Healthcare Analytics 3 (2023) 100202

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

Development and performance analysis of machine learning methods for predicting depression among menopausal women

Md. Mamun Ali [a], Hussein Ali A. Algashamy [b], Enas Alzidi [c], Kawsar Ahmed [d,e,*], Francis M. Bui [d], Shobhit K. Patel [f], Sami Azam [g], Lway Faisal Abdulrazak [h], Mohammad Ali Moni [i,**]

[a] Department of Software Engineering (SWE), Daffodil International University (DIU), Daffodil Smart City (DSC), Birulia, Savar, Dhaka 1216, Bangladesh
[b] Medical laboratory techniques, Al-Rafidain University College, Baghdad, Iraq
[c] Computer Technologies Engineering, Al-Turath University College, Baghdad, Iraq
[d] Department of Electrical and Computer Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK S7N 5A9, Canada
[e] Group of Biophotomatiχ, Department of Information and Communication Technology, Mawlana Bhashani Science and Technology University, Santosh, Tangail 1902, Bangladesh
[f] Department of Computer Engineering, Marwadi University, Rajkot, Gujarat, 360003, India
[g] College of Engineering, IT and Environment, Charles Darwin University, Casuarina, NT 0909, Australia
[h] Department of Computer Science, Cihan University Sulaimaniya, Sulaimaniya 46001, Kurdistan Region, Iraq
[i] Artificial Intelligence & Digital Health, School of Health and Rehabilitation Sciences, Faculty of Health and Behavioural Sciences, The University of Queensland, St Lucia, QLD 4072, Australia

ARTICLE INFO

Keywords:
Artificial intelligence
Depression
Machine learning
Menopause
Quality of life
Random Forest

ABSTRACT

Menopause is an obligatory phenomenon in a woman's life. Some women face mental and physical issues during their menopausal period. Depression is one of the issues some women struggle with during their menopausal period. The scarcity of specialists, lack of knowledge, and awareness is the motivating factor in this research to predict depression among menopausal women and enhance their quality of life. The prediction of depression symptoms among menopausal women with machine learning techniques is promising and challenging in artificial intelligence. This study develops a system with significant accuracy using a supervised machine-learning approach. Various classification algorithms are used to determine the best-performing classifier by evaluating multiple parameters, including accuracy, sensitivity, specificity, precision, recall, F-Measure, Receiver Operating Characteristic (ROC), Precision–Recall Curve (PRC), and Area Under the Curve (AUC). We found that Random Forest and XGBoost classifiers are the performers with 99.04% accuracy employing the 14 most significant features.

1. Introduction

Menopause is a natural phenomenon for every woman [1]. Menopause is considered a period of a woman's life, when her natural menstrual gets stopped, most possibly, ovaries will cease to produce estrogen and progesterone [2]. The lack of a 12-month continuous period is the key criterion for menopause [3]. A woman becomes disabled to get pregnant when it takes place. The range of age is 45 to 55 years old and the average age for menopause is 50 years according to China and Western countries [4–6]. But perspective to Bangladesh, the mean age of menopausal women is 45 years with the range of natural menopausal period being 40 to 50 years [7]. A woman loses almost one-third of her life in menopause with the heightening of her lifespan [8].

Menopause has different phases and stages. According to research conducted in Bangladesh on menopausal women, there are three kinds of menopausal phases [9]. One is "premenopausal women" another one is "postmenopausal women", and another one is perimenopausal premenopausal. Those women's menstrual bleeding was regular for the last 12 months were considered as premenopausal women and those women who did not face menstrual bleeding in the last year were listed as postmenopausal women [10]. The women whose age range is 40 to 50 years old are listed as natural menopausal women [7]. Those women who faced irregular menstrual periods within 12 months or last menses that happened 3 months earlier are known as premenopausal [3]. The result of menopause is a product of social, physical, and psychological transition, which causes depression [11,12]. Every female with

* Corresponding author at: Department of Electrical and Computer Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK S7N 5A9, Canada.
** Corresponding author.
E-mail addresses: mamun35-274@diu.edu.bd (M.M. Ali), husseinali1@ruc.edu.iq (H.A.A. Algashamy), inas.fadel@turath.edu.iq (E. Alzidi), kawsar.ict@mbstu.ac.bd, k.ahmed.bd@ieee.org, k.ahmed@usask.ca (K. Ahmed), francis.bui@usask.ca (F.M. Bui), shobhitkumar.patel@marwadieducation.edu.in (S.K. Patel), sami.azam@cdu.edu.au (S. Azam), lway.faisal@sulicihan.edu.krd (L.F. Abdulrazak), m.moni@uq.edu.au (M.A. Moni).

https://doi.org/10.1016/j.health.2023.100202
Received 23 February 2023; Received in revised form 14 May 2023; Accepted 25 May 2023

menopause confronts multiple mental and also physical ailments [3,8]. Depression is one of the major mental issues among them, which causes physical illness later.

Depression is a kind of mood disorder, which causes major disability among women [13]. The depression rate among women is almost double of men, which is almost one-fifth of total women [14]. Major depression is increasing in the world and it is expected to be the major cause of disease burdens by 2030. It is currently one of the main causes of depression for women worldwide [15]. The incidence of depressive syndrome among women who are going through menopause is 5.9%–23.8% [16–18]. According to research, the chance of depressed mood is 2–3 times over in perimenopausal women compared to premenopausal women [19–21]. Each year, almost 73 million adult women experience major depression [22].

Since menopause causes depression and depression results in mental and physical ailments. It is one of the major concerns for menopausal women. Due to a lack of awareness, and education, most men do not take it into account, which causes them to see mental and physical health issues. It is Menopausal women can get depression in early stage if they get aware and can detect in the early stage that she is depressed due to menopause. Besides, it is also seen that sometimes psychiatrists, psychologists, and clinicians are not sure whether a patient is depressed or not due to less severity of some symptoms. So, a model or system will make it easy to diagnose depression among menopausal women. From that perspective, we have conducted the study to find a machine learning-based model for the initial diagnosis of depression among menopausal women.

2. Background

Prediction of depression among menopausal women is a complex and complicated task in the medical field. But very few studies have been carried out to predict depression among women going through menopause. Menopausal signs were not examined extensively in Southeast Asia as were in Western countries [23]. The literature on menopausal is very low to develop countries like Bangladesh. X. Li et al. in 2015 develop a system implementing artificial neural networks to predict menopausal symptoms [24]. They measured the sharpness of menopausal syndromes by the Kupperman Menopausal Index (KMI). This study is focused on the severity of the symptoms. But depression is one of the major symptoms, which causes different psychological or physical disorders. It does not predict the depression or risk of depressive symptoms specifically.

In 2017, Zheng et al. proposed a scoring system based on symptoms such as hot flashes and sweating among menopausal women to predict the risk of depressive syndrome [13]. In this study, they considered two symptoms mainly to predict depression. But Mamun Ibn Bashar et al. showed that eight factors are significantly correlated to depression among pre- and post-menopausal women, which includes both physical and mental symptoms [3]. There are more symptoms related to depressive disorder. Menopausal women can experience different forms of symptoms. In 2011, Chuni and Sreeramareddy showed that menopausal signs have an incidence of more than 50 percent [25]. J. T. Bromberger et al. in 2015 showed that quite a significant risk component for middle-aged women is severe depressive disorder [26]. Juang et al. in 2015 stated that in East Asian cultures, women who are post- and perimenopausal with hot sweats are associated with depression and anxiety syndromes [27].

In 2022, Handing et al. conducted a regional study to find out the predictors of depression among middle-aged and older citizens of Europe [28]. Their findings showed that menopause is one of the key factors for depression among menopausal women. According to some other recent studies, 67% of menopausal women believe that menopause is responsible for obesity, and 65% of menopausal women experience lifestyle changes that manage menopause [29–32]. 44% of them had mild hot flashes and sweating episodes, 23% of women experienced severe symptoms, and 45% of women had mild sleep issues [33–35]. 36% were suffering from mild depression, and 30% had no symptoms. 29% of women had osteoporosis, 46% of women have joint pain 25% of women made diet dietary modifications during menopause [36–39].

Since a very limited number of researches have been conducted to predict depression among menopausal women, predicting the risk factors is not clear and specific. Besides, the discussion mentioned above indicates that depression prediction among menopausal women is very crucial. From that perspective, the study is conducted to predict depression among menopausal women.

3. Materials & Methods

3.1. Data and sample design

A survey was conducted to collect samples among rural women in Tangail, the second-largest district by the population under the Dhaka division in Bangladesh. Besides, it is the largest and one of the central cities in the Dhaka division. A total of 346 samples were assembled from individual women through interviews, which were performed face to face, based on the standard questionnaires. The interview was held between a well-trained interviewer and the participant. The dataset belongs to 25 attributes, which are mentioned in Table 1 in detail. The questionnaires of the dataset were defined following two studies conducted by Ahmed and et al. in 2016 and Bashar et al. in 2017 [3,7]. They showed that these attributes are highly associated with depression and quality of life among menopausal women. Based on these questionnaires, the attributes of the dataset are designed. The questionnaires are represented in Table 2.

3.2. Methods

Fig. 1 demonstrates the methodology of this research work. The entire methods of the study are described as follows sequence step by step. It provides a clear understanding of the working flowchart.

3.2.1. Preprocessing

A classifier is inefficient in some cases in processing the raw data due to some features [40]. Raw data is sometimes incomplete and noisy. Sometimes, it is inconsistent also [41,42]. Raw data is very sensitive to inconsistency or noise, which affects the outcome of the analysis progress badly. So, data preprocessing is very important for data mining [43,44]. Data preprocessing is associated to clean noisy and inconsistent data, handling missing values, reduction of dimensionality, attribute selection, etc. [45,46]. So, the preprocessing phase is very important to prepare a dataset for making a better prediction in data mining and machine learning.

