Classification of Parkinson's disease using machine learning algorithms: a comparative study

Dr.Mohamed Mamoun Abdulerahim

Faculty of computer and technology, Alzaiem Alazhari University

Dr.Ashraf Osman Ibrahim

Faculty of computer and technology, Alzaiem Alazhari University

Abstract.

This study aims to diagnose Parkinson's disease using machine learning, five different classification techniques were applied to dataset obtained from UCI Machine Learning Repository.We conducted five experiments, using tow class perceptron, tow class decicion forest , tow class boosted decicion tree , tow class support vector machine and tow class decicion Jungle, The boosted decicion tree technique achieved the best results, as the classification accuracy was 89.7 While the perceptron ranked last with the worst result with 75.9 accuracy The paper also discussed the reasons for the superiority of algorithms over others, at the end of the project, we created a web page and it was linked to the classification model so that the data could be entered and the diagnosis would be made directly.

General Terms :**Parkinson Disease**,Artificial **N**eural **N**etworks**,D**ecicion**F**orest **, B**oosted **D**ecicion**T**ree**, S**upport **V**ector **M**achine **.**

Keywords: Parkinson's disease diagnosis,Dataminig, Ensemble methods.

تشخيص مرض الشلل الرعاشي باستخدام خوارزميات تعلم الالة : دراسة مقارنة

د. محمد مامون عبدالرحيم – كلية علوم الحاسوب وتقانة المعلومات – جامعة الزعيم الأزهري
د.أشرف عثمان إبراهيم - كلية علوم الحاسوب وتقانة المعلومات – جامعة الزعيم الأزهري

المستخلص:

هدفت هذه الدراسة إلى تشخيص مرض الشلل الرعاشي(باركنسون) باستخدام خورزميات تعلم الآلة ،حيث يعتبر المرض من الامراض المعقدة والتي يصعب تشخيصها في المراحل الاولى , حيث تم تطبيق خمس خوارزميات تصنيف مختلفة على مجموعة البيانات التي تم الحصول عليها من UCI Machine Learning Repository ، وأجرينا خمس تجارب ، حيث قمنا باستخدام الشبكات العصبية ، وشجرةالقرار ، وشجرةالقرار المعززة ، وآلة المتجهات الداعمة غابة القرار ، حققت خوارزمية شجرةالقرار المعززة أفضل النتائج ، حيث كانت دقة التصنيف 89.7 بينما احتلت الشبكة العصبية المرتبة الأخيرة مع أسوأ نتيجة بدقة 75.9.

الكلمات المفتاحية : مرض الشلل الرعاشي , الشبكات العصبية الاصطناعية , مجموعة القرار , شجرة القرار المعزز , الة المتجهات الداعم.

1.Introduction

Parkinson's disease is a progressive disorder that affects the nervous system and affects movement. Symptoms begin gradually, sometimes with tremors that may not be noticed in one hand, and then worsen over time. Although tremors are common, the disorder causes stiffness and slows down , To this day, there is no cure for the late stages of the disease, but it is easier to control the symptoms in its early stages if it diagnosed early,Therefore, there is an urgent need to discover and diagnose this disease in its early stages, as it is mysterious and invisible ,The computer made a very big breakthrough in the field of diagnosing complex diseases)1(,In recent decades, machine learning has made a fundamental shift in diagnosing diseases in general, and complex and mysterious diseases in particular , As there are a large number of academic and non-academic centers and institutions collected huge patient dataand this data was made freely available to researchers, With the development day after day, the computer, along with various other sciences, such as applied mathematics and dataminigcontributed to the diagnosis of a very large number of complex diseases such as Parkinson's disease, Alzheimer's disease And cancer,

Where researchers have developed a number of techniques and methods to diagnose problems in general and are used by doctors with technical support from computer expert.

The main contributions of this study are as follows We used three optimized methods against two traditional methods without any improvement. The experiment demonstrated the superiority of the improved methods in classifying a complex disease such as Parkinson compared to the unimproved methods. We also created a web page to enter the disease data and the case will be diagnosed immediately.

This paper organized as follow: In Section 1, rntroductio and related , Material and methods is explained in Section 2, the experimental design is described in detail and The experimental results and a discussion are presented in Section 3. Finally, Section 4 cona cludes the paper.

2.Related Works

In the last few decades, the research about algorithms enhancement has been increased extremely specially with algorithms that make classification in the medical filed because the great need for accurate results. Here we will survey a related works in Parkinson Disease classification using different AI Techniques.

