A Comprehensive Study Of Prediction Of Parkinson's Disease Using Machine Learning

1st Kirti Wanjale
Computer Engineering Department
VIIT (SPPU University)
Pune, India
kirti.wanjale@viit.ac.in

2nd Madhavi Nagapurkar
Computer Engineering Department
VIIT (SPPU University)
Pune, India
madhavi.nagapurkar@viit.ac.in

3rd Parag Kaldate
Computer Engineering Department
VIIT (SPPU University)
Pune, India
parag.kaldate@viit.ac.in

4th Onkar Kumbhar
Computer Engineering Department
VIIT (SPPU University)
Pune, India
onkar.kumbhar@viit.ac.in

5th Subhranil Bala
Computer Engineering Department
VIIT (SPPU University)
Pune, India
subhranil.bala@viit.ac.in

***Abstract*—Parkinson's sickness is caused due to a nervous breakdown of dopamine secreting cells due to which several techniques have been used to detect its severity and its existence. Therefore this research area has got enormous attention to work upon. Until now several procedures have been used by various researchers, some of them which include tracking through the UPDRS Scale, Using inbuilt sensors of the smartphones, using various physical tasks, etc. This paper gives a significance of all the above-said methods as well as introduce to some more methods, which take unique considerations and assumptions, which are applied via multiple machine learning models and displaying their inter-model comparisons. It also covers the advancements and challenges. We conclude with a thorough study and detailed impact on a basket full of methods concerning several publications, journals/venues, and subtopics.**

***Index Terms*—Machine Learning, Parkinson, Smartphones, Random Forest, Neural Networks**

I. Introduction

Parkinson's disease (PD) is a self developing disorder caused due to the degeneration of the neurons in the medial section of the brain which leads to the issue in the human nervous system, hence it is found to be the second most ordinary Neuro degenerative disease after Alzheimer's disease. It is affected due to the declining levels of dopamine in the brain. The primary symptoms found in PD patients are tremor, muscle stiffness, bradykinesia, hypokinesia(lack of movement). It is found that the following reasons may also lead to Parkinsonism such as drugs and a few rare conditions like multiple cerebral infarction, supra-nuclear palsy and multiple system atrophy etc. Parkinsonism is prevalently a disease with the disorder in the movements with a few problems such as depression including dementia which leads to improper mobility of movements in action. The studies have shown that maximum patients with PD show rupture in vocal speech for the detection of PD. So during the initial phase, the diagnosis could proceed with speech as a field of study in detecting PD apart from tremor, muscle stiffness, lack of movements, etc. Since the speech following disorders is noticed – Viable pitch, iteration of words and syllables, exalt in speech, unintended pauses, and strident breath. Although, various diseases portray similar symptoms which makes it tedious to diagnose PD at an early stage with more accuracy. It is observed that the disorder in the speech is found five years before the diagnosis of the patient with PD.

II. Data Assembling

A. Gathering the Data

In [1] the data utilized was collected at the time of mPower study, which was an observational study regarding PD. It was entirely based on a smartphone app. The participants(inclusive of both PD and generals) were required to execute 4 tests on the smartphone which were walking, voice recording, tapping game and memory game 3 times per day on their own. Data in [2] was generated from 40 participants(22 control and 18 PWP) using a smartphone and a wired head-worn omnidirectional microphone in .wav format. It comprises the users to record vowel /a/ and consonant /m/ while following protocols. Specifically for vowel 'a' demographic data and vocal phonation records are extracted from mPower database in [3].In [4], Data is collected by making the participants read the first paragraph of 'Rainbow Passage' which includes 3 pauses - Mandatory Pause Phrase as MPP, Optional Pause Phrase as OPP and No Pause Phrase as NPP. Tri-Axial Accelerations from the accelerometer sensor is extracted though the LG Optimus S in [5] by making the 20 users inclusive of 10 PD and 10 generals walk for 20 steps from point A to point B and back to point A for gait and sway stand for 30 seconds for 4 times a day - one before taking levodopa, one after 1 hour of taking it, one in mid-afternoon and one before bed. In [6] data is gathered from 3 clinics that perform 3 tasks first is the phonation task, second is the diadochokinetic task and third is the reading task within a time duration of 4 min using a portable device. In [7], 100 Spanish speakers from Columbia

