

Model Training Report: Track 1

Configuration Details

Target Variable

target

Selected Features

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

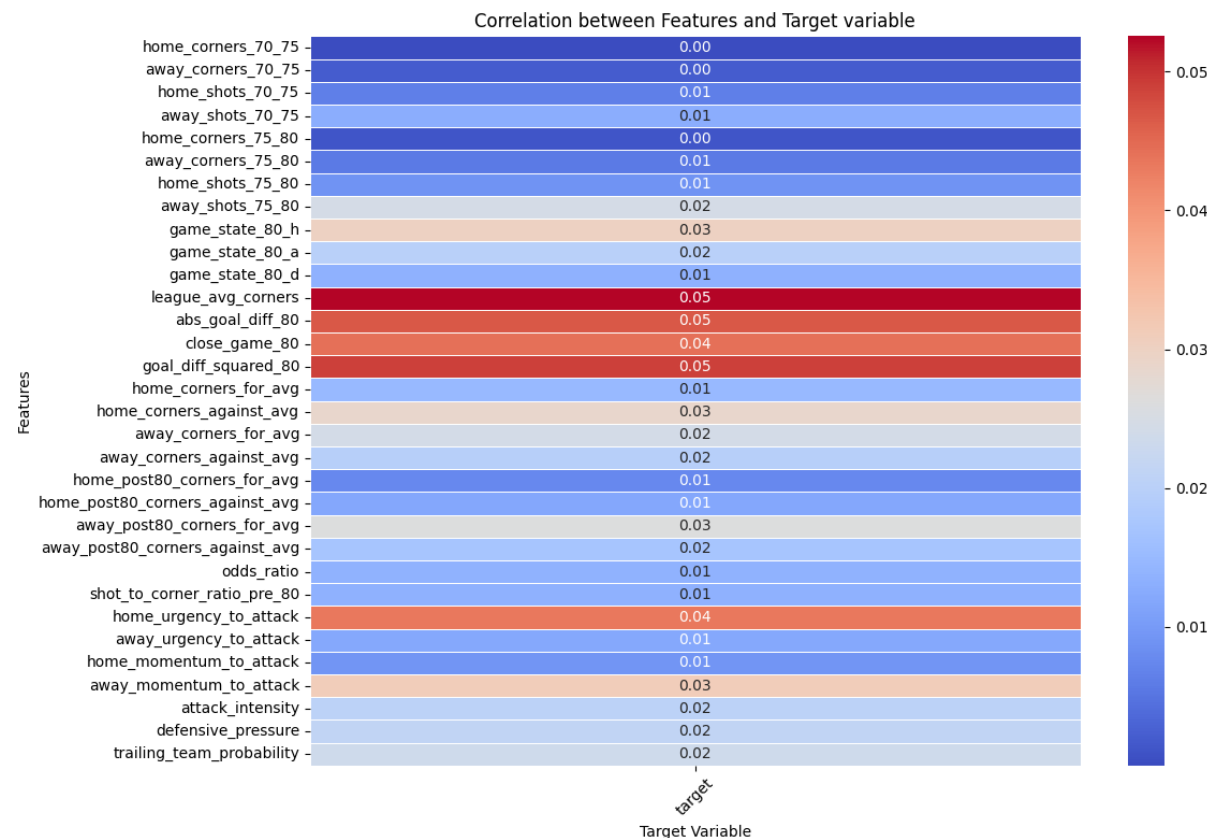
Constructed Features

game_state_80_h, game_state_80_a, game_state_80_d, league_avg_corners,
abs_goal_diff_80, close_game_80, goal_diff_squared_80, home_corners_for_avg,
home_corners_against_avg, away_corners_for_avg, away_corners_against_avg,
home_post80_corners_for_avg, home_post80_corners_against_avg,
away_post80_corners_for_avg, away_post80_corners_against_avg, odds_ratio,
shot_to_corner_ratio_pre_80, home_urgency_to_attack, away_urgency_to_attack,
home_momentum_to_attack, away_momentum_to_attack, attack_intensity,
defensive_pressure, trailing_team_probability

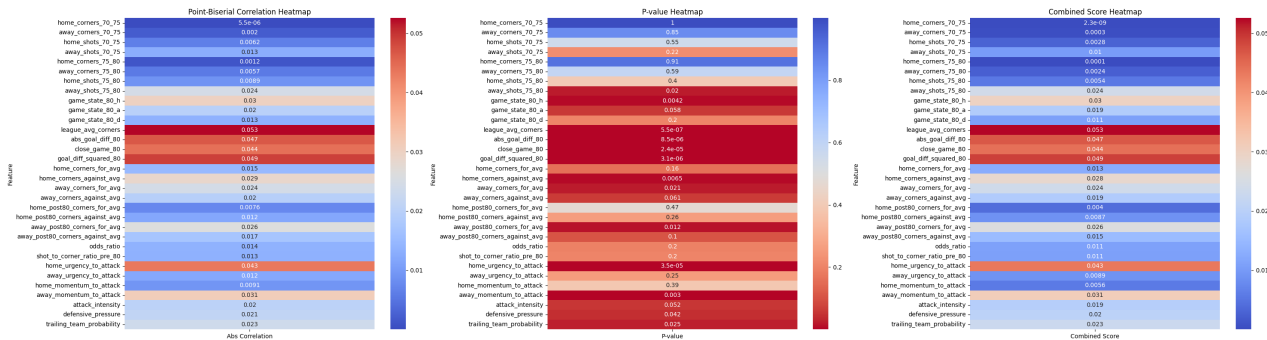
Models Trained (& Hyperparameters)

