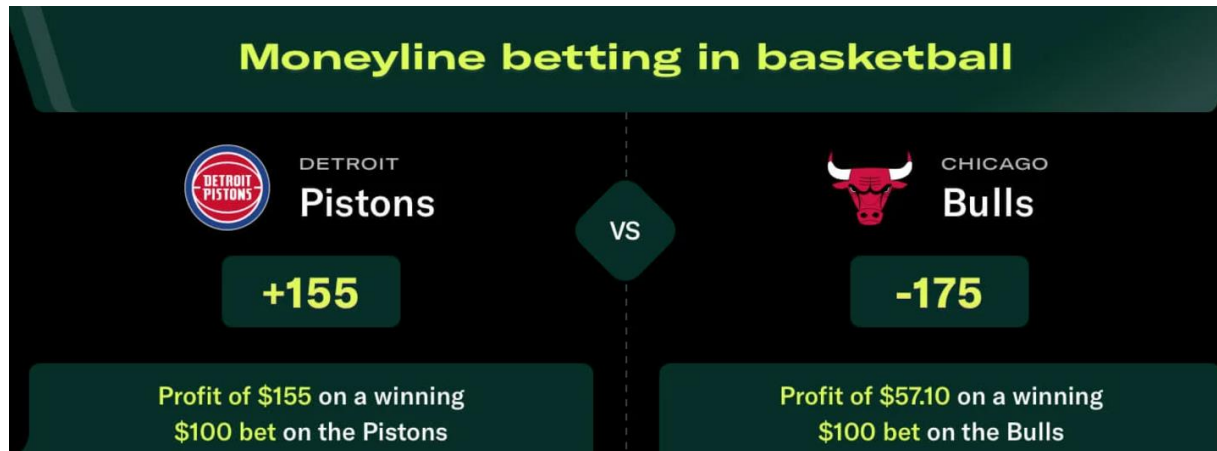# *MILPlays*

Mixed-Integer Linear Parlays
An Algorithmic Sports Betting Approach

Presented by:

Aidan Bagley

Introduction

- May 2018, SCOTUS struck down the "Professional and Amateur Sports Protection Act" [1]

- $426 billion has been wagered [2]

Example NBA Moneyline Wager   [3]

Positive Odds:
- Team is not favored to win
- You would <u>win more money</u> than you put in

Negative Odds:
- Team is favored to win
- You would <u>win less money</u> than you put in

- A professional sports a long-term winning percentage ~55% [4]
- How can optimization techniques be applied?

[1] https://www.congress.gov/bill/102nd-congress/senate-bill/474
[2] https://www.legalsportsreport.com/sports-betting/revenue/
[3] https://www.professionalgambler.org/winning-percentages
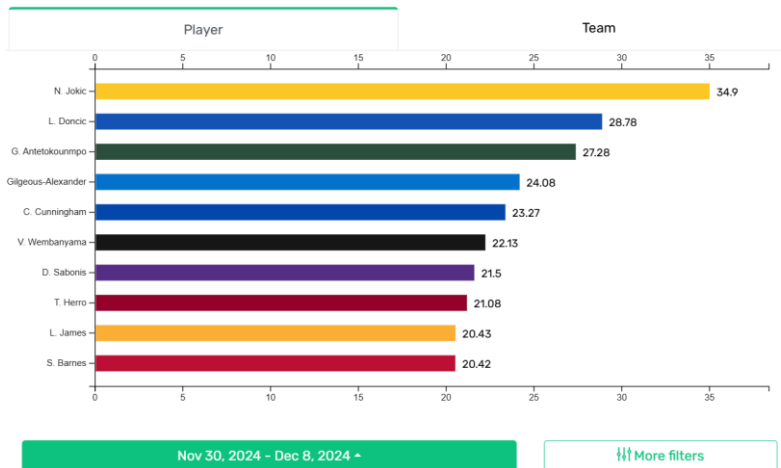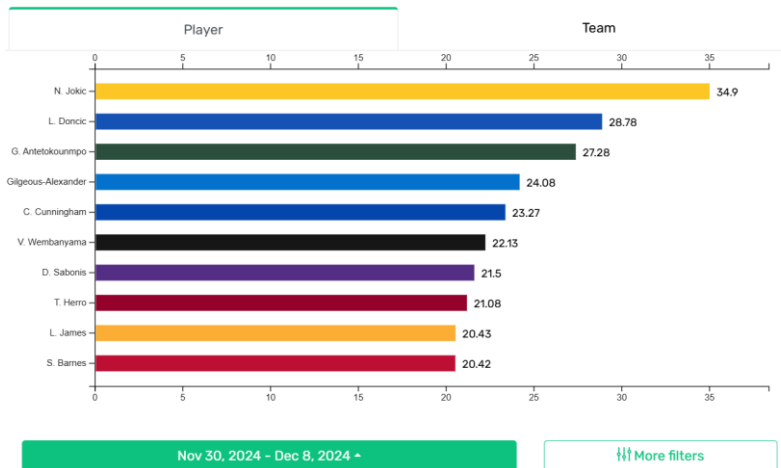[4] https://www.legalsportsreport.com/how-to-bet/moneyline/

Background and Related Work

- To increase expected returns for (NBA) sports wagering:

Data Science/Statistics

Advanced metrics:
- Game score
- Offensive/Defensive efficiency
- Defensive win shares
- Player efficiency rating

Player								Team
N. Jokic								34.9
L. Doncic							28.78	
G. Antetokounmpo						27.28		
Gilgeous-Alexander					24.08			
C. Cunningham					23.27			
V. Wembanyama				22.13				
D. Sabonis				21.5				
T. Herro				21.08				
L. James				20.43				
S. Barnes				20.42				

Nov 30, 2024 - Dec 8, 2024 ▲   ⇅ More filters

Player Average Game Score 2024
Source: https://viziball.app/nba/en

Background and Related Work

- To increase expected returns for (NBA) sports wagering:

