

# Model Training Report: Track 2

---

## Configuration Details

---

### Target Variable

target

### Selected Features

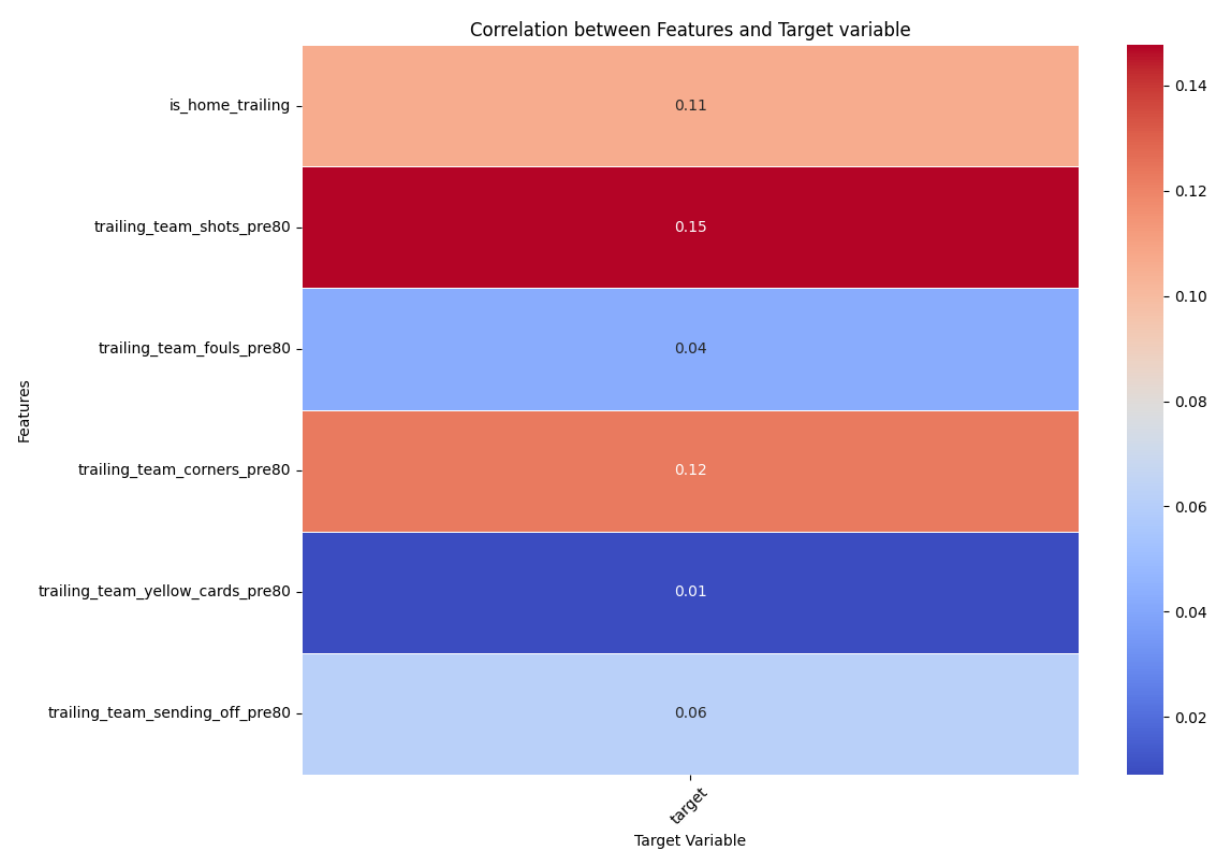
is\_home\_trailing, trailing\_team\_shots\_pre80, trailing\_team\_fouls\_pre80,  
trailing\_team\_corners\_pre80, trailing\_team\_yellow\_cards\_pre80,  
trailing\_team\_sending\_off\_pre80

### Constructed Features

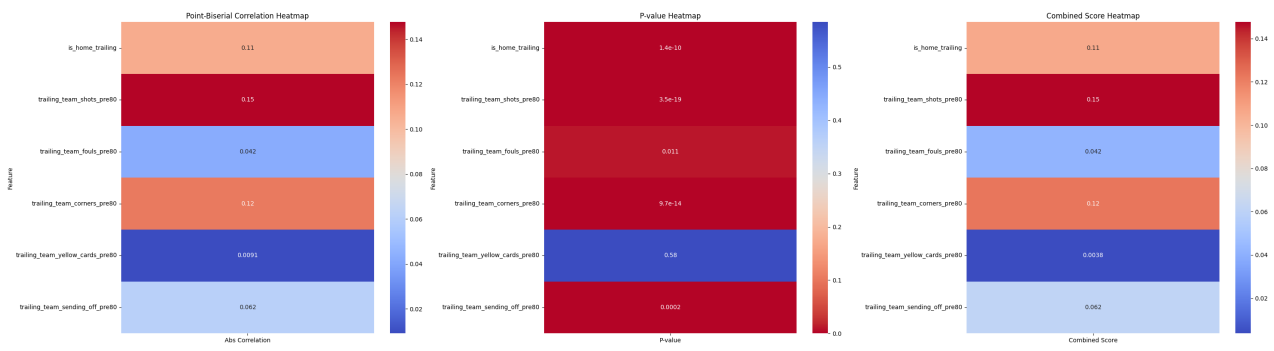
### Models Trained (& Hyperparameters)

