

Model Training Report: Track 2

Configuration Details

Target Variable

target

Selected Features

is_home_trailing, trailing_team_shots_pre80, trailing_team_fouls_pre80, trailing_team_corners_pre80, trailing_team_yellow_cards_pre80, trailing_team_sending_off_pre80, leading_team_corners_70_75, trailing_team_corners_70_75, leading_team_shots_70_75, trailing_team_shots_70_75, leading_team_corners_75_80, trailing_team_corners_75_80, leading_team_shots_75_80, trailing_team_shots_75_80

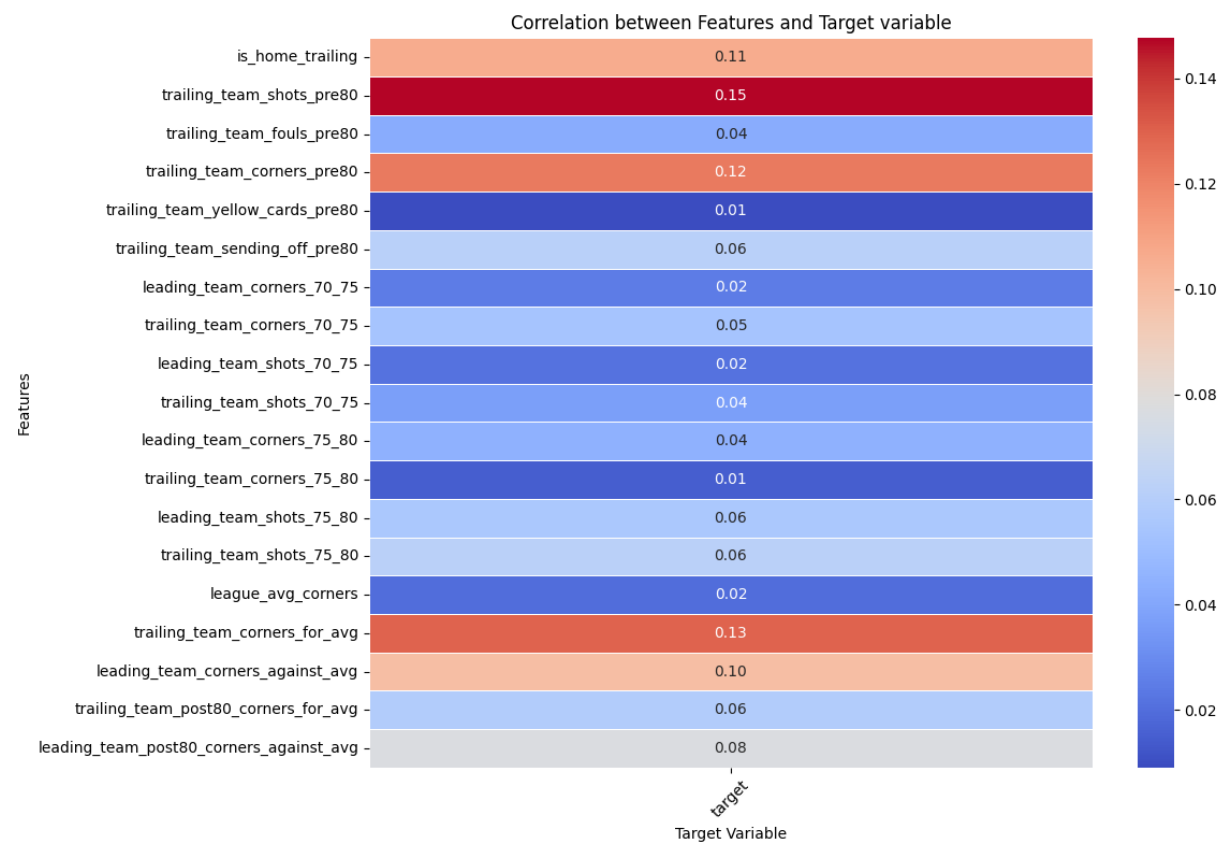
Constructed Features

league_avg_corners, trailing_team_corners_for_avg, leading_team_corners_against_avg, trailing_team_post80_corners_for_avg, leading_team_post80_corners_against_avg

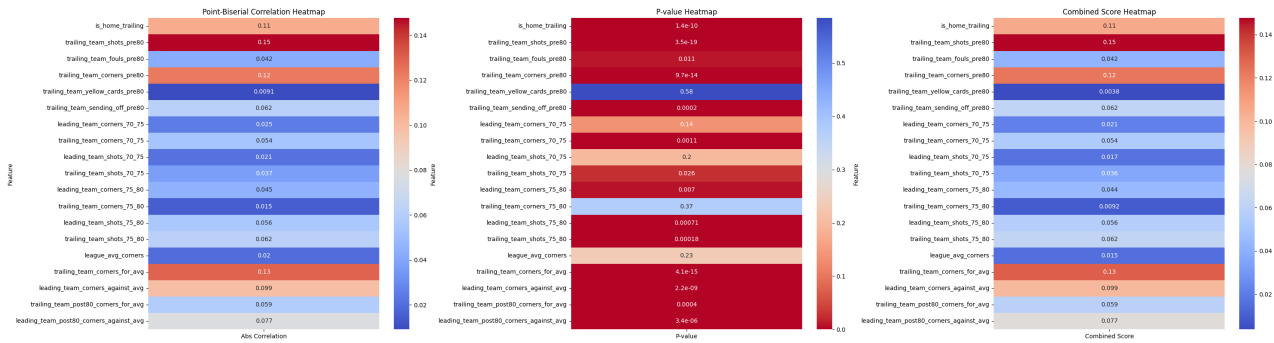
Models Trained (& Hyperparameters)

