KIET Group of Institutions, Ghaziabad

Computer Science and Information Technology

Summer Internship Report

On

DATA HANDLING WITH PYTHON

Summer Internship at Mini-Project

(J.A.R.V.I.S AI ASSISTANT)

Aug-Sept

(2022)

Submitted By: *Vidhi*

    COURSE : *B.Tech*

    BRANCH : *Computer Science*

    SECTION: *C*

Index for Internship Report:

Following essentials are to be taken in to consideration during preparation of internship report:

- Acknowledgement**.**
- Certificate.
- Abstract.
- Introduction of Project Internship.
- Tasks assigned.
- Technical learning during delivery of task.
- Conclusion of Internship.
- Future scope of work.
- Certificate of moocs with verification link.
- Literature review report**.**
- Daily Log report.

ACKNOWLEDGEMENT:

I've got this golden opportunity to express my kind gratitude and sincere thanks to my director sir, Dr.Amik Garg, KIET Group of Institutions of Engineering and Technology, and **Prof. Ms. Neha Shukla** of "**Dept. Computer Science**" for their kind support and necessary counselling in the preparation of this project report. I'm also indebted to each and every person responsible for the making up of this project directly or indirectly.

I must also acknowledge or deep debt of gratitude each one of my colleague who led this project comeout in the way it is. It's my hard work and untiring sincere efforts and mutual cooperation to bring out the project work. Last but not the least, I would like to thank my parents for their sound counselling and cheerful support. They have always inspired us and kept our spirit up.

VIDHI
B.Tech and CS branch
5rd Semester
University Roll No: 2000290120188

CERTIFICATE

This is to certify that the internship project report entitled **"Data Handling With Python"** submitted by **Ms. Vidhi** in the Department of **Computer Science** of KIET Group of Institutions, Ghaziabad, affiliated to Dr. A. P. J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India, is a record of candidate summer internship. She has successfully completeted his internship under my supervision and guidance and is worthy of consideration for the same.

Signature of Supervisor:
Supervisor's Name:
Date:

ABSTRACT

I have done my Internship in Data Handling With Python, mini project. Internship cum training is an opportunity where we will be learning the new things application of knowledge gained. In this program we have undergone for 3-4 weeks of training cum internship. The purpose of this program is to enhance our knowledge and use them to build the application.

The first week was spent, getting to know Python. We were introduced to Python's philosophy which emphasizes code readability, allowing programmers to express concepts in fewer lines of code, making Python a more simplistic language than others. In the next few weeks, we were taught the fundamental concepts of data handling with python programming. After getting well versed with the concepts we were given practice questions on daily basis to apply our knowledge and learn more.

The final week was spent in understanding Mini project. The project is about " AI ASSISTANT".

INTRODUCTION OF PROJECT INTERNSHIP:

TITLE OF PROJECT: J.A.R.V.I.S AI ASSISTANT

My project is " AI ASSISTANT" which is an application program that understand natural language voice cammand and completes task of users.

Intelligent virtual assistant recognize what user speaks and convert into digital data that software can understand.

An **intelligent virtual assistant** (**IVA**) or **intelligent personal assistant** (**IPA**) is software agent that can perform tasks or services for an individual based on commands or questions. The term "chat bot" is sometimes used to refer to virtual assistants generally or specifically accessed by online chat. In some cases, online chat programs are exclusively for entertainment purposes. Some virtual assistants are able to interpret human speech and respond via synthesized voices. Users can ask their assistants questions, control home automation devices and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands.

There are currently several Python libraries that allow for this functionality and are continuously maintained and have new features added to them.

For "AI ASSISTANT", we will need the following Python library:

1.pyttsx3

2.SpeechRecognizer

3.Os

4.Datetime

5.Wikipedia.

