

Model Training Report: Track 1

Configuration Details

Target Variable

target

Selected Features

goal_diff_80, home_shots_pre80, away_shots_pre80, home_fouls_pre80, away_fouls_pre80, home_corners_pre80, away_corners_pre80, home_yellow_cards_pre80, away_yellow_cards_pre80, home_sending_off_pre80, away_sending_off_pre80, home_corners_70_75, away_corners_70_75, home_shots_70_75, away_shots_70_75, home_corners_75_80, away_corners_75_80, home_shots_75_80, away_shots_75_80, odd_h, odd_d, odd_a

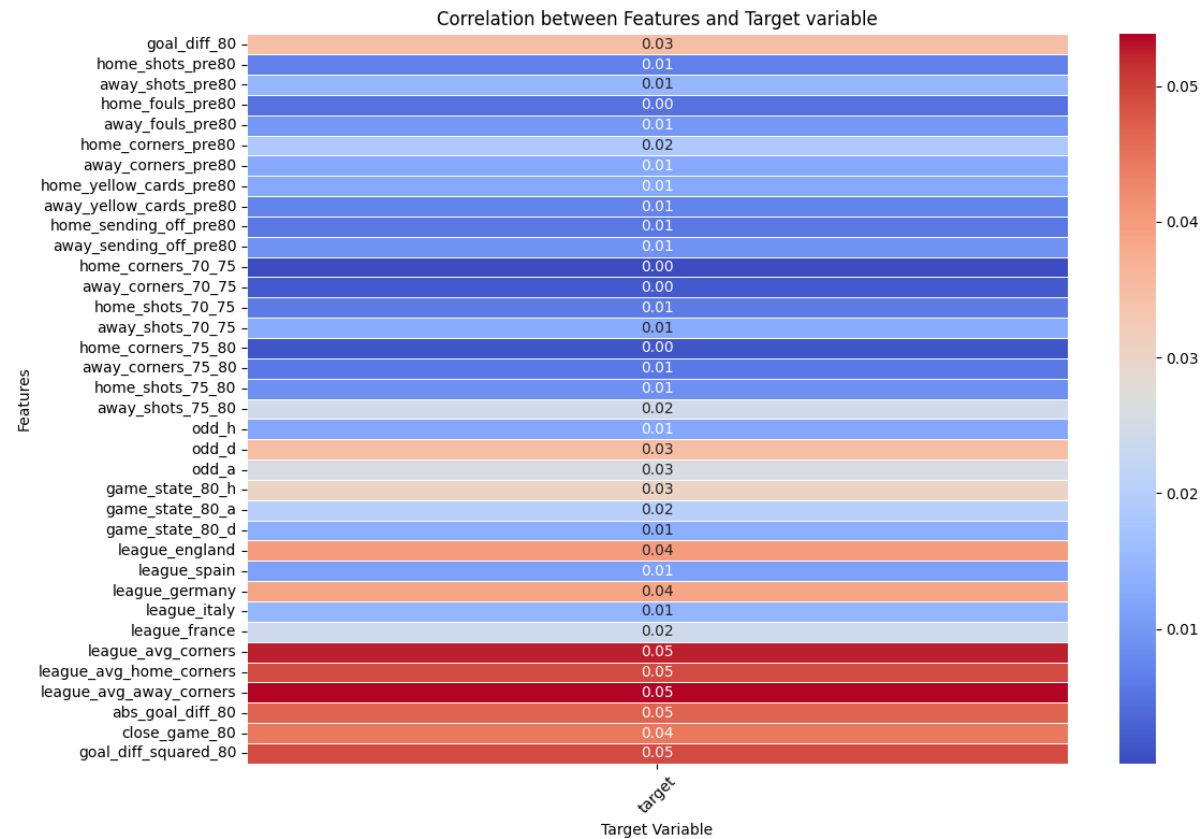
Constructed Features

game_state_80_h, game_state_80_a, game_state_80_d, league_england, league_spain, league_germany, league_italy, league_france, league_avg_corners, league_avg_home_corners, league_avg_away_corners, abs_goal_diff_80, close_game_80, goal_diff_squared_80

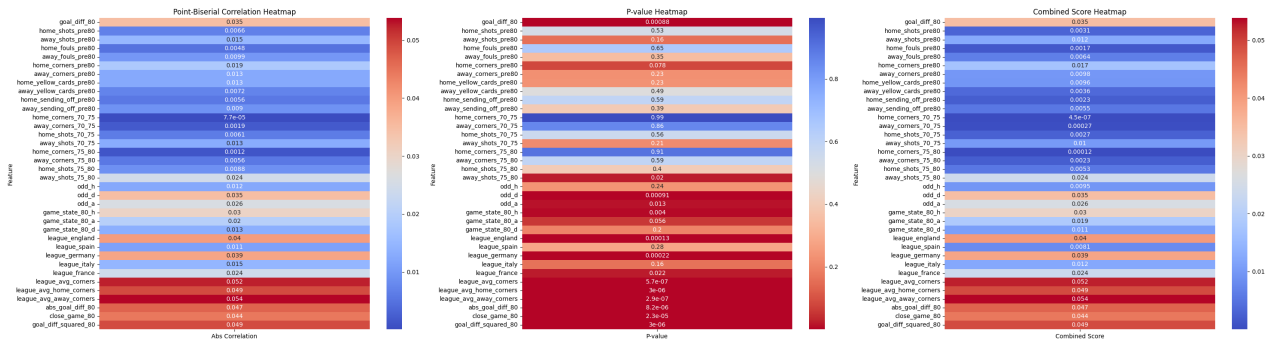
Models Trained (& Hyperparameters)

- **random_forest:** {'bootstrap': True, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 200}
- **logistic_regression:** {'solver': 'liblinear', 'penalty': 'l2', 'C': 1.0}
- **svc:** {'C': 1.0, 'kernel': 'rbf', 'gamma': 'scale'}
- **xgboost:** {'n_estimators': 50, 'max_depth': 3, 'learning_rate': 0.1, 'subsample': 0.8}

