are related to each other without the loss of information.

The main idea is to diagnose PDD independently using other PDD and 11 demographic variables as independent variables. This approach will help address the loss of information in MLC that can affect the comorbidities (simultaneous diagnosis which is an advantage of MLC over SLC). Comorbidities in PDD or mental illness is common; however, the relation of a disorder with another does not infer causation. The relationship can be better viewed as features or risk factors that can be easier to interpret in SLC. Variables with the highest risk (feature importance in this aspect) compensate for the entropy prevalent in the diagnosis of mental illness since risk factors may be similar for different PDD [40,41]. In addition, the comparison between the order of importance of independent variables against a target variable using feature selection methods and the order of feature importance after classification results have not feature in the literature in this context. Hence, investigation of the correlation between the orders will help in deeper understanding of the nature of the PDD. As seen later in this article, the order of contribution of the risk factors differed for each PDD which could not be obtained if comorbidities are done using MLC.

The line of argument profoundly presented here is using other PDDs and demographic factors as predictors rather than as a comorbid. Most researchers that use sociodemographic, psychiatric or other index variables in predicting PDD often yield inconsistent results because of latent factors such as genetics, environmental and phenotypes [42,43]. In some instances, PDDs may be in a state of comorbidity and can only be detected by data models that classify the hidden disorder as an independent variable rather than a co-dependent variable [44]. Some of the causes may not be captured in the measuring tool or data instrument, but the presence of a significant predictor(s) can link to a further probe which can lead to the discovery of the latent cause. As a result, emphasis on comorbidity reduced performance of the data mining models, an example can be seen in Ref. [45].

Low performance scores have been reported for comorbidities; F1 score of 68.02% for psychiatric disease comorbid in a study in China has been reported despite of the huge dataset used [46]. Conceptualizing comorbidity may be a daunting task [47] and concerns over heterogeneity within the target classes could lead to misclassification and diagnosis [48]. Hence, the emphasis should be on the accuracy of classification to reduce hospital readmission of patients, reduction of risks inherent in the management and treatment of psychotic disorders and improving the overall quality of life. Accurate predictive models are highly sought after to achieve the aims. But accuracy of the models must be investigate to ensure that high accuracy and precision is not due to overfitting. Hence, performance metrics on the assessment of the difference between the training and test results will help to boost confidence on the results.

Consequently, the objectives of this paper are as follows:

a). Apply feature selection method before classification. The feature selection will eliminate some of the independent variables.
b). To classify each of the 5 PDD independently using the 11 demographic variables using 10 Machine learning models and output the classification accuracy (CA), precision and recall.
c). To classify each of the 5 PDD independently using both the 11 demographic variables and the remaining 4 PDD as co-predictors using 10 Machine learning models and output the classification accuracy (CA), precision and recall.
d). To assess the performance of the classification in (b) and (c) without using feature selection.
e). To assess the performance of the results in (b) and (c) testing on test data versus testing on the training data.
f). To investigate the effect of including other PDD versus excluding them as independent variables in classifying PDD.

2. Methods

2.1. Data

The data was historical PDD patients from 2010 to 2014 admitted in Yaba Psychiatric Hospital, Lagos State, Nigeria. The data was published as a data article [49]. The summary of the data is presented in Table 1.

The sample size of 500 patients consists of 267 females and 233 males who were tested for five PDDs namely Insomnia, Schizophrenia, MBD, VD and BD. The data on the survival and treatment were not made available due to poor data records. Eleven demographic variables as presented in Table 1 and are the independent variables. They are sex which is categorical (male and female); the age is continuous and ranges between 6 and 86; religion is categorical (Christianity, Islam and others); occupation is categorical

Table 1
The data summary.

	Details
Subjects	Psychotic patients
Source	Yaba Psychiatric Hospital, Yaba, Lagos State, Nigeria
Year	2010 to 2014
Sample size	500
Sex	267 females, 233 males
Age range	Between 6 and 86 years
Religion	222 Christianity, 219 Islam, 59 Others
Occupation	144 Artisan, 73 Civil servant, 21 Force, 46 Retired, 120 Student, 96 Unemployed
Marital Status	219 Single, 281 Married
Divorce	440 No, 60 Yes
History in Family	231 No, 269 Yes
Hereditary	279 No, 221 Yes
Loss of Parents	202 No, 298 Yes
Head Injury	406 No, 94 Yes
Spiritual Consultation	153 No, 347 Yes
Insomnia	297 Negative, 203 Positive
Schizophrenia	75 Negative, 425 Positive
VD	154 Negative, 346 Positive
MBD	282 Negative, 218 Positive
BD	299 Negative, 201 Positive

(artisan, civil servant, force, retired, student and unemployed) and marital status is categorical (single and married). The remaining six variables (divorce status, history of PDD in family, hereditary, loss of parents; head injury and spiritual consultation) are all categorical with the same binary responses (yes and no).

Five target variables are the five PDD, which are categorical with the same binary diagnostic outcomes (negative and positive). Negative and positive connotes the absence and presence of PDD respectively for a given patient.

2.2. Data analysis tools and methods

SPSS version 24 was used in the cross-tabulation and Chi-square test of independence test of the four target variables. Orange software (https://orangedatamining.com) was used in the feature selection and classification analysis, and the eleven independent variables were used to classify each of the four target variables separately. Secondly, some of the target variables were included as independent variables and was used to classify each other.

The classification was done using the following ML models: Adaptive boosting (AD), Gradient Boosting (GB), Neural Network (NN), k-Nearest Neighbor (kNN), Naïve Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), Classification Tree (Tree), Logistic Regression (LR) and Stochastic Gradient Descent (SGD). These ML models are very popular in classification and regression problems and are readily available in Python, IBM SPSS Modeler, Orange Software, RapidMiner, KEEL, KNIME, WEKA, R and other data mining software. The major goal of classification models is to predict the class or category of a given input. Details on classification models can be seen in Ref. [50].

As discussed in the introduction section, SLC was chosen because of the following;

(i). It is easier to implement than MLC because SLC avoids the ambiguity that can arise in multi-label scenarios, where an instance can belong to multiple classes simultaneously [51].
(ii). It is fast and hence, reduced computational complexity, hence fewer computational resources and memory [52].
(iii). It has fewer odds than MLC in producing false positives [53].
(iv). SLC models are less susceptible to class imbalance [54].
(v). SLC models are less likely to overfit, as they are only predicting a single label [55].

