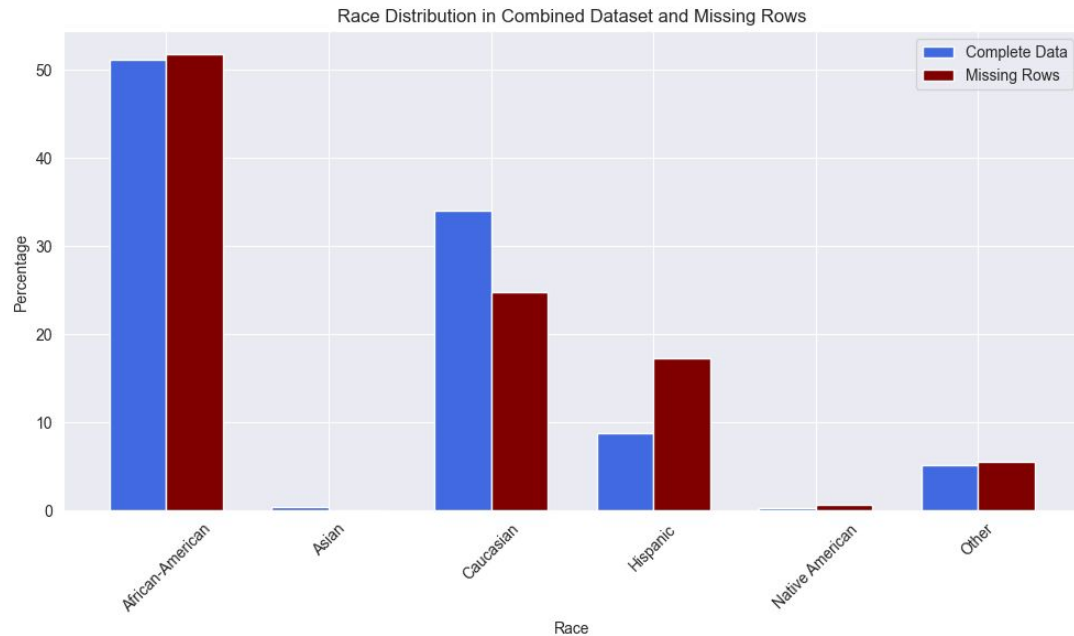# INNOCENT UNTIL PROVEN GUILTY

The Case of Markee Singletary

Choosing a Model

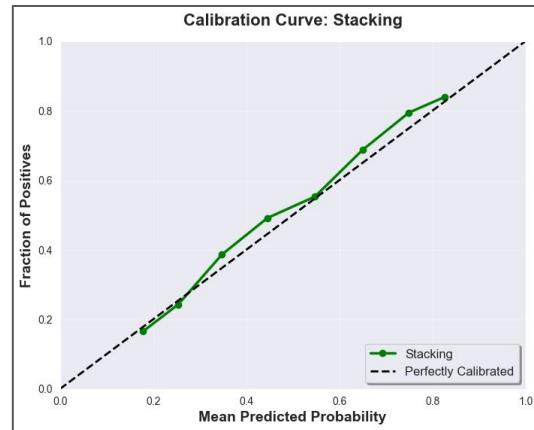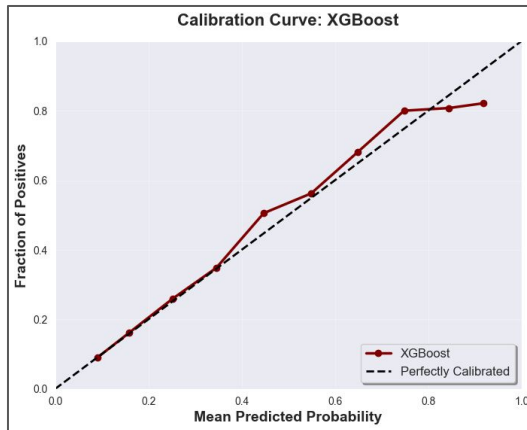