

3 epochs newBalancedData:

Text

Description automatically generated51%

5 epochs newBalancedData:

Text

Description automatically generated47%



3 epochs newBalancedData:

Text

Description automatically generated 51.2%

5 epochs newBalancedData:

Text

Description automatically generated58%



3 epochs newBalancedData:

Text

Description automatically generated62%

5 epochs newBalancedData:

Text

Description automatically generated51%

NEW TEST ON CONSTRAINING PROBABILITY:



Np argmax

Text, letter

Description automatically generated

<0.6

<0.55

Text, letter

Description automatically generated

<0.52

<0.57

Testing other transformers with test dataset:

Hate-speech-CNERG/dehatebert-mono-english

Text, letter

Description automatically generated

mrm8488/distilroberta-finetuned-tweets-hate-speech

Text, letter

Description automatically generated

Hate-speech-CNERG/bert-base-uncased-hatexplain

Text, letter

Description automatically generated

Narrativaai/deberta-v3-small-finetuned-hate\_speech18

Text, letter

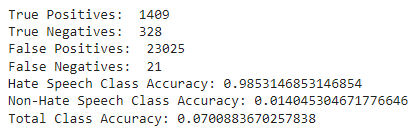
Description automatically generated

Sakil/distilbert\_lazylearner\_hatespeech\_detection

Text, letter

Description automatically generated

monologg/koelectra-base-v3-hate-speech



Testing other transformers with train dataset:

Hate-speech-CNERG/dehatebert-mono-english

Text

Description automatically generated

mrm8488/distilroberta-finetuned-tweets-hate-speech

Text, letter

Description automatically generated

Hate-speech-CNERG/bert-base-uncased-hatexplain

Text

Description automatically generated

Narrativaai/deberta-v3-small-finetuned-hate\_speech18

Text, letter

Description automatically generated

Sakil/distilbert\_lazylearner\_hatespeech\_detection

Text, letter

Description automatically generated

monologg/koelectra-base-v3-hate-speech

Text

Description automatically generated

With ensemble: Text

Description automatically generated

BERT2212: Text

Description automatically generated

BERT2211: Text

Description automatically generated

DISTILBERT: Text

Description automatically generated

NBSVM: Text

Description automatically generated

GRU: Text

Description automatically generated

