

Model Training Report: Track 1

Configuration Details

Target Variable

target

Selected Features

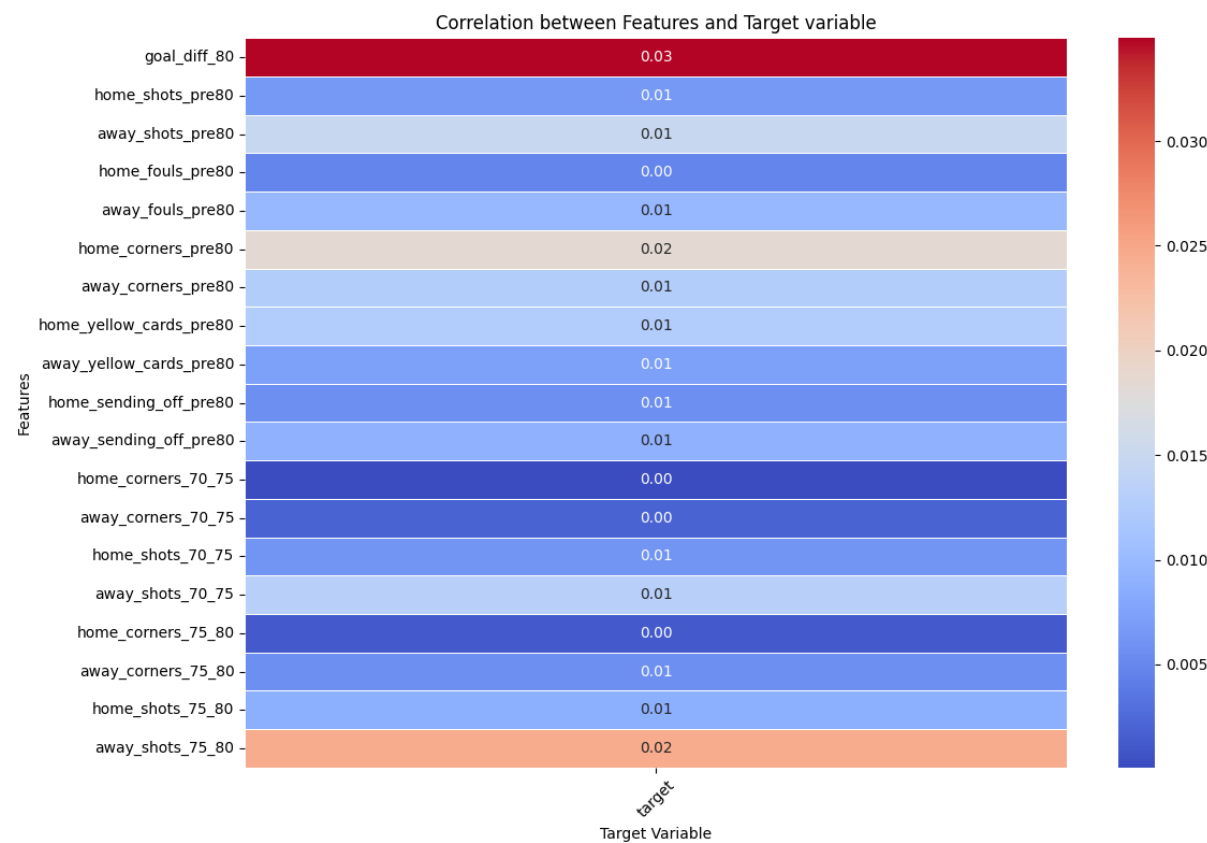
goal_diff_80, home_shots_pre80, away_shots_pre80, home_fouls_pre80, away_fouls_pre80, home_corners_pre80, away_corners_pre80, home_yellow_cards_pre80, away_yellow_cards_pre80, home_sending_off_pre80, away_sending_off_pre80, home_corners_70_75, away_corners_70_75, home_shots_70_75, away_shots_70_75, home_corners_75_80,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

Constructed Features

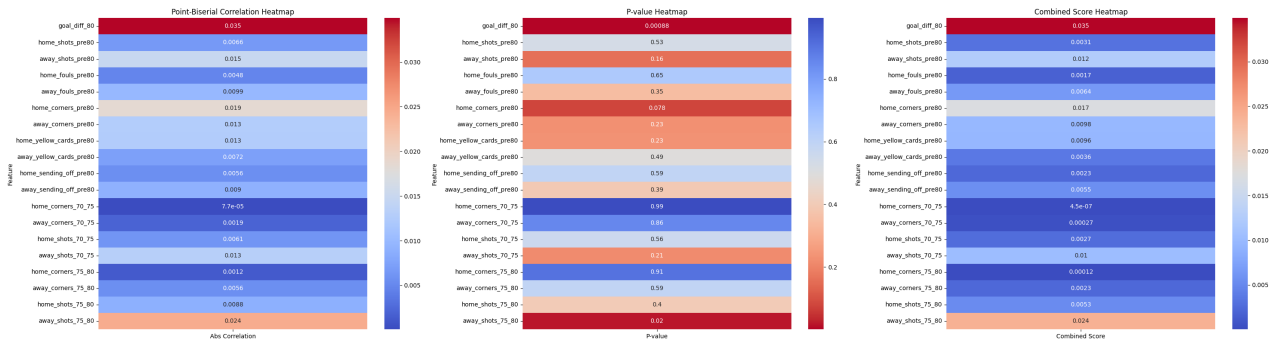
Models Trained (& Hyperparameters)

