INTERNSHIP REPORT

INTERNSHIP REPORT
ON
DATA SCIENCE USING PYTHON

K.Sai Ananya
Roll No.-22951A66J8

DATA SCIENCE USING PYTHON

AN **INTERNSHIP REPORT**

**Bachelor of Technology in
Computer Science and Engineering(AI&ML)** *by*

**K.SAI ANANYA
Roll no-22951A66J8**

Department of CSE(AI&ML)

INSTITUTE OF AERONAUTICAL ENGINEERING
(Autonomous)
Dundigal, Hyderabad – 500 043, Telangana

September, 2023

DECLARATION

I certify that

a. The work contained in this report is original and has been done by me under the guidance of my supervisor (s).

b. The work has not been submitted to any other Institute for any degree or diploma.

c. I have followed the guidelines provided by the Institute for preparing the report.

d. I have conformed to the norms and guidelines given in the Code of Conduct of theInstitute.

e. Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the report and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Place:Dundigal,Hyderabad   **Signature of the Student**

Date: **Roll No.22951A66J8**

CERTIFICATE

राष्ट्रीय लघु उद्योग निगम—तकनीकी सेवा केन्द्र
THE NATIONAL SMALL INDUSTRIES CORPORATION LTD.
TECHNICAL SERVICES CENTRE

(भारत सरकार का उद्यम / A Government of India Enterprises)
ई.सी.आई.एल एक्स रोड, कुशाईगुडा, हैदराबाद — 500062, तेलंगाना, भारत
E.C.I.L X Road, Kushaiguda, Hyderabad - 500062, Telangana, India.

क्रमांक / S.No. 205368

दिनांक / Date: 31/07/2023

Certificate

This is to certify that Mr. / Ms. __Kanna Sai Ananya__
son/daughter of Mr. __Kanna Venugopal__ pursuing __BTech__ in __AI&ML__ from
(College Name) __Institute of Aeronautical Engineering, Hyderabad__
Roll No. __22951A66J8__ has successfully completed the __Field Project__
entitled/in the area of __Data Science Using Python__

_____ under
our guidance. It is a bonafide work carried out by her/him from __17/07/2023__ to __31/07/2023__

He/She has completed the assigned module as per the requirements within the time frame
During the above period, the trainee's conduct was found __Good__

Project Coordinator

Centre Head

वह प्रमाण पत्र होलोग्राम लगा होने के साथ ही मान्य होगा / This Certificate shall be valid only with affixed hologram

APPROVAL SHEET

This INTERNSHIP REPORT entitled **Data Science using Python** by **K.Sai Ananya** is approved for the award of the Degree Bachelor of Technology in **Computer Science And Engineering(AI&ML)**

Supervisor   **Head of the Department**

Date:

Place:

ABSTRACT

Keywords: Big Data, predictive modeling, machine learning, statistical analysis, data visualization, programming, and data manipulation.

The goals, approaches, and conclusions of an internship at the National Small Industries Corporation (NSIC) with a focus on the use of Python programming in data science are summarized in this abstract. The goal of the internship was to give participants hands-on experience with Python libraries and tools for analysis, visualization, and predictive modeling.

An in-depth introduction to Python programming basics, covering data structures, control flow, and functions, was given at the start of the internship. Important libraries including NumPy, Pandas, Matplotlib, and Scikit-learn were presented to the participants; these are necessary for a variety of data science projects.

Participants worked on practical projects and exercises throughout the internship to solidify their grasp of Python for data science. These projects included feature engineering, exploratory data analysis, machine learning model creation, and data cleaning and preprocessing.

Acquainting participants with real-world datasets and industry-relevant problems was one of the internship's main goals. In order to accomplish this, participants worked on projects and case studies drawn from a variety of industries, including social media, e-commerce, healthcare, and finance.

The internship also placed a strong emphasis on problem-solving and collaborative learning. In order to promote a collaborative and encouraging environment that is conducive to skill development, participants were encouraged to work in teams to tackle complex data science tasks.

All things considered, the internship offered a thorough and hands-on introduction to the use of Python in data research. By gaining practical experience and honing the skills necessary to address real-world data challenges, participants set themselves up for success in the data science industry.

Internship Report Brief

Introduction:

The National Small Industries Corporation (NSIC) internship focused on equipping participants with practical skills in utilizing Python for data science applications. This report provides a comprehensive overview of the internship program, including objectives, methodologies, and outcomes.