In (2) this work Dr. R. Geetha Ramani and G. Sivagami maked a survey of dataming techniques to classify Parkinson Disease. This study was a comparative study to classify healthy and people with Parkinson Disease, they applied the feature relevance on the training dataset and they used Various Feature selection algorithm , the classification algorithm is used as a black box to find the best subset of attributes. They got the dataset for the disease from UCI , The training dataset comprises of 197 instances with 22 characteristic features. Thet compared between following algorithms : Binary Logistic Regression , ID3, C4.5, Classification and regression Tree, K- Nearest Neighbor , Random Tree (Rnd Tree) , Partial Least Square Regression and Support Vector Machine. The main goal was to know the algorithm with highest accuracy . after

perform the experiment . we found that The Random Tree Algorithm classifies the Parkinson Disease dataset accurately and provides the 100% with zero error rate. In)3(AnchanaKhemphila and Veera Boonjing used the Multi-Layer Perceptron (MLP) with-Back-Propagation learning algorithm and the feature selection algorithm to classify the presence of Parkinson disease ,the also took dataset from the University of California at Irvine (UCI) machine learning repository . In this study, the number of attributes has been reduced from 22 to 16 attrbutes , they used Weka tool v3.6.6 , the problem was the inputs are continuous real numbers ,they slove that by by partitioning the range of values into a finite number of subsets , they removed innecessary features by using use Information gain filter which do not need to take from patients , this study show that feature selection contribute to increase efficiency of computation and classification accuracy and simplise the complexity by reduce data. Firstly they used ANN without information gain based on feature selection function , in this case accuracy was 80.769% , after using information gain based on feature selection function the accuracy increased up to 83% , so we deduct that using information gain based on feature selection function with ANN contribute to enhance the classification process ,In)4(K. Witt, A. Nühsman and G. Deuschl worked on the advanced Parkinsion Disease by Intact artificial grammar learning this case study was conducted on patients with cerebellar degeneration and advanced Parkinson's disease , 21 Parkinson's disease (9 women and 12 men) , 10 patients with cerebellar degeneration (CD) and 15 control subjects were tested on artificial grammar learning , The mean age of PD patients was 58.3 years , in the time of experimentation all patients were stable obtaining anti-parkinsonian medication . The artificial grammar learning test originally described as a paper-and-pencil test was used in a computer adopted version , In the training phase of the artificial grammar learning task , controls were able to repeat 96.67% of the letter strings in the correct order , The mean for CD patients is 58.39%(S.E.M.: 2.44%); the

mean for all PD patients is 61.48% . this experimentation show that patients with cerebellar degeneration were less successful in reproducing the letter strings during the training phase of the task While patients with Parkinson's were doing their best in this experiment .in)5(they tried to discriminate PD subjects from healthy and test the accuracy of novel algorithms , those algorithms applied on speech signals as data set. These data was received from The National Center for Voice and Speech (NCVS) , dataset consist of 43 samples (17 females and 26 males , 10 healthy controls and 33 PWP) , persons in this case study had age range of 46 to 72 years. signal to noise ratio (SNR) was used to estimate measures quantifies noise . they used four feature selection algorithms to make dysphonia measures , those four feature selection applied with stasticalclassifer like random forests and support vector machines to map these feature subsets to a binary classification. this study aimed to analyze the speech signals ,also compute jitter and shimmer of features , The motivation for these features is that the vocal fold vibration pattern is periodic pattern is considerably disturbed in pathological cases whereas nearly periodic in healthy voices, in this study the proposed dysphonia measures which have not been previously used in this application (as auther said) , all the dysphonia measures in this study were computed anew using the algorithms described in "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson"s disease symptom severity" , they used used nonlinear SVMs for mapping features to the response , with the best results reporting approximately 93% classification accuracy on a subset of 22 features and they Proved that they can achieve almost 99% accuracy using 10 dysphonia measures .in)6(this paper compared different types of classification methods for for effective diagnosis of Parkinson disease , Where they were looking for the most effective way to diagnose this serious disease and distinguish the healthy people from the affected , they were applied four classification methods for this purpose, These methods are

respectively Neural Networks, DMneural, Regression and Decision Tree , they used SAS Enterprise programe to conduct data pre-processing and they used SAS Enterprise Miner 5.2 program to analyze and recognize the PD by combining several classification methods with model comparison node , the PD database composed of 23 columns and 197 rows , The data is in ASCIICSV format, The dataset was divided into two categories training dataset and testing dataset , They did not specify the source of the evidence (dataset) in this paper, This Experiment shows that The backpropagation learning algorithm has been used in the feed-forward, single hidden layer neural network with 10 nerounintial weights were randomly selected , after conduct the Experiment we see that neural network is best classifier its obtainedhigher classification score and gained 92.9 classification accuracy while Regression gained 88.6 classification accuracy and DMNeural classifier and Decision tree classifier gained 84.3 classification accuracy. In)7(Hui-Ling Chen , Chang-Cheng Huang and Su-Jing Wang have established an effective diagnostic system for Parkinson's disease , they used fuzzyk-nearest neighbor approach (FKNN) for detection of this dangerous dieses , But before that they compared (FKNN) with support vector machines (SVM) based approaches , the aim of this study was to improve accuracy for detection of Parkinson's disease So it can be said that it is a comparative study before everything , the PD data set was taken from from UCI machine learning repository , the data set is a results of various medical tests carried out on a patient and composed of a range of biomedical voice measurements from 31 people, 23 with PD thos people have ages ranged from 46 to 85 years , the data set had no missing values , The proposed diagnosis system work in tow stages ,in the first stage they used PCA approach to Exclude the redundant features and thus enhancefurther the classification performance , in the second stage they run the FKNNmodel wich start training on the training set , finally the optimal FKNN model Will be used to conduct the classification task ,

