PERFORMANCE ANALYSIS OF ELMAN AND MODIFIED ELMAN NEURAL NETWORK TO PREDICT ASD SUB GROUB FROM GNOME DATA SEQUENCE

S. Padmapriya[1], S. Murugan[2]

1. Assistant Professor, Department of Computer Science, SRM Trichy Arts and Science College, Trichy.
2. Associate Professor, Department of Computer Science, Nehru Memorial College, Puthanampatti.

[1]callpadmapriya@gmail.com

ABSTRACT

Recent research work shows the connection between gene patterns and Autism Spectrum Disorder (ASD). A particular pattern of gene sequence shows the possibilities of having a specific type of ASD. Classification plays a significant role within the data processing approach and it aims to create a classifier model to explain and establish the necessar y knowledge categories and provides a much better understanding of data at large. In this paper, a completely unique intelligent classifier model is intended using ELMAN neural network. In this research work a gene based appearance in Autism Spectrum Disorder is analysed with the goal of selecting the most qualified genes and performing classification. The objective was achieved by introducing a new layer in Elman Neural Network. The result shows the most discriminative identification of desired gene sequence data. The gene subset was selected by modified Elman neural network.

Keywords: Autism spectrum Disorder, Classification, Elman Neural Network.

I.INTRODUCTION:

Data mining is the term used to describe the process of extracting knowledge from a database [5]. It involves the use of powerful data analysis tools to find previously unknown, valid patterns and relationships in large data sets. The tools include sophisticated, mathematical algorithms and deep learning methods consequently, data mining consists of more than collecting and managing data, it also includes prediction and analysis. A number of innovations in technology and business processes have devoted to a growing interest in data mining in both the public and private sectors. Some of these changes include the development of computer networks, which can be used to connect databases, enhanced search-related techniques such as neural networking and algorithms, the spread of the distributed computing

model, allowing users to access centralized data resources from the desktop, and an increased ability to combine data from various sources into a single search source.

II. NEUTRAL NETWORK:

An artificial neural network (ANN), is simply known as “neural network" (NN), is a mathematical model or computational model based on biological neural networks. Neural network (NN) [1] is a nonlinear and adaptive information processing system based on the intelligent computation of the computer network simulating biological neural network, which processes and memorizes information by simulating cranial nerve and consists of large interconnect processing unit. [2]

It consists of an interconnected group of artificial neurons and processes information using a connectionist approach for computation. In most cases an ANN is a dynamic system that changes its structure based on external or internal information which flows through the network during the learning phase.

Neural network represents a brain metaphor for information processing. These models are biologically inspired rather than an exact model of how the brain actually functions. Neural networks have been shown to be very favourable systems in many forecasting applications and business classification applications neural network represents a brain metaphor for information processing. The resulting model from neural computing is often called an Artificial Neural Network (ANN) [6]. ANN has the ability to mapping high nonlinear system, associable memory and abstractly generalization. Neural networks have been used in many business applications for speech, pattern recognition, prediction, forecasting and classification. Neural networks are computational models for data processing and are particularly useful for detecting the fundamental relationship among a set of variables or patterns in the data.

III. The Perception

The perceptron is a mathematical model of a biological neuron. The simplest and oldest model of Neuron takes some inputs, sums them up, applies activation function and passes them to output layer [3]. The dendrite in actual neurons receives electrical signals from the axons of other neurons, and the electrical signals are represented as numerical values in the perceptron. At the synapses between the dendrite and axons, electrical signals are modulated in various amounts. These are modelled in the perceptron by multiplying every input value by a value called the weight. An actual neuron fires an output signal only when the total strength

of the input signals exceeds a certain threshold. The model shows observable fact in a perceptron by calculating the weighted sum of the inputs to represent the total strength of the input signals, and applying a step function on the sum to determine its output. The output is fed to other perceptrons as in biological neural network.

