Journal of Computer & Robotics 13(2), 2020, 1-12

Computer
& Robotics

Reliability Measurements in Depression Detection Using a Data Mining Approach Based on Fuzzy-Genetics

Mohammad Nadjafi [a,*], Sepideh Jenabi [b], Adel Najafi [c], Ghasem Kahe [d]

[a] *Aerospace Research Institute (Ministry of Science, Research and Technology)*

[b] *Computer engineering,Science and Research Branch, Islamic Azad University,Tehran, Iran*

[c] *Researcher and Instructure of Computer Science, Tehran, Iran*

[d] *Aerospace Research Institute (Ministry of Science, Research and Technology),Tehran,Iran*

Received 09 October 2020; Accepted 09 December 2020

Abstract

Developing a reliable data mining method is one of the most challenging issues in the features of advanced computer-based systems. Model reliability in depression disorder detection is the determining p-value or confidence limit for accuracy score. In this regard, data mining evaluation metrics provide a path to knowledge discovery and feature extraction is an important process for discovering patterns in data by exploring and modeling big data. The present paper discusses the data mining approach about detection in depression disorder characterized by symptoms such as sadness, feeling empty, anxiety, and sleep symptoms as well as the loss of initiation and interest inactivity. In this survey, a unique dataset containing sensor data collected from patients with depression was used. For each patient, sensor data were measured over several days. In this regard, the represented dataset could be useful for a better understanding of the relationship between depression and motor activity. On the other hand, to overcome the uncertainties raised from wearable sensors (that caused a significant amount of error in similar previous studies using conventional learning methods such as SVM, LR, NB), and also to increase the efficiency and accuracy of the results and to develop a reliable decision-making framework, the evolutionary hybrid machine learning method (fuzzy-genetic algorithm) will be used. The results show the high accuracy of the proposed method compared to other existing methods.

Keywords: Data Mining Evaluation Metrics, Reliability, Fuzzy Genetic Algorithm, Feature Extraction, Depression Disorder.

1.Introduction

According to the World Health Organization (WHO), depression is a common mental disorder worldwide affecting more than 300 million people regardless of their age. Long-term depression can lead to suicide if people do not receive specialized help at critical times. Depending on the diagnostic criteria for major depressive disorder (MDD), symptoms of depression include clinical depression, depressed mood or loss of interest in daily activities that last for more than two weeks, weight loss or weight gain, change in sleep cycle, Loss of energy is the loss of motivation for activity and mobility, disability and suicidal thoughts. Also, if these symptoms cause significant discomfort or damage to any of the vital areas (social, occupational, educational, etc.), it can be an important reason for clinical diagnosis. Symptoms of Generalized Anxiety Disorder (GAD) are excessive anxiety, restlessness, easy tiredness, difficulty concentrating, and irritability, sleep disturbance, and muscle tension. MDD and GAD have common symptoms such as sleep disorders, fatigue, and depression-focused problems, people tend to move slowly and have mild reactions. Anxious people are more likely to be at risk and afraid of their future. According to the American Association for Anxiety and Depression (ADAA), about half of those diagnosed with clinical depression are likely to have anxiety disorders, and these two disorders can cause each other. According to the World Health Organization, one

* Corresponding Author Email: m.nadjafi@ari.ac.ir

of the important factors involved in the diagnosis and treatment of depression is the lack of professionals, social shame, misdiagnosis, etc. [1].

If the accuracy of the diagnosis of depression is low or if it is diagnosed for some reason, people also do not receive the necessary treatment, so the life of the individual is impaired, so diagnosing depression is an important issue that can help improve the quality of life of the patient. A review article [2] examines many of the important work done in the field of mental illness, and the authors have used this article for reference. Traditional p-value/Fischerian Frequentist - Central Limit Theorem" statistics, "resistance" to accepting in the thinking process of predictive analytics is the issue. Many of the medical residents were "resistant to learn"; also, having been "indoctrinated" by medical journal articles using P-value statistics. Insist that any medical resident research paper including data mining modelling / analysis be anticipated; thus the paper was divided into two parts: 1) the presentation of the traditional classification evaluation, which usually was primarily only descriptive because the evaluation metrics used for comparisons; and 2) the DM/PA analyses for reliability measures.

Developing a reliable and safe system is one of the most important features of advanced computer-based systems [3]. The software is often responsible for controlling the behaviour of mechanical and electrical components as well as interactions between components in systems. Therefore, considering software safety and fault detection are essential in software development. Authors in [3] introduces an approach to engineering evidence that examines the software in its lifecycle according to the principles of software safety and system safety engineering. Their approach ensures that software risks are identified and documented in the software lifecycle, after which the risks are reduced to an acceptable level in terms of safety according to the proposed methods. The presented approach was applied to a real master case with positive results, namely the Data and Command Unit. Industrial companies seek to manufacture products of higher quality which can be achieved by increasing reliability, maintainability and thus the availability of products. Typically, maintenance activities are aimed to reduce failures of industrial machinery and minimize the consequences of such failures. So the industrial companies try to improve their efficiency by using different fault detection techniques. One strategy is to process and analyse previous generated data to predict future failures. The purpose of [4] is to detect wasted parts using different data mining algorithms and compare the accuracy of these algorithms.