- **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}

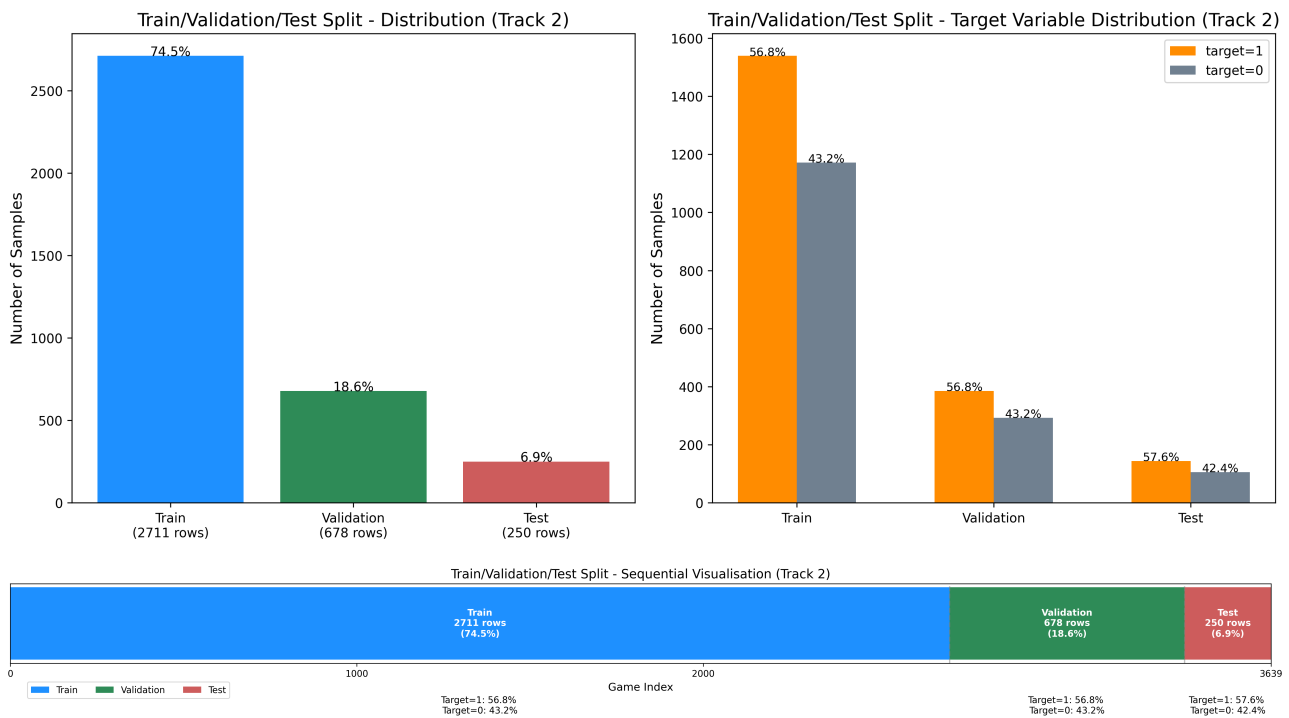
# Feature Correlation



# Point-Biserial Correlation



# Dataset Train/Validation/Test Split Visualisation



# random\_forest Model

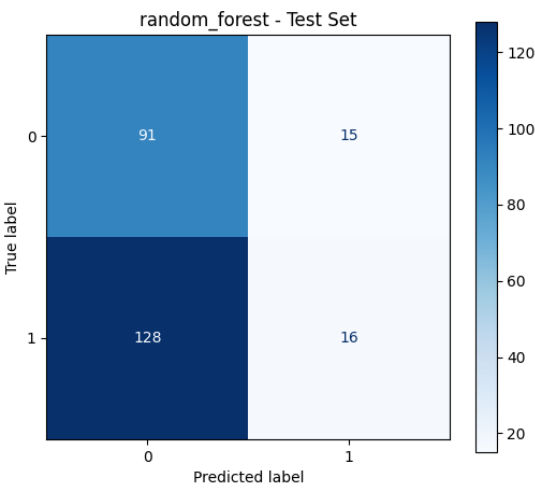
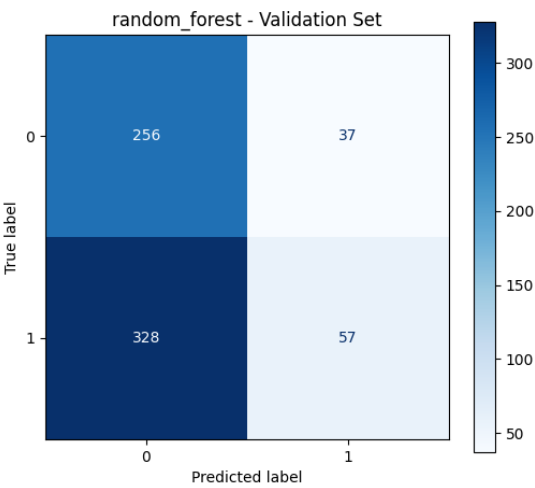
## Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_fouls_pre80	0.32111	trailing_team_sending_off_pre80	0.02282
trailing_team_shots_pre80	0.27024	is_home_trailing	0.02689
trailing_team_corners_pre80	0.21002	trailing_team_yellow_cards_pre80	0.14891
trailing_team_yellow_cards_pre80	0.14891	trailing_team_corners_pre80	0.21002
is_home_trailing	0.02689	trailing_team_shots_pre80	0.27024
trailing_team_sending_off_pre80	0.02282		
trailing_team_shots_pre80	0.27024		
trailing_team_corners_pre80	0.21002		

## Optimal Threshold

- **Optimal threshold:** 0.83
- **Expected Precision:** 0.606
- **Expected Recall:** 0.148

## Classification Reports



---

### Classification Report (Validation Set)

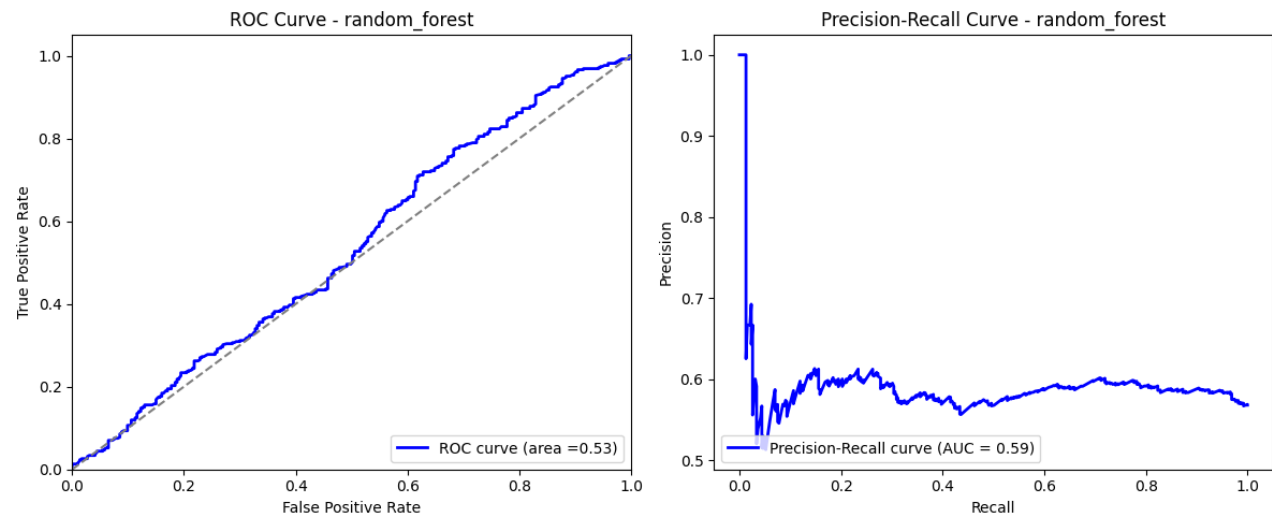
	precision	recall	f1-score	support
0	0.44	0.87	0.58	293
1	0.61	0.15	0.24	385
accuracy			0.46	678
macro avg	0.52	0.51	0.41	678
weighted avg	0.53	0.46	0.39	678

