

Abrasion of Current Weather of a City Using Variant Python Libraries and Weather Application Programming Interface (API)

H. Sumesh Singha¹, Dr. Bhuvana J²

¹Master of Computer Application, ²Associate Professor,
^{1,2}Jain Deemed-to-be University, Bengaluru, Karnataka, India

ABSTRACT

Web allows a wide extent of realities and information source set up by people. However, it will comprise of a gigantic arrangement of divergent and debilitated coordinated information. In metropolitan urban areas of India, there are different institution that measures climate such as The Energy Resource and Institute, Indian Institute of Tropical Meteorology and many more. There are also different sites that also provide accurate weather data of cities in India. For this research, I am using web scraping method to collect current weather report for a specific city in India from two different sources. The fundamental target of Web Scraping is to extract data from one or numerous sites and cycle it into basic structures such as JSON, text, CSV files. This incorporates web crawler and data extractor. Web crawler slithers all the connections present in a site page and stores them in an information base though data extractor extricates information from these put away connections. The information gathered by this web scraping procedure will shape a data set or information distribution centre that can be utilized for additional exploration on information digging for climate determining in metropolitan urban areas. These data stored in the database can be utilize to enhance the application and also meteorologist can study the weather pattern from the data and alert the people of the states from upcoming natural disasters.

KEYWORDS: Web scraping, web crawler, data extractor, meteorologist.

1. INTRODUCTION

The weather influences nearly all that activities we do in our day by day lives. It influences what we wear and what kinds of exercises that we do. In numerous callings, it influences either we should work, at what time to work, where should we work or we have to work for the whole work time in our daily lives. Weather have become an important part for us people. There are different factors that affect the weather such as temperature, windspeed, latitude, longitude, humidity and many others. The world wide web is a significant wellspring of data for some experts in different areas. It contains helpful and futile, organized and non-organized data, in various types of format, and from diverse sources. In India there are multiple organizations, such as The Energy Resource and Institute, Indian Institute of Tropical Meteorology and many more which produces daily weather reports. Apart from these organizations, there are multiple sites like www.wunderground.com, www.openweathermap.org that provides real-time current weather data reports. Every day the weather changes and the data generated for each day is stored in the database. These data can be structured or unstructured in nature. These data that are produced in large volume is known as Big Data. Handling such massive quantity of data is very complicated for a normal person. Scientist and researchers use various types of technologies to handle such data. For our research, we are going to use these datasets to generate weather of a city in our application. These datasets can be

How to cite this paper: H. Sumesh Singha | Dr. Bhuvana J "Abrasion of Current Weather of a City Using Variant Python Libraries and Weather Application Programming Interface (API)" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-1, December 2020, pp.593-596, URL: www.ijtsrd.com/papers/ijtsrd38037.pdf



IJTSRD38037

Copyright © 2020 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



study the weather and climate as well as to enhance the web application.

2. OVERVIEW OF WEB SCRAPING

Web Scraping, additionally called Web Harvesting, or Web data extraction is a way towards separating information from a site. On the off chance that you have ever duplicated data from a site and glued it into an Excel accounting page or Word doc then in fact you are scraping a site. Scraping manually every information from a site is very complicated and time consuming. Rather than sitting at a PC hitting Ctrl+C and Ctrl+V for every information you need, you set a web scraping code running that will separate your ideal information from the site and store it in your favoured document design very quickly. Today, web scraping is utilized in everything from land to web-based business. From the activity perspective, a web scraping look like manual reorder task. But there is some variation that is this task is finished in a coordinated and programmed way, by a virtual PC specialist. The initial step of web scraping script is to make a https request to the target website for the data of a specific URL. This step is accomplished by "Scraper". Once the site returns the html file the "scraper" extracts the information from that HTML file. For this research we are extracting the current weather report which is accomplished by a process called "parsing". The last advance is for the script to store this information in a CSV excel sheet, JSON or in an information base so it tends to be utilized physically or in another program.

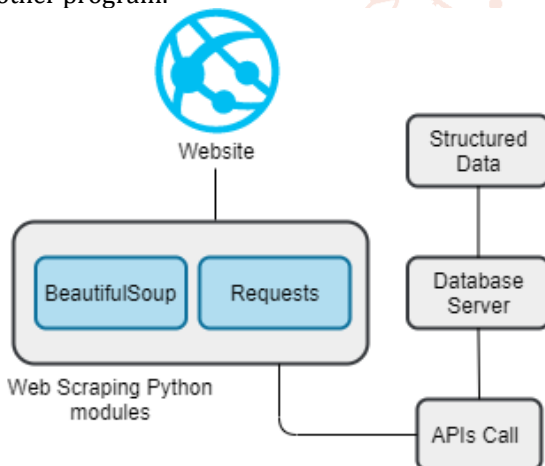


Figure 1. Basic architecture of Web Scraping

3. APPLICATIONS OF WEB SCRAPING

Applications of web scraping increases day by day. The number of applications accessible today is more than it was yesterday and in future these applications will keep on increasing. Below there are a-few number of fields where web scraping is mostly used:

- Weather forecasting.
- Comparison of prices of different product.
- Analysing advertisement.
- Web indexing & rank checking
- Analysing of graphical datasets.
- Optimizing search engines.
- Content production.

