

Model Training Report: Track 1

Configuration Details

Target Variable

target

Selected Features

goal_diff_80, 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

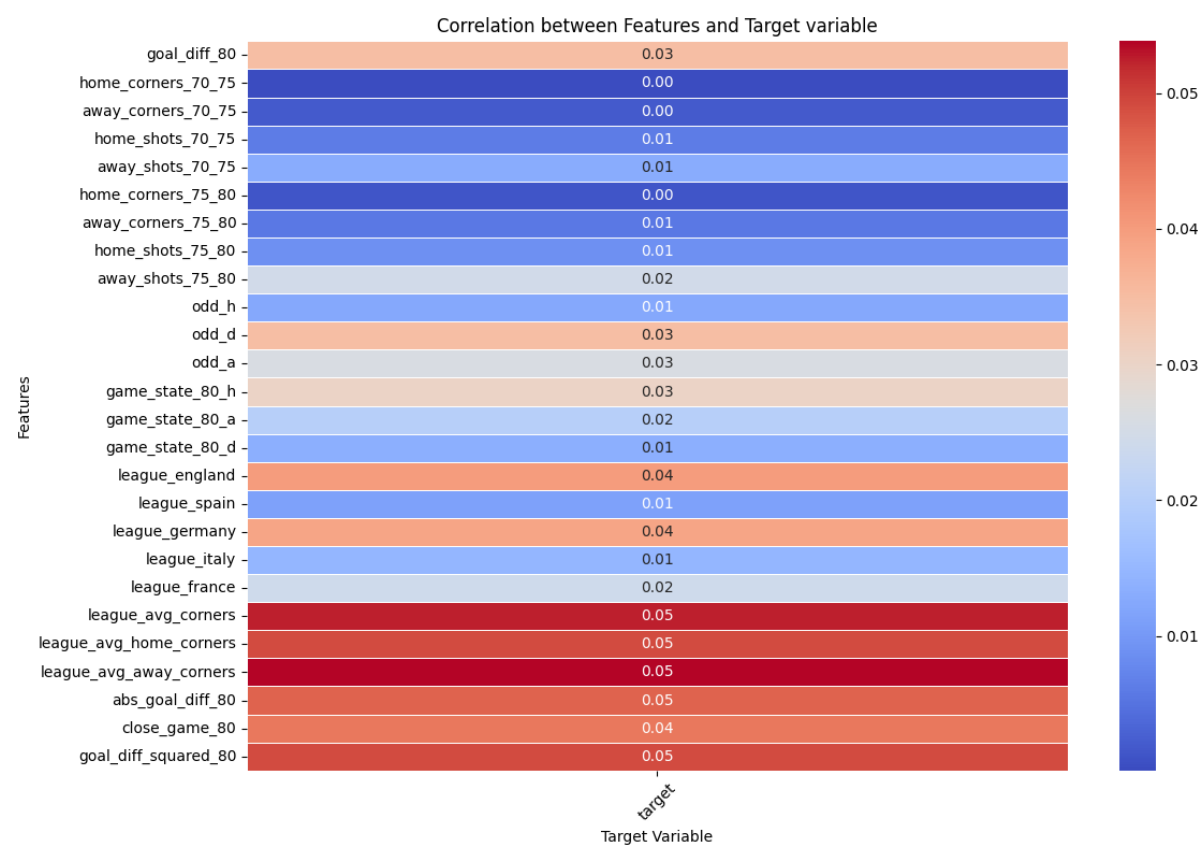
Constructed Features

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

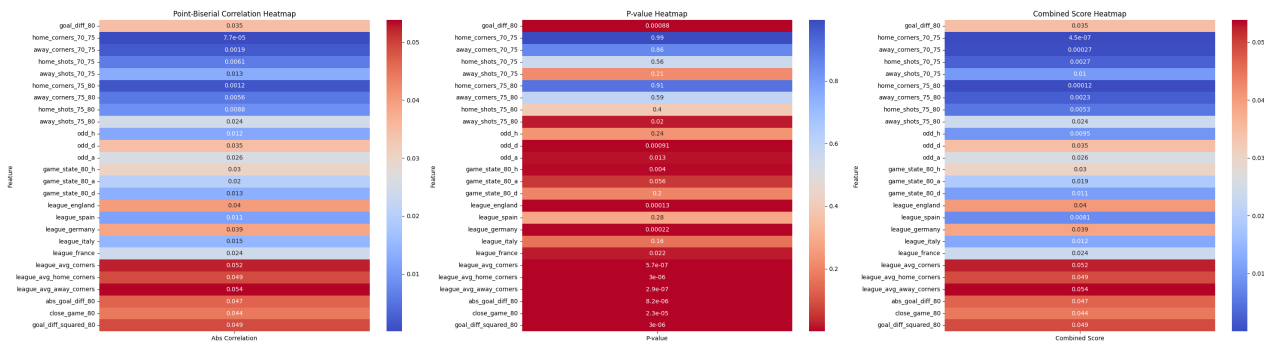
Models Trained (& Hyperparameters)

