Advancing Pediatric Heart Transplantation: Predictive Models for Donor Selection

DTSA 5506: DATA MINING PROJECT

Executive Summary

- Current pediatric heart donor selection lacks data-driven precision
- Using the ORCHID dataset this project developed ML models to improve donor selection
- Correlations
  - Strong Positive: Liver/kidney recovery, brain death diagnosis
  - Negative: Natural or cardiovascular death, anuria
- Brain death is the strongest predictor
- Modest donor size differences
- Random Forest slightly outperforms Logistic Regression (LR), handling non-linear relationships better
- Recall and F1 metrics suggest strong and stable minority class detection, essential for clinical utility.

Problem Statement

- Current donor selection for pediatric heart transplantation lacks data-driven precision
- This project aims to develop classification models to improve donor selection

Related Work

Focus Area	Notable Issues	Key Sources
Donor Selection	Age, ischemia time, infection status, weight matching	Singh et al. (2020); Singh et al. (2019); Conway et al. (2020)
Ethical Considerations	Obesity, intellectual disability, resource allocation, xenotransplantation	Berkman & Wightman (2021); Hurst et al. (2024); Wilkens et al. (2020); Bearl (2019)
Outcome Predictors	Pre-transplant support, diagnosis, panel reactive antibody, BMI	Mantell & Elizer (2025); Donné et al. (2021); Canter et al. (2007)
Registry & Trends	Global disparities, median donor age, mortality trends	Singh et al. (2020); AHA Guidelines (2007)
Clinical Innovations	Xenotransplantation trials, high-risk acceptance algorithms	Hurst et al. (2024); Nabzdyk et al. (2024)

Proposed Work

- Dataset: https://physionet.org/content/orchid/2.0.0/
- ORCHID Dataset Overview
  - 133,101 deceased donor referrals
  - 8,972 organ donations
  - 13 states covered
- Referrals by OPO
  - OPO 1: 32,148 potential donors
  - OPO 2: 16,144 potential donors
  - OPO 3: 12,516 potential donors
  - OPO 4: 33,641 potential donors
  - OPO 5: 15,738 potential donors
  - OPO 6: 22,914 potential donors

Proposed Work

- All data was de-identified in accordance with HIPAA standards using structured data cleansing and date shifting.ORCHID consists of ten tables (CSV files), linked via PatientID, and grouped into three categories:

- OPO Referrals
  - Referral information: patient demographics & cause of death
  - Process data: timestamps for every action (next-of-kin approach, authorization, procurement, death modes)
  - Outcomes: binary flags for approached, authorized, procured; plus per-organ recovery results

Proposed Work

- OPO Events
  - ChemistryEvents: blood chemistry (kidney panel, LFTs, electrolytes)
  - CBCEvents: complete blood count with differential
  - ABGEvents: arterial blood gas + ventilator settings
  - SerologyEvents: presence/absence of donation-relevant antigens/antibodies
  - CultureEvents: infection culture results (blood, urine, other)
  - HemoEvents: hemodynamics over time_event_start → time_event_end (or point measurements)
  - FluidBalanceEvents: fluid intake/output over time_event_start → time_event_end
- OPO Deaths Captures referring-hospital death data (asystole, brain death timestamps, etc.)

Proposed Work

- Preprocessing
  - Evaluate missing data
  - One hot encoding
- Exploratory Data Analysis
  - Evaluate for anomalies
  - Chi Square
  - Correlations
- Modeling
  - Supervised ML for classifcation
    - Consider XGboost
    - Consider Linear Regression
    - Consider SVM
    - Consider RandomForest

