International Journal of Advances in
Soft Computing and Intelligent Systems (IJASCIS)
2022, Vol 01, Issue 02, 30-40

Science Transactions © 2022

Original Paper

DEVELOPMENT AND TESTING OF AUTOMATION TOOLS USING PYTHON

Dhairya Dutt[1], Akshat Jain[2], Digvijaysinh Chudasama[3], Jainil Raval[4], Keyuri Kariya[5], Mann Mehta[6], Sneh Patel[7], Sachin Gajjar[8]

1Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec031@nirmauni.ac.in

2Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec007@nirmauni.ac.in

3Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec029@nirmauni.ac.in

4Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec108@nirmauni.ac.in

5Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec051@nirmauni.ac.in

6Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec070@nirmauni.ac.in

7Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
19bec099@nirmauni.ac.in

8Department of Electronics and Communication Engineering, Nirma University, Ahmedabad, Gujarat, India.
sachin.gajjar@nirmauni.ac.in

Abstract

In this paper, algorithms are developed and tested to automate tasks like web scraping, extraction of stock price data and its analysis and Bulk Email generation. Web scraping is a special method used for the extraction of data from the World Wide Web (WWW) and saving it to a file system or database for various types of data analysis. It is sometimes referred to as web extraction or web harvesting. The Hypertext Transfer Protocol (HTTP) is frequently used to scrape web data. Web scraping is widely recognized as an effective and powerful technique for gathering significant volumes of data due to the massive amount of heterogeneous data continuously generated on the WWW. In this paper algorithms are designed for storing and searching details of a product in the Flipkart website, predicting stock market prices, and sending automated mails to a number of recipients. Algorithms are coded in Python using libraries like Tkinter for GUI purpose, beautifulSoup for extracting data from HTML and XML format, urlLib for URL handling and TimeSeries for timestamps and tracing open, high, low, close, volume of the equity for stock market pricing and analysis.

Keywords
Software tools for Automation, database, World Wide Web, Python, Web scraping, Email generator

1. Introduction

Python is an interpreted, object-oriented, high-level, dynamically semantic artificial language [1]. It is particularly desirable for Rapid Application Development and to be used as a scripting or glue language to tie existing components together. This is because it has high-level built-in data structures, and supports dynamic typing and dynamic binding. Python's support for modules and packages allows the program to be modular and code can be reused.

The extensive standard library and the python interpreter are freely distributable and available in source or binary form for a number of processors. With python several computational tasks can be automated without human intervention [1].

A number of tasks that were formerly done manually have now become more productive, dependable, and/or quick with the help of computer automation. Compared to other programming languages, python offers a pleasant, accessible syntax that's simple to read and grasp. Additionally, because it's open-source, an unlimited selection of tools, libraries, frameworks, and support is readily available [1]. Because of this python is one of the most preferable languages for tasks related to computer automation. The rest of the paper is organised as follows. Section 2 presents the literature survey. Tasks that are automated are discussed in Section 3 and Algorithms developed for the Automation are discussed in Section 4. Finally the paper is concluded in Section 5.

2. Literature Review

2.1. Python and the Other Languages

Python is an open-source programming language that supports a variety of automation related frameworks for unit, end-to-end, and integration testing. Majority of the other languages like JavaScript strongly focus on front end development and testing of automation framework [2].

The Python frameworks PyTest, Robot, Nose 2, Behave, Lettuce, and Testify are used by more than 2 million websites worldwide [3]. Microsoft offers automated testing tools like MSTest, NUnit, and xUnit.Net [4]. They enable a large number of automated testers to do cross-border, unit, and functional testing in order to strictly ensure consistency and extensibility throughout code execution.

While the frameworks of the above languages are slow, Python's framework is fast. Hence Pythons with its library support is best suited for developing Automation tools.

2.2. Library Support for Python

- For Web scraping, Python includes libraries that can assist in integrating major protocols like HTTPs, FTP, and SSL. These libraries enable quick and accurate data extraction in the form of news from many sources [5].
- For Machine Learning Scikit-learn (Sklearn) and other libraries are available in python [6].
- For Data Science libraries available are Pandas and NumPy [7].
- For Game development PySoy, a 3D gaming engine is available in Python 3. Further, PyGame, a game development toolkit is also available [8].

- For Audiovisual applications, TimPlayer and Clay are the libraries available in Python [9]-[11].
- For GUI and User Interface, Tkinter library is available. Some more toolkits include wxWidgets, Kivy, and PYQT [12].

