Early Prediction of Autism Spectrum Disorder in Toddlers

MSc Research Project
Data Analytics

Shivani Vaidya
Student ID: X22210580

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Shivani Prakash Vaidya		
Student ID:	x22210580…		
Programme:	Data Analytics……	**Year:**	**…**2023-2024
Module:	MSc Research Project…		
Supervisor:	Dr. Catherine Mulwa …		
Submission Due Date:	12/08/2024**…**		
Project Title:	Early Prediction of Autism Spectrum Disorder in Toddlers		
Word Count:	…9178……… **Page Count**……: 25		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Shivani Prakash Vaidya …

Date: 12/08/2024…

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Early Prediction of Autism Spectrum Disorder in Toddlers

Shivani Vaidya
x22210580

Abstract

Autism spectrum disorder is one of the most complex neurological development disorders significantly affecting social and communication skills. In this research, deep learning and machine learning techniques are utilized for early prediction of autism spectrum disorder in toddlers, using the Autism Spectrum Quotient screening method for data gathering. The three deep-learning models and one machine-learning model are implemented, with the Multilayer Perceptron, Convolutional Neural Network, and Long Short-Term Memory model exhibiting superiority in terms of accuracy and performance parameters as compared to the machine-learning model, Random Forest. A novel feature of the study is the feature importance analysis carried out using the SHapley Additive exPlanations technique for calculating the individual contribution of certain demographic, genetic, and environmental factors such as ethnicity, gender, and jaundice, in the dataset. The findings of this research are beneficial for the healthcare sector as early and precise prediction can minimize long-term treatment costs and reduce mental stress among toddlers and their families.

Keywords- Autism spectrum disorder, Autism Spectrum Quotient screening method, Multilayer Perceptron model, Convolutional Neural Network model, Long Short- Term Memory model, Random Forest model, Feature important analysis, Shapley additive explanations technique

1 Introduction

Autism spectrum disorder has been ranked as one of the most rapidly growing neurological development disorders. Over the last decade, the diagnosis as well as acceptance rate of autism spectrum disorder has significantly escalated, as observed from the information documented by various organizations like Disease Control and Prevention (Maenner, 2023). Autism spectrum disorder demands costly medical treatment from the initial diagnostic stage, prolonged care, and support services that can be financially burdensome for the families of affected children. Hence the early diagnosis of autism in children plays a vital role in reducing severe health implications and designing affordable intervention programs (Tang et al., 2014). Autism research is pivotal for the deep understanding of life-threatening risk factors associated with this disorder.

Diagnosis of autism is complex as the cognitive and emotional development stages of each child differ depending upon various factors such as age, weight, genetics, intelligence quotient, etc. A detailed comprehensive study of such factors is important for the early detection and construction of support programs, clinical interventions, etc to reduce discrimination, lack of acceptance, and rude behavior towards autistic children.

1.1 Background and Motivation

Autism spectrum disorder creates hurdles in communication, social gatherings, and interactions. Repetitive behaviors and restricted interests can be seen in autistic children. For autism detection, there is no specific clinical test designed globally. In the autism spectrum, different types of autism such as pervasive development disorder, childhood disintegrative disorder, and Asperger syndrome are included. Such a spectrum is collectively known as autism spectrum disorder (Hodges et al., 2020) (Mercadante and Schwartzman, 2006). The symptoms of autism can usually be observed between 18-24 months i.e. at the toddler stage (Tanner and Dounavi, 2021).

Identification of autism at a toddler stage enables children to enhance their social, communication, and interaction skills and avail of specialized educational curriculums designed by taking the unique needs of autistic children into account (Werkhoven et al., 2022). Families of autistic children benefit from appropriate resources, and intervention programs developed for such kids. Accuracy plays a very crucial role in this research as misdiagnosis will lead to inappropriate treatment, mental and physical health of children may get hampered and integrity and trust in healthcare systems will be reduced. Hence the main motivation of this research is early and accurate diagnosis of autism for timely initiation of required treatment. The study aims to increase public awareness and social acceptance regarding this subject.

1.2 Research Question, Research Objectives and Contributions

Traditional methods for autism diagnosis like behavioral analysis performed with the help of video and image data lack precision and require additional processing time. This aspect was taken as the focal point while constructing the research question. In addition, the research question attempted to determine the importance of different environmental and demographic factors in autism spectrum disorder.

RQ: *How well can deep learning techniques predict autism spectrum disorder in toddlers by using the Autism Quotient screening method, and evaluate the importance of various demographic and environmental factors contributing to autism spectrum disorder?*

This is vital so that healthcare professionals and stakeholders will develop early intervention and treatment programs to improve the quality of life of autistic toddlers and their families.

To solve the research question, the objectives presented in Table 1 were implemented.

Table 1: Research Objectives

Serial Number	Objective Description	Methodology Utilized
1	A critical review of the literature on existing autism prediction methods.	A systematic review using scientific journals, published papers
2	Data Collection and Pre-processing	The Autism Spectrum Quotient method
3	Design, implementation, and evaluation of deep learning and machine models	Multilayer Perceptron model, Convolutional Neural Network model, Long Short- Term Memory model, Random Forest model Feature Important analysis model
4	Model Optimization	Hyperparameter tuning
5	Comparison of developed Models	Comparative analysis based on accuracy
6	Comparison of Developed Models versus Existing Models	Comparative analysis based on accuracy

Project Contribution:

The most prominent contribution of this research is the development of a deep learning-based prediction approach for early and accurate detection of autism in toddlers. To improve the efficiency of the predictive system, a deep analysis of environmental, demographic, and genetic factors was performed. The early prediction of autism in toddlers carried out is beneficial in minimizing the inferiority complex amongst autistic children by providing them with appropriate support services, and interventions and improving the quality of life of such children.

This further report document is divided into 7 sections. The first section introduces the topic, background and motivation, research question, objectives, and project contribution. The second section, related work asses the previous work done in that area and critically evaluates it. The third section, methodology, and design specification explain the different stages of methodology through a process flow diagram. In the fifth section, the description of the implementation performed followed by results obtained after implementation are analyzed. The discussion section contains a comparative analysis and discusses the strengths and limitations of the research. While the last section summarizes the findings, and potential implications and provides an outline for future work

2 Related Work

The literature reviewed in this section assesses various machine learning and deep learning techniques utilized for autism prediction. Comparative analysis is conducted to study, compare, and analyze results, performance parameters, and limitations of these techniques and identify research gaps. The most prominent motive of this review is to decide the approach for enhancing the accuracy of autism prediction in toddlers. In addition, this review provides insights into the investigation done in determining the influence of environmental, demographic, and genetic factors that cause autism in children. This related work also throws light on the research conducted in evaluating the advantages of early detection. The literature reviewed is further divided into five sub-sections.

