Nursery School Application Selection using Decision Tree-based Learning Mode
Mini Project 1 Code: NSDT

Group 2

Members:
Naveen Gupta - 17CH30054
Harsh Choudhary - 17CH30040
Shivang Gupta - 17CS10050

1. Decision Tree Model without Pruning

A. Gini Method

Max depth of the Decision tree is: 8
The total numbers of leaf nodes are: 701

Classification report for test data:

	precision	recall	f1-score	support
not_recom	0.99	1.00	0.99	962
priority	0.97	0.97	0.97	965
spec_prior	0.98	0.98	0.98	949
very_recom	0.84	0.72	0.77	86
accuracy			0.98	2962
macro avg	0.94	0.92	0.93	2962
weighted avg	0.98	0.98	0.98	2962

B. Entropy Method

Max depth of the Decision tree is: 8
The total numbers of leaf nodes are: 738
Classification report for test data:

	precision	recall	f1-score	support
not_recom	0.98	1.00	0.99	962
priority	0.97	0.96	0.97	965
spec_prior	0.98	0.97	0.97	949
very_recom	0.84	0.76	0.80	86
accuracy			0.97	2962
macro avg	0.94	0.92	0.93	2962
weighted avg	0.97	0.97	0.97	2962

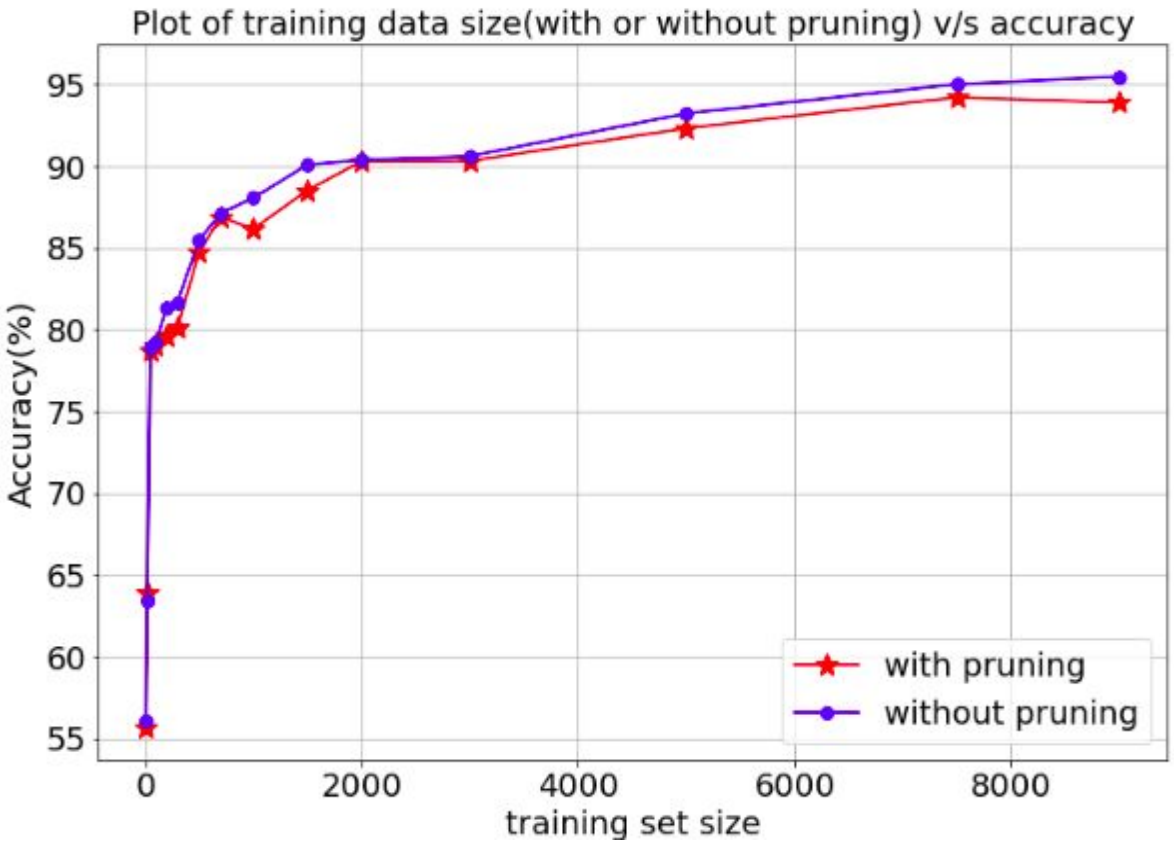
C. Sklearn library- Entropy method

Classification report for test data:

	precision	recall	f1-score	support
not_recom	1.00	1.00	1.00	962
priority	0.99	0.99	0.99	965
recommend	0.00	0.00	0.00	0
spec_prior	0.99	0.99	0.99	949
very_recom	0.98	0.94	0.96	86
accuracy			0.99	2962
macro avg	0.79	0.79	0.79	2962
weighted avg	0.99	0.99	0.99	2962

2. Comparing Decision Tree With and Without Pruning using a Validation set

The graph below shows the decision tree's accuracy with and without pruning on the validation set, using Kfold cross-validation on training data, starting from 10 examples to the entire data set.



3. Revised Decision Tree Model with Pruning

Classification report for test data:

	precision	recall	f1-score	support
not_recom	0.99	1.00	0.99	962
priority	0.95	0.96	0.95	965
spec_prior	0.98	0.96	0.97	949
very_recom	0.73	0.72	0.73	86
accuracy			0.96	2962
macro avg	0.91	0.91	0.91	2962
weighted avg	0.96	0.96	0.96	2962