Joann Guyoncour

MAIS 202 – Deliverable 1

Project Title: Clash Royale Win Prediction Model

1. Choice of Dataset

I have chosen a **Clash Royale Games dataset** from Kaggle (Clash Royale Battles). This dataset contains battle data from 500 million Clash Royale matches, including player IDs, the cards used in each match, the number of crowns obtained, and the player's trophy count. A dataset with precise stats for each card could be added for more precise information (Clash Royale Card Stats). The dataset is large-scale and well-structured allowing us to analyze how deck composition influences match outcomes. The same data collection methodology could be used to extract more recent data from the latest Clash Royale season (Clash Royale Analysis).

2. Methodology

a. Data Preprocessing

- **Feature Selection:**
  - Player 1 & Player 2 decks: List of 8 cards per player.
  - Number of crowns: Represents how decisive the win was.
  - Elixir average of deck: Affects the pacing of the game.
  - Additional card characteristics?

b. Machine Learning Model

The goal is to predict the probability of Player 1 winning a match, which is a binary classification problem with the probability being a confidence measure. From my understanding and ChatGPT's insight, this problem could be tackled with:
- a logistic regression: interpretable, fast to train, and provides probability scores but may struggle to capture nonlinear relationships.
- a random forest model: handles nonlinearity well, provides feature importance insights.

c- Evaluation metric

A confusion matrix with accuracy/precision-recall/logistic loss (classification problems) could do the job.

3. Application:

Create a win prediction model accessible through a web-based interface where users can input two Clash Royale decks and receive a win probability prediction.