

A Novel Model for Classification of Parkinson's Disease: Accurately Identifying Patients for Surgical Therapy

Farhan Mohammed¹, Xiangjian He¹, Jinjun Chen² and Yiguang Lin³

¹ Global Big Data Technologies Centre, University of Technology Sydney, Sydney, Australia
Farhan.Mohammad@student.uts.edu.au, Xiangjian.He@uts.edu.au

² Swinburne Data Science Research Institute, Swinburne University of Technology, Melbourne, Australia
jchen@swin.edu.au

³ School of Life Sciences, University of Technology Sydney, Sydney, Australia
Yiguang.Lin@uts.edu.au

Abstract

Parkinson's disease (PD) is a neurodegenerative disorder and a global health problem that has no curative therapies. Surgery is a well-established therapy for controlling symptoms of advanced PD patients. This paper proposes a streamlined model to classify PD and to identify appropriate patients for surgical therapy. The data was gathered from the Parkinson's Progressive Markers Initiative consisting of 1080 subjects. Multilayer Perceptron (MLP), Decision trees, Support Vector Machine and Naïve Bayes are used as classifiers. MLP achieves the highest accuracy as compared to other three classifiers. The dataset used in our experiments is from the Parkinson Progressive Markers Initiative. With feature selection, it is observed that the same classification accuracy is achieved with 60% of the attributes as that using all attributes. It is demonstrated that our classification model for PD patients produces the most accurate results and achieves the highest accuracy of 98.13%.

1. Introduction

Parkinson's disease (PD) is a chronic and progressive neurodegenerative disorder mainly affecting movement. The cardinal symptoms of PD are tremor, rigidity, postural instability and slow movement [1]. Approximately ten million people worldwide are currently living with PD [2]. The cause remains largely unknown. Many researchers believe that the disease results from an interaction between genetic and environmental factors that leads to progressive degeneration of neurons in susceptible regions of the brain. PD involves the malfunction and death of vital neurons located in the area of the brain called basal

ganglia. Some of these neurons produce dopamine, a chemical that sends messages to the part of the brain that controls movement and coordination. As PD progresses, the amount of dopamine produced in the brain further decreases, thereby unabling the patient to control movement normally [3].

Mostly PD occurs in the ages of 50-60 and is difficult to identify in the earlier stage. The prevalence of PD increases in aged population. As more and more people live longer globally, the number of people living with PD worldwide is also rising. Currently, there are no laboratory tests that have diagnostic values for PD. In early stages, many diagnostic methods, such as blood tests, brain imaging techniques such as magnetic resonance image (MRI), positron emission tomography (PET scan), and single photon emission computed tomography (SPECT), are used to exclude other medical conditions such as stroke or brain tumors that imitate symptoms of PD [4]. Amongst others, one of the methods for the diagnosis of PD is detecting and analyzing voice disorders by using acoustic tools that record the changes in pressure at lips or inside the vocal tract. It has been found [5] that some features in the voices of the patients with PD can be used as discriminatory measures to differentiate PD by adopting data mining techniques.

When it is mild, PD is generally well controlled by drugs; however, as the disease progresses, pharmacotherapy often fails to provide adequate symptom relief and sometimes causes disabling complications. Additional treatment approaches, such as surgery treatment including deep brain stimulation (DBS), pallidotomy, and thalamotomy, were used to manage problems of advanced PD [6, 7]. The efficacy of surgical therapy such as DBS has been well established through randomized controlled studies involving a large

number of PD patients. The success of DBS is firstly dependent on the selection of appropriate patients as not all PDs are suitable for surgical therapy.

Selection of appropriate patients, in clinical practice, is based on many factors including the age of the patient, disease stage, disease duration, comorbidities, and responsiveness to levodopa medication. It is usually required to form an interdisciplinary team consisting of neurologist, neurosurgeon, psychiatrist, neuropsychologist, rehabilitation specialist, and sometimes a social worker, to discuss these factors and then decide if the patient is suitable for surgical treatment [8, 9]. Therefore, there is an urgent need to streamline the process of selecting appropriate PD patients for surgical treatments. In this study, we propose an effective model to classify PD and select suitable patients for surgery using data mining algorithms and feature selection based on information gain with PD patients.

Data Mining is defined as the nontrivial extraction of implicit, previously unknown, and potentially useful information from generic data. The use of classifier systems in disease diagnosis is increasing. Technological advances in the field of Artificial Intelligence (AI) have led to the emergence of expert systems and Decision Support Systems for medical applications. Moreover, in the last few decades, computational tools have been designed to improve the experiences and abilities of doctors and medical specialists in taking decisions regarding diagnosis and treatment about their patients. However, expert systems and different AI techniques for classification have the potential of being good supportive tools. Classification systems can help in increasing accuracy and reliability of diagnoses and minimizing possible errors, as well as making the diagnoses more time efficient [10].

