PREDICTING ADHD

Model development to predict both an individual's sex and their ADHD diagnosis using functional brain imaging data of children and adolescents, and their socio-demographic, emotions, and parenting information.

Sheila Acar [1, 2,†]

1 Northwestern University School of Professional Studies
Master of Science in Data Science Program

2 www.linkedin.com/in/sheila-acar

† sheilamaeacar@gmail.com

Abstract

ADHD, attention-deficit/hyperactivity disorder, is a common neurodevelopmental disorder, yet females are frequently underdiagnosed due to differences in symptom presentation. While males typically exhibit hyperactivity and impulsivity, females more often display inattention and internalizing behaviors. These symptoms, combined with co-occurring conditions and compensatory masking strategies, contribute to delayed or missed diagnoses. This leads to significant long-term consequences for females. Existing diagnostic criteria, largely based on male presentations, further exacerbate this disparity.

This study, developed for the 2025 WiDS Datathon, aims to improve ADHD classification and address diagnostic bias by leveraging functional brain imaging data, socio-demographic factors, emotional assessments, and parenting information. Unlike previous studies that focus solely on functional MRI connectome matrices or EKG data, this approach integrates behavioral and neuroimaging features to enhance prediction accuracy.

Multiple machine learning models, including Random Forest, Support Vector Machines, Logistic Regression, and Graph Neural Networks, were trained to classify ADHD diagnosis and sex. Feature selection methods, such as Recursive Feature Elimination and SHAP-based analysis, were applied to identify the most predictive variables. Results show that Random Forest outperformed deep learning models like Graph Neural Networks, achieving the highest accuracy at 90% for ADHD classification and 82.71% for gender classification when using SHAP-based feature selection.

This research highlights the importance of incorporating behavioral and neuroimaging data for improving sex-specific ADHD detection. The predictive model developed in this study, along with future advancements, has the potential to improve early ADHD diagnosis, support timely interventions, personalize treatment approaches, and enhance long-term outcomes for females with ADHD.

Table of Contents

Introduction and Problem Statement

ADHD is one of the most common childhood disorders. In the United States, at least 11% of children are diagnosed with ADHD, with males making up the majority of the cases. In childhood, ADHD is diagnosed three to sixteen times more frequently in males compared to females, yet in adulthood, nearly equivalent numbers of males and females receive diagnose (Babinski 2021). In a recent Swedish population registry study involving 37,591 females identified with a diagnosis of ADHD, findings show that females were diagnosed approximately four years later than males with the disorder.

Females with ADHD often go undiagnosed in childhood due to differences in symptom presentation. While males typically exhibit hyperactivity and impulsivity, females more commonly display inattention and internalizing behaviors such as daydreaming, disorganization, anxiety, and emotional distress. These symptoms are less disruptive, making ADHD harder to recognize. The issue is further compounded by co-occurring conditions, including anxiety and depression, as well as compensatory masking strategies that help females adapt but mask their struggles. As a result, ADHD is diagnosed far more frequently in males, and diagnostic criteria remain largely based on male symptoms, which contribute to underdiagnosis and misdiagnosis in females (Young et al. 2020).

The consequences of undiagnosed ADHD in females are severe. Without proper diagnosis and intervention, females are more likely to experience chronic anxiety, depression, and social withdrawal (Quinn 2005). Long-term impacts include higher rates of self-harm, early engagement in sexual behavior, intimate partner violence, academic underachievement, unintended pregnancy, and employment difficulties. Many struggle academically not because of cognitive deficits, but due to unrecognized executive functioning challenges. As academic and social demands increase with age, coping mechanisms often fail, leading to greater distress and functional impairment.

This research is a contribution to the 2025 WiDS Datathon. The goal is to develop a predictive model for ADHD diagnosis and sex classification using functional brain imaging data of children and adolescents, along with socio-demographic, emotions, and parenting information. Datasets and support for this challenge are provided by the Healthy Brain Network (HBN), the signature

scientific initiative of the Child Mind Institute, and the Reproducible Brain Charts project (RBC).

Unlike previous studies that focus solely on EKG or functional MRI connectome matrices, this study integrates behavioral or demographic data, to enhance ADHD prediction accuracy and reduce diagnostic bias. Such a robust predictive tool can reduce misdiagnosis and provide earlier support females with ADHD during childhood. This could lead to timely interventions, improved academic and societal outcomes, and a better understanding of sex-specific ADHD presentations. In addition, such a tool can assist professionals in making more accurate assessments, reduce bias in diagnostics, and may contribute to the development of personalized treatment strategies. Improvement in ADHD diagnosis in females in childhood can significantly enhance the lives of those affected, as ADHD is a highly treatable condition. Outcomes can be improved with medications, lifestyle adjustments, and appropriate accommodations.

Several preprocessing methods and models were tested and developed to predict two separate target variables. Ultimately, the model with highest accuracy may serve as a promising tool for diagnosing females with ADHD in childhood.

Literature Review

In their 2022 study, Salman et al. evaluated the performance of various machine learning models in predicting ADHD diagnoses. The models they tested include decision trees, random forests, support vector machines, and multilayer perceptrons, using data from individuals with ADHD and control groups. Their results showed that the support vector machine model performed the best, achieving a 91% accuracy rate. This study highlights the potential of machine learning to improve ADHD diagnosis and suggests that AI could play a valuable role in clinical settings for more reliable identification. While Salman et al. focused on model accuracy, Thomas et al. addressed bias mitigation strategies.

Thomas et al. explored how machine learning can be used to predict ADHD in school-aged children by analyzing education and health data from over 56,000 students in South London. They tested multiple models and found that random forest and logistic regression performed the best, with AUC scores of 0.86 in population samples and 0.72 in clinical samples. One key aspect of their study was addressing potential biases in predictions. They applied a fairness algorithm, and reduced biases related to

ethnicity and language without sacrificing accuracy. This research shows the potential of using routinely collected data to build scalable, accurate, and more equitable models for identifying children at risk of ADHD, which could help improve early diagnosis and resource allocation. Like their study, the dataset for this research includes sociodemographic data, which could provide valuable insights into potential disparities in ADHD diagnosis. Incorporating these features into machine learning models allows for evaluating the impact of demographic factors on predictions.

Data

The data provided by the WiDS competition includes both the training and test datasets. The training data consists of three types of information for over 1,200 subjects: target variables (ADHD diagnosis and sex), functional MRI connectome matrices, and socio-demographic information, emotional, and parenting information. The test data includes unseen data frames for more than 300 subjects. The datasets include both quantitative and categorical data.

The quantitative data includes the Laterality Index (Score), color vision test score, and responses from the Strengths and Difficulties Questionnaire and the Alabama Parenting Questionnaire - Parent Report. The Strengths and Difficulties Questionnaire (SDQ) is a brief behavioral screening tool that assesses emotional and behavioral difficulties in children and adolescents aged 2 to 17. The SDQ assesses emotional and behavioral problems while also highlighting strengths. It consists of five scales: pro-social behavior, hyperactivity, emotional problems, conduct or behavioral problems, and peer problems. Parenting data is derived from responses to the “Alabama Parenting Questionnaire”, a copyrighted 42-item measure that evaluates key parenting dimensions linked to conduct problems and delinquency in youth. These dimensions include Positive Reinforcement, Parental Involvement, Inconsistent Discipline, Poor Monitoring and Supervision, and Harsh Discipline.

The categorical variables include the year of enrollment, indicating when a participant joined the study; the site of phenotypic testing, showing where assessments were conducted; ethnicity and race of the child; scan location, specifying where scans took place; and parental background details. Parental background is captured through Parent 1 and Parent 2 level of education, which reflect their highest educational attainment,

and Parent 1 and Parent 2 occupation, which categorize their respective job roles.

Finally, the datasets include functional MRI connectome matrices. Functional MRI (fMRI) data is used to measure brain activity by detecting changes in blood oxygenation levels over time. From this data, time-series signals are extracted for different brain region, which represent fluctuations in neural activity. These time-series signals are then analyzed to assess the functional relationships between regions by computing correlations, which indicate how synchronously different areas of the brain are activated. The resulting correlation values are used to construct functional MRI connectome matrices**,** which map the strength and patterns of connectivity across the brain and provide insights into neural networks and their interactions.

