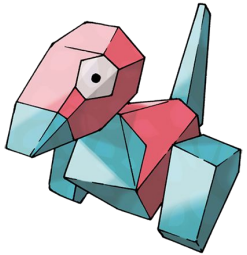
By Audrey, Chelsea, and Thora

Legendary Pokemon: Gotta Catch ‘em All

(with Machine Learning)

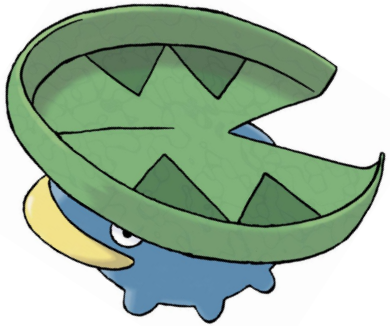







Project Goal

Predict if a Pokemon is legendary for new generations



What is Pokémon ?

Pokémon is a Japanese media franchise that depicts trainers and fictional monsters that can battle in video games, animated television series, card games, movies, mobile apps and other media.

What are “Legendary” Pokémon ?

(Virtually Gods)

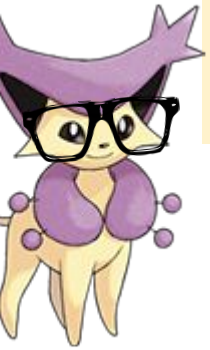
Legendary Pokémon are extremely rare and powerful. They often have to do with legends of creation and destruction in their native regions.






The Data

- Mythical Pokemon and Ultra Beasts considered Legendary Pokemon
- Made up of various statistics for Pokemon up to the seventh generation, including:
  - Base egg steps
  - Base happiness
  - All base stats
  - All base stats total
  - Capture rate
  - Experience growth
  - Percent male



Scrubbing the Data / Information We Didn't Use

- Legendaries can be all:
  - Types (electric, fairy, grass)
  - Weights
  - Heights

Memorization Issue

- Useless categories for pokemon generalization:
  - Name
  - Pokemon Classification
  - Pokedex Number
  - Abilities

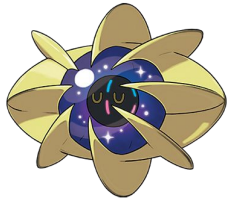
Challenges

- Categories we had issues with:
  - Capture Rate
  - Percentage Male

Jirachi: 2.4 pounds

Kyogre: 776 pounds

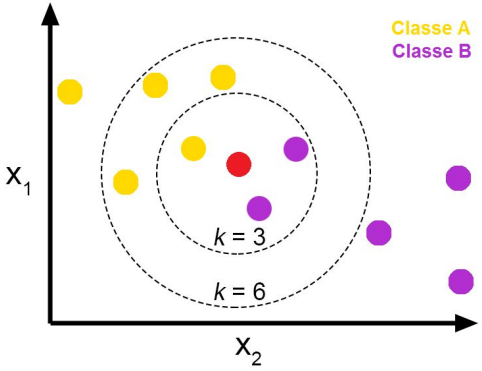
All Legendaries!!



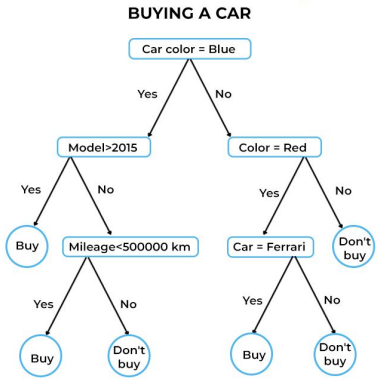
Cosmoem: 2,204 pounds

What We Ended Up Doing

- Tested several models (*k* nearest neighbor, decision tree, svm, etc.)
- Best was the decision tree
  - Roughly 99% testing accuracy across several training/testing splits
- Decision trees have the advantage of being easy to tell what is going on behind the scenes

