International Conference

Computational Intelligence & Data Science

February 4, 2021

East West College of Engineering,
Yelahanka New Town, Bengaluru

In collaboration with Universal Innovators

International Conference

Computational Intelligence & Data Science

February 4, 2021 | East West College of Engineering, Yelahanka New Town, Bengaluru

BACKGROUND

International Conference on Computational Intelligence and Data Science (ICCIDS-2021) is organized with the objective of bringing together scientists, professors, research scholars, students and industrial experts in the field of Data Science and its various Applications to a common forum. Overall, the conference will provide the researchers and attendees with prospects for national and international collaboration and networking. ICCIDS-2021 will be held at East West College of Engineering, Yelahanka New Town, Bangalore. All the accepted papers (after double blinded peer review) are published in Elsevier SSRN.

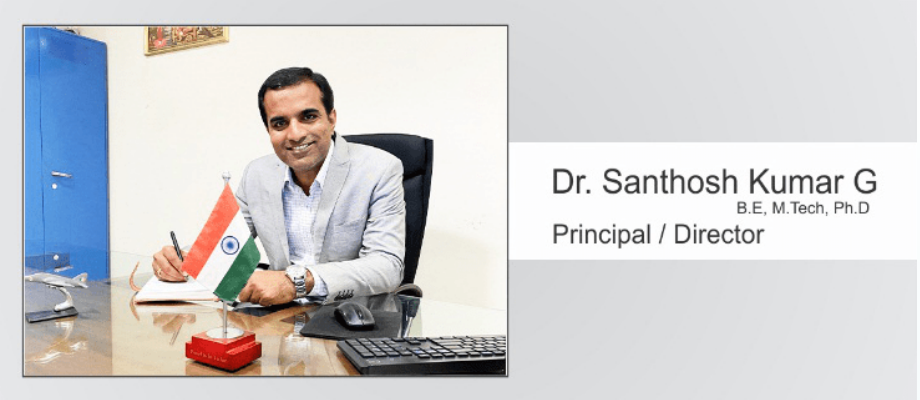
OBJECTIVE OF THE CONFERENCE

The primary objective of the conference is to provide opportunity for academicians, industry experts, practitioners, professionals, researchers and policy makers from different fields to engage in discussion based on issues related to dynamic and challenging economic environment. It will also provide a platform to get acquainted with latest developments and trends in the economy and business environment coupled with their implications for the organizations. The forum will facilitate interaction among members inside and outside their own respective disciplines to enjoy the fellowship of other professionals and scholars in the field.

FOCUS AREAS

- To facilitate discussion on Computational Intelligence & Data Science that will expedite nation's growth and stability.
- To bridge the gap between academic wisdom and practical knowledge.
- To develop a series of focused research topics that will benefit the academic world.

Director General Chair:



Chief Patron:

Patrons:

CONFERENCE CHAIR:

Dr. Aravind K U
Professor & Head of Department – M.E
Member, RESEARCH AND INNOVATIVE CELL (RIC)
East West College of Engineering Yelahanka New Town, Bangalore

Convener:

Dr.Piyush Kumar Pareek,
Professor & Head of Department of Computer Science and Engineering
East West College of Engineering, Bengaluru

ADVISORY COMMITTEE

- Dr Ashish Khanna, Computer Science and Engineering Dept., Maharaja Agrasen Institute of Technology, GGSIPU, New Delhi
- Dr Deepak Gupta, Computer Science and Engineering Dept., Maharaja Agrasen Institute of Technology, GGSIPU, New Delhi
- Dr.E.Madhusudhana Reddy, Professor, Department of Computer Science and Engineering , Guru Nanak Institutions Technical Campus(AUTONOMOUS), Telangana, India.
- Dr Raghavi K Bhujang, Head of Computer Applications and Information Science, ISBR Business School, BLORE
- Dr. Manoj H M, CSE, Affiliation: BMS Institute of Technology and Management, Bangalore
- Dr Dinesha H A, Prof & Head, CSE, SGBIT, VTU, BELAGAVI
- Dr.C.P.S.Prakash, Department:Mechanical engineering,Affiliation: Principal, Dayananda Sagar College of Engineering
- Dr Rajnish Ratna Gedu college of Business Studies Royal University of Bhutan
- Dr. P Priyanga, Associate professor , Computer science and engineering , Global Academy of Technology, Bangalore
- Dr S.Venkatesan , Prof &Dean , School of Computer,Information &Mathematical Sciences , BSA Crescent institute of science and technology , Chennai
- Dr Latha p h, Hod dept of ISE,Vtu, Sait
- Dr R V Praveena Gowda, PhD , Principal,G Madegowda Institute of Technology, Karnataka
- Dr.Sunanda Dixit , Computer Science and Engineering Department,BMS Institute of Technology and Management, Bangalore
- Dr. Sonal Amit Jain , Computer Science Engineering, JK Lakshmipat University, Jaipur
- Dr.Sangeetha V , Information Science and Engineering , SVIT , Blore
- Prof. Chandrasekar Shastry,Dean PG Studies, Faculty of Engineering & Technology,Jain (Deemed to be) University,Bangalore, Karnataka.
- Prof.Santosh Kumar J , Associate Professor,Department of CSE,K S School of Engineering and Management , Holiday Village Road, Vajarahalli Village, Mallasandra, off, Kanakapura Rd, Bengaluru
- Dr. Raghavendra S , Associate Professor,Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST DEEMED TO BE UNIVERSITY. BENGALURU

TECHNICAL PROGRAMMING COMMITTEE

- Prof. Nishanth.P,H.O.D,Department of Aeronautical Engineering,East West College of Engineering, Bangalore.
- Prof. Umapathi M A,H.O.D ,Civil Engineering Department,East West College of Engineering,Bangalore.
- Prof. Rashmi H N,H.O.D,Department of Electronics and Communications Engineering,East West College of Engineering,Bangalore.
- Prof. Venkatesh G,H.O.D,Department of Basic Science and Humanities,East West College of Engineering,Bangalore.
- Dr.Geetha.N,Training and Placement Officer,East West College of Engineering,Bangalore.
- Dr. Babu N,Department of Mathematics,East West College of Engineering,Bangalore.
- Dr. Gufran Ahmad Ansari, Prof. & Head, Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai
- Dr.Narasimha Murthy M S , CSE , EWIT,Blore

CALL FOR PAPER AND SUGGGGESTED THEMES

Original papers on the following indicative topics, but not limited to, are welcome to be submitted for the conference:

- Artificial Intelligence and Machine Learning.
- Data Mining and Text mining.
- Big Data Analytics and Metrics.
- Cloud Computing.
- Mobile Computing.
- Natural Language Processing and Machine Translation.
- Parallel and Distributed Algorithms.
- Pattern Recognition and Analysis.
- Modeling Systems and Software Engineering
- Ubiquitous and High-Performance Computing.
- Information, network and applications security.
- Access control, web filters and firewalls.
- Intrusion detection and incident response.
- Vulnerability assessment and mitigating attack.
- Social engineering and phishing attacks
- Information security management and governance.
- Software security and secure coding.
- Cognitive Radio and Cognitive Networks.
- Communication Architecture.
- Control Systems and Applications.
- Real-Time Networking.
- Satellite and Space Communications.
- Smart Grid Systems and Intelligent Control.
- Smart Communication System Models.
- Industrial Networks and Automation.
- Cognitive Approach for Robotics.
- Mobile Robots and Intelligent Autonomous Systems.
- Human-Machine Interfaces.
- Image and Video processing.
- Object recognition.
- Speech processing.
- Medical imaging.
- Renewable Energy and Power System
- Power Electronics
- Electromagnetic and Remote Sensing
- Integrated Systems, Circuits and VLSI Design
- Nano Technology & Photonics
- MEMS, NEMS
- Systems, Controls
- Robotics
- HVAC and HVDC Systems

ACCEPTANCE AND PUBLICATION

After a double-blind peer review, qualifying Regular Papers may be accepted as either Full Papers or Short Papers.

- All accepted and presented papers of the conference will be included in ELSEVIER-SSRN digital library (ISSN 1556-5068).
- The papers must be part of the worldwide scholarly discourse in the field covered by the library. The reviews will be done to make sure papers are relevant for the chosen classifications to ensure subscribers receive relevant content.
- NO extra fee is charged from authors for inclusion in the digital/eLibrary. Most of the articles in the digital/eLibrary will have the full text of the article available online in open access so authors can view/download their papers at NO extra cost.
- NO subscription is required on the part of any organization, school, or individual author as a prerequisite to submission or download.
- Authors will grant a non-exclusive, revocable license that allows providing services to users.
- The proceedings page would be permanently available, not need to pay an additional fee to keep the papers available on the digital library.

INDEXING

Post-conference, SSRN proceedings will be made available to the following indexing services for possible inclusion:

- ISI Conference Proceedings Citation Index - ISI Web of Science
- Google Scholar
- Scopus
- DBLP

Depending on the focus of the particular indexing services, they may decide to include or not. If included one can expect it in 10-12 months. DBLP and Google Scholar are fast.

*Some high-quality papers will be considered for Scopus Indexed Springer Proceedings (The selection of paper in Springer/SSRN will be made while sending the acceptance notification).

https://online-conf.com/iccids-home.html

3rd INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND DATA SCIENCE
February 4, 2021|EAST WEST COLLEGE OF ENGINEERING

ELSEVIER

AN APPROACH TO ANALYSIS ON COVID-19 DATA THROUGH WEB SCRAPER AND VOICE ASSISTANT – A SURVEY

Mr. Pradeep Kumar KG[1] , *Ms. Harshitha C Prabhu*[2] , *Ms. Chaitrali Salian*[3] , *Ms.Sameeksha Shetty*[4] *and Ms. Shreelatha*[5]

Department of CSE, A J Institute of Engineering and Technology, Mangaluru (KA)

putturpradeep@gmail.com, harshithaprabhu11@gmail.com, chaitralisalian30@gmail.com

sameekshashetty90@gmail.com, shreelatha104@gmail.com

ABSTRACT:

Corona virus pandemic has been recognized as a global threat across the world and many methods were adopted for the prevention of this disease. This pandemic has caused global and economic disruption which resulted in numerous Covid-19 cases across the world. To know the number of cases and keep a track of this pandemic situation we need to collect the live data sets from the worldwide corona virus records. This can be achieved by the technique of Web Scraping which enables the extraction of live data sets from a specific platform. It facilitates the user to access the World Wide Web wherein specific data is gathered, copied from the web and then it is stored in a central local database then provides ways to retrieve and analyse the data. This estimate is to design a platform where you can obtain the live data sets and have a compact knowledge about the present scenario. This is an elementary approach to scrap the live data sets through a user interface from the Worldometer Covid-19 data set with the aid of a Google voice assistant. To implement this scheme, we use the Python programming language. To effectuate this task, we acquire the process of making API calls to the Worldometer Covid-19 website and simultaneously we will make use of the regular expressions to extract the data from the web page. However, this action includes a series of tactics that has to be recognized analysed and accomplished sequentially. Initially the input is given by the user in the form of speech. Then the required contents are searched and matched with the user's input. If the contents match then with the help of a Google voice assistant result is obtained which is the output in turn.

An Approach to Analysis on Covid-19 Data Through Web Scraper and Voice Assistant – A Survey

***Mr. Pradeep Kumar KG*[a*], *Ms. Harshita C Prabhu B*[b], *Ms. Chaitraly*[b], *Ms. Sameeksha Shetty*[b], *Ms. Shreelatha*[b]**

[a] *Faculty, Department of CSE, A J Institute of Engineering and Technology, Kottara Chowki, Mangaluru 575006 INDIA*

[b] *Final Year BE, Department of CSE, A J Institute of Engineering and Technology, Kottara Chowki, Mangaluru 575006 INDIA*

* *putturpradeep@gmail.com*

Abstract:

Corona virus pandemic has been recognized as a global threat across the world and many methods were adopted for the prevention of this disease. This pandemic has caused global and economic disruption which resulted in numerous Covid-19 cases across the world. To know the number of cases and keep a track of this pandemic situation we need to collect the live data sets from the worldwide corona virus records. This can be achieved by the technique of Web Scraping which enables the extraction of live data sets from a specific platform. It facilitates the user to access the World Wide Web wherein specific data is gathered, copied from the web and then it is stored in a central local database then provides ways to retrieve and analyse the data. This estimate is to design a platform where you can obtain the live data sets and have a compact knowledge about the present scenario. This is an elementary approach to scrap the live data sets through a user interface from the Worldometer Covid-19 data set with the aid of a Google voice assistant. To implement this scheme, we use the Python programming language. To effectuate this task, we acquire the process of making API calls to the Worldometer Covid-19 website and simultaneously we will make use of the regular expressions to extract the data from the web page. However, this action includes a series of tactics that has to be recognized analysed and accomplished sequentially. Initially the input is given by the user in the form of speech. Then the required contents are searched and matched with the user's input. If the contents match then with the help of a Google voice assistant result is obtained which is the output in turn.

