**Twitter Dataset API Documentation**

**Preview:**

To extract the historical tweets for amazon locker and access points according to different categories and parameters such as timespan, Language, geographic location, start and end date and maximum number of tweets.

1. **Tweepy-**

**->**Tweepy is a Python Library that can be used to access the Twitter API. There are various functionality can be performed by using the Tweepy.

**->**It can search tweets and extract information like Timespan, Username, Follower Count, Hashtags, language, location etc.

1. **GetOldTweets3**

**->**It is created as a project and allows to extract the tweets that are older and historic. It also allows to get larger number of tweets according to various filter.

**->**There are various Twitter Criteria such as -> **setUsername(), setSince(), setUntil(),setQuerySearch()** that can be used to filter the tweets.

1. **Twitterscraper**

**->**Twitterscraper is a great way of scrapping the useful information from twitter that is not bounded by the time limit, no of tweets and also helps to scrap lot of information per tweet such as **Tweet timestamp, Tweet No. of replies, Tweet No. of likes** etc

**->**With Twitterscraper we can use twitter advance search option where we can search according to different filters and use the URL generated to scrap the data.

**Difference and My Findings**

|  |  |  |
| --- | --- | --- |
| **Tweepy** | **GetOldTweets3** | **TwitterScraper** |
| Easy Approach to get Twitter data. Objective functions are already created | It has to be implemented by Python Library with various Models and creator functions. | It is simple script to scrape for tweets. |
| Cannot get historic Tweets that are older than past 1 week | Can get historic tweets but was not able to filter it by language as there is no setLanguage() criteria. | It is able to solve the problems that are faced in Tweepy and GetOldTweets and by using advance search , tweets are filtered . |
| Twitter restricts volume to tweets over 3000 | Maximum number is decided by the API request send per minute | It scrapes using internet speed or bandwidth. |
| Is Good for real-time data analysis. | Is Overall good for trending tweets | Is good and gives satisfying results. |