3.2.2. Imputation of missing values

At first, the dataset is loaded into Weka (Data Mining Tools, Version 3.8.3). Then, it is observed that the dataset contains some missing values. So, a filter called ReplaceMissingValues under the attribute of the unsupervised filter was applied to replace missing values. The filter replaces missing values using mean or mode [47]. The mean is used in terms of numeric values. On the other hand, the mode is used to replace nominal values. It is the most popular and used filter to replace missing values [48].

3.2.3. Outlier detection and removal

Outlier is indeed a piece of data that does not represent regular behavior as other data point shows [49]. Outlier influences the Machine Learning or Data Mining Process for forecasting results and discovering efficient approaches to relevant issues [50]. So, making the dataset free of outliers and extreme value is an essential task to get superior

Table 1
Explanation of dataset.

Attribute name	Description	Data type
age	Age of the participant	Numeric
menopause	Menopausal Status (0 = No, 1 = Yes)	Binary
peww	Physical Exercise Without Work (0 = No, 1 = Yes)	Binary
hf	Hot flashes (0 = No, 1 = Yes)	Binary
ost	Osteoporosis (0 = No, 1 = Yes)	Binary
diabet	Type 2 Diabetes (0 = No, 1 = Yes)	Binary
chd	Coronary Heart Disease (0 = No, 1 = Yes)	Binary
heart_beat	Heart Beating Quickly (0 = No, 1 = Yes)	Binary
tense	Feeling Tensed or Nervous (0 = No, 1 = Yes)	Binary
sleep	Sleeping problem (0 = No, 1 = Yes)	Binary
excitable	Excitable (0 = No, 1 = Yes)	Binary
concentration	Difficulty in Concentrating (0 = No, 1 = Yes)	Binary
tired	Feeling Tired or Lacking in Energy (0 = No, 1 = Yes)	Binary
sweat	Sweat at Night (0 = No, 1 = Yes)	Binary
menstruation	Physical Problem During Menstruation (0 = No, 1 = Yes)	Binary
irritability	Irritability (0 = No, 1 = Yes)	
pressurehead	Pressure or Tightness in Head (0 = No, 1 = Yes)	Binary
tingling	Tingling (0 = No, 1 = Yes)	Binary
headaches	Headaches (0 = No, 1 = Yes)	Binary
pain	Muscle and joint pain (0 = No, 1 = Yes)	Binary
breath	Breathing Difficulties (0 = No, 1 = Yes)	Binary
knowledge	Knowledge about Menopause (0 = No, 1 = Yes)	Binary
pilknowledge	Knowledge about Pill (0 = No, 1 = Yes)	Binary
agediff	Age difference from their husband	Numeric
depression	Depression present or not (1 = Depressed, 0 = Not Depressed)	Binary

Table 2
Explanation of question.

Sl. No.	Question	Attributes used in dataset	Answer
1	Name of the participant	Not used in the dataset	
2	Age of the participant	age	Age in number
3	Current menopausal status of the participant	menopause	(0 = No, 1 = Yes)
4	Does the participant do physical exercise without work?	peww	(0 = No, 1 = Yes)
5	Does the participant experience hot flashes? It means a sudden feeling of heat that seems to come from nowhere and spreads throughout the body.	hf	(0 = No, 1 = Yes)
6	Does the participant experience osteoporosis?	ost	(0 = No, 1 = Yes)
7	Does the participant have diabetes or not?	diabet	(0 = No, 1 = Yes)
8	Does the participant have coronary heart disease or not?	chd	(0 = No, 1 = Yes)
9	Does the participant experience heart beat quickly?	heart_beat	(0 = No, 1 = Yes)
10	Does the participant fee tense or nervous?	tense	(0 = No, 1 = Yes)
11	Does the participant experience sleep problem?	sleep	(0 = No, 1 = Yes)
12	Does the participant feel excitable?	excitable	(0 = No, 1 = Yes)
13	Does the participant experience difficulty in Concentrating?	concentration	(0 = No, 1 = Yes)
14	Does the participant feel tired or lacking in energy?	tired	(0 = No, 1 = Yes)
15	Does the participant experience sweat at night?	sweat	(0 = No, 1 = Yes)
16	Does the participant experience physical problem during menstruation?	menstruation	(0 = No, 1 = Yes)
17	Does the participant experience irritability?	irritability	(0 = No, 1 = Yes)
18	Does the participant feel pressure or tightness in head?	pressurehead	(0 = No, 1 = Yes)
19	Does the participant experience tingling?	tingling	(0 = No, 1 = Yes)
20	Does the participant experience headaches?	headaches	(0 = No, 1 = Yes)
21	Does the participant experience muscle and joint pain?	pain	(0 = No, 1 = Yes)
22	Does the participant experience any breathing difficulties?	breath	(0 = No, 1 = Yes)
23	Does the participant have knowledge about menopause?	knowledge	(0 = No, 1 = Yes)
24	Does the participant have knowledge about pill of menopause?	pilknowledge	(0 = No, 1 = Yes)
25	Age difference from their husband	agediff	Number
26	The participant is depressed or not.	depression	(0 = No, 1 = Yes)

prediction output in the Machine Learning approach. There is a filter in Weka known as InterQuartileRange to detect outlier and extreme values. Outlier and extreme values can be found by applying the same filter. To detect outliers, the dataset is classified into three quartiles. The first quartile is represented by Q_1, the second quartile le Q_2, and the third quartile by Q_3. The value of Inter Quartile Range (*IQR)* is measured following the equation $IQR = Q_3–Q_1$. Then, B_{min}, lower boundary, and B_{max} upper boundary were calculated by implementing the following equations [51]:

$$B_{min}= Q_1-1.5 * IQR \tag{1}$$

$$B_{max}= Q_3+1.5 * IQR \tag{2}$$

Here, the instance is known as an outlier, which shows a value lower than B_{min} and greater than B_{max}. A value in a dataset is considered an extreme value, which value is either very small or very large. After removing outlier and extreme value, tune-up of data from either positive or negative case be moved more than the other groups. It causes an imbalance in the class attributes, which is responsible for poor accuracy. So, it is an essential task to make the dataset balanced if the dataset is imbalanced. A filter called Synthetic Minority Oversampling Technique (SMOTE) is performed to balance the imbalanced dataset.

3.2.4. Decision tree

The decision tree is one of the pioneer classification algorithms in the machine learning and data science field. It makes a tree perform

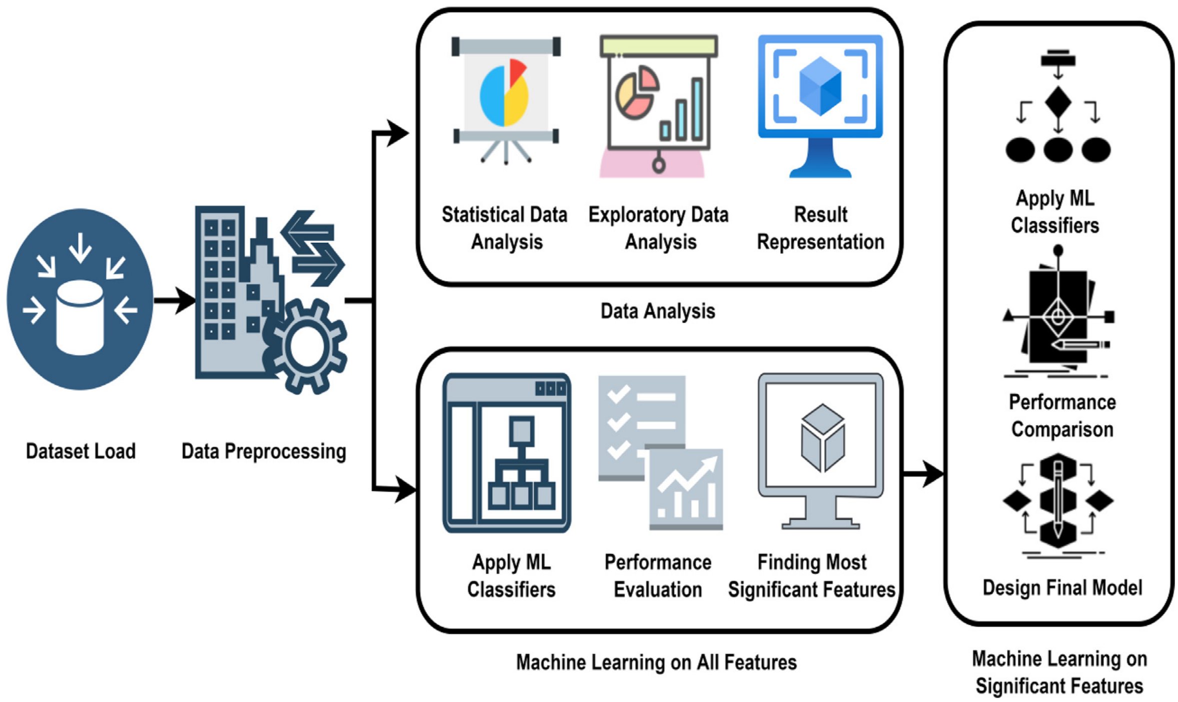

Fig. 1. Pipeline of the research methodology.

classification where a root node and a number leaf nodes are found. At first, the value of the Gini index or entropy is calculated. Both Gini and entropy are measures of the impurity of a node. The value of Gini and entropy are calculated by the following equations [51]:

$$Gini = 1 - \sum_{i=1}^{n} P^2 \left(C_i\right) \tag{3}$$

$$Entropy = \sum_{i=1}^{n} -p\left(c_i\right) \log_2 \left(p\left(c_i\right)\right) \tag{4}$$

where $p\left(c_i\right)$ is the possibility of class c_i in a node. The information gain *(IG)* value is calculated to split a node. *IG* measures the reduction in entropy. Higher *IG* indicates a lower entropy. The *IG* is found using the following equation [40]:

$$IG\left(S, A\right) = H\left(S\right) - H\left(S|A\right) \tag{5}$$

Here *IG (S, A)* refers to the information for the dataset *S* for the variable *A* for a random variable, *H(S)* is the calculated *entropy* for the dataset and *H (S | A)* is the conditional entropy for the dataset given the variable *A*. Based on these parameters, a tree is built to decide the category of an instance. Thus, a decision tree works and classifies. It is one of the most efficient and well-performed classification algorithms.

