IT-302

Machine Learning and Deep Learning Approaches for Brain Disease Diagnosis

Research Project

LAKSHAY KAMRA(2K19/MC/063) & LAKSHITA SINGH(2K19/MC/065)

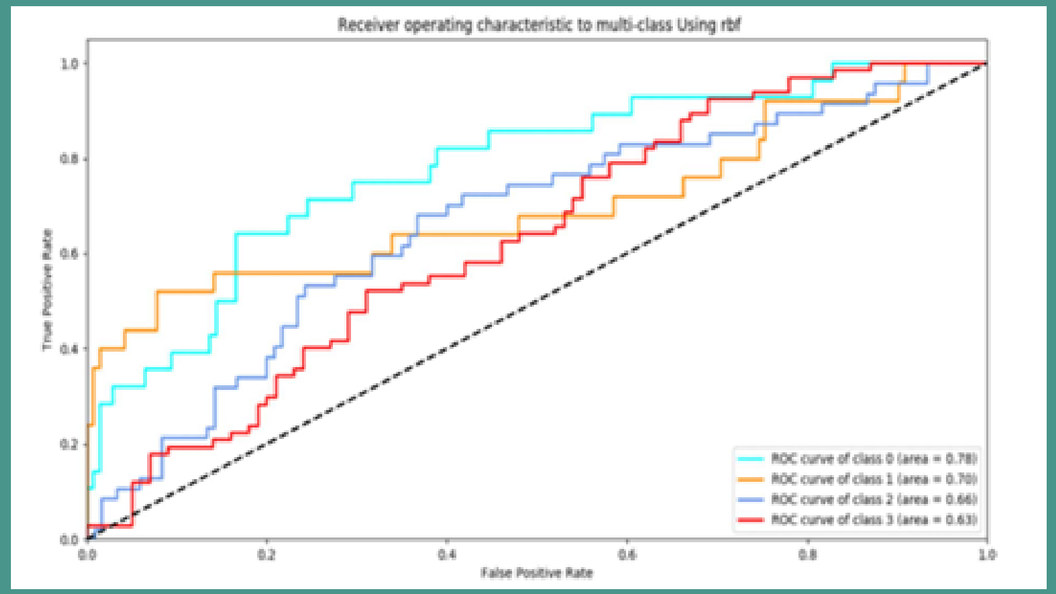
INTRODUCTION

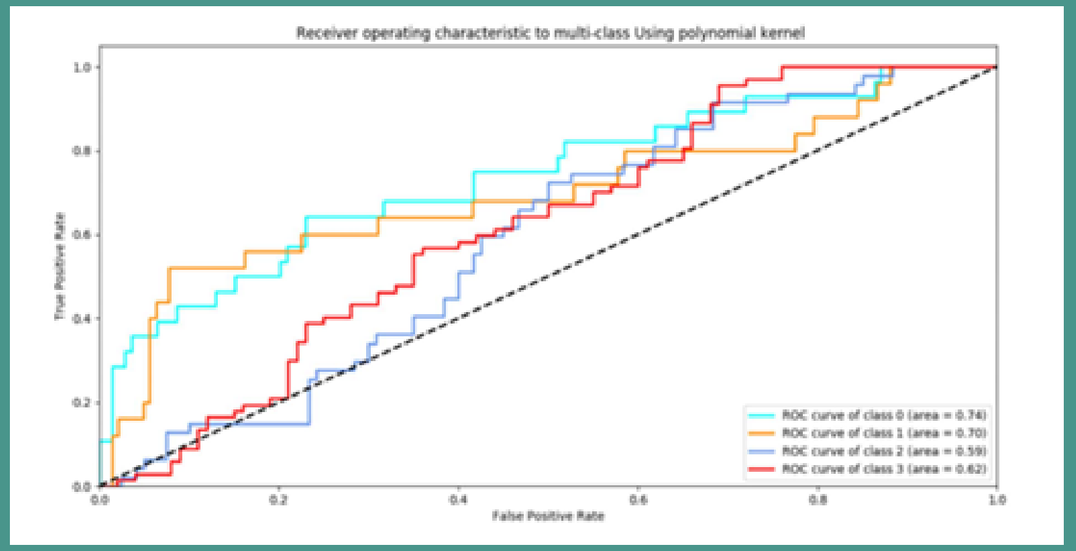
Machine learning and deep learning approaches are being successfully applied to different fields of life sciences for the past years. Medicine is one of the most suitable application domains for these techniques since they help model diagnostic information based on causal and/or statistical data and therefore reveal hidden dependencies between symptoms and illnesses. In this paper we give a detailed overview of the recent machine learning research and its applications for predicting cognitive diseases, especially the Alzheimer's disease. The study aims to seek out risk-disease genes using machine learning approaches. Various algorithms like Support Vector Machines (SVM) and Decision Trees have been used along to find a suitable hypothesis for classification of candidate genes as Alzheimer's Disease (AD) associated or unassociated.

