**Diagnosis of Autism with Optimized Machine Learning Model: EDA of Genetic and Personal Characteristics Dataset**

**Onyedikachi A. Erete**

**Student ID: B00858757**

**School of Computing, Engineering and Intelligent Systems**

**Supervisor: Dr Saugat Bhattacharyya**

**Magee Campus**

***Abstract-*** The study seeks to establish the use of an appropriate machine learning model to detect autism among patients. Autism disorder is a condition that limits the cognitive functionalities of an individual. Detection of this disorder is usually not fast and early enough, creating many challenges for sufferers. Developing machine learning models appropriate for the detection of autism is important. The data was extracted from the Kaggle website. Python programming language was used for the analysis, modelling and deployment. “Logistic Regression”, “Support Vector Machine (SVM)”, “Gaussian Naive Bayes”, Extra Trees Classifier, Random Forest, Decision Tree, LightGBM, XGBoost, and CatBoost algorithms” were used for modelling. The AUC score was used to determine the accuracy of probability while the confusion matrix was used to ascertain the classification model for the study. In this case, we go with Random Forest Classifier with a classification accuracy of 90% after optimization and a 0.925731 AUC Score. The Random Forest Classifier can hence be used for autism prediction for patients.

**Keywords- autism, machine learning, patients, models**

**I. INTRODUCTION**

1. *Problem Statement*

Autism Spectrum Disorder (ASD) (often known as "autism") is a mental condition that limits a person's linguistic, cognitive, and social abilities. With a prevalence incidence of 1.85 per cent, ASD affects around 1% of the global population. Autism’s aetiology or cure is relatively unknown. Hence, parents frequently struggle to determine whether their child has ASD, making early diagnosis of an autistic child extremely challenging. Because the signs of autism change as a child grows, test diagnoses on children aged 2-3 years are not as trustworthy as diagnostic tests on children aged 4-5 years. The worrying element is that an autistic person's ability to reach life milestones is dependent on early diagnosis. [1] It is influenced mostly by a mix of genetic and environmental variables. Each autistic person has distinct abilities and flaws due to autism's spectrum nature. Autism ranges from highly proficient to profoundly hindered learning, reasoning, and problem-solving. [1]

High-quality early intervention has been shown in studies to increase learning, communication, social skills, and underlying brain development. However, the diagnostic process can take years. [2]

Early diagnosis should let the patient's family take proactive, effective steps to ensure normalcy. It can help healthcare professionals and patients' families finance required treatment and therapy, reducing delayed diagnosis costs. Autism disorder is a communication and behaviour disorder which affects children and adults. A mix of hereditary and environmental factors plays a major role. [2]

Autism is a spectrum illness, as a result, each autistic individual has a unique set of abilities and limitations. Research shows that brain development and associated improvement can be done via the intervention of high-quality care. Despite this, the diagnostic procedure can take years. [3] Many times, the use of a questionnaire and behavioural exams have been used in previous times to diagnose autism in learners, but this has been time-consuming and costly for parents and clinicians. It is important to look for alternative measures that would help in early diagnosis to remediate and deal with autism among children. Using machine-learning algorithms, this work seeks to make autism diagnosis a faster and more accurate procedure. It will allow therapy to be delivered quickly before the conditions begin to worsen. This study then seeks to predict autism using various machine learning models and compare results using AUC ROC. It will then pick the model that best captures the rare cases of those with autism, the model with the highest recall rate (sensitivity) in classifying the rare cases out of the models used. The uniqueness of this study would be in further employing the use of Streamlit sharing to deploy the developed model.

1. *Identified Dataset*

Data was extracted from Kaggle via the link  <https://www.kaggle.com/c/autism-prediction/data>. The survey responses of more than 700 individuals who participated in the study by filling out an application form are included in the dataset. Because there are labels indicating whether a person has been given a diagnosis of autism or not, machine learning models can estimate the chance of a person having autism. This enables healthcare practitioners to focus their resources. The dataset includes the train and test data all in CSV format. [3]

The data extracted that will be used for the study contains the following information: “ID - ID of the patient.; A1\_Score to A10\_Score –“Score on Autism Spectrum Quotient (AQ) 10-item screening tool”.; Age – “Age of the patient in years”; Gender – “Gender of the patient” ; Ethnicity – “Ethnicity of the patient”; Jaundice - Whether the patient had jaundice at the time of birth; Autism - Whether a nuclear family member has been confirmed with autism.; Country\_of\_Res – “Country of residence of the patient.”; Used\_App\_Before – “Whether the patient has undergone a screening test before.”; Result - Score for the AQ1-10 screening test.; Age\_Desc - Age of the patient; Relation - Relation of the patient who completed the test; Class/ASD - Classified result as 0 or 1. Here 0 represents No and 1 represents Yes. This is the target column.”

The data is extracted from ML Olympiad\_ Autism prediction challenge data hosted by TFUG Chennai and TFUG Mysure. The goal of the data is to assist to improve autism screening through the prediction of the likelihood of having the condition.

