Blake Masters
CSC 466
Lab 2
February 2, 2025

Lab 2 Report

Total Results Tabled

Mode	Dataset	Ideal Threshold	CV-Accuracy	Gain/Ratio	Info Gain(s)	Ratio(s)
Student	Iris	0.1	0.94	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]
Sci-Kit	Iris	0.1	0.92	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]
Student	Nursery	0.1	0.985	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]
Sci-Kit	Nursery	0.1	0.6625	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]
Student	Letter Recognition	0.1	0.8815	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]
Sci-Kit	Letter Recognition	0.1	0.2348	Gain	[0.1,0.3,0.5]	[0.1,0.3,0.5]

Notes: My model is performing better at a near alarming rate. I have verified the accuracy scores for SK, and I am attributing the large difference due to likely misuse of the framework and functional libraries. While my LR took about 80 minutes, Sci-Kit took less than 10 seconds, indicating there is much to be desired for efficiency of runtime of my cross validation techniques.

We can see in the datasets that Nursery has 100% miss categories with “recommend” in both implementations due to the tree structure missing by threshold. We also see this with “very-recommend” in SK learn below. Iris is pretty well clustered, so it doesn't miss by much more than 10%.

Conclusions

While the trees had similar structures, they were not identical. Again, the Nursery trees show the clearest display of difference due to the threshold missing a large portion of very_recommends while my tree managed to catch it with a large quantity. Other small “op” discrepancies in values were present to fractional amounts, but this swung some points even in iris-virginica. I have not compared the Letter Recognition Trees due to the scale and impossibility of reading beyond surface level in a reasonable amount of time. In the general case, we see my sci-kit learn implementation’s accuracy drop harshly with larger datasets as my slow dataframe based approach holds strong, dropping less than 10% across over 100x the size of the datasets, even showing improvement in Nursery over Iris, unlike the sk-learn implementation.

Note the low accuracy of the sci-kit model for Letter Recognition and the spotty diagonal built with some letters like ‘R’ being detected strongly as ‘D’ or ‘N’ and ‘M’ falling under ‘W’. This makes physical sense as the letters appear with similar structural tendency, but show a massive decrease in accuracy compared to my model

Cross Validation Results

given the following hyperparams for cross validation

```
hyperparams = {
    "InfoGain": [0.1, 0.3, 0.5],
    "Ratio": [0.1, 0.3, 0.5]
}
```

Iris

```
bemaster@127x07:~/kdd/lab2 $ python3 crossVal.py Data/Iris/iris.data.csv hyperparams.json

Best Model Hyperparameters:
Splitting Metric: Gain, Threshold: 0.1
Overall Cross-Validation Accuracy: 0.9400
Confusion Matrix:
Iris-setosa: {'Iris-setosa': 50, 'Iris-versicolor': 0, 'Iris-virginica': 0}
Iris-versicolor: {'Iris-setosa': 0, 'Iris-versicolor': 46, 'Iris-virginica': 4}
Iris-virginica: {'Iris-setosa': 0, 'Iris-versicolor': 5, 'Iris-virginica': 45}
```

Nursery

```
bemaster@127x07:~/kdd/lab2 $ python3 crossVal.py Data/Nursery/nursery.csv hyperparams.json

Best Model Hyperparameters:
Splitting Metric: Gain, Threshold: 0.1
Overall Cross-Validation Accuracy: 0.9850
Confusion Matrix:
not_recom: {'not_recom': 4320, 'priority': 0, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}
priority: {'not_recom': 0, 'priority': 4155, 'recommend': 0, 'spec_prior': 56, 'very_recom': 55}
recommend: {'not_recom': 0, 'priority': 0, 'recommend': 0, 'spec_prior': 0, 'very_recom': 2}
spec_prior: {'not_recom': 0, 'priority': 47, 'recommend': 0, 'spec_prior': 3997, 'very_recom': 0}
very_recom: {'not_recom': 0, 'priority': 24, 'recommend': 10, 'spec_prior': 0, 'very_recom': 294}
```

