

DATA MINING (CS F415)

K-NN AND NAIVE BAYES IMPLEMENTATION

1)Team details:-

ROHITH SARANGA (ID: 2017A7PS0034H)

KAUSHIK PERIKA (ID: 2017A7PS0207H)

MAHESH BABU (ID: 2017A7PS0235H)

BHARGAV NIMMAGADDA (ID: 2017A7PS1574H)

2) Dataset Used:

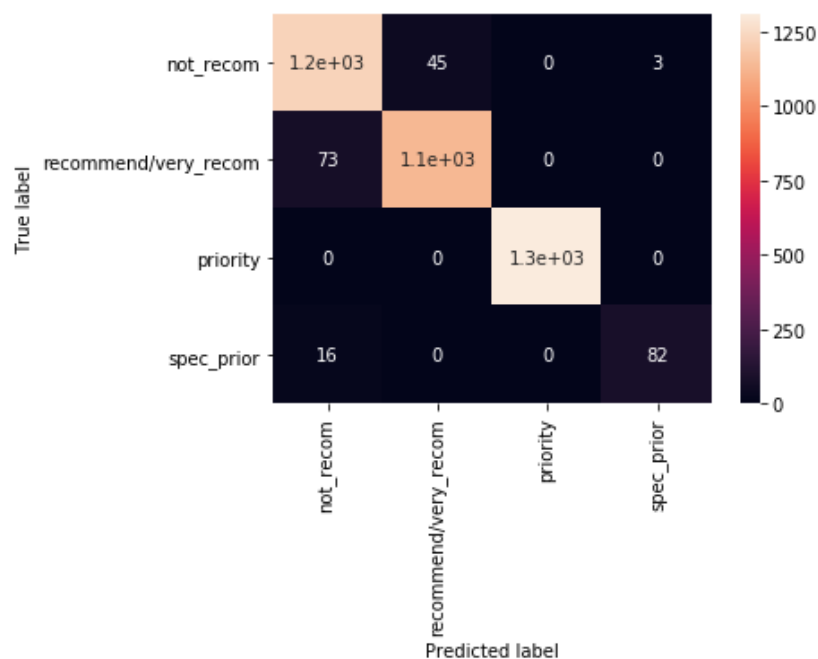
Nursery dataset (<https://archive.ics.uci.edu/ml/datasets/nursery>)

3) Pre-processing done:

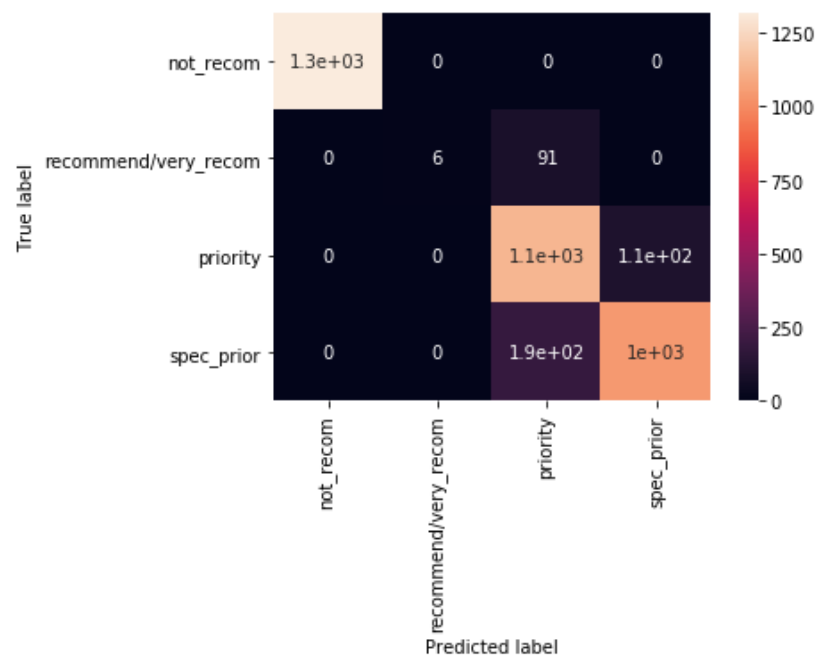
The dataset had very few rows of class " and thus was clubbed with class ". And the columns with string values were mapped to integer values using the 'map' function. And the numeric values thus obtained were standard scaled in case of K-NN and Min-Max scaled in case of Naïve Bayes.

4) Confusion Matrix:-

KNN:-



NAÏVE BAYES:-



5)Accuracy, Precision and Recall values:-

KNN:- Accuracy: 96%

	precision	recall	f1-score	support
0	0.95	0.94	0.94	1222
1	1.00	1.00	1.00	1284
2	0.93	0.95	0.94	1281
3	0.95	0.82	0.88	101
micro avg	0.96	0.96	0.96	3888
macro avg	0.96	0.93	0.94	3888
weighted avg	0.96	0.96	0.96	3888

NAÏVE BAYES:- Accuracy: 0.91%

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1325
1	1.00	0.11	0.19	103
2	0.82	0.92	0.87	1257
3	0.91	0.87	0.89	1203
micro avg	0.91	0.91	0.91	3888
macro avg	0.93	0.72	0.74	3888
weighted avg	0.92	0.91	0.90	3888