ARTIFICIAL NEURAL NETWORKS BASED-PREDICTION OF AUTISM SPECTRUM DISORDER

I. Ucuz, and A. Uzun Cicek

Abstract*— *Aim: Autism Spectrum Disorders (ASD) is one of the important neurodevelopmental disorders. This study aimed to perform artificial-intelligence-based modelling based on the prenatal-perinatal factors, family history, and developmental characteristics, which are emphasized as risk factors for ASD in the literature. *Materials and Methods:* The study was designed with retrospective management and data from 136 children with ASD and 143 healthy children were included. *Results:* According to the findings of the MLP model, the five most important factors were the mean age of first words (months), the mean age of head control (months), the mean age of sitting without support (months), history of autism in the family, and the mean paternal age at pregnancy (years), respectively. Overall percentages of the training and testing samples were 91.4% and 88.0%. AUC for the model was 0.922 for the separation of the autism and control groups. *Conclusion:* The proposed model is able to successfully differentiate patients with autism spectrum disorders from healthy individuals and identify factors associated with the disease.

Keywords— Artificial neural networks, autism, prenatal risk factors, perinatal risk factors

1. INTRODUCTION

AUTISM Spectrum Disorders (ASD) is one of the important neurodevelopmental disorders that manifest itself with deficits in social communication-interaction and limited-repetitive behavioural patterns, whose symptoms start at the early developmental period and cause clinically significant impairment in social or other important areas of functioning [1]. Although symptoms can be seen even in the 6-12 month period, the diagnosis age is usually around 24 months [2]. Studies on the prevalence of ASD have indicated that its prevalence varies according to the countries and the diagnostic methods used; however, the main emphasis is that the incidence is gradually increasing [3]. In general, the prevalence in developed countries is reported to be 1-5% [4]. A study comparing the prevalence of ASD in China and western countries found a 1% prevalence rate [5].

ASD is an important disorder that has many burdens for both the individual and society [1]. Given that the dramatic increase in its prevalence rates, it is important to know the situations that pose a risk for ASD. Several factors have been suggested in the etiology. In the meta-analysis study conducted by *Tick et al.*, it was emphasized that ASD is hereditary at a rate of 64-91%, and environmental factors, as well as genetic features, have an important effect on the increase in ASD prevalence [6]. Among environmental factors, prenatal characteristics are shown as one of the most important factors emphasized in ASD etiology. A meta-analysis study by *Gardener et al.* has determined that advanced parental age at pregnancy, mother's use of medication in the prenatal period, presence of gestational diabetes, history of bleeding in pregnancy, being the first-child are the risk factors associated with ASD [7].

Artificial neural networks (ANNs) are computing systems created on the computer to automatically perform data samples regarding the abilities and features of the human brain, such as generating new information through learning and discovering new information [8]. Artificial intelligence-based systems have found wide use in many fields of medical sciences as well as other fields [8-13]. This study aimed to perform artificial-intelligence-based modelling based on the prenatal-perinatal factors, family history, and developmental characteristics, which are emphasized as risk factors for ASD in the literature. In this way, it is purposed to determine the contribution of possible factors on ASD development by addressing many risk factors and developmental variables together.

2. MATERIAL AND METHODS

2.1. Study Procedure

This study was designed retrospectively and conducted on the data of 136 patients diagnosed with ASD (case group) at the Department of Child and Adolescent Psychiatry in Inonu University Faculty of Medicine between October 2015-September 2020, and the data of 143 healthy children (control group) matched with the case group for age and gender. ASD was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) criteria [14]. Children with known chronic medical illness(es) and non-psychiatric drug use were excluded from the study.

