Predicting Wheeling in MTG Drafts with Random Forests

Ryan Saklad

05/01/2024

Overview

This project investigates the factors influencing card selection strategies and outcomes in Magic: The Gathering (MTG) drafts. Utilizing a comprehensive dataset from 17lands, an MTG draft tracker, we explore the impact of draft type, player rank, and the sequence of card selections on the probability of cards returning to the drafter in subsequent rounds. Through extensive data cleaning, exploratory analysis, and machine learning modeling, we uncover insights into optimal drafting strategies and the dynamics of the draft process. Our findings contribute to a deeper understanding of strategic decision-making in MTG and offer practical implications for players and game designers alike.

The dataset consists of over 20,000 unique drafts across multiple MTG sets. For each draft, detailed information is provided including the cards available in each pack, the card selected by the player, relevant player stats, and the draft outcome. This rich data enables a thorough examination of the complex interplay between player skill, card evaluation, signaling, and probability that shapes the draft experience.

Applying a Random Forest classifier to predict the likelihood of a card wheeling based on the cards in the pack, we achieved promising initial results with an accuracy of 0.86, precision of 0.91, recall of 0.77, and F1 score of 0.83. These findings lay the groundwork for more sophisticated modeling approaches that incorporate a wider range of draft features. Ultimately, this line of research aims to equip players with data-driven insights to optimize their drafting decisions and to inform the design of limited environments that facilitate engaging and skill-testing gameplay.

Summary

This project aims to investigate the factors influencing card selection strategies and outcomes in Magic: The Gathering (MTG) drafts. Utilizing a comprehensive dataset from 17lands, an MTG draft tracker, I explore the impact of draft type, player rank, and the sequence of card selections on the probability of cards returning to the drafter in subsequent rounds. Through extensive data cleaning, exploratory analysis, and machine learning modeling, I seek to uncover insights into optimal drafting strategies and the dynamics of the draft process. The findings of this research contribute to a deeper understanding of strategic decision-making in MTG and offer practical implications for players and game designers alike.

Project Description

Motivation

Magic: The Gathering drafts present players with a complex strategic challenge, requiring not only a deep understanding of the card pool but also the ability to read signals and anticipate opponents' decisions. Skilled players often cite the ability to identify open colors, read signals, and navigate the draft based on the flow of packs as key factors in their success. However, there is currently limited empirical research quantifying the impact of these skills and how they interact with the structure of the draft itself. This project seeks to address this gap by leveraging the extensive data collected by 17lands to uncover the key determinants of desirable outcomes like wheeling powerful cards.

Aims

The primary objectives of this research are:

- Quantify the influence of player rank on card selection patterns and outcomes
- Investigate how the sequence of picks within a pack affects the probability of wheeling
- Compare card wheeling rates across different types of drafts (e.g., Premier vs. Quick)
- Develop predictive models to estimate the likelihood of a card wheeling based on observable features

- Generate actionable insights for both players looking to improve their draft strategies and game designers aiming to create engaging limited environments

Dataset

The dataset comprises over 20,000 unique draft events recorded by 17lands, spanning multiple MTG sets. For each draft, the data includes:

- Anonymized player ID and rank
- Draft type (e.g., Premier, Quick)
- Set name
- Pack number and pick number for each selection
- Contents of each pack
- Card selected by the player
- Outcome of the draft (e.g., number of match wins)

This detailed, longitudinal data enables a thorough examination of the complex factors shaping draft decisions and outcomes. The dataset is sourced directly from the 17lands website and is continuously updated with new draft data and sets.

Methods

Data Cleaning and Preprocessing

The raw data from 17lands underwent cleaning and preprocessing to ensure its suitability for analysis. Rows with missing or inconsistent information were removed to maintain data integrity. Rank categories were standardized to facilitate grouping and comparison.

Pack contents and pick information, originally stored as concatenated strings, were split and expanded into individual columns, with each card represented as a binary indicator variable. This transformation allowed for more intuitive analysis and modeling of card selections.

The cleaned data was stored in a SQLite database to enable efficient querying and retrieval. The database schema was designed to normalize the data and establish proper relationships between drafts, packs, picks, and outcomes.

Exploratory Data Analysis

Exploratory data analysis (EDA) was conducted to gain insights into the patterns and relationships within the dataset. Key visualizations included (but are not limited to):

- Distribution of drafts by player rank and set
- Breakdown of draft types by player rank
- Visualization of the frequency of cards being picked at each position within a pack
- Comparison of match win rates across player ranks and sets

These visualizations, along with summary statistics, helped inform feature selection for predictive modeling. They will be covered in more detail in the Conclusion section.