4. LITERATURE REVIEW

There are many approaches of web scraping. In paper [1] many different approaches are explained. In this paper i am going to explain only the basics of those approaches.

a) Mimicry Approach

In this approach the scrapper is preconfigured with the location of information gathered from the web page. This approach is well organized, yet it is less adjusted with regards to handle different heterogeneous sites.

b) Weight Measurement Approach

This methodology depends on a nonexclusive calculation which investigations the Document Object Model tree of a page and measures the heaviness of words in each branch.

c) Differential Approach

This approach explains that the content of the page will vary only from the body of the web page where both of the page belongs form the same web page.

d) Machine Learning Approach

The main motive of this approach is to physically examined sites pages on a huge sample and train an algorithm on it.

Also, paper [1] has explained different categories and tools for web scraping. Progression of data retrieval have increased significantly over the period of time. Information retrieval (IR) was used for web searches even before internet came into appearance. This IR was common since as ahead of the schedule as 1960 in business and knowledge applications. The capacity limit, accuracy, handling force of IR have dramatically increased over the period of time. Such turn of events and headway in the field of IR have likewise brought many changes in the progression of techniques of querying from the physical library-based approaches. After the creation of world wide web, the content of pages was very few and doesn't need IR to scrape information from them. Despite the fact that this strategy experiences a disadvantage of being restricted to thoughtfully homogenous writings documents permits the archives to be filed by the manner in which they are referred to in different articles. The most well-known data recovery is the ad-hoc querying where a query finds for a bunch of static reports. This information retrieval technique was used by commercials search engines like Google and Alta Vista. The downside of this strategy is that the precision is very low. Agent paradigm is another type of promising innovation for data retrieval. IR system became much more scalable, adaptable, interoperable with the addition of agent in it.

5. PROBLEM STATEMENT

There are multiple web applications that provides real-time weather report. Obviously, these applications would not deliver false weather report but still there are some misbelief whether the weather report is accurate or not. For this research I am creating a uncomplicated and straightforward web application for my academic purpose which will display current weather of a city. My project guide came with a problem and asked me whether the report displayed in my application is accurate or not? To overcome this problem, I came with an idea and thought to build a script that will scrape weather information from two different sites.

6. PROPOSED SYSTEM

Different Strategies and tools are used by web scraping technology to scrape website data and export them into our local database in the form of JSON, CSV, XML etc. Python provides different libraries or modules that can be used for scraping datasets from the HTML web documents such as

Requests, BeautifulSoup, Scrapy, Selenium. To overcome the problem of my project guide my application will use Requests to make an API call to www.openweathermap.org and display the report in the web application whereas BeautifulSoup will scrape HTML page from www.wunderground.com and exhibits the topical weather temperature of a city on the console. This application is created only for academic purpose. This application will show temperature in a graphical representation along with the other weather factors like windspeed, humidity weather description any many more also it has a very simple user interface with only a text field to enter the city name and search button. This application will be deployed on a cloud service i.e. www.pythonanywhere.com which provides three months trail hosting service for your web-based python application. This cloud service is fully automated, easy and Simple to use.

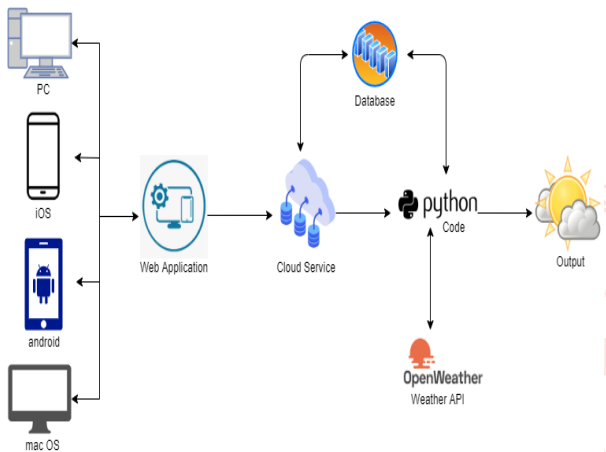


Figure 2. Proposed System Architecture

7. METHODOLOGY

BeautifulSoup is one of the important python libraries which is used to handle XML, HTML and other types of mark-up languages. Whereas Requests library is used to send HTTPs request without the need of manually querying string to the URL. Inside the request there is a module urllib3 which keep-alive the URL and allows constant polling of data. Following principles are followed to run the application:

- First the application is accessed by a web browser.
- Then we need to entered the city name for the weather report.
- After that the script sends a Https Get request to the targeted site(www.openweathermap.org) using the request module.
- Next the Scripts check whether the city exist in the world or not from the database server. Depending on the city existence a successful or error memo is flashed on the application.
- Similarly, Beautifulsoup parse the weather temperature from the target site(www.wunderground.com) which is a HTML page and display it in the console.
- If any updates are required then a POST request is send using the Python requests module.

8. SEQUENCE DIAGRAM

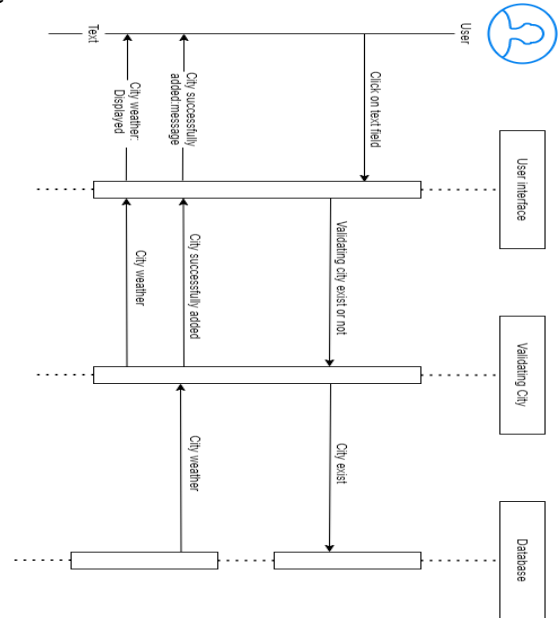


Figure 3. Sequence Diagram

9. MODULE DIAGRAM

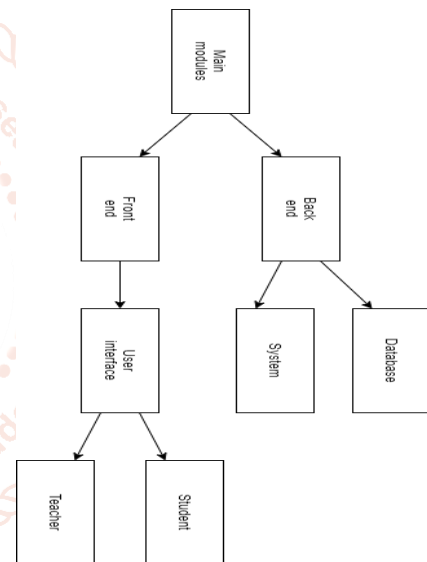


Figure 4. Module Diagram

10. RESULTS

```
PS C:\Users\Sys\Desktop\weather_forecast> flask run
* Serving Flask app "app.py" (lazy loading)
* Environment: development
* Debug mode: on
* Restarting with stat
* Debugger is active!
* Debugger PIN: 103-878-889
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```

Figure 5. Running the Application on Localhost

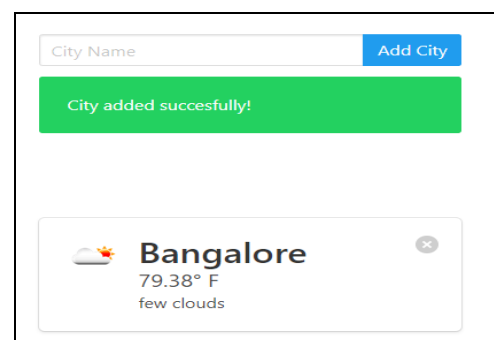


Figure 6. Weather Report of a City (from www.openweathermap.org)

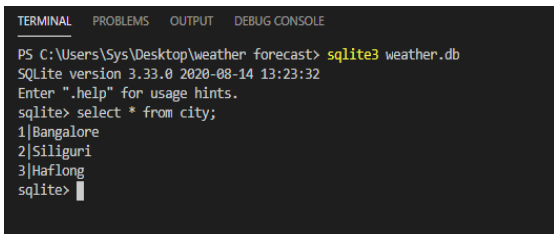


Figure 7. Cities names in the Database

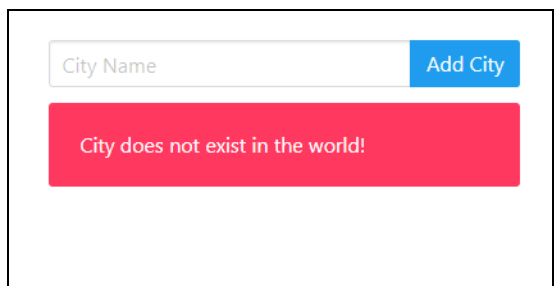
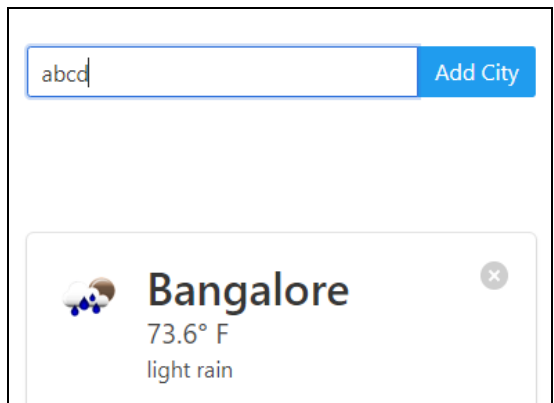


Figure 8. Error Message of Invalid City

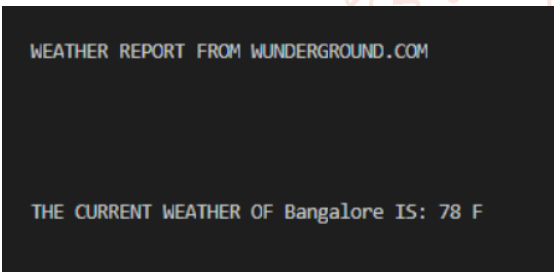


Figure 9. weather report of a city (from www.wunderground.com)

11. CONCLUSION AND FUTURE ENHANCEMENT

With the help of this application generating current weather reports become much easier. With minimalistic user interface anyone can use this application and observe the current weather of their city. There are very few chances for this application to malfunction as it used APIs for generating the weather report. In future the database can be modified to stored not only the names of the cities but also the weather factors that affect the climate. Using that data, meteorologist can study the weather pattern and generate report for upcoming weeks, months and years. This will help to alert people about the natural disasters at an early stage so that