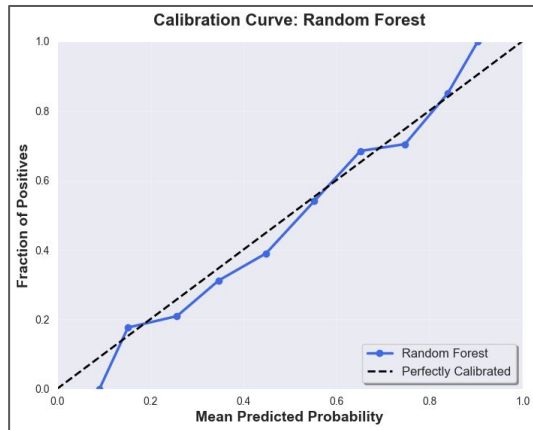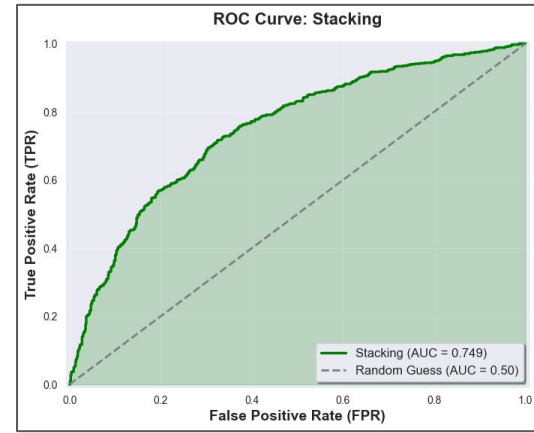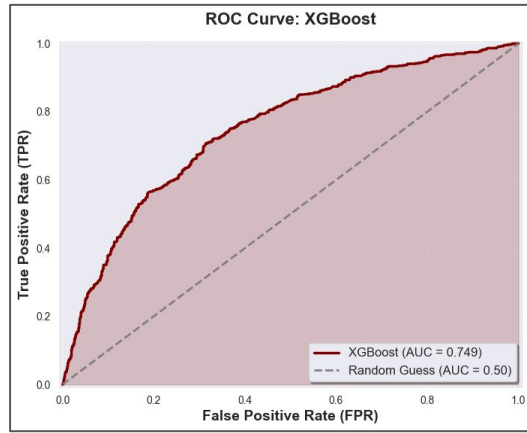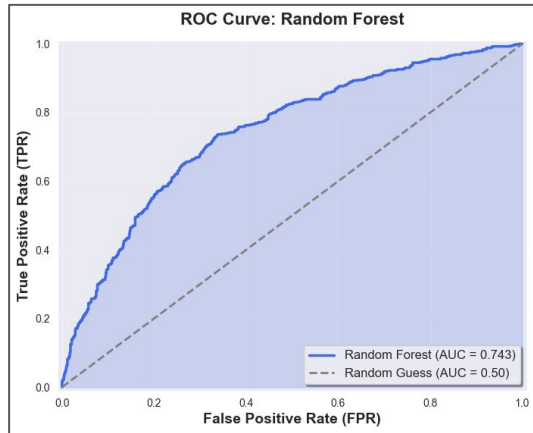# Missing Data

