1. SVM baseline		**2. SVM optimised**	

Trained and tested on normal dataset

	precision	recall	f1-score	support
NOT	0.7885	0.9500	0.8617	620
OFF	0.7257	0.3417	0.4646	240
accuracy			0.7802	860
macro avg	0.7571	0.6458	0.6632	860
weighted avg	0.7710	0.7802	0.7509	860
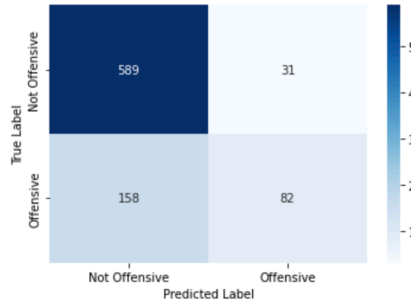
	precision	recall	f1-score	support
NOT	0.8174	0.9387	0.8739	620
OFF	0.7432	0.4583	0.5670	240
accuracy			0.8047	860
macro avg	0.7803	0.6985	0.7204	860
weighted avg	0.7967	0.8047	0.7882	860
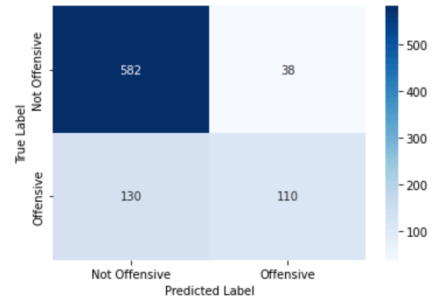
Trained and tested on masked dataset

	precision	recall	f1-score	support
NOT	0.7577	0.9887	0.8579	620
OFF	0.8627	0.1833	0.3024	240
accuracy			0.7640	860
macro avg	0.8102	0.5860	0.5802	860
weighted avg	0.7870	0.7640	0.7029	860
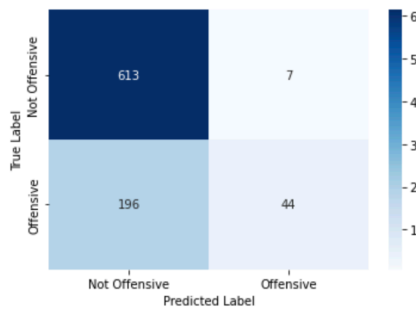
	precision	recall	f1-score	support
NOT	0.8036	0.9306	0.8625	620
OFF	0.6972	0.4125	0.5183	240
accuracy			0.7860	860
macro avg	0.7504	0.6716	0.6904	860
weighted avg	0.7739	0.7860	0.7664	860
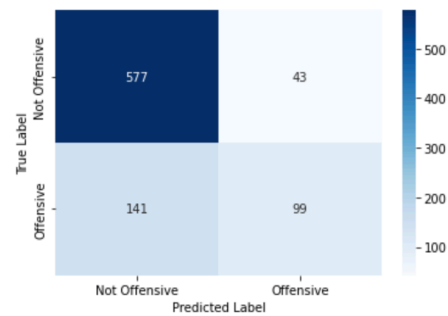
Trained with normal and tested on masked dataset

	precision	recall	f1-score	support
NOT	0.7896	0.9806	0.8748	620
OFF	0.8667	0.3250	0.4727	240
accuracy			0.7977	860
macro avg	0.8281	0.6528	0.6738	860
weighted avg	0.8111	0.7977	0.7626	860
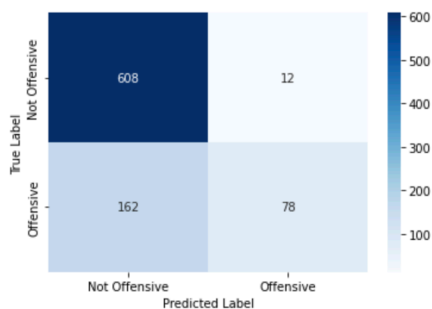
	precision	recall	f1-score	support
NOT	0.7691	0.9726	0.8590	620
OFF	0.7763	0.2458	0.3734	240
accuracy			0.7698	860
macro avg	0.7727	0.6092	0.6162	860
weighted avg	0.7711	0.7698	0.7235	860
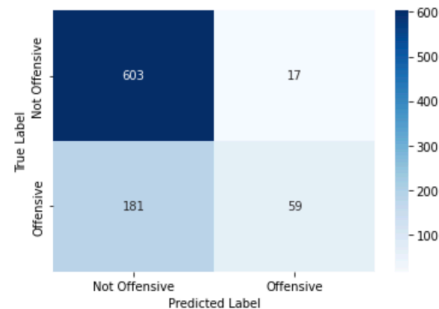
3. LSTM   # 4. BERT