Keywords

Web Scraping, Worldometer, Covid-19, API Calls, Google Voice Assistant

1. Introduction

Web Scraper is a process of fetching and mining for the essential data by scrawling through a web page. Web scrapers function in an ideal way wherein the content of the page maybe parsed, searched or reformatted. Then the data which is collected is copied to a spreadsheet or it can also be stored in a database for further analysis. With the end goal of analysis, the data needs to be segregated into different advances further on, for example, beginning with its specification collection, organizing process, cleaning process, re-dissecting, applying different models and various algorithms and the eventual outcome. There are two ways of extracting data from websites, first one is manual extraction technique and the second one is automated extraction technique. Web scrapers assemble site information similarly to how a human would do that is the scraper goes onto a webpage of the site, gets the pertinent data, and push ahead to the following website page. Each website has an alternate structure that is the reason web scrapers are generally built to search through a website. Web scraping can assist in obtaining any sort of information that is intended. We would then have the option to retrieve, analyze and utilize the information in the manner we need. So web scraping streamlines the way towards deriving information, speeds it up via automation and makes simple to access the extracted data by offering it in a CSV pattern. Web scraping commonly extracts a lot of data from websites for example, monitoring interests of consumers, monitoring price i.e. value observing, advancing AI models, monetary information accumulation, tracking news, and so forth. Hence no doubt that web scraping is a programmed technique to acquire a lot of data from websites. Web pages contain information in a unstructured format in a HTML design which is then changed over into organized information in a spread sheet or a data set so it tends to be utilized for different applications. Web scraping requires two sections in particular the crawler and a scraper. The crawler is a man-made AI algorithm that parses the web to look through the specific information needed by following links over the internet. The scraper, is a particular tool made to extract data from the sites.

Data analysis is a method to extract solutions to the problems by interrogation and interpretation of data. The web scraper program is planned to be thorough for all significant information from various online stores and mining, and gathering it into the new site. A web scraper works as an API to extract data from a website here it is Worldometer in which data is accessible free of cost to end users. Web data scraping techniques are used by extensive number of people used in exploration and business for making content or offering reactions

to grow the precision of business for promoting that empowers people to convey assets in progressing and building up the business. The project mainly covers the fundamentals of web scraping, voice assistance using python. It makes use of a tool called Parse Hub. Parse Hub is a free web scraper that is very powerful and is easy to use. This tool allows you to scrap the web merely just by clicking on the elements you would like to get out from that website. We will be scraping from the very famous website for statistics which is Worldometer that is regularly updated with the coronavirus information. The Parse Hub is clearly very efficient and uses an Artificial Intelligence technology to understand which elements you would want. The process of web scraping can be carried out by the help of regular expressions and data visualization. The method of extraction of data may seem easier such as getting the data from textbook which appears as ready to use in a clean format. But in reality, it involves a lot of steps which has to be performed. It's not as easy as it appears to be. In this case obtaining data means getting our hands dirty which simply means pulling or scraping data from the web. The implementation is done by the programming language Python. It has great tools for extraction of data. It requests library for retrieving content from a web page, and bs4 (BeautifulSoup) for extracting the relevant information. These 2 libraries are often used wherein first a GET request is made to a website. Then a BeautifulSoup object is created from the content that is returned and parses it using several methods. Regular expressions are used for matching patterns. The basic idea for applying regular expressions is to define a pattern that is to be matched in a string and then search in that string to return the pattern matched. Regular expressions come up while parsing string information. Regular Expression is a unique succession of characters that causes match or search different strings or strings set, utilizing a specific syntax to match a pattern. Pattern matching is search for substring in a given string. Regular expressions are used in validation, extracting and classification. Writing and understanding regular expressions requires to have knowledge and experience. If a single character written wrong would cause different pattern to be matched. Regular Expressions are basically a highly specific programming language that is embedded in Python and made accessible through the re module. Utilizing this little expressions, we can indicate the principles for the arrangement of strings that needs to be matched; this set may contain English sentences, or email patterns, TeX commands, number matching and many more. Python offers two diverse operations dependent on these expressions: Match that checks for a matching pattern just towards the beginning of a string and Search that checks for a matching pattern that can exist anywhere in the string. The Python module re offers full help for Perl-like expressions for pattern matching in Python. The re module raises the special case re-error if a mistake happens while accumulating or utilizing regular expressions. Pattern matching is checking whether a particular grouping of characters, sentences or tokens which exists among the given data. For this situation regular expressions permits us to locate the particular patterns of words and obtain the data we need more effectively than physically looking for explicit characters in the website. Pattern matching is used for checking a given group of tokens for the presence of certain pattern. A traditional approach to perform pattern matching is to look for a sub string in a given string. It can be got via looking a "sub string or pattern" through the given string. Contrasting and the other algorithms used for pattern matching RE can uphold in a more extensive way to describe patterns. A Regular Expression is called as regex or regexp, which is string and its layout depicts a bunch of strings. With these strategies the execution of this project is worked out. The web scraper program is anticipated for comprehensive for all critical information from various online sites, collecting and mining. The tool used for scraping is programmed to derive a lot of sensible data from the web.

2. Literature Survey

Python has rich set of libraries that are available to extract digital contents from the internet. Amongst the libraries available, the following three are popularly known: BeautifulSoup, LXml and RegEx. A statistical study was carried out on the datasets that was available, it shows that RegEx was capable to deliver the requested data on an average of 153.6 ms. But RegEx has that drawbacks of limited data extraction for web pages with more of HTML inner tags. Due to this demerit RegEx is used to perform only moderate complex data extraction. Some of the other libraries like BeautifulSoup and LXml are able to extract content of the web pages under complex environment which yielded a response rate of 457.66 ms and 203 ms respectively. These libraries are based on the Document Object Model (DOM) proved to be accessible libraries. Web scrappers are the one by which regional languages available in social media are influenced and hence modern content grading system are developed. The survey conducted in this paper proves the overwhelming performance of RegEx under varying situations [1].

The main goal of data analysis is to obtain the information which is useful from data and take decisions on the basis of data analysis. Web scraping refers to collecting data from the web. Web scraping is also popularly known as data extraction. For the purpose of analysis the data can be divided into several steps such as cleaning, organizing etc. Scrapy is most popularly used open source for collecting the data that is needed by the user. The main purpose of using scrapy is to scrape the data from its sources. Scrapy, which is a web crawler and based on a python programming language, is very helpful to get the data which we require by making use of URLs which are necessary for scraping data from its sources. Web scraper is an API which is useful to get data from a website. Scrapy provides all the tools which are necessary for extracting the data from a website and then process the data as per the needs of the user and store the data in specific format as specified by the users [2].

Regular expressions can be used in extraction, classification and validation of data. In order to understand the concept of regular expressions one must have good knowledge on it. Our main aim is to learn about how the regular expression is changing with time. If the scope of regular expressions increases as time goes on then we need to match more strings outside the regular expressions language.

