Classification of Autism Spectrum Disorder: Deep Learning Approach

Sevdanur Genç[1*], Duygu Bağcı Daş[2]

[1]*Department of Computer Technology, Taşköprü Vocational School, Kastamonu University, Kastamonnu, Turkey*
[2]*Department of Computer Technology, Bergama Vocational School, Dokuz Eylül University İzmir, Turkey*

***Abstract*—Autism is a complex developmental disorder that manifests itself as a life-long neuropsychiatric disorder in the first years of life, manifested by significant delays and deviations in the area of interaction and communication and restrictive interests. Autistic individuals may have problems in social skills, language development, and behavior. These problems are usually communicating with other people, making friends and difficulties in doing what is said. It is estimated that besides genetic causes, environmental reasons are also effective in the development of autism. Today it is certain that there is not a single factor that causes autism. Autism is a complex disorder that occurs when multiple factors come together. Nowadays, many researchers have worked on more effective solutions to these complex disorders. For this purpose, classification estimations have been made using machine learning methods on various dataset that have been used in the literature. Deep learning is another approach that has expanded machine learning and artificial intelligence scope. Deep Learning is a special kind of machine learning. It learns the examined world in the form of hierarchical concepts that are nested, defining each concept as an association with simpler concepts. At this point, classifications become very strong and flexible. In this study, it has been analyzed the datasets of Autism Spectrum Disorder using deep learning based classification approach which is a sub-branch of machine learning. As a result of the analyzes, it has been observed that the deep learning approach in test data gives better results than the other machine learning methods.**

***Keywords*—classification, machine learning, artificial intelligence, deep learning, autism spectrum disorder.**

I. INTRODUCTION

Autism is a complex developmental disorder that manifests itself as a neuropsychiatric disorder in the early years of life and is characterized by significant delay and deviation in the field of interaction and communication, and restrictive areas of interest. Individuals who suffer from autism can struggle in social skills, language development, and attitude. These issues are generally about communicating with other people, being a friend and performing what has been said. It has been thought that besides genetic causes, environmental causes also affect autism. Today, it has been certainly known that there is not just a single parameter that causes autism. Autism is a complex disorder that occurs in which many factors are gathered.

Nowadays, many researchers have been studying on such disorders to come up with more effective solutions. To achieve such a goal, many classification predictions have been performed by using machine learning methods on various dataset. On the other hand, deep learning which expands the study field of machine learning and artificial intelligence can also be utilized to accomplish such an aim.

Deep learning [1] is a special kind of machine learning. It learns the investigated environment as nested hierarchical concepts in which each concept have been defined as the relationship with simpler concepts. Therefore, classifications have become more powerful and flexible.

The reason for using deep learning in problem solutions is that the result in multi-layered network designs can be reached in the best and optimum way. Important parameters like dataset size, learning rate, momentum coefficient, epoch, activation functions, layer amount and hidden layer amount usually need to design a model [2].

Artificial deep neural networks have various sub-architectures. Among these architectures, Deep Neural Networks are networks in which data have been conducted forwardly from input to output. Convolutional Deep Neural Networks, on the other hand, are suitable for computer vision and speech recognition. In addition to those architectures, Recurrent Neural Networks and Deep Belief Networks are used frequently. Recurrent Neural Networks are utilized in mostly language modeling. The sub-classes of this architecture are appropriate to classifying time series.

Thabtah [3] has utilized a machine learning approach for a classification problem that is related with autism spectrum disorder. In that study, datasets have been published as a free mobile app which has contained test questions. Those tests, which are answered by patients and their relatives, have control individual's behavioral conditions that only indicate autistic characteristics [4]. In addition, it has been foresighted that Decision Tree and Support Vector Machine (SVM) algorithms can be applied for classification problems.

In this study, analyses have been performed on dataset about autism spectrum disorder by using deep learning based classification approach which is one of the sub-branch of machine learning. Analyses have indicated that deep learning approach gives healthier results in both training and test data when compared to other machine learning methods.

Manuscript received December 09, 2018; accepted July 08, 2019.
*Corresponding author: sgenc@kastamonu.edu.tr

II. MATERIAL AND METHOD

Three datasets about autism spectrum disorder that belongs to the adults, adolescents, and children have been accessed from UCI Machine Learning Repository and classification studies have been performed by using deep learning algorithms on these datasets [6][7][8].

