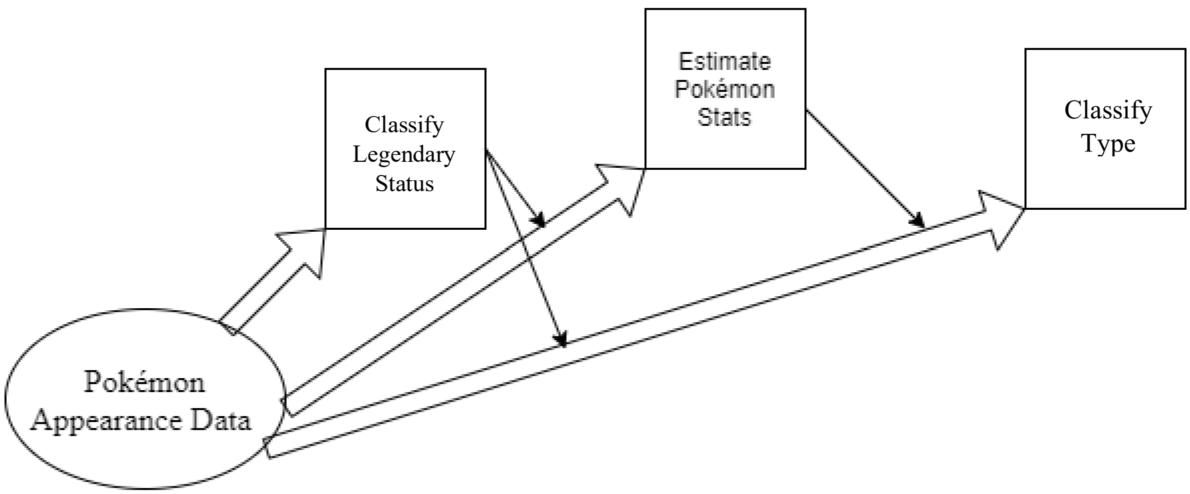
AI Pokédex Replacement

IST 707 Wesley Stevens

Introduction

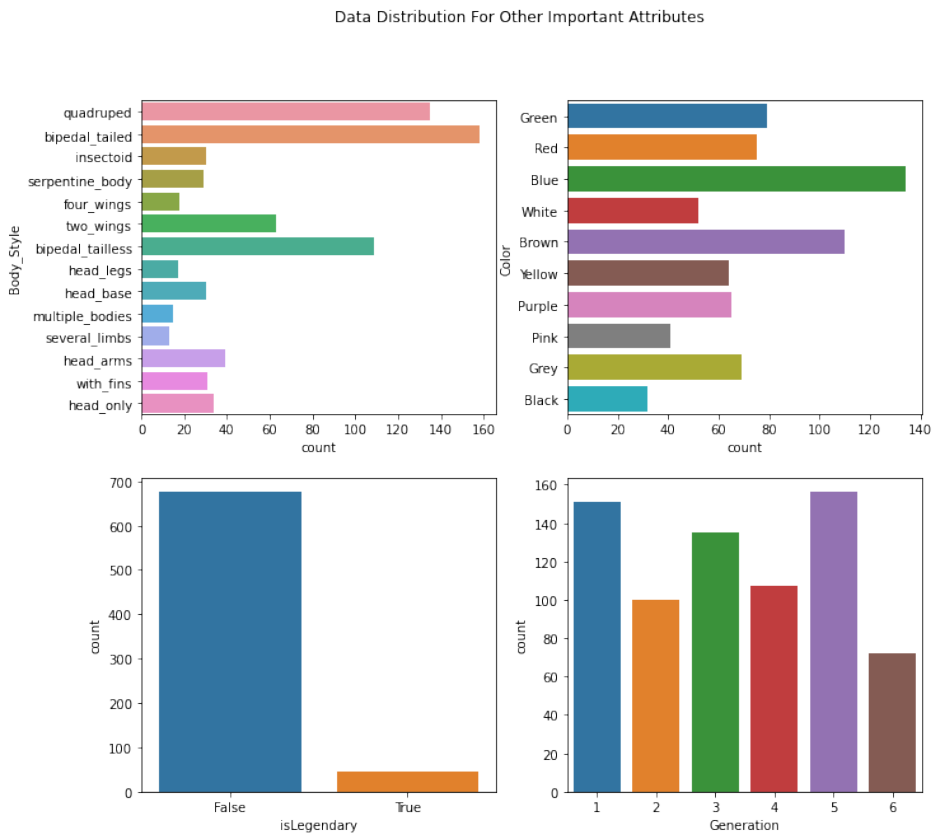
The world of Pokémon is large and diverse and those who deal with Pokémon are concerned about two things: battling and identification. Historically, Pokémon collectors have needed to rely upon a complete Pokémon scanning and recording system called the Pokédex. The Pokédex is usually carried about by collectors and takes a lot of time and energy to scan and index encountered Pokémon. We propose a machine learning solution to rapidly identify the type, estimate the stats, and identify the Legendary status of an encountered Pokémon. The proposed model will determine Legendary stats, estimate statistics, and then identify the type of Pokémon, using the gained information from each step in the next evaluation as shown in the figure below.