Data sampler was used to divide the data into two; 70% for training and 30% for testing (cross-validation).

The goal of applying feature selection methods is to identify the most relevant and informative features from a given dataset, thereby reducing complexity and improving model interpretability. ReliefF was selected among other feature selection methods because of its robustness and sensitive to interactions. The Relief F algorithm ranks the order of importance of the variables and all the variables are have zero or negative values were excluded. Other available feature selection methods for classification are information gain, Gini ratio, Gini, ANOVA, Chi-square, Fast Correlation Based Filter (FCBF).

Three performance metrics were used to evaluate the performance of the ML models. The three are CA, precision and recall. Confusion matrices were connected to output instances of correct and incorrect classifications. Feature explanation widget was connected to graphically depict the order of effects of the independent variables in the classification of the target variables.

T-test, analysis of variance (ANOVA) and corresponding posthoc was performed using SPSS version 24, to determine the mean significance of the proposed metrics.

2.3. Evaluation metrics

Four new evaluation metrics that measures the extent of which cross validation affects the sensitivity, accuracy and specificity of machine learning classification results were proposed.

2.4. Ethical approval

Not applicable. Data from a published data article was analyzed using machine learning methods. The data article was fully cited (integrity), and is an observational study of health records without contacts with the psychotic patients (low risk). The data contains no

Table 2
Cross-tabulation of the diagnostic result of the four target variables.

		Positive						
		0	1	2	3	4	5	Total
Negative	**0**						48	48
	1					88		88
	2				149			149
	3			150				150
	4		55					55
	5	10						10
Total		10	55	150	149	88	48	500

demographic, phenotypical or genotypic variables that can be linked to the patients (confidentiality). Moreover, the original dataset was approved by the Health Review Board of the University of Ilorin Teaching Hospital, Nigeria.

3. Results

3.1. Cross-tabulation

Cross-tabulation was performed to show the extent of the prevalence of PDD among the 500 patients. The result as presented in Table 2 showed that only 10 (2%) tested negative for all the five PDD, 55 (11%) tested positive for only one PDD; 150 (30%) tested positive for two PDD; 149 (29.8%) tested positive for three PDD; 88 (17.6) tested positive for four PDD and 48 (9.6) tested positive for all the five PDD. Furthermore, it can be deduced that 490 (98%) of the patients has at least a PDD while only 10 (2%) were not diagnosed with any of the five PDD.

Chi-square test of independence showed a significant relationship among the diagnostic outcomes of the five PDDs (Chi-square = 2500, degrees of freedom = 25, p-value <0.0001). This showed that the target variables have imbalance classes. The imbalance classes will have a minimal effect since single factor classification was used.

3.2. Feature selection and classification analysis

Ten ML models available in Orange Software were adopted to classify the target variable using the 11 independent variables only without including the remaining PDD. The ML models are Adaptive boosting (AD), Gradient Boosting (GB), Neural Network (NN), k-Nearest Neighbor (kNN), Naïve Bayes (NB), Random Forest (RF), Support Vector Machine (SVM), Classification Tree (Tree), Logistic Regression (LR) and Stochastic Gradient Descent (SGD).

The raw results of the application of ReliefF as a feature selection are provided in **Supplementary Data** which also contains the classification results (test on test data versus test on training data) for feature versus without feature selection.

3.3. Experiments

Firstly, the eleven independent variables were used to classify the five PDD independently without including any PDD as a predictor. The result of the evaluation metrics was presented in no order for the classification of insomnia, Schizophrenia, VD, MBD and BP. The order of the performance of the ML models are presented in the **Supplementary Data.**

Secondly, the eleven independent variables and four PDD were used to classify each of the five PDD independently. Hence the other four PDD serves as part of independent variables in the classification of the remaining one. In these instances, the independent variables increase to 15.

Feature Importance is done to determine the order of importance or the impact of the independent variables in classifying the five PDD. Accordingly, the order is in decreasing feature importance of the independent variables in diagnosing the PDD, and the top ones are selected for all the cases.

The feature importance did not indicate the nature (positive or negative) impact of the variables, but only indicates the strength of importance of the variables in classifying or predicting the target variables. Generally, feature importance refers to a class of techniques for assigning scores to independent variables to a predictive model that indicates the relative importance of each feature (predictor) when making a classification [56,57]. The process of feature selection involves identifying the most informative features via discrete ranking that are relevant to the target variable, and then removing the irrelevant or redundant features that may introduce noise or bias into the model. The ones that are relevant are ranked higher than those that are redundant. Including the redundant will cause reduce accuracy, waste computation time and reduced the interpretability (principle of parsimony).

Comparison between the order of importance of independent variables against a target variable using feature selection methods

Table 3
Analysis of the relationship of the variables between ones obtained via Feature importance and ones with or without feature selection.

Target	Without Feature selection			With Feature selection		
	NOIV	R	R Square	NOIV	R	R Square
Insomnia I	11	0.509	0.259	8	0.333	0.111
Schizophrenia I	11	0.354	0.126			
VD I	11	0.464	0.215	7	0.464	0.215
MBD I	11	0.482	0.232	5	0.6	0.36
BD I	11	0.182	0.033	7	−0.464	0.216
Insomnia II	15	0.107	0.011	11	0.564	0.318
Schizophrenia II	15	0.071	0.005			
VD II	15	0.307	0.094	13	0.033	0.001
MBD II	15	0.068	0.004	13	−0.236	0.056
BD II	15	0.607*	0.368	11	0.218	0.047

*($p < 0.05$); NOIV = Number of Independent variables; I means that only the demographic factors were used as predictors; II means that other PDD were used as predictors.

and the order of feature importance after classification results are presented in Table 3. Most importantly, the rank of the feature importance was rearranged to correspond with the ones obtained using ReliefF for the cases of with or without feature selection. In addition, for feature selection, excluded variables in ReliefF were excluded in the classification and hence, do not appear in the feature importance. Spearman rank correlation (R) was used to quantify the relationship while R square also known as coefficient of determination was used to measure the extent of the relationship. R square close to zero and one connotes worst and best fits respectively.