- **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}

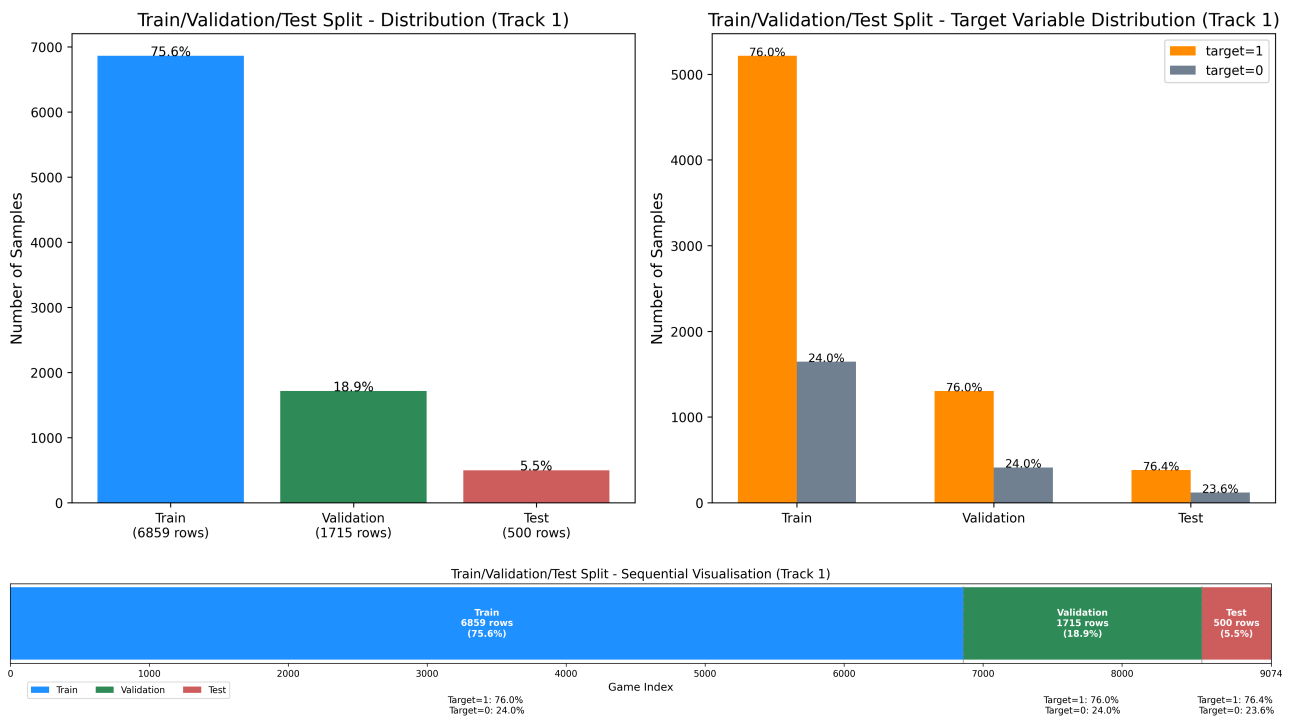
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

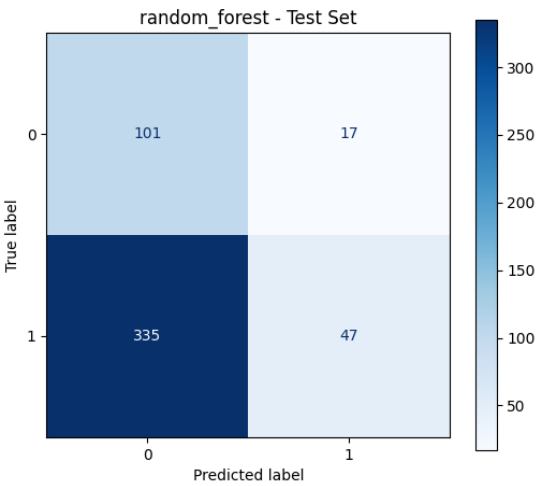
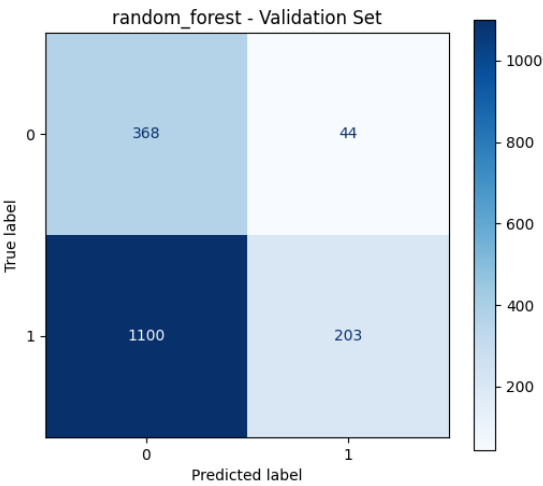
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
home_shots_pre80	0.10485	home_sending_off_pre80	0.01021
away_fouls_pre80	0.10088	away_sending_off_pre80	0.01309
home_fouls_pre80	0.09976	away_corners_70_75	0.01970
away_shots_pre80	0.09955	away_corners_75_80	0.02016
home_corners_pre80	0.08191	home_corners_70_75	0.02271
away_corners_pre80	0.07757		
goal_diff_80	0.06790		
away_yellow_cards_pre80	0.06203		

Optimal Threshold

- **Optimal threshold:** 0.83
- **Expected Precision:** 0.822
- **Expected Recall:** 0.156

Classification Reports



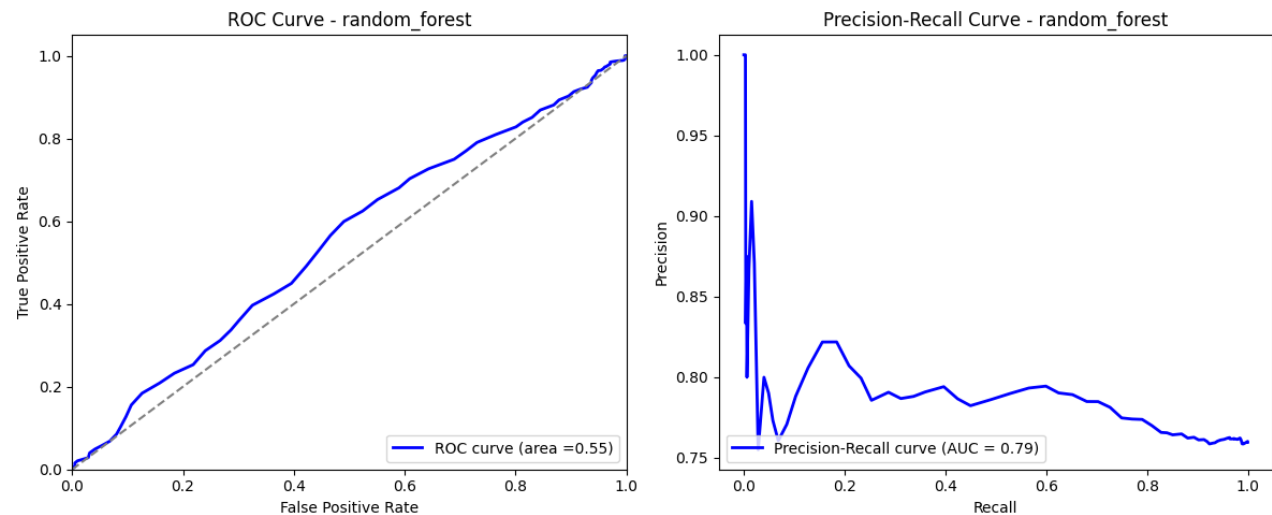
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.25	0.89	0.39	412
1	0.82	0.16	0.26	1303
accuracy			0.33	1715
macro avg	0.54	0.52	0.33	1715
weighted avg	0.68	0.33	0.29	1715