The authors of [5] remind that the fact that up to 85% of the medical research journal articles are incorrect in their use of traditional p-value statistics. The use of data mining / predictive analytics in medical research and healthcare delivery is much more easily understandable, and accurate, thus this “resistance to its use” needs to be rectified if they as a society are to make headway in providing accurate diagnoses and accurate treatment plans for the individual patient- authors think that all realization accomplishing these goals is essential to bringing medical health care costs down to reasonable levels. Why are exploratory graphics and visualization important? Basic knowledge regarding crime trends and patterns represents an important first step in information-based approaches to resource allocation, and other operational decisions. Moreover, in order to effectively characterize data so as to reveal important associations and create accurate and reliable models, it is important to understand the data and be able to generate samples that are relatively homogenous, except with regard to those unique features that have value from a modelling perspective [6].

The remainder of the paper is structured as follows: Section 2 provides a history of the subject under study. Section 3 deals with the concepts of reliability based on data mining. Section 4 provides the mathematical expression of the subject and the required formulation for the proposed framework is presented in section 5. In the following, the attained results are tabulated and demonstrated in section 6 and then the conclusion and future work are given.

2. Research Background

Today, the main challenge of studying emotional disorders is to master the patient's emotional changes. Research into user emotion recognition mainly consists of three categories: emotion recognition based on audio and video signals, physiological signals, and multitasking data [7]. Continuous monitoring of a person's stress level is essential for recognizing and managing personal stress [8]. Some physiological markers are widely used to assess stress, including galvanic skin reaction, various features of heart rate patterns, blood pressure and respiratory activity [9] Suicide is another consequence of fear of mental illness. While there are significant differences in suicide rates among many countries, suicide is among the top 15 causes of death worldwide [10]. Early detection and accurate diagnosis of depression are essential criteria for optimizing the method of selecting and achieving results, thus reducing the economic and psychosocial costs that result from hospitalization, lost productivity, and suicide [11].

In the last two decades, the use of data mining techniques in various disciplines has increased dramatically. [12] Its big. Data mining involves automated learning algorithms for learning, extracting and identifying useful information and subsequent knowledge from large databases [13]. For the past 10 years, data mining techniques have been used in medical research, especially in neuroscience and biomedical sciences. Recently, psychiatry has taken advantage of these approaches to gain a better understanding of the genetic makeup of mental illness [14]. Table 1 shows studies on data mining techniques and algorithms in patients with depression [15, 16]. Other some summarized researches are provided in Table.

Table 1
Summarizes the work done on depression [15, 16].

Description	Technique and Algorithm	Year of publication and author
Making depressive terms and calculating the probability of depression	Ontology, Bayesian Networks	Chang, Hung & Juang, 2013, [17]
Model used to evaluate and predict acceleration-based depression, with feature selection techniques	AdaBoost by DT Classification	Thanathamathee, 2014, [18]
Examine the most common symptoms of depressed patients as well as their scenarios	Associative analysis, repeated pattern tree	Ghafoor, Huang & Liu, 2015, [19]
Build predictive models, linkages between reading habits, and depression tendencies	KNN, SVM, Naïve Bayesian, linear regression, logistic regression	Hou et al., 2016, [20]
Predicting general anxiety disorder among women	Random forest	Husain et al., 2016, [21]
EEG-based research, searching for high-frequency bands and areas of the brain that are most often associated with mild depression.	Bayes Net, SVM, Regression Logistic, KNN, random forest. Best First Choice (BF), Greedy Stepwise (GSW), Genetic Search (GS), Prior Linear Selection (LFS), and Rank Search (RS) were applied based on correlation feature selection (CFS) for feature selection.	Li et al., 2016, [22]
Evaluation of Neurophysiological Characteristics of Elderly Participants with Depression	Regression tree, a linear combination of covariate variables.	Nie, Gong & Ye, 2016, [23]
The most common symptoms of depressed patients, as well as their scenarios	Random forest, Random tree, Multilayer Perceptron (MPL network), SVM	Spyrou et al., 2016, [24]
Offer passive infrared sensors to monitor the daily living activities of the elderly only	Neural Networks, DT C4.5, Bayesian Networks, SVM	Kim et al., 2017, [25]

Information technology has the power to analyse diseases in the field of mental health and help us advance knowledge faster [26]. With the help of these technologies, patients can receive very personal therapies, help therapists in decision-making based on evidence, provide scientists with the opportunity to search for new knowledge that shows the real causes of mental illness, and also develop methods. Therapies will be more effective. Genetic algorithm, known as one of the random optimization methods, was invented by John Holland in 1967. Later, with the efforts of Goldberg 1989, [27] this method found its place, and today, due to its capabilities, it is well-positioned among other methods. GA is used in various fields of Data Mining to obtain optimal solutions for better performance of the data required for accurate decision making and processing. Genetic algorithms are understood for classification techniques using a uniform general population method and a uniform population-inspired operator. Initial populations provide genetic variation with mutation biomarkers by eliminating the stochastic method using the subsequent convergence operator [28]. Inappropriate diagnosis of mental health illnesses leads to wrong treatment and causes irreversible deterioration in the client's mental health status including hospitalization and/or premature death.

The study [29] constructs the semi-automated system based on an integration of the technology of genetic algorithm, classification data mining and machine learning. The goal is not to fully automate the classification process of mentally ill individuals, but to ensure that a classifier is aware of all possible mental health illnesses could match patient's symptoms. The classifier/psychological analyst will be able to make an informed, intelligent and appropriate assessment that will lead to an accurate prognosis. The analyst will be the ultimate selector of the diagnosis and treatment plan.