3.2.5. Random forest

Random Forest (RF) is a type of ensemble classification algorithm that derives from many decision trees [45,52]. Every tree relies on the measure of a random vector, which is sampled freely and by the equivalent distribution for whole trees that exist in the forest [53]. RF is a machine-learning supervised algorithm, which is executed to classify and predict the actual result based on previous data (training data). It is taken into account as the best performer classifier because of its massive number of trees that exist in the forest which provides better outcomes compared to decision trees [54]. Basically, at the training stage, each of these trees is trained freely and the result of the training of all of these trees is assembled by calculating the average [45]. A class attribute value is predicted by every tree that exists in the forest for each data at the testing stage. When a class attribute value is predicted for every tree, then the final result for every single test data is decided based on majority voting [55,56]. Which classes attribute value achieves the votes from the majority that attribute value is considered as the final predicted value and assigned to the test data [57].

The algorithm works as follows [54]:

Step 1: Randomly choose x attributes from total y attributes, where $x \ll y$.

Step 2: Encompassed by the x attributes, compute the node "k" dealing with the ideal intersect point.

Step 3: Divide the parent node into child nodes dealing with the ideal split.

Step 4: Follow again 1 to 3 steps while l number of nodes is created.

Step 5: Shape the forest by redoing steps 1 to 4 to create t number of trees following t number of times.

At first, x attributes are selected among the total y attributes. In the second step, in every tree randomly choose x attributes to trace the parent node exploiting the ideal split approach. The next step computes the child node by the same method for the dataset. Correspondingly, the trees are constructed from the parent node, while the entire child nodes are brought out from the attributes. This randomly formed tree builds the random forest, which is applied to predict the unknown results. The algorithm is exploited as a regression and classification method [28]. It provides better performance results and can handle noise, extreme value, and an outlier in the data. It is implemented in this study because it has less possibility of over-fitting and it has previously provided better classification accuracy [53,56,58–60]. Besides, RF is considered the most preferable algorithm in maximum clinical research [61,62].

3.2.6. XGBoost

XGBoost is an abbreviation for Extreme Gradient Boosting. It is a sort of supervised ensemble machine-learning method. The method is primarily built on the decision tree notion, which is based on a gradient-boosting framework [23]. It is used to address issues including regression, classification, ranking, and user-defined prediction. It is a fully modular tree-boosting algorithm that is widely used in machine learning. XGBoost offers a significant capacity to deliver solutions to real-world situations with minimal resources. XGBoost is expected to be similar to tree boosting (also known as GBDT, GBM) to quickly and accurately solve a wide range of machine learning issues [24]. XGBoost has become one of the most effective classification algorithms in the field of machine learning and data science in recent years.

3.2.7. Adaboost classifier

AdaBoost, also known as Adaptive Boosting, is a boosting technique-based machine learning approach that is utilized as an ensemble method. It is named Adaptive Boosting because the weights are added to each instance, with larger weights applied to mistakenly identify instances. While classifying, it creates a model by assigning equivalent weights to all data objects. It, therefore, gives more weight to points that were incorrectly recognized. In the following model, all of the points with higher weights are given more attention. This implies that each consecutive model will receive a weighted input. It will continue to train models until a lower error occurs. All of these separate models are insufficient to accurately identify the data objects and are commonly referred to be weak learners. After incorporating all of these weak learners, the final model is able to properly classify all of the data objects. This final model is referred to as a strong learner. This strong learner is known as an Adaboost classifier.

3.2.8. Performance analysis

There are various types of classifiers available in machine learning. All of these classifiers are exploited to build specific machine learning-based systems. All classifiers have their own implementation. The efficiency of a machine learning-based model depends on a suitable classifier. Every classifier has some pros and cons. The accuracy of the classifiers varies based on methodology, data types, and dataset. Every classifier provides different accuracy in different methods and datasets. It is a very important task in machine learning to find out a suitable classifier for a specific model. The suitable classifier is chosen based on the performance of that particular classification algorithm. Various characteristics and parameters are taken into consideration to evaluate the performance of a classifier. The rest of the section will describe the characteristics and variables used to assess a classifier and model.

3.3. Evaluation metric by class

Different classifiers were applied to find out the best performer classifier for the proposed system. Every classifier produces a confusion matrix. The confusion matrix represents the number of correctly and incorrectly classified instances. Based on this confusion matrix, the different statistical result is calculated.

Based on the result of different classifiers' confusion matrices, different measures were considered to compare and evaluate the proposed classifier from other algorithms. These measures are mentioned as follows with the calculation formula [63,64]:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{6}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{7}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

$$\text{True Positive Rate} = \frac{TP}{TP+FN} \tag{9}$$

$$\text{False Positive Rate} = \frac{FP}{FP+TN} \tag{10}$$

True Positive, represented by *TP*, is a positive case, which is correctly identified by the classification algorithm [65]. If the prediction result by the classifier is "Yes" and the actual result is "Yes" then it is categorized into *TP* [66]. True Negative, represented by *TN*, is a negative case, which is identified correctly [65]. False Positive, demonstrated by *FP*, is a negative case, which is incorrectly identified as a positive case [65]. If the actual result is "No" but the predicted result by the classifier is "Yes" is categorized into *FP* [54]. On the other hand, False Negative, represented by *FN*, is a positive case but miscategorized into negative cases [65]. The percentage of correctly classified cases is said to be accurate [64].

For more precise results, three more parameters are brought into consideration such as precision, recall, F-Measure, and Matthews Correlation Coefficient (MCC). The described equations are applied to get this variable's result [67,68].

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{12}$$

$$\text{F-Measure} = \frac{2*precision*recall}{precision+recall} \tag{13}$$

$$MCC = \frac{TP*TN - FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{14}$$

Precision is a component of the positive cases between the collected instances [54]. The Recall is a tiny portion of the correct instances that have been well recovered over the overall number of application instances. The F-Measure (or f-score) is known to be founded on the two-fold precision–recall times divided by the amount of precision as well as recall [68]. It is a correlation coefficient between actual and predicted values binary classifications, taking values from −1 representing inverse forecasting (total difference of opinion) to 1 representing perfect prediction (full agreement) [69]. It is considered to be a balanced metric which can be used even in scenarios of imbalanced data. Nowadays, the MCC is widely used in scientific research [70–72].

3.4. Area under the ROC curve (AUROC)

The Receiver Operating Characteristic (ROC) curve is a visual the area under the ROC curve (AUROC) has been a well-known measure of the performance of the classification algorithm [73–75]. The AUC is a combined indicator of sensitivity (TP/(TP+FN)) and specificity (TN/(TN+FP)) over the broad spectrum of potential threshold values [76]. It refers to the possibility that the classification algorithm will be randomized to rank chosen positive instances stronger than the randomly picked negative cases [77]. The AUC of a classification algorithm is estimated by the following equation [78].

$$R = \frac{k - p(p+1)/2}{p*n} \tag{15}$$

where p and n refer to the number of positive and negative cases respectively. Besides, $k = \sum r_i$ where r_i is used to represent the rank of ith positive cases in the list of rank.

Some prior studies have shown how AUC does better than other measurement metrics [79–81]. Most importantly, AUC is a statistically better performer than accuracy [80].

3.5. Precision–Recall Curve (PRC curve)

The Precision–Recall curve refers to a graphical view of precision and recall, where the x-axis represents recall (sensitivity) and y-axis represents precision (specificity). It is a widely used parameter to evaluate the machine-learning model, especially where data is imbalanced. It is considered as more accurate than the ROC curve when binary classification is performed [81]. It provides a better understanding of the performance of any machine learning system.

4. Results and discussion

In this study, four different supervised machine learning algorithms were applied to the dataset, such as Decision tree, AdaBoost Classifier, XGBoost, and Random Forest Classifier. Different parameters of these algorithms were analyzed and compared to get the best performer classification algorithm among them. The result and performance of all the applied classifiers have been discussed in this section.

Table 3
Statistical representation of all characteristics of all patients in the dataset.