Note: The best performing ML models were used to obtain the rank of the feature importance.

From Table 3, the number of independent variables after application of ReliefF was shown. However, using all the variables are selected in the two cases of classifying Schizophrenia. This is an indication that all the 11 and 15 independent variables in both cases are included in the classification and hence no correlation could be performed. Only two negative correlations were obtained while only one significant positive correlation were obtained. This has shown that the ranking of variables using feature selection does not always predict the rank or order of importance of the variables after classification. The low values of the R square indicate that feature selection does not imply classification and vice versa.

3.4. Proposed evaluation metrics

Four different experiments yielded four different but similar performance metrics. In all the cases, 10 ML models yielded different values of CA, precision and recall as could be seen in the **Supplementary Data.**

First is classification without feature selection and testing on training data (TrW). Second is classification without feature selection and testing on test data (TeW). Third is classification with feature selection and testing on training data (TrF). Fourth is classification with feature selection and testing on training data (TeF).

The use of testing on test data is often preferred result if overfitting is suspected. An example of classification of Insomnia for the four experiments are shown in Fig. 1 (CA) and Fig. 2 (precision).

Figs. 1 and 2 showed that Adaptive Boosting is the best performing ML model while Naïve Bayes is the least. Generally, TrW and TeF yielded the best and worst results respectively for the 10 ML models. However, it appears that the four results of the experiments converges at Adaptive Boosting. Other plots of the remaining PDD can be produced and the numerical values of the performance metrics of the 10 ML models can be assessed in the Supplementary Data.

However, individual analysis of the models will be cumbersome and SLC produces more precise results than MLC. Hence, the authors proposed four different evaluation metrics which are a modification of the RMSE. The metrics is aim as assessing the performance of CA and precision of the 10 ML models between two experiments results. The proposed metrics are to measure the following and presented in **Equations (1)–(3)** and **4** respectively.

A. Difference between Test on Test Data and Test on Training Data without Feature selection

$$R_1 = \sqrt{\frac{\sum_{i=1}^{N} (ML_{TrW} - ML_{TeW})^2}{N}} \tag{1}$$

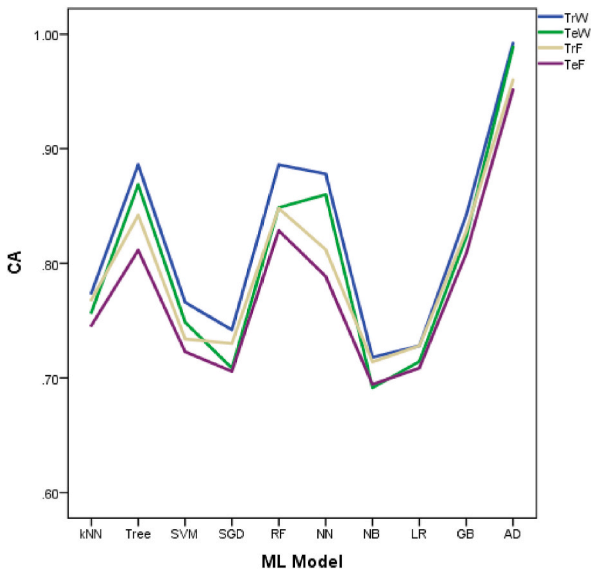

Fig. 1. Comparison of the CA of the 10 ML models for the 4 experiments (Insomnia I).

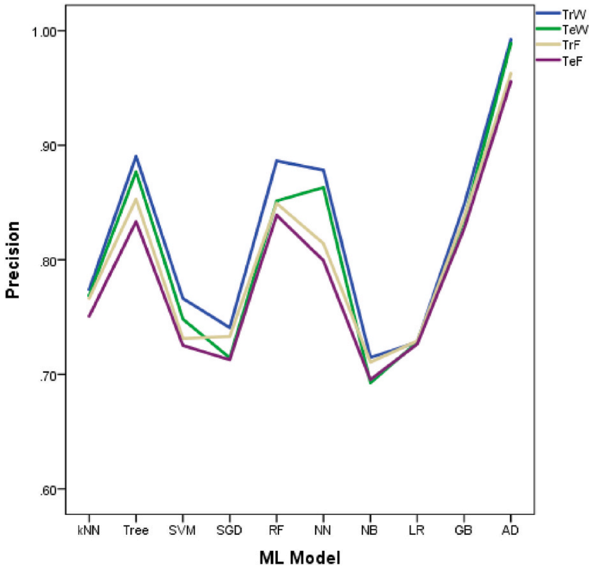

Fig. 2. Comparison of the precision of the 10 ML models for the 4 experiments (Insomnia I).

here ML_{TrW} and ML_{TeW} are the CA and precision of the TrW and TeW for the 10 ML models.

B. Difference between Test on Test Data and Test on Training Data with Feature selection

$$R_2 = \sqrt{\frac{\sum_{i=1}^{N} (ML_{TrF} - ML_{TeF})^2}{N}} \tag{2}$$

here ML_{TrF} and ML_{TeF} are the CA and precision of the TrF and TeF for the 10 ML models.

C. Difference between the classifications results of with and without feature selection for Test on Training Data

$$R_3 = \sqrt{\frac{\sum_{i=1}^{N} (ML_{TrW} - ML_{TrF})^2}{N}} \tag{3}$$

D. Difference between the classifications results of with and without feature selection for Test on Test Data

$$R_4 = \sqrt{\frac{\sum_{i=1}^{N} (ML_{TeW} - ML_{TeF})^2}{N}} \tag{4}$$

Table 4
Summary of the result of the four performance metrics for CA.

Target	R_1	R_2	R_3	R_4
Insomnia I	0.022374913	0.020709063	0.031899843	0.033454203
Schizophrenia I	0.014648535			
VD I	0.005107817	0.009150889	0.025783716	0.022223356
MBD I	0.008304904	0.009006574	0.051722336	0.0450759
BD I	0.018280356	0.025854065	0.028642626	0.035144019
Insomnia II	0.002317634	0.006310341	0.009423375	0.0058554
Schizophrenia II	0.010114447			
VD II	0.003398679	0.01043972	0.008049845	0.01146423
MBD II	0.008631693	0.008602325	0.017955501	0.019959142
BD II	0.002469405	0.003683388	0.005403702	0.00404061

Table 5
Summary of the result of the four performance metrics for Precision.

