

Parkinson disease classification: a comparative analysis on classification techniques

Nazri Mohd Naw¹, Mokhairi Makhtar², Zehan Afizah Afip³, Mohd Zaki Salikon⁴,
Khairul Amin Mohamad Sukri⁵

^{1,3,4,5}Soft Computing and Data Mining Centre (SMC), Faculty of Computer Science and Information Technology,
Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia

²Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia

Article Info

Article history:

Received Oct 12, 2019

Revised Dec 13, 2019

Accepted Dec 27, 2019

Keywords:

Artificial neural networks
(ANN)

Decision tree

Parkinson's disease (PD)

Comparative analysis

ABSTRACT

Parkinson's disease (PD) among Alzheimer's and epilepsy are one of the most common neurological disorders which appreciably affect not only live of patients but also their households. According to the current trend of aging social behaviour, it is expected to see a rise of Parkinson's disease. Even though there is no cure for PD, a proper medication at the early stage can help significantly in alleviating the symptoms. Since, the traditional method for identifying PD is rather invasive, expansive and complicated for self-use, there is a high demand for using classification method on PD detection. This paper compares the performance of Neural Network and decision tree for classifying and discriminating healthy people for people with Parkinson's disease (PD) by distinguishing dysphonia. The simulation results demonstrate that Neural Network outperformed decision tree by giving accurate results with 87% accuracy as compared to decision tree with only 84% accuracy in determining the classification of healthy and people with Parkinson's.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Nazri Mohd Naw,
Soft Computing and Data Mining Centre (SMC),
Faculty of Computer Science and information Technology,
Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia.
Email: nazri@uthm.edu.my

1. INTRODUCTION

Neurodegenerative disorder of central nervous system such as Parkinson's disease (PD) causes full or partial loss motor control, mental processing, speech, behaviour and other vital functions [1]. Nearly 10 million people all around the world suffers from PD which makes it the second most common neurological health problem among the elderlies [2]. Even though various drug therapies exist to lessen the complications caused by the disorder, the diagnosis and treatment are still invasive [3]. Where in medical procedure invasive means that there is break in the skin is created and there is contact with the mucosa, or skin break, or internal body cavity beyond a natural or artificial body orifice [4-6]. Indirectly, this complicates the diagnosis and treatment process for the patients who are suffering from PD.

Recently, the popularity of classification methods had gained big attention by medical application in using those methods to classify and discriminate healthy people for people from Parkinson's disease (PD) [7-8]. Furthermore, the popularity of classification methods such as artificial neural networks (ANN) and decision tree (DT) are widely used in many applications such as manufacturing, chemistry as well as in medical. In order to ease the situation, a non-invasive method of diagnosis was been developed. The diagnosis will be done through the collection of speech data from the subjects [9]. Since, tele-diagnosis and tele-monitoring systems that based on speech signals are cost cheaper and easy to self-use. Therefore,

it is the best candidate to be chosen for PD diagnosis [10]. Besides that, voice impairment is one of the earliest indicators for PD. Thus, the advancement in technology and new methods, the rising for voice measurement method for PD diagnosis using classification techniques had gained attention from researchers in data mining [11-13].

The previous study on the prediction of Parkinsonism used speech pattern analysis applications where the selected subsets of features were mapped to UPDRS (Unified Parkinson's Disease Rating Scale) using a combination of regression and classification techniques [14-15]. In addition, support vector machines (SVM) were also used to distinguish the samples belonging to subjects. Unlike the traditional methods, nonlinear speech analysis algorithms are used with standard PD metrics like UPDRS. The use of new method caused a slight difference in estimation (about 2 points) compared to the traditional clinical method [16-17].

Sakar et al. [18] applied a mutual information-based feature selection algorithm with the permutation test for evaluating the relevance and the statistical correlation between the features and PD-score. Utilising this method and collected data, a computer-aided data collection, storage and analysis to simplify the diagnostic and treatment process of PD were designed to fulfil the purpose of this study. In 2012, the Multi-Layer Perceptron (MLP) with Back-Propagation learning algorithm are been used to classify to effective diagnosis Parkinsons disease (PD) by Anchana K. and Veera B [18]. The results showed that MLP with BP accurately classify Twenty-Two attributes with 82.051% in training and 83.33% in validation.