3. Tasks Automated In The Work

3.1. Web Scraping

Web Scraping is an automated technique to retrieve copious volumes of data from websites [5]. Web scraping assists in gathering and storing unstructured data found on the Internet. The major steps in the web scraping process are:

- Locating URL
- Page Inspecting
- Locating the Data to be Extracted
- Running the algorithm to retrieve the data
- Storing data in the necessary format

3.2 Extraction of Stock Price Data

A Python library named Pandas is used for the data frame, Matplotlib for plotting the graph and Alpha vantage is used to get the stock prices from the API key. API key is used to get the stock prices of Microsoft which has its ticker key as MSFT (Microsoft corporation). The dataset is printed at specific time intervals. Next, the dataset of the prices are converted into the pandas data frame. Using pandas data frame the stock prices data are converted into the csv file format and stored in excel file [14].

3.3 Bulk Email Generator

An algorithm that sends personalised email to several recipients using the same template is developed. It can be used to send a single email or email in bulk, it shows the status of the email whether the email has been sent to all the users or not, password based login.

4. Automation Tools Used In Python

4.1 Automation Tool for Web Scraping

The GUI developed has the facility for searching and storing the details of a product from a given website. Figure 1 shows the interface where the user can enter the name of the excel file and the link of a website for web scraping. Figure. 2 shows the interface where a user can search for a particular product present in the excel file. The search results are shown in Figure 3. This task is made to extract the product name, price of the product, and its specifications.

Python library urlLib is used to fetch URLs (Uniform Resource Locator). It can retrieve URLs with the urlopen function. Next, BeautifulSoup library is used that can extract data from XML and HTML files. It can carry out the task of traversing, searching, and modifying the parse tree for the preferred parser. Pandas library is used for data analysis and storing data in CSV format.

Figure 1. Interface for entering csv file

Figure 2. Interface for searching a specific

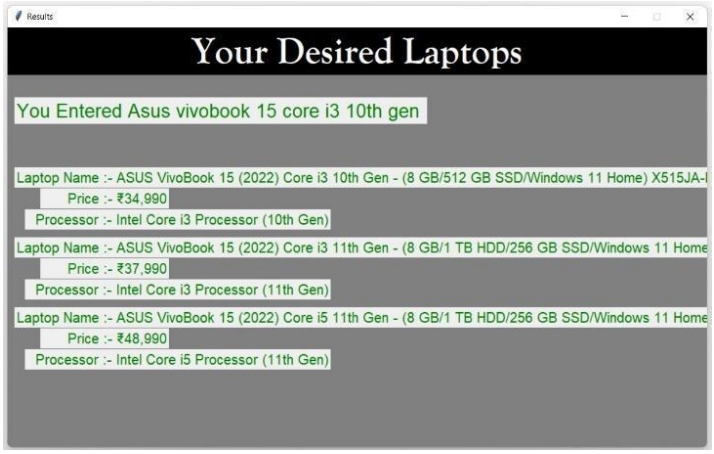


Figure 3. Interface of the filtered results

The automation technique is shown in the form of flow chart in Figure 4. The algorithm first fetches data from the website, for the current product details in the example of Flipkart website. This is done by the urlLib library which uses GET, POST, and PUT methods. It uses the data in JSON format. It has access to all of the data on the web page. Next, the

beautifulSoup library extracts the product name, its price, its specification and stores it in variables of python. Next, with Pandas library, the filtered data viz. product name, price, specifications are stored in a
CSV file. For searching for a product, users can search for one which is present in the CSV file. For that user has to enter the name of the product, name of the CSV file. With Natural Language Processing [13] technique, the name of the product gets auto-filled when it passes the threshold of matching of characters of 40 %.