This research is important because advanced PD patients who do not respond to drug treatments requires surgical therapy to control PD symptoms. However, the selection of appropriate patients for surgery is complicated, costly and involves several discussions by an interdisciplinary team. This in turn is time consuming resulting in significant delay in the treatment. Using data mining algorithms and feature selection, this research introduces a streamline model to classify PD and identify appropriate patients for surgery. This will provide the team an insight as to which PD patient to select for surgery based on their values after feature selection process.

The main contributions of this research are:

- Development of a classification model that would simplify and streamline the process of PD classification and selection of suitable PD patients

for surgery. The performance of the model is supported with a high accuracy result of 98.13% using MLP as the classifier. This specific path of research, i.e., using data mining techniques to classify and streamline PD patients for surgery, is novel and has not been addressed before. Most of the previous research only addressed the diagnosis of PD.

- A feature reduction technique used to determine if all features are necessary to include for achieving a high or equal accuracy. Information Gain (IG) technique was used and it was discovered that with 60% of the attributes an equal accuracy result was achieved to that with all attributes. The advantage of this technique is that it will reduce the number of tests required to get the data, thereby reducing cost and time for selecting PD patients for surgery.
- A comparison study was performed to understand which features contribute the most to the classification of PD patient for surgery thus gaining knowledge on which test to prioritize in clinical practice.

In the next section, we will discuss on data mining and its related techniques along with the background of PD classification using data mining techniques. Section 3 will describe our proposed approach for classification. In Section 4, our experiments and results will be presented. Section 5 demonstrates the discussion in which the results will be interpreted in depth with a highlight on the most important features in our model and its relevance to the previous work. A comparison table comparing our research work with the state-of-the-art work is also provided. The conclusion is presented in Section 6.

2. Related Work

Data mining has been defined by previous studies as the process of analyzing large databases to detect meaningful patterns and rules [11, 12]. Data mining is a step within the Knowledge Discovery in Databases process that involves using data analysis and discovery algorithms to yield particular patterns (or models) based on the data. The main activities of data mining are as follows.

- Classification: Classification involves identifying profiles of classes in terms of their attributes and determines which of the predefined classes a new item belongs. For example, given particular classes of patients with different

medical treatment responses, the classification is used to identify the form of treatment to which a new patient is most likely to respond [13, 14].

- Clustering: Clustering involves identifying a set of classes in which particular items are grouped according to their characteristics. Clustering is best used to identify groups of items that are similar. For example, based on a patient data set, clustering can be used to identify subgroups of patients with similar treatment schemas [15].
- Association: Association involves the identification of relationships between items in which the presence of one pattern implies the presence of another pattern. For example most patients who receive prescriptions for medication A also receive prescriptions for medication B [16].

2.1. Classification of PD Using Data Mining Techniques.

Most of the research in the classification of PD is done by detecting and analyzing voice disorders by using acoustic tools that record the changes in pressure at lips or inside the vocal tract. One of the earliest classifications of PD using data mining was done by Max Little from University of Oxford [17]. In his study, he developed a software that learns to detect differences in voice patterns, in order to spot distinctive clues associated with PD. By using machine learning, a large amount of data was collected in order to know if someone has PD or not. A new measure of dysphonia, pitch period entropy (PPE) was introduced, which is robust to many uncontrollable effects including noisy acoustic environments and normal, healthy variations in voice frequency. He collected sustained phonations from 31 people, 23 with PD, and performed classification using a kernel support vector machine (SVM).

Shaikh and Chhabra [18] conducted a study to explore WEKA filters on data mining algorithm Naïve Bayes, which was used to classify PD patients from healthy controls. The datasets were obtained from University of California Irvine (UCI) machine learning repository. At first, Naïve Bayes classifier was used and the parameters were noted down. Then, the supervised Discrete attribute filter from WEKA was used on both data sets, which divides the input values of the datasets to a range of values and its parameters were noted down. Similarly, another filter (unsupervised) 'Numeric Transform' was used on both datasets and the results were noted down. Naïve Bayes is used again to compare

the performance measures of accuracy, sensitivity and specificity.

Khemphila and Boonjing [10] conducted PD classifications with reduced number of attributes. They used IG to determine the attributes from patients. Then, they used Artificial Neural Networks (ANN) to classify PD in the given dataset. The dataset was taken from the UCI machine learning repository created by Max Little. Their experiment was done using WEKA. They divided the dataset into two sets: training and testing dataset. At first, IG was not used and classification was done using ANN directly. The accuracy was of 91.45% of the training set and 80.76% of the testing set. Then, by using IG, the attributes were sorted according to feature of importance for classification. After implementing IG, the attributes were reduced to 16 and then ANN was implemented again. The accuracy rate was 82.05% for the training set and 83.33% for testing set. The authors concluded that with IG the results showed a higher accuracy rate for ANN.