- **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}

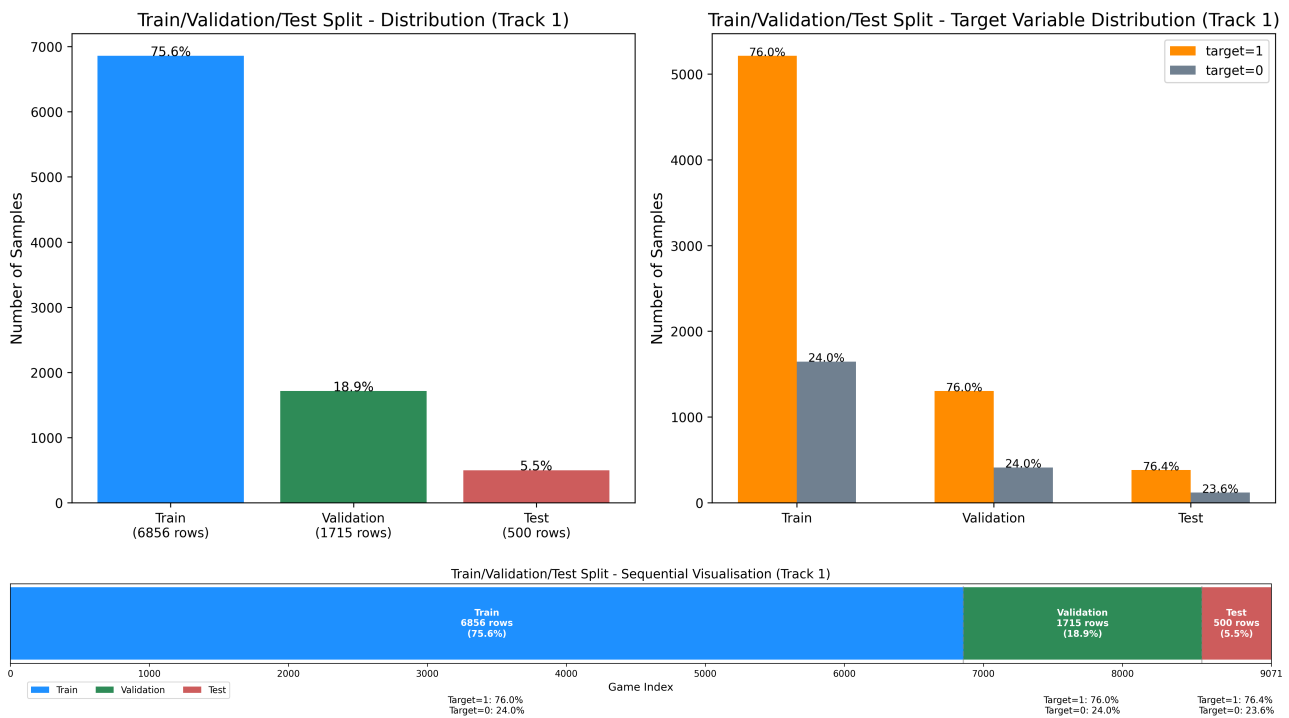
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

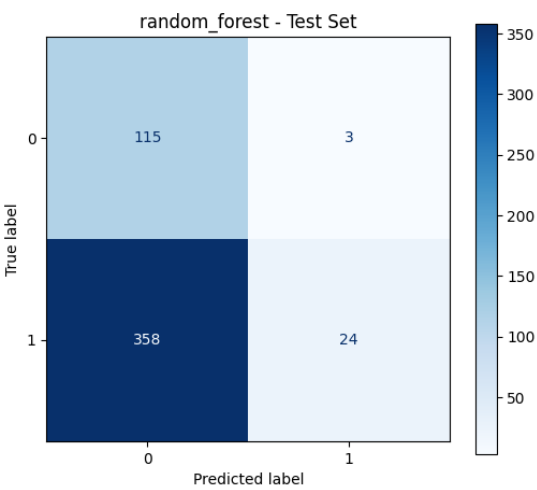
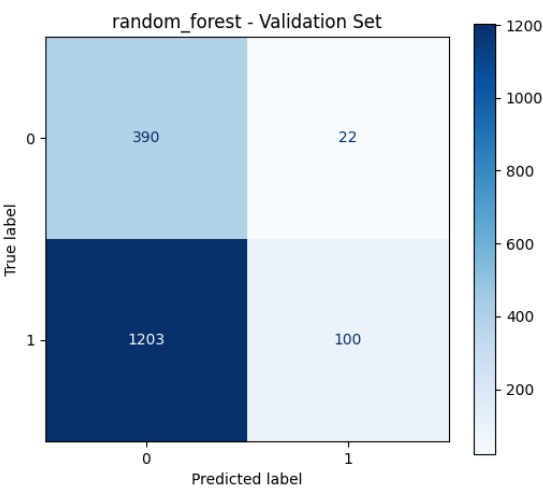
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
home_corners_against_avg	0.06358	game_state_80_d	0.00225
away_post80_corners_for_avg	0.06233	close_game_80	0.00367
away_corners_for_avg	0.06198	game_state_80_h	0.00399
shot_to_corner_ratio_pre_80	0.06186	game_state_80_a	0.00401
away_corners_against_avg	0.06113	away_corners_70_75	0.00778
home_corners_for_avg	0.06090		
home_post80_corners_for_avg	0.06044		
home_post80_corners_against_avg	0.06018		

Optimal Threshold

- **Optimal threshold:** 0.86
- **Expected Precision:** 0.818
- **Expected Recall:** 0.062

Classification Reports



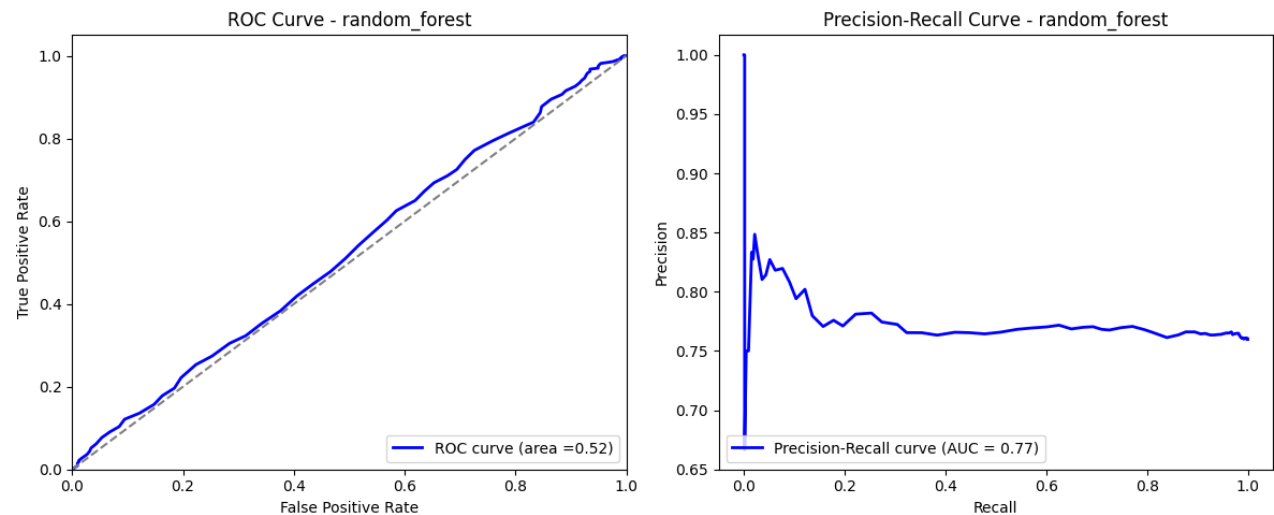
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.24	0.95	0.39	412
1	0.82	0.08	0.14	1303
accuracy			0.29	1715
macro avg	0.53	0.51	0.26	1715
weighted avg	0.68	0.29	0.20	1715