Race Distribution in Combined Dataset and Missing Rows

**Missing not at random - related by race
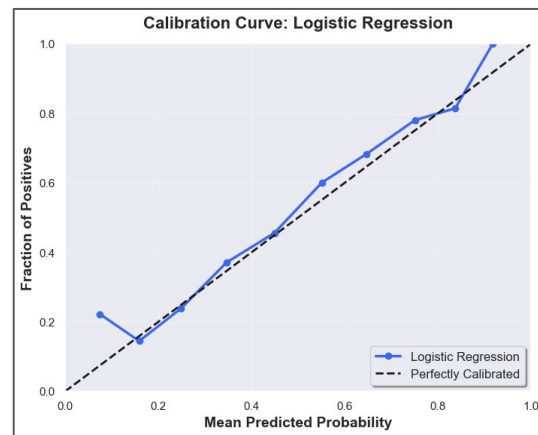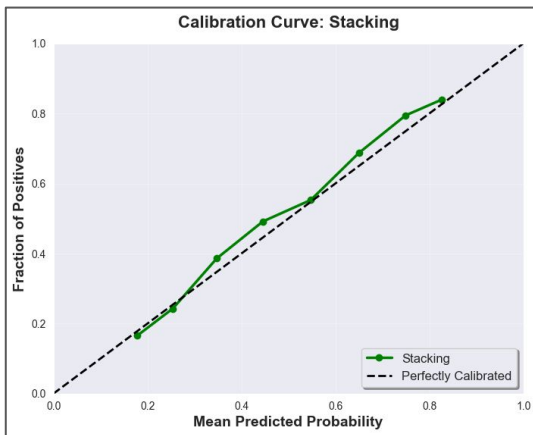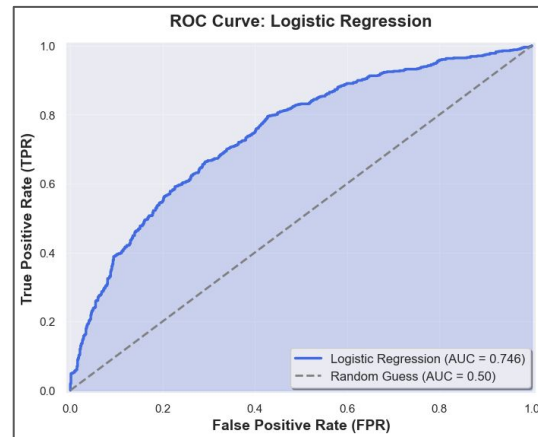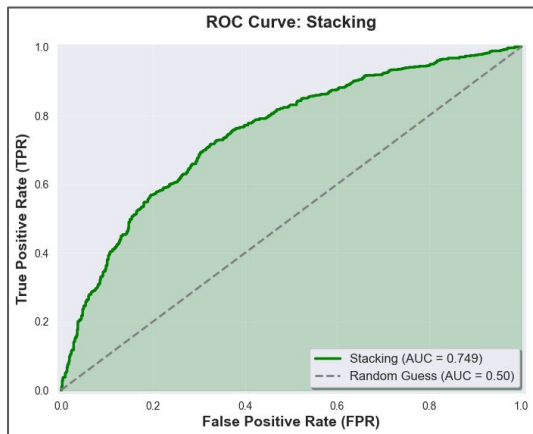K-Nearest Imputation**