- **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}

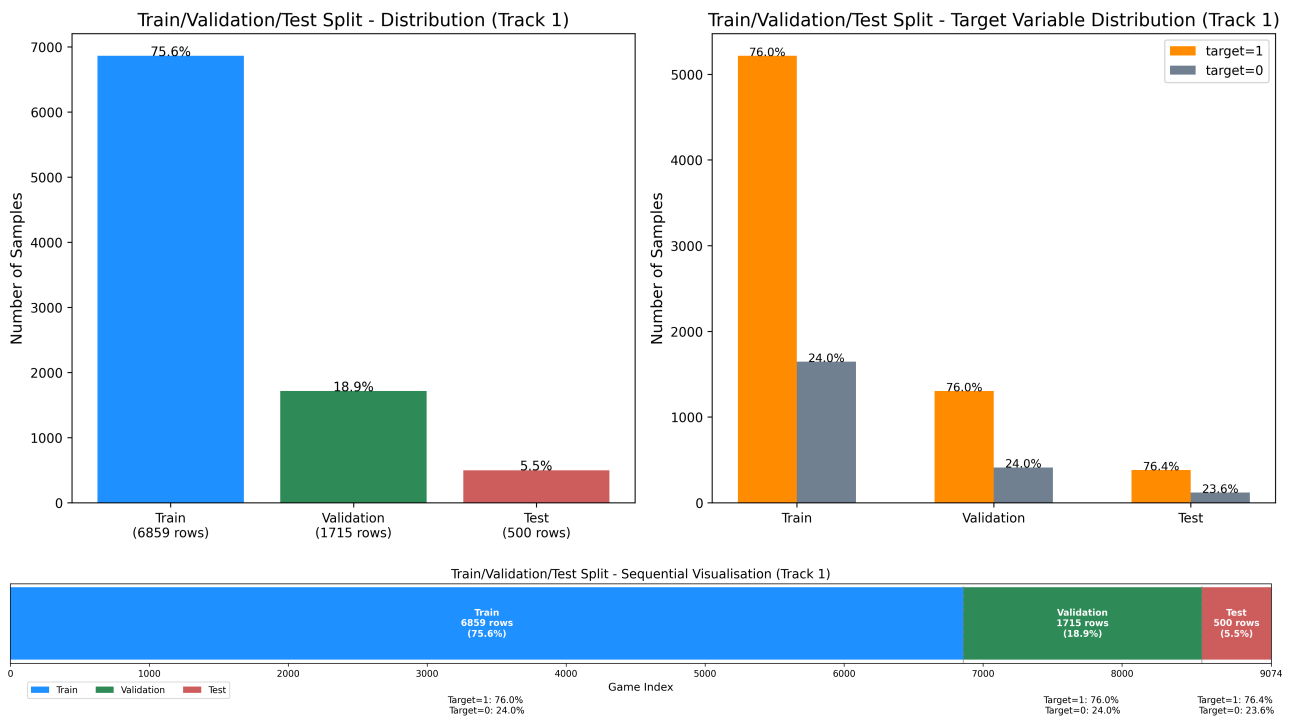
Feature Correlation



Point-Biserial Correlation



Dataset Train/Validation/Test Split Visualisation



random_forest Model

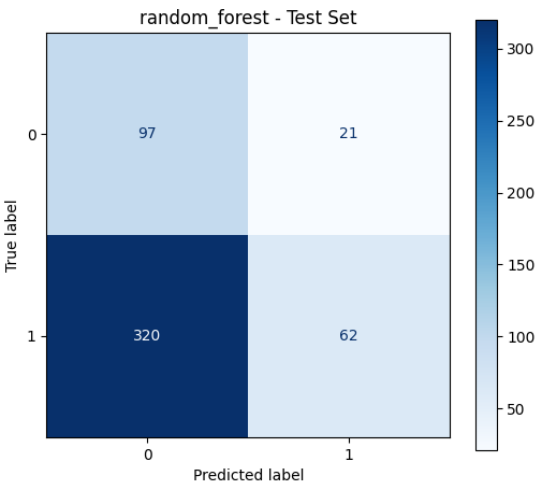
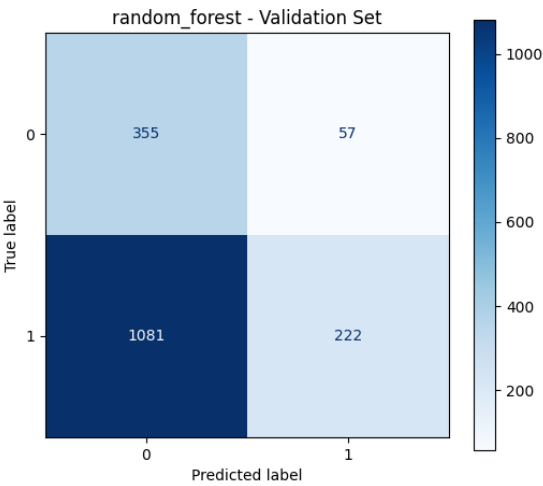
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
odd_a	0.18648	league_england	0.00285
odd_h	0.18337	league_france	0.00415
odd_d	0.17210	league_spain	0.00435
home_shots_70_75	0.05188	league_germany	0.00435
home_shots_75_80	0.05113	game_state_80_d	0.00513
away_shots_70_75	0.04752		
away_shots_75_80	0.04481		
home_corners_75_80	0.03013		

Optimal Threshold

- **Optimal threshold:** 0.88
- **Expected Precision:** 0.796
- **Expected Recall:** 0.170

Classification Reports



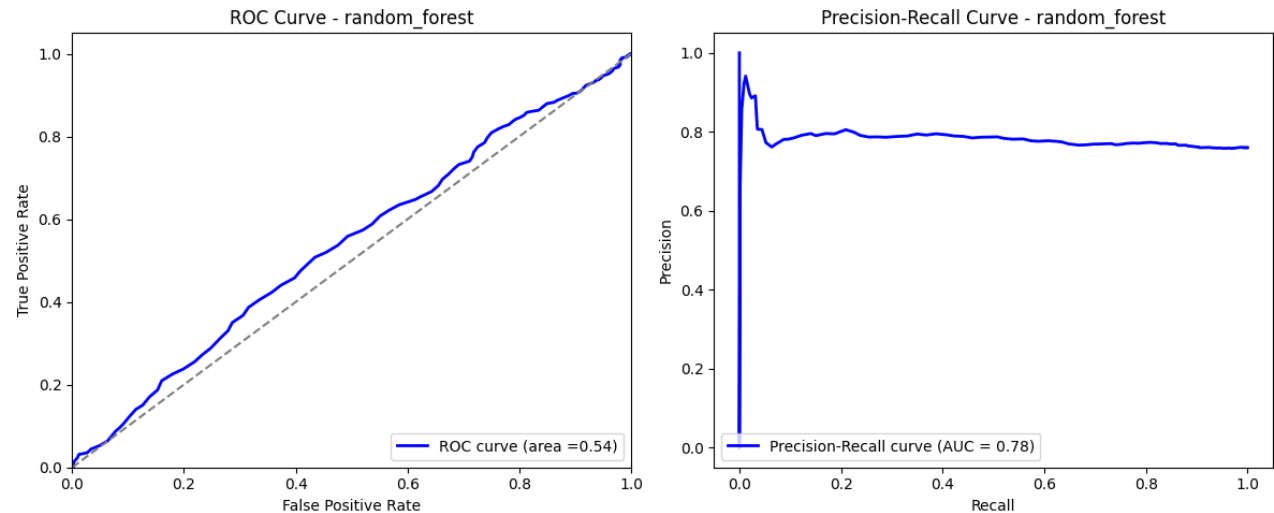
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.25	0.86	0.38	412
1	0.80	0.17	0.28	1303
accuracy			0.34	1715
macro avg	0.52	0.52	0.33	1715
weighted avg	0.66	0.34	0.31	1715

