

# Report for NLP Disaster tweets

## Problem statement

Natural-Language-Processing-with-Disaster-Tweets

**Aim:** Prediction of whether a tweet is about a real disaster or not.

## Requirements

Tweet's datasets, Machine Learning model, Python

## Description:

In this project, there are two similar datasets that include id, text, location, keyword and target (in train.csv only). One dataset is titled "train.csv" and the other is titled "test.csv".

"target" column denotes whether a tweet is about a real disaster (1) or not (0) and the "text" column contains the tweets.

Training data is use for training the model and the test data is use for the submission on the Kaggle.

Machine Learning models used for training are Naive Bayes classifier, Random Forest Classifier, Logistic Regression, SVM and KNN.

I have used three Machine Learning models which are Support Vector machine classifier (SVM), Random Forest Classifier, Multinomial Naive Bayes and KNN classifier to solve the problem.

Counter Vectorizer from sklearn is used for data pre-processing. For training the model text column is used which contains the tweets.

## Results:

Maximum accuracy reached to 80% by Multinomial Naive Bayes model.

.

## For Split Train Test Data:

Model	Accuracy
SVM	79%
KNN	71%
Random Forest Classifier	80%
Multinomial Naïve Bayes	80%