2.1 Critique Review of Traditional Machine Learning Techniques in ASD Prediction for Toddlers

The utilization of traditional machine learning for the prediction of autism has significantly developed since the beginning of 2000 with the rise of the Artificial intelligence era. Five different machine learning models K-Nearest Neighbours, Random Forest Support Vector Machine, Decision Tree, and Gaussian Naive Bayes were implemented by (Islam et al., 2020) on data gathered through standard screening tools for autism prediction. In those scenarios, KNN achieved the highest accuracy of approximately 95%, while random forest gave 93% accurate results. The remaining three models reached up to 90% accuracy. Although the accuracy achieved using KNN was high due to its ability to manage non-linear relationships in data, it is computationally intensive. The random forest model outperforms the remaining models due to accuracy and robustness behavior against overfitting. However, it is not ideal for prediction as demands many computational resources. The unique approach was taken into consideration by (Akter et al., 2019) which deals with data collected from the eye-tracking method which uses specialized clinical equipment for monitoring toddler's eye movements and collects scan paths. The SVM machine learning model was applied to pre-processed data to obtain results. He stated that mean decline was observed in eye fixation of toddlers aged 2-6 months by analyzing evaluation parameters. This approach is relevant when image data is present. However, for successful implementation using this approach, feature selection, extraction, and data transformation should be carried out carefully and in-depth. Along similar lines, research was performed by (Minissi et al., 2022) by analyzing social visual attention data using supervised machine learning algorithms. An accuracy of around 88.51% was achieved using SVM. This study highlighted the need for a more advanced and robust methodology that will enhance efficiency, accuracy, and extensive feature engineering. The research was conducted by (Karpagam and Gomathi alias Rohini, 2022) and utilized a logistic regression model on the children dataset. A mean accuracy of 90% was obtained after using chi-square and information gain techniques to carry feature selection process. The requirement for fine-tuning model parameters for enhancing prediction efficiency is underscored through this analysis. The logistic regression model is not adaptable for handling complex data patterns

hence not a suitable option for handling complex autism data. The video data was processed by (Wu et al., 2021) for Autism prediction in children using a random forest algorithm.
Video frames of the toddler's behavior were converted into features and processed before applying the machine learning model. Even though Random Forest outperformed SVM in terms of accuracy by giving 95% accurate results over SVM which gave 90% accuracy, this approach lacks feasibility and requires highly complex preprocessing and computational costs.

2.2 Critical Review of Deep Learning Techniques in ASD Prediction for Toddlers

Cutting-edge computation power provided by deep learning algorithms is a boon for autism spectrum disorder prediction as it can efficiently discover intricate data patterns while handling large chunks of data. Autism prediction is a complex task that requires extremely accurate results which can be obtained by applying various deep learning models. Mittal et al. (2022) developed an enhanced convolutional neural network model using a facial image dataset collected from the Kaggle website. Techniques were implemented for hyperparameter optimization of the I-CNN model, batch normalization, and dropout. An accuracy of 97% was achieved indicating the high potential for processing complex image data like facial expressions of toddlers. However irrespective of the results obtained, this approach has severe ethical implications as data privacy is very crucial as the identity of the toddler, and family might get leaked. On the grounds of confidentiality and data protection issues, this approach is not ideal. Along similar lines, Poornima and Kousalya (2022) worked on a 2-D CNN model to evaluate the emotional behavior of toddlers with the help of audio data. The extraction of parameters like Mel Frequency Cepstral Coefficient and Mel Spectrograms (Shanmugam and Radhakrishnan, 2024) was done from toddlers' voices and later emotions were categorized into eight different categories. This model outperformed traditional models like SVM, and random forest in terms of accuracy as it gave 96% accuracy. Complex voice patterns and emotional variation were perfectly captured and analyzed using this model. Yet data privacy issue comes into the picture in this case as well. In addition, misdiagnoses are highly possible as introverted or reserved toddlers might identify as autistic. A combination of different deep learning techniques was applied to behavioral data acquired from video analysis performed by (Akter et al., 2023) to predict autism. This approach gave noteworthy results regarding robustness, efficiency, and accuracy. Comparative analysis of this behavioral video data helped in determining the most efficient deep learning model out of CNN, ANN, and RNN architectures. Despite its merits, this approach needs huge, labelled data and lacks in feasibility perspective due to the requirement for a higher dimension clinical setup which is very expensive. A multilayer perceptron model was utilized by (Shanmugam and Radhakrishnan, 2024) for classifying autistic patients based on EEG data. This study proved highly beneficial in distinguishing between autism and other control groups. The accuracy achieved using this approach was 96.20%. However, this approach heavily on labelled training data available for analysis. Hence research niche lies in utilizing deep learning techniques for autism prediction which provides highly accurate results in less time.

2.3 Comparison of Traditional machine learning models vs deep learning Models

It is evident from Table 2 designed, that the comparison of traditional machine learning models with deep learning models shows a clear pattern that states that deep learning techniques consistently give highly accurate results irrespective of the data collection method. The accuracy parameter is crucial for this research as misdiagnosis can cause severe implications and the quality of life of the toddler and his family might be directly impacted in different aspects. Misidentification between multiple developmental disorders and autism can lead to inappropriate treatment which can affect the mental, and physical health of toddlers negatively and will prove inefficient as the treatment provided is not aligned with the specific needs if the correct developmental disorder is not identified.

Table 2: Comparison of traditional machine learning models versus deep learning models

Serial Number	Author	Model Utilized	Data collection technique	Accuracy	Year
1	Islam et al.	KNN	Standardized Screening Tools	95%	2020
2	Islam et al.	Random forest	Standardized Screening Tools	93%	2020
3	Minissi et al.	SVM	Eye tracking	88.51%	2021
4	Karpagam and Rohini,	Logistic regression	Standardized Screening Tools	90%	2022
5	Mittal et al.	I-CNN	Facial Image data	97%	2022
6	Poornima and Kousalya	2-D CNN	Voice data	96%	2022
7	Shanmugam and Radhakrishnan,	MLP	EEG data	96.20%	2024

2. 4 Critique review evaluation of environmental, Demographic, and Genetic factors in ASD in toddlers

To enhance diagnostic accuracy and understand thoroughly the other factors like demographics, genetics, etc contributing to autism, more personalized data collection procedures came into the picture. The review performed (Chaste and Leboyer, 2012) analyzed various genetic studies on autism in toddlers and observed the high heritability in the majority of the studies. In addition, they emphasized finding out the correlation between different environmental factors such as several complications, infections, exposure to toxins during pregnancy, and autism in newborns. Even though this is groundbreaking due to its crucial findings, it's completely relied on secondary data. This research failed to provide new empirical data hence despite being theoretically strong this study is not ideal in terms of feasibility.

(Botsas et al., 2023) performed a similar study by examining the prenatal development periods, the environmental exposures during this phase, and the association between its critical implications and the risk of autism in children. Exposures related to the quality of air, and nutrition levels during the prenatal phase were studied by utilizing longitudinal data collection techniques. Yet the chances of induction of inaccuracies and different biases are possible in this scenario. A comprehensive data analysis performed by (Frazier et al., 2014) proved beneficial in evaluating the relationship between variables such as parental age, ethnical background, and socioeconomic status of the parents, and children. Data was gathered from various cohort studies and different clinical tests conducted on data. He stated that the older the age of the parent, the risk of developing autism in children increases due to the presence of high de novo mutations. This study helped understand the impact of social and economic factors causing autism. Despite that, this approach is not suitable for analyzing large chunks of data .

2.5 Critique Review of Benefits of Early Detection and Cost Reduction

Early diagnosis of autism spectrum disorder has proven favorable not only due to the delivery of multiple development outcomes for children and their families but also because significant cost-savings were achieved owing to early detection. The study performed by (Elder et al., 2017) highlighted the importance of early diagnosis of Autism in the case of children ranging between 1-3 years. The benefits of timely initiation of different treatment programs and their impact on the initial developmental stages of autistic kids were thoroughly analyzed in this research. The results obtained from that study revealed that the quality of life of children below four years of age who received appropriate clinical treatment is far better as compared to those kids who were diagnosed later. On a similar track, (Zhou et al., 2021) the study indicated a reduction in emotional and financial stress that families of autistic children go through as an effect of early detection of autism. Both studies are helpful for an in-depth understanding of the merits of early detection of autism, still, they are not efficient due to the generalizability of their findings.

In this way, this review focused on analyzing different machine learning, and deep learning models applied by different authors for Autism prediction. Results obtained through different models were compared and the limitations of the research studied. In addition, this review takes previous research done evaluating the individual contribution of different genetic, demographic, and genetic factors into consideration. The review throws light on the social, and economic advantages of early detection of autism spectrum disorder. This review is crucial for a detailed analysis of different data collection techniques and methodologies utilized in the existing research for the selection of the most appropriate data collection technique. The research niche is to develop a technically integrated deep learning approach for early and accurate autism prediction.