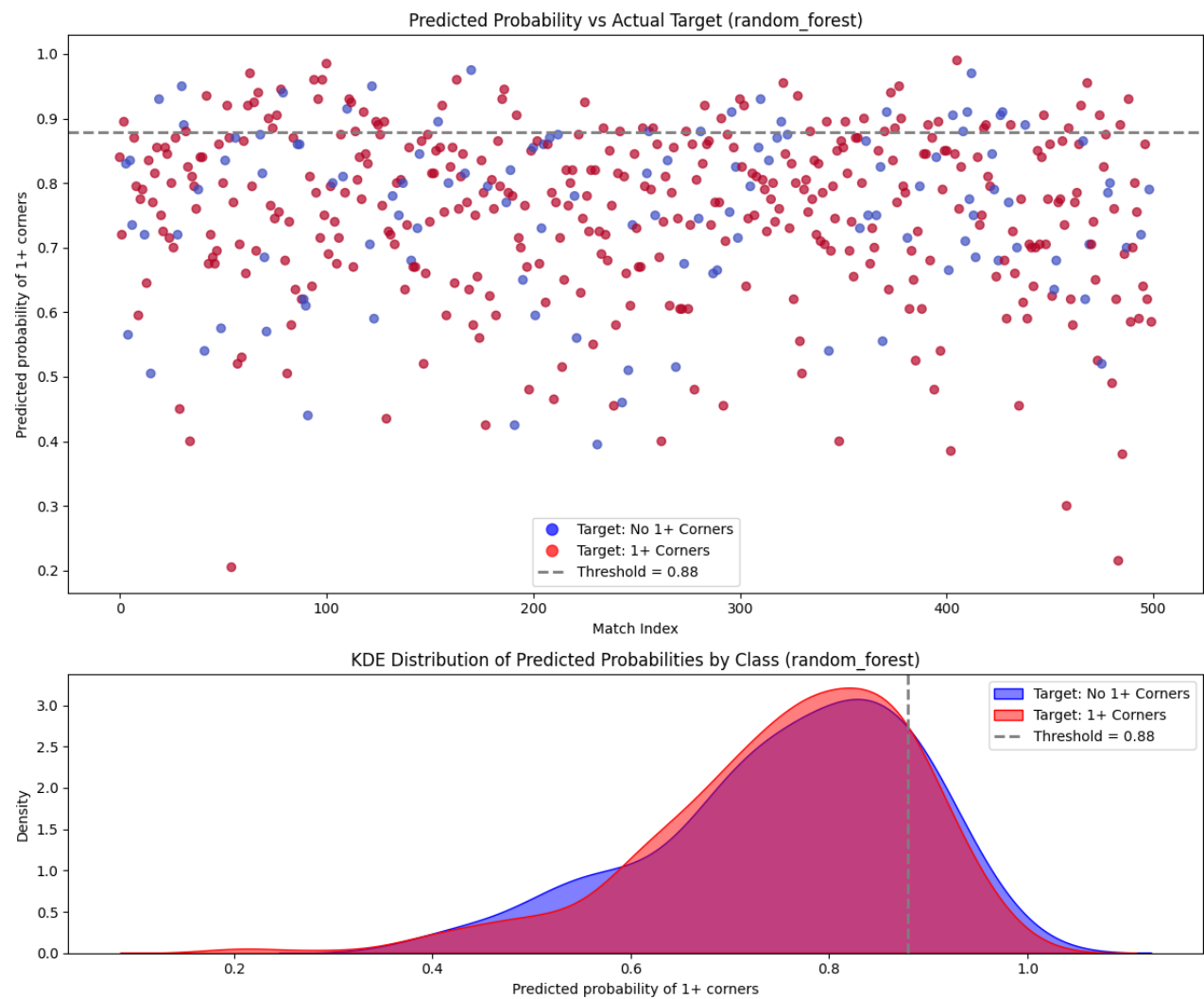
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.23	0.82	0.36	118
1	0.75	0.16	0.27	382
accuracy			0.32	500
macro avg	0.49	0.49	0.31	500
weighted avg	0.63	0.32	0.29	500

ROC and Precision-Recall Curves



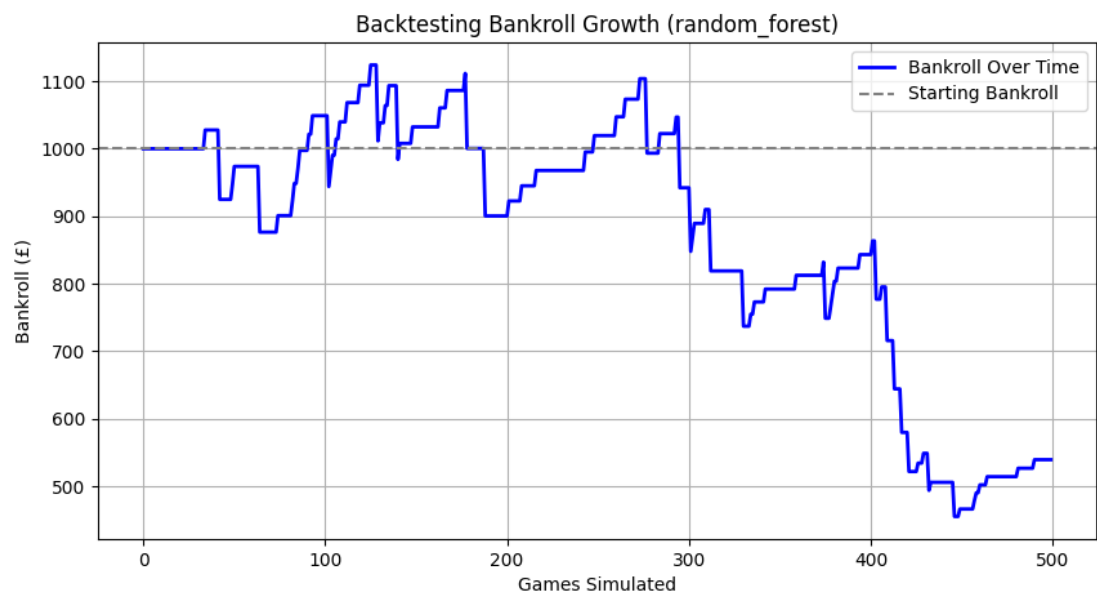
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £539.35
- 📊 Total Staked: £6949.50
- 📈 Total Profit: £-460.65
- 🇩🇪 ROI: -46.07%
- ✅ Win rate: 75.00% over 80 bets
- 📉 Sharpe Ratio: -0.11
- 🎯 Edge Over Bookies: -4.78%



