Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 132 (2018) 1788–1794

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2018)

Predicting Severity Of Parkinson's Disease Using Deep Learning

Srishti Grover, Saloni Bhartia, Akshama, Abhilasha Yadav, Seeja K. R.*

Department of Computer Science and Engineering
Indira Gandhi Delhi Technical University for Women, Kashmere Gate, Delhi, 110006, India

Abstract

Parkinson's disease is a progressive and chronic neurodegenerative disorder. As the dopamine-generating neurons in parts of the brain become damaged or die, people begin to experience difficulty in speaking, writing, walking, or completing other simple tasks. These symptoms grow worse over time, thus resulting in the increase of its severity in patients. In this paper, we have proposed a methodology for the prediction of Parkinson's disease severity using deep neural networks on UCI's Parkinson's Telemonitoring Voice Data Set of patients. We have used 'TensorFlow' deep learning library of python to implement our neural network for predicting the severity. The accuracy values obtained by our method are better as compared to the accuracy obtained in previous research work.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

Keywords: Parkinson's Disease; Deep Neural Networks; TensorFlow; UPDRS

* Corresponding author. Tel.: +91-9968217403
E-mail address: seeja@igdtuw.ac.in

1877-0509 © 2018 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/3.0/)
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018).
10.1016/j.procs.2018.05.154

1. Introduction

Parkinson's disease is a neurodegenerative disorder, resulting in progressive degeneration of functions related to motor abilities of the patients due to the damage of dopamine-producing brain cells. The various symptoms of this disease are shaking, difficulty in movement, behavioural problems, dementia and depression. The prime motor symptoms, together, are referred to as "Parkinsonism", or a "Parkinsonian Syndrome". The changes in patients' voice is one of the common symptoms which can be identified analysing the patients' voice data. Patient's voice tends to stutter and progressively becomes affected as the disease becomes more severe.

Deep learning has become a popular technique for effectively analysing unstructured data like speech and audio signals. Deep neural networks use multiple layers of neurons stacked together to create classification and feature selection models. In this paper, deep learning is used to analyse voice data of the patient to classify it into "severe" and "not severe" classes. The evaluation metrics used in this research are the two UPDRS (Unified Parkinson's Disease Rating Scale) scores - total UPDRS and motor UPDRS scores. The motor UPDRS evaluates the motor ability of the patient on a scale of 0-108, whereas total UPDRS provides a higher range of the score scale ranging from 0-176.

2. Literature Survey

A lot of research has been done to predict Parkinson's disease in a patient, but less work has been reported to predict its severity. These works have used various machine learning techniques. In a survey by Das et al. [1] on the application of various classification techniques in diagnosing the Parkinson's disease(PD), neural network was found as the better classifier compared to regression and decision tree. In most of the reported research, the features extracted from speech signals [6,7,15] are used for predicting the severity of PD.Genain et al. [2] used Bagged decision trees to predict the PD severity from voice recordings of patients and found an improvement of 2% accuracy. Maleket al.[3] used 40-features dataset and identified 9 best features using Local Learning Based Feature Selection (LLBFS) to classify PD subjects into four classes (Healthy, Early, Intermediate & Advance), based on their UPDRS score. Cole et al.[4] explored the use of dynamic machine learning algorithms for identifying the severity of tremors and Dyskinesia from the data collected from wearable sensors. Angeles et al.[5] developed a sensor system to record kinetic data from the arm in order to assess symptom severity changes during Deep Brain Simulation Therapy. Nilashiet al.[8] proposed a new hybrid intelligent system using Adaptive neuro fuzzy inference system(ANFIS) and Support Vector Regression(SVR) for predicting the PD progression. Chen et al.[9] proposed a PD diagnostic system using PCA for feature extraction and Fuzzy KNN for classification.Polat[10] proposed a model using Fuzzy C-Means (FCM) clustering and KNN to diagnose the PD. Åström and Koker[11] designed a PD prediction system using parallel feed forward Neural Network and then the output is compared against a rule-based system for making the final decision. Li et al.[12], suggested a fuzzy based nonlinear transformation method where PCA is used for feature extraction and SVM for PD prediction. Hariharanet al.[13] proposed a hybrid intelligent system using clustering, feature reduction and classification methods for accurate PD diagnosis.