Methods

Data Processing

To begin exploratory data analysis, the datasets were combined and cleaned, with several preprocessing steps applied to ensure data quality and consistency. First, 11 missing values in the *Ethnicity of Child*(PreInt_Demos_Fam_Child_Ethnicity) variable were removed. Additionally, 360 missing values in *Age at Time of MRI Scan* (MRI_Track_Age_at_Scan) were replaced with the median based on the distribution's skewness of 0.53. Visualizations were then used to examine the distribution of categorical, quantitative, and target variables. Figures 1-13 in the appendix present these distributions, providing insights into the data characteristics. The most notable imbalance was observed in the target variables, illustrated in Figure 14. This imbalance could potentially bias the model's predictions. To mitigate this, Synthetic Minority Over-sampling Technique (SMOTE) was applied to generate synthetic samples for the underrepresented class and ensure a more balanced distribution to improve model performance.

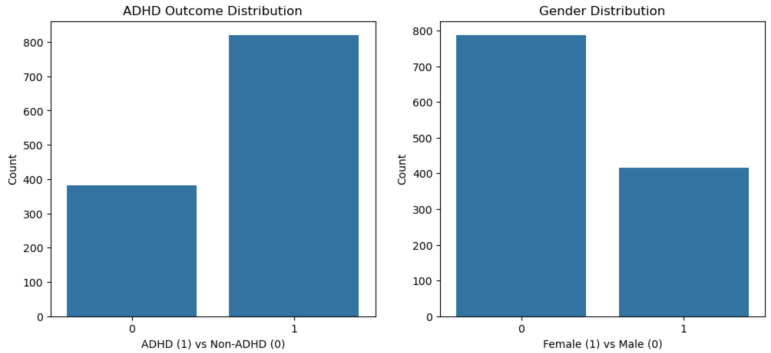

Figure 14. Imbalance in target variable distribution, highlighting the disparity between class frequencies before applying resampling techniques.

Two different strategies were applied to the connectivity matrices to analyze brain connectivity patterns. First, using NetworkX, the connectivity matrices were converted into graphs by filtering out weak connections, below 0.03, to reduce noise and emphasize stronger functional relationships between brain regions. The brain's functional connectivity for an individual participant is visualized in Figure 15, where nodes represent brain regions and edges indicate significant correlations in fMRI-derived time-series activity. Key graph-theoretic features, network density, clustering coefficient, degree centrality, and average path length, were extracted to quantify connectivity patterns. These features were then used as input for traditional machine learning models including random forests, logistic regression, and support vector machines (SVMs).

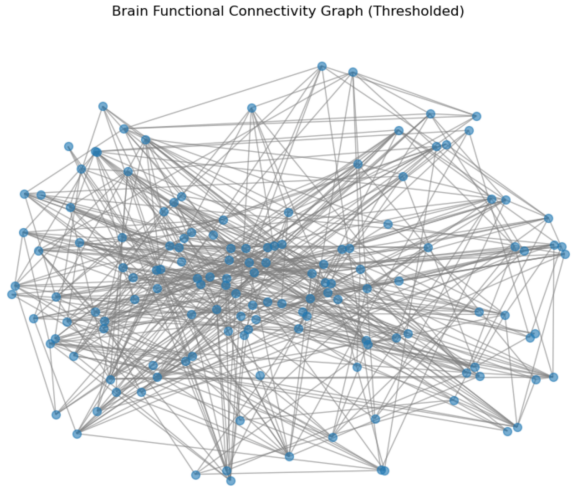

Figure 15. Thresholded brain functional connectivity graph for a single participant. Nodes represent brain regions, and edges indicate significant correlations in fMRI-derived time-series activity.

Another strategy applied is Graph Neural Networks (GNNs), which preserved the full graph structure rather than summarizing it with statistical features. GNNs leveraged spatial dependencies between brain regions and learned meaningful node and edge representations to capture deeper connectivity patterns. This approach allowed models to adaptively learn relationships within the brain network.

Model Development

The initial experiments focused on evaluating traditional machine learning methods, including random forests, logistic regression, and SVMs. The models used features extracted from NetworkX, which captured the structural properties of brain connectivity graphs and served as input for classification models. To improve model performance, Recursive Feature Elimination (RFE) was applied with 10, 15, and 20 features to systematically rank and remove the least important predictors and retain only the most informative ones. Each subset was then used to train the same traditional models which allows for a comparative assessment of how feature reduction impacted classification accuracy. Next, SHAP (SHapley Additive exPlanations)-based feature selection was tested using the 10 most impactful features identified by SHAP values. This method provided insights into feature importance by quantifying each predictor's contribution to the model's decisions. Despite testing different feature subsets, adding more features did not improve performance for SHAP models.

Beyond traditional machine learning, Graph Neural Networks (GNNs) were implemented to fully utilize the connectivity matrices while preserving their graph structure. Unlike feature-based models, GNNs processed connectivity graphs directly and captured spatial dependencies between brain regions through graph convolutional layers. A series of experiments were conducted to optimize the GNN architecture: (1) a base model established a benchmark, (2) a learning rate scheduler dynamically adjusted the learning rate for better convergence, (3) dropout regularization was added to reduce overfitting, (4) a deeper architecture with four convolutional layers (64 units each) was tested for enhanced feature extraction, and (5) a shallower but wider model with two convolutional layers (256 units each) was evaluated for capturing complex patterns. These

experiments were designed to improve GNN's ability to model functional connectivity and improve classification performance.

Results

The results indicate that Random Forest outperformed all other models in both ADHD and gender classification across different approaches. In traditional machine learning models, Random Forest achieved the highest accuracy, with 90% for ADHD and 79% for gender, which makes it a strong baseline. SVM and Logistic Regression both had an ADHD accuracy of 84%, but SVM had lower gender accuracy of 65% compared to Logistic Regression of 67%. Among these models, Logistic Regression was the simplest, while SVM provided a more balanced trade-off.

Feature selection techniques, such as RFE and SHAP-based selection, were implemented to improve model performance. With RFE, Random Forest maintained the highest ADHD accuracy, improving from 0.7571 with 10 features to 0.7834 with 20 features, while SVM and Logistic Regression peaked at 0.6849 and 0.7024, respectively**.** A similar trend was observed for gender accuracy. The SHAP-based approach further improved results, with Random Forest achieving 87.75% ADHD accuracy and 82.71% gender accuracy, outperforming SVM and Logistic Regression, which both had slightly lower ADHD accuracy of around 83.37% and 83.81% and significantly lower gender accuracy of 70.02% and 68.71%, respectively.

In contrast, GNNs did not perform as well as traditional models. The ADHD accuracy remained around 68.65% across most architectures, while gender accuracy hovered around 66.73%. Incorporating dropout layers significantly reduced accuracy to 50% for ADHD and 49.96% for gender which suggests that the deep learning approach struggled to capture meaningful patterns in this dataset. Overall, Random Forest consistently demonstrated superior predictive performance, with feature selection, specifically SHAP, enhancing accuracy further. While GNNs offer a promising direction, their lower performance suggests that traditional machine learning models remain more effective for this predictive task.

Discussion

This study evaluated the performance of Random Forest, SVM, Logistic Regression, and GNNs for predicting ADHD and gender

classification using functional MRI data. Random Forest consistently outperformed other models, achieving the highest accuracy across different evaluation settings. Feature selection techniques, such as RFE and SHAP-based selection, had minimal impact on improving ADHD classification accuracy which suggests that reducing the number of features did not significantly enhance model performance. These findings may be attributed to the distributed nature of ADHD-related brain activity, where no single set of features strongly dominates prediction. ADHD appears to be linked to a broader network of subtle neural interactions, meaning that weak but collectively important signals might still contribute to classification. Since RFE systematically removes the least important features, it may have discarded important but weak signals that are essential for modeling ADHD. This may explain why RFE did not improve accuracy.