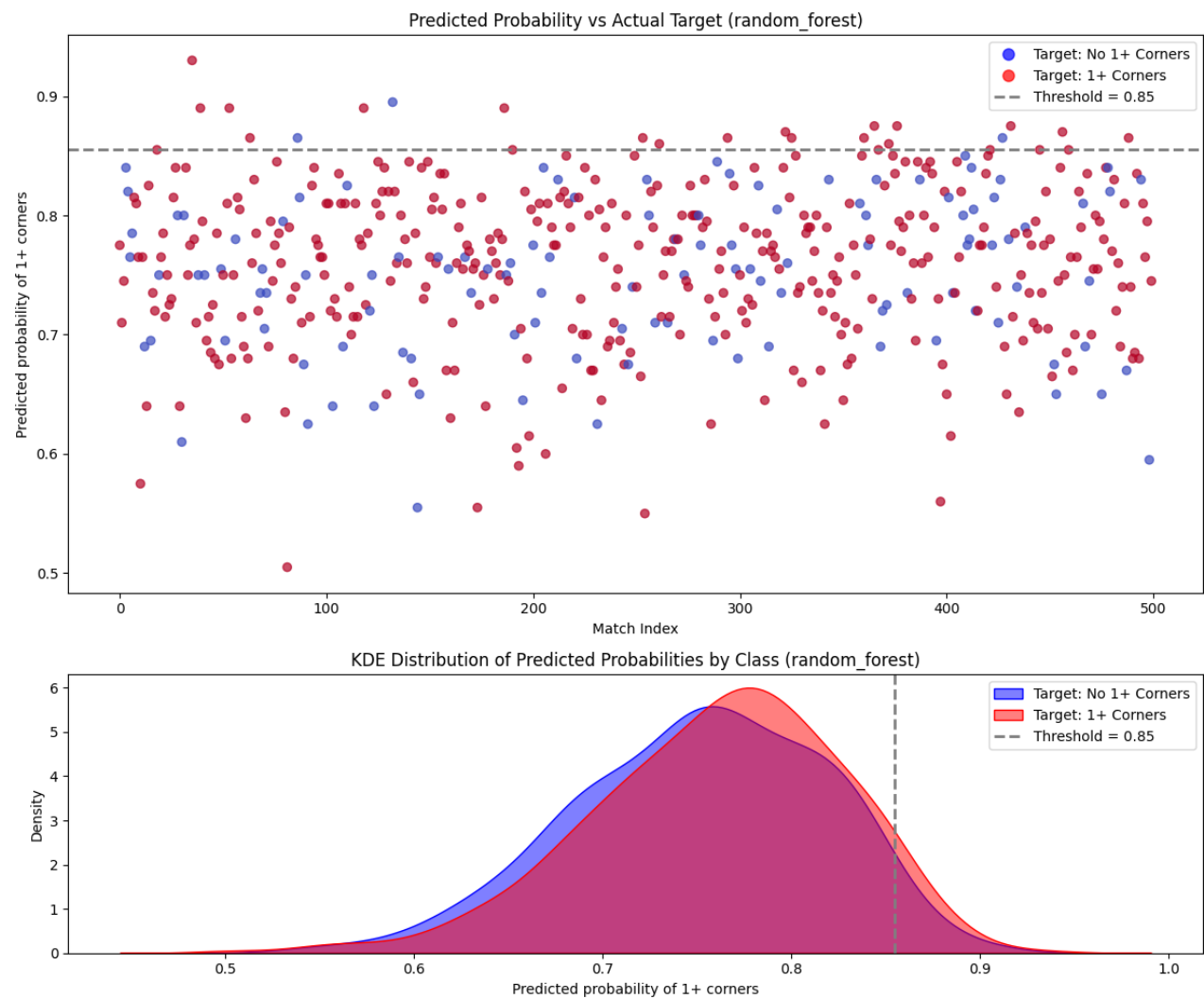
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.24	0.97	0.39	118
1	0.89	0.06	0.12	382
accuracy			0.28	500
macro avg	0.57	0.52	0.25	500
weighted avg	0.74	0.28	0.18	500