- **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}

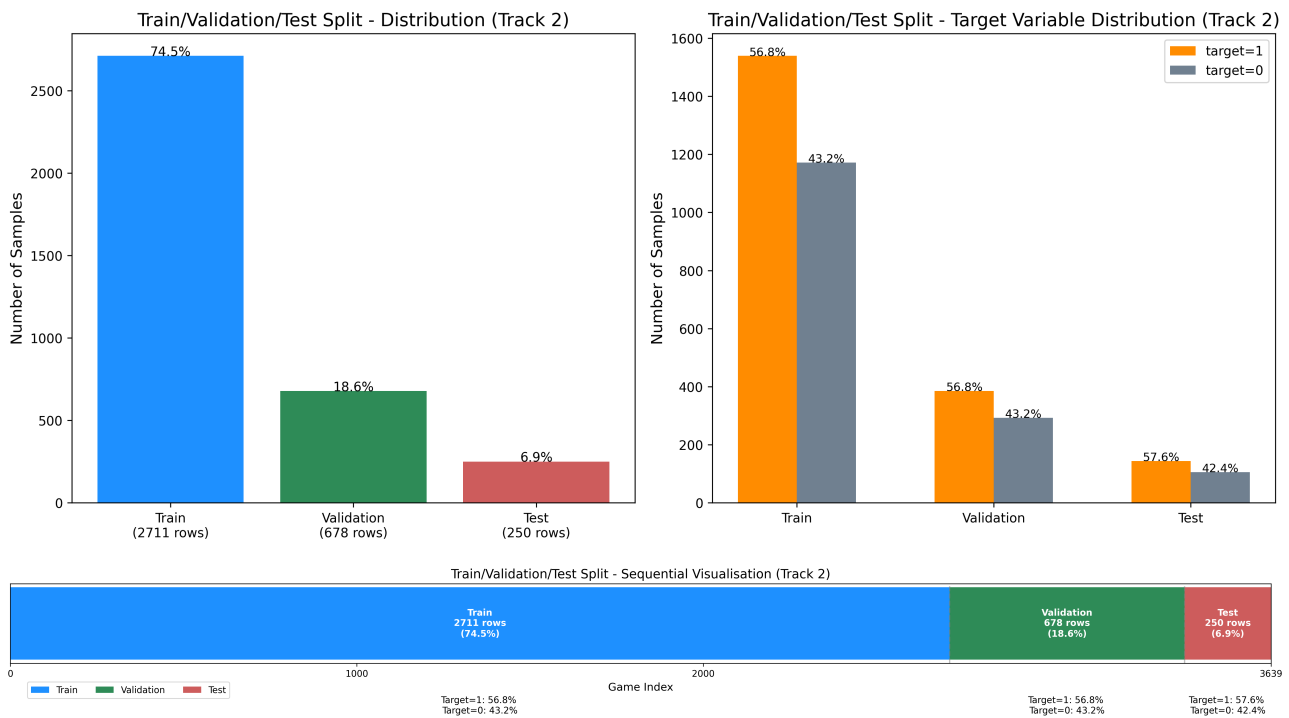
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

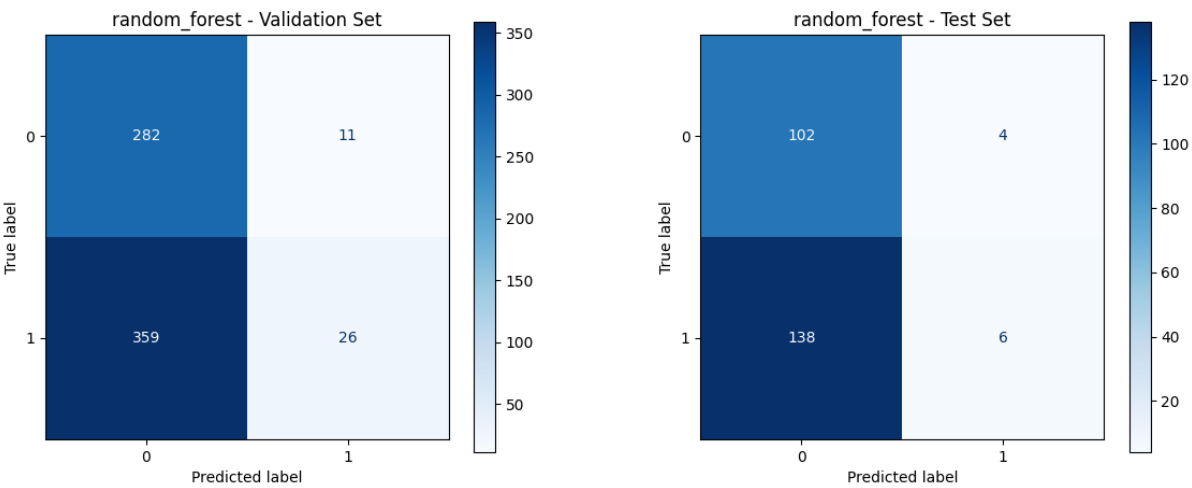
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_corners_for_avg	0.12411	trailing_team_sending_off_pre80	0.01243
leading_team_corners_against_avg	0.11609	is_home_trailing	0.01467
leading_team_post80_corners_against_avg	0.11427	leading_team_corners_75_80	0.01553
trailing_team_post80_corners_for_avg	0.11373	trailing_team_corners_70_75	0.01650
trailing_team_shots_pre80	0.08490	leading_team_corners_70_75	0.01686
trailing_team_fouls_pre80	0.07849		
trailing_team_corners_pre80	0.06789		
trailing_team_yellow_cards_pre80	0.04830		

Optimal Threshold

- **Optimal threshold:** 0.74
- **Expected Precision:** 0.703
- **Expected Recall:** 0.068

Classification Reports



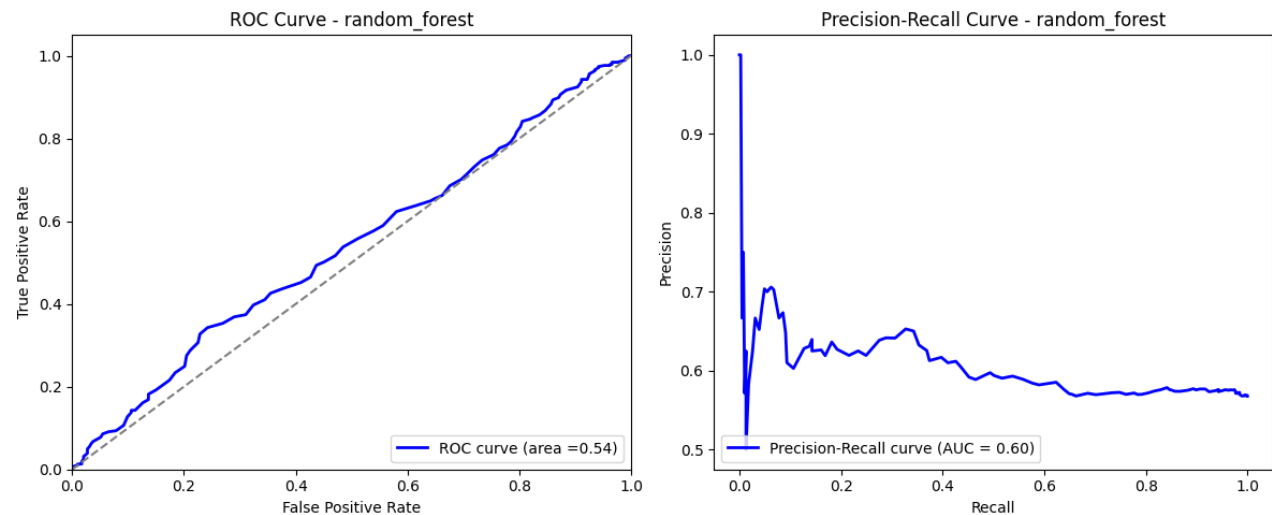
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.44	0.96	0.60	293
1	0.70	0.07	0.12	385
accuracy			0.45	678
macro avg	0.57	0.51	0.36	678
weighted avg	0.59	0.45	0.33	678