6.Webbrowser

+++++

```python
import pyttsx3
import datetime
import speech_recognition as sr
import wikipedia
import webbrowser
import os


engine = pyttsx3.init('sapi5')
engine.setProperty('rate', 190)

voices =engine.getProperty('voices')
#print(voices[0].id)
engine.setProperty('voice',voices[0].id)


def speak(audio):
    engine.say(audio)
    engine.runAndWait()

def wishMe():
    hour = int(datetime.datetime.now().hour)
    if hour>=0 and hour<12:
        speak("Good Morning!")

    elif hour>=12 and hour<18:
        speak("Good Afternoon!")

    else:
        speak("Good Evening!")

    speak("This is David! Please tell me how may I help you")


def takeCommand():
```

```python
    elif hour>=12 and hour<18:
        speak("Good Afternoon!")

    else:
        speak("Good Evening!")

    speak("This is David! Please tell me how may I help you")


def takeCommand():
    #it takes microfone intput from user

    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        r.pause_threshold=1
        audio = r.listen(source)

    try:
        print("Recognizing...")
        query=r.recognize_google(audio,language='en-in')
        print(f"user said:{query}\n")

    except Exception as e:
        #  print(e)
        print("Say that again please...")
        return "None"

    return query




if __name__=="__main__":
    wishMe()
    while True:
```

```python
        while True:
            query = takeCommand().lower()

        #logic for executing based on query
            if 'wikipedia' in query:
                speak('Searching Wikipedia...')
                query =query.replace("wekipedia","")
                results=wikipedia.summary(query,sentences=2)
                speak("According to Wikipedia")
                print(results)
                speak(results)

            elif 'open youtube' in query:
                webbrowser.open("youtube.com")

            elif 'open google' in query:
                webbrowser.open("google.com")

            elif 'open kiet' in query:
                webbrowser.open("https://www.kiet.edu")

            elif 'play spotify' in query:
                webbrowser.open("https://open.spotify.com/playlist/0JgZlSZSN7iY0t3Jy5TNuF")

            elif 'the time' in query:
                strtime=datetime.datetime.now().strftime("%H:%M:%S")
                speak(f"sir, the time id {strtime} ")

            elif 'open code' in query:
                codepath="C:\\Users\\vidhi\\AppData\\Local\\Programs\\Microsoft VS Code\\Code.exe"
                os.startfile(codepath)

            elif 'play music' in query:
                music_dr="C:\\Users\\vidhi\\OneDrive\\Desktop\\songs"
                songs=os.listdir(music_dr)
                os.startfile(os.path.join(music_dr,songs[1]))

            elif 'quit' in query:
```

```python
                os.startfile(codepath)

            elif 'play music' in query:
                music_dr="C:\\Users\\vidhi\\OneDrive\\Desktop\\songs"
                songs =os.listdir(music_dr)
                os.startfile(os.path.join(music_dr,songs[1]))

            elif 'quit' in query:
                exit()
```

TASK ASSIGNED

This is a project-based course for beginners to become professional Python programmers . Comprising 30 lectures spanning just above 30 hours in total length, this course is ideal for people who have zero experience in the programming world. The course is replete with quizzes and exercises that help with understanding the theoretical and practical aspects of Python programming. The program starts with a complete introduction to Python basics and gradually builds up from there.

We were allotted assignments after every lecture on specific topic in python with additional activities and quizzes on daily basis and also we were making log report on each day of internship.

We read research papers and wrote summary on it.

And at the last week, we were indulged in making mini project .

TECHNICAL LEARNING DURING DELIVERY

Technical learnings are-

- Data types and variables.
- Introduction to OOPs Programming.
- Object Oriented Programming System.
- OOPS Principles.
- Define Classes.
- Creating Objects.
- Class variables and Instance Variables Constructors.
- Basic concept of Object and Classes.
- Access Modifier.
- Modules

<u>CONCLUSION</u>

I believe the trial has shown conclusively that it is both possible and desirable to use Python as the principal teaching language:

1) It is Free (as in both cost and source code).
2) It is a flexible tool that allows both the teaching of traditional procedural programming and modern OOP; It can be used to teach a large number of transferable skills.
3) It is a real-world programming language that can be and is used in academia and the commercial world.
4) It appears to be quicker to learn and, in combination with its many libraries, this offers the possibility of more rapid student development allowing the course to be made more challenging and varied.
5) It is trivial to install on a Windows PC allowing students to take their interest further. For many the hurdle of installing a Pascal or C compiler on a Windows machine is either too expensive or too complicated and most importantly, its clean syntax offers increased understandingand enjoyment for students.

