Classification Report
SVM

Set Exploration

The exploration of the dataset led to the following dates

Dataset	Hatespeech-Extend-Dataset
Size	32,374 sets
Features	1: `c_text` 2: `c(c_text)` 3: `c(emojis)` 4: `sentiment_neg` 5: `sentiment_pos` 6: `sentiment_neu` 6: `sentiment_comp` 6: `profanity_score`
Label	1: `hatespeech` 2: `target`
Label distribution (Label 1)	`0: 13,728` `1: 18,646`

Label distribution (Label 2)	`person: 4,834`
	`group: 3,513`
	`public: 2,018`
	`NaN: 22,009`

The data set was then examined for the distribution and frequency of occurring terms for the processing of the text data. Background noise was removed.

Classification Metrics

After training the classifier in a predefined test-train split, the following metric was determined to evaluate the classifier performance.

Dataset	Hatespeech-Extend-Dataset
training set size	0.8
testing set size	0.2
Hatespeech Metrics	
5fold-CV-Score	0.852± 0.005337
execution time	76.769s
Target Metrics, Binary approach	
5fold-CV-Score	0.731± 0.015
execution time	21.414s
Target Metrics, Concatenated approach	
5fold-CV-Score	0.687± 0.009
execution time	1,328.920s

Examined best performing model score

–Hatespeech –
The SVM classifier achieved a **placement of 1** in hatespeech classification, which ranks it **1/6 in the overall ranking**.

–Target –
The SVM classifier achieved a **placement of 5** in target classification, which ranks it **5/6 in the overall ranking**.

Hatespeech classification Ranking		
1 st.	SVM	0.852
2 nd.	Logistic Regression	0.840
3 rd.	Ridge Classifier	0.836
4 th.	NaiveBayes	0.833
5 th.	RandomForest Classifier	0.826
6 th.	DecisionTree Classifier	0.768

Target classification Ranking		
1 st.	RandomForest Classifier	0.760
2 nd.	Ridge Classifier	0.740
3 rd.	DecisionTree Classifier	0.737
4 th.	Logistic Regression	0.732
5 th.	SVM	0.731
6 th.	NaiveBayes	0.695