Developmental characteristics, prenatal-perinatal period risk factors, and family history of children included in both the case and control groups were used as input parameters in artificial intelligence modelling. The developmental characteristics considered in the study are the age of onset of the first words (months), age of providing head control (months), age of starting to walk (months), and age of starting to sit without support (months). Variables of the maternal and paternal age at pregnancy, family history of autism, number and order of siblings, socioeconomic level of the family, maternal and paternal education level, place of residence, and parental

✉ **İlknur UCUZ,** Inonu University Faculty of Medicine, Department of Child and Adolescent Psychiatry, Malatya, Turkey, (ilknur.ucuz@inonu.edu.tr)

Ayla UZUN CİCEK, Cumhuriyet University, Faculty of Medicine, Department of Child and Adolescent Psychiatry, Sivas, Turkey, (aylauzun@cumhuriyet.edu.tr)

Manuscript received Sep 9, 2020; accepted Oct 30, 2020.
Digital Object Identifier:

consanguinity were employed as family history. The presence of the mother's history of regular drug use, excessive tea and coffee consumption, smoking, hypertension, gestational diabetes, an endocrine disease other than diabetes mellitus, radiation exposure, threatened miscarriage, severe infection requiring inpatient treatment such as respiratory tract or urogenital during pregnancy, and vitamin supplements status and nutritional problems in pregnancy were included in the category of prenatal risk factors. Variables including the month and season of the child's birth, birth time (weeks), birth weight, birth type, birthplace, neonatal jaundice, history of blood incompatibility and blood transfusion, birth complication(s) and seizure, and nulliparity were assigned to the model as perinatal factors. Artificial intelligence-based modelling was made for the diagnosis of ASD by giving the selected variables as input parameters to the system.

3. Data Analysis

IBM SPSS Statistics 25.0 program was used for data analysis. Qualitative data were expressed as number (n) and percentage (%), quantitative data were given as mean ± standard deviation (SD). Normality was analyzed using the one-sample Kolmogorov-Smirnov test. Statistical comparisons were performed with the chi-square test, Mann Whitney U test, independent t-test as appropriate. A p-value of less than 0.05 was accepted as statistically significant.

4. Modelling and Performance Evaluation

Nearly 70 percent and 30 percent of the entire dataset were used for training and testing stages in the development of MLP ANN models, respectively. For the model, the rescaling procedure for quantitative variables was standardized, the number of the hidden layer was 1, the number of units in the hidden layer was 6, the hidden layer activation function was a hyperbolic tangent, the number of units in the output layer was 2, the activation function of the output layer was softmax and error function was cross-entropy. The gradient descent approach optimized the models' hyperparameters. The performance of the model was evaluated with the accuracy, cross-entropy error, and area under the receiver operating curve metrics.

5. Results

The mean age of the children was 45.9±11.1 (months) in the ASD group and 46.5±11.3 (months) in the control group (p=0.677). Of the participants, 99 (72.8%) children in the ASD group, and 108 (75.5%) children in the control group were male (p=0.602). While the mean maternal age at pregnancy was 29.1±7.4 (years) in the ASD group and 28.9±5.6 (years) in the control group (p=0.752), the mean paternal age at pregnancy was 32.7±7.8 (years) in the ASD group 31.4 ± 5.7 (years) (p= 0.324). The mean age of providing head control of children in the ASD group was 3.8±0.9 (months), the mean age of sitting unsupported was 9.0±1.7 (months), the mean age of walking was 14.4±2.4 (months), and the mean age of the first words was 22.8±7.6 (months). In the control group, these values were 2.6±0.8 (months) for head control, 7.9±1.4 (months) for sitting unsupported, 12.7±1.6 (months) for walking, 15.6±3.3 (months) for first words, respectively (all p-values <0.001). While children with ASD were born mostly in July (15.4%) and summer (39.7%), children in the control group were born mostly in May (14.7%) and spring (28.0%) (p=0.060 and 0.023, for the month at birth and the birth season, respectively).

Sociodemographic characteristics such as place of residence, parental education level, family income level, number and order of siblings, parental consanguinity, and some variables in pregnancy including threatened miscarriage, hypertension, presence of endocrine disease other than diabetes mellitus, regular drug use, radiation exposure, nutritional problems, vitamin supplement intake, excessive tea, and coffee consumption during pregnancy were similar between the two groups (all p-values >0.05). Also, there was no statistically significant difference between the case and control groups for some data on birth (birthplace, birth time and weight, birth complication(s), history of blood incompatibility, and blood transfusion) (all p-values> 0.05).