FUTURE SCOPE OF WORK

The Future Scope of a Python Developer is very bright The world is getting digitized. Python has become the core language as far as the success of these technologies is concerned. Let's dive into the technologies which use python as a core element for research, production and further developments. The future is all about automating processes and utilizing the heaps of data to make intelligent decisions. This puts to the forefront technologies such as Artificial Intelligence (AI), machine and deep learning, Internet of Things (IoT), etc.

Python through its high-performance toolkits and Libraries, it has been successfully contributing to analyzing a large number of data sets across computer clusters.

listed below are a few best features of Python that make it unique in itself-

- Open-Source
- Easy and simple to Learn
- Simplify Complex Software Development
- Multiple Programming Paradigms
- Readable and Maintainable Code
- It has a high demand-supply ratio
- Preferable language for Artificial Intelligence and Machine learning

CERTIFICATE OF INTERNSHIP

CERTIFICATE OF COMPLETION

THIS AWARD CERTIFIES THAT

vidhi pandey

HAS SUCCESSFULLY COMPLETED

Live Fundamental Concept Classes Batch - 18 August(KIET)

Minimum 80% Attendance Required	-	6310d676e8b561f4e50cf80e	1 Sep 2022
COURSE CREDITS	**SCORE**	**CERTIFICATE ID**	**DATE**

Validate your certificate by visiting this link: https://mycourse.app/jQ147LduahSqCc3q7

LITERATURE REVIEW REPORT

#RESEARCH PAPER 1

1. A Software Player for Providing Hints in Problem-Based Learning According to a New Specification

Nowadays, learning through computers is performing a standardization process, and proper tools that are compliant with such standards are needed for the successful adoption. Several experiments have shown that the provision of hints during problem solving has been a successful strategy in the learning process. Next, there is a list of the software player features that are in accordance with the hints model requirements. The combination of all the features is allowed. The corresponding justification for each feature is provided. Problems can be presented to students as part of assessments. Each problem can have an indefinite number of hints related to it. he hints related to a problem can be available at the beginning or they can become available only as a result of an incorrect answer to the problem. The atomic elements of hints can be text or other problems. An indefinite level of hints hierarchy can be achieved. Text Feedback is allowed in whichever problem. Hint Sequencing is allowed. Hint Grouping is allowed. Students can select a maximum number of hints from a list based on meta-information. Definition of different hints depending on the student aim is allowed. Hints for assessments and sections can be provided. An XML binding of the full hinting data model has been provided. Different authoring tools can generate XML files that are in accordance with the hinting model, while different players can load and interpret these XML files according to the semantics with independence from the authoring tool that created it. In this way, interoperability between different hinting systems and reusability of different hinting descriptions can be achieved. We have implemented the first player that is full compliant to this new hints specification, and this software module can interpret and run these XML files with the help of the XTutor system. At present, we are using the XTutor hints player in classroom in order to let students interact with created hinting material. From quantitative and qualitative data analysis from the experiences, we can infer that the tool is very useful in the learning process and positive results have been obtained.

1. **Computer programming with Python for industrial and system engineers:**

Programming requires students to break a large problem into smaller pieces and then solve them piece by piece. A basic understanding of programming is a necessity in nearly every engineering field. It is even more so in the future of technological growth and development. first question that needed to be answered was which programming language should be taught. After a thorough survey of the popularity and practicality of a variety of programming languages, Python was found to be the best alternative. Python greatly reduces the complexity of problem solving by hiding the intricate "what is really happening under the hood" side of computing. Python is a high-level language that is consistently ranked among the most popular programming languages by TIOBE Index , IEEE Spectrum , and GitHut . In a recent survey by the Association for Computing Machinery. Python surpassed Java as the top introductory teaching language. Eight of the top ten universities in the United States use Python to introduce programming. The significance of an open-source alternative. Notevery one has access to a "free" Matlab license, and Matlab may not be available to students when they leave the university. Python offers them a free, always available option. The easy installation of Python distributions. Recent Python distributions such as Anaconda and Enthought Canopy provide "oneclick" installers that make the establishment of scientific and engineering computing environment much easier than installing Matlab. The nice integration with lecture notes. This is a bonus point for teachers because they have been increasingly integrating code into lecture notes to enhance students learning. This integration is not easy with Matlab, but it is straight forward with Python through Jupyter Notebook. Lessons learned from the instructor's and the students' perspectives are presented. End course surveys helped in understanding both the successes as well as areas that need improvement. This paper identifies that the learning performance is slightly different for female and male students, which may be caused by the difference in sample sizes. There are also reasonable differences according to students' class standing and attendance. The introduction of this course into the undergraduate curriculum has helped fillan observed deficit in the ISE program. Students learned how to apply Python programming skills to solve diverse, but relevant, ISE problems. Students have been better prepared for future employment and academic opportunities through establishment of this base in computer programming. In accordance with the great interest among students in learning how Python can be used to mine Twitter and Facebook data, the plan is to integrate one mor e chapter

