



Simplifying Autism Spectrum Disorder Screening

Syeda Tahera Ahmed; Abhishek Dutt; Atisha Rajpurohit; Matene Zamaninia
Georgetown University, MS.



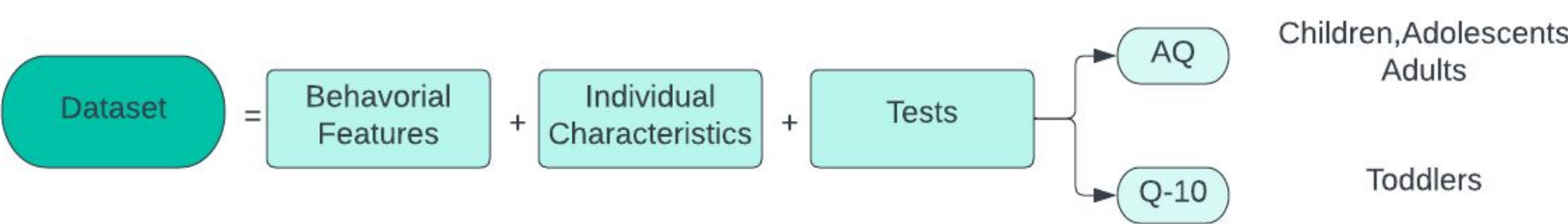
Introduction

Autism spectrum disorder (ASD) is a neurological and developmental disorder that is associated with challenges in communication, social interaction, and repetitive behavior patterns. It is preferred it is diagnosed at an early stage, to understand how to help the child. Unfortunately, waiting times for an ASD diagnosis are lengthy and procedures are not cost effective. This project focuses on understanding which behavioral aspect is indicative of ASD for toddlers and would advise a formal clinical diagnosis.

