NBA Spread Prediction

By:
Bryce Anthony , Archie Frank, Michael d'Amore, John Miller, Mervin McDougall

Problem Definition

Goal: To predict the NBA winning team and the point spread using a neural network.

- The surge of online sports betting industry
- What is a spread?
- Win-loss prediction
- More informed betting decisions

Data Sources and Scraping

- Basketball Reference -https://www.basketball-reference.com/
- NBA Stats - https://www.nba.com/stats
- Data Scrape (python) and CSV download

Advanced Stats:

- PER (Player Efficiency Rating) – All-in-one basketball rating, that boils down all of a player's contributions into one number. Using a detailed formula it rates every player's statistical performance.(It has drawbacks)
- WS (Win Shares) - The estimated amount of wins a contributed by a player
- AST% (Assist Percentage) - An estimate of the percentage of teammate field goals a player assisted on while on the court
- VORP (Value over Replacement Player) -A box score estimate of the points per 100 TEAM possessions that a player contributed above a replacement level player, translated to an average team and proportional to an 82 game season
- BMP (Box Plus Minus) – Estimate the player's contribution in points above league average per 100 possessions played.
- OBMP (Offensive Box Plus Minus) - a statistic that measures a player's impact on team offense
- DBMP (Defensive Box Plus Minus) - a statistic that measures a player's impact on team defense

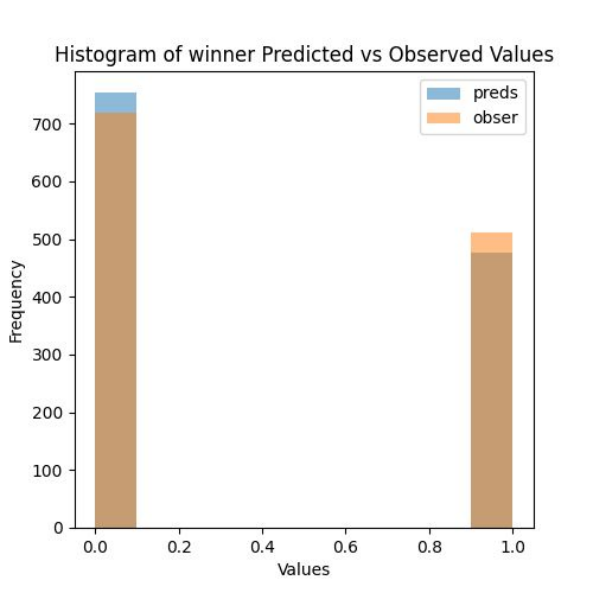
Exploratory Data Analysis

- Arena, Home Team, and Away Team data was converted to sparse matrices
- Spread was not included but was derived from home team points and away team points
- Correlated features identified during the EDA were removed. Eg. home_fta, which is correlated with home_fta_pct

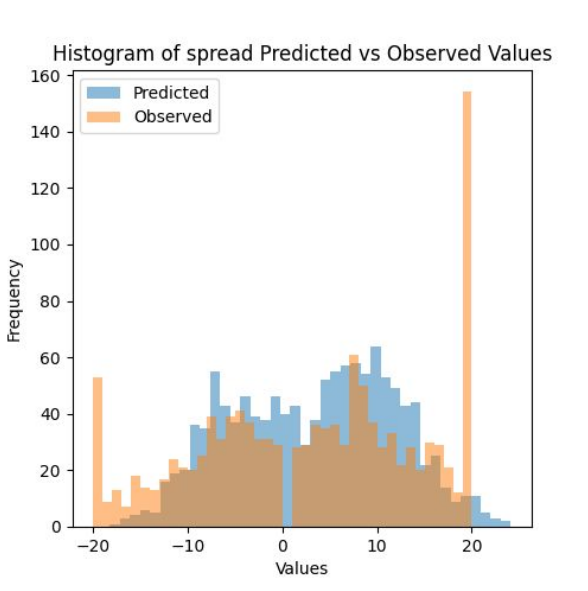
Training and Experimentation

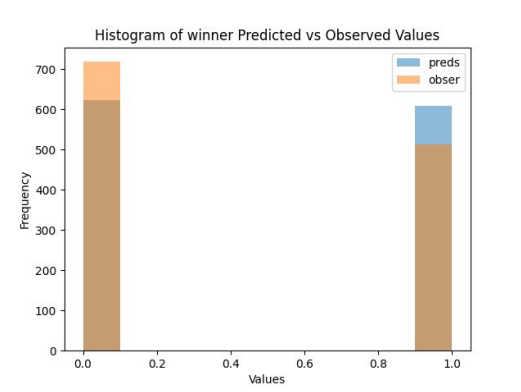
- Gelu. (best response)
- SeLu
- Swish
- Mish
- ReLu
- Lazy ReLu
- Soft Max
- Soft Plus

