About AI in Lower Judiciary

Usecase – Outcome Extraction from Case Orders pdfs of Lower Judiciary

Problem Statement - To identify the Case Outcome from Case Order pdfs as accepted or rejected to be able to update the database with case outcomes.

Text Mining Process - It was attempted to extract text corpus from the case pdf files and to search for most important Terms/ phrases/sentences etc.. in the documents by creating a Term document Matrix. Exploratory datya analysis of Term frequencies were carried out.

Machine Learning Algorithm Used

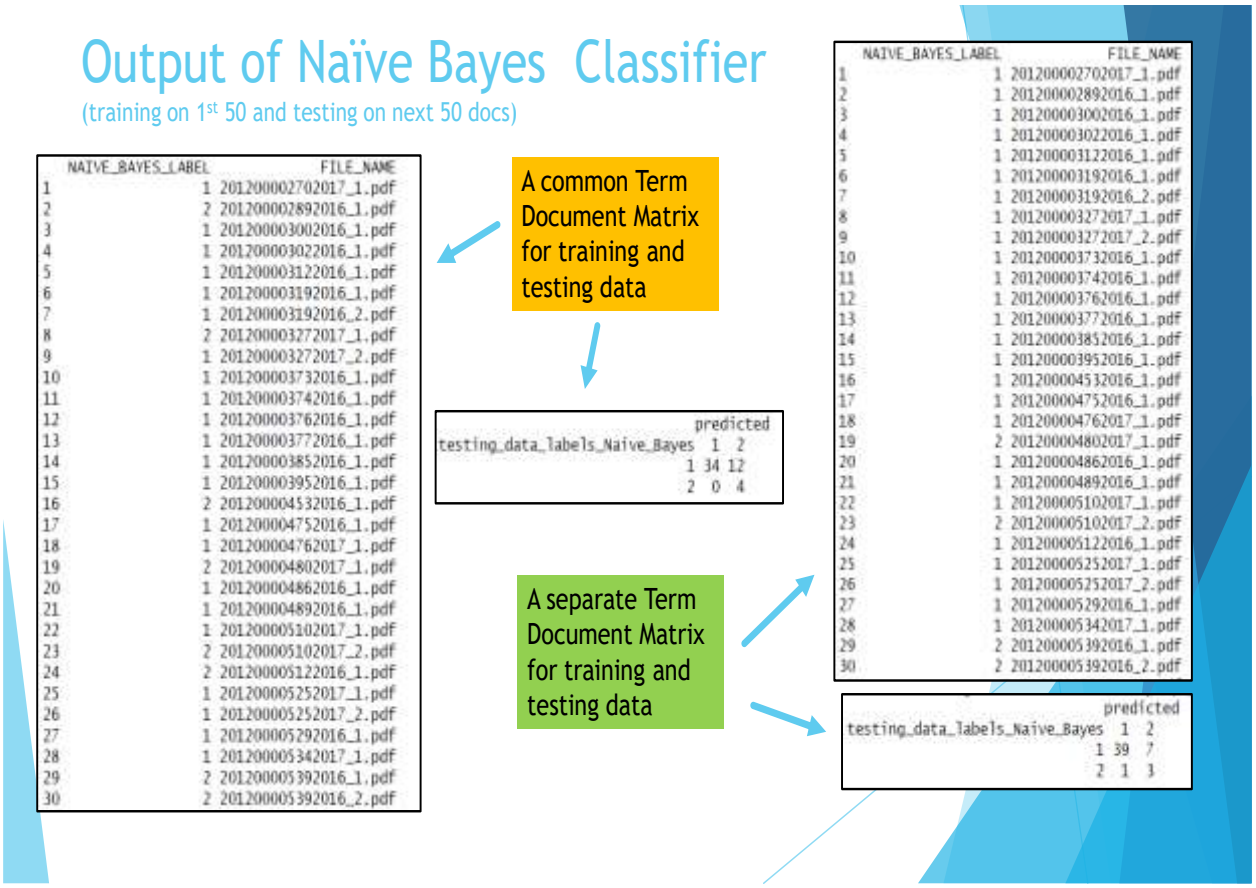
1. For Clustering of documents on basis of appearance of accept or reject word in document k-means algorithm was used with k=2 clusters.
2. For Document Classification Naive Bayes algorithm and Support Vector Machine (SVM) Algorithm were tried using labeled data for training.