### Classification Report (Test Set)

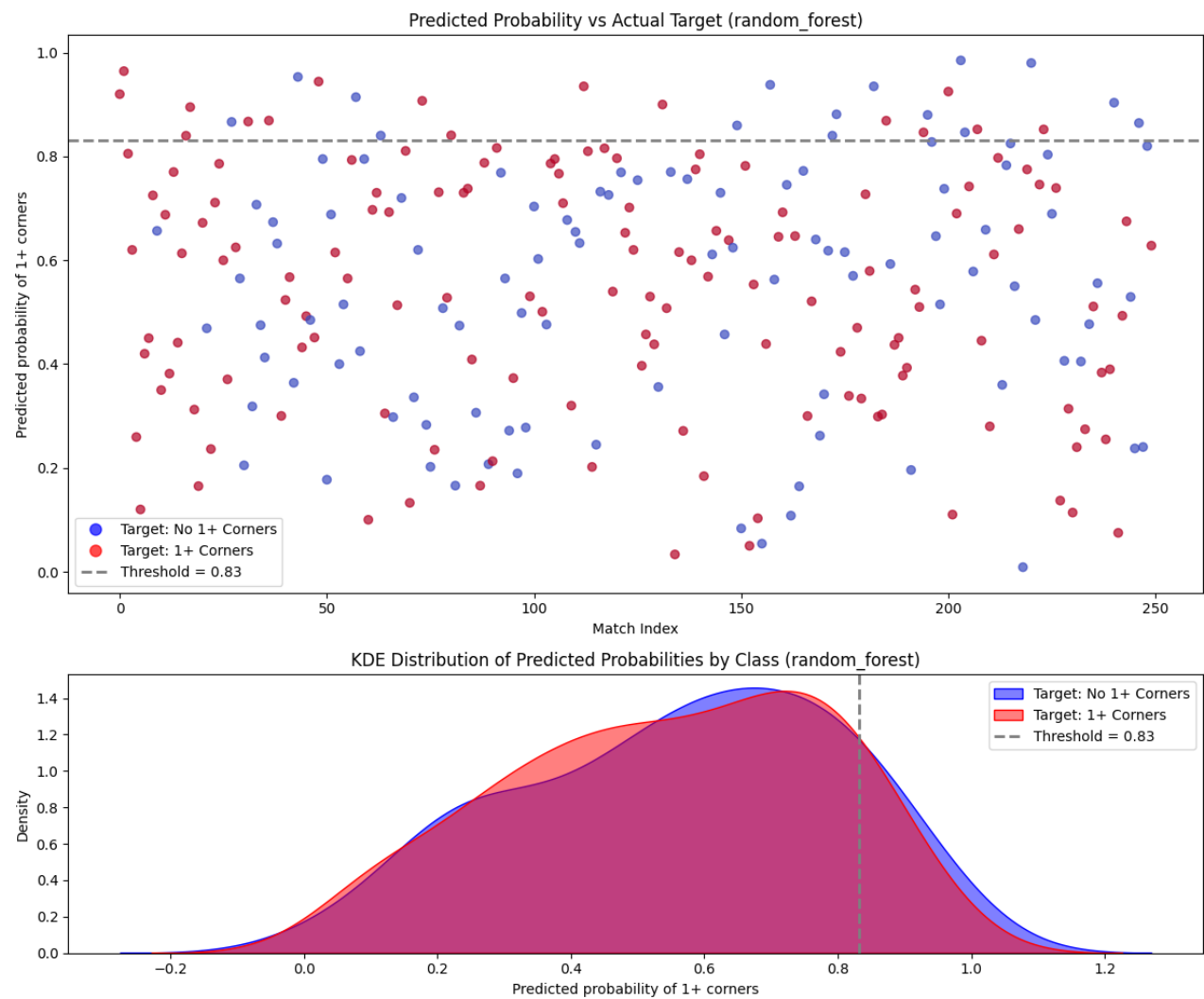
	precision	recall	f1-score	support
0	0.42	0.86	0.56	106
1	0.52	0.11	0.18	144
accuracy			0.43	250
macro avg	0.47	0.48	0.37	250
weighted avg	0.47	0.43	0.34	250

---

### ROC and Precision-Recall Curves



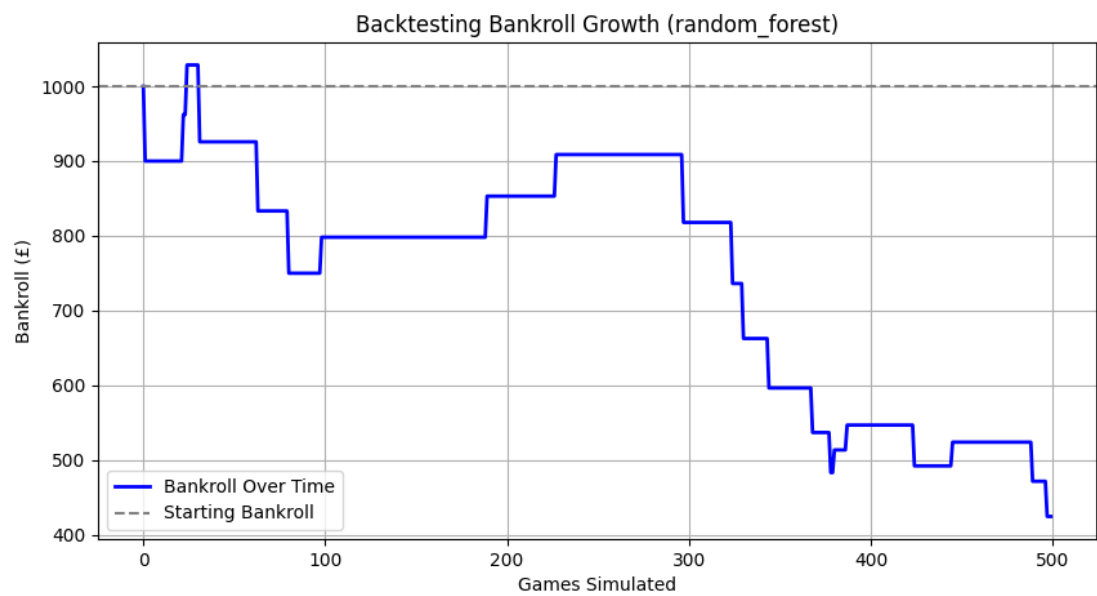
# Prediction Distribution Graphs



# Backtesting Results

--- Overall Summary ---

- 🏦 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £424.20
- 📊 Total Staked: £1533.85
- 📈 Total Profit: £-575.80
- 🇮🇹 ROI: -57.58%
- ✅ Win rate: 38.10% over 21 bets
- 📉 Sharpe Ratio: -0.44
- 🎯 Edge Over Bookies: -22.09%