In contrast, SHAP-based selection proved beneficial for gender classification, likely because gender-related differences in brain connectivity are more localized to specific brain regions rather than being widely distributed like ADHD-related patterns. By isolating the most impactful predictors, SHAP allowed models to focus on the most relevant features for distinguishing male and female participants, leading to higher classification accuracy.

While traditional machine learning models performed well, GNNs underperformed across all configurations. The limited accuracy improvements suggest that deep learning models may require larger datasets or more complex architectures to effectively capture meaningful patterns in brain connectivity data. In addition, the significant drop in performance when applying dropout layers suggests that the GNN models may have struggled with over-regularization which prevented them from learning critical connectivity patterns. Unlike feature-based models, GNNs rely on graph structure and node relationships, so simplified connectivity graphs or a lack of sufficient labeled data may have further limited their effectiveness. Future work should explore graph augmentation techniques, deeper architectures, and additional training data to enhance GNN performance.

Overall, Random Forest remains the most reliable model for ADHD classification, with only marginal gains from feature selection methods. SHAP-based selection proved beneficial for improving gender classification. The results highlight the importance of considering the distributed nature of ADHD-related brain connectivity and the value of feature selection in tasks where localized patterns are more relevant, such as gender classification.

Conclusions

This study addressed the underdiagnosis of ADHD in females, a critical issue resulting from diagnostic criteria that are largely based on how ADHD presents in males. Females with ADHD often exhibit more subtle and internalized symptoms, such as inattention, anxiety, and disorganization, rather than the hyperactive and impulsive behaviors commonly seen in males. These differences, along with co-occurring conditions and compensatory masking strategies, contribute to delayed or missed diagnoses in females which lead to long-term academic, social, and mental health consequences. A robust predictive tool for ADHD that accounts for sex-specific presentations can help reduce diagnostic bias, enable earlier intervention, and support personalized treatment strategies.

This study explored various machine learning models, including Random Forest, SVM, Logistic Regression, and GNNs, to classify ADHD and gender using functional MRI data and socio-demographic factors. Random Forest consistently achieved the highest accuracy, making it the most effective model for ADHD classification. Random Forest consistently achieved the highest accuracy, with 90% for ADHD classification and 79% for gender classification, making it the most effective model. Feature selection techniques, such as RFE and SHAP-based selection, had minimal impact on ADHD prediction, likely because ADHD-related neural activity is distributed across multiple brain regions rather than localized to specific areas. However, SHAP-based selection improved gender classification accuracy, which reinforces the idea that sex-based neural differences may be more regionally specific.

While GNNs showed lower accuracy than traditional models, their ability to capture spatial dependencies in brain connectivity suggests that further optimizations, such as larger datasets, alternative graph structures, and hyperparameter tuning, could enhance their effectiveness.

Overall, this study highlights the importance of using machine learning to refine ADHD diagnosis, particularly in females, and demonstrates that traditional models like Random Forest remain the most effective for this task.

Directions for Future Work

Future research should focus on expanding datasets, refining deep learning models, improving feature selection, and

addressing biases to enhance ADHD diagnosis. One key area for improvement is expanding and diversifying datasets. Current ADHD datasets may be demographically imbalanced, which can introduce sampling bias and limit generalizability. Incorporating multi-site neuroimaging data, longitudinal studies, and larger, more diverse populations can help improve model performance and fairness. To address class imbalance, future work should also explore other methods beyond SMOTE to ensure better representation of underdiagnosed groups.

While Random Forest demonstrated strong predictive performance, GNNs underperformed, likely due to over-regularization, limited labeled data, and suboptimal graph representations. Future research should investigate alternative GNN architectures and hyperparameter tuning to better capture spatial dependencies in brain connectivity.

Feature selection results showed that RFE had minimal impact on ADHD classification. Future studies should explore alternative feature selection techniques to refine model interpretability without losing important predictive information.

By incorporating larger datasets, advanced deep learning techniques, and refined feature selection strategies, future studies can develop more accurate and clinically useful ADHD diagnostic tools.

Acknowledgements

The WiDS Datathon Global Challenge was developed in collaboration with the Ann S. Bowers Women's Brain Health Initiative (WBHI), Cornell University, and UC Santa Barbara. This project benefited from datasets and support provided by the Healthy Brain Network (HBN), the signature scientific initiative of the Child Mind Institute, and the Reproducible Brain Charts (RBC) project.

Data Availability

The datasets used in this study were provided by the Healthy Brain Network (HBN), the signature scientific initiative of the Child Mind Institute, and the Reproducible Brain Charts (RBC) project. These datasets were accessed through the WiDS Datathon 2025 competition on Kaggle (WiDS Datathon 2025). Data access is subject to the competition's terms and conditions, and any future use should comply with the respective data-sharing policies.

Code Availability

The code used for data processing, analysis, and modeling in this study was developed as part of the WiDS Datathon 2025 competition. Access can be provided upon request.

References

Babinski, D. E. 2024. "Sex differences in ADHD: Review and priorities for future research." *Current Psychiatry Reports*26: 151–156. https://doi.org/10.1007/s11920-024-01492-6.

Hinshaw, Stephen P., and Emily Rosenthal. 2021. "A 2021 review of ADHD: Key findings and future directions." *ADHD Girls and Women.* https://adhdgirlsandwomen.org/wp-content/uploads/2022/05/Hinshaw_2021_Review.pdf.

Quinn, Patricia O. 2005. "Treating adolescent girls and women with ADHD: Gender-specific issues." *Journal of Clinical Psychology* 61 (5): 579–87. https://doi.org/10.1002/jclp.20121.

Salman, S. A., Z. Lian, M. T. Ahvanooey, et al. 2022. "Prediction of attention deficit hyperactivity disorder using machine learning methods." *IEEE Access.* https://ieeexplore.ieee.org/document/10574571/.

Young, Susan, Nicola Adamo, Bryndís B. Ásgeirsdóttir, et al. 2020. "Females with ADHD: An expert consensus statement taking a lifespan approach providing guidance for the identification and treatment of attention-deficit/hyperactivity disorder in girls and women." *BMC Psychiatry* 20: 404. https://doi.org/10.1186/s12888-020-02707-9.