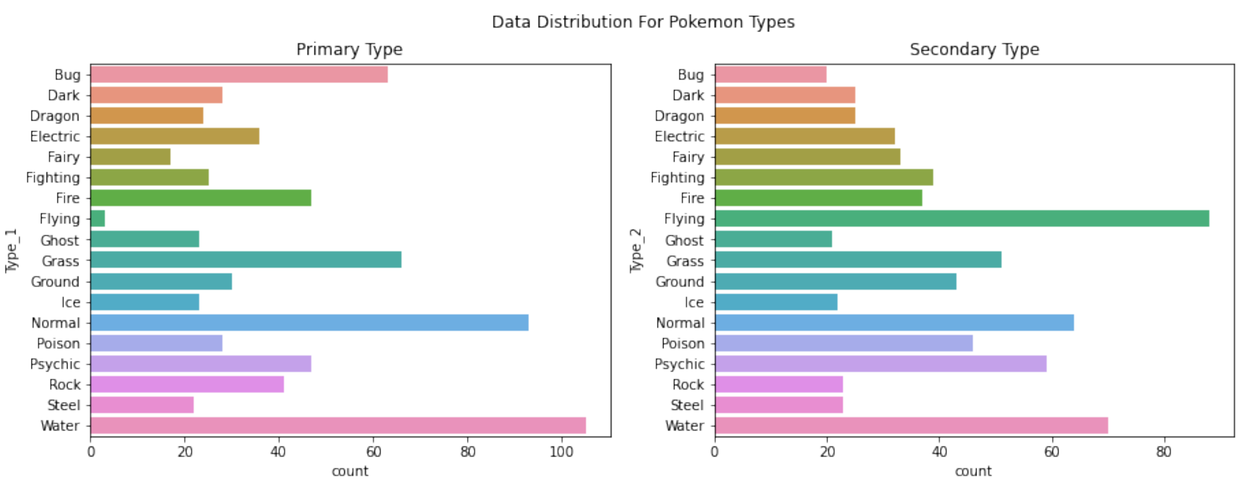


Exploratory Data Analysis & Preprocessing

The dataset of known Pokémon contained over 700 different Pokémon and contained missing values in 2 of the columns (Egg_Group_2 and Pr_Male) [1]. We dropped these columns due to the number of missing values, and dropped all other identification attributes (ID, number, name, etc.). We then normalized the continuous variables, which included the Pokémon's base stats (HP, Speed, etc.) and the Height and Weight. The rest of the variables were either binary or nominal, and thus were one-hot encoded. A visualization of the distribution of some of the visual attributes are shown in the first figure below. Pokémon base stats range from 1 to 255, have a mean of 125 and a standard deviation of 31 [2].