# logistic\_regression Model

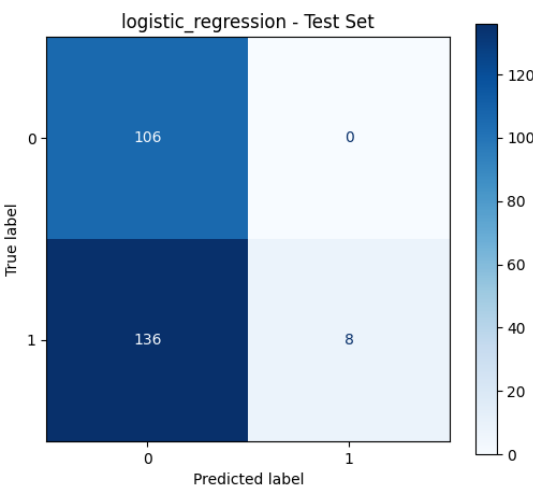
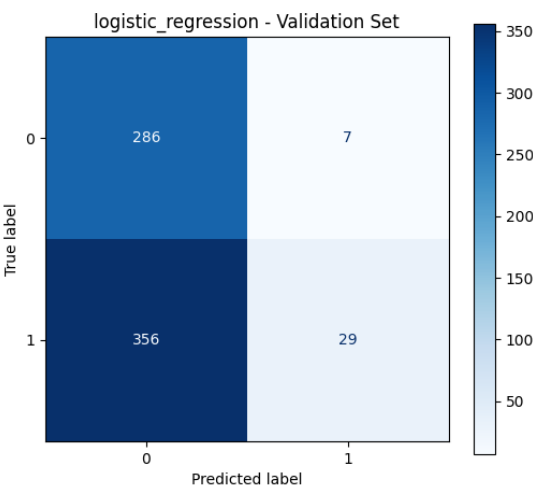
## Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	1.33703	trailing_team_sending_off_pre80	-0.84484
trailing_team_corners_pre80	0.62086	trailing_team_fouls_pre80	-0.50576
trailing_team_yellow_cards_pre80	0.34581	is_home_trailing	0.27212
is_home_trailing	0.27212	trailing_team_yellow_cards_pre80	0.34581
trailing_team_fouls_pre80	-0.50576	trailing_team_corners_pre80	0.62086
trailing_team_sending_off_pre80	-0.84484		
trailing_team_corners_pre80	0.62086		
trailing_team_yellow_cards_pre80	0.34581		

## Optimal Threshold

- **Optimal threshold:** 0.67
- **Expected Precision:** 0.806
- **Expected Recall:** 0.075

## Classification Reports





---

### Classification Report (Validation Set)

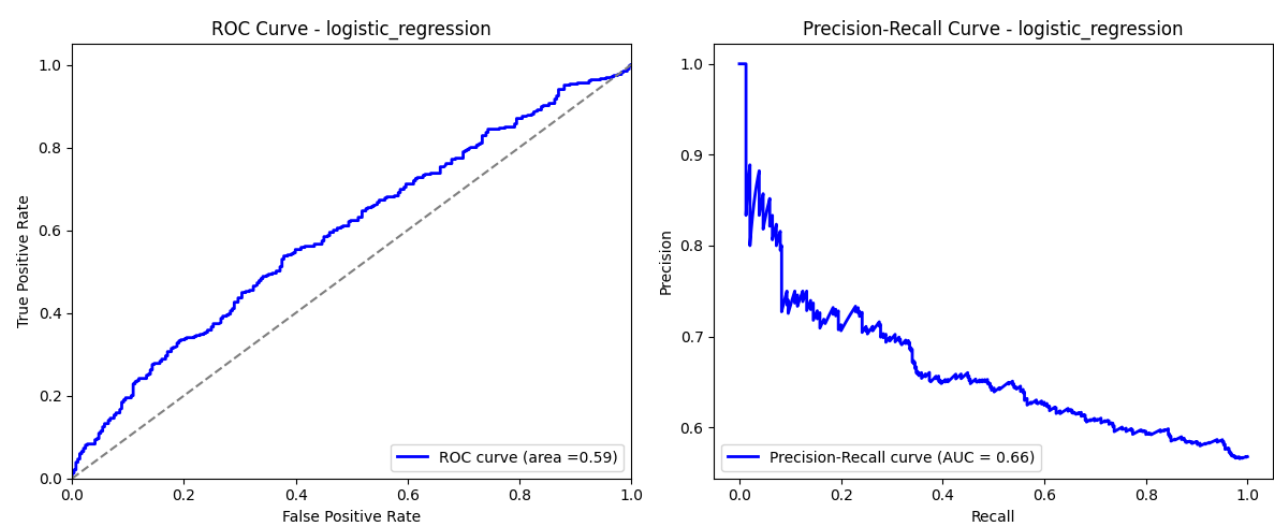
	precision	recall	f1-score	support
0	0.45	0.98	0.61	293
1	0.81	0.08	0.14	385
accuracy			0.46	678
macro avg	0.63	0.53	0.37	678
weighted avg	0.65	0.46	0.34	678

### Classification Report (Test Set)

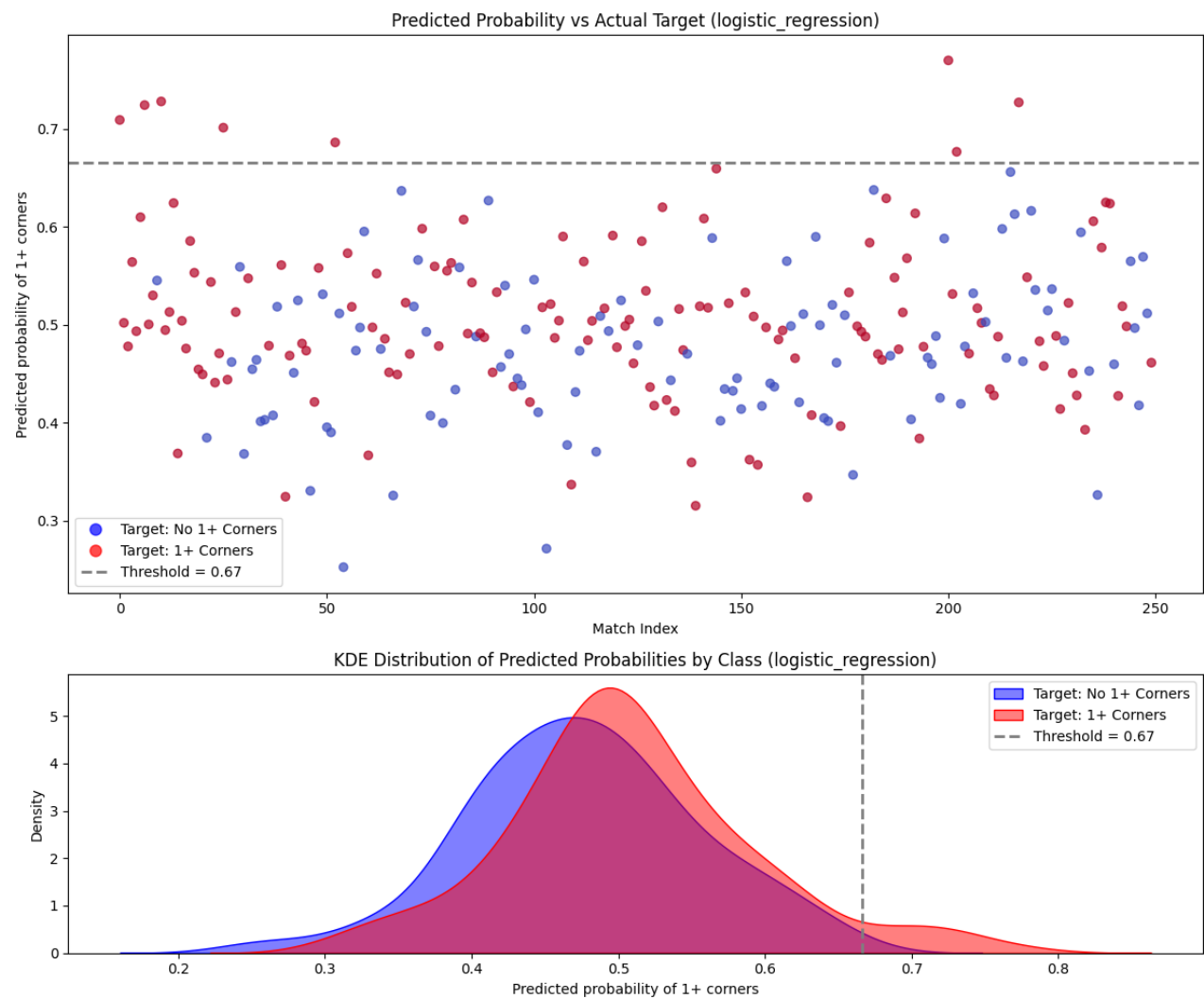
	precision	recall	f1-score	support
0	0.44	1.00	0.61	106
1	1.00	0.06	0.11	144
accuracy			0.46	250
macro avg	0.72	0.53	0.36	250
weighted avg	0.76	0.46	0.32	250