Otherwise if its importance does not increase as passes then we need to focus on matching strings inside language of regular expressions. It is a language which can be used in describing a set of strings which can be matched by it, and generally there can be numerous ways for expressing this. The regular expressions can be generated from huge number of strings which are labelled or it can be generated by using the previously existing ones. Usually the editing in regular expressions involves changing various features, and usually the adding of new features is more when compared to deletion of the existing features. Based on the range of the features which are added newly and the features which are removed from the existing ones we can construct mutation operators in new regular expression which will be generated and it also provides guidelines for the process of mutation generations. The editing in regular expressions however do not affect escape characters such as backslash, dot, question mark etc. In general, the process of defining operators in mutation depends mainly on the changes which are likely to be present among features of regular expressions and also depend on changes which are most likely to be present inside regular expressions features. Mutation testing: The ability of the generic set of the strings in exposing all the faults which are possible to be present in the regular expression which is user tested [3].

The internet mainly considers web pages which covers a large quantity of explanatory substances which include text, audio, graphical, video etc. This process, The Web Scraping which mainly deals with the collection of raw data from the website. It is a process wherein you extract the automation of data very quickly. The process enables us to extract a particular data requested by the user. The most popular used method consists of constructing individual web data scraper by making use of any known language .The tools which can be used for the purpose of scraping data are regular expression and rvest. Major applications of web scraping are research, marketing, combination of data mining. There are three methods which are implemented for the process of web scraping in a manual way. The software's which can be used in scraping of data from web are frameworks, libraries and desktop based environments. Regular expressions play a very vital role in web scraping. This method involves automatic data extraction in spite of extracting it manually. Here data from websites are extracted and stored. Web scraping is used on aspects which can be about comparison, detection and integration. Tools used for web scraping are rvest and regular expressions. The three methods of web scraping can be divided as syntactic, semantic and computer view website analysing. This process has made many processes simpler by scraping and extraction of information through the web. This paper has given the information about various aspects of web scraping [4].

Speech to text mechanism has been implemented in numerous devices in the recent years. It provides a quicker alternative for typing text while freeing hands of the user to perform various other tasks. Moreover, they offer a new mode of accessibility choices for people having different types of disabilities, enhancing their involvement in a world that is moving towards an era which is touchscreen - dominated. This paper proposes that a high compression frequency and a rise in the unambiguousness of messages transferred can be achieved by combining automatic speech transcription in addition to speech synthesis technique as an alternative for typical voice codes. To evaluate the likelihood of an alternative voice codec along with speech transcription and its synthesis, this paper aims at evaluating the unambiguousness of messages interchanged by this method against 8.0kHz and 8 bit PCM speech samples are compressed when using A-law codec. A bunch of pre-recorded samples of speech are encoded utilizing both the techniques and consequently introduced to a bunch of members to be interpreted in a notebook given by the analysis conductor. This evaluation was established on the Semantically Unpredictable Sentences (SUS) test. After record stage, data were normalized and assessed for their Word Error Rates (WER) and Levenshtein distances. These algorithms think about first sentence against an expression translated by members and return a number that increments relatively with the quantity of changes needed in-order to change the previous into the one that is last mentioned. The numbers are then gathered by encoding cycle and afterward arrived at the average to measure the degrees of understandability they had accomplished. These outcomes show a comparable degree of coherence between the compression techniques tested, which has a normal (7.0%) diminishing in the accurately translated words between the best and the most noticeably terrible gathering. This paper reasons that it's possible to utilize automated transcription of speech and combination of voice as speech codec, providing enormous amount of savings of data transmission utilization for the client at the expense of some degree of clarity in its communication. Some of them have additionally advanced Skype's functionalities to give a speech specialized instrument to individuals with extreme speech impairment utilizing a speech to speech pipeline, that was made using a speech to text transcription step, a machine interpretation engine and a text to speech synthesizer. The outcomes additionally suggested this pipeline was adequately incredible to execute continuously. There are additionally VoIP application built up that depends on similarly created speech-to-speech pipeline to give an interpretation device to VoIP application which in addition to a tool like drawing board used to provide directions in many languages to the tourists. The utilization of a speech to speech pipeline isn't new in this field, as it was been utilized on different occasions to give computerized interpretation administrations and assistive abilities to platforms as Software-as-a-Service (SaaS). Because of complex speech transcription issue, developers adopt to third party cloud technologies developed by innovation goliaths like Google and Microsoft execute both the synthesis and speech transcription. Due to this it is expected by them to transfer and download the streams of speech to the service provider's servers before any of the changes is been done to the source material. Such streams may speak to a piece of the data used by these arrangements, that could be reduced if both the transcription stages and synthesis were being executed locally, in the cell phone itself. On the far edge, it was possible to change the content stream back to voice samples using the Text To Speech API

recently included among the standard android library, which was tried to work in flight mode and could be easily be integrated to the testing application [5].

Web scratching is data collection method by means of a program that communicates through an API (Application Programming Interface). The web scratching strategy is generally done by creating programs that consequently queries the web server, requesting for data (ordinarily HTML and different types of pages), at that point parses the data to extract required info. Web scraping utilizes different programming and innovation strategies, for example, natural language parsing, data security and data analysis. Approaches used for making weather forecasts are profoundly reliant on the observed data and some of the strategies and techniques used in forecasting the weather. Data is important in the direction of gauging data not just at a point for strengthening the weather forecast analysis. It is required at certain points for measuring data to see the movement of wind, atmosphere, direction of clouds, and so on. Hence, the weather data is substantial. Numerous gadgets and strategies are accessible which are generally free and simple to utilize. This paper, talks about the subject of data mining that uncovered scientific classification in web mining and contrast it and web scraping. The beginning phases of the analysis stage is considering the structure of HTML doc from all sites that will be scratched. This cycle is done to sort elements and data to be extracted or stored. The subsequent stage is to make a crawler program made with Python script utilizing the BeautifulSoup and the Requests library. The consequences of web scraping are being stored in an Excel file. Creating a task Scheduler that runs scripting data scraping intermittently each 60 minutes. Auto assignment scheduler scraps data through the website and save into the outcomes file. The following step of web scraping to extract the crawled data. Extraction step is finished with the assistance of instruments Pentaho Kettle. From the obtained data, data cleaning is taken place to eliminate superfluous data, for example, units of stored variables. Change data to change configurations and information structures varying (e.g., city information, time arrangement, date and more). And do documents scraped into one record to encourage the analysis process obtained as per the needs during application development. Make insights of weather data and analysis of acquired by the necessities of utilization advancement. Information statistics are made utilizing Python. Information can be proceeded varying by doing the process of grouping the data. For analysing, data can be introduced in the form of table, charts and graphs. Likewise, the histogram can be introduced in every day, month to month and twelve-monthly structures as needed by utilizing the group function by Date variable. After the information sufficiently gathered (over one year), at that point the following phase of exploration will be made climate forecast utilizing the methodology of Data Mining and Machine Learning. Additionally, the future consequences of information gathered by web scrapping procedures will be used for the study of relevant weather patterns used in decision making. And the subsequent information is extremely definite and can be utilized further for data analysis. Limitations of this research are recently created analytics data in the form of statistical data that has been running (one month). The data collection process will keep on running to deliver the climate dataset of the urban areas in South Sumatra. Once the dataset is gathered, significant research will be undergone to foresee the climate just as climate pattern analysis decision making in transportation and agriculture [6].