Rashidah et al. [19] used the Multilayer Feedforward Neural Network (MLFNN) with Back-propagation (BP) algorithm in early detection and diagnosis of PD. The result shows that network can be used in diagnosis and detection of PD due to the good performance, which is 83.3% for sensitivity, 63.6% for specificity, and 80% for accuracy. Furthermore, the popularity of artificial neural networks (ANN) in data mining field had gained great attention from researchers to incorporate the technique into different kind of applications and classification problems [20-27].

The objective of this paper is to compare the performance of classification methods in classifying a person for PD. This classification could be done by using a collection of speech data from healthy and patient with Parkinson's. In order to achieve this, traditional and novel method of measurement must be replaced with a more efficient method such as pitch period entropy (PPE) [28]. The introduction of new method which is sensitive to the changes in speech specific to PD with data mining's classification technique can be used to distinguish healthy for patient with Parkinson's [29]. Decision boundary in the feature space created by these vectors that are based on subsequent voice measures can be utilised to predict the class of each subject (healthy or PD) [1].

The remaining of this paper is organized as follows. Section 2 presents the methodology used to perform the data mining task along with the dataset and the evaluation metrics. Section 3 presents the results and finally Section 4 concludes with some direction for future work.

2. RESEARCH METHOD

All selected data mining techniques for this paper used the Cross Industry Standard Process for Data Mining (CRISP-DM) approach [30], which describes the process of data mining in six phases as visualized in Figure 1. The method used to analyse the data is broken down to six stages:(1) Data Selection; (2) Data Pre-Processing; (3) Data partition; (4) Parameter Settings; (5) Data Simulation; (6) Result and Analysis. Microsoft Studio Azure were utilised to implement the proposed model's architecture and algorithm on a system sporting AMD Ryzen™ 1500X CPU @ 3.8GHz, 8GB RAM, x64 based processor running on 64-bit Windows Operating System.

2.1. Data Selection

In getting the measurement of accuracy, it is a very vital thing and it must be done with extra cautions to sustain the research integrity. Selecting a proper datasets is crucial in order to make sure that the results that going to analysed is reliable and accurate. This research has selected 17 features from the datasets which then filtered to 10 features to decrease the complexity of the findings [1]. Those 10 selected features are as follows in Table 1.

2.2. Data Pre-Processing

In this stage the acquired raw data will be transformed into something more standardized and understandable format. Most of the collected data are incomplete where most of the data are lacking in attribute values, lacking certain attributes of interest, or containing only aggregate data that full with noisy that containing errors or outliers. Furthermore, the noisy data is inconsistent which containing discrepancies in codes or names.

There are some tasks and preventive steps that involve in data preprocessing which include:

- (a) Data cleaning: fill in missing values, smooth noisy data, identify or remove outliers, and resolve inconsistencies.
- (b) Data integration: using multiple databases, data cubes, or files.
- (c) Data transformation: normalization and aggregation.
- (d) Data reduction: reducing the volume but producing the same or similar analytical results.
- (e) Data discretization: part of data reduction, replacing numerical attributes with nominal ones.