---

### ROC and Precision-Recall Curves



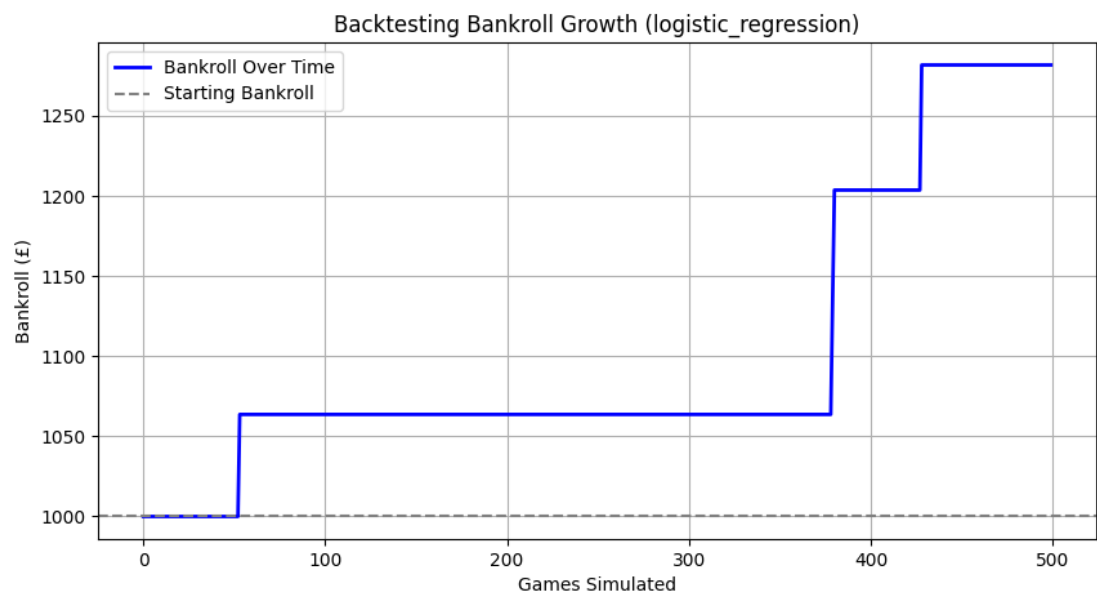
# Prediction Distribution Graphs



# Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £1281.78
- 📁 Total Staked: £440.00
- 📈 Total Profit: £281.78
- 🇬🇧 ROI: 28.18%
- ✅ Win rate: 100.00% over 4 bets
- 📊 Sharpe Ratio: 66.49
- 🎯 Edge Over Bookies: 39.81%



# svc Model

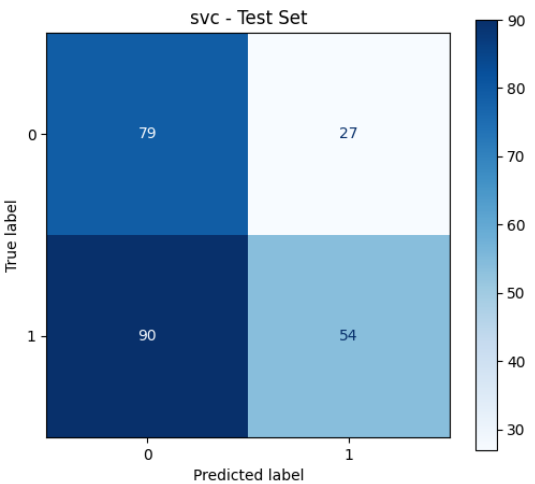
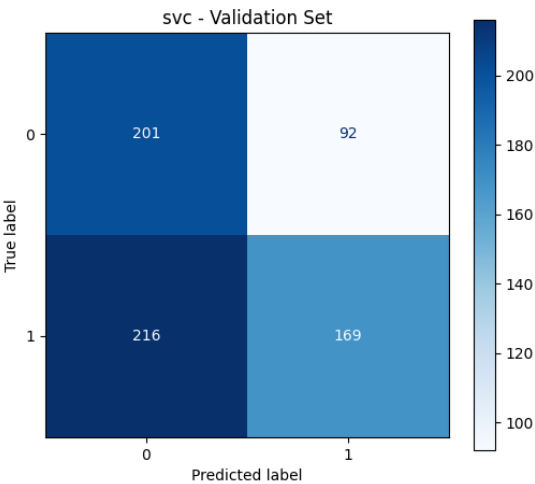
## Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
is_home_trailing	0.00000	trailing_team_sending_off_pre80	0.00000
trailing_team_shots_pre80	0.00000	trailing_team_yellow_cards_pre80	0.00000
trailing_team_fouls_pre80	0.00000	trailing_team_corners_pre80	0.00000
trailing_team_corners_pre80	0.00000	trailing_team_fouls_pre80	0.00000
trailing_team_yellow_cards_pre80	0.00000	trailing_team_shots_pre80	0.00000
trailing_team_sending_off_pre80	0.00000		
trailing_team_shots_pre80	0.00000		
trailing_team_fouls_pre80	0.00000		

## Optimal Threshold

- **Optimal threshold:** 0.62
- **Expected Precision:** 0.648
- **Expected Recall:** 0.439