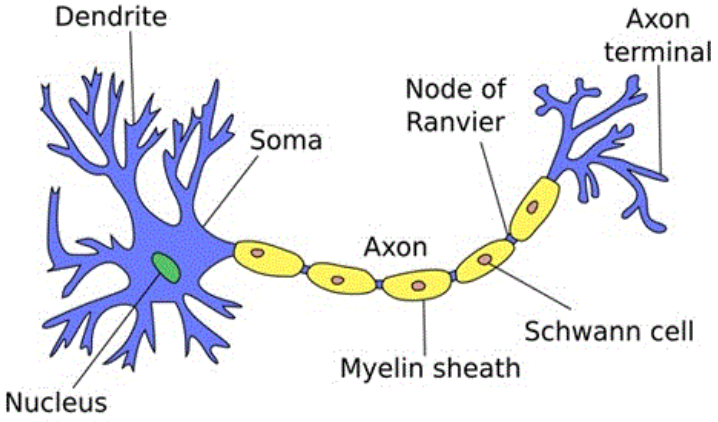

Fig1: Biological Neuron

IV: CLASSIFICATION

The problem of data classification has numerous applications in a wide variety of mining applications. This is because the problem attempts to learn the relationship between a set of feature variable and a target variable of interest. Since many practical problems can be expressed as associations between feature and target variables, this provides a broad range of applicability of this model. Given a set of training data points along with associated training labels, determine the class label for an unbalanced instance. Classification algorithms typically contain two phases.

Training Phase: In this phase, a model is constructed from the training instances.

Testing Phase: In this phase, the model is used to assign a label to an unbalanced test instance. Applications of classification places a vital role in medical disease diagnosis and in biological data analysis.

Medical Disease Diagnosis:

In recent years, the use of data mining methods in medical technology has gained increasing traction. The features may be extracted from the medical records, and the class labels correspond to whether or not a patient may pick up a disease in the future. In these cases, it is desirable to make disease predictions with the use of such information.

Biological Data Analysis:

Biological data is often represented as discrete sequences; in which it is desirable to predict the properties of particular sequences. In some cases, the biological data is also expressed in the form of networks

Classification plays a major role in the data mining approach and it aims to build a classifier model to describe and identify the important data classes and provides a better understanding of data at large. In this paper, a novel intelligent classifier model is designed employing. ELMAN neural model being recurrent neural network architecture explores the classification process with its faster convergence.

V:AUTISM SPECTRUM DISORDER(ASD)

Autism spectrum disorder (ASD) is a neuro developmental disorder that is defined by weakened social interactions, impaired verbal and non-verbal communication and repetitive actions [1]. Autism spectrum disorders (ASD) are a group of related brain-based disorders that affect a child's behaviour, social, and communication skills. ASD is of complex disorders of brain development. These disorders are characterized, in varying grades, by difficulties in social interaction, verbal and nonverbal communication and repetitive behaviours. Autism seems to own its roots in very primary brain development [4]. However, the most apparent signs of autism and symptoms of autism tend to appear between 2 and 3 years of age. Both children and adults with autism normally show difficulties in verbal and non-verbal communication, social interactions, and leisureliness or play activities.

Although neuropsychiatric disorders have a well-established genetic background, their specific molecular foundations remain elusive. This has prompted many investigators to design10 studies that identify explanatory biomarkers, and then use these biomarkers to predict clinical outcomes. One approach involves using machine learning algorithms to classify patients based on blood. Autism spectrum disorders (ASDs) are devastatingneurodevelopmental disorders characterized by deficits inSocial communication and interaction across multiplecontexts as well as restricted, repetitive patterns of interests andbehaviour.

VI: ELMAN Network

Elman neural network [2] is a feedback neural network, which is optimized based on the research of back propagation (BP) neural network [3] in 1990 by Elman; it is to add a

connecting layer to the hidden layer of feed forward network as time delay operator so as to memorize and, therefore, make the system have the characteristic of time varying and have stronger global stability

Elman Network is introduced by Jeffrey L. Elman to make the RNN learn from time-varying pattern input sequences. Elman network has four types of layers, they are Input Layers, Hidden Layers, Context Layers and Output Layers. The network architecture of Elman network is given in Figure 1.

