Project name

Malignant-Comments-Classifier.

Submitted by: **VIGNESH.R**

BUSINESS PROBLEM

In most of the online conversation platforms, social media users often face abuse, harassment, and insults from other users. Due to which, many users stop expressing their ideas and opinions. Platforms struggle to facilitate conversations effectively.

Use of Machine Learning to solve the problem

With all the latest advancements of AI-ML, the task is to build an efficient model that is capable of detecting various levels/categories of toxicity like threats, obscenity, insults, and identity-based hate for a given comment. Thus helping online conversations/discussion become more productive and respectful.

The problem at hand is not a multi-class classification, but it is a multi-label classification. In other words, a given comment may belongs to none or many levels of toxicity.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis is the most important process in any machine learning project. EDA brings out the useful insights, summaries from the data, thus which will help in solving the business problem with logical and strategical approach.

2. Shape of the dataset

Train data set contains 159571 rows and 8 columns.

Out of 8 columns, 6 columns are class labels and 1 column is for comment and one more column is id to identify a row uniquely.

3. Check for duplicates and missing values

There are no duplicates and missing values in the dataset.

4. Individual counts of Toxic and Non-Toxic comments in each class-label

Count of comments that are non-toxic(class-0) and toxic (class-1) under each class-label individually i.e. without considering the relation between each-other labels.

5. Count of Toxic & Non-Toxic comments in complete dataset

Count of comments that are completely non-toxic i.e. class-0 in all the class labels and toxic under at least one class labels

- Data-set contains 89.83% Non-toxic comments and 10.17% Toxic comments.

- **Word Cloud for Toxic labels in each class**

- Word Cloud for the most common words in toxic comments per each class label.

CONCLUSION

After checking the board of the actual competition, standard Machine Learning approaches yield a maximum score of 0.9560494, irrespective of any approach. In order to get a large mar gin over this score one has to employ Deep Learning (DL) techniques.