Autism Spectrum Disorder Prediction using ML

Presented by Caroline Araujo

Table of Contents

01 ASD
A brief overview

02 ML Models
Models used in this project

03 Avg Score
How well the model works

04 MySQL
Questions asked with SQL

05 Streamlit app
How to turn this into a product?

06 Conclusions
What's next?

Autism

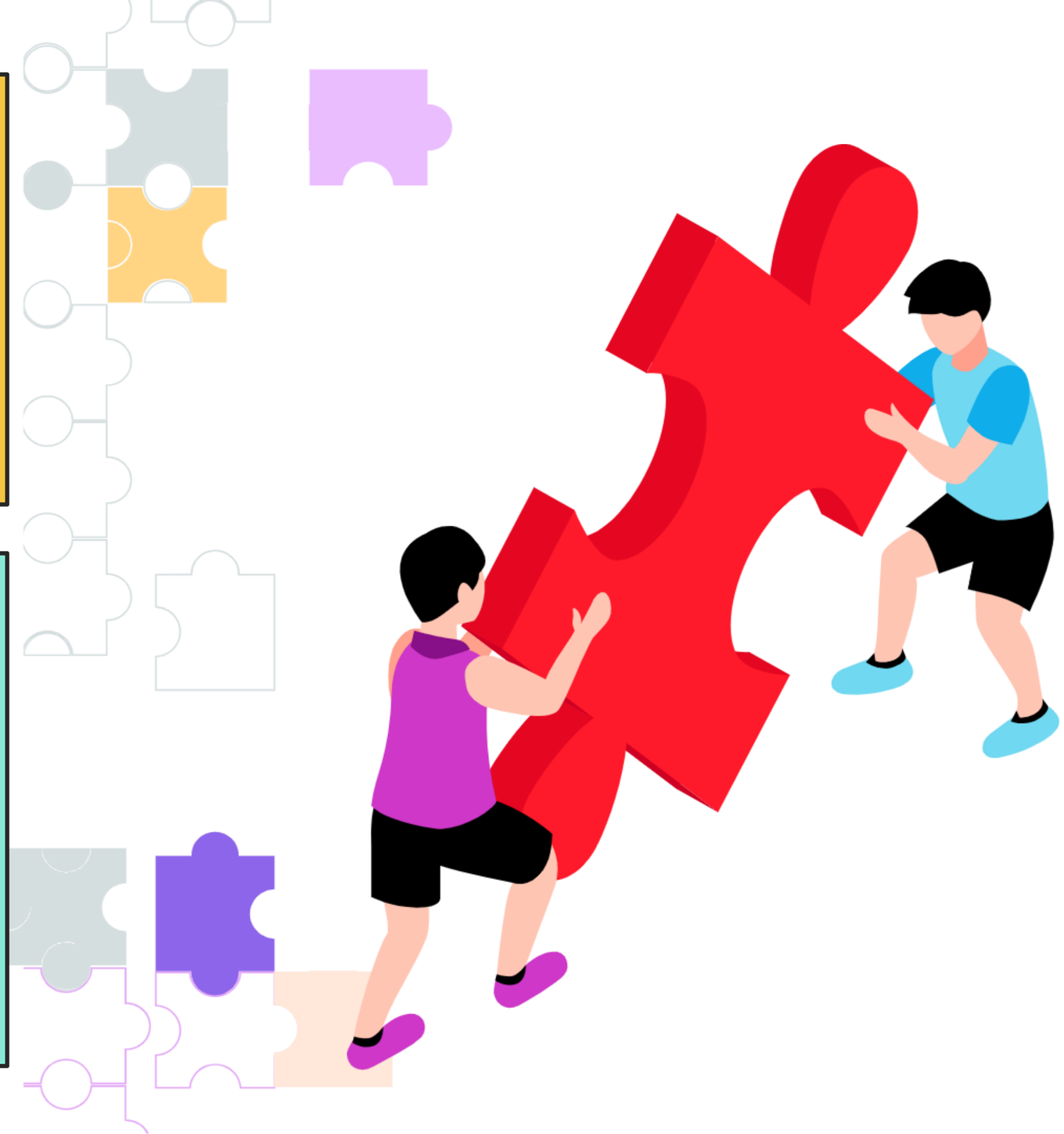
A variety of developmental differences in social interaction, communication, and behavior.

