Early Detection of Autism Spectrum Disorder Using Deep Learning

Gavin Dsa
Shreyaskumar Bhat

Abstract

Autism Spectrum Disorder (ASD) is a developmental disorder that can cause significant social, communication, and behavioral challenges. Early diagnosis and intervention can help people with ASD reach their full potential, but current diagnostic procedures are time-consuming and ineffective. This study aimed to improve the diagnosis of autism by training and testing several state-of-the-art machine learning models on an Autism Spectrum Disorder dataset from the University of California, Irvine. The models were used to quantitatively identify the most significant indicators of autism in toddlers, Adolescents, and Adults. The use of neural networks is shown here with various optimization strategies to classify cases more likely to have ASD. This can be used for screening people that can be further diagnosed by medical professionals.

Introduction

Autism spectrum disorder (ASD) is a lifelong developmental disability that affects communication and behavior. It is the fastest-growing developmental disability, and prevalence rates continue to rise incredibly. Current estimates indicate that about 1.5% of the world's population is on the autism spectrum. Therefore, it is important for our classifier to work with high accuracy and precision scores since screening is an important factor in detecting ASD. The misclassification rates need to be low since not detecting an ASD can be dangerous, and treating it in time is also important. At the same time, it is important to note that classifying a case that does not have ASD as someone that has ASD can be worse since they do not need to undergo treatment; therefore, its important for these algorithms to be precise. In the current dataset, we have ten questions that are asked of the patients and other metrics such as gender, ethnicity, jaundice detection, etc. Based on these factors, we are trying to screen individuals for ASD and recommend further diagnoses for the same.

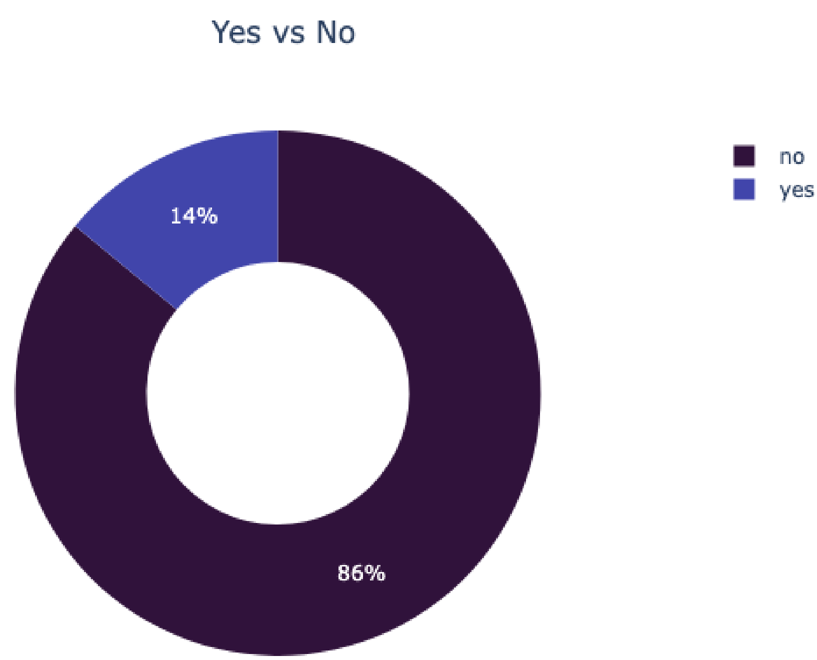
Background

The paper titled "An Accessible and efficient autism screening method for behavioral data and predictive analyses" by Fadi Thabtah, published in the Journal of Medical Systems in 2018 focuses on developing an accessible and efficient method for autism screening using behavioral data and predictive analysis. The proposed method uses a machine learning algorithm to analyze data from the Autism-Spectrum Quotient (AQ) questionnaire, which measures certain behavioral traits commonly associated with autism. The algorithms mainly used were Logistic Regression and Naive Bayes, which are simplistic and not as robust as complex neural networks that are bound to give higher accuracies and consider hidden characteristics that can be detected in case of future research findings. To make higher-end neural networks, we have started with creating basic models and further tried to optimize these for each dataset well as a case where all the datasets have been combined.

