Pathological Voice Signal Analysis Using Machine Learning Based Approaches

Yahia Alemami[1] & Laiali Almazaydeh[2]

[1] Department of Computer Science, Al-Hussein Bin Talal University, Maan, Jordan

[2] Department of Software Engineering, Al-Hussein Bin Talal University, Maan, Jordan

Correspondence: Laiali Almazaydeh, Department of Software Engineering, Al-Hussein Bin Talal University, Maan, Jordan. Tel: 962-776-245-243. E-mail: laiali.almazaydeh@ahu.edu.jo

Abstract

Voice signal analysis is becoming one of the most significant examination in clinical practice due to the importance of extracting related parameters to reflect the patient's health. In this regard, various acoustic studies have been revealed that the analysis of laryngeal, respiratory and articulatory function may be efficient as an early indicator in the diagnosis of Parkinson disease (PD). PD is a common chronic neurodegenerative disorder, which affects a central nervous system and it is characterized by progressive loss of muscle control. Tremor, movement and speech disorders are the main symptoms of PD. The diagnosis decision of PD is obtained by continued clinical observation which relies on expert human observer. Therefore, an additional diagnosis method is desirable for most comfortable and timely detection of PD as well as faster treatment is needed. In this study, we develop and validate automated classification algorithms, which are based on Naïve Bayes and K- Nearest Neighbors (KNN) using voice signal measurements to predict PD. According to the results, the diagnostic performance provided by the automated classification algorithm using Naïve Bayes was superior to that of the KNN and it is useful as a predictive tool for PD screening with a high degree of accuracy, approximately 93.3%.

Keywords: pattern recognition, Parkinson disease, voice signal, Naïve bayes, KNN

1. Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative disorders, which affecting older people, with most cases occurring after the age of 50 (Janvin et al., 2006). This kind of disease affects a central nervous system which causes progressive loss of muscle control with the distinctive signs include shiver or shake in the hand, arm, leg, jaw, face, bradykinesia and Difficulty in swallowing, chewing, breathing, speech (Noth et al., 2011).

Many people with PD might experience additional problems like cognitive problems, emotional changes, thinking difficulties and sleep disorders (Brooks, 2012).

The genetic factors considers as the main factors which is might be the cause of PD. However, other factors could be associated with the relationship of developing PD (Nagal & Singla, 2012).

A patient with PD has a higher risk of tribulation dementia compared to the healthy person. However, the patient could be benefit from the scheduled treatment. It has been reported that the problems with the PD patient could be improved with planned treatment based on regular physical exercise, especially problems related to difficulty in speech, mobility and strength (Nagal & Singla, 2012).

A diagnosis of Parkinson's disease is subjective based on signs and symptoms review of the patient. Because there is no definitive test for the PD detection (Jankovic, 2008), therefore, new reliable methods based on clinical criteria for diagnosis and screening of PD are needed, in order to have a major benefit of the treatment on PD to be more effective.

Pattern recognition process could be very useful in medical domain for building computer-aided diagnosis systems that support final diagnosis decision. In pattern recognition process given unknown objects are assigned to prescribed known classes. However, a universally best pattern recognition system has never existed. The best system relies heavily on the set of features used for the classification task as well as the training method itself

(Bishop, 2006). Therefore, it might worth for employing different recognition methods and to keep the best-fitting method that gives the best accuracy.

In this work, we develop an efficient algorithm for automatic classification of voice signal features to detect abnormalities in speech in order to predict PD. We use Naïve Bayes and KNN to classify voice signal into normal and PD.

Several methods have been focused on presenting automatic methods for the identification of PD disease. The work in (Joan et al., 2010) proposed the analysis of neurons in brain in order to detect PD.

Different acoustic, articulatory and respiratory studies have underlying the pattern of voice and speech disorder characteristics in PD. Perceptually, Person with PD is characterized by hoarse voice quality, imprecise articulation and stress in his / her voice and speech (Ramig et al., 2008).