Research

Genetic Influences
Environmental Influences
(Increase Risk)

Risk Factors

Family History
Advanced Parental Age
Pregnancy Complications
Birth Complications
Pregnancy Spacing

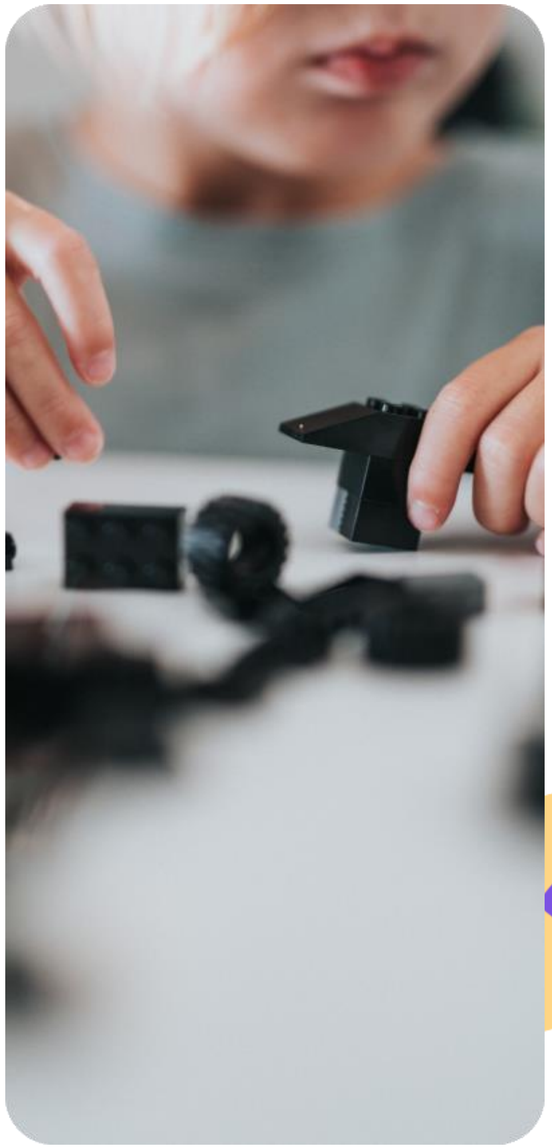


Diagnostic Factors

- Developmental History
- Social interaction
- Behavior
- Social-Communication Skills
- Developmental Skills
- Observation & Assessment
- Caregiver Report

Symptoms

- Speech & Nonverbal Communication
- Making Eye Contact
- Understanding Expressions & Feelings
- Interacting with Peers
- Initiating "Pretend" or "Imaginary" Play
- Sensory Processing
- Sensitivities (sounds, touch, movement & changes)
- Routine or Repetitive Behaviors
- Hyper Focus



Strenghts

1

Learning & Remembering (Facts, Events, Stories)

2

Visual Problem Solving (Math Problems, Puzzles, How Things Work)

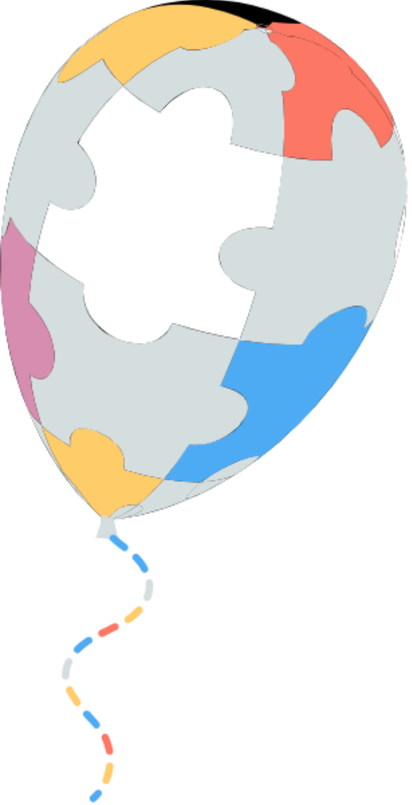
3

Following Routines & Schedules

4

Thinking from a different perspective

The dataset



UC Irvine Machine Learning Repository

Datasets | Contribute Dataset | About Us | Search datasets... | Login

Autistic Spectrum Disorder Screening Data for Children

Donated on 12/23/2017

Children screening data for autism suitable for classification and predictive tasks

Dataset Characteristics	Subject Area	Associated Tasks
Multivariate	Health and Medicine	Classification
Feature Type	**# Instances**	**# Features**
Integer	292	20

Dataset Information

Additional Information

see attached file for variables' description

Has Missing Values?

Yes

DOWNLOAD (39.1 KB)