Results -

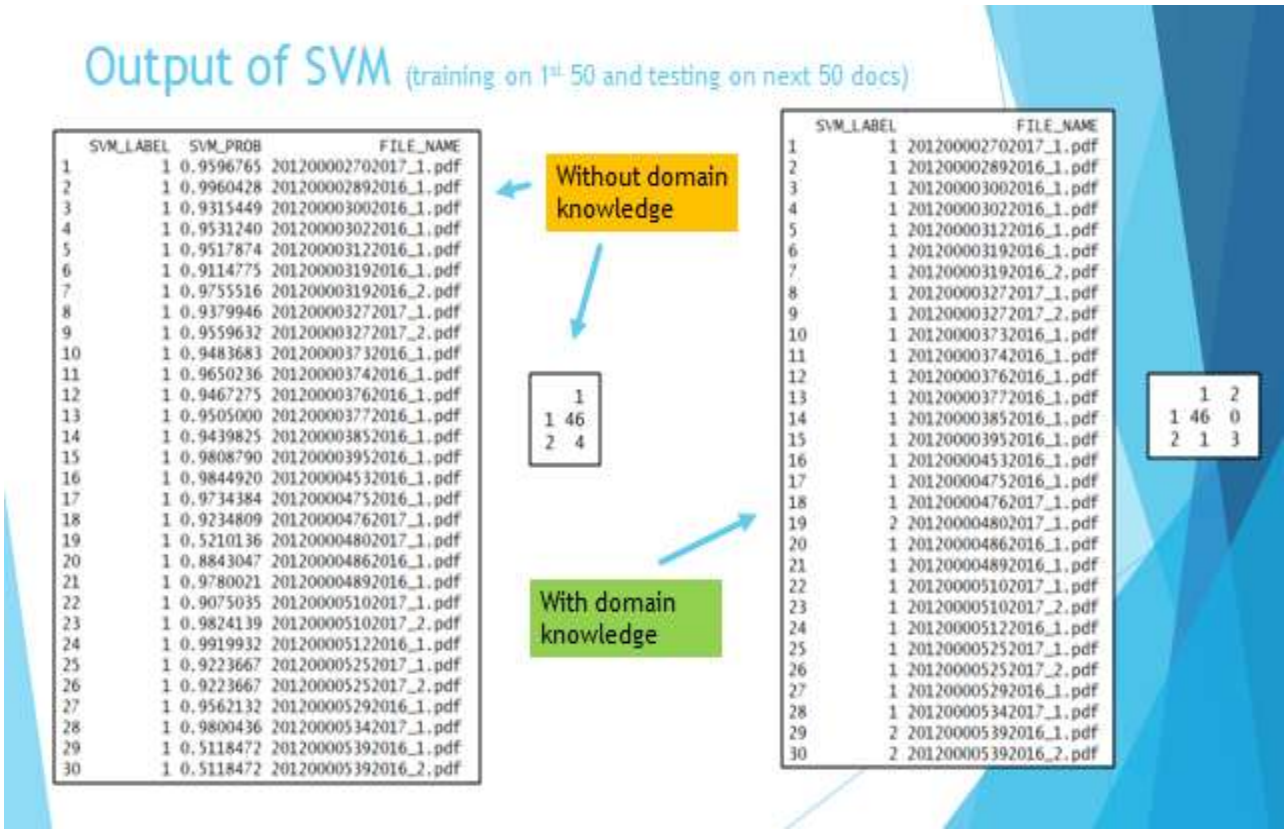
1. For Unsupervised Learning using k _means 92% accuracy was obtained.

```
## 201200000012016_1.pdf 201200000052016_1.pdf 201200000052017_1.pdf
##                     1                     1                     1
## 201200000052017_2.pdf 201200000062016_1.pdf
##                     1                     1
```

2. For Naive Bayes using probabilities of outcome prediction accuracy was 76%.



3. For Supervised Learning using SVM algorithm 98% accuracy was obtained.



	SVM_LABEL	SVM_PROB	FILE_NAME
1	1	0.9596765	201200002702017_1.pdf
2	1	0.9960428	201200002892016_1.pdf
3	1	0.9315449	201200003002016_1.pdf
4	1	0.9531240	201200003022016_1.pdf
5	1	0.9517874	201200003122016_1.pdf
6	1	0.9114775	201200003192016_1.pdf
7	1	0.9755516	201200003192016_2.pdf
8	1	0.9379946	201200003272017_1.pdf
9	1	0.9559632	201200003272017_2.pdf
10	1	0.9483683	201200003732016_1.pdf
11	1	0.9650236	201200003742016_1.pdf
12	1	0.9467275	201200003762016_1.pdf
13	1	0.9505000	201200003772016_1.pdf
14	1	0.9439825	201200003852016_1.pdf
15	1	0.9808790	201200003952016_1.pdf
16	1	0.9844920	201200004532016_1.pdf
17	1	0.9734384	201200004752016_1.pdf
18	1	0.9234809	201200004762017_1.pdf
19	1	0.5210136	201200004802017_1.pdf
20	1	0.8843047	201200004862016_1.pdf
21	1	0.9780021	201200004892016_1.pdf
22	1	0.9075035	201200005102017_1.pdf
23	1	0.9824139	201200005102017_2.pdf
24	1	0.9919932	201200005122016_1.pdf
25	1	0.9223667	201200005252017_1.pdf
26	1	0.9223667	201200005252017_2.pdf
27	1	0.9562132	201200005292016_1.pdf
28	1	0.9800436	201200005342017_1.pdf
29	1	0.5118472	201200005392016_1.pdf
30	1	0.5118472	201200005392016_2.pdf

	SVM_LABEL	FILE_NAME
1	1	201200002702017_1.pdf
2	1	201200002892016_1.pdf
3	1	201200003002016_1.pdf
4	1	201200003022016_1.pdf
5	1	201200003122016_1.pdf
6	1	201200003192016_1.pdf
7	1	201200003192016_2.pdf
8	1	201200003272017_1.pdf
9	1	201200003272017_2.pdf
10	1	201200003732016_1.pdf
11	1	201200003742016_1.pdf
12	1	201200003762016_1.pdf
13	1	201200003772016_1.pdf
14	1	201200003852016_1.pdf
15	1	201200003952016_1.pdf
16	1	201200004532016_1.pdf
17	1	201200004752016_1.pdf
18	1	201200004762017_1.pdf
19	2	201200004802017_1.pdf
20	1	201200004862016_1.pdf
21	1	201200004892016_1.pdf
22	1	201200005102017_1.pdf
23	1	201200005102017_2.pdf
24	1	201200005122016_1.pdf
25	1	201200005252017_1.pdf
26	1	201200005252017_2.pdf
27	1	201200005292016_1.pdf
28	1	201200005342017_1.pdf
29	2	201200005392016_1.pdf
30	2	201200005392016_2.pdf