

WHAT IS TEXT MODERATION ?

Text moderation is the process of blocking, approving or reviewing content submissions based on the policies and thresholds of a certain site. This **type of content moderation** is used to enhance human moderation in an environment that require partners, employees and consumers who can generate text content.

This includes blog or forum comments, reviews, chat rooms, discussions boards and tweets. The main goal is to protect your users and brand without affecting user experience

FUNCTION :

Each text of the input sentence along with the whole input sentence is passed into a toxicity/sentiment classification model.

Upon receiving the output, if the respective input seems toxic or of negative sentiment level, then that word is passes into spacy library that outputs all the synonyms of that toxic word.

Then each synonym is arranged in increasing order of toxicity by separately passing them into the toxicity classification model / sentiment classification model.

So , if the top n words have toxicity less than the input toxic word, then using them instead of that word may improve the overall toxicity level and shift it towards positive by some value.

Library :

Spacy is a natural language processing library of python having following advantages :

- spaCy is designed to help you do real work — to build real products, or gather real insights. The library respects your time, and tries to

avoid wasting it. It's easy to install, and its API is simple and productive. We like to think of spaCy as the Ruby on Rails of Natural Language Processing.

- spaCy excels at large-scale information extraction tasks. It's written from the ground up in carefully memory-managed Cython. Independent research in 2015 found spaCy to be the fastest in the world. If your application needs to process entire web dumps, spaCy is the library you want to be using.
- spaCy is the best way to prepare text for deep learning. It interoperates seamlessly with TensorFlow, PyTorch, scikit-learn, Gensim and the rest of Python's awesome AI ecosystem. With spaCy, you can easily construct linguistically sophisticated statistical models for a variety of NLP problems.