Target	R_1	R_2	R_3	R_4
Insomnia I	0.018337643	0.013376927	0.030409134	0.028558522
Schizophrenia I	0.015293759			
VD I	0.005137953	0.009968433	0.025997972	0.022182974
MBD I	0.008845655	0.009483422	0.051862703	0.045050914
BD I	0.013925117	0.021717699	0.029264056	0.036255312
Insomnia II	0.002082988	0.006086024	0.009387255	0.005656229
Schizophrenia II	0.010707874			
VD II	0.004168119	0.010771259	0.007849775	0.011423947
MBD II	0.008573501	0.008167627	0.016405115	0.017884112
BD II	0.002315796	0.003477053	0.005370151	0.00400991

Generally, there are no much differences between the pairs; TrW and TeW, TrF and TeF, TrW and TrF, and TeW and TeF. The use of feature selection and cross validation may not greatly affects the accuracy of the SLC.

in all the four proposed metrics, N = 10, which is the number of the ML models. The metrics are similar to the RMSE which ranges from 0 to 1. Values close to zero signifies that there is a significant better fit between the two models.

The results of the metrics are presented in Table 4 (CA) and Table 4 (Precision) (see Table 5).

3.5. Statistical evaluation of the proposed metrics

Analysis of variance (ANOVA) was performed and varying mean differences were obtained for R_1 to R_4 and are presented in Table 6. Least significance difference (LSD) was the posthoc tool adopted. The last six columns of Table 6 are the multiple comparison result of the LSD.

As seen in Table 6, the ANOVA models are all significant at $p < 0.05$.There are significant mean difference between the pairs (R_1, R_3) and (R_1, R_4) for CA, and (R_1, R_3), (R_1, R_4) and (R_2, R_3) for the precision metric. On the other hand, there appears to be no significant mean difference between the pairs of others as shown in Table 6. The result amplify the similarity between having feature selection or not and cross-validation or not.

Finally, *t*-test was conducted to investigate the mean difference between the results in CA and precision assuming equal variances. The result, as presented in Table 7, showed the absence of significance mean difference between the results obtained for CA and precision. The Levene's test for equality of variance were not also significant, which is an indication that the assumptions of the test was not violated.

3.6. Summary of the feature importance

The summary of the feature importance for the two major groups (with or without feature selection) is presented in Table 8 which represented the top 4 features in decreasing order of contribution to the classification.

Hence, age appears to be the leading feature in diagnosing PDD in almost all the results. BD is the leading predictor of insomnia while insomnia is the leading predictor of BD. The results of with or without feature selection are almost the same especially for insomnia I and VD II.

Table 6
Summary of the ANOVA of the four performance metrics for CA and precision.

ML metric	F	R_1 & R_2	R_1 & R_3	R_1 & R_4	R_2 & R_3	R_2 & R_4	R_3 & R_4
CA	2.933*	−0.00215	−0.01280*	−0.01259*	−0.01060	−0.01043	0.00021
Precision	3.418*	−0.00144	−0.01313*	−0.01244*	−0.01169*	−0.01099	0.00069

*p value < 0.05.

Table 7
Summary of the T-Test of the four performance metrics between CA and precision.

Metric	F	T
R1	0.359	0.222
R2	1.057	0.406
R3	0.002	0.038
R4	0.002	0.105

*p value < 0.05.

Table 8
Summary result of the Feature importance with or without Feature selection.

Without Feature	1st	2nd	3rd	4th	With Feature	Ist	2nd	3rd	4th
Insomnia I	Age	Religion	LOP	OCC	Insomnia I	Age	Religion	LOP	OCC
Schizophrenia I	Age	LOP	SC	Religion	Schizophrenia I	Age	LOP	SC	Religion
VD I	SC	Age	HIF	Hereditary	VD I	SC	Age	Hereditary	HIF
MBD I	Age	Sex	LOP	HIF	MBD I	Age	Sex	OCC	MS
BD I	Age	LOP	Religion	HIF	BD I	Age	Religion	HIF	Sex
Insomnia II	BD	OCC	Age	MBD	Insomnia II	BD	Age	OCC	SC
Schizophrenia II	Age	SC	VD	LOP	Schizophrenia II	Age	SC	VD	LOP
VD II	SC	Age	HIF	Insomnia	VD II	SC	Age	HIF	Insomnia
MBD II	Age	Sex	OCC	HIF	MBD II	Age	Sex	OCC	LOP
BD II	Insomnia	OCC	LOP	HI	BD II	Insomnia	OCC	MS	Hereditary

LOP = Loss of Parents; HIF = History in Family; OCC = Occupation; SC = Spiritual Consultation; HI = Head Injury; MS = Marital Status.

Table 9
Model comparison with [37,38].

Target	[37]	This paper (with FS	This paper (without FS)	[38]
Insomnia II	Multi label	0.9874	0.9851	0.7929
Schizophrenia II	Multi label	0.9396	0.9485	0.9250
VD II	Multi label	0.8851	0.8923	0.8536
MBD II	Multi label	0.8137	0.8097	0.7786
BD II	Multi label	0.9868	0.9851	0.8143
Average	0.513448	0.9225	0.9242	0.8329

FS = Feature Selection.

3.7. Model comparison

The authors in Ref. [37] worked with the same data used in this paper in MLC. In their paper, the five PDD was used as a multi-label dependent variable. The classification resulted in 32 instances that are from (0, 0, 0, 0, 0) to (1, 1, 1, 1, 1) while in SLC, only 5 instances are possible. In MLC, more than one label (PDD) can be predicted while SLC allows only one at a time. The comparison was done by finding the arithmetic mean of the 60 models used by the authors **(**Table 4 of [37]**)** and compared with the average of the 10 ML models used to predict each PDD with or without feature selection **(Supplementary Data)** and the result is presented in Table 9. Direct comparison is not possible because the labels of the dependent variables are different.

