

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

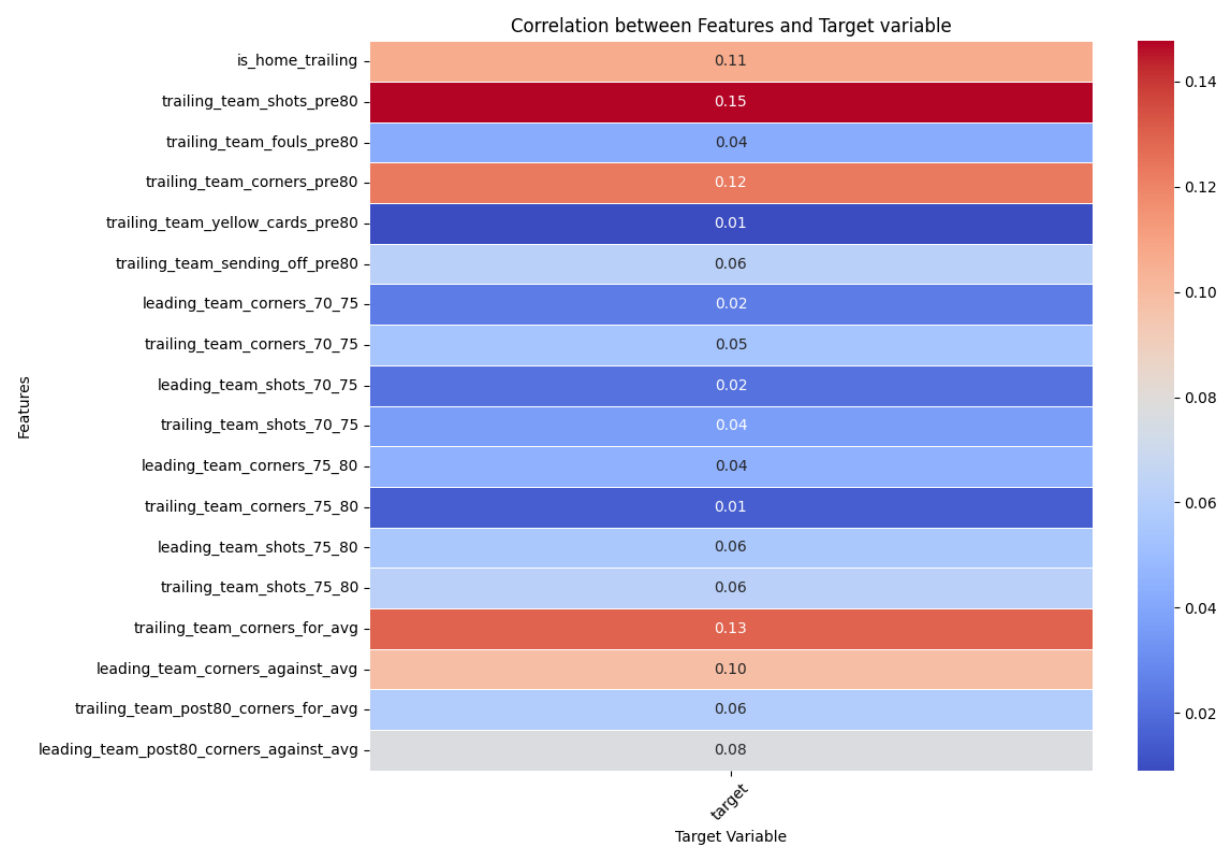
Constructed Features

trailing_team_corners_for_avg, leading_team_corners_against_avg,
trailing_team_post80_corners_for_avg, leading_team_post80_corners_against_avg

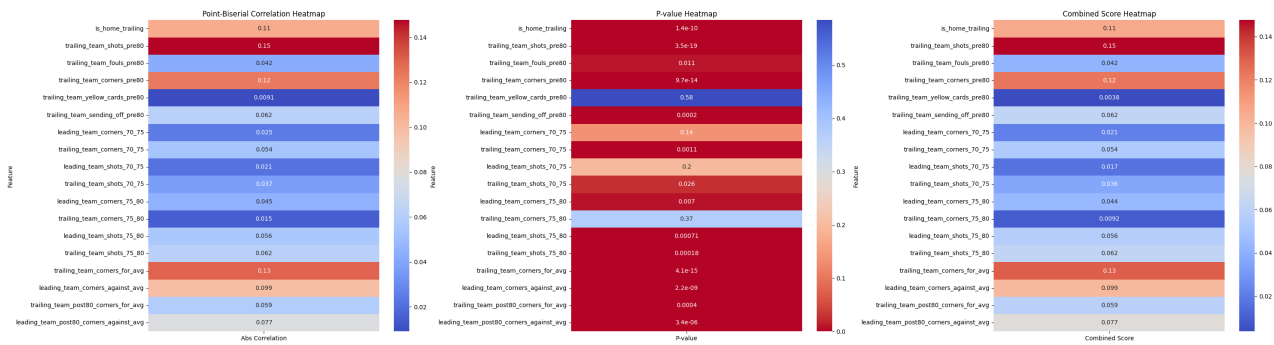
Models Trained (& Hyperparameters)

- **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}

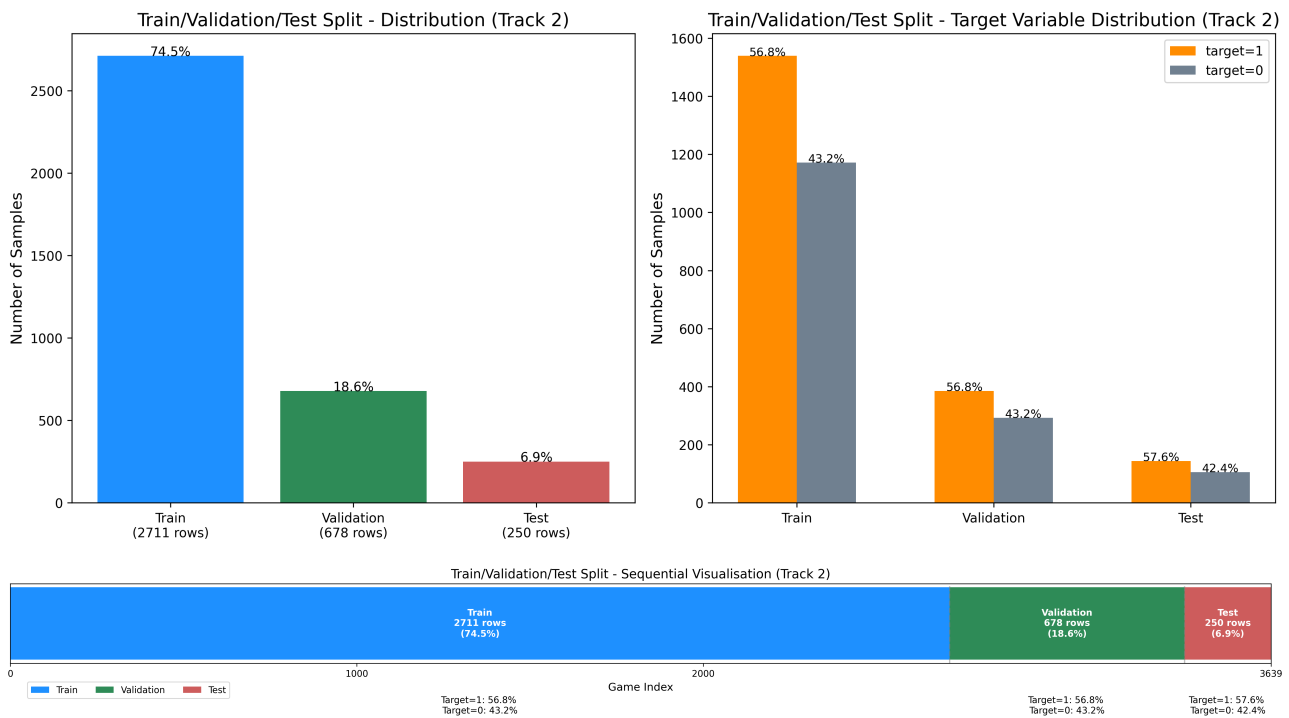
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