This model can simply be summarized as follows , parkinsion data set is normalized by scaling to range 0 and 1 then they conduct the features reductions to reduce feature space after that the divided the data set into training data set and test data set(each timeone of the 10 subsets is used as the test set and the remaining 9 subsetsare used as a training set) using 10-fold cross validation to find the optimal parameter combination (k,m) that making up the FKNN model , the 10-fold CV was used to evaluate the classification accuracy, After the experiment a comparison was made between FKNN model and prominent support vector machine (SVM) classifier In order to validate the superiority of the proposed FKNN-based diagnosis system , FKNN classifier was implemented on an Intel Quad-Core Xeon 2.0 GHz CPU using MATLAB language under Windows Server 2003 environment , The comparison result was in favor of proposed FKNN-based diagnosis system wichgaind best classification accuracy 96.07% while SVM classifier gaind accuracy 93.47 , This experiment also compared the result with a number of methods in previous studies , for example Parallel NN gaind accuracy 91.20 , Gravitational searchalgorithm + OPF gained accuracy 84.01 , Particle swarm optimization + OPF gained accuracy 73.53 , Fuzzy entropy measures + similarity gained accuracy 85.03 , ANN gained accuracy 92 , Improved mRVMs gained accuracy 89.47 , Pre-selection filter + Exhaustive search + SVM gained accuracy 91.4 and Dirichlet process mixtures gaind accuracy 87.7.

3.Material and methods

2.1Parkinson Disease

Since the case study in this research is the diagnosis of Parkinson's disease, so we will briefly review some information about the disease , Most of us at one time or another have noticed our hands tremble when we are holding a drink, or our teeth chatter when it is very cold out, Sometimes though a small tremor is one of the first signs of Parkinson's disease. Parkinson's disease is a

chronic, progressive, movement disorder that affects muscle control and balance. These symptoms gradually worsen over time, and generally other secondary symptoms start to occur, Parkinson's disease is caused by a gradual deterioration in the function of nerve cells, or neurons, located in a specific area of your brain. Your brain contains millions of neurons, which process and transmit information. It receives information via neurons from all parts of your body, processes it, and then sends out instructions to other parts of your body to produce specific actions. The information travels along neurons as electrical impulses. When a nerve impulse reaches the junction between two neurons, a chemical neurotransmitter is released that stimulates the impulse in the next neuron so that it continues on to its destination. The primary symptoms of Parkinson's are related to movement, and vary from one person to the next. Sometimes at the onset of Parkinson's, you may not even be aware that you have any symptoms. Gradually though, and usually over a period of years, symptoms do become noticeable, and are often of three distinct types: tremors , bradykinesia and rigidity)8(,Parkinson's disease can be diagnosed in many ways , One of these ways is Vocal Tests through sounds which may be letters (sustained vowels) or words or even sentences Which is common or by testing the movement of walking through sensors placed in a certain way in the feet and then recorded this observation data as happened in one of the papers discussed in the previous chapter , Or by brain images which are filmed by CT scan radiation but what is rare is treated with this type of data because it is inaccurate.

2.2 **Artificial Neural Network**

Artificial neural network (AAN) is a model that simulates the natural neural network, which takes the same principle of work, it is a complex mathematical algorithm rather not subject to fixed laws, As is well known the human brain can be described as a biological neural network, Brain consists of a dense tangled group of neurons. Each nerve cell, called a neuron .the human brain consist

of 10 million neurons and 60 trillion complexity of an engagement point (nervous) to each other. Since the biological neural network consists of the cell body (soma), Dendrites (inputs), axon (outputs) and synapses. Soma collects the signals which receives from Dendrites. Dendrites are sensors that receives signals from another nervous cells, these signals will be sent via Axon through synapses. Synapses is a point that connects axon to the rest of the neurons Artificial neural network is a simulation of real neural network, we can say it's a program that processes data like (a part of) the nervous system. Its learns from examples like the brain, this process of learning called "training". Where the network is training and each time the weights are adjusted until reaching the optimal solution. Simply neuron receives many signals as inputs and then doing some processors and then automatically send the signals that have been processed to the next neuron as outputs. These outputs are or may be the ultimate solution,to calculate the output neuron, it uses a function called activation functions. These functions are sign function, step function, sigmoid function and linear function[9].