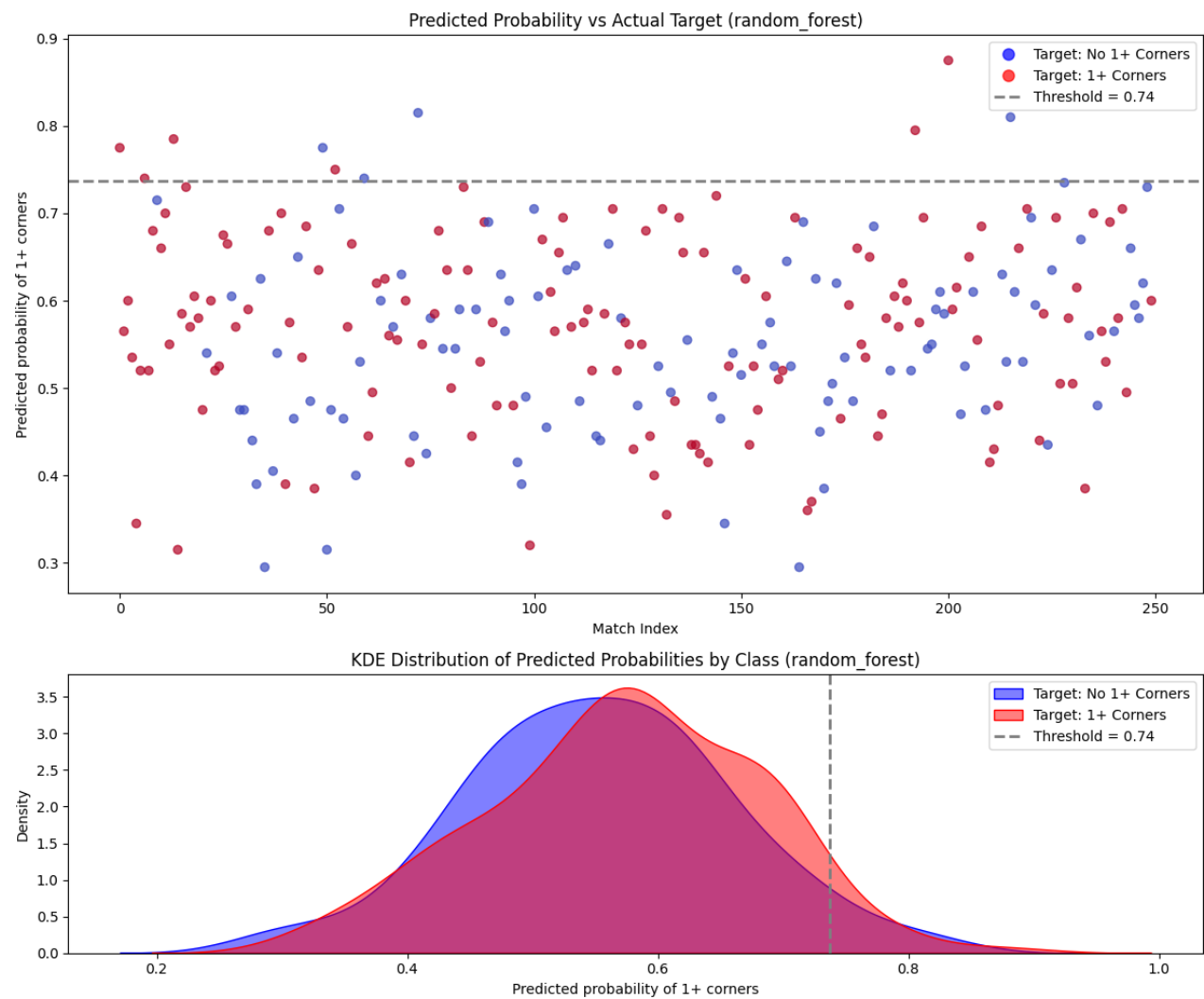
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.42	0.96	0.59	106
1	0.60	0.04	0.08	144
accuracy			0.43	250
macro avg	0.51	0.50	0.33	250
weighted avg	0.53	0.43	0.29	250

ROC and Precision-Recall Curves



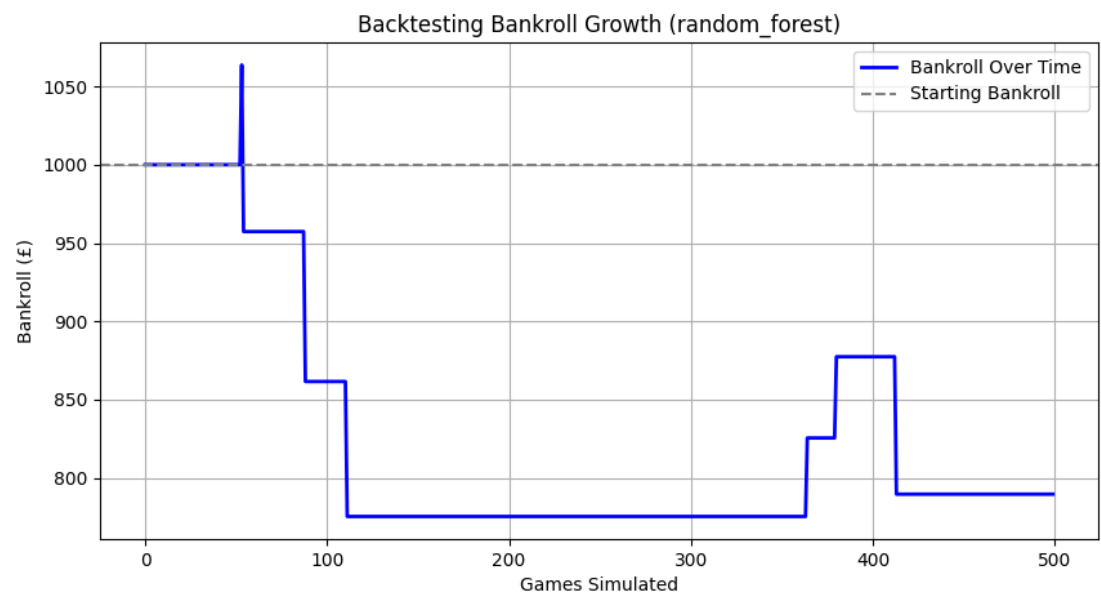
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £789.67
- 📊 Total Staked: £636.09
- 📈 Total Profit: £-210.33
- 📉 ROI: -21.03%
- ✅ Win rate: 42.86% over 7 bets
- 📉 Sharpe Ratio: -0.34
- 🎯 Edge Over Bookies: -17.33%