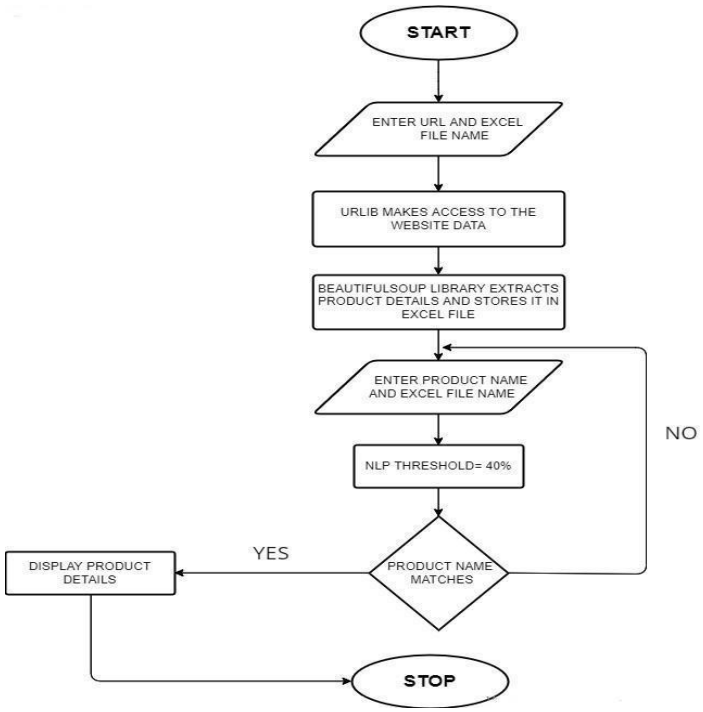


Figure 4. Flow chart of the task Web Scraping Using Python Applied on Flipkart

4.2 Automation Tool to extract Stock Price Data

Web scraping can be used to get stock prices. This is done using Python libraries such as alpha-vantage, matplotlib, pandas. These libraries are used to get the stock data and store into CSV file format for further processing. The prices of the stock are visualised through graphs with details like opening, closing, highest, and lowest of the stock price. Further stock prices can be predicted. If the change in stock price goes above the threshold value tool can give an alert and necessary actions can be taken on a basis of the alert [14].

Flow chart of the tool developed is shown in Figure 5. The software uses pandas python libraries for the data frame, matplotlib for plotting the graph and alpha vantage to get the stock prices from the API key generated from the alpha vantage. In the next step API key is used to get the stock prices of Microsoft which has its ticker key as MSFT. As shown in Figure 6. Stock information viz opening, closing, low, volume of stock is shown with the timestamps. This information is converted into the panda's data frame and further to CSV format and stored in excel file.

Next as shown in Figure 7. The closing price of the Microsoft stock is plotted by using the

matplotlib library. Then the percentage change of the stocks is also calculated from the closing price and if the percentage change is greater than the threshold the alert message is popped by the tool to alert the user.

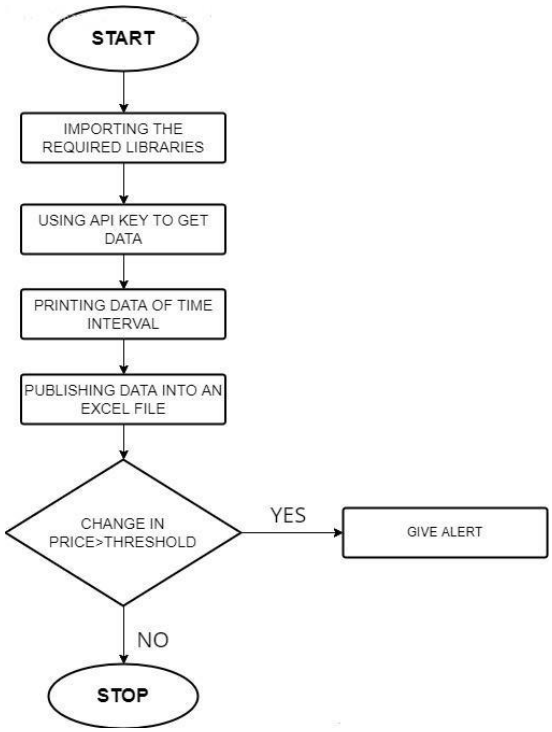
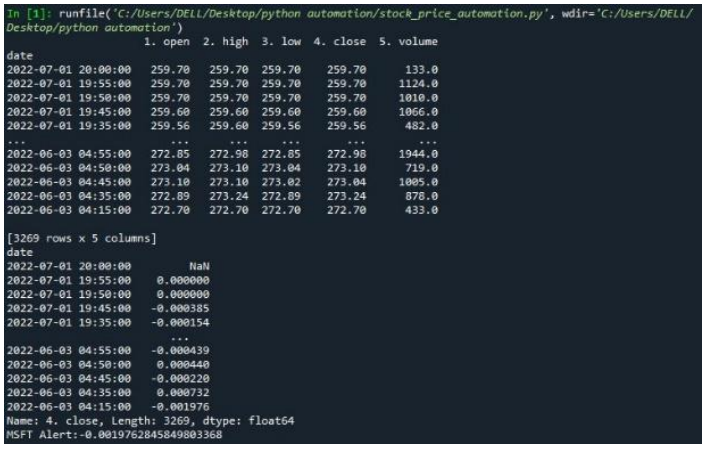