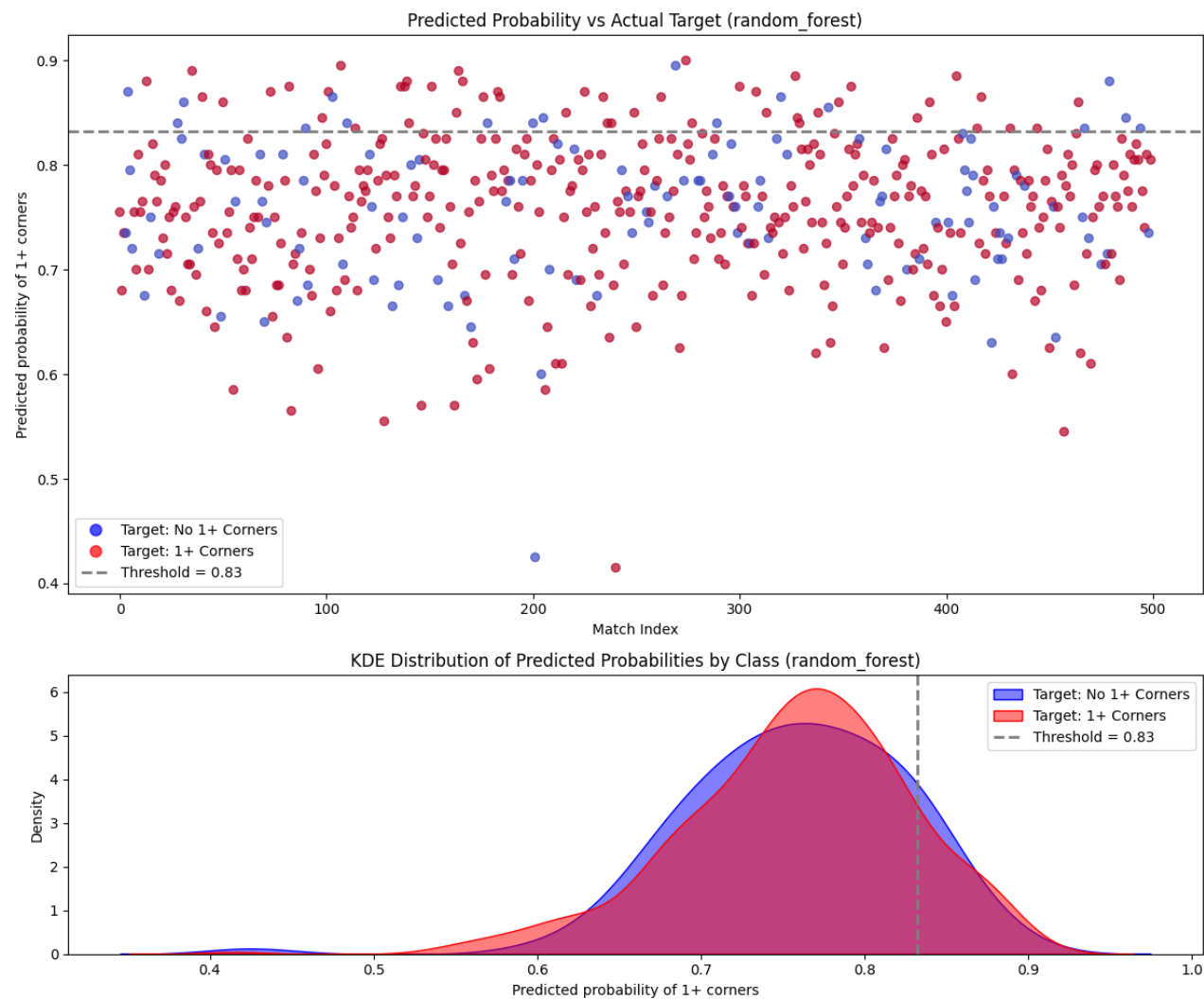
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.23	0.86	0.36	118
1	0.73	0.12	0.21	382
accuracy			0.30	500
macro avg	0.48	0.49	0.29	500
weighted avg	0.62	0.30	0.25	500