The study in (Canter, 1963) (Ho et al., 2001) (Fralle & Cohen, 1995) have reported a reduced frequency range in the speech of patient with PD. (Rosen et al., 2006) showed acoustic signatures that have phonetic variation with PD patients, where the patients and healthy people had conversational speech for two minutes. Also, the reductions in vocal sound pressure level (vocSPL) with PD patients were investigated in (Ramig et al., 2008) (Canter, 1963) (Metter & Hanson, 1986). In most recent, Fox and Raming founded that vocSPL was 2-4 decibel lower across a variety of speech tasks with PD patients (Ramig et al., 2008) (Constantinescu, 2010).

The study in (Little et al., 2007) proposed the analysis of complex nonlinear aperiodicity, non-Gaussian randomness and aero acoustic of the sound in order to evaluate voice disorder.

In (Little et al., 2009) pitch period entropy (PPE) measure was used to discriminate PD patients from healthy people. PPE considers as robust measure towards variations in speech frequency and noisy environment. The speech sounds from 23 patients with PD and 8 healthy people were collected, then the classification was performed using support vector machine, where the accuracy was 91.4%.

(Eichhorn et al., 1996) performed detection of PD through kinematic parameters analysis of handwriting movements. Subjects were asked to draw concentric circles using a digitizing tablet, then changes of different parameters such as mean peak acceleration and mean peak velocity were tested.

In our work, the speech signal are saved and processed off-line by an automated system we developed using MATLAB to compute different features measurements that can be used to increase the clinical usefulness of the PD diagnosis system.

In the following section, the methodology of our proposed system is described. Section 3, demonstrates the results of our system. Then, we conclude our paper in section 4, and highlight some directions for future research.

2. Method

The proposed system consists of two main parts which are features extraction and features classification. Features extraction process is based on statistical measures. Features classification process is learned through different classifiers, Naïve bayes and KNN.

2.1 Subjects

In this work we used Parkinson speech dataset with multiple types of sound recordings. This is available from the University of California, Irvine (UCI) machine learning repository website (http://archive.ics.uci.edu). UC Irvine machine learning repository maintains a growing collection of biomedical datasets from healthy subjects and patients as a service to the machine learning community.

The PD dataset consists of a range of voice signal measurements from 41 patients with PD and 52 healthy individuals who were recruited at the Department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University (Sakar et al., 2013). Multiple speech recordings of each patient is gathered and stored. From all subjects, 26 a wide variety of voice samples, including number from 1 to 10, nine words, four short sentences, and sustained vowels "a", "o", and "u" were recorded (Sakar et al., 2013).

2.2 Features Extraction

After collecting the speech dataset, a series of features are extracted from the voice samples. In this context, a group of 26 linear and time-frequency based features are parsed from the dataset considering the previous works (Little et al., 2009), (Sakar & Kursun, 2010). However, the mathematical fundamentals of these features presented by (Teixeira et al., 2011) and (Boersma, 1993).

These features are detailed as follows, respectively (http://archive.ics.uci.edu), (Sakar et. al, 2013):

- Frequency features 1-5: Jitter (local), Jitter (local, absolute), Jitter (rap), Jitter (ppq5), and Jitter (ddp). These features depend on the calculation of all periods by the waveform-matching.

- Amplitude features 6-11: Shimmer (local), Shimmer (local, dB), Shimmer (apq3), Shimmer (apq5), Shimmer (apq11), and Shimmer (dda). These features refer to the average absolute difference between the amplitude of a period and its closest neighbors.

- Harmonicity features12-14: Autocorrelation, Noise-to-Harmonic, Harmonic-to-Noise. These features related to the ratio of voice signal frequency to the reference signal frequency.

- Pitch features 15-19: Median pitch, Mean pitch, Standard deviation, Minimum pitch, Maximum pitch. These features represent the degree of lowness or highness of a tone.

- Pulse features 20-23: Number of pulses, Number of periods, Mean period, Standard deviation of period. The pulse features related to a single period.

- Voicing features 24-26: Fraction of locally unvoiced frames, Number of voice breaks, Degree of voice breaks. These features are identified with the absence or presence of periodic vibration of the acoustic chords.

As an example, figure 1 shows the main representation of jitter and shimmer features in voice signal (Teixeira et al., 2013).

The extracted features and downloaded data are fed into two different classifiers. The classification methodology is based on the Naïve Bayes classifier and K-nearest neighbor algorithm (KNN) with different cross-validation methods and accuracy, specificity and sensitivity evaluation metrics are reported.