## Classification Reports



---

### Classification Report (*Validation Set*)

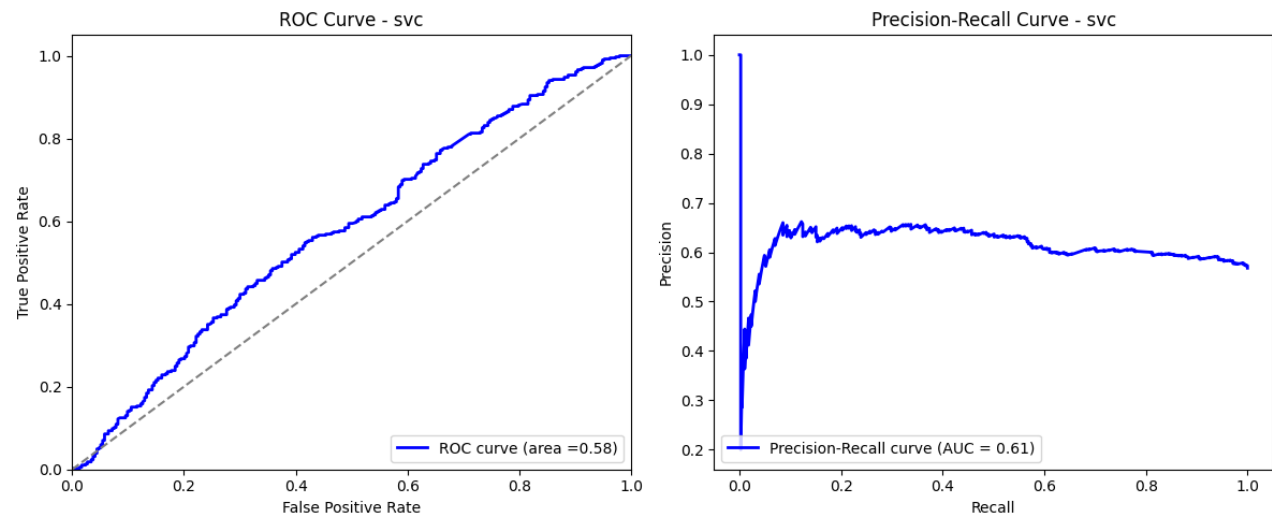
	precision	recall	f1-score	support
0	0.48	0.69	0.57	293
1	0.65	0.44	0.52	385
accuracy			0.55	678
macro avg	0.56	0.56	0.54	678
weighted avg	0.58	0.55	0.54	678

### Classification Report (*Test Set*)

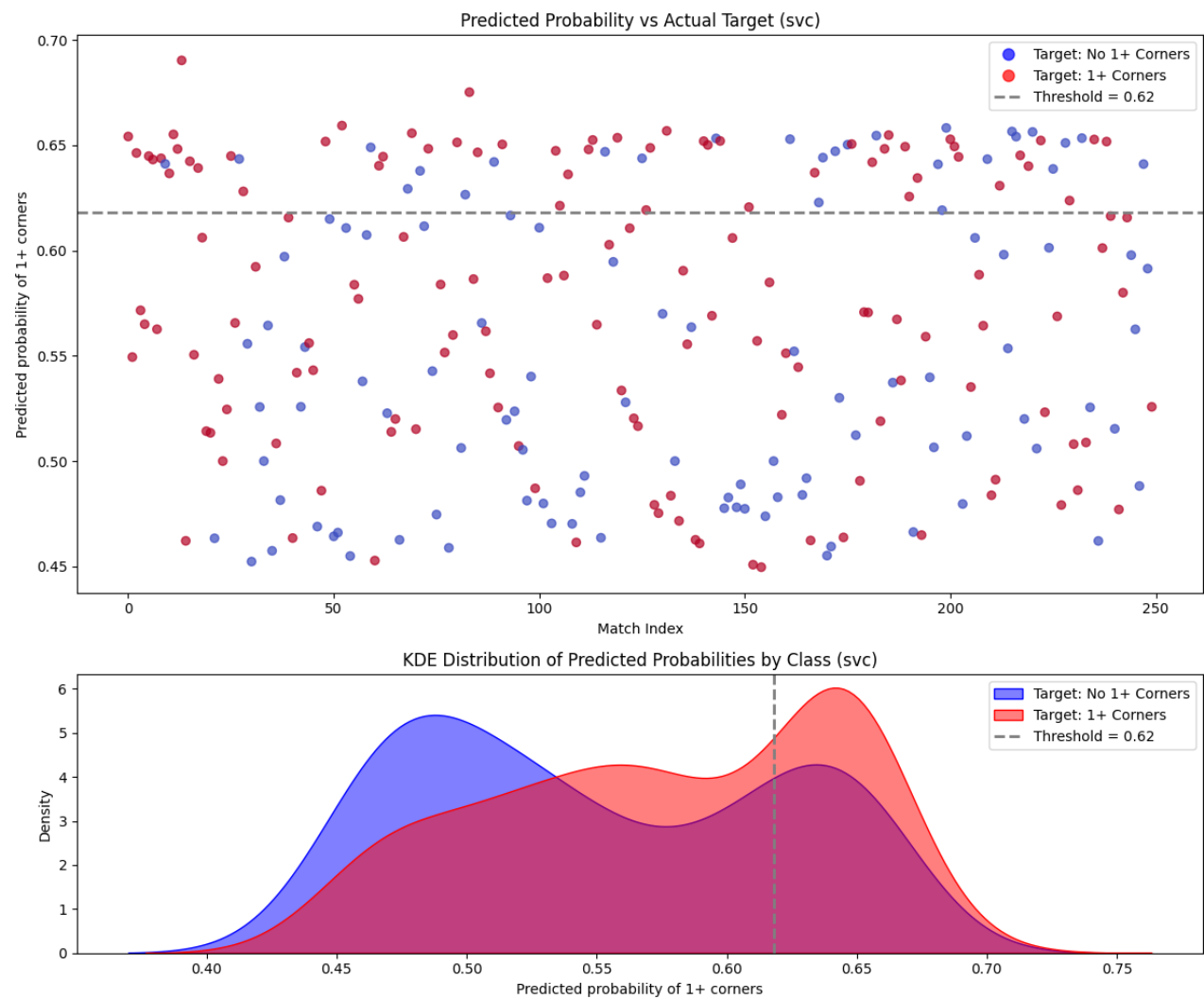
	precision	recall	f1-score	support
0	0.47	0.75	0.57	106
1	0.67	0.38	0.48	144
accuracy			0.53	250
macro avg	0.57	0.56	0.53	250
weighted avg	0.58	0.53	0.52	250