Data Science/Statistics	Predictive Modeling

Advanced metrics:
- Game score
- Offensive/Defensive efficiency
- Defensive win shares
- Player efficiency rating

Classification Methods
- Logistic Regression
- Ensemble classifiers
- Multi-layer perceptron classifiers

Player Average Game Score 2024
Source: https://viziball.app/nba/en

Multi-Layer Perceptron Architecture
Source:
https://spotintelligence.com/2024/02
/20/multilayer-perceptron-mlp/

Background and Related Work

- To increase expected returns for (NBA) sports wagering:

Data Science/Statistics

Advanced metrics:
- Game score
- Offensive/Defensive efficiency
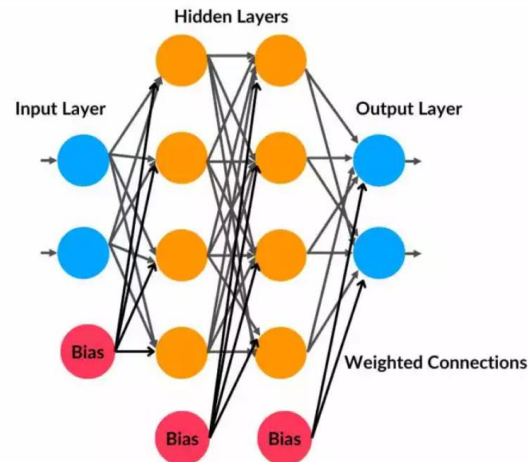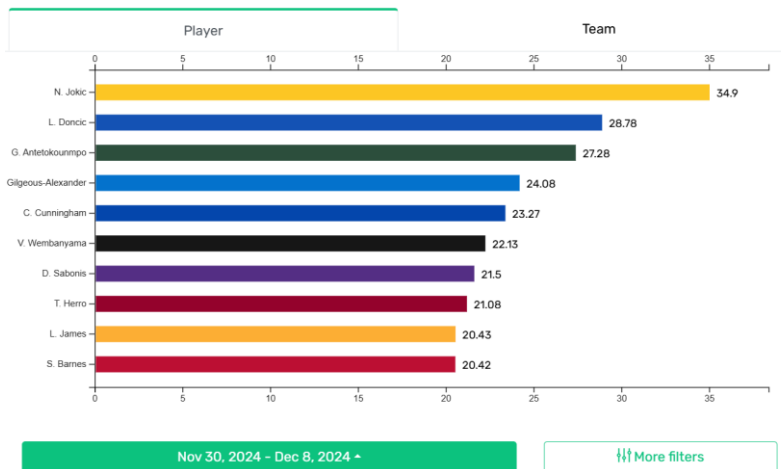- Defensive win shares
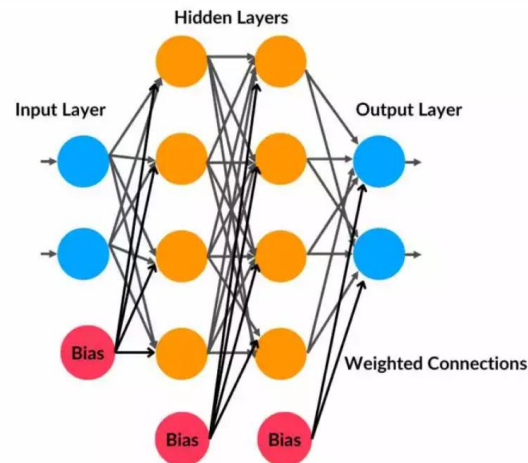- Player efficiency rating

Player Average Game Score 2024
Source: https://viziball.app/nba/en

Predictive Modeling

Classification Methods
- Logistic Regression
- Ensemble classifiers
- Multi-layer perceptron classifiers

Multi-Layer Perceptron Architecture
Source:
https://spotintelligence.com/2024/02/20/multilayer-perceptron-mlp/

Anomaly Identification

Are the odds mispriced?

"By developing models that can accurately predict match outcomes and compare them with the odds offered by bookmakers, bettors can identify instances where the odds are mispriced, allowing them to place bets with a positive expected value."

Galekwa, René & Tshimula, Jean & Tajeuna, Etienne & Kyamakya, Kyandoghere. (2024). A Systematic Review of Machine Learning in Sports Betting: Techniques, Challenges, and Future Directions. 10.48550/arXiv.2410.21484.

Problem Formulation: Probability of Win

- Consider observing either a win or loss on game i: $y_i \in \{0,1\}$

$\vec{x}$	Feature
x_1	Home Game Win Percentage
x_2	Away Game Win Percentage
x_3	Total Win Percentage
x_4	Offensive Efficiency
x_5	Rolling Offensive Efficiency
x_6	Rolling Scoring Margin
x_7	Number of Rest Days

1. A probability of *home team A* winning game: $\hat{p}_i \in \{0,1\} \quad f(\hat{p}_i) = \begin{cases} \text{w}in, & \hat{p}_i > 0.5 \\ \text{loss}, & \hat{p}_i \leq 0.5 \end{cases}$

2. A characteristic feature set and weighting vector: $\vec{x}, \vec{\theta} \in \mathcal{R}^n$

TABLE I: Data Features

$$L(\vec{x}_i, \vec{\theta}) = P(y_i | \vec{x}_i, \vec{\theta}) \qquad (1) \quad \text{Likelihood function}$$

$$P(y_i | \vec{x}_i, \vec{\theta}) = \sigma(\vec{x}_i, \vec{\theta}) = \frac{1}{1 + \exp(-\vec{x}_i^T \vec{\theta})} \qquad (2) \quad \text{Sigmoid function}$$

$$\max_{\theta}(L(\vec{x}, \vec{\theta})) = \max_{\theta}\left(\sum_{i=1}^{n} P(y_i | \vec{x}_i, \vec{\theta})^{y_i} \cdot P(y_i | \vec{x}_i, \vec{\theta})^{1-y_i}\right)$$