3. Proposed Methodology

The proposed methodology for predicting the Parkinson's disease severity using deep learning is outlined in Fig. 1. In first step, the voice data of PD patients is collected for analysis. Then the collected data is normalized using min-max normalization. In the next step deep neural network is designed with input layer, hidden layers and output layer. The number of neurons in the input layer is fixed as the number of attributes in the input data. The output layer contains two neurons corresponds to the two classes – "severe" and "non-severe".The normalized data is fed into the constructed deep neural network for training and testing.

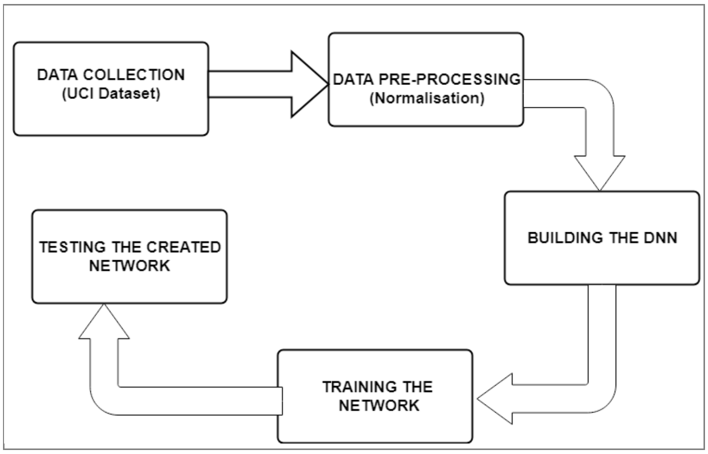

Fig. 1. Proposed deep learning framework

4. Implementation

The proposed model is implemented in a system with Intel Core i5-5200U CPU @2.20GHz and 8 GB RAM. The Python library, TensorFlow (tf.estimator) [14] is used to implement the Deep Neural Network.

4.1. Data Collection

We have used ParkinsonsTelemonitoring Voice Data Set[15]from UCI Machine Learning Repository. The dataset comprises of biomedical voice measurements of 42 patients. The various attributes of the data are subject number, subject age, subject gender, time interval, Motor UPDRS, Total UPDRS and 16 biomedical voice measures. The datasetcontains 5,875 voice recordings of these patients. The data is in ASCII CSV format. On an average, there are approximately 200 recordings collected from each patient (identification can be done through the first attribute - subject number).

4.2. Data Preprocessing

We have normalised the dataset into a range of 0-1 using min-max normalization. The normalization was done column wise using equation (1):

$$Normalized\ value\ of\ x = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Where x= column value, min(x) = minimum value for that column and max(x) = maximum value for that column.

4.3. Building the DNN

The total-UPDRS score in the dataset ranges from a minimum value of 5.0377 to a maximum value of 54.992, while it ranges from minimum of 5.0377 to maximum of 39.511 in case of motor-UPDRS scores. We created the train and test datasets by splitting the normalized dataset into parts of 80% (for train dataset) and 20%(for test dataset). Further, separate train and test datasets were created for both motor UPDRS score and total UPDRS score, keeping each of these scores as the output variable in their corresponding files.

The normalized values of 16 biomedical voice measures namely Jitter (%), Jitter (Abs), Jitter RAP, PPQ5, DDP, Shimmer(dB), Shimmer: APQ3, APQ5, Shimmer: APQ11, Shimmer DDA, NHR, HNR, RPDE, DFA, PPE are selected as features for classification. The output classes are ‘non-severe’ and ‘severe’. We have defined the range for the various metrics for severe and non-severe classes as shown in Table 1 due to the limitation of values in the dataset.

Table 1. Severity Class range

Metric	Severe	Non-severe
Total-UPDRS	Above 25	0-25
Motor-UPDRS	Above 20	0-20

The algorithm takes in the input dataset and creates an input pipeline, and defines iterators over it, these are variables which helps in scanning over data set. The defined algorithm also provides the functionality to shuffle the dataset in order to provide randomness. After defining the input pipeline, the second step is to feed the input data into the training model, this is done with the help of lambda function. The model after getting data performs training, evaluation and prediction. The training is done by defining arrays of hidden layer with the pre-initialised weights to the layer, which creates and saves model, in the processing system. And in the end, we perform evaluation of the resultant DNN classifier.