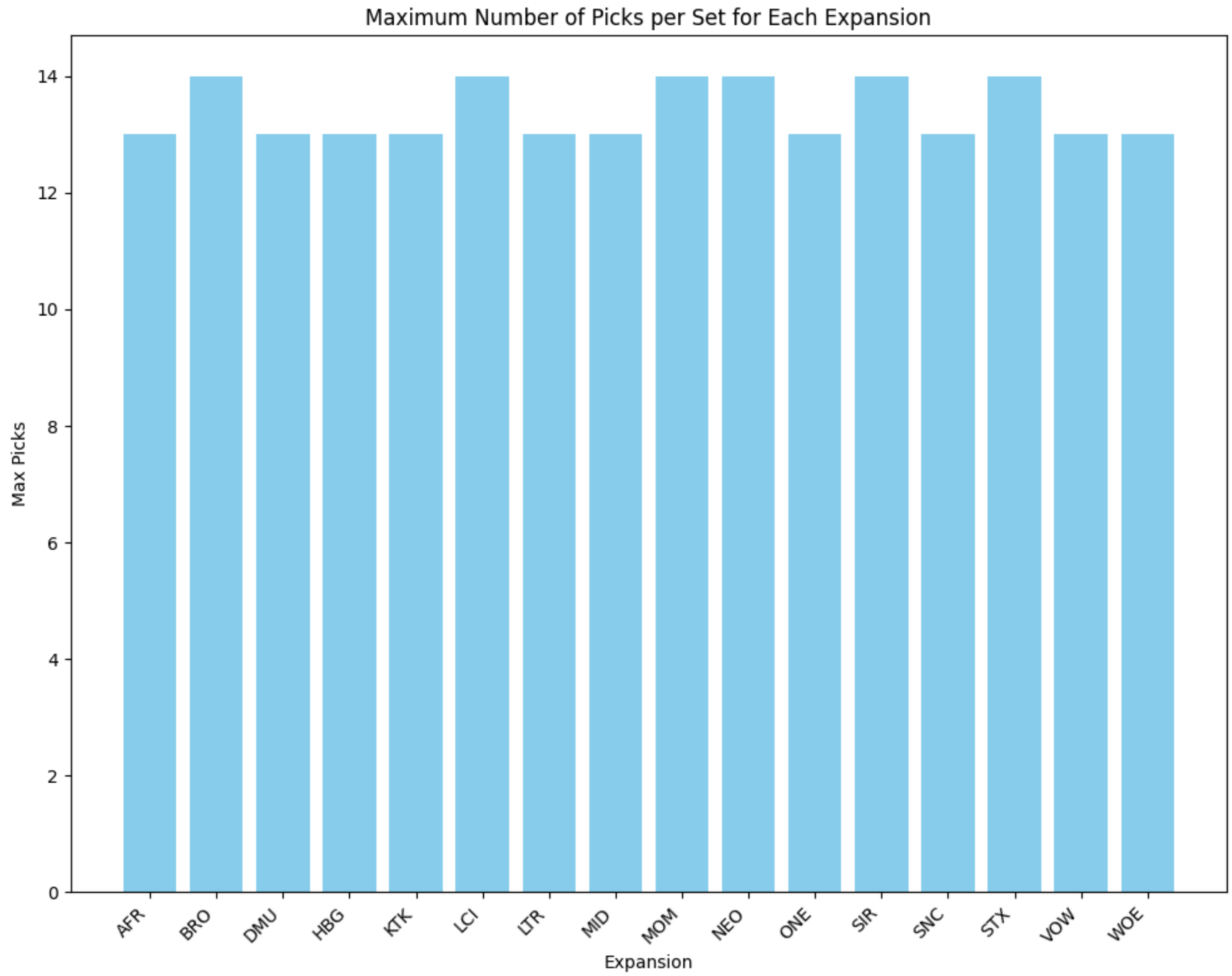
Predictive Modeling

A binary classification approach was employed to predict the probability of a card wheeling, with the target variable being whether the card was seen again in a later pack (1) or not (0).
The dataset was split into training and testing subsets. Various classification algorithms were evaluated, and hyperparameter tuning was performed to optimize model performance.
Model performance was assessed using metrics such as accuracy, precision, recall, and F1 score. The final selected model was a random forest classifier.
To assess generalizability, the model was evaluated on subsets of the data, such as specific player ranks or sets.