3 Methodology Approach and Data Pre-Processing

This chapter explains the methodology and design specification utilized in this research. The Knowledge Discovery in Database methodology builds a comprehensive framework for autism prediction in toddlers. The structured approach of methodology facilitates the systematic flow for conducting multiple processes like data acquisition, data pre-processing, etc. Every step of this methodology was meticulously planned and implemented to ensure that the developed autism predictive model was accurate. All the stages of this methodology were integrated to offer valuable insight and helpful for early detection.

The methodology approach shown in Figure 1 outlines each building block of the methodology utilized for predicting autism spectrum disorder in toddlers.

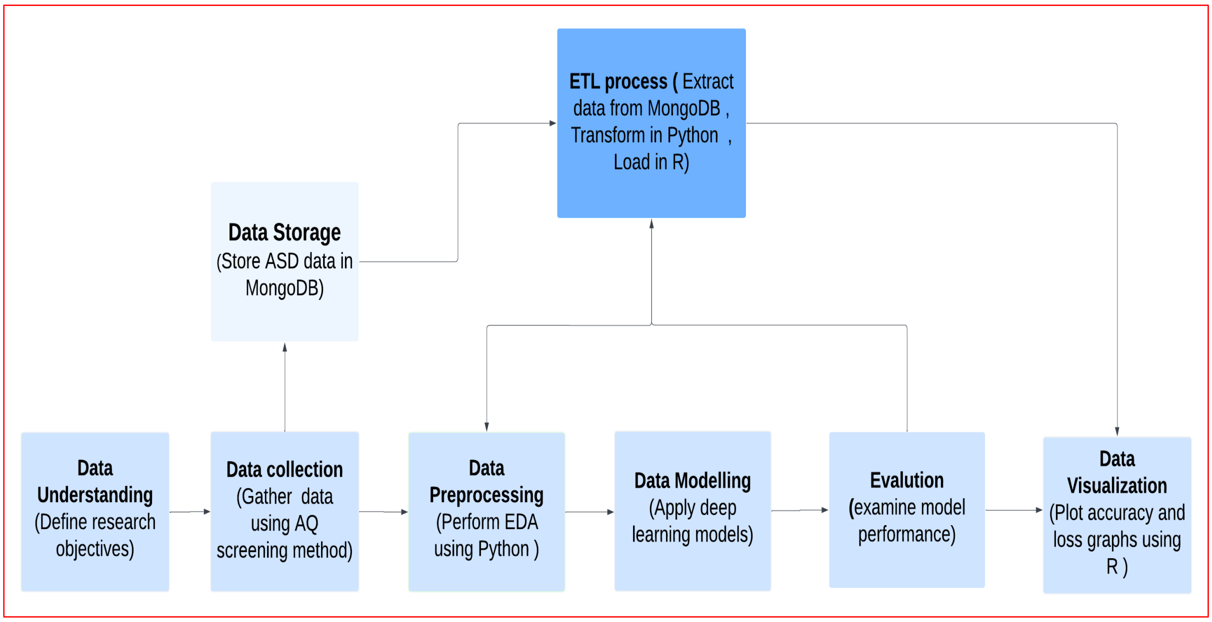


Figure 1 : Methodology Approach

Data understanding: The data understanding phase was vital for establishing a solid foundation for the analysis. The requirement gathering was carried out in this stage which involved selecting metrics for performance evaluation of autism prediction. Different data sources and data collection methods were considered and analyzed in depth in this phase. All data privacy and protection rules and regulations were considered in every step of this methodology. Ethical considerations play a very crucial role in this research.

Data Collection: Gathering relevant and good-quality data for analysis was the most crucial step of this methodology. For data-gathering purposes, the Autism Spectrum Quotient screening method was selected.

The Autism Spectrum Quotient screening method

The Autism Spectrum Quotient screening method was a diagnostic tool developed at the Autism Research Centre in Cambridge UK by clinical psychologist Simon Baron-Cohen and his team (Baron-Cohen et al., 2001).In this screening method, a set of particular questions was designed to evaluate the existence of autistic traits, and patterns in toddlers between ages 1 to 3 years. The questionnaire designed in the screening method examines multiple factors like social, and communication skills, imagination abilities, speaking abilities, etc. Each question is marked on a point 4 scale. In that score, the degree of agreement or disagreement is considered to check the severity of autism in children. The mapping of A1-A10 scores as per the quantitative checklist published by the University of Cambridge for Autism in Toddlers (Q-CHAT) (Allison et al., 2021) is given in Table 3.

Table 3: Q-chat score quantitative checklist

Attribute Name	Mapping to Q-chat Screening Scores Descriptor.
A1_Score	Response of toddler when his/her name is called.
A2_Score	Ease of establishing eye contact with toddlers.
A3_Score	Pointing abilities of the child to express what they want.
A4_Score	Pointing abilities of the child with family to exhibit shared attention.
A5_Score	Engagement of the child in pretend play (talking with toys, caring for dolls)
A6_Score	The ability of the child to follow the direction of another individual's gaze.
A7_Score	The compassion of the child towards a family member who is visibly upset.
A8_Score	The speaking ability of the toddler.
A9_Score	The ability of the child to do simple gestures (e.g. wave goodbye)
A10_Score	Abnormal staring abilities that could indicate disengagement.

The screening method can be administered quickly through a clinical setting or with the help of simple questionnaires making it a suitable choice for this research. A broad range of signs and symptoms that assess communication, eye contact, and gestures can be studied through this screening method. Compared to other data collection methods like video analysis, neuroimaging, and facial analysis it requires fewer clinical resources hence it is cost-efficient compared to other data gathering methods. It safeguarded data privacy aspects like participants' privacy, identity, and confidentiality. Hence, this autism spectrum quotient screening method was the optimal choice for this research (Ruzich et al., 2015).

Data storage: Initially, the autism spectrum disorder dataset was stored in the MongoDB server. MongoDB provided an efficient document-based storage format that was advantageous for storing complex datasets like the autism dataset.

Extract Transform Load: The Extract, Transform, Load process was performed in this research to ensure the data was well prepared for modeling and further analysis. This was one of the most prominent building blocks of this methodology as it assessed the quality and consistency of data thoroughly in the process. Hence it is used in many constructing real-world prediction models. The stored autism data was extracted from the MongoDB server and transformed using Python. Data preprocessing was done followed by data modeling and evaluation of the results was conducted and subsequently data was loaded in R for data visualization. Python provides extensive support and thousands of libraries for data analysis hence utilized for transformation, data modeling, and evaluation in this study. For evaluation of results and data visualization purposes, R language was used due to its powerful statistical analysis capabilities.

Data Preprocessing: In the data preprocessing step data was cleaned initially and transformed to prepare it for further analysis before the application of deep learning techniques. The "ASD extended optimized" dataset from the Data World website was chosen for this research after careful consideration and analysis of data requirements, and the research objectives. This dataset contains 11,705 records and 21 columns. Attributes like A1-A10 score, age, gender, ethnicity, jaundice, and country of residence are present in the dataset. A well-defined rubric was taken into account to understand the mapping of the A1-A10 score values present in the dataset, ranging between 0 to 2. Score 0 indicates the absence of the trait whereas score 1 represents the moderate presence of the autistic trait and score 2 denotes the high presence of the autistic trait.

Variables like A1-A10 are numeric variables whereas gender, ethnicity, jaundice, country of residence, etc are categorical attributes. The selected dataset consists of A1-A10, and Q-chat scores which were captured by thoroughly assessing the behavior, and responses of toddlers in particular situations. In addition, environmental, and demographic attributes like gender, ethnicity, jaundice at birth, a family member with autism, age of a toddler were present in the research dataset. The validity and reliability of the predictions were enhanced by a selection of a diverse dataset that covered multiple aspects to detect autism in toddlers through a careful data collection process. Hence this dataset was considered for research as it perfectly aligned with all the desired criteria for achieving the designed research objectives.