Pokémon may have one or two types and certain types are much more common as a primary type than secondary type as shown in the second figure below. Thus, for this classification, we built two models: one to identify the first (primary) type and one to classify the secondary type. Pokémon with only one type will be considered to have the same primary and secondary types. Since we are doing our process in stages, the dataset for each section contains a slightly different number of columns, all with 721 observations. For each evaluation, the results will be the result of a 7-fold cross validation on a 75-25 split. The final dataset for the legendary classification model contains 29 attributes, the final dataset for the statistics classification contains 35 attributes, and the type classification dataset contains 37 attributes.





Methods

While we experimented with many algorithms for each of the three tasks, including Naïve Bayes, K-Means, XGBoost, SVM, and ANNs, we will only report on the top performing algorithms. For each method, we set aside training, testing, and validation datasets. We also experimented with feature selection using scikit-learn's RFE feature selection tool to overcome the curse of dimensionality.

For the XGBoost classifier, we experimented with the number of estimators and several lambda coefficient values. The optimal values were the default values for XGBoost in Python.

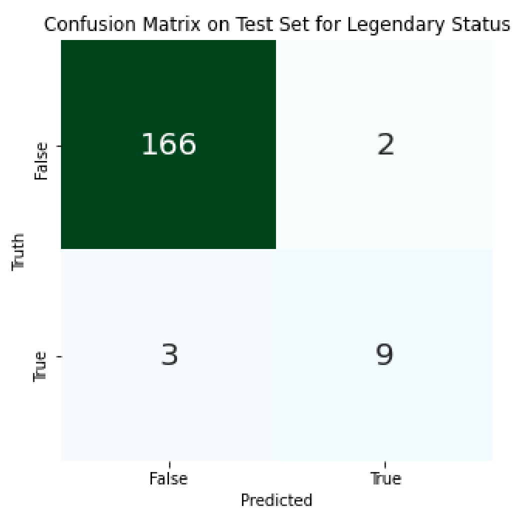
For the SVM classifier and regressor, we did a parameter sweep along kernel types, regularization parameters, and degree of polynomial for the polynomial kernel.

For the ANN, we experimented with a variety of depths, activation functions, number of nodes, and normalization functions. We even evaluated performance with a RNN architecture. The best performing architecture, however, was a single layer feed forward network with a Layer normalization and a Softplus activation function. The hidden layer contained 26 nodes. Additionally, we used the RMSprop optimizer with a learning rate of 0.0015 and a Cosine Annealing lr scheduler. We also used early stopping on the validation Cross Entropy loss with a patience of 15 and a batch size of 8.

Results

Legendary Pokémon Classification

For this initial classification, we only used visual attributes of the Pokémon to do the classification. We found that the XGBoost model gave the best results at 97.2% accuracy, 90% precision, 86.9% recall, and 88.4% F-measure. The confusion matrix on the test set is shown below. This is much better than a baseline model of predicting always non-Legendary, which would give you an accuracy of 90%. Thus, we say that we can accurately predict a Pokémon's legendary status.



Pokémon Statistics Regression

We evaluate each model's performance by the MSE metric since MSE is a good standard indicator of how close a regression output is to the truth. The smaller an MSE value, the better.

As a baseline, each of the Pokémon's statistics can be represented with a normal distribution with a standard deviation of 31. Thus, the acceptable range of MSE is from 0 – 1.1.

For regression on each of the statistics, we built 6 models for each of the 6 different statistic attributes (HP, Speed, Attack, Defense, Special Attack, and Special Defense) since each statistic is very different in possible ranges. We chose to try the XGBoost, Linear Regression, and SVM regression algorithms to do this task. The comparison results are showcased in the table below.