logistic_regression Model

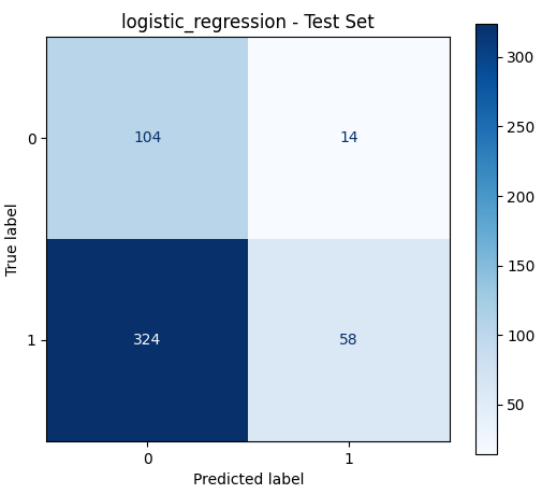
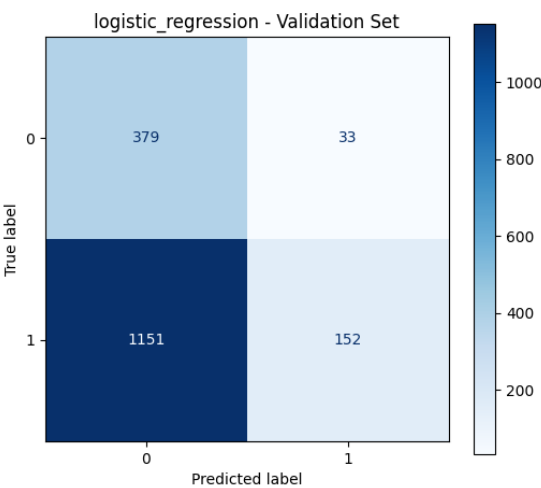
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
away_shots_75_80	0.50221	abs_goal_diff_80	-0.97341
home_shots_75_80	0.31705	odd_h	-0.65807
league_england	0.15500	odd_a	-0.26889
game_state_80_a	0.12282	away_shots_70_75	-0.22689
league_avg_corners	0.09442	odd_d	-0.19507
league_avg_away_corners	0.08863		
home_shots_70_75	0.07818		
league_avg_home_corners	0.07720		

Optimal Threshold

- **Optimal threshold:** 0.57
- **Expected Precision:** 0.821
- **Expected Recall:** 0.116

Classification Reports



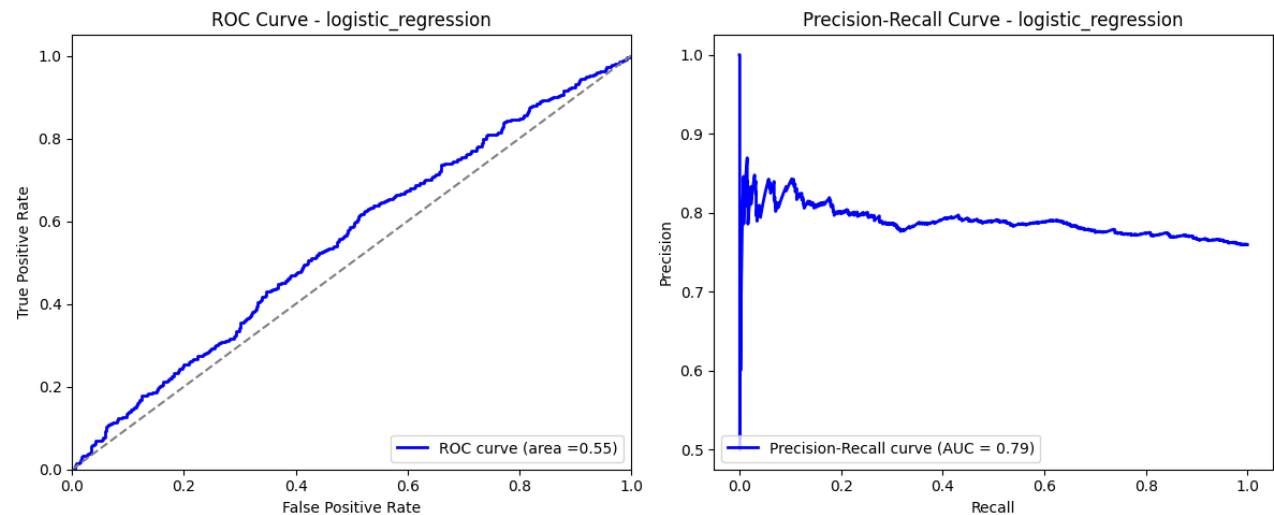
Classification Report (Validation Set)

	precision	recall	f1-score	support
0	0.25	0.92	0.39	412
1	0.82	0.12	0.20	1303
accuracy			0.31	1715
macro avg	0.53	0.52	0.30	1715
weighted avg	0.68	0.31	0.25	1715

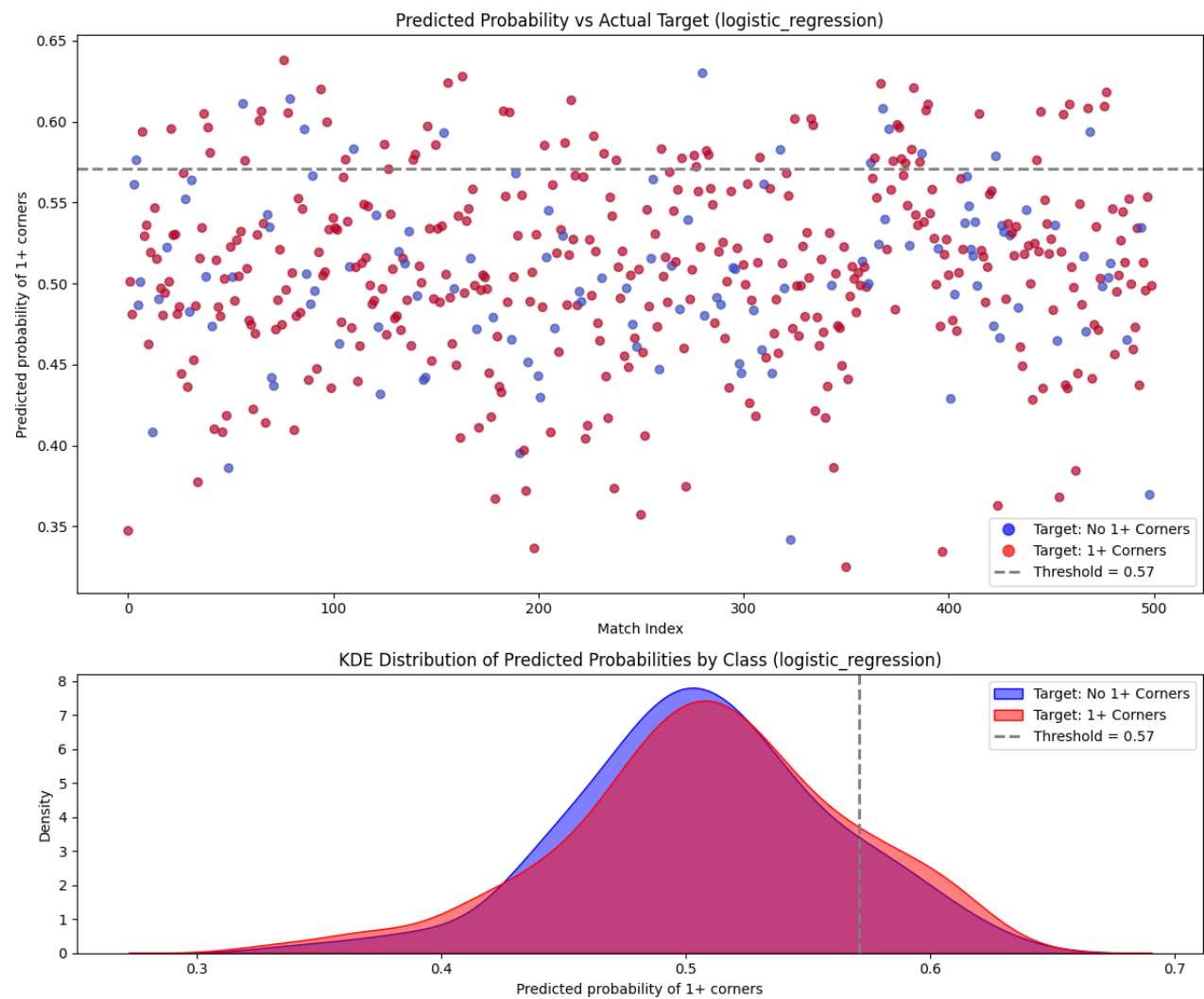
Classification Report (Test Set)