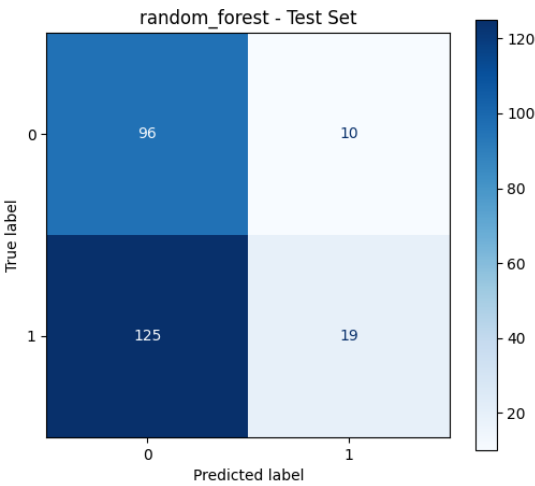
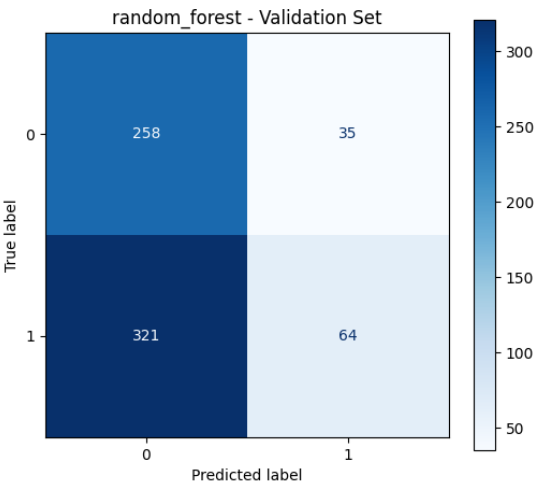
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_corners_for_avg	0.13174	trailing_team_sending_off_pre80	0.01312
leading_team_post80_corners_against_avg	0.12378	is_home_trailing	0.01522
leading_team_corners_against_avg	0.12369	leading_team_corners_75_80	0.01561
trailing_team_post80_corners_for_avg	0.11599	leading_team_corners_70_75	0.01653
trailing_team_shots_pre80	0.08782	trailing_team_corners_70_75	0.01724
trailing_team_fouls_pre80	0.08280		
trailing_team_corners_pre80	0.06993		
trailing_team_yellow_cards_pre80	0.04984		

Optimal Threshold

- **Optimal threshold:** 0.69
- **Expected Precision:** 0.646
- **Expected Recall:** 0.166

Classification Reports



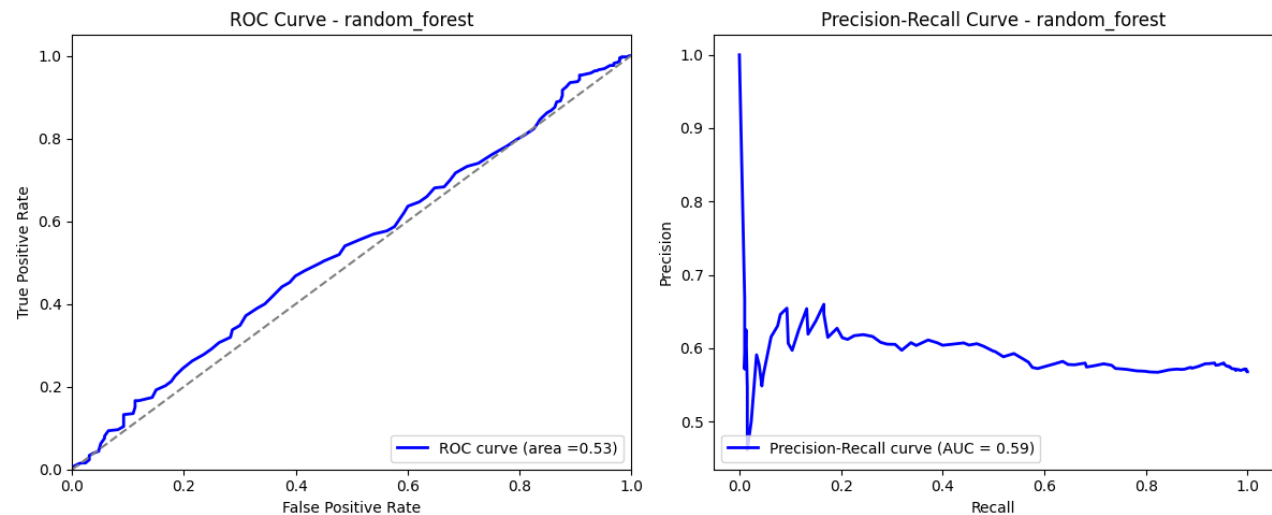
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.45	0.88	0.59	293
1	0.65	0.17	0.26	385
accuracy			0.47	678
macro avg	0.55	0.52	0.43	678
weighted avg	0.56	0.47	0.41	678