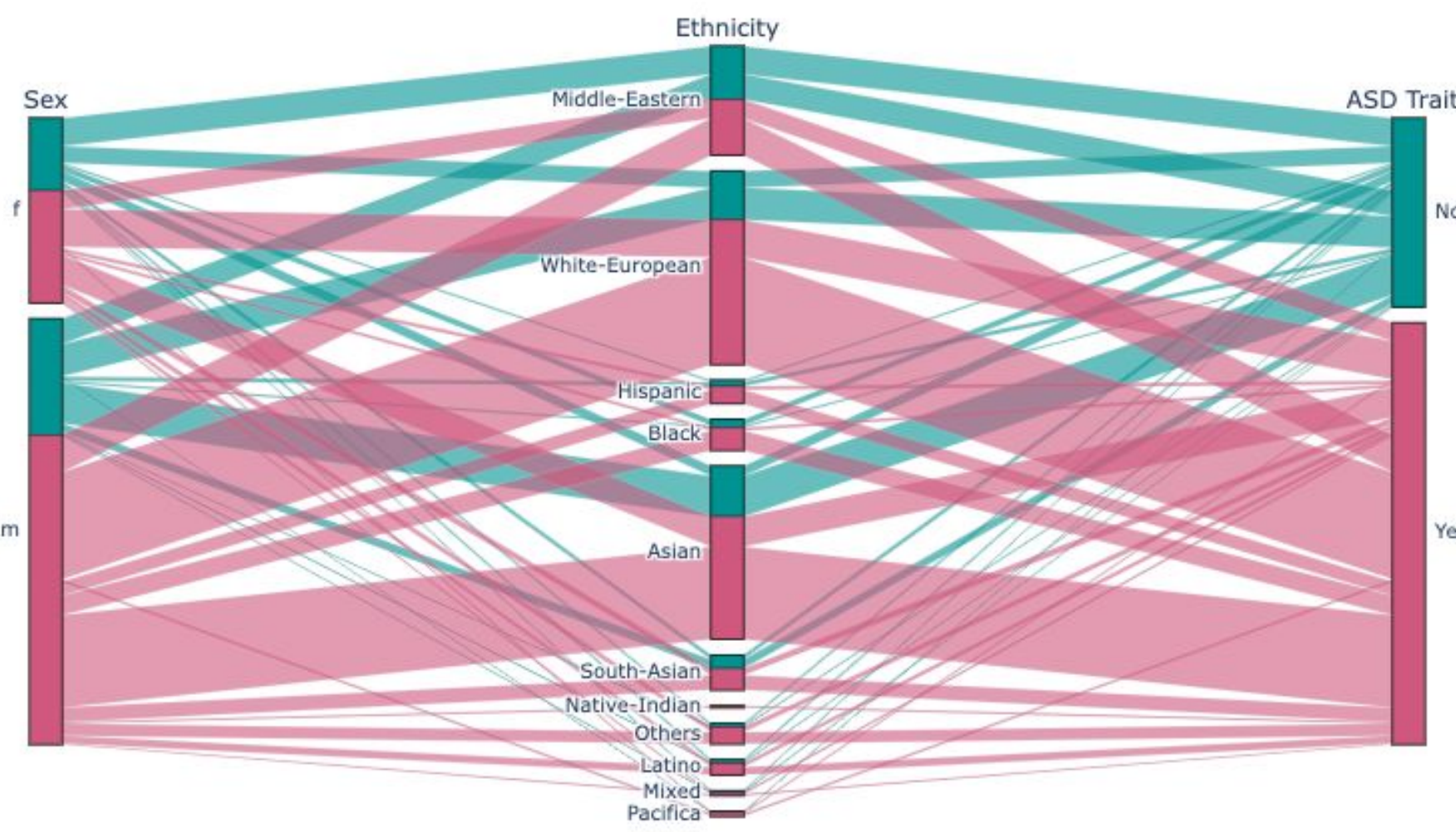
Analysis

About the Data

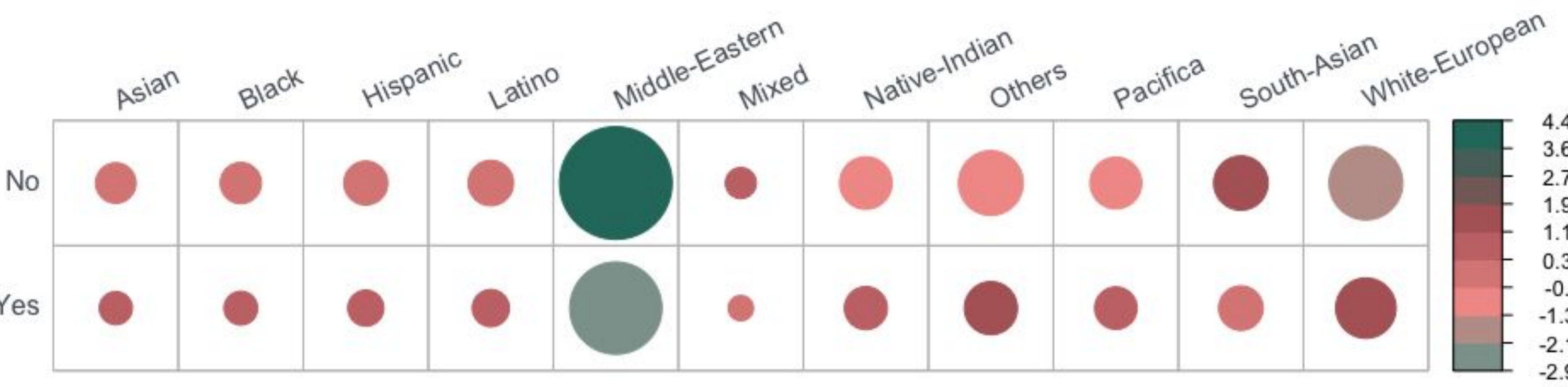
The dataset was developed by Dr. Fadi Fayez Thabtah to screen Autism Spectrum Disorders. The dataset contains 10 questions about behavioral features and a few demographic characteristics.