IMPORT IN PYTHON

CITE

0 citations

9670 views

Creators

Fadi Thabtah

DOI

10.24432/C5659W

Notes

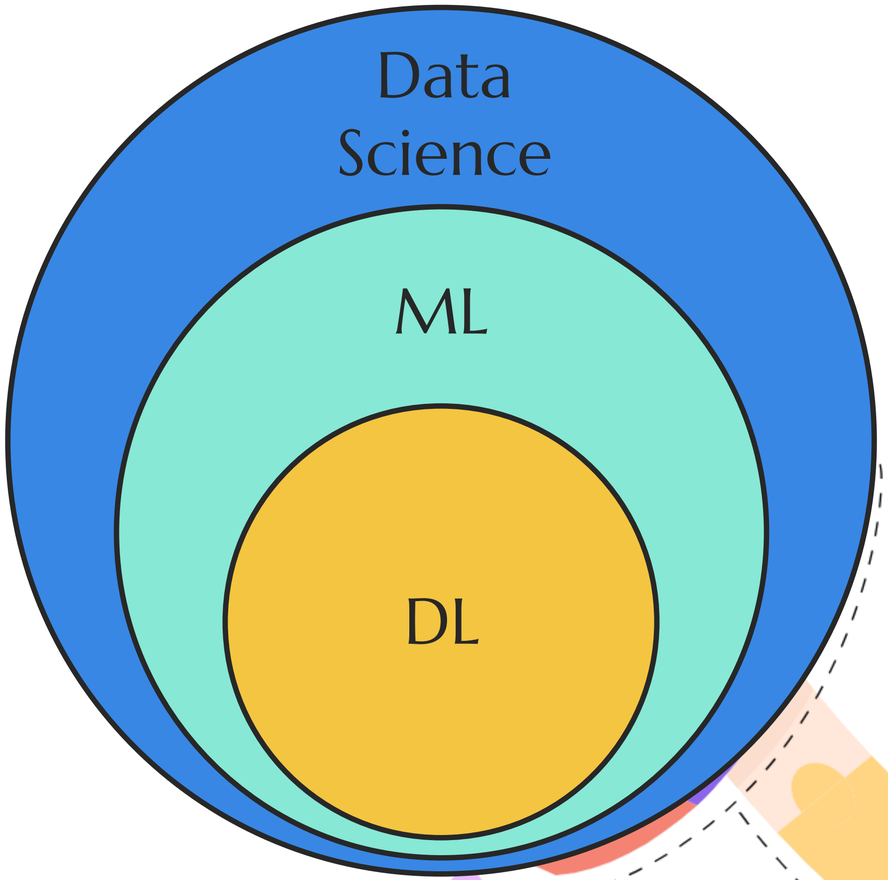
419 and 420 are duplicates?

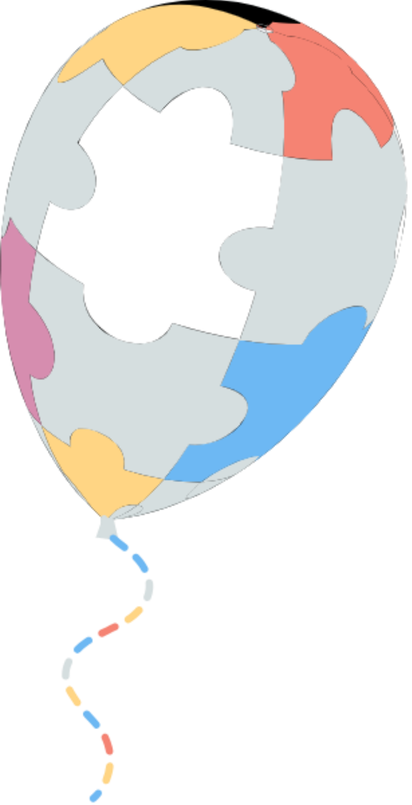
License

ML Models

The models used in this project were:

- **Logistic Regression**
- **Support Vector Classifier (SVC)**
- **Naive Bayes (GaussianNB and MultinomialNB)**
- **MLPClassifier (Neural Network)**
- **SGDClassifier (Stochastic Gradient Descent)**
- **KNeighborsClassifier (K-Nearest Neighbors)**
- **Decision Tree Classifier**
- **Random Forest Classifier (with hyperparameter tuning)**
- **Gradient Boosting Classifier (with hyperparameter tuning)**
- **LGBMClassifier**
- **XGBoost Classifier**