Similarly, the author in Ref. [38] used the same data and the model performance evaluation for single-label classification is performed by keeping the feature variables the same but changing the target to represent a symptom. The same methodology was used in this paper. However, the authors produced results for imbalanced and balanced dataset however, the comparison is limited to imbalanced class because this paper did not consider the effect of class imbalance. Also, the authors used deep learning while this paper used ordinary machine learning classification models. The comparison is the average of the CA of the models used in this paper and the ones reported in (Table 1 of [38]. The result is the last column of Table 8.

On the average, the result of this paper appears to perform better than others.

5. Discussion and conclusion

In the preliminary result obtained from this study, it can be deduced from the cross-tabulation, that 490 (98%) of the patients have at least a PDD, while only 10 (2%) were not diagnosed with PDD. This implies a high prevalence of PDD in Nigeria since the sample represents the larger population of PDD patients. The high prevalence was because the data was the diagnostic results of patients undergoing treatment. This contradicts the lower prevalence reported in Refs. [58,59], which included random samples that contains non PDD patients.

By harnessing the power of SLC, the study succeeded in efficiently detecting related PDDs without loss of information, a notable advantage over MLC. The use of ML has been applied to classify five PDDs using 11 demographic variables. Furthermore, in the second experiment, PDD was added as independent variables to classify a single PDD and the experiment was conducted five times for both cases. The aim is to classify the present (positive) or the absence (negative) of PDD using independent variables. The evaluation metrics showed better results than the ones obtained in Refs. [37,38], which is an evidence of better performance by using single factor classification. Although, this approach cannot be used in simultaneous diagnosis of PDD. The strength of this article is that the near-perfect diagnosis result would be very useful since misdiagnosis of PDD can results to severe health consequences, injuries and death. The argument here goes in favor of precision than comorbidity since it is often difficult to establish comorbidity in mental illness because some of the ailments have similar risk factors such as age. Although age is often cited as a risk factor of PDD, it often interacts with a myriad of other factors, including but to limited to genetic and environmental variables, substance use, socioeconomic factors and overall public health [60]. Accurate diagnosis as seen in this work will greatly help to reduce the burden on mental healthcare professional and optimize the allocation of limited resources in management of mental illness as suggested in Ref. [61]. A decision support system that implements these results especially for low resources settings is recommended [62,63].

This work has revealed the tolerance level of the cross validation in ML classification. Testing on test data reduces the accuracy and precision of ML classification but is preferred because testing on training data may lead to overfitting and inaccurate estimates. The proposed metrics in this work has shown that the use of feature selection and testing on train data could be tolerated to a certain level, although further research is needed in this direction. The application of feature selection in this work helps to magnify the relevant and informative features and removing redundant issues, leading to more accurate and meaningful interpretability [64] and creating avenue for the proposal of new performance metrics as obtained in this paper. The lack of significant correlation between feature selection and feature importance is an area that merits more research because the former is supposed to be somewhat similar to the later. Further research is urgently needed to output an acceptable bound where cross validation may be acceptable. The proposed metric is the extension of the RMSE but took into account the peculiarities of cross-validation (testing versus training).

The order of importance of the risk factors (independent variables) in the accurate diagnosis of PDD is another contribution of this research. The prominent feature of age as a predictor of mental illness does not come as a surprise since it has been reported that the prevalence of PDD increases with increasing age [65]. Most PDD has age as their major risk factor. Bipolar disorder and insomnia were the leading risk factors between each other which corroborates the earlier findings that patients with BP tend to experience high rates of insomnia [66–68]. Spiritual consultation, which can also be viewed as a form of superstition was found to be a major risk factor for vascular dementia, a disease that has age as a major risk factor [69]. Again, this is unsurprising as Nigeria is a very religious country where PDD are considered as spiritual attacks instead of mental diseases [70]. In this study, the findings on spiritualism are consistent with [71], who noted that it can be measured and used to improve the conditions of those living with VD. In this case, the data used for this research used yes or no to determine if the patients have had spiritual consultation before coming to the psychiatric hospital for diagnosis. The order of importance of the risk factors could be useful in the treatment, management, and counselling of PDD patients. Because of the sensitive nature of PDD, efforts should be made to conduct a comprehensive assessment of PDD patients to identify and rank the specific risk factors. Treatment and interventions should on the basis of the ranks. However, caution should be exercised since the risk factors can change over time [72].

In conclusion, the near perfect result obtained from some of the SLC machine learning models present hope of isolating PDD patients from the general population in a low income setting. MLC model can help in detecting comorbidities but is plagued with lower precision, high false positives and diagnosis. The high risk of PDD demands accurate or precise diagnosis than comorbidities, which this paper supports with the analysis of the result. In addition, simple questionnaire-like decision support system could be used to diagnose PDD while further medical tests is used for confirmation. This will save lives, resources and dangers associated with the manifestations of mental illness. In addition, the incorporation of PDDs as predictors for classifying others yielded remarkable improvements in the performance metrics, as demonstrated by the innovative metrics introduced in this paper.

6. Limitation

The only major limitation of this approach is it is could be complex to implement in decision support system if the aim is to diagnose more than one PDD at an instance. This is because different lines of codes are to be written for the diagnosis of the individual PDDs.

Author contribution statement

Hilary I. Okagbue, PhD; Ogochukwu A Ijezie; Paulinus O Ugwoke; Temitope M Adeyemi-Kayode, PhD; Oluranti Jonathan, PhD: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors appreciate the efforts of the anonymous reviewers toward this publication. The support from Covenant University, Nigeria is also deeply appreciated.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e19422.

