

DETECTING PARKINSONS DISEASE USING MACHINE LEARNING

Literature Survey

Team ID : PNT2022TMID05458

Team Leader : KARTHIKEYAN C (921319205057)

Team member : MATHESH M (921319205071)

Team member : RAJESH KUMAR B (921319205109)

Team member : LOKHU PRASANTH A
(921319205065)

Several strategies are recorded for early detection of PD based on the different ML techniques. But accuracy in detection and classifying within the time is very important or else, it causes development of more symptoms. There are different kinds of data, brain MRI images, Voice data, posture images, sensor captured data, handwritten data, using which we can predict whether person is having PD or not. Out of all those, speech or voice data helps in identifying PD accurately.

Eduardo Tolosa et al proposed a twofold fully automatic approach with 3D images has shown promising results in their experimentation.

Max A. Little et al presented a new dysphonia measure, pitch period entropy (PPE) and used a kernel support vector machine and has achieved classification accuracy of 91%.

RAINER SCHOE NWEILER et al identified a different approach which used voice analysis with ANN and got good results but observed that cost-effectiveness remains to be a challenge.

Marius Ene et al suggested NN based approach with three types of internal methods and discriminated persons having PD with healthy persons.

Ipsita Bhattacharya et al identified the ROC curve variation and identified that values of TP and FP rates show changes while increase in the CV folds.

Freddie Åström et al proposed unique approach of parallel neural networks and then outcome of each neural network is assessed by using a rule-based system for the decision. During the training process, data that is not yet learned of each neural network is collected and applied in

the training set of the later neural network. This helped to increase prediction accuracy.

Athanasios Tsanas et al developed a novel algorithm based on speech signals but it's questionable as most of the features are not considered here, only 10 features are used.

Hui-Ling Chen et al proposed FKNN centered system using a 10-fold cross validation method.

Mohammad S Islam et al has compared various ML techniques based on their performance accuracies in determining whether person is having PD or not and mentioned that new classifier may be built to get better accuracies.

Bo Peng et al suggested Computer Aided Analysis with image data and used BrainLab software for processing the images and calculate thickness of the cortex, volume of gray matter, and surface area of the cortex on each region of interest (ROI). Use of Multilevel ROI-based features improved the classification performance.

Derya Avci and Akif Dogantekin proposed another approach using Genetic Algorithm-Wavelet Kernel-Extreme Learning and achieved good accuracy results.

R Prashanth identified that multimodal features can be used to predict PD in earlier stage.

Satyabrata Aich proposed a unique approach by using Genetic algorithm and PCA as feature selection methods and applied seven ML algorithms for classification, that saved time and productivity while doing pattern classification with two categories such as PD and not PD.

Leandro A. Passos compared ResNet-50, Optimum-Path Forest (OPF) classifier with Support Vector Machines (SVM) and Bayes and achieved 96% of identification rate.

Deepak Gupta followed a different approach cuttlefish algorithm and used for feature selection,

different fitness functions approximations are used to improve cuttlefish algorithm and is termed as Optimized cuttlefish algorithm (OCFA). Decision tree and K-Nearest Neighbor classifiers are applied and achieved 94% of accuracy in detecting PD effected patients .

(i) Multiple Feature Evaluation Approach (MFEA) of a multi-agent system (ii) Implementation of five classification schemas which are Decision Tree, Random Forests, Neural Network, Naïve Bayes and Support Vector Machine on the Parkinson's diagnosis before and after applying their approach, and (iii) Author approach witnessed the following average rate of accuracies : Decision Tree achieved accuracy of 10.51%, Naïve Bayes shown 15.22%, Neural Network is found with 9.19%, Random Forests and SVM performed with 12.75% and 9.13% respectively.

