DS 720: Project Proposal

Pokémon TCG Cube Analysis Tool

1 BACKGROUND INFO

The Pokémon Trading Card Game, TCG for short, has been around since the 90s and has a rich history. The traditional way to play the game has players construct decks from their personal collection before playing, but a popular alternative way to play has players drafting cards, much like a fantasy football draft, to construct decks from a limited and shared pool. This limited pool is called a cube. It consists of anywhere from 360 to more than a thousand carefully handpicked cards. The process of creating a cube is a very engaging, but tedious effort. I plan to fix this issue.

2 RESEARCH OBJECTIVE

I propose that I should design and implement a tool for analyzing Pokémon TCG cubes in order to streamline the process of creating and balancing a cube.

2.1 RELEVANCE

I spend a lot of time poring over large amounts of Pokémon cards and contemplating if they would work well inside a cube. The process requires many considerations when balancing the cube. A tool for analyzing the many different aspects of the cards within a cube would expedite the process.

2.2 APPLICATIONS

Other Pokémon TCG players would be able to use the tool similarly to facilitate the process of cube design.

2.3 FUTURE INVESTIGATIONS

If given more time, I would like to use large language models to further analyze individual cards with a cube and their power levels among other cards to predict any outliers that could cause issues in design.

3 POTENTIAL DATA SOURCES

As for data sources, I have found the PokémonTCG API (https://pokemontcg.io/) which has a built in Python SDK. This will provide me with specific information about individual cards.

For the cubes (card collections) themselves, I have found CubeKoga (https://cubekoga.net) which can manage cube information and export the collections for analysis.

3.1 Preliminary Summary Stats

Using a prebuilt cube from CubeKoga, I queried the PTCG API for each card and found the API was able to correctly return information about 100% of the cards. However, the API documentation states cards are only updated to the Sword and Shield format, that is information doesn't include newer released cards, but I haven't found any issues querying those cards.

4 Intended Approach

To complete the project, I will undertake the following steps:

1. Creating a system for dynamically preprocessing CubeKoga card data and converting it to PTCG SDK card data
2. Using ML techniques to summarize the following data:
   a. card category (For trainers, consistency or utility, for Pokémon (consistency / utility / attacker / pre-evolution)
   b. card relative power level
3. Implementing the following visualizations:
   a. Distribution of card supertypes
   b. Distributions of card subtypes within each super type (trainers and Pokémon)
   c. Distribution of attack energy costs within "line topper" Pokémon
   d. Distribution of attack damage within "line topper" Pokémon
   e. Distribution of HP within "line topper" Pokémon grouped by Stage
   f. Relation of attack energy cost to attack damage with "line topper" Pokémon, grouped by stage
   g. Distribution of card usage category (For trainers, consistency or utility, for Pokémon (consistency / utility / attacker / pre-evolution) grouped by super type
   h. Distribution of relative power level of cards

4.1 Potential Techniques/Types of Modeling

I would like to use ML techniques to summarize data about individual cards. The most important of that being the card usage category and the card relative power level. These summaries could be obtained in several ways:

- Card power level:
  - Using neural networks to analyze information about card HP, attack cost, and damage to predict relative power level
- Card category:
  - Using LLMs to categorize based on card text data
  - Using unsupervised learning on card text data
  - Classification based on specific keywords within text like search, shuffle, draw, and the lack there of