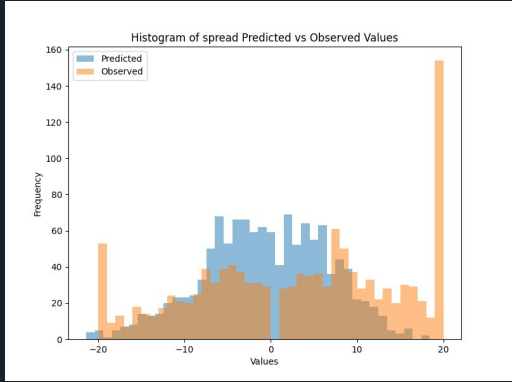
- Adam
- AdaMax
- Nadam (best response)
- Batch normalization
- Dropout training
- Transfer Learning
- Bucketization of outcome
- Feature selection



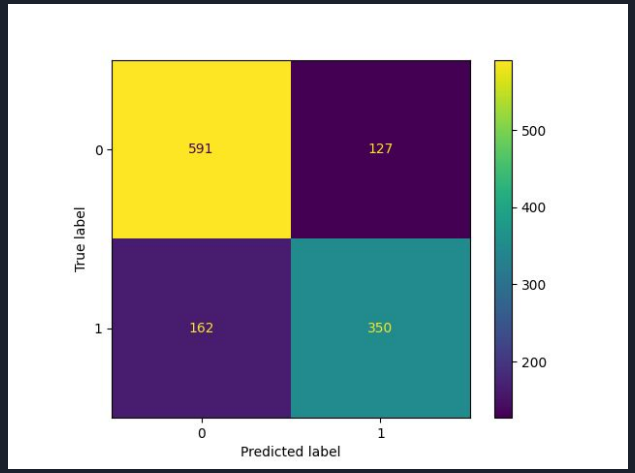
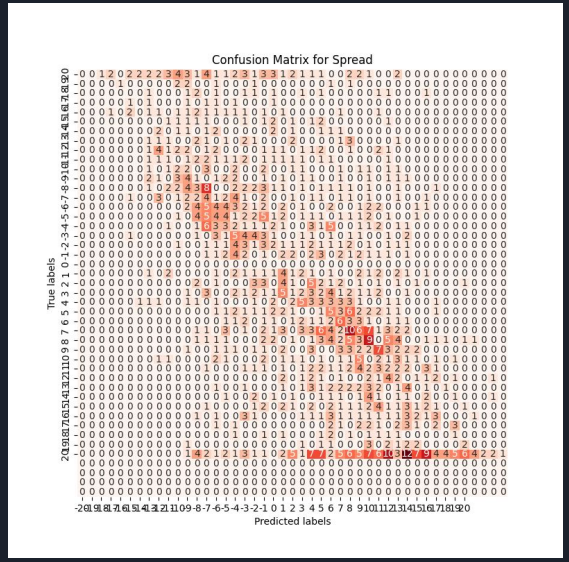
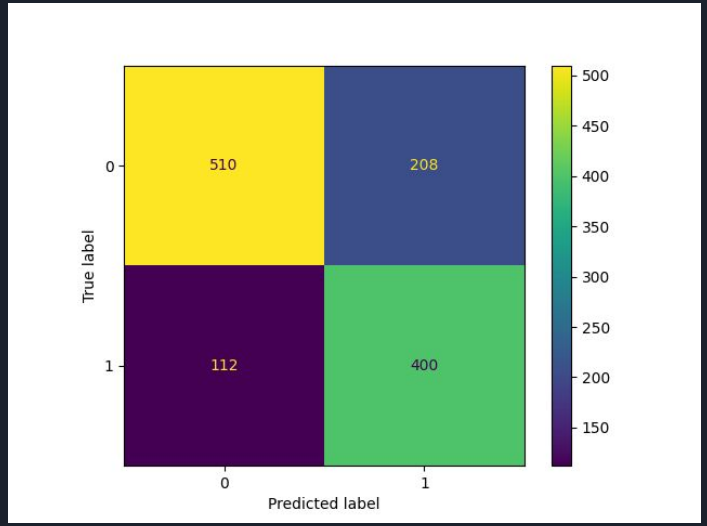
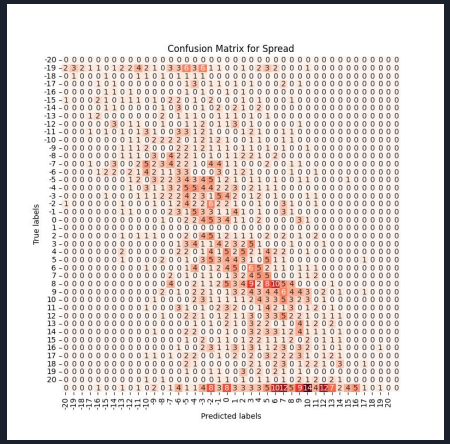
diction but our spread is not ac
hat predicting the spread goes
lepict findings having trained o







Train Results:
Accuracy: 76.50 %
F1-score : 80.35 %
Precision score: 78.49%:
Recall score : 78.48%

0 = home team win
1= home team loss

Test Results:
Accuracy: 73.98
F1-score : 76.12%
Precision score: 81.99%:
Recall score : 81.9%