Web scrapping using Tweeter API

Step 1: Create a Twitter (X) Developer Account

1. Go to **https://developer.x.com/**
2. Sign in with your Twitter account.
3. Apply for a **Developer Account**.
4. Choose **Free**, **Basic**, or higher plan (Free plan has search limits).
5. Create a **Project** and **App**.
6. Get the following credentials:
    - **API Key**
    - **API Key Secret**
    - **Bearer Token**
    - **Access Token**
    - **Access Token Secret**

2. Install Required Python Library

The official Twitter API uses OAuth2 and OAuth1. A good wrapper is **tweepy**.

```
pip install tweepy
```

3. Authenticate with Twitter API Using Bearer Token

```
import tweepy

BEARER_TOKEN = "YOUR_BEARER_TOKEN"

client = tweepy.Client(bearer_token=BEARER_TOKEN)
```

4. Extract Recent Tweets by Keyword

Example: Search recent tweets containing the word "Nepal"

```
query = "Nepal -is:retweet lang:en"   # minus retweets, only English tweets


tweets = client.search_recent_tweets(
   query=query,
```

```
    max_results=100,  # Free tier max is usually 100 per request
    tweet_fields=["created_at", "author_id", "public_metrics"]
)
```

5. Save Extracted Tweets to CSV

```
import pandas as pd

data = []

for t in tweets.data:
    data.append({
        "tweet_id": t.id,
        "author_id": t.author_id,
        "created_at": t.created_at,
        "text": t.text,
        "retweets": t.public_metrics["retweet_count"],
        "replies": t.public_metrics["reply_count"],
        "likes": t.public_metrics["like_count"],
        "quotes": t.public_metrics["quote_count"]
    })

df = pd.DataFrame(data)
df.to_csv("tweets_dataset.csv", index=False, encoding="utf-8")
print("Dataset saved as tweets_dataset.csv")
```

6. Extract Tweets From a Specific User

Example: Get tweets from @elonmusk

```
import tweepy
client = tweepy.Client(bearer_token="YOUR_BEARER_TOKEN")
user = client.get_user(username="elonmusk")

tweets = client.get_users_tweets(
    id=user.data.id,
    max_results=5,
    tweet_fields=["created_at", "public_metrics"]
```

```
)

for t in tweets.data:
    print(t.text)
    print(t.created_at)
    print(t.public_metrics)
    print("----")
```