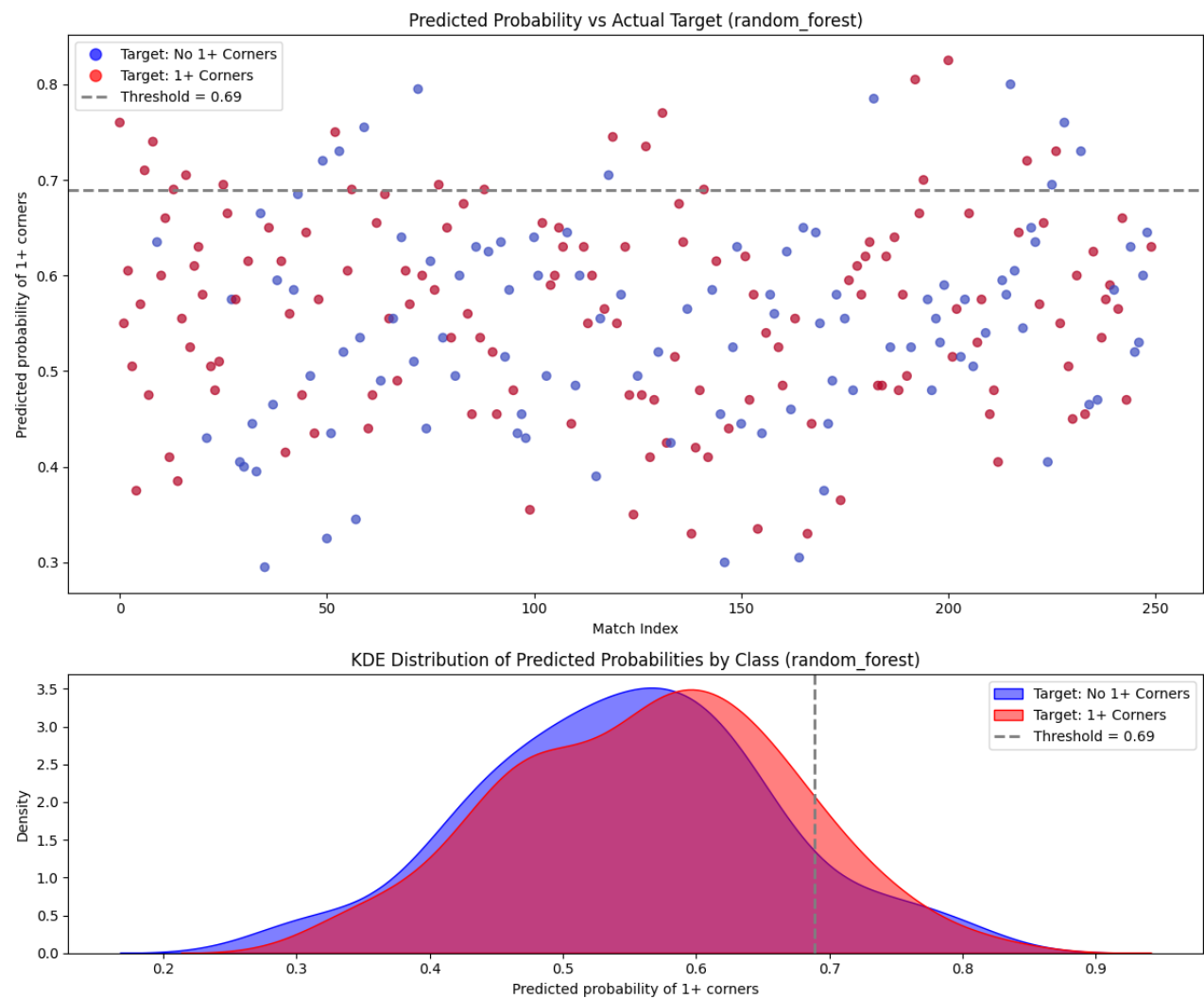
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.43	0.91	0.59	106
1	0.66	0.13	0.22	144
accuracy			0.46	250
macro avg	0.54	0.52	0.40	250
weighted avg	0.56	0.46	0.38	250

ROC and Precision-Recall Curves



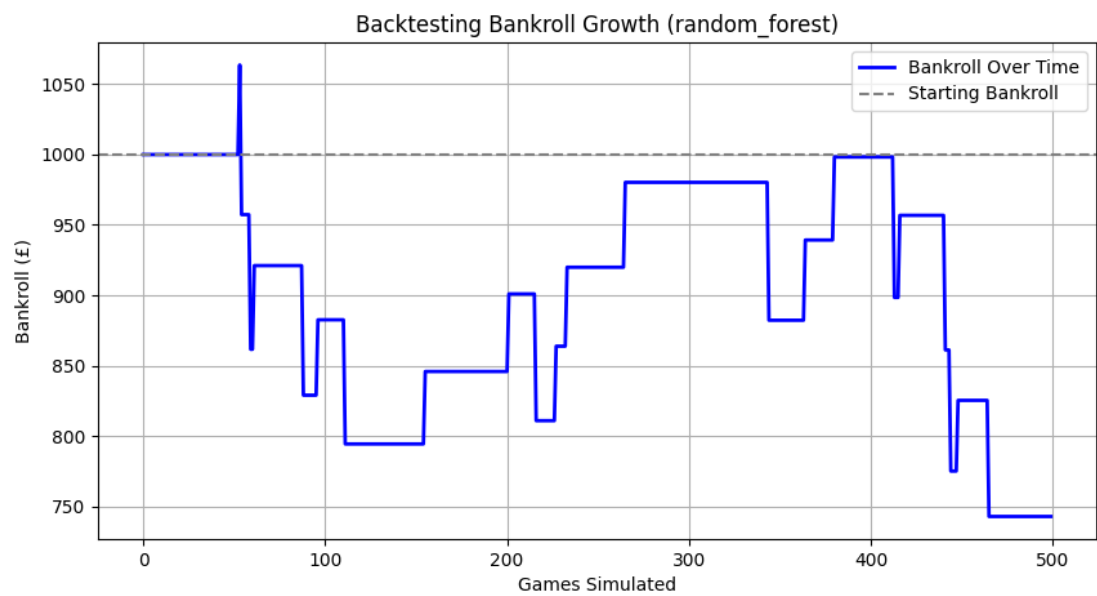
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £742.75
- 📊 Total Staked: £1976.74
- 📈 Total Profit: £-257.25
- 🇮🇹 ROI: -25.73%
- ✅ Win rate: 54.55% over 22 bets
- 📉 Sharpe Ratio: -0.12
- 🎯 Edge Over Bookies: -5.64%