2.3 Naïve Bayes Classifier

In this research, we apply a Naïve Bayes as a classifier to identify the diagnostic performance of PD using voice signal features.

Naïve Bayes classifier is used in supervised learning method and it is based on 'probability' concept to classify new entities. Meanwhile, it assigns a new observation to the most probable class.

The process of classification based on two steps, as follows:

1. Training step: Using the training samples, the method computes the probability distribution of that sample.

2. Prediction step: For test sample, the method computes the posterior probability of that unknown instance. The posterior is predicting that the sample belonging to each class according to the largest posterior probability, which is called Maximum A Posterior (MAP).

However, Naïve Bayes classifier is one of the most practical learning methods and it has been proven to be extremely successful in assisting medical specialties in medical diagnosis applications.

2.4.K-Nearest Neigbor (KNN) Classifier

In this work, in order to identify the diagnostic performance of PD, we use KNN as a classification method. KNN is one of the pattern recognition techniques used to build a classification model. In classification context, this means that samples with similar input values likely belong to the same class. Conceptually, in KNN the concept of "nearness" is used to classify new entities. It classifies an instance by finding its nearest neighbors and picking the most popular class among the neighbors.

The value of K determines the number of nearest neighbor to consider. Thus, when K =1, then only the closest training neighbor is examined to predict the class of the new sample. When K =2, then the 2 nearest neighbors are considered.

As with KNN, the classification of a data sets is dependent on their similarity with neighbors, some measures of similarity are needed to determine how close two samples are together. This is necessary to determine which samples are the nearest neighbors. Distance measures such as distance are commonly used. Other distance measures that can be used, including Manhattan and hamming distance (Suguna & Thanushkodi, 2010).

KNN can generate complex decision boundaries allowing for complex classification decisions to be made. However, it can be susceptible to noise, because classification decisions are made using only information about a few neighboring points instead of the entire dataset.

3. Results

Here there are 83% training samples and each sample have 26 feature values and the size of testing dataset is 17% samples and 26 features for each testing sample. Both classifiers, Naïve Bayes and KNN return the class level of testing samples identified by the help of training samples, each classifier based on its classification methodology, i.e. for Naïve Bayes based on "probability" and for KNN based on "nearness".

We evaluated the classification performance of both classifiers configurations on the test set. Sensitivity, specificity for testing and accuracy are computed. A confusion matrix is generated for both classifiers.

Sensitivity reflects the rate of PD positive subjects correctly classified while specificity is equal to the true negative rate. The classification accuracy here refers to the ratio of correct decisions (i.e., true positive plus true negatives) to the total number of cases (Almazaydeh et al., 2012).

Figure 2 and 3 shows the confusion matrix for testing set classification of Naïve Bayes and KNN, respectively. The confusion matrix shows the total percent of correctly classified cases and the total percent of misclassified cases.

The results show very good KNN with K=1 diagnostic performance of 80% and high Naïve Bayes diagnostic performance with an accuracy of 93.3% correct detection rate (sensitivity 87.5%, and specificity 100%). As for treatment decisions, a higher specificity is preferable in order not to inflict further investigation or treatment on patients without the disease (False positive).

Power analysis could be performed after the experimental work to tell you what you should have known earlier and to estimate the sample size required for the experiment. Using the power analysis with a confidence of 95 percentile showed that the data set size will be 57, but we have used larger size of 93.

From the experimental results for this particular problem with the used data set which is large enough, we could conclude as a result that the Naïve Bayes has a better performance than KNN, and Naïve Bayes using voice signal measurements is a practical and useful screening test to estimate whether patients have Parkinson disease or not.

4. Conclusion

In this work, we studied the possibility of the detection of Parkinson disease from the voice signal variation patterns. We further developed a model using the voice signal features and evaluated its effectiveness. From the experimental results, we conclude that the recognition rate of KNN comes out to be 80%, but, with Naïve Bayes the classification has the highest recognition rate which is 93.3%. So we can say that it is better for recognition.

The developed system was confirmed that voice signal measurement is a practical and useful screening test to estimate whether patients have PD or not.