In the study by Gök [19], experimentation on classifying PD from healthy controls was done using the dysphonic symptoms (vocal characteristics) of PD. The dataset used was developed by Max Little. The experiment consists of two phases. First phase is feature selection (FS). In FS, relevant features were selected from 22 features using linear SVM for discovering informative features. Thus out of 22, 10 were selected. Then, those 10 were ranked again based on their information gain with respect to class. Six classifiers were selected to perform the comparison analysis. The accuracy of the competing algorithms was evaluated by means of using both the selected 10 and all 22 features. The best result was obtained by ensemble of k-NN on PD dataset using only the 10 selected features.

Prashanth, et al. [20] performed a study on early PD classification using Single Photon Emission Computed Tomography (SPECT) imaging. The data was gathered from the PPMI database. They used support vector machines as their classifier for developing automatic classification and prediction/prognostic models for early PD. In their study, they found that the accuracy of SVM with radial basis function provided the highest accuracy.

Hirschauer, et al. [21] focused on the diagnosis of PD using Enhanced Probabilistic Neural Network (EPNN). A comprehensive computer model was presented using motor, non-motor and neuroimaging features from the PPMI dataset. The model was tested for differentiating PD patients from those with scans without evidence of dopaminergic deficit (SWEDD). The results were compared to four other commonly used algorithms: the probabilistic neural network (PNN), support vector machines (SVM), k -nearest neighbor (k -NN) and

classification tree (CT). Based on the results EPNN had the highest accuracy.

Prashanth, et al. [22] worked on the detection of early PD through multimodal features and machine learning. The authors used non-motor features of rapid-eye-movement sleep behavior disorder (RBD) along with other features like cerebrospinal fluid (CSF) and dopaminergic imaging markers from 183 healthy normal and 401 early PD subjects. These data were gathered from PPMI and the classifiers used were Naïve Bayes, Support Vector Machine (SVM), Boosted trees and Random Forests. The dataset was divided in a way so that 70% was used for training and the rest 30% was used for testing. Based on the results SVM provided the highest accuracy.

3. Proposed Approach

This study utilizes real patients' clinical, medical and surgical data to develop a practical model for PD classification and for selection of suitable PD patients for surgical treatment. Our proposed approach is outlined in Figure 1.

3.1. Data collection

The data was collected from the Parkinson's Progression Markers Initiative (PPMI) repository [23]. The repository is updated regularly as the longitudinal follow-up of study subjects continues and the data generated from different analysis techniques are completed. For this study, a new dataset is developed

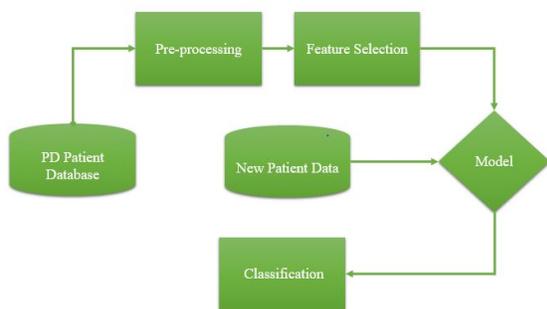


Figure 1. Proposed approach architecture.

by gathering different types of data from PPMI such as medication, health, surgical, PD symptoms, MRI scan results and post medication data. This provided us with 1080 patients, of which 40 patients opted for surgery, with 40 attributes for each patient. The dataset was further divided into a training dataset (70%) and a testing dataset (30%). The training dataset consisted of records of 758 patients and the testing dataset consisted

of records of 322 patients each with 40 attributes, where 23 out of 322 patients were of surgical patients.

3.2. Pre-processing

During the pre-processing stage, the dataset was analyzed to see if it had any missing values or redundant values. To address the issue of missing values WEKA filter Remove Missing Values was used [24]. For redundant data, the latest record of the data registered in the database was used erasing old data. We then used two datasets for our experiments. Experiment 1 has all 40 attributes. For Experiment 2, the dataset, after feature selection, was split-up into 10 datasets each consisting of 10% of the attributes.

3.3. Feature Selection

Feature selection is the process of selecting a subset of relevant features (variables, predictors) for use in model construction. Feature selection techniques were used to avoid overfitting and improve model performance. They can be used to provide faster and more cost-effective models. They were also used to gain deeper insight into the underlying processes that generated the data [25]. Thus, with this model, selection of appropriate patients for surgery treatment will be highly accurate and effective. The objective is to reduce the attributes in order to identify the most important feature that contributes to the classification. In our work, we select the filter methods because they are moderately robust against the overfitting problem. We use Information Gain (IG) technique as this technique computes the information gain of a feature with respect to class [26]. The Ranker filter is also used along with IG. Ranker helps in ranking the features based on their information gain with respect to class.