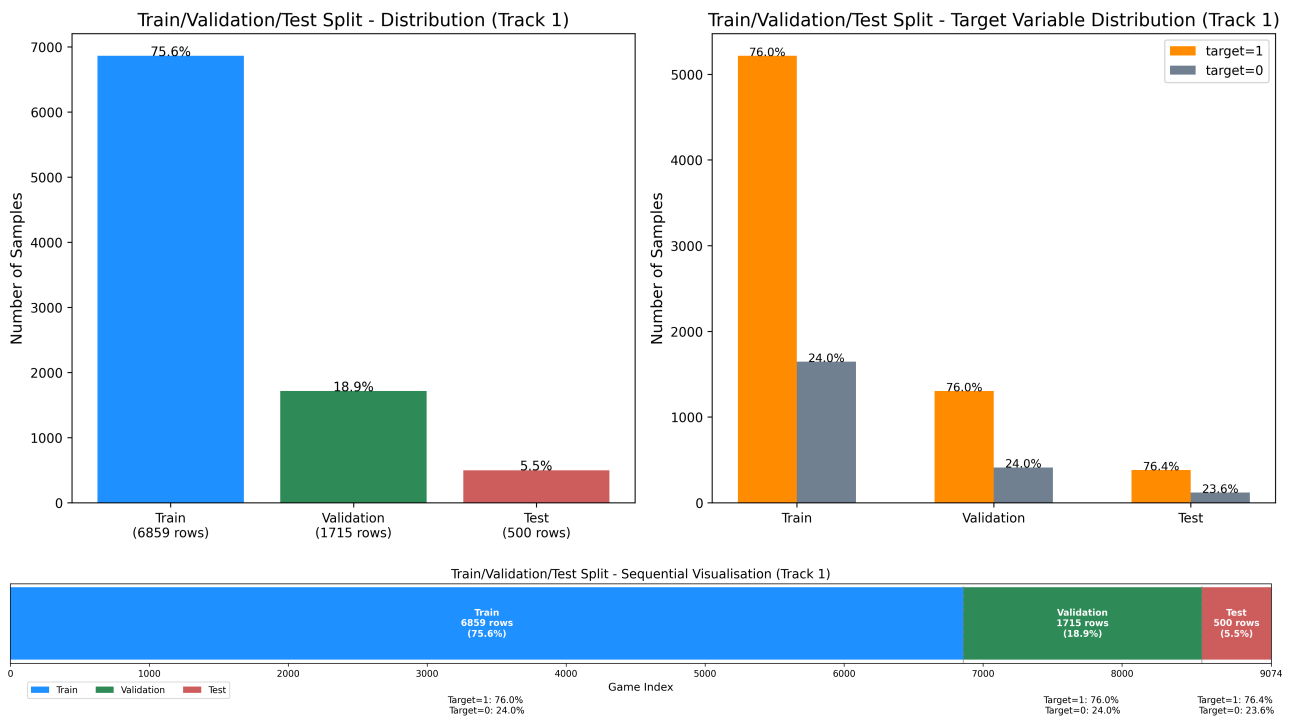
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

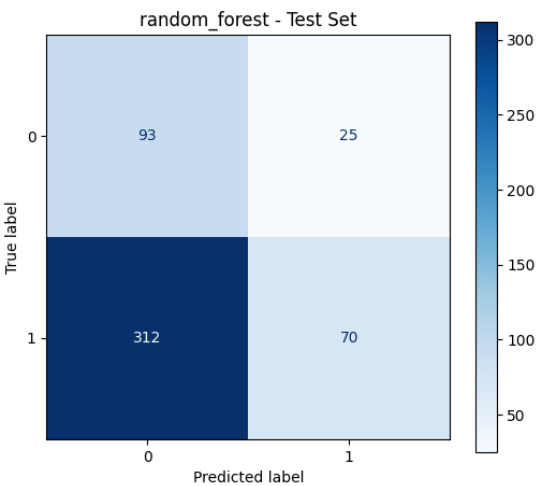
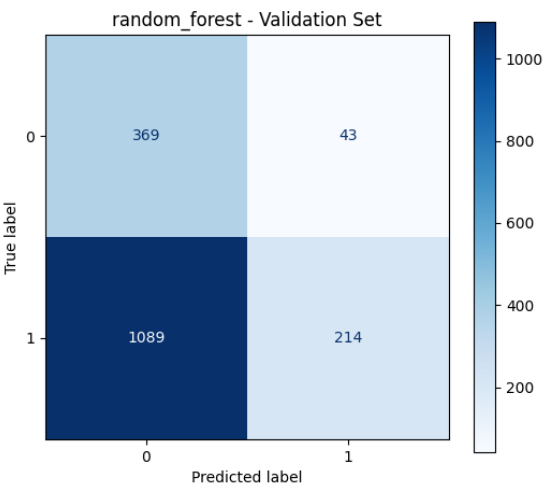
Feature Importance

| Top 8 Features | Importance | Bottom 5 Features | Importance |
|--------------------|------------|-------------------|------------|
| odd_a | 0.08567 | league_england | 0.00266 |
| odd_h | 0.08547 | league_france | 0.00428 |
| odd_d | 0.08130 | league_germany | 0.00433 |
| home_shots_pre80 | 0.06809 | league_spain | 0.00438 |
| away_fouls_pre80 | 0.06528 | close_game_80 | 0.00525 |
| home_fouls_pre80 | 0.06424 | | |
| away_shots_pre80 | 0.06393 | | |
| home_corners_pre80 | 0.05538 | | |

Optimal Threshold

- **Optimal threshold:** 0.83
- **Expected Precision:** 0.833
- **Expected Recall:** 0.164

Classification Reports



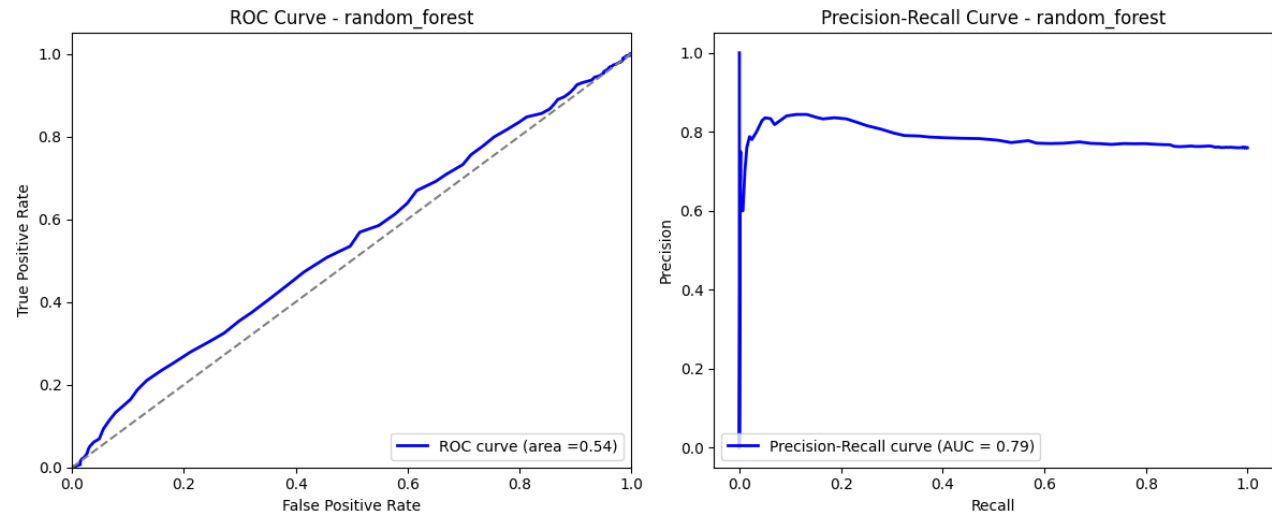
Classification Report (Validation Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.25 | 0.90 | 0.39 | 412 |
| 1 | 0.83 | 0.16 | 0.27 | 1303 |
| accuracy | | | 0.34 | 1715 |
| macro avg | 0.54 | 0.53 | 0.33 | 1715 |
| weighted avg | 0.69 | 0.34 | 0.30 | 1715 |