Outliers

- 1.5 IQR Rule

- Log transformations

Choosing the Model - Performance Metric

Choosing the Model - Performance Metric
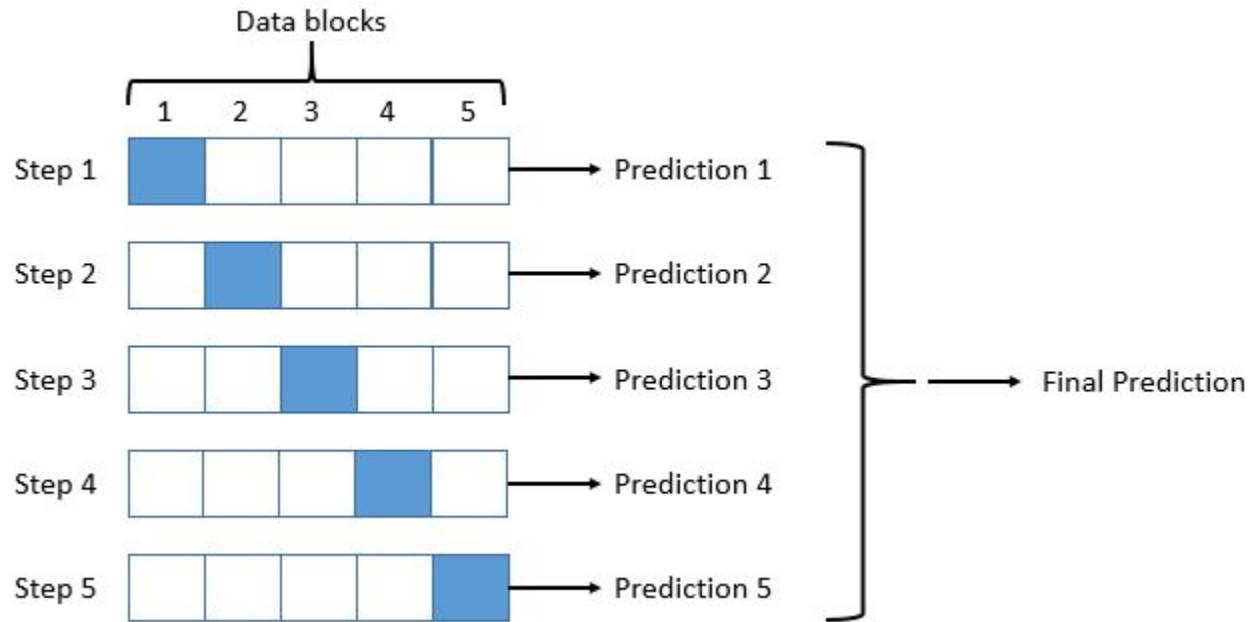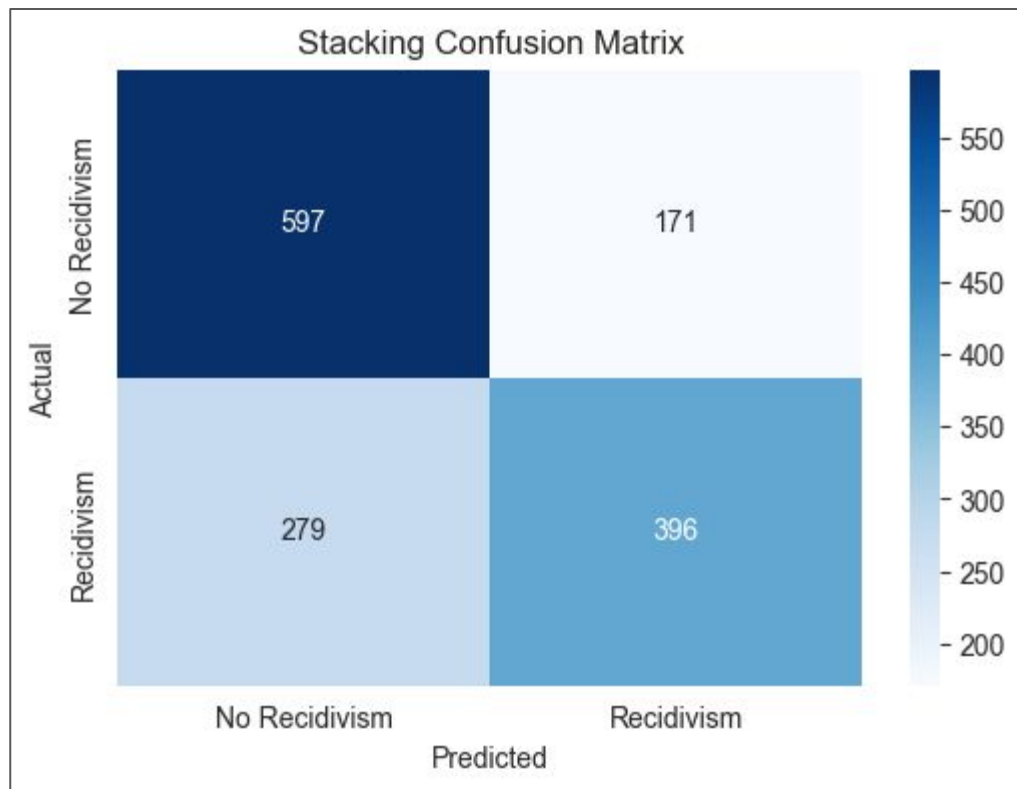
Chosen Model: Stacking Model