Exploratory Data Analysis



Representing the flow of data from Sex to Ethnicity and ASD Trait



Significance of Ethnicity for determining presence of the ASD Trait

Statistical Methods

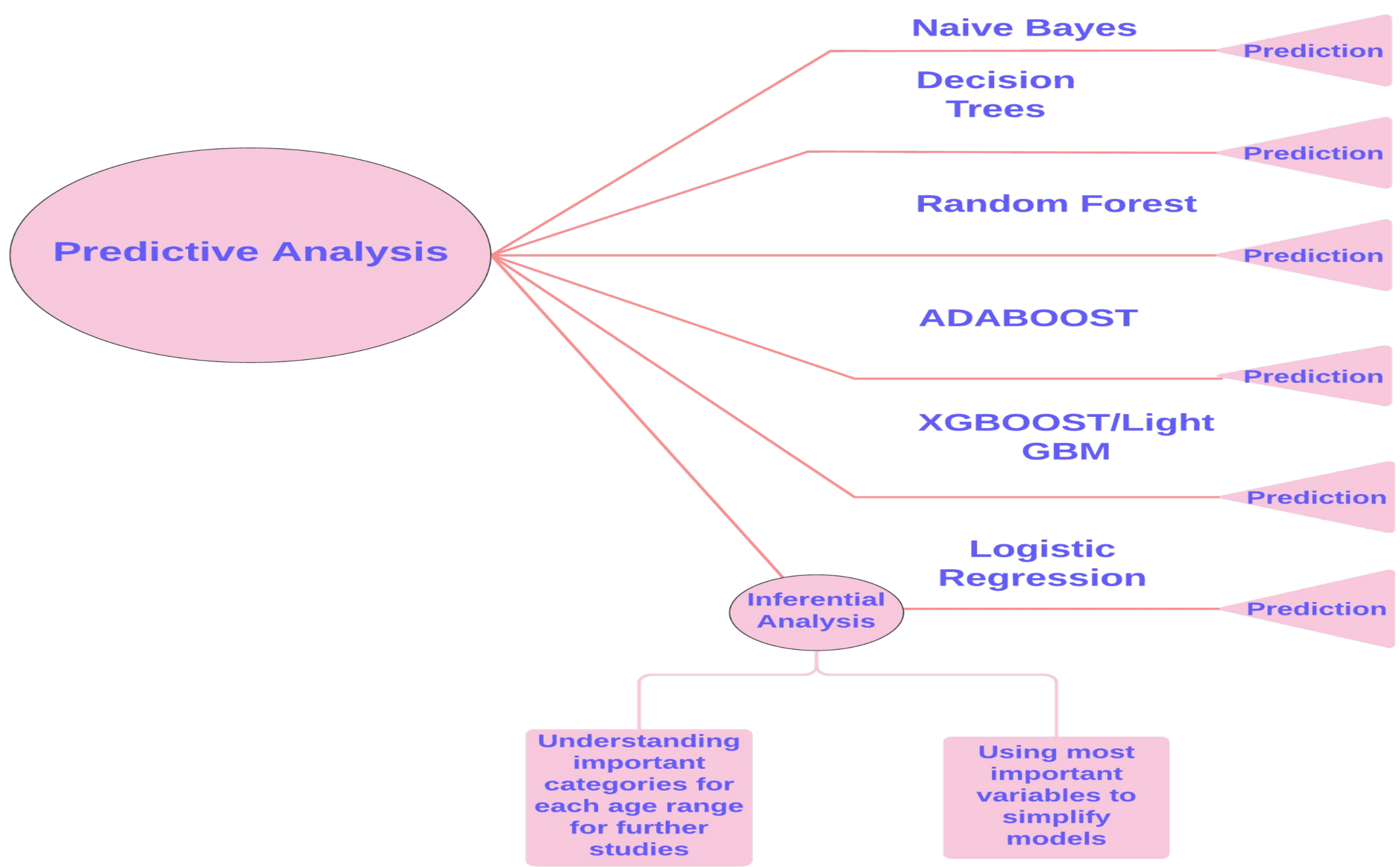
Predictive Analysis

The ultimate aim of the project is to understand which variables have better predictive ability in each age group. In this case, all of the 21 variables are kept under consideration.