Depression is considered to be a chronic mood disorder. In paper [30] attempts to mathematically model how psychiatrists clinically perceive the symptoms and then diagnose depression states. According to the Diagnostic and Statistical Manual (DSM)-IV-TR, fourteen symptoms of adult depression have been considered. A load of each symptom and the corresponding severity of depression are measured by the psychiatrists (i.e. the domain experts). Using the Principal Component Analysis (PCA) out of fourteen symptoms (as features) seven has been extracted as latent factors. Using these features as inputs, a hybrid system consisting of Mamdani's Fuzzy logic controller (FLC) on a Feed-Forward Multilayer Neural Net (FFMNN) has been developed. The output of the hybrid system was tuned by a back propagation (BPNN) algorithm. Finally, the model is validated using 302 real-world adult depression cases and 50 controls (i.e. normal population).

The study concludes that the hybrid controller can diagnose and grade depression with an average accuracy of 95.50%. Mental health is measured with a high degree of impairment, such as affective disorder that leads to depression and anxiety disorders. By improving the quality of treatment, costs can be significantly reduced and this quality can be improved by introducing data mining tools and techniques into mental health [34]. The research in this study requires motion data, so the brief data collection approach is described below, focusing on the detection of depression and anxiety using extracted motion data. Wearable sensors that measure different parts of people's activities are commonplace in technology today.

The data generated is usually collected by measuring the number of daily steps or calories burned since continuous heart rate and activity recording. The stopwatch is a wearable sensor in a wrapper like a wristwatch mounted on the wrist [35]. The purpose of this study was to extract predicted information in this area, build models of clinical information, and use them to predict diagnostic results on unseen medical data. It is intended to support clinical decision making [36]. Predictive data extraction methods can be applied to the construction of decision models for medical purposes such as prognosis, diagnosis and treatment planning that can be interpreted in clinical systems as systematic support components [37].

Feature selection involves specifying the most important features (attributes) of each data (sample) used for training. A feature selection algorithm attempts to select a small subset of features that have the most resolutions and impact. In problems described by hundreds or thousands of features, feature selection, with the great advantage of computational time required to obtain an efficient solution, allows us to reduce the number of variables [38].

The research method presented is based on a comparison with [39]. In this study, the analysis is presented on a unique dataset containing sensor data collected from patients with depression. The dataset consists of recordings of locomotors activity of 23 depressed patients and 32 healthy individuals. They used machine learning to classify patients as depressed and non-depressed. Validation is used to evaluate the algorithms. The best results so far are F1 0.73 and MCC 0.44. The overall findings indicate that sensor data contains information that can be used to determine a person's depression status. In general, the results of the authors' methods are presented in Table 2. Based on the [39], for 10-folded cross validation, the best performing classifiers are bolded.

Table 2
Validation metrics [39]

Classifier	Class	PREC	REC	ACC	SPEC
Nearest Neighbors	Non Depressed	0.878	0.669	0.675	0.705
Nearest Neighbors	Weighted average	**0.752**	0.678	0.675	0.696
Linear SVM	Depressed	0.577	0.721	0.727	0.734
Linear SVM	Non Depressed	0.836	0.734	0.727	0.721
Linear SVM	Weighted average	0.735	**0.729**	**0.727**	0.726
RBF SVM	Depressed	0.546	0.732	0.724	0.724
RBF SVM	Non Depressed	0.853	0.724	0.724	0.732
RBF SVM	Weighted average	0.733	0.727	0.724	**0.729**

Nevertheless, most research and studies have been done in the field of depression disorder detection, there are some drawbacks in the exact and reliable detection. In order to cope aforementioned problems, in this paper a comprehensive method will be introduced to solved the problems. The proposed framework is basis on the hybrid fuzzy-evolutionary algorithms and neural network concept used for training part can be shown better and more flexible than any other algorithms because of simple learning method. In this regard, the feature extraction process is basis on the homogenization, compression and dimension reduction led to more comparative data. And the loss function is basis on the MAE (mean absolute error) between the predictions and the actual observations more resistant than MSE (mean squared error).

3. Reliability in Data Mining

All the algorithms presented in data mining for improvement are a confirmation that need to provide highly reliable algorithms. But how is reliability expressed in data mining? What factors are important to determine the reliability of data mining? The Reliability of a working system has been realized generally by probability functions, Mathematical Expectation and simulation Processes based on Weibull's Failure Laws and other Failure Laws. But to trust a working system in anticipation of good performance process, the Mathematical computations can be transformed to computerization methods for which a detailed analysis both on the mathematical and software applications, are preferred [31]. A Hazard model for initial failure but late working in a good condition with no further failure on a hardware system has been proposed and mathematically the complex valued expression for Mean time to Failure and Variance are derived for experimental execution and the outputs prove and concur with Data-Mining results.

Failure is the state [31] of becoming unserviceable in the case of a machine or non-achievement of an aim or goal on the part of an individual. But it is exciting to find a system or person to fail at the start for some time interval and picks up for a nice stretch of work without failure to achieve a particular Endeavour. This transition status is found in physical condition like walking, playing, eating and learning etc. Hence a prediction process is essential by scientific approaches including Computer Software, Statistical, Mathematical and Data Mining approaches and they are proposed in detail in this paper after perusal of present status quo of findings on this entity and a Mathematical analysis of calculus methods are executed.