While the rate of positive family history for autism was 44.1% in the case group, it was 5.6% in the control group (p <0.001). 48.5% of the children with ASD and 34.3% of the control group children were the first children (p=0.016). While the nulliparity rate in the mothers of those in the ASD group was 49.3%, this rate was 34.3% in the control group (p=0.011). The smoking rate of the mothers during pregnancy was 28.7% in the ASD group and 16.8% in the control group (p=0.018). Gestational diabetes history was 8.8% in the ASD group and 2.8% in the control group (p=0.030). While 23.5% of the ASD group had a history of severe infection requiring inpatient treatment during pregnancy, this value was 11.2% in the control group (p=0.006). 66.2% of the children with ASD and 82.5% of the children in the control group were born by normal vaginal delivery (p=0.002). 72 (57.1%) of the children with ASD and 59 (41.3%) of the children in the control group had a history of neonatal jaundice (p=0.009). Besides, 44 (34.9%) of children with ASD had a history of seizure/convulsion in 6 (4.2%) of healthy children (p <0.001).

Independent variable importance values of the MLP model is demonstrated in Table 1. According to the findings of the MLP model, the five most important factors were the mean age of first words (months), the mean age of head control (months), the mean age of sitting without support (months), history of autism in the family, and the mean paternal age at pregnancy (years), respectively. Importance and normalized importance values for each variable are presented in descending order (Table 1).

The classification results of the training and testing samples for the model are summarized in Table 2. Overall percentages of the training and testing samples were 91.4% and 88.0%.

 ISSN: 2548-0650 http://dergipark.gov.tr/jcs

TABLE I
INDEPENDENT VARIABLE IMPORTANCE VALUES OF THE MLP MODEL

Variable	Importance	Normalized Importance
Age of first words (months)	0.189	100.0%
Age of head control (months)	0.140	74.3%
Age of sitting without support (months)	0.060	31.8%
Paternal age at pregnancy (years)	0.057	30.1%
Positive family history of ASD	0.055	29.0%
Maternal age at pregnancy (years)	0.047	25.1%
Month at birth	0.029	15.5%
Regular drug use during pregnancy	0.028	14.9%
Number of sibling(s)	0.025	13.1%
Threatened miscarriage	0.022	11.4%
Family income level	0.021	11.2%
Excessive tea and/or coffee consumption during pregnancy	0.021	11.2%
Birth type	0.021	10.9%
Birth place	0.020	10.8%
Smoking during pregnancy	0.020	10.8%
Order of sibling(s)	0.020	10.5%
Other endocrine problems other than diabetes during pregnancy	0.019	10.2%
Gestational age at birth (years)	0.019	10.1%
First-born	0.019	10.1%
Paternal education level	0.018	9.5%
History of neonatal jaundice	0.017	9.1%
Vitamins, minerals, and supplements in pregnancy	0.017	8.8%
Gender	0.015	7.7%
Birth season	0.014	7.2%
Seizure/convulsion history	0.012	6.6%
Blood incompatibility	0.011	5.8%
Birth complication(s)	0.010	5.2%
Severe infection in pregnancy	0.010	5.1%
Birth weight	0.010	5.1%
Place of residence	0.009	4.9%
Hypertension in pregnancy	0.008	4.3%
Nulliparity	0.004	2.4%
Gestational diabetes	0.003	1.8%
Age of starting to walk (months)	0.002	1.3%
Maternal education level	0.002	1.3%
Nutritional problems during pregnancy	0.002	1.2%
Parental consanguinity	0.002	1.0%
Blood transfusion	0.001	0.5%
Radiation exposure in pregnancy	0.000	0.1%

TABLE II
THE CLASSIFICATION RESULTS OF THE TRAINING AND TESTING SAMPLES FOR THE MODEL

Sample	Observed	Predicted		
		Autism	Control	Percent Correct
Training	Autism	76	15	83.5%
	Control	1	94	98.9%
	Overall Percent	41.4%	58.6%	91.4%
Testing	Autism	28	7	80.0%
	Control	3	45	93.8%
	Overall Percent	37.3%	62.7%	88.0%

Table 3 indicates the model summary information for both training and testing stages. Cross entropy errors were 54.226 for the training sample and 26.631 for the testing sample. Besides, AUC for the model was 0.922 for the separation of the autism and control groups.

