Toxic Comment Classification

You are provided with a large number of Wikipedia comments which have been labeled by human raters for toxic behavior. The types of toxicity are:

Content

The dataset consists of:

comment\_text: The text of the Comment

class: The respective class to which the tweet belongs. There are 5 classes -:

* toxic
* severe\_toxic
* obscene
* threat
* insult

identity\_hateBusiness Objective:

1. Need to get sentiment analysis of tweets gauge its impact and type

Architecture level analysis:

1. Data transformation/Text processing using R/Python
2. Need to get sentiments Analysis and Emotion mining with some charts like histogram, Density plot, Barplot, pie-plot etc.
3. Deployment through R Shiny or Flask/ Streamlit

**Milestones: Phase 1**

|  |  |  |
| --- | --- | --- |
| **Milestone** | **Duration** | **Task start - End Date** |
| Kick off and Business Objective discussion | 1 day |  |
| Data set Details | 1 Week – 1 ½ week |  |
| EDA | 1 Weeks – 1 ½ week |  |
| Model Building | 1 Week – 1 ½ week |  |
| Model Evaluation | 1 week |  |
| Feedback |  |
| Deployment | 1 Week |  |
| Final presentation | 1 day |  |

Protocols:

1. All participants should add here to agreed timelines and timelines will not be extended
2. All the documentation – Final presentation and R/python code to be submitted before the final presentation day
3. All the participants must attend review meetings