3.4. Model

The developed model aims to provide an accurate classification of PD patients and effectively assess if the patients requires surgical treatment. When a new patient arrives, all of his/her information is stored in the database. By this stage, several features would have been extracted through feature selection technique. Therefore, the data collected from new patients consist of those features only. Their features are applied to the model to know whether they are applicable for surgery. Figure 2 illustrates how the model works. Figure 3 provides the list of features of the dataset.

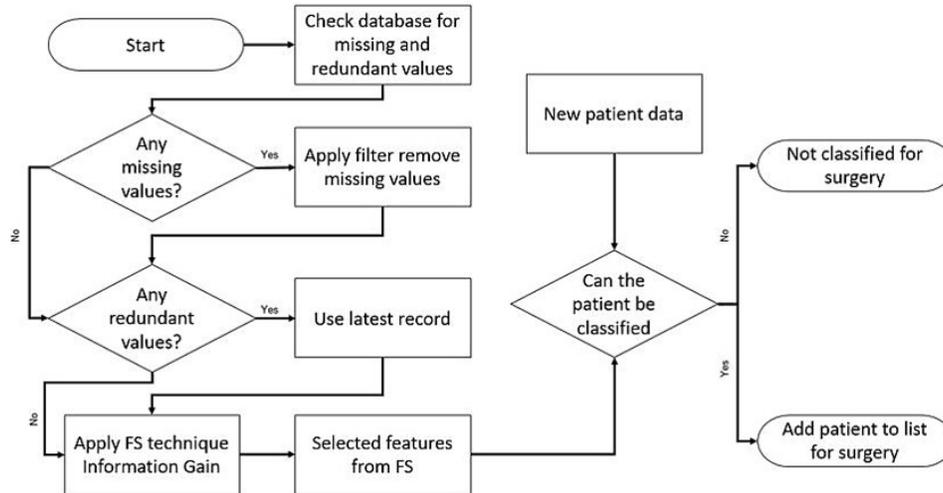


Figure 2. Flowchart of the model

3.5. Classification

In this step, the new patient will be classified into whether the patient is suitable to undergo a surgery or not. This would help the specialists to undertake proactive steps so that the patient can get the right treatment.

4. Experiments and results

All our experiments were performed using the WEKA platform. The data was gathered from PPMI. The PPMI repository contains the full set of clinical and biological data collected as part of an ongoing study along with bio-specimen analysis results. The PPMI database is a multi-centre international study involving subjects from different countries adding diversity in the database thus making it more robust. PPMI is being sponsored by the Michael J Fox Foundation [27]. PPMI consists of 1286 subjects (533 females and 753 males). Of the 1286 subjects, 215 subjects belong in the control group. For our study, we gathered data from various datasets in PPMI presenting us with 1080 subjects with 40 attributes.

4.1. Experiment 1: General classification of PD patients for surgery using different classifiers

The aim of this experiment is to develop a model to provide an accurate classification of a PD patient so that it can suggest if the patient is indicated for a surgical option. Since we are trying to classify which patients require surgery, there are two classes for this research,

i.e. class 0 is for those that do not need surgery and class 1 requires surgery.

In this experiment, Naïve Bayes, Support Vector Machine (SVM), J48 (Decision Tree) and Multilayer Perceptron (MLP), that is a form of a neural network algorithm from WEKA, were used to construct the model [28]. Table 1 presents classification results

For performance measurement, a 10-fold cross validation of the test dataset is done and the confusion matrix is obtained to estimate four measures: Accuracy, Precision, F-measure and Sensitivity. They are

Table 1. General Classification of PD Patients for Surgery Using Different Classifiers.

Learning Machine	Model Accuracy	Precision	Sensitivity (weighted average)	F-Measure
Naive Bayes	93%	0.862	0.929	0.894
Decision Tree (J48)	94.7%	0.95	0.95	0.932
Multilayer Perceptron	98.13%	0.982	0.981	0.980
Support Vector Machine	93%	0.862	0.929	0.894