logistic_regression Model

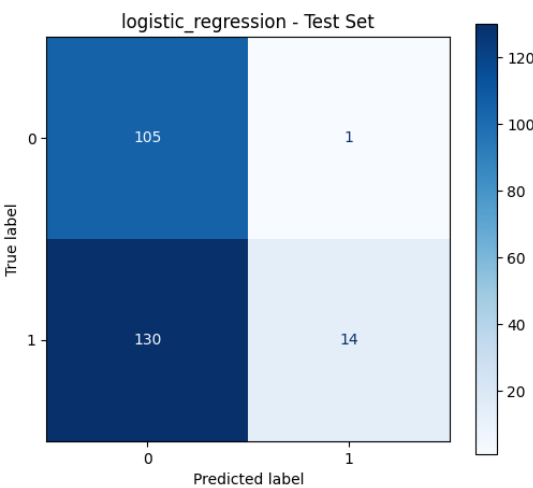
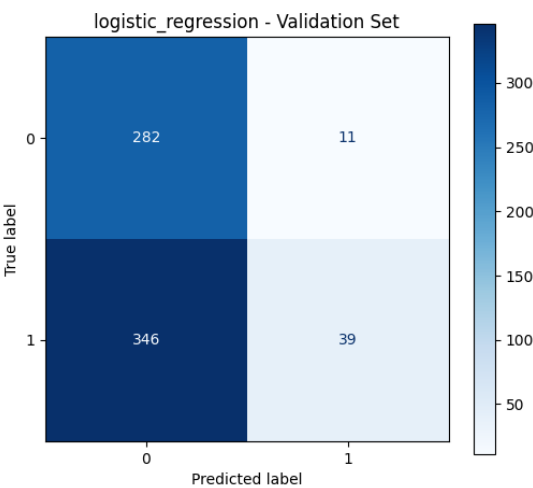
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	1.18748	trailing_team_sending_off_pre80	-0.81368
trailing_team_corners_for_avg	1.10856	leading_team_shots_75_80	-0.65342
leading_team_post80_corners_against_avg	1.00762	trailing_team_fouls_pre80	-0.47166
trailing_team_corners_pre80	0.53239	trailing_team_corners_75_80	-0.34591
trailing_team_yellow_cards_pre80	0.27687	leading_team_corners_75_80	-0.33760
trailing_team_corners_70_75	0.19445		
trailing_team_shots_75_80	0.16543		
leading_team_corners_against_avg	0.12351		

Optimal Threshold

- **Optimal threshold:** 0.67
- **Expected Precision:** 0.784
- **Expected Recall:** 0.104

Classification Reports



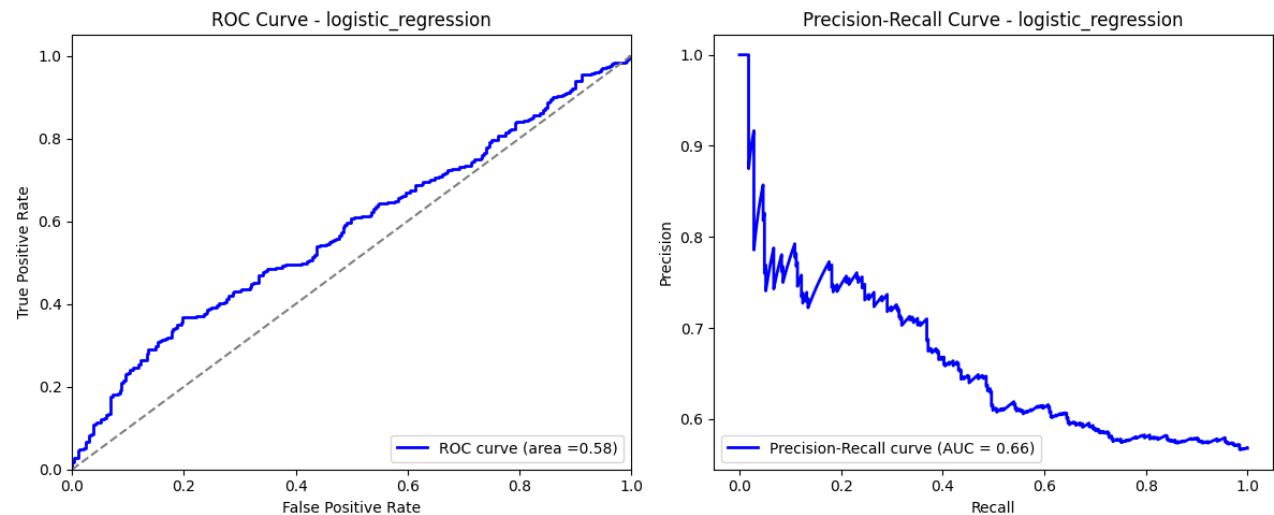
Classification Report (Validation Set)

	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	0.40	678
weighted avg	0.64	0.47	0.37	678

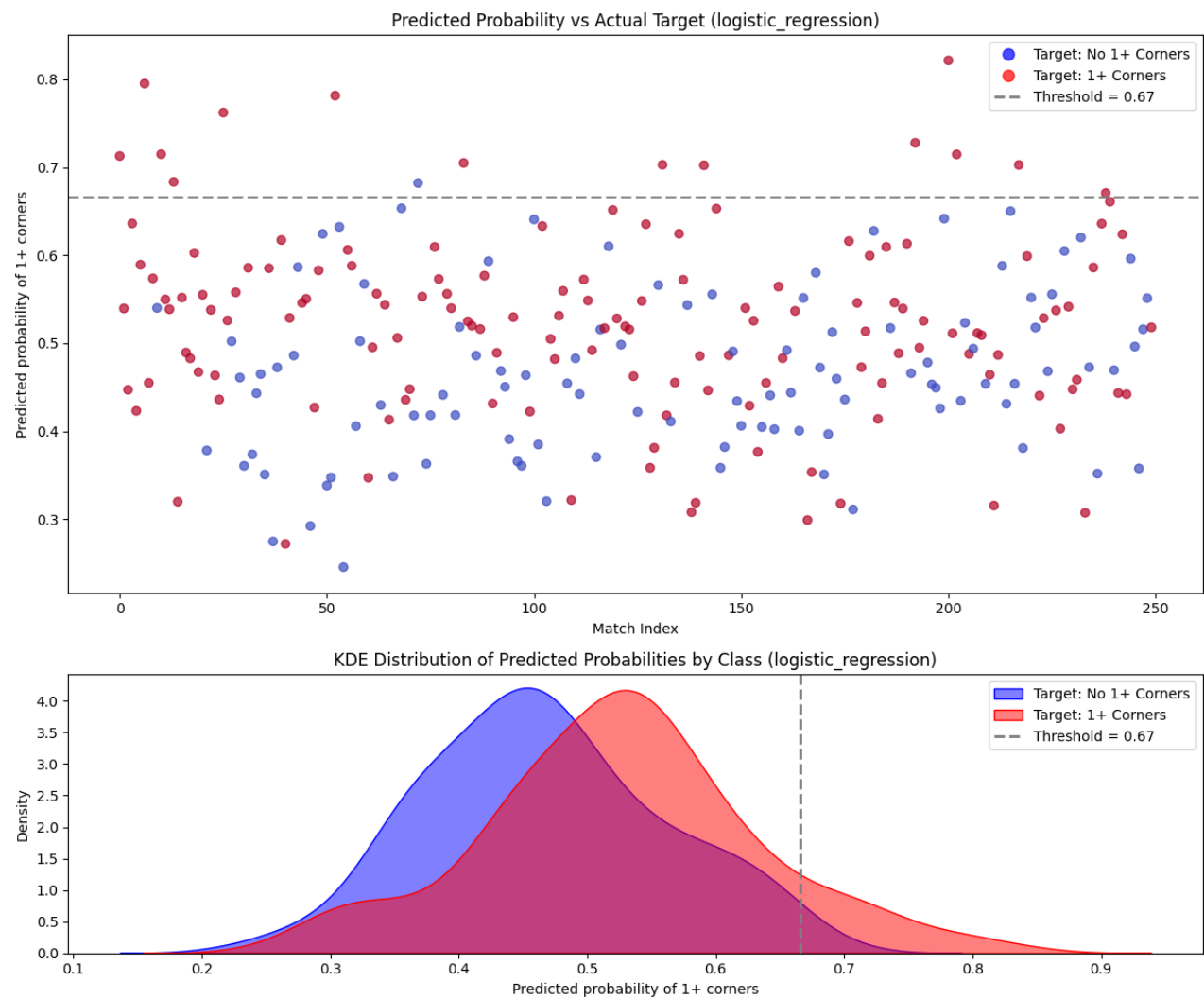
Classification Report (Test Set)

