Muhammad Aisy Iffat
FC14025

MLBB Draft Win Probability Predictor

3 January 2026

Problem Statement

The “Why”:

- MLBB players often struggle to evaluate whether a draft composition has a higher probability of winning before the match starts.

Stakeholders:

- Casual players
- Competitive / ranked players

Impact:

- Poor draft decisions can lead to lower win rates, wasted time, and inefficient team compositions.

Data Overview

Source:

- sample_101_Mlbb_Heroes.csv

Granularity:

- One row represents one hero and its attributes

Size:

- 100 rows, 18 columns

Target Variable:

- Win_rate

OBJECTIVES & KEY QUESTIONS

Objectives:

- Analyze MLBB hero statistics
- Build a model to estimate win probability
- Deploy an interactive draft prediction application

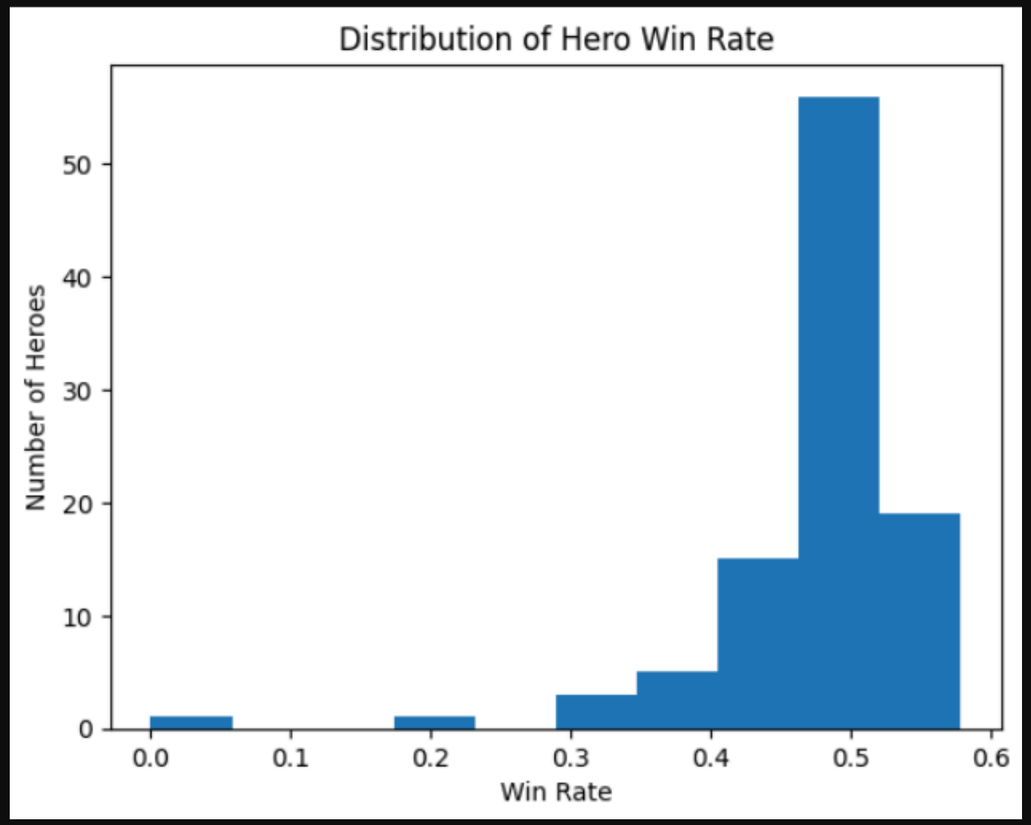
Key Questions:

- How do hero statistics affect win rate?
- Can draft-level stat differences predict match outcome?