1. *Existing Literature*

Reference [1] investigated the utilization of AI optimization to predict autism in babies. The research was done to improve autism diagnosis via training and testing different AI models utilizing the Autism range dataset from the University of California.[1] The machine learning techniques were utilized to decide the main determinants of autism in little children quantitatively. The Neural Network and Random Forest classifier were used to develop a model. The models were then trained to determine the presence of autism in young children.[5] The study also used feature selection via the LightGBM parameter to ascertain the highest predictor of autism. As seen in the study, both models functioned well, but the neural network model marginally outperformed the random forest classifier pre-optimization. Furthermore, it was observed that the “neural network model” did excellently with the test set, which is due to the extremely high accuracy of the XGBoost model's baseline performance on the data. Although there was not much room for improvement given the pre-optimized Random Forest Classifier's already extremely high accuracy, the accuracy of the Random Forest model utilizing Grid Search was somewhat boosted. The study performed excellently in finding the best model by performing hyperparameter optimization techniques to improve the probability of the accuracy of obtaining an accurate classifier, although all the models used in the work performed excellently well. The study revealed that behavioural characteristics are stronger predictors of autism than physical characteristics. The study suggested clinical testing to validate the machine learning models in conjunction with other machine learning models. The study also suggested other methods of feature selection.

It has been explained that genetic and personal characteristics are the most easily accessible and useful features for machine learning (ML) algorithms for discovering novel and obscure data forms to aid with ASD predictions.[5] Families will be able to begin early therapeutic steps as a result of this. Despite this, the data's high complexity makes prediction challenging. Also, the use of machine learning approaches to uncover traits in newborns linked to ASD diagnoses to improve the predictability of early diagnosis. [1] It was observed that different machine learning models used in predicting accuracy and efficiency give much more space for improvement.[1] This shows a better opportunity for statistical analysis to gain more knowledge about significant features that will seek to provide information and more accurate predictive models on autism.

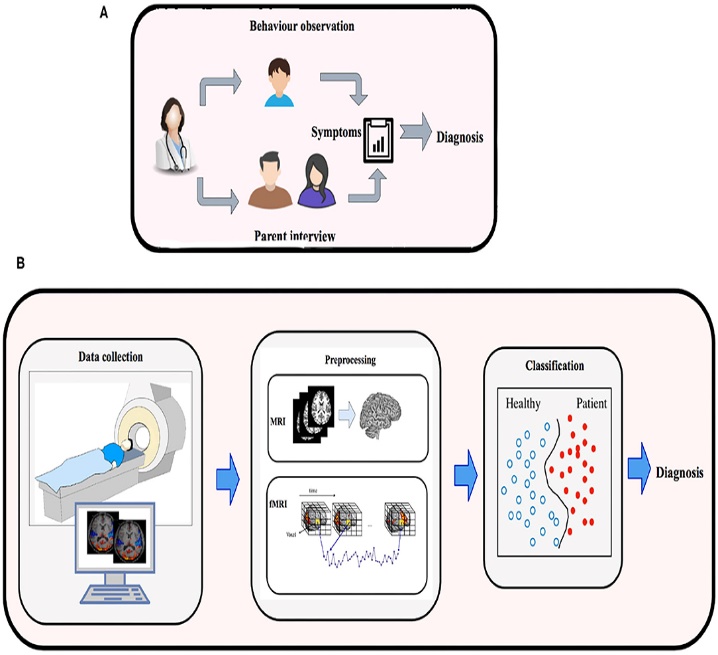


Fig. 1: Autism Spectrum Disorder Detection

“Using tabular data and conventional machine learning models to analyze 1D characteristic data to predict autism is another method for autism detection using machine learning.” [5] Regression Trees, SVM, and K-Nearest Neighbour Classifiers, for example, have been used in conjunction with patient feedback on the “Autism Quotient 10” test for adults to classify and adolescents with autism with an accuracy of about 90%.” [5,6]. The scientific standard for tabular data collection is the Autism Quotient 10 exam, which uses questions created especially to uncover characteristics that could or could not be related to autism. The Autism Quotient 10 test, however, has been discovered to be erroneous and not reliable when evaluated manually by medical professionals [7,8]. More precisely, even though the questions on the Autism Quotient 10 test are still good indicators of ASD, Misdiagnosis due to human interpretation is possible. This shows that machine learning is a particularly significant technique in diagnosing ASD using such behavioural traits, and it highlights the need for research that boosts machine learning model accuracy above the level of current studies (90%). It should be noted that the current study concentrates on a broad age span, from infants to adults. As quick therapy intervention at a young age will improve long-term outcomes compared to adults diagnosed with ASD, we chose to concentrate on toddlers for this study.

