Abstract

This paper intends to provide an evaluation of the different pre-processing techniques that can aid a classifier in the classification of the Parkinson Disease (PD) dataset. PD is a chronic and progressive moment disorder caused due to the malfunction and death of vital nerve cells in the brain. The key indications of the chronic malady in the central nervous system can be best captivated from the Mentation, Activities of Daily Life (ADL), Motor Examination, and Complications of Therapy. The speech symptom which is an ADL is a common ground for the progress of the PD. A comprehensive study on the application of different pre-processing techniques is carried out on the PD dataset obtained from the UCI website. For classifying the PD dataset we employed the ANN based MLP classifier. With the objective of improving the prediction accuracies of the healthy, and people with Parkinson disease on the PD dataset this study highlights the fact that the combination of several pre-processing techniques namely Discretization, Resampling, and SMOTE can best aid in the classification process. This study is unique in the sense that we have not come across any similar studies in the Data Mining literature.
References


Index Terms

Computer Science  Artificial Intelligence

Keywords

ANN, MLP, Discretization, Resampling, SMOTE, Classification