Initially, data was cleaned by handling null and missing values to preserve the data integrity by application of column mean, method. Redundancy in data was minimized by the identification and removal of duplicate records. The Interquartile range method was utilized for the detection and removal of outliers in the autism data.

Removal of outliers was very important in this analysis to ensure accurate statistical analysis. Data was normalized after outliers' removal using StandardScaler. Numerical features in the data like A1-A10 were standardized to achieve a statistical mean of zero along with a standard deviation of one. After data normalization, data encoding was performed to convert categorical variables in the data like gender, country of residence, etc. into more suitable forms for further analysis. The binary target variable Class/ASD was label encoded on the other hand remaining categorical variables were hot encoded to eliminate multicollinearity.

Dimensionality Reduction: To minimize the dimensionality of the data, Principal Component Analysis was implemented. Principal component analysis was advantageous for this research as computational resources were reduced and the speed of the process enhanced automatically enhancing model performance. In this analysis, cleaned data was transformed into 2 principal components.The risk of overfitting was significantly reduced due to principal component analysis.

Modeling: In this stage, deep learning models Multilayer Perceptron, Convolutional Neural Network, and Long Short-Term Memory model along with Random Forest machine learning were applied to transformed data. Python libraries like TensorFlow and Keras were utilized for the implementation of deep learning models. Hyperparameter tuning was performed after the implementation of deep learning techniques to enhance the accuracy of the model. A random Search optimization technique was used for hyperparameter tuning.

Evaluation: The statistical performance and efficiency of the applied deep learning models were critically examined in this evaluation phase. A comprehensive assessment of the deep learning model was done through key performance parameters. The most prominent performance parameter of this research was accuracy. In addition, the other confusion matrix parameters like precision, recall, and F1 score were computed to get efficient evaluation.

Feature important analysis: Feature important analysis was conducted in this study to calculate the importance of every feature in the dataset in autism diagnosis. The SHapley Additive exPlanation technique was chosen for this purpose. The impact of a variety of features on the target variable (Class/ASD) of data was evaluated using this technique.The most prominent benefit of this feature importance analysis was the individual contribution of factors like ethnicity, gender, jaundice, and Q-chat scores were calculated accurately

Data Visualization: Data visualization was carried out using R language due to its effective data visualization and statistical evaluation abilities. During the modeling phase, accuracy and loss that occurred at the end of every epoch were recorded and stored in a CSV file. That CSV file was loaded in R and used for plotting accuracy and loss plots for training and validation data sets.

4 Design Specification

This chapter provides insights into the workflow diagram and all techniques and tool stacks used for the implementation. The complete workflow of the developed autism prediction framework was outlined using the process flow diagram in Figure 2.

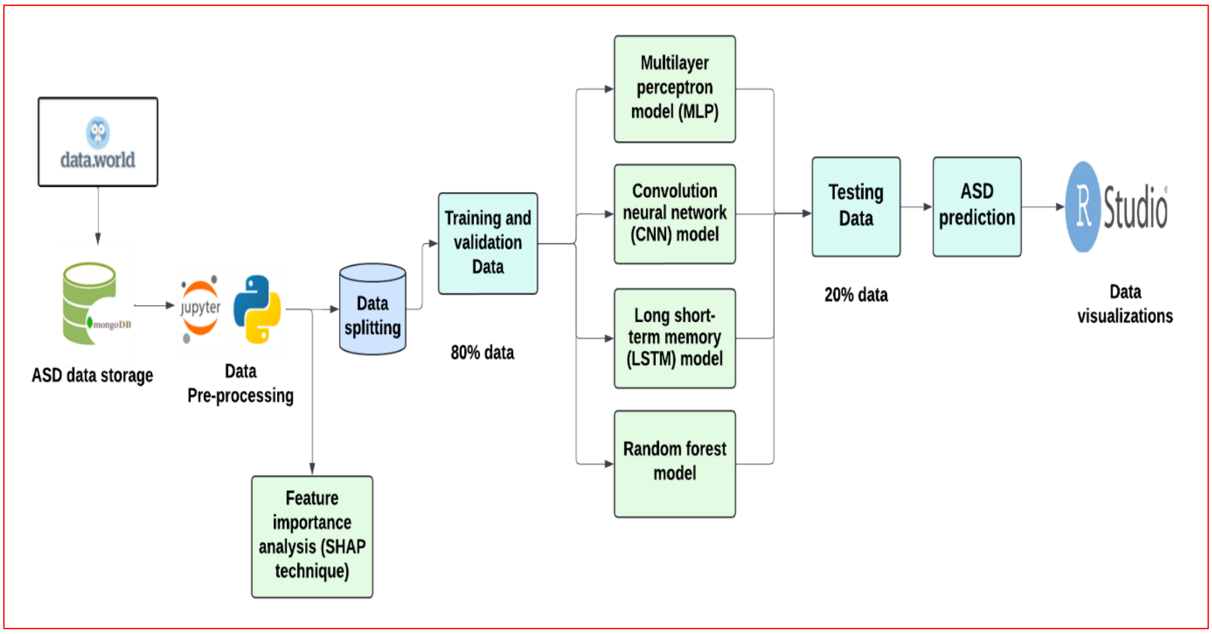


Figure 2: Design Process Flow

Initially, data was collected from the Data World website and stored in the MongoDB server. Then extraction of stored data was done in Jupyter Notebook where data was pre-processed and transformed for modelling. The SHapley Additive exPlanation technique was utilized on the pre-processed data for conducting feature importance analysis. Then pre-processed data was divided into two sets which were training, and testing data sets. From the pre-processed data, 80% of the data was used as training data and 20% data as testing data. Out of total training data, 10% of training data was utilized as validation data. Fine-tuning of the model was done using this validation data. Evaluation of model performance was conducted iteratively using validation data. Data splitting was performed to build a more reliable, robust deep learning framework. The applicability of the constructed model in real time was enhanced due to data splitting as it evaluates the performance of deep learning architectures on unseen data. Complex deep learning architectures Multilayer perceptron, Convolution neural networks, and Long short-term memory models were applied to training and validation data. In addition, the random forest model which is a machine learning model was implemented on training data. The statistical performance of the trained models was evaluated with the help of testing data for autism prediction in toddlers. In the end, the data was loaded into R studio for data visualization.

5 Implementation, Evaluation, and Results of Deep Learning and Machine Learning Models

This chapter describes the implementation, and evaluation of results obtained by application of deep learning models. In this section, three deep-learning architectures Multiplayer perceptron, Convolution Neural network, Long Short -Term Memory, and Random Forest machine learning model were implemented. Results obtained using deep learning techniques and machine learning models were compared in this phase. The results obtained through feature importance analysis were thoroughly analyzed in this stage.

5.1 Experiment 1 - Multilayer Perceptron Model

For solving complex classification problems like autism prediction, the multi-layer perceptron classifier is beneficial as it belongs to the feed-forward artificial neural network class. This model contains input, output, and one or more hidden layers. Due to efficiency in handling complex data relationships by acquiring weights as well as biases that reduce error amongst actual and predicted results, the MLP model was considered for this research. In addition, highly accurate results are generated by the implementation of this model. A total of four layers were present in the MLP architecture used in this scenario. The two hidden layers were incorporated into this MLP architecture. The pre-processed features of the autism dataset like A1-A10 scores, age, gender, etc. were received by the input layer. The number of neurons in the input layer was correlated with the total number of features obtained after data transformation. The input layer acted as an entry point for this MLP architecture. The first hidden layer in this architecture was composed of 50 neurons on the other hand second layer was comprised of 100 neurons. In both layers 'tanh' or 'relu' activation functions were used and with the help of hyperparameter tuning, intricate patterns in the data were captured. Only a single neuron was present in the output layer of the architecture along with the sigmoid activation function. The sigmoid function was given probability values between 0 and 1 suggesting the likelihood of autism in toddlers. Hyperparameter tuning was performed using the random search method and played a vital role in the optimization of the MLP model by exploring different hyperparameter combinations like the number of neurons, learning rates, activation functions, etc. The best hyperparameters were selected for training the MLP model. Adam optimizer was utilized for this purpose. In this case, the MLP model was trained with 200 epochs. Model performance was boosted with an increase in the number of epochs.