logistic_regression Model

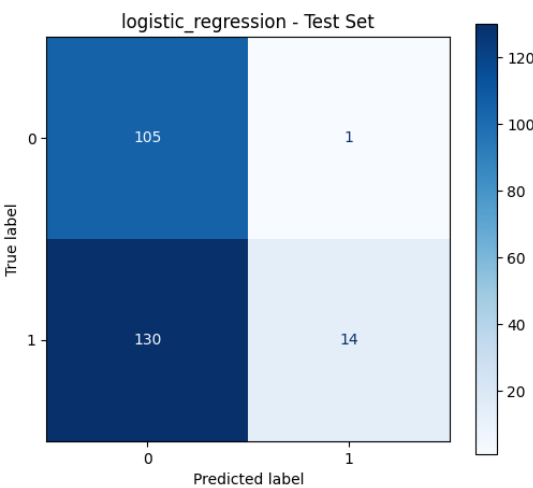
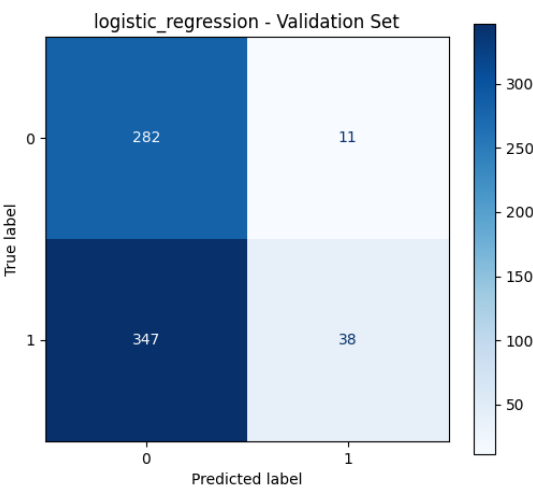
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	1.18654	trailing_team_sending_off_pre80	-0.81453
trailing_team_corners_for_avg	1.11504	leading_team_shots_75_80	-0.65217
leading_team_post80_corners_against_avg	1.00804	trailing_team_fouls_pre80	-0.47780
trailing_team_corners_pre80	0.53272	trailing_team_corners_75_80	-0.34660
trailing_team_yellow_cards_pre80	0.28305	leading_team_corners_75_80	-0.33903
trailing_team_corners_70_75	0.19493		
trailing_team_shots_75_80	0.16530		
leading_team_corners_against_avg	0.13120		

Optimal Threshold

- **Optimal threshold:** 0.67
- **Expected Precision:** 0.776
- **Expected Recall:** 0.099

Classification Reports



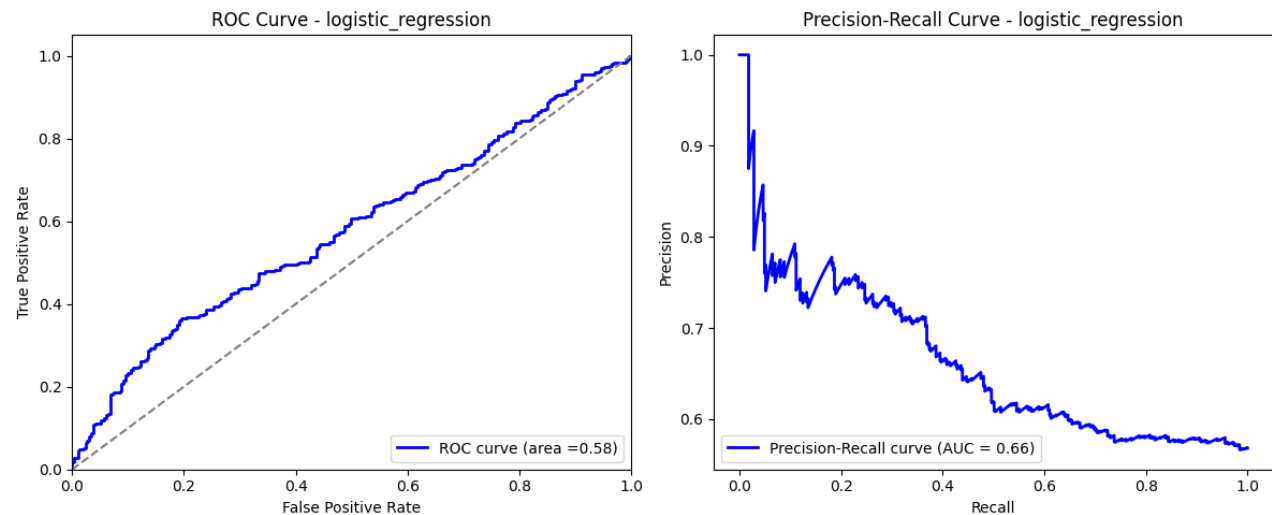
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.45	0.96	0.61	293
1	0.78	0.10	0.18	385
accuracy			0.47	678
macro avg	0.61	0.53	0.39	678
weighted avg	0.63	0.47	0.36	678

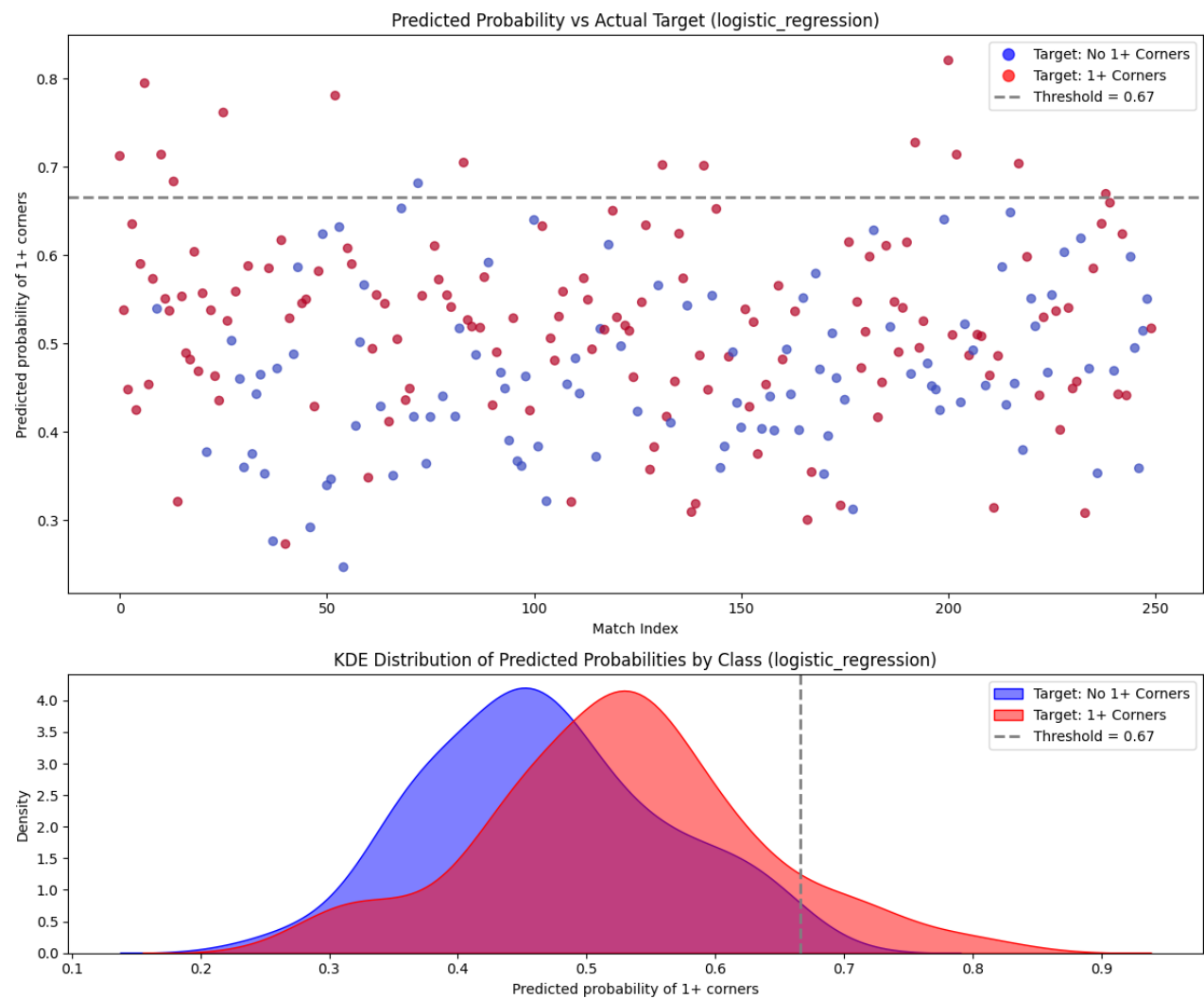
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.45	0.99	0.62	106
1	0.93	0.10	0.18	144
accuracy			0.48	250
macro avg	0.69	0.54	0.40	250
weighted avg	0.73	0.48	0.36	250