Trained and tested on normal dataset

Classification Report:

	precision	recall	f1-score	support
0	0.818	0.829	0.824	620
1	0.543	0.525	0.534	240
accuracy			0.744	860
macro avg	0.681	0.677	0.679	860
weighted avg	0.742	0.744	0.743	860

Classification Report on test set:

	precision	recall	f1-score	support
0	0.876	0.932	0.903	620
1	0.790	0.658	0.718	240
accuracy			0.856	860
macro avg	0.833	0.795	0.811	860
weighted avg	0.852	0.856	0.852	860

LSTM confusion matrix (normal): True Not Offensive / Predicted Not Offensive = 514, Predicted Offensive = 106; True Offensive / Predicted Not Offensive = 114, Predicted Offensive = 126.

BERT confusion matrix (normal): True Not Offensive / Predicted Not Offensive = 578, Predicted Offensive = 42; True Offensive / Predicted Not Offensive = 82, Predicted Offensive = 158.

Trained and tested on masked dataset

Classification Report:

	precision	recall	f1-score	support
0	0.832	0.845	0.838	620
1	0.583	0.558	0.570	240
accuracy			0.765	860
macro avg	0.707	0.702	0.704	860
weighted avg	0.762	0.765	0.764	860

Classification Report on test set:

	precision	recall	f1-score	support
0	0.804	0.950	0.871	620
1	0.756	0.400	0.523	240
accuracy			0.797	860
macro avg	0.780	0.675	0.697	860
weighted avg	0.790	0.797	0.774	860

LSTM confusion matrix (masked): True Not Offensive / Predicted Not Offensive = 524, Predicted Offensive = 96; True Offensive / Predicted Not Offensive = 106, Predicted Offensive = 134.

BERT confusion matrix (masked): True Not Offensive / Predicted Not Offensive = 589, Predicted Offensive = 31; True Offensive / Predicted Not Offensive = 144, Predicted Offensive = 96.

Trained with normal and tested on masked dataset

Classification Report:

	precision	recall	f1-score	support
0	0.782	0.848	0.814	620
1	0.497	0.388	0.436	240
accuracy			0.720	860
macro avg	0.639	0.618	0.625	860
weighted avg	0.702	0.720	0.708	860

Classification Report on test set:

	precision	recall	f1-score	support
0	0.830	0.968	0.894	620
1	0.854	0.487	0.621	240
accuracy			0.834	860
macro avg	0.842	0.728	0.757	860
weighted avg	0.837	0.834	0.817	860

LSTM confusion matrix (normal-trained, masked-tested): True Not Offensive / Predicted Not Offensive = 526, Predicted Offensive = 94; True Offensive / Predicted Not Offensive = 147, Predicted Offensive = 93.

BERT confusion matrix (normal-trained, masked-tested): True Not Offensive / Predicted Not Offensive = 600, Predicted Offensive = 20; True Offensive / Predicted Not Offensive = 123, Predicted Offensive = 117.