---

### ROC and Precision-Recall Curves



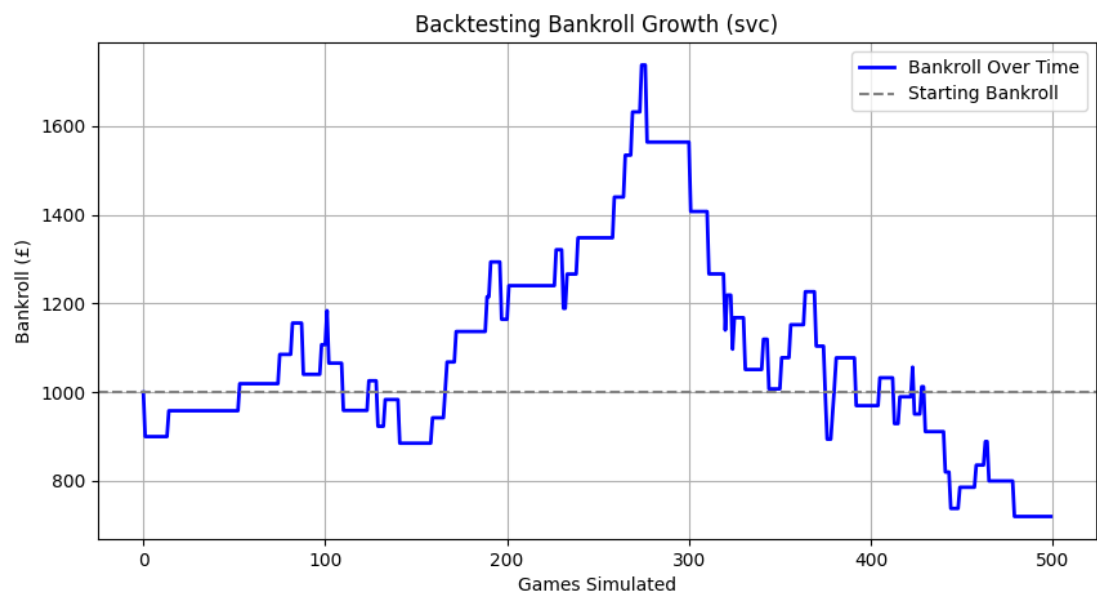
# Prediction Distribution Graphs



# Backtesting Results

--- Overall Summary ---

- 🏠 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £720.00
- 📊 Total Staked: £7006.75
- 📈 Total Profit: £-280.00
- 📉 ROI: -28.00%
- ✅ Win rate: 59.38% over 64 bets
- 📉 Sharpe Ratio: -0.02
- 🎯 Edge Over Bookies: -0.81%



# xgboost Model

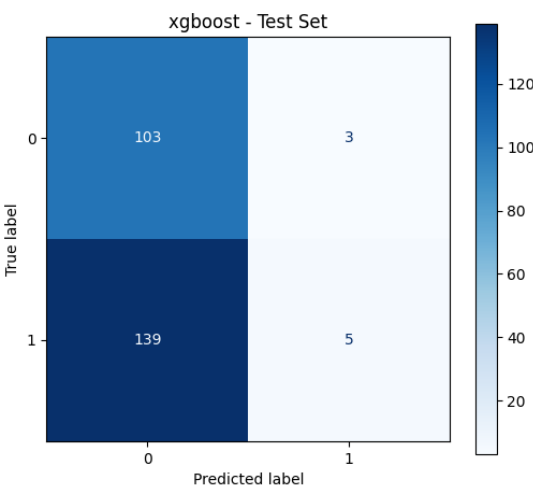
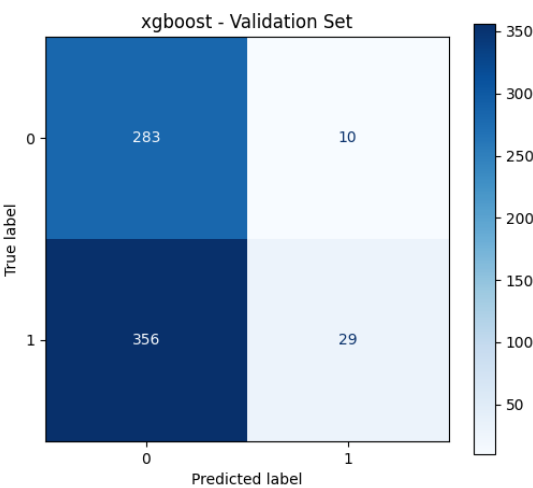
## Feature Importance

Top 8 Features	Importance	Bottom 5 Features	Importance
trailing_team_shots_pre80	0.22472	trailing_team_yellow_cards_pre80	0.12250
is_home_trailing	0.18292	trailing_team_fouls_pre80	0.14924
trailing_team_sending_off_pre80	0.16440	trailing_team_corners_pre80	0.15623
trailing_team_corners_pre80	0.15623	trailing_team_sending_off_pre80	0.16440
trailing_team_fouls_pre80	0.14924	is_home_trailing	0.18292
trailing_team_yellow_cards_pre80	0.12250		
is_home_trailing	0.18292		
trailing_team_sending_off_pre80	0.16440		

## Optimal Threshold

- **Optimal threshold:** 0.74
- **Expected Precision:** 0.744
- **Expected Recall:** 0.075

## Classification Reports





---

### Classification Report (*Validation Set*)

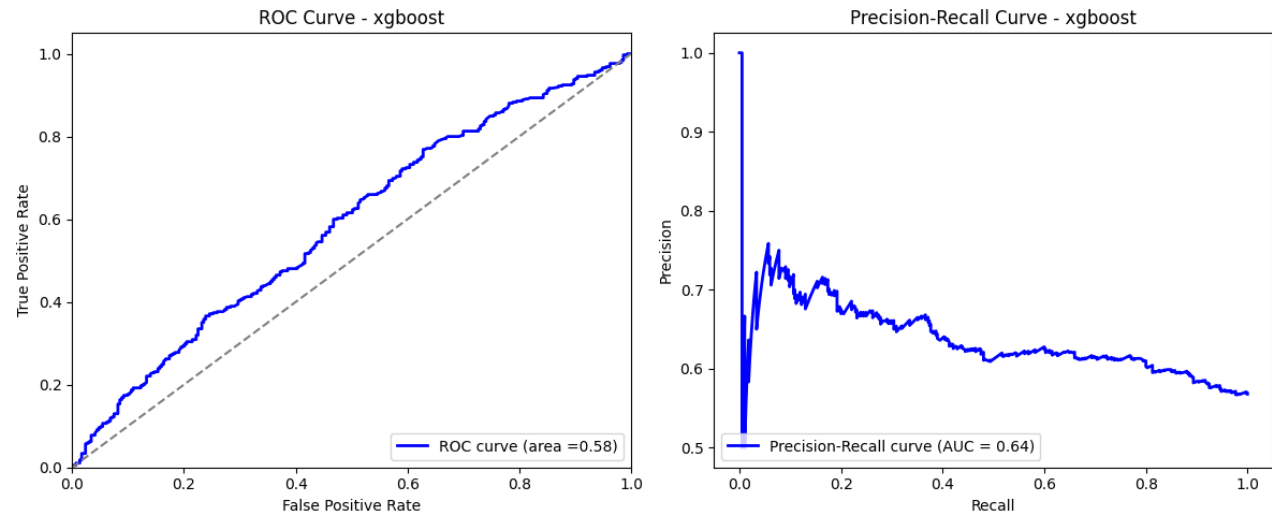
	precision	recall	f1-score	support
0	0.44	0.97	0.61	293
1	0.74	0.08	0.14	385
accuracy			0.46	678
macro avg	0.59	0.52	0.37	678
weighted avg	0.61	0.46	0.34	678

