Classification Models For Legendary Pokemon

Alex Ryan

May 9, 2018

Introduction

This is a project that focuses on classifying Non-Legendary Pokemon from Legendary Pokemon. I will explore the data with visualizations and explanations. I will also utilize knn, naive bayes, and xgbtree. These models will attempt to distinguish between these two types of pokemon.

Data Dictionary

The Pokemon dataset was taken from Kaggle. The pokemon included come from Generations 1-6. The dataset includes attributes such as id number, name, Type.1, Type.2, total, HP, Attack, Defense, Sp.Atk, Sp.Def, Speed, Generation, and Legendary. Added features will be discussed in later sections.

Variable Name:	Variable Description:
'id'	Identification number of pokemon
'name'	Name of Pokemon
'Type.1'	Primary property trait of pokemon
'Type.2'	Secondary property trait of pokemon
'total'	Total stats of pokemon
'HP'	Total Health of pokemon
'Attack'	Total Attack of pokemon
'Defense'	Total Defense of pokemon
'Sp.Atk'	Total Special Attack of pokemon
'Sp.Def'	Total Special Defense of pokemon
'Speed'	Total Speed of pokemon
'Generation'	The generation number of pokemon
'Legendary'	A true or false value that defines whether the pokemon is a legendary or not