Letter Recognition

```
bemaster@127x07:~/kdd/lab2 $ python3 crossVal.py Data/LR/letter-recognition.data.csv hyperparams.json

Best Model Hyperparameters:
Splitting Metric: Gain, Threshold: 0.1
Overall Cross-Validation Accuracy: 0.8815
Confusion Matrix:
```

```
bemaster@127x07:~/kdd/lab2 $ python3 crossVal.py Data/LR/letter-recognition.data.csv hyperparams.json

Best Model Hyperparameters:
Splitting Metric: Gain, Threshold: 0.1
Overall Cross-Validation Accuracy: 0.8815
Confusion Matrix:
A: {'A': 754, 'B': 0, 'C': 0, 'D': 2, 'E': 0, 'F': 0, 'G': 2, 'H': 0, 'I': 1, 'J': 2, 'K': 1, 'L': 9, 'M': 1, 'N': 0, 'O': 1, 'P': 0, 'Q': 6, 'R': 2, 'S': 2, 'T': 0, 'U': 1, 'V': 0, 'W': 0, 'X': 0, 'Y': 4, 'Z': 1}
B: {'A': 1, 'B': 641, 'C': 0, 'D': 11, 'E': 7, 'F': 4, 'G': 5, 'H': 10, 'I': 3, 'J': 3, 'K': 3, 'L': 2, 'M': 1, 'N': 4, 'O': 5, 'P': 5, 'Q': 2, 'R': 21, 'S': 9, 'T': 3, 'U': 2, 'V': 11, 'W': 2, 'X': 6, 'Y': 1, 'Z': 4}
C: {'A': 1, 'B': 0, 'C': 650, 'D': 1, 'E': 6, 'F': 6, 'G': 27, 'H': 2, 'I': 0, 'J': 1, 'K': 5, 'L': 5, 'M': 1, 'N': 1, 'O': 5, 'P': 1, 'Q': 5, 'R': 7, 'S': 0, 'T': 5, 'U': 3, 'V': 0, 'W': 0, 'X': 2, 'Y': 1, 'Z': 1}
D: {'A': 1, 'B': 14, 'C': 0, 'D': 702, 'E': 0, 'F': 1, 'G': 4, 'H': 17, 'I': 4, 'J': 3, 'K': 3, 'L': 4, 'M': 2, 'N': 8, 'O': 4, 'P': 7, 'Q': 6, 'R': 8, 'S': 3, 'T': 3, 'U': 3, 'V': 1, 'W': 0, 'X': 2, 'Y': 2, 'Z': 3}
E: {'A': 0, 'B': 6, 'C': 9, 'D': 1, 'E': 676, 'F': 3, 'G': 11, 'H': 1, 'I': 0, 'J': 2, 'K': 11, 'L': 5, 'M': 1, 'N': 0, 'O': 1, 'P': 3, 'Q': 3, 'R': 3, 'S': 8, 'T': 4, 'U': 0, 'V': 0, 'W': 0, 'X': 10, 'Y': 3, 'Z': 7}
F: {'A': 1, 'B': 2, 'C': 1, 'D': 3, 'E': 4, 'F': 654, 'G': 0, 'H': 4, 'I': 9, 'J': 5, 'K': 0, 'L': 2, 'M': 1, 'N': 1, 'O': 0, 'P': 37, 'Q': 1, 'R': 2, 'S': 9, 'T': 12, 'U': 1, 'V': 5, 'W': 1, 'X': 6, 'Y': 12, 'Z': 2}
G: {'A': 3, 'B': 7, 'C': 13, 'D': 6, 'E': 15, 'F': 4, 'G': 662, 'H': 8, 'I': 1, 'J': 1, 'K': 3, 'L': 4, 'M': 1, 'N': 0, 'O': 5, 'P': 0, 'Q': 19, 'R': 6, 'S': 6, 'T': 2, 'U': 1, 'V': 2, 'W': 4, 'X': 0, 'Y': 0, 'Z': 0}