Model Performance

Internal Validation
Cross validation technique: 5-fold cross validation

Model Performance

Overall Accuracy: 68.6%

False Positive Rate: 22.3%

False Negative Rate: 41.3%

Comparison of Model against Compas
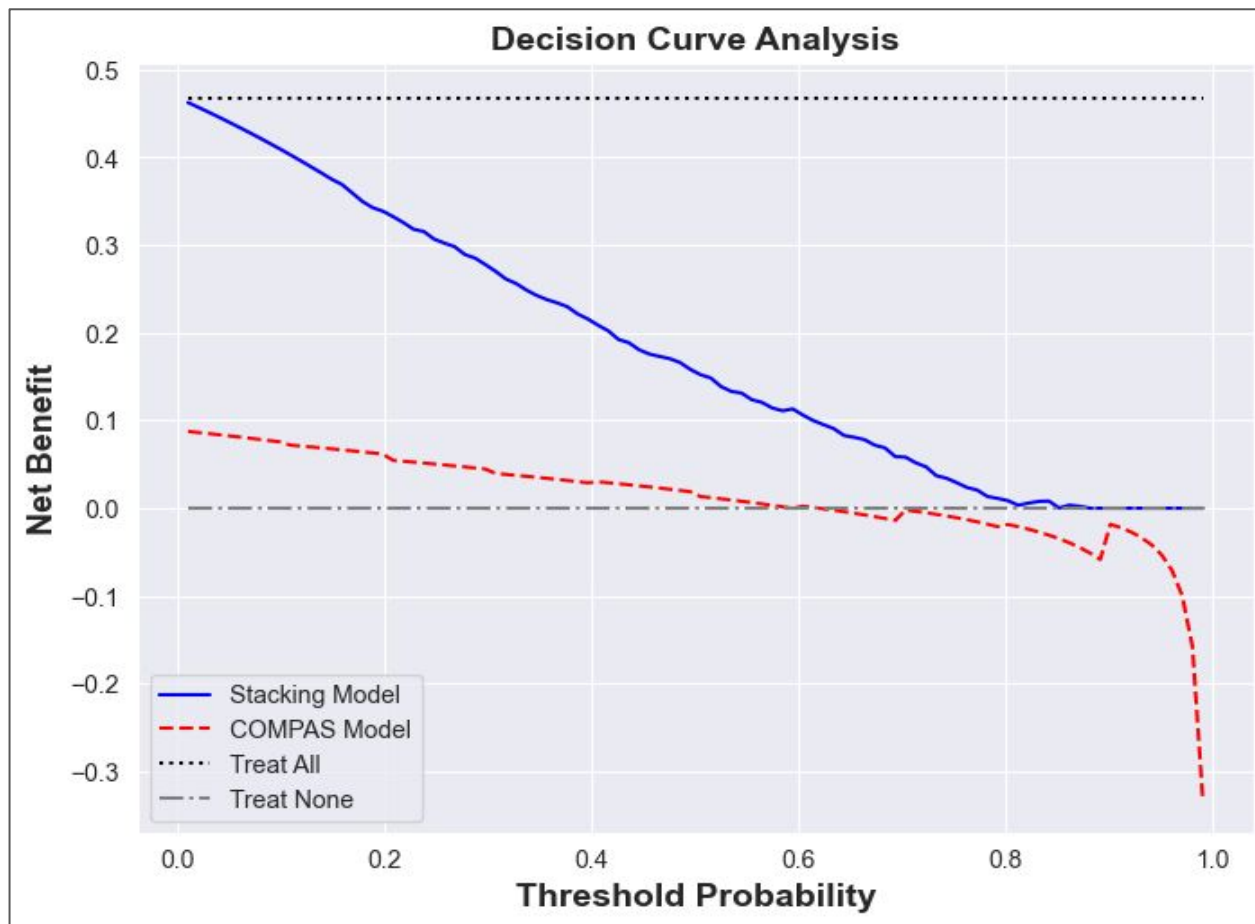

Compas Confusion Matrix
