The Journal of
Mathematics and Computer Science

Available online at
http://www.TJMCS.com

The Journal of Mathematics and Computer Science Vol. 4 No.2 (2012) 257 - 263

A Decision Support System for Parkinson's Disease Diagnosis using Classification and Regression Tree

A.H. Hadjahmadi*[1], and Taiebeh J. Askari[2]

[1]Faculty Member, Department of Computer Engineering Vali-e-Asr University of Rafsanjan, Iran,
hadjahmadi@vru.ac.ir

[2]Faculty Member, Islamic Azad University, Iranshahr Branch, Iranshahr, Iran,
t.askari@iauiranshahr.ac.ir

Received: January 2012, Revised: April 2012
Online Publication: June 2012

ABSTRACT

Parkinson's disease (PD) is a progressive disorder of the nervous system that affects movement. It develops gradually, often starting with a barely noticeable tremor in just one hand. But while tremor may be the most well-known sign of Parkinson's disease, the disorder also commonly causes a slowing or freezing of movement. Parkinson's disease is the second most common Neurodegenerative action only surpassed by Alzheimer's disease. However, a proper diagnosis at an early stage can result in significant life saving. A system for automated medical diagnosis would enhance the accuracy of the diagnosis and reduce the cost effects.

The present study compares the accuracy of several machine learning methods including Bayesian Networks, Regression, Classification and Regression Trees (CART), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) for proposing a decision support system for diagnosis of parkinson's disease. The proposed system achieved an accuracy of 93.7% using classification and regression tree. Sensitivity analysis via classification and regression tree was also used to find importance of input variables.

Key words: Parkinson's disease, Clinical Decision Support System, Classification and Regression Tree, Sensitivity analysis.

* Corresponding Author

INTRODUCTION

Parkinson's disease is one of a larger group of neurological conditions called motor system disorders. Historians have found evidence of the disease as far back as 5000 B.C. It was first described as "the shaking palsy" in 1817 by British doctor James Parkinson. Because of Parkinson's early work in identifying symptoms, the disease came to bear his name. Parkinson's disease is the second most common neurodegenerative action only surpassed by Alzheimer's disease [3].

In the normal brain, some nerve cells produce the chemical dopamine, which transmits signals within the brain to produce smooth movement of muscles. In Parkinson's patients, 80 percent or more of these dopamine-producing cells are damaged, dead, or otherwise degenerated. This causes the nerve cells to fire wildly, leaving patients unable to control their movements.

Though full-blown Parkinson's can be crippling or disabling, experts say early symptoms of the disease may be so subtle and gradual that patients sometimes ignore them or attribute them to the effects of aging. At first, patients may feel overly tired, "down in the dumps," or a little shaky. Their speech may become soft and they may become irritable for no reason.

Nowadays, with the development of computer technology, data mining has been widely used in various fields. Clinical decision support systems (CDSSs) form a significant part of the field of clinical knowledge management technologies through their capacity to support the clinical process and use of knowledge, from diagnosis and investigation through treatment and long-term care.

In the current work we compare the accuracy of several machine learning methods including Bayesian Networks, Regression, Classification and Regression Trees (CART), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) for proposing a clinical decision support system for diagnosis of parkinson's disease.

In this paper after a brief study of clinical decision support systems, classification and regression trees, support vector machines, and sensitivity analysis in Section II and the experimental results are shown in section III with detail.

A. *CLINICAL DECISION SUPPORT SYSTEM*

Clinical decision support systems (CDSSs) form a significant part of the field of clinical knowledge management technologies through their capacity to support the clinical process and use of knowledge, from diagnosis and investigation through treatment and long-term care.

Artificial intelligence (AI) systems are intended to support healthcare workers with tasks that rely on the manipulation of data and knowledge. Expert systems are the commonest type of CDSS in routine clinical use. They contain medical knowledge about a very specifically defined task. Their uses include: alerts and reminders, diagnostic assistance, therapy critiquing and planning, prescribing decision support, information retrieval, image recognition and interpretation. Reasons for the failure of many expert systems to be used clinically include dependence on an electronic medical record system to supply their data, poor human interface design, failure to fit naturally into the routine process of care, and reluctance or computer illiteracy of some healthcare workers. Many expert systems are now in routine use in acute care settings, clinical laboratories, educational institutions, and incorporated into electronic medical record systems. Some

CDSS systems have the capacity to learn, leading to the discovery of new phenomena and the creation of medical knowledge. These machine learning systems can be used to: develop the knowledge bases used by expert systems, assist in the design of new drugs, and advance research in the development of models from experimental data. Benefits from CDSS include improved patient safety, improved quality of care, and improved efficiency in health care delivery [4].

