

# An ensemble feature selection framework for early detection of Parkinson's disease based on feature correlation analysis

Sarfraz Masood<sup>1</sup>, Khwaja Wisal<sup>1</sup>, Om Pal<sup>2</sup>, and Chanchal Kumar<sup>1</sup>

<sup>1</sup>Jamia Millia Islamia

<sup>2</sup>Ministry of Electronics and Information Technology

March 9, 2021

## Abstract

Parkinson's disease (PD) is a highly common neurological disease affecting a large population worldwide. Several studies revealed that the degradation of voice is one of its initial symptoms, which is also known as dysarthria. In this work, we attempt to explore and harness the correlation between various features in the voice samples observed in PD subjects. To do so, a novel two-level ensemble-based feature selection method has been proposed, whose results were combined with an MLP based classifier using K-fold cross-validation as the re-sampling strategy. Three separate benchmark datasets of voice samples were used for the experimentation work. Results strongly suggest that the proposed feature selection framework helps in identifying an optimal set of features which further helps in highly accurate identification of PD patients using a Multi-Layer Perceptron from their voice samples. The proposed model achieves an overall accuracy of 98.3%, 95.1% and 100% on the three selected datasets respectively. These results are significantly better than those achieved by a non-feature selection based option, and even the recently proposed chi-square based feature selection option.

## Hosted file

A\_new\_EBFMSM\_for\_early\_detection\_of\_PD\_based\_on\_feature\_correlation\_analysis\_Updated.pdf  
available at <https://authorea.com/users/400329/articles/512663-an-ensemble-feature-selection-framework-for-early-detection-of-parkinson-s-disease-based-on-feature-correlation-analysis>