Appendix A

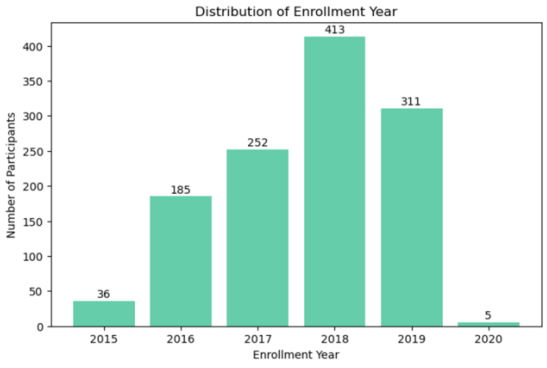

Figure 1. Distribution of Enrollment Year

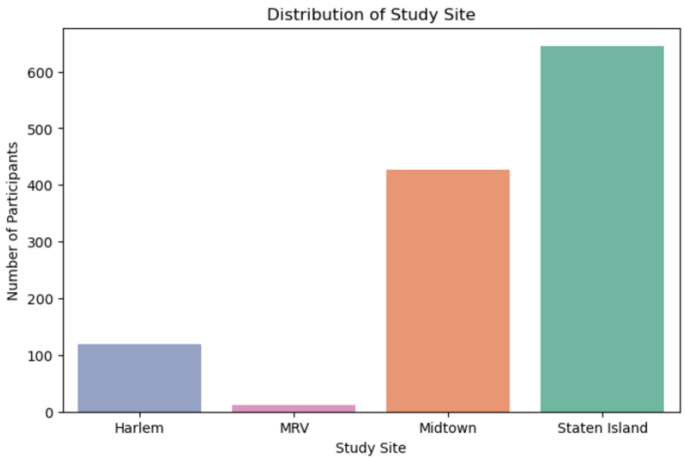

Figure 2. Distribution of Study Site

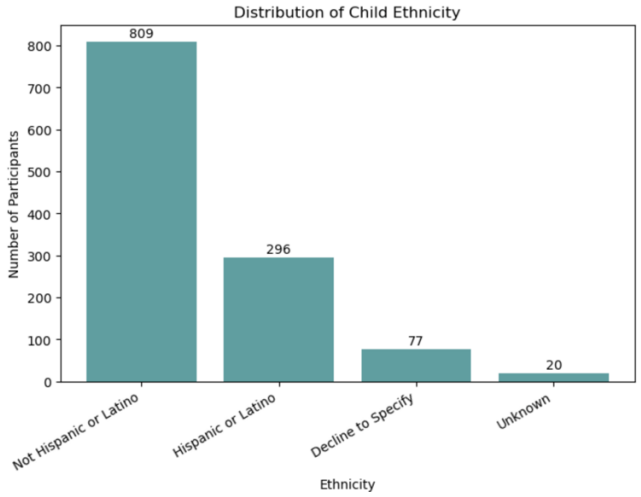

Figure 3. Distribution of Child Ethnicity

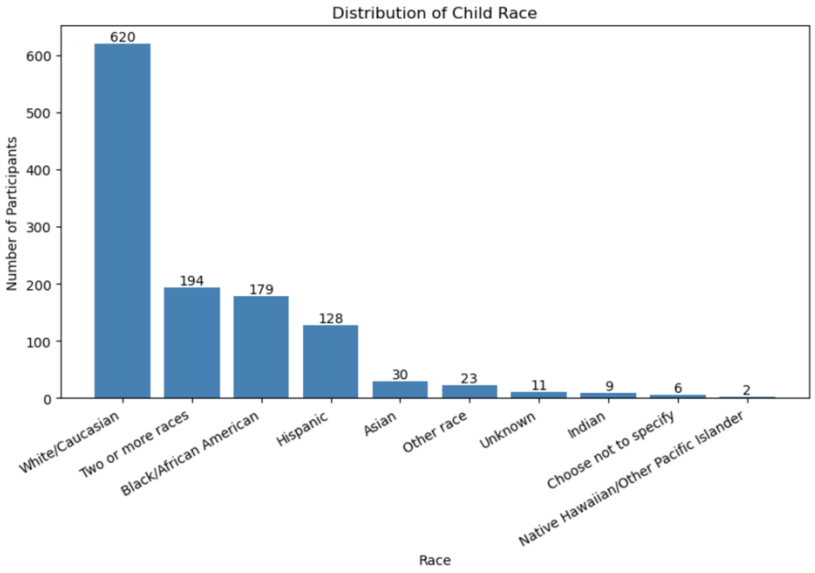

Figure 4. Distribution of Child Race

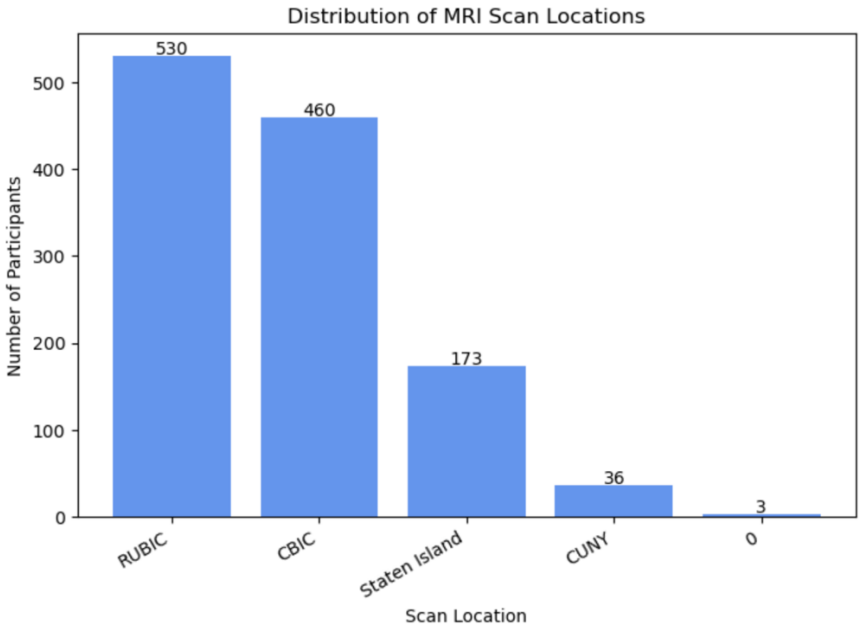

Figure 5. Distribution of MRI Scan Locations

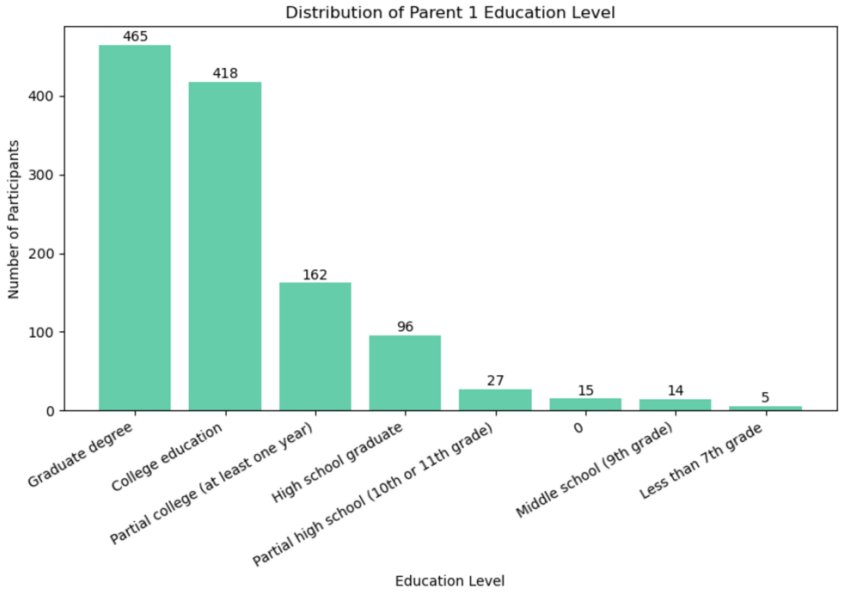

Figure 6. Distribution of Parent 1 Education Level

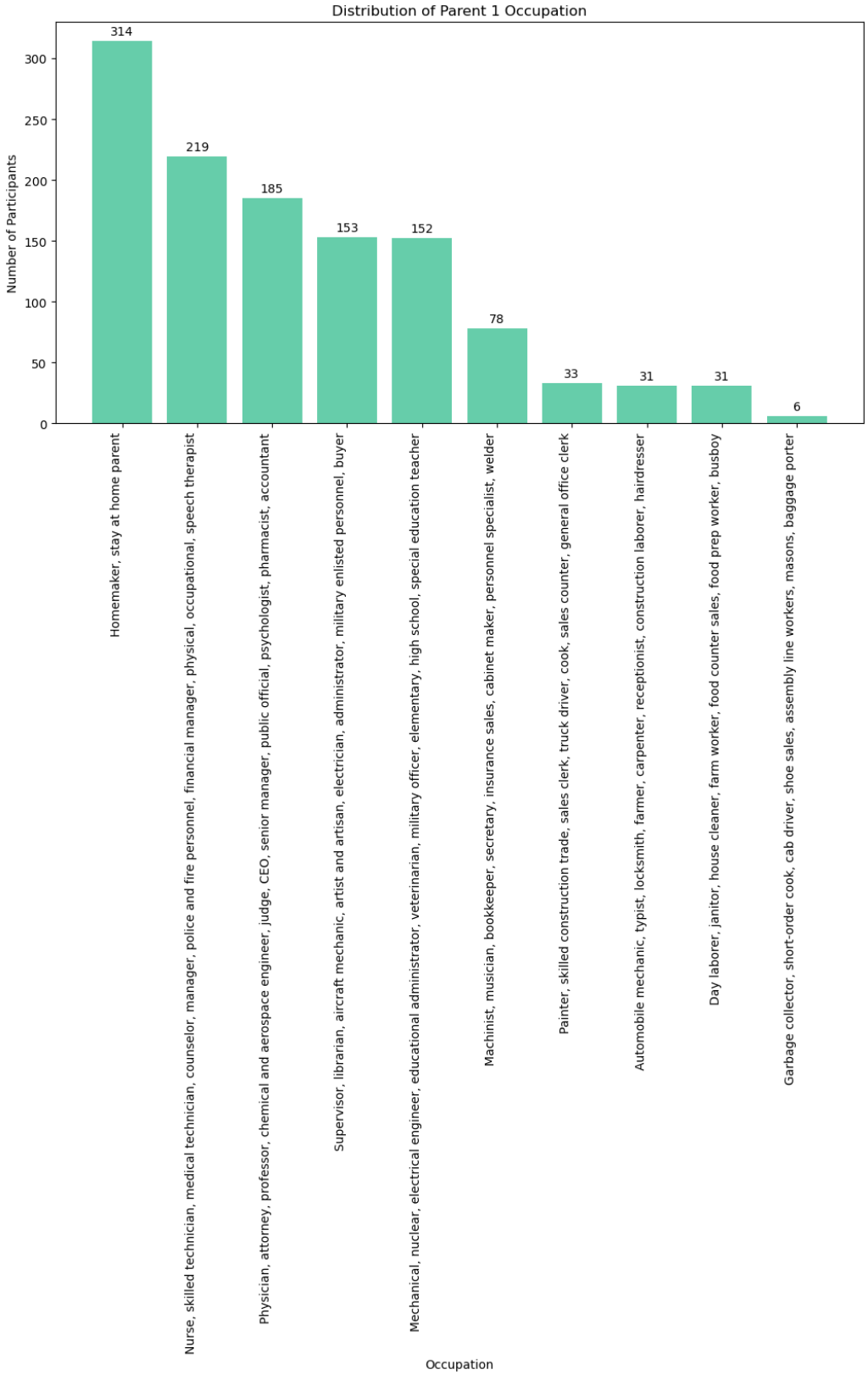

Figure 7. Distribution of Parent 1 Occupation

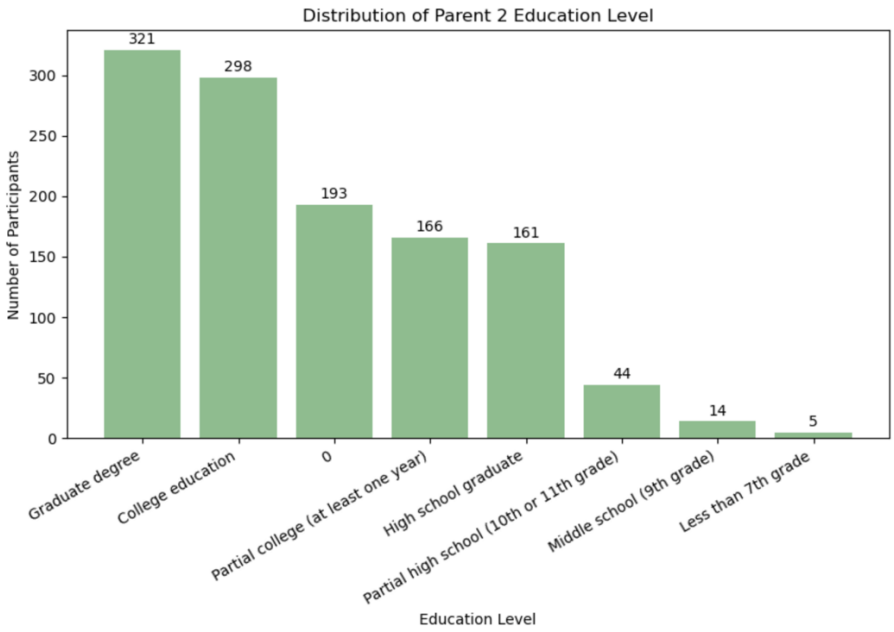

Figure 8. Distribution of Parent 2 Education Level

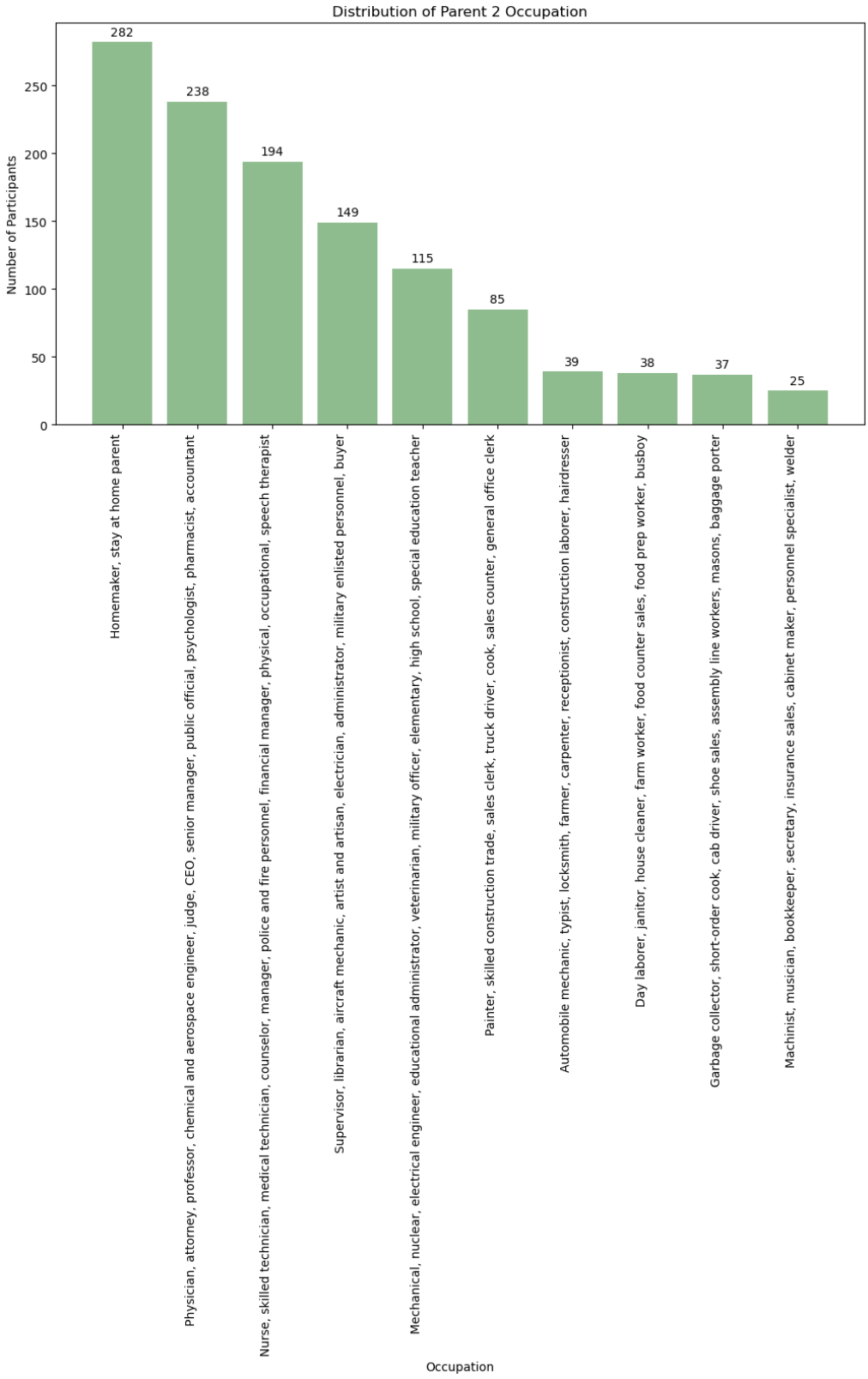