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £1272.76

🎲 Total Staked: £2839.80

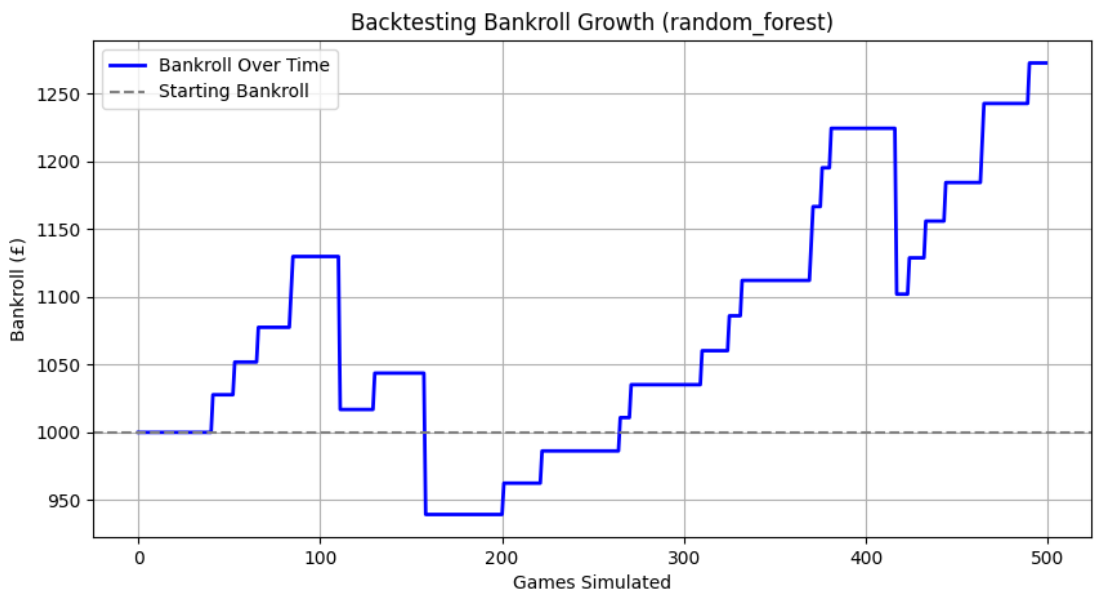
📈 Total Profit: £272.76

🇮🇹 ROI: 27.28%

✅ Win rate: 88.46% over 26 bets

📊 Sharpe Ratio: 0.25

🎯 Edge Over Bookies: 8.69%



logistic_regression Model

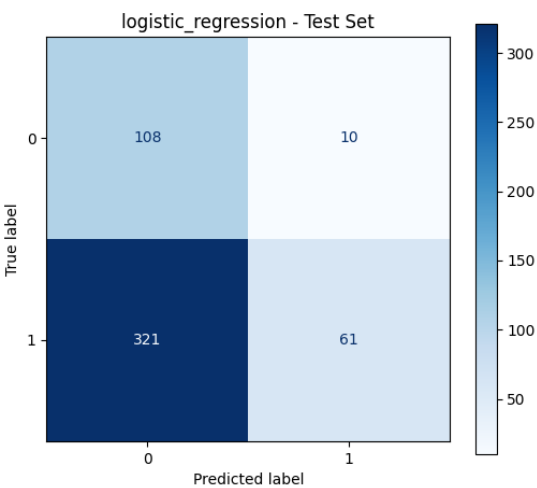
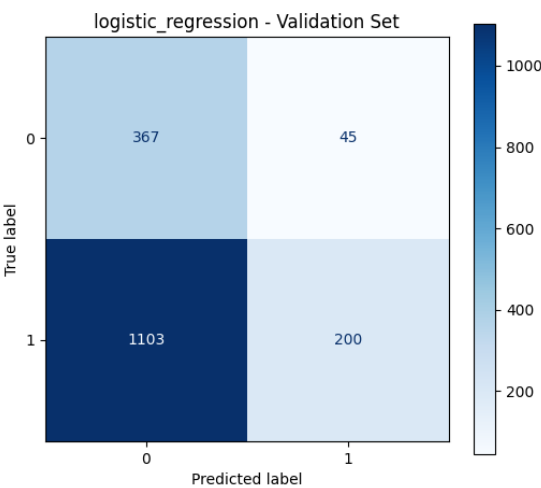
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
home_urgency_to_attack	1.11730	game_state_80_d	-1.03036
away_momentum_to_attack	0.93273	abs_goal_diff_80	-0.67038
away_urgency_to_attack	0.87547	attack_intensity	-0.34529
home_momentum_to_attack	0.73410	goal_diff_squared_80	-0.31094
defensive_pressure	0.49261	trailing_team_probability	-0.28643
away_post80_corners_for_avg	0.41276		
away_shots_75_80	0.38485		
home_shots_75_80	0.37936		

Optimal Threshold

- **Optimal threshold:** 0.57
- **Expected Precision:** 0.816
- **Expected Recall:** 0.153

Classification Reports



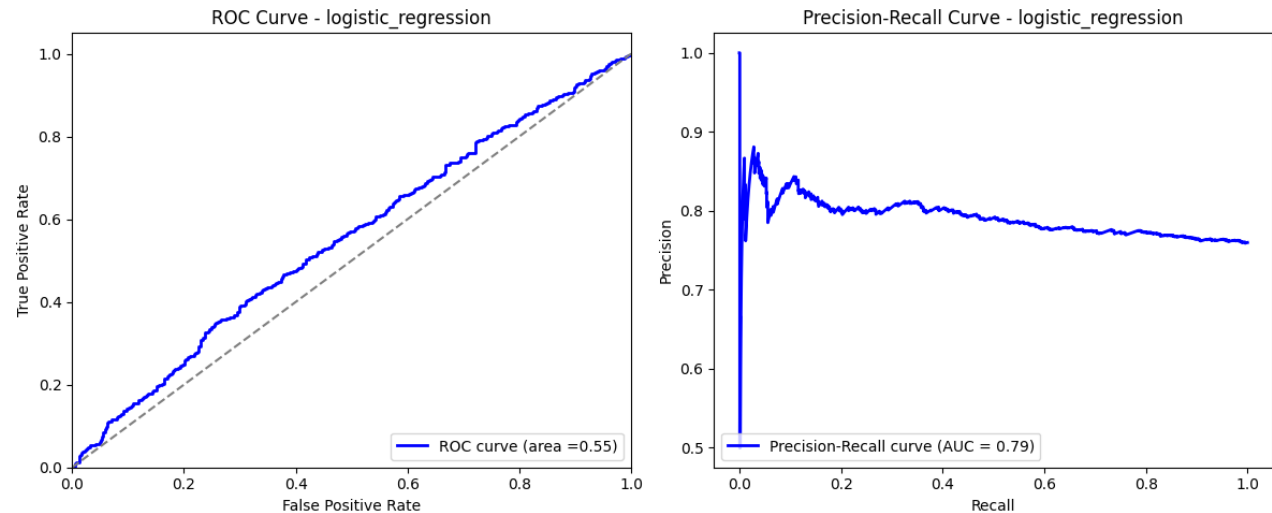
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.25	0.89	0.39	412
1	0.82	0.15	0.26	1303
accuracy			0.33	1715
macro avg	0.53	0.52	0.32	1715
weighted avg	0.68	0.33	0.29	1715

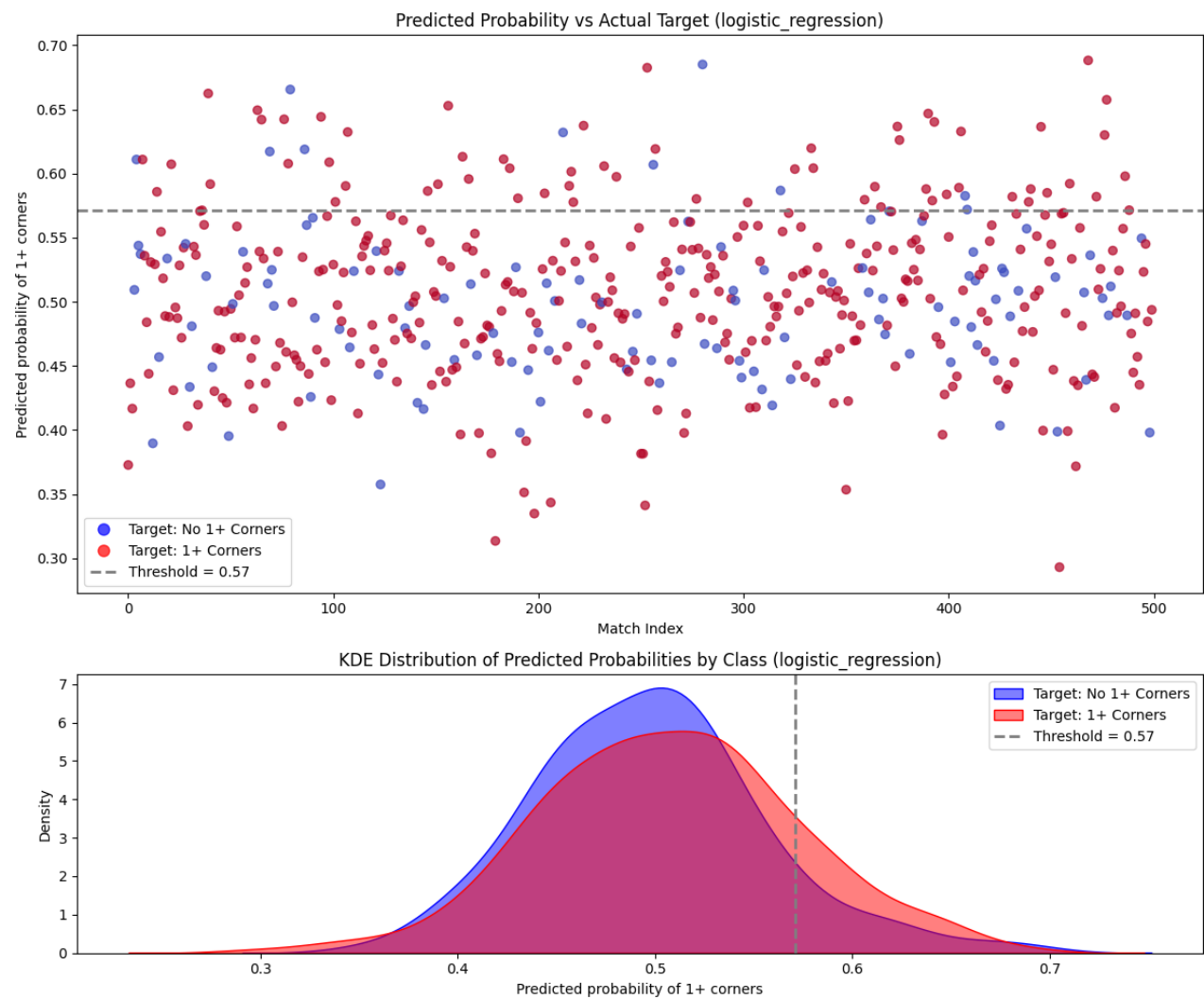
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.25	0.92	0.39	118
1	0.86	0.16	0.27	382
accuracy			0.34	500
macro avg	0.56	0.54	0.33	500
weighted avg	0.72	0.34	0.30	500