(25 men/25 women for both PD and general) each performing 42 speech tasks containing 24 isolated words, 10 sentences, 1 reading text, 1 monologue and rapid repetition of syllables /pa-ta-ka, /pa-ka- ta/ and /pe-ta-ka. In [8] Exceptionally little datasets (normally under 60 PD cases) are utilized in most studies performed so far with different achievement, utilizing 'a' phonation from 33 PD and 10 HC subjects,using 'e' phonation from 20 PD and 20 HC subjects,using 'i' phonation from 50 PD and 50 HC subjects.In [9] the data used in this study is the voice recording originally done at the Oxford University by Max Little. The voices of 31 peoples among which 23 are suffering from PD and remaining have their normal character. 195 entries are extracted from the voice data of 31 people in which 147 are of PD and the remaining 48 are normal.Data in [10]were collected through mPower where an iPhone app that uses digital bio-markers and health data by recording the voices of participants with and without PD. And a unique id is assigned to each entry in the dataset.

III. Data Pre-Processing and Model Building

A. *Patrick Schwab et ał*

In [1] for improving computational performance the input signals were down-sampled for walking, tapping and voice test by factors of 10,10 and 4 respectively. After downsampling, the size of sensor data was truncated to fixed lengths for all test types. For the walking, voice, memory and tapping tests the fixed lengths were 300,250,25 and 300. In the matter of the voice test, instead of passing the raw voice signal, Mel-Frequency cepstral was passed which was extracted from the audio. The accelerometer data of tapping and walking test was standardized by using mean as 0 and variance as 1. The approach to distinguish people with and without PD has two stages. In the first stage, specialized predictive models are trained to predict the diagnosis from the signal data given for all types of tests. For the second phase, the specialized model outcomes, and metadata are aggregated to form a single diagnostic prediction using EAM. CNN is used in walking, tapping and voice test and RNN is used in memory tests show the outputs of the hierarchical neural network mechanism which was employed on a user having PD who performed 18 tests. The timeline depicts the tasks performed by the user in the given order.

B. *R. Viswanathan, et ał*

In [2] the database containing the /a/ and /m/ phonation undergoes feature extraction and 60 features are extracted. The SVM classifier with Radial Basis Function as RBF Kernel is used to classify the patients as PWP and control subjects using the reduced feature set. The validation method that was used is to leave out validation. To find the efficiency of binary classifier Matthews Correlation Coefficient (MCC) was calculated. To determine the feature distribution Anderson-Darling normality test was used. Association strength in the reduced features and UPDRS-III score was calculated by the Spearman Correlation coefficient. The ROC curves for both the models which were based on /a/ and /m/ phonation are provided.

C. *Monica Giuliano et ał*

In [3] PCA was done 76.7% variance and 62 components were reduced to 33 components, further PCA was done again with 69.5% variance. K-means clustering was implemented on 9-factor PCA. ANOVA was also implemented on the same. An MLP model was generated, trained and tested with 5 voice parameters.the ROC curve for the MLP model which is generated using 5 voice parameters as well as including age and gender.the ROC curve for the LR model which is generated using 5 voice parameters as well as including age and gender.

D. *Goberman A.M et ał*

Data is extracted prior to 30 minutes earlier before their normal schedule of levodopa-carbidopa medication in the morning in [4].All the patients were almost 8 hr delay from their proper schedule during which the levels of dopamine were the least.Also, the data is taken after taking the proper practice of the required tasks.data is recorded using quality cassette tapes.Patients were made to read the famous Rainbow passage with 3 conditions MPP, OPP and NPP.first one is mandatory to pause phrase which is at "beautiful colors. These take", next one optional pause phrase which is located at "path high above and its," and the third one is no pause phrase at "boiling pot of gold". after this Each phrase is broken down to 8 segments (i.e. 24 segments). Each segment is either latency or a period. With 24 such segments, 12 relative time ratios were created. With the ratios for each pause phrase - mean, standard deviation and range were calculated as shown. Now after calculating all the 12 ratios, for reliability check, Pearson-product moment correlation (PPMC) analysis is used to calculate the interjudge and intrajudge reliability.Now the Second test which is applied on these 12 ratios were the Man Whitney U test (alpha level = 0.05) and the effect size analysis test to find the correlation between independent and dependent variables.