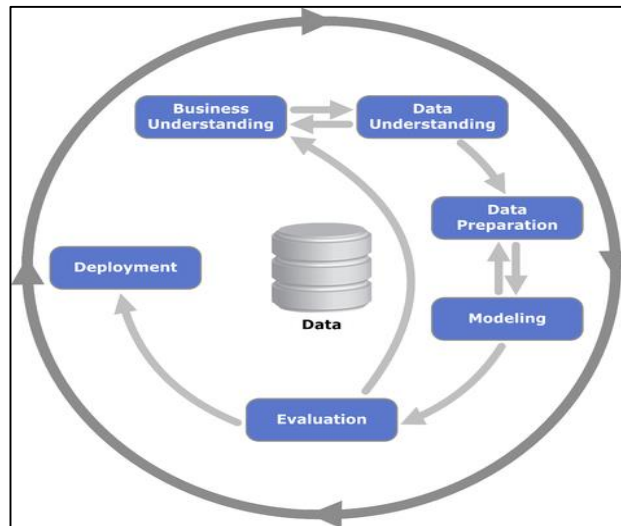


Figure 1. CRISP-DM process model for data mining

Table 1. Attribute Information

Attribute	Type	Description
MDVP: Jitter(Abs)	Integer	Kay Pentax MDVP absolute jitter in microseconds
Jitter: DDP	Integer	Average absolute difference of differences between cycles, divided by the average period
MDVP: APQ	Integer	Kay Pentax MDVP 11-point Amplitude Perturbation Quotient
Shimmer: DDA	Integer	Average absolute difference between consecutive differences between the amplitudes of consecutive periods
NHR	Integer	Noise-to-Harmonics Ratio
HNR	Integer	Harmonics-to-Noise Ratio
RPDE	Integer	Recurrence Period Density Entropy
DFA	Integer	Detrended Fluctuation Analysis
D2	Integer	Correlation dimension
PPE	Integer	Pitch period entropy

2.3. Data Partition

The next stage is to divide data into two partitions. There is no single condition or no rule of thumb for data partition. In this paper, the data were divided into 80% and 20%, each for training and testing respectively. Since the collected data consists of 195 vocal measurements samples from 31 participants therefore this make it 156 samples used for training and 39 for testing purposes.

2.4. Parameter Settings

Both algorithms selected from classification techniques has its own parameter setting that need to be adjusted before the final simulation. Furthermore, both algorithms can be difficult to tune. If the network hyperparameters are poorly chosen, the network may learn slowly, or perhaps not at all. The learning rate is one of, if not the most important hyperparameter. If this is too large or too small, your network may learn very poorly, very slowly, or not at all. Typical values for the learning rate are in the range of 0.1 to 0.9. Moreover, adjusting weights are so crucial and must be setting neither too big nor too small.

Data distribution also plays an important rule in training. Data need to be scall properly based on two conditions such as:

- a) Continuous values: the best values should be in the range of -1 to 1, 0 to 1 or distributed normally with mean 0 and standard deviation 1. However, the values do not have to be exact, but ensuring the inputs are approximately in this range can help during training by scaling down large inputs and scale up small inputs. (b) Discrete classes for classification problems for the output generally use a one-hot representation. For examples if there are 3 classes, then the data will be represented as [1,0,0], [0,1,0] or [0,0,1] for each of the 3 classes respectively. An epoch is defined as a full pass of the data set. If too few epochs are used it do not give network enough time to learn good parameters. However, if there are too many then it might overfit the training data. One way to choose the number of epochs is to use early stopping. The following table are the parameter used for each of the algorithm as shown in Table 2 and Table 3.

Table 2. Decision Tree Parameter

Parameter	Value
Maximum Number of Leaves Per Tree	20
Maximum Number of Samples Per Leaf Node	10
Learning Rate	0.2

Table 3. Neural Network Parameter

Parameter	Value
Number of Hidden Nodes	4
Learning Rate	0.5
Number of Learning Iteration (epochs)	100
Initial Learning Weights	0.5
Type of Normalizer	Min Max Normalizer

2.5. Data Simulation

The results are obtained by running some simulation experiments with the Parkinson's Disease Data on WEKA. Two classification algorithms had been selected and there are 10 trials of simulation were run for each algorithm. The main purpose of the simulation is to minimize the error rate of each training algorithm. The first phase is on training where the data will go through the simulation process to perform available output. Then next simulation is on the data that been reserved for testing which will be used to evaluate the efficiency of the algorithm. During testing, the inputs are provided from the data to predict the output. Lastly, the parameters of the algorithms are adjusted until the accuracy of the selected algorithm achieved a satisfying level of accuracy. The simulation has to be ran with many times in order to achieve a higher accuracy level. The higher the accuracy the algorithm the more precise the prediction will be.

3. RESULTS AND ANALYSIS

The performance of both classification techniques are evaluated on Parkinson's disease Data. Those two algorithms are (a) decision tree and (b) artificial neural networks (ANN). The simulation experiments were performed on a 1.66 GHz AMD Processor with a 2GB of RAM. For performing simulations WEKA software was used. The decision tree is compared with artificial neural networks based on:

- Receiver Operating Characteristic (ROC) curve where the true positive rate (Sensitivity) is plotted in function of the false positive rate. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold.
- Precision and recall where precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances, while recall (also known as sensitivity) is the fraction of relevant instances that have been retrieved over the total amount of relevant instances.
- F1 score or F-measure is a measure of a test's accuracy where it considers both the precision p and the recall r of the test to compute the score: p is the number of correct positive results divided by the number of all positive results returned by the classifier, and r is the number of correct positive results divided by the number of all relevant samples (all samples that should have been identified as positive).