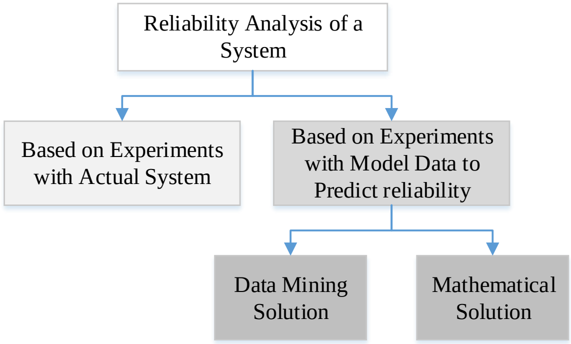

Fig. 1. Flow Diagram to predict reliability of the system [31]

The definition of reliability is based on the definition of the occurrence of failure. There are several methods for modelling reliability. Methods based on statistics and statistical information resulting from performance and determining the number of failures as well as the study of failure physics. To calculate the reliability of each component based on available statistical data, a model for failure rate is selected and its parameters are estimated based on available data or estimated by simulation and engineering knowledge and experience of experts. In engineering methods, determining the model and mechanism of failure is very important.

The two main functions in investigating failure behaviour are failure density functions (f) and hazard rate (z). In a given period of time, the density of failure is equal to the ratio of the occurrence of failures that occurred in that period to the total initial population of elements, divided by the length of the interval. In other words, this function indicates the average speed (or overall speed of failures). But the hazard rate can be considered as the instantaneous rate of failure, which is equal to the ratio of the number of failures in the interval to the number of healthy elements at the beginning of the interval, divided by the interval length.

4. Formulation

4.1. Classification Evaluation Metrics

Classification evaluation metrics score generally indicates how correct we are about our prediction [32]. The higher the score, the better our model is.

The parameters used in the classification algorithms are given in the table 3, 4, and 5 respectively:

Table 3

Type of input parameters

True positive (**TP**)	An instance for which both predicted and actual values are positive
True negative (**TN**)	An instance for which both predicted and actual values are negative.
False Positive (**FP**)	False Positive: An instance for which predicted value is positive but actual value is negative.
False Negative (**FN**)	False Negative: An instance for which predicted value is negative but actual value is positive.

Table 4

Prediction values

Actual Values	**Positive**	**Negative**
Positive	Number of True Positives	Number of False Negatives
Negative	Number of False Positives	Number of True Negatives

Table 5

Evaluation metrics in classification models

Metrics	**Definition**	**Formula**
Accuracy	Accuracy can be defined as the percentage of correct predictions made by our classification model.	(TP+TN)/number of rows in data
Precision	Precision indicates out of all positive predictions, how many are actually positive. It is defined as a ratio of correct positive predictions to overall positive predictions.	TP/TP+FP
Recall	Recall indicates out of all actually positive values, how many are predicted positive. It is a ratio of correct positive predictions to the overall number of positive instances in the dataset.	TP/TP+FN
F1 score	When avoiding both false positives and false negatives are equally important for our problem, we need a tradeoff between precision and recall. This is when we use the f1 score as a metric.	$\frac{tp}{tp + \frac{1}{2}(fp + fn)}$

4.2. Faulty & Reliability Detections

As mentioned before in a given period of time, the density of failure is equal to the ratio of the occurrence of failures that occurred in that period to the total initial population of elements, divided by the length of the interval. However, in datamining methods like proposed algorithm there is not existence of time period so the fault density is equal to the ratio of the faulty detection that occurred in the execution fold with respect to the total dataset`s initial population. Based on the general definition of reliability analysis using fault detection on data mining, the following relationships are existing:

$$R(t) = 1 - F(t) \quad (1)$$

Where R (t) is the reliability of prediction values detection process, and F (t) represent the fault ratio. In order to detect depression status in patients, The Cronbach's alpha test as a statistical test is used for reliability confidence of data. One of the most useful formula for estimating the validity of these types of tests or scales regardless of right or wrong answers is called the Kuder-Richardson method:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma^2}\right) \quad (2)$$

Where K represents input parameters, σ_i^2 is the variance of each parameter, and σ^2 is the total variance.

5. Proposed Method

In this study as mentioned above a unique dataset containing sensor data collected from patients with depression (Simula site). The dataset consists of activity recordings for 23 depressed patients and 32 healthy individuals, each file contains recorded activity per minute interval and this leads to very long record for processing. As a result, using improved approach for data analysis is necessary. The most important idea in this study is the pre-processing stage (feature-extraction stage), which consists of three basic stages according to Figure 2.

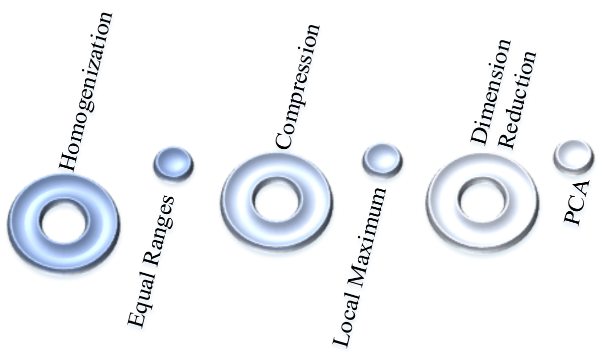

Fig. 2. Pre-processing steps

5.1. Proposed Feature Extraction

• First step, the alignment is performed, which involves determining equal ranges for the subjects under study (i.e.: same length records, for example 2 weeks' activity).