ROC and Precision-Recall Curves



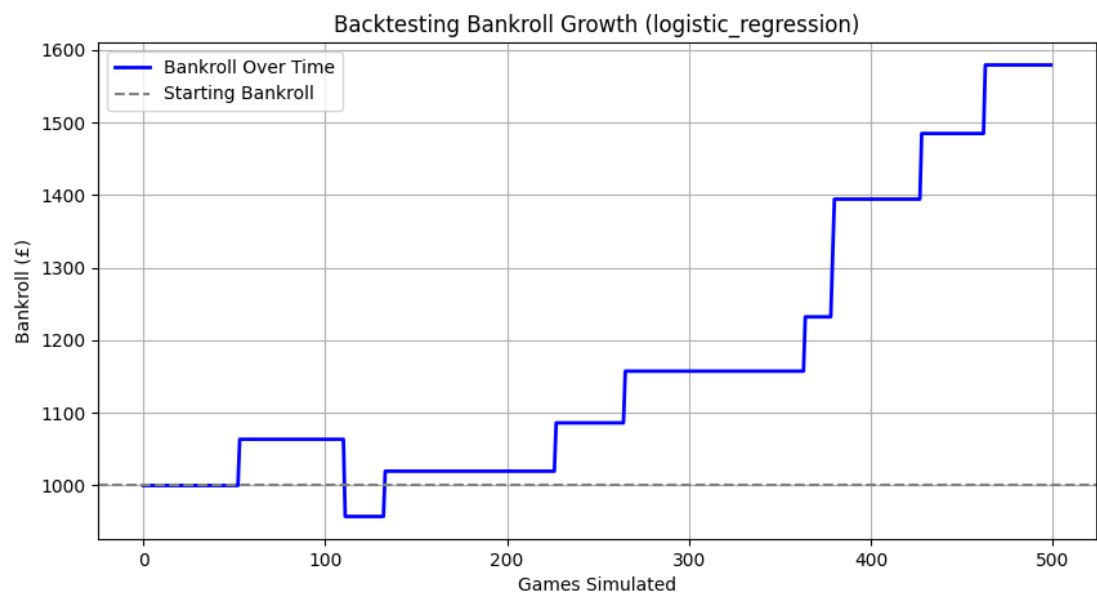
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £1579.49
- 📁 Total Staked: £1170.89
- 📈 Total Profit: £579.49
- 📊 ROI: 57.95%
- ✅ Win rate: 90.00% over 10 bets
- 📉 Sharpe Ratio: 0.92
- 🎯 Edge Over Bookies: 29.81%



svc Model

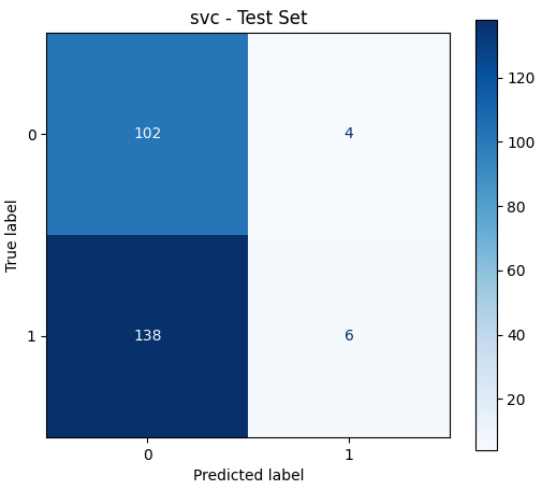
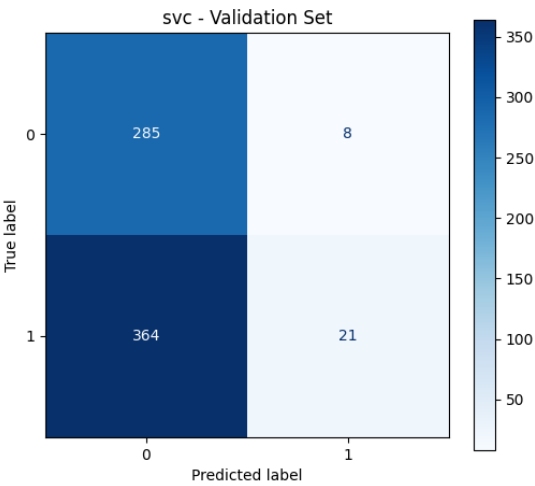
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
is_home_trailing	0.00000	leading_team_post80_corners_against_avg	0.00000
leading_team_corners_75_80	0.00000	trailing_team_fouls_pre80	0.00000
trailing_team_post80_corners_for_avg	0.00000	trailing_team_corners_pre80	0.00000
leading_team_corners_against_avg	0.00000	trailing_team_yellow_cards_pre80	0.00000
trailing_team_corners_for_avg	0.00000	trailing_team_sending_off_pre80	0.00000
league_avg_corners	0.00000		
trailing_team_shots_75_80	0.00000		
leading_team_shots_75_80	0.00000		

Optimal Threshold

- **Optimal threshold:** 0.71
- **Expected Precision:** 0.724
- **Expected Recall:** 0.055