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £589.44

🎲 Total Staked: £5029.21

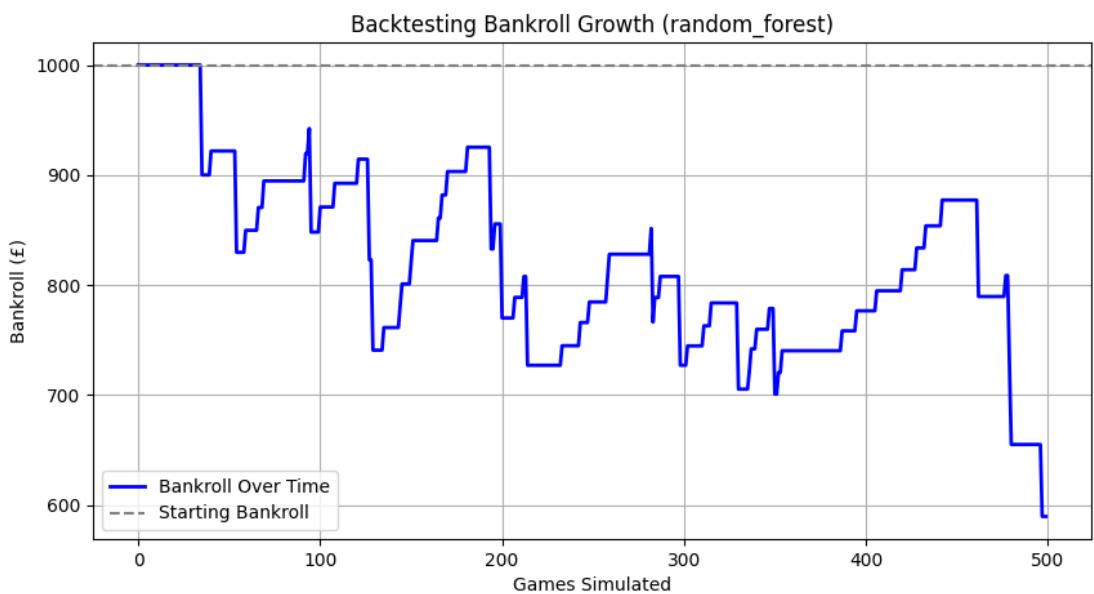
📈 Total Profit: £-410.56

🇮🇹 ROI: -41.06%

✅ Win rate: 74.19% over 62 bets

📉 Sharpe Ratio: -0.12

🎯 Edge Over Bookies: -5.58%



logistic_regression Model

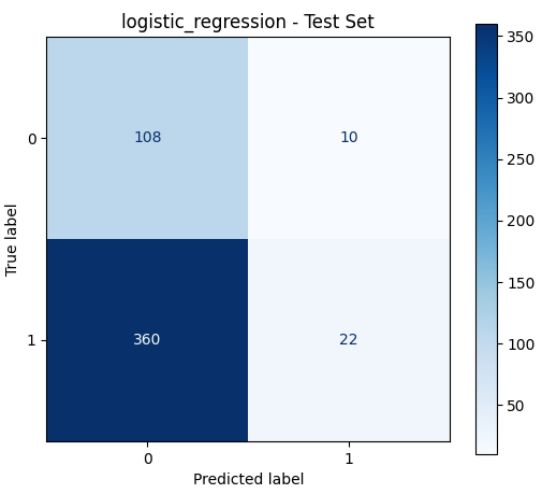
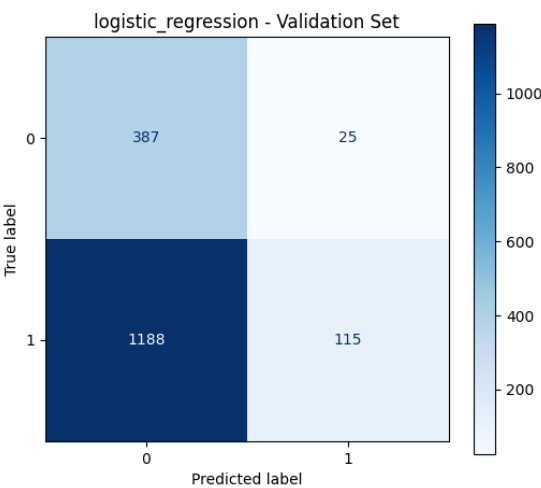
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
away_shots_75_80	0.42965	goal_diff_80	-0.73217
home_corners_pre80	0.32319	away_shots_70_75	-0.30252
home_shots_75_80	0.21567	home_fouls_pre80	-0.28350
home_yellow_cards_pre80	0.21060	away_fouls_pre80	-0.10267
home_shots_pre80	0.15527	home_corners_75_80	-0.05695
away_yellow_cards_pre80	0.14210		
away_corners_pre80	0.12133		
away_shots_pre80	0.10991		

Optimal Threshold

- **Optimal threshold:** 0.55
- **Expected Precision:** 0.816
- **Expected Recall:** 0.085

Classification Reports



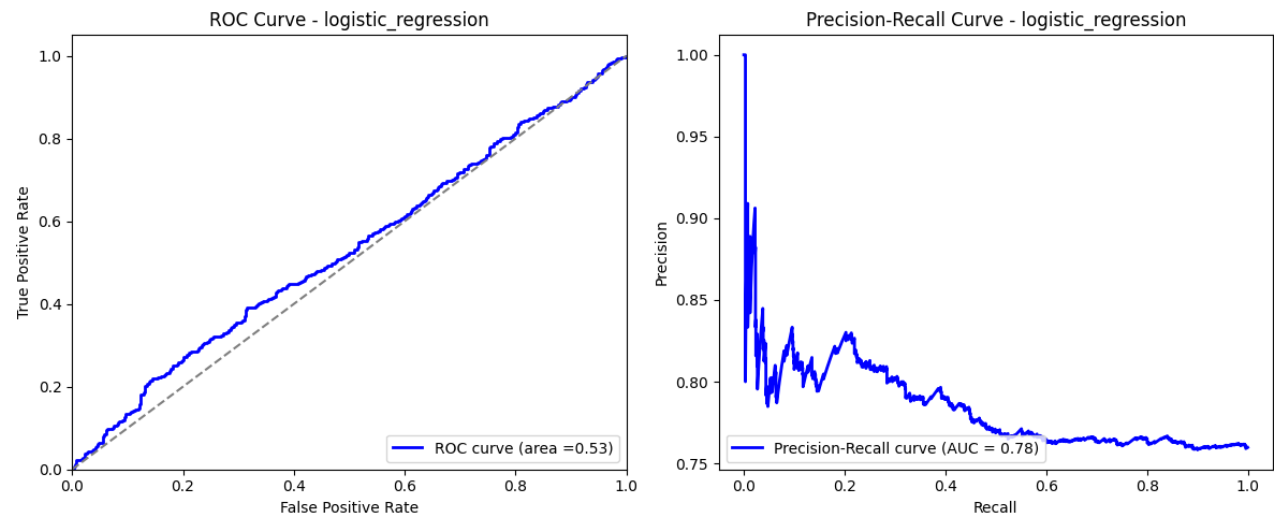
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.25	0.94	0.39	412
1	0.82	0.09	0.16	1303
accuracy			0.29	1715
macro avg	0.53	0.51	0.27	1715
weighted avg	0.68	0.29	0.21	1715

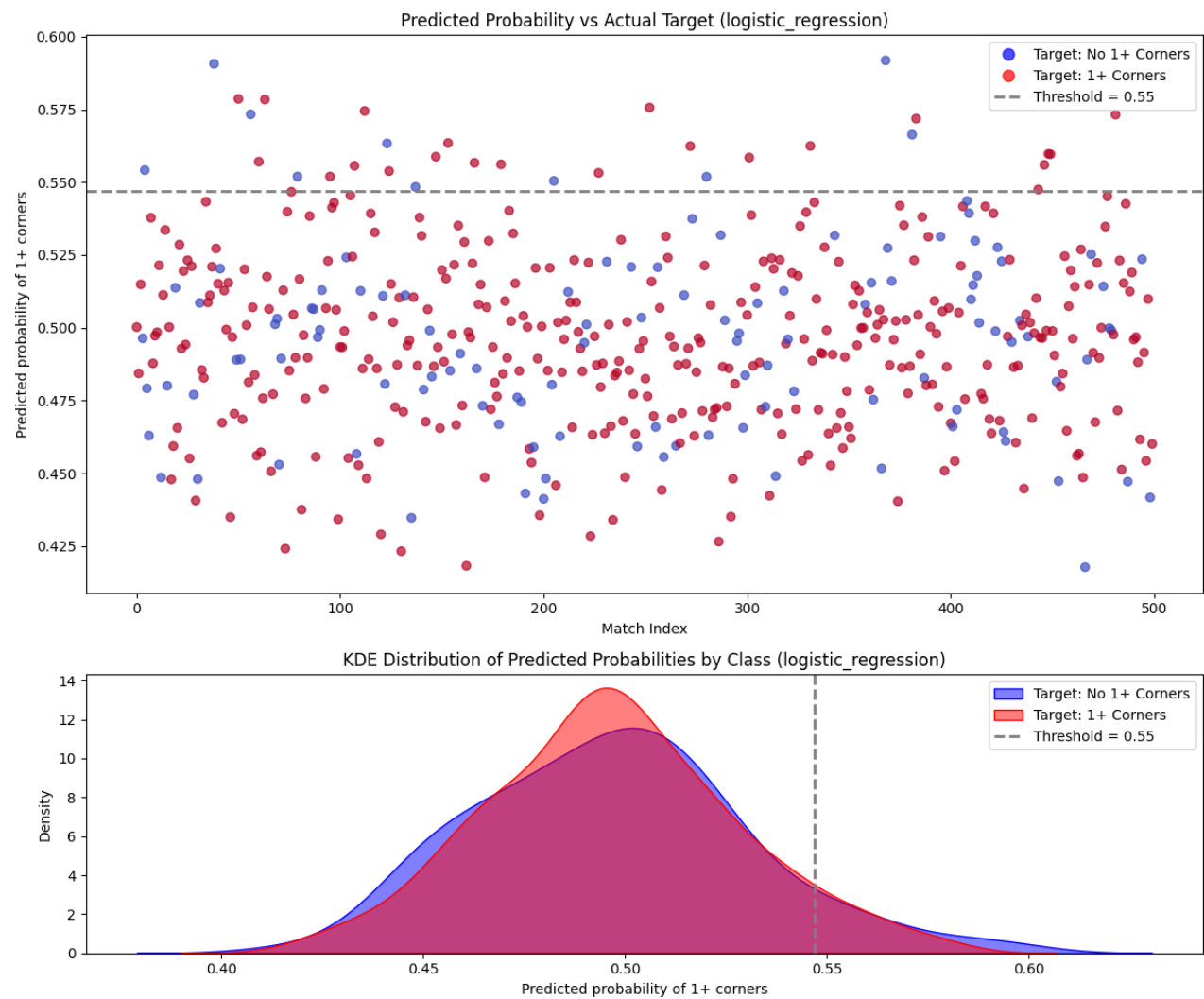
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.23	0.92	0.37	118
1	0.69	0.06	0.11	382
accuracy			0.26	500
macro avg	0.46	0.49	0.24	500
weighted avg	0.58	0.26	0.17	500