• Second step, the compression is expressed, using local optimizations as in Figure 5 (local maximum).

• Third step, reducing the dimension. PCA is used to supply the data required for the proposed algorithm.

As can be seen in Figure 3, the choice of window length is influential in determining the local optimum. The data were reduced to an acceptable and effective level after the pre-processing operation. So after feature-extraction stage the dataset is ready to fuzzification. This fuzziness is best characterized by its membership function. In other words, assume that the membership function represents the degree of truth in fuzzy logic, so the local maximum values in the previous stage can be used for fuzziness.

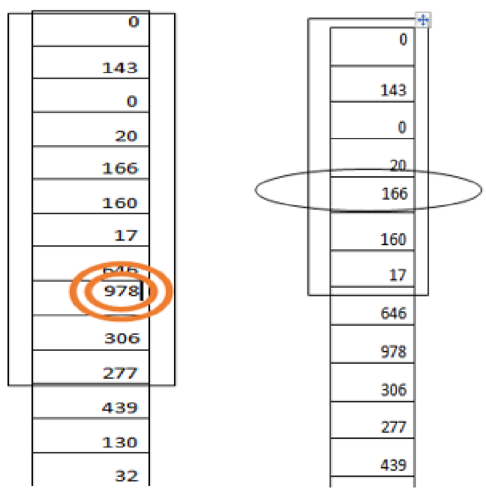

Fig. 3. Compression-Maximum.

5.2. Proposed Supervised Machine Learning

The proposed supervised machine learning is as follows: (figure 4)

- First step: The dataset fuzzification process (transforming a crisp quantity into a fuzzy quantity).
- Second step: The dataset is divided into training and testing groups.

- Third step: The training data is evaluated by one of the evolutionary algorithms selected in this research genetic algorithm (figure 8-9).
- Fourth step: When the stop condition is applied there are faced with a bunch of fuzzy rules, so can be evaluated these rules with test data.
- Fifth step: Defuzzification centroid method (convert the fuzzy results into crisp results).
- Sixth step: Comparing the actual labels of the test data with the labels obtained from the algorithm after new labelling.

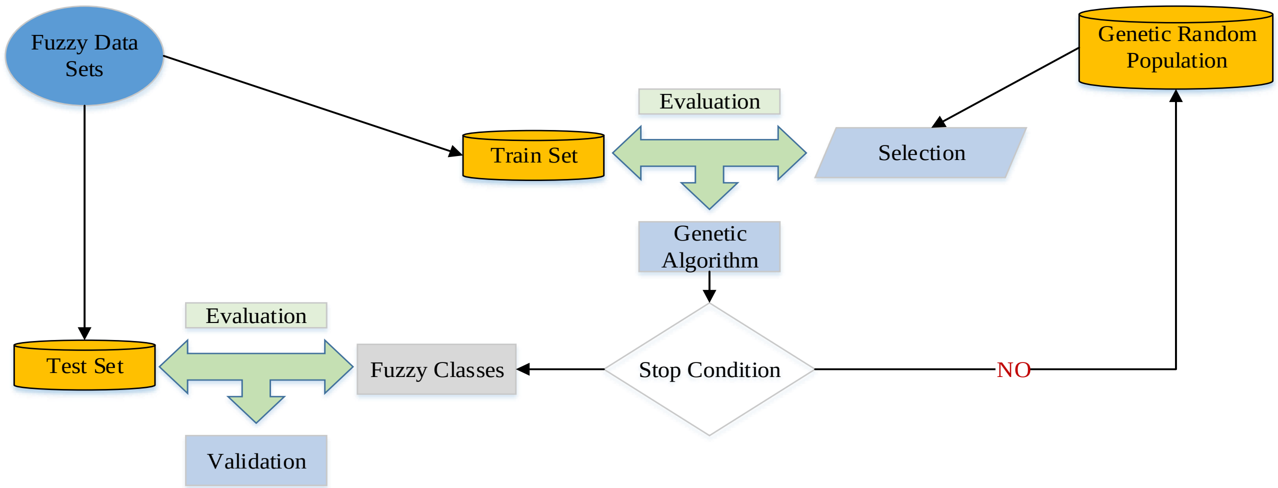

Fig. 4. Proposed Machine Learning Method

The proposed method for determining the fitness function is similar to the regression procedure in that when a record is generated from the initial random population, the difference is calculated with each individual patient and stored as a fitting value until the end of this fitting best. Approximate for all patients.

$$\text{Fittness} = \sum_{i=1}^{N} \text{sum}\left(\text{abs}\left(x - \text{population}\ (i)\right)\right) \qquad (3)$$

In this respect, N is the total number of training records, x vector (random population member) and patient population in the Excel file.

In the proposed approach, the goal is to reduce the difference intra-groups (i.e., the variations of data from the mean line for each group are minimized), and increase the inter-groups differences (i.e., increase the line spacing between the two groups. In other words, the two groups become more distinct). In this way, according to the labeled data, we get the rules (pattern) of each class of healthy and sick people. The entry is compared with the two rules (pattern) obtained and is placed in the category which has the least difference.