In another study, the autistic disorder was assessed using a machine learning approach.[9] The study utilized National Institute for Health Research (NHS) in Europe. The data used contained 704 samples with 21 attributes with 10 questions related to behavioural traits which can help to detect autism. Different machine learning algorithms on WEKA were used. This was used to preprocess the data and perform the data classification. Cross-validation of the data was also done as the data was divided into the training and test dataset. The models used were Naïve Bayes, K-star, Random Trees and Random Forest. The accuracy rates for the different classifiers were presented and the Random Forest was seen with the best accuracy at 100%. Hence, it was concluded that ASD in children will be best predicted using the Random Forest Classifier. [9]

**II. METHODOLOGY**

Four important processes will be explored to gain insights and to find solutions to pertinent issues in the analysis. These include descriptive, predictive, diagnostic and prescriptive analysis.

The major steps that will be involved in building the autism models include: Identifying the business problem, organising our data set, Data Exploration and Visualization, Data Cleaning and Transformation, Statistical Inference, Feature Selection, creating our training and validation data, evaluating the model, Making predictions on our test data and making decisions from the model to solve the business problem.

For the study, we are facing a rare case issue whereby 76.875% of the total number of patients don't have Autism and 23.125% of the total number of patients have Autism.

*A. Importing and Installing Libraries*

Python programming language will be used as it is an open-source language that is widely used for data science tasks. Jupyter notebook via anaconda will be used to access the files. The necessary libraries like pandas, numpy, matplotlib, seaborn, and sklearn were imported for the analysis. For the model building, algorithms such as Logistic Regression, SVM, Extra Trees Classifier, Random Forest Classifier, Catboost Classifier, LightGBM Classifier and XGBoost were used.

*B. Data Exploration*

The training and test data were examined for missing values in all the variables. The datasets were complete with no missing values. However, some variables were put in the right format before modelling. It was observed that we have 800 rows and 22 columns in the training dataset while the test dataset has 200 rows and 21 columns. From the training dataset, we have 12 integer variable types, 2 float variable types, and 8 object variable types. The gender variable has 415 females and 385 males. 12 ethnicity types were represented in the ethnicity variable, with 196 respondents having jaundice. It was also seen that 117 had family members with autism and most of the respondents were 18 years or more. Two variables patient ID and Age description were dropped as they do not in any way have much significance for the analysis.

*a) Descriptive Analysis*: Descriptive analysis was performed to get an overview of the available data. It was observed that from the Autism Spectrum Quotient, A1-A10 scores are answers to questions that patients were asked. From the descriptive analysis of the data set, it is observed that all records in 2,3,4,5,6 and 9 all follow the same trend as more patients agree than disagree. Further analysis of the gender, jaundice presence, autism, app usage and relation data revealed that there were more female than male patients. Most of the patients do not have jaundice as most of the patients' close relatives do not have autism. It was also revealed that most of the patients have not used the screening apps before while most of the records are from individual representations.

Furthermore, the number of patients from each country was explored to know the nationality of the respondents and it was observed that most of the patients are from the USA. (Appendix I). The descriptive analysis of the age of the respondents was analysed and it was seen that the minimum age is 9.56, the median age is 25.47 and the highest age recorded is 72.40. The maximum, minimum and median results are 13.39, -2.59 and 6.89 respectively. It was observed that some negative values were observed in the distribution of the results. Further analysis of the zero values showed that all results below zero belong to the no autism class. The result variable was not used as a predictor to avoid data leakage problems in the model.

*b) Bivariate Analysis*: This was done to ascertain the multicollinearity of the variables. This is to ensure a better model performance such that variables that are highly correlated are combined or reduced as they are likely to produce the same results. Results of the analysis were presented in Appendix II.

Scatterplots were used to show the correlation between variables as seen in Appendix III. A linear relationship was observed between age and results. It was shown that there is a very low percentage of patients with Autism when results are less than 10, irrespective of age. Also, there is a very high percentage of patients with Autism when results are greater than 10 irrespective of age. Screening test result distribution was determined as seen in Appendix IV

*c)Diagnostic Analysis*: Diagnostic analytics uses data to assess the causes of trends and correlations of different variables. It can be seen as the next logical process after identifying trends via descriptive analytics.

In this section, we'll look at how different categorical variables relate to whether a patient has autism or not by utilizing the Chi-Square method. This was done by selecting some features and formulating a hypothesis

**Formulating our hypothesis**

H0 - The features do not significantly affect the Autism class. H1 - The features checked have a significant effect on the Autism class.

If the p-value is > 0.05, the features do not significantly affect the Autism class and vice versa. The correlation was carried out using a chi-squared test based on a p-value as shown in Figure 1.

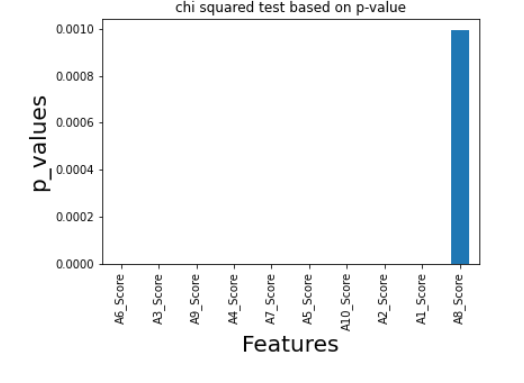


Fig.2: Chi-squared based on p-value

All the categorical features selected have a p-value < 0.05 and 0.01 significance levels. Hence, we accept H1.