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £598.43

🎲 Total Staked: £2090.89

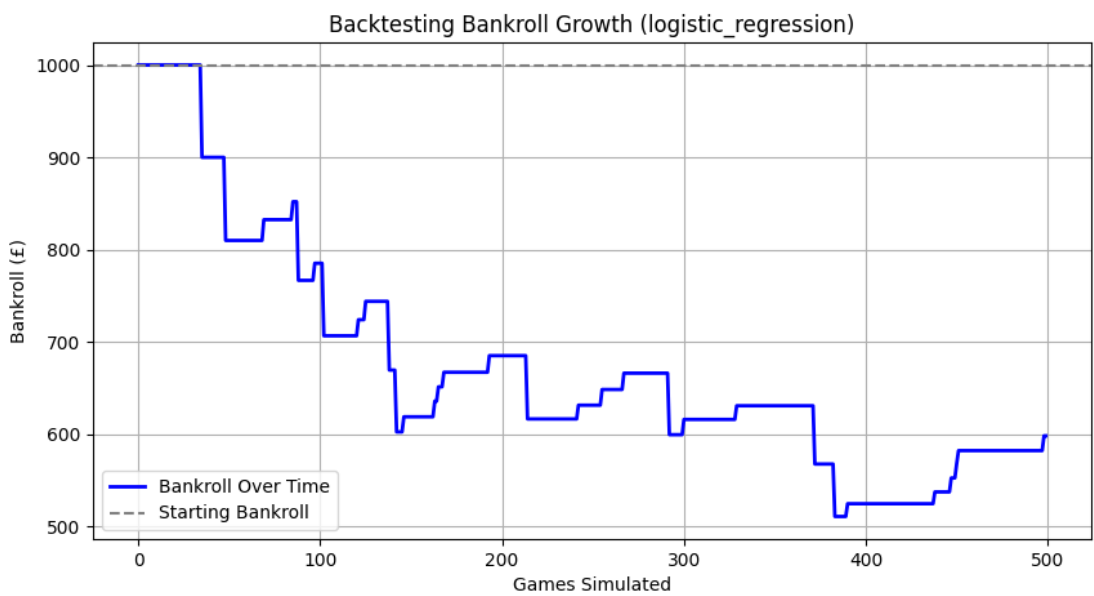
📈 Total Profit: £-401.57

🇮🇹 ROI: -40.16%

✅ Win rate: 67.74% over 31 bets

📉 Sharpe Ratio: -0.24

🎯 Edge Over Bookies: -12.03%



svc Model

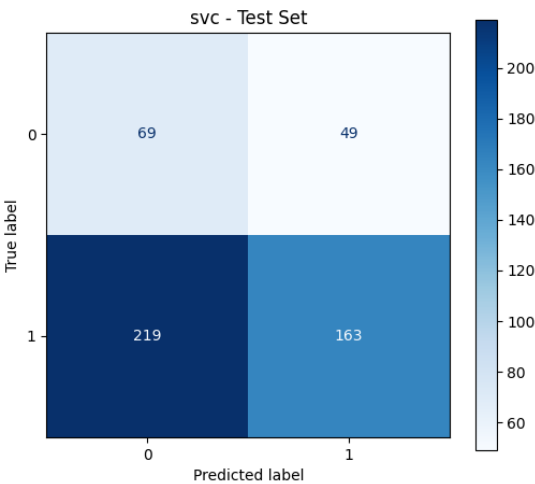
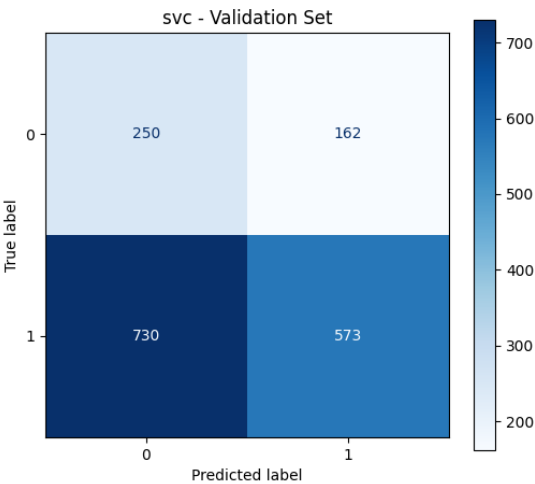
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
goal_diff_80	0.00000	away_shots_75_80	0.00000
away_sending_off_pre80	0.00000	away_shots_pre80	0.00000
home_shots_75_80	0.00000	home_fouls_pre80	0.00000
away_corners_75_80	0.00000	away_fouls_pre80	0.00000
home_corners_75_80	0.00000	home_corners_pre80	0.00000
away_shots_70_75	0.00000		
home_shots_70_75	0.00000		
away_corners_70_75	0.00000		

Optimal Threshold

- **Optimal threshold:** 0.76
- **Expected Precision:** 0.773
- **Expected Recall:** 0.457