Exploratory Data Analysis

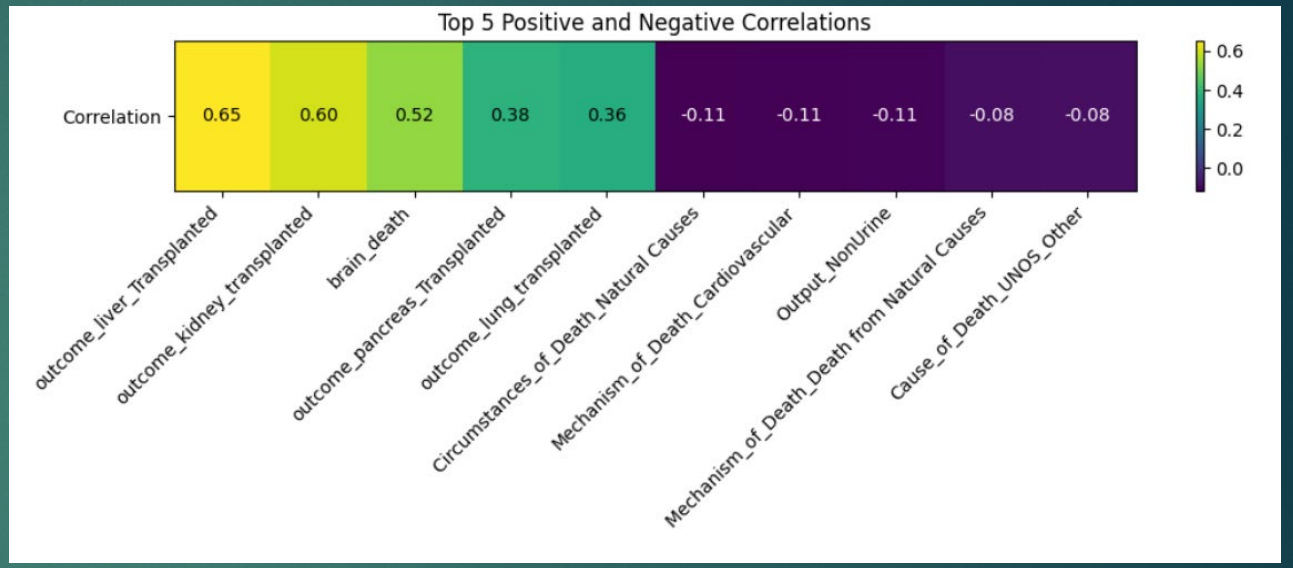
- Total potential pediatric heart donors: 1347
- Total transplanted: 441
- Not transplanted: 906

Exploratory Data Analysis

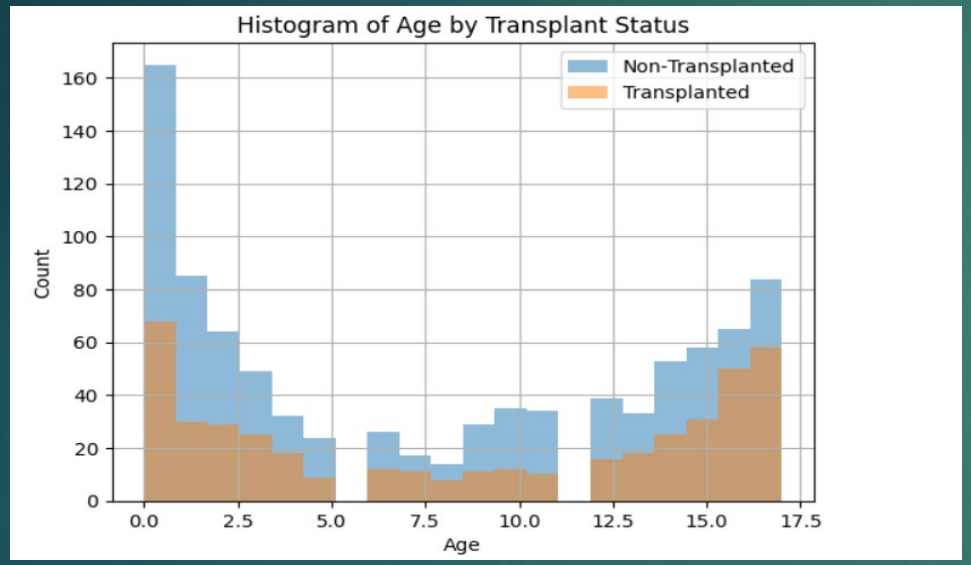
Strong Positive Correlation
- Liver transplanted
- Kidney transplanted
- Brain death diagnosis