The DNN Classifier was built using TensorFlow with Keras as the backend. Our neural network contains 16 units in the Input layer, 10, 20, 10 neurons in each of the 3 hidden layers respectively.The network was further trained with 1000 and 2000 steps respectively.

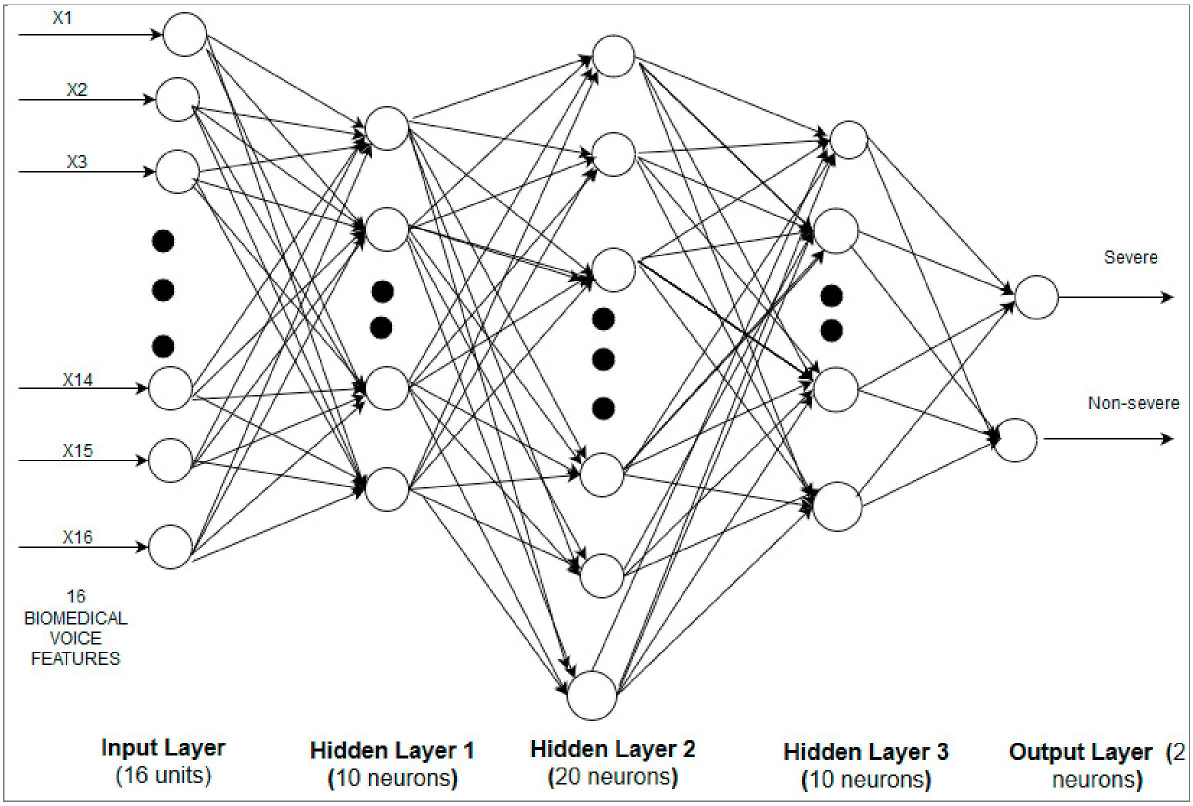

Fig. 2. Proposed Deep Neural Network

5. Experiments and Results

5.1. Experiment 1 (PD Severity prediction on the basis of Total UPDRS Score)

The input dataset is the 16 biomedical voice features and the output variable is Total UPDRS score. The classification accuracy obtained is 94.4422% and 62.7335% for train and test dataset respectively.
We compared our result with that of the work by Nilashiet al.[8] , since they evaluated their model on the same UCI Parkinson's Telemonitoring Voice Data set. They used ANFIS and SVR to predict the Parkinson's Disease advancement. Their work produced an average accuracy of 47.2% for the Total UPDRS score. The performance comparison of classifiers for Total UPDRS score is shown in Fig.3.