TABLE III
THE MODEL SUMMARY INFORMATION FOR BOTH TRAINING AND TESTING STAGES

Model Summary		
Training	Cross-Entropy Error	54.226
	Percent Incorrect Predictions	8.6%
Testing	Cross-Entropy Error	26.631
	Percent Incorrect Predictions	12.0%

6. DISCUSSION

In this study, 39 variables were included in the modelling for ASD prediction in MLP modelling, and variables contributing to the diagnosis of ASD were determined according to their contribution percentages. Thus, among the variables selected for ASD diagnosis, it was found that the age of onset of the first words, age of providing head control, age of starting to sit without support, history of autism in the family, and age of the father at pregnancy was the most important five variables. On the other hand, it was determined that the presence of nutritional problems during pregnancy, parental consanguinity, the history of blood transfusion during the neonatal period, and the history of radiation exposure in pregnancy affected the least effect on the diagnosis of ASD.

Machine learning is a field of research that creates prediction models within the framework of accuracy and originality from data sets by integrating artificial intelligence, mathematics, and other branches of science. In this way, it is ensured that the data give the best estimation outcome without too much human intervention [15]. Regarding the relevant literature in ASD, it is seen that artificial intelligence-based machine learning systems

are used for different purposes in ASD [15-20]. In a meta-analysis study aiming to determine the accuracy of machine learning algorithms in+ diagnosing ASD from brain magnetic resonance imaging (MRI) studies, 40 studies were included. This study has found that the sensitivity is 0.83 and the specificity is 0.84 for the structural MRI (sMRI) subgroup meta-analysis, and researchers have emphasized that further research is required to determine the potential benefits of machine learning in the clinical setting [16]. In a study conducted on the diagnosis of ASD using machine learning methods from 3-minute home videos, 30 behavioural characteristics including eye contact, expressive language, emotion expression, echolalia, joint attention were evaluated, and eight different machine learning methods were used for ASD prediction. Based on the results of the study, it has been suggested that modelling can be used in the diagnosis of ASD [17]. In addition, it is seen that artificial intelligence-based systems can be used for ASD in different areas such as determining the best distinguishing features in assessment tools used in the diagnosis of ASD [21], the prediction of stereotypical movements, which are one of the basic symptoms in ASD, with different techniques [22], analysis of neonatal leukocyte epigenomic markers for ASD prediction [23], developing a deep convolutional neural network that can help children with ASD recognize facial expressions accurately using mobile devices [24], autism classification based on logistic regression analysis [25].

In particular, clinical features and neuroimaging techniques have been used in artificial intelligence-based models for the diagnosis of ASD. Studies examining models based on different environmental features are extremely limited. In our study, modelling was performed based on various developmental characteristics and prenatal-perinatal risk factors. When the relevant literature in this field is examined; in a meta-analysis study conducted with 37,634 children with ASD, several factors such as mother and father's age above 35, some ethnic origins, history of hypertension and diabetes during pregnancy, birth complications, cesarean delivery, low birth weight were associated with ASD. In contrast, this study has indicated that umbilical cord entanglement, premature rupture of membranes, 5th minute Apgar score of less than 7, and respiratory tract infection are not among the factors that increase the risk in ASD [26]. In the study of *Modabbernia et al.*, it was found that vaccination, maternal smoking, thimerosal exposure, and assisted reproductive techniques were not associated with ASD risk. In parallel with our study, advanced parental age is associated with an increased risk of ASD. This study has also shown that while birth complications related to trauma or ischemia and hypoxia are strongly related to ASD, other factors related to pregnancy, such as maternal obesity, maternal diabetes, and cesarean delivery, are less associated with ASD [27]. In our study, among many prenatal-perinatal variables, the paternal age at pregnancy was one of the most effective factors in diagnosing ASD. In addition, the results of our study demonstrated that the age of the first words, the age of obtaining head control, the age of sitting without support, and the family history of autism are more dominant characteristics for the diagnosis of ASD than other prenatal-perinatal factors. Among these variables, the two most predictive factors are the age of onset of the first words and the time to provide head control.