H: {'A': 3, 'B': 9, 'C': 3, 'D': 39, 'E': 0, 'F': 0, 'G': 7, 'H': 581, 'I': 0, 'J': 6, 'K': 25, 'L': 2, 'M': 2, 'N': 2, 'O': 11, 'P': 4, 'Q': 2, 'R': 13, 'S': 7, 'T': 0, 'U': 4, 'V': 3, 'W': 2, 'X': 4, 'Y': 4, 'Z': 1}
I: {'A': 1, 'B': 3, 'C': 1, 'D': 5, 'E': 0, 'F': 11, 'G': 1, 'H': 1, 'I': 683, 'J': 24, 'K': 1, 'L': 1, 'M': 0, 'N': 1, 'O': 0, 'P': 8, 'Q': 0, 'R': 1, 'S': 6, 'T': 1, 'U': 1, 'V': 0, 'W': 0, 'X': 5, 'Y': 0, 'Z': 0}
J: {'A': 3, 'B': 5, 'C': 0, 'D': 3, 'E': 1, 'F': 5, 'G': 0, 'H': 3, 'I': 22, 'J': 666, 'K': 2, 'L': 5, 'M': 0, 'N': 1, 'O': 3, 'P': 3, 'Q': 3, 'R': 2, 'S': 5, 'T': 0, 'U': 1, 'V': 0, 'W': 0, 'X': 6, 'Y': 1, 'Z': 7}
K: {'A': 3, 'B': 3, 'C': 3, 'D': 4, 'E': 8, 'F': 2, 'G': 2, 'H': 21, 'I': 0, 'J': 5, 'K': 632, 'L': 2, 'M': 3, 'N': 2, 'O': 0, 'P': 1, 'Q': 2, 'R': 25, 'S': 2, 'T': 4, 'U': 1, 'V': 0, 'W': 0, 'X': 12, 'Y': 0, 'Z': 2}
L: {'A': 5, 'B': 2, 'C': 6, 'D': 0, 'E': 8, 'F': 1, 'G': 5, 'H': 3, 'I': 3, 'J': 2, 'K': 4, 'L': 705, 'M': 0, 'N': 1, 'O': 3, 'P': 0, 'Q': 2, 'R': 4, 'S': 3, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 3, 'Y': 1, 'Z': 0}
M: {'A': 2, 'B': 0, 'C': 1, 'D': 1, 'E': 1, 'F': 1, 'G': 2, 'H': 4, 'I': 0, 'J': 0, 'K': 3, 'L': 0, 'M': 746, 'N': 12, 'O': 2, 'P': 0, 'Q': 1, 'R': 3, 'S': 2, 'T': 0, 'U': 2, 'V': 4, 'W': 4, 'X': 0, 'Y': 1, 'Z': 0}
N: {'A': 0, 'B': 3, 'C': 1, 'D': 14, 'E': 0, 'F': 1, 'G': 2, 'H': 7, 'I': 1, 'J': 1, 'K': 1, 'L': 2, 'M': 7, 'N': 702, 'O': 7, 'P': 0, 'Q': 0, 'R': 11, 'S': 0, 'T': 0, 'U': 11, 'V': 3, 'W': 6, 'X': 2, 'Y': 1, 'Z': 0}
O: {'A': 3, 'B': 1, 'C': 3, 'D': 21, 'E': 1, 'F': 5, 'G': 6, 'H': 6, 'I': 1, 'J': 3, 'K': 3, 'L': 0, 'M': 1, 'N': 5, 'O': 646, 'P': 1, 'Q': 20, 'R': 3, 'S': 4, 'T': 5, 'U': 6, 'V': 3, 'W': 4, 'X': 0, 'Y': 0, 'Z': 2}
P: {'A': 1, 'B': 2, 'C': 0, 'D': 2, 'E': 2, 'F': 21, 'G': 4, 'H': 9, 'I': 6, 'J': 1, 'K': 1, 'L': 1, 'M': 0, 'N': 5, 'O': 3, 'P': 719, 'Q': 2, 'R': 3, 'S': 2, 'T': 0, 'U': 0, 'V': 4, 'W': 3, 'X': 3, 'Y': 5, 'Z': 4}
Q: {'A': 3, 'B': 6, 'C': 4, 'D': 2, 'E': 6, 'F': 1, 'G': 12, 'H': 2, 'I': 1, 'J': 3, 'K': 3, 'L': 2, 'M': 2, 'N': 1, 'O': 18, 'P': 4, 'Q': 681, 'R': 9, 'S': 1, 'T': 2, 'U': 3, 'V': 1, 'W': 1, 'X': 3, 'Y': 6, 'Z': 6}