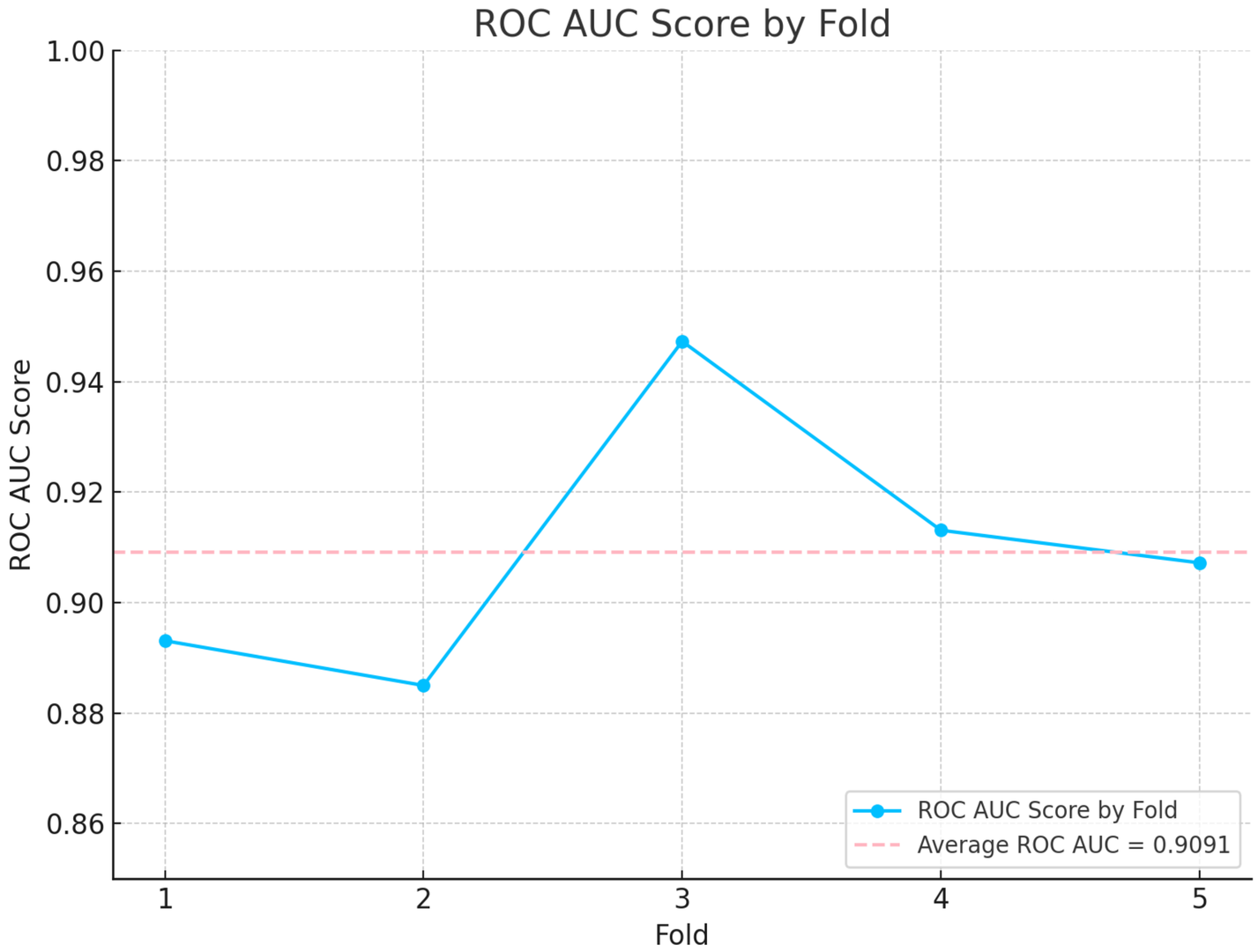




ML Models

```python
# Step 12: Define classifiers
seed = 123
model1 = LogisticRegression(max_iter=500, random_state=seed)
model2 = SVC(random_state=seed)
model3 = GaussianNB()
model4 = MLPClassifier(random_state=seed, max_iter=500)
model5 = SGDClassifier(random_state=seed)
model6 = KNeighborsClassifier()
model7 = DecisionTreeClassifier(random_state=seed)
model8 = RandomForestClassifier(random_state=seed, class_weight="balanced")
model9 = GradientBoostingClassifier(random_state=seed)
model10 = LGBMClassifier(random_state=seed)
model11 = XGBClassifier(random_state=seed, use_label_encoder=False)
```

ROC AUC Score



- **Fold 1/5 - Score: 0.8931**
- **Fold 2/5 - Score: 0.8850**
- **Fold 3/5 - Score: 0.9473**
- **Fold 4/5 - Score: 0.9131**
- **Fold 5/5 - Score: 0.9072**