Categorical feature					
Feature	Category	All participants (%) N = 346	Participant's condition		P value
			Normal	Depressed	
menopause					0.539
	Yes	151 (43.64)	33 (21.85)	118 (78.15)	
	No	195 (56.36)	170 (87.18)	25 (12.82)	
peww					<0.001
	Yes	81 (23.41)	61 (75.31)	20 (24.69)	
	No	265 (76.59)	142 (53.58)	123 (46.42)	
hf					
	Yes	208 (60.12)	89 (42.79)	119 (57.21)	<0.001
	No	138 (39.88)	114 (82.61)	24 (17.39)	
ost					<0.001
	Yes	38 (10.98)	13 (34.21)	25 (65.79)	
	No	308 (89.02)	190 (61.69)	118 (38.31)	
diabet					<0.001
	Yes	336 (97.11)	197 (58.63)	139 (41.37)	
	No	10 (2.89)	6 (60)	4 (40)	
chd					<0.001
	Yes	312 (90.17)	169 (54.17)	143 (45.83)	
	No	34 (9.83)	34 (100)	0 (0)	
heart_beat					<0.001
	Yes	224 (64.74)	91 (40.625)	133 (59.375)	
	No	122 (35.26)	112 (91.80)	10 (8.20)	
tense					0.072
	Yes	120 (34.68)	40 (33.33)	80 (66.67)	
	No	226 (65.32)	163 (72.12)	63 (27.88)	
sleep					0.018
	Yes	174 (50.29)	83 (47.70)	91 (52.30)	
	No	172 (49.71)	120 (69.77)	52 (30.23)	
excitable					0.009
	Yes	177 (51.16)	81 (45.76)	96 (54.24)	
	No	169 (48.84)	122 (72.19)	47 (27.81)	
concentration					<0.001
	Yes	89 (25.72)	25 (28.09)	64 (71.91)	
	No	257 (74.28)	178 (69.26)	79	
tired					<0.001
	Yes	25 (7.23)	3 (12)	22 (88)	
	No	321 (92.77)	200 (62.31)	121 (37.69)	
sweat					<0.001
	Yes	195 (56.36)	84 (43.08)	111 (56.92)	
	No	151 (43.64)	119 (78.81)	32 (21.19)	
menstruation					<0.001
	Yes	260 (75.14)	127 (48.85)	133 (51.15)	
	No	86 (24.86)	76 (88.37)	10 (11.63)	
irritability					0.004
	Yes	107 (30.92)	67 (62.62)	40 (37.38)	
	No	239 (69.08)	136 (56.90)	103 (43.10)	
pressurehead					0.146
	Yes	162 (46.82)	69 (42.59)	93 (57.41)	
	No	184 (53.18)	134 (72.83)	50 (27.17)	
tingling					<0.001
	Yes	210 (60.69)	84 (40)	126 (60)	
	No	136 (39.31)	119 (87.50)	17 (12.5)	
headaches					0.590
	Yes	150 (43.35)	58 (38.67)	92 (61.33)	
	No	196 (56.65)	145 (73.98)	51 (26.02)	
pain					<0.001
	Yes	95 (27.46)	52 (54.74)	43 (45.26)	
	No	251 (72.54)	151 (60.16)	100 (39.84)	

(*continued on next page*)

Table 3 (*continued*).

Categorical feature					
Feature	Category	All participants (%) N = 346	Participant's condition		P value
			Normal	Depressed	
breath					<0.001
	Yes	288 (83.24)	154 (53.47)	134 (46.53)	
	No	58 (16.76)	49 (84.48)	9 (15.52)	
knowledge					<0.001
	Yes	293 (84.68)	153 (52.22)	140 (47.78)	
	No	53 (15.32)	50 (94.34)	3 (5.66)	
pilknowledge					<0.001
	Yes	89 (25.72)	84 (94.38)	5 (5.62)	
	No	257 (74.28)	119 (46.30)	138 (53.70)	

Numerical feature							
Feature	All patient		Participant's condition				P value
			Normal		Depressed		
	Mean (Median)	IQR (STD)	Mean (Median)	IQR (STD)	Mean (Median)	IQR (STD)	
Age (Years)	44.64 (50)	37.0 (20.83)	54.10 (58)	12 (16.92)	31.19 (23)	1 (18.36)	<0.001
Agediff (Years)	11.14 (10)	2 (2.64)	11.52 (10)	3 (2.98)	10.59 (10)	0 (1.97)	0

Table 4
Performance comparison of applied algorithms.

Algorithm	Accuracy (%)	Precision	Recall	F1 score	Log loss
Decision tree	94.23	0.94	0.94	0.94	1.99
AdaBoost classifier	95.19	0.95	0.95	0.95	1.66
Random forest	95.19	0.95	0.95	0.95	1.66
XGBoost	97.11	0.97	0.97	0.97	0.996

a. Statistical Data Analysis

Table 3 states the statistical representation of all the features of the dataset including numeric and binary. Table 3 explains that 78% of menopausal women are facing depression. Of the women who do not face hot flashes, menstruation, and breathing issues 82.61%, 88.37% and 84.48% respectively are not facing depression. Besides, 94.34% of women without having menopausal knowledge experience depression during the menopausal period. On the other hand, 94.38% of women who do not face depression during the menopausal period have pill knowledge. Most menopausal women face depression duration menopause with osteoporosis, heartbeat, tension, concentration, pressure or tightness in the head, tingling, and headache issues. In addition to that, it is found that the average age and age difference is 31.19 and 10.59 years for women facing depression. Finally, it can be decided that having pill knowledge is crucial for women who are passing through the menopausal period.

b. Exploratory Data Analysis

Fig. 3(A) represents a kernel density estimate (KDE) plot for the experiment of the density of participants according to age distribution. It is found that age groups from 15 to 30 years, and 50 to 70 years are denser than other age groups for the participants who are experiencing depression. However, the age group from 50 to 70 years is denser for the participant who is not facing depression. Fig. 3(B) represents the KDE plot for the experiment of the density of participants according to the distribution of age difference of the participant from their husbands. It is found that the age difference of 9 to 11 years is denser for both groups of participants.

c. Performance Analysis

In Section 3.2.5 under Section 3, it is shown that a confusion matrix is derived from every classifier. From the result of the confusion matrix, the performance for all classifiers is estimated based on different parameters. The rest of the section will analyze the performance of all applied classifiers for tracing the best-performer classifiers for the proposed system.

Table 4 demonstrates the outcome of different kinds of performance metrics with values of the classification algorithms such as Decision tree, AdaBoost Classifier, Random Forest, and XGBoost Classifier., which is found exploiting these algorithms on the menopausal dataset. These performance parameters are used to evaluate the suitable classification algorithm compared with other classifiers. Table 3 represents that RF is the best-performing classifier with 97.11% accuracy and 0.97 precision, recall, and F1 score along with 0.996 log loss.

Fig. 4 visualizes the feature importance values of each applied classification algorithm. It also represents the most significant risk factors with high values. According to the figure, age, tiredness, and menopause are the most significant risk factors. Besides, ost, knowledge, headaches heart beats, tingling, and excitability are also important symptoms for having depression among menopausal women.

Table 5 demonstrates the most significant risk factors, which is 14 in number. These risk factors are identified by the feature importance score of each applied classification algorithm. The most frequent risk factors are collected for further machine learning analysis. According to the table, age, ost, knowledge heartbeat, and others are represented in the last column of Table 5.

Table 6 represents the outcome of different kinds of performance metrics with values of the applied classification algorithms such as Decision tree, AdaBoost Classifier, Random Forest, and XGBoost Classifier., which is found exploiting these algorithms on the 14 most significant features of the menopausal dataset. It is observed that Random Forest and XGBoost are the best-performing classifier with 99.04% accuracy and 0.99 precision, recall, and F1 score along with 0.332 log loss.

Fig. 5 depicts the performance comparison between the performance of the entire dataset and the 14 most significant features. The figure illustrates the performance of accuracy, precision, recall, F1 score, and log loss. According to the figure, the 14 most significant features have produced a better performance by every performance metric for the detection of depression among menopausal women. So, these 14 most significant features are able to detect depression among menopausal women.

AUROC is one of the major characteristics, which measures the performance of a classification algorithm. It is mostly used characteristics to evaluate the performance of any classifiers. Fig. 6(A) and (B) illustrate the AUROC of different applied classifiers for all the features and 14 most significant features respectively to visualize the performance of applied classifiers graphically. Fig. 6(A) demonstrates that the Random Forest classifier is showing a superior result to any other result. Random Forest classifier covers 99.7% area under the ROC curve, which is the best result for all the features. On the other hand, the Random Forest classifier is the best-performing classifier with 100%