	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



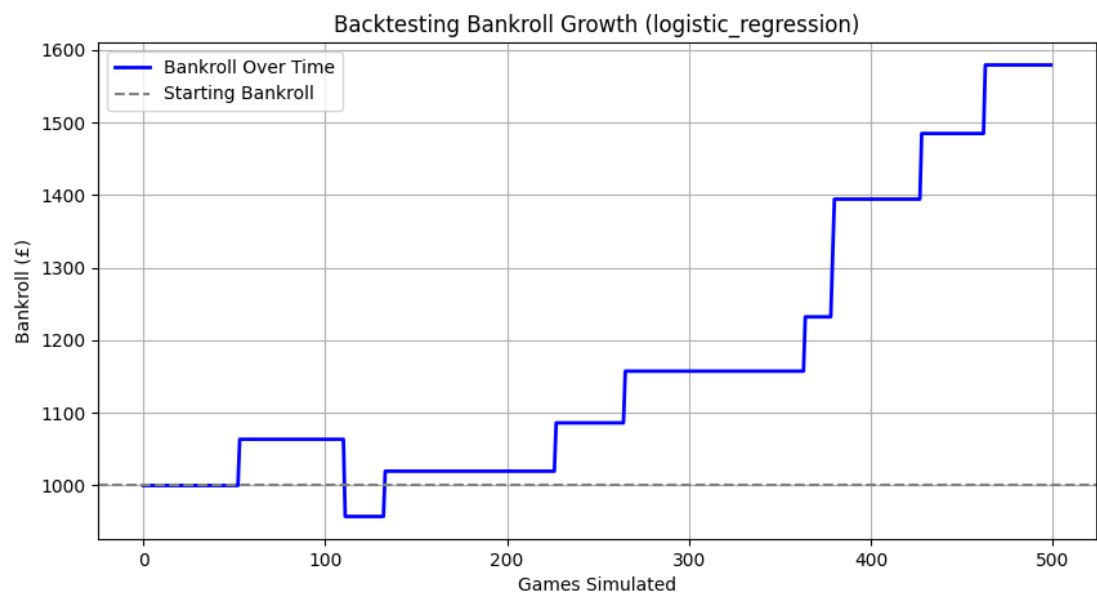
Prediction Distribution Graphs



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

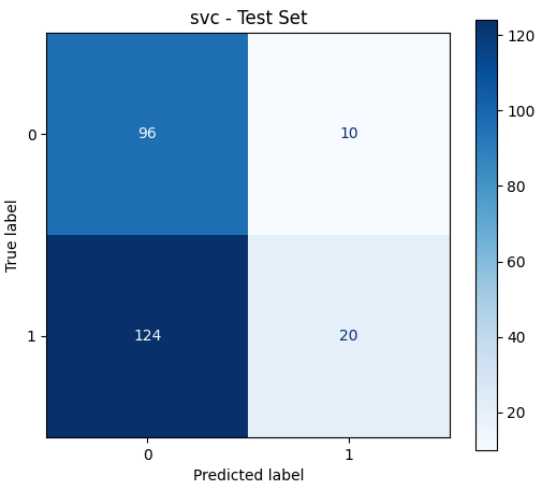
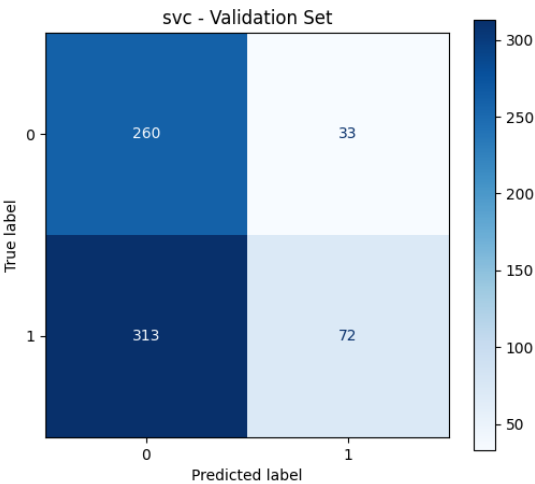
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
is_home_trailing	0.00000	leading_team_post80_corners_against_avg	0.00000
trailing_team_shots_pre80	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
trailing_team_shots_75_80	0.00000		
leading_team_shots_75_80	0.00000		
trailing_team_corners_75_80	0.00000		

Optimal Threshold

- **Optimal threshold:** 0.67
- **Expected Precision:** 0.689
- **Expected Recall:** 0.190