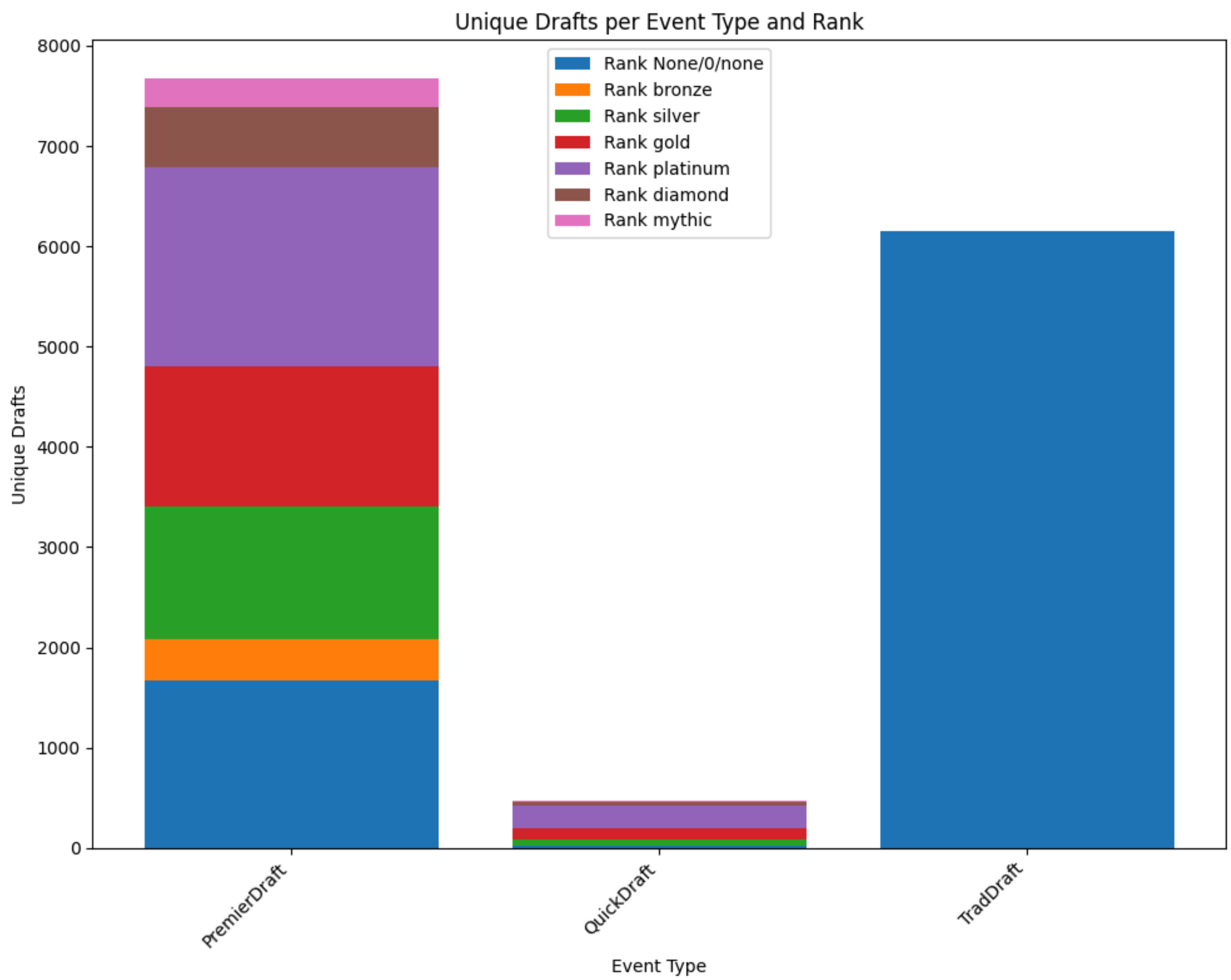
Conclusion

Visualizations

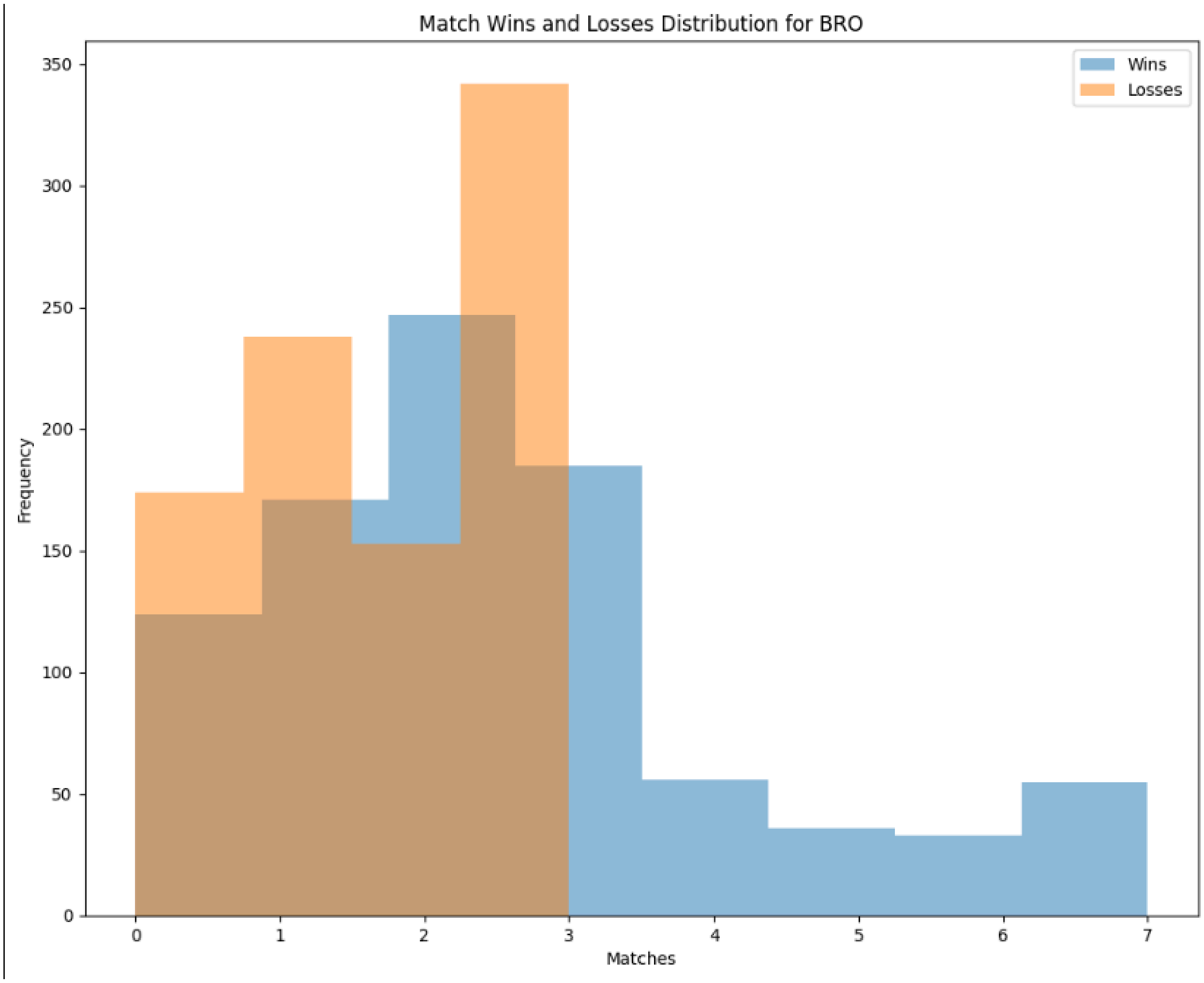
While numerous visualizations were created, some were more relevant than others, and will be highlighted here. For a full list of the visualizations performed, check the notebook.



The model was robust at handling various sets, despite the number of total picks in the draft varying between 13 and 14. In older sets (not included in this data), that number fluctuates between 8 and 15+, so it is important that it can handle different pack sizes on the more limited data.



There is a significant disparity between the types of drafts and the ranks within them. This leads to an imbalanced data set, although the nature of Random Forests and the similarities between draft formats greatly mitigates this concern.



Due to the nature of Arena MTG drafts, where a player is out after 3 losses, the distribution is highly irregular. This may affect the drafting process by encouraging players to shoot for the moon

Results

As for results, the model performed quite well, despite the Naive approach. It reached a final accuracy of 86.46%, precision of 90.70%, recall of 77.12%, and a F1-score of 83.37%. While these results are lower than prior tests, this is because a smaller percentage of data from each set was used to train each model to save on computational costs. If data was increased, results would likely be markedly better. Additionally, results may have been worse than the original set due to differences between the sets, such as greater variability in what cards were likely to wheel.

Limitations and Conclusions

The dataset covers a subset of MTG sets (with a limited amount of data per set) and may not fully represent all draft environments. The predictive model focuses primarily on card wheeling likelihood and does not directly incorporate factors such as deck composition, drafting preference, and various other potentially relevant factors. The current approach treats each pick independently, not accounting for potential interdependencies within a draft. Additionally, it uses card names only due to computing power constraints, rather than holistically viewing everything. Fine-tuning a LLM is a promising direction for much better understanding of the draft holistically, which would also allow it to easily understand game data, synergies, and more.

Future work could address these limitations by incorporating additional data sources, refining feature engineering, and exploring advanced modeling techniques. Qualitative analysis of player interviews or surveys could provide valuable context to complement the quantitative findings.