Hate Speech Classification Report

Dataset : Hate speech containing 24783 tweets classified as 0-Hate, 1-offensive, 2- neutral.

Feature Design –

Pre-processing steps

- Removal of out of dictionary words such as tags and urls
- Tokenization
- Stop words removal
- Generation of n-grams
- Removing <2% frequency of words.
- Lemmatization (resulted in decrease of accuracy), hence removed it from pre-processing step.

Feature Space

TD-IDF of vector of size `(24783, 7086)`

Truncated SVD `(24783, 100)` – decided 100 feature size by generating different truncated SVD and doing thread off evaluation of feature size and accuracy for couple of algorithms.

Training and test datasets

Test size = 10%

Kfold split = 5

Algorithms –

1. Logistic Regression with c = 0.001; 0.01; 0.1; 1; 10; 100; 100

Best scores with c – 0.01

	precision	recall	f1-score	support
0	0.45	0.57	0.50	164
1	0.96	0.91	0.93	1905
2	0.82	0.94	0.88	410
accuracy			0.89	2479
avg	0.74	0.81	0.77	2479
weighted avg	0.91	0.89	0.90	2479

For other c values just accuracy

C	ACCURACY
0.001	0.83
0.01	0.89
0.1	0.87
1	0.86
10,100,1000	0.85

Note : with Truncated SVD the best accuracy was 0.87, but result for hate classification was really off.

2. k-NN classifier

with k = [1 , 3, 5] and p = [1,2, ∞]
trained model with truncated svd with 100 features.

Best scores with k=5 and p=2

	precision	recall	f1-score	support
0	0.44	0.20	0.28	164
1	0.88	0.94	0.91	1905
2	0.73	0.65	0.68	410
accuracy			0.84	2479
avg	0.68	0.60	0.62	2479
weighted avg	0.82	0.84	0.83	2479

K	P	ACCURACY
1	1	0.82
1	2	0.80
1	∞	0.76
3	1	0.85
3	2	0.84
3	∞	0.78
5	1	0.85
5	2	0.84
5	∞	0.80

Note : Though the accuracy was better for k=3,5 and p=1, the precision,recall,f1-score were much better with k=5, p=2 for hate classification.

3. kernelized SVM:

gamma {0.01, 0. 1, 1, 10, 100 } C {0.01, 0.1, 1, 10, 100}
trained model with truncated svd with 100 features.

best score with gamma = 0.01 and c = 100

	precision	recall	f1-score	support
0	0.41	0.35	0.38	164
1	0.89	0.91	0.90	1905
2	0.75	0.75	0.75	410
accuracy			0.84	2479
avg	0.68	0.67	0.68	2479
weighted avg	0.83	0.84	0.83	2479

4. KDE based bayes classifier

trained model with truncated svd with 100 features
bandwith: [0.1,1,10]}
distance : euclidean

Gaussian Kernel for smoothing

	precision	recall	f1-score	support
0	0.30	0.22	0.25	164
1	0.87	0.90	0.89	1905
2	0.65	0.64	0.64	410
accuracy			0.81	2479
avg	0.61	0.58	0.59	2479
weighted avg	0.80	0.81	0.80	2479

Epanechnikov kernel for smoothing

	precision	recall	f1-score	support
0	0.07	0.01	0.02	164
1	0.78	0.98	0.87	1905
2	0.57	0.07	0.13	410
accuracy			0.76	2479
avg	0.47	0.35	0.34	2479
weighted avg	0.70	0.76	0.69	2479

TopHat kernel for smoothing

	precision	recall	f1-score	support
0	0.00	0.00	0.00	164
1	0.77	0.98	0.86	1905
2	0.50	0.06	0.10	410
accuracy			0.76	2479
avg	0.42	0.34	0.32	2479
weighted avg	0.68	0.76	0.68	2479

5. Gaussian based Bayes classifier. (Naïve Bayes)

	precision	recall	f1-score	support
0	0.11	0.35	0.16	164
1	0.90	0.70	0.79	1905
2	0.57	0.62	0.59	410
accuracy			0.67	2479
avg	0.52	0.56	0.51	2479
weighted avg	0.79	0.67	0.72	2479