ROC and Precision-Recall Curves



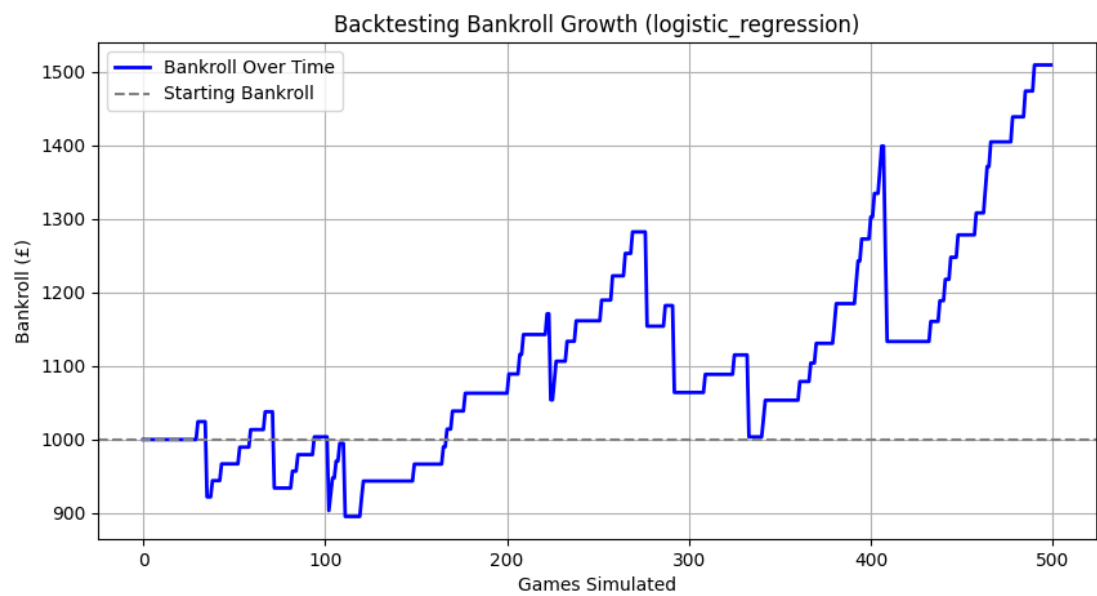
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £1509.54
- 📊 Total Staked: £7965.21
- 📈 Total Profit: £509.54
- 📊 ROI: 50.95%
- ✅ Win rate: 85.92% over 71 bets
- 📊 Sharpe Ratio: 0.16
- 🎯 Edge Over Bookies: 6.14%



svc Model

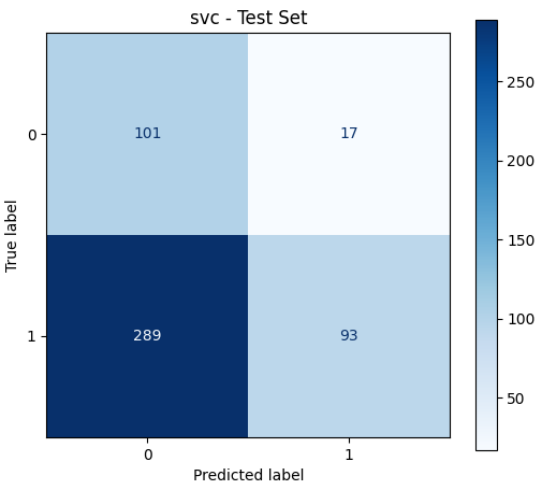
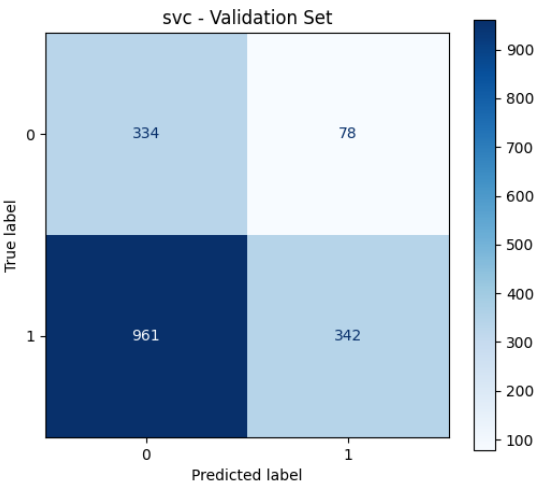
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
home_corners_70_75	0.00000	trailing_team_probability	0.00000
away_corners_70_75	0.00000	home_shots_70_75	0.00000
defensive_pressure	0.00000	away_shots_70_75	0.00000
attack_intensity	0.00000	home_corners_75_80	0.00000
away_momentum_to_attack	0.00000	away_corners_75_80	0.00000
home_momentum_to_attack	0.00000		
away_urgency_to_attack	0.00000		
home_urgency_to_attack	0.00000		

Optimal Threshold

- **Optimal threshold:** 0.78
- **Expected Precision:** 0.814
- **Expected Recall:** 0.259