titled "Network for Social Media Mining." Networkx is a library that can create and manipulating networks with complex structures and dynamics. An introduction to the basic s of the agent-based model (ABM), an important system engineering technique, will be integrated into this chapter as well. ABM can be used to simul ate the actions and interactions among individuals with a view to assessing their effects on the system as a whole.

How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future

The world has encountered the most cruicial tragedy of all times aka covid19. The Covid-19 pandemic represents one of the greatest global health emergencies of the last few decades with indelible consequences for all societies throughout the world. Covid-19 pandemic is a major global challeng to public health along with devastating social, economic and political impact. Millions of people have been infected, frequently requiring continuous assistance and monitoring. Smart healthcare technologies and Artificial Intelligence algorithms constitute promising solutions useful not only for the monitoring of patient care but also in order to support the early diagnosis, prevention and evaluation of Covid-19 in a faster and more accurate way.

The data generated on a daily basis through the internet of things along with classical datasets, can help us better understand the dynamics of infectious diseases, their progression, response to treatment and transmission if used cautiously which can make AI storger to fight covid 19 efficiently. . The availability of sufficient amount of high-quality data is crucial for a successful implementation of AI in the management of Covid-19.

AI-based tools are being evaluated and used in preliminary screening of SARS-CoV-2 infected individuals. Through the AI-based learning frameworks, individuals are being screened for any eventual infections leading to their categorizations as low, moderate, or high risk individuals.

AI tools may contribute significantly in better understanding the nature of the etiological virus (SARS-CoV-2), designing vaccines, identifying therapeutic targets, diagnosis of the disease and predicting new outbreaks spots based on the available digital data.

AI-based approaches are being implemented in the diagnosis of Covid-19. Some of the AI-based tools have demonstrated to perform efficient classifcation of the level of severity of Covid-19 by using radiographic images such as X-ray and CT-scans. For the diagnosis of Covid-19, supervised learning algorithms such as convolutional neural network (CNN), linear discriminant analysis, Naive Bayes, support vector machines, decision trees, logistic regression and random forest are being used.

The vaccine development strategies against the pathogenic human CoVs including SARS and MERS have targeted either the whole virus or its structural proteins nucleoprotein (N), spike protein (S) or membrane protein (M). AI and machine learning based approach may play a critical role in vaccine design by contributing to the entire process using in silico-based analysis, prediction and validation.

Also during covid period It is also observed that some initial false-negative rRT-PCR results, when complemented by AI-based deep learning computer- aided diagnostic system, were later

confirmed positive. Thus, a more reliable diagnosis of Covid-19 can be achieved by complementary combined AI-based system, such as the application of CT derived deep leaning methods along with rRT-PCR Three AI models are used to generate the probability of a patient being COVID-19 (+): the first is based on a chest CT scan, the second on clinical information and the third on a combination of the chest CT scan and clinical information.

Machine learning in cybersecurity: A review

Today's world is highly network interconnected of small personal devices (e.g., smartphones) as well as large computing devices or services (e.g., cloud computing or online banking), and thereby each passing minute millions of data are being generated, processed, exchanged, shared, and utilized to yield outcomes in specific applications. Thus, securing the data, machines (devices), and user's privacy in cyberspace has become an utmost concern for individuals, business organizations, and national governments.So a advanced cybersecurity is required to secure data which can be achieved only through Machine learning. ML algorithms are vulnerable to attacks both in the training and testing phases, which usually leads to remarkable performance decreases and security breaches.

Machine learning has become a vital technology for cybersecurity. Machine learning provide solutions to various cybersecurity attack.. As with development of technology ,cyber attacks become more common and easier.A malicious hacker can manipulate the input data, exploiting specific data and compromise the entire security. With machine learning, cybersecurity systems can analyze patterns and learn from them to help prevent similar attacks and respond to changing behavior.