	precision	recall	f1-score	support
0	0.24	0.88	0.38	118
1	0.81	0.15	0.26	382
accuracy			0.32	500
macro avg	0.52	0.52	0.32	500
weighted avg	0.67	0.32	0.29	500

ROC and Precision-Recall Curves



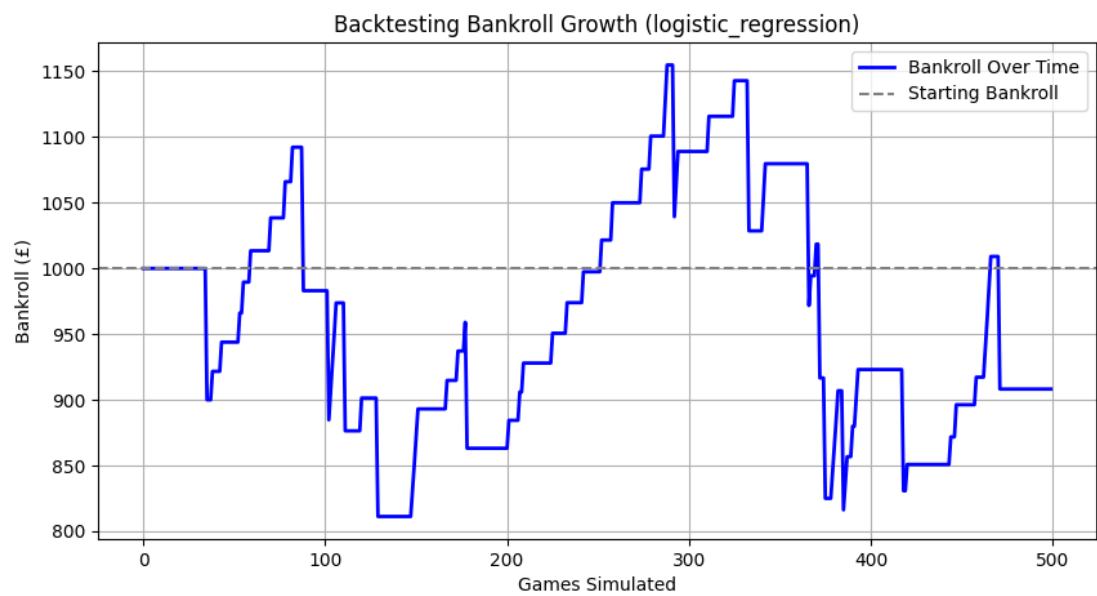
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £908.19
- 📊 Total Staked: £6775.54
- 📈 Total Profit: £-91.81
- 📊 ROI: -9.18%
- ✅ Win rate: 80.28% over 71 bets
- 📊 Sharpe Ratio: -0.00
- 🎯 Edge Over Bookies: 0.51%



svc Model

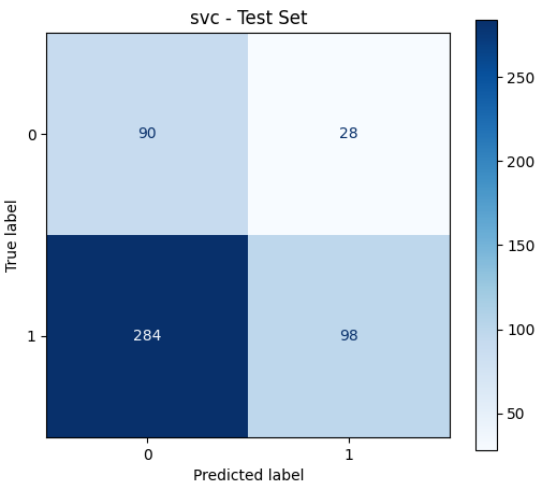
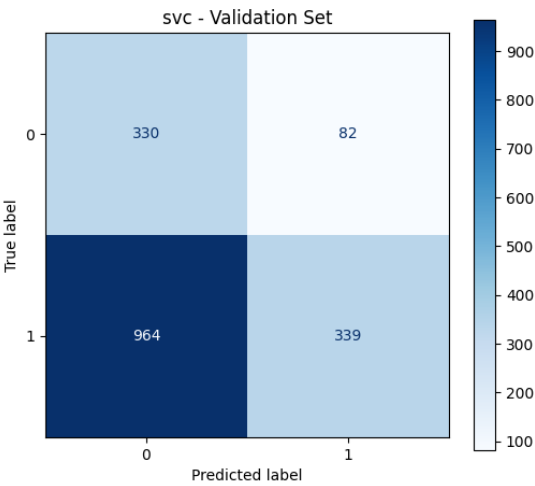
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
goal_diff_80	0.00000	goal_diff_squared_80	0.00000
home_corners_70_75	0.00000	away_corners_70_75	0.00000
close_game_80	0.00000	home_shots_70_75	0.00000
abs_goal_diff_80	0.00000	away_shots_70_75	0.00000
league_avg_away_corners	0.00000	home_corners_75_80	0.00000
league_avg_home_corners	0.00000		
league_avg_corners	0.00000		
league_france	0.00000		

Optimal Threshold

- **Optimal threshold:** 0.78
- **Expected Precision:** 0.803
- **Expected Recall:** 0.256

