



Scales-based classification of schizophrenics and controls

Ling Yun Yeow¹; Geetha Soujanya Chilla¹; Qian Hui Chew²; Kang Sim²; Bhanu Prakash KN¹

¹Institute of Bioengineering and Bioimaging, A*STAR, Singapore, 138667

²Institute of Mental Health, Singapore, 539747



Introduction

Schizophrenia is a major psychotic illness which causes one to lose touch with reality. There is a range of assessment tools which can be used to assess aspects such as functioning, cognitive capabilities and quality of life of people. These tests include Wide Range Achievement Test (WRAT3), Brief Assessment of Cognition in Schizophrenia (BACS) and brief form of World Health Organization Quality of Life (QOL). The use of features from these 3 tests would be the focus of this study.

Initial exploratory data analysis using t-Distributed Stochastic Neighbor Embedding (t-SNE) showed some degree of separation between healthy controls (HC) and patients with schizophrenia (SZ) when all the features of the 3 tests were used (Fig 1). This prompted further investigation by utilizing machine learning (ML) methods to identify the more important psychometric features followed by classification between SZ and HC. From the performance of the classifiers, we aim to determine how well these psychometric scores can be used to delineate SZ from HC.

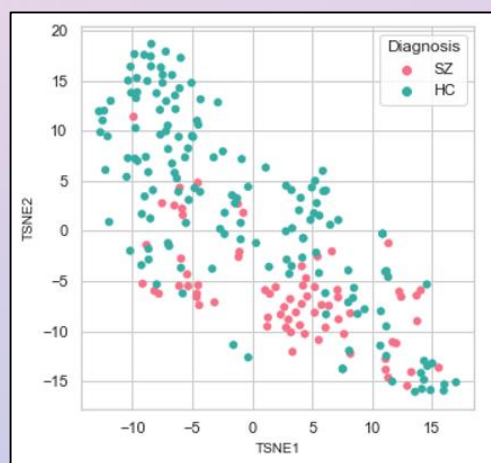


Fig. 1. 2D t-SNE plot of all between SZ and HC.

Methodology

Data: 72 SZ & 150 HC with the corresponding WRAT3, QOL and BACS scores

Data pre-processing: QOL scores were standardized to 0-100 range and grouped into four domains (physical health, psychological, social relationships, and environment) and two standalone items (overall QOL and general health). Composite BACS scores were generated and standardized based on age and gender using seven BACS items. Finally, we have a total of eight features and robust scaling was performed before feature selection and model instantiation.

Feature selection: Embedded (random forest (RF) and gradient boosted (GB) with tree importance) methods were initially used to identify important features and output showed overall QOL and general health features of QOL to be of lesser importance compared to other features. Following this, wrapper-based methods using Recursive feature elimination (RFE) with different base estimators including support vector classification (SVC), decision tree (DT), RF, logistic regression (LR), GB and perceptron were tested. Output features were consistent with that of the embedded methods. In the end, WRAT3, 4 domains of QOL and composite BACS scores were used as features for classification.

Methodology (Cont'd)

Data processing and evaluation: Data was stratified into train-test with ratio of 70:30. Grid-search with stratified 5 folds was performed for each classifier – logistic regression (LR), K nearest neighbors (KNN), linear SVC, Decision Tree (DT), Random Forest (RF), adaptive boosting, and Gradient Boosting (GB) – and performance was evaluated using AUC, accuracy, sensitivity, specificity and weighted F1 score.

Results

Using the six features for classification, mean training accuracies ranged from 79-85% whereas testing accuracies, AUC, sensitivity, specificity and F1 scores were from 76-90%, 80-91%, 69-89%, 86-100% and 77-90% respectively. Fig 2 shows the training and testing accuracy. Fig 3 shows the ROC curves for all the classifiers. Overall, even though AUC for RF is highest at 91% and specificity was 100%, the sensitivity was lower at 82% as compared to other classifiers such as KNN and GB, which showed an AUC of 89%, sensitivity of 89% and specificity of 91%.