E. *Siddharth Arora et ał*

In [5] after the data extraction process, feature extraction is done and more than enough features are extracted. some of them are Mean, Standard deviation,25th percentile,75th percentile, etc.After this, Random Classifier is applied which provides too good sensitivity almost in the range of 98% - 99% and specificity of 97% - 98%.To make sure that random classifier is working as per our needs random classifier and conditional random classifier is used.also, avg of sensitivity and specificity are known as balanced accuracy is calculated and 10 fold cross-validation is repeated 100 times to make the model predict more accurately.

F. *Lireza Bayestehtashka et ał*

Data in [6] paper is collected from 3 clinics in total from 168 subjects.using openSIMILE in a total of 1582 features were extracted.tasks including the phonation task, in this the subject is made to keep its voice steady for 10s.The second task is Diadochokinetic in which the subject is made to repeat the defined order of syllables /pa/ta/ka as clearly and fast as possible.In the third task, subjects were made to read three passages ”The North Wind and the Sun”, ”The Rainbow Passage”, and ”The Grandfather Passage”.After this feature extraction process is implemented as features like monotonous pitch, reduced loudness, monotonous loudness, etc.Then 3 forms of regularization models are applied which are ridge regression with L2 norm, the L1 norm in lasso regression and hinge loss function in support vector machine and their corresponding Mean Absolute Errors are calculated.Visualizations are shown.

G. *Zoltan Galaz et ał*

In [7] after data collection, the feature extraction phase is initiated and using multiple tools like Neurological Disease Analysis Tool (NDAT) and PRAAT acoustic software analysis tool approximately 715 features were extracted. some of them were HNR,NHR,SNR,GME,EMD,APQ3,APQ5,PPE,PPQ, etc.Now after this, feature selection phase is done in which embedded methods are used out of the 3 classifications. Guided Regularized Random Forest (GRRF) is used and plots for each task were separately done .After that 10 fold, 100 times cross-validation is done. Also the comparison between true UPDRS and predicted UPDRS is done.

H. *Vaiciukynas E et ał*

In [8]The information taken is supported by phonation and text-dependent speech discourse modalities for the accomplishing objective. Here, phonation can be characterized as the articulating of a short sentence or a vowel. The recording is done through AC(acoustic cardoid) and smartphone(SP) microphones. The parting of the modalities into voiced and unvoiced parts is to get extra modalities. The machine learning approach utilized here is random forest(RF) and EER(Equal error rate) is utilized to identify executionn and the expense of probability proportion. Essentia and YAAFE produce less and diverse EER values for AC and SP separately. As shown in the visualization of the AC(Acoustic cardoid) microphone as meta-RF proximity matrix by the t- SNE and OOB detection performance by the DET curves. Hence the mix of all the capabilities declines the EER significantly more for PD screening.

I. *Freddie Astrom et ał*

In [9], to separate voice and background noise from the voice data, the massive parallel neural network is what has led to the success of the prediction of the task. The Backpropagation neural network is a topology used in this approach which depends on the ‘Steepest dependent training algorithm’ that is further modified to the ‘Levenberg-Marquardt Training algorithm’. This algorithm is used to minimize the error rate and has more accuracy compared to that of the Steepest dependent training algorithm. • The parallel type of neural network approach follows two steps which are- Step 1.Parallel feed-forward neural network with Levenberg- Marquardt Training algorithm: Count of the parallel type neural networks should be an odd number and result is predicted by considering majority of output. Step 2.Rule-based system: This framework depends on a democratic choice plan and expands the power of expectation. Visualizations of the AVEC ROC curves and GeMaps ROC curves can be seen respectively.