(3) Max likelihood function

$$\max_{\theta}(\mathcal{L}(\vec{x}, \vec{\theta})) = \max_{\theta}\left(\sum_{i=1}^{n} y_i \log \sigma(\vec{x}_i, \vec{\theta}) + \right.$$

$$\left. (1 - y_i) \log(1 - \sigma(\vec{x}_i, \vec{\theta}))\right) \qquad (4) \quad \text{Max log-likelihood function}$$

$$\theta^* \leftarrow \arg\min_{\theta}\left(-\sum_{i=1}^{n} y_i \log \sigma(\vec{x}_i, \vec{\theta}) + \right.$$

$$\left. (1 - y_i) \log(1 - \sigma(\vec{x}_i, \vec{\theta}))\right) \qquad (5) \quad \text{Argmax negative log-likelihood function}$$

$$\hat{p}_i \leftarrow \sigma(\vec{x}_i, \vec{\theta}^*) \qquad (6) \quad \text{Optimal probability of win}$$

Problem Formulation: Decision Making

- Consider observing either a win or loss on game i: $y_i \in \{0,1\}$

1. A probability of *home team A* winning game: $\hat{p}_i \in \{0,1\}$ $f(\hat{p}_i) = \begin{cases} \text{win}, & \hat{p}_i > 0.5 \\ \text{loss}, & \hat{p}_i \leq 0.5 \end{cases}$ $\hat{p}_i \leftarrow \sigma(\vec{x}_i, \vec{\theta}^*)$ (6)

2. A characteristic feature set and weighting vector: $\vec{x}, \vec{\theta} \in \mathcal{R}^n$

3. Sports book money lines for *team A and team B*: $M_i \in \{[-\infty, -100], [100, \infty]\}$

4. Wager amount b_i: $b_i \geq 0$

$$W = \arg\max_{b_i, k_i} \quad (\mathrm{E}[\vec{b}]^T \vec{k})$$

$$R_i(M_i) = \begin{cases} \frac{M_i}{100}, & \textbf{if } M_i > 0 \\ \frac{100}{|M_i|}, & \textbf{if } M_i < 0 \end{cases}$$ (7) Moneyline multiplier

$$P_i(b_i, M_i) = b_i \cdot R_i(M_i)$$ (8) Profit function

$$\mathrm{E}[b_i] = \hat{p}_i \cdot P_i(b_i, M_i) - (1 - \hat{p}_i) \cdot b_i$$ (9) Weighted expected value of profit

$$\text{subject to} \quad \sum_{i=1}^{n} k_i = 5$$
$$\sum_{i=1}^{n} b_i k_i \leq B \qquad (10)$$
$$b_i \geq 0, \quad \forall i = 1, \ldots, n$$
$$k_i \in \{0, 1\}, \quad \forall i = 1, \ldots, n$$

Non-Linear Program maximizing expected profit

Problem Formulation: Decision Making

- Consider observing either a win or loss on game i: $\quad y_i \in \{0,1\}$

1. A probability of *home team A* winning game: $\quad \hat{p}_i \in \{0,1\} \quad f(\hat{p}_i) = \begin{cases} \text{win}, & \hat{p}_i > 0.5 \\ \text{loss}, & \hat{p}_i \leq 0.5 \end{cases} \qquad \hat{p}_i \leftarrow \sigma(\vec{x}_i, \vec{\theta}^*) \qquad (6)$

2. A characteristic feature set and weighting vector: $\quad \vec{x}, \vec{\theta} \in \mathcal{R}^n$

3. Sports book money lines for *team A and team B*: $\quad M_i \in \{[-\infty, -100], [100, \infty]\}$

4. Wager amount b_i: $\quad b_i \geq 0$

$$R_i(M_i) = \begin{cases} \frac{M_i}{100}, & \textbf{if } M_i > 0 \\ \frac{100}{|M_i|}, & \textbf{if } M_i < 0 \end{cases}$$

$\qquad (7) \quad$ Moneyline multiplier

$$P_i(b_i, M_i) = b_i \cdot R_i(M_i) \qquad (8) \quad \text{Profit function}$$

$$\mathrm{E}[b_i] = \hat{p}_i \cdot P_i(b_i, M_i) - (1 - \hat{p}_i) \cdot b_i \qquad (9) \quad \text{Weighted expected value of profit}$$

$$z_i = b_i \cdot k_i \qquad (11) \quad \text{Slack variable}$$

$$W = \arg \max_{z_i, b_i, k_i} \sum_{i=1}^{n} \mathrm{E}[\vec{z}] \, \vec{1}^T$$

$$\text{subject to} \quad \sum_{i=1}^{n} k_i = 5$$

$$\sum_{i=1}^{n} z_i \leq B \qquad (12)$$

$$z_i \leq b_i, \quad \forall i = 1, \ldots, n$$
$$z_i \leq B \cdot k_i, \quad \forall i = 1, \ldots, n$$
$$z_i \geq 0, \quad \forall i = 1, \ldots, n$$
$$b_i \geq 0, \quad \forall i = 1, \ldots, n$$
$$k_i \in \{0, 1\}, \forall i = 1, \ldots, n$$

Mixed Integer Linear Program maximizing expected profit

Solution Approach:

nba_api

An API Client Package to Access the APIs of NBA.com

`nba_api` is an API Client for `www.nba.com`. This package intends to make the APIs of NBA.com easily accessible and provide extensive documentation about them.