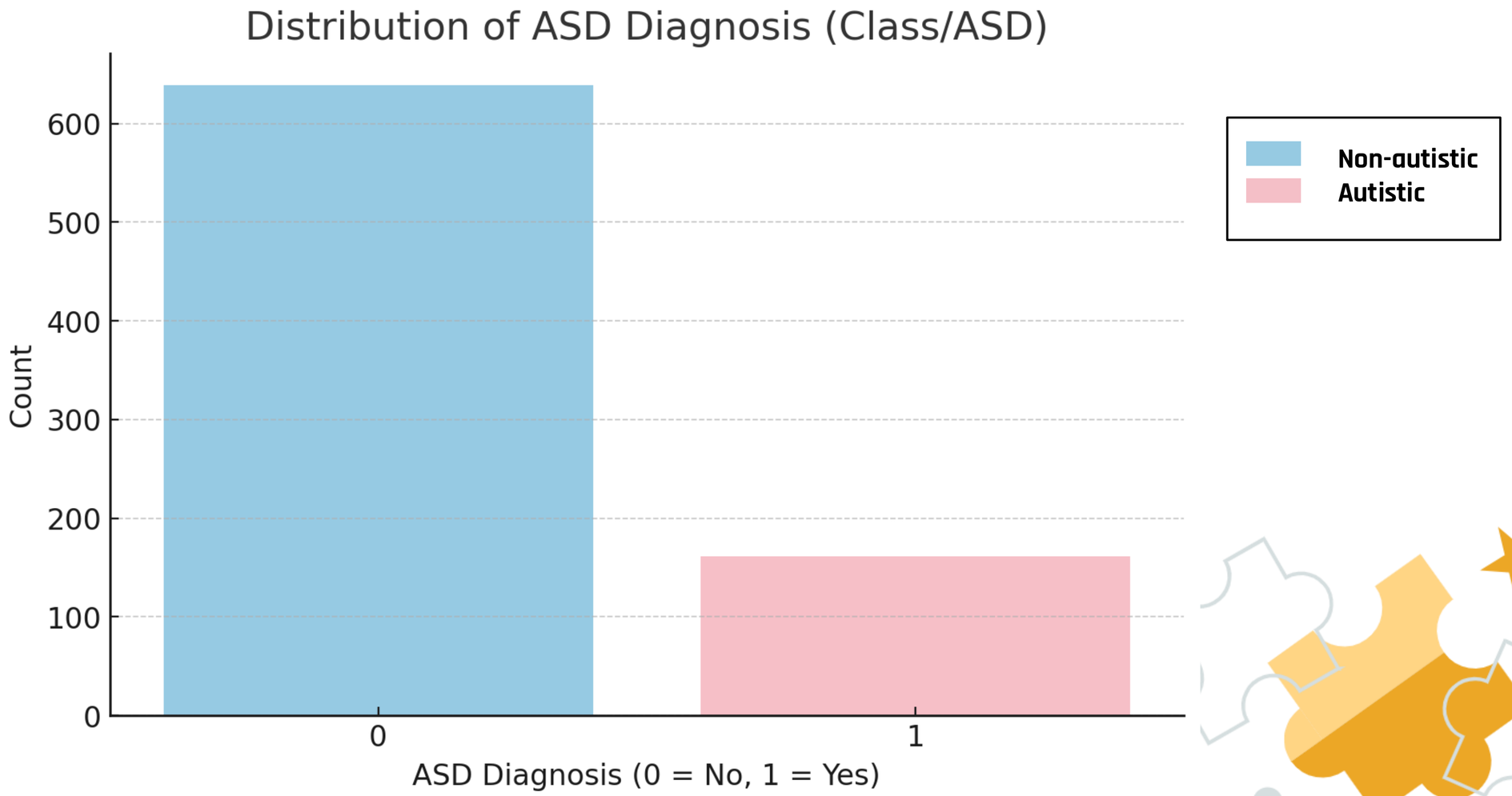
Avg scores - 0.9091

90%

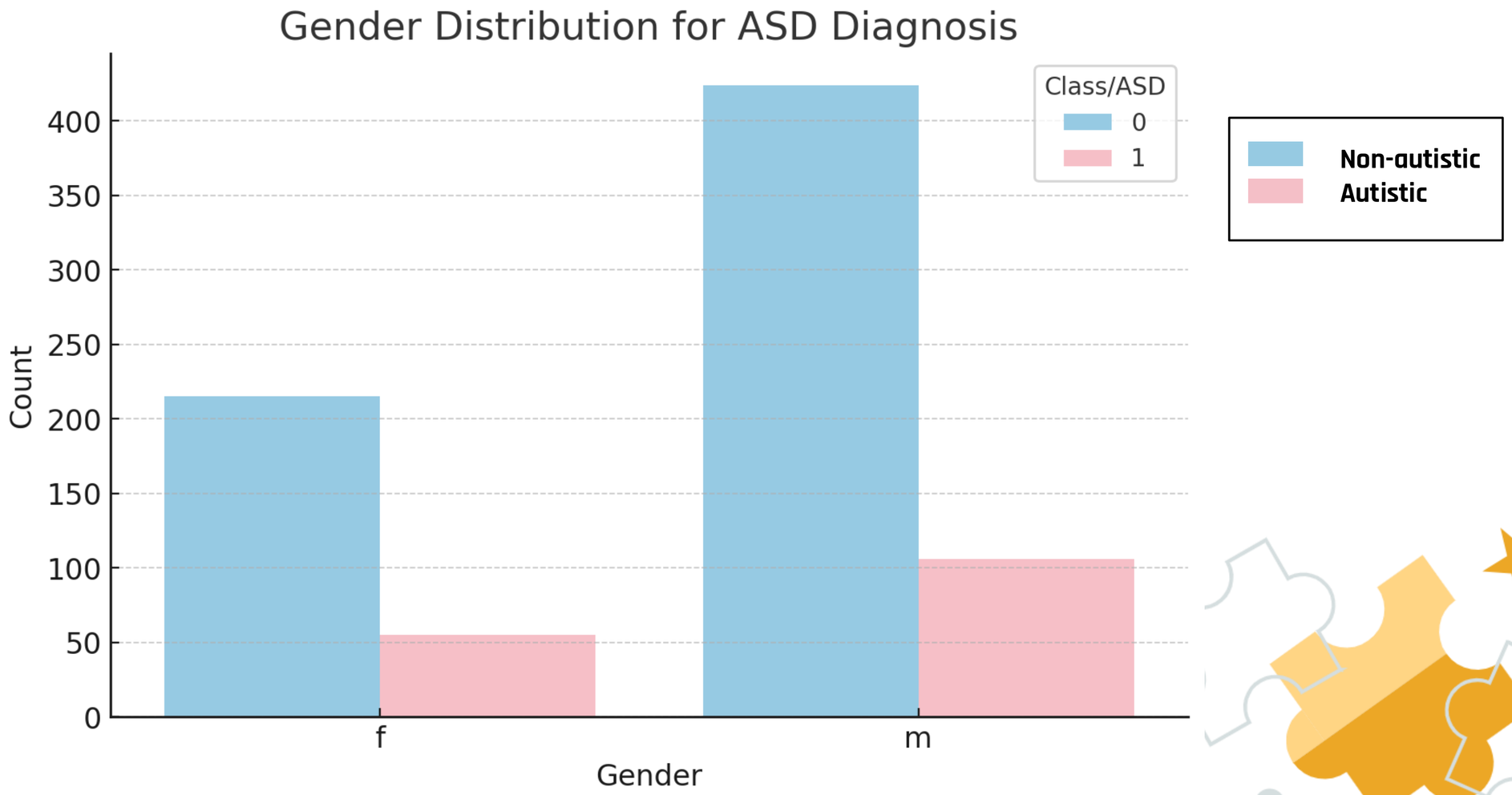
Average Score

We can say that the model is very good at correctly identifying classes. This