There were 4000 datasets that been collected from UCI repository which consist of 17 input features but only 10 were chosen for more simplified yet effective prediction. All datasets are split into 80:20 where 80% of data were used for training purposes while the remaining 20% were used for testing purposes. The reason why we need more training data is because more training data will give better model and help researchers to be able to evaluate different models and make direct comparisons on the same subset of the dataset. If the same portion of datasets for model selection and testing are used then the model will be subject to model selection bias. That is, the model could overfit and may not generalize well to unseen data. Furthermore, by splitting data into 80:20 it will have high probability chance of getting all the target class detailed observation into the training dataset. This will help us for modeling.

For artificial neural networks, three layer feed forward neural network which consisted of an input layer, a hidden layer, and an output layer was used where the hidden neurons were kept fixed to 5 neurons

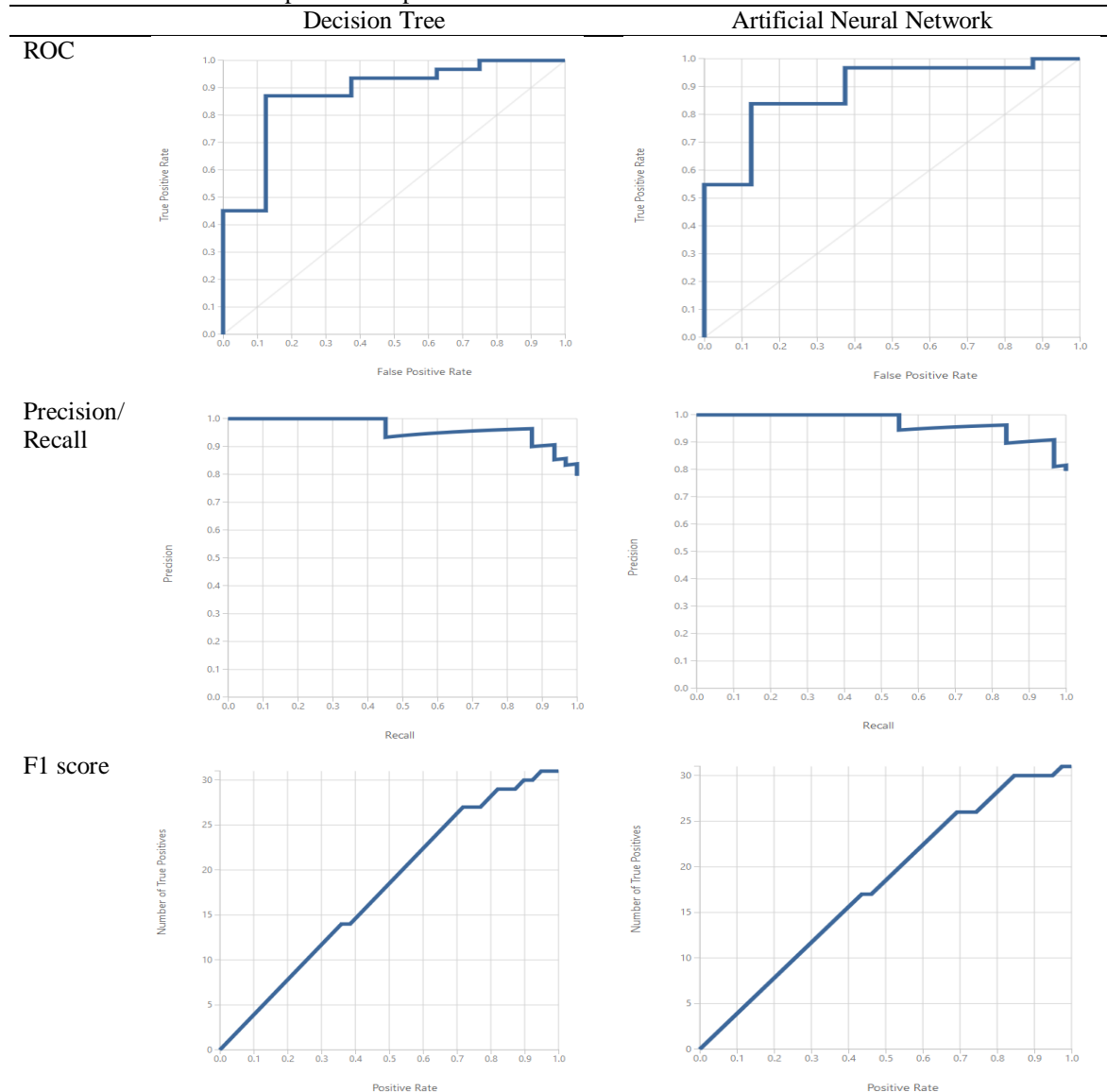
only. In the network structure the biases nodes were applied and the log sigmoid was placed as the activation function for the hidden and output layers nodes.