R: {'A': 2, 'B': 22, 'C': 3, 'D': 9, 'E': 0, 'F': 3, 'G': 4, 'H': 14, 'I': 2, 'J': 2, 'K': 19, 'L': 3, 'M': 0, 'N': 6, 'O': 3, 'P': 2, 'Q': 5, 'R': 644, 'S': 5, 'T': 1, 'U': 2, 'V': 2, 'W': 2, 'X': 3, 'Y': 0, 'Z': 0}
S: {'A': 6, 'B': 18, 'C': 0, 'D': 4, 'E': 11, 'F': 8, 'G': 9, 'H': 10, 'I': 11, 'J': 3, 'K': 2, 'L': 5, 'M': 0, 'N': 0, 'O': 3, 'P': 3, 'Q': 7, 'R': 6, 'S': 615, 'T': 6, 'U': 1, 'V': 2, 'W': 0, 'X': 4, 'Y': 3, 'Z': 11}
T: {'A': 1, 'B': 2, 'C': 5, 'D': 6, 'E': 6, 'F': 15, 'G': 7, 'H': 1, 'I': 1, 'J': 1, 'K': 1, 'L': 7, 'M': 0, 'N': 1, 'O': 2, 'P': 1, 'Q': 3, 'R': 1, 'S': 5, 'T': 702, 'U': 0, 'V': 5, 'W': 0, 'X': 3, 'Y': 16, 'Z': 4}
U: {'A': 2, 'B': 1, 'C': 3, 'D': 7, 'E': 0, 'F': 0, 'G': 6, 'H': 5, 'I': 0, 'J': 4, 'K': 2, 'L': 0, 'M': 9, 'N': 25, 'O': 4, 'P': 1, 'Q': 4, 'R': 0, 'S': 0, 'T': 2, 'U': 729, 'V': 4, 'W': 2, 'X': 2, 'Y': 1, 'Z': 0}
V: {'A': 0, 'B': 12, 'C': 2, 'D': 0, 'E': 0, 'F': 7, 'G': 5, 'H': 2, 'I': 0, 'J': 1, 'K': 0, 'L': 1, 'M': 7, 'N': 2, 'O': 2, 'P': 10, 'Q': 0, 'R': 0, 'S': 1, 'T': 2, 'U': 2, 'V': 682, 'W': 17, 'X': 0, 'Y': 9, 'Z': 0}
W: {'A': 0, 'B': 0, 'C': 2, 'D': 1, 'E': 0, 'F': 1, 'G': 1, 'H': 3, 'I': 0, 'J': 0, 'K': 2, 'L': 0, 'M': 5, 'N': 4, 'O': 6, 'P': 3, 'Q': 1, 'R': 1, 'S': 0, 'T': 1, 'U': 3, 'V': 8, 'W': 706, 'X': 0, 'Y': 3, 'Z': 1}
X: {'A': 1, 'B': 8, 'C': 0, 'D': 4, 'E': 9, 'F': 2, 'G': 0, 'H': 8, 'I': 3, 'J': 3, 'K': 21, 'L': 5, 'M': 0, 'N': 2, 'O': 0, 'P': 2, 'Q': 0, 'R': 6, 'S': 10, 'T': 1, 'U': 0, 'V': 0, 'W': 0, 'X': 697, 'Y': 1, 'Z': 4}
Y: {'A': 3, 'B': 1, 'C': 1, 'D': 1, 'E': 1, 'F': 11, 'G': 3, 'H': 4, 'I': 2, 'J': 4, 'K': 1, 'L': 0, 'M': 1, 'N': 4, 'O': 0, 'P': 6, 'Q': 2, 'R': 0, 'S': 3, 'T': 19, 'U': 3, 'V': 8, 'W': 1, 'X': 2, 'Y': 702, 'Z': 3}
Z: {'A': 3, 'B': 4, 'C': 0, 'D': 9, 'E': 14, 'F': 3, 'G': 3, 'H': 0, 'I': 0, 'J': 2, 'K': 0, 'L': 1, 'M': 0, 'N': 1, 'O': 3, 'P': 2, 'Q': 9, 'R': 2, 'S': 13, 'T': 2, 'U': 1, 'V': 0, 'W': 0, 'X': 3, 'Y': 5, 'Z': 654}
```

