Naïve Bayes Classification of Pokémon Types

Richa Patel

In this project, I implemented a Naïve Bayes classifier from scratch to predict Pokémon types (Normal, Water, or Flying) based on features such as HP, Attack, and Defense.

The dataset was sourced from the given link and was divided into 275 training instances and 22 test instances. The classifier was designed to use prior probabilities and likelihood estimates to make predictions.

Methodology:

The classifier follows the fundamental principles of **Naïve Bayes classification**, assuming that each feature is independent given the class. The main steps include:

a) Prior Probabilities (P(Type))

We calculate the probability of each Pokémon type occurring in the dataset:

$P(Type) = Number\ of\ instances\ of\ Type\ /\ Total\ instances\ in\ training\ set$

b) **Likelihood (P(Feature | Type)) with Laplace Smoothing**

Since HP, Attack, and Defense are categorical (High, Mid, Low), we compute the likelihood of each feature value within each class using:

$$P(Feature\ |\ Type) =$$

$$Total\ class\ instances + 1\ /Number\ of\ unique\ feature\ values$$

$$Count\ of\ feature\ in\ class$$

c) Prediction

For each Pokémon in the test set, we calculate:

$$P(Type\ |\ Features) = P(Type) \times P(HP\ |\ Type) \times P(Attack\ |\ Type) \times P(Defense\ |\ Type)$$

The class with the **highest probability** is selected as the predicted type

Results:

The classifier successfully predicted Pokémon types based on **HP, Attack, and Defense**. Sample outputs are:

Pokémon	Predicted Type	Probability
Pidgeotto	Normal	0.0157
Zubat	Flying	0.0165
Suicune	water	0.0294

Conclusion

This project demonstrates how **Naïve Bayes** can be applied to **categorical data classification**. The model successfully classified Pokémon types, and **Laplace Smoothing** ensured that all feature values contributed effectively to the predictions.

For future improvements:

- **Using more features** (e.g., Speed, Special Attack).

- **Applying other classification models** (e.g., Decision Trees, k-NN) for comparison.

Overall, the classifier effectively predicts Pokémon types based on limited data and provides a solid foundation for future improvements.