There are 1100 samples in total and 20 features that includes 10 behavioral and 10 individual characteristics which belong to most distinctive autistic characteristics. For classification, KNIME, which is a software that recognizes data and designs data science workflows to make them reusable, have been utilized [9].

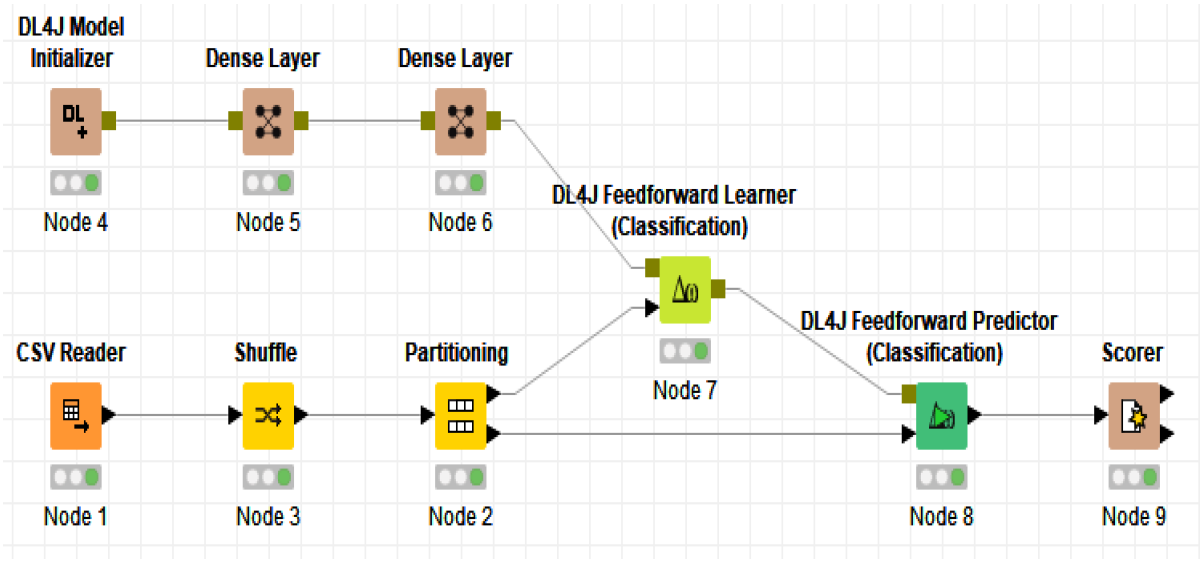

Fig. 1. The Deep Learning Algorithm created on the dataset by KNIME software.

Deep learning extension has been used with KNIME [10]. The workflow of the Deep Learning Algorithm created by KNIME software is designed as shown in figure 1. If we examine this workflow step by step, firstly, as shown in figure 2, sorted CSV files in the dataset which are existed in the first phase of application are read and samples are shuffled randomly. This randomly shuffled dataset is separated as training and test dataset.

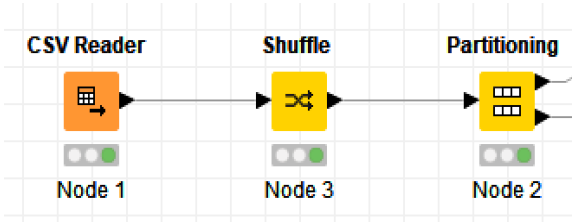

Fig. 2. Preparation phase of dataset for model

In the second phase which has been shown in figure 3, the Neural Network model is built. Necessary layers are determined. A linear operation is designed in which each input is connected to an output with a weight via layers.

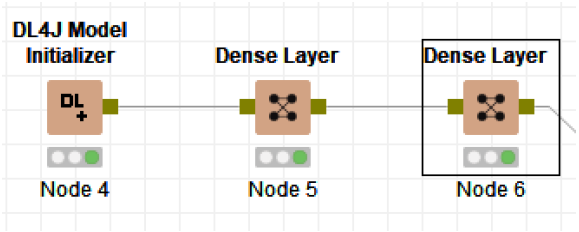

Fig. 3. Constructing of Neural Network model

Lastly, a model is created for training dataset as it is shown in figure 4.