Classification Reports



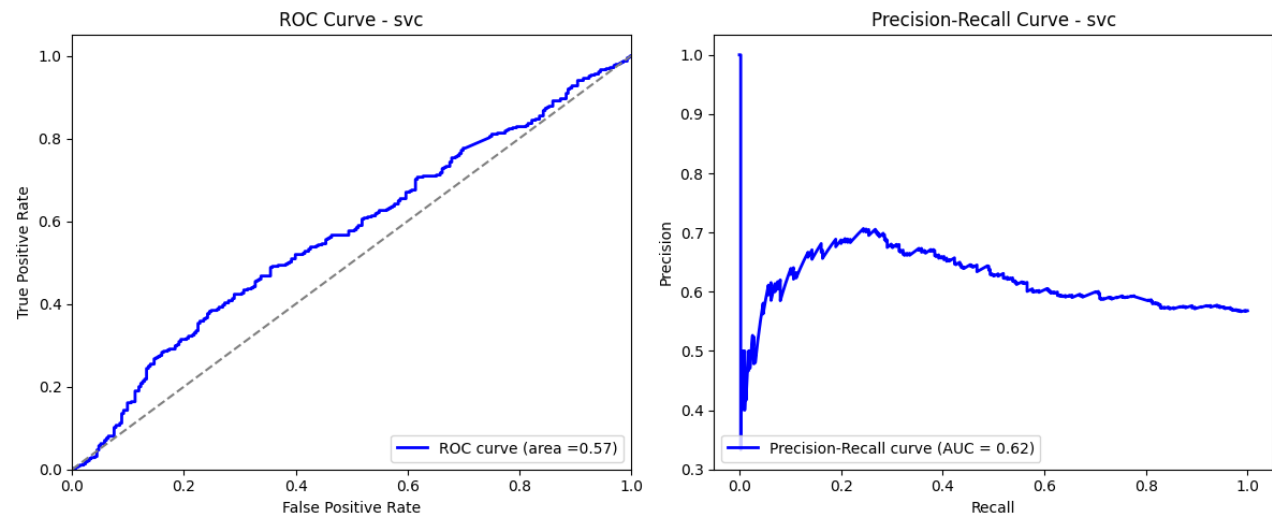
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.45	0.89	0.60	293
1	0.69	0.19	0.29	385
accuracy			0.49	678
macro avg	0.57	0.54	0.45	678
weighted avg	0.59	0.49	0.43	678

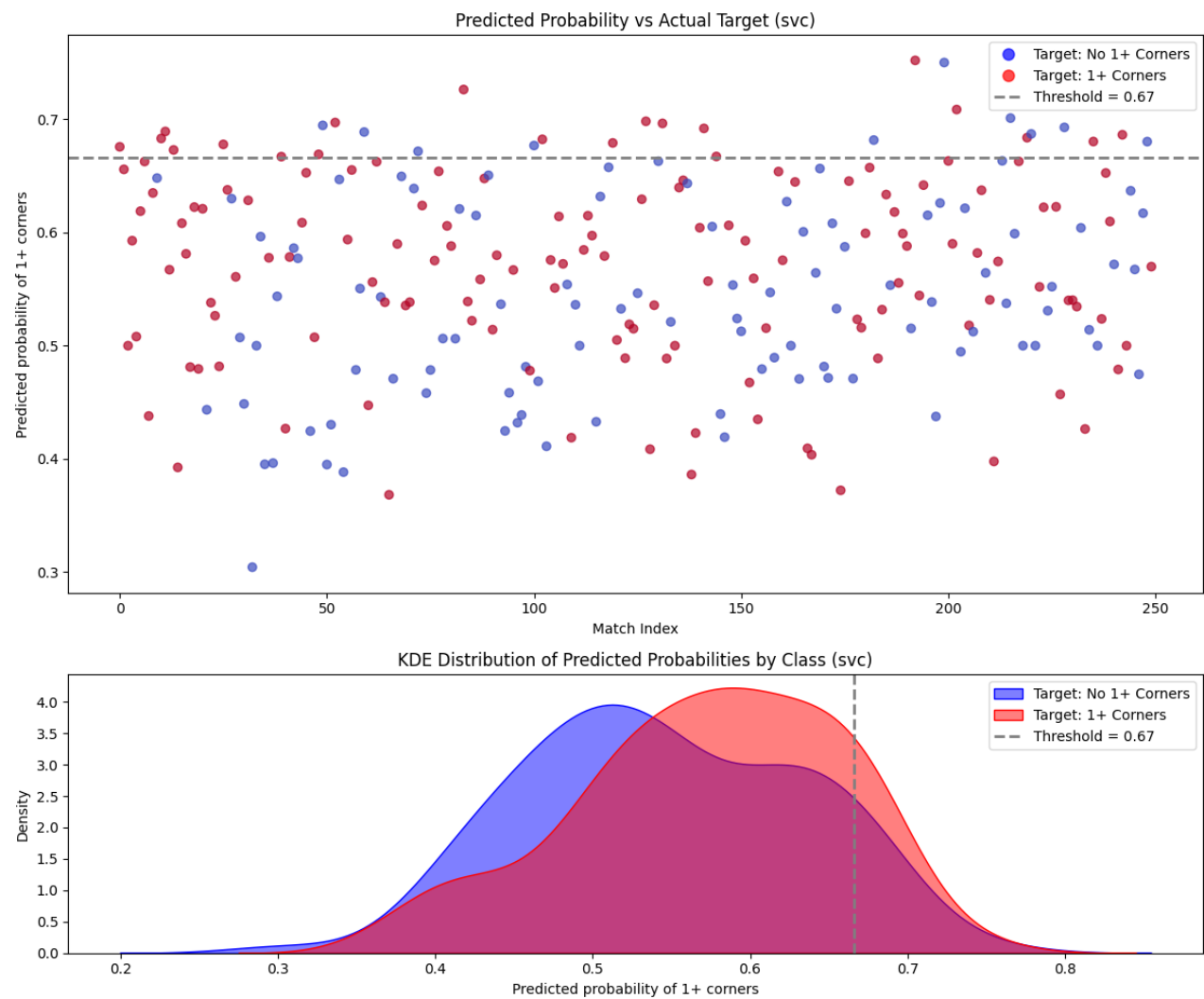
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.44	0.91	0.59	106
1	0.67	0.14	0.23	144
accuracy			0.46	250
macro avg	0.55	0.52	0.41	250
weighted avg	0.57	0.46	0.38	250

