A REPORT ON SUMMER INTERNSHIP

Name of the Student   **:** Harika Vavilapalli

Name of the College   **:** Vignan's Institute of Information Technology

Registration Number   **:** 22L31A4359

Period of Internship   : 4 Weeks

From   **:** 29th April 2024

Year   **:** II

Name and Address of the Intern Organization: EISYSTEMS SERVICES 110, First Floor, Express Greens Plaza, Vaishali, Sector 1, Ghaziabad, Uttar Pradesh 201010

An Internship Report on

MACHINE LEARNING INTERNSHIP

Submitted in accordance with the requirement for the degree of

Bachelor of Technology
in
Computer Science & Engineering – (Artificial Intelligence)

By
HARIKA VAVILAPALLI

Under the Guidance of
Mrs B. Pavani

VIGNAN'S INSTITUTE OF INFORMATION TECHNOLOGY (AUTONOMOUS)
www.vignaniit.com (Approved by AICTE & Affiliated to JNTUK)   Estd – 2002
Beside VSEZ, Duvvada, VISAKHAPATNAM-530049. A.P.
Phone : 0891- 2755222 / 333 / 444 **::** Fax : 0891 - 2752333 **::** E-Mail : vignaniit@yahoo.com

DEPARTMENT OF ADVANCED COMPUTER SCIENCE & ENGINEERING

May 2024

DECLARATION

I, HARIKA VAVILAPALLI Reg. No **: 22L31A4359,** of the Department of **Computer Science and Engineering-Artificial Intelligence** do hereby declare that I have completed the mandatory internship from EIsystems under the Faculty Guideship of Mrs. B. Pavani, Department of Computer Science and Engineering-Artificial Intelligence, Vignan's Institute of Information Technology.

Signature of the Student

CERTIFICATE

This is to certify that Harika Vavilapalli Reg. No.22L31A4359 has completed her Internship in EIsystems under my supervision as a part of partial fulfillment of the requirement for the Degree of in the Department of Computer Science and Engineering-Artificial Intelligence, Vignan's Institute of Information Technology. This is accepted for evaluation.

Signature of Guide   **Head of the department**

Mrs. B. Pavani   Mrs. K. Swathi

Assistant Professor   Assistant Professor

CERTIFICATE

Certificate of Internship

This is to certify that

VAVILAPALLI HARIKA

did his/her internship with us between

30 APRIL TO 29 MAY OF 2024 (4 WEEKS)

As a part of internship, he/she completed training on

MACHINE LEARNING

followed by submitting a report on

NEURAL NETWORK MODEL FOR REAL TIME CREDIT CARD FRAUD DETECTION

Mayur Dev Sewak
Head, Internship & Operations
Eisystems Services

Mallika Srivastava
Head, Training Delivery
Eisystems Technologies

EISystems

Robokwik

Date of Issue : 20-Jun-24
Certificate ID : EIS/AICT/24SP-0551
AICTE ORID   : CORPORATE65a015a9925f071704990121
Internship Name / Phase ID : EISYS-4/8/12/16 2024

Validate at robokwik.com/certificatevalidation

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation for the opportunity to complete my internship as a **Machine Learning intern** at EIsystems from 29-04-2024 to 30-05-2024. I am also grateful to Vignan's Institute of Information Technology for providing me with the necessary skills and knowledge that laid the foundation for this enriching experience. I am deeply grateful to the faculty coordinator Mrs. B. Pavani for support and guidance throughout my internship.

CONTENTS

CHAPTER 1: EXECUTIVE SUMMARY

This internship report provides an overview of my internship experience as a Machine Learning intern at EIsystems. The report highlights the learning objectives and outcomes achieved and summarizes the activities performed during the internship period.

Learning Objectives and Outcomes:

1. Objective: Foundations of Python
2. Objective: Exception Handling
3. Objective: Object oriented programming
4. Objective: Data mining and preprocessing
5. Objective: Numpy,pandas,model deployement

Internship Organization:

EIsystems

Information technology is making possible substantial changes in the organization and effectiveness of manufacturing activities. Equipment and stations within factories, entire manufacturing enterprises, and webs of suppliers, partners, and customers located throughout the world can be more effectively connected and integrated through the use of information technology. Information technology provides the tools to achieve goals that are widely regarded as critical to the future of manufacturing: rapid shifts in production from one product to another; faster implementation of new concepts in products and faster delivery of products to customers; more intimate interactions with customers, who more directly and completely specify what they need; fuller utilization of capital; and streamlining of operations to focus on what is essential to a business and to eliminate unnecessary activities. As this list suggests, technology is a critical enabler, but its development and implementation will be shaped by organizational, managerial, and human resources concerns. Because of these concerns, manufacturers have had difficulty getting the most out of the technology that exists today. Sensitivity to these concerns is essential to the successful development and implementation of the technologies associated with visions of manufacturing for the 21st century.

CHAPTER 1: ACTIVITY LOG FOR WEEK-1

Day	Date	Brief description of the daily activity	Learning Outcome
Day-1	01-05-2024	Introduction to Machine Learning	Understand what ML is and how it differs from traditional programming Learn the basic working principles and types of ML Identify real-world applications of ML in industries
Day-2	02-05-2024	Introduction to Python for ML Topics Covered:	Set up the Python development