Classification Reports



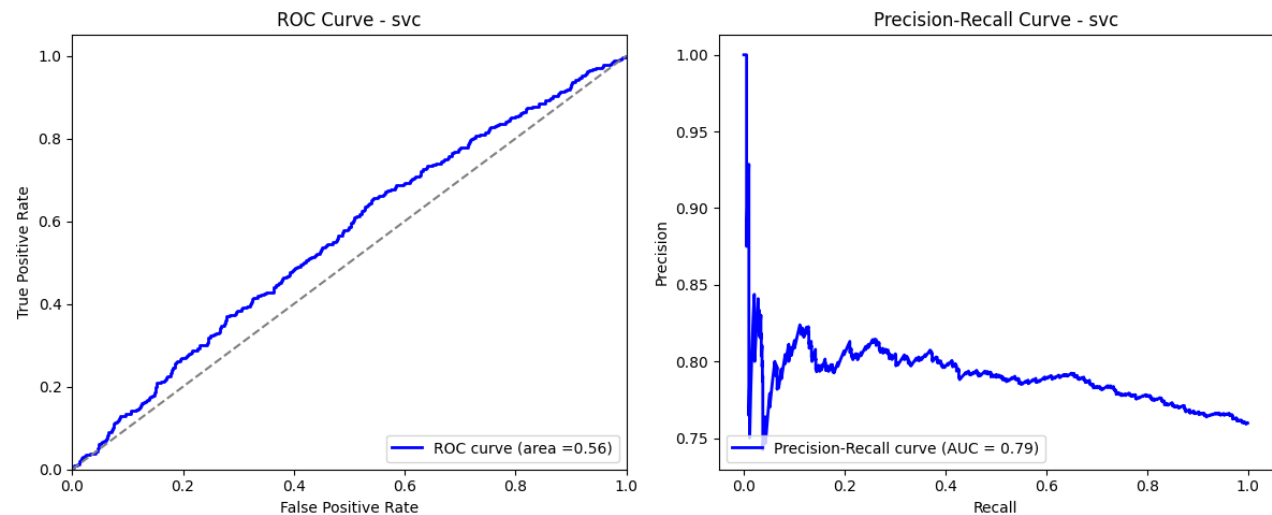
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.26	0.81	0.39	412
1	0.81	0.26	0.40	1303
accuracy			0.39	1715
macro avg	0.54	0.54	0.39	1715
weighted avg	0.68	0.39	0.40	1715

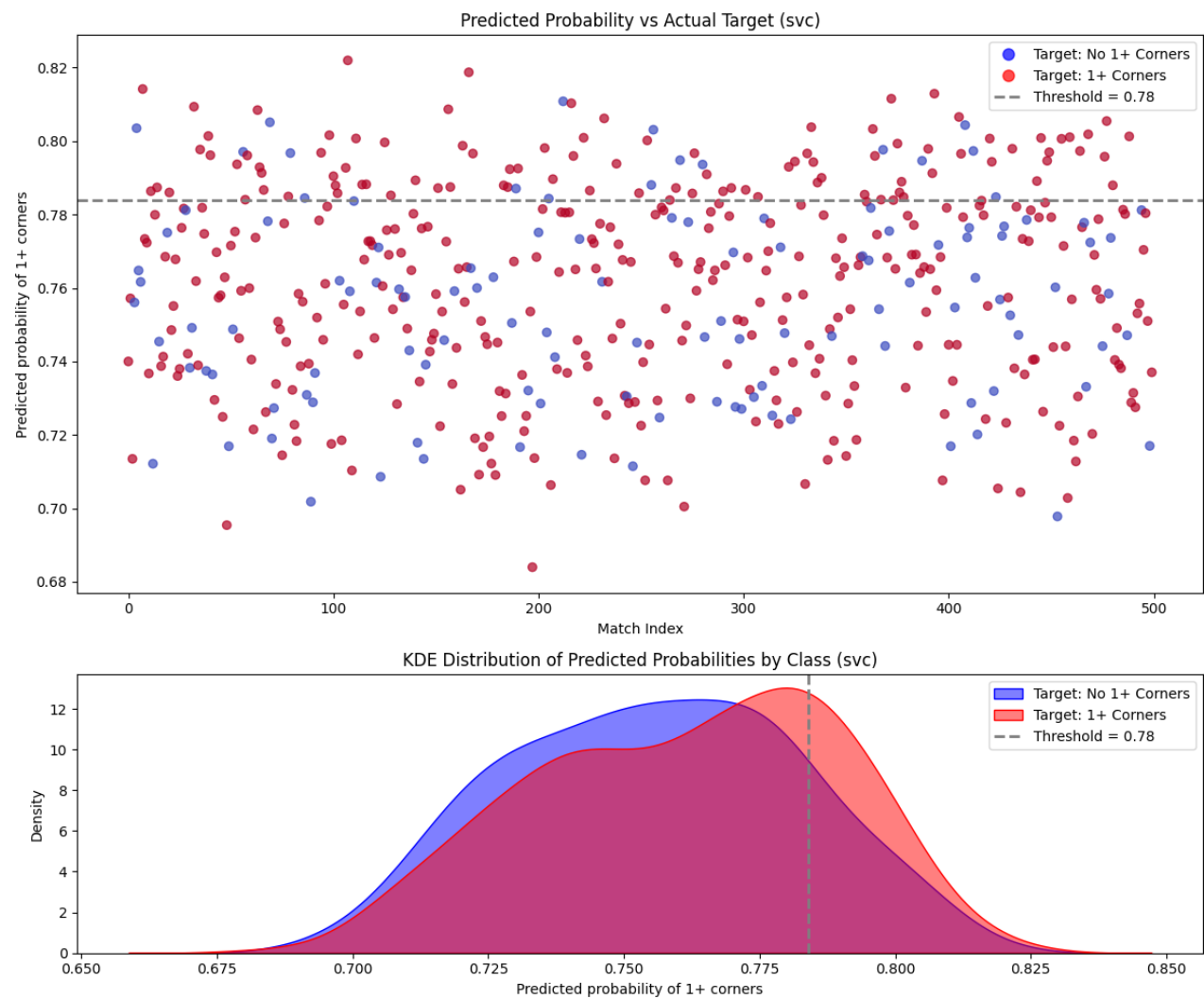
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.26	0.86	0.40	118
1	0.85	0.24	0.38	382
accuracy			0.39	500
macro avg	0.55	0.55	0.39	500
weighted avg	0.71	0.39	0.38	500