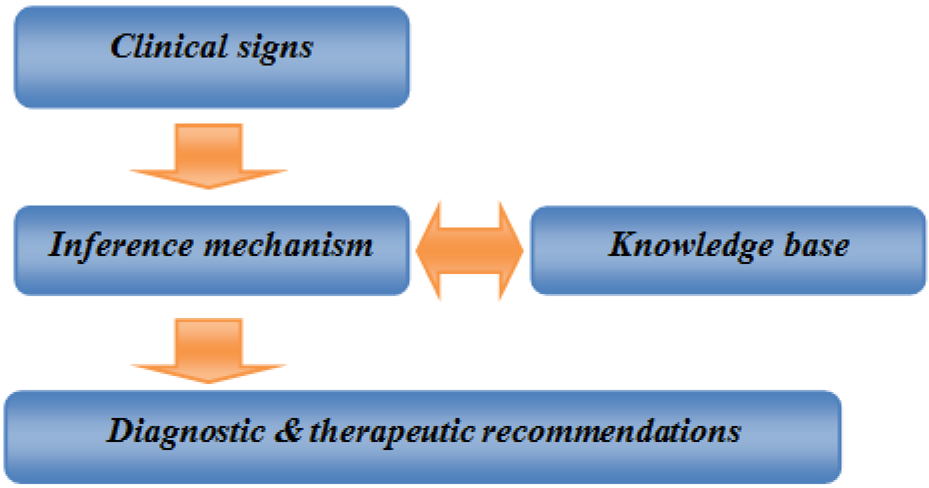

Fig 1: The General Model of CDSSs.

According to the general CDSS model described in Fig 1, the system users interact with the CDSS in an iterative fashion, selectively entering clinical signs, and using the CDSS output recommendations to assist with the diagnostic and therapeutic decision-making processes [1].

B. CLASSIFICATION AND REGRESSION TREE

Classification and Regression Trees is a classification method which uses historical data to construct so-called decision trees. Decision trees are then used to classify new data. In order to use CART we need to know number of classes a priori.

CART methodology was developed in 80s by Breiman, Freidman, Olshen, and Stone in their paper "Classification and Regression Trees" (1984). Solomon et al. (2006) said that the decision tree model is a powerful and popular tool to classify and predict data patterns since rules are generated more straightforward and relatively easy to be interpreted.

The decision tree begins with a root node *t* derived from whichever variable in the feature space minimizes a measure of the impurity of two sibling nodes. The measure of impurity at node *t*, denoted by $i(t)$, is as shown in the following equation (1),

$$i(t) = -\sum_{j=1}^{k} p(w_j|t) \log p(w_j|t) \tag{1}$$

Where $p(w_j|t)$ is the proportion of patterns x_i allocated to class w_j at node t. Each non terminal node is then divided into two further nodes, t_L and t_R, such that p_L, p_R are the proportions of entities passed to the new nodes t_L, t_R respectively. The best division is that witch maximizes the difference given in (2):

$$\Delta i(s,t) = i(t) - p_L i(t_L) - p_R i(t_R) \tag{2}$$

The decision tree grows by means of the successive sub-divisions until a stage is reached in which there is no significant decrease in the measure of impurity when a further additional division s is implemented. When this stage is reached, the node t is not sub-divided further, and automatically becomes a terminal node. The class w_j associated with the terminal node t is that which maximizes the conditional probability $p(w_j|t)$.

C. SUPPORT VECTOR MACHINE

Support vector machine is a powerful data mining technique for classifying data. The support vector machine is a training algorithm for learning classification and regression rules from data. SVM was developed form statistical learning theory and was first suggested by Vapnik in the 1960 for data classification. The basic idea of Support Vector Machines is to map the original data *X* into a feature space *F* with high dimensionality through a non linear mapping function and construct an optimal hyper-plane in new space.

In [2] Meyer, Leisch and Hornik concluded that for classification, simple statistical procedures and ensemble methods proved very competitive, mostly producing good results “out of the box” without the inconvenience of delicate and computationally expensive hyper-parameter tuning. As to the dispersion characteristics, the results for SVMs compared to other methods showed average precision. However, there is no evidence that the SVM results were particularly affected by either model bias or model variance [5].