Recently, cyber-attackers cause huge blackouts which is cause huge loss for country economy and security. This malicious attacker have become threat to control system, hence we need to secure them, Machine learning algorithms can be used to have an automatic intrusion detection system that can secure control systems and critical infrastructures (CI). Intrusion detection in some systems are :

SCADA systems are essential for managing and monitoring CI such as electric power generation, transmission and distribution or a water treatment plant, for better performance now they are connected to IT network ,which increase the threat and risk of being hacked. Hence here is a need to develop a intrusion detection systems (IDS), which are used to identify an attack and initiate proper alerts that may help take appropriate actions.

VANET is an emerging technology in modern transportation systems for providing safety and valuable information. It provides the benefits of road safety and travelers comfort while protecting driver's privacy from different types of attack .This detection can be done more accurately if there is a collaboration among vehicles in VANET, but the major problem is collaborative learning can leads to privacy data can be compromised.A CIDS is developed that helps the vehicles to share the information without exchanging the training data.

A serious and common threat is malware refers to a program that is inserted into a system, usually covertly, with the intent of compromising the confidentiality, integrity, or availability of the victim's data, applications, or operating system or of otherwise annoying or disrupting the victim. Combating malware is very important for software /privacy data security. The advantage of machine learning techniques is that it will not only detect known malware but also deliver knowledge for the detection of new malwar.

In short, machine learning can make cybersecurity [simpler, more proactive, less expensive and far more effective](#).

1. <u>Grammar engineering for multiple front-ends for Python</u>

Brian A. Malloy

Grammars for programming languages can be conducted in either a theoretical context, such as the study of parsing algorithms, or a software development context, where grammars are seen as a software artefact. The second context is particularly relevant to the practice of software engineering, since many of the tools used to analyse, measure and translate software are grammar-based systems. One problem that arises in grammar engineering is the construction of grammar ware for the various dialects or versions of a language. The rules consist of terminal symbols that represent tokens in the language and nonterminal symbols that are used to rewrite the rules used in the derivation of the program. Grammars can be expressed in many formalisms; however, in this paper, we use the term grammar to refer to either a CFG or a grammar in EBNF format. When necessary, we refer specifically to either CFG or the EBNF format. A parser is a software component that accepts a program as input and determines if that program is syntactically correct with respect to the grammar used to build the parser, which we refer to as the reference grammar. A compiler writer, like any software developer, can profit from the use of software development tools including parser generators such as yacc, bison, antlr, or menhir.[19, 20] The bison and menhir parser generators are generalised forms of yacc; however, bison cannot accept a grammar in EBNF format but rather the grammar must be transformed into a machine readable format known as bison syntax or bison grammar. In a bison grammar, a non-terminal or terminal symbol is represented as an identifier, similar to a C identifier. Terminal symbols are also known as token types and can be represented as character literals. In this paper, we refer to a grammar that has been transformed from EBNF into bison syntax as a yaccified grammar. we report on the construction of a front end for Python based on integrating our parsers with a flex-based scanner. We provide empirical evidence of the success of this strategy using a test suite of Python programs. While the use of language-preserving XBGF transformations ensures that we do

not change the language specified by the Python grammar from that found at python.org, the use of a test suite serves to validate its integration with the scanner and scaffolding code in the front end. The use of the verified XBGF transformations, supported by analysis and implementation in Rascal, is a good example of "lightweight" formal methods. Since the transformations were verified in advance, they removed the burden of verification when applying individual instances of the transformations. Even though the transformations were verified, their implementation still posed a threat to internal validity, and other parts of the system, such as the scanner, were not subject to verification. Thus, assembling and using a large test suite as described in Section 6.1 was vital to ensuring the full integrity of the system.The error messages and modularisation constructs of menhir were highly effective for debugging the grammar, and we hope that some version of these features will eventually find their way into other tools, including bison. In future work we intend to reexamine the relationship between the transformations for the different versions of the grammars. At the moment, scripts written in the transformation language are just a list of individual transformation operations. If this transformation language could be enhanced with structuring and modularisation constructs it would greatly aid with the process of maintaining multiple versions of parsers