The Web Scraping morals and methodology depend on the working of how the web scraper is planned. The three strategies in the web scraping method are, it draws the ideal links from the web page, and then data retrieval process is done to get the information and then information is put in a csv file. Execution completes by utilizing the Python Programming language. Data analysis is the strategy wherein the answers for issues extracted by means of cross examination and interpretation of data. This paper manages the significant parts of Web Scraping. It also enhances and upgrades the indispensable uses of web scraping. It returns to the current framework for use, ideas, classes and its assets through recognition of its limits. The various approaches used in this paper for web scarping involves - :Differential approach, Weight measurement approach, Mimicry approach and Machine learning approach.Browser extensions utilized in this paper are Data Scraper, Agenty, Spider and so forth. There is an extension called Data Miner which allows data extraction in a csv document. The plugin Cloump U-Scraper, Outwit Hub are some powerful extensions. Applications which are utilized as key elements in quest for job search engine. There are some other applications wherein client share links to study which is quiet decent and widespread. Search engine Blekko gives an unfathomable amount of data. Internal representation of data gathers it through web. The significance of web scraping has happened to be a fast method for gathering information in kind of sectors. The platform NewsOne retrieves by collecting data from many official resources. The application is additionally utilized in the quest for writing and promoting. This article delivers the best in class in Web Scraping. It centers around the various methodologies and classes of the web scraping techniques and tools. It likewise features the uses of Web Scraping [7].

This system provides voice navigation and GPS tracking for visually handicapped people. In case of any emergency situations an alert message is sent to the caretakers of that particular person through a model named GSM. In general, visually impaired people face lot of difficulties in day to day life such as crossing streets, railway tracks etc. It is difficult for the caretakers of the person who is visually impaired to know about their current locations, by making use of this system it becomes easy for the caretakers of the visually impaired person to track live location to know where they are actually travelling. Voice recognition modules help the visually impaired to travel through the roads by themselves, even when they come across any obstacle in front of them this will immediately inform them to stop. This system helps in directing them in right direction and also sharing their live locations with their close ones. It helps them to travel in correct directions like right or left , down or up is also guided by the system immediately. This can be used by people of all age groups such as kids and elders by giving live location to their close ones so

as to ensure their security. Panic alerts to caretaker's is sent to the caretakers when the visually impaired or handicapped person passes through less secure areas. Live tracking is done by caretaker's that is when there is any emergency situation being faced by the visually impaired person locations are sent as SMS to the near ones so that they can come to their rescue on time. This system helps in making the visually impaired independent and allows them to travel to their destinations being independent [8].

The voice assistant plays a vital role in home healthcare. For the people who face difficulty in using phones or keyboards this system helps in providing communication verbally rather than typing their health issues, thereby providing solution for their difficulty to type on keyboard. The patients can know about their health conditions and also monitor health by making use of voice assistants. In this paper the voice assistant is used for providing information on patient's heart disease. The medical cloud collected data and then processes the data to find out the data which is more important and then stores in in the EHR of the patient. The aim of paper is to solve difficulties which the patient might face which may be caused while using wearable health sensors at their house. The obstacles used by the patient might include the readings of the sensors being misunderstood etc. The voice service provides various functionalities like checking heart condition of the patient, appointment setup and also provide reminder as though when is the next therapy. The voice application sends a request to the web service to send the data related to the heart condition of the patient, the web service then sends the requested data from the EHR of the patient which will be stored in the medical cloud. In case, there is any emergency condition in the report then the patient can send the file through the web service to the doctor concerned. The voice system is designed to set up an appointment with the doctor by direct communication. The web service acts as layer for communication with medical cloud. This paper mainly helps the patient to know about his heart condition by making use of communication [9].

In this paper we are using Machine learning to explain prediction of stock. Here Random Forest Regression algorithm is taken into consideration while predicting the stock markets. This algorithm is used to obtain market data stock at current time by making use of web scraping with stock data. Several individual trees for training are generated by training machine random forest. Prediction from all trees combines and gives final prediction; these are known as Ensemble techniques. Data collection: This is the first step in any project. In this paper, the data is collected by web scraping using BeautifulSoup. The data to be used for prediction is data combined from Nifty50 and the company whose stock price needs to be determined. Recursion: Here model predicts stock price. This includes price which is currently predicted combined with previously used data for predicting further stock price. Later, the price of predicted stock will be added to data set automatically. Hence, above process will continue to predict the price. For predicting future stock price RFCR algorithm was helpful. The training of this model is done in such a way that it can handle problem related to real time stock [10].

In the petroleum information retrieval system which is used by the petroleum exploration and development researchers, First step is to use an open source framework-Scrapy in order to create web crawler that crawls over the data that businesses work on. Second step the K-MEANS algorithm will be used in-order to cluster the crawled information. Hence, the useful information will be extracted and presented on the system. The results of this method shows that the proposed retrieval system is agile, efficient, improves accuracy and automation level of work. This information retrieval system developed using Scrapy reduces a ton of workload from the oil exploration and development personnel improves speed and accuracy of the retrieval information. Python is widely used as a design language for the development of web crawlers. Python has got unique advantages while developing a web crawler. Python crawler has become mainstream due to its ability of being rich open-source library and exceptional code encapsulation property. SCRAPY is an open-source web crawler built using the python language. This tool quickly reads and crawls websites to extract structured data from a web page. Web crawlers are a precise application or a script that mechanically extracts explicit content in a page as indicated by a rule for matching. The most common uses of web crawler is search engine, obtaining information from the Internet and downloading website pages. The purpose of the web crawler is to download information from the Internet onto a local backup. The crawler starts crawling from at least one given URL sets, it first acquires a URL and duplicates the content of that URL, at that point it extracts the other required URL in the page and places it into a set of queue, it rehashes this cycle until crawls all the website pages. Python offers a very rich standard library for network protocols, like the most basic crawler library, such as urllib, urllib2, and so on. Adding on to it, Python environment has rich third party toolkits, such as powerful Scrapy, Requests, BeautifulSoup, and various other network tool libraries.