5. RoBERTa

6. DeBERTa

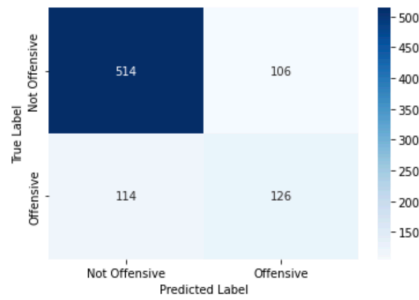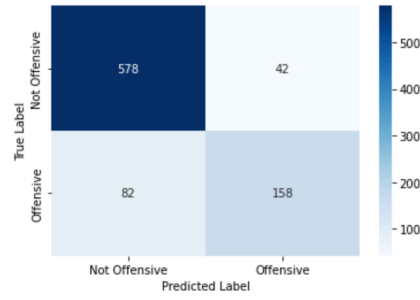*Trained and tested on normal dataset*

```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.888     0.895     0.892       620
           1      0.723     0.708     0.716       240

    accuracy                          0.843       860
   macro avg      0.806     0.802     0.804       860
weighted avg      0.842     0.843     0.843       860
```
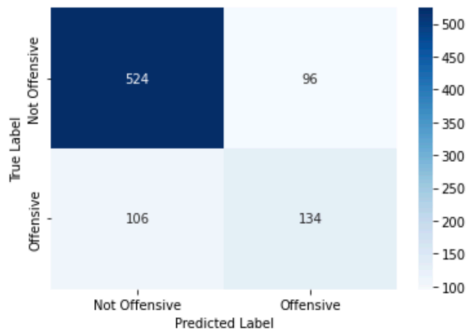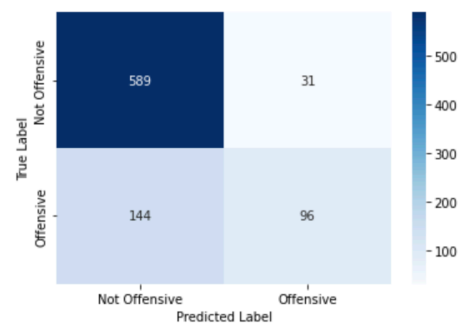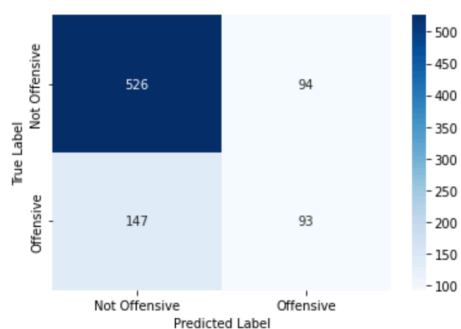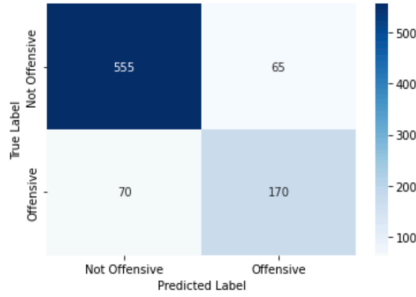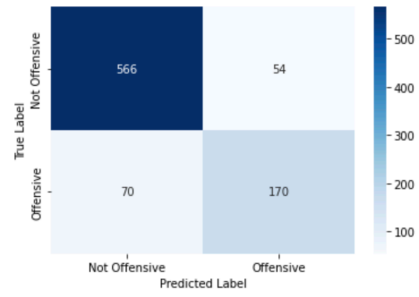


```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.890     0.913     0.901       620
           1      0.759     0.708     0.733       240

    accuracy                          0.856       860
   macro avg      0.824     0.811     0.817       860
weighted avg      0.853     0.856     0.854       860
```
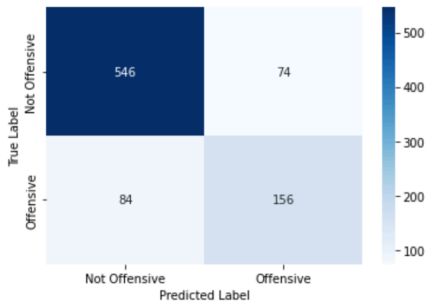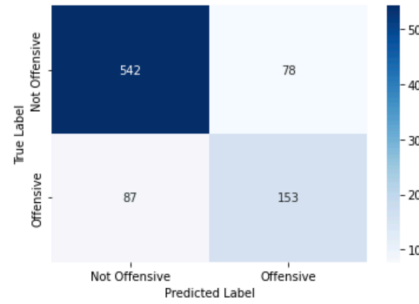

Trained and tested on masked dataset

```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.867     0.881     0.874       620
           1      0.678     0.650     0.664       240

    accuracy                          0.816       860
   macro avg      0.772     0.765     0.769       860
weighted avg      0.814     0.816     0.815       860
```



```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.862     0.874     0.868       620
           1      0.662     0.637     0.650       240

    accuracy                          0.808       860
   macro avg      0.762     0.756     0.759       860
weighted avg      0.806     0.808     0.807       860
```
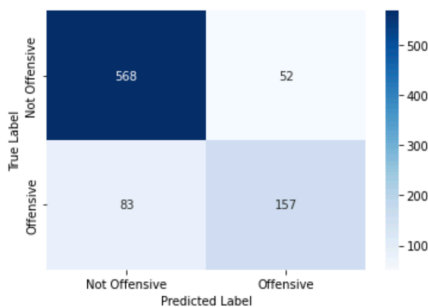

Trained with normal and tested on masked dataset

```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.873     0.916     0.894       620
           1      0.751     0.654     0.699       240

    accuracy                          0.843       860
   macro avg      0.812     0.785     0.797       860
weighted avg      0.839     0.843     0.840       860
```



```
Classification Report on test set:
              precision    recall  f1-score   support

           0      0.897     0.910     0.903       620
           1      0.758     0.729     0.743       240

    accuracy                          0.859       860
   macro avg      0.827     0.819     0.823       860
weighted avg      0.858     0.859     0.858       860
```