6. Result

The study dataset in the Simula Machine Learning site study has two named folders (control and condition) for healthy and depressed individuals, respectively. In cross-validation, the initial data are randomly divided into two subsets D1, D2, which are mutually incompatible and their sizes are divided into different iterations. The training and testing process of K is performed repeatedly and at the end of the mean, minimum and maximum parameters calculated in each division are shown in Table 6. However, the number of records in that dataset is very low beside that obtained results has significant improvement for both depressed and non-depressed cases.

Table 6

Results of Evaluation Criteria for Different cross-folds

Accuracy			**Precision**			**Recall**			**f-measure**			**Training set size**
Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	
0.812	0.925	1	0.801	0.925	1	0.816	0.927	1	0.8090	0.926	1	50
0.812	0.881	0.93	0.754	0.869	0.944	0.791	0.882	0.954	0.76	0.876	0.940	60
0.812	0.875	1	0.812	0.877	1	0.817	0.878	1	0.8149	0.878	1	70
0.812	0.906	1	0.809	0.902	1	0.783	0.913	1	0.7960	0.907	1	80

The evaluation phase and the results consist of two parts (improvement and development). First, the results of data mining metrics are compared with the results in Table 2 and according to Table 6, in the worst case (min column) the proposed method performed better than the results in Table 2 with at least 4% in each of the evaluation metrics values. (Improvement)

In the next section, the concept of reliability in data mining and Cronbach's alpha is used to show the reliability of the proposed method. As showed in table 7 faulty detections calculated for any cross-folds and worst case of faulty detection is bold in precision column value. However, all results have good success values. (Development)

Table 7

Results of Failure rate for Different cross-folds

Accuracy	**Precision**	**Recall**	**f-measure**	**Failure**
0.075	0.075	0.073	0.074	50
0.119	**0.131**	0.118	0.124	60
0.125	0.123	0.122	0.122	70
0.094	0.098	0.087	0.093	80

After calculation of faulty detections for folds, Alpha Cronbach values used to determine reliability range. As the result of table 8 is presented good performance for proposed method.

Table 8

Calculate Cronbach's alpha for a set of psychometric measurements

Standardized Alpha	**Unstandardized Alpha**
0.991863615819825	0.991593031375099
0.990791153785860	0.990186632085540
0.960177934616145	0.955547298566368
0.963922846194650	0.959720929873564

As a Cronbach's reliability and alpha coefficient calculation reliability summarized in table 9, and shows that with every consideration in proposed method results are high reliable.

Table 9

Cronbach's reliability ranges

Reliability	**Alpha Coefficient**
High	**Alpha >= 0.9**
Good	0.9 > Alpha >= 0.8
acceptable	0.8 > Alpha >= 0.7
unsure	0.7 > Alpha >= 0.6
weak	0.6 > Alpha >= 0.5
Unacceptable	0.5 > Alpha

Also, figure 5 shows that differences between validations results (ACC, ReC, PreC and F-val respectively stand for accuracy, recall, precision, f-measure) changes very slowly, this figure shows every slice of training folds.