Classification Reports



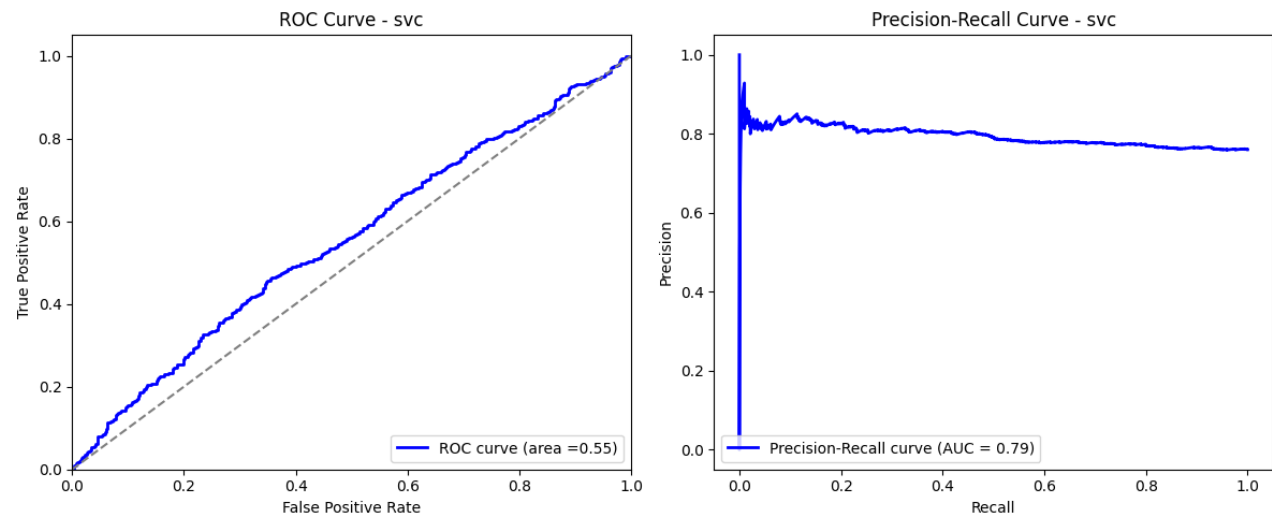
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.26	0.80	0.39	412
1	0.81	0.26	0.39	1303
accuracy			0.39	1715
macro avg 0.53 0.53 0.39 1715 weighted avg 0.67 0.39 0.39 1715				

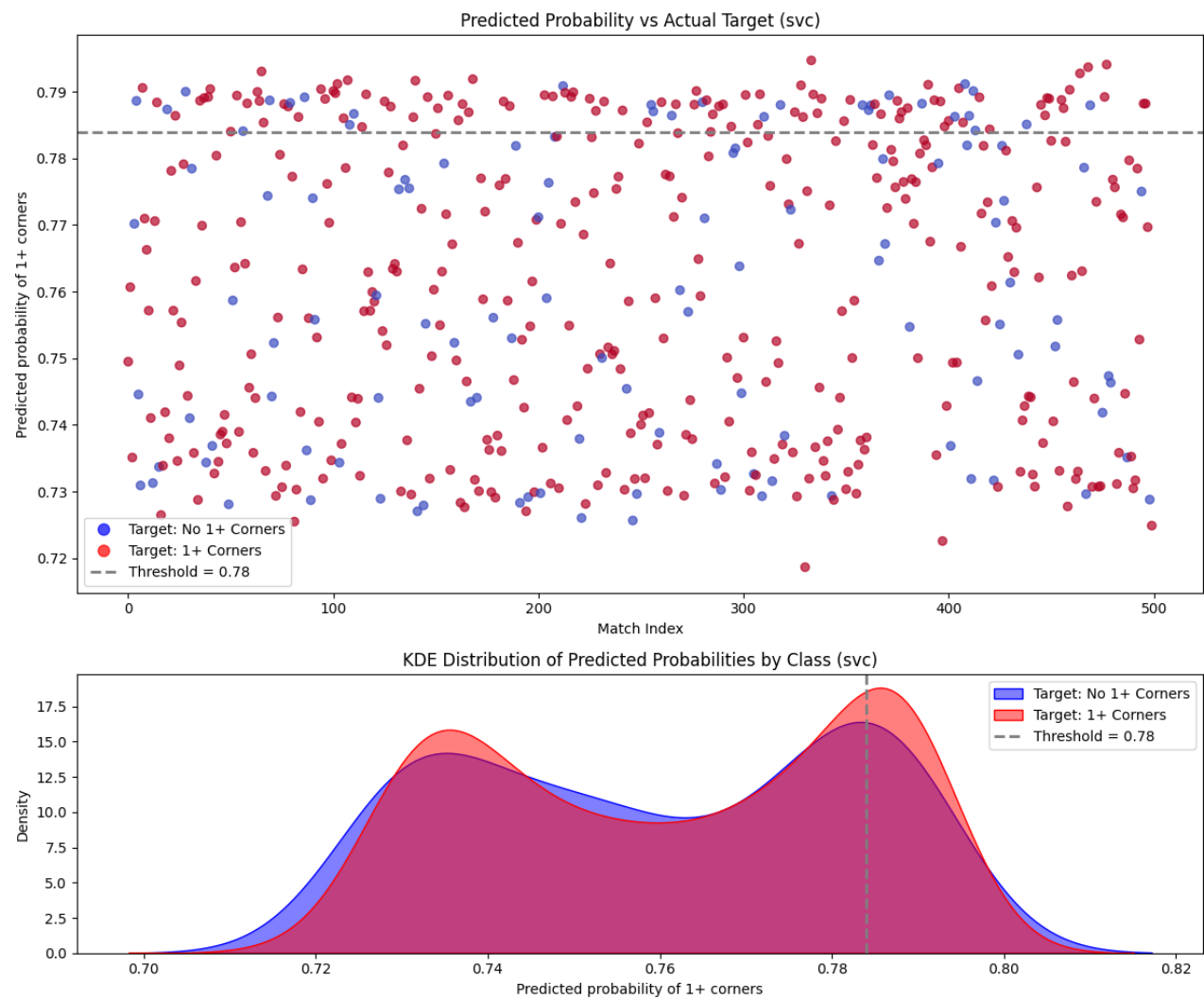
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.24	0.76	0.37	118
1	0.78	0.26	0.39	382
accuracy			0.38	500
macro avg 0.51 0.51 0.38 500 weighted avg 0.65 0.38 0.38 500				