Figure 5. Flow chart for Extraction of Stock Price Data



```
In [1]: runfile('C:/Users/DELL/Desktop/python automation/stock_price_automation.py', wdir='C:/Users/DELL/
Desktop/python automation')
                     1. open  2. high  3. low  4. close  5. volume
date
2022-07-01 20:00:00   259.70   259.70  259.70    259.70      133.0
2022-07-01 19:55:00   259.70   259.70  259.70    259.70     1124.0
2022-07-01 19:50:00   259.70   259.70  259.70    259.70     1010.0
2022-07-01 19:45:00   259.60   259.60  259.60    259.60     1066.0
2022-07-01 19:35:00   259.56   259.60  259.56    259.56      482.0
...                      ...      ...     ...       ...        ...
2022-06-03 04:55:00   272.85   272.98  272.85    272.98     1944.0
2022-06-03 04:50:00   273.04   273.10  273.04    273.10      719.0
2022-06-03 04:45:00   273.10   273.10  273.02    273.04     1005.0
2022-06-03 04:35:00   272.89   273.24  272.89    273.24      876.0
2022-06-03 04:15:00   272.70   272.70  272.70    272.70      433.0

[3269 rows x 5 columns]
date
2022-07-01 20:00:00         NaN
2022-07-01 19:55:00    0.000000
2022-07-01 19:50:00    0.000000
2022-07-01 19:45:00   -0.000385
2022-07-01 19:35:00   -0.000154
                          ...
2022-06-03 04:55:00   -0.000439
2022-06-03 04:50:00    0.000440
2022-06-03 04:45:00   -0.000220
2022-06-03 04:35:00    0.000732
2022-06-03 04:15:00   -0.001976
Name: 4. close, Length: 3269, dtype: float64
MSFT Alert:-0.0019762845849803368
```

Fig. 6. Output that shows price prediction

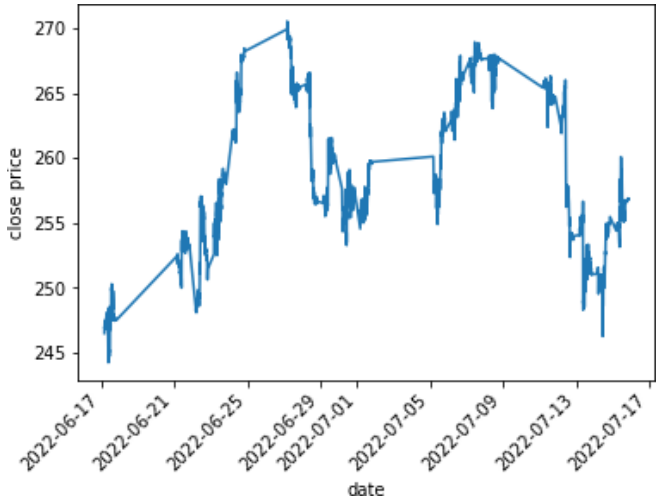


Figure 7. Graph between Close Price vs Date

4.3 Automation Tool to send Bulk Email

This tool can be used to send customised bulk emails to specific recipients using secure Simple Mail Transfer Protocol (SMTP). The addressing and person specific content in the email will be personalized while the remaining of the content will be common for all the emails. The list of recipients can be specified using the excel file. Figure 9 and 10 shows the interface for the tool. Figure 11, shows the email sent by the tool.

For the Tool for Bulk Email generation libraries like, smtplib, pandas, ImageTk, Tkinter and SMTP are used. SMTP library is used for setting up the format of Email content, encrypt it, and transfer email messages between the mail servers. IMAP library is used to retrieve emails that have been sent [15]. Using these libraries the emails can be sent and received via any electronic device that has Internet connectivity and SMTP or ESMTP daemons running in it.