The most crucial parameter of the analysis, which was accuracy in the scenario was 96.71%. The precision and recall achieved in the case of class 0 was 0.98 which shows that 98% of cases predicted were correct. Similarly, 98% of non-ASD class instances were recognized correctly. On the other hand, precision and recall for class 1 were 0.93 and 0.94 respectively. In addition, f1 score values were 0.94 and 0.98 for ASD and non-ASD classes showing that the number of false positive values was less. The reliability of predictions was increased due to the high F1 score value. The values of true positives, true negatives, false positives, and false negatives were showcased using the confusion matrix in Figure 3.

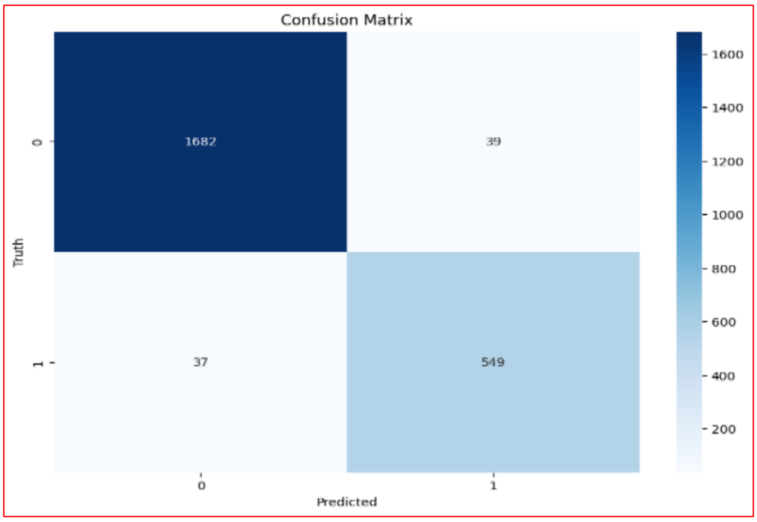


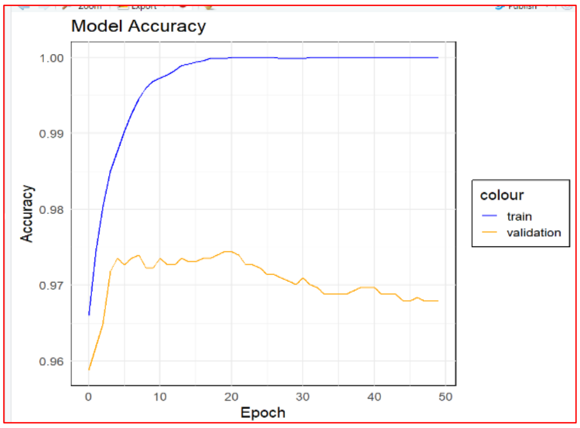
Figure 3: Confusion Matrix

To evaluate model performance confusion matrix. The confusion matrix parameters like accuracy, precision, and F1 score for each class were computed and shown using a classification report provided in Figure 4.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.98      0.98      1721
           1       0.93      0.94      0.94       586

    accuracy                           0.97      2307
   macro avg       0.96      0.96      0.96      2307
weighted avg       0.97      0.97      0.97      2307


Accuracy Score of the MLP model:
Accuracy: 96.71%
```

Figure 4: Classification Report

The accuracy and lost plots for training and validation data were plotted in R for 50 epochs to visualize and interpret results as represented in Figure 5 and Figure 6 respectively. The accuracy graph indicated that the model's accuracy for training and validation datasets increased with an increasing number of epochs. In the case of validation data, the initial rise was seen and then it began to stabilize and slight fluctuations were observed after 10 epochs. It was evident from the graph that the model fits well for training data and further training cannot significantly enhance its generalization ability. On the other hand, the difference between actual and predicted values was studied thoroughly with the help of a loss graph. For training data, a steady decrease was observed whereas for validation data an initial decline was experienced before being stabilized and then slight variations were seen. It can be concluded from the loss graph that further training will cause overfitting.



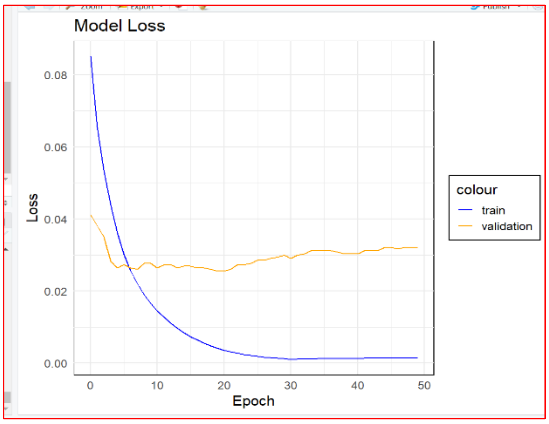


Figure 5: Accuracy plot for training and validation data

Figure 6: Loss plot for training and validation data

5.2 Experiment - Convolutional Neural Network Model

Convolutional Neural Network was used for autism prediction. It had several layers like the input layer, convolution layer, pooling layer, dropout layer, flatten layer, and two dense layers. Pre-processed features from the data were fetched to the input layer. In this case, data reshaping was done to fit the data as per input requirements like samples, features, and time steps of which Conv1D layer. The convolution layer had a kernel of size two, 64 filters, and a ReLu activation function. Spatial hierarchies in the data were recognized through this layer. Pooling layer down-sampled input presentation, reduced dimensionality, and computation load while keeping key features. The pool size of this layer was 2. The overfitting was prevented by introducing a dropout layer with a 0.5 rate. The flattened layer made pooled features compatible with a dense layer which were fully connected by transforming them into a single vector. Out of two dense layers, the first layer had 50 neurons along with a ReLU activation function for learning complex patterns while the second layer had 2 neurons and a SoftMax activation function which gave the probability distribution of the output classes according to the number of categories present in the target attribute. This CNN model was implemented using a categorical cross-entropy loss function and Adam optimizer. This model was preferred in the case of multi-class classification. Hyperparameter tuning was performed, and ideal parameters were identified by training model for 50 epochs with a batch size of 32.

In this scenario, the precision, recall, and f1 score values for the non-ASD class (0) were 0.98 which highlights very a small number of false positives, and false negatives in non-ASD cases. Similarly for the ASD class, the precision and recall values were 0.96 and the f1 score was 0.95. It can be stated that the model detected ASD cases with great accuracy while weighing trade-offs between recall and precision values. The combined accuracy obtained using this model was 97.23%. High accuracy was achieved by the model showing the robustness and efficiency of the applied CNN model.

The evaluation matrix was obtained using a classification report as shown in Figure 7.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.98      0.98      1721
           1       0.94      0.96      0.95       586

    accuracy                           0.97      2307
   macro avg       0.96      0.97      0.96      2307
weighted avg       0.97      0.97      0.97      2307


Accuracy Score of the CNN model :
Accuracy: 97.23%
```

Figure 7: Classification Report

The accuracy and loss plots for the CNN model are represented in Figure 8 and Figure 9 respectively. The accuracy graph plotted for 50 epochs depicts that the training accuracy represented by a blue line was elevated consistently while validation accuracy shown by the orange line increased but with slight variations and then stabilized at around 97%. Robust generalization abilities of the model were demonstrated using this model. In the model loss graph training loss declined abruptly before flattening showing reduced error. Similarly, validation loss also decreased but with a higher variability, and then it became stable indicating convergence of the model.