means that, on average, it gets it right more than 90% of the time when distinguishing between the classes we are predicting.

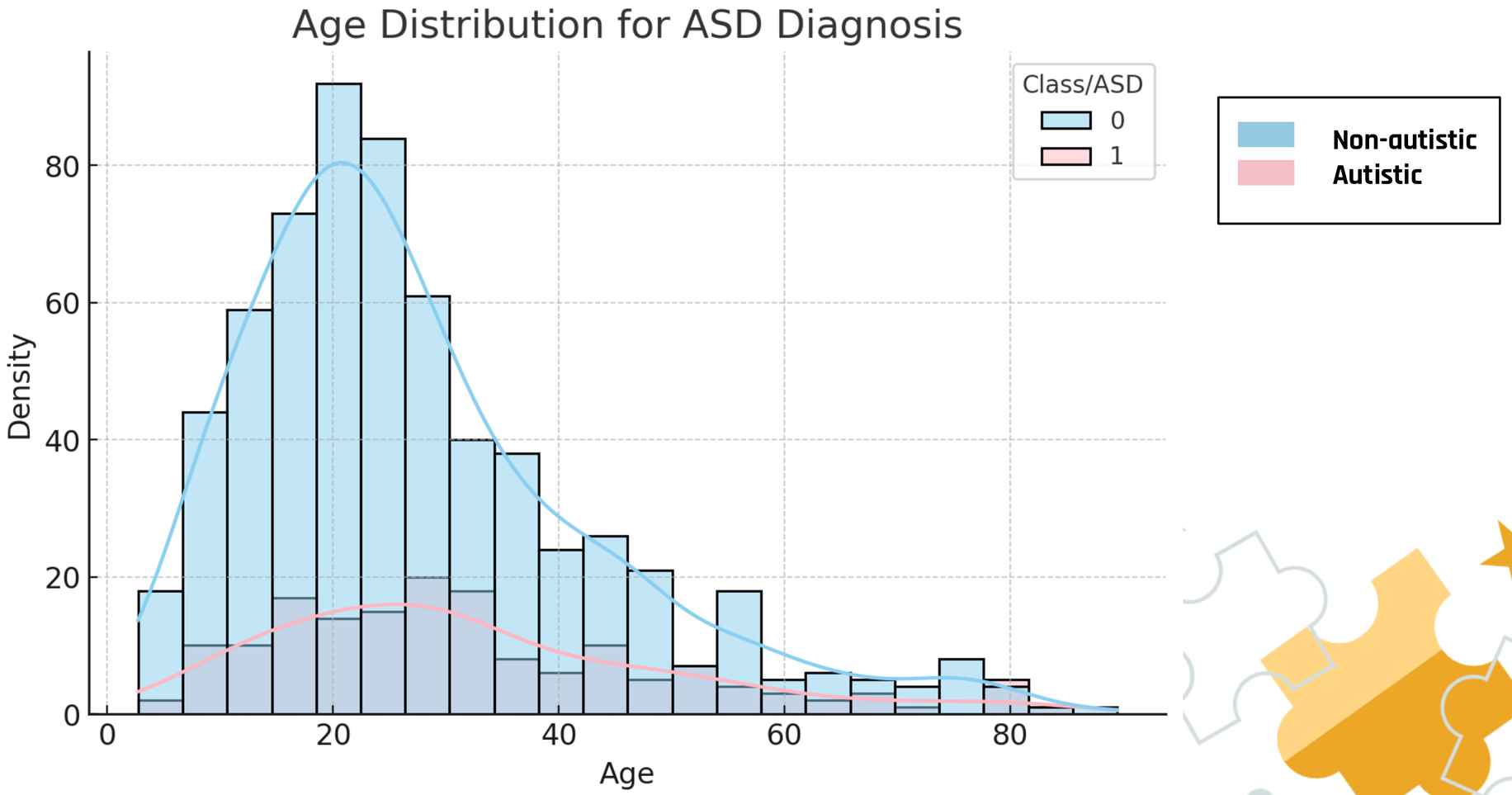
1. How many users have autism?