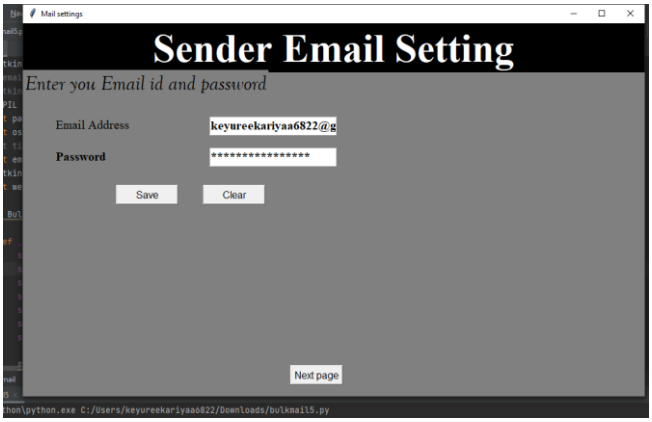


Figure 9. Home interface of the Bulk Email Sender

Figure 10. Interface of Bulk Email generator

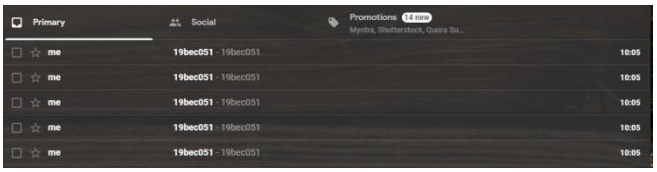


Figure 11. Output that shows receiver's interface

5. Conclusion

This paper discusses the development and testing of Automation tools for web scraping, stock price extraction from websites and bulk email generation. Python and the supported libraries are used to develop the tools. The web scraping tool is successfully tested on the Flipkart website, stock price extraction tool is tested on Microsoft shares and Bulk email generator is tested for sending program invitations to students.

Acknowledgements

The authors would like to thank the department of Electronics and Communication for providing such great resources and their constant support from start to end. Also, the authors would like to show our gratitude towards Dipesh Panchal and Chandni Patel - our mentors at "eInfochips," for sharing their corporate experience and guiding us throughout the process.

References

[1] Srinath, K. R. "Python–the fastest growing programming language." In- ternational Research Journal of Engineering and Technology 4.12 (2017): 354-357.

[2] Decan, Alexandre, Tom Mens, and Maelick Claes. "On the topology of package dependency networks: A comparison of three programming language ecosystems." Proceedings of the 10th European Conference on Software Architecture Workshops. 2016.

[3] Haben, Guillaume, et al. "A replication study on the usability of code vocabulary in predicting flaky tests." 2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR). IEEE, 2021.

[4] Chaczko, Zenon, et al. "Design of unit testing using xUnit. net." 2014 Information Technology Based Higher Education and Training (ITHET). IEEE, 2014.

[5] Ghazvinian, Holbert. ”Viswanathan.“.” Simple Web Scraping.” Inter- net: https://seanholbert. wordpress. com/2011/07/15/scrappy-simple-web- scraping (2015).

[6] Sodhi, Pinky, Naman Awasthi, and Vishal Sharma. ”Introduction to machine learning and its basic application in python.” Proceedings of 10th International Conference on Digital Strategies for Organizational Success. 2019.

[7] Kadiyala, Akhil, and Ashok Kumar. ”Applications of python to evaluate environmental data science problems.” Environmental Progress Sustain- able Energy 36.6 (2017): 1580-1586.

[8] McGugan, Will. Beginning game development with Python and Pygame: from novice to professional. Apress, 2007.

[9] Giannakopoulos, Theodoros. ”pyaudioanalysis: An open-source python library for audio signal analysis.” PloS one 10.12 (2015): e0144610.

[10] Machado, Felipe, Norberto Malpica, and Susana Borromeo. ”Parametric CAD modeling for open source scientific hardware: Comparing Open- SCAD and FreeCAD Python scripts.” Plos one 14.12 (2019): e0225795.

[11] Shmueli, Galit, et al. Data mining for business analytics: concepts, techniques and applications in Python. John Wiley Sons, 2019.

[12] Moriasi, Daniel N., et al. ”SWAT-LUT: A desktop graphical user inter- face for updating land use in SWAT.” JAWRA Journal of the American Water Resources Association 55.5 (2019): 1102-1115.

[13] Selvapandian, D., R. Udaya Nirmala Mary, and C. Karthikeyan. ”Arti- ficial intelligence in online shopping using natural language processing (NLP).” Journal of Critical Reviews 7.4 (2019): 2020.