The significance for gender and autism was also tested and the results showed that we can accept the alternative hypothesis (H1) for gender and the target variable (class) at a 0.002 p-value. Family autism and Jaundice and the target variable (class) significantly affect each other since their p-values are less than 0.05. For other categorical features like ethnicity, contry\_of\_res and relation that do not have statistical significance with the target variables since their p-values are greater than 0.05, we must note that their interaction might have a good potential to better classify autism. Hence, we keep them as predictors except for contry\_of\_res because of the high cardinality effect.

* 1. *DATA MODELLING*

For the modelling, different algorithms were used to build the models for predicting whether a person has autism or not. The models were used on the test data for predictions. For yes and no representations in our variables, we employ ordinal encoding. Consequently, yes=1 and no=0.

We checked for high cardinality features before the transformation. This was done to examine ethnicity; country of residence and relation overlap in training and test data. We then convert the data using nominal encoding (PD.get\_dummies). This was carefully done because if we use one-hot encoding for country\_of\_res, it will cause an imbalance of representation in the train and test data and for this reason, we exclude it from our independent variables. The next phase is feature selection where we drop the variables that are not needed for the modelling in the train and test set. This feature selection type is classified as the most relevant feature.

D. *BUILDING MACHINE MODEL*

The algorithms that will be used for building the model are “Logistic Regression”, “Support Vector Machine”, “Gaussian Naive Bayes” (GaussianNB), Extra Trees Classifier, “Random Forest Classifier”, “Decision Tree Classifier”, “Light Gradient Boosting Machine” (LightGBM), “Extreme Gradient Boosting” (XGBoost) and boost.” An explanation of how each of the algorithms work is given in Appendix V [10, 11]:

The validation set assesses predicted accuracy more objectively than the training set. Validation set data are used to identify predictors or estimate model parameters since they are comparable to future records. After training on the training set, accuracy is tested utilizing validation set prediction errors.

E*. ACCURACY MEASUREMENT*

Most accuracy measurements are generated from the confusion or classification matrix. This matrix explains a classifier's accurate and inaccurate dataset classifications. The confusion matrix's rows and columns represent predicted and actual classifications. Applying different classifiers to training data generates a confusion matrix for a two-class (0/1) problem. The confusion matrix calculates true and misclassification rates. If we have a large enough dataset, our estimations will be accurate. We use the validation data's confusion matrix to estimate future classification errors. We randomly select records to create training and validation sets. The data is divided into 80% training set and 20% validation set. We build a classifier utilizing a training set and using it for the validation set. This produces anticipated validation set classifications. A confusion matrix is created. We can summarize training data findings in a confusion matrix, but it's not useful for new data owing to overfitting.

**IV. RESULTS**

The AUC score for the accuracy was used to determine the accuracy of probability. AUC is the probability that a positive (autistic) example is to the right of a negative one. AUC=0-1 A model with 100% erroneous predictions has an AUC of 0.0.

AUC has two advantages:

1. Scale-invariant and it gauges rankings, not absolute values.
2. AUC is threshold-invariant. It measures model predictions regardless of threshold.

From the AUC scores shown below, the Random Forest Algorithm has the best performance on the validation set of our model while the XGBoost Classifier has the least performance.

Table 1: Model and AUC Sco

|  |  |  |
| --- | --- | --- |
| S/N | Model | AUC Score |
| 1 | Random Forest | 0.925731 |
| 2 | SVC | 0.923973 |
| 3 | Logistic Regression | 0.923533 |
| 4 | Catboost | 0.912766 |
| 5 | Extra Trees | 0.907053 |
| 6 | Gaussian | 0.880026 |
| 7 | Decision Tree | 0.877499 |
| 8 | LightGBM | 0.876291 |
| 9 | XGBoost | 0.874313 |

The ROC Curve for each Model was plotted as shown in Appendix VI. ROC curves are used for plotting a model's specificity and sensitivity. ROC plots sensitivity and specificity as the cut-off value fall from 1 to 0. Alternately, 1-specificity can be plotted on the x-axis, with 0 on the left and 1 on the right. Better performance is shown by top-left curves. The diagonal comparison curve shows the average performance of an uninformed guessing classifier which shows the performance of the predictors or outcome variables. This guessing classifier guesses that a proportion α of the records is 1 and therefore assigns each record an equal probability P(Y = 1) = α. In this case, on average, a proportion α of the 1s will be correctly classified (Sensitivity = α), and a proportion α of the 0s will be correctly classified (1 - Specificity = α). As we increase the cutoff value from 0 to 1, we get the diagonal line Sensitivity = 1 - Specificity: Note that the naive rule is one point on this diagonal line, where α = proportion of actual 1s. A known metric to explain a ROC curve is "area under the curve (AUC)”, "which ranges from 1 (perfect discrimination between classes) to 0.5 (no better than random guessing)”.