As a future work, we plan to incorporate this work into a real-time monitoring system that acquires and analyze the speech signal of subjects rather than analyzing the off-line signal.
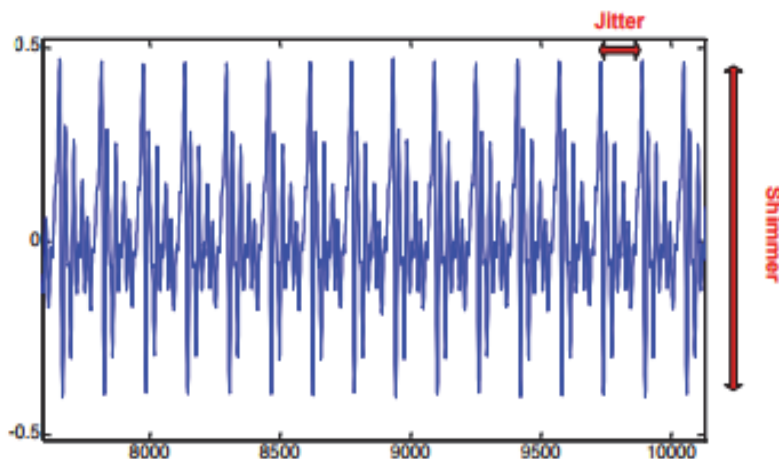

Figure 1. Representation of jitter and shimmer features in voice signal (Teixeira et al., 2013)
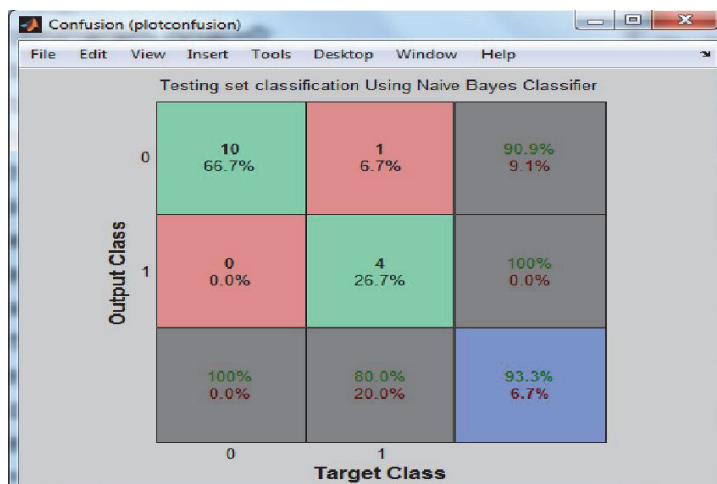
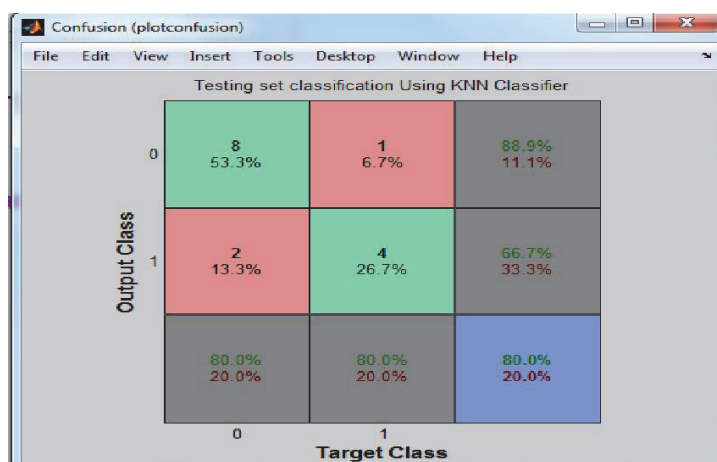Figure 2. Confusion matrix of Naïve Bayes for testing set classification

Figure 3. Confusion matrix of K-NN for testing set classification

References

Almazaydeh, L., Faezipour, M., & Elleithy, K. (2012). A Neural Network System for Detection of Obstructive Sleep Apnea through SpO$_2$ Signal Features. *International Journal of Advanced Computer Science and Applications, (IJACSA), 3*(5), 7-11.

Bishop, C. (2006). Pattern recognition and machine learning. Springer.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonic-to-noise ratio of a sample sound. IFA Proceedings, 97-110.