J. *T.J.Wroge et ał*

In [10], Required voice data which is non-invasive and recorded easily using Smartphones. In this approach, several Machine Learning methods are applied to predict results where mPower dataset is used. Voice data is pre-processed using the PyAudioAnalysis library in python and cleaned using Voicebox’s Voice Activation Detection algorithm method from the Audio-Visual Emotion recognition Challenge (AVEC) were used fro preliminary audio analysis and method of Minimum Redundancy Maximum Relevance (mRMR) are applied to AVEC audio features. GeMaps extracts some lower level features. OpenSMILE is used to extract AVEC GeMaps features, these both sets of features contain Mel cepstrum frequency coefficients(s) (MFCCs) which provides information regarding the sound frequencies. MFCCs offers a mean to detect the effects. A set of many decision tree classifiers were used to classify the dataset including standard decision trees, random forest, gradient boosted decision trees and extra tree classifiers. separations create a classification accuracy over the training set that is applied to the testing set to assess generalization. Random forests use arbitrary mixing of the data to create different subsets of the training data. Extra three classifiers depend on stochastic methods that create shallower but wider decision trees. The algorithm iteratively modifies the previous classification state by creating another classifier for the training set. This process repeats to en masse classifiers to classify the training sets. Another classifier is the Support Vector Machine as SVM, which accurately perform non-linear classification via the kernel trick.

IV. Figures and Tables

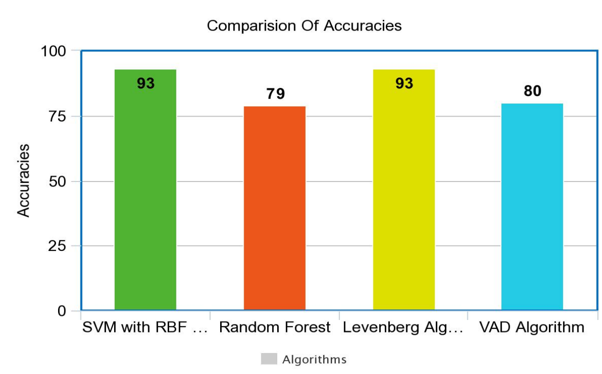

Fig. 1: correlation of accuracies for the papers having creators as R.Viswanathan et ał, EvaldasVaiciukynas et ał, Freddie Astrom et ałand T.J.Wroge et ał

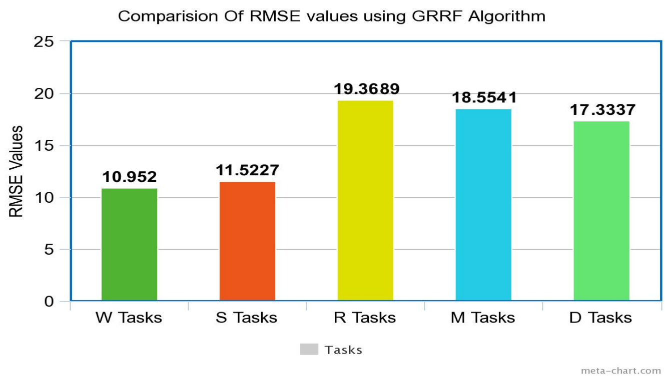

Fig. 2: correlation of RMSE values for the paper written by Zoltan Galaz et al

LIMITATIONS

one of the dominant limitations is the scarce amount of data of patients with PD. Also sometimes there is a possibility of misdiagnosis due to the similarity of symptoms for other neurological disorders. Data usually collected is for the age group between 45-85 years but accuracy reduces if we include data of younger age groups. If the prediction is done through the UPDRS scale then it involves the patients to perform various physical tasks at the clinics under medical supervision sometimes might lead to manual errors. The model that is trained using the data of PWP for a particular region might not work at par for other regions due to changes in accent and other voice-related features etc. Data recorded during the input phase through various recording devices might contain noise and other disturbances which might hamper the accuracy. Under no supervision, if the patients don't perform their assigned tasks properly then it leads to the generation of fallacious data. It also shows the strongly correlated coefficients of the features in the huge data-set and the prevalent correlations may lead to ambiguous predictions.Thus, to overcome this issue we apply, PCA (Principle Component Analysis) and other techniques.

