**Using Machine Learning to Predict the Outcome of VGC Pokémon Battles**

**Overview**

The Video Game Championships, or VGC, is the official format of Nintendo’s Pokémon tournament battles. Our goal for the project is to create a machine learning model which can predict the winner of a battle based on information available in an open team sheet.

**Challenges**

Pokémon VGC is a complex game to model. During team building, players choose a team from over 1000 Pokémon, each with unique stats, moves, abilities, and types. Pokémon have 4 different moves to use per turn, and the same Pokémon can perform differently between matches depending on training (EV distribution), items, movesets, and synergy with other Pokémon. VGC is played in a bring-6 pick-4 doubles format, expanding the total options available per turn and before-battle multiplicatively.

**Dataset-Specific Challenges**

1. While we would ideally like to model official tournament battles, the largest public source of in depth battle data comes from Pokémon Showdown, an online battle simulator. Our data is therefore biased towards *Pokémon Showdown* players, who may play differently to those in official tournaments.

2. Rule changes called ‘regulations’ occur every 3 to 4 months, massively influencing the popularity and effectiveness of different Pokémon and strategies. The current regulation is called Regulation I and will end in September 2025. Our findings will apply only until that date, though we plan to make our code modular with the capacity to model past and future regulations with minor alterations.

3. Some online battles end abruptly by forfeit or disconnection. These are encoded automatically in our dataset as losses for the forfeiting / disconnecting player, but may not be informative to our.model if the battles were not properly engaged in by both players. In preprocessing, we removed games lasting only one turn, and games in the bottom 10th percentile of turns (4 turns) in which a player disconnects or forfeits.

**Similar Projects**

The [Pokémon Battle Predictor](https://www.pokemonbattlepredictor.com/home) similarly used in-game data to [model](https://www.reddit.com/r/stunfisk/comments/h9t99t/pokemon_battle_predictor_a_machine_learning/?utm_source=share&utm_medium=web3x&utm_name=web3xcss&utm_term=1&utm_content=share_button) win probabilities for VGC players on a turn-by-turn basis. That project evolved into creating an agent with impressive results against human players as well as a browser extension for players. Our project is different in that it focuses on the 2 vs 2 double battle format. It is also different in that predictions will be made before a single turn has been played, and not from any choices or events that occur mid-battle. That project used [TensorFlowJS](https://www.tensorflow.org/js/), and we will consider TensorFlow, though it would be in Python.

**Part A (Supervised Learning)**

Data consists of 10,000+ battle logs scraped from Pokémon Showdown. We will run logistic regression, random forests, and eventually neural networks, to predict the probability of a player winning given their and their opponent’s Pokémon team composition. Models will be evaluated on ROC-AUC score and F1-Score. Feature importances will also be reported. Final visualizations will use official Pokémon artwork through [PokeAPI](https://pokeapi.co/) calls (in [Pokepaste](https://pokepast.es/6d4003745e66f0ed) style).

**Part B (Unsupervised Learning)**

Dataset is the same as in the supervised portion. Unsupervised learning will identify common team compositions, which will be added as features to improve our supervised predictions. Categorical features will be represented in a sparse bag of words matrix – with rows being teams, and columns being one hot-encoded feature variables. PokeAPI will be used to supplement battle data with data on the base stats (e.g. health, attack, defence) of each Pokémon used on each team. Clustering will then be done with K-means, with ideal cluster count assessed using inertia and silhouette scores. Clusters will then be labeled and evaluated for quality using domain knowledge of team archetypes. Visualizations will use t-SNE to show cluster spacing, and graphics created using [PokeAPI with](https://pokeapi.co/) official Pokémon art.

**Team Planning:**

By the end of **week 3**, the two of us will have collaborated on similar parts of exploratory data analysis. By the end of **week 4**, Sean will have experimented with model parameters in order to gain insights via supervised learning. Jackson will have discovered insights into common clusters. By the end of **Week 6,** our model will be at least 95% ready for production. We will focus the last two weeks on final model touchess and cleaning up visualizations. We would like to dedicate the last week to presentation creation.