### Classification Report (*Test Set*)

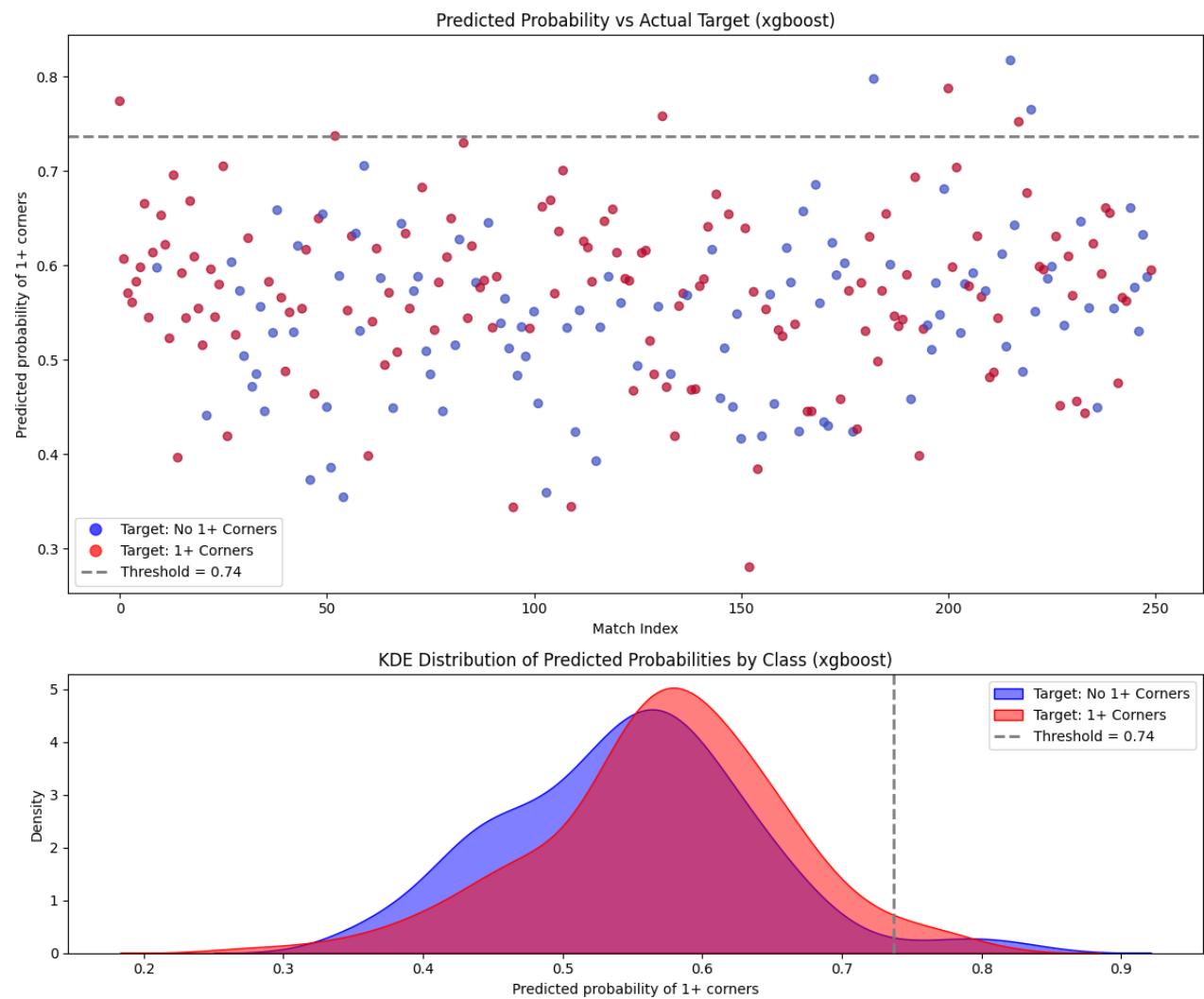
	precision	recall	f1-score	support
0	0.43	0.97	0.59	106
1	0.62	0.03	0.07	144
accuracy			0.43	250
macro avg	0.53	0.50	0.33	250
weighted avg	0.54	0.43	0.29	250

---

### ROC and Precision-Recall Curves



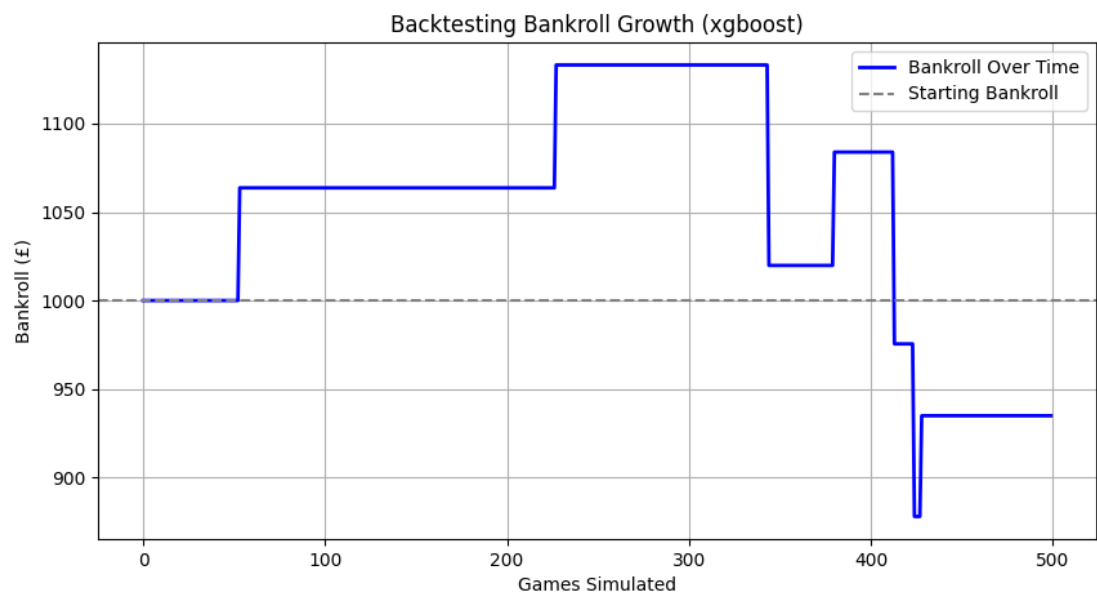
# Prediction Distribution Graphs



# Backtesting Results

--- Overall Summary ---

- 🏦 Initial Bankroll: £1000.00
- 💰 Final Bankroll: £934.89
- 📊 Total Staked: £715.39
- 📈 Total Profit: £-65.11
- 📊 ROI: -6.51%
- ✅ Win rate: 57.14% over 7 bets
- 📊 Sharpe Ratio: -0.07
- 🎯 Edge Over Bookies: -3.05%



# Backtesting Results Summary