Figure 9. Distribution of Parent 2 Occupation

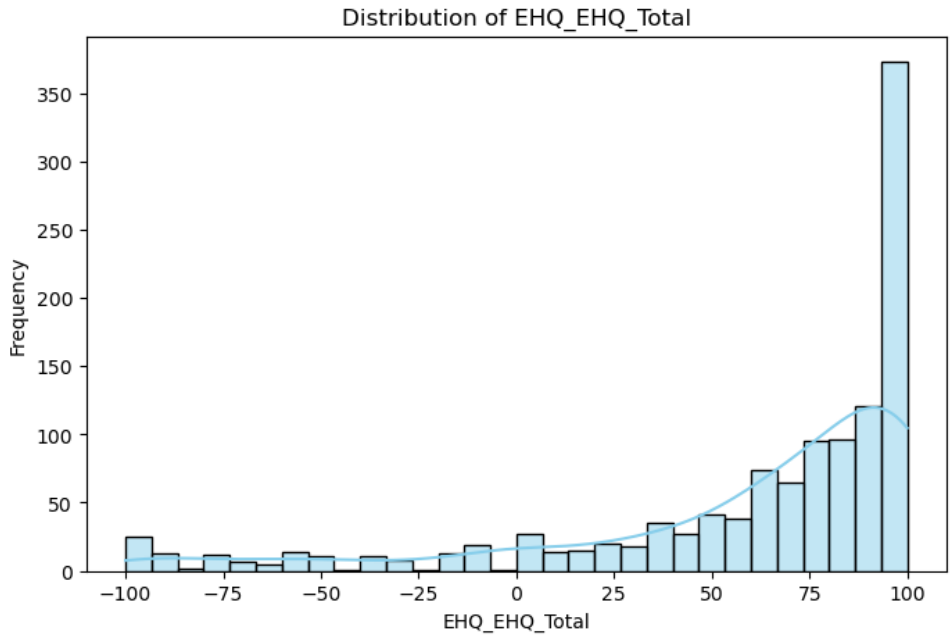

Figure 10. Distribution of EHQ_EHQ_Total

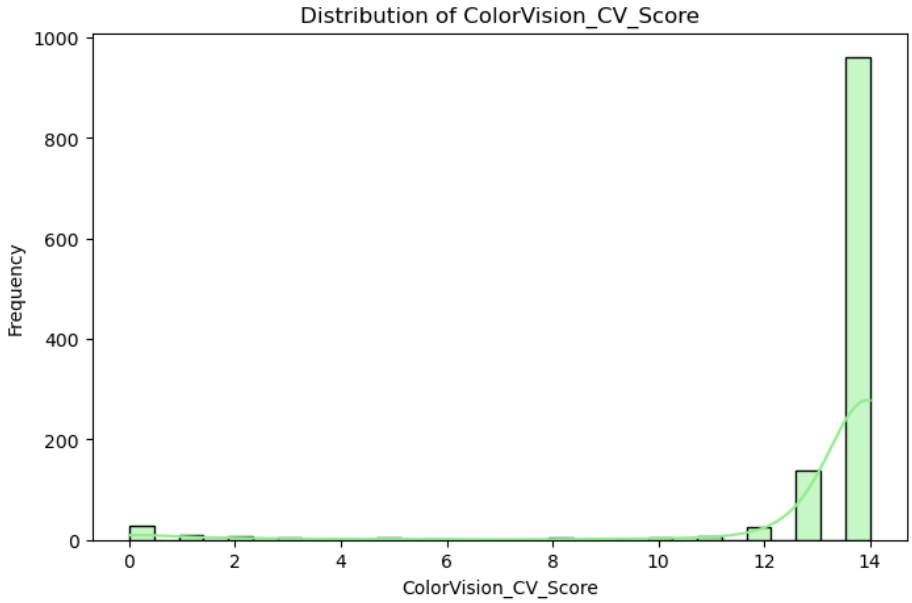

Figure 11. Distribution of ColorVision_CV_Score

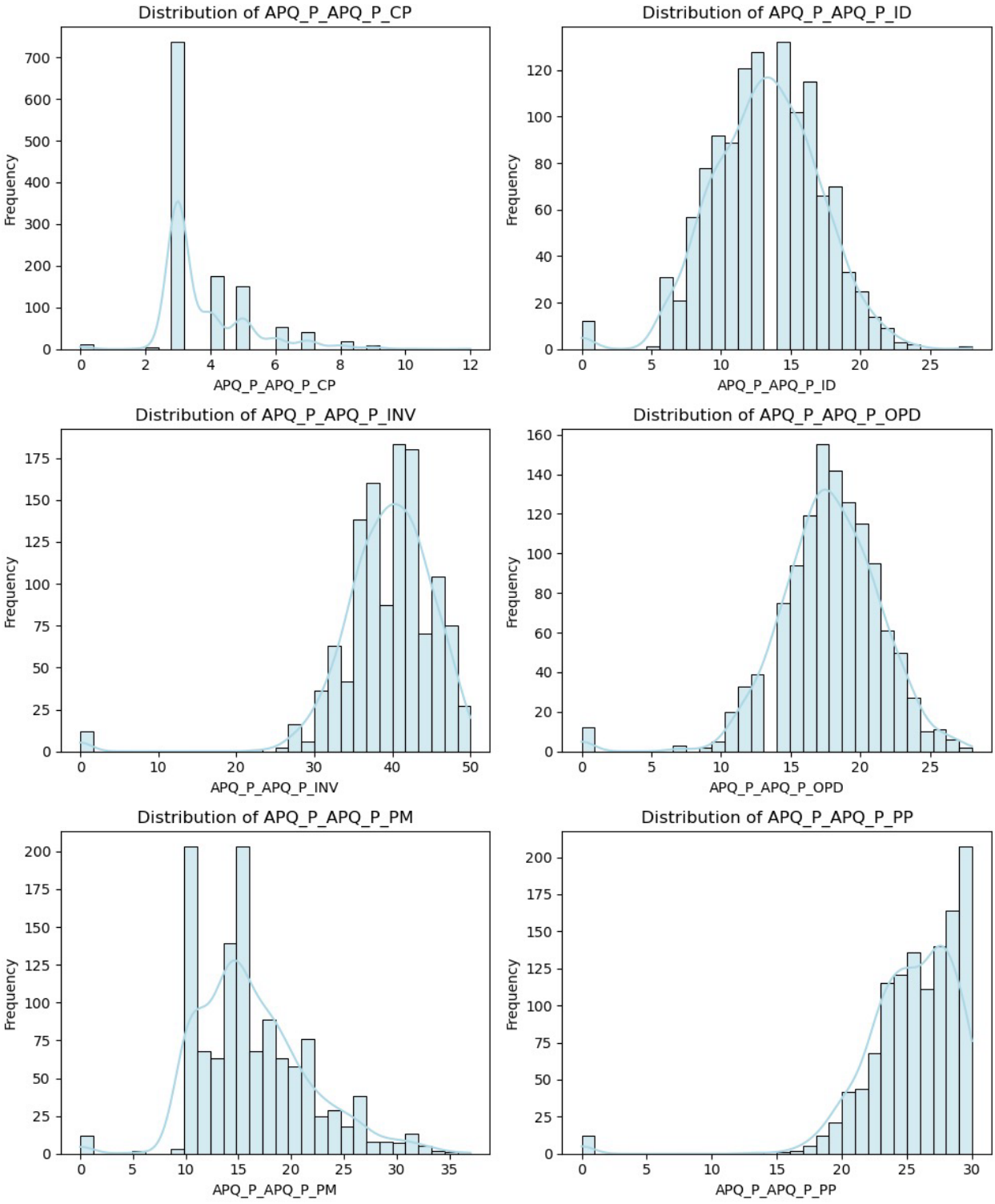

Figure 12. Distributions for APQ

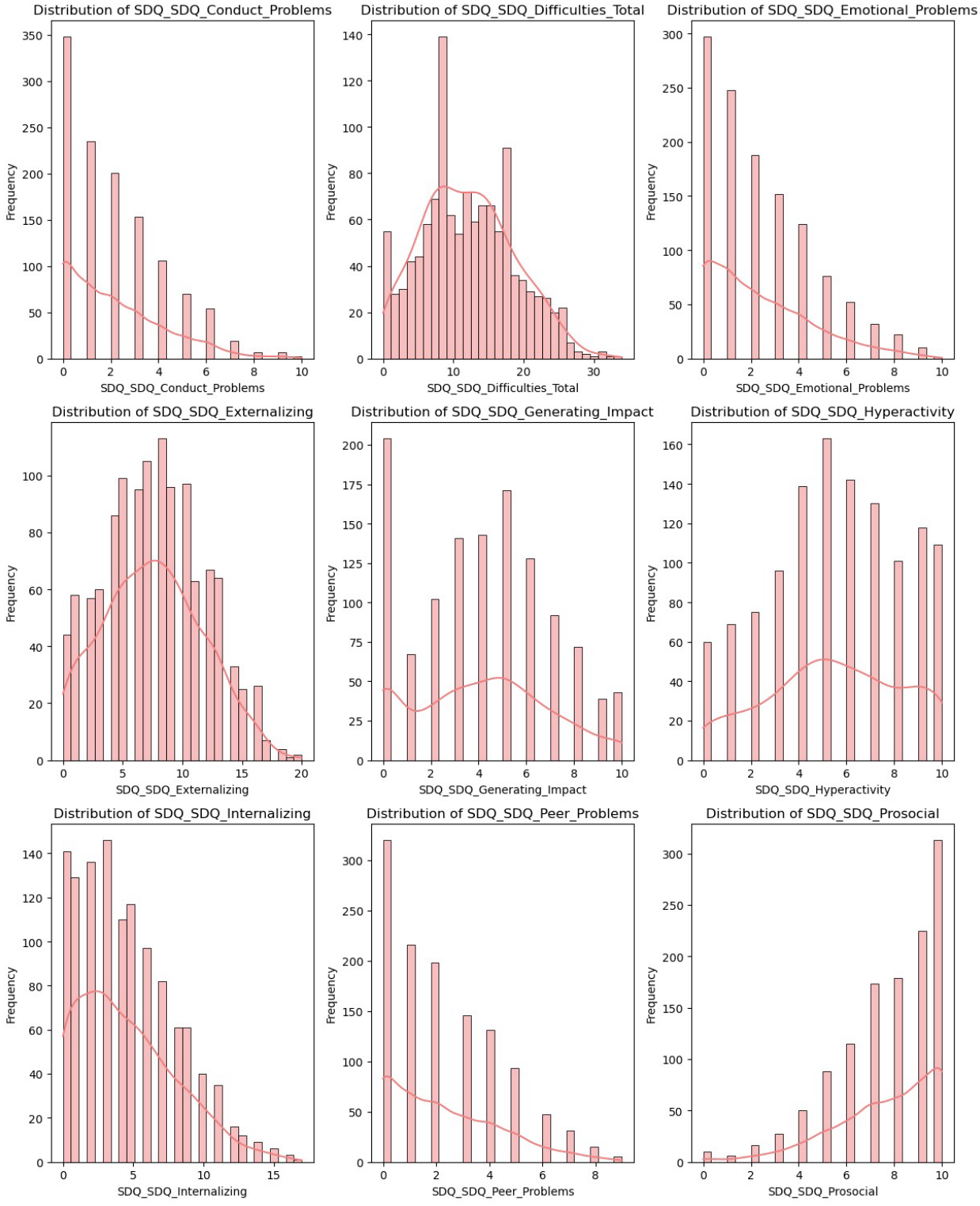

Figure 13. Distributions for SDQ

Model_Performance_Comparison

Model Type	ADHD Accuracy	Gender Accuracy	Best For
Random Forest (Raw Data)	90.0	79.0	Strongest ML baseline
SVM (Raw Data)	84.0	65.0	Balanced trade-off
Logistic Regression (Raw Data)	84.0	67.0	Simpler model

Figure 16. Performance of Traditional Machine Learning Models for ADHD and Gender Classification