METHODOLOGY

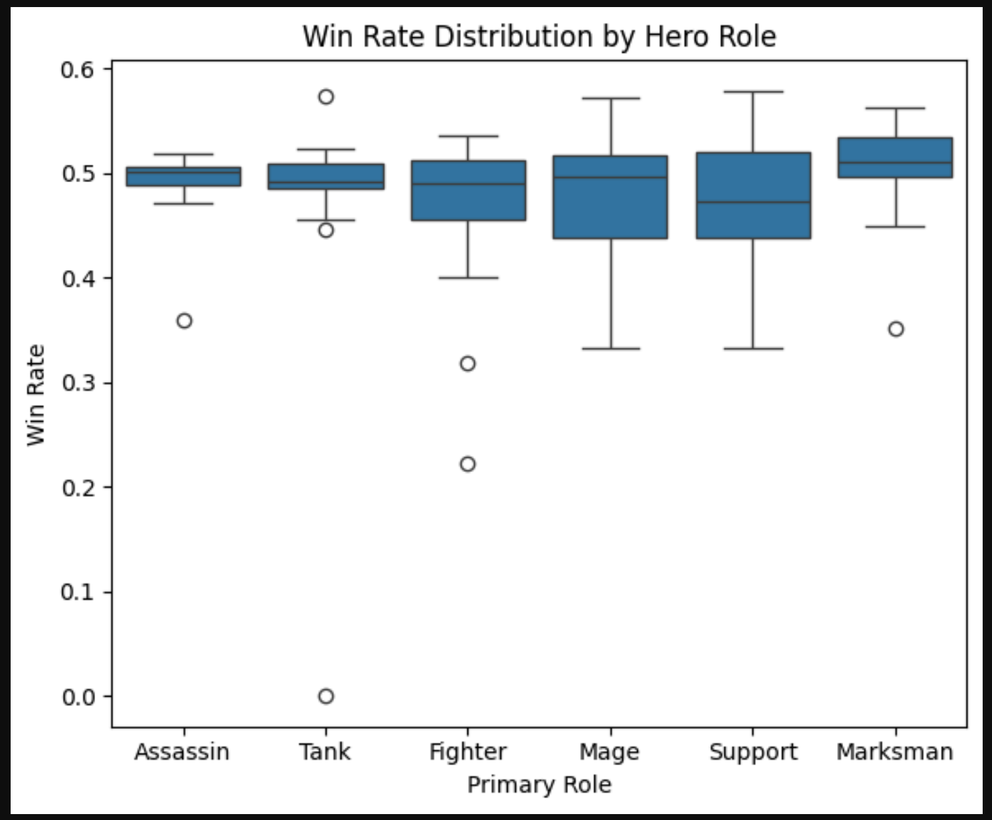
- Data Collection: MLBB hero statistics were collected from a CSV dataset.
- Data Cleaning: Irrelevant columns were removed and missing values were handled.
- EDA: Data distribution and relationships were analyzed using statistical plots.
- Feature Engineering: Hero statistics were aggregated at the draft level.
- Model Training: Linear Regression was trained using a train-test split.
- Deployment: The model was deployed as a local Streamlit application.

EDA KEY FINDINGS



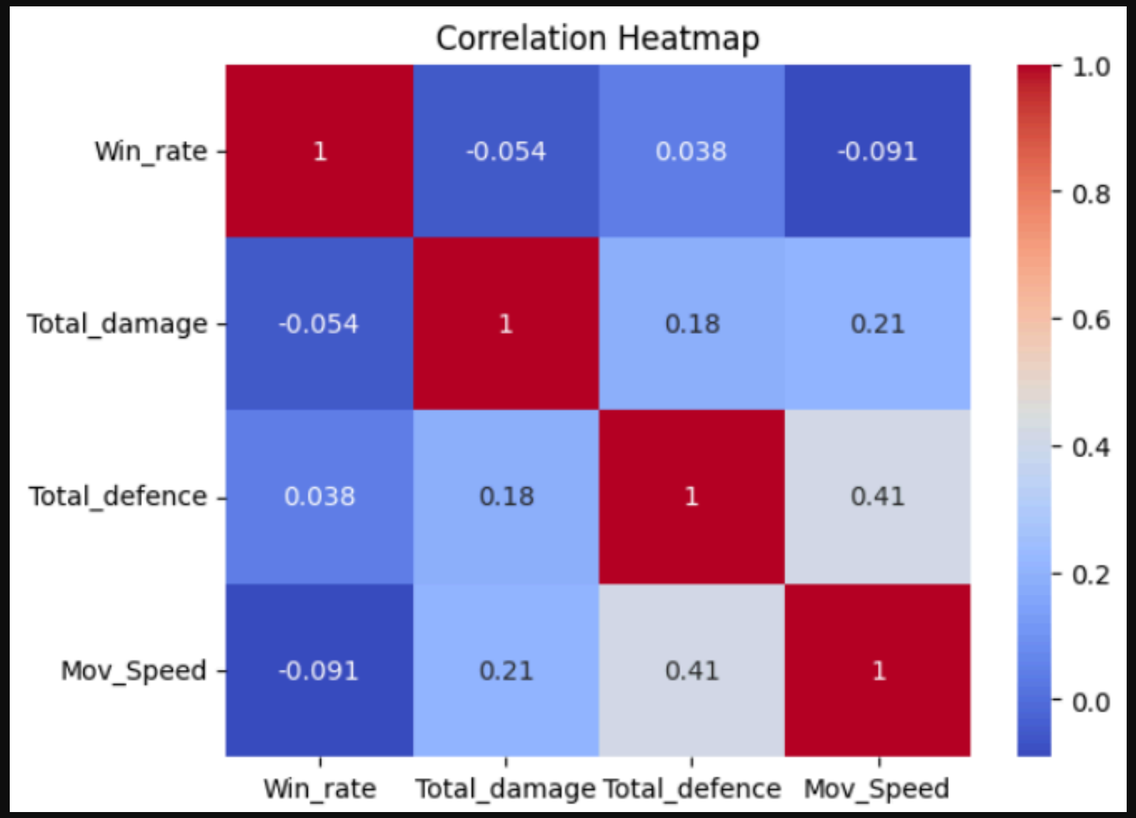
- Evidence: Win rate values are centered around the mid-range.
- Interpretation: Most heroes have similar performance levels.
- Action: Win rate is suitable as a regression target.