ROC and Precision-Recall Curves



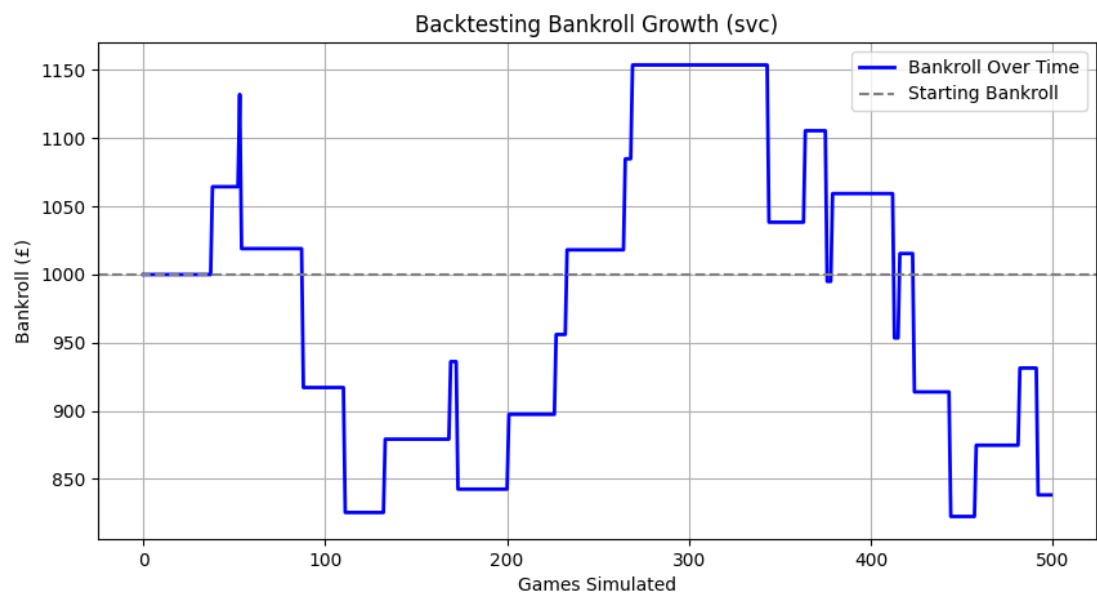
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £838.26
- 📁 Total Staked: £2343.59
- 📈 Total Profit: £-161.74
- 📊 ROI: -16.17%
- ✅ Win rate: 58.33% over 24 bets
- 📉 Sharpe Ratio: -0.05
- 🎯 Edge Over Bookies: -1.86%



xgboost Model

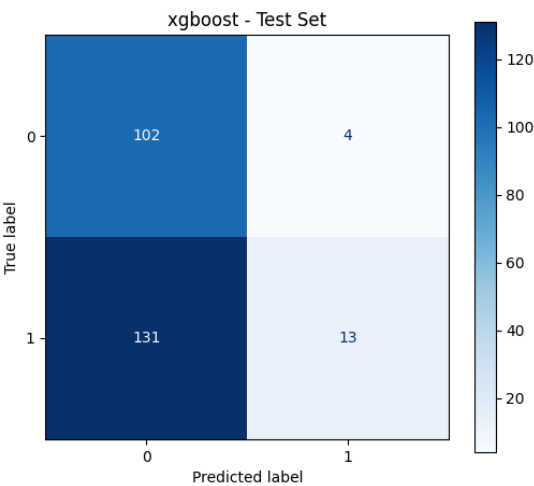
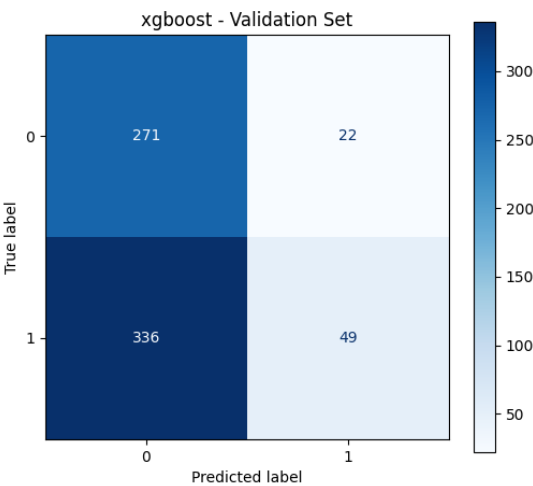
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	0.10944	trailing_team_corners_70_75	0.02534
trailing_team_corners_pre80	0.07278	leading_team_corners_70_75	0.04158
trailing_team_corners_for_avg	0.06804	trailing_team_corners_75_80	0.04182
trailing_team_sending_off_pre80	0.06514	is_home_trailing	0.04347
leading_team_shots_75_80	0.06189	leading_team_corners_75_80	0.04471
leading_team_post80_corners_against_avg	0.05874		
trailing_team_shots_75_80	0.05831		
trailing_team_shots_70_75	0.05586		

Optimal Threshold

- **Optimal threshold:** 0.71
- **Expected Precision:** 0.690
- **Expected Recall:** 0.127

Classification Reports



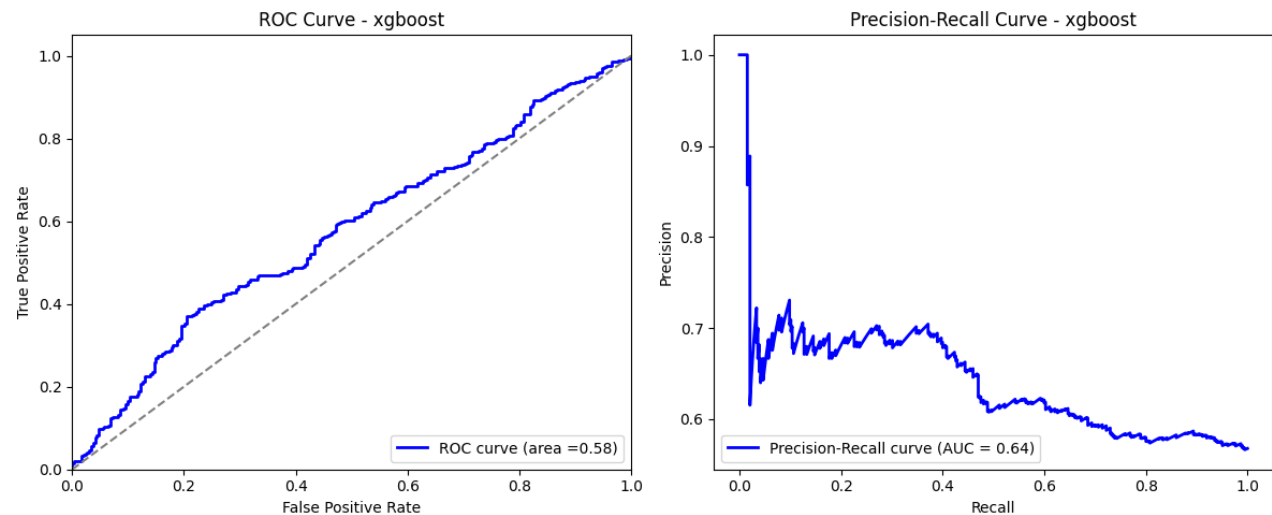
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.45	0.92	0.60	293
1	0.69	0.13	0.21	385
accuracy			0.47	678
macro avg	0.57	0.53	0.41	678
weighted avg	0.58	0.47	0.38	678

