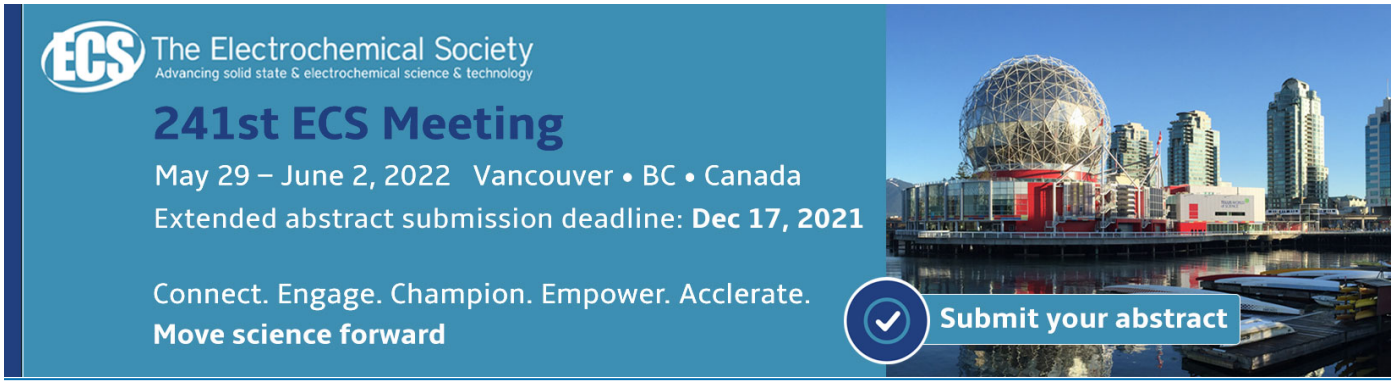
Enhanced of Autism Spectrum Disorder Using Grey Relational Analysis and Supervised Learning for Classification



IOP Conf. Series: Materials Science and Engineering **864** (2020) 012083 doi:10.1088/1757-899X/864/1/012083

Enhanced of Autism Spectrum Disorder Using Grey Relational Analysis and Supervised Learning for Classification

Zuriahati Mohd Yunos[1], Wafa Idrus[1], Siti Mariyam Shamsuddin[1], Mohd Saiful Izwaan Saadon[2], Suhaila M Yusuf[1]

[1]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia.
[2]Faculty of Maritime Studies, Universiti Malaysia Terengganu, Malaysia

E-mail: zuriahati@utm.my

Abstract. Autism spectrum disorder (ASD) is complicated to be diagnosed and many study had shown that machine learning technique have been proven to accurately diagnose ASD. However, there were also some drawbacks in the results obtained and one of it is related to the lower accuracy upon implementation. Thus, a feature selection method namely Grey Relational Analysis (GRA) is proposed to enhance the performance of the machine learning technique in the classification of ASD as it was proven too often produce high accuracy results. GRA is used to select relevant features and rank them from the highest to the lowest. The data used was the adult autism data, which consists of 608 data with 16 features. The machine learning techniques used are Support Vector Machine (SVM) and Artificial Neural Network with Multi Layer Perceptron (ANN-MLP) to classify ASD. From results obtained, the integration of GRA and machine learning techniques have managed to produce a high accuracy of more than 90%. The SVM gave the good accuracy of 98.1%, while ANN produce of 98.36%.

1. Introduction

Autism Spectrum Disorder (ASD) is a neuro-developmental disorder which is characterized by a persistent deficit in social interaction and communication, as well as a restricted, repetitive patterns of behaviour, interests or activities [8].

Another cause of ASD is contributed by several factors [7]. These factors were the outcome of the genetic and environmental risk factors as well as their interaction [7]. Currently, ASD diagnosis is based on clinical assessments such as the observations on the individual's behaviour and intellectual abilities [4]. This includes certain factor such as comorbidity and time consuming [6]. In addition, the conventional statistical techniques existed when computing was not an option [8]. Thus, there is a lack in using the conventional statistical techniques when it is applied in classification [8].

The results produce a less reliable outcome, when the technique using a small samples and produce incurable assumptions that related to its distribution [9]. Thus, the application of machine learning for the early detection of ASD may be useful for an early intervention, other than it being useful to the physicians to diagnose the patients with similar disorder [3]. Other than that, by utilising machine learning may assist in improving the diagnostic and behavioural sciences such as ASD [2].