[14] Jansen, S. (2020). Machine Learning for Algorithmic Trading: Predictive models to extract signals from the market and alternative data for systematic trading strategies with Python. Packt Publishing Ltd.

[15] Mohammed, S., Mohammed, O., Fiaidhi, J., Fong, S., Kim, T. H. (2013). Classifying unsolicited bulk email (UBE) using python machine learning techniques. International Journal of Hybrid Information Tech- nology, 6(1), 43-56

Authors

Mr Dhairya Dutt

Mr Dhairya Dutt, currently in final year pursuing btech degree in electronics and communication. He has had 2 internships so far, one at 'einfochips' (2 months) and another at 'techtable'(4 months). His interests includes cs fundamentals and solving real life problems

Mr Akshat Jain

Mr Akshat Jain is currently in final year pursuing B.Tech degree from Nirma University. He has a major in Electronics and communication engineering and a minor in Computer and Science. He worked as an SDE intern at Einfochips, an Arrow-based company. His area of interest includes software development, machine learning and Data scientist and how to implement it to solve real world life problems. He has presented/published two papers in international conferences in the area of communication and won one National level Matlab Hackathon.


Mr Digvijaysinh Chudasama

Mr Digvijaysinh Chudasama, a final-year Bachelor of Technology (B.Tech) student at the Institute of Technology, Nirma University, Ahmedabad. His major is Electronics and Communication Engineering and minor in Computer Science Engineering. He worked as an SDE Intern at eInfochips, an Arrow-based company,and worked on various automation tasks using Python. He is an active student who appreciates taking on problems and coming up with creative solutions. He is keen to work on new applications and technologies in order to best assist the organisation and contribute to a technologically evolved world. He wishes to make the best use of my engineering expertise and knowledge.


Mr Jainil Raval

Mr Jainil Raval is in my final year pursuing my B.Tech degree in Electronics and communication Engineering from Nirma university. He has completed a two month internship from EInfochips. His area of interest includes machine learning and web development and to implement it to solve real life problems.


Miss Keyuri Kariya

Miss Keyuri Kariya from Nirma University undergraduate student majoring in Electronics and Communication and minoring in Computer Science and Engineering. She had completed a two-month internship with Einfochips on the topic of automation using Python. She is a passionate web developer who wants to improve their skills and build a solid career in computer science and problem solving. Her interests and hobbies include website development and writing stories.


Mr Mann Mehta

Mr Mann Mehta is currently pursuing bachelors in technology from Nirma university, Ahmedabad. He had done a two month internship at Einfochips, an arrow based company. He is a passionate philomath and a keen imaginator and wants to boost his skills and build a strong career in computer science and solving real life problems. His hobbies and interests include website development, writing and sports.


Mr Sneh Patel

Mr Sneh Patel, currently in final year pursuing btech degree in electronics and communication. He has had 2 internships so far, one at 'einfochips' (2 months). He is interested in Artificial intelligence and Machine learning.

Dr Sachin Gajjar

Dr Sachin Gajjar is working as an Associate Professor in Electronics and Communication Engineering Department. He has an experience of more than 20 years in the field of teaching, research and industry. He did his B E in Electronics and Communication Engineering in 1999, MTech in Computer Technology in 2007 from the Indian Institute of Technology (IIT), Delhi and PhD from Nirma University in 2016. He joined Nirma University in the year 2002. He was Lecturer in Computer Engineering at L D College of Engineering, Ahmedabad during the year 2000 to 2002. He has also worked with Ferromatik Milacron, Ahmedabad as Software Engineer in the year 2000 and with Videocon International Limited, Gandhinagar in the year 1999. He has published more than ten research papers in the area of Wireless Sensor Networks, Networked Embedded Systems, and Computer Networks in international refereed journals and presented/published more than twenty five papers in international conferences/proceedings with h-index of 11 and i-10 index of 13. He has successfully executed three research projects funded by Nirma University, GUJCOST and IEEE-Standards Association in the area of Wireless Sensors and Actuator Networks. He is currently executing a project funded under ISRO-Respond Programme. He has guided more than ten PG dissertations. He is currently guiding one PhD student. He is actively involved in consultancy, research, and testing in the area of Embedded Systems. His research interests include the IoT, wireless sensor networks, and embedded systems.