- Features:
  - The NBA_API provides historical game by game data

Example Data Features

$\vec{x}$	Feature
x_1	Home Game Win Percentage
x_2	Away Game Win Percentage
x_3	Total Win Percentage
x_4	Offensive Efficiency
x_5	Rolling Offensive Efficiency
x_6	Rolling Scoring Margin
x_7	Number of Rest Days

TABLE I: Data Features

- Model Evaluation: Logistic Regression vs Multi-Layer Perceptron
  - Accuracy
  - Classification report: precision, recall, F1-Score
  - Cross entropy loss
  - (Area Under) Receiver Operating Characteristic Curve

Solution Approach:

- Features:
  - The NBA_API provides historical game by game data


```
For Season = 2020-21, Game ID = 22001066 -> 2021-05-16, Home Team ID = 1610612737 -> Atlanta Hawks, Away Team ID = 1610612745 -> Houston Rockets

  Input:

  HOME_LAST_GAME_OE      ,HOME_LAST_GAME_HOME_WIN_PCTG ,HOME_NUM_REST_DAYS     ,HOME_LAST_GAME_AWAY_WIN_PCTG ,HOME_LAST_GAME_TOTAL_WIN_PCTG ,HOME_LAST_GAME_RO
  0.5555                 ,0.6857                       ,3.0                    ,0.4444                       ,0.5633                        ,9.3333

  AWAY_LAST_GAME_OE      ,AWAY_LAST_GAME_HOME_WIN_PCTG ,AWAY_NUM_REST_DAYS     ,AWAY_LAST_GAME_AWAY_WIN_PCTG ,AWAY_LAST_GAME_TOTAL_WIN_PCTG ,AWAY_LAST_GAME_RO
  0.6129                 ,0.25                         ,2.0                    ,0.2285                       ,0.2394                        ,-2.0

  Output:
  HOME_W
  1
```

Example Data Features

- Model Evaluation: Logistic Regression vs Multi-Layer Percep
  - Accuracy
  - Classification report: precision, recall, F1-Score
  - Cross entropy loss
  - (Area Under) Receiver Operating Characteristic Curve

- Wagering
  - Using best model compare expected vs actual return

DATE	AWAY	HOME	AWAY_MONEYLINE	HOME_MONEYLINE
2024-11-27	Hawks	Cavaliers	350	-455
2024-11-27	Bulls	Magic	320	-410
2024-11-27	Trail Blazers	Pacers	360	-470
2024-11-27	Clippers	Wizards	-470	360
2024-11-27	Rockets	76ers	-218	180
2024-11-27	Heat	Hornets	-175	145
2024-11-27	Knicks	Mavericks	-170	142
2024-11-27	Kings	Timberwolves	130	-155
2024-11-27	Pistons	Grizzlies	275	-345
2024-11-27	Raptors	Pelicans	130	-155
2024-11-27	Lakers	Spurs	-125	105
2024-11-27	Nuggets	Jazz	-535	400
2024-11-27	Nets	Suns	320	-410
2024-11-27	Thunder	Warriors	-180	150
2024-12-02	Heat	Celtics	575	-950
2024-12-02	Pelicans	Hawks	325	-540
2024-12-02	Lakers	Timberwolves	230	-285
2024-12-02	Nets	Bulls	250	-310
2024-12-02	Bucks	Pistons	-162	126
2024-12-02	76ers	Hornets	-205	170
2024-12-02	Wizards	Cavaliers	950	-1650
2024-12-02	Magic	Knicks	190	-230
2024-12-02	Pacers	Raptors	-130	110
2024-12-02	Jazz	Thunder	750	-1200
2024-12-02	Grizzlies	Mavericks	150	-180
2024-12-02	Spurs	Suns	205	-250
2024-12-02	Warriors	Nuggets	170	-205
2024-12-02	Rockets	Kings	-130	110
2024-12-02	Trail Blazers	Clippers	310	-395
2024-12-05	Mavericks	Wizards	-750	525
2024-12-05	Nuggets	Cavaliers	140	-166
2024-12-05	Hornets	Knicks	850	-1450
2024-12-05	Thunder	Raptors	-410	320
2024-12-05	Bulls	Spurs	110	-130
2024-12-05	Suns	Pelicans	-115	-105
2024-12-05	Kings	Grizzlies	150	-180
2024-12-05	Rockets	Warriors	-170	142
2024-12-05	Magic	Bucks	170	-205
2024-12-05	Mavericks	Thunder	170	-205

- Predicted home wins were correct 65%
- 77% of home wins were identified

- Predicted home losses were correct 63%
- 48% of home losses were identified

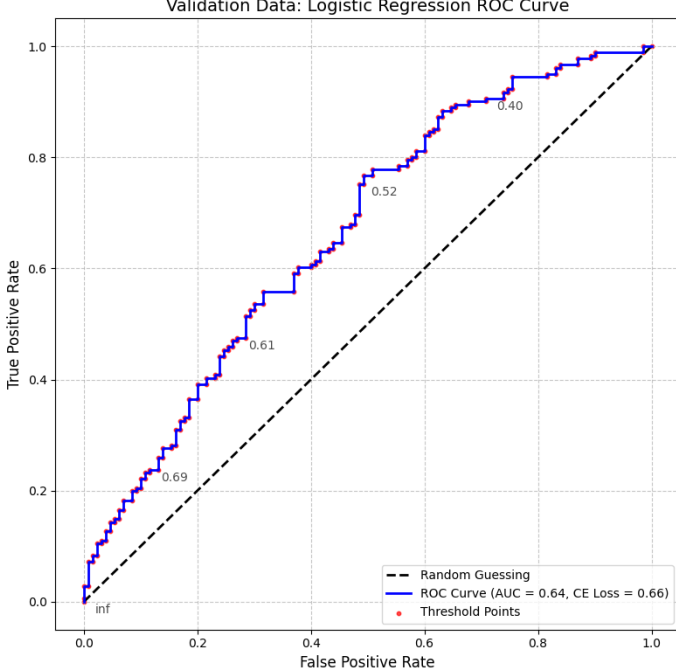TABLE II: Training Data: Logistic Regression Report