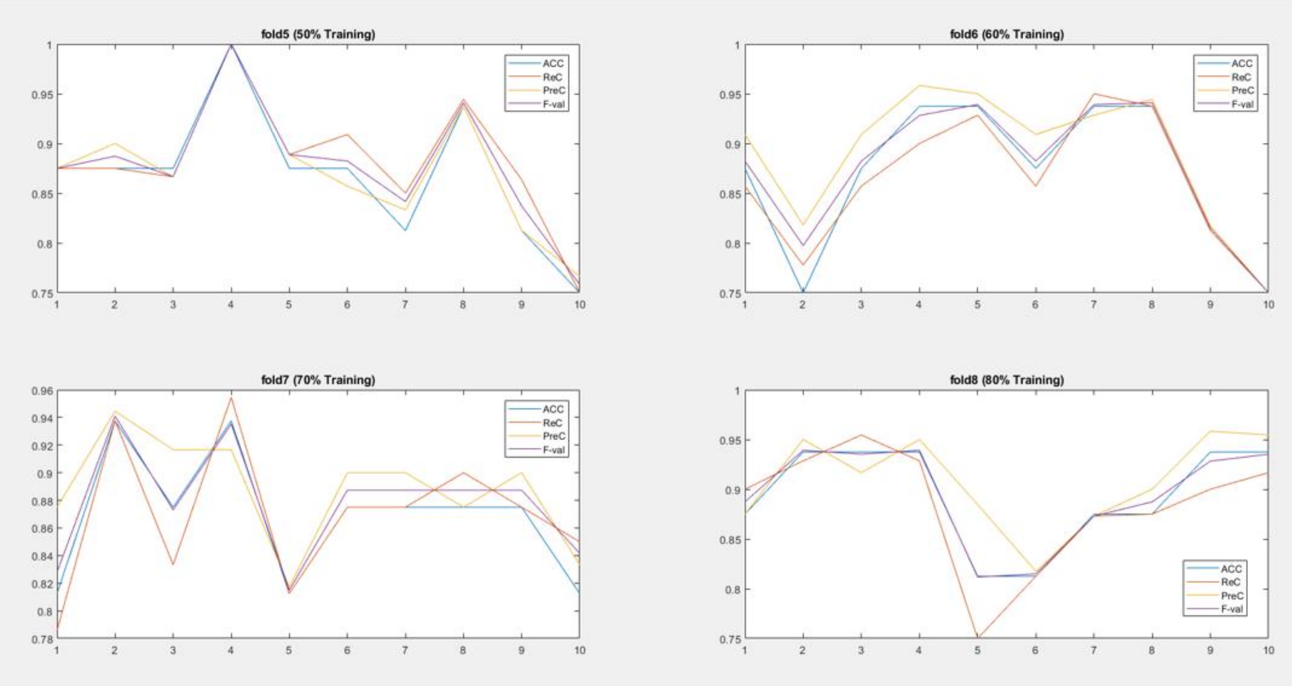

Fig. 5. Evaluation metrics in classification models

7. Conclusion and Future Work

The aim of this study was to improve the accuracy and reliability in the diagnosis of depression by presenting a new method of data mining. According to the performance and results obtained in this research, by presenting a new algorithm in machine learning, it has the best score among all researches so far, which is shown in Table 6. The proposed method mentioned in the extraction of features is one of the reasons for the success of this research compared to other works, in addition to the successful extraction of features, we used extra-innovative algorithms for training, which also increased the acceptable accuracy of the work Designing a reliable methods in software has been discussed for a long time, and therefore the authors of this article, by entering the discussion of reliability in data mining, have a look at the design of reliable predictors and diagnostics in the future.

References

[1] Depression: Overview, http://www.WHO.int/newsroom /factsheets/detail/depression.

[2] Susel G. A & Isabel de la T. & Sofiane H. & Miguel L.-C. & Diego C. B. & Lola M. N. & Manuel F. “Data Mining Algorithms and Techniques in Mental Health: A Systematic Review” Journal of Medical Systems (2018) 42: 161 https://doi.org/10.1007/s10916-018-1018-2

[3] Shahrzad Oveisi, Mohammad Ali Farsi, Mohammad Nadjafi, Ali Moeini “A New Approach to Promote Safety in the Software Life Cycle”, Journal of Computer & Robotics 12 (1), 2019 77-91

[4]Golriz Amooee, Behrouz Minaei-Bidgoli, Malihe Bagheri-Dehnavi, “A Comparison Between Data Mining Prediction Algorithms for Fault Detection (Case study: Ahanpishegan co.)” IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 3, November 2011 ISSN 1694-0814

[5] McCue, C. Data Mining and Predictive Analytics. Data Mining and Predictive Analysis, 31–48. doi:10.1016/b978-0-12-800229-2.00003-1, 2015.

[6] Larose, Daniel T., and Chantal D. Larose, Discovering knowledge in data: an introduction to data mining. Vol. 4. John Wiley & Sons, 2014.

[7] Susel G. A & Isabel de la T. & Sofiane H. & Miguel L.-C. & Diego C. B. & Lola M. N. & Manuel F. “Data Mining Algorithms and Techniques in Mental Health: A Systematic Review” Journal of Medical Systems (2018) 42: 161 https://doi.org/10.1007/s10916-018-1018-2

[8] Yang, S., Zhou. P., Duan, K., Hossain, M. S., and Alhamid, M. F., "emHealth: Towards emotion health through depression prediction and intelligent health recommender system". Mob. Netw. Appl.; 1-11, 2017.

