**NLP Project Report**

**Grp members:- Chirag Bhandakkar and Vipra Nagaich**

**Topic:-**

**Choice of Topic :-**

The topic that we have chosen is “Sentimental Analysis of Qoura

college reviews dataset ”.

It involves analyzing the reviews of a certain domain , like

College reviews, School Reviews etc. on Qoura and telling people

whether the College/ School is GOOD, BAD or AVG.

**Reason :-**

The Reason why we chose this topic is because we searched the

Internet but never came close to discovering a software or a

bot that scrapes the internet , analyzes the scraped content and

gives us a final verdict on a particular school or College.

**Problem Statement:-**

So , the problem statement states that we have to perform “Sentimental Analysis of Qoura college reviews dataset .”.

Thus, it demands that we make a stable framework wherein we

have to divide the overall task into three smaller but integral

frameworks.

* Framework 1 :

A framework which will be a GUI in which the user will have a proper interface to enter the name of College/School to get analysis about

* Framework 2 :

A framework which will search and scrape the entire internet for the

reviews and will do so based on predefined search criteria.

* Framework 3 :

A framework which will analyze the searched/ scraped data and

then categorize the gathered info into GOOD, BAD or AVG.

* Framework 4 :

Finally, we would require a framework which will collect the final   
 result and output from framework 3 and then display that result

to the user in a neat, streamlined manner.

**Approach and Solution:-**

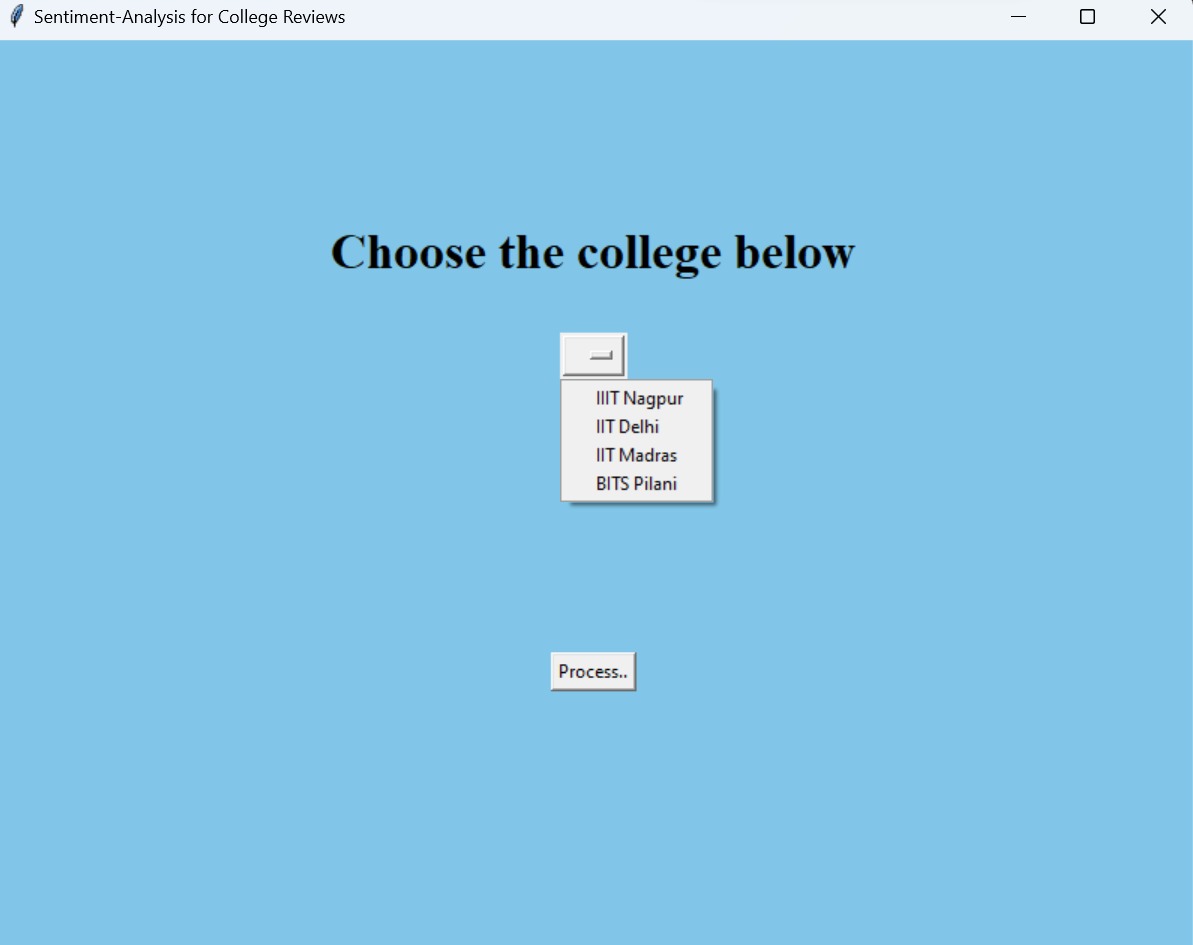
* The approach here is to use the concept of Web Scrapper , which will basically a python bot that will use the concept of web scrapping. In this we utilize and take help of a crawler.
* Then we will search the entire internet for the keyword that we wish to look for , like, the name of a particular college/ school.
* Further, we will use a predefined algorithm for the web crawler to give data which only pertains to a set no of words (like for eg: only sentences which contain two or more occurrences of the keyword, and which contain atleast 20-30 words.)
* Finally, we will use our personalized python algorithm and framework to perform Sentimental Analysis using NLP. Then the output will be displayed by the help of framework 4.