ROC and Precision-Recall Curves



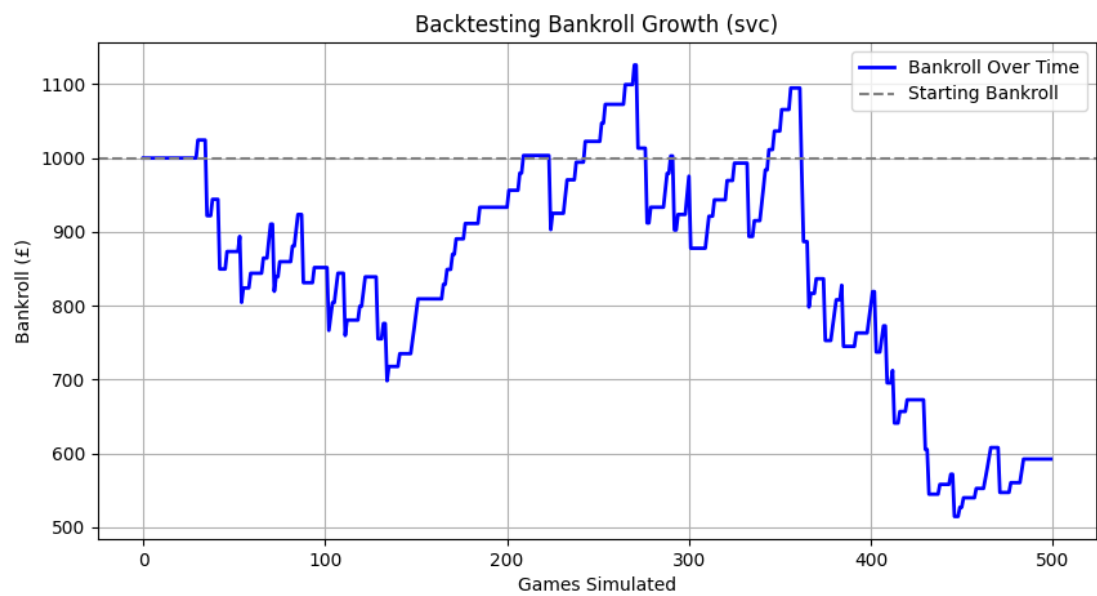
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £592.45
- 📁 Total Staked: £10203.53
- 📈 Total Profit: £-407.55
- 🇮🇹 ROI: -40.76%
- ✅ Win rate: 77.87% over 122 bets
- 📉 Sharpe Ratio: -0.06
- 🎯 Edge Over Bookies: -1.91%



xgboost Model

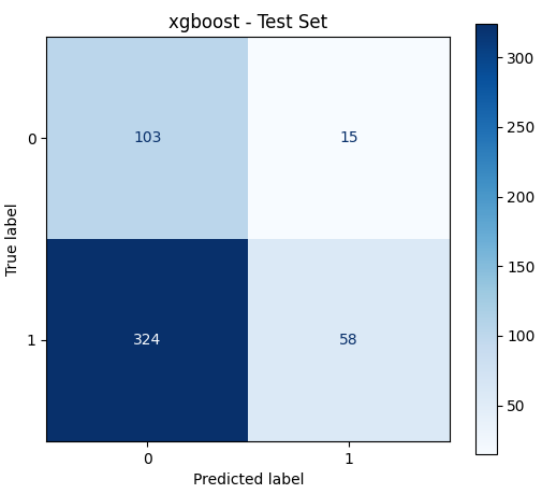
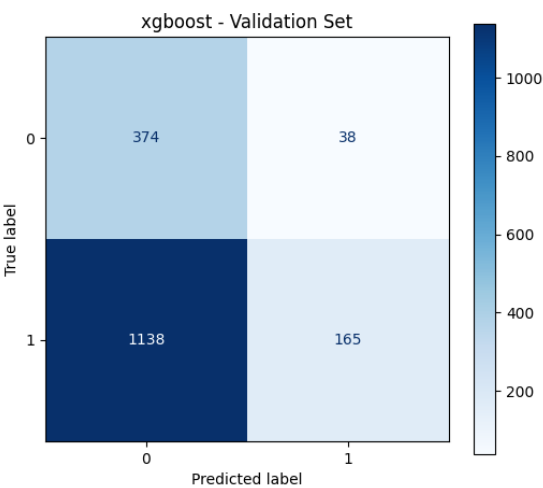
Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
abs_goal_diff_80	0.08402	goal_diff_squared_80	0.00000
league_england	0.07772	game_state_80_h	0.00000
league_avg_corners	0.06402	league_spain	0.00000
away_corners_75_80	0.06047	league_avg_home_corners	0.00000
game_state_80_d	0.05924	game_state_80_a	0.00000
odd_a	0.05630		
odd_h	0.05601		
away_corners_70_75	0.05584		

Optimal Threshold

- **Optimal threshold:** 0.81
- **Expected Precision:** 0.808
- **Expected Recall:** 0.129

Classification Reports



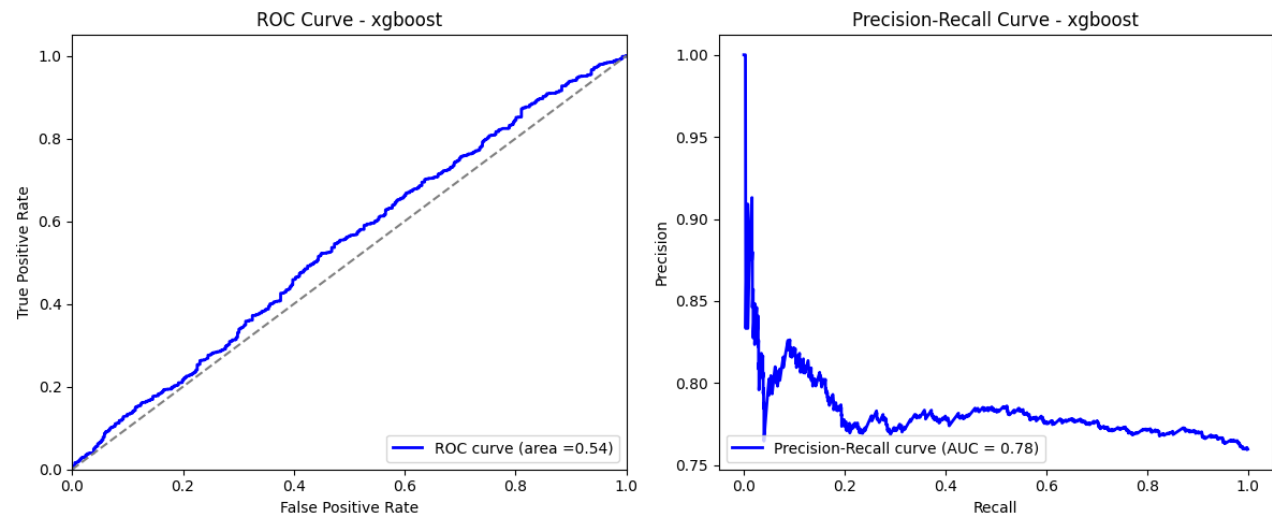
Classification Report (*Validation Set*)

	precision	recall	f1-score	support
0	0.25	0.91	0.39	412
1	0.81	0.13	0.22	1303
accuracy			0.31	1715
macro avg	0.53	0.52	0.30	1715
weighted avg	0.68	0.31	0.26	1715