2. Methodolgy

2.1. Feature Selection Using Grey Relational Analysis

Grey Relational Analysis (GRA) is a feature selection and consists of three main steps, which are; the data pre-processing, the calculation of Grey Relational Coefficient (GRC), and the calculation of Grey Relational Grade (GRG). These steps are crucial to rank each of the features in order to determine their significance. There are three formula used in GRA, which are equation (1) shows the formula for data normalization, equation (2) shows the formula used to calculate GRC, and equation (3) shows the formula used to calculate GRG.

$$X_i^*(k) = \frac{x_i^{(o)} - minx_i^{(o)}(k)}{maxx_i^{(o)}(k) - minx_i^{(o)}(k)} \tag{1}$$

Where,

$x_0(k)$ is the referential series for attribute k

$x_i(k)$ is the comparable series for attribute k

$$\gamma(x_0(k), x_i(k)) = \frac{min_{i,k} |x_0(k) - x_i(k)| + \zeta \, .max_{i,k} \, | x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta . max_{i,k} \, | x_0(k) - x_i(k)|} \tag{2}$$

Where,

$x_0(k)$ is the referential series for attribute k

$x_i(k)$ is the comparable series for attribute k

ζ is the identification or distinguishing coefficient = 0.5

$$\alpha_i = \frac{1}{N} \sum_{k=1}^{N} \gamma_i(k) \tag{3}$$

2.2. Support Vector Machine

Support Vector Machine (SVM) is related to classification problems and there have been plenty of accomplishment in a variety of applications [8]. The purpose of using SVM is to solve the kernel function and problem in quadratic programming. The first step to be done is the parameter tuning which the parameter values will significantly affect the results of the classification [8]. In this study, the suitable parameters were determined from a set of fixed values by using the GridSearchCV function in Python 3. For the classification of ASD, the classes are divided into non-autistic labelled as 0 and autistic labelled as 1. The type of data used in this study is a non-linearly separable data.

2.3. Artificial Neurall Network

This study emphasized on the feed-forward neural network with Multi Layer Perceptron (MLP) and the input passes only in the forward direction. For the implementation of ANN-MLP in Python, the dataset is split into training and testing set by using the train-test-split method. In this study, the data was trained using the *GridSearchCV* function in Python. Testing set is a set of data used only to test the performance of a fully-specified classifier. The method used to search the hidden nodes is via trial and error method. Various number of iterations are used via trial and error method until an optimal condition is achieved or until the agent stopped its attempt.

There are two approaches for this method, which are the forward and backward approach. The forward approach was chosen and for starter, a small number of hidden neurons were selected. Then, the neural network was trained to get the best parameters using the *GridSearchCV* function in Python and then tested using *mlpClassifier* function. The number of hidden nodes was added an iteration after another. This process is iterated until the best result is obtained.

3. Results and Discussions

3.1. Feature Ranking

Table 1 shows the selected features according to their GRG values and only features with GRG values larger than 0.7 were selected to be trained and tested on the classifier models. Then, the features is fed into the classifier as input during the model implementation.

In SVM, the most important step is the parameter tuning. The purpose is to find the best combinations of these parameters for each number of features with the highest accuracy score by using trial and error method. Meanwhile, the same process is repeated in ANN-MLP to determine the best parameter of momentum and learning rate values. Table 2 and 3 shows the best parameter for SVM based on accuracy and the best parameters for ANN-MLP.

Table 1. Selected features with their respective GRG values

Notation	Feature Name	GRG	Rank
F9	Answer code based on screening method	0.89	1
F6	Answer code based on screening method	0.89	1
F5	Answer code based on screening method	0.83	2
F4	Answer code based on screening method	0.80	3
F3	Answer code based on screening method	0.80	3
F7	Answer code based on screening method	0.79	4
F2	Answer code based on screening method	0.78	5
F16	Relation	0.76	6
F10	Answer code based on screening method	0.75	7
F12	Gender	0.70	8

Table 2. The best parameter for SVM

C value	Gamma value	Kernel function
1.0	0.8	Radial basis function (RBF)

Table 3. The best parameters for ANN-MLP

Parameter	Values
Activation function	*logistic*
Alpha value	0.1
Number of maximum iterations	80
Learning rate	*adaptive*
Momentum	0.9 (default)
Solver	*lbfgs*
Hidden nodes	16

3.2. Classification Results

In order to make a comparison between the performances of the classifier models, measurement metrics such as accuracy, recall, precision and AUC of the models were also observed. The general results were obtained from each iteration of the respective classifiers based on the number of selected features. The detailed results is tabulated in Table 4 and is summarized in figure 1.

