CPSC 5700
Project Problem Statement
Pokémon Classifier
Group 2 - Pabi Dhaliwal, Eric Nunn, Heetae Yang
10/21/2019

Problem Statement:

Professor Oak's laboratory is designing a Pokédex to give amateur Pokémon trainers ability to identify wild Pokémon. The professor needs to identify Pokémon given a single image given by the trainer. By identifying the Pokémon, the trainer will be able to determine its basic information, weaknesses, and have a higher chance of catching them all. Potential outcomes could be identifying specific Pokémon from the given picture and categorize Pokémon.

Data Sources:

The data source we are using is a pre-labeled dataset from Kaggle.com which consists of 60 images for each Pokémon. We will need to resize all the images, so they all have the same dimensions.

Algorithms:

- Convolutional Neural Network (CNN) - Machine learning model for our image classification.
- Python - Programing languages we will use to build our CNN model.
- Keras - High-level machine learning library (TensorFlow back-end)
- Pandas - Used for cleaning and prepping the data as well creating graph data for analysis.
- Matplotlib - Plotting the graphs.
- Jupyter Notebook- Sharing code and notes among the group.

Risks:

- Not enough data to train the model. Not enough unique pictures of the Pokémon
- Inaccurate dataset (mis-labeled photo, duplicate pictures, irrelevant pictures)
- Low accuracy of CNN model data
- Inaccurate recording of CNN model performance

Challenge:

- Finding the optimal CNN architecture.
- Getting our CNN model to 90+ percent accuracy.
- Being able to recognize Pokémon in different poses, outfits, or variations.