calculated by:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad (1)$$

$$Precision = \frac{TP}{TP + FP}, \quad (2)$$

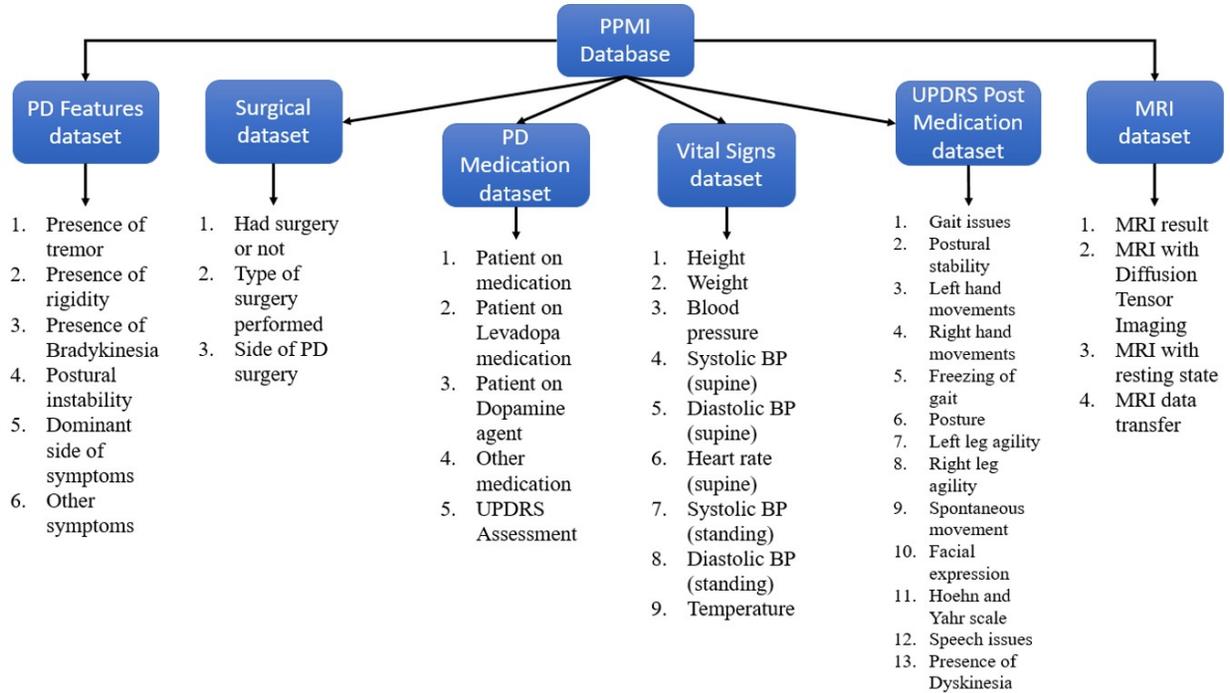


Figure 3. List of features

$$Sensitivity = \frac{TP}{TP + FN}, \quad (3)$$

$$F - measure = 2 * \frac{Precision * Sensitivity}{Precision + Sensitivity}, \quad (4)$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives. For this experiment, the weighted average value of sensitivity is used because it provides an overall performance of the model by taking into account the results of both classes. As a result, MLP had the highest accuracy of 98.13% followed by J48 at 94.7%, SVM and Naïve Bayes at 93% each.

4.2. Experiment 2: Classification of PD patients for surgery after using feature selection

The aim of this experiment is to study the effect of feature selection in the accuracy of the classification. Feature selection is one of the dimensionality reduction techniques for reducing the attribute space of a feature set. More precisely, it determines how many features should be enough to give moderate accuracy.

For feature selection, we used IG filter from WEKA platform. This filter acts as an attribute evaluator as it evaluates attributes according to their information gain.

This process also uses the Ranker filter to rank the features based on their information gain with respect to class. For this experiment, we used the same dataset that was used in Experiment 1. Ten datasets are built depending on the number of selected features. The first dataset contains only 10% of the total attributes. Then, each time, total feature selections is increased by 10%. Therefore, dataset 1 contains 10% of all attributes, dataset 2 contains 20%, dataset 3 contains 30%, ..., and dataset 10 contains 100% of all attributes.

We used MLP as the classifier because it had the highest accuracy from Experiment 1. The MLP model is made up of three hidden layers with 500 nodes in each layer. The default activation function, i.e., Approximate Sigmoid function, is used in our model. Each feature-reduced dataset is used for a 10-fold cross-validation for evaluation. It is observed that selecting 60% of the attributes gives the same high accuracy as that using all attributes. This infers that not all features may be necessary to attain a highly accurate classification. The sensitivity is measured for the test set, using the 10-fold cross validation. For this experiment, the results of sensitivity are provided with respect to surgery class only. These results are listed so that we can measure how the proposed algorithm can correctly classify PD patients for surgery.

A 95% confidence interval is also measured. The main purpose of the confidence interval is to provide a

Table 2. Classification of PD Patients for Surgery Using Feature Selection.

% of Features selected	# of Features selected	Correctly classified for surgery	Incorrectly classified for surgery	Correctly classified for non-surgery	Incorrectly classified for non-surgery	Sensitivity (with respect to surgery)	Precision (95% confidence interval)	Surgery vs Non-surgery accuracy
10%	4	0	23	299	0	0	0.92±0.034	92.85%
20%	8	5	18	299	0	0.217	0.94±0.031	94.4%
30%	12	11	12	299	0	0.478	0.96±0.024	96.6%
40%	16	14	9	299	0	0.609	0.97±0.021	97.2%
50%	20	16	7	298	1	0.696	0.97±0.018	97.5%
60%	24	17	6	299	0	0.739	0.98±0.018	98.13%
70%	28	15	8	299	0	0.652	0.97±0.015	97.5%
80%	32	17	6	299	0	0.739	0.98±0.015	98.13%
90%	36	17	6	299	0	0.739	0.98±0.015	98.13%
100%	40	17	6	299	0	0.739	0.98±0.015	98.13%

range of values for an estimated parameter rather than a single point value. All of these results are presented in Table 2. The confidence interval is calculated by:

$$95\%CI = const * \sqrt{error * \frac{1 - error}{n}}, \quad (5)$$

where const is a constant value (which is 1.96 in our case) corresponding to the probability (i.e., 95%), error is the classification error and n is the sample size.