The classification accuracy for Logistic Regression Model was 88.75%. The classification accuracy for the Support Vector Classifier was 86.88%. The classification accuracy for Gaussian Classifier was 65.00%. The classification accuracy for the Extra Tree Classifier was 85.00%. The classification accuracy for the Random Forest Classifier was 84.38%. The classification accuracy for the Decision Tree Classifier was 85.00%. The classification accuracy for the CatBoost Classifier was 86.25%. The classification accuracy for the LightGBM Classifier was 84.38%. The classification accuracy for the XGBoost Classifier was 84.38%. (See Appendix VI)

Hyper-Parameter optimization using Random Search CV was performed on the Random Forest Classifier based on its initial performance and the accuracy score was further increased to 90.0% using n\_estimators 1000, min\_samples\_split 13, min\_samples\_leaf 6 and max\_depth 9.

Out of the 200 records in our test set, our model was able to classify 37 individuals as those with Autism while 163 were non-autistic using a cutoff value of 0.5.

1. Model Deployment

The optimized Random Forest model was saved as a pickle file for deployment which will be used for predicting whether a patient is autistic or not via a web application framework called Streamlit. It allows for turning application development time from days into hours (See Appendix VIII).

**V. DISCUSSION**

The Logistic Regression model has high accuracy, but in medical research, we focus on the model with high specificity (Recall). In this case, we go with Random Forest Classifier with an initial classification accuracy of 84.38% and 0.925731 AUC Score. In a real-life scenario especially in the health sector, we take care of the rare class using metrics like ROC AUC. The algorithm was then fit with all the training data. We then run a prediction with the test data. For the test data, we generate the propensities for classification for each of the records. Most classification algorithms start by estimating the likelihood that a record is grouped into each class. The term "propensities" also applies to these probabilities. Propensities are often used to rank-order the records based on their likelihood of belonging to an interesting class or as an intermediate level for generating anticipated class membership (classification). If overall classification accuracy (across all classes) is important, classify the record with the highest likelihood. We shall concentrate on the class known as "autistic" in the autism test record because it is of particular importance. In classifiers with two classes, the default cutoff value is 0.5. As a result, records are categorized as 1 if the likelihood that they belong to the class of autistic people is more than 0.5. However, it is feasible to employ a threshold that is either greater or lower than 0.5. The number of records classified as 1 will decrease if the cutoff is greater than 0.5, whereas the number of records classified as 1 will increase if the cutoff is lower than 0.5. Usually, the misclassification rate will increase in both scenarios. In a real-life scenario, it is possible to say we have more healthy people than those with autism.

Based on the promising results of the Random Forest classifier, we can anticipate that the prediction can be used to monitor autism among patients.

**VI. CONCLUSION**

This study developed accurate ASD screening models to help parents diagnose their children. Some families and adult patients don't know ASD symptoms; therefore, instances aren't treated early. AI and ML are used in most life research, and their usage in medical diagnosis is a pioneering step in utilizing accessible data for improvement and progress.

The transformation of this study into a web application will give families a fast and uncomplicated ASD scan tool, increasing accessibility and early detection. This study could be used in the Ministry of Health and Prevention's health system, as data are available for all needed personnel. Educational institutions can be given an adaptable approach for early detection using the application. This study could also apply machine learning and AI models in health settings that record patient data for different illnesses and signs to aid in early diagnosis.

1. **REFERENCES**

[1] Singh. A., Farooqui, Z., Sattler, B., Usua, S. Helde, M., 2021. Using Machine Learning Optimization to Predict Autism in Toddlers, Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore.

[2] Alsuliman, M.; Al-Baity, H.H. Efficient Diagnosis of Autism with Optimized Machine Learning Models: An Experimental Analysis on Genetic and Personal Characteristic Datasets. Appl. Sci. 2022, 12, 3812. <https://doi.org/10.3390/app12083812>

[3] Altman, N. S. (1992). An introduction to kernel and nearest-neighbour nonparametric regression. Am. Stat. 46, 175–185. DOI: 10.2307/2685209.

[4]<https://www.kaggle.com/c/autismprediction/data>.

[5] Alteneiji, M., Alqaydi, L., and Tariq, M., Autism Spectrum Disorder Diagnosis using Optimal Machine Learning Methods, International Journal of Advanced Computer Science and Applications, vol. 11, no. 9, 2020

[6] Chowdhury, K., and Iraj, M., Predicting Autism Spectrum Disorder Using Machine Learning Classifiers, Proceedings of the 2020 International Conference on Recent Trends in Electronics, Information, Communication &Technology, pp. 324-327, 2020.

[7]Jia, R., Steelman, Z., and Jia, H., Psychometric Assessments of Three Self-Report Autism Scales (AQ, RBQ-2A, and SQ) for General Adult Populations, Journal of Autism and Development Disorders, vol. 49, pp. 1949-1965, 2019.

[8] Taylor, E., Livingston, L., Clutterbuck, R., and Shah, P., Psychometric concerns with the 10-item Autism-Spectrum Quotient (AQ10) as a measure of trait autism in the general population, Experimental Results, vol. 1, 2020.