ADHD_Accuracy_Results

Model	Accuracy (ADHD) 10 Features	Accuracy (ADHD) 15 Features	Accuracy (ADHD) 20 Features
Random Forest	0.7571	0.7593	0.7834
SVM	0.6586	0.6761	0.6849
Logistic Regression	0.7002	0.6805	0.7024

Figure 17. Performance of Traditional Machine Learning Models for ADHD Classification with Recursive Feature Elimination (RFE)

Gender_Accuracy_Results

Model	Accuracy (Gender) 10 Features	Accuracy (Gender) 15 Features	Accuracy (Gender) 20 Features
Random Forest	0.7571	0.7593	0.7834
SVM	0.6586	0.6761	0.6849
Logistic Regression	0.7002	0.6805	0.7024

Figure 18. Performance of Traditional Machine Learning Models for Gender Classification with Recursive Feature Elimination (RFE)

Model_Accuracy_Summary

Model	Accuracy (ADHD)	Accuracy (Gender)
Random Forest	0.8775	0.8271
SVM	0.8337	0.7002
Logistic Regression	0.8381	0.6871

Figure 19. Performance of Traditional Machine Learning Models for ADHD and Gender Classification with SHAP-Based Feature Selection

Model_Performance_Table

Model Type	ADHD Accuracy	Gender Accuracy
Base Model	68.65%	66.73%
Learning Rate Scheduler	68.65%	66.73%
Dropout	50.00%	49.96%
4 Layers with 64 units	68.70%	66.77%
2 Convolution Layers with 256 units	68.65%	66.73%

Figure 20. Performance of Traditional Machine Learning Models for ADHD and Gender Classification Using Graph Neural Networks

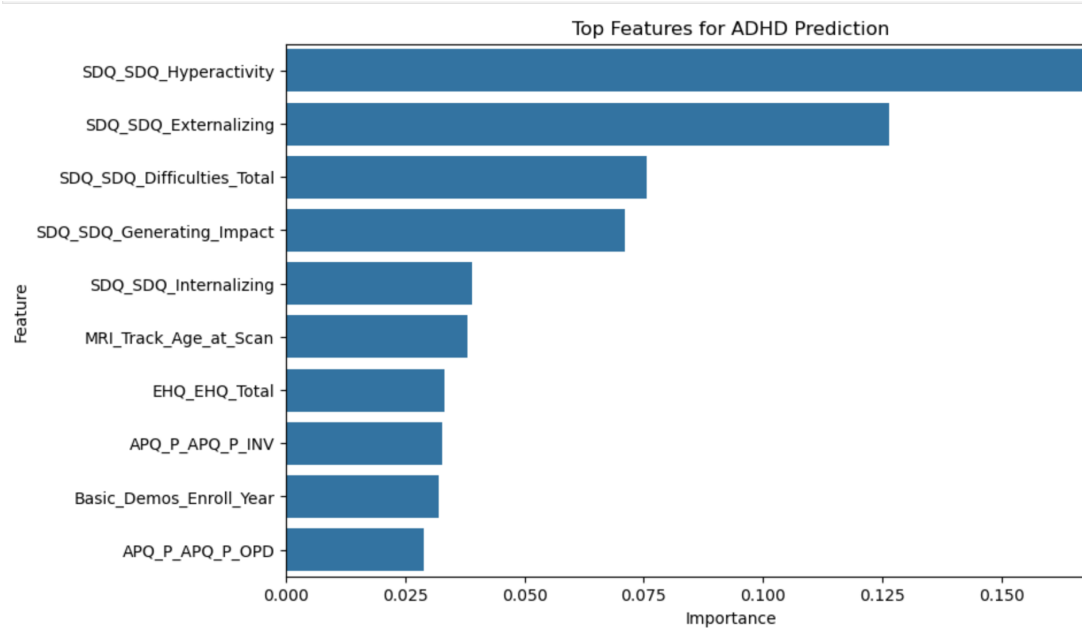

Figure 21. Top Features for ADHD Prediction Identified Using Recursive Feature Elimination (RFE)