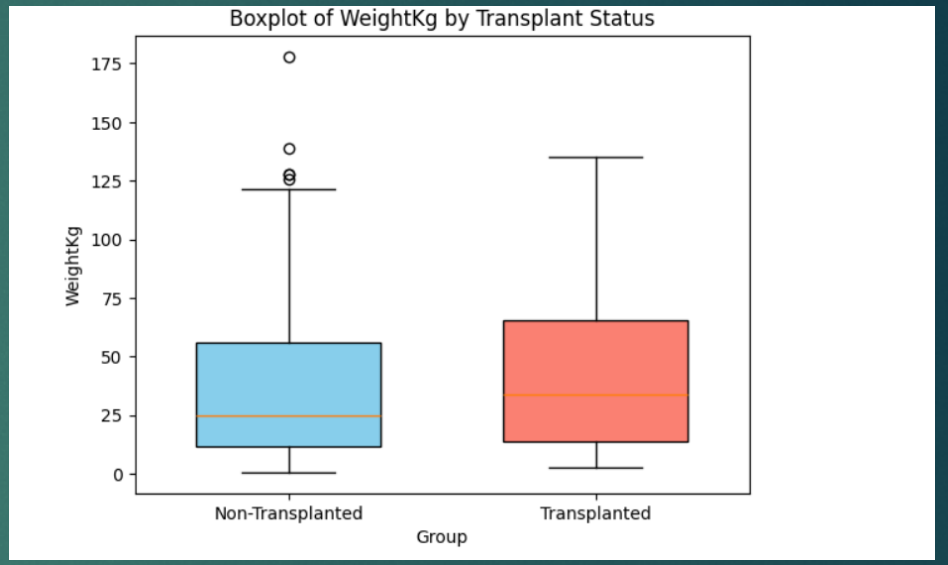
Moderate Positive Correlation
- Pancreas transplanted
- Lung transplanted

Weak Negative Correlation
- Death by natural causes
- Death by cardiovascular causes
- Non-urine output