References

[1] L. Cook, Mental Health in Australia: A Quick Guide, Parliament of Australia, Canberra, ACT, Australia, 2019, pp. 6–10.
[2] H. Herrman, E. Jané-Llopis, The status of mental health promotion, Publ. Health Rev. 34 (2012) 6, https://doi.org/10.1007/BF03391674.
[3] H.A. Whiteford, A.J. Ferrari, L. Degenhardt, V. Feigin, T. Vos, The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010, PLoS One 10 (2015), e0116820, https://doi.org/10.1371/journal.pone.0116820.
[4] M. Mansourian, S. Khademi, H.R. Marateb, A comprehensive review of computer-aided diagnosis of major mental and neurological disorders and suicide: a biostatistical perspective on data mining, Diagnostics 11 (2021) 393, https://doi.org/10.3390/DIAGNOSTICS11030393.
[5] S. Saxena, Y. Setoya, World health organization's comprehensive mental health action plan 2013–2020, Psych Clin. NeuroSci. 68 (2014) 585–586, https://doi.org/10.1111/pcn.12207.
[6] I. Sharpe, C.M. Davison, A scoping review of climate change, climate-related disasters, and mental disorders among children in low-and middle-income countries, Int. J. Environ. Res. Publ. Health 19 (2022) 2896.
[7] M.O. Owolabi, M. Leonardi, C. Bassetti, J. Jaarsma, T. Hawrot, A.I. Makanjuola, .J.D. Roa, The neurology revolution, Lancet Neurol. 21 (2022) 960–961.
[8] H. Oh, A. Nagendra, M. Besecker, L. Smith, A. Koyanagi, J.S.H. Wang, Economic strain, parental education and psychotic experiences among college students in the United States: findings from the Healthy Minds Study 2020, Early Interven. Psych. 16 (2022) 770–781.
[9] A.Y. Mikhaylov, A.V. Yumashev, E. Kolpak, Quality of life, anxiety and depressive disorders in patients with extrasystolic arrhythmia, Arch. Med. Sci. 18 (2022) 328.
[10] D. Bzdok, A. Meyer-Lindenberg, Machine learning for precision psychiatry: opportunities and challenges, Biol. Psych. Cogn. Neurosci. Neuro. 3 (2018) 223–230, https://doi.org/10.1016/j.bpsc.2017.11.007.
[11] S. Vieira, X. Liang, R. Guiomar, A. Mechelli, Can we predict who will benefit from cognitive-behavioural therapy? A systematic review and meta-analysis of machine learning studies, Clin. Psychol. Rev. 97 (2022), 102193.
[12] D.B. Dwyer, P. Falkai, N. Koutsouleris, Machine learning approaches for clinical psychology and psychiatry, Annu. Rev. Clin. Psychol. 14 (2018) 91–118, https://doi.org/10.1146/annurev-clinpsy-032816-045037.
[13] C. Benjet, G. Borges, S. Miah, Y. Albor, R.A. Gutiérrez-García, A. Zavala Berbena, .P. Mortier, One-year incidence, predictors, and accuracy of prediction of suicidal thoughts and behaviors from the first to second year of university, Depress. Anxiety 39 (2022) 727–740.
[14] J.W. Lai, C.K. Ang, U.R. Acharya, K.H. Cheong, Schizophrenia: a survey of artificial intelligence techniques applied to detection and classification, Int. J. Environ. Res. Publ. Health 18 (2021) 6099, https://doi.org/10.3390/ijerph18116099.
[15] W. Yin, L. Li, F.X. Wu, Deep learning for brain disorder diagnosis based on fMRI images, Neurocomputing 469 (2022) 332–345.
[16] K. Kim, J.I. Ryu, B.J. Lee, E. Na, Y.T. Xiang, S. Kanba, S.C. Park, A machine-learning-algorithm-based prediction model for psychotic symptoms in patients with depressive disorder, J. Personalized Med. 12 (2022) 1218.
[17] P. Lu, L. Hu, N. Zhang, H. Liang, T. Tian, L. Lu, A two-stage model for predicting mild cognitive impairment to Alzheimer's disease conversion, Front. Aging Neurosci. 14 (2022), 826622.
[18] A. Tyagi, V.P. Singh, M.M. Gore, Towards artificial intelligence in mental health: a comprehensive survey on the detection of schizophrenia, Multimed. Tool. Appl. 82 (2023) 20343–20405.
[19] G.S. Suri, G. Kaur, S. Moein, Machine learning in detecting schizophrenia: an overview, Intel. Auto, Soft Comput. 27 (2021) 723–735, https://doi.org/10.32604/IASC.2021.015049.
[20] Y.T. Jo, S.W. Joo, S.H. Shon, H. Kim, Y. Kim, J. Lee, Diagnosing schizophrenia with network analysis and a machine learning method, Int. J. Methods Psychiatr. Res. 29 (2020), e1818, https://doi.org/10.1002/mpr.1818.
[21] O. Karasch, M. Schmitz-Buhl, R. Mennicken, J. Zielasek, E. Gouzoulis-Mayfrank, Identification of risk factors for involuntary psychiatric hospitalization: using environmental socioeconomic data and methods of machine learning to improve prediction, BMC Psychiatr. 20 (2020) 401, https://doi.org/10.1186/s12888-020-02803-w.