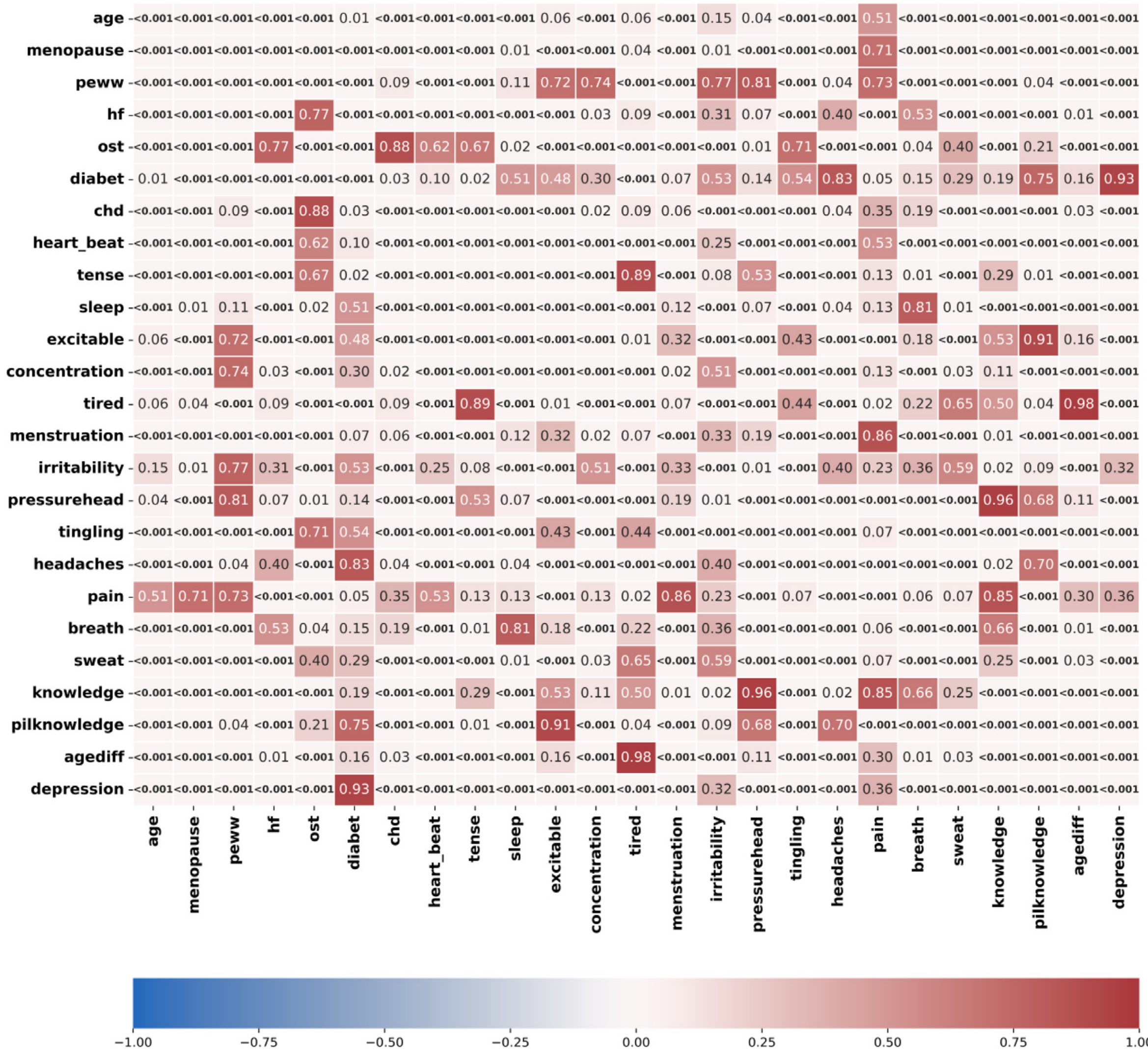

Fig. 2. illustrates a heatmap, which is generated based on the p-value of all the features of the dataset used in this study. Heatmap is a graphical representation of different data into a map or a figure. Fig. 2 represents the statistical significance among all the attributes graphically with p values. It is a statistical representation of the dataset. The figure provides a clear understanding of the dataset statistically.

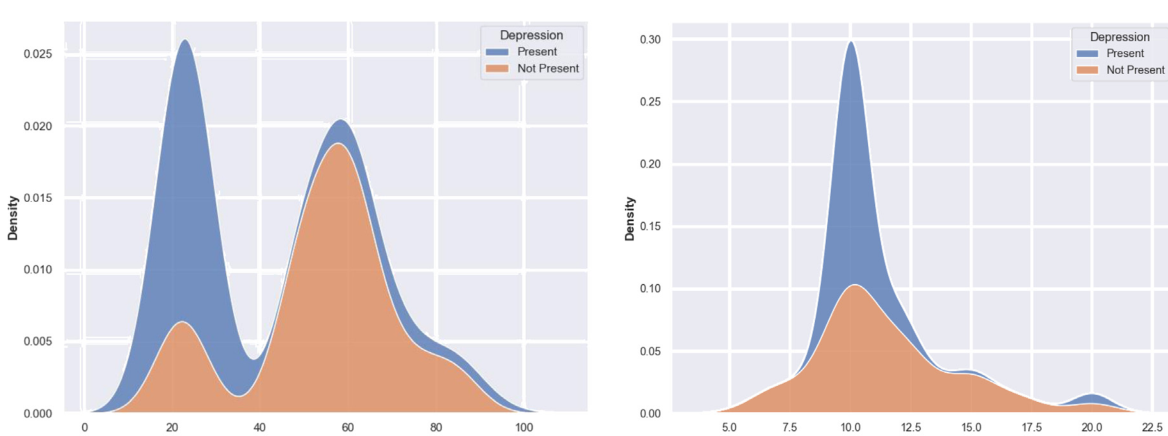

Fig. 3. KDE plot for age distribution.

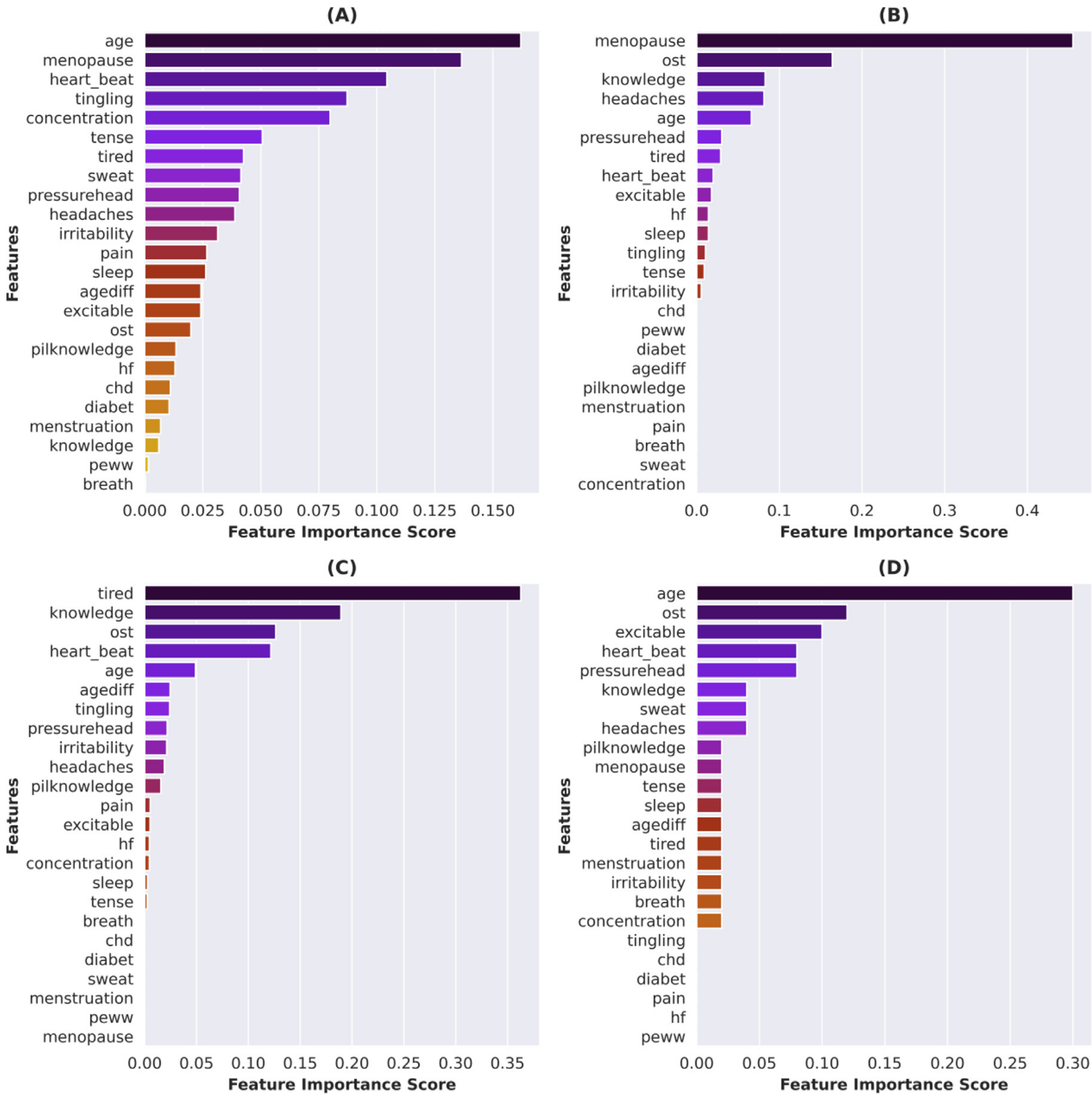

Fig. 4. Visual representation of ranking important risk factors based on feature importance score finding for all the applied algorithms. (A) Random forest, (B) Decision tree, (C) XGBoost, (D) Adaboost classifier.

Table 5
14 most important risk factors for depression among menopausal women according to all the applied algorithms.

Feature ranking	Random forest	Decision tree	XGBoost	Adaboost classifier	Collected most significant features
1	age	menopause	tired	age	age
2	menopause	ost	knowledge	ost	ost
3	heart_beat	knowledge	ost	excitable	knowledge
4	tingling	headaches	heart_beat	heart_beat	heart_beat
5	concentration	age	age	pressurehead	pressurehead
6	tense	pressurehead	agediff	knowledge	pilknowledge
7	tired	tired	tingling	sweat	tired
8	sweat	heart_beat	pressurehead	headaches	headaches
9	pressurehead	excitable	irritability	pilknowledge	excitable
10	headaches	hf	headaches	menopause	headaches
11	irritability	sleep	pilknowledge	tense	tense
12	pain	tingling	pain	sleep	sleep
13	sleep	tense	excitable	agediff	agediff
14	agediff	irritability	hf	tired	irritability

performance for the 14 most significant features, which is illustrated in Fig. 6(B).