D. SENSITIVITY ANALYSIS

Sensitivity analysis can be defined as the study of how the variation in the output of a model can be apportioned, (qualitatively or quantitatively) to different sources of input variation. It can provide valuable information regarding the structure of the model, and its reliance upon the input variables, or lack thereof (Saltelli 2000). The sensitivity of an input variable or parameter is an indication of the effect that a variation of that input will have on the output; an input variable of higher sensitivity will result in a greater variation of the output and vice versa. The sensitivity of a variable illustrates the care that modelers must take to obtain and employ an appropriate value for the variable, but can also signify its importance in relation to its dependency by the model structure (Saltelli et al. 1999).

PARKINSON DATASET

The data set provided by the UCI repository of machine learning databases for this study consists of 195 sustained vowel phonations from 31 male and female subjects, of which 23 were diagnosed with PD. The time since diagnoses ranged from 0 to 28 years, and the ages of the subjects ranged from 46 to 85 years (mean 65.8, standard deviation 9.8). Averages of six phonations were recorded from each subject, ranging from 1 to 36 seconds in length. The phonations were recorded in an IAC sound-treated booth using a head-mounted microphone (AKG C420) positioned at 8 cm from the lips. The microphone was calibrated using a Class 1 sound level meter (B&K 2238) placed 30 cm from the speaker. The voice signals were recorded directly to computer using CSL 4300B hardware (Kay Elemetrics), sampled at 44.1 kHz, with 16 bit resolution [6,7]. The data set was randomly divided into a training set (70%) and a test set (30%).

RESULTS AND ANALYSIS

The main aim of this research is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD. It is a binary classification problem. The feature information of Parkinson Dataset is shown in Table 1.

Table1. Feature information of Parkinson Dataset

Input feature	Discrption
MDVP:Fo(Hz)	*Average vocal fundamental frequency*
MDVP:Fhi(Hz)	*Maximum vocal fundamental frequency*
MDVP:Flo(Hz)	*Minimum vocal fundamental frequency*
MDVP:Jitter(%),Jitter(Abs), RAP, PPQ, 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*
RPDE,D2	*Two nonlinear dynamical complexity measures*
DFA	*Signal fractal scaling exponent*
spread1,spread2,PPE	*Three nonlinear measures of fundamental frequency variation*
Status	*Health status of the subject (one) - Parkinson's, (zero) - healthy*

In this experiment comparing of accuracy for machine learning classifiers such as … was measured by using our Parkinson data set. The same training and test sets were utilized in all experiments. The performance results of classifications are shown in Fig 2.

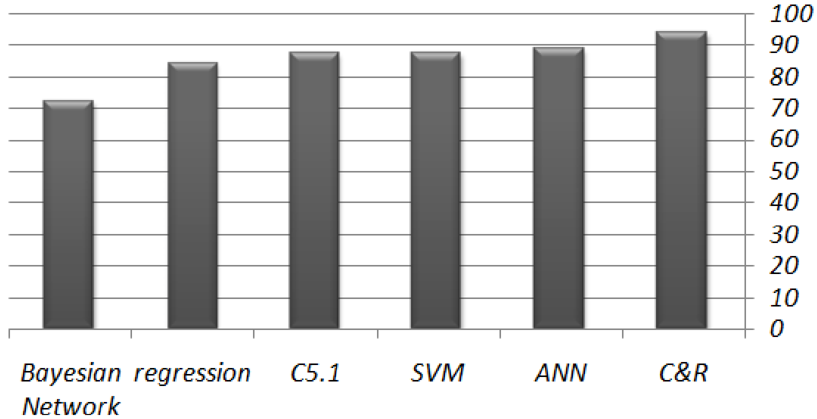

Fig 2: The classification results.

By comparing all the methods the classification and regression tree method with accuracy of 93.75% is the best classifier. Artificial neural network and support vector machine are the next best classifiers.
Computing of the variable importance is an important issue in many applied problems complementing variable selection by interpretation issues. Sensitivity analysis via classification and regression tree was also used to find importance of input variables and the results are shown in Fig 3.