Objectives:

- To familiarize participants with Python programming fundamentals.
- To introduce participants to key Python libraries for data manipulation, analysis, and visualization.
- To provide hands-on experience in applying Python for data preprocessing, exploratory data analysis, and machine learning.
- To engage participants in collaborative projects addressing real-world data science challenges.

Methodologies:

Curriculum Design: The internship curriculum was structured to cover Python programming basics, followed by in-depth exploration of data science libraries such as NumPy, Pandas, Matplotlib, and Scikitlearn.

Hands-on Projects: Participants engaged in practical exercises and projects aimed at applying Python for data manipulation, visualization, and machine learning. Projects included real-world datasets from various domains.

Collaborative Learning: Participants worked in teams to tackle complex data science challenges, fostering collaboration and problem-solving skills.

Mentorship: Experienced mentors provided guidance and support throughout the internship, offering assistance with technical challenges and project development.

Outcomes:

Enhanced Python Proficiency: Participants developed a strong foundation in Python programming and gained proficiency in utilizing Python libraries for data science tasks.

Practical Skills: Participants acquired practical skills in data preprocessing, exploratory data analysis, visualization, and machine learning model development using Python.

Collaborative Experience: The internship fostered a collaborative learning environment, enabling participants to work effectively in teams and leverage collective expertise.

Real-world Application: Participants applied their skills to real-world datasets, gaining insights into industry-relevant data science challenges and solutions.

Challenges:

Technical Complexity: Understanding and implementing Python libraries and tools for data science can be challenging, especially for interns with limited programming experience. The complexity of algorithms and techniques used in data analysis and machine learning may require additional time and effort to comprehend.

Data Quality Issues: Real-world datasets often contain missing values, outliers, or inconsistencies, which can affect the accuracy and reliability of analysis. Cleaning and preprocessing data to ensure its quality can be time-consuming and may require advanced techniques.
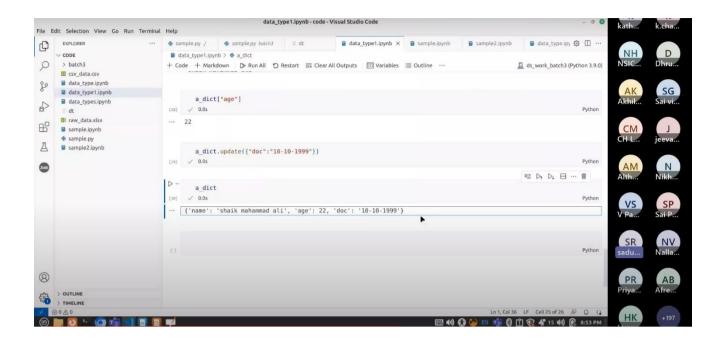
Algorithm Selection: Choosing the appropriate algorithms and models for a given problem is crucial in data science. Interns may face challenges in selecting the most suitable algorithms, understanding their assumptions and limitations, and tuning their parameters for optimal performance.
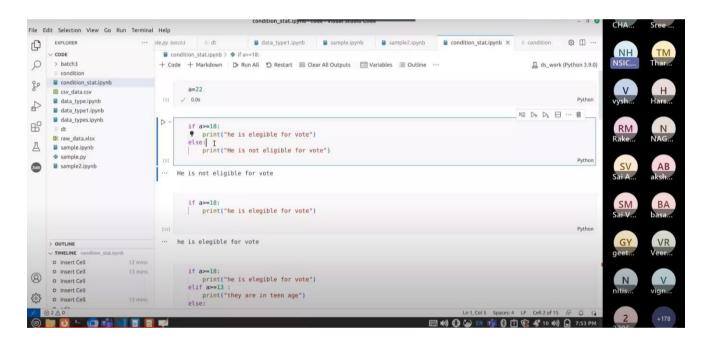
Resource Constraints: Limited access to computational resources such as high-performance computing clusters or cloud services can impede the scalability and efficiency of data analysis and modeling tasks, particularly when working with large datasets.

Conclusion:

The NSIC internship provided participants with a valuable opportunity to explore data science using Python. By combining theoretical knowledge with hands-on experience and collaborative projects, participants emerged with enhanced skills and readiness to tackle data science challenges in diverse domains. The internship served as a stepping stone for participants' careers in the dynamic field of data science.

Screenshots of Virtual Training.

THANK YOU