**Predicting tweet sentiments**

**Problem statement**

You work in an event management company. On Mother's Day, your company has organized an event where they want to cast positive Mother's Day related tweets in a presentation. Data engineers have already collected the data related to Mother's Day that must be categorized into positive, negative, and neutral tweets.

You are appointed as a Machine Learning Engineer for this project. Your task is to build a model that helps the company classify these sentiments of the tweets into positive, negative, and neutral.

**Data description**

This data set consists of six columns:

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| id | ID of tweet |
| original\_text | Text of tweet |
| lang | Language of tweet |
| retweet\_count | Number of times retweeted |
| original\_author | Twitter handle of Author |
| sentiment\_class | Sentiment of Tweet (Target) |

The data folder consists of two .csv files. The details are as follows:

* train.csv: 3235 x 6
* test.csv: 1387 x 5