| Class	Precision	Recall	F1-Score	Support
0	0.63	0.48	0.54	511
1	0.65	0.77	0.70	636
Weighted Average	0.64	0.64	0.63	1147
Additional Metrics:				
Accuracy			0.64	
Cross-Entropy Loss			0.64	
AUC-ROC			0.66	
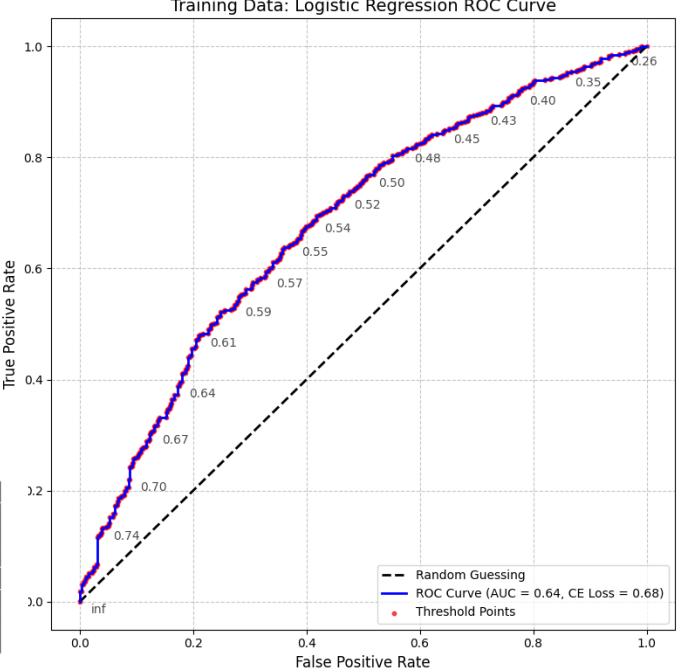
- Predicted home wins were correct 68%
- 77% of home wins were identified

- Predicted home losses were correct 61%
- 49% of home losses were identified

TABLE III: Validation Data: Logistic Regression Report

Class	Precision	Recall	F1-Score	Support
0	0.61	0.49	0.54	130
1	0.68	0.77	0.72	181
Weighted Average	0.65	0.66	0.65	311
Additional Metrics:				
Accuracy			0.66	
Cross-Entropy Loss			0.64	
AUC-ROC			0.66	
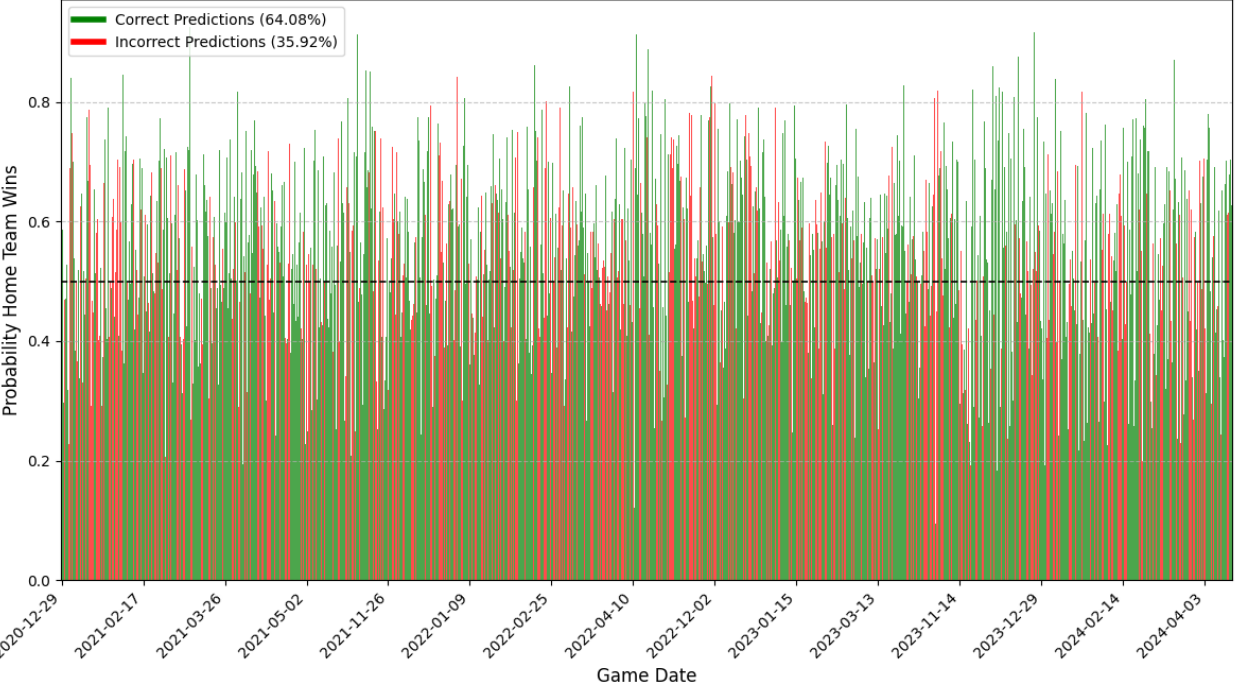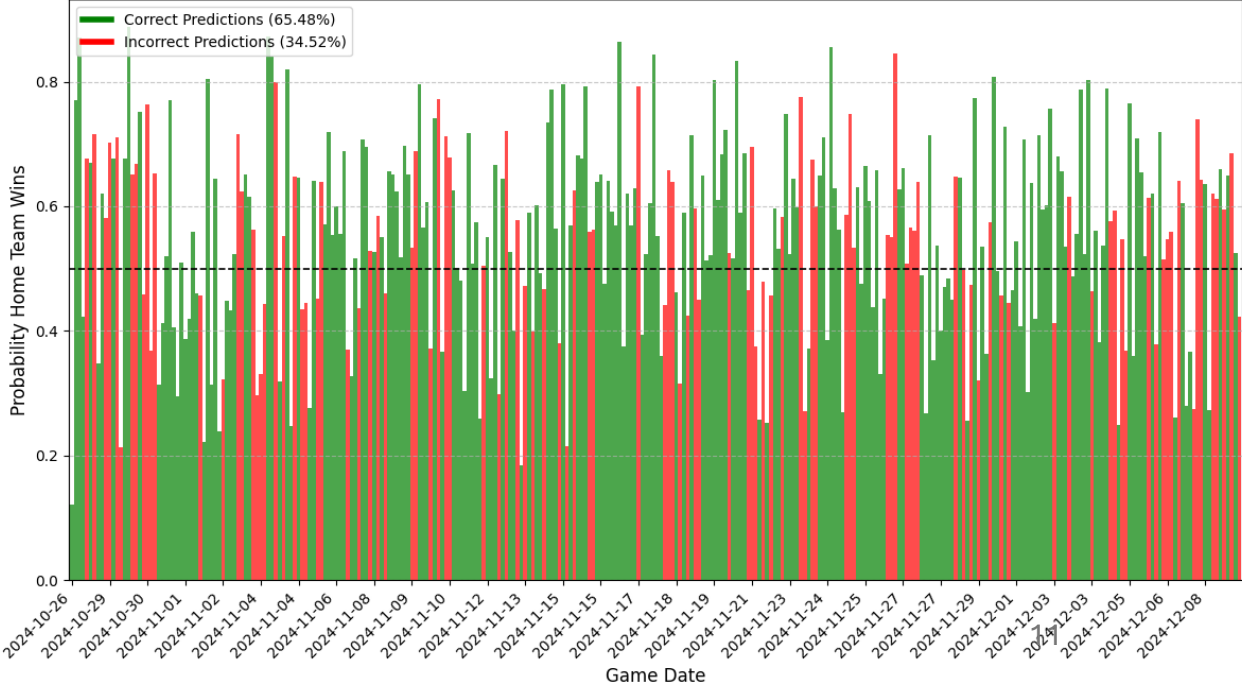
- Predicted home wins were correct 60% (LR: 65%)
- 58% (LR: 77%) of home wins were identified