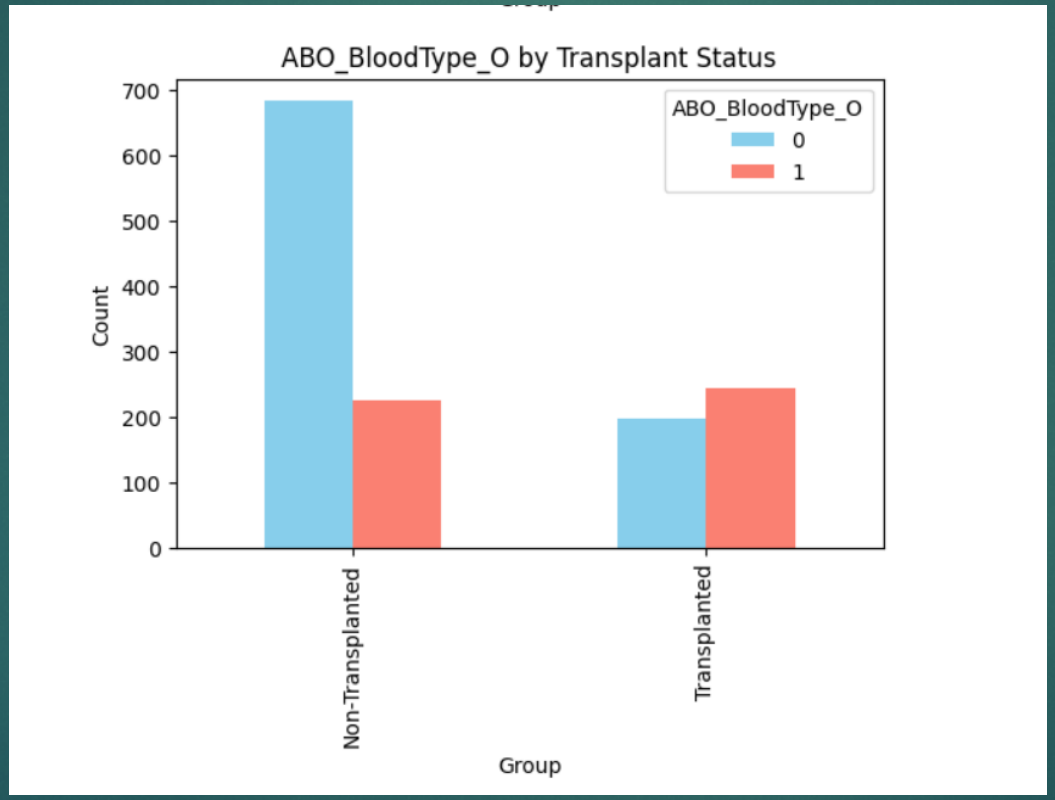


Exploratory Data Analysis





Exploratory Data Analysis



Exploratory Data Analysis

- Significant Findings: Transplanted vs. Non-Transplanted Donors
  - Brain death status is the strongest discriminator
    - 99.4% of non-transplanted donors were *not brain-dead*
    - 51.6% of transplanted donors *were brain-dead*
    - Highly significant difference ($p < 1 \times 10^{-107}$)
  - Referral type associated with transplant likelihood
    - Tissue referrals more common in transplanted donors (34.9% vs. 22.6%, $p = 3 \times 10^{-4}$)
    - Eye referrals also higher in transplanted donors (35.4% vs. 22.7%, $p = 8 \times 10^{-5}$)
    - Suggests multi-tissue engagement may enhance heart procurement

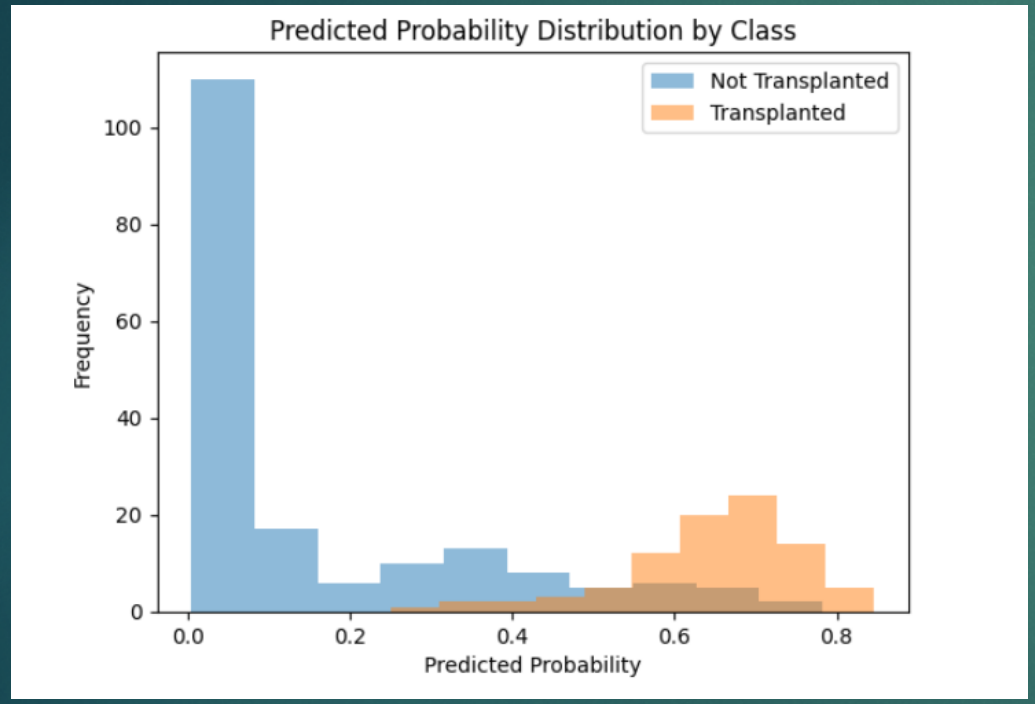
Exploratory Data Analysis

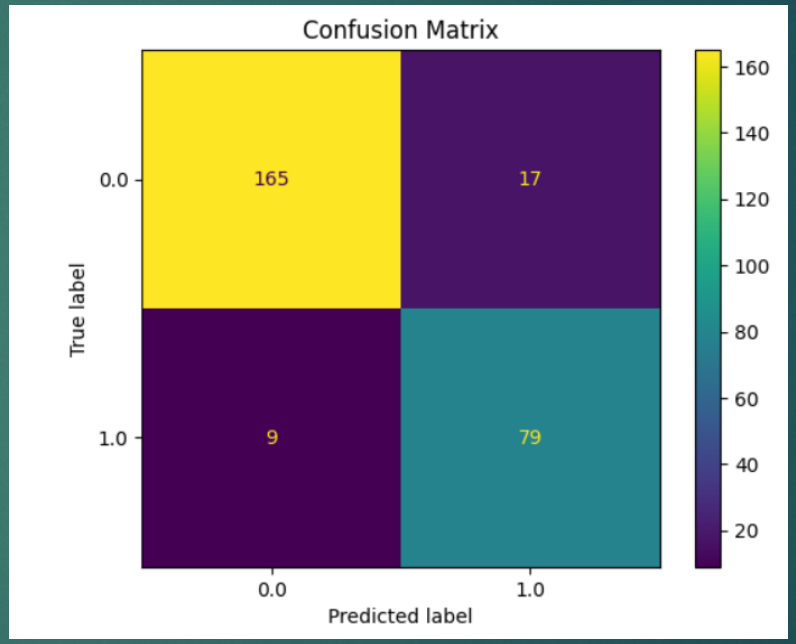
- Microbiology testing more frequent in transplanted donors
  - Right bronchial Gram stain positive in 73% of transplanted vs. 50% non-transplanted ($p = 6 \times 10^{-5}$)
  - May reflect more thorough evaluation, not increased infection
- Modest differences in donor size
    - Age: 8.7 vs. 7.7 years ($p \approx 0.005$)
    - Height: 50.95 in vs. 48.4 in ($p \approx 0.03$)
    - Weight: 41.5 kg vs. 34.8 kg ($p \approx 1 \times 10^{-4}$)