Classification Reports



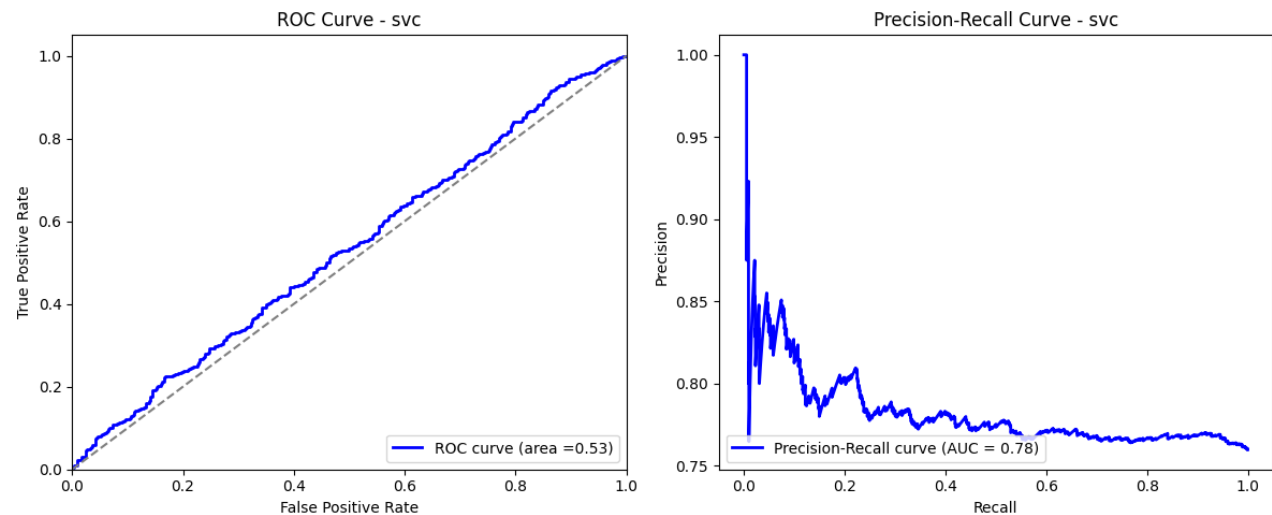
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.26	0.61	0.36	412
1	0.78	0.44	0.56	1303
accuracy			0.48	1715
macro avg	0.52	0.52	0.46	1715
weighted avg	0.65	0.48	0.51	1715

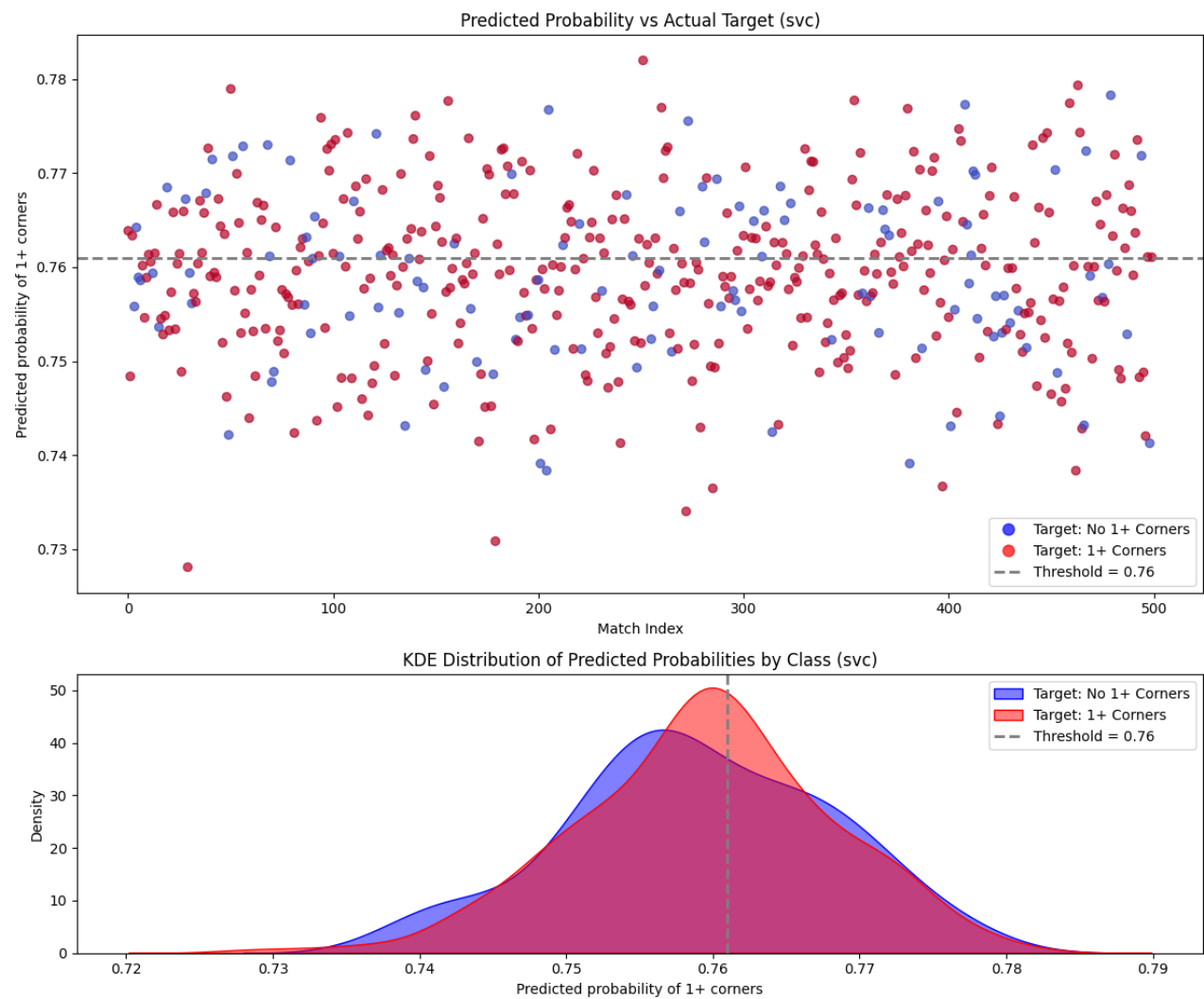
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.24	0.58	0.34	118
1	0.77	0.43	0.55	382
accuracy			0.46	500
macro avg	0.50	0.51	0.44	500
weighted avg	0.64	0.46	0.50	500