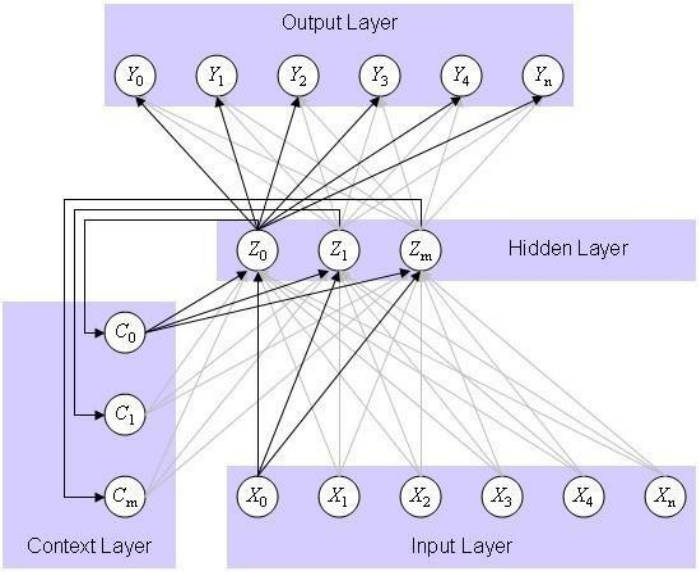

Figure2: Elman Neural Network

Standard Elman Network contains four different types of layers. The Input layer, hidden layers and output layers are very similar to the Artificial Neural Network environment. Context layer is added in the Elman network to produce more relevant results by which previous iteration results are permitted to control the current outputs to a certain degree.

VII: Modified ELMAN Network:

In the process of finding ASD classifications from Genome sequences, a lengthy pattern of human gene sequence has to be verified for several alleged occurrences of predefined genome patterns. A new layer named participant polypeptide is introduced in modified Elman network, in which a set of polypeptide/proteins is uesd in ASD findings. The sigmoid function is also altered to adopt new participant polypeptides in this layer. Each node in the context layer gets an influence from each other node in the participant polypeptide layer which shifts-up the sigmoid function to fire the connection in detection of a ASD gene sequence occurrence. Proposed CRNNC Elman network is illustrated in Figure 2.

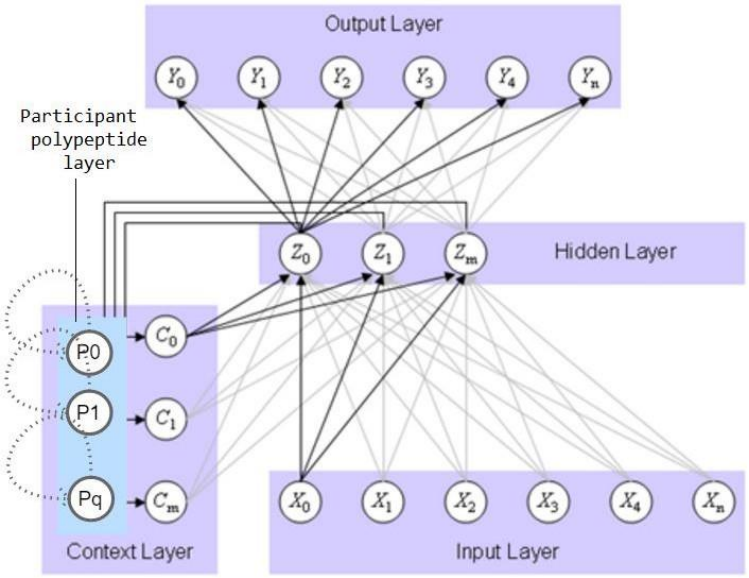

Figure3: Modified Elman Neural Network

These changes are adopted in proposed Elman Network model to find the earliest detection of a particular gene sequence based on its correlation to the ASD classification. The detection time is inversely proportional to the correlation of the gene sequence and ASD subgroup. The activation function curve of proposed CRNNC Elman network is given in Figure4.

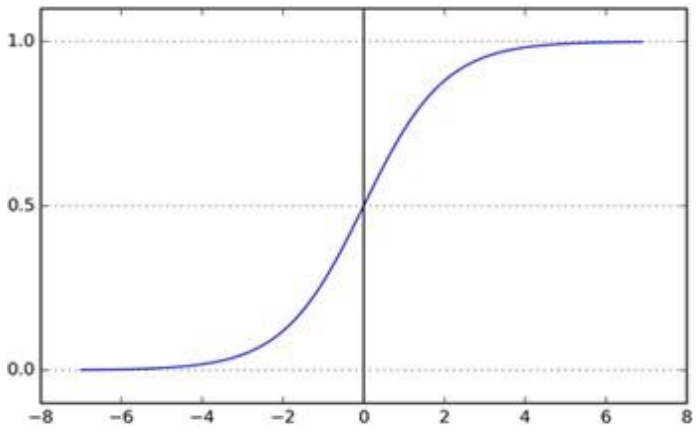

Figure 4: Activation function curve

VIII:RESULTS AND ANALYSIS

The experiments are conducted by bifurcating the entire dataset into 10 different parts to evaluate the intermediate performances of existing methods and proposed method. Accuracy, Sensitivity, Specificity, F1-Score and processing time are measured for all methods in each

time period to get the complete analysis. The parameters which are taken into consideration hare are classification accuracy sensitivity and specificity, F-score and processing time.