5. Discussion

In the medical domain, an unmet medical need in the management of PD is how to simplify the complicated process of selecting suitable PD patients for surgical treatment, a breakthrough in recent years. Therefore, our objective is to develop a practical classification model that can accurately identify PD patients for surgery based on data mining algorithms combined with feature selection using clinical data. The model will enable physicians to, reliably and accurately, classify PD patients using available clinical parameters without the time consuming and costly medical team meetings and discussions.

In the current study, we have developed a novel model to streamline the way to classify PD and to identify appropriate PD patients for surgery. The model is remarkably reliable and accurate with an accuracy of 98.13%.

From Experiment 2, it can be noted that not all features/attributes contribute to high accuracy of classification. Using only 24 (60%) attributes, the classification accuracy of PD patients for surgery is similar to the result of classification using all the features. The values of sensitivity results differ in the two experiments. The weighted average of sensitivity, in Experiment 1, takes into account the sensitivity values

of both classes (0 and 1), multiplies them with the instances classified into their respective classes and divides the total with the total number of instances. However, for Experiment 2, the values of sensitivity with respect to only the surgery class are provided so that we can measure how the proposed algorithm can correctly classify PD patients for surgery.

Furthermore, each of the 24 features does not make equal contribution to the classification of PD patients. The top features that contribute more are presented in Table 3. This was generated by using the correlation

Table 3. Top seven features' details.

Attribute code	Category	Description
DYSKPRES	Motor Assessments	Presence of dyskinesia
NHY	Motor Assessments	Hoehn and Yahr scale
NP3SPCH	Motor Assessments	Speech issues after medication
NP3GAIT	Motor Assessments	Gait disorder after medication
NP3PSTBL	Motor Assessments	Postural stability after medication
MRIWDTI	MRI Result	MRI with Diffusion Tensor Imaging (DTI)
DOMSIDE	PD Features	Dominant side of PD

filter from WEKA. The correlation filter evaluates the worth of a feature by measuring the Pearson's correlation between the attribute and the class [29].

Understanding the importance of the top features in their contribution to classification is important because it provides us with an idea of how the features interact with the class for classification. For example, the Hoehn and Yahr scale (HY) is a widely used clinical rating scale, which defines broad categories of motor

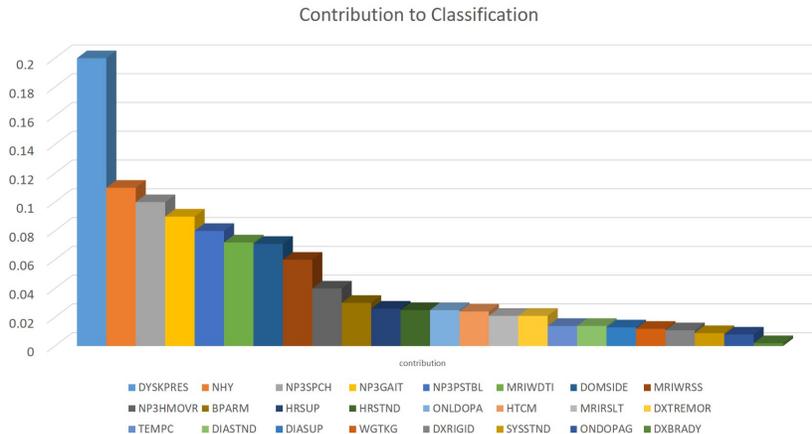


Figure 4. Contribution of features to the classification of PD patients for surgery

function in PD. Among its advantages, it is simple and easily applied. It captures typical patterns of progressive motor impairment, which can be applied whether or not patients are receiving therapy. Progression in HY stages has been found to correlate with motor decline, deterioration in quality of life, and neuroimaging studies of dopaminergic loss [30]. As shown in Table 2,

among the 23 patients that have undergone surgery, the proposed model with 60% of the attributes, correctly classifies 17 of them. Out of those 17 patients, nine patients can be correctly classified by the seven attributes listed in Table 3. Figure 4 provides an indication of how much contribution each feature has made to the classification accuracy. Figure 5 shows a comparison of classification accuracies using the top seven features and the rest of the features.