ROC and Precision-Recall Curves



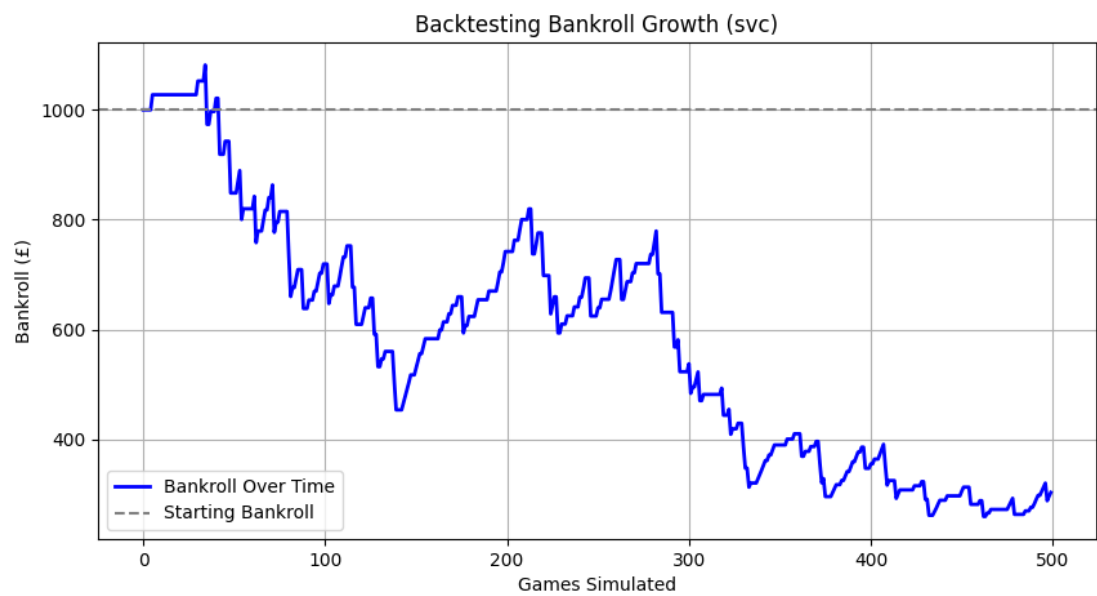
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £303.44
- 📁 Total Staked: £11174.77
- 📈 Total Profit: £-696.56
- 🇮🇹 ROI: -69.66%
- ✅ Win rate: 76.47% over 204 bets
- 📉 Sharpe Ratio: -0.08
- 🎯 Edge Over Bookies: -3.30%



xgboost Model

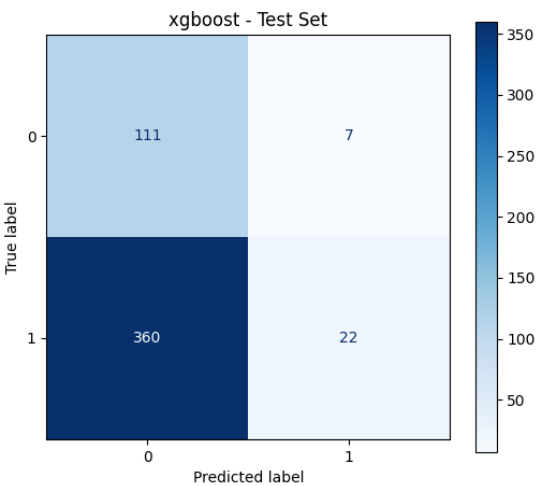
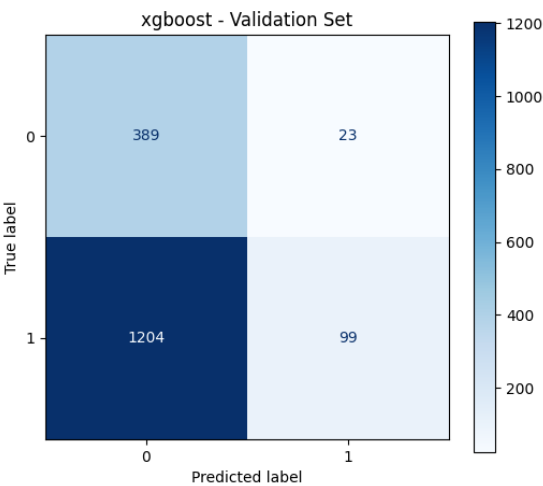
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
goal_diff_80	0.07444	home_sending_off_pre80	0.02721
home_corners_70_75	0.06246	home_corners_75_80	0.03587
home_yellow_cards_pre80	0.06131	away_corners_70_75	0.04004
away_yellow_cards_pre80	0.05748	away_corners_75_80	0.04482
home_shots_pre80	0.05691	away_shots_75_80	0.04818
home_corners_pre80	0.05634		
home_fouls_pre80	0.05630		
away_fouls_pre80	0.05626		

Optimal Threshold

- **Optimal threshold:** 0.81
- **Expected Precision:** 0.811
- **Expected Recall:** 0.076

Classification Reports



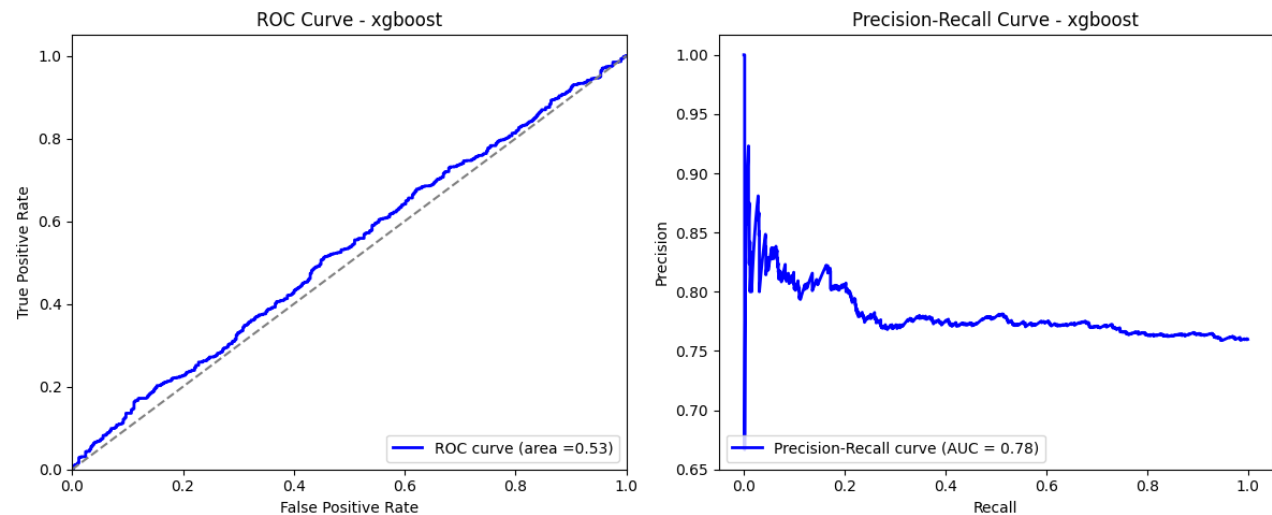
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.24	0.94	0.39	412
1	0.81	0.08	0.14	1303
accuracy			0.28	1715
macro avg	0.53	0.51	0.26	1715
weighted avg	0.68	0.28	0.20	1715