Evaluation

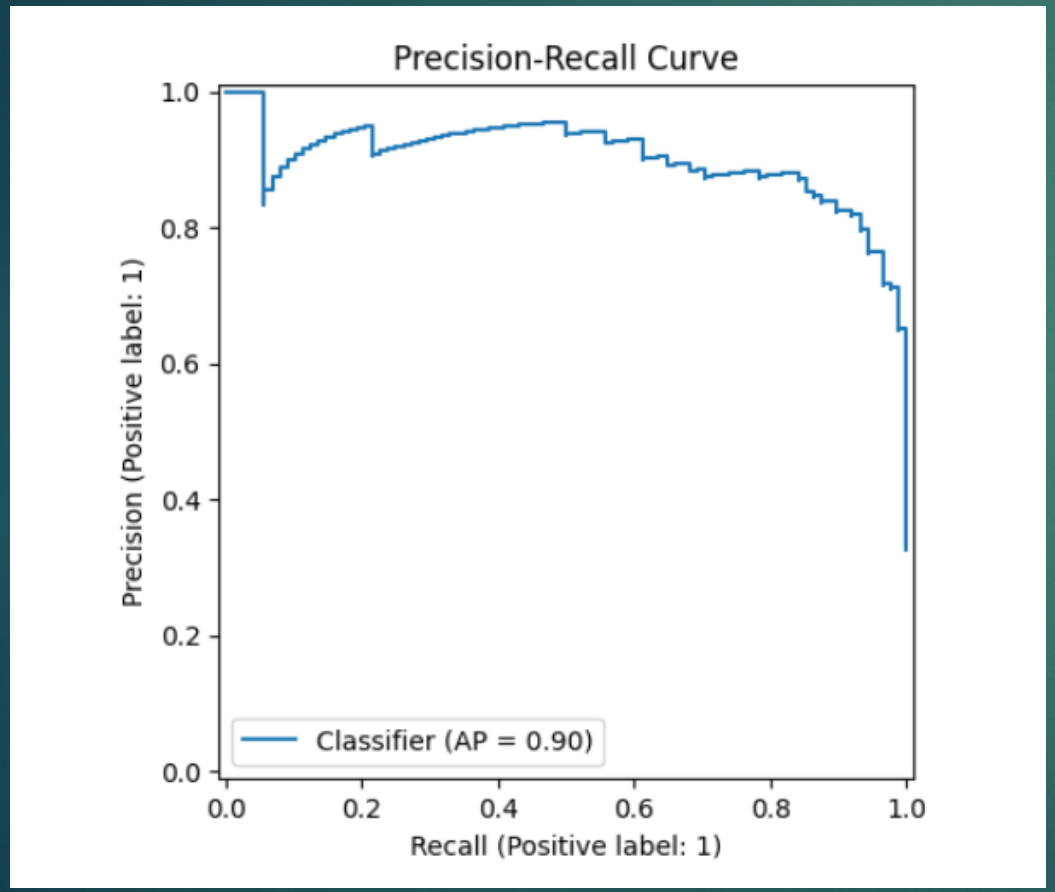
Model	Precision	Recall	F1-Score	Accuracy
Logistic Regression (LR)	0.86	0.85	0.86	0.91
LR + RFE	0.85	0.85	0.85	0.90
Support Vector Machine	0.80	0.89	0.84	0.89
Random Forest	0.84	0.86	0.85	0.90
Random Forest (Tuned)	0.83	0.92	0.87	0.91