- Predicted home losses were correct 53% (LR: 63%)
- 54% (LR: 48%) of home losses were identified

TABLE IV: Training Data: Neural Network Report

Class	Precision	Recall	F1-Score	Support
0	0.53	0.54	0.49	533
1	0.60	0.58	0.59	614
Weighted Average	0.57	0.56	0.57	1147
Additional Metrics:				
Accuracy			0.56	
Cross-Entropy Loss			2.84	
AUC-ROC			0.60	
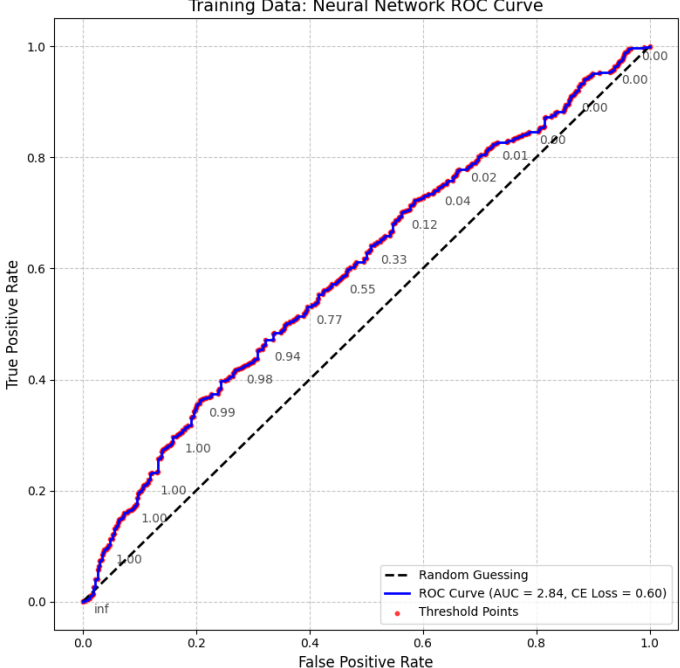
- Predicted home wins were correct 61% (LR: 68%)
- 55% (LR: 77%) of home wins were identified

- Predicted home losses were correct 45% (LR: 61%)
- 52% (LR: 49%) of home losses were identified