2.3**DecicionForest algorithm**

Random Decision Forests is a learning method used for Classification and regression tasks , this algorithm works on building many decision treesDuring the training process,the prediction process is improved as each decision tree performs the prediction process based on the given historical data , a large number of trees are being created to form the tree forest and tree predictions are combined,this algorithm is usually used for economic forecasts as well as forecasting weather conditions and forecasting bank profits and losses , The idea of creating a decision tree is very simple. Let's explain an example of how to create a single decision tree , Imagine that we have a set of data that we want to separate based on its characteristics. For example 110000 , if we have the following statement and we want to specify the red color that is underlined

by a line , The tree will be as follows :

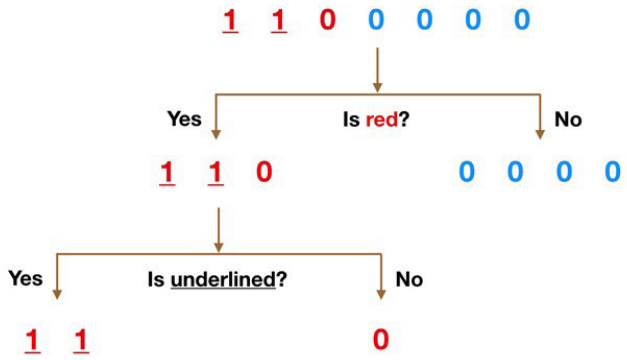


Fig 1:Simple Decision Tree Example

As shown in the picture above, the data will first be isolated based on the red color, so the red color was placed in the left branch and the blue color in the right branch. In the next step, the isolation was done by the characteristic of the line under the letter, so the data underneath the line in a branch and the other in another branch simply became This simple tree is able to isolate data according to these two characteristics,decicion Forest algorithm operate as an ensemble ("use multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone"))11()10(.

2.4Boosted Decicion Tree

A decision tree takes a set of input features and splits input data recursively based on those features , The data is split based on a value of one of the input features at each node (interior nodes) and evry node has Terminal nodes called (Leaves),Trees can be built with branches splitting into many sub-branches, The upgrade process (Boosting) is a method ofcombining weak trees into a powerful and stable classifier by combining many learners with fixed error rate, with the goal of improving the prediction process , There are many different ways of iteratively adding learners to minimize a loss function such as AdaBoost , Gradient Boosting and XGBoost, the accuracy of the algorithm is improved by fitting

the previous trees But that might be alead to less coverage risk , The idea of the algorithm can be summarized in the following steps :

Step 1 : Train classifier C1 on N events

Step 2 : Train C2 on new N-sample, half of which misclassified by C1

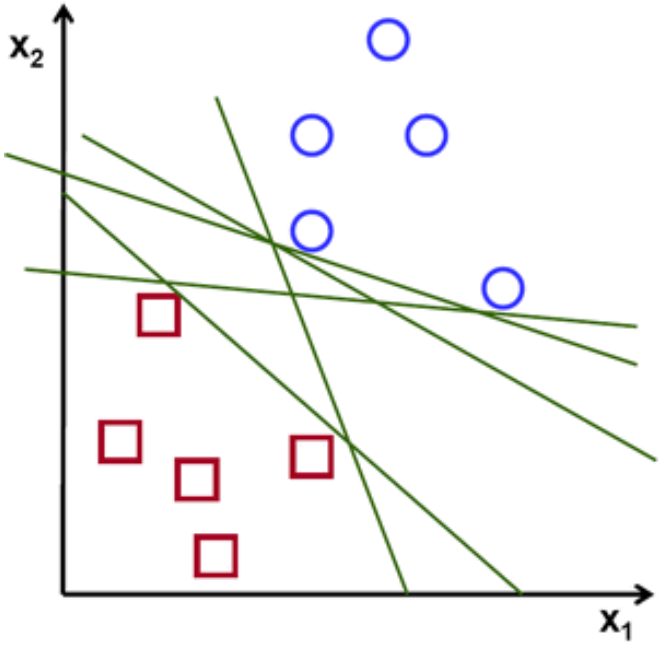
Step 3 : Build C3 on events where C1and C2disagree

Step 4 : Boosted classifier: MajorityVote (C1,C2,C3).

Boosting Decicion Tree is also type of ensemble method, It should be noted that the Boosting process differs from "Bagging " , it used when our goal is to reduce the variance of a decision tree by creating several subsets of data from training sample chosen randomly with replacemen .)13)(12(

2.5**Binary support vector machine**

SVM is a linear algorithm that separates or isolates two types of data through a line at a level(Decicion line) and performs classification by finding the hyperplane that maximizes the margin between the two classes See figure no 2 :



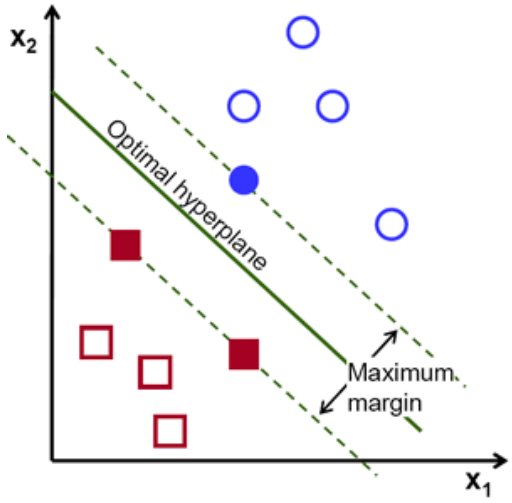


Fig 2: Possible hyperplanes

The algorithm looks for a line that gives biggest value to the marginwith maximum distance between data points of both classes , The higher the margin distance, the greater the confidence in separating future data , hyperplanes ideally deals with two types of data (two class) and it draws a line and cannot deal with a third type of data, svm performs classification , regression and clustering tasks ,Svm algorithm can be described in the following steps :

1. Define an optimal hyperplane (maximize margin)
2. Extend the above definition for non-linearly separable problems.
3. Map data to high dimensional space where it is easier to classify with linear decision surfaces.