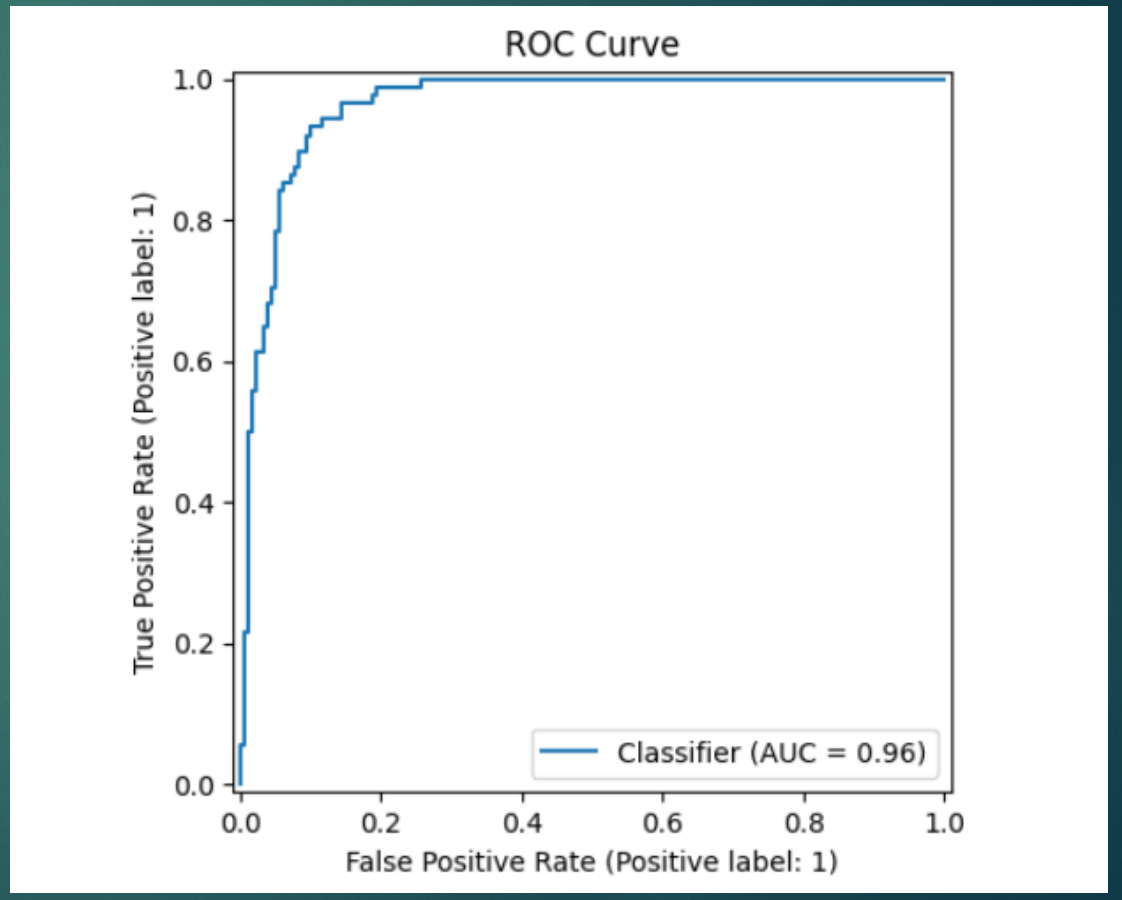
Evaluation: CM and Probability Distribution



