

CS561 - ARTIFICIAL INTELLIGENCE LAB
ASSIGNMENT-4: DECISION TREE

Group Name: 1801cs31_1801cs32_1801cs33

Students:

<u>Names</u>	<u>Roll No.</u>	<u>Batch</u>
M Maheeth Reddy	1801CS31	B.Tech.
M Nitesh Reddy	1801CS32	B.Tech.
Nischal A	1801CS33	B.Tech.

1. 10 Fold Cross-validation score with Gini index as the metric
 - a. Gini Index
 - b. Precision Score = 0.8033050015590296
 - c. Recall Score = 0.756537826734936
 - d. F Score = 0.7717747826525434
2. Results on the test set
 - a. GINI Index

	precision	recall	f1-score	support
ABBR	0.86	0.67	0.75	9
DESC	0.76	0.97	0.85	138
ENTY	0.68	0.72	0.70	94
HUM	0.92	0.85	0.88	65
LOC	0.89	0.70	0.79	81
NUM	0.99	0.81	0.89	113
accuracy			0.82	500
macro avg	0.85	0.79	0.81	500
weighted avg	0.84	0.82	0.82	500

b. Cross-Entropy

	precision	recall	f1-score	support
ABBR	0.86	0.67	0.75	9
DESC	0.66	0.97	0.79	138
ENTY	0.69	0.50	0.58	94
HUM	0.90	0.86	0.88	65
LOC	0.88	0.73	0.80	81
NUM	0.98	0.81	0.88	113
accuracy			0.79	500
macro avg	0.83	0.76	0.78	500
weighted avg	0.81	0.79	0.78	500

c. Misclassification Error

	precision	recall	f1-score	support
ABBR	0.86	0.67	0.75	9
DESC	0.77	0.83	0.79	138
ENTY	0.57	0.80	0.66	94
HUM	0.92	0.85	0.88	65
LOC	0.92	0.69	0.79	81
NUM	0.98	0.79	0.87	113
accuracy			0.79	500
macro avg	0.83	0.77	0.79	500
weighted avg	0.82	0.79	0.80	500

3. Feature Ablation Study (Reporting using Gini impurity as a metric)

a. No length feature

	precision	recall	f1-score	support
ABBR	0.86	0.67	0.75	9
DESC	0.76	0.97	0.85	138
ENTY	0.68	0.72	0.70	94
HUM	0.92	0.85	0.88	65
LOC	0.89	0.70	0.79	81
NUM	0.99	0.81	0.89	113
accuracy			0.82	500
macro avg	0.85	0.79	0.81	500
weighted avg	0.84	0.82	0.82	500

We can see that the “length” feature does not affect the classification results much. We get nearly the same F1 scores both with and without the length feature. We have verified this result on all three metrics Gini, misclassification, and cross-entropy

b. No POS tag feature

	precision	recall	f1-score	support
ABBR	0.86	0.67	0.75	9
DESC	0.73	0.98	0.84	138
ENTY	0.60	0.63	0.61	94
HUM	0.87	0.85	0.86	65
LOC	0.88	0.65	0.75	81
NUM	1.00	0.77	0.87	113
accuracy			0.79	500
macro avg	0.82	0.76	0.78	500
weighted avg	0.81	0.79	0.79	500

We can see that there is a significant decrease in the overall accuracy and the F1 score of the test set when the “POS tag” feature is not used. Hence, we see that the POS tag feature is very important for this problem.

- c. We also posit that the unigram, bigram, and trigram features are important just like the POS tag feature. As per empirical observations, the length feature is not as important as the other features.

4. Error Analysis:

We carry an error analysis experiment to see how many samples which are misclassified by GINI are correctly classified by Cross-Entropy and Misclassification Error.

Here are the outputs: -

```
Entropy correctly classifies 10 as compared to GINI metric
```

```
Misclassification error correctly classifies 10 as compared to GINI metric
```