Accuracy:

Accuracy is one of the prime evaluating factors of any classification algorithm. Accuracy is calculated as $\frac{(TP+TN)}{(TP+TN+FP+FN)}$, where TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative. Accuracy is directly proportional to the quality of a classification algorithm. The measured values of accuracy for existing methods and proposed method are given in Table

S.NO	ELMAN	MOD.ELMAN
1	91.91	94.54
2	91.72	95.43
3	94.46	96.11
4	94.46	96.11
5	91.21	96.01
6	90.87	96.28
7	91.27	97.41
8	93.08	95.49
9	93.87	94.43
10	91.77	95.1

Table 1: Accuracy(%)

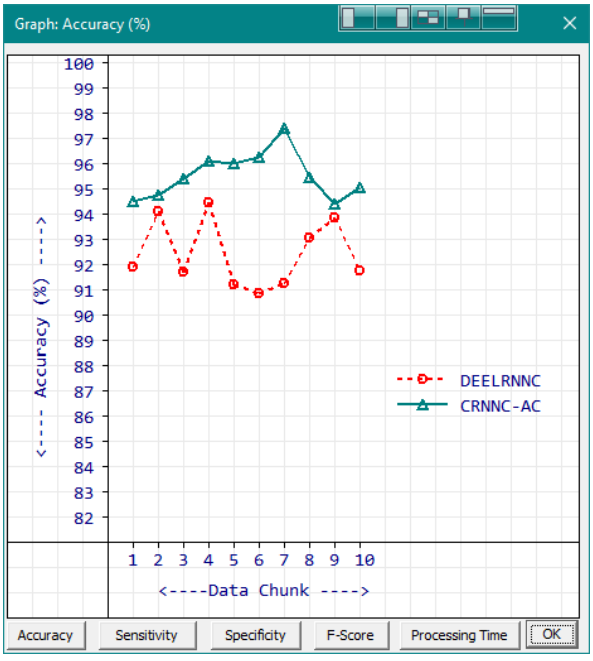

Figure 5: Accuracy(%)

IX: CONCLUSION

The gene expression values of Autism Spectrum Disorder(ASD) were successfully analysed with the goal of improving the selection process. This was accomplished by introducing a new layer called participant polypeptide layer in standard Elman Network. It was noted that the expression of genes potentially associated with the ASD varies greatly among the observation hence the existing standard Elman Neural Network approach does not provide a reliable result. The classification accuracy in predicting the ASD gene sequence from this modified Elman Network experimentally proved to be good than that of the existing method.

X: REFERENCES:

[1]. Selection and classification of gene expression in autism disorder: Use of a combination of statistical filters and a GBPSOSVM algorithm

[2]. Study on Optimized Elman Neural Network Classification Algorithm Based on PLS and CA, Computational Intelligence and Neuroscience Volume 2014, Article ID 724317, 13 pageshttp://dx.doi.org/10.1155/2014/724317

[3]. J. L. Elman, "Finding structure in time," *Cognitive Science*, vol.14, no. 2, pp. 179–211, 1990.

[4]. Published international journal titled" A Novel Feature Selection Method for pre-processing the ASD Dataset" in International journal on pure and Applied Mathematics (scopusIndexed) Volume 118 No.8 2018,17-25,ISSN 1311-8080(printed Version)1314-3395(on-line Version) Special Issue

[5]. Priyanka Gaur, "Neural Networks in Data Mining", International Journal of Electronics and Computer Science Engineering, ISSN- 2277-1956.

[6].M.CharlesArockiaraj, "Applications of Neural Networks in Data Mining" International Journal of Engineering and Science, Vol.3, Issue 1 (May 2013), PP 08-11 ISSN (e): 2278-4721, ISSN (p):2319-6483.