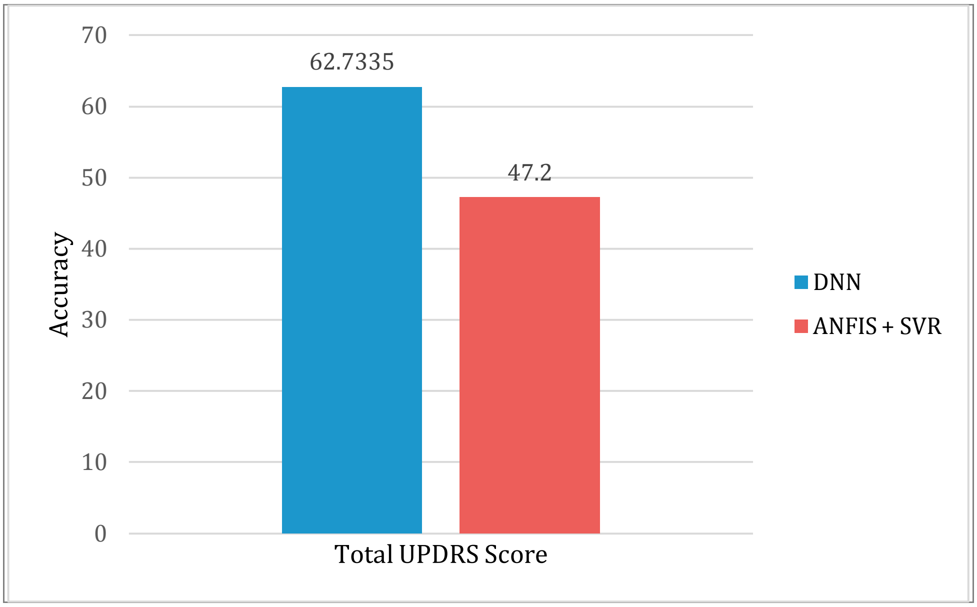

Fig. 3. Accuracy Comparison for Total UPDRS Score

5.2.Experiment 2(PD Severity prediction on the basis of Motor UPDRS Score)

In this experiment, the input dataset is the 16 biomedical voice features and the output variable is Motor UPDRS score. The classification accuracy obtained is 83.367% and 81.6667% for train and test dataset respectively. In comparison to this, the methodology proposed by Nilashiet al.[8] produced an average accuracy of 44.3% for the Motor UPDRS score.The performance comparison of classifiers for Motor UPDRS score is shown in Fig.4.