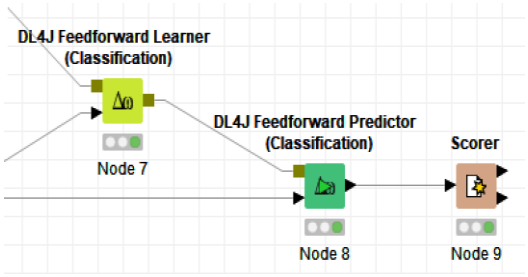

Fig. 4. Training dataset modeling

Using test dataset on this created model, classification is evaluated and results are saved in the system. In this study, while Forward Feed Network method has been used as a model [11], Stochastic Gradient Descent [12] and Negative Log Likelihood [13] has been used as optimization algorithms.

III. EXPERIMENTAL RESULTS

In order to reach the information about whether that the adults, adolescents, and children have autism spectrum disorder or not by the results of test questions which were answered by those people via the mobile application, the utilization of machine learning methods have been foresighted from early studies. By using deep learning algorithm, which is one of the classification methods, the dataset of autism spectrum disorder has been analyzed. According to analyses results performance indicators like accuracy, sensitivity, specificity, and F-measure has been calculated.

Experimental results of test data and performance measurements of these results have been given in table 1.

According to the results, estimations about whether those adults, adolescents, and children who have those 10 behavioral and 10 individual characteristics have this

sickness or not by 100%, 91.67%, and 93.33% respectively. Experimental results have clearly indicated that utilizing deep learning on the classification of autism spectral disorder is effective.

TABLE I. RESULTS OF DEEP LEARNING ALGORITHM FOR RELATED DATASET.

Data Set	Accuracy	Sensitivity	Specificity	F-Measure
Adults	%100	1.0	1.0	1.0
Adolescents	%91.67	1.0	0.875	0.889
Children	%93.33	1.0	0.8947	0.9167

IV. DISCUSSIONS AND CONLUSIONS

In this study, the suitability of deep learning methods which is still controversial on small datasets has been investigated by an example. Feed forward deep neural network has been applied on a dataset which consists of 1100 samples with 20 features. By utilizing Stochastic Gradient Descent optimization, oscillations have reduced and learning speed has been improved. Experimental results have clearly indicated that Deep Learning architecture can be also successful on small databases.

In Deep Learning methods, it should be considered about hyper-parameter groups like optimization algorithms, weight values, and the dimension of the dataset, momentum ratio, and learning rate. Changing of several parameters can lead the model to independent models. Therefore, when evaluating these parameters, it is necessary to take them into account together.

Training of the model can take days in Deep Learning methods. So, an appropriate combination of hyper-parameters improves efficiency. Basically, Deep Learning architecture trains model by forming optimization approaches. Additionally, Deep learning architecture can fabricate meaningful results from all datasets in which artificial algorithms are used when the artificial intelligence architecture of deep learning is considered.

REFERENCES

[1] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436.

[2] Derin Öğrenme Uygulamalarında En Sık kullanılan Hiper-parametreler : https://medium.com/deep-learning-turkiye/derin-ogrenme-uygulamalarinda-en-sik-kullanilan-hiper-parametreler-ece8e9125c4 [Accessed: 08.12.2018]

[3] Thabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. Proceedings of the 1st International Conference on Medical and Health Informatics 2017, pp.1-6. Taichung City, Taiwan, ACM.

[4] Thabtah, F. (2017). ASDTests. A mobile app for ASD screening. www.asdtests.com [Accessed December 20th, 2017].

[5] Thabtah, F. (2017). Machine Learning in Autistic Spectrum Disorder Behavioral Research: A Review. To Appear in Informatics for Health and Social Care Journal. December, 2017 (in press).

[6] Autism Screening Adult Data Set: https://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult [Accessed: 08.12.2018].

[7] Autistic Spectrum Disorder Screening Data for Adolescent Data Set: https://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent [Accessed: 08.12.2018].

[8] Autistic Spectrum Disorder Screening Data for Children Data Set - https://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children [Accessed: 08.12.2018].

[9] Knime. Software available at: http://www.knime.org/ [Accessed: 08.12.2018].

[10] Deep Learning Tutorial MNIST - https://www.knime.com/nodeguide/analytics/deep-learning [Accessed: 08.12.2018].

[11] Corani, Giorgio. "Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning." Ecological Modelling 185.2-4 (2005): 513-529.

[12] Bottou, Léon. "Large-scale machine learning with stochastic gradient descent." Proceedings of COMPSTAT'2010. Physica-Verlag HD, 2010. 177-186.

[13] Friedman, Jerome H. "Stochastic gradient boosting." Computational Statistics & Data Analysis 38.4 (2002): 367-378.