The Precision–Recall curve is another major characteristic of a machine learning model, which represents the performance of the model. The following figure describes the precision–recall curve of all the applied classifiers. Fig. 6(B) and (D) represent the precision–recall curve for all the features and 14 most features respectively. The curve analysis demonstrates that the performance of the AdaBoost classifier is

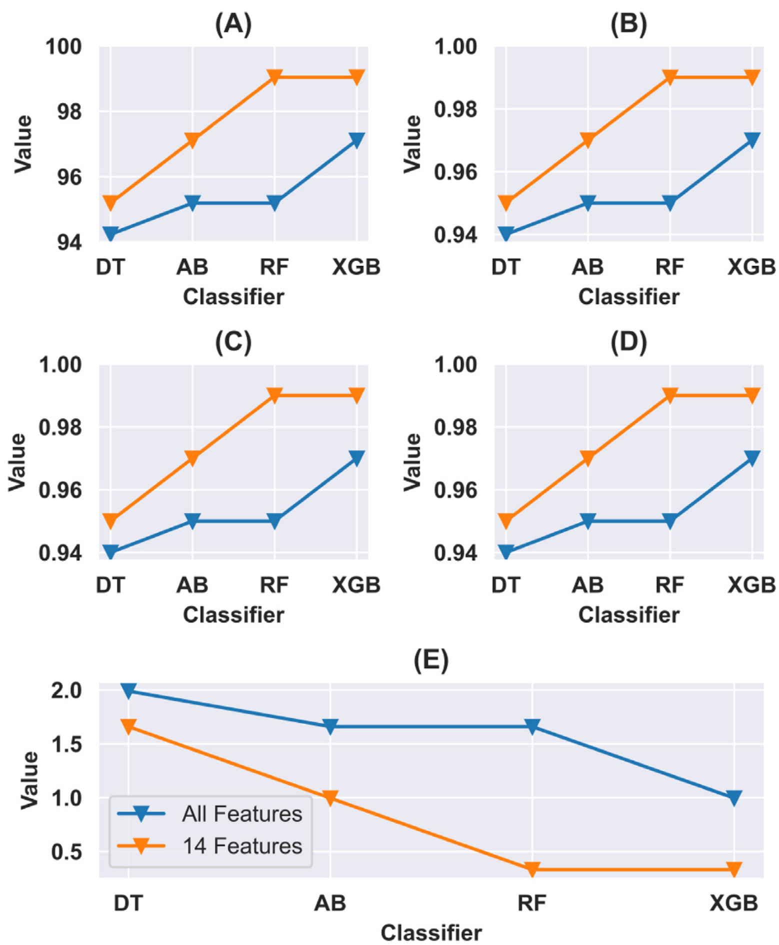

Fig. 5. Visual representation of performance comparison between performance of all the features and 14 most significant features. (A) Accuracy, (B) Precision, (C) Recall, (D) F1 score (E) Log loss.

Table 6
Performance of all the applied algorithms for 14 features.

Algorithm	Accuracy	Precision	Recall	F1 score	Log loss
Decision tree	95.19	0.95	0.95	0.95	1.66
AdaBoost classifier	97.11	0.97	0.97	0.97	0.996
Random forest	99.04	0.99	0.99	0.99	0.332
XGBoost	99.04	0.99	0.99	0.99	0.332

not satisfactory since they are showing the least performance compared to the other classifiers for both 6(B) and (D). Though the Random Forest classifier is providing the best outcome for both cases, 14 most features produced better performance than the performance produced by all the features. So, Random Forest is the most suitable classification algorithm in this case with the 14 most significant features.

In summary, we gathered a menopausal dataset with respect to depression. The collected dataset was preprocessed as necessary. For a better understanding of the dataset, EDA and statistical data analyses were employed on the dataset. To develop the machine learning model, we selected four machine learning algorithms: Adaboost Classifier, XGBoost Classifier, Decision Tree, and Random Forest. These algorithms were applied to the processed dataset and compared all the algorithm's performances based on accuracy, sensitivity, specificity, precision, recall, F1 score, MCC, ROC, and PRC. In addition, the ROC and Precision–Recall curves were produced to test the models. Then we found the most significant risk factors with respect to each applied algorithm, which is 14 in number. Then we again applied machine learning algorithms to the most significant features and found significant performances of all the applied algorithms. Where Random Forest and XGBoost classifiers generated the best performance with 99.04% accuracy. So, Random Forest and XGBoost classifiers can be applied to predict depression among menopausal women employing the 14 most significant features mentioned in Table 5. This study will help psychiatrists, psychologists, and clinicians to predict depression among menopausal women. Sometimes, due to less severity of symptoms, it is not possible to predict depression among menopausal women by psychiatrists, psychologists, and clinicians. Since depression causes serious mental and physical ailments, it is very crucial to predict whether a menopausal woman is depressed or not. From this point of view, this study will play a crucial role to predict depression and ensuring quality of life among menopausal women. The proposed model can be deployed on mobile application, web applications and other software.

5. Conclusion

The depression of menopausal women unquestionably disrupts the quality of life. Depression is also a fundamental issue that induces severe physical and mental ailments. Most of the women do not take this into consideration because of a lack of awareness and knowledge. In addition, limitations of psychiatric diagnosis and high costs, which are other reasons to skip the issue. In order to solve this kind of problem, machine learning plays a key role in predicting and detecting numerous issues in the age of scientific development. From that perspective, a prediction system is developed in this study to predict depression among menopausal women. Numerous machine learning classification algorithms are employed on a raw dataset. From all of these algorithms, a superior performer algorithm is chosen for the proposed method based on accuracy, sensitivity, specificity, precision,

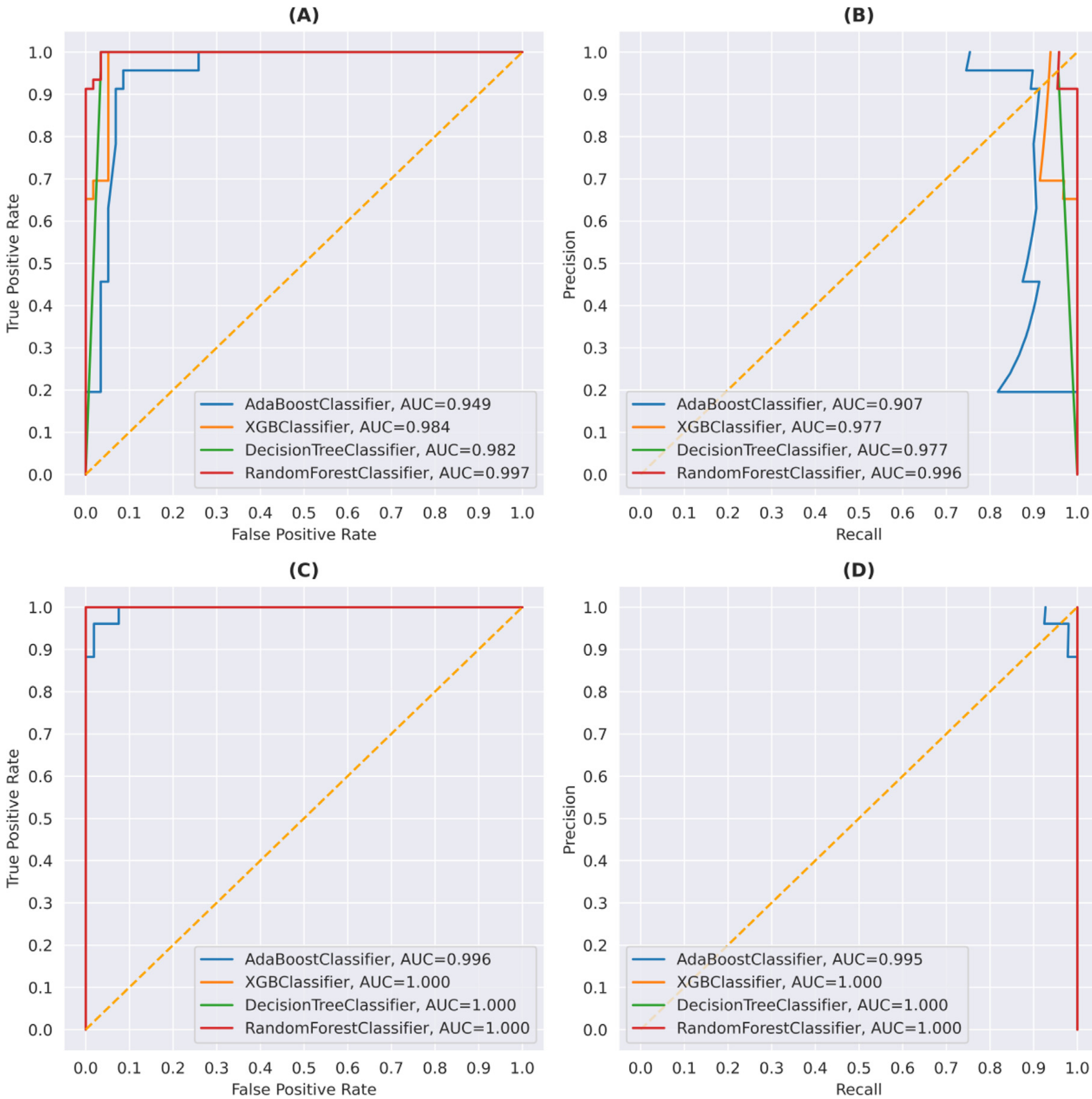

Fig. 6. Area under curve (AUC). (A) Area under ROC (AUROC) curve for all the features, (B) Area under Precision–Recall (AUPRC) curve for all the features. (C) Area under ROC (AUROC) curve for 14 most significant features, (D) Area under Precision–Recall (AUPRC) curve.

recall, and F-Measure. It is observed that Random Forest and XGBoost show significantly better outcomes than other classifiers with the 14 most significant features. This study aims to provide a simple and user-friendly prediction system, which is easy to use.