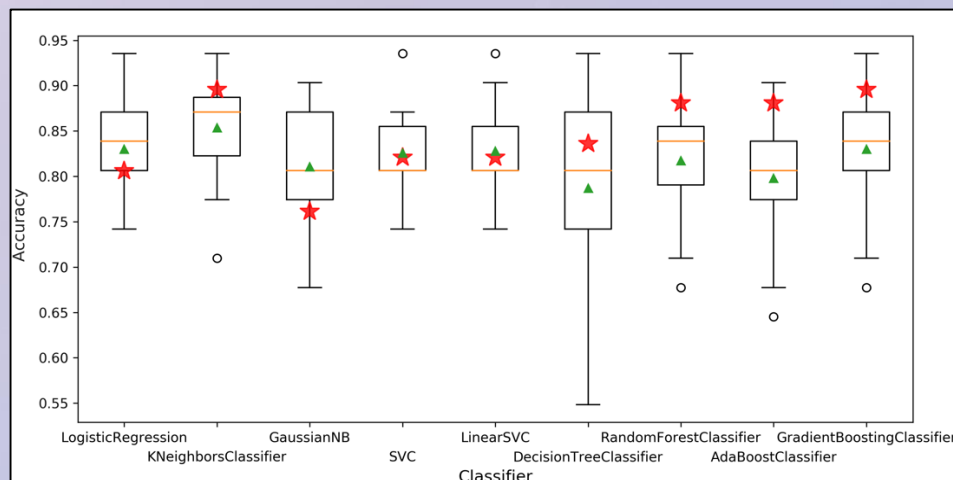


Fig. 2. Training and testing accuracies across the classifiers. Boxplot shows the training accuracies obtained from 3 5-fold repeats of each classifier.

▲ represents the mean of the training accuracy and ★ represents the testing accuracy of each classifier.

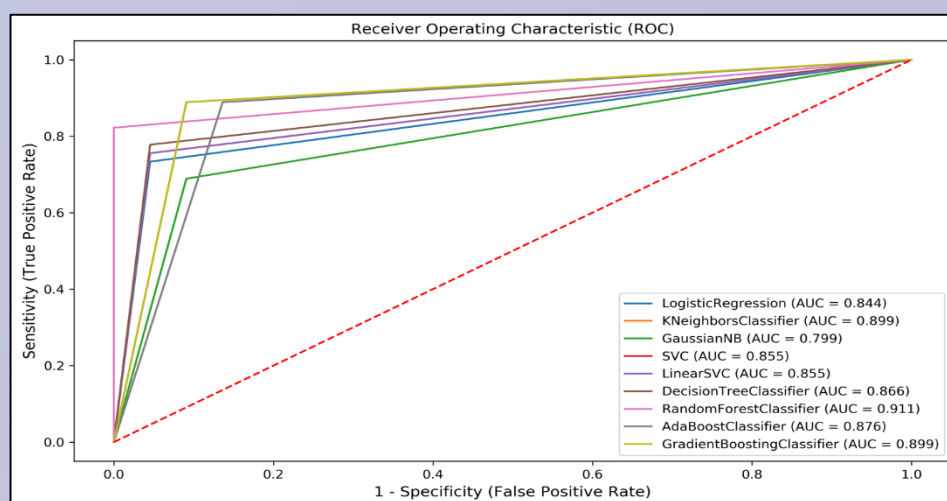


Fig. 3. ROC curves of all the classifiers, with AUC values indicated.

Conclusion

We demonstrated ML-based classification of SZ and HC using reading ability (WRAT3), four domains of QOL and cognitive functioning (BACS). Overall, both KNN and GB with tuned hyper-parameters had the best classification performance. This study serves as a proof-of-concept support for further use of functional and cognitive measures in differentiating between subtypes of the illness.