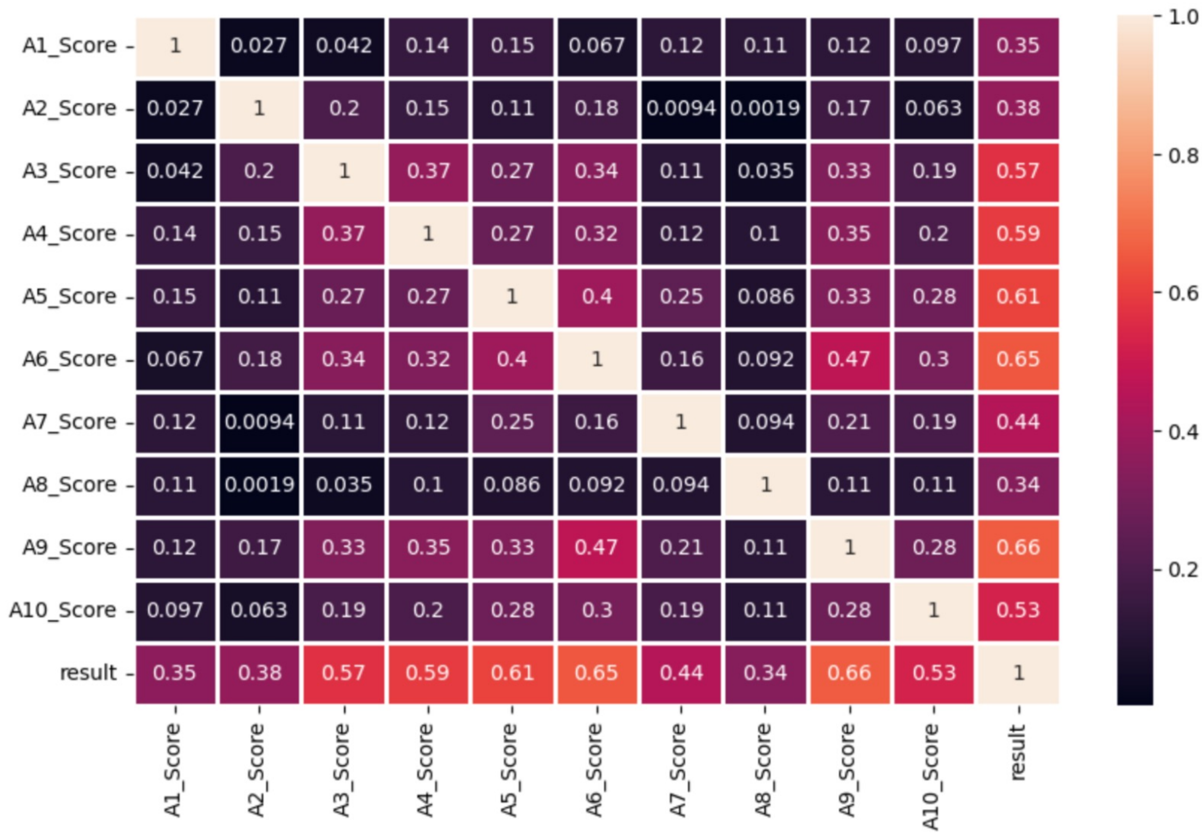
Approach

Exploratory Data Analysis:

The combination of the toddlers, adolescents, and adults datasets had a massive imbalance and required Random Oversampling to balance the classes. The master dataset also has 21 features and 1098 data points.



We also checked for correlation in the dataset and ensured that the ones we use are not highly correlated.