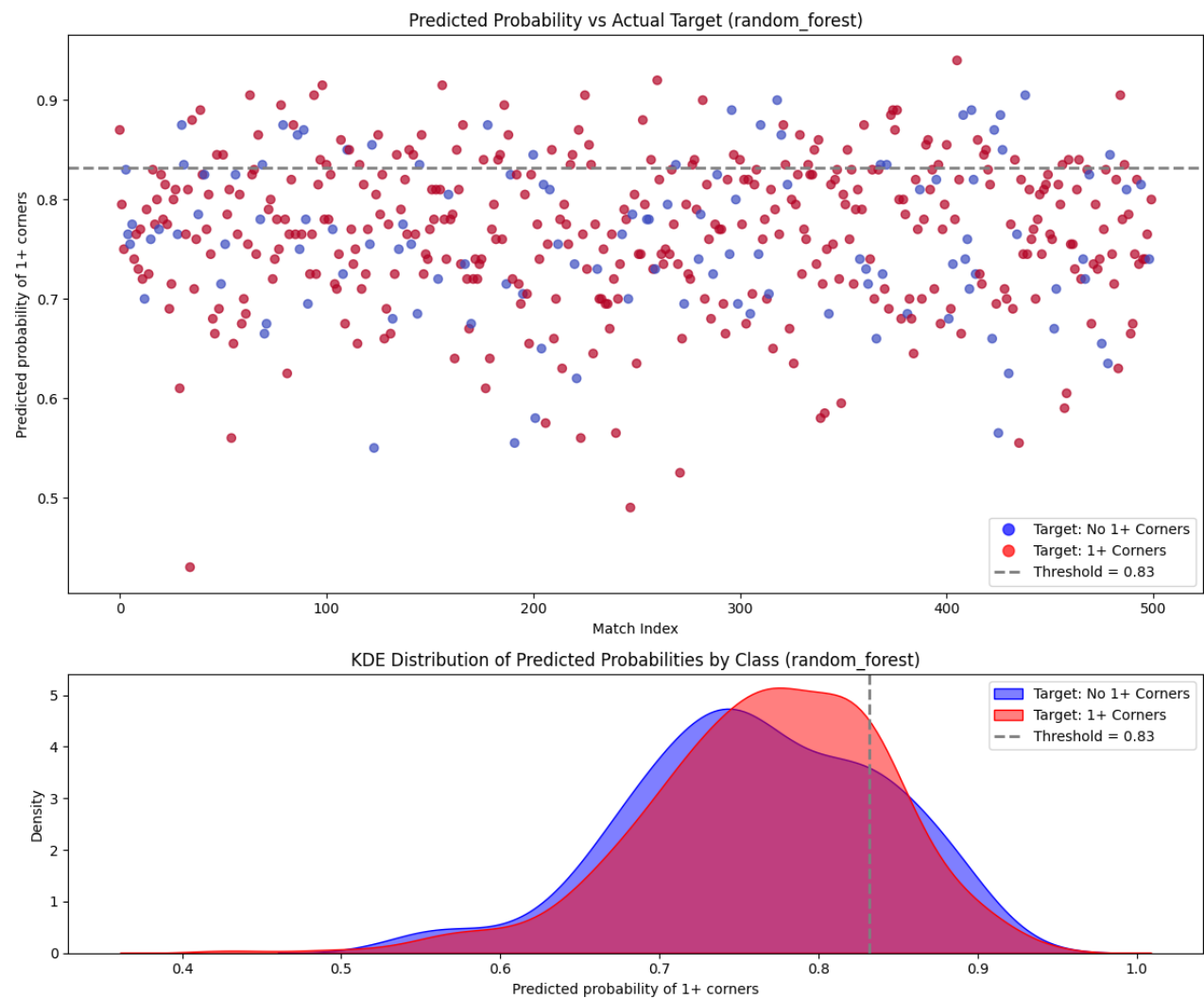
Classification Report (Test Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.23 | 0.79 | 0.36 | 118 |
| 1 | 0.74 | 0.18 | 0.29 | 382 |
| accuracy | | | 0.33 | 500 |
| macro avg | 0.48 | 0.49 | 0.32 | 500 |
| weighted avg | 0.62 | 0.33 | 0.31 | 500 |