Inferential Analysis

This is built on top of the predictive analysis. The most important variables from the predictive analysis were identified from the Logistic Regression variable importance plot.

Once the most important features are identified, the Machine Learning Models are built using the top 3 variables, per age group.

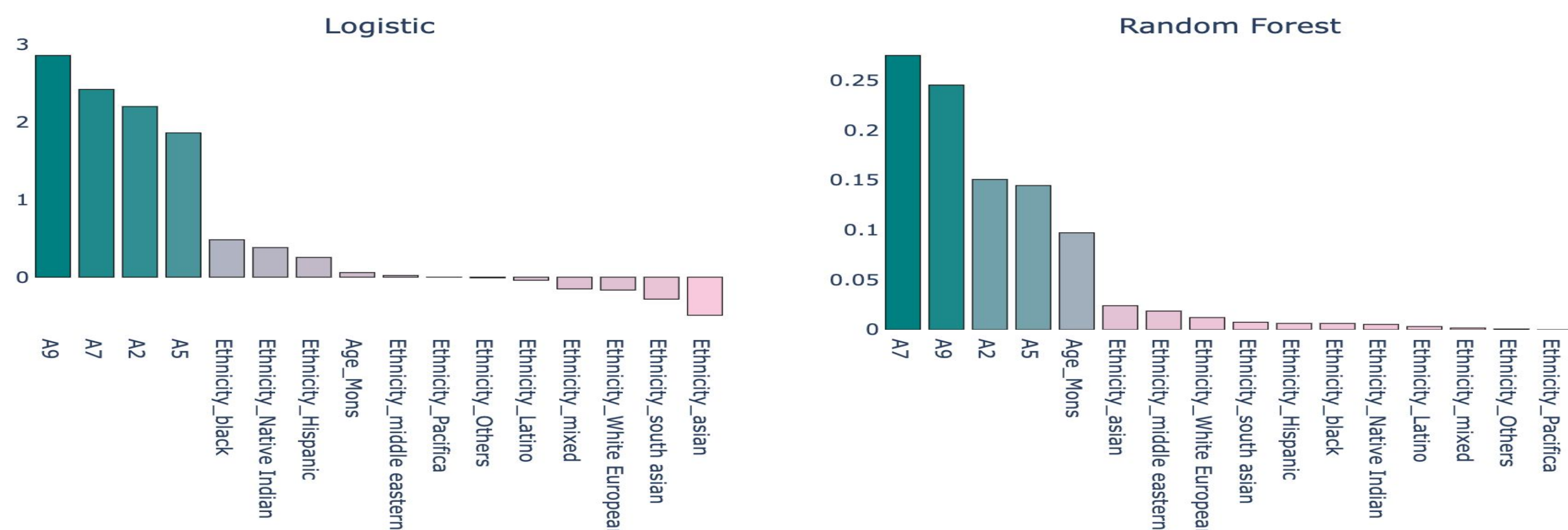


Understanding the important cognitive and social abilities serves two purposes; First, it allows us to reduce the number of questions needed with minimal decrease in prediction accuracy from 10 to 3. Secondly, understanding the important variables per age group allows us to focus on these areas in further studies and question design.