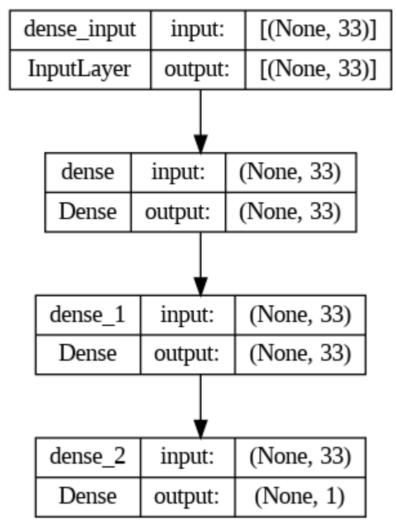


Data Preprocessing and Cleaning:

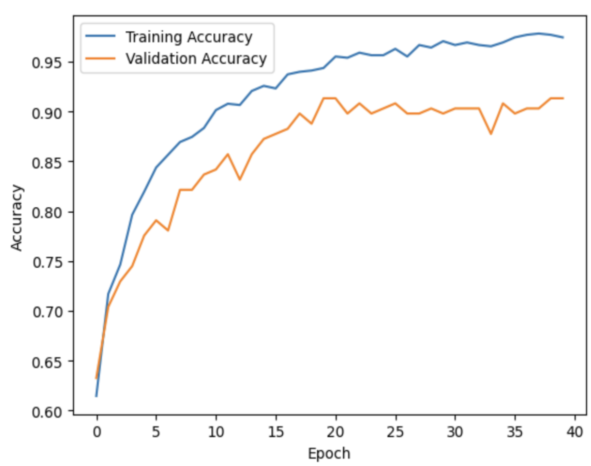
Since the dataset was highly imbalanced, we used random oversampling to balance the dataset for classification. We further filled null values for features like ethnicity and relation with their null values since there were 95 such cases, and completely excluding them would be a big loss, given the number of data points we have. We imputed these with their corresponding mode values as they are more likely to be there. We hot-encoded the categorical variables and created new columns for the neural network to comprehend and used the ‘Autism’ column as our target variable, which was created by further diagnoses of the patients. We also standardized the dataset so that no feature was given more importance than the other and are treated equally. We then divided the dataset into the train, validation, and test sets to feed it to the neural networks. We picked 80% of the dataset for training, 10% for validation, and the remaining 10% for testing. We did the same for all 4 cases. We then used TensorFlow to make tensors, which makes it easier for handling our data when performing transformations while dealing with neural networks.

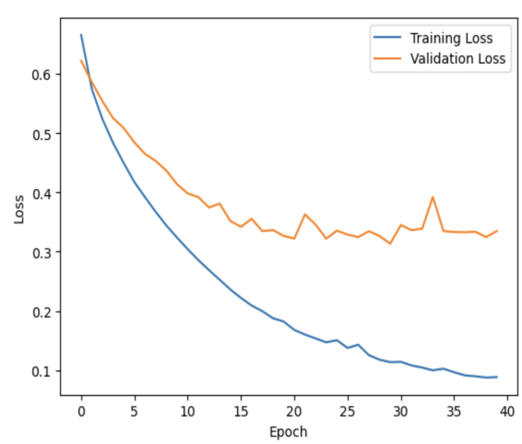
Model Implementation:

We created base models for the 4 cases that we are considering: The adult dataset, the Adolescents dataset, the Toddler dataset, and the combined dataset. The architecture for these base models was set as:



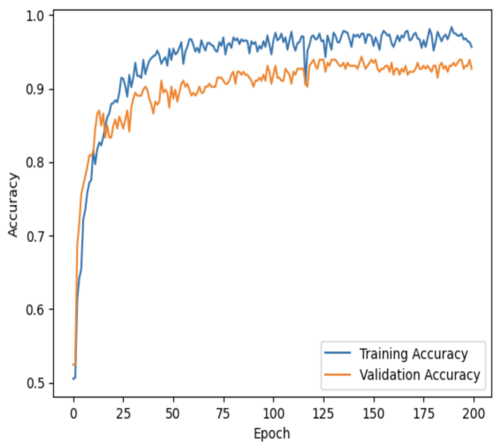
Here we have one input layer, one input layer with ReLU activation, a hidden layer with ReLU activation, and an output layer with a sigmoid activation function. Along with this, we used the Adam optimizer and binary cross entropy along with accuracy as the metric for evaluation.