EDA KEY FINDINGS



- Evidence: Win rate varies across different hero roles.
- Interpretation: Certain roles perform more consistently than others.
- Action: Role-based differences support using hero stats in draft evaluation.

EDA KEY FINDINGS



- Evidence: Damage, defence, and movement speed correlate with win rate.
- Interpretation: Multiple attributes influence hero effectiveness.
- Action: These features were selected for model training.

MODELING APPROACH

Algorithm:

- Linear Regression

Why:

- To predict continuous win rate values based on aggregated draft statistics.

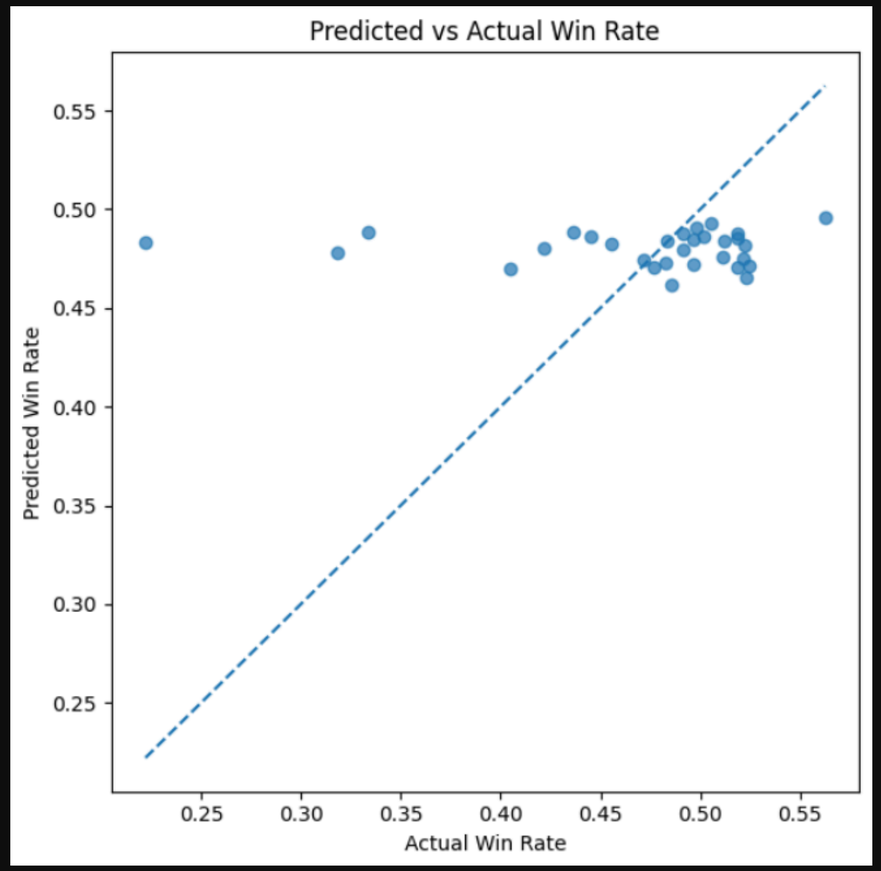
Validation:

- Train-test split (80% training, 20% testing)

Feature Engineering:

- Aggregate total damage, defence, and movement speed
- Compute ally vs enemy draft differences

RESULTS & EVALUATION



Metrics :

- MAE: 0.04624414232006068
- RMSE: 0.07142764009669994
- R2: -0.03492917418200525

"So What?" Insight :

By predicting win probability from draft compositions, the model helps players identify weaker drafts early and adjust hero selections before the match starts.

PROJECT DEMO

Flow : Hero Selection → Model → Win Probability

Link : https://project-mlbb-draft-predictor.streamlit.app/

MEASURE OF SUCCESS

Target Metric:
Root Mean Squared Error (RMSE)
Result:
The model achieved a low RMSE value, indicating that the predicted win rates are close to the actual win rates.
RMSE: 0.07142764009669994

Business KPI Success Criteria:

- The model can distinguish between stronger and weaker draft compositions.
- Predicted win rates change meaningfully when draft heroes are changed.
- Players can use the prediction as guidance during the draft phase.

CHALLENGES & LIMITATIONS

Incomplete hero statistics :
Handled by using only consistently available core features.

Uneven data coverage across heroes :
Addressed by aggregating hero stats at the draft level.

Limited dataset size :
Evaluated using a simple train-test split.

FUTURE WORK

- Use match-level draft data instead of aggregated hero-level statistics.
- Incorporate hero synergy and counter-pick information into draft features.
- Include role balance features in draft evaluation.
- Expand the dataset to include more heroes and updated statistics.

TECH STACK

- Language: Python
- Libraries: Pandas, NumPy, Scikit-learn
- Infrastructure: Streamlit

Q&A SESSION

THANK YOU