[22] M. Montazeri, M. Montazeri, K. Bahaadinbeigy, M. Montazeri, A. Afraz, Application of machine learning methods in predicting schizophrenia and bipolar disorders: a systematic review, Heal. Sci. Rep. 6 (2023) e962.
[23] M. Bracher-Smith, E. Rees, G. Menzies, J.T. Walters, M.C. O'Donovan, M.J. Owen, V. Escott-Price, Machine learning for prediction of schizophrenia using genetic and demographic factors in the UK biobank, Schizo. Res. 246 (2022) 156–164.
[24] P.L. Ballester, M.T. Romano, T. de Azevedo Cardoso, S. Hassel, S.C. Strother, S.H. Kennedy, B.N. Frey, Brain age in mood and psychotic disorders: a systematic review and meta-analysis, Acta Psychiatr. Scand. 145 (2022) 42–55.
[25] G. Koppe, A. Meyer-Lindenberg, D. Durstewitz, Deep learning for small and big data in psychiatry, Neuropsychopharmacology 46 (2021) 176–190, https://doi.org/10.1038/s41386-020-0767-z.
[26] D. Sadeghi, A. Shoeibi, N. Ghassemi, P. Moridian, A. Khadem, R. Alizadehsani, U.R. Acharya, An overview of artificial intelligence techniques for diagnosis of Schizophrenia based on magnetic resonance imaging modalities: methods, challenges, and future works, Comput. Biol. Med. 146 (2022), 105554.
[27] J. Oh, B.L. Oh, K.U. Lee, J.H. Chae, K. Yun, Identifying schizophrenia using structural MRI with a deep learning algorithm, Front. Psych. 11 (2020) 16, https://doi.org/10.3389/fpsyt.2020.00016.
[28] R.J. Janssen, J. Mourão-Miranda, H.G. Schnack, Making individual prognoses in psychiatry using neuroimaging and machine learning, Biol. Psych. Cogn. Neurosci. Neuro. 3 (2018) 798–808, https://doi.org/10.1016/j.bpsc.2018.04.004.
[29] S. Davis, J. Zhang, I. Lee, M. Rezaei, R. Greiner, F.A. McAlister, R. Padwal, Effective hospital readmission prediction models using machine-learned features, BMC Heal. Serv. Res. 22 (2022) 1415.
[30] S. Kwakernaak, K. van Mens, W. Cahn, R. Janssen, Group Investigators. Using machine learning to predict mental healthcare consumption in non-affective psychosis, Schizo. Res. 218 (2020) 166–172, https://doi.org/10.1016/j.schres.2020.01.008.
[31] D. Stamate, A. Katrinecz, D. Stahl, S.J. Verhagen, P.A. Delespaul, J. van Os, S. Guloksuz, Identifying psychosis spectrum disorder from experience sampling data using machine learning approaches, Schizo. Res. 209 (2019) 156–163, https://doi.org/10.1016/j.schres.2019.04.028.
[32] J. Benoit, H. Onyeaka, M. Keshavan, J. Torous, Systematic review of digital phenotyping and machine learning in psychosis spectrum illnesses, Harvard Rev. Psych. 28 (2020) 296–304, https://doi.org/10.1097/HRP.0000000000000268.
[33] G. Starke, E. De Clercq, S. Borgwardt, B.S. Elger, Computing schizophrenia: ethical challenges for machine learning in psychiatry, Psychol. Med. 51 (2021) 2515–2521, https://doi.org/10.1017/S0033291720001683.
[34] S. Vieira, Q.Y. Gong, W.H. Pinaya, C. Scarpazza, S. Tognin, B. Crespo-Facorro, A. Mechelli, Using machine learning and structural neuroimaging to detect first episode psychosis: reconsidering the evidence, Schizo. Bull. 46 (2020) 17–26, https://doi.org/10.1093/schbul/sby189.
[35] N. Koutsouleris, D.B. Dwyer, F. Degenhardt, C. Maj, M.F. Urquijo-Castro, R. Sanfelici, PRONIA Consortium, Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression, JAMA Psychiatr. 78 (2021) 195–209, https://doi.org/10.1001/jamapsychiatry.2020.3604.
[36] R. Jacobucci, K.J. Grimm, Machine learning and psychological research: the unexplored effect of measurement, Perspect. Psychol. Sci. 15 (2020) 809–816.
[37] S.O. Folorunso, S.G. Fashoto, J. Olaomi, O.Y. Fashoto, A multi-label learning model for psychotic diseases in Nigeria, Inform. Med. Unlocked 19 (2020), 100326, https://doi.org/10.1016/j.imu.2020.100326.
[38] I. Elujide, S.G. Fashoto, B. Fashoto, E. Mbunge, S.O. Folorunso, J.O. Olamijuwon, Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases, Inform. Med. Unlocked 23 (2021), 100545, https://doi.org/10.1016/j.imu.2021.100545.
[39] F. Ge, L.W. Di Zhang, H. Mu, Predicting psychological state among Chinese undergraduate students in the COVID-19 epidemic: a longitudinal study using a machine learning, Neuropsychiatric Dis. Treat. 16 (2020) 2111–2118, https://doi.org/10.2147/NDT.S262004.