Results

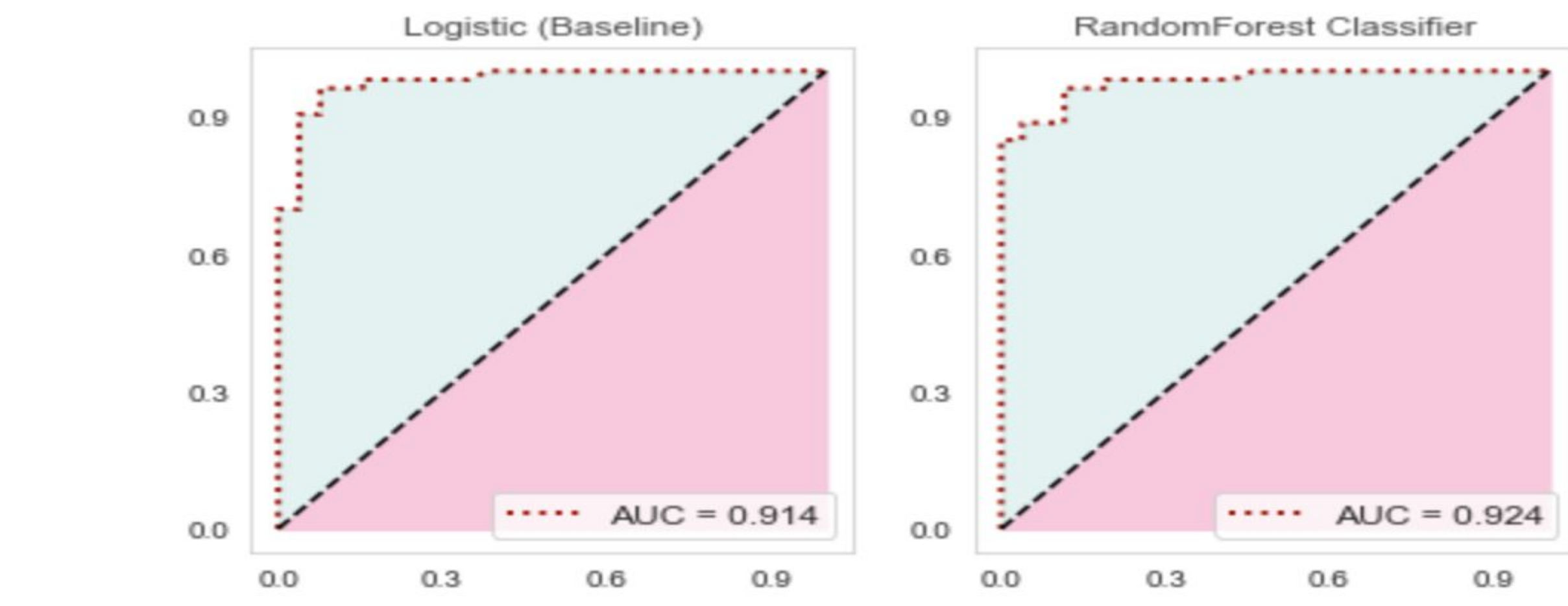
Model	Accuracy	Recall	Precision	F1-Score	ROC_AUC	Kappa_Metric
Logistic (Baseline)	0.9114	0.9057	0.96	0.932	0.9144	0.8051
RandomForest Classifier	0.9114	0.8868	0.9792	0.9307	0.9242	0.8087

RF model predicts if someone has ASD traits 97% of the times. It also correctly identifies 88% of the times patients actually have ASD traits. The tradeoff in this case is well handled. Question 9,7,5 and 2 are shown to have the greatest importance.