Oil Information Retrieval based on K-means

- Setting up a dictionary of petroleum oil proficient phrasing, for example, petrel, rock type. Over the overview, the articles that the money managers centre around and clarify the keywords for each kind of article. Thusly, it frames a dictionary of oil proficient phrasing.
- Counting the occurrences of words showing up in each article. Thusly, organized data is been obtained which is in the form of a sparse matrix. The attributes of the dataset is words from dictionary and row data is quantity of words that show up in the article.
- K-means algorithm is used for clustering of these documents. Clustering is used in the application of business area it includes: gathering various reports, music or films as per various subjects, or discovering clients with comparable interests dependent on regular buy conduct, and building suggestion engine. Pick the typical clustering K-MEANS algorithm.
- Extracting the meaning of the article in-order to shape a briefing. Some of the few extraction methodologies are:

First one is extracting the main contents from abstract. Second one is conclusion parts other is extracting contents from highly recurrence part of the keywords. For articles which has summary and conclusion, first method is chosed. Else utilize the subsequent one. The derived data is gathered to understand classification of data.

- After clustering, data is divided into several large plates, each of the plate describes summary of each article.

Text clustering is a vital method used for text mining, it partitions text into a few important bunches depending on specific measures. The similarities between the writings in a similar bunch must be higher than that of the text likeness between various other clusters, in-order to accomplish powerful gathering of text data association and the management. Text classification method has grown so quickly, and many algorithms of clustering have been produced. Clustering algorithms can be generally divided into two methods: 1) Hierarchical method and 2) Partition method. K-MEANS algorithm chooses K initial cluster centres, it emphasizes iteratively until the target function is reached. Here to represent the cluster the centroid is considered. This paper speaks about a framework for information retrieval system in oil exploration and development. Firstly, it uses scrapy to build a web based crawler. Secondly, it uses the K-MEANS algorithm for clustering the crawled documents. Finally, it extracts the key information and forms a briefing based on clustering results. The results of operation displays that the proposed retrieval system is accurate, nimble and convenient, it improves the operational efficiency and accuracy of the personnel. This framework has got a few deficiencies in the system but depends more on proficient experience and has got places to improve, for example maintaining the dictionary and count of the word frequency [11].

In this system, the driver interacts with assistant system which is very necessary for safety. Several functionalities of ADAS have various driver alerts that are through Audio, visual, and vibrational means which gives safety alerts to the drivers. The action of the driver for the given alert depends on his moods; his mood decides whether he understands the alert. The voice alert is effective human alert but its static nature makes it less effective for the driver to understand alerts when he has a low level of concentration. Here adaptive voice alert system is developed according to the drivers emotional characteristics. This system will change or adjust the voice alert according to the drivers mood. Voice alert is an important component of ADAS functionalities. Voice alert gives a warning to the driver about when his concentration is lost or when there is some event which is not recognized which has to be brought to the notice of the driver. This alert includes both present and predicted events. The functionality of the alert is to inform the driver about the speed which needs to be maintained in various situations and also inform him about braking whenever someone crosses the road immediately. Depending on the mood of the driver the voice alert is generated, and the information content in alert is not changed, that is if the driver is happy means that he will follow the voice alerts but when he is angry the voice alerts are elaborated to make him understand, therefore to provide automotive companies to achieve zero accidents result by using autonomous cars [12].

Since the number of documents in offline and online are increasing daily the task of searching for a particular document among these has become a difficult task. In regular expression rule based approach can be used to find the information needed from documents which may be present in any format. Python is a programming language which consists of regular expressions by means of libraries. Regular expressions can be denoted by using re. The regular expression can be used to search for a particular character in string, breaking of given string into substrings, and also for replacing particular part of string. The search pattern matching algorithm takes set of m characters that has to be searched as input. The search pattern matching algorithm traverses from left to right while trying to match each character with characters in the string which we need to compare and returns successful if the patterns match otherwise it returns unsuccessful search, if the search is unsuccessful and the text has not ended then the pattern is adjusted to right by one position and the traversal starts again from left to right. Regex are used in compilers, editors, matching of patterns etc. Information extraction is regarding the position in which the item we require is present. The re package is used to determine whether the regular expression is similar and is matching with the entire string in python language [13].

Lately, with the expanding improvement of Artificial Intelligence, Big Data and Cloud processing, and so on, the data on the Internet has been blasting, so how to get target data productively and rapidly has become an earnest issue to be solved. This article centres around the data collection and acquisition issue of shipping with respect to the job hunting information under network environment. In this paper, two sorts of data collection techniques for shipping job hunting dependent on web crawler have been proposed. In view of the Python standard libraries and Scrapy crawler framework, comparing web crawler program is planned and executed to scrape the target data from target site and eventually store the gathered data into a local file. Through the measure of information crawled and tedious comparative analysis, the outcome exhibits that the information collection strategy based on the Scrapy crawler system is easy to work, effectively extensible, featuring being targeted, with high effectiveness and fast speed in gathering shipping job hunting data. The gathered information cannot just assist researchers for conducting data mining analysis, but also provides data support for following up on the shipping job hunting database. Web crawler looks for the site page through URL (Uniform resource locator), and returns the concerned

information to clients straightforwardly. Hence, clients don't have to get to data by browsing web page, which can save time and energy and improve the exactness of information collected too. The primary objective in scraping is to extract organized data from unstructured or semi-organized pages for further data mining and analysis. Notwithstanding the data administration sites of organizations related with shipping, there are some specific shipping service sites that provide recruitment hunting data. In this paper, the www.hy163.com site is taken as target site, which is one of China's leading site in team job hunting and enrolment, gaining practical experience in group enlistment, work chasing, enrolment, preparing, assessment, oceanic news and other coordinated team data. During the crawling process first he Scrapy engine sends a request, second the scheduler sends initial URL to the downloader. Next, the downloader sends requests to server and the response downloads contents from the web, and afterward present the content to web spider. From that point the spider will parse the page content. Spider will analyse the results in two ways one is to get new URL, and later the Scrapy engine sends request to the scheduler once again, begins another round of crawling, and continuously repeats the process, the other is to get the necessary information, which will be sent to the project pipeline for additional handling. Project pipeline is answerable for post-handling, for example, data cleaning, confirmation, sifting, deduplication and storage capacity, lastly the scrapped data will be regularly composed into a nearby document or put away into a database [14].

