Final Project Proposal

Novel Analysis of Magic the Gathering Decks

Philip Nelson and Brigham Michaelis

# Motivation

Magic the gathering is an incredibly popular trading card game with over 35 million players! ([link](https://www.dicebreaker.com/games/magic-the-gathering-game/news/mtg-commander-audience-tripled#:~:text=Its%20last%20major%20info%2Ddump,to%2Dplay%20release%20on%20PC.)) In the game, 2 players attempt to defeat their opponents by dealing them 20 or more damage throughout a series of turns. Players can attack their opponents and defend themselves by playing cards from their hand. These cards could be magical creatures, artifacts, spells, or abilities. The first player to successfully deal 20 damage to their opponent is the winner. Players create their own custom decks using any combination of the 20,000 available cards. There are four rules that govern building a deck: ([link](https://magic.wizards.com/en/content/standard-formats-magic-gathering#:~:text=Your%20deck%20must%20be%20at,deck%20in%20your%20hands%20unassisted.))

* Your deck must be at least 60 cards.
* Up to fifteen cards may be included in your sideboard, if you use one.
* Include no more than four copies of any individual card in your main deck and sideboard combined (except basic lands).
* There's no maximum deck size, as long as you can shuffle your deck in your hands unassisted.

A common method for gauging the effectiveness of a deck is called goldfishing. Goldfishing, or playing against a Goldfish, is the practice of playing a deck without an opponent. This involves drawing a starting hand and proceeding to continue to play until the goldfish, who does nothing to stop you, is defeated. This method is limited in its effectiveness because it does not model a true gameplay scenario.

The other major source of deck ranking comes from looking at decks that are currently being played in daily events. For example, on mtggoldfish.com they provide a score called META% which shows the percentage of players that won three or four out of four matches in Daily Events recently with a certain deck. This is great for playing the current epoch but does not aid players who want to see how their decks stack up for past epochs.

# Project Description

Over the next two weeks we will work to develop a novel method of Magic the Gathering deck analysis. Our goal is to take a deck and assign it a numeric value which reflects its rank compared to other decks. We will use a dataset from kaggle with basic card information and scrape data from various fan websites to build a dataset that we can mine for deck statistics.

Some of the data science methods we will use to get us started are data correlation (p value test) between various card attributes like the mana cost of a card in combination with how powerful its abilities are can drive a cards rank or a deck’s rank. We will use linear regression to rank a deck compared to known decks. To drive our analysis we will attempt to answer questions like “do any of these methods provide hints or clues to the underlying rank of cards and or decks?” Another question we could ask is, “are there correlations between card/deck rank and mana cost?” Standardization of the card data may also help to flush this out. Another method that we could explore is training a neural network to the inputs of cards or various known decks. The inputs could be the mana cost of the card, its type, its potential damage, etc, then using it to generate decks.

# Schedule

April 12-15: Get data from kaggle, munge and analyze.

April 16-25: Using deck data from kaggle, mtggoldfish, mtgmeta, and more, devise a novel method to rank a deck. (This may involve several failed attempts)

April 26-30: Write 6 page report.

May 1: Submit project

# Responsibilities

This project is going to require collaboration during each task listed in the schedule. It will not work to segregate the work in a manner that one of us ignorant or unaware of some aspect of the data or some application of a modeling technique. We anticipate that we will work together on data munging and the novel method creation. We will break up the websites we want to scrape so we can generate our data set more quickly, then work together to create our deck ranking model. As we encounter work or tasks that need to be accomplished we will dynamically balance the load.

# Potential Problems and Solutions or Alternatives

The first issue that we foresee is that our novel deck ranker could suck. However it will not be for lack of trying. It is anticipated that multiple attempts will have to be made before we are able to (even naively) come up with a useful deck ranker. All of our various approaches and ideas will be explored and documented. With the limited amount of time and large amount of data we are going to be looking at, it may be possible to come up with a naive card ranker with all the data we will have available to us. Furthermore, it may be that we could significantly add to the dataset on Kaggle to create and submit a new dataset. Finally, if our previous goals fail, we will be able to ask and answer some very interesting questions with regards to Magic the Gathering.

# Any preliminary results

We have only started gathering data but not yet started building our novel deck ranking analysis.