Gotta Type ‘em All – An Analysis of Classification Algorithms

Tyler Rose and Seth Dippold

**Abstract**

We will use scikit-learn in Python to classify the data using SVM, ID3 Decision Trees, and Naïve Bayes. The dataset we will use includes every Pokémon from six generations with their respective stats and types. This data has been acquired from Kaggle which used many different valid sources. In addition to the scikit-learn documentation, we will use the documents provide in class and “An Empirical Comparison of Supervised Learning Algorithms” from Cornell University. The goal of this project is to determine which classification algorithm can more accurately predict the type of a Pokémon according to its stats: Attack, Defense, Special Attack, Special Defense, Hit Points, and Speed.

Kaggle Pokémon Dataset - <https://www.kaggle.com/abcsds/pokemon>-+

Scikit-learn - <http://scikit-learn.org/>

Comparison of algrithms - <https://www.cs.cornell.edu/~caruana/ctp/ct.papers/caruana.icml06.pdf>