**Topic Background:**

Many social media, news, e-commerce websites allow users to share their comments with others. This type of information sharing helps people stayed informed, have better understanding of each other and brings others that are geographically remote closer together.

**Problem:**

While some people share their opinions in a meaningful and respectful manner, some are offensive and hurtful towards others. This type of toxic comments devalues the communication channel and ruins the experience when sharing information with others. Furthermore, it may even deter people from visiting or commenting on the website.

**Objective:**

Develop an algorithm to identify whether comments are toxic. Client can use this algorithm to flag, filter and remove these comments from their website posts.

**Data Set:**

The data set contains a large number of Wikipedia comments hosted in the below location.

< <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>>

The data set contains a training data set with about 159500 comments and a test data set with about 153000 comments. The target variables are sub-divided into the following categories and labeled.

• Toxic

• Severe\_toxic

• Obscene

• Threat

• Insult

• Identity\_hate

The objective is to create a model that takes comments as input and output prediction of the probability of each type toxicity.