The paper comprises of data where in the web spiders or web scrapers are software programs that can be utilized to bring out data that is accessible online. The process where data is been explored and collected from this huge span of data source this is web harvesting and the software used to perform this task is web spider. Majority of the data on web today is in HTML documents which are seen by people on the browser. This paper intends to extract online product's data for inflation and CPI study. A strategy to scrap this information and analyse price of vegetables and fruits is proposed in this paper with a goal to estimate momentary expansion. Numerous devices were utilized during the usage, for example, Selector gadget which encourages the system to catch important information. There are several modules which specifies the functionalities for better understanding the methodology and briefly summarises the paper. Web Scrapper module gathers the data particularly requested by user or which belongs to a specific sector. Web crawling methodologies incorporate live and documented with the end goal of information extraction. During the entire process a part of the issues were that lone the depictions of the page were available in the web recorded data. This process includes powerful site data extraction. This application is intended to get some specified domain-based data so that the user can get the whole text content which is found in the web by the technique of crawling. Some of the other challenges anticipated are incomplete site previews in stored data where storing the previous pages were limited. This is complexity based data extraction. Uses of Artificial Intelligence could be adopted and actualized as proposed future work [15].

The regular expression helps the developers to use and hence developers tend to make errors. Automatic Checking of Regular Expression takes a regular expression and makes 11 checks on the regular expression. The validations are built on common mistakes. It checks for wrong usage of character sets [], wildcards (.), and line anchors (^, $).In this user has to enter a RE and tick the check button, any harms are stated to the user with wrong portions of the RE is highlighted. Error in Regular expression can take the string which has to be excluded hence the program gets crashed resulting an incorrect string. ACRE has a set of 11 checkers. The checker concentrate on mistake generated while developing a RE. Once fault occurs the part of RE with fault is highlighted. Mostly some characters, punctuation marks should be excluded. The paper here tells about the automatically checks regular expressions tool. There are total of 826 regular expressions, out of which automatic checking of regular expression has found 283 buggies of RE and 46 has wrong alarms. In upcoming time more checkers can be added to automatic checking of regular expression. [16].

The initial technological world was a world of fewer requirements of technology yet with some purpose. There was a need to upgrade the technological world with new and advanced technology as all the data analysing processes were not handled and performed smoothly due to lack of necessary requirements. By a small research on the improvements about the technology and the updates in the developing technological world, this paper reviews about the clear ideas that were used for understanding the process of web scraping and the developmental ideas along with its applications. The web data extraction of fewer amount of raw data was performed easily and there were no drawbacks as such, so that the actions were left incomplete. Then the advancements in technology emerged and it changed the present world drastically. Effectuating the construction of data extraction wherein the data retrieval and other vital operations were carried out with rapid processes. It became easy to analyse and understand the functioning of a web scraper wherein massive data selection and extraction were involved. The new applications applied to make a simpler way to access data also was a beneficial method to make the process more reliable and efficient. Certain automated tools which were used in order to have a better platform and smooth process to carry out the data extraction procedures. While extracting data the earlier technologies did not have a well-developed process so as to satisfy user requirements. The factors of legacy, compatibility and performance have to be considered while mastering the web data extraction processes with the aid of scraping tools. To have security during the process certain guidelines and privacy features are applied while accessing proprietary contents. Data

harvesting involves the simple process by selecting the real time datasets initially and then retrieval of specific data that satisfies users requirements from official and relevant websites and finally the data is sent for testing and if it matches output is produced accordingly. The numerous tools and advanced developments which were utilised for the betterment of web data extraction process are spotlights of this paper. The web data retrieval has made it possible to perform the action through massive data sources with the involvement of real-world entities implemented in some applications. All these facilities have made the web scraping process simpler and easier. Besides being a technical process, it has overcome all the drawbacks which were observed in the initial stages and has shown a great improvement by the significant role of human resources. The change of action might have taken a considerable time but the drastic change in technology was observed from the initial stages itself. Implementations of these ideas were practiced and experimented in e-learning applications which functions in all platforms irrespective of the domains [17].

The paper gives actual time visualization of speech information for teaching students studying music therapy. This gives the quality of singing information and information on vocal tract shape. This provides discussion on this tool for teaching students to understand the music therapy. Here authors have developed a package of actual sound, speech and time examination for students to discover speech science fundamentals; these are called "production of speech and auditory view study kernel of the next generation". There are two speech specific tool, they are speech production simulators. Authors have found that speech tools are beneficial for awareness of student's body voice relations. The actual time compatibility is essential for creating visual feedback of rich speech information's. SPEECH PRODUCTION TOOLS: It is necessary to imagine the resulting picture that gives reason for communication. A set of UI is used for this purpose. This part gives two tools to assist learners get the difficult body speech relation. This paper provides actual time response of speech information to help students understand about music therapy to get essential singing skill. Here authors are concerned about song treatment as a support to orange technology concentrating on supporting therapists teaching by giving essential actual time, communicating tools [18].

In this paper the given pattern in datasets are extracted. Gene expression dataset is considered in this case. Depending on the abnormal pattern it will help in finding out abnormality of person. In this case we use MATLAB software for experimentation. Here we examine the gene expression data for abnormal pattern. We use RE established pattern matching technique for searching abnormal patterns. Regular expression: Pattern matching is searching for a substring in a given string. A RE is called regex or regexp. The steps for finding a pattern to identify the corresponding string matched with a regex's pattern are called pattern matching. The advantage of RE is, it deletes the entire pattern from any file. Regular expression is mainly used in creating different third party libraries. In this paper the irregular pattern is provided as an input to regular expression and regular expression hunts for the availability of input pattern in dataset. The data set is pre-processed and regularized in order to have growth in effectiveness of identifying and matching of pattern. During this experiment empty gene data, (Ramesh Mand, 2018) false gene data and other useless data are removed. As a result it is clear that this approach is well suited in gene expression data for pattern matching. [19].

Earlier days even though there were some limited facilities which would help the people staying nearby with a private health care centre, it was not possible to function ideally. Competing with today's world and the developing technology all the functions and processes have been digitalised. There are many smartphone applications existing today as a means of healthcare assistance. Most of the industries have been developed with the advanced technological world and have undergone major transformations. The number of people accessing the healthcare applications are increasing day by day. Since the world has become a digitalised platform, right from basic health treatment to medical requirements people rely on smartphone healthcare applications. However, during the development and invention of these applications the initiators faced several challenges, and also have easily dealt with them. This research paper describes about diabetes, a health condition which was used to predict by the aid of a smartphone application. It explains the various methods utilised for the prediction of diabetes based on the symptoms and some specific conditions. The development of any particular application requires a good amount of knowledge about how to design the front end so that its compatible on any platform and also the cloud to store the back-end data for analysing and future retrieval purposes. With the provided user data and their conditions, it can be easily predicted. A particular programming language also has to be used to implement the functionalities of the application. The paper deals with the various methods that were implemented and metrics used for design of an ideal application so as to give better performance, satisfy user requirements, good quality and a liable platform which can be accessed by anyone. The application developed must be of greater use to the general public and needs of people must be considered rather than the needs of the community only then the application will be successfully launched and can be utilised on a longer run [20].