CRediT authorship contribution statement

Md. Mamun Ali: Analyzed the data, Wrote the manuscript. **Hussein Ali A. Algashamy:** Helped perform the experimental analysis with constructive discussion. **Enas Alzidi:** Helped perform the experimental analysis with constructive discussion. **Kawsar Ahmed:** Provided the idea, Designed the experiments, Analyzed the data, Wrote the manuscript. **Francis M. Bui:** Provided the idea, Designed the experiments, Helped perform the experimental analysis with constructive discussion, Supported by funding. **Shobhit K. Patel:** Helped perform the experimental analysis with constructive discussion. **Sami Azam:** Helped perform the experimental analysis with constructive discussion. **Lway Faisal Abdulrazak:** Helped perform the experimental analysis with constructive discussion. **Mohammad Ali Moni:** Provided the idea, Designed the experiments.

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.

Data availability

The data that has been used is confidential.

Acknowledgment

All authors discussed the results and contributed to the manuscript.

Funding

The work is funded in part from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] L. Al-Mehaisen, O. Al-kuran, Impact of menopause on quality of life: A cross sectional study in menopausal females in the north of Jordan, Med. Sci. Discov. 5 (6) (2018) 234–238.

[2] F. Jafari, M.H. Hadizadeh, R. Zabihi, K. Ganji, Comparison of depression, anxiety, quality of life, vitality and mental health between premenopausal and postmenopausal women, Climacteric 17 (6) (2014) 660–665.

[3] M.I. Bashar, K. Ahmed, M.S. Uddin, F. Ahmed, A.A. Emran, A. Chakraborty, Depression and quality of life among postmenopausal women in Bangladesh: A cross-sectional study, J. Menopausal Med. 23 (3) (2017) 172–181.

[4] World Health Organization, Research on the Menopause in the 1990s: Report of a WHO Scientific Group, 1996.

[5] W. Rössler, V. Ajdacic-Gross, A. Riecher-Rössler, J. Angst, M.P. Hengartner, Does menopausal transition really influence mental health? Findings from the prospective long-term Zurich study, World Psychiatry 15 (2) (2016) 146–154.

[6] N. Chuni, C.T. Sreeramareddy, Frequency of symptoms, determinants of severe symptoms, validity of and cut-off score for Menopause Rating Scale (MRS) as a screening tool: a cross-sectional survey among midlife Nepalese women, BMC Women's Health 11 (1) (2011) 30.

[7] K. Ahmed, P. Jahan, I. Nadia, F. Ahmed, Assessment of menopausal symptoms among early and late menopausal midlife Bangladeshi women and their impact on the quality of life, J. Menopausal Med. 22 (1) (2016) 39–46.

[8] S. Sharma, V.R. Tandon, A. Mahajan, Menopausal symptoms in urban women, Alcohol 4 (2007) 3–41.

[9] S. Jesmin, A.S. Islam, S. Akter, M.M. Islam, S.N. Sultana, N. Yamaguchi, O. Okazaki, M. Moroi, M. Hiroe, S. Kimura, T. Watanabe, Metabolic syndrome among pre-and post-menopausal rural women in Bangladesh: result from a population-based study, BMC Res. Notes 6 (1) (2013) 157.

[10] M.R. Soules, S. Sherman, E. Parrott, R. Rebar, N. Santoro, W. Utian, N. Woods, Executive summary: stages of reproductive aging workshop (STRAW), Climacteric 4 (4) (2001) 267–272.

[11] M.H. Cheng, S.J. Lee, S.J. Wang, P.H. Wang, J.L. Fuh, Does menopausal transition affect the quality of life? A longitudinal study of middle-aged women in Kinmen, Menopause 14 (5) (2007) 885–890.

[12] J.L. Fuh, S.J. Wang, S.J. Lee, S.R. Lu, K.D. Juang, Quality of life and menopausal transition for middle-aged women on Kinmen island, Qual. Life Res. 12 (1) (2003) 53–61.

[13] Y. Zheng, Y. Zhou, J. Hu, J. Zhu, Q. Hua, M. Tao, Prediction of risk of depressive symptoms in menopausal women based on hot flash and sweating symptoms: a multicentre study, Clin. Interv. Aging 12 (2017) 1993.

[14] M. Hickey, D.A. Schoenaker, H. Joffe, G.D. Mishra, Depressive symptoms across the menopause transition: findings from a large population-based cohort study, Menopause 23 (12) (2016) 1287–1293.

[15] P.R. Albert, Why is depression more prevalent in women? J. Psychiatry Neurosci.: JPN 40 (4) (2015) 219.

[16] N. Santoro, C.N. Epperson, S.B. Mathews, Menopausal symptoms and their management, Endocrinol. Metab. Clin. 44 (3) (2015) 497–515.

[17] K.E. Campbell, L. Dennerstein, S. Finch, C.E. Szoeke, Impact of menopausal status on negative mood and depressive symptoms in a longitudinal sample spanning 20 years, Menopause 24 (5) (2017) 490–496.

[18] M. de Kruif, A.T. Spijker, M.L. Molendijk, Depression during the perimenopause: A meta-analysis, J. Affect. Disord. 206 (2016) 174–180.

[19] L.S. Cohen, C.N. Soares, A.F. Vitonis, M.W. Otto, B.L. Harlow, Risk for new onset of depression during the menopausal transition: the Harvard study of moods and cycles, Arch. Gen. Psychiatry 63 (4) (2006) 385–390.

[20] J.T. Bromberger, K.A. Matthews, L.L. Schott, S. Brockwell, N.E. Avis, H.M. Kravitz, S.A. Everson-Rose, E.B. Gold, M. Sowers, J.F. Randolph Jr., Depressive symptoms during the menopausal transition: The Study of Women's Health Across the Nation (SWAN), J. Affect. Disord. 103 (1–3) (2007) 267–272.

[21] J.T. Bromberger, H.M. Kravitz, Mood and menopause: findings from the Study of Women's Health Across the Nation (SWAN) over 10 years, Obstet. Gynecol. Clin. 38 (3) (2011) 609–625.

[22] C.N. Soares, Mood disorders in midlife women: understanding the critical window and its clinical implications, Menopause 21 (2) (2014) 198–206.

[23] S.A. Rahman, S.R. Zainudin, V.L. Mun, Assessment of menopausal symptoms using modified Menopause Rating Scale (MRS) among middle age women in Kuching, Sarawak, Malaysia, Asia Pac. Fam. Med. 9 (2010) 5.

[24] X. Li, F. Chen, D. Sun, M. Tao, Predicting menopausal symptoms with artificial neural network, Expert Syst. Appl. 42 (22) (2015) 8698–8706.

[25] N. Chuni, C.T. Sreeramareddy, Frequency of symptoms, determinants of severe symptoms, validity of and cut-off score for Menopause Rating Scale (MRS) as a screening tool: a cross-sectional survey among midlife Nepalese women, BMC Women's Health 11 (1) (2011) 30.

[26] J.T. Bromberger, L. Schott, H.M. Kravitz, H. Joffe, Risk factors for major depression during midlife among a community sample of women with and without prior major depression: are they the same or different? Psychol. Med. 45 (8) (2015) 1653–1664.

[27] K.D. Juang, S.J. Wang, S.R. Lu, S.J. Lee, J.L. Fuh, Hot flashes are associated with psychological symptoms of anxiety and depression in peri-and post-but not premenopausal women, Maturitas 52 (2) (2005) 119–126.

[28] E.P. Handing, C. Strobl, Y. Jiao, L. Feliciano, S. Aichele, Predictors of depression among middle-aged and older men and women in Europe: A machine learning approach, Lancet Reg. Health-Eur. 18 (2022) 100391.

[29] M.B. Waqar, S. Noreen, B. Rizwan, A. Khan, I. Bashir, R. Tanveer, T. Akram, M. Tariq, I. Khalid, M. Asif, Knowledge, attitude, and practices regarding menopause among elderly women attending tertiary health care centers in Lahore, Pakistan: Menopause among elderly women, Pak. J. Health Sci. 1 (2023) 2–17.

[30] B. Shakya, P.K. Pokharel, B.K. Yadav, P.M. Shyangwa, I.S. Paudel, H.P. Pokharel, Prevalence of depression among menopausal women in a rural community of Morang, Nepal, Nepal Med. Coll. J. 24 (1) (2022) 30–39.

[31] S.Y. An, Y. Kim, R. Kwon, G.Y. Lim, H.R. Choi, S. Namgoung, S.W. Jeon, Y. Chang, S. Ryu, Depressive symptoms and suicidality by menopausal stages among middle-aged Korean women, Epidemiol. Psychiatr. Sci. 31 (2022) e60.

[32] F. Saeed Elsayed, F. Mohammed Ahmed, S. Saad Moustafa, F. Abdallah Shams Eldin, Correlation between menopausal symptoms, depressive symptoms and marital satisfaction among menopausal women, Int. Egypt. J. Nurs. Sci. Res. 3 (2) (2023) 419–437.

[33] A. Kułak-Bejda, K. Krajewska-Ferishah, A. Szyszko-Perłowska, N. Waszkiewicz, Risk assessment of depression amongst women during menopause before and during the COVID-19 pandemic, Int. J. Environ. Res. Public Health 20 (1) (2022) 596.

[34] S. Álvarez-Pardo, J.A. de Paz, E.M. Romero-Pérez, J.M. Tánori-Tapia, P.A. Rendón-Delcid, J.J. González-Bernal, J. Fernández-Solana, L. Simón-Vicente, J. Mielgo-Ayuso, J. González-Santos, Related factors with depression and anxiety in mastectomized women breast cancer survivors, Int. J. Environ. Res. Public Health 20 (4) (2023) 2881.