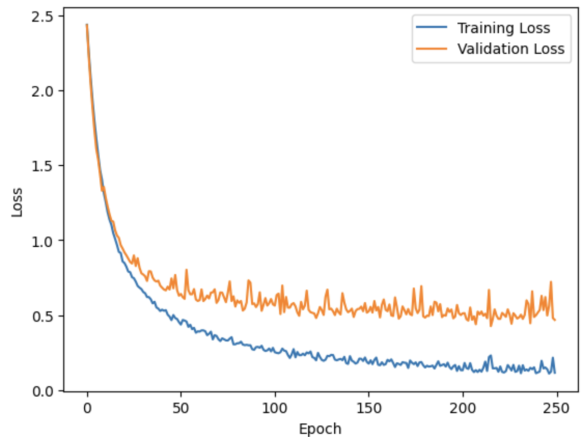




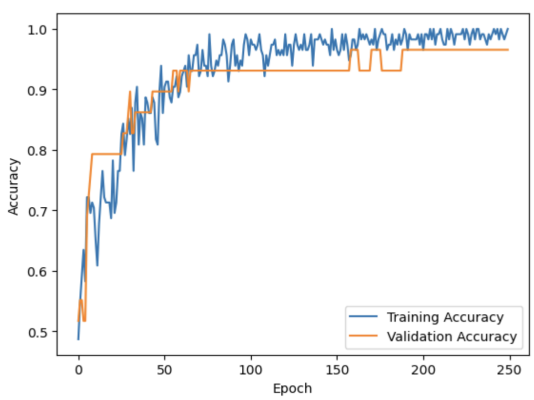
As you can see from the above plots, the model is not performing to the best of its ability and has scope for improvement on these datasets.

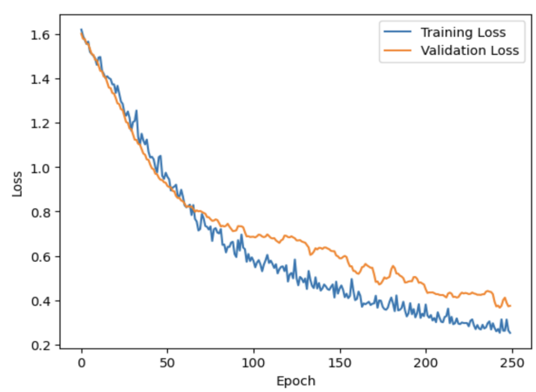
For the Adult dataset, we have used The first dense layer has 512 neurons and uses the ReLU activation function, It also applies L2 regularization with a strength of 0.03 and 5 hidden layers with ReLU activation functions and dropout layers to reduce overfitting. The final dense layer has a single neuron with a sigmoid activation function. To understand the performance of the network we plot the following:





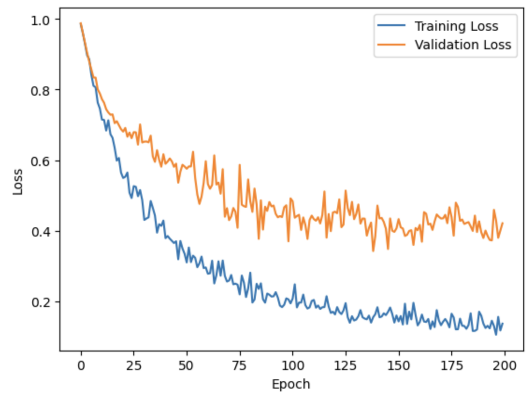
For the Adolescents dataset, we have used a neural network model that comprises an input layer, four hidden layers, and an output layer. The input layer takes the input of the shape of training data which depends on the number of features in the input data. The hidden layers have ReLU activation functions and are followed by dropout layers with a rate of 0.2. The output layer is a single neuron with a sigmoid activation function and outputs a probability value between 0 and 1. To understand the performance of the network we plot the following:



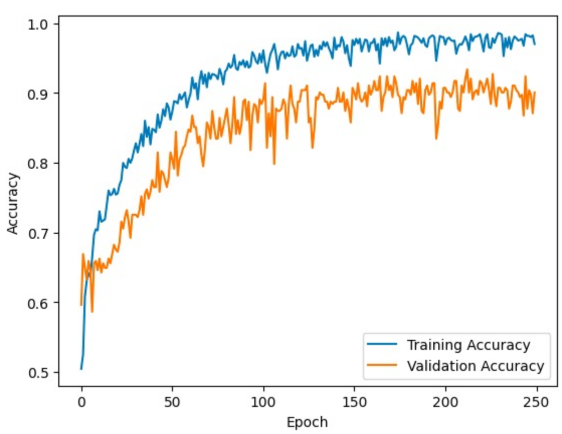


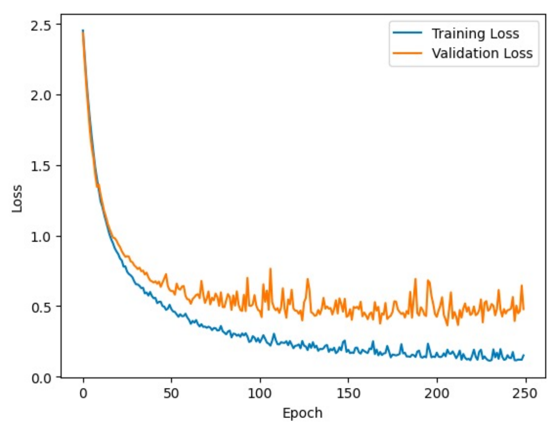
For the Toddlers dataset, the input layer has 32 neurons since it takes shape of the input set, which depends on the number of features in the input data. The first hidden layer has 32 neurons, uses the ReLU activation function, and applies L2 regularization with a strength of

0.01. The subsequent hidden layers consist of 128, 256, and 32 neurons, respectively, and also use the ReLU activation function. All the hidden layers also have a dropout layer after them with a rate of 0.2. The output layer consists of a single neuron with a sigmoid activation function that outputs a probability value between 0 and 1. To understand the performance of the network we plot the following:



For the combined dataset, the first layer is a Dense layer with 512 units, ReLU activation function, and an input dimension that is the same as the combined dataset. The kernel_regularizer argument specifies L2 regularization with a regularization strength of 0.03. There are 8 hidden layers with ReLU activation functions, which are commonly used for deep neural networks, along with dropout layers for regularization to prevent overfitting. The output layer is a Dense layer with 1 unit and a sigmoid activation function. To understand the performance of the network, we plot the following:





Results:

The accuracies and misclassification rates for the neural networks we have built for all 4 datasets are as follows:

	Toddlers	Adolescents	Adults	Combined Age Group
Accuracy (%)	93.88	97.22	96.70	93.38
Misclassification Rate	0.1020	0.0278	0.1020	0.1190

Conclusion

As we can see, the accuracies are higher when we implement the models separately on different age demographics, rather than combining the age groups and combining a master dataset. The misclassification rates are also lower in each dataset and indicate the same. Rather than creating a model where one size fits all, it is better to create separate models in such a scenario. Detection of ASD is important since early detection can lead to earlier intervention and treatment, which can be more effective in improving outcomes for children with ASD. Research has shown that early intervention can improve language and communication skills, social interaction, and adaptive behavior in children with ASD. A more accurate classification for screening can hence reduce the risk of ASD in the future and medical professionals can further inspect these cases if needed.

Future Directions

The present study covers an important aspect of a patient's physical characteristics. However, to enhance the comprehensiveness of the research, it is recommended to incorporate additional physical characteristics such as speech, motor activities, and other relevant physical data. In addition, to build a more robust classifier, a larger and more diverse dataset is required. These steps can be taken to further improve the accuracy and reliability of the results, making the study more valuable to the scientific community.

References

1. Tabtah, 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.

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

3. Thabtah, F. (2017). Machine Learning in Autistic Spectrum Disorder Behavioural Research: A Review. Informatics for Health and Social Care Journal. December, 2017 (in press)

4. A Machine Learning Approach in Autism Spectrum Disorders: From Sensory Processing to Behavior Problems: https://www.frontiersin.org/articles/10.3389/fnmol.2022.889641/full