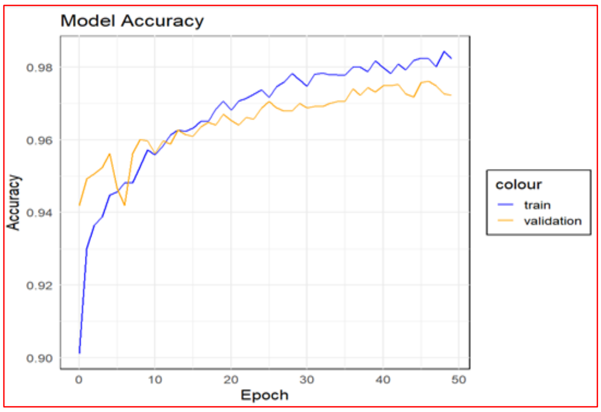


Figure 8: Accuracy plot for training and validation data

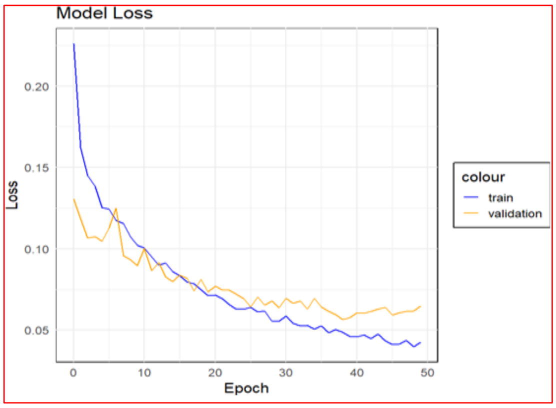


Figure 9: Loss plot for training and validation data

5.3 Experiment - Long short-term memory model

The long short-term memory model with 4 layers was utilized for autism prediction. The first layer had 50 neurons for capturing dependencies in data. For the implementation of the LSTM model 3D input was needed. Therefore, reshaping was performed accordingly. The second layer, which was the dropout layer, had a dropout rate of around 0.2 to prevent overfitting by assigning a specific fraction of input units to 0 during the training phase. The LSTM layer was the most crucial layer of this architecture which had 50 LSTM units for retaining data and information over long periods. Thus, it is considered the ideal choice was capture temporal dependencies in sequential data. The last layer, which was the dense layer had a single neuron along with a sigmoid function. Hyperparameter optimization was carried out to determine the best parameters. Initially, 20 epochs were used for the training model with the help of optimal settings. Adam optimizer and binary cross-entropy loss function were used for model compilation. As the problem was a binary classification problem the binary cross entropy function was taken into consideration.

The evaluation matrix was computed for the LSTM model. The precision obtained for the non-ASD class (0) was 0.97 whereas that for the ASD class was 0.95. It was evident that the model correctly predicted the high percentage of ASD classes without misidentifying the non-ASD classes. Similarly, the recall values for non-ASD and ASD classes were 0.98 and 0.91 respectively. In this case as recall value for the ASD class was less suggesting that a few ASD classes were misclassified as non-ASD classes. However high recall values of the non-ASD class showed that the model was efficient in recognition of true negatives. The F1 scores calculated for non-ASD and ASD classes were 0.98 and 0.93 respectively which expressed the balanced nature of the matrix in effective dealing with the trade-off that took place between recall and precision. The overall accuracy achieved by the LSTM model was 96.49%. The high predictive performance of the LSTM model was denoted by high accuracy.

The classification report for the LSTM model is shown in Figure 10.

```
Classification Report:
              precision    recall  f1-score   support

         0.0       0.97      0.98      0.98      1721
         1.0       0.95      0.91      0.93       586

    accuracy                           0.96      2307
   macro avg       0.96      0.95      0.95      2307
weighted avg       0.96      0.96      0.96      2307


Accuracy Score of the LSTM model :
Accuracy: 96.49%
```

Figure 10: Classification Report

The accuracy graph and loss graphs were plotted in Figure 11 and Figure 12 for the LSTM model for 20 epochs. The training accuracy was elevated initially in a rapid manner and then stabilized at 98% after the 10^{th} epoch. On the other hand, validation accuracy improved quickly and leveled off slightly below the training accuracy. The consistent performance of the model for both training and validation data showed the model's effectiveness. The loss graph for both training and validation data depicted a sharp decline for epochs highlighting that the model was quick in reducing error for both the sets. After the fifth epoch, the loss was stabilized at a relatively low value showcasing further advancement.

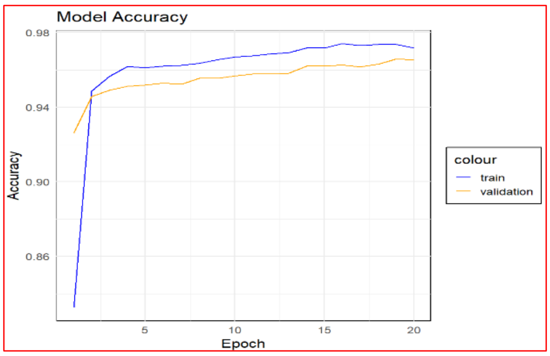


Figure 11: Accuracy plot for training and validation data

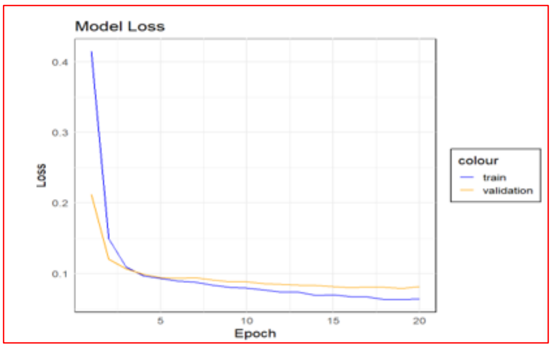


Figure 12: Loss plot for training and validation data

5.4 Experiment - Random Forest Model

The machine learning model was implemented for autism prediction and comparative analysis was done to thoroughly analyze and compare results obtained using deep learning techniques and machine learning techniques. A random forest classifier was applied to the pre-processed data. One hot encoding was performed to convert the target variable into binary format. After that data splitting was done in which 80% data was utilized as training data and 20% data as testing data. Training data was used to train the model while testing data was used for making predictions. A confusion matrix was printed to evaluate the matrix.

The overall accuracy achieved by the model was 85.05%. In the case of this machine learning method, data quality issues like noise, data imbalance, insufficient training data, and inability to handle and capture complex data patterns were the most prominent reasons behind the low accuracy achieved random forest model. The recall value for the positive class was very low which was 0.45, indicating that the model lacked in correctly identifying positive cases. In addition, the confusion matrix showed that the number of false positives and false negatives was relatively high in number indicating high chances of misdiagnosis that can give rise to severe issues like unwanted stress among families and toddlers, and loss of faith regarding healthcare and clinical practices. The imbalance in the dataset was visible and denoted the overfitting issue for specific features that don't perform well for unseen data.

The confusion matrix for the random forest model is shown in Figure 13.

```
Classification Report:
              precision    recall  f1-score   support

         0.0       0.84      0.99      0.91      1721
         1.0       0.93      0.45      0.60       586

    accuracy                           0.85      2307
   macro avg       0.88      0.72      0.76      2307
weighted avg       0.86      0.85      0.83      2307


Accuracy Score of the Random Forest model: 85.05%

Confusion Matrix:
[[1701   20]
 [ 325  261]]
```

Figure 13: Classification Report

5.5 Experiment – Feature Importance Analysis using SHapley Additive exPlanation technique

The feature important analysis performed using the SHapley Additive exPlanation technique enabled healthcare providers to enhance autism screening tools, and intervention programs and allocate resources sensibly. The most influential factors in the dataset contributing to autism were determined using this analysis. While carrying out this study all the ethical aspects were taken into consideration. Sensitive and discriminating remarks against any community, ethnic group, gender, etc were avoided instead the main crux of this study was present in assessing individual contributions of all the features involved in the dataset like A1-A10 score, gender, age, jaundice, etc. Mean absolute SHAP values and percentage contribution of each feature of the dataset were calculated for evaluation of the contribution of individual features of the dataset. Along with direct effects and exchanges with other features SHAP values determined the influence of each feature using mean absolute values. The average impact of each feature was calculated by mean absolute values while the relative importance of each feature was represented by comparing the SHAP value of every feature with the total SHAP value in the percentage calculation.