Regressor Comparison using the MSE metric

	HP	Speed	Attack	Defense	Sp Attack	Sp Defense
XGBoost	0.007582	0.029	0.02118	0.014497	0.03314	0.014675
Linear Regression	0.007038	0.025981	0.023484	0.012285	0.02905	0.014259
SVM Regression	0.007597	0.024273	0.024655	0.012877	0.031679	0.014614

As we see in the table, each model does exceptionally well at building a regression model for each task. It seems that Linear Regression is better suited for regressing on HP, Defense, Special Attack, and Special Defense. It seems that XGBoost is best suited for regressing on Attack, and the SVM Regressor is best for predicting Speed. Since each Pokémon can have a range of stats, and with how small these MSE values are, we can say with confidence that these regression models can accurately predict the statistics of a Pokémon within the expected error tolerance.

Pokémon Type Classification

For most of the methods we tried (namely ANN, XGBoost, and Kmeans), classification of these types proved extremely difficult, scoring at best 40% accuracy with the ANN. While still better than random guessing, these models are not usable in a production setting. Thus we were pleasantly surprised when our results for this classification with the SVM gave better results. We used an RBF kernel with a C value of 10000. The statistics and confusion matrices containing our results are shown in the figures and tables below. We were able to achieve 83.3% accuracy on primary type classifications and 82.1% accuracy on secondary type classification, which is more than double the best classification score we could get with the other models.

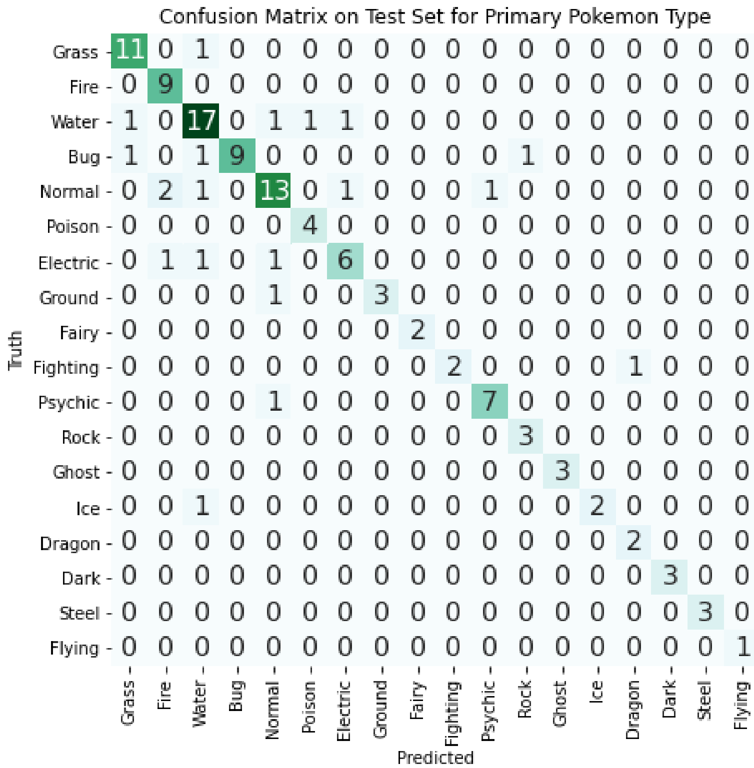
Classification Statistics for Primary Pokémon Type