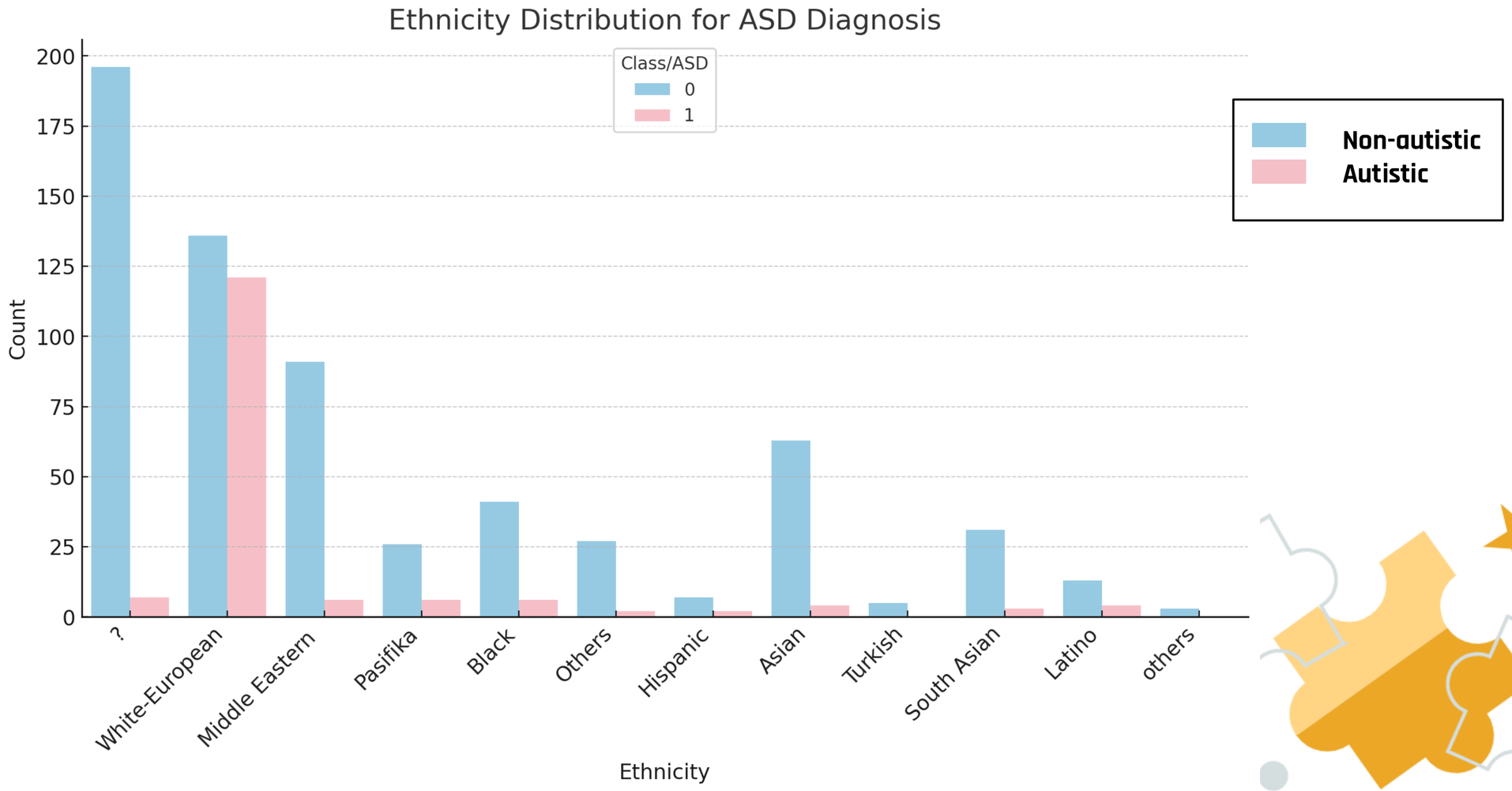
2. What is the gender distribution among individuals with autism?