S.No	Author Name	Year	Methodology	Input data	Performances
1	Ali H. Al-Fatlawi et al	2016	Deep belief network, Restricted Boltzmann Machines ,Back propagation algorithm	Voice data	Acc: 94%
2	Marius Ene et al	2008	Probabilistic neural network (PNN)	Speech samples	Accuracies ranging between 79% and 81%
3	David Gil A, Magnus Johnson B	2009	ANNs and SVMs	Speech	90%
4	Chien-Wen Cho et al	2009	Principal component analysis with linear discriminant analysis.	Voice samples	95.49%
5	Max A. Little et al	2009	SVM	Voice recordings	classification performance of 91.4%
6	Resul Das et al	2010	Neural Networks, DMneural, Decision Tree and Regression	Speech	Score of 92.9% is achieved
7	C. Okan Sakar & Olcay Kursun	2010	SVM	Speech data	classification accuracy:92.75%
8	Zachary C.Lipton et al	2016	Long Short-Term Memory (LSTM-RNN) with forget gate, MLP	Voice data	Several accuracies are compared.
9	Ipsita Bhattacharya et al	2010	Used LibSVM for classifying along with random split of the dataset, and determine accuracy for the different kernel functions	speech	Improved average accuracy achieved.
10	Freddie Åström et al	2011	Used a different neural network to minimize the probability of outcome with error	Voice data	Total nine parallel neural networks are arranged and achieved development of 8.4%

					for the prediction of PD compared to single network
11	Athanasios Tsanas et al	2012	Speech signal processing algorithms, RF,SVM	Voice signals	99%
12	Indrajit Mandal et al	2017	Multinomial logistic regression, rotation forest together with SVM and PCA, ANN, boosting methods	Speech	100% accuracy achieved with sparse multinomial logistic regression and linear logistic regression, observed sensitivity:0.983 and specificity: 0.996
13	Hui-Ling Chen et al	2013	FKNN,SVM	Speech	96.07% obtained by the FKNN dependent system using a 10-fold CV method
14	Tarigoppula V.S Sriram et al	2013	SVM,KNN,NB,RF	Voice data	Random Forest shown better accuracy
15	Mohammad S Islam et al,2014	2014	SVM, Random Tree and Feedforward Back-propagation built Artificial Neural Network.	Speech	90% recognition accuracy
16	Oana Geman et al	2015	SVM,DNN	Voice data	90% accuracy achieved
17	Bo Penga et al	2015	t-test, SVM, and Minimum Redundancy and Maximum Relevance.	Speech impairment data	Proposed method used multilevel ROI-based features and is observed better classification accuracy..
18	Othman Ibrahim , Mehrbakhsh Nilashi, & Ali Ahani	2016	PCA is used for feature selection, EM, ANFIS and Support Vector Regression (SVR).	Voice data	SVM:AUC-0.9623 ANFIS:AUC-0.848
19	Hui-Ling Chen et al	2016	Extreme learning machine and kernel ELM	Speech samples	10- fold cross validation through 10 runs achieved 96.47% accuracy
20	Derya Avci and Akif Dogantekin et al	2016	Genetic Algorithm, wavelet kernel and Extreme Learning Machines(ELM).	Voice data	96.81%.
21	Thomas J. Hirschauer	2015	EPNN (Enhanced Probabilistic Neural Network	Speech	98.6 %
22	Lígia Sousa et al	2019	DNN, KNN,PCA (for optimizing feature set)	Voice samples	93.4% for the binary classification,84.7% for multiclass classification.
23	Leandro A. Passos	2018	ResNet-50 , Optimum-Path Forest (OPF) classifier	HandPD dataset, speech	96% of identification rate using speech samples.
24	Deepak Gupta	2018	Optimized cuttlefish algorithm ,Decision tree, KNN	Speech data and Handwritten data are used to evaluate the proposed model.	94%
25	Shreya Bhat	2018	Along with advanced machine learning methods, Neuroimaging modalities also used	Image data, speech, ,MRI, EEG	(Various implementations are discussed)
26	Hariharan et al	2014	Gaussian mixture with PCA and LDA. SVM classifier	Speech data	100%
27	Zhang et al	2017	Stacked autoencoders, KNN	Speech	In the range of 94-98%
28	Oung et al	2018	Classifiers used are KNN, PNN, ELM classifiers.	Motion and Speech	KNN:93.26% PNN: 95.22% ELM: 95.93%

29	Hlavnicka et al	2017	Zero-crossing rate, variance of autocorrelation function.	Speech	Accuracy: 71.30% Sensitivity: 56.70% Specificity: 80%
30	Salama A. Mostafa	2018	Decision Tree, Random Forests, Naïve Bayes, Support Vector Machine and Neural Network.	Voice data	Avg rate of improved accuracies achieved are: Decision Tree: 10.51%, Random Forests: 12.75% Naïve Bayes:15.22%, Support Vector Machine: 9.13%, Neural Network: 9.19%
31	Rainer schonweiler et al	2000	Artificial neural networks, Regression tree	Voice data	Various combinations of methods applied and achieves improved accuracies.