" However, there are situations where a nonlinear region can separate the groups more efficiently SVM handles this by using a kernel function (nonlinear) to map the data into a different space where a hyperplane (linear) cannot be used to do the separation ")15)(14).

2.6**Decision Jungle**

Decision Jungle Algorithm is a development of Decision Forest algorithmor recent extension to decision forests ,It can be said

in a nutshell its Basically a faster, lower memory decision tree , as there are some limitations in Decision Forest algorithm that the number of nodes will increase exponentially with depthBecause memory is a limited resource, the rapid growth of trees limits its depthand that definitely affects accuracy , Unlike traditional decision trees that allow only one path per node, the decision Jungle outgrow this limitation by allowing multiple paths from the root to each leaf , decision jungle consists of an ensemble of decision directed acyclic graphs (DAGs), during training the node is split and the node is merged by reducing the same objective function and require less memory while considerably improving generalization , Although this is at the expense of training time, by allowing tree branches to merge, this leads to better generalization performance than a decision tree, This improved algorithm is more used with multi-class classification.

2.7**Dataset Description**

The dataset was created by Max Little of the University of Oxford)16(, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals. The original study published the feature extraction methods for general voice disorders. This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Each column in the table is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column). The main aim of the data is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD. The data is in ASCII CSV format. The rows of the CSV file contain an instance corresponding to one voice recording. There are around six recordings per patient, the name of the patient is identified in the first column, The following table shows the attributes)18(, see table no 1

Attribute	Description
Name	ASCII subject name and recording number
MDVP	Fo(Hz) - Average vocal fundamental frequency
(MDVP:Fhi(Hz	Maximum vocal fundamental frequency
(MDVP:Flo(Hz	Flo(Hz) - Minimum vocal fundamental frequency
MDVP Jitter(%),MDVP:Jitter(Abs),MDVP:RAP,MDVP:P-- PQ,Jitter:DDP	Several measures of variation in fundamental frequency
MDVP	Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA - Several measures of variation in amplitude
NHR,HNR	Two measures of ratio of noise to tonal components in the voice
status	Health status of the subject (one) - Parkinson's, (zero) – healthy
RPDE,D2	Two nonlinear dynamical complexity measures
DFA	Signal fractal scaling exponent
spread1,spread2,PPE	Three nonlinear measures of fundamental frequency variation

Table 1: Attribute information

4.Classification Model

The classification model reads Parkinson's disease dataset wich described in the previous paragraph, The next step normalizes the data to eliminate data distortions and make them balanced,Then the model divides the data into training data by 80%, and this is common, and test data by 20% , Then the training process is conducted according to the chosen classifier with the train model , after the training process is over, the success of the process is tested using test data using a Score Model , Finally, the results of the training and testing process are summarized using the evaluation model , See figure no 3.