In this analysis, ethnicity turned out to be the most prominent by contributing 31.74% to the autism prediction. The age factor was the second most impactful in this analysis with a 5.95% contribution followed by behavioral features A1-A10 contributing to the range of 2% to 6 %. Out of these behavioral factors, A9 was the most contributing factor with a score of 5.53% on the other hand A8 score was the least contributing factor with a score of 2.69%, Genetic factors like the male gender contributed 4.58% to the overall predictions. The self-relation feature contributed 4.60 % and other factors like family history, and jaundice gave a contribution in the range of 0.5% to 2 %. The feature important plot is shown in Figure 13.

Feature	Importance	Percentage
ethnicity	0.248731	31.742884
age	0.046616	5.949151
A9_Score	0.043300	5.525928
A3_Score	0.043123	5.503274
A6_Score	0.042584	5.434511
A5_Score	0.041993	5.359094
A4_Score	0.036627	4.674369
relation_Self	0.036042	4.599649
gender_m	0.035854	4.575607
A7_Score	0.029943	3.821277
A1_Score	0.027696	3.534588
A10_Score	0.027361	3.491771
A2_Score	0.025247	3.222046
A8_Score	0.021050	2.686409
austim_yes	0.016620	2.121008
jundice_yes	0.015636	1.995512
relation_Parent	0.013049	1.665311
relation_Relative	0.010488	1.338460
relation_Health care professional	0.010366	1.322925
used_app_before_yes	0.006872	0.876941
relation_Others	0.004382	0.559284

Figure 14: Feature Importance Plot

The feature importance vs percentage contribution bar chart is provided in Figure 13. The bar chart was plotted to represent the percentage contribution of features based on SHAP values. On the x-axis, features were plotted while percentage contribution was plotted on the y-axis. A clear overview of the individual contribution of all features was provided through this visual representation.

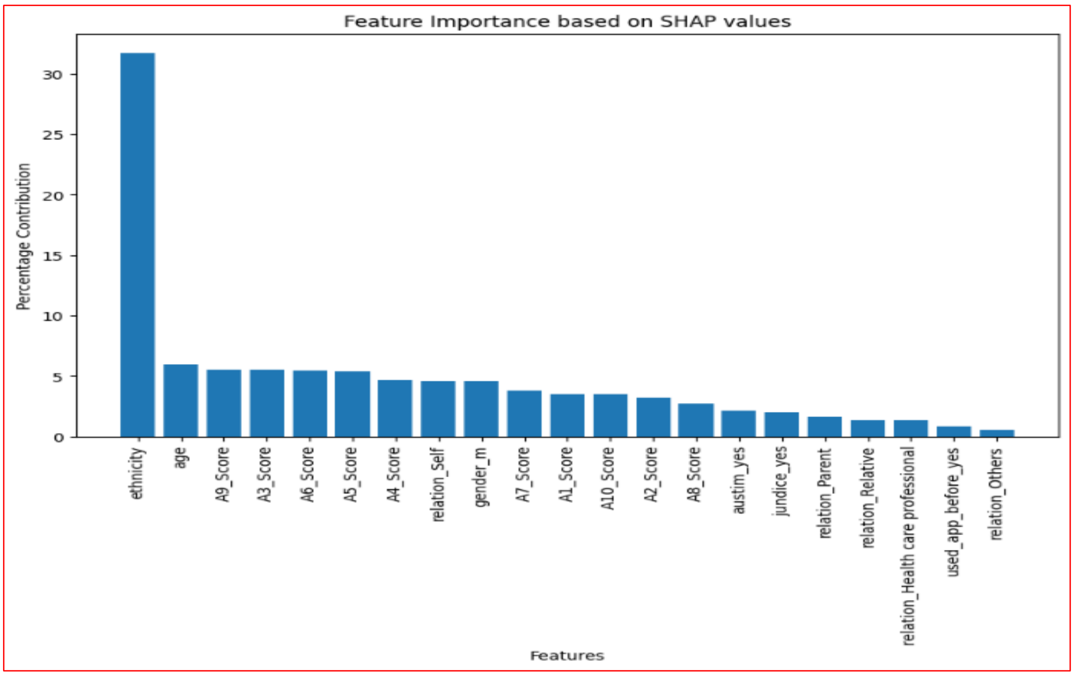


Figure 15: Feature Importance Bar chart

6 Discussion and Comparisons

This chapter includes comparisons performed to determine the most accurate model among applied deep learning and machine learning models and the most accurate model between developed and existing models respectively. The strengths and limitations of this research are also discussed in this section.

6.1 Discussion and Comparison of Developed Deep Learning versus Machine Learning Model

The comparative analysis was conducted to understand the efficiency of all applied techniques and to evaluate the superior techniques applied in this research. Among the applied models the CNN model gave the highest accuracy of around 97.23% followed by MLP and LSTM models which were 96.79% and 96.49% respectively. The highly accurate results delivered by the deep learning techniques indicate the resilience and efficiency of the deep learning techniques in the prediction of autism in toddlers. The results obtained using these deep learning techniques surpassed the results achieved using traditional machine learning techniques. The random forest model, a traditional machine-learning model, had an accuracy of 85.05% which is much less compared to the deep-learning techniques applied.

A noteworthy difference between the accuracy obtained by deep learning models and that of accuracy achieved by machine learning models suggests that machine learning models like random forest are not efficient in handling complex data analytical tasks like autism prediction. This comparative analysis is shown in Figure 5.

Table 5: Comparison of developed deep learning model versus machine learning model

Model Name	Accuracy
Multiplayer Perceptron Model	96.71%
Convolution Neural Network Model	97.23%
Long Short-Term Memory Model	96.49%
Random Forest Model	85.05%

6.2 Discussion and Comparison of Developed Deep Learning and Machine Learning Model versus Existing Models

The comparative analysis carried out was crucial for determining the merits of developed deep learning and machine learning models with the existing models. For evaluation of the generalizability of the developed models, distinct datasets and extraction methods were utilized. It can be stated from the comparison that the Multilayer Perceptron and Convolutional

Neural Network model developed in this research outperformed the existing versions. The Multilayer Perceptron model achieved an accuracy of 96.79% which was 96.20% in the existing model. Similarly, the Convolution neural network model showed 97.23 % accuracy as compared to 97% accuracy obtained through the existing Convolution Neural Network model. The random forest model underperformed in this research by giving an accuracy of only 85.05% which was around 93% in the earlier case. The comparison is given in Table 6.

Table 6: Comparison of developed model accuracy versus existing model accuracy

Model Name	Developed Model Accuracy	Existing Model Accuracy
Multiplayer Perceptron Model	96.71%	96.20%
Convolution Neural Network Model (1-D)	97.23%	97%
Random Forest Model	85.05%	93%

High precision was achieved in this study due to the use of deep learning techniques while the confidentiality of the toddlers and their respective families was maintained, and the risk of misdiagnosis was drastically reduced. Highly accurate results were obtained in this research in less computation time. The feature important analysis performed using SHAP added novelty to this study as the traditional research on this topic dealt with finding the contribution of behavior causing ASD. The feature importance analysis done in this study was successful in determining the individual contribution of demographic, environmental, and genetic factors like ethnicity, age, gender, and jaundice present in the data. The clarity and interpretability offered by feature important analysis will be advantageous for healthcare professionals in understanding the impact of different factors on autism. The study's outcomes provide clear evidence of the high predictive performance of the implemented deep learning techniques.