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £425.37

🎲 Total Staked: £6773.07

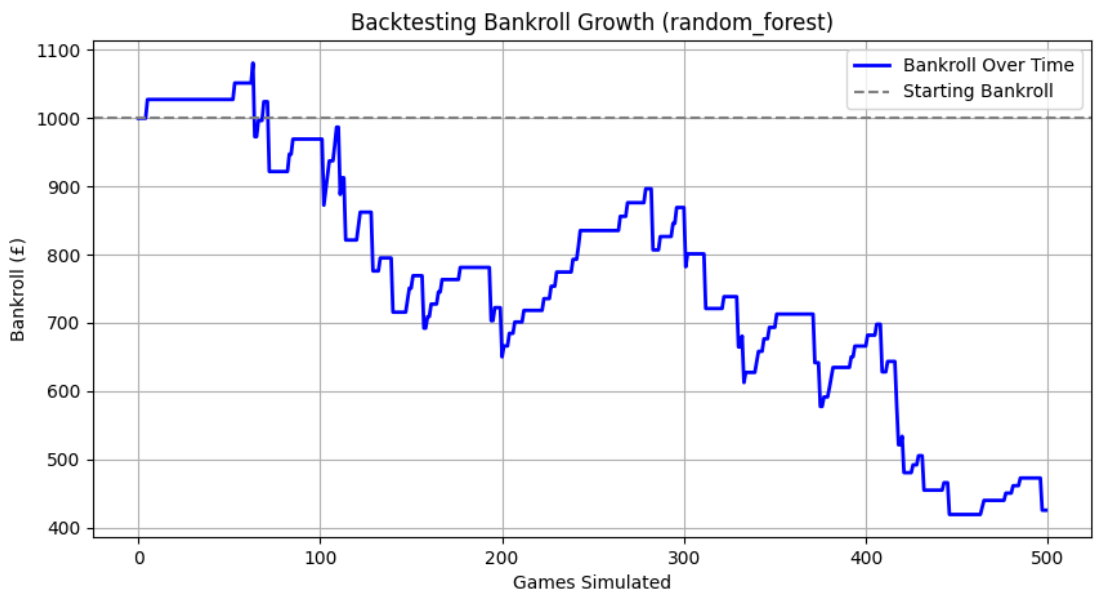
📈 Total Profit: £-574.63

🇮🇹 ROI: -57.46%

✅ Win rate: 73.91% over 92 bets

📉 Sharpe Ratio: -0.14

🎯 Edge Over Bookies: -5.86%



logistic_regression Model

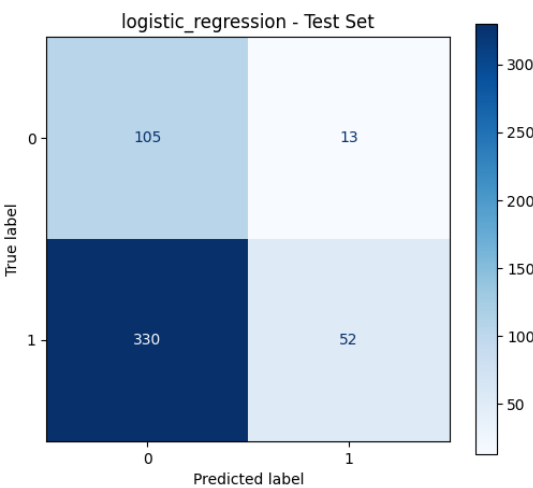
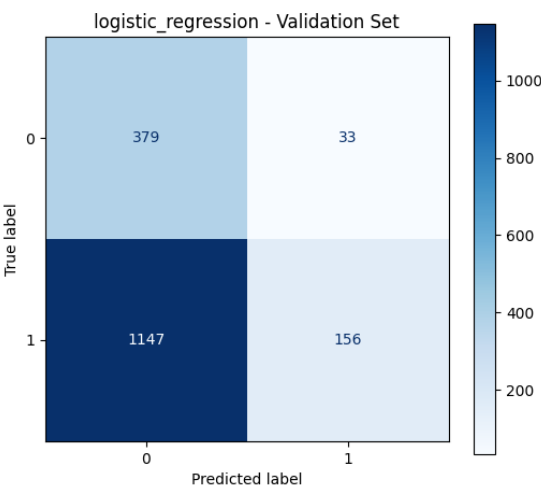
Feature Importance

| Top 8 Features | Importance | Bottom 5 Features | Importance |
|-------------------------|------------|-------------------|------------|
| away_shots_75_80 | 0.45198 | abs_goal_diff_80 | -0.97436 |
| home_shots_pre80 | 0.32872 | odd_h | -0.59493 |
| away_shots_pre80 | 0.27922 | odd_a | -0.29989 |
| home_shots_75_80 | 0.23116 | away_shots_70_75 | -0.28545 |
| home_corners_pre80 | 0.17042 | odd_d | -0.25521 |
| league_england | 0.15447 | | |
| home_yellow_cards_pre80 | 0.09027 | | |
| game_state_80_a | 0.08703 | | |

Optimal Threshold

- **Optimal threshold:** 0.57
- **Expected Precision:** 0.825
- **Expected Recall:** 0.120

Classification Reports



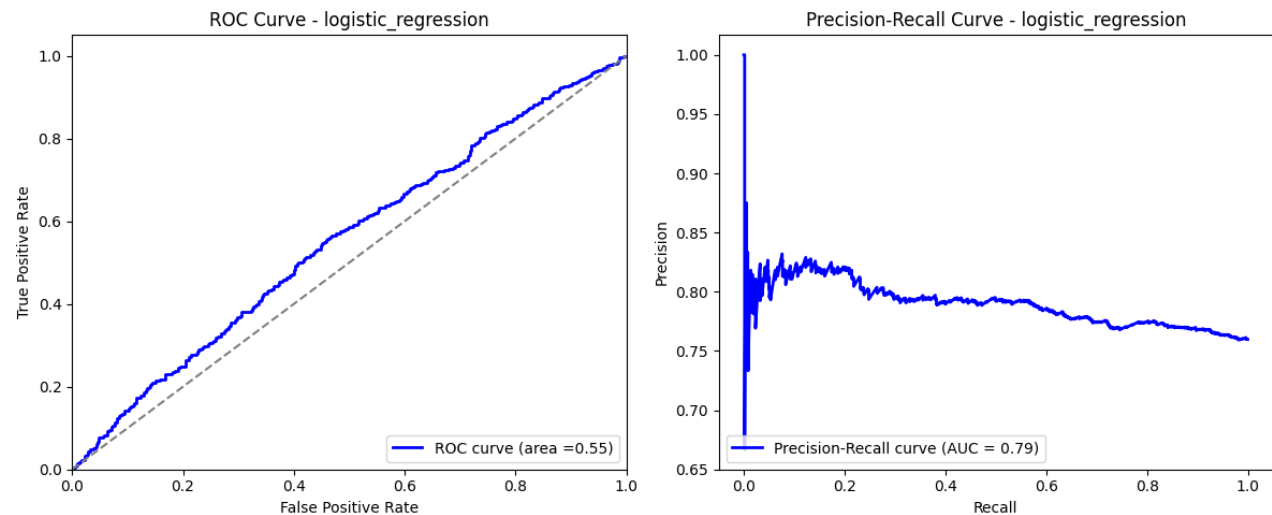
Classification Report (Validation Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.25 | 0.92 | 0.39 | 412 |
| 1 | 0.83 | 0.12 | 0.21 | 1303 |
| accuracy | | | 0.31 | 1715 |
| macro avg | 0.54 | 0.52 | 0.30 | 1715 |
| weighted avg | 0.69 | 0.31 | 0.25 | 1715 |

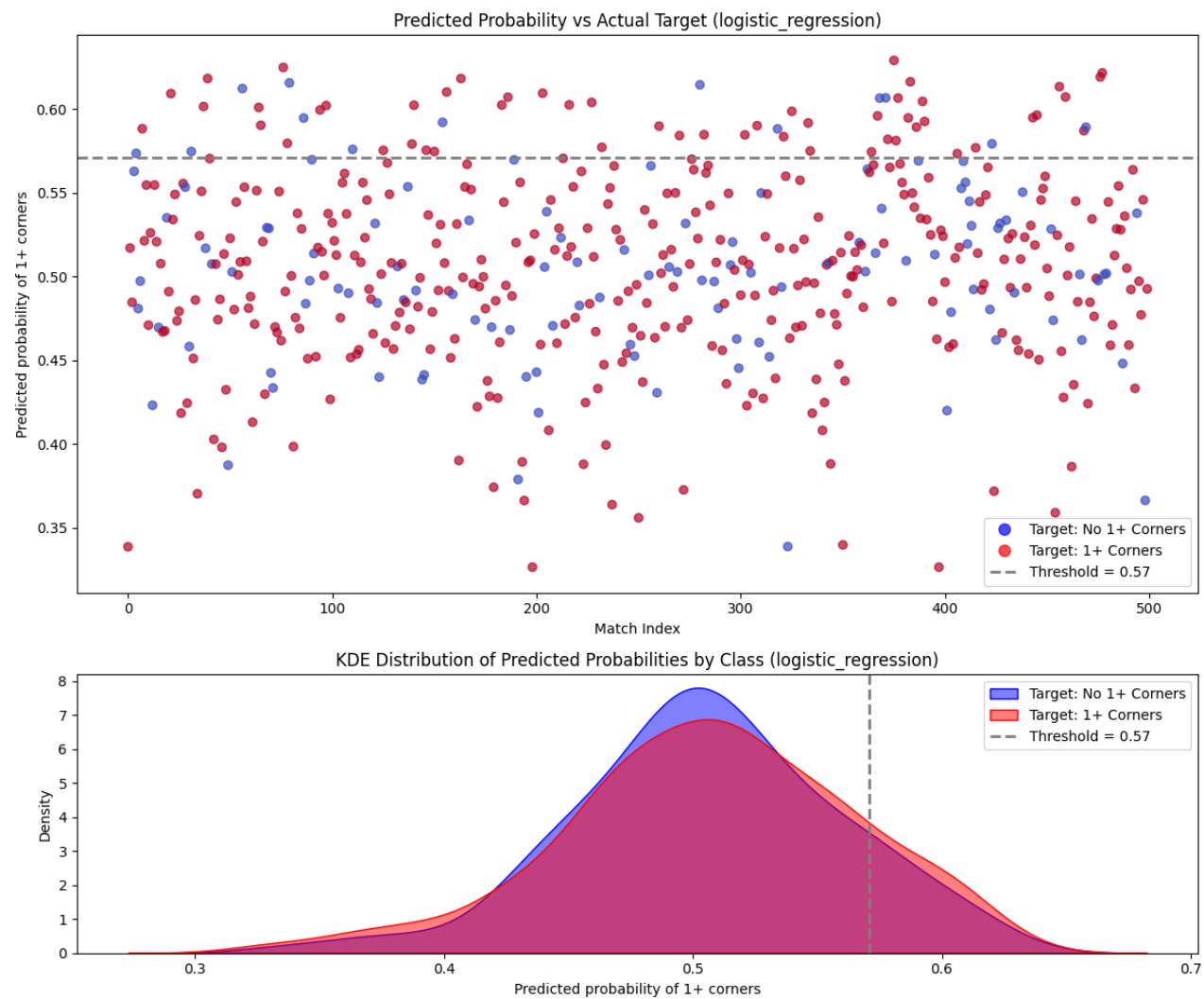
Classification Report (Test Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.24 | 0.89 | 0.38 | 118 |
| 1 | 0.80 | 0.14 | 0.23 | 382 |
| accuracy | | | 0.31 | 500 |
| macro avg | 0.52 | 0.51 | 0.31 | 500 |
| weighted avg | 0.67 | 0.31 | 0.27 | 500 |