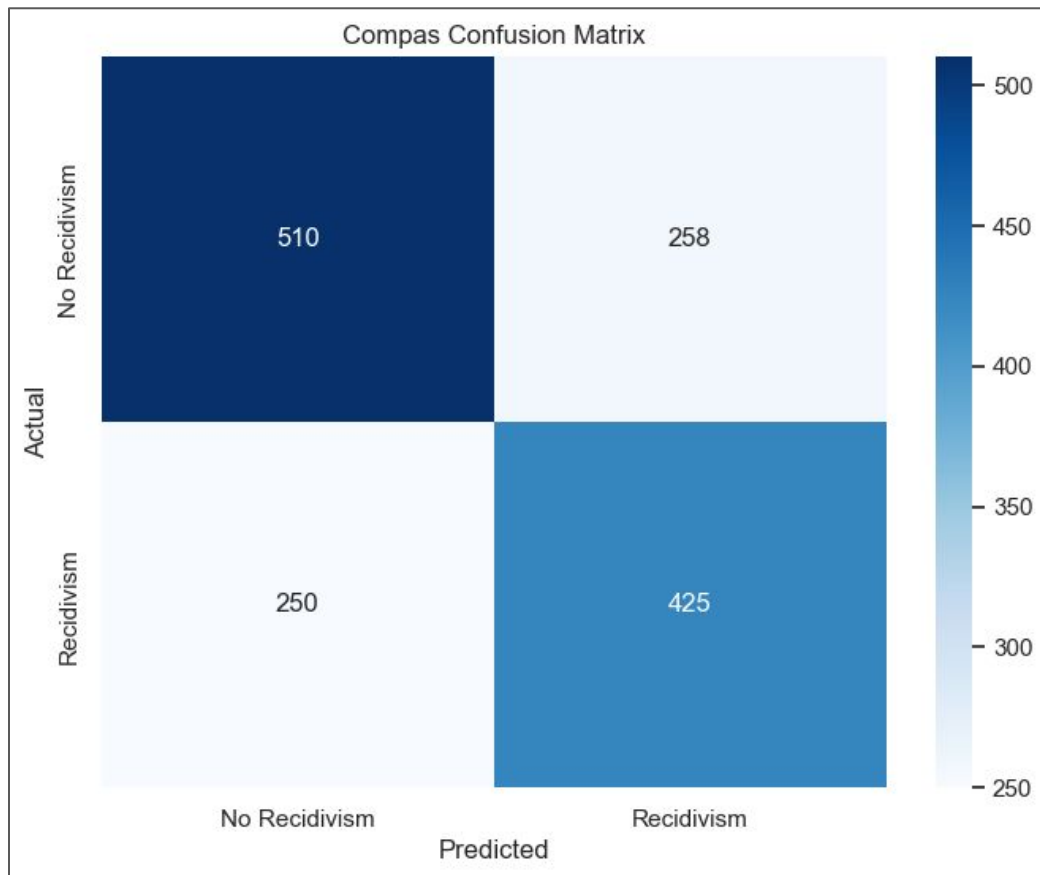
Overall Accuracy: 65%

False Positive Rate: 34%

False Negative Rate: 37%

Comparison of Model against Compas

Performance Metric	Stacking Model	Compas Model
Accuracy	68.60%	65%
False Positive Rate	22.30%	34%
False Negative Rate	41.30%	37%