ASD is one of the important neurodevelopmental disorders that significantly impair the quality of life of individuals in different social environments such as school [1, 28]. Therefore, making the diagnosis process as early as possible also provides significant benefits for the treatment and follow-up process [29]. However, one of the most important obstacles to early diagnosis is the lack of any definitive markers for diagnosis. This problem suggests the importance of determining the risk factors that would bring ASD to mind and lead to the diagnosis. The strength of this study is that it reveals the importance of several developmental characteristics for diagnosis before clinical signs fully appear. Despite our study has many strengths, it also has some limitations. Since the data included in the modelling are recruited retrospectively, some data may be missing or subject to recall bias. Future longitudinal studies with a larger sample size are of great importance in supporting the findings of this study.

REFERENCES

[1] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele, "Autism spectrum disorder," The Lancet, vol. 392, pp. 508-520, 2018.

[2] K. Sanchack and C. A. Thomas, "Autism spectrum disorder: Primary care principles," American family physician, vol. 94, pp. 972-979, 2016.

[3] E. Fombonne, "The rising prevalence of autism," Journal of Child Psychology and Psychiatry, vol. 59, pp. 717-720, 2018.

[4] K. Lyall, L. Croen, J. Daniels, M. D. Fallin, C. Ladd-Acosta, B. K. Lee, et al., "The changing epidemiology of autism spectrum disorders," Annual review of public health, vol. 38, pp. 81-102, 2017.

[5] V. Courchesne, A.-A. S. Meilleur, M.-P. Poulin-Lord, M. Dawson, and I. Soulières, "Autistic children at risk of being underestimated: school-based pilot study of a strength-informed assessment," Molecular Autism, vol. 6, p. 12, 2015.

[6] C. A. Labarrere, J. Woods, J. Hardin, G. Campana, M. Ortiz, B. Jaeger, et al., "Early prediction of cardiac allograft vasculopathy and heart transplant failure," American Journal of Transplantation, vol. 11, pp. 528-535, 2011.

[7] H. Gardener, D. Spiegelman, and S. L. Buka, "Prenatal risk factors for autism: comprehensive meta-analysis," The British journal of psychiatry, vol. 195, pp. 7-14, 2009.

[8] D. Rahul, "Machine learning in medicine," Circulation, vol. 132, pp. 1920-1930, 2015.

[9] A. Ari and D. Hanbay, "Deep learning based brain tumor classification and detection system," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 26, pp. 2275-2286, 2018.

[10] M. C. Çolak, C. Çolak, H. Kocatürk, S. Sagiroglu, and İ. Barutçu, "Predicting coronary artery disease using different artificial neural network models/Koroner arter hastaliginin degisik yapay sinir agi modelleri ile tahmini," Anadulu Kardiyoloji Dergisi: AKD, vol. 8, p. 249, 2008.

[11] E. Güldoğan, T. Zeynep, and C. Çolak, "Classification of Breast Cancer and Determination of Related Factors with Deep Learning Approach," The Journal of Cognitive Systems, vol. 5, pp. 10-14.

[12] E. Güldoğan, T. Zeynep, A. Ayça, and C. Çolak, "Performance Evaluation of Different Artificial Neural Network Models in the Classification of Type 2 Diabetes Mellitus," The Journal of Cognitive Systems, vol. 5, pp. 23-32.

[13] J. Stewart, P. Sprivulis, and G. Dwivedi, "Artificial intelligence and machine learning in emergency medicine," Emergency Medicine Australasia, vol. 30, pp. 870-874, 2018.

[14] A. Association, "American Psychiatric Association's Diagnostic and statistical manual of mental disorders (DSM-V)," 2013.

[15] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," Informatics for Health and Social Care, vol. 44, pp. 278-297, 2019.