ROC and Precision-Recall Curves



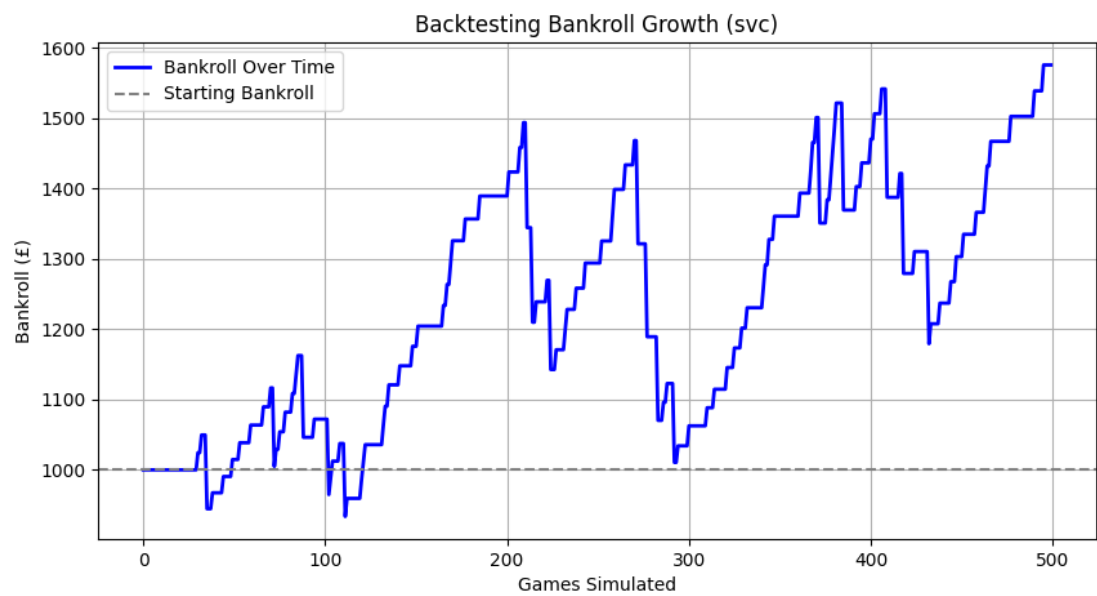
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £1575.68
- 📁 Total Staked: £13417.20
- 📈 Total Profit: £575.68
- 🇮🇹 ROI: 57.57%
- ✅ Win rate: 84.40% over 109 bets
- 📉 Sharpe Ratio: 0.12
- 🎯 Edge Over Bookies: 4.63%



xgboost Model

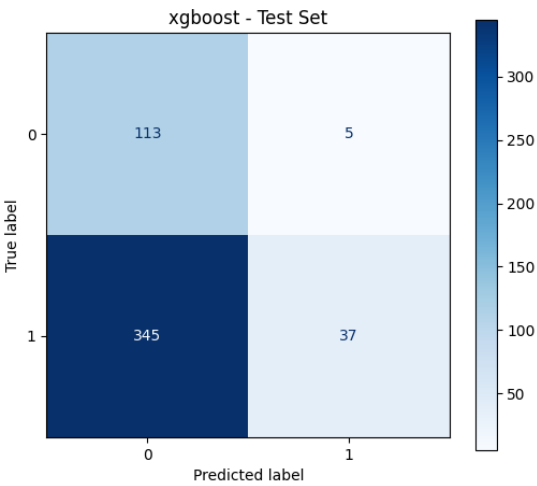
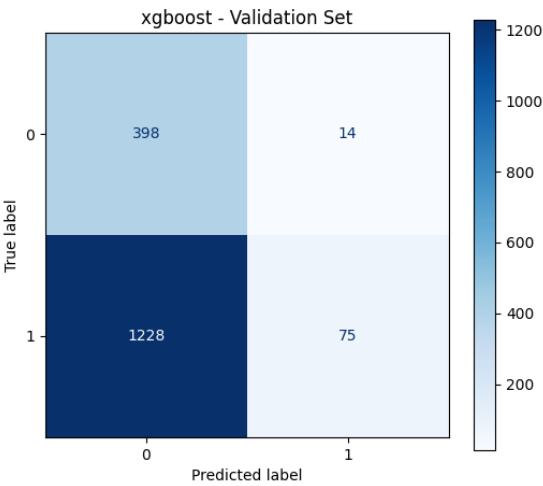
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
away_corners_75_80	0.05530	home_shots_75_80	0.00000
away_urgency_to_attack	0.05264	game_state_80_h	0.00000
league_avg_corners	0.05130	game_state_80_d	0.00000
home_corners_70_75	0.04513	close_game_80	0.00000
abs_goal_diff_80	0.04482	goal_diff_squared_80	0.00000
trailing_team_probability	0.04454		
home_urgency_to_attack	0.04327		
away_post80_corners_for_avg	0.04258		

Optimal Threshold

- **Optimal threshold:** 0.83
- **Expected Precision:** 0.837
- **Expected Recall:** 0.059

Classification Reports



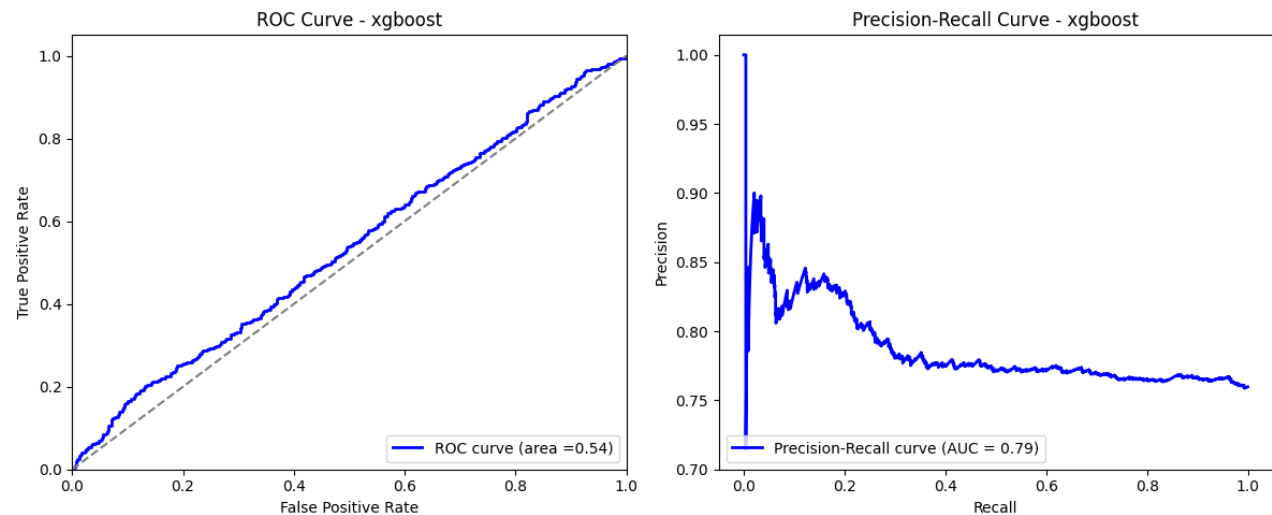
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.24	0.97	0.39	412
1	0.84	0.06	0.11	1303
accuracy			0.28	1715
macro avg	0.54	0.51	0.25	1715
weighted avg	0.70	0.28	0.18	1715