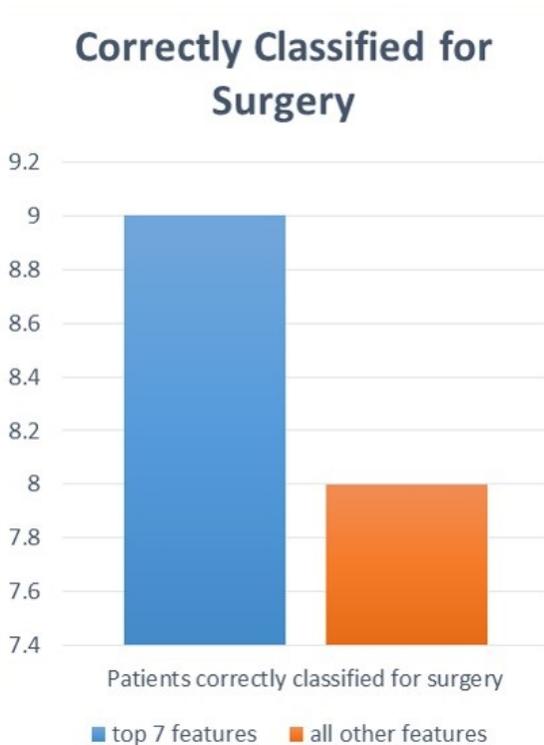


Figure 5. Top seven features vs all other features classification

5.1. Comparison of current study with other research work

In this section, we compare the performance among the approaches proposed in the paper and the related works, and demonstrate the comparison results in Table 4.

In this study, data on subjects that had been diagnosed into various stages of PD were gathered. Based on the data, we have classified whether the subject needs to undergo surgery. To the best of our knowledge, we are the first to use data mining techniques to streamline the process of selecting suitable PD patients for surgery with potential important clinical applications.

Similar work has not been done previously as most of the other related existing approaches focus on only pure classification and diagnosis of PD patients during early stages.

The significant contribution from the present study is that our experiments have given high performance based on PPMI, a large database that is one of the large-scale and standard databases publicly available for PD. Furthermore, most of the related works have a limitation of either having a small dataset or small sample size as detailed in Table 4.

Table 4. Summary and comparisons of previous related works on PD classification.

Research works	Dataset used	Attributes in dataset	Sample size	Classifier used	Accuracy	Feature selection	Accuracy after feature selection
Little et al. [17]	Developed own dataset	Vocal attributes	31	Support Vector Machine	91.4%	No	N/A
Rustempasic and Can [5]	Max Little dataset	Vocal attributes	31	Fuzzy C-Means clustering	80.88%	No	N/A
Khemphila and Boonjing [10]	Max Little dataset	Vocal attributes	31	Artificial Neural Network (ANN)	80.76%	Yes	83.33%
Shaikh and Chabra [18]	Max Little dataset	Vocal attributes	31	Naive Bayes	69.23%	Yes	78.46%
Gok [19]	Max Little dataset	Vocal attributes	31	k-Nearest Neighbor (k-NN)	N/A	Yes	98.46%
Prashanth et al. [20]	PPMI	Striatal binding ratio (SBR) values	493	Support Vector Machine	96.1%	No	N/A
Hirschauer et al. [21]	PPMI	Non-motor and motor functions, olfactory and cognitive function, autonomic function and SBR	666	Enhanced Probabilistic Neural Network (EPNN)	92.5%	No	N/A
Prashanth et al. [22]	PPMI	Non-motor functions, CSF and dopaminergic markers	584	Support Vector Machine	96.4%	No	N/A
Suganya and Sumanthi [31]	Private dataset	Vocal attributes	195	Artificial Bear Optimization (ABO)	97.5%	No	N/A
Current paper	PPMI	MRI result data, surgical data, PD symptoms, motor functions, vital health data, medication data and post medication data	1080	Multilayer Perceptron (MLP)	98.13%	Yes	98.13%

6. Conclusion

In this study, we have proposed and developed a data mining and feature selection based model for accurate PD classification and selection of suitable PD patients for surgery. Potentially, this model could fill the unmet medical need of streamlining the complicated process of selecting suitable PD patients for surgery.

In Experiment 1, we have developed a novel PD classification model with several classifiers. After comparing several alternative classifiers, we have found that MLP consistently outperforms the others in most experiments with the highest PD classification accuracy of 98.13%. Experiment 2 has identified the most important attributes required for such classification by using feature selection. One of the key findings is that, using only 60% of the attributes, MLP classification with IG has produced a remarkably high accuracy (98.13%) indicating that a smaller number of clinical parameters is sufficient for reliable and accurate diagnosis of PD.

The results from our experiments have demonstrated that the model we developed can be a useful tool in clinical practice for accurate classification of PD and selection of appropriate PD patients for surgery. Our model has also provided a better understanding of features that contribute to reliable and accurate PD classification indicating that not all features are required for the accurate and efficient PD classification.

7. Acknowledgements

This research was partly supported by Precision Medicine and Healthcare (Beijing) Science and Technology Co. Ltd (UTS project ID PRO17-2960), and Intergenepharm Pty Ltd (UTS project ID PRO17-2867).