TABLE V: Validation Data: Neural Network Report

Class	Precision	Recall	F1-Score	Support
0	0.45	0.52	0.48	130
1	0.61	0.55	0.58	181
Weighted Average	0.55	0.54	0.54	311
Additional Metrics:				
Accuracy			0.54	
Cross-Entropy Loss			3.12	
AUC-ROC			0.55	
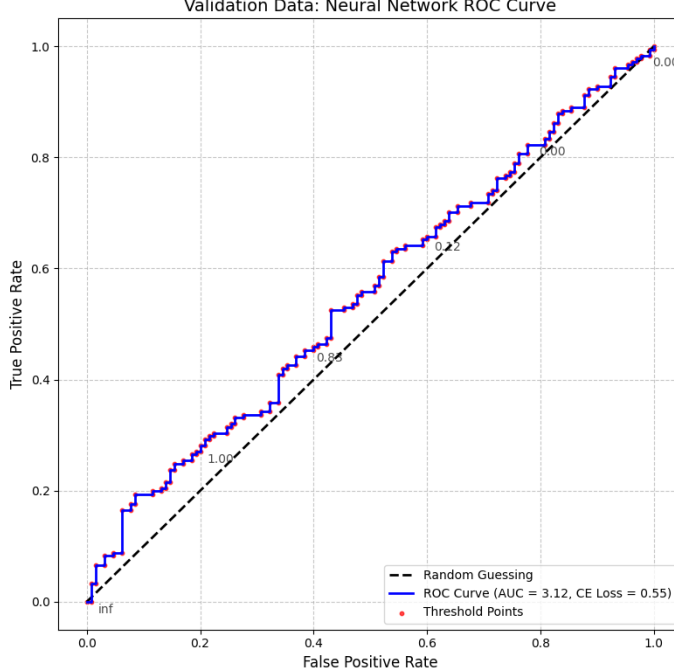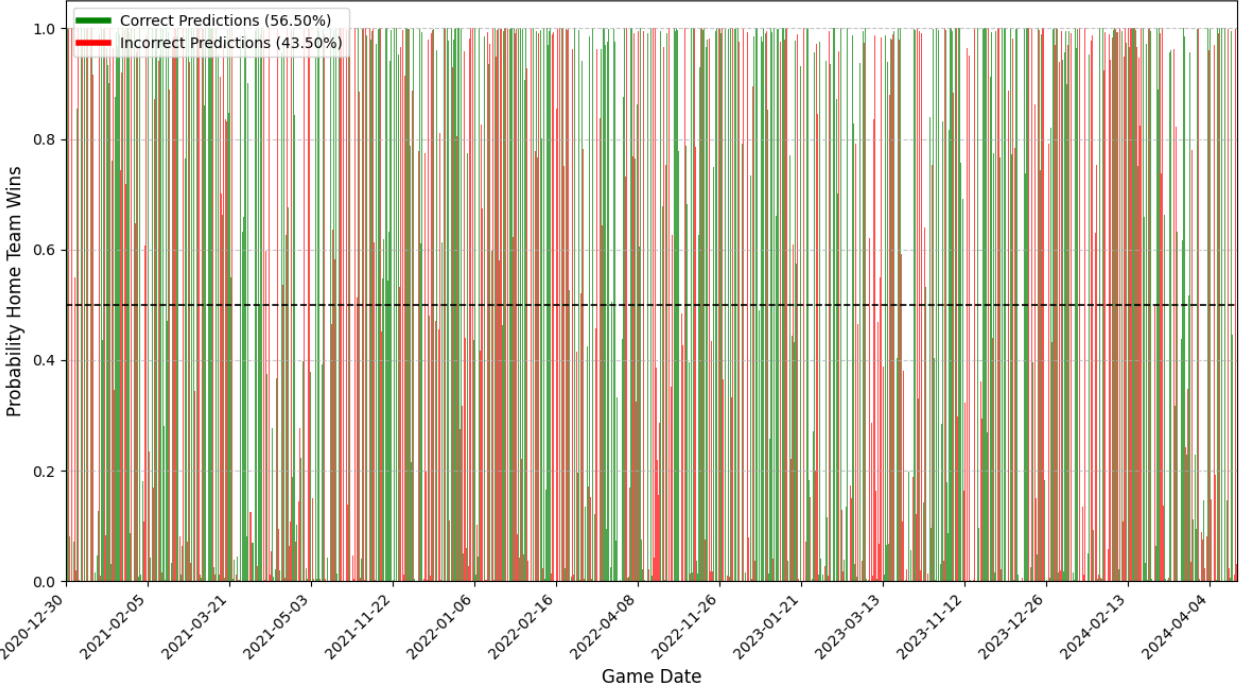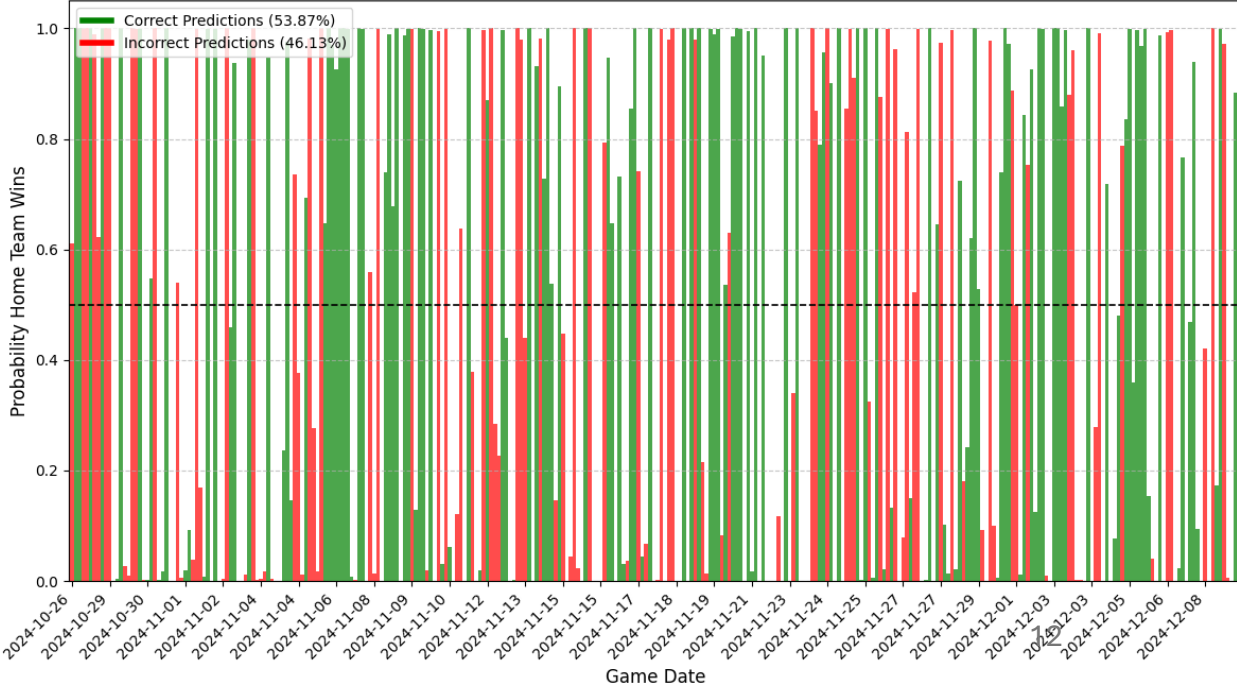
MILPlay Results: -$353.33 ☹

