Machine Learning Plan for ADHD and SEX Prediction.

Overview

- **Model Selection**
- **Bias Detection and Mitigation**
- **Explainability**

1. Model Selection

I suggest the model given that the dataset explored has Structured data, Demographic Information and Continuous fMRI measurements:

A. **Random Forests**:

- Excellent for handling both numerical and categorical features.
- Built-in features are important for understanding key predictive factors.

2. Bias Detection and Fairness

- **Fairness Metrics:** To evaluate the model's fairness using metrics such as:
  - **Demographic Parity**, **Equal Opportunity**, and **Disparate Impact**
- **Bias Mitigation Strategies**
  - **Class Balancing**: Adjust class weights to mitigate the effects of imbalanced data (e.g., fewer females diagnosed with ADHD).
- **Evaluation Metrics**
  - **Accuracy, Precision, and Recall, F1-Score**
- **Visualization and Validation**
  - **Cross-Validation**: Use k-fold cross-validation for reliable performance estimates.

3. Explainability and Feature Importance

I will use tools like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) for complex models to ensure interpretability and transparency. Ensures that socio-demographic features do not lead to biased decisions.

Conclusion

This plan aims to improve how we predict ADHD. It would address critical aspects like fairness, spotting bias, and being clear on decisions. By using regular machine learning models that offer flexibility and accuracy. Tools like SHAP and LIME will show how the model makes decisions, helping clinicians better understand and trust the system.