References

- [1] P. Durga, V. S. Jebakumari, and D. Shanthy, "Diagnosis and classification of parkinsons disease using data mining techniques," *International Journal of Advanced*

- Research Trends in Engineering and Technology*, vol. 3, pp. 86–90, 2016.
- [2] A. Elkouzi, “Understanding parkinson: Statistics.” <http://parkinson.org/Understanding-Parkinsons/Causes-and-Statistics/Statistics>, 2018.
 - [3] R. G. Ramani and G. Sivagami, “Parkinson disease classification using data mining algorithms,” *International Journal of Computer Applications*, vol. 32, 2011.
 - [4] H. Karimi Rouzbahani and M. R. Daliri, “Diagnosis of parkinsons disease in human using voice signals,” *Basic and Clinical Neuroscience*, vol. 2, 2011.
 - [5] 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.
 - [6] “Surgery for parkinson’s disease.” <https://www.webmd.com/parkinsons-disease/guide/parkinsonssurgical-treatments>, 2018.
 - [7] A. Mohammed, R. Bayford, and A. Demosthenous, “Toward adaptive deep brain stimulation in parkinson’s disease: a review,” *Neurodegenerative Disease Management*, vol. 8, 2018.
 - [8] T. Pringsheim, N. Jette, A. Frolkis, and T. D. Steeves, “The prevalence of parkinson’s disease: A systematic review and meta-analysis,” *Movement Disorders*, vol. 29, no. 13, pp. 1583–1590, 2014.
 - [9] A. Wagle Shukla, P. Zeilman, H. Fernandez, J. A. Bajwa, and R. Mehanna, “Dbs programming: an evolving approach for patients with parkinsons disease,” *Parkinsons Disease*, 2017.
 - [10] A. Khemphila and V. Boonjing, “Parkinsons disease classification using neural network and feature selection,” *World Academy of Science, Engineering and Technology*, vol. 64, 2012.
 - [11] M. J. Berry and G. Linoff, *Mastering data mining*. Wiley New York, 2000.
 - [12] G. Shmueli, N. R. Patel, and P. C. Bruce, *Data mining for business intelligence: Concepts, techniques, and applications in Microsoft Office Excel with XLMiner*. John Wiley and Sons, 2011.
 - [13] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
 - [14] S. Bouktif, E. M. Hanna, N. Zaki, and E. A. Khouza, “Ant colony optimization algorithm for interpretable bayesian classifiers combination: application to medical predictions,” *PloS one*, vol. 9, 2014.
 - [15] M. Verma, M. Srivastava, N. Chack, A. K. Diswar, and N. Gupta, “A comparative study of various clustering algorithms in data mining,” *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, 2012.
 - [16] D. T. Larose and C. D. Larose, *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, 2014.
 - [17] 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.
 - [18] T. A. Shaikh and A. Chhabra, “Effect of weka filters on the performance of the naive bayes data mining algorithm on arrhythmia and parkinsons datasets,” 2014.
 - [19] M. Gök, “An ensemble of k-nearest neighbours algorithm for detection of parkinson’s disease,” *International Journal of Systems Science*, vol. 46, no. 6, pp. 1108–1112, 2015.
 - [20] 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.
 - [21] T. J. Hirschauer, H. Adeli, and J. A. Buford, “Computer-aided diagnosis of parkinsons disease using enhanced probabilistic neural network,” *Journal of Medical Systems*, vol. 39, no. 11, p. 179, 2015.
 - [22] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, “High-accuracy detection of early parkinson’s disease through multimodal features and machine learning,” *International Journal of Medical Informatics*, vol. 90, pp. 13–21, 2016.
 - [23] “Parkinson’s progressive markers initiative.” <http://www.ppmi-info.org/>, 2017.
 - [24] T. C. Sharma and M. Jain, “Weka approach for comparative study of classification algorithm,” *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 4, pp. 1925–1931, 2013.
 - [25] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, 2007.
 - [26] “Feature selection package - algorithms - information gain.” <http://featureselection.asu.edu/old/documentation/infogain.htm>, 2017.
 - [27] “The michael j fox foundation for parkinson’s research.” <https://www.michaeljfox.org/>, 2017.
 - [28] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, S. Y. Philip, et al., “Top 10 algorithms in data mining,” *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1–37, 2008.
 - [29] Q. Zou, J. Zeng, L. Cao, and R. Ji, “A novel features ranking metric with application to scalable visual and bioinformatics data classification,” *Neurocomputing*, vol. 173, pp. 346–354, 2016.
 - [30] R. Bhidayasiri and D. Tarsy, *Movement disorders: a video atlas*. Springer Science Business Media, 2012.
 - [31] P. Suganya and C. Sumathi, “A novel metaheuristic data mining algorithm for the detection and classification of parkinson disease,” *Indian Journal of Science and Technology*, vol. 8, no. 14, 2015.