Project Proposal

To Find Legendary Pokemon Using Machine Learning

Siddhant Acharya

Pokemon is a renowned video game played by millions of people across the world. Pokemons are different characters having different abilities that battle out with each other. Pokemons are of different types as well. Most of them are normal (with limited abilities and strength and with no special powers), but the legendary (with special powers and abilities) ones are rare. In this project, I’ll be finding out these legendary Pokemons from the whole Pokedex (An index of all the pokemons) using Machine Learning algorithms like Random Forest Algorithm and Logistic Regression.

Goals:

1. Create a dataset of all the Pokemons.
2. Develop a model which can decide whether the pokemon is legendary or not?
3. Build a User Interface (UI) for the users to interact with the web app.
4. Deploying the web app on an open-source Cloud Platform for easier access.

Improvised goals:

1. Since the pokemon datasets are not that large, I might add additional features to it like the following:
   1. Battle-out two pokemons and see which one can win the battle.
   2. How different power selection can affect the winning rate of the pokemon?

Timelines:

1. April 4th: Start the project's detailed research work - Completed
2. April 7th: Start with the Project Setup and model building -Completed
3. April 20th: Finalize the Project model along with testing -Completed

Improvised Timelines:

1. April 25th: Gather more data on the powers of the pokemons.
2. April 30th: Try adding more features to the app.
3. May 2nd: Construct the UI of the web app.
4. May 6th: Deploy the App on the Cloud.
5. May 7th: Start writing the final report and start working on the final presentation.

Google search for similar projects:

1. Predicting Play Calls Using a Random Forest Classifier.
2. Project of game prediction based on LeBron’s stats using 3 machine learning models.
3. Logistic Regression project for loan default prediction.
4. Regression project to build a song popularity predictor.
5. Inventory stock demand forecasting.
6. Project for customer ad clicks.

Google Scholar search results:

1. Gradient boosting machine learning algorithms.
2. OpenStreetMap quality assessment using unsupervised machine learning methods.
3. Monsters, metaphors, and machine learning.
4. Augmented reality and machine learning incorporation using YOLOv3 and ARkit.
5. Using Machine Learning to Forecast and Improve clinical outcomes and health aging using sensor data.
6. Feature Engineering and Selection.

Possible Complications and Points of Failures:

1. Pokemon as a game is ever evolving and with each succeeding generation of pokemons passing a new dataset of pokemons is generated. Updation of the dataset is required.
2. A lot of the Pokemons have a few similarities with the legendary pokemons concerning their abilities. So, the model must take care of this edge case.
3. Sparse datasets could turn out to be an issue as Random Forest Algorithm is not that good at handling sparse data.
4. Many decisions can also make the model a bit slow.