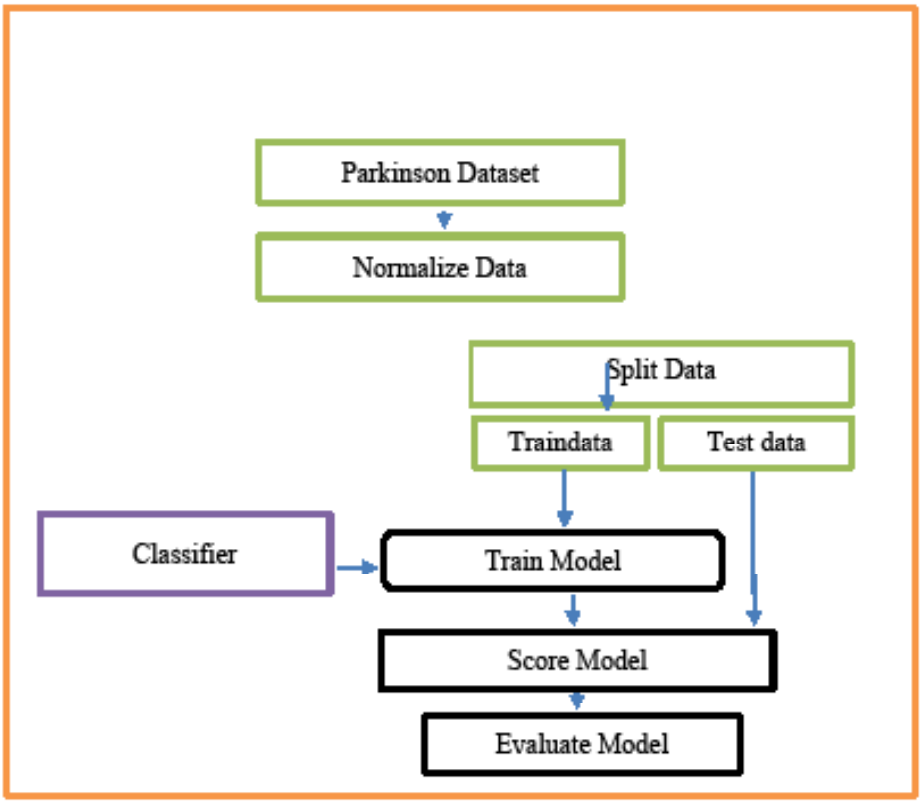


Fig3 :classification Model

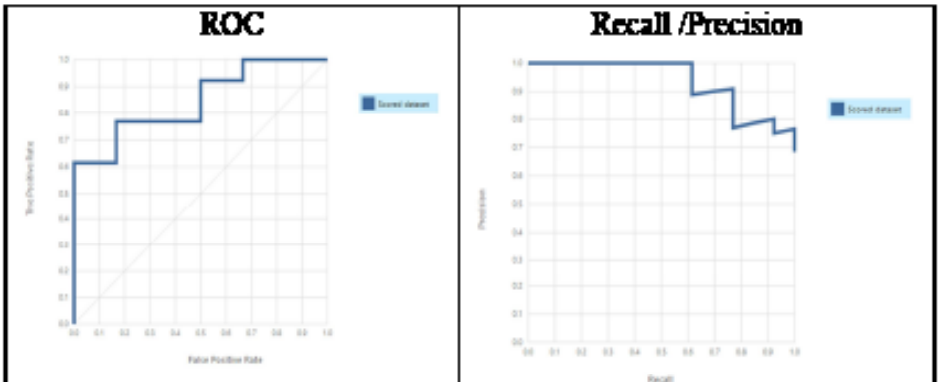
3.1 Results & Discussion

In this paragraph, we will explain the results of carrying out the experiment above using the five methods of the paperusing Microsoft Azure Machine Learning Studio, where the results were

as follows:

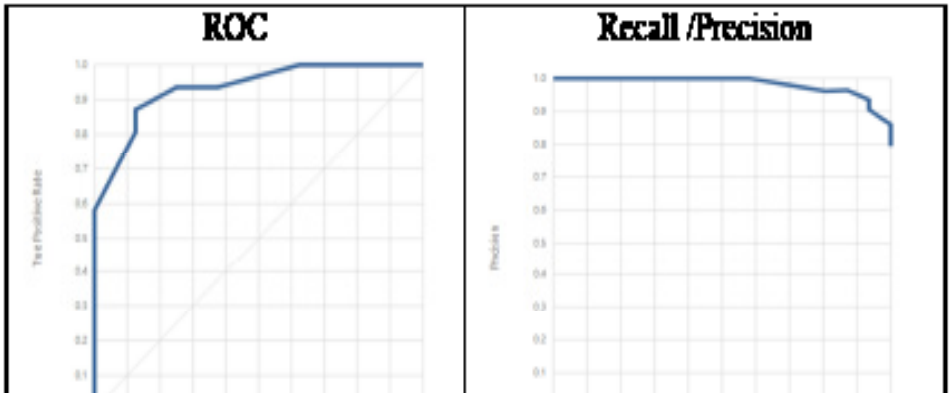
1. Tow-class perceptron :

True positive	False positive	True negative	False negative
32	4	10	12
Accuracy	Precision	Recall	F1 score
0.759	0.889	0.762	0.821



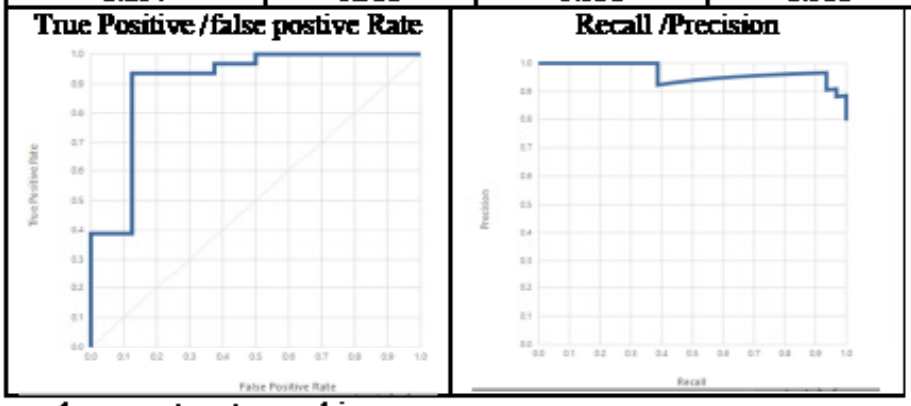
2.Tow-class decicion forest

True positive	False positive	True negative	False negative
31	6	0	2
Accuracy	Precision	Recall	F1 score
0.846	0.838	1.000	0.912