Cross Validation Sci-Kit Learn Results

Iris

```
bemaster@127x07:~/kdd/lab2 $ python3 crossValSKL.py Data/Iris/iris.data.csv hyperparams.json
Using target column: species
Warning: Ignoring Ratio hyperparameters; using InfoGain only.

Best Model Hyperparameters (Scikit-Learn):
Splitting Metric: Gain (sklearn: 'entropy')
Threshold (min_impurity_decrease): 0.1
Overall CV Accuracy: 0.9200
Confusion Matrix:
Iris-setosa: {'Iris-setosa': 50, 'Iris-versicolor': 0, 'Iris-virginica': 0}
Iris-versicolor: {'Iris-setosa': 0, 'Iris-versicolor': 46, 'Iris-virginica': 4}
Iris-virginica: {'Iris-setosa': 0, 'Iris-versicolor': 8, 'Iris-virginica': 42}
```

Nursery

```
bemaster@127x07:~/kdd/lab2 $ python3 crossValSKL.py Data/Nursery/nursery.csv hyperparams.json
Using target column: class
Warning: Ignoring Ratio hyperparameters; using InfoGain only.

Best Model Hyperparameters (Scikit-Learn):
Splitting Metric: Gain (sklearn: 'entropy')
Threshold (min_impurity_decrease): 0.1
Overall CV Accuracy: 0.6625
Confusion Matrix:
not_recom: {'not_recom': 4320, 'priority': 0, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}
priority: {'not_recom': 0, 'priority': 4266, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}
recommend: {'not_recom': 0, 'priority': 2, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}
spec_prior: {'not_recom': 0, 'priority': 4044, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}
very_recom: {'not_recom': 0, 'priority': 328, 'recommend': 0, 'spec_prior': 0, 'very_recom': 0}

Final model trained on the entire dataset.
```