[9] Jena, L., and Kamila, N. K., "A model for prediction of human depression using Apriori algorithm". 2014 Int. Conf Inf Technol.; 240-244, 2014.
[10] Jung, Y., and Yoon, Y. I., "Multi-level assessment model for wellness service based on human mental stress level". Mu/timed. Tools Appl. 76(9):11305-11317, 2017.
[11] Morales, S., Barros, J., Echavarri, 0., Garcia, F., Osses, A., Moya, C. et al., "Acute mental discomfort associated with suicide behavior in a clinical sample of patients with affective disorders: Ascertaining critical variables using artificial intelligence tools". Front. Psychiatr: 8:7, 2017.
[12] Mohammadi, M., Al-Azab, F., Raahemi, B., Richards, G., Jaworska, N., Smith, D. et al., "Data mining EEG signals in depres sion for their diagnostic value". BMC Med. Inform. Decis. Mak. 15(1):108, 2015.
[13] J. F. Dipnall, J. A. Pasco, M. Berk, L. J. Williams, S. Dodd, F. N. Jacka and e. al, "Fusing data mining, machine learning and tradi tional statistics to detect biomarkers associated with depression," PLoS One, vol. 11(2), pp. 1-23, 2016.
[14] M. Pirooznia, F. Seifuddin, J. Judy, P. Mahon, J. Potash and P. Zandi, "Data mining approaches for genome-wide asscociation of mood disorders," Psychiatr: Genet, vol. 22(2), pp. 55-61, 2012.
[15] H. Ni, X. Yang, C. Fang, Y. Guo, M. Xu and Y. He, "Data mining-based study on sub-mentally healthy state among residents in eight provinces and cities in China," J Tradit. Chinese Med, vol. 34(4), pp. 511-517, 2014.
[16] Sepideh Jenabi, Ghasem Kahe, “A Review of Depression Detection by Data Mining Methods”, 8th National Conference on Computer Science and Engineering and Information Technology, Iran. https://www.civilica.com/Paper-CECCONF08-CECCONF08_031.html
[17] Chang YS, Hung WC, Juang TY. "Depression diagnosis based on ontologies and bayesian networks". Proc - 2013 IEEE Int. Conf. Syst. Man., Cybern. SMC. ;3452-3457, 2013.
[18] Thanathamathee, P., "Boosting with feature selection technique for screening and predicting adolescents’ depression". 2014 4th Int Conf Digit Inf Commun Technol Its Appl DICTAP. ;23-27, 2014.
[19]Ghafoor, Y., Huang, Y. P., and Liu, S. I., "An intelligent approach to discovering common symptoms among depressed patients". Soft. Comput. 19(4):819-827, 2015.
[20]Hou, Y., Xu, J., Huang, Y, and Ma, X., "A big data application to predict depression in the university based on the reading habits". 2016 3rd Int. Conf Syst. Inform.,! CSA! ;1085-1089, 2016.
[21] Husain, W., Xin, L. K., Rashid, N. A., and Jothi, N., "Predicting generalized anxiety disorder among women using random forest approach". 2016 3rd Int ConfComput Inf Sci. ;37-42, 2016.
[22] Li, X., Hu, B., Sun, S., and Cai, H., "EEG-based mild depressive detection using feature selection methods and classifiers". Comput. Methods Programs Biomed. 136:151-161, 2016.
[23] Nie, Z., Gong, P., and Ye, J., "Predict risk ofrelapse for patients with multiple stages of treatment of depression". Proc. 22Nd ACM SIGKDD Int. Conf Knowl. Discov. Data Min.; 1795-1804, 2016.
[24] Spyrou, I. M., Frantzidis, C., Bratsas, C., Antoniou, I., and Bamidis, P. D., "Geriatric depression symptoms coexisting with cognitive decline: A comparison of classification methodologies". Bio med. Sign. Process Contrl. 25:118-129, 2016.
[25] Kim, J. Y., Liu, N., Tan, H. X., and Chu, C.H., "Unobtrusive mon itoring to detect depression for elderly with chronic illnesses". IEEE Sens. J 17(17):5694- 5704, 2017.
[26] Hadzic, M., Hadzic, F., and Dillon, T. S., Mining of patient data: "Towards better treatment strategies for depression". Int. J Funct. Inform. Personal. Med. 3(2):122-143, 2010.
[27] Goldberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley.
[28] Ranno Agarwal “Genetic Algorithm in Data Mining”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 9, September 2015
[29] Ghassan A.; Clay G.; Naser E.; Su Y.; Rajasree H. N.; Israa A. “Intelligent data mining and machine learning for mental health diagnosis using genetic algorithm” 2015 IEEE International Conference on Electro/Information Technology (EIT). DOI: 10.1109/EIT.2015.7293425
[30] Subhagata C. “A neuro-fuzzy approach for the diagnosis of depression” https://doi.org/10.1016/j.aci.2014.01.001
[31] K.S.Chandrasekaran, V.Mahalakshmi , M R Anantha Padmanaban ,”Measuring Performance Reliability of a system using Data Mining” , 2019 11th International Conference on Advanced Computing (ICoAC) 978-1-7281-5286-8/IEEE 10.1109/ICoAC48765.2019.246834, ©2020
[32] Mansah, “A Tour of Evaluation Metrics for Machine Learning”, https://www.analyticsvidhya.com/blog /2020 / 11/a-tour-of-evaluation-metrics-for-machine-learning/#, November 24, 2020.
[33] M. Finkelstein, “Failure Rate Modeling for Reliability and Risk” Springer Science & Business Media, 2008 M11 7
[34] Sepideh Jenabi, Ghasem Kahe, “A Review of Depression Detection by Data Mining Methods”, 8th National Conference on Computer Science and Engineering and Information Technology, Iran. https://www.civilica.com/Paper-CECCONF08-CECCONF08_031.html

[35] Sadeh, Avi; Acebo, Christine (May 2002). "The role of actigraphy in sleep medicine". Sleep Medicine Reviews. 6 (2): 113–124. doi:10.1053/smrv.2001.0182.

[36] S. Yoon, B. Taha and S. Bakken, "Using a data mining approach to discover behavior correlates of chronic disease: A case study of depression," Stud. Health Technol. Inform, vol. 201, pp. 71-78, 2014

[37] C. Lee, C. P. Lam and M. Masek, "Rough-fuzzy hybrid approach for identification of bio-markers and classification on Alzheimer's disease data," Proc. 2011 llth IEEE Int Conf Bioinforma Bioeng BIBE, pp. 84-91, 2011.

[38] Guyon, I., and Elisseeff, A. (2003) An intro¬duction to variable and feature selection, Journal of Machine Learning Research 3, 1157-1182.

[39] Enrique Garcia-Ceja, Michael Riegler, Petter Jakobsen, Jim Tørresen, Tine Nordgreen, Ketil J. Oedegaard, Ole Bernt Fasmer, Depresjon: "A Motor Activity Database of Depression Episodes in Unipolar and Bipolar Patients", In MMSys'18 Proceedings of the 9th ACM on Multimedia Systems Conference, Amsterdam, The Netherlands, June 12 - 15, 2018.