**Challenges and Obstacles:-**

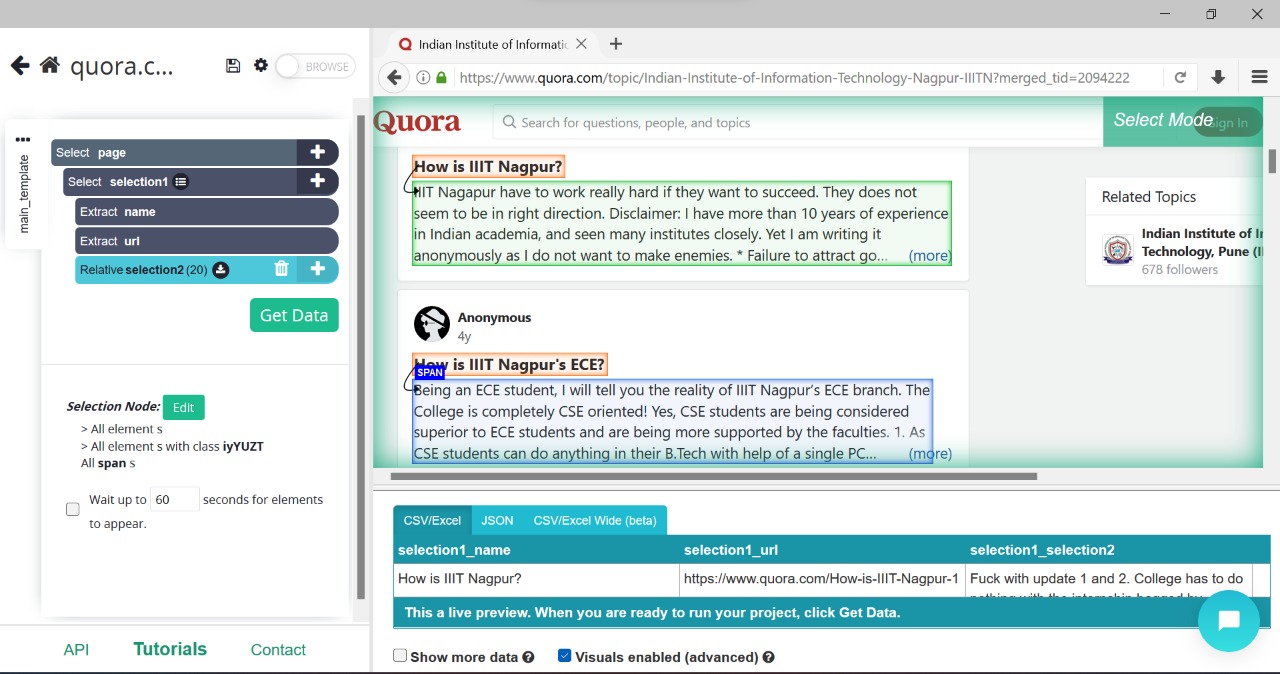
* API of Qoura is hard to find.
* Web Crawler might not find all the data pertaining to the criteria that we have found.
* GUI integration.