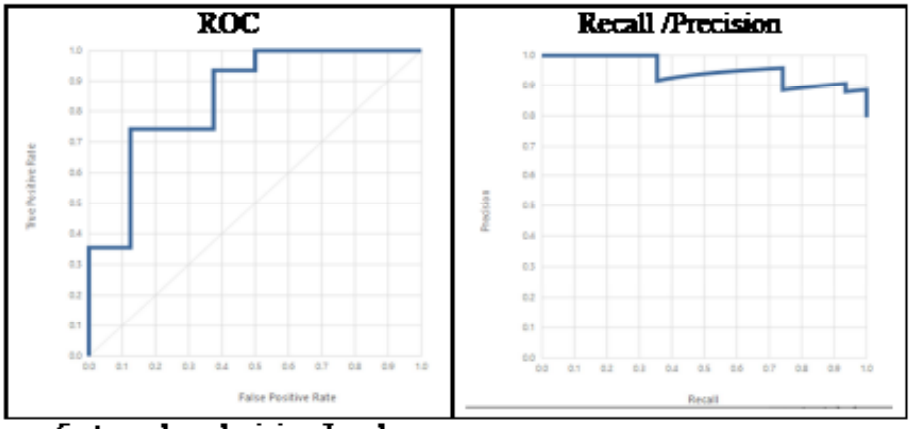
3. **tow-class boosteddecicion tree**

True positive	False positive	True negative	False negative
29	2	6	2
Accuracy	Precision	Recall	F1 score
0.897	0.935	0.935	0.935



4. **support vector machine**

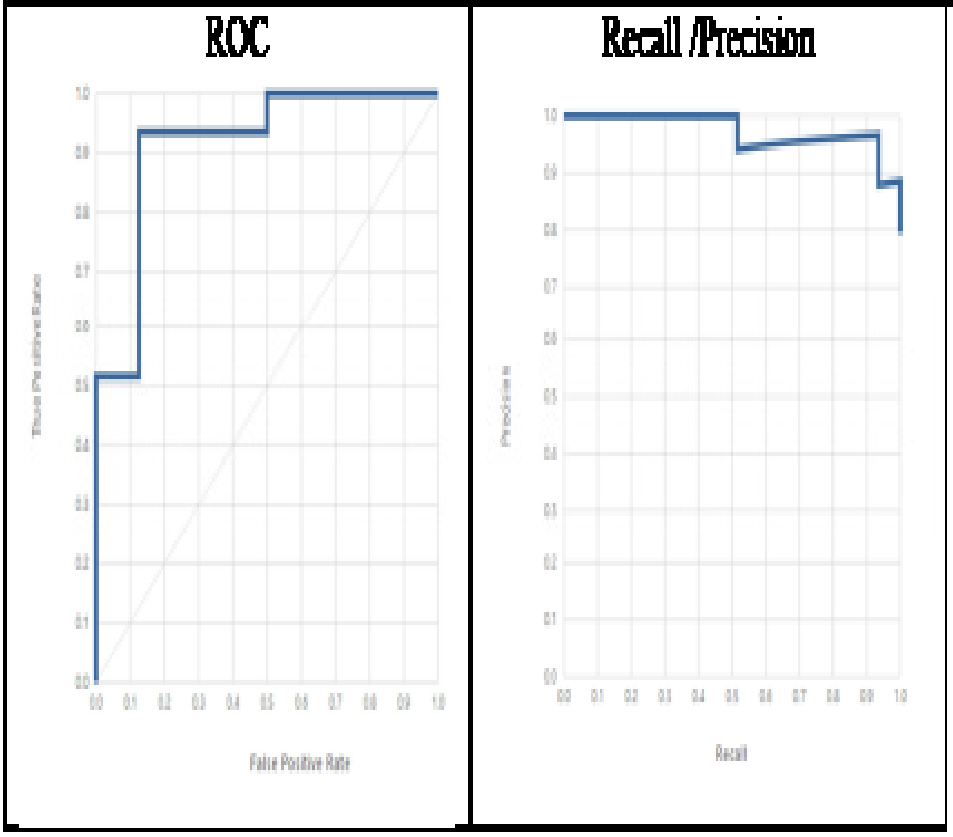
True positive	False positive	True negative	False negative
27	3	5	4
Accuracy	Precision	Recall	F1 score
0.821	0.900	0.871	0.885



5. **tow-class decicion Jungle**

True positive	False positive	True negative	False negative
29	3	5	2
Accuracy	Precision	Recall	F1 score

0.842	0.906	0.935	0.921



Through the five experiments, we notice the superiority of boosted decicion treealgorithm , as it achieved the highest classification accuracy, reaching approximately 90%, decicion Jungle algorithm ranked second Strictly very good classification accuracyreached 87%, decicion forest algorithm ranked third with classification accuracy not relatively bad reached 84.6% while support vectors machine ranked fourthwith Not good classification accuracy compared to the first three reached 82%, Finally The PerceptronNeuralNetwork achieved the worst score with classification accuracy reached 76% as the worst result score among the five methods.

3.2Web service :

A web service has been created and linked to the classification model , the patient will enter the data and get immediate results see figure no 4.