Although the study is superior in several aspects, the consideration of limitations of the study is needed. The primary drawback of the research is the absence of generalizability as the selected data set specifically captured data regarding a particular demographic population. Additionally, the volume of the data taken into account was insufficient to strengthen the findings of the research. Secondly, the more computational resources used for the optimization of deep learning models can create accessibility concerns in rural areas.

7 Conclusion and Future Work

The main motive of the research was to address the research question of determining how effectively deep learning techniques can be utilized for the early detection of autism spectrum disorder in toddlers along with the evaluation of the individual contribution of demographic, environmental, and genetic factors in the data. It can be concluded from the evaluation of the results obtained that the research question has been fully addressed in this research. All the designed research objectives were implemented successfully. It is evident that all the implemented deep learning models showcased highly accurate results with the Convolution Neural Network stated as the most accurate deep learning model with an accuracy of 97.23%. While the Random Forest model proved to be the least accurate model with an accuracy of 82.05%. The significant difference between the predictive performance of deep learning and machine learning models for analyzing complex autism data was highlighted in this research through the comparative analysis. Similarly, the superiority of developed models to that of existing models in terms of accuracy was explained. The low false positive values for all applied deep-learning models indicate the reliability of the developed models used for autism spectrum disorder prediction in toddlers. The most prominent aspect of this research is the feature importance analysis carried out which plays a vital role in evaluating the individual contribution of different demographic, genetic, and environmental factors in the data. More efficient and customized care can be provided to autistic toddlers suing feature important analysis. It can be concluded from the results obtained through feature importance analysis that ethnicity is the most contributing factor to autism in toddlers contributing 31.74%. In addition, age, gender, and jaundice such factors can be considered notable contributors to the prediction of autism along with the behavioral A1-A10 scores obtained through the autism quotient screening method. However large, more diverse, and complex datasets, advanced methodology is necessary to understand thoroughly the broader applicability of the results obtained through this study.

More complex, large, and diverse data should be taken into consideration in the future to enhance the generalizability of the predictions. Potential bias in the data should be minimized by taking various populations and participants of different socio-economic backgrounds into account. Future research should concentrate on developing an advanced deep learning-based framework that will provide optimized and highly accurate results and will predict autism across all age groups. In the future, results obtained through this AI-driven approach should be cross-verified or assessed by collaborating with healthcare practitioners in domains like genetics, psychology, gynaecology, etc. Future research should be done to develop cost-efficient and extremely accurate AI-driven diagnostic tools for autism spectrum disorder.

Acknowledgment

I want to extend my sincere appreciation to the National College of Ireland and the for the guidance and assistance. I want to express my heartfelt gratitude to my supervisor, Dr. Catherine Mulwa for her unwavering support and guidance which was extremely beneficial in enhancing my research project.

References

Akter, S., Shahriar, H., Cuzzocrea, A., 2023. Autism Disease Detection Using Transfer Learning Techniques: Performance Comparison Between Central Processing Unit vs Graphics Processing Unit Functions for Neural Networks, in: 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC). Presented at the 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE, Torino, Italy, pp. 1084–1092. https://doi.org/10.1109/COMPSAC57700.2023.00164

Akter, T., Satu, M., Khan, Md.I., Ali, M., Uddin, S., Lio, P., Quinn, J., Moni, M.A., 2019. Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders. IEEE Access PP. https://doi.org/10.1109/ACCESS.2019.2952609

Allison, C., Matthews, F.E., Ruta, L., Pasco, G., Soufer, R., Brayne, C., Charman, T., Baron-Cohen, S., 2021. Quantitative Checklist for Autism in Toddlers (Q-CHAT). A population screening study with follow-up: the case for multiple time-point screening for autism. BMJ Paediatr Open 5, e000700. https://doi.org/10.1136/bmjpo-2020-000700

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., n.d. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians.

Botsas, G., Koidou, E., Chatzinikolaou, K., Grouios, G., 2023. Environmental Influences on Individuals with Autistic Spectrum Disorders with Special Emphasis on Seasonality: An Overview. Children 10, 1851. https://doi.org/10.3390/children10121851

Chaste, P., Leboyer, M., 2012. Autism risk factors: genes, environment, and gene-environment interactions. Dialogues Clin Neurosci 14, 281–292.

Elder, J.H., Kreider, C.M., Brasher, S.N., Ansell, M., 2017. Clinical impact of early diagnosis of autism on the prognosis and parent–child relationships. Psychol Res Behav Manag 10, 283–292. https://doi.org/10.2147/PRBM.S117499

Frazier, T.W., Youngstrom, E.A., Embacher, R., Hardan, A.Y., Constantino, J.N., Law, P., Findling, R.L., Eng, C., 2014. Demographic and clinical correlates of autism symptom domains and autism spectrum diagnosis. Autism 18, 571–582. https://doi.org/10.1177/1362361313481506

Hodges, H., Fealko, C., Soares, N., 2020. Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. Transl Pediatr 9, S55–S65. https://doi.org/10.21037/tp.2019.09.09

Islam, S., Akter, T., Zakir, S., Sabreen, S., Hossain, M.I., 2020. Autism Spectrum Disorder Detection in Toddlers for Early Diagnosis Using Machine Learning, in: 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). Presented at the 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), IEEE, Gold Coast, Australia, pp. 1–6. https://doi.org/10.1109/CSDE50874.2020.9411531

Karpagam, C., Gomathi alias Rohini, S., 2022. An Experiment on Logistic Regression Analysis to Detect Autism Spectrum Disorder, in: 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT). Presented at the 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), pp. 1–4. https://doi.org/10.1109/ICEEICT53079.2022.9768505

Maenner, M.J., 2023. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. MMWR Surveill Summ 72. https://doi.org/10.15585/mmwr.ss7202a1

Minissi, M.E., Chicchi Giglioli, I.A., Mantovani, F., Alcañiz Raya, M., 2022. Assessment of the Autism Spectrum Disorder Based on Machine Learning and Social Visual Attention: A Systematic Review. J Autism Dev Disord 52, 2187–2202. https://doi.org/10.1007/s10803-021-05106-5

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., Baron-Cohen, S., 2015. Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Molecular Autism 6, 2. https://doi.org/10.1186/2040-2392-6-2

Shanmugam, S., Radhakrishnan, M., 2024. Autism Classification and Identification of Significant Brain Lobe Using Cepstral Coefficients. Traitement du Signal 41, 23–34. https://doi.org/10.18280/ts.410102

Tang, F., McMahan, R.P., Allen, T.T., 2014. Development of a low-cost tactile sleeve for autism intervention, in: 2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) Proceedings. Presented at the 2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) Proceedings, pp. 35–40. https://doi.org/10.1109/HAVE.2014.6954328

Tanner, A., Dounavi, K., 2021. The Emergence of Autism Symptoms Prior to 18 Months of Age: A Systematic Literature Review. J Autism Dev Disord 51, 973–993. https://doi.org/10.1007/s10803-020-04618-w

Werkhoven, S., Anderson, J.H., Robeyns, I.A.M., 2022. Who benefits from diagnostic labels for developmental disorders? Dev Med Child Neurol 64, 944–949. https://doi.org/10.1111/dmcn.15177

Wu, C., Liaqat, S., Helvaci, H., Chcung, S.S., Chuah, C.-N., Ozonoff, S., Young, G., 2021. Machine Learning Based Autism Spectrum Disorder Detection from Videos, in: 2020 IEEE International Conference on E-Health Networking, Application & Services (HEALTHCOM). Presented at the 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM), IEEE, Shenzhen, China, pp. 1–6. https://doi.org/10.1109/HEALTHCOM49281.2021.9398924

Zhou, W., Wu, K., Chen, S., Liu, D., Xu, H., Xiong, X., 2021. Effect of Time Interval From Diagnosis to Treatment on Economic Burden in Families of Children With Autism Spectrum Disorder. Front Psychiatry 12, 679542. https://doi.org/10.3389/fpsyt.2021.679542