FUTURE SCOPE

The system that we propose should contain features like the data generation process that should be done under proper supervision with minimum noise involvement. It should avoid the patients the overhead of taking their time from their daily routine to visit the clinics and perform physical tasks. The System should be developed in such a way that the final prediction is made by combining the results from all the tasks. For the global outreach of the application, the model should be trained with region wise data. Constant research should be done to improve the performance of the application and for these new algorithms, hybrid algorithms can be used even with huge data-set.

CONCLUSION

The gravity of the literature is to aid researchers to apprehend methodologies for predicting the patients with PD which encompasses all the techniques and numerous approaches conducted by various experts and researchers to

TABLE I: Summary Of Machine Learning Based Methods for Parkinson Disease Prediction

Author Name	***Machine Learning Method***	***Data Description***	***Performance***
Patrick, Walter Karlen	CNN and RNN	accelerometer and gyroscope record, Voice,positions and timestamps of the participant's taps	AUC of 0.85 (95% CI : 0.81, 0.89), an AUPR of 0.87(95% CI: 0.82, 0.91) and a sensitivity at 95% specificity of 43.0 (95% CI: 0.19, 0.54)
R. Viswanathan,P. Khojasteh,B. Aliahmad,S.P. Arjunan,S. Ragnav,P. Kempster,Kitty Wong, Jennifer Nagao and D. K. Kumar	SVM classifier with Radial Basis Function (RBF) kernel	recording of sustained vowel /a/ and sustained consonant /m/	Accuracy rate of 93%
Monica Giuliano,Alfonsa García-López,Silvia Pérez,Fransisco Díaz Pérez,Osvaldo Sposito,Julio Bossero	Principal Components Analysis (PCA),K-means clustering,ANOVA, 2 layer MLP	Voice recording of the phonation of the vowel /a/	AUC=0.826. ACCURACY=0.768 (cut value=0.5).
Alexander M. Goberman, Jessica McMillan	Mann-Whitney U test (alpha level = 0.05) and effect size (h2) analysis	Voice dataset generated through passage reading	MPP : 0.021 to 0.069, OPP:0.006 to 0.035, NPP: 0.000 to 0.142
Siddharth Arora, Vinayak Venkataraman, Sean Donohue, Kevin M. Biglan, Earl R. Dorsey, Max A. Little	random forest classifier,Random Classifier,Conditional Random Classifier	tri-axial accelerometer sensor data	sensitivity and specificity : RF - 98.5% and 97.5%,RC - 50.0% and 50.2%, CRC - 67.7% and 32.6%
Alireza Bayestehtashka, Meysam Asgaria, Izhak Shafrana, James McNames	three forms of regularization, L2-norm in ridge regression, L1-norm in lasso, and hinge loss function in support vector machine	voice dataset is generated by performing 3 tasks - the phonation task,the Diadochokinetic task and the reading task	MAE - Chance 7.5 Lasso 6.9 Ridge 5.9 Linear SVR 5.9
Zoltan Galaz, Zdenek Mzourek, Jiri Mekyska, Zdenek Smekal, Tomas Kiska Irena Rektorovat, Juan Rafael Orozco-Arroyave and Khalid Daoudi	guided regularized random forest (GRRF), RF regression algorithm (10-fold validation with 100 repetition)	Voice dataset	RMSE - W tasks 10.9520, S tasks 11.5227, R tasks 19.3689, M tasks 18.5541, D tasks 17.3337
Evaldas Vaiciukynas, Antanas Verikas, Adas Gelzinis, Marija Bacauskiene	Random forest	Speech recording	79.17%

Freddie Astrom, Rasit Koker	Levenberg Marquardt training algorithm	Voice recordings	75.4% baseline accuracy – 93.5% best accuracy
T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi	VoiceBox's Voice Activation Detection (VAD) algorithm for data cleanning,ANN,SVM, RF, decision tree classifiers	Voice recordings	Accuracy 75% to 86% in various methods

achieve the objective using the UPDRS Scale, inbuilt sensors of the smartphones, using various physical tasks. The literature survey also includes the succinct form of machine learning techniques used, as well as its performance accuracy and data acquisition forms in a concise table. It also comprehends the complications of acquiring the data-set which may be ambiguous or erroneous, to process and solve it to carry out the required task. This paper gives a substance of different strategies just as acquaint with some more techniques, which take novel contemplation and presumptions. The additional motive behind this literature is also to cite the study performed by many and to conduct more research to overcome the challenges and future study to enhance in this eld. We conclude with a scrupulous and comprehensive study on plenty of methods from numerous reputed publications, journals/venues, and subtopics to achieve the aftermath of the methodologies.