ROC and Precision-Recall Curves



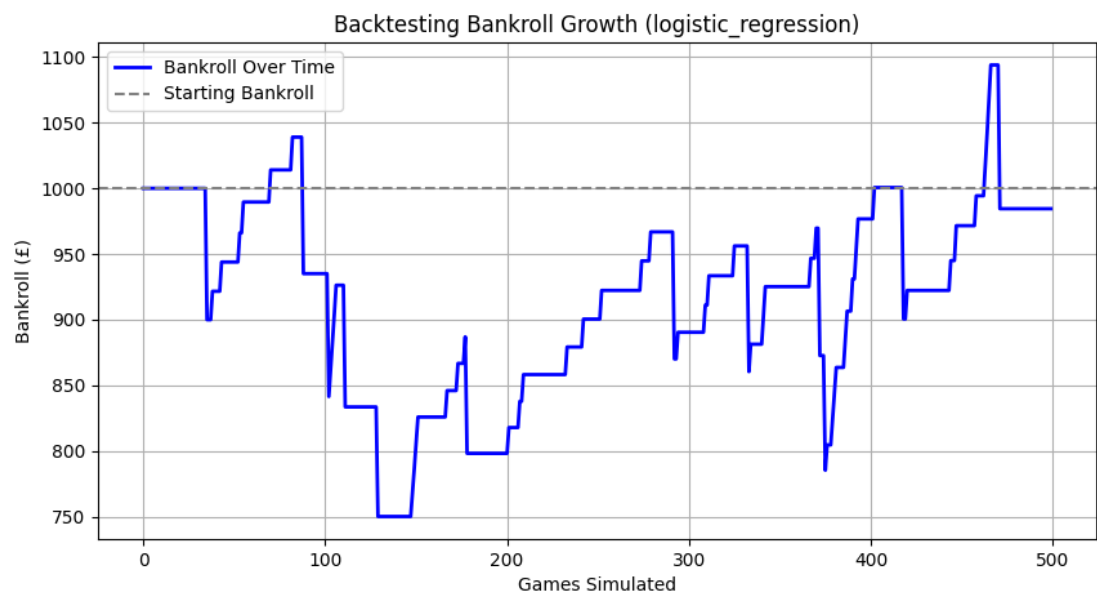
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £984.48
- 🎰 Total Staked: £5807.55
- 📈 Total Profit: £-15.52
- 📊 ROI: -1.55%
- ✅ Win rate: 81.25% over 64 bets
- 📉 Sharpe Ratio: 0.02
- 🎯 Edge Over Bookies: 1.47%



svc Model

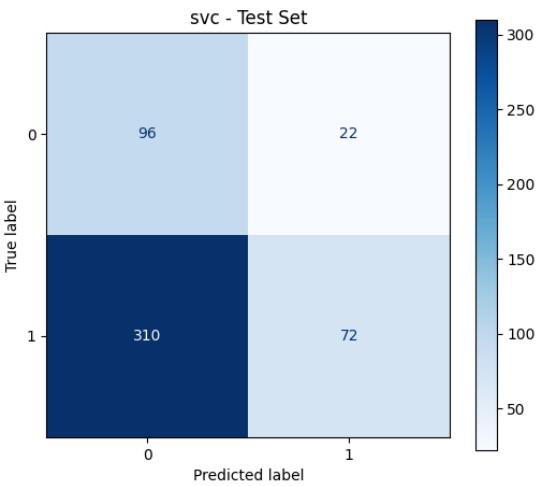
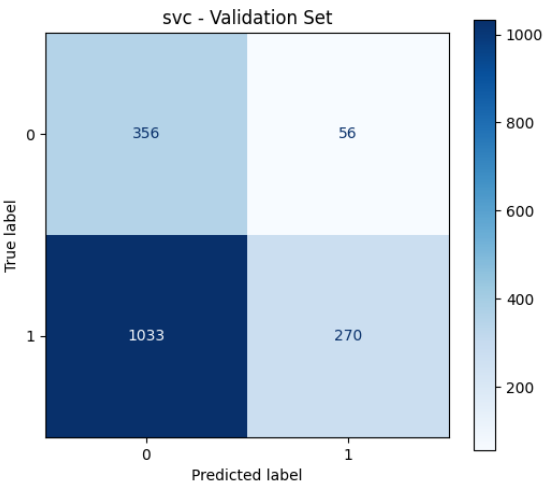
Feature Importance

| Top 8 Features | Importance | Bottom 5 Features | Importance |
|------------------|------------|----------------------|------------|
| goal_diff_80 | 0.00000 | goal_diff_squared_80 | 0.00000 |
| home_shots_pre80 | 0.00000 | home_corners_75_80 | 0.00000 |
| odd_d | 0.00000 | away_shots_70_75 | 0.00000 |
| odd_a | 0.00000 | home_shots_70_75 | 0.00000 |
| game_state_80_h | 0.00000 | away_corners_70_75 | 0.00000 |
| game_state_80_a | 0.00000 | | |
| game_state_80_d | 0.00000 | | |
| league_england | 0.00000 | | |