Comparison of Model against Compas

Decision Curve Analysis

Model Interpretability - Feature Importance

Stacking Model: Feature Importance
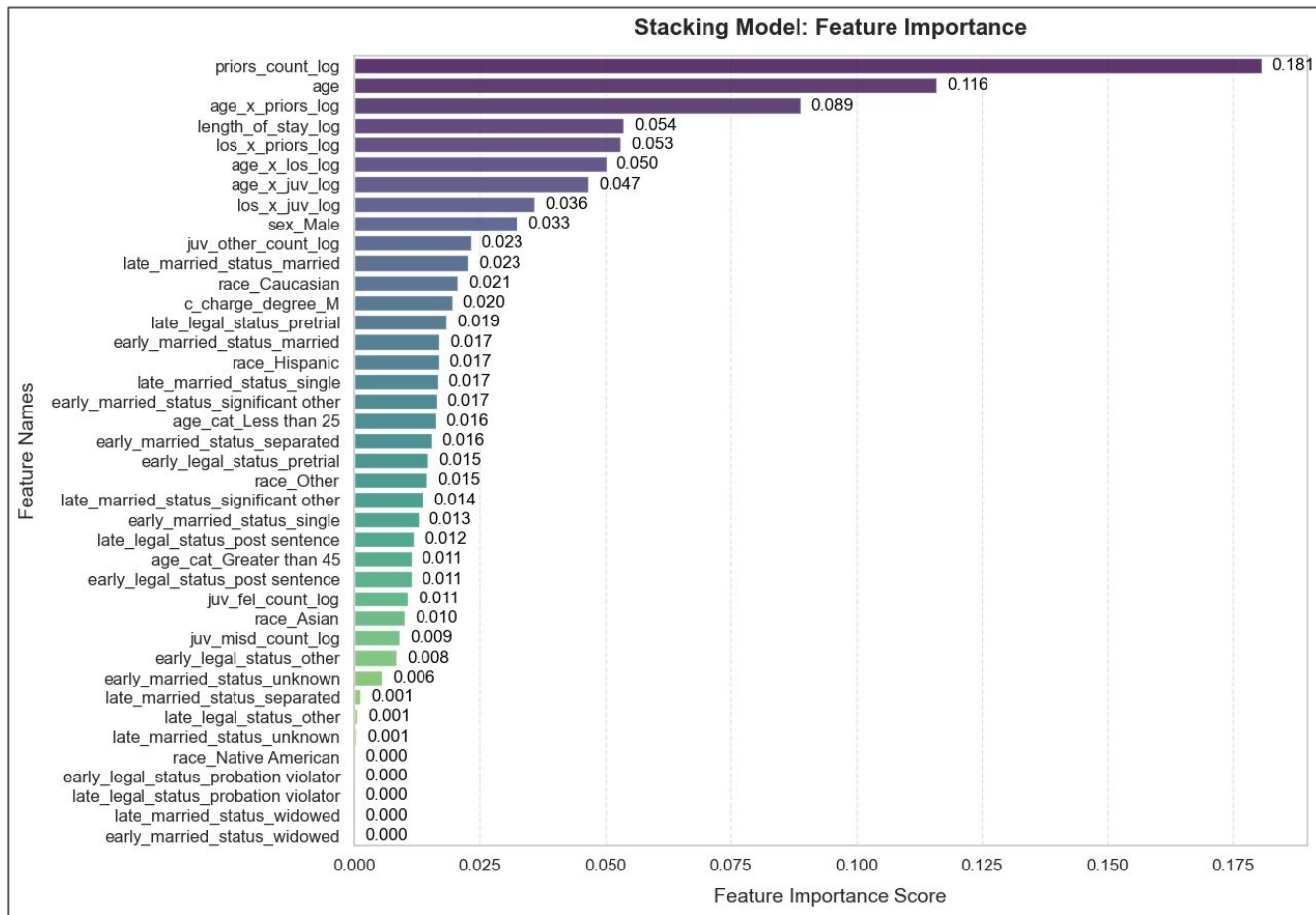
Model Interpretability - Partial Dependence Plots
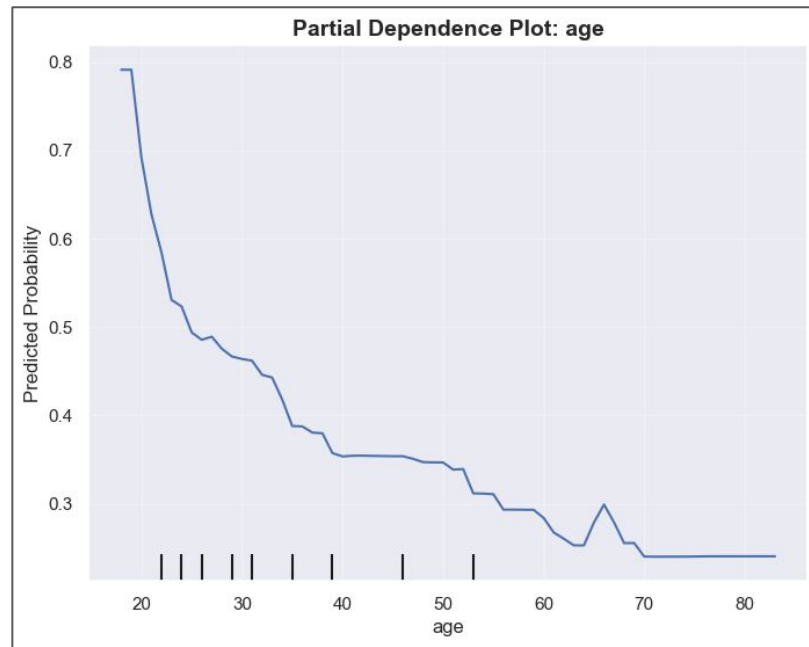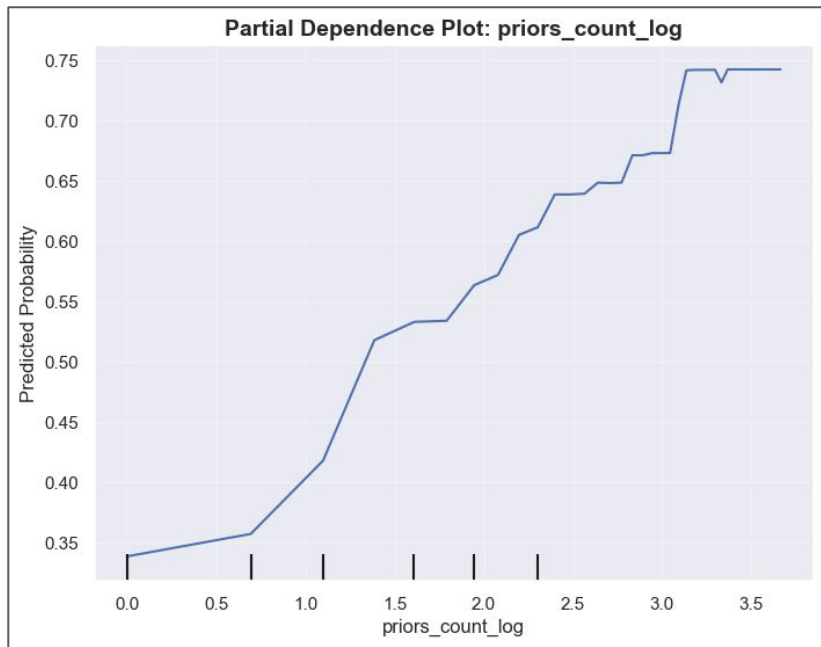
Model Fairness

INNOCENT UNTIL PROVEN GUILTY

Performance Metric	Stacking Model	Compas Model
Accuracy	68.60%	65%
False Positive Rate	22.30%	34%
False Negative Rate	41.30%	37%

Primary Fairness Metric: False Positive Rate

Final Recommendation

Stacking Model

Predicted Score: 0.282277

Risk Category: Low

Compas Model

Decile Score: 2

Risk Category: Low

Recommendation: Grant Parole.