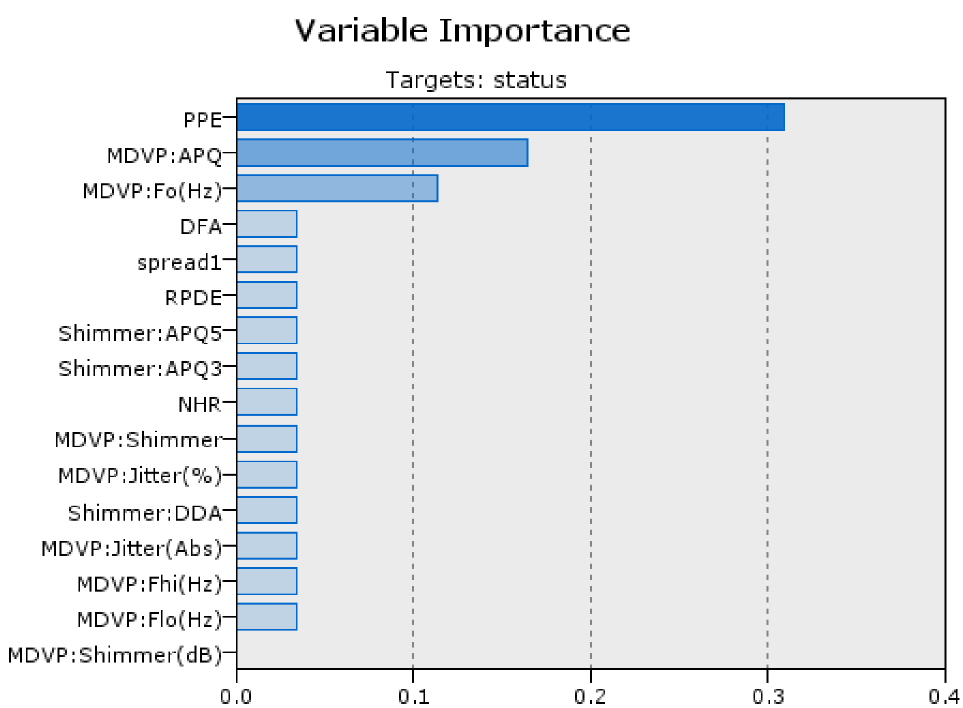

Fig 3: Input variables importance.

Fig 3 demonstrates that PPE is the most important variable and, "MDVP:APQ" and "MDVP:F0" are the next. It can be observed that the "MDVP:Shimmer" has no influence in this classification problem.

CONCLUSION

We have proposed an efficient and powerful method for Parkinson's disease (PD). Parkinson's disease is the second most common Neurodegenerative action only surpassed by Alzheimer's disease. However, a proper diagnosis at an early stage can result in significant life saving. A system for automated medical diagnosis would enhance the accuracy of the diagnosis and reduce the cost effects.

The present study compares the accuracy of several machine learning methods including Bayesian Networks, Regression, Classification and Regression Trees (CART), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) for proposing a decision support system for diagnosis of parkinson's disease. The proposed system achieved an accuracy of 93.7% using classification and regression tree. Sensitivity analysis via classification and regression tree was also used to find importance of input variables.

REFERENCE

[1] J. B. Yang, D. L. Xu, and G. Kong, Clinical Decision Support Systems: a Review of Knowledge Representation and Inference under Uncertainties. *International Journal of Computational Intelligence Systems*, Vol.1(2), pp. 159-167. 2008.

[2] D. Meyer, F. Leisch, and K. Hornik, The support vector machine under test, *Neurocomputing* Vol(55), pp. 169–186, 2003.

[3] D. Gil, M. Johnso, Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines, *Global Journal of Computer Science and Technology*, Vol(9), No 4, pp 63-71, 2009.

[4] E. Turban, and J. E. Aronson, Decision Support Systems And Intelligent Systems, Sixth Edition, ISBN 0-13-089465-6, 2000.

[5] B. Scholkopf, and A. J. Smola, Learning with Kernels. MIT Press, Cambridge, MA, ISBN 0-262-19475-9, 2002.

[6] 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, *BioMedical Engineering*, pp. 6-23, 2007.

[7] M.A. Little, P.E. McSharry, E.J. Hunter, L.O. Ramig, Suitability of dysphonia measurements for telemonitoring of Parkinson's disease, *IEEE Transactions on Biomedical Engineering*, Vol(56), pp. 1015-1022, 2008.