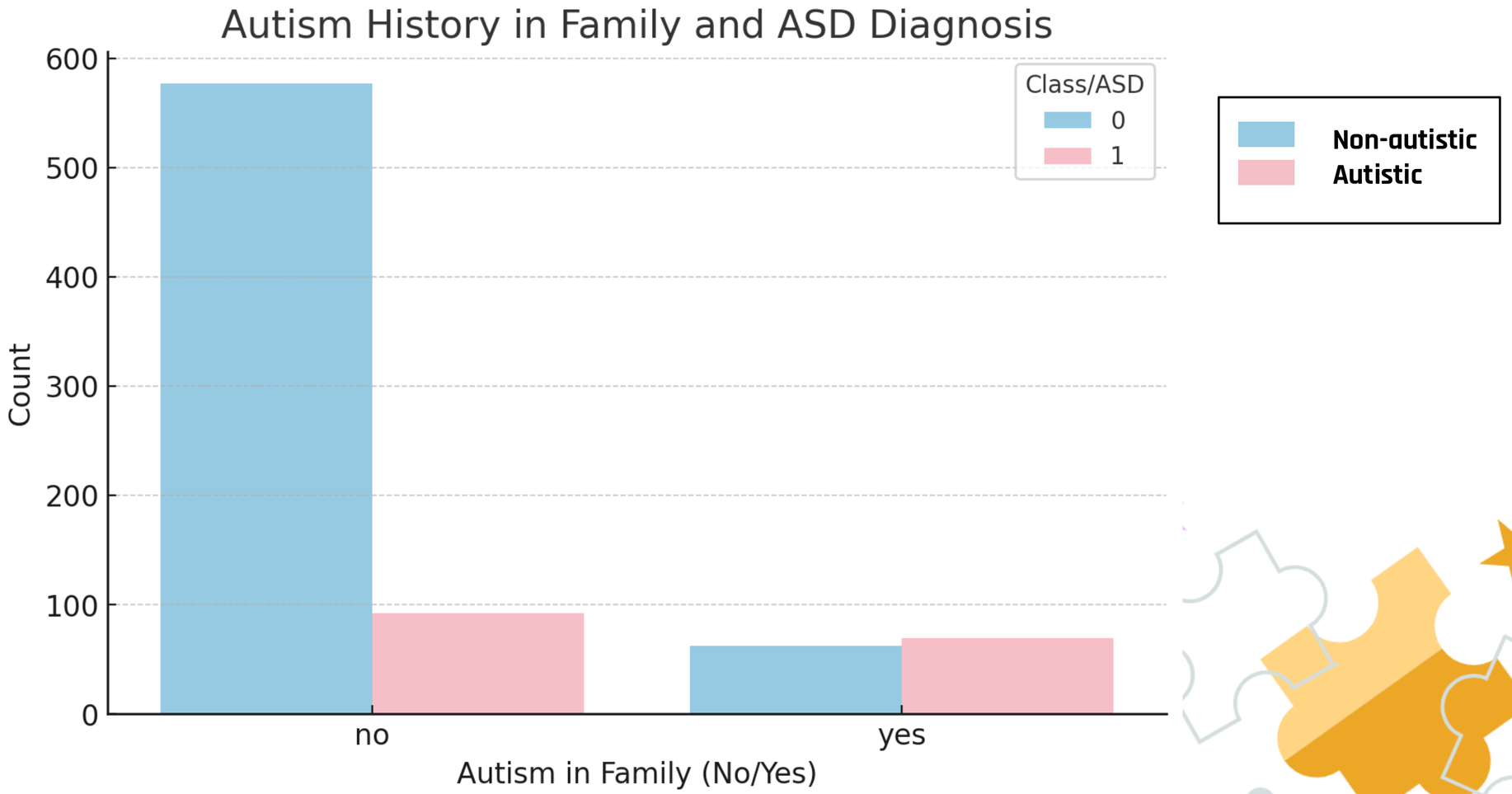
3. What is the average age of individuals with autism versus those without?



4. What are the main ethnicities among individuals diagnosed with autism?



5. Is there a link between a family history of autism and a diagnosis of ASD?



Streamlit app

“I don’t have a different brain; I have a brain that thinks differently.”

— Temple Grandin, *Thinking in Pictures*

References

What is Autism? | Cincinnati Children's: **https://www.youtube.com/watch?v=hwaaphuStxY**

Autistic Spectrum Disorder Screening Data for Children: **https://archive.ics.uci.edu/dataset/419/autistic%2Bspectrum%2Bdisorder%2Bscreening%2Bdata%2Bfor%2Bchildren**

Speeding Autism Diagnosis, Improving Outcomes Using Machine Learning: **https://today.duke.edu/2019/07/speeding-autism-diagnosis-improving-outcomes-using-machine-learning**

Data Analysis and Classification of Autism Spectrum Disorder: **https://www.researchgate.net/publication/338435130**

APPDA Lisbon: **https://appda-lisboa.org.pt/autismo/diagnostico**

Diagnostic Approach and Intervention in Autism Spectrum Disorder in Pediatric and Adult Age (DGS - Portugal): **https://normas.dgs.min-saude.pt/wp-content/uploads/2019/09/Abordagem-Diagnostica-e-Intervencao-na-Perturbacao-do-Espetro-do-Autismo-em-Idade-Pediatrica-e-no-Adulto 2019.pdf**

Thank you for your attention!