Intention:

In order to efficiently identify and remove these posts and/or messages which violate the terms of service, we need some trained models to support spotting them out

Tasks:

First three I adopted them from existing code and improved them.

For ensemble one, I created it on my own

LSTM:

Bidirectional (Past and future)

NBSVM:

TF-IDF: filter some frequent but meaningless words (e.g., I, this, etc.)

# TF: the frequency of THE word

# IDF: log(a/b), a is number of document, b is number of document has THE word

N-Gram: a contiguous sequence of n items from a given sample of text or speech

CNN:

S: length of a sentence (may be limited)

V: length of a vector for step 1.

Row order based on order of the sentence