[35] Y. Liu, X. Yang, Y. Xu, Y. Wu, Y. Zhong, S. Yang, Cognitive function and depressive symptoms among Chinese adults aged 40 years and above: The mediating roles of IADL disability and life satisfaction, Int. J. Environ. Res. Public Health 20 (5) (2023) 4445.

[36] P.T. Thuy, T.T.H. Oanh, N.T.A. Nhan, N.T.Y. Hoai, Depression among menopausal women: A cross-sectional study in Vietnam, Age, 40 (44) 56.

[37] S. Khatak, M. Gupta, S. Grover, N. Aggarwal, Depression among peri-and post-menopausal women during COVID-19 pandemic in Chandigarh, North India: A study from community, J. Mid-Life Health 13 (3) (2022) 233.

[38] S. Theis, S.J. Baumgartner, H. Janka, A. Kolokythas, C. Skala, P. Stute, Quality of life in menopausal women in the workplace–a systematic review, Climacteric 26 (2) (2023) 80–87.

[39] R.M. Hables, A.A. Moussa, Association between perceived social support, attitude and depression among menopausal women, Tanta Sci. Nurs. J. 25 (22) (2022) 11–22.

[40] I. Ahmad, M. Basheri, M.J. Iqbal, A. Rahim, Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection, IEEE Access 6 (2018) 33789–33795.

[41] K. Ahmed, A.A. Emran, T. Jesmin, R.F. Mukti, M.Z. Rahman, F. Ahmed, Early detection of lung cancer risk using data mining, Asian Pac. J. Cancer Prev. 14 (2013) 595–598.

[42] K. Ahmed, T. Jesmin, M.Z. Rahman, Early prevention and detection of skin cancer risk using data mining, Int. J. Comput. Appl. 62 (4) (2013).

[43] S. Asaduzzaman, S. Chakraborty, M.G. Hossain, M.I. Bashar, T. Bhuiyan, S.S. Chandan, K. Ahmed, B.K. Paul, Hazardous consequences of polygamy, contraceptives and number of childs on cervical cancer in a low incoming country: Bangladesh, Cumhur. Üniv. Fen-Edeb. Fak. Fen Bilim. Derg. 37 (1) (2016) 74–84.

[44] K. Ahmed, S. Asaduzzaman, M.I. Bashar, G. Hossain, T. Bhuiyan, Association assessment among risk factors and breast cancer in a low income country: Bangladesh, Asian Pac. J. Cancer Prev. 16 (17) (2015) 7507–7512.

[45] S.M.M. Hasan, M.A. Mamun, M.P. Uddin, M.A. Hossain, Comparative analysis of classification approaches for heart disease prediction, in: 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), IEEE, 2018, pp. 1–4.

[46] J.H. Joloudari, E. Hassannataj Joloudari, H. Saadatfar, M. GhasemiGol, S.M. Razavi, A. Mosavi, N. Nabipour, S. Shamshirband, L. Nadai, Coronary artery disease diagnosis; ranking the significant features using a random trees model, Int. J. Environ. Res. Public Health 17 (3) (2020) 731.

[47] M.D.R.V. Gimpy, Missing value imputation in multi attribute data set, Int. J. Comput. Sci. Inf. Technol. 5 (4) (2014) 1–7.

[48] L. Peng, L. Lei, A review of missing data treatment methods, Intell. Inf. Manag. Syst. Technol. 1 (3) (2005) 412–419.

[49] M.R. Rahman, T. Islam, T. Zaman, M. Shahjaman, M.R. Karim, F. Huq, J.M. Quinn, R.D. Holsinger, E. Gov, M.A. Moni, Identification of molecular signatures and pathways to identify novel therapeutic targets in alzheimer's disease: Insights from a systems biomedicine perspective, Genomics (2019).

[50] Four Techniques for Outlier Detection, https://www.kdnuggets.com/2018/12/four-techniques-outlier-detection.html.

[51] M.S. Satu, S.T. Atik, M.A. Moni, A Novel Hybrid Machine Learning Model To Predict Diabetes Mellitus.

[52] S. Uddin, A. Khan, M.E. Hossain, M.A. Moni, Comparing different supervised machine learning algorithms for disease prediction, BMC Med. Inform. Decis. Mak. 19 (1) (2019) 1–16.
[53] A. Liaw, M. Wiener, Classification and regression by randomForest, R News 2 (3) (2002) 18–22.
[54] H. David, S.A. Belcy, Heart disease prediction using data mining techniques, ICTACT J. Soft Comput. 9 (1) (2018).
[55] A. Cutler, D.R. Cutler, J.R. Stevens, Random forests, in: Ensemble Machine Learning, Springer, Boston, MA, 2012, pp. 157–175.
[56] Y. Liu, Y. Wang, J. Zhang, New machine learning algorithm: Random forest, in: International Conference on Information Computing and Applications, Springer, Berlin, Heidelberg, 2012, pp. 246–252.
[57] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[58] D.F. Schwarz, I.R. König, A. Ziegler, On safari to random jungle: a fast implementation of random forests for high-dimensional data, Bioinformatics 26 (14) (2010) 1752–1758.
[59] K.R. Gray, P. Aljabar, R.A. Heckemann, A. Hammers, D. Rueckert, Alzheimer's Disease Neuroimaging Initiative, Random forest-based similarity measures for multi-modal classification of Alzheimer's disease, NeuroImage 65 (2013) 167–175.
[60] M. Xu, K.G. Tantisira, A. Wu, A.A. Litonjua, J.H. Chu, B.E. Himes, A. Damask, S.T. Weiss, Genome wide association study to predict severe asthma exacerbations in children using random forests classifiers, BMC Med. Genet. 12 (1) (2011) 90.
[61] E. Mosca, R. Alfieri, I. Merelli, F. Viti, A. Calabria, L. Milanesi, A multilevel data integration resource for breast cancer study, BMC Syst. Biol. 4 (1) (2010) 76.
[62] R. Genuer, J.-M. Poggi, C. Tuleau-Malot, VSURF: an R package for variable selection using random forests, R J. 7 (2) (2015) 19–33.
[63] H.M. Le, T.D. Tran, Lang Van Tran, Automatic heart disease prediction using feature selection and data mining technique, J. Comput. Sci. Cybern. 34 (1) (2018) 33–48.
[64] R. Chitra, V. Seenivasagam, Heart disease prediction system using supervised learning classifier, Bonfring Int. J. Softw. Eng. Soft Comput. 3 (1) (2013) 01–07.
[65] J. Han, M. Kamber, J. Pei, Data Mining Concepts and Techniques, third ed., Morgan Kaufmann, 2011.
[66] J. Coe, Performance comparison of Naïve Bayes and J48 classification algorithms, Int. J. Appl. Eng. Res. 7 (11) (2012) 2012.
[67] Z. Lei, Y. Sun, Y.A. Nanehkaran, S. Yang, M.S. Islam, H. Lei, D. Zhang, A novel data-driven robust framework based on machine learning and knowledge graph for disease classification, Future Gener. Comput. Syst. 102 (2020) 534–548.
[68] X. Luo, F. Lin, Y. Chen, S. Zhu, Z. Xu, Z. Huo, M. Yu, J. Peng, Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features, Sci. Rep. 9 (1) (2019) 1–13.
[69] K. Roy, S. Kar, R.N. Das, Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Academic Press, 2015.
[70] Ni Huang, Insuk Lee, Edward M. Marcotte, Matthew E. Hurles, Characterising and predicting haploinsufficiency in the human genome, PLoS Genet. 6 (10) (2010) 1–11.
[71] G. Liu, T.R. Mercer, A.M.J. Shearwood, S.J. Siira, M.E. Hibbs, J.S. Mattick, O. Rackham, A. Filipovska, Mapping of mitochondrial RNA-protein interactions by digital RNase footprinting, Cell Rep. 5 (3) (2013) 839–848.
[72] L. Shi, G. Campbell, W.D. Jones, F. Campagne, Z. Wen, S.J. Walker, Z. Su, T.M. Chu, F.M. Goodsaid, L. Pusztai, J.D. Shaughnessy Jr., The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models, Nature Biotechnol. 28 (8) (2010) 827.
[73] M.S. Pepe, The statistical evaluation of medical tests for classification and prediction, Medicine (2003).
[74] M. Mahdizadeh, E. Zamanzade, Smooth estimation of the area under the ROC curve in multistage ranked set sampling, Statist. Papers (2020) 1–24.
[75] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.
[76] Chongomweru Halimu, Asem Kasem, S.H. Shah Newaz, Empirical comparison of area under ROC curve (AUC) and mathew correlation coefficient (MCC) for evaluating machine learning algorithms on imbalanced datasets for binary classification, in: Proceedings of the 3rd International Conference on Machine Learning and Soft Computing, 2019, pp. 1–6.
[77] D.J. Hand, R.J. Till, A simple generalisation of the area under the ROC curve for multiple class classification problems, Mach. Learn. 45 (2) (2001) 171–186.
[78] D. Li, Y. Tian, Survey and experimental study on metric learning methods, Neural Netw. 105 (2018) 447–462.
[79] C.X. Ling, J. Huang, H. Zhang, AUC: a statistically consistent and more discriminating measure than accuracy, in: Ijcai, Vol. 3, 2003, pp. 519–524.
[80] B. Zhu, B. Baesens, S.K. vanden Broucke, An empirical comparison of techniques for the class imbalance problem in churn prediction, Inform. Sci. 408 (2017) 84–99.
[81] T. Saito, M. Rehmsmeier, The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets, PLoS One 10 (3) (2015) e0118432.