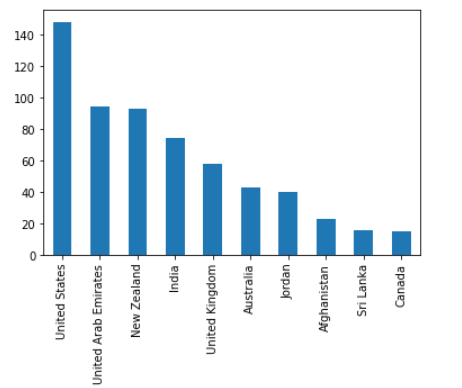
[9] Raj, S., and Masood, S., Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques, Procedia Computer Science, vol. 167, pp. 994-1004, 2020.

[10] Singh, A., Thakur, N.& Sharma, A., A review of supervised machine learning algorithms, *3rd* International Conference on Computing for Sustainable Global Development(INDIACom), 2016, pp.1310-1315.

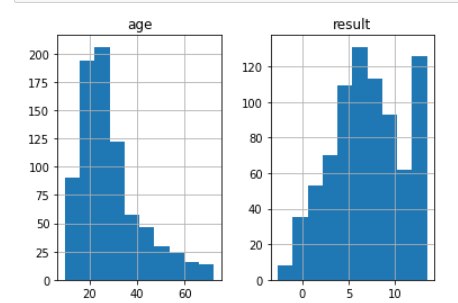
[11] Sathya, R. & Abraham, A., Comparison of Supervised and Unsupervised Learning “Algorithms for Pattern Classification, International Journal of Advanced Research in Artificial Intelligence,2(2), 2013.”

**APPENDIX**

APPENDIX I

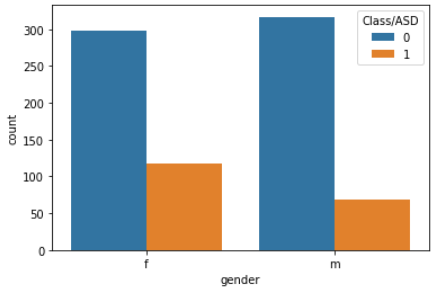


Countries and Number of patients

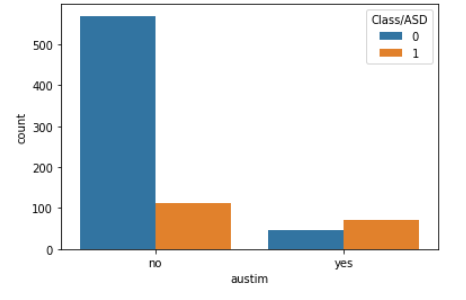


Age Distribution of respondents

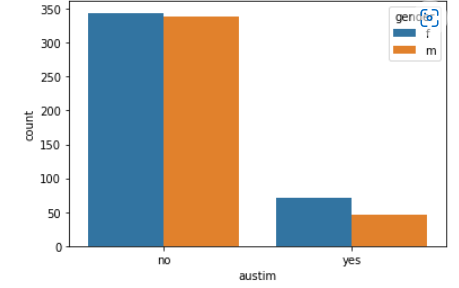
APPENDIX II



Gender and Autism

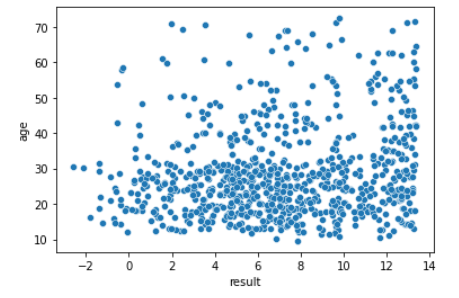


Family members and Autism

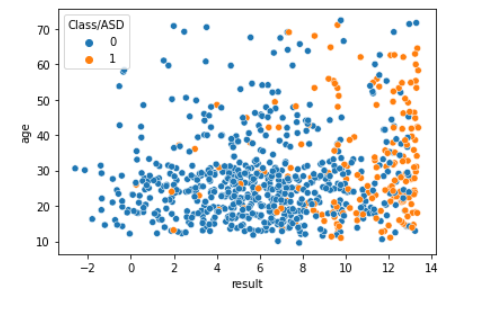


Gender of family members and Autism

APPENDIX III



Scatterplot of Age and Results



Autism and patients less/greater than 10.

APPENDIX IV

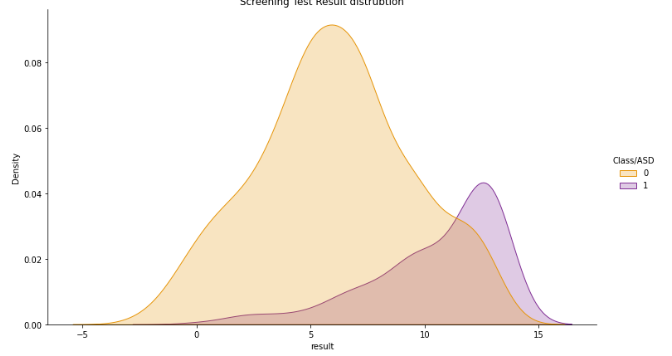


Figure 1: Screening Results

The chart below shows that the age distribution of those without Autism is lower than those with Autism.