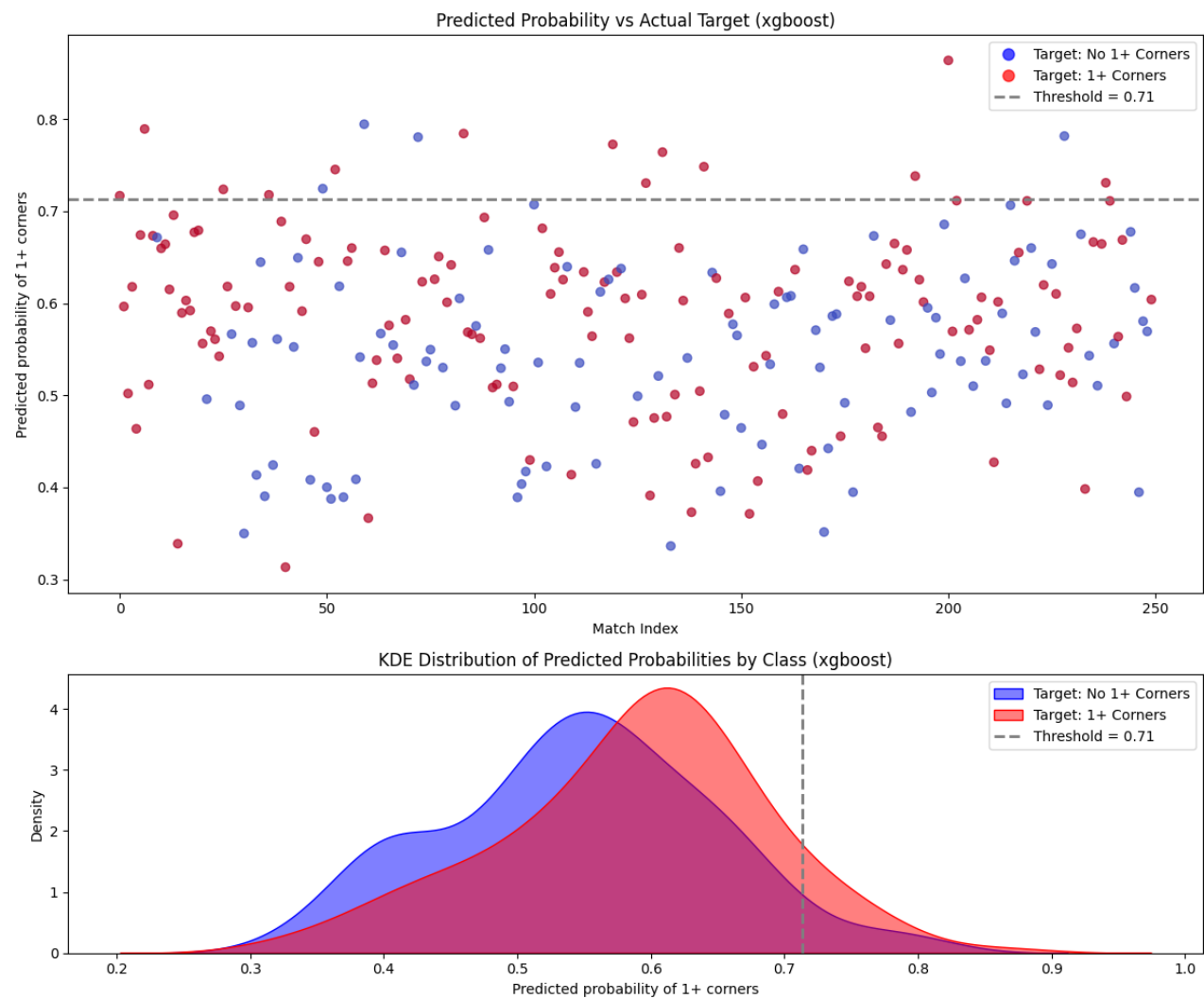
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.44	0.96	0.60	106
1	0.76	0.09	0.16	144
accuracy			0.46	250
macro avg	0.60	0.53	0.38	250
weighted avg	0.63	0.46	0.35	250

ROC and Precision-Recall Curves



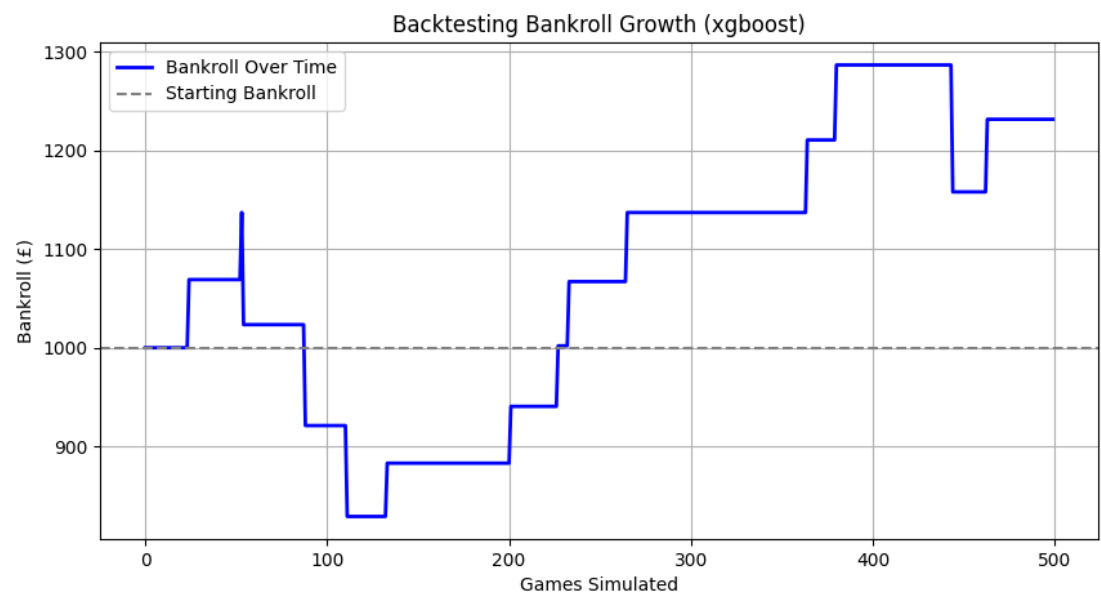
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £1231.53
- 📁 Total Staked: £1466.39
- 📈 Total Profit: £231.53
- 🇮🇹 ROI: 23.15%
- ✅ Win rate: 71.43% over 14 bets
- 📉 Sharpe Ratio: 0.23
- 🎯 Edge Over Bookies: 11.24%



Backtesting Results Summary