Classification Reports



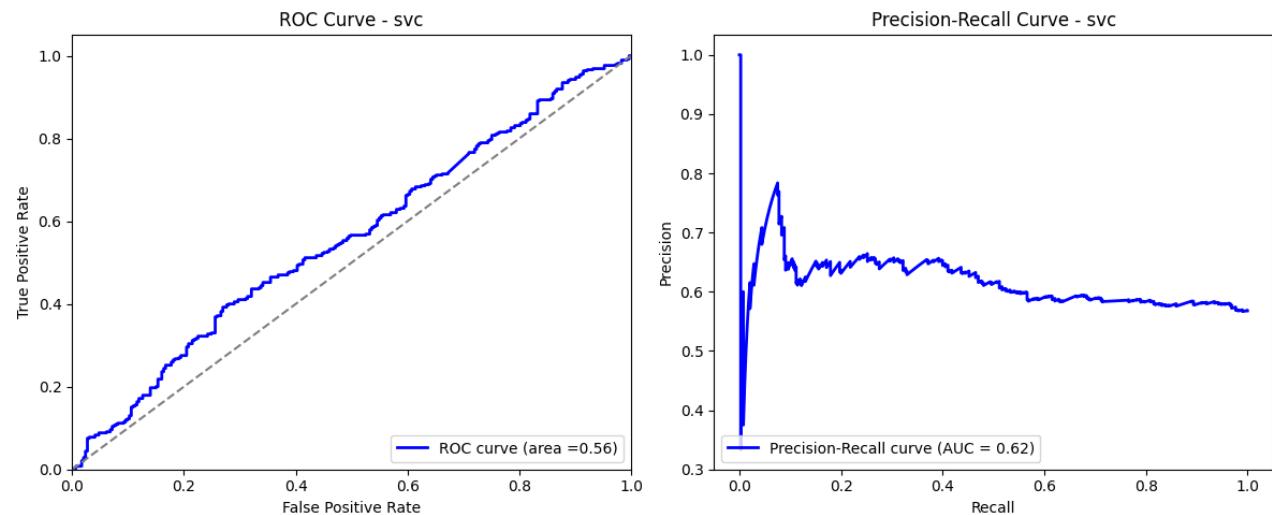
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.44	0.97	0.61	293
1	0.72	0.05	0.10	385
accuracy			0.45	678
macro avg	0.58	0.51	0.35	678
weighted avg	0.60	0.45	0.32	678

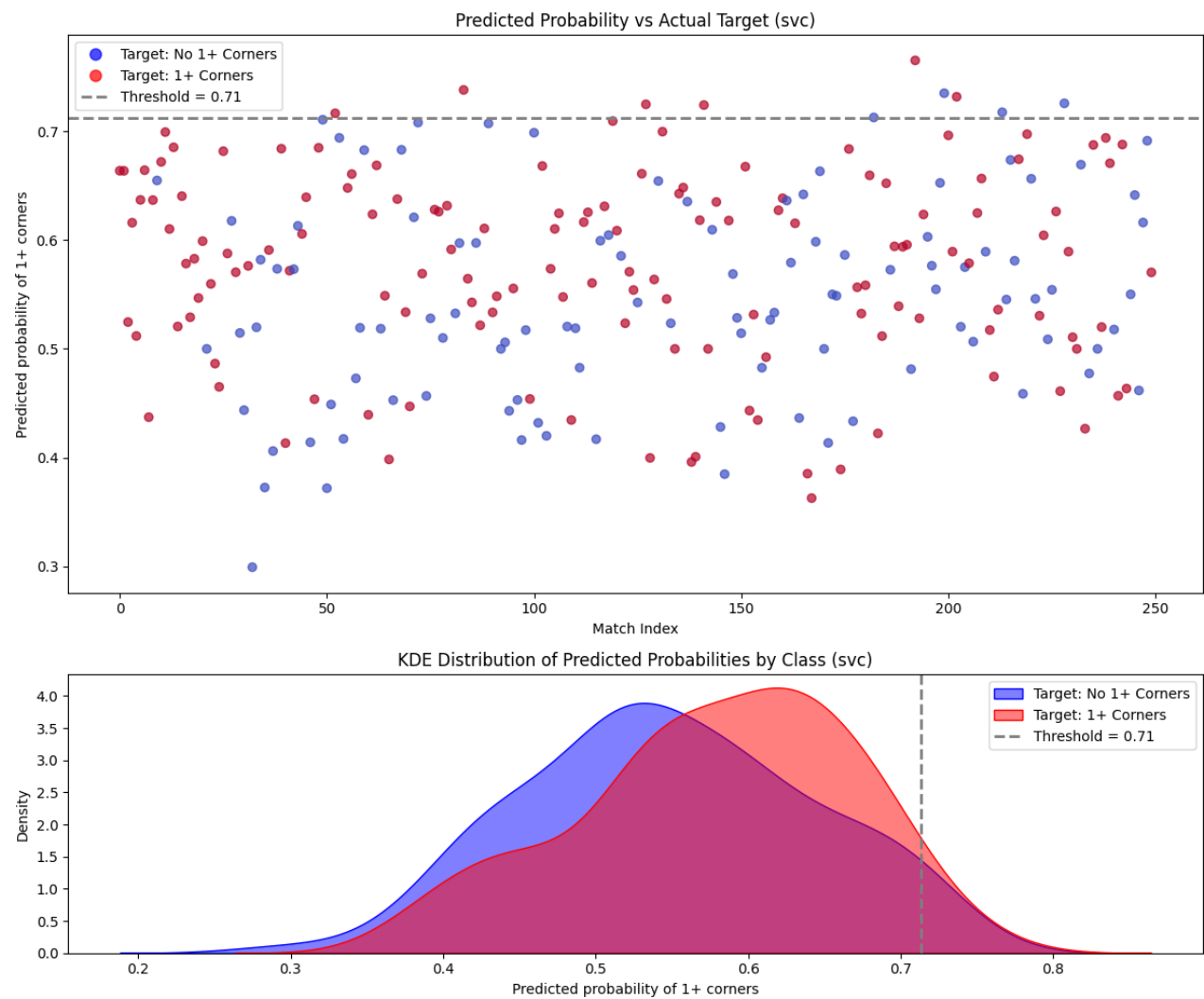
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.42	0.96	0.59	106
1	0.60	0.04	0.08	144
accuracy			0.43	250
macro avg	0.51	0.50	0.33	250
weighted avg	0.53	0.43	0.29	250