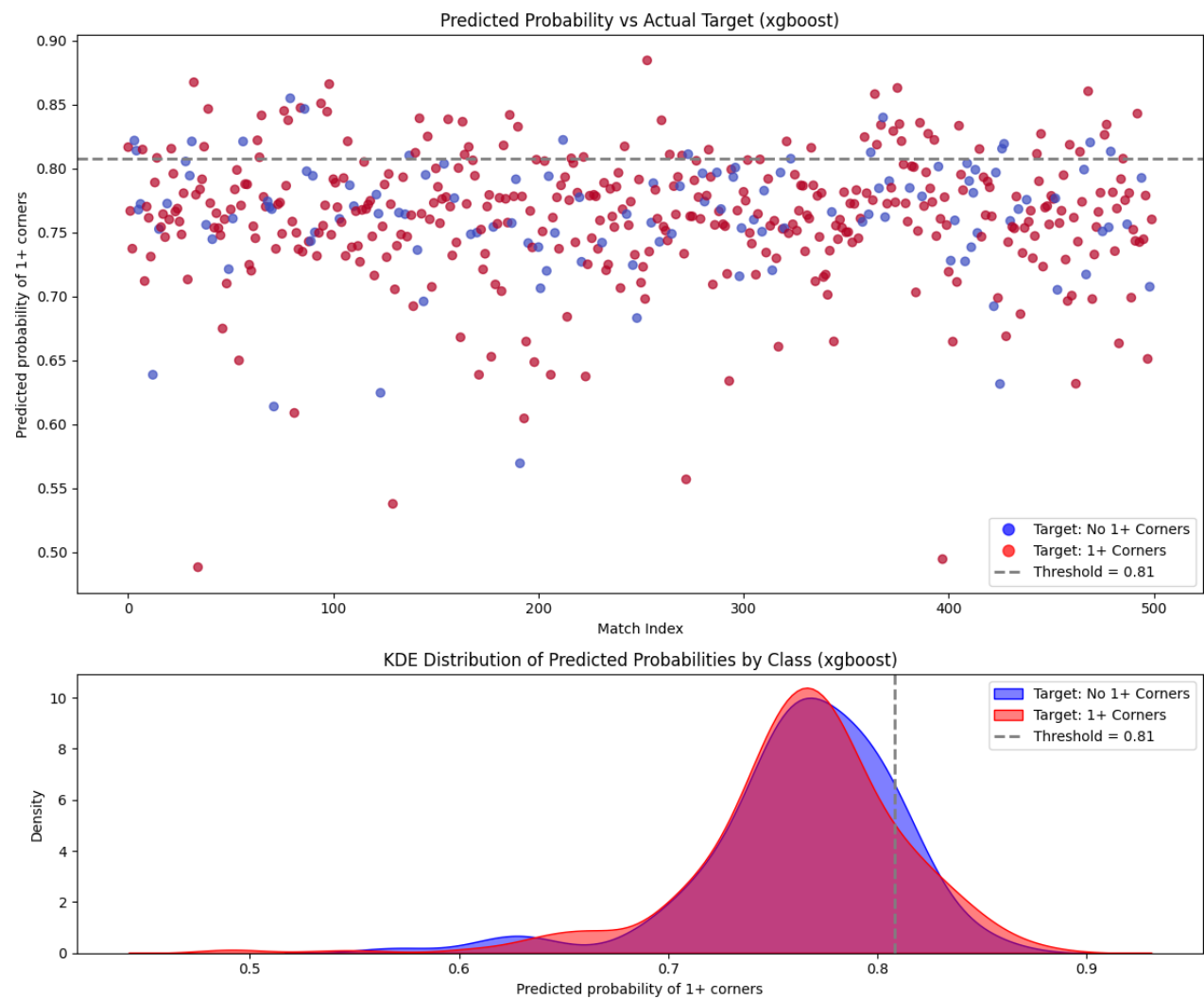
Classification Report (*Test Set*)

	precision	recall	f1-score	support
0	0.24	0.87	0.38	118
1	0.79	0.15	0.25	382
accuracy			0.32	500
macro avg	0.52	0.51	0.32	500
weighted avg	0.66	0.32	0.28	500

ROC and Precision-Recall Curves



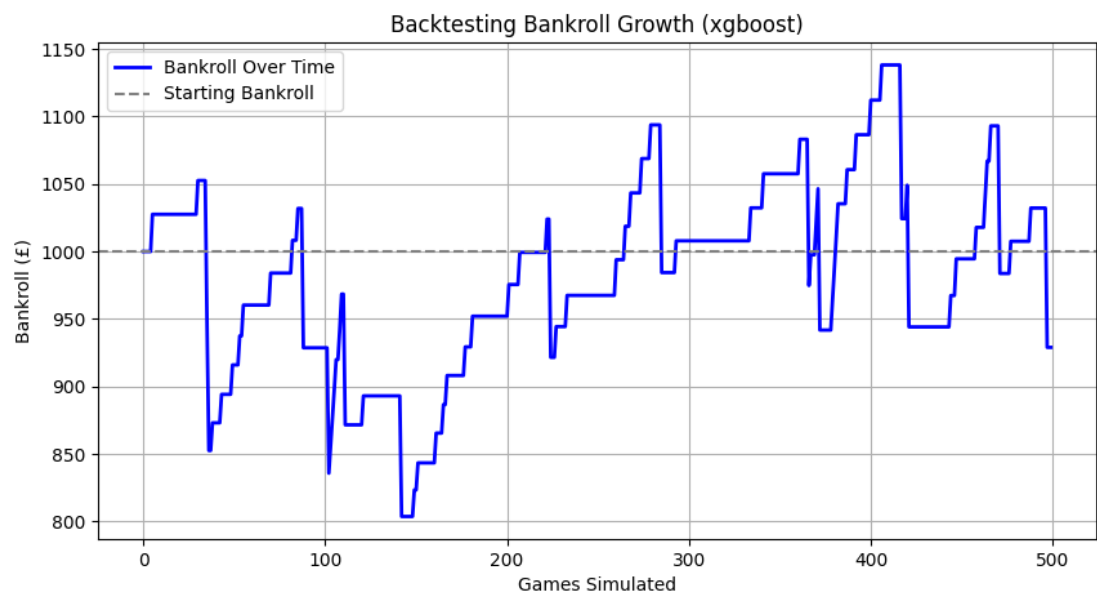
Prediction Distribution Graphs



Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £928.99
- 📊 Total Staked: £7030.50
- 📈 Total Profit: £-71.01
- 📊 ROI: -7.10%
- ✅ Win rate: 80.56% over 72 bets
- 📊 Sharpe Ratio: 0.01
- 🎯 Edge Over Bookies: 0.78%



Backtesting Results Summary

