Input: {Training Dataset (X\_train, y\_train), Testing Dataset (X\_test, y\_test), classifiers ()}

Output: { MultiLabelModel\_result, i.e., MultiLabel model files and Result file path }

1. Initializing with multiple classifiers

2. Develop machine learning models

3. Monitor classification report {Accuracy, Loss, Zero\_one\_loss, Hamming\_loss} on training and testing dataset for MultiLabel model

4. Save classification report and machine learning model

5. Select top 3 model () with best optimal accuracy score

6. Fit MultiLabel classifier model MLC = ()

7. Evaluate final predictions using voting method:

We proposed an improved algorithm for MultiLabel classification problem. The proposed algorithm takes 3 input parameters, Training Dataset, Testing Dataset and a list of classifiers and MultiLabel model result file with result file path. Initiating with a list of classifiers in step 1 after that multipal machine learning models are built in step 2. In step 3, classification report {Accuracy, Loss, Zero\_one\_loss, Hamming\_loss} are generated on the basis of training and testing dataset and in step 4, classification report and machine learning models are saved. In step 5, the top 3 models are selected on the basis of best optimal accuracy score. In step 6, fit the model on optimal MultiLabel classifier and in step 7, evaluate final predictions using voting method.

Machine learning Life cycle

Gathering Data

Data preparation

Data Wrangling

Analyse Data

Train the model

Test the model

Deployment

Gathering Data :

The initial stage of the machine learning life cycle is data gathering. This step's objective is to locate and collect all data-related difficulties.

The different data sources must be identified in this stage since data can be gathered from a variety of sources, including databases, logs, iot devices, the internet, and so on. The output's effectiveness of the model depends on the quantity and quality of the data collected. More data means a more precise prediction, thus more data is better.

The worldweatheronline site API are used as a data source to collect the data for different timelapse. The data is collected from 2010 to 2021 with 4227 records.

Data preparation :