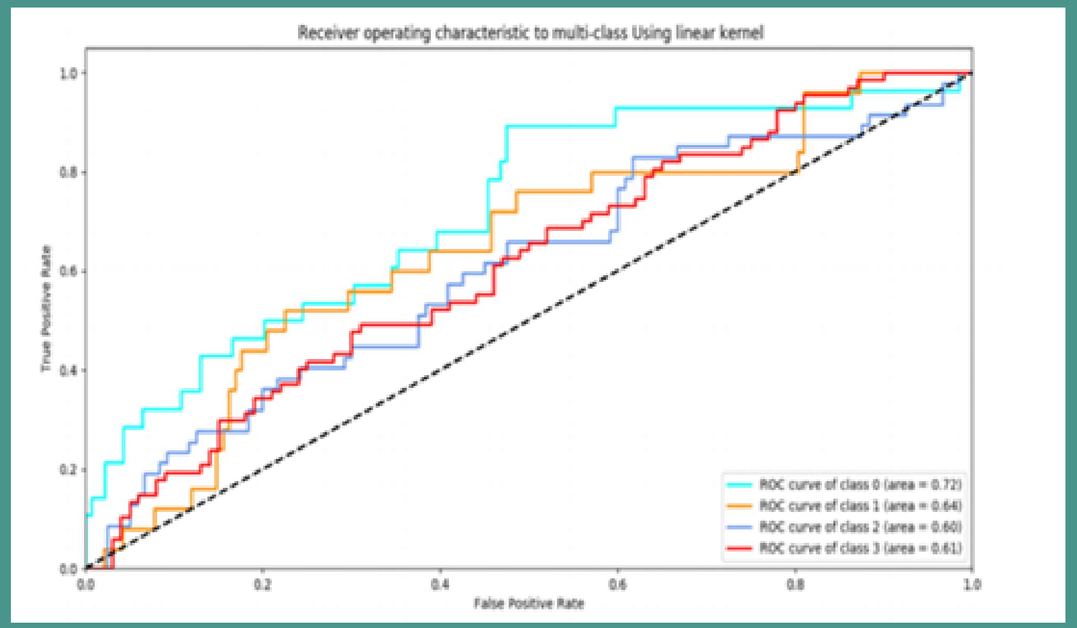
IMPLEMENTATION

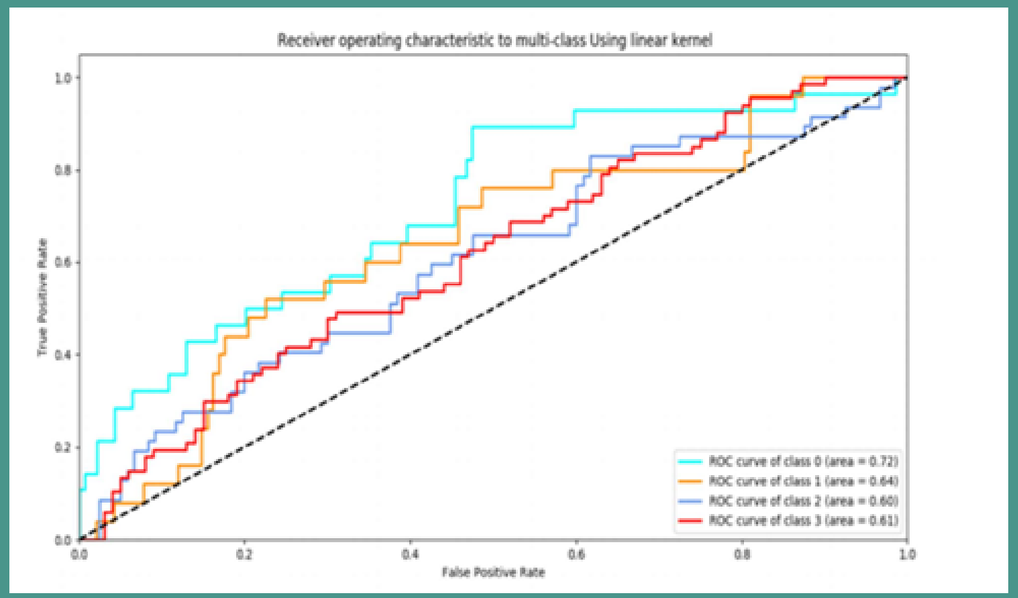
The dataset was evaluated using different kernels as well as other parameters. For instance, while applying the SVM method, Gaussian (Radial), Linear, Polynomial kernels were used to evaluate the best possible case. Cross-validation has been used to obtain optimal results. The results for 2-cross, 5- cross, 10-cross-validation have been displayed below. The Receiver Operating Characteristic (ROC) curve has also been generated for better visualization. Parameter C(=1/λ) has also been adjusted to avoid overfitting.

ROC CURVES


Receiver operating characteristic to multi-class Using rbf
True Positive Rate
False Positive Rate
ROC curve of class 0 (area = 0.78)
ROC curve of class 1 (area = 0.70)
ROC curve of class 2 (area = 0.66)
ROC curve of class 3 (area = 0.63)


Receiver operating characteristic to multi-class Using polynomial kernel
True Positive Rate
False Positive Rate
ROC curve of class 0 (area = 0.74)
ROC curve of class 1 (area = 0.70)
ROC curve of class 2 (area = 0.59)
ROC curve of class 3 (area = 0.62)


Receiver operating characteristic to multi-class Using linear kernel
True Positive Rate
False Positive Rate
ROC curve of class 0 (area = 0.72)
ROC curve of class 1 (area = 0.64)
ROC curve of class 2 (area = 0.60)
ROC curve of class 3 (area = 0.61)


Receiver operating characteristic to multi-class Using linear kernel
True Positive Rate
False Positive Rate
ROC curve of class 0 (area = 0.72)
ROC curve of class 1 (area = 0.64)
ROC curve of class 2 (area = 0.60)
ROC curve of class 3 (area = 0.61)

Method used	Accuracy
SVM Library - R	84.56%
Radial Kernel (C=1)	82%
Radial Kernel (C=2)	86%
Linear Kernel	80%
Polynomial Kernel	65%
Sigmoid Kernel	57%
Decision Tree	88.29%

The algorithms were evaluated on micro average, macro average, and weighted average of their accuracy, precision, F-1 score and support results on the four predicted classes as shown