Optimal Threshold

- **Optimal threshold:** 0.78
- **Expected Precision:** 0.826
- **Expected Recall:** 0.200

Classification Reports



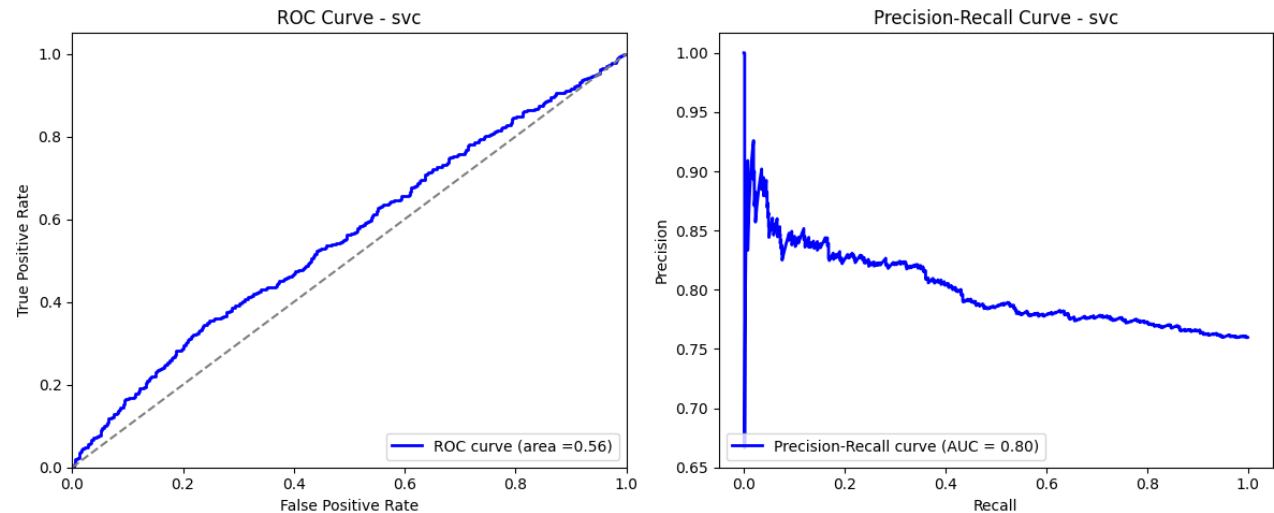
Classification Report (*Validation Set*)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.26 | 0.86 | 0.40 | 412 |
| 1 | 0.83 | 0.21 | 0.33 | 1303 |
| accuracy | | | 0.37 | 1715 |
| macro avg | 0.54 | 0.54 | 0.36 | 1715 |
| weighted avg | 0.69 | 0.37 | 0.35 | 1715 |

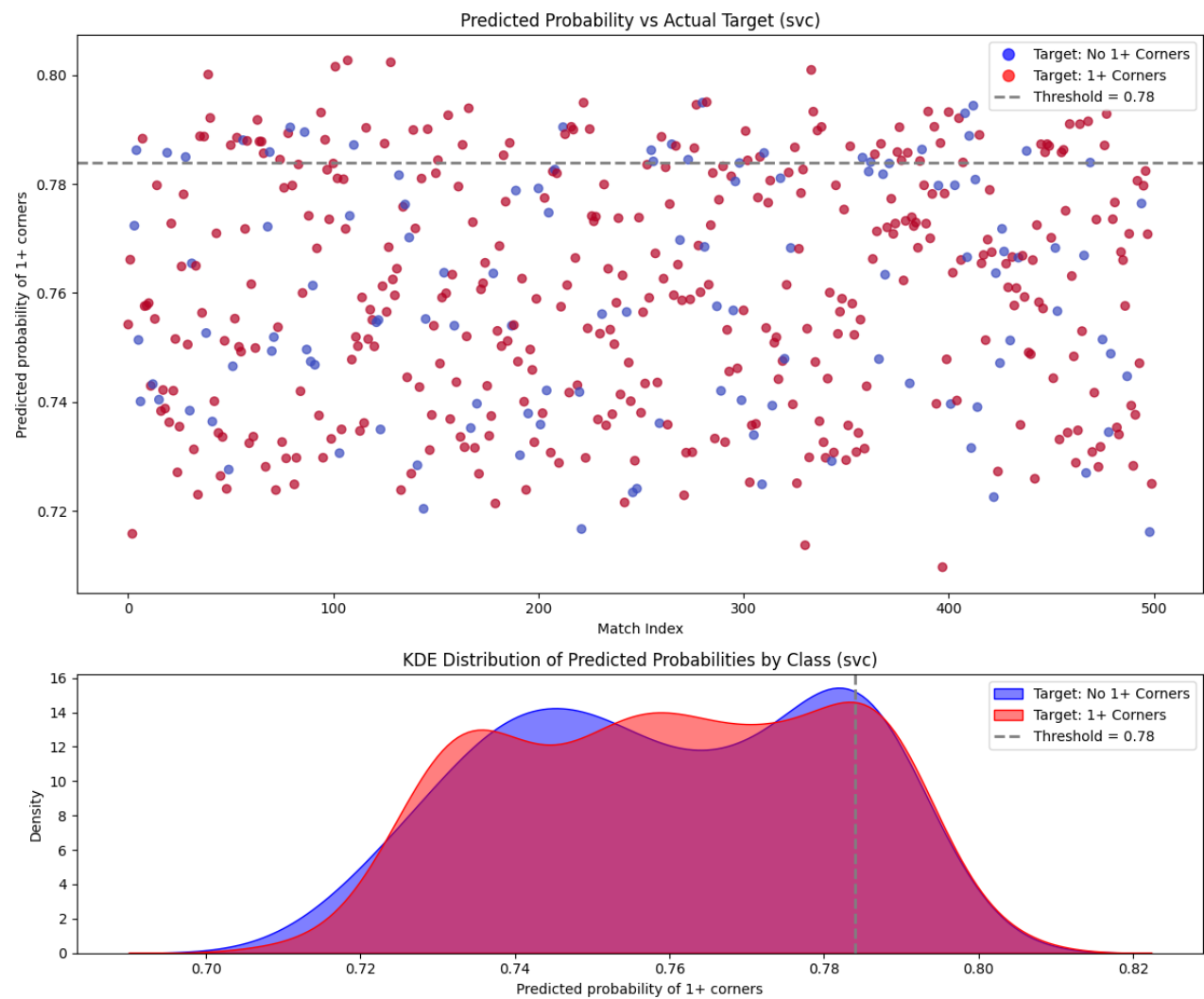
Classification Report (*Test Set*)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.24 | 0.81 | 0.37 | 118 |
| 1 | 0.77 | 0.19 | 0.30 | 382 |
| accuracy | | | 0.34 | 500 |
| macro avg | 0.50 | 0.50 | 0.33 | 500 |
| weighted avg | 0.64 | 0.34 | 0.32 | 500 |