3. Proposed Work

In this approach, by utilizing the Worldometer COVID-19 dataset to fetch the data on Covid-19 updates and also for optimizing the search results. The use ParseHub, a tool that is free of cost for web scraping and a GUI for manually selecting HTML tags from the web page. Scraping the user requested data which is either in the form of voice or text by making use of regular expression by pattern matching.

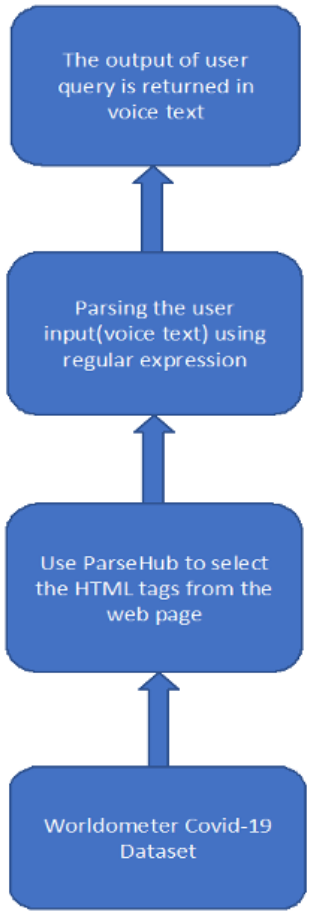


Figure 1: The Workflow of Proposed System

4. Conclusion

Use of Python programming also emphasizes on understanding the application of pattern matching and regular expressions for web scraping. The Worldometer Coronavirus dataset can be gathered from true reports, straightforwardly from Government's correspondence channels via local media sources when considered as reliable. A group of experts and analysts who approve information from a consistently developing list of more than 5,000 sources may be feasible for this site to accumulate data in an efficient manner. The input given by the user is analysed and scraped from the website and output is produced using Google voice assistant in the form of speech. The output could be produced either in the form of text or speech. This approach is simple and straight forward to extract the number of cases in a particular state or country, the number of deaths in a particular country or state and the number of recovered patients all over the globe, country or a state.

REFERENCES

Thivaharan. S, Srivatsun. G and Sarathambekai. S, "A Survey on Python Libraries Used for Social Media Content Scraping", Proceedings of the International Conference on Smart Electronics and Communication (ICOSEC 2020) IEEE Xplore Part Number: CFP20V90-ART; ISBN: 978-1-7281-5461-9, PP: 361-366.

David Mathew Thomas and Sandeep Mathur, "Data Analysis by Web Scraping using Python", Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology [ICECA 2019] IEEE Conference Record # 45616; IEEE Xplore ISBN: 978-1-7281-0167-5, PP: 450-454.

Peipei wang, Gina R. Bai and Kathryn T. Stolee, "Exploring Regular Expression Evolution", SANER 2019, Hangzhou, China Research Papers IEEE, PP: 502-513.

Vidhi Singrodia, Anirban Mitra and Subrata Paul, "A Review on Web Scrapping and its Applications", 2019 International Conference on Computer Communication and Informatics (ICCCI -2019), Jan. 23 – 25, 2019, Coimbatore, INDIA.

Rafael Dantas, Dr Chris Exton and Dr Andrew Le Gear, "A Communications using a speech-to-text-to-speech pipeline", 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob),

Fatmasari, Yesi Novaria Kunang and Susan Dian Purnamasari, "Web Scraping Techniques to Collect Weather Data in South Sumatera", International Conference on Electrical Engineering and Computer Science (ICECOS) 2018, PP: 386-390.

Rabiyatou DIOUF, Edouard Ngor SARR, Ousmane SALL, Babiga Birregah, Mamadou BOUSSO and Sény Ndiaye MBAYE, "Web Scraping: State-of-the-Art and Areas of Application", 2019 IEEE International Conference on Big Data (Big Data), PP: 6040-6042.

Priscilla Ebenezer.R, Vishnu priya.M and Nivetha.B, "GPS Navigation with Voice Assistance and Live Tracking for Visually Impaired Travelers", IEEE 6th International Conference on smart structures and systems ICSSS 2019.

Dimitri Dojchinovski, Andrej Ilievski and Marjan Gusev, "Interactive home healthcare system with integrated voice assistant", MIPRO 2019, May 20-24, 2019, Opatija Croatia, PP: 284-288.

Bikrant Bikram P. Maurya, Ayush Ray, Aman Upadhyay, Dr. Bhupesh Gour and Dr. Asif Ullah Khan, "Recursive Stock Price Prediction With Machine Learning And Web Scrapping For Specified Time Period", 2019 IEEE.

Yili Ren and Yiting Ren, "A Framework of Petroleum Information Retrieval System Based On Web Scraping With Python", 2018 IEEE.

Sarala S M , Sharath Yadav DH and Asadullah Ansari, "Emotionally Adaptive Driver Voice Alert System for Advanced Driver Assistance System (ADAS) Applications", International Conference on Smart Systems and Inventive Technology (ICSSIT 2018) IEEE Xplore Part Number: CFP18P17-ART; ISBN:978-1-5386-5873-4, PP: 509-512.

Ramesh Mande, Kalyan Chakravarti Yelavarti and JayaLakshmi G, "Regular Expression Rule-Based Algorithm for Multiple Documents Key Information Extraction", International Conference on Smart Systems and Inventive Technology (ICSSIT 2018) IEEE Xplore Part Number: CFP18P17-ART; ISBN:978-1-5386-5873-4, PP:262-265.

Dongcheng Peng, Tieshan Li, and Yang Wang, "Research on Information Collection Method of Shipping Job Hunting Based on Web Crawler". 8th

International Conference on Information Science and Technology June 30-July 6, 2018; Granada, Cordoba, and Seville, Spain, PP:57-62.

Ranjani Murali, "An intelligent web spider for online e-commerce data extraction", Second International Conference on Green Computing and Internet of Things (ICGCIoT), 2018 IEEE, PP:332-339.

Eric Larson, "Automatic Checking of Regular Expressions", 2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation, PP: 225-234.

Shreya Upadhyay, Vishal Pant, Shivansh Bhasin and Mahantesh K Pattanshetti "Articulating the Construction of a Web Scraper for Massive Data Extraction", 2017 IEEE.

Eri Haneishi and Kaori Hagiwara, "Real Time Feedback Of Singing Voice Information For Assisting Students Learning Music Therapy".

L.Sharmila, U.Sakthi, A.Geethanjali and Suresh Sagadevan, " Regular Expression Based Pattern Matching for Gene Expression Data to Identify the Abnormality Gnome", 2017 Second International Conference on Recent Trends and Challenges in Computational Models, PP: 301-305.

Amruta Kulkarni, Deepa Kalburgi and Poonam Ghuli, "Design of Predictive Model for Healthcare Assistance Using Voice Recognition", 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017, PP: 61-64.