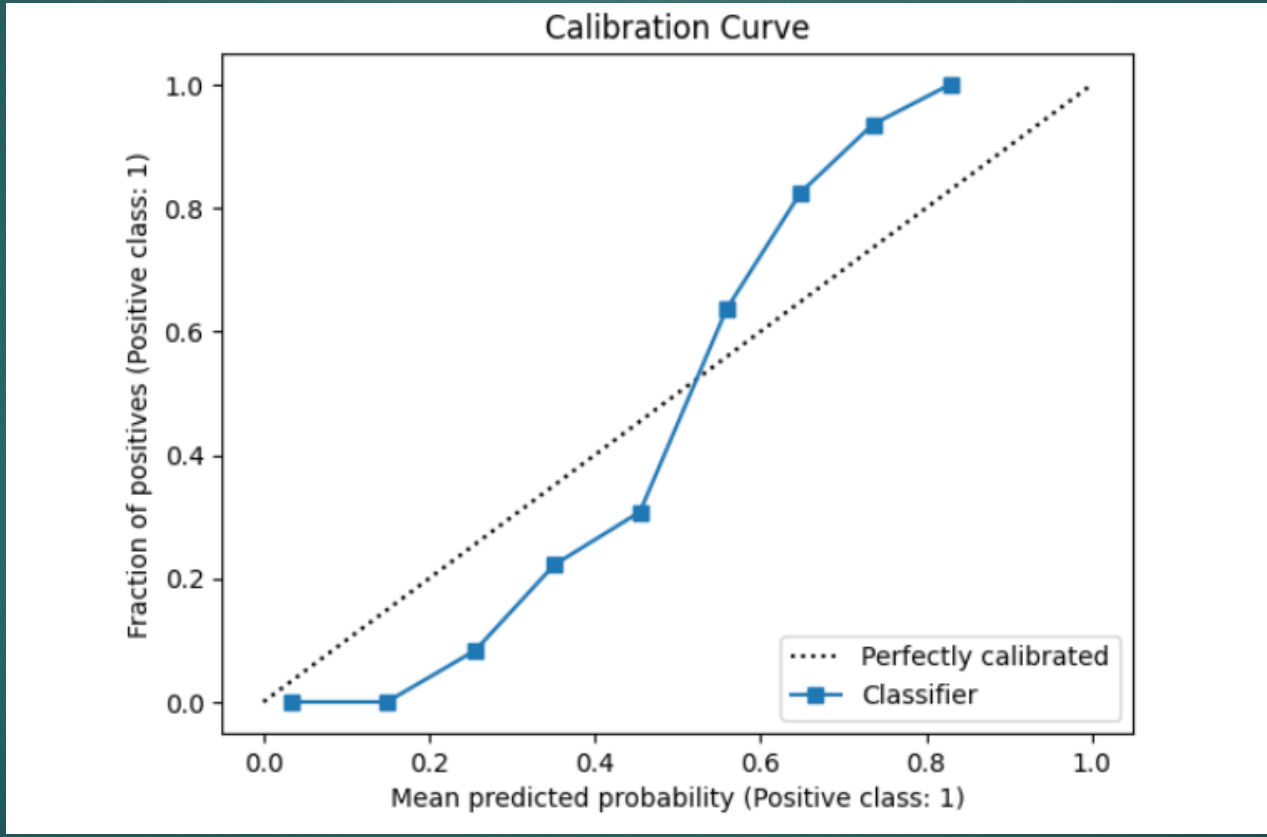


Evaluation: ROC and PR





Evaluation: Calibration



Evaluation

- Precision Recall plot
  - Good model for identifying positives
- ROC plot
  - Excellent discrimination ability
  - High sensitivity and specificity
- Calibration plot
  - Model is well-calibrated at higher probabilities
  - Slight underconfidence at lower probabilities
- Predicted Probabilities Distribution
  - Clear class separation, with some overlap
  - Model is reasonably confident in most predictions.

Evaluation: Cross Validation

- Cross-Validation Results (K = 5)
  - Mean Precision: 0.753 (± 0.065)
  - Mean Recall: 0.907 (± 0.036)
  - Mean F1-Score: 0.820 (± 0.039)
  - Mean Accuracy: 0.869 (± 0.034)

- Accuracy: Some overfitting but the SD suggests reasonable stability
- Precision: Moderate consistency
- Recall: Strong, stable detection of the minority class
- F1: Good balance between precision and recall

Evaluation

- LR and RF were fairly close in predicting the outcome variable
- There are slightly more false positives compared to LR
- Reason for RF over LR -> The ensemble nature captures non-linear patterns better than LR

Timeline

- Week 1: Completed
  - Data filtering, initial data preprocessing
- Week 2: Completed
  - Finish preprocessing including missing data and one-hot encoding
  - Exploratory data analysis including corrections, chi square, box plots, etc
- Week 3-4: Completed
  - Explore machine learning models
  - Evaluation of final model
- Week 5: Completed
  - Complete final write-up