[16] S. J. Moon, J. Hwang, R. Kana, J. Torous, and J. W. Kim, "Accuracy of Machine Learning Algorithms for the Diagnosis of Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Brain Magnetic Resonance Imaging Studies," JMIR mental health, vol. 6, p. e14108, 2019.

[17] Q. Tariq, J. Daniels, J. N. Schwartz, P. Washington, H. Kalantarian, and D. P. Wall, "Mobile detection of autism through machine learning on home video: A development and prospective validation study," PLoS medicine, vol. 15, p. e1002705, 2018.

[18] A. Pratap and C. Kanimozhiselvi, "Predictive assessment of autism using unsupervised machine learning models," International Journal of Advanced Intelligence Paradigms, vol. 6, pp. 113-121, 2014.

[19] M. Duda, R. Ma, N. Haber, and D. Wall, "Use of machine learning for behavioral distinction of autism and ADHD," Translational psychiatry, vol. 6, pp. e732-e732, 2016.

[20] S. H. Lee, M. J. Maenner, and C. M. Heilig, "A comparison of machine learning algorithms for the surveillance of autism spectrum disorder," PloS one, vol. 14, p. e0222907, 2019.

[21] D. P. Wall, J. Kosmicki, T. Deluca, E. Harstad, and V. A. Fusaro, "Use of machine learning to shorten observation-based screening and diagnosis of autism," Translational psychiatry, vol. 2, pp. e100-e100, 2012.

[22] N. M. Rad, S. M. Kia, C. Zarbo, T. van Laarhoven, G. Jurman, P. Venuti, et al., "Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders," Signal Processing, vol. 144, pp. 180-191, 2018.

[23] R. O. Bahado-Singh, S. Vishweswaraiah, B. Aydas, N. K. Mishra, A. Yilmaz, C. Guda, et al., "Artificial intelligence analysis of newborn leucocyte epigenomic markers for the prediction of autism," Brain research, vol. 1724, p. 146457, 2019.

[24] M. I. U. Haque, "A Facial Expression Recognition Application Development Using Deep Convolutional Neural Network For Children With Autism Spectrum Disorder To Help Identify Human Emotions," 2019.

[25] F. Thabtah, N. Abdelhamid, and D. Peebles, "A machine learning autism classification based on logistic regression analysis," Health information science and systems, vol. 7, p. 12, 2019.

[26] C. Wang, H. Geng, W. Liu, and G. Zhang, "Prenatal, perinatal, and postnatal factors associated with autism: a meta-analysis," Medicine, vol. 96, 2017.

[27] A. Modabbernia, E. Velthorst, and A. Reichenberg, "Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses," Molecular autism, vol. 8, p. 13, 2017.

[28] P. Titelman, Differentiation of self: Bowen family systems theory perspectives: Routledge, 2014.

[29] S. Mishra, D. Joshi, R. Ribeiro, and S. Anand, "Kinematics-coordinated walking pattern based on embedded controls," Journal of medical engineering & technology, vol. 34, pp. 329-334, 2010.

BIOGRAPHIES

İIknur UCUZ, graduated from Çukurova University Faculty of Medicine in 2005 and received her speciality education in the Department of Child and Adolescent Psychiatry at Atatürk University between 2009-2014. Since 2017, she has been working as a faculty member at İnönü University Faculty of Medicine, Department of Child and Adolescent Psychiatry. She is interested in anxiety disorders, autism spectrum disorders, attention deficit and hyperactivity disorder, artificial intelligence and machine learning in children and adolescents.

Ayla UZUN ÇİÇEK, who graduated from Cumhuriyet University Faculty of Medicine in 2007, received her speciality education in the Department of Child and Adolescent Psychiatry at Akdeniz University between 2009-2013. Between 2013-2015, he served as a specialist in Rize Recep Tayyip Erdoğan Training and Research Hospital. Since 2015, she has been working as a faculty member at Cumhuriyet University Faculty of Medicine, Department of Child and Adolescent Psychiatry. She deals with mood disorders, anxiety disorders, autism spectrum disorders, attention deficit and hyperactivity disorder in children and adolescents.