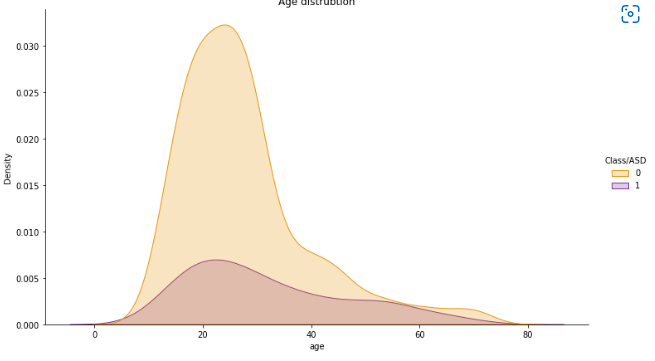
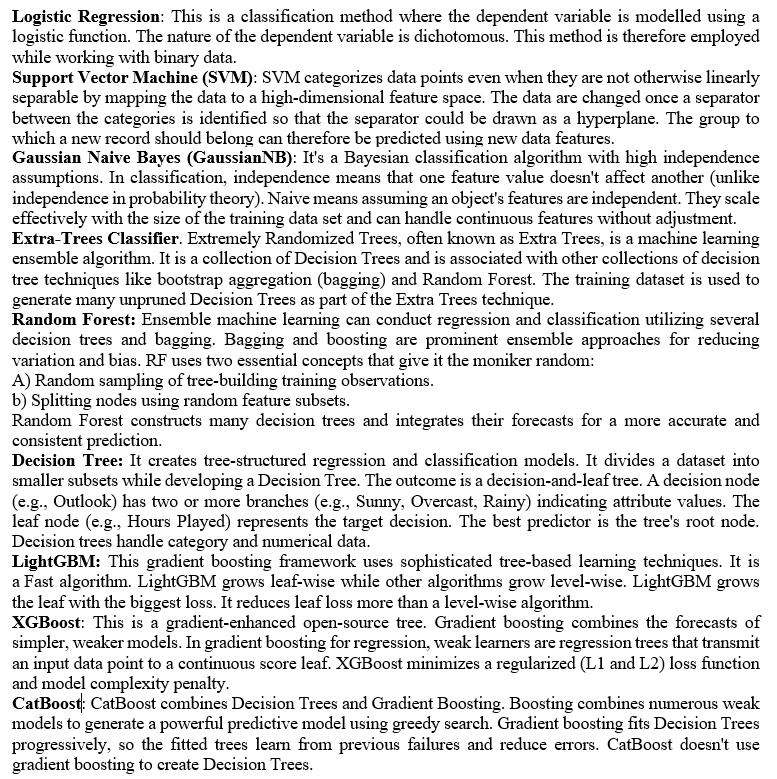
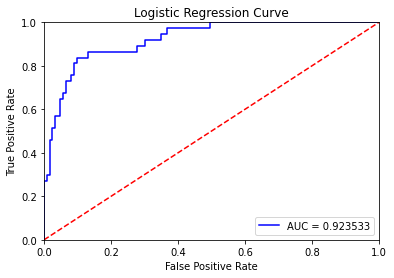