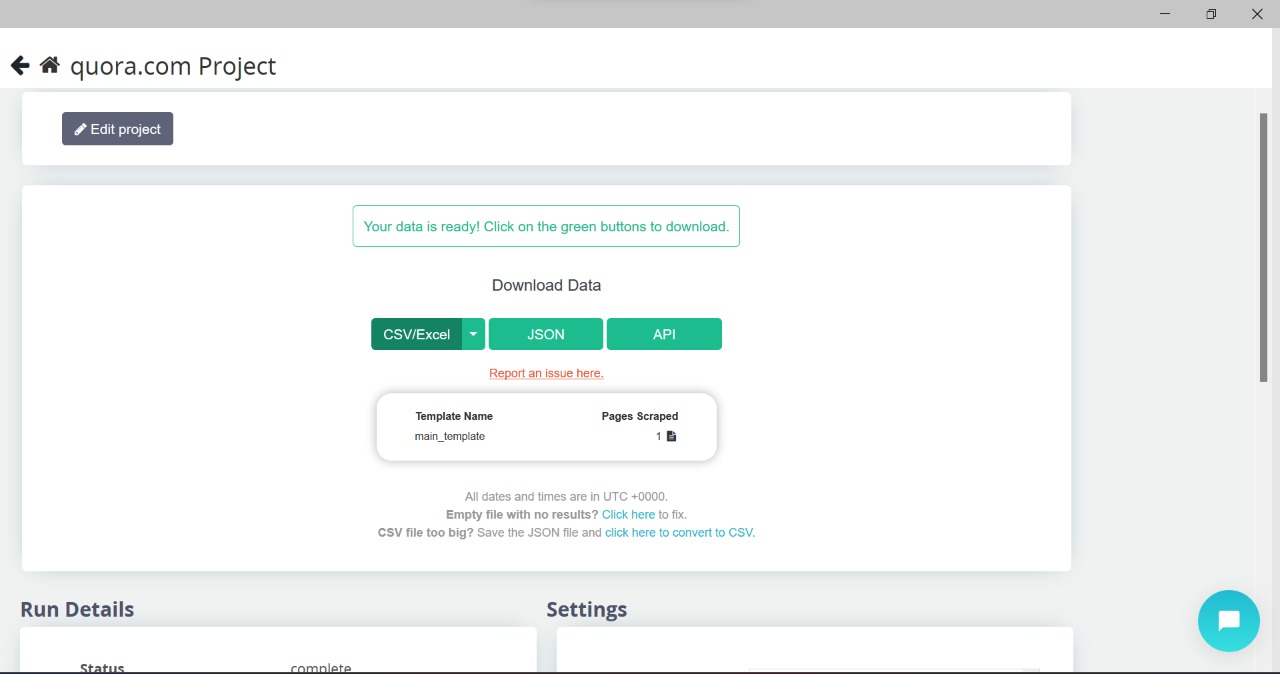
**Output Results :-**

GUI:-



Scrapping :-





Sentiment Analyzer:-