- Accuracy : The ability to differentiate between nonautistic and autistic individuals. It calculates the proportion of true positive and true negative in all evaluated cases [1].
- Recall : Also known as TPR or sensitivity which measures a correctly classified autistic individuals

- Precision : Precision is the number of True Positive (TP) over the number of TP and the number of False Positive (FP), also called Positive Predictive Value (PPV) [5].
- AUC : A measure of separability of different classes. The higher the value of AUC, the better the model performance is at separating classes or in this case, the better the model at distinguishing between non-autistic and autistic adults. As the value is closer to 1, it shows that the model is a good model. However, if the AUC value is equal to 0.5, it shows that the model has no ability to separate the classes at all.

Based on the results in table 4, it can be concluded that both SVM and ANN-MLP achieved high accuracy in the classification of ASD with more than 90% of accuracy. Table 2 show that SVM outperformed ANN-MLP with accuracy of 98.91%, and AUC of 0.99 (highlighted in bold). It indicate that an AUC value closer to 1 is said to be a good model. Thus, SVM is much better at classifying autism classes compared to ANN as it was found that SVM provided more advantages during classification. However, both classifier performances also show a good classification results as both produced a result of more than 90% in all the performance measurement used in this study. Furthermore, there were only a slight difference in both classifiers performance. This indicates that SVM and ANN-MLP are relevant supervised learning in binary classification.

Table 4. Comparison of performance measurements

Classifier	Accuracy	Recall	Precision	AUC
SVM	**98.91%**	98.39%	98.39%	**0.99**
ANN-MLP	98.36%	98.39%	96.83%	0.98

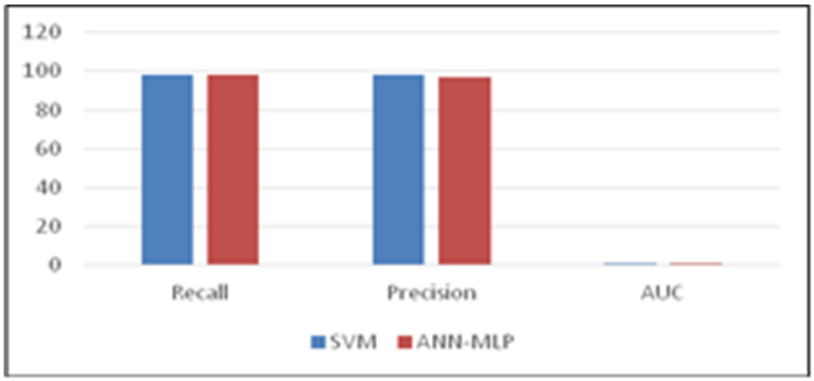


Figure 1. Summary of the overall detailed classifier performances

4. Conclusions

In conclusion, GRA feature ranking was more precise by giving results of the most and the least significant features. Besides that, GRA is able to improve the performance of SVM and ANN classifiers. In addition, SVM is a good classifier with an accuracy of 98.91% and AUC of 0.99 closer to 1. Lastly, SVM and ANN-MLP are relevant supervised learning in binary classification.

Acknowledgements

This work was financially supported by a UTMER grant, cost center number 17J47. The authors would like to thank anonymous reviewer for providing constructive and generous feedback. The authors also would like to thank Research Management Centre (RMC) of Universiti Teknologi Malaysia, for the

research activities and Applied Industrial Analytics (Ali@s) for the support and motivation in making this research a success.

References

[1] Baratloo A, Safari S, Elfil M and Negida A 2015 *Emergency* **3** p 170.
[2] Bone D, Goodwin M S, Black M P, Lee C, Narayanan S, Angeles L and Fu K. 2016 *J. of Autism and Developmental Disorders* **45** pp 1121–36.
[3] Bekerom B V D 2012 *Proceeding of 20th Student Conference in IT*.
[4] Close H A, Lee L-C, Kaufmann C N and Zimmerman A W 2012 *Pediatrics* **129** pp 305-16.
[5] Guyon, Isabelle and Elisseeff A 2003 *J. on Machine Learning Research Special Issue on Variable and Feature Selection* **3** pp 1157-82.
[6] Katuwal G J 2017 PhD Thesis. Rochester Institute of Technology .
[7] Park H R, Lee J M, Moon H E, Lee D S, Kim B N, Kim J and Paek S H 2016 *Experimental Neurobiology* **25** p 1.
[8] Hassibi K 2016 Machine Learning vs. Traditional Statistics: Different philosophies, Different *Data Science Central*.