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £609.58

🎲 Total Staked: £6630.49

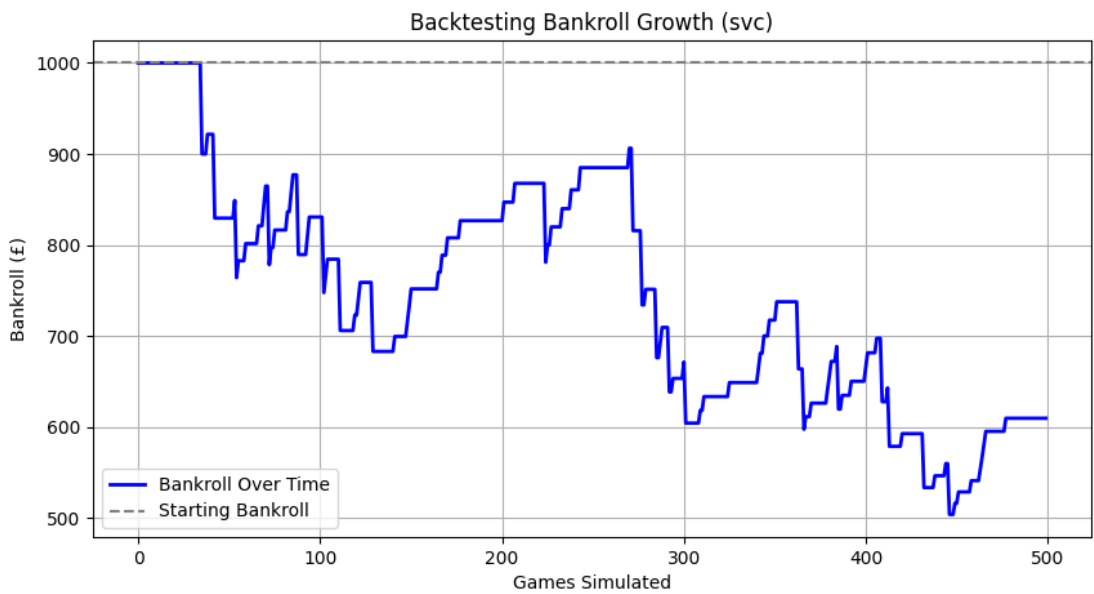
📈 Total Profit: £-390.42

🇮🇹 ROI: -39.04%

✅ Win rate: 77.17% over 92 bets

📉 Sharpe Ratio: -0.07

🎯 Edge Over Bookies: -2.60%



xgboost Model

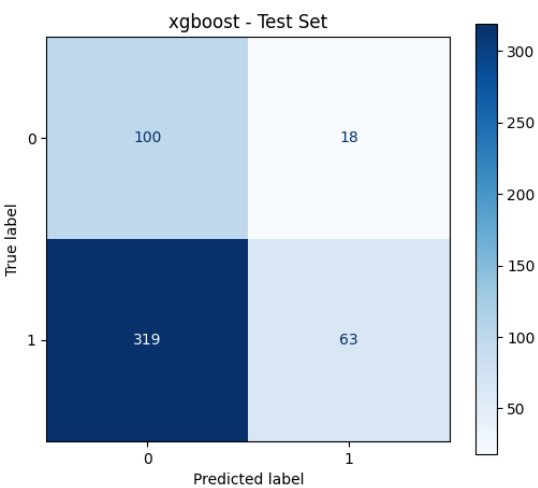
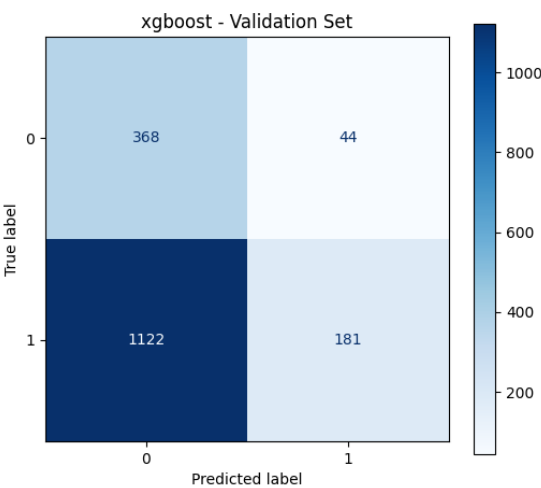
Feature Importance

| Top 8 Features | Importance | Bottom 5 Features | Importance |
|------------------------|------------|-------------------------|------------|
| away_sending_off_pre80 | 0.05902 | goal_diff_squared_80 | 0.00000 |
| league_england | 0.05112 | close_game_80 | 0.00000 |
| abs_goal_diff_80 | 0.05094 | league_avg_home_corners | 0.00000 |
| league_spain | 0.04169 | league_france | 0.00000 |
| odd_a | 0.04133 | home_sending_off_pre80 | 0.00000 |
| league_avg_corners | 0.04041 | | |
| goal_diff_80 | 0.03977 | | |
| home_shots_pre80 | 0.03755 | | |

Optimal Threshold

- **Optimal threshold:** 0.81
- **Expected Precision:** 0.804
- **Expected Recall:** 0.139

Classification Reports



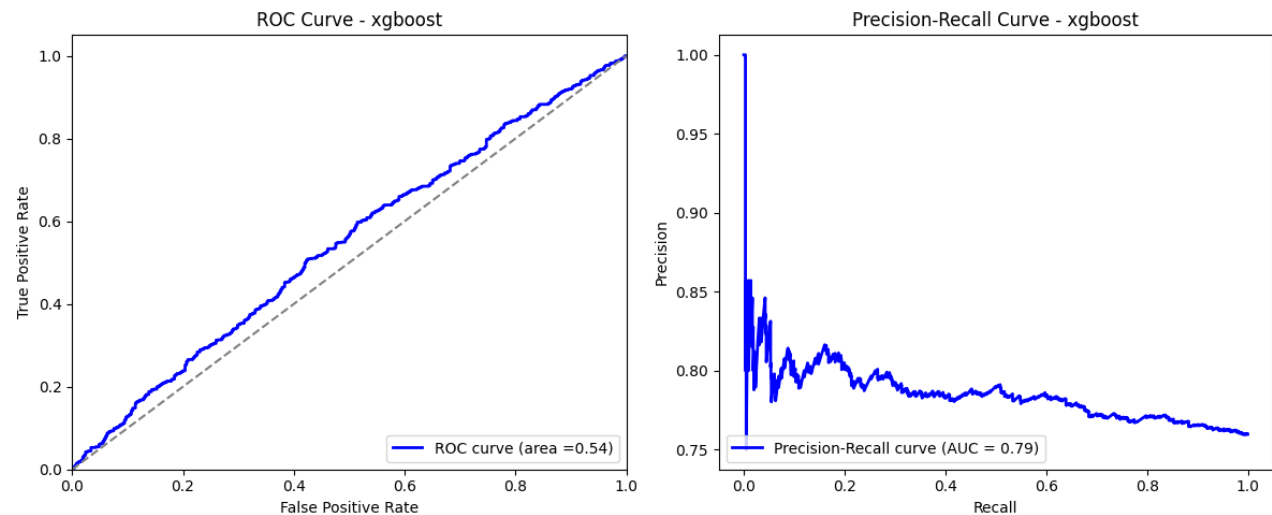
Classification Report (Validation Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.25 | 0.89 | 0.39 | 412 |
| 1 | 0.80 | 0.14 | 0.24 | 1303 |
| accuracy | | | 0.32 | 1715 |
| macro avg | 0.53 | 0.52 | 0.31 | 1715 |
| weighted avg | 0.67 | 0.32 | 0.27 | 1715 |