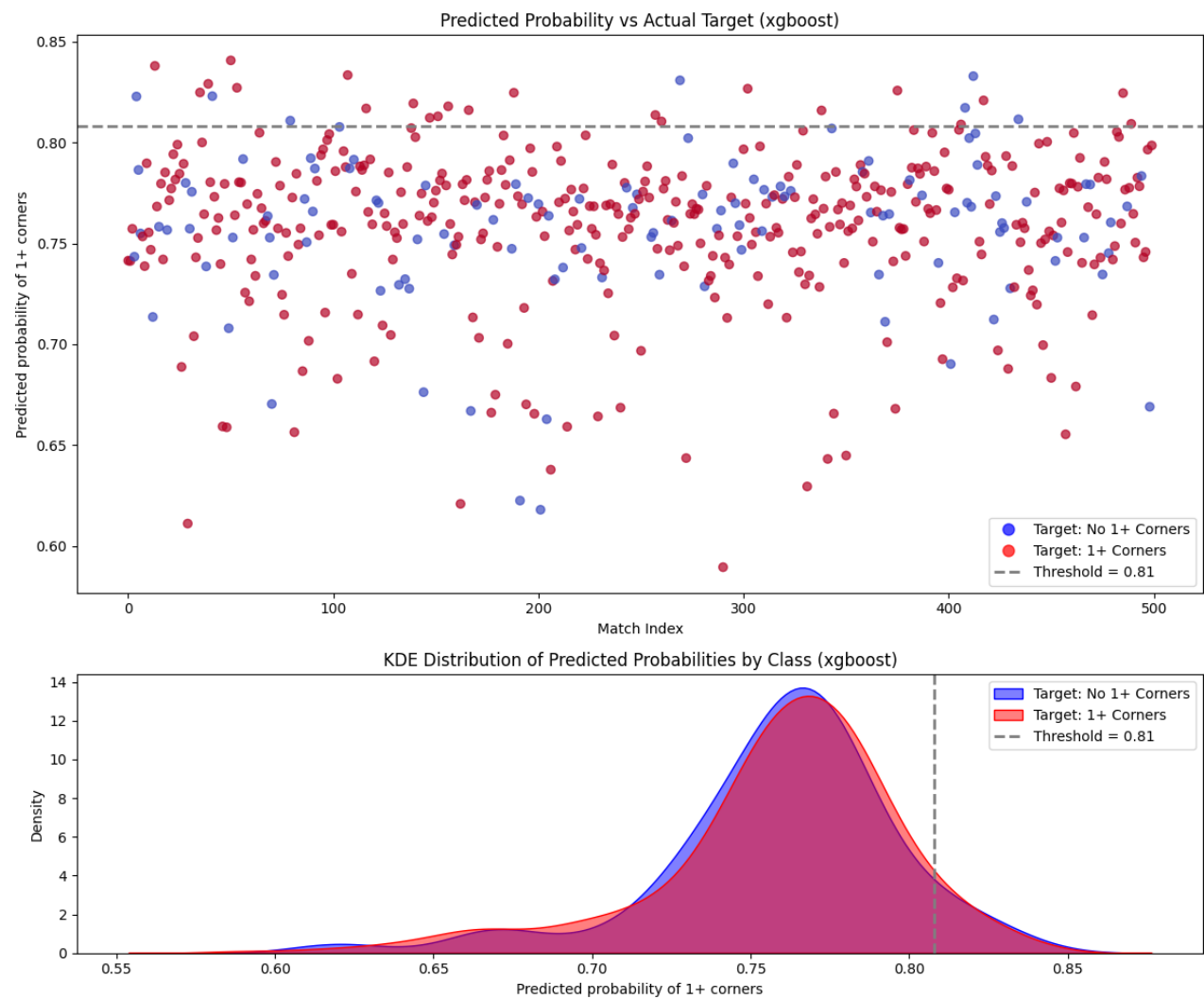
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.24	0.94	0.38	118
1	0.76	0.06	0.11	382
accuracy			0.27	500
macro avg	0.50	0.50	0.24	500
weighted avg	0.64	0.27	0.17	500

ROC and Precision-Recall Curves



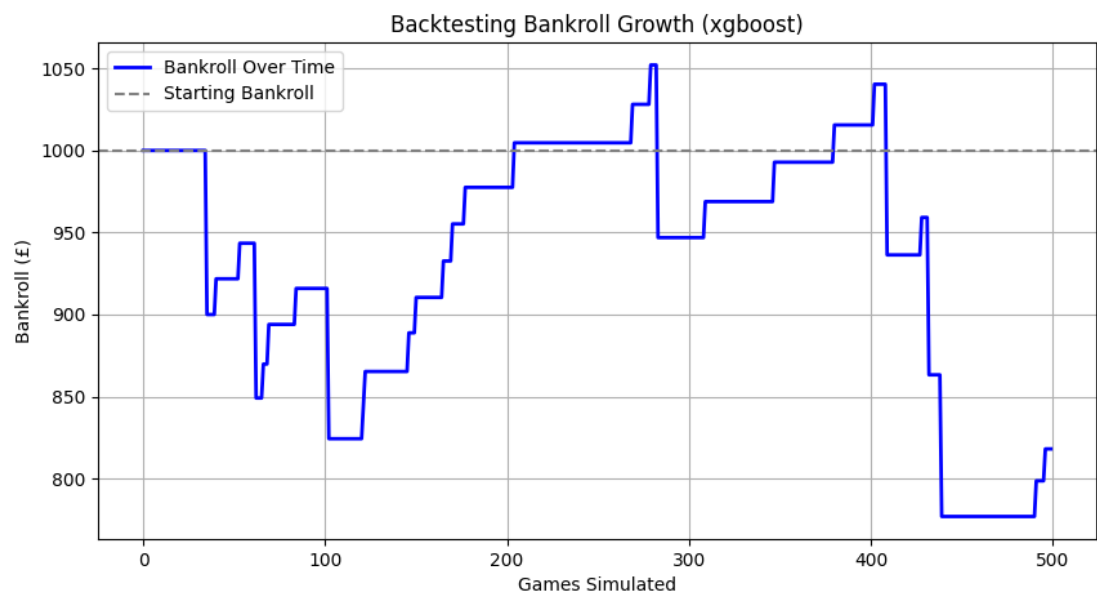
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £818.11
- 📁 Total Staked: £2687.65
- 📈 Total Profit: £-181.89
- 🇮🇹 ROI: -18.19%
- ✅ Win rate: 75.86% over 29 bets
- 📉 Sharpe Ratio: -0.10
- 🎯 Edge Over Bookies: -3.91%



Backtesting Results Summary