The other model of supervised learning is Decision Tree model. The model is constructed from a high level, decision tree induction goes through 4 main steps to build the tree: (a) begin with training dataset, which should have some feature variables and classification or regression output, (b) determine the “best feature” in the dataset to split the data in order to define “best feature” later, (c) split the data into subsets that contain the possible values for this best feature. This splitting basically defines a node on the tree i.e each node is a splitting point based on a certain feature from the data. (d) recursively generate new tree nodes by using the subset of data created from step 3. The sstep keep splitting until reach a point where it reach the optimum, by some measure, maximum accuracy while minimising the number of splits / nodes. Then the Two-Class Boosted Decision Tree was used to split into two classes which are ‘Yes’ and ‘No’. Lastly, Evaluate Model module is used to evaluate the trained model which is based on the score that generated by Score Model module. Tweaks are made to the parameters until a higher accuracy score is achieved.

3.1. Comparison Evaluation Result between Decision Tree and Artificial Neural Network

The performance of Artificial Neural Network and Decision Tree model is analyse and the results are presented in Table 4 where it shows the comparison graph of ROC, Precision or Recall and F1 score between Decision Tree and Artificial Neural Network.

Table 4. Comparison Graph between Decision Tree and Artificial Neural Network



From Table 4, it shows that Artificial Neural Network gives a much better result as compared to Decision Tree. This is because a neural network is known as universal approximator and it can deal with different kinds of non-linear problems as compared to decision tree. Furthermore, with a sufficient amount of data and time and the right ANN architecture you should be able to approximate whatever function generated the data with an arbitrary amount of accuracy. In addition decision tree is not good for online learning (your data is coming continuously and you need to update your model) since any data includes some exceptional situation will force your tree to be fall apart and need to be constructed again. The performance of Neural Network also outstanding as shown in Table 5 where Artificial Neural Network shows better performance matrix than Decision Tree in term of precision and accuracy.

Table 5. Comparison of Performance Matrix between Decision Tree and Neural Network

Performance Matrix\ Model	Decision Tree	Artificial Neural Network
Accuracy	0.846	0.870
Precision	0.903	0.963
Recall	0.903	0.897
F1 Score	0.903	0.839

4. CONCLUSION

Two popular data mining techniques known as Decision Trees or also called classification trees and Neural Networks were selected for comparison. These two data mining techniques are very different from the way they look to the way they find relationship within variables. The neural network is an assembly of nodes, looks somewhat like the human brain. While the decision tree is an easy to follow top down approach of looking at the data. Their performances are compared in classifying and discriminating healthy people for people with Parkinson's disease (PD) by distinguishing dysphonia. The simulation results demonstrate that Artificial Neural Network is more efficient and outperformed decision tree by giving accurate results in determining the classification of healthy and people with Parkinson's. Artificial Neural Network achieves 87% accuracy on a data set while the decision tree model only achieved 84% accuracy on the same data set. Therefore, artificial neural networks are the best fitted model for this datasets since the model is that most accurately fits on the selected datasets.

ACKNOWLEDGEMENTS

The authors would like to express the deepest appreciation to the Ministry of Education Malaysia for financially supporting this Research under Trans-disciplinary Research Grant Scheme (TRGS - Vote number T003). This work is partially supported by UniSZA (Grant No. R0008). Additional supports in terms of facilities were also provided by Centre Soft Computing and Data Mining Centre (SMC), FSKTM, Universiti Tun Hussein Onn Malaysia (UTHM).