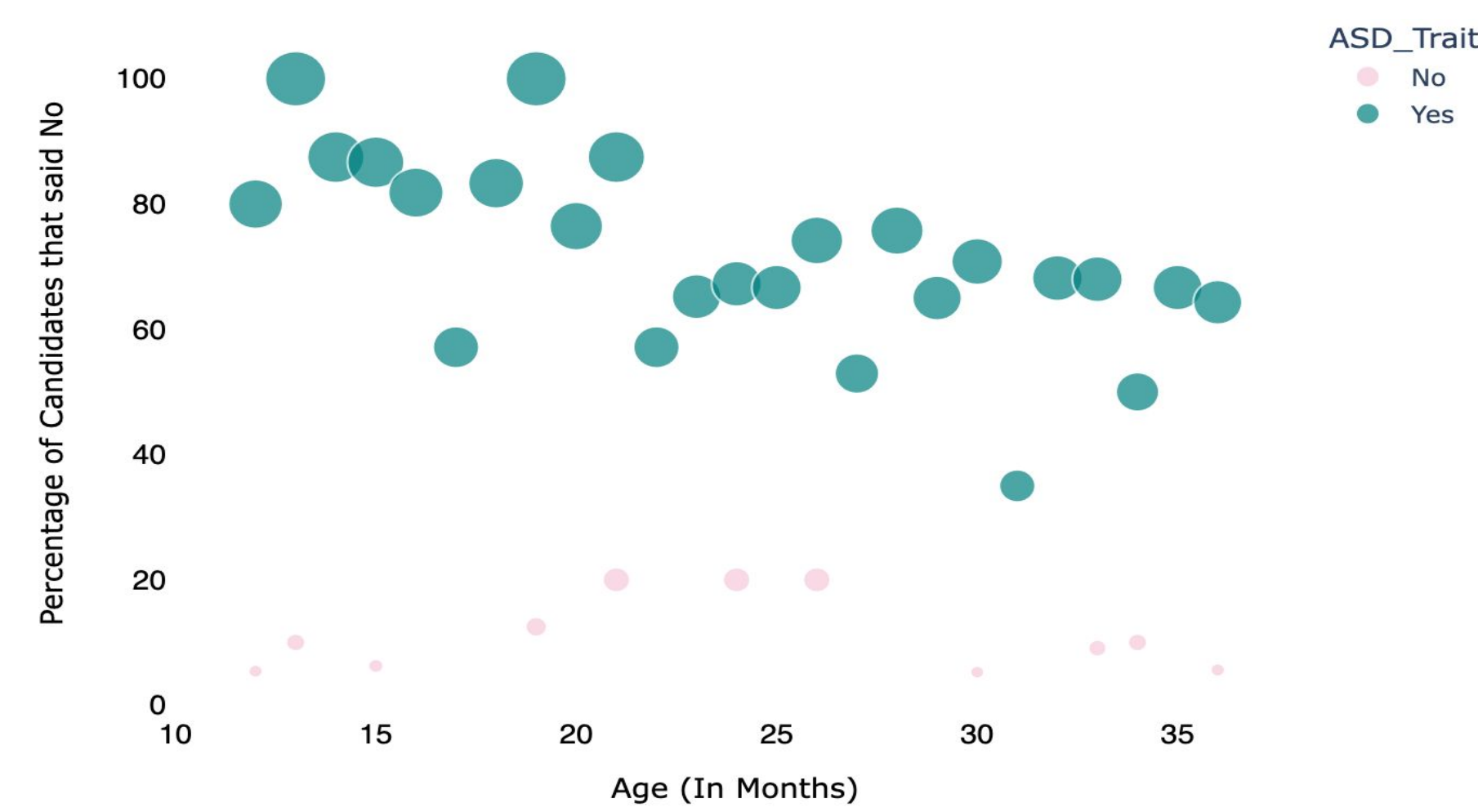


Feature Importance according to Logistic and Random Forest Classifiers

The models are able to distinguish the patients with ASD traits and those who don't 92% of the times.



ROC- AUC Curves for Logistic and Random Forest Classifiers



Question 9 and ASD traits

A high percentage of children were classified as having ASD when a candidate answered No to whether their child uses simple gestures to communicate.

Conclusions

From the analysis, we can conclude that by just using three questions along with a few demographic characteristics, we can achieve high accuracies in predicting ASD cases using machine learning algorithms. This helps in providing early attention to cater to the needs of toddlers with ASD.

We can also conclude that questions in the category of communication are more important in predicting ASD in toddlers, but as the individual grows older, the questions in the category of social interaction become more and more important.

The limitation of this study is that it does not take into account the help an individual might need after the diagnosis.

References

- 1) Thabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. Proceedings of the 1st International Conference on Medical and Health Informatics 2017, pp.1-6. Taichung City, Taiwan, ACM.
- 2) Thabtah, F. (2017). ASDTests. A mobile app for ASD screening. www.asdtests.com [accessed December 20th, 2017].
- 3) Thabtah, F. (2017). Machine Learning in Autistic Spectrum Disorder Behavioural Research: A Review. Informatics for Health and Social Care Journal. December, 2017 (in press)