[40] H. Salem, A. Ruiz, S. Hernandez, K. Wahid, F. Cao, B. Karnes, S. Beasley, M. Sanches, E. Ashtari, T. Pigott, Borderline personality features in inpatients with bipolar disorder: impact on course and machine learning model use to predict rapid readmission, J. Psychiatr. Pract. 25 (2019) 279–289, https://doi.org/10.1097/PRA.0000000000000392.
[41] A. Bayes, M.J. Spoelma, D. Hadzi-Pavlovic, G. Parker, Differentiation of bipolar disorder versus borderline personality disorder: a machine learning approach, J. Affect. Disord. 288 (2021) 68–731, https://doi.org/10.1016/j.jad.2021.03.082.
[42] J. Edgcomb, T. Shaddox, G. Hellemann, J.O. Brooks, High-risk phenotypes of early psychiatric readmission in bipolar disorder with comorbid medical illness, Psychosomatics 60 (2019) 563–573, https://doi.org/10.1016/j.psym.2019.05.002.
[43] A. Kushki, E. Anagnostou, C. Hammill, P. Duez, J. Brian, A. Iaboni, R. Schachar, R. Crosbie, P. Arnold, J.P. Lerch, Examining overlap and homogeneity in ASD, ADHD, and OCD: a data-driven, diagnosis-agnostic approach, Transl. Psychiatry 9 (2019) 318, https://doi.org/10.1038/s41398-019-0631-2.
[44] K. Srinivasan, N. Mahendran, D.R. Vincent, C.Y. Chang, S. Syed-Abdul, Realizing an integrated multistage support vector machine model for augmented recognition of unipolar depression, Electronics 9 (2020) 647, https://doi.org/10.3390/electronics9040647.
[45] D. Morel, K.C. Yu, A. Liu-Ferrara, A.J. Caceres-Suriel, S.G. Kurtz, Y.P. Tabak, Predicting hospital readmission in patients with mental or substance use disorders: a machine learning approach, Int. J. Med. Inform. 139 (2020), 104136, https://doi.org/10.1016/j.ijmedinf.2020.104136.
[46] M. Guo, Y. Yu, T. Wen, X. Zhang, B. Liu, J. Zhang, R. Zhang, Y. Zhang, X. Zhou, Analysis of disease comorbidity patterns in a large-scale China population, BMC Med. Genom. 12 (2019) 177, https://doi.org/10.1186/s12920-019-0629-x.
[47] C. Li, D.A. Gheorghe, J.E. Gallacher, S. Bauermeister, Psychiatric comorbid disorders of cognition: a machine learning approach using 1175 UK Biobank participants, Evid. Based Ment. Health. 23 (2020) 140–145, https://doi.org/10.1136/ebmental-2020-300147.
[48] M. Kapadia, M. Desai, R. Parikh, Fractures in the framework: limitations of classification systems in psychiatry, Dialog. Clin. Neurosci. 22 (2020) 17–26, https://doi.org/10.31887/DCNS.2020.22.1/rparikh.
[49] A.O. Adejumo, N.A. Ikoba, E.A. Suleiman, H.I. Okagbue, P.E. Oguntunde, O.A. Odetunmibi, O. Job, Quantitative exploration of factors influencing psychotic disorder ailments in Nigeria, Data Brief 14 (2017) 175–185, https://doi.org/10.1016/j.dib.2017.07.046.
[50] A. Tharwat, Classification assessment methods, Appl. Comp. Inform. 17 (2021) 168–192, https://doi.org/10.1016/j.aci.2018.08.003.
[51] D. Goštautaitė, L. Sakalauskas, Multi-label classification and explanation methods for students' learning style prediction and interpretation, Appl. Sci. 12 (2022) 5396.
[52] C. Li, L. Sun, D. Peng, S. Subramani, S.C. Nicolas, A multi-label classification system for anomaly classification in electrocardiogram, Heal, Inform. Sci. Syst. 10 (2022) 19.
[53] S. Singh, A. Majumdar, Non-intrusive load monitoring via multi-label sparse representation-based classification, IEEE Trans. Smart Grid 11 (2019) 1799–1801.
[54] A.N. Tarekegn, M. Giacobini, K. Michalak, A review of methods for imbalanced multi-label classification, Pattern Recogn. 118 (2021), 107965.
[55] T. Burgert, M. Ravanbakhsh, B. Demir, On the effects of different types of label noise in multi-label remote sensing image classification, IEEE Trans. Geosci. Rem. Sens. 60 (2022) 1–13.
[56] Z. Zhou, G. Hooker, Unbiased measurement of feature importance in tree-based methods, ACM Trans. Know. Disc. Data 15 (2021) 1–21, https://doi.org/10.1145/3429445.
[57] F. Curia, Cervical cancer risk prediction with robust ensemble and explainable black boxes method, Health Technol. 11 (2021) 875–885, https://doi.org/10.1007/s12553-021-00554-6.
[58] O.E. Amoran, T.O. Lawoyin, O.O. Oni, Risk factors associated with mental illness in Oyo State, Nigeria: a Community based study, Ann. Gen. Psych. 4 (2005) 1–6, https://doi.org/10.1186/1744-859X-4-19.
[59] V.O. Lasebikan, A. Ejidokun, O.A. Coker, Prevalence of mental disorders and profile of disablement among primary health care service users in Lagos Island, Epidemiol. Res. Int. 2012 (2012), 357348, https://doi.org/10.1155/2012/357348.
[60] M.J. Taylor, J. Martin, Y. Lu, I. Brikell, S. Lundström, H. Larsson, P. Lichtenstein, Association of genetic risk factors for psychiatric disorders and traits of these disorders in a Swedish population twin sample, JAMA Psychiatr. 76 (2019) 280–289.
[61] L. Stone, E. Waldron, H. Nowak, Making a good mental health diagnosis: science, art and ethics, Austr. J. Gen. Pract. 49 (2020) 797–802, https://doi.org/10.31128/AJGP-08-20-5606.
[62] H.I. Okagbue, P.I. Adamu, P.E. Oguntunde, E.C.M. Obasi, O.A. Odetunmibi, Machine learning prediction of breast cancer survival using age, sex, length of stay, mode of diagnosis and location of cancer, Health Technol. 11 (2021) 887–893, https://doi.org/10.1007/s12553-021-00572-4.
[63] H.I. Okagbue, P.E. Oguntunde, E.C.M. Obasi, P.I. Adamu, A.A. Opanuga, Diagnosing malaria from some symptoms: a machine learning approach and public health implications, Health Technol. 11 (2021) 23–37, https://doi.org/10.1007/s12553-020-00488-5.
[64] B. Zhang, J. Peng, H. Chen, W. Hu, Machine learning for detecting Wilson's disease by amplitude of low-frequency fluctuation, Heliyon 9 (2023), e18087.
[65] P.N. Ogbonna, P.N. Iheanacho, N.P. Ogbonnaya, C.J. Mbadugha, I. Ndubuisi, P.C. Chikeme, Prevalence of mental illness among adolescents (15–18 years) treated at Federal Neurospsychiatric Hospital, Enugu Nigeria, from 2004 to 2013, Arch. Psych. Nurs. 34 (2020) 7–13, https://doi.org/10.1016/j.apnu.2019.12.008.
[66] A.G. Harvey, D.A. Schmidt, A. Scarnà, C.N. Semler, G.M. Goodwin, Sleep-related functioning in euthymic patients with bipolar disorder, patients with insomnia, and subjects without sleep problems, Amer. J. Psych. 162 (2005) 50–57, https://doi.org/10.1176/appi.ajp.162.1.50.
[67] K.A. Kaplan, A.G. Harvey, Behavioral treatment of insomnia in bipolar disorder, Amer. J. Psych. 170 (2013) 716–720, https://doi.org/10.1176/appi.ajp.2013.12050708.
[68] L. Palagini, M. Miniati, D. Marazziti, V. Sharma, D. Riemann, Association among early life stress, mood features, hopelessness and suicidal risk in bipolar disorder: the potential contribution of insomnia symptoms, J. Psych. Res. 135 (2021) 52–59, https://doi.org/10.1016/j.jpsychires.2020.12.069.
[69] S. Yun, M. Maxfield, Correlates of dementia-related anxiety: self-perceived dementia risk and ageism, Educ. Gerontol. 46 (2020) 563–574, https://doi.org/10.1080/03601277.2020.1790103.
[70] O. Agofure, O.R. Okandeji-Barry, I.S. Ume, Knowledge and Perception of Mental Disorders among relatives of mentally ill persons in a rural community in South-South Nigeria, J. Com. Med. Prim. Heal. Care. 31 (2019) 66–77.
[71] H. Scott, The importance of spirituality for people living with dementia, Nurs. Stand. 30 (2016) 41–50, https://doi.org/10.7748/ns.30.25.41.s47.
[72] S.B. Guessoum, J. Lachal, R. Radjack, E. Carretier, S. Minassian, L. Benoit, M.R. Moro, Adolescent psychiatric disorders during the COVID-19 pandemic and lockdown, Psych. Res. 291 (2020), 113264.