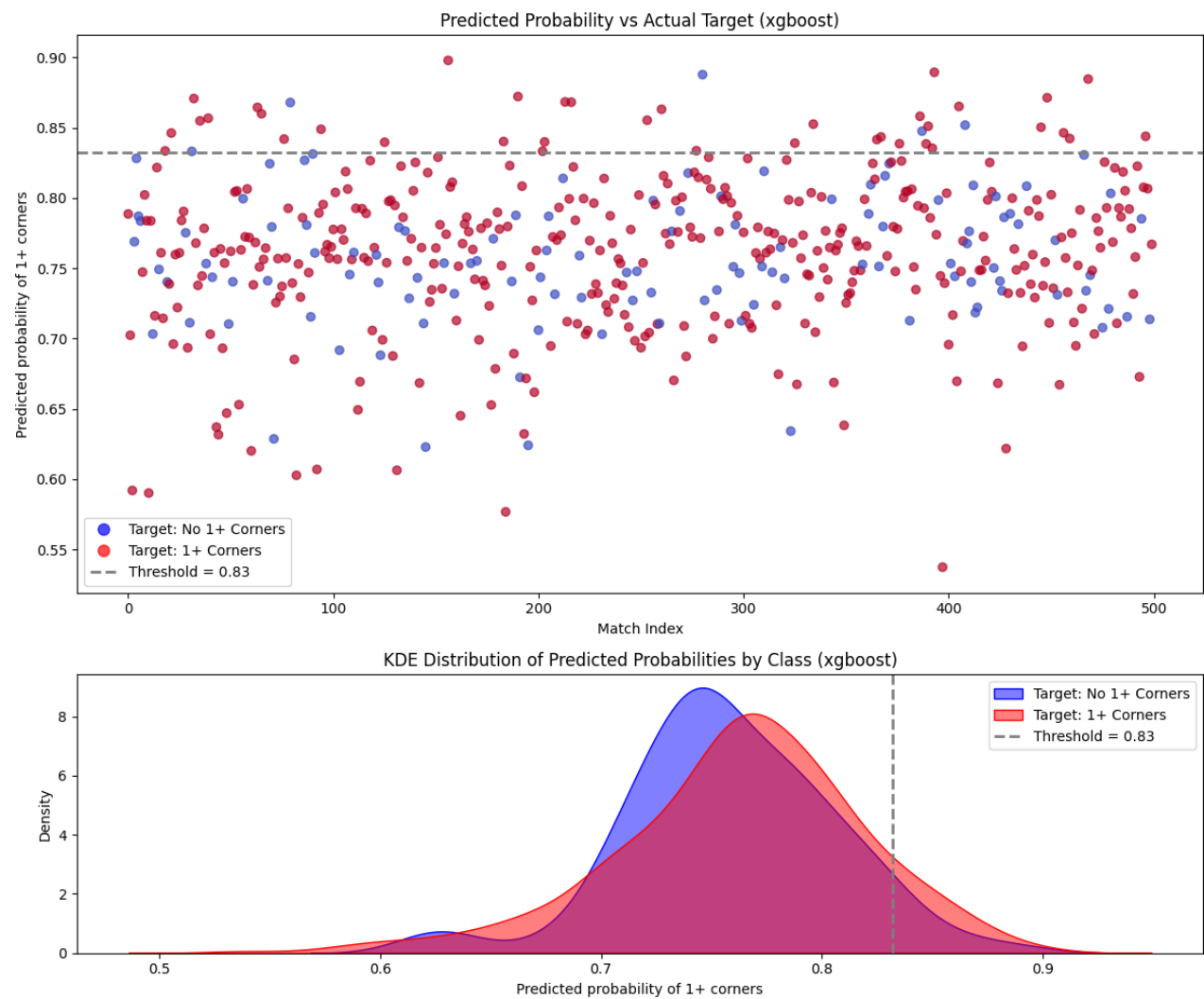
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.25	0.96	0.39	118
1	0.88	0.10	0.17	382
accuracy			0.30	500
macro avg	0.56	0.53	0.28	500
weighted avg	0.73	0.30	0.23	500

ROC and Precision-Recall Curves



Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

🏠 Initial Bankroll: £1000.00

💰 Final Bankroll: £1600.79

🎲 Total Staked: £5345.84

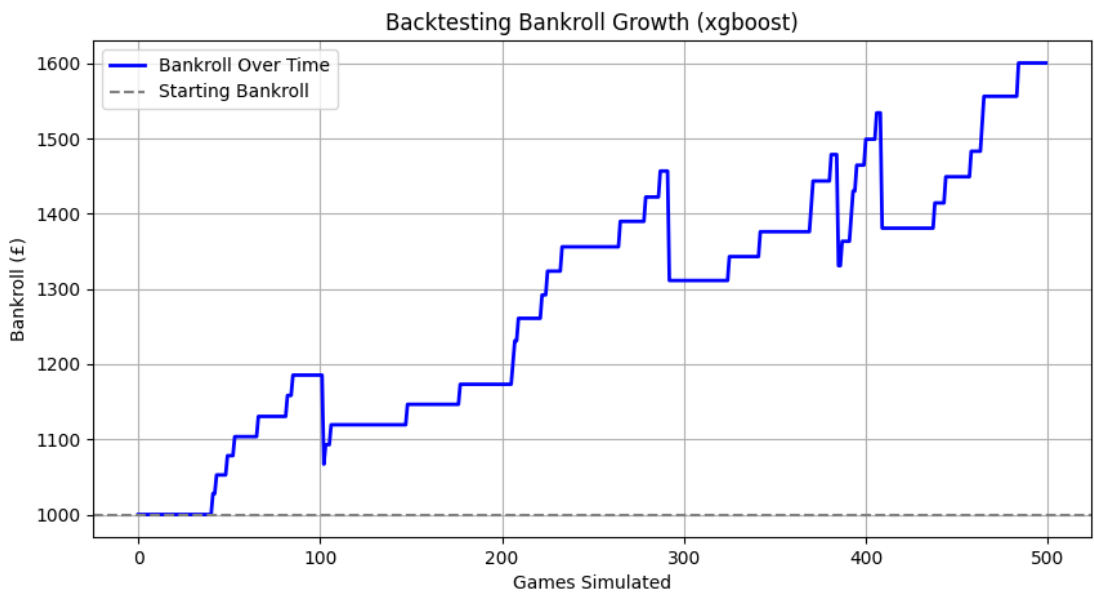
📈 Total Profit: £600.79

🇮🇹 ROI: 60.08%

✅ Win rate: 90.24% over 41 bets

📊 Sharpe Ratio: 0.33

🎯 Edge Over Bookies: 10.47%



Backtesting Results Summary