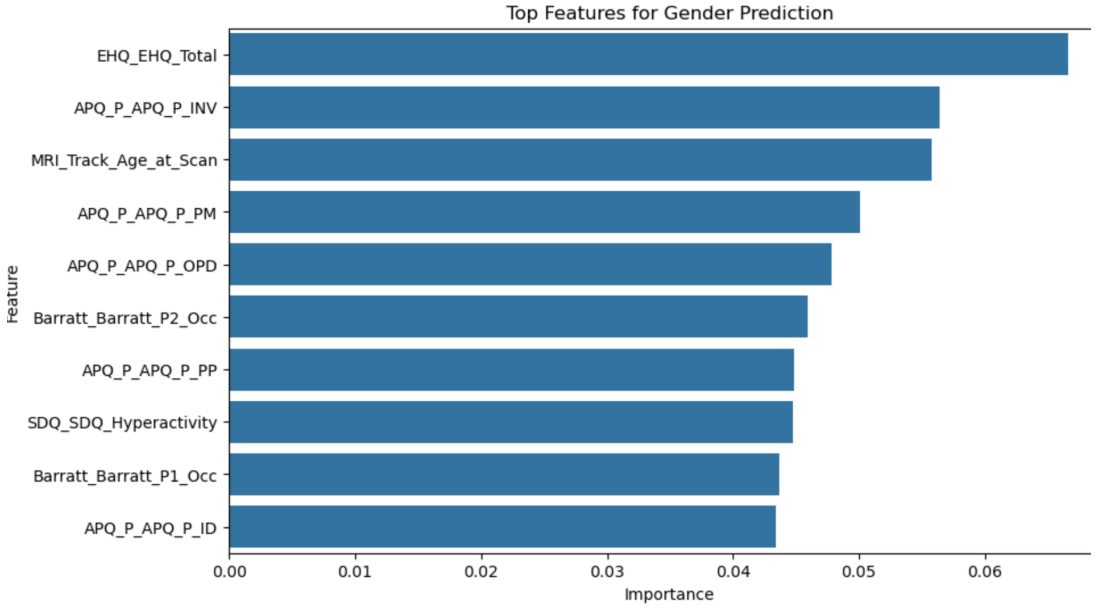

Figure 22. Top Features for Gender Prediction Identified Using Recursive Feature Elimination (RFE)

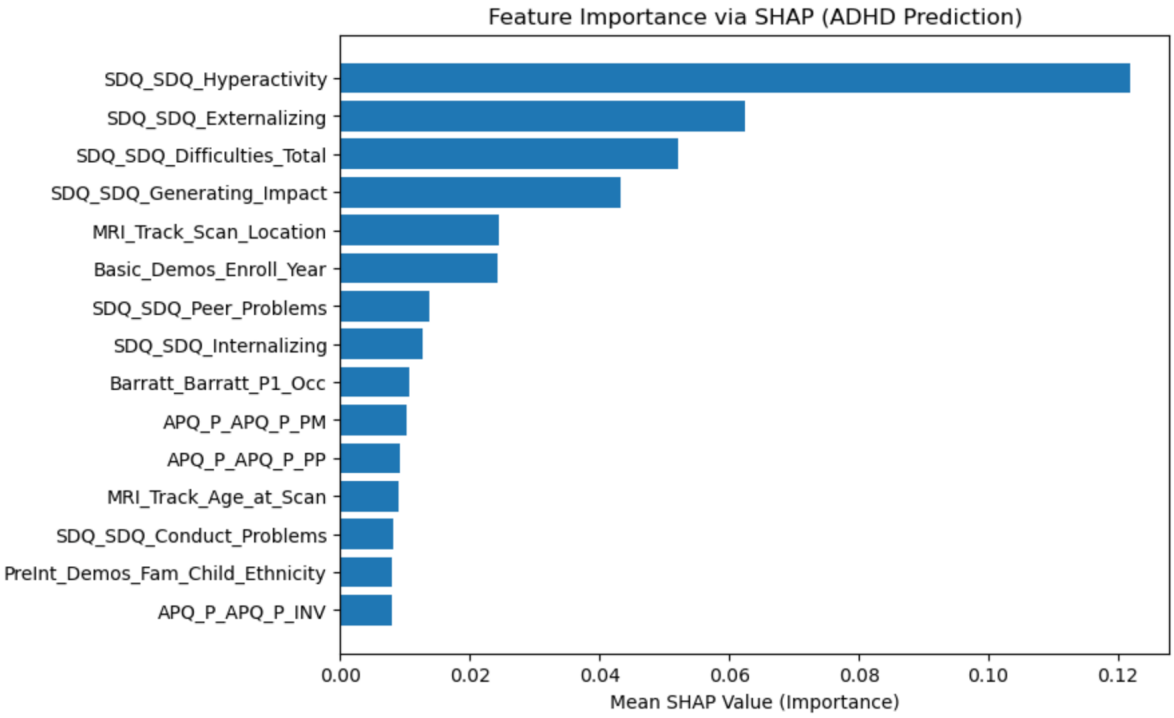

Figure 23. Top Features for ADHD Prediction Identified Using SHAP

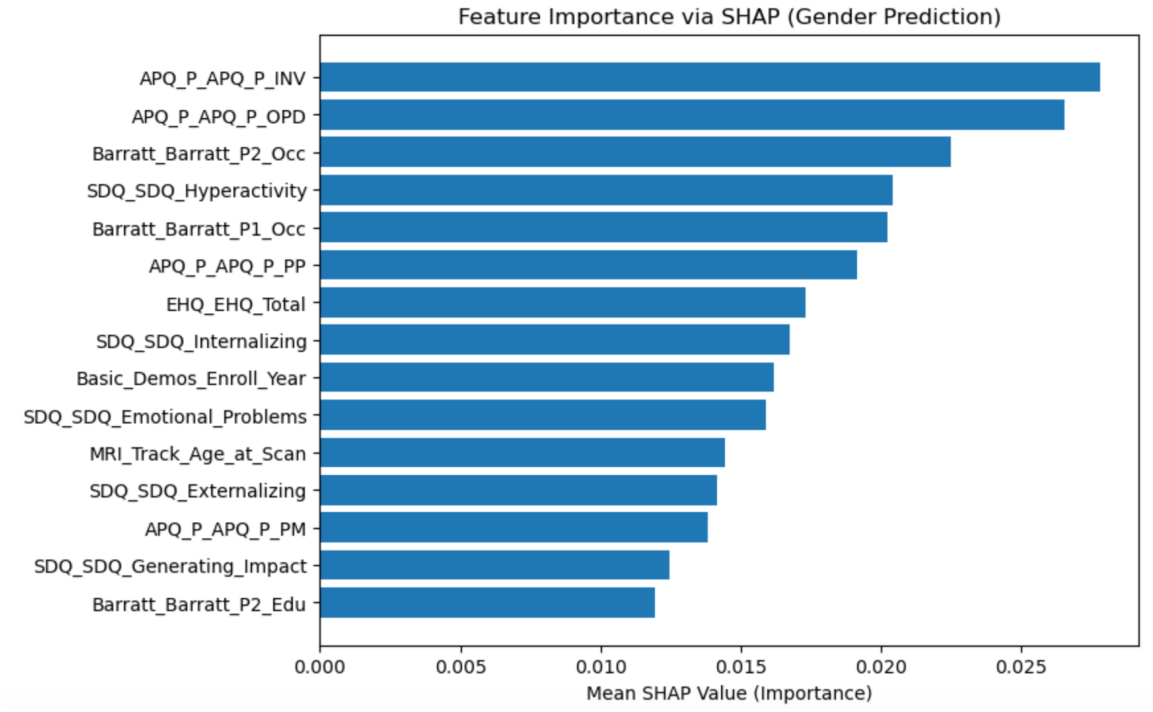

Figure 24. Top Features for Gender Prediction Identified Using SHAP