DATE	AWAY	HOME	Away ML	Home ML	Predicted Home Win	Probability Home Win	Wager	Optimal Wager	Potential Return	Actual Return
11/27/2024	Hawks	Cavaliers	350	-455	TRUE	0.845	Yes	$ 17.86	$ 3.92	$ (17.86)
11/27/2024	Bulls	Magic	320	-410	TRUE	0.714	No	$ -	$ -	$ -
11/27/2024	Trail Blazers	Pacers	360	-470	TRUE	0.627	No	$ -	$ -	$ -
11/27/2024	Clippers	Wizards	-470	360	FALSE	0.267	No	$ -	$ -	$ -
11/27/2024	Rockets	76ers	-218	180	FALSE	0.4	No	$ -	$ -	$ -
11/27/2024	Heat	Hornets	-175	145	TRUE	0.56	Yes	$ 892.86	$1,294.64	$ (892.86)
11/27/2024	Knicks	Mavericks	-170	142	TRUE	0.537	Yes	$ 17.86	$ 25.36	$ 25.36
11/27/2024	Kings	Timberwolves	130	-155	TRUE	0.551	Yes	$ 17.86	$ 11.52	$ (17.86)
11/27/2024	Pistons	Grizzlies	275	-345	TRUE	0.661	No	$ -	$ -	$ -
11/27/2024	Raptors	Pelicans	130	-155	FALSE	0.49	Yes	$ 17.86	$ 23.21	$ 23.21
11/27/2024	Lakers	Spurs	-125	105	TRUE	0.639	Yes	$ 17.86	$ 18.75	$ (17.86)
11/27/2024	Nuggets	Jazz	-535	400	FALSE	0.353	No	$ -	$ -	$ -
11/27/2024	Nets	Suns	320	-410	TRUE	0.566	No	$ -	$ -	$ -
11/27/2024	Thunder	Warriors	-180	150	TRUE	0.508	Yes	$ 17.86	$ 26.79	$ (17.86)
							Total	6	$1,000.00	$1,404.20	$ (915.71)

DATE	AWAY	HOME	Away ML	Home ML	Predicted Home Win	Probability Home Win	Wager	Optimal Wager	Potential Return	Actual Return
12/2/2024	Heat	Celtics	575	-950	TRUE	0.714	No	$ -	$ -	$ -
12/2/2024	Pelicans	Hawks	325	-540	TRUE	0.757	Yes	$ 937.50	$ 173.61	$ 173.61
12/2/2024	Lakers	Timberwolves	230	-285	TRUE	0.594	Yes	$ 62.50	$ 21.93	$ 21.93
12/2/2024	Nets	Bulls	250	-310	TRUE	0.601	No	$ -	$ -	$ -
							Total	2	$1,000.00	$ 195.54	$ 195.54

DATE	AWAY	HOME	Away ML	Home ML	Predicted Home Win	Probability Home Win	Wager	Optimal Wager	Potential Return	Actual Return
12/3/2024	Bucks	Pistons	-162	126	FALSE	0.488	No	$ -	$ -	$ -
12/3/2024	76ers	Hornets	-205	170	TRUE	0.615	Yes	$ 909.09	$1,545.45	$ (909.09)
12/3/2024	Wizards	Cavaliers	950	-1650	TRUE	0.803	No	$ -	$ -	$ -
12/3/2024	Magic	Knicks	190	-230	TRUE	0.535	No	$ -	$ -	$ -
12/3/2024	Pacers	Raptors	-130	110	TRUE	0.523	Yes	$ 22.73	$ 17.48	$ 17.48
12/3/2024	Jazz	Thunder	750	-1200	TRUE	0.788	No	$ -	$ -	$ -
12/3/2024	Grizzlies	Mavericks	150	-180	TRUE	0.556	No	$ -	$ -	$ -
12/3/2024	Spurs	Suns	205	-250	TRUE	0.656	Yes	$ 22.73	$ 9.09	$ 9.09
12/3/2024	Warriors	Nuggets	170	-205	FALSE	0.463	Yes	$ 22.73	$ 11.09	$ (22.73)
12/3/2024	Rockets	Kings	-130	110	FALSE	0.412	Yes	$ 22.73	$ 17.48	$ (22.73)
12/3/2024	Trail Blazers	Clippers	310	-395	TRUE	0.68	No	$ -	$ -	$ -
							Total	5	$1,000.00	$1,600.60	$ (927.97)

DATE	AWAY	HOME	Away ML	Home ML	Predicted Home Win	Probability Home Win	Wager	Optimal Wager	Potential Return	Actual Return
12/5/2024	Mavericks	Wizards	-750	525	FALSE	0.249	No	$ -	$ -	$ -
12/5/2024	Nuggets	Cavaliers	140	-166	TRUE	0.709	Yes	$ 31.25	$ 18.83	$ 18.83
12/5/2024	Hornets	Knicks	850	-1450	TRUE	0.765	No	$ -	$ -	$ -
12/5/2024	Thunder	Raptors	-410	320	FALSE	0.36	No	$ -	$ -	$ -
12/5/2024	Bulls	Spurs	110	-130	TRUE	0.547	No	$ -	$ -	$ -
12/5/2024	Suns	Pelicans	-115	-105	FALSE	0.369	Yes	$ 31.25	$ 29.76	$ (31.25)
12/5/2024	Kings	Grizzlies	150	-180	TRUE	0.654	Yes	$ 31.25	$ 17.36	$ 17.36
12/5/2024	Rockets	Warriors	-170	142	TRUE	0.52	Yes	$ 906.25	$1,286.88	$ 1,286.88
							Total	4	$1,000.00	$1,352.82	$ 1,291.81

- $1000 wagered on 4 different days
- Must wager of ½ available games
- Minimum wager is ($1000/(2 x # Games))

Conclusions, Future Work, References

Conclusions:

- Both models had over 50% accurate predictions in testing and validation sets

- Decision making chooses one massive wager and minimum for the rest

- Did not translate to monetary result … yet

Future Work:

- More features!
    - Injuries * Player Importance
    - Historical frequency analysis

- Evaluate different modeling methods/parameters

- Minimum risk constraints

References:

https://medium.com/@theresearchlab/create-an-nba-win-loss-model-w-68-precision-d1c6a21f0ded