Letter-Recognition

```
Best Model Hyperparameters (Scikit-Learn):
Splitting Metric: Gain (sklearn: 'entropy')
Threshold (min_impurity_decrease): 0.1
Overall CV Accuracy: 0.2348
Confusion Matrix:
```

```
bemaster@127x07:~/kdd/lab2 $ python3 crossValSKL.py Data/LR/letter-recognition.data.csv hyperparams.json
Using target column: lettr
Warning: Ignoring Ratio hyperparameters; using InfoGain only.

Best Model Hyperparameters (Scikit-Learn):
Splitting Metric: Gain (sklearn: 'entropy')
Threshold (min_impurity_decrease): 0.1
Overall CV Accuracy: 0.2348
Confusion Matrix:
A: {'A': 0, 'B': 7, 'C': 0, 'D': 53, 'E': 0, 'F': 0, 'G': 29, 'H': 0, 'I': 0, 'J': 326, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 10, 'Q': 168, 'R': 0, 'S': 0, 'T': 0, 'U': 55, 'V': 0, 'W': 0, 'X': 141, 'Y': 0, 'Z': 0}
B: {'A': 0, 'B': 54, 'C': 0, 'D': 562, 'E': 4, 'F': 0, 'G': 33, 'H': 0, 'I': 0, 'J': 3, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 60, 'Q': 36, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 14, 'Y': 0, 'Z': 0}
C: {'A': 0, 'B': 0, 'C': 0, 'D': 1, 'E': 209, 'F': 0, 'G': 298, 'H': 0, 'I': 0, 'J': 33, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 8, 'Q': 90, 'R': 0, 'S': 0, 'T': 36, 'U': 0, 'V': 0, 'W': 0, 'X': 61, 'Y': 0, 'Z': 0}
D: {'A': 0, 'B': 71, 'C': 0, 'D': 559, 'E': 4, 'F': 0, 'G': 14, 'H': 0, 'I': 0, 'J': 48, 'K': 0, 'L': 0, 'M': 1, 'N': 0, 'O': 0, 'P': 60, 'Q': 16, 'R': 0, 'S': 0, 'T': 0, 'U': 17, 'V': 0, 'W': 0, 'X': 15, 'Y': 0, 'Z': 0}
E: {'A': 0, 'B': 0, 'C': 0, 'D': 1, 'E': 292, 'F': 0, 'G': 207, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 0, 'Q': 84, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 184, 'Y': 0, 'Z': 0}
F: {'A': 0, 'B': 7, 'C': 0, 'D': 57, 'E': 0, 'F': 0, 'G': 29, 'H': 0, 'I': 0, 'J': 4, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 281, 'Q': 20, 'R': 0, 'S': 0, 'T': 369, 'U': 4, 'V': 0, 'W': 1, 'X': 3, 'Y': 0, 'Z': 0}
G: {'A': 0, 'B': 4, 'C': 0, 'D': 36, 'E': 103, 'F': 0, 'G': 326, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 20, 'Q': 270, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 14, 'Y': 0, 'Z': 0}
H: {'A': 0, 'B': 12, 'C': 0, 'D': 137, 'E': 29, 'F': 0, 'G': 80, 'H': 0, 'I': 0, 'J': 17, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 48, 'Q': 131, 'R': 0, 'S': 0, 'T': 0, 'U': 202, 'V': 0, 'W': 0, 'X': 78, 'Y': 0, 'Z': 0
}
I: {'A': 0, 'B': 7, 'C': 0, 'D': 73, 'E': 47, 'F': 0, 'G': 10, 'H': 0, 'I': 0, 'J': 466, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 53, 'Q': 17, 'R': 0, 'S': 0, 'T': 16, 'U': 0, 'V': 0, 'W': 0, 'X': 66, 'Y': 0, 'Z': 0}
J: {'A': 0, 'B': 8, 'C': 0, 'D': 88, 'E': 1, 'F': 0, 'G': 11, 'H': 0, 'I': 0, 'J': 552, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 40, 'Q': 14, 'R': 0, 'S': 0, 'T': 13, 'U': 4, 'V': 0, 'W': 0, 'X': 16, 'Y': 0, 'Z': 0}
K: {'A': 0, 'B': 8, 'C': 0, 'D': 29, 'E': 135, 'F': 0, 'G': 50, 'H': 0, 'I': 0, 'J': 14, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 10, 'Q': 150, 'R': 0, 'S': 0, 'T': 0, 'U': 119, 'V': 0, 'W': 0, 'X': 224, 'Y': 0, 'Z': 0
}
L: {'A': 0, 'B': 1, 'C': 0, 'D': 5, 'E': 56, 'F': 0, 'G': 18, 'H': 0, 'I': 0, 'J': 472, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 3, 'Q': 85, 'R': 0, 'S': 0, 'T': 0, 'U': 7, 'V': 0, 'W': 0, 'X': 114, 'Y': 0, 'Z': 0}