REFERENCES

- [1] J.Jankovic, "Parkinson's disease: Clinical features and diagnosis," *J. Neurol. Neurosurgery Psychiatry*, vol. 79, no. 4, pp. 368–376, 2007.
- [2] Parkinson Derneği. (2011). [Online]. Available: (<http://www.parkinsonderneği.org/Icerik.aspx?Page=parkinsonnedir&ID=5>)
- [3] National Collaborating Centre for Chronic Conditions, Parkinson's disease, London, U.K.: Royal College of Physicians, 2006.
- [4] M. Politis, K. Wu, S. P. G. B. Molloy, K. E. Chaudhuri, and P. Piccini, "Parkinson's disease symptoms: the patient's perspective," *Movement Disorders*, vol. 25, no. 11, pp. 1646–1651, 2010.
- [5] K. P. Bhatia, P. Bain, N. Bajaj et al., "Consensus statement on the classification of tremors. Task force on tremor of the International Parkinson and Movement Disorder Society," *Movement Disorders*, vol. 33, no. 1, pp. 75–87, 2018.
- [6] Dr. R.Geetha Ramani, G.Sivagami, Shomona Gracia jacob " Feature Relevance Analysis and Classification of Parkinson's Disease TeleMonitoring data Through Data Mining", *International Journal of Advanced Research in Computer Science and Software Engineering*, vol 2, Issue 3, March 2012.
- [7] Dr. R.Geetha Ramani and G.Sivagami "Parkinson Disease Classification using Data Mining Algorithms", *International Journal of Computer Applications (IJCA)*, Vol-32, No.9, October 2011.
- [8] Farhad Soleimani Gharehehpoogh, Peyman Mohammadi, "A Case Study of Parkinson's Disease Diagnosis Using Artificial Neural Networks", *International Journal of Computer Applications*, Vol 73, No.19, July 2013.