Figure 2: Age Distribution and Autism

APPENDIX V



APPENDIX VI

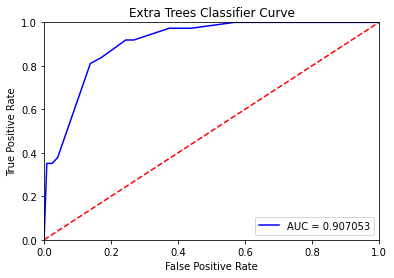


Graphical user interface

Description automatically generated with medium confidence

Line chart

Description automatically generated



Graphical user interface

Description automatically generated with medium confidence

Chart, line chart

Description automatically generated

Graphical user interface

Description automatically generated with low confidence

A picture containing line chart

Description automatically generated

Chart, line chart

Description automatically generated with medium confidence

APPENDIX VII

**Logistic Regression**

Prediction

Actual 0 1

0 112 11

1 7 30

**Support Vector Classifier**

Prediction

Actual 0 1

0 111 12

1 9 28

**Gaussian NB Classifier**

Prediction

Actual 0 1

0 70 53

1 3 34

**Extra Tree Classifier**

Prediction

Actual 0 1

0 106 17

1 7 30

**Random Forest Classifier**

Prediction

Actual 0 1

0 104 19

1 6 31

**Decision Tree Classifier**

Prediction

Actual 0 1

0 106 17

1 7 30

**CatBoost Classifier**

Prediction

Actual 0 1

0 111 12

1 10 27

**LightGBM Classifier**

Prediction

Actual 0 1

0 110 13

1 12 25

**XGBoost Classifier**

Prediction

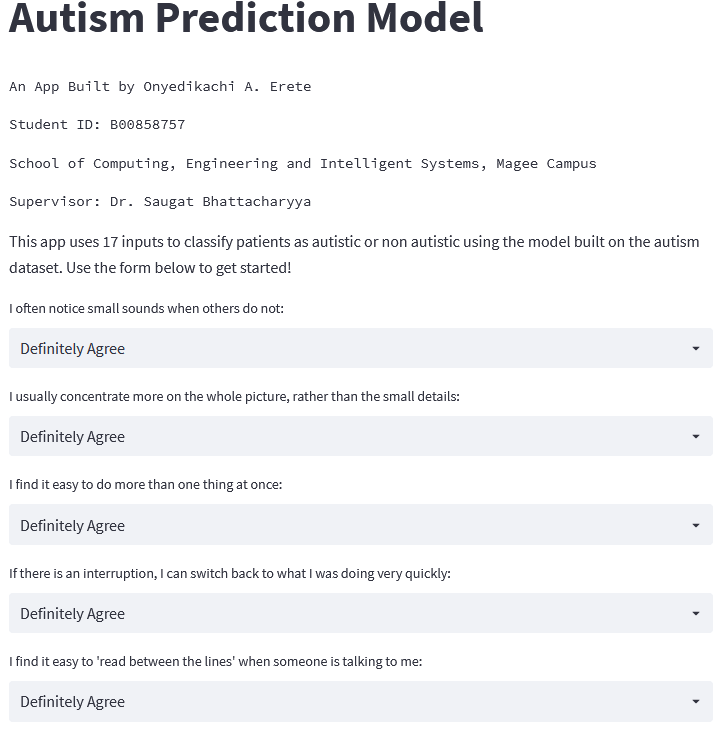
Actual 0 1

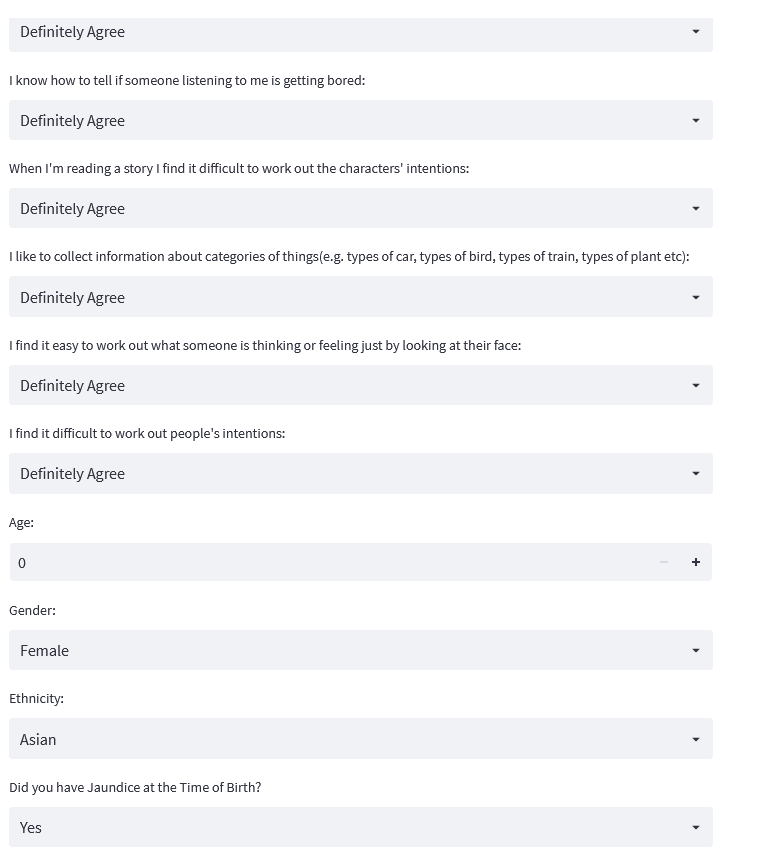
0 110 13

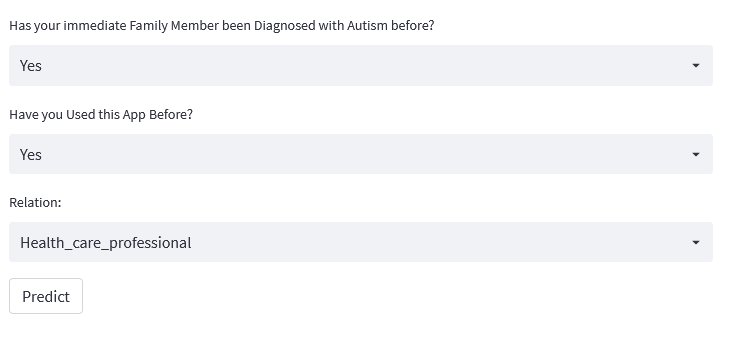
1 12 25

**APPENDIX VIII**

[**https://onyedikachierete-autism-masters--autism-model-deployment-gj5610.streamlitapp.com/**](https://onyedikachierete-autism-masters--autism-model-deployment-gj5610.streamlitapp.com/)

****

****

****