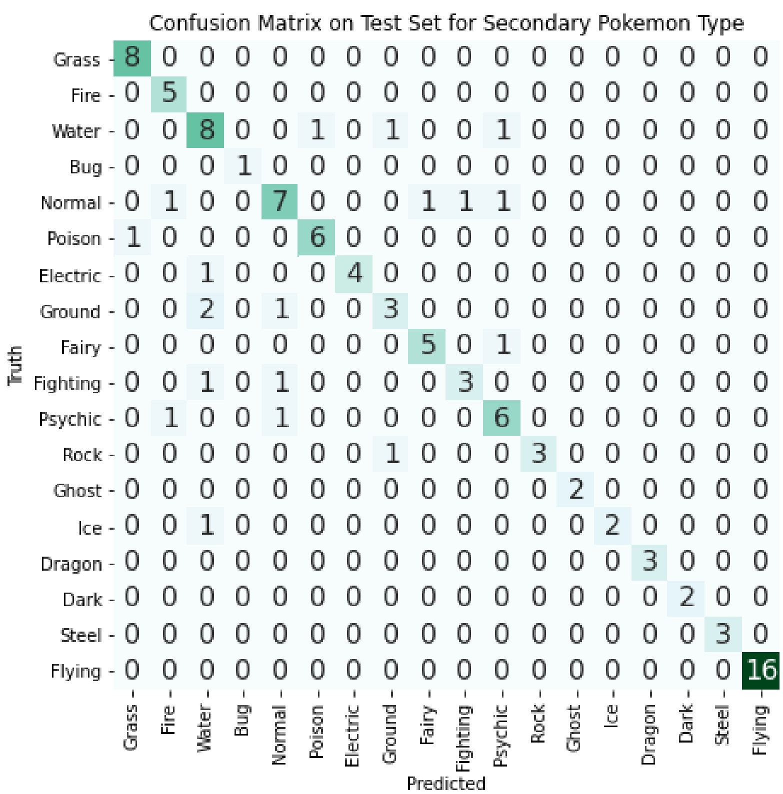
	Precision	Recall	F-Measure
Average	88.8%	87.9%	87.3%
Weighted Average	83.3%	85.0%	83.4%

Classification Statistics for Secondary Pokémon Type

	Precision	Recall	F-Measure
Average	86.8%	84.0%	84.8%
Weighted Average	82.1%	83.3%	82.2%

It is important to have two separate models for the primary and secondary type classifications since Pokémon heavily rely on more common types as their primary type. We tested this hypothesis and found it to be true in our experiments. The best we could get a model trained on the primary types to predict secondary types was 73.8% accuracy with a similar precision, recall, and F-measure.





Temporal Complexity

The question still remains if this pipeline is actually faster than using a Pokédex. We took an experienced Pokédex user and timed him to determine how fast one could scan a Pokémon with a Pokédex machine [3]. Then we timed our model pipeline and determined that, as seen in the comparison table below, our pipeline is indeed faster by 5.37 seconds. That is approximately 42 times faster than the previous method.

Pokedex	AI Method
5.5 seconds	0.13 seconds

Conclusion

The Pokédex is a cumbersome instrument that takes a lot of time and energy to evaluate Pokémon and record its attributes, inherent statistics, and types. In this work, we have shown that everything the Pokédex does can accurately be either observed by the user or estimated by the 9 statistical models we have built in a pipeline. This pipeline is 42 times faster than the Pokédex scanning. The accuracy of the predicted attributes varies at each stage, but all are done at a rate of greater than 82%, and in most cases, greater than 90%.

The advantage of this novel product is that a Pokémon's attributes can be accurately estimated and recorded in a Pokédex database without the user needing to capture the Pokémon and feed it through the Pokédex. This saves on time for both collectors who want to know what varieties of Pokémon live in the world, and Pokémon battlers that want to find the best Pokémon in the world to fight with. With this tool, Pokémon enthusiasts can now save money spent on Pokéballs and other related equipment, save energy for themselves from carrying the heavy Pokédex and electricity spent using it, and save time identifying wild Pokémon.

References

[1] "Pokemon with stats," [Online]. Available: https://www.kaggle.com/abcsds/pokemon. [Accessed 9 2020].

[2] "Base Stats," Bulbapedia, [Online]. Available: https://bulbapedia.bulbagarden.net/wiki/Base_stats. [Accessed 14 11 2020].

[3] "New pokemon go account! How to complete your pokedex in one day? Before generation 2 comes," [Online]. Available: https://www.youtube.com/watch?v=yvboiG26vVY&ab_channel=SlytheNeko. [Accessed 14 11 2020].