REFERENCES

[1] Patrick Schwab,Walter Karlen "PhoneMD:Learning to Diagnose Parkinson's Disease from Smartphone Data" Published in AAAI 2018 DOI:10.1609/aaai.v33i01.33011118.

[2] R. Viswanathan,P. Khojasteh,B. Aliahmad,S.P. Arjunan,S. Ragnav,P. Kempster,Kitty Wong, Jennifer Nagao and D. K. Kumar "Efficiency of Voice Features based on Consonant for Detection of Parkinson's Disease" Published in IEEE Life Sciences Conference… 2018 DOI:10.1109/lsc.2018.8572266.

[3] Monica Giuliano,Alfonsa García-López,Silvia Pérez,Fransisco Díaz Pérez,Osvaldo Spositto,Julio Bossero "Selection of voice parameters for Parkinson s disease prediction from collected mobile data" 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA).

[4] Goberman A.M. and Mc Millan J, "Relative speech Timing in Parkinson's disease," commun Science Discord 2005,32:22-29.

[5] Siddharth Arora, Vinayak Venkataraman, Sean Donohue, Kevin M. Biglan5, Earl R. Dorsey, Max A. Little, "HIGH ACCURACY DISCRIMINATION OF PARKINSON'S DISEASE PARTICIPANTS FROM HEALTHY CONTROLS USING SMARTPHONES" 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

[6] Alireza Bayestehtashka, Meysam Asgaria, Izhak Shafrana,James McNames "Fully automated assessment of the severity of Parkinson's diseasefrom speech" Computer Speech and Language 2015 DOI:10.1016/j.csl.2013.12.001.

[7] Zoltan Galaz, Zdenek Mzourek, Jiri Mekyska, Zdenek Smekal, Tomas Kiska Irena Rektorovat, Juan Rafael Orozco-Arroyave:j:§ and Khalid Daoudi "Degree of Parkinson's Disease Severity Estimation Based on Speech Signal Processing" 2016 39th International Conference on Telecommunications and Signal Processing (TSP).

[8] Vaiciukynas E, Verikas A, Gelzinis A, Bacauskiene M (2017) "Detecting Parkinson's disease from sustained phonation and speech signals".Published on October 5, 2017 PLoS ONE 12(10): e0185613. https://doi.org/10.1371/journal.pone.0185613.

[9] Freddie Astrom, Rasit Koker "A parallel neural network approach to prediction of Parkinson's Disease" Published in Expert Systems with Applications 38 (Sep-2011) 12470–12474.

[10] T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi, "Parkinson's Disease Diagnosis Using Machine Learning and Voice," 2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, 2018, pp. 1-7. doi: 10.1109/SPMB.2018.8615607.

[11] Daiga Heisters, "Parkinson's: symptoms, treatments and research." Published in British journal of nursing 2011, DOI:10.12968/bjon.2011.20.9.548.

[12] M. Hariharan a , Kemal Polat , R. Sindhu, "A new hybrid intelligent system for accurate detection of Parkinson's disease" Published in Computer Methods and programs in biomedicine 10.1016/j.cmpb.2014.01.004.

[13] Athanasios Tsanas , Max A. Little, Member, IEEE, Patrick E. McSharry, Senior Member, IEEE, and Lorraine O. Ramig "Accurate Telemonitoring of Parkinson's Disease Progression by Noninvasive Speech Tests" Published in IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 57, NO. 4, APRIL 2010.

[14] Athanasios Tsanas, Max A. Little, Patrick E. McSharry and Lorraine O. Ramig "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity"17 November 2010https://doi.org/10.1098/rsif.2010.0456.

[15] Dávid Sztahó, Miklós Gábriel Tulics1, Klára Vicsi "Automatic Estimation of Severity of Parkinson's Disease Based on Speech Rhythm Related Features" 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2017) • September 11-14, 2017 • Debrecen, Hungary.