Brooks, D. (2012). Parkinson's disease: diagnosis. *Parkinsonism & related disorders, 18*, S31-S33.

Canter, G. J. (1963). Speech characteristics of people with parkinson's disease: I. Intensity, pitch and duration. J. Speech Hear. *Diord, 28*, 221-229.

Constantinescu, G. (2010). Home-based speech treatment for Parkinson's disease delivered remotely: A case report, 100-104.

Eichhorn, E. (1996). Computational analysis of open loop handwriting movements in Parkinson's disease: A rapid method to detect dopamimetic effects. *Movement disorders, 11*(3), 289-297.

Fralle, V., & Cohen, H. (1995). Prosody in Parkinson's disease: relations between duration, amplitude and fundamental frequency range," Presented at: International Neuropsychological Society 23rd Annual Meeting.

Ho, A. K., Iansek, R., & Bradshaw, J. L. (2001). Motor instability in parkinsonian speech intensity.

Neuropsychiatr. Neuropsychol. Behav. *Neurol, 14*, 109-116.

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *J. Neurol Neurosurg Psychiatry, 79,* 368-376.

Janvin, C., Larsen, J., Aarsland, D., & Hugdahl, K. (2006). Subtypes of mild cognitive impairment in Parkinson's disease: progression to dementia. *Movement Disorders, 21*(9), 1343-1349.

Joan, K., Tara, M., Whitehill, L., & So, S. (2010). Intonation Contrast in Cantonese Speakers with Hypokinetic Dysarthria Associated with Parkinson's disease. *J. Speech Lang Hear Res, 53*, 836-849.

Little, M. A., McSharry, P. E., Roberts, S. J., Costello, D. A. E., & Moroz, I. M. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Biomed. Eng. Online, 6*(23). https://doi.org/10.1186/1475-925X-6-23

Little, M., McSharry, P., Hunter, E., Spielman, J., & Ramig, L. O. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE Transactions on Biomedical Engineering, 56*(4), 1010–1022.

Metter, E. J., & Hanson, W. R. (1989). Clinical and acoustical variability in hypokinetic dysarthria. J. Commun. Disord, 19, 347-366.

Nagal., A., & Singla, R. (2012). Parkinson's Disease: Diagnosis, Therapeutics & Management. WebmedCentral, 1-21.

Noth, E., Stemmmer, G., & Ruzickova, H. (2011). Detection of persons with Parkinson's disease by acoustic, vocal, and prosodic analysis. IEEE workshop on Automatic Speech Recognition and Understanding (ASRU), 478-483.

Ramig, L., Fox, C., & Sapir, S. (2008). Speech treatment for parkinson's disease. Expert Reviews. *Neurotherapeutics, 8*(2), 299-311.

Rosen, K., Kent, D., Delaney, A., & Duffy, R. (2006). Parametric Quantitative Acoustic Analysis of Conversation Produced by Speakers with Dysarthria and Healthy Speakers. *J. Speech Lang Hear Res., 49,* 395-411.

Sakar, B., Isenkul, M., Sakar, C., Sertbas, A., Gurgen, F., Delil, S., Apaydin, H., & Kursun, O. (2013). Collection and Analysis of a parkinson Speech Dataset with Multiple Types of Sound Recotdings. *IEEE Journal of Biomedical and Health Informatics, 17*(4), 828-834.

Sakar, C., & Kursun, O. (2010). Telediagnosis of Parkinson's disease using measurements of dysphonia. *J. Med. Syst., 34*(4), 591–599.

Suguna, N., & Thanushkodi, K. (2010). An improved k-nearest neighbor classification using genetic algorithm. *International Journal of Computer Science Issues, 7*(2), 18-21.

Teixeira, J., Ferreira, D., & Carneiro, S. (2011). Análise acústica vocal - determinação do Jitter e Shimmer para diagnóstico de patalogias da fala. In 6º Congresso Luso-Moçambicano de Engenharia. Maputo, Moçambique.

Teixeira, J., Oliveira, C., & Lopes, C. (2013). Vocal acoustic analysis-jitter, Shimmer and HNR parameters. *Procedia technology, 9,* 1112-1122.

Copyrights