- [9] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1010–1022, Apr. 2009.
- [10] Max A. Little, Member, IEEE, Patrick E. McSharry, Eric J. Hunter, Jennifer Spielman, and Lorraine O. Ramig, "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease (2009)." *IEEE Transactions on Biomedical Engineering*, vol. 56, 2009
- [11] Tripti Kapoor, R.K.Sharma, "Parkinson's Disease Diagnosis Using Mel-Frequency Cepstral Coefficients and Vector Quantization", *International Journal of Computer Applications*, Vol-4, No.3, Jan 2011.
- [12] H. Karimi Rouzbahani and M. R. Daliri, "Diagnosis of parkinsons disease in human using voice signals," *Basic and Clinical Neuroscience*, vol. 2, 2011.
- [13] I. Rustempasic and M. Can, "Diagnosis of parkinsons disease using fuzzy c-means clustering and pattern recognition," *Southeast Europe Journal of Soft Computing*, vol. 2, no. 1, 2013.
- [14] A. Khemphila and V. Boonjing, "Parkinsons disease classification using neural network and feature selection," *World Academy of Science, Engineering and Technology*, vol. 64, 2012.
- [15] S. Bouktif, E. M. Hanna, N. Zaki, and E. A. Khousa, "Ant colony optimization algorithm for interpretable bayesian classifiers combination: application to medical predictions," *PloS one*, vol. 9, 2014.
- [16] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, 2009
- [17] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, "Automatic classification and prediction models for early parkinsons disease diagnosis from spect imaging," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3333–3342, 2014.
- [18] Betul Erdogdu Sakar, M. Erdem Isenkul, C. Okan Sakar, Ahmet Sertbas, Fikret Gurgun, Sakir Delil, Hulya Apaydin, and Olcay Kursun, "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings," 2013.
- [19] Anchana Khemphila and Veera Boonjing, "Parkinsons Disease Classification using Neural Network and Feature Selection", *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering* Vol: 6, No: 4, 2012.
- [20] Rashidah. Funke Olanrewaju, Nur Syarafina Sahari, Aibinu A. Musa and Nashrul Hakiem, "Application of Neural Networks in Early Detection and Diagnosis of Parkinson's disease", 2014 *International Conference on Cyber and IT Service Management (CITSM)*, page 78-82, 2014.
- [21] S. Lahmiri, "Wavelet transform, neural networks and the prediction of S & P price index: a comparative paper of back propagation numerical algorithms," *Business Intelligence Journal*, Vol. 5, Issue 2): 235-244, 2012.
- [22] N. M. Nawawi, R. S. Ransing, M. N. M. Salleh, R. Ghazali, and N. AbdulHamid, "An improved back propagation neural network algorithm on classification problems," *Communications in Computer and Information Science (CCIS)*, Vol.118: 177-188, 2011.
- [23] J. N. D. Gupta, and R. S. Sexton, "Comparing back propagation with a genetic algorithm for neural network training," *The International Journal of Management Science*, Vol. 27: 679–684, 1999.
- [24] N. M. Nawawi, R. Ghazali, and M. N. M. Salleh, "The development of improved back-propagation neural networks algorithm for predicting patients with heart disease," *LNCS*, Vol. 6377: 317–324, 2010.
- [25] Patil TR, Sherekar MS, "Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification," *International Journal of Computer Science and Applications*, 6(2), 256-261, 2013.
- [26] Mehdi RamezaniFard, B. S. Mousavi, "Digital image classification by optimised fuzzy system", *Indonesian journal of Electrical Engineering and Computer Science (IJECS)*, Vol. 14, No. 3, pp: 1196-1202, 2019.
- [27] Jesmeen M. Z. H, J. Hossen, S. Sayeed, CK Ho, Tawsif K, Armanur Rahman, E.M.H. Arif, (2019), "A Survey on Cleaning Dirty Data Using Machine Learning Paradigm for Big Data Analytics", *Indonesian journal of Electrical Engineering and Computer Science (IJECS)*, Vol. 10, No. 3, pp: 1234-1243
- [28] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. E. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomed. Eng. Online*, vol. 6, no. 23, 2007, doi: 10.1186/1475-925X-6-23.
- [29] D. A. Rahn, M. Chou, J. J. Jiang, and Y. Zhang, "Phonatory impairment in Parkinson's disease: Evidence from nonlinear dynamic analysis and perturbation analysis," *J. Voice*, vol. 21, pp. 64–71, 2007.
- [30] Singhal S, Jena M, "A study on WEKA tool for data preprocessing, classification and clustering," *International Journal of Innovative Technology and Exploring Engineering*, 2(6), 250-253, 2013.

BIOGRAPHIES OF AUTHORS



Nazri Mohd Nawawi received his B.S. degree in Computer Science from University of Science Malaysia (USM), Penang, Malaysia. His M.Sc. degree in computer science was received from University of Technology Malaysia (UTM), Skudai, Johor, Malaysia. He received his Ph.D. degree in Mechanical Engineering department, Swansea University, Wales Swansea. He is currently a lecturer in Software Engineering Department at Universiti Tun Hussein Onn Malaysia (UTHM). His research interests are in optimization, data mining techniques and neural networks.



Mokhairi Makhtar is an Associate Professor of Computing from Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin. He received his Ph.D. in Computer Science from the University of Bradford in 2012. His research interests include Machine Learning, Data Mining and big data analytics for toxicology, education, health and business applications.



Zehan Afizah Afip currently is currently a lecturer in Software Engineering Department at Universiti Tun Hussein Onn Malaysia (UTHM). He received his bachelor degree in computer science from Universiti Teknologi Malaysia (UTM) Malaysia. He did his master degree in in computer science from Universiti Utara Malaysia (UUM). His research areas are soft computing, scheduling, and data base system



Mohd Zaki Salikon currently is currently a lecturer in Software Engineering Department at Universiti Tun Hussein Onn Malaysia (UTHM). He received his bachelor degree in computer science from Universiti Teknologi Malaysia (UTM) Malaysia. He did his master degree in in computer science from Universiti Utara Malaysia (UUM). His research areas are soft computing, scheduling, and data base system



Khairul Amin Mohamad Sukri currently is currently a lecturer in Information Security and Web Technology Department at Universiti Tun Hussein Onn Malaysia (UTHM). He received his bachelor degree in computer science from Universiti Teknologi Mara (UiTM) Malaysia. He did his master degree in in computer science from Universiti Teknologi Malaysia (UTM). His research areas are information security, computer and mobile security and network intrusion detection system.