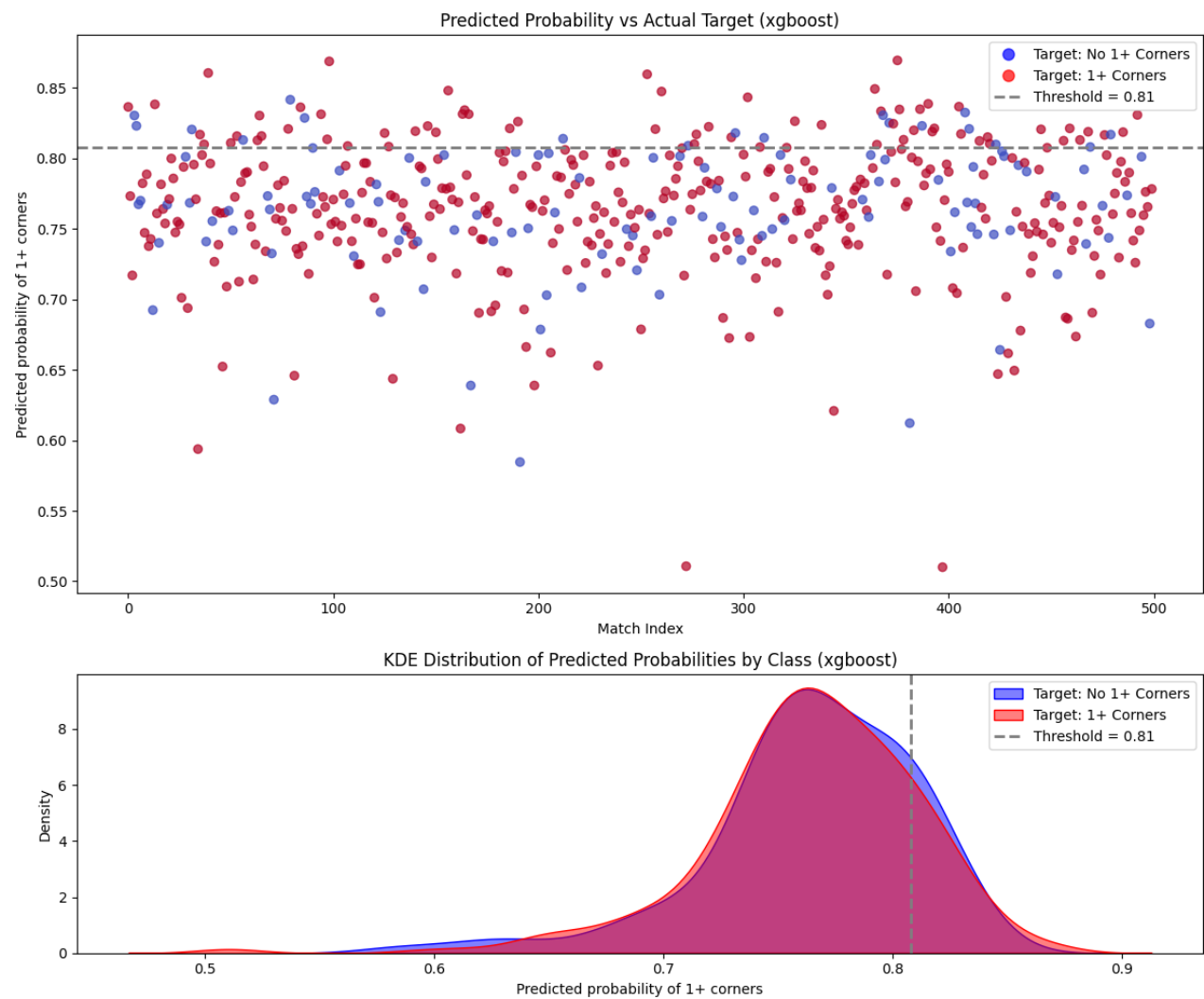
Classification Report (Test Set)

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.24 | 0.85 | 0.37 | 118 |
| 1 | 0.78 | 0.16 | 0.27 | 382 |
| accuracy | | | 0.33 | 500 |
| macro avg | 0.51 | 0.51 | 0.32 | 500 |
| weighted avg | 0.65 | 0.33 | 0.30 | 500 |

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £760.63

📊 Total Staked: £7456.60

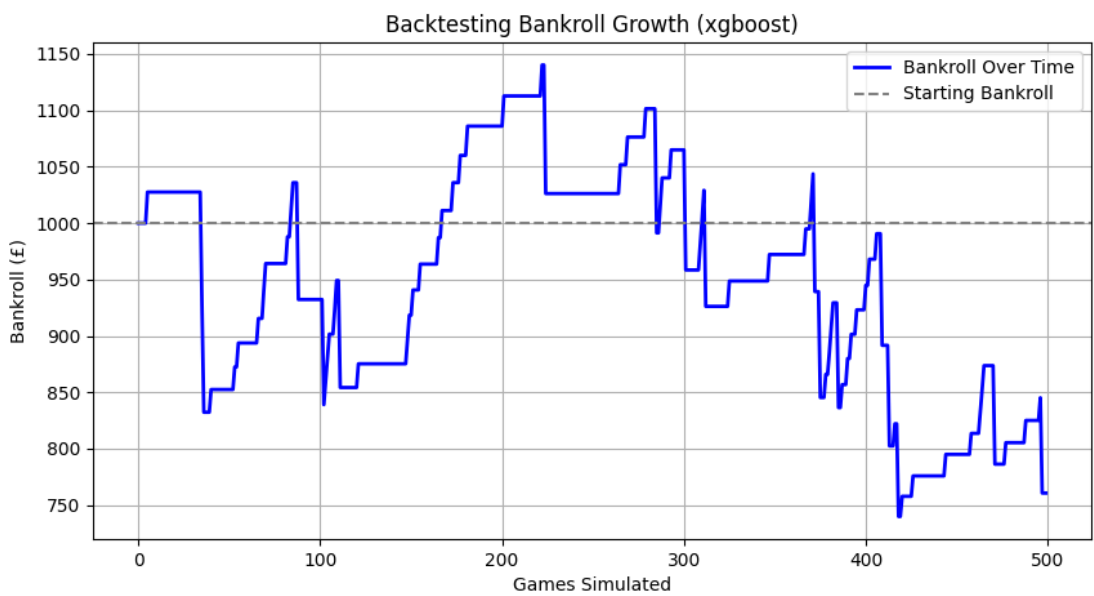
📈 Total Profit: £-239.37

📉 ROI: -23.94%

✅ Win rate: 78.75% over 80 bets

📉 Sharpe Ratio: -0.04

🎯 Edge Over Bookies: -1.03%



Backtesting Results Summary

