Development of Machine Learning Algorithms to Diagnose Parkinson's Disease

Shenoy, Jay

Speech impairment has long been linked to Parkinson's disease. Individuals with Parkinson's are often unable to produce sustained phonations and suffer from vocal tremor, hoarseness, and breathiness. In this paper, two logistic regression classifiers and an artificial neural network (ANN) were trained to detect Parkinson's disease using a de-identified data set from the University of Oxford containing 195 observations. 10-fold cross validation was performed to produce classification accuracies that were representative of their real-world values. The ANN had the best performance overall, with an accuracy of 94.21% and a sensitivity of 95.86%. In addition, the speech signals that had the greatest influence on the diagnosis, as determined from the standardized coefficients of the logistic regression models, were those pertaining to detrended fluctuation analysis (DFA) and pitch period entropy (PPE). Further research must be performed to improve the accuracies of these models before they can be introduced into clinical settings.