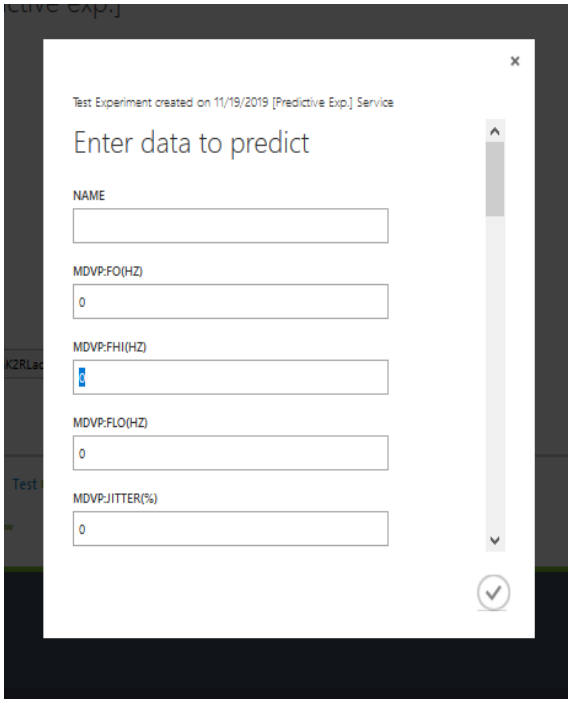


Fig 4: web service snapshot

4Conclusion and future work

It is noted through the results above that the superiority of the improved algorithms and this is very normal thing and expected, Whereas, the decicion forest algorithm is an improvement and development of the regular decision tree , Likewise the decision jungle algorithm is a development of the decision forest algorithm, so it result was better than decision forest , also the improved decision tree was the best result because it came after a series of improvements and uses Ensemble While both support vectors machine and Perceptron achieved unsatisfactory resultsBecause they were used without an optimized model,Therefore, we find that the researchers have done studies to develop the work of perceptron by obtaining the best values for the structure of the neural network by obtaining the hyperparameter, which is obtained by hybridizing some algorithms with each other

References

(1) https://www.khanacademy.org/science/health-and-medicine/nervous-system-diseases/parkinsons-disease/a/what-is-parkinsons-disease , [9/12/2019 ,04: 19 PM].

(2) Dr. R. Geetha Ramani, G. Sivagami, "Parkinson Disease Classification using Data Mining Algorithms". Volume 32– No.9, October 2011.

(3) Anchana Khemphila and Veera Boonjing ,"Parkinsons Disease Classification using Neural Network and Feature selection " ,World Academy of Science, Engineering and Technology International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering Vol:6, No:4, 2012.

(4) K. Witt, A. Nühsman, G. Deuschl , "Intact artificial grammar learning in patients with cerebellar degeneration and advanced Parkinson's disease " ,Neuropsychologia 40 (2002) 1534–1540.

(5) Athanasios Tsanas*, Max A. Little, Patrick E. McSharry, Senior Member, IEEE, Jennifer Spielman, Lorraine O. Ramig ,"Novel speech signal processing algorithms for high-accuracy classification of Parkinson"s disease " ,TBME-00887-2011.R1.

(6) (Resul Das ,Department of Informatics, Firat University, 23119 Elazig, Turkey,"comparison of multiple classification methods for diagnosis of Parkinson disease " ,Expert Systems with Applications 37 (2010) 1568–1572.

(7) Hui-Ling Chen a, Chang-Cheng Huang a, Xin-Gang Yu b, Xin

Xu c, Xin Sun d, Gang Wangd, Su-Jing Wang d , "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach",Expert Systems with Applications 40 (2013) 263–271.

(8) https://www.khanacademy.org/science/health-and-medicine/nervous-system-diseases/parkinsons-disease/a/what-is-parkinsons-disease ,[28/10/2019 , 8:21 PM] .

(9) (21)https://towardsdatascience.com/understanding-neural-networks-19020b758230 ,[16/9/2019 ,3: 19 PM].

(10) https://towardsdatascience.com/understanding-random-forest-58381e0602d2,[11/12/2019 ,7: 13 PM].

(11) L. Breiman, "Random forests",Machine Learning, 45 (1), 5–32, 2013

(12) Y. Freund, "Boosting a weak learning algorithmmajority",Information andcomputation. 121(2):256–285, 1995.

(13) https://www.quora.com/What-are-boosted-decision-trees , ,[6/12/2019 ,8: 22 PM].

(14) https://www.geeksforgeeks.org/classifying-data-using-support-vector-machinessvms-in-python/ , ,[6/12/2019 ,10: 14 PM].

(15) https://www.analyticsvidhya.com/blog/201709//understaing-support-vector-machine-example-code/ ,[7/12/2019 ,8: 16 PM].

(16) http://primo.ai/index.php?title=Decision_Jungle, [7/12/2019 ,10: 27 PM].

(17) https://dl.acm.org/citation.cfm?id=2999638 , [9/12/2019 ,04: 19 PM].

(18) https://archive.ics.uci.edu/ml/datasets/parkinsons, [18/10/2019 ,09: 34 PM].