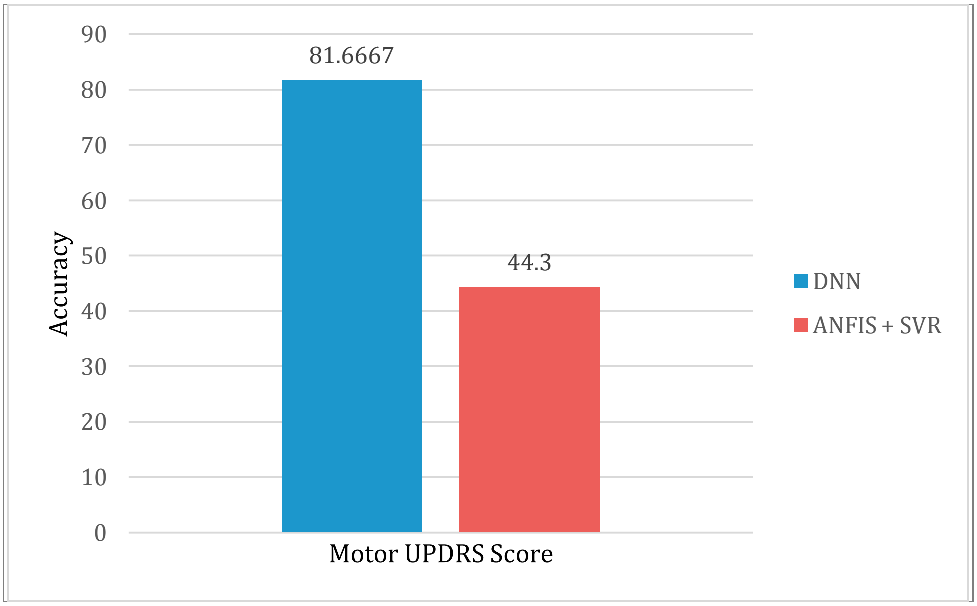

Fig. 4. Accuracy Comparison for Motor UPDRS Score

6. Conclusion

In this paper, we have implemented a deep neural network to predict the severity of Parkinson's disease. The proposed DNN model achieved better accuracy compared to other existing techniques. It is also found that the classification based on motor UPDRS score is better than the classification based on total UPDRS score and hence it can be concluded as a better metric for Severity Prediction. Although we have used a dataset of 5875 instances, the accuracy of our approach can be further improved by implementing it on a larger dataset, having more number of instances of each severity class as well as on a combined database of patients' voice data and other patient attributes like gait and handwriting features.

References

[1] Das R. (2010) "A comparison of multiple classification methods for diagnosis of Parkinson disease". *Expert Systems With Applications*"; **37**:1568-1572.

[2] Genain N, Huberth M, Vidyashankar R. (2014) "Predicting Parkinson's Disease Severity from Patient Voice Features."

[3]Benmalek E, Elmhamdi J, Jilbab A. (2015) "UPDRS tracking using linear regression and neural network for Parkinson's disease prediction." *International Journal Of Emerging Trends & Technology In Computer Science (IJETTCS)*; **4**:189-193.

[4] Cole B, Roy S, De Luca C, Nawab S. (2014) "Dynamical Learning and Tracking of Tremor and Dyskinesia From Wearable Sensors." *IEEE Transactions On Neural Systems And Rehabilitation Engineering*; **22**:982-991.

[5] Angeles P, Tai Y, Pavese N, Wilson S, Vaidyanathan R. (2017) "Automated Assessment of Symptom Severity Changes during Deep Brain Stimulation (DBS) Therapy for Parkinson's Disease." *International Conference On Rehabilitation Robotics (ICORR)*.

[6] Galaz Z, Mzourek Z, Mekyska J, Smekal Z, Kiska T, Rektorova I et al. (2016) "Degree of Parkinson's Disease Severity Estimation Based on Speech Signal Processing." *39th International Conference On Telecommunications And Signal Processing (TSP).*

[7] Asgari M, Shafran I. (2010) "Extracting Cues from Speech for Predicting Severity of Parkinson's Disease." *IEEE International Workshop On Machine Learning For Signal Processing.*

[8] Nilashi M, Ibrahim O, Ahani A. (2016) "Accuracy Improvement for Predicting Parkinson's Disease Progression". *Scientific Reports*; **6**. 34181

[9] Chen H, Huang C, Yu X, Xu X, Sun X, Wang G et al. (2013) "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach." *Expert Systems With Applications*; **40**:263-271.

[10] Polat K. (2012) "Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering." *International Journal Of Systems Science*; **43**:597-609.

[11] Åström F, Koker R. (2011) "A parallel neural network approach to prediction of Parkinson's Disease." *Expert Systems With Applications*; **38**:12470-12474.

[12] Li D, Liu C, Hu S. (2011) "A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets." *Artificial Intelligence In Medicine*; **52**:45-52.

[13] Hariharan M, Polat K, Sindhu R. (2014) "A new hybrid intelligent system for accurate detection of Parkinson's disease." *Computer Methods And Programs In Biomedicine*; **113**:904-913.

[14] Getting Started with TensorFlow. https://www.tensorflow.org/versions/r1.1/get_started/get_started

[15] Tsanas A, Little M, McSharry P, Ramig L. (2010) "Accurate Telemonitoring of Parkinson's Disease Progression by Noninvasive Speech Tests." *IEEE Transactions On Biomedical Engineering*; **57**:884-893.