they can take required precautions and safety measures. Also, this will help farmers to decide whether the weather is suitable for planting, watering, and other harvesting activities.

REFERENCE

- [1] R. DIOUF, E. N. SARR, O. SALL, B. BIRREGAH, M. BOUSSO and S. N. MBAYE, "Web Scraping: State-of-the-Art and Areas of Application," IEEE International Conference on Big Data (Big Data), pp. 6040-6042, 2019.
- [2] F. Y. N. Kunang and S. D. Purnamasari, "Web Scraping Techniques to Collect Weather Data in South Sumatera," INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING AND COMPUTER SCIENCE (ICECOS), 2018.
- [3] M. S. Parvez, S. S. Rajendra, K. S. A. Tasneem and K. R. Bodke, "Analysis Of Different Web Data Extraction Techniques," International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, 2018.
- [4] P. Ashiwal, S. R. Tandan, P. Tripathi and R. Miri, "Web Information Retrieval Using Python and BeautifulSoup," International Journal for Research in Applied Science & Engineering Technology (IJRASET), pp. 335-339, 2016.
- [5] S. Singh and R. Jain, "Weather report on metropolitan cities in India using web scraping technique," International Journal of Advance Research, Ideas and Innovations in Technology, pp. 1068-1070, 2019.
- [6] A. V. Saurkar, K. G. Pathare and S. G. Gode, "An Overview On Web Scraping Techniques And Tools," International Journal on Future Revolution in Computer Science & Communication Engineering, pp. 363-367, 2018.
- [7] B. G. Dastidar, D. Banerjee and S. Sengupta, "An Intelligent Survey of Personalized Information Retrieval using Web Scraper," IJ. Education and Management Engineering, 2016.
- [8] R. Samya and R. Rathipriya, "Predictive Analysis for Weather Prediction using Data Mining with ANN: A Study," International Journal of Computational Intelligence and Informatics, pp. 150-154, 2016.
- [9] D. Karthikeyan and M. Shivaranjani, "A Review of Weather Forecasting Using Data Mining Techniques," International Journal Of Engineering And Computer Science, pp. 19784-19788, 2016.
- [10] V. Singrodia, A. Mitra and S. Paul, "A Review on Web Scraping and its Applications," International Conference on Computer Communication and Informatics (ICCCI -2019), 2019.