M: {'A': 0, 'B': 2, 'C': 0, 'D': 50, 'E': 0, 'F': 0, 'G': 37, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 55, 'N': 0, 'O': 0, 'P': 11, 'Q': 65, 'R': 0, 'S': 0, 'T': 0, 'U': 2, 'V': 0, 'W': 536, 'X': 34, 'Y': 0, 'Z': 0}
N: {'A': 0, 'B': 4, 'C': 0, 'D': 47, 'E': 3, 'F': 0, 'G': 11, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 56, 'N': 0, 'O': 0, 'P': 41, 'Q': 23, 'R': 0, 'S': 0, 'T': 0, 'U': 78, 'V': 0, 'W': 504, 'X': 16, 'Y': 0, 'Z': 0}
O: {'A': 0, 'B': 18, 'C': 0, 'D': 103, 'E': 43, 'F': 0, 'G': 254, 'H': 0, 'I': 0, 'J': 54, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 37, 'Q': 224, 'R': 0, 'S': 0, 'T': 0, 'U': 13, 'V': 0, 'W': 0, 'X': 7, 'Y': 0, 'Z': 0}
P: {'A': 0, 'B': 0, 'C': 0, 'D': 15, 'E': 0, 'F': 0, 'G': 15, 'H': 0, 'I': 0, 'J': 9, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 684, 'Q': 17, 'R': 0, 'S': 0, 'T': 61, 'U': 1, 'V': 0, 'W': 0, 'X': 1, 'Y': 0, 'Z': 0}
Q: {'A': 0, 'B': 4, 'C': 0, 'D': 39, 'E': 44, 'F': 0, 'G': 294, 'H': 0, 'I': 0, 'J': 7, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 10, 'Q': 323, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 62, 'Y': 0, 'Z': 0}
R: {'A': 0, 'B': 40, 'C': 0, 'D': 303, 'E': 6, 'F': 0, 'G': 138, 'H': 0, 'I': 0, 'J': 1, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 9, 'Q': 206, 'R': 0, 'S': 0, 'T': 0, 'U': 2, 'V': 0, 'W': 0, 'X': 53, 'Y': 0, 'Z': 0}
S: {'A': 0, 'B': 17, 'C': 0, 'D': 282, 'E': 101, 'F': 0, 'G': 97, 'H': 0, 'I': 0, 'J': 3, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 70, 'Q': 99, 'R': 0, 'S': 0, 'T': 1, 'U': 1, 'V': 0, 'W': 0, 'X': 77, 'Y': 0, 'Z': 0}
T: {'A': 0, 'B': 1, 'C': 0, 'D': 34, 'E': 34, 'F': 0, 'G': 21, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 121, 'Q': 24, 'R': 0, 'S': 0, 'T': 506, 'U': 20, 'V': 0, 'W': 1, 'X': 34, 'Y': 0, 'Z': 0}
U: {'A': 0, 'B': 4, 'C': 0, 'D': 22, 'E': 25, 'F': 0, 'G': 69, 'H': 0, 'I': 0, 'J': 6, 'K': 0, 'L': 0, 'M': 6, 'N': 0, 'O': 0, 'P': 6, 'Q': 53, 'R': 0, 'S': 0, 'T': 0, 'U': 558, 'V': 0, 'W': 48, 'X': 16, 'Y': 0, 'Z': 0}
V: {'A': 0, 'B': 0, 'C': 0, 'D': 0, 'E': 0, 'F': 0, 'G': 19, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 2, 'N': 0, 'O': 0, 'P': 70, 'Q': 48, 'R': 0, 'S': 0, 'T': 299, 'U': 276, 'V': 0, 'W': 18, 'X': 32, 'Y': 0, 'Z': 0}
W: {'A': 0, 'B': 2, 'C': 0, 'D': 8, 'E': 1, 'F': 0, 'G': 39, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 72, 'N': 0, 'O': 0, 'P': 48, 'Q': 46, 'R': 0, 'S': 0, 'T': 0, 'U': 3, 'V': 0, 'W': 529, 'X': 4, 'Y': 0, 'Z': 0}
X: {'A': 0, 'B': 14, 'C': 0, 'D': 108, 'E': 160, 'F': 0, 'G': 24, 'H': 0, 'I': 0, 'J': 7, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 72, 'Q': 144, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 258, 'Y': 0, 'Z': 0}
Y: {'A': 0, 'B': 1, 'C': 0, 'D': 6, 'E': 3, 'F': 0, 'G': 25, 'H': 0, 'I': 0, 'J': 1, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 212, 'Q': 41, 'R': 0, 'S': 0, 'T': 469, 'U': 3, 'V': 0, 'W': 0, 'X': 25, 'Y': 0, 'Z': 0}
Z: {'A': 0, 'B': 17, 'C': 0, 'D': 127, 'E': 252, 'F': 0, 'G': 61, 'H': 0, 'I': 0, 'J': 0, 'K': 0, 'L': 0, 'M': 0, 'N': 0, 'O': 0, 'P': 49, 'Q': 19, 'R': 0, 'S': 0, 'T': 0, 'U': 0, 'V': 0, 'W': 0, 'X': 209, 'Y': 0, 'Z': 0}

Final model trained on the entire dataset.
```