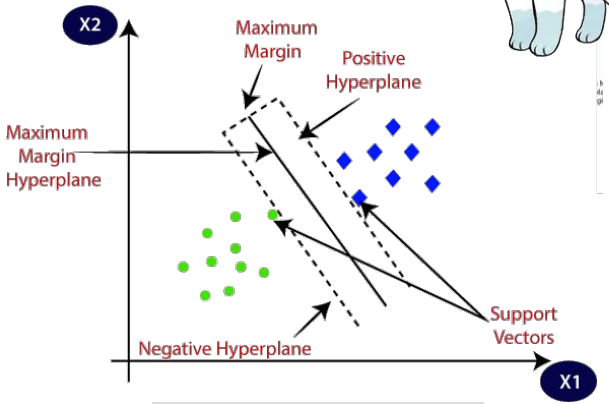


k Nearest Neighbor: 98.5%



Decision Tree: 99.5%

Random Forest: 98.5%

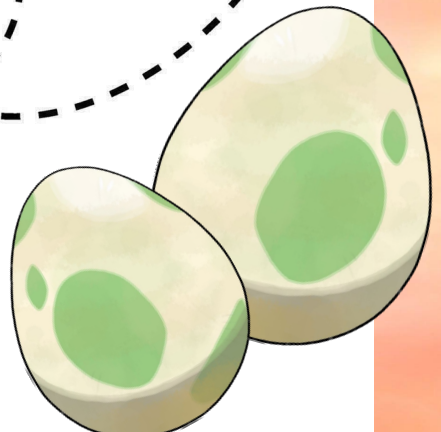


SVM: 99%

What We Found

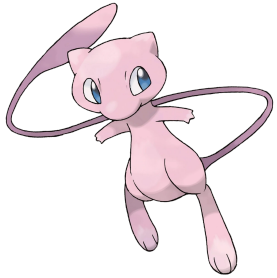
Egg steps are very important!

Then capture rate and total base stats…



Our Final Decision Tree

- Very broad because of low amount of nodes
- Highest accuracy of models using given data set



Testing the Model on New Generations (8 & 9)

- The decision tree model was 100% accurate at predicting if a new Pokemon was legendary!!
- It even worked for paradox Pokemon!

Paradox definition:

Pokemon found in Area Zero with high stats that are not classified as legendary pokemon

Confusion Matrix:

	Actually Legendary	Actually Non-Legendary
Tree Predicted Legendary	18	0
Tree Predicted Non-Legendary	0	12

Only One Legendary Pokemon Slipped Through

Why?

Unknown steps to hatch egg:

Pecharunt's: ???	Average: ~ 5,000

Insane base stats:

Pecharunt's: 600	Average: ~ 436

Very low capture rate:

Pecharunt's: 3	Average: ~ 100

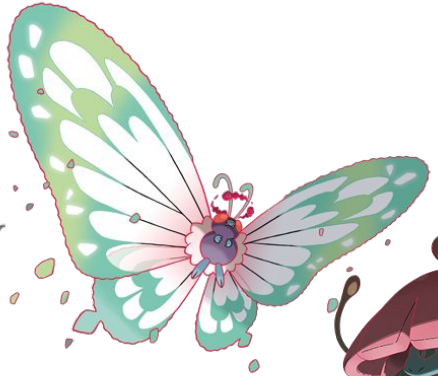
Generation 9: Pecharunt

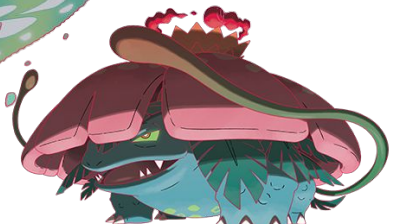
Technically it still worked:

If put through our decision tree as zero steps it would have still been classified as a legendary

What We May Do in the Future

- Our model works on almost 100% of current Pokemon
- Fix both Type: Null and Silvally
- In the future, Pokemon that the model will not correctly classify may be introduced (Weird classes like ultra beast)
- Historical Bias: The data set was last updated in 2016





Fin