ROC and Precision-Recall Curves



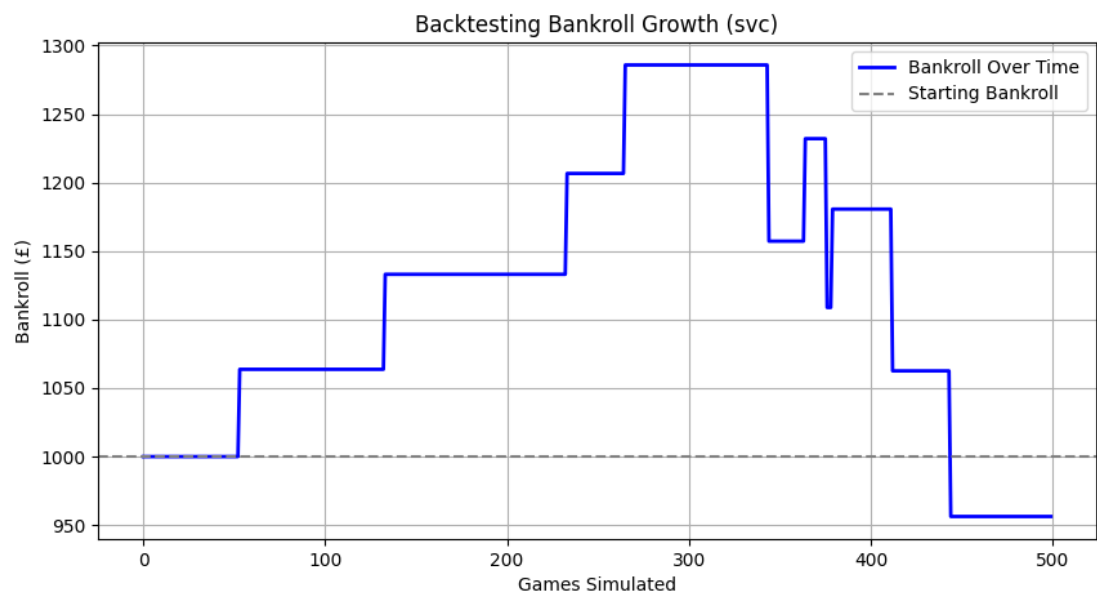
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £956.30
- 📁 Total Staked: £1143.02
- 📈 Total Profit: £-43.70
- 🇮🇹 ROI: -4.37%
- ✅ Win rate: 60.00% over 10 bets
- 📉 Sharpe Ratio: -0.01
- 🎯 Edge Over Bookies: -0.19%



xgboost Model

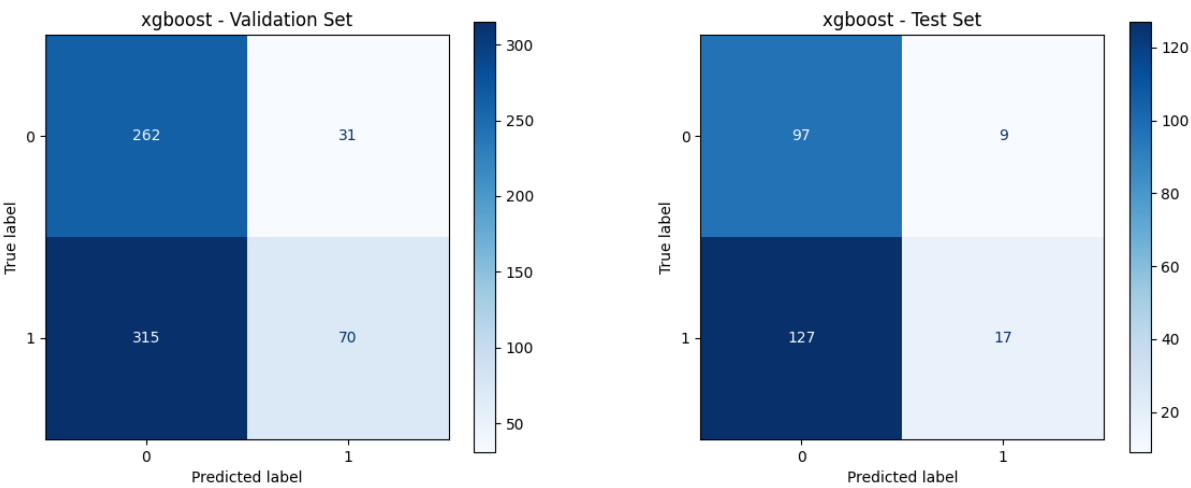
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	0.10686	trailing_team_corners_70_75	0.03024
trailing_team_corners_pre80	0.06324	trailing_team_corners_75_80	0.03764
trailing_team_corners_for_avg	0.06255	leading_team_corners_70_75	0.03989
leading_team_post80_corners_against_avg	0.05814	leading_team_corners_75_80	0.04050
leading_team_shots_75_80	0.05800	trailing_team_yellow_cards_pre80	0.04229
trailing_team_sending_off_pre80	0.05795		
trailing_team_shots_70_75	0.05478		
is_home_trailing	0.05421		

Optimal Threshold

- **Optimal threshold:** 0.69
- **Expected Precision:** 0.690
- **Expected Recall:** 0.179

Classification Reports



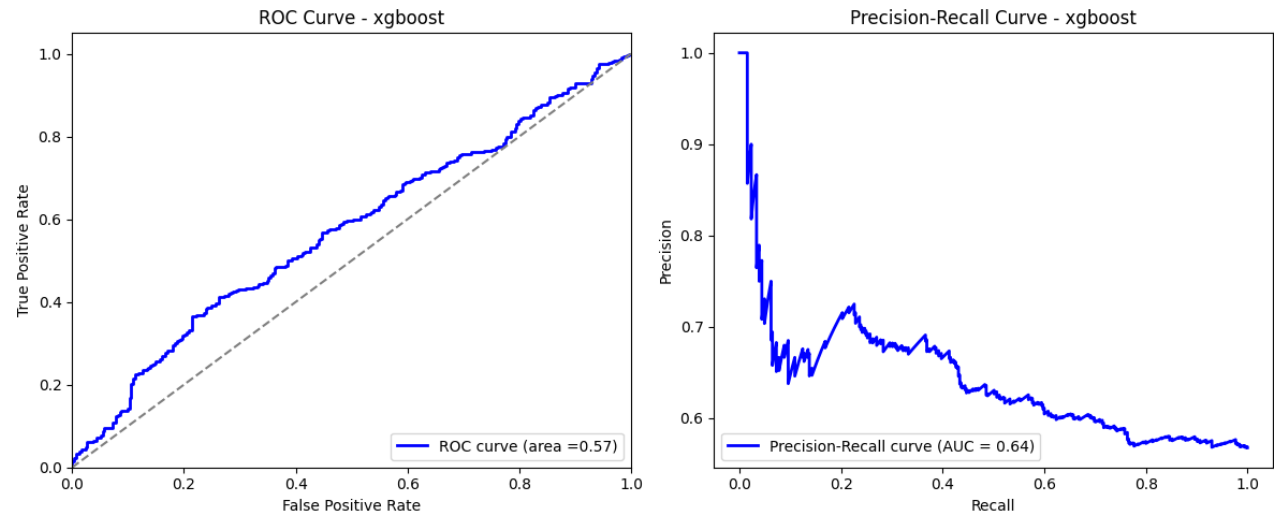
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.45	0.89	0.60	293
1	0.69	0.18	0.29	385
accuracy			0.49	678
macro avg	0.57	0.54	0.45	678
weighted avg	0.59	0.49	0.42	678

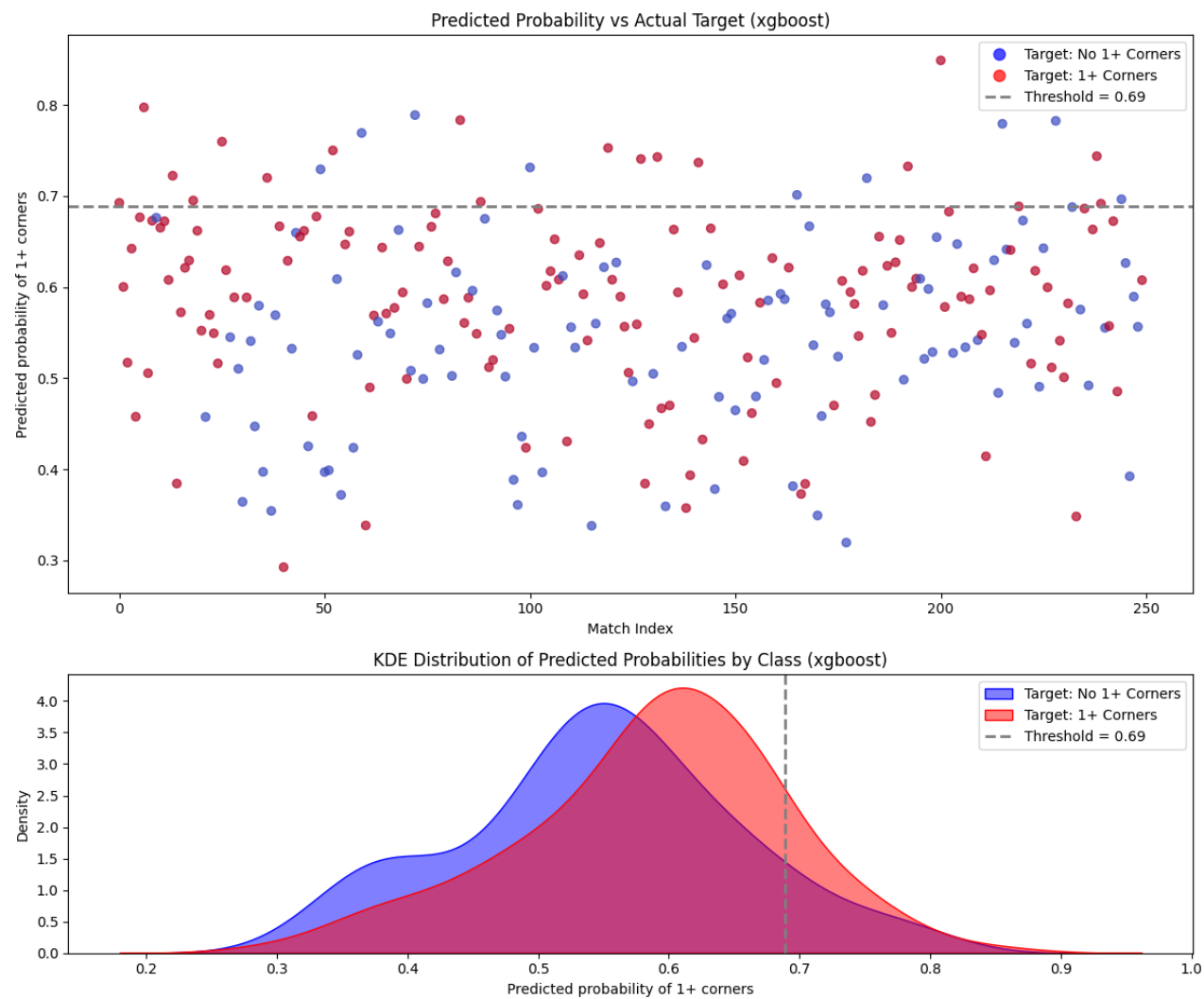
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.43	0.92	0.59	106
1	0.65	0.12	0.20	144
accuracy			0.46	250
macro avg	0.54	0.52	0.39	250
weighted avg	0.56	0.46	0.36	250

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £824.69

🎲 Total Staked: £1990.93

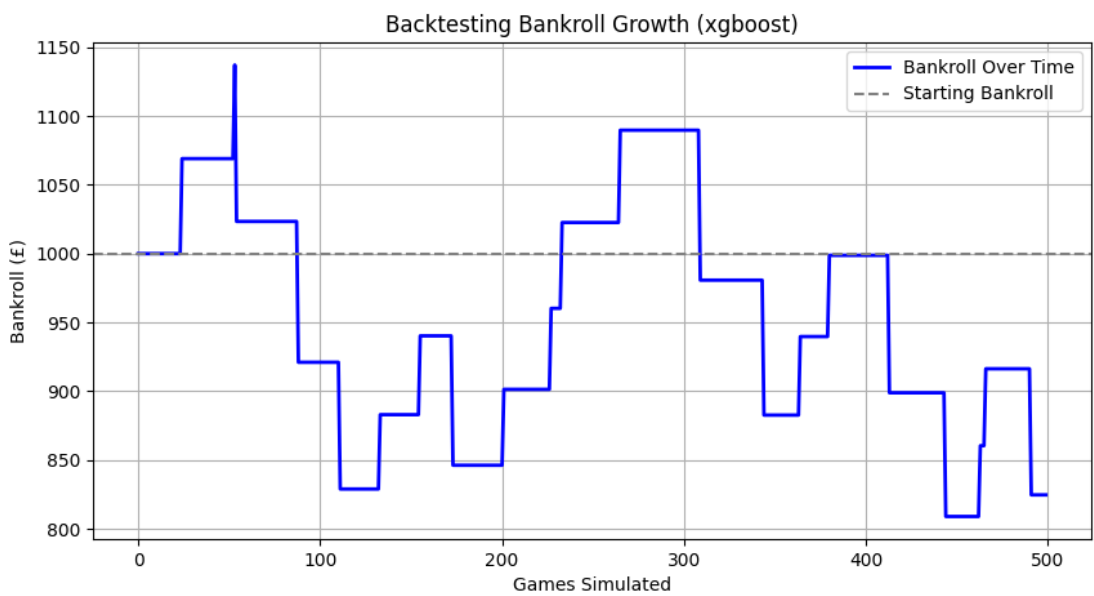
📈 Total Profit: £-175.31

📊 ROI: -17.53%

✅ Win rate: 57.14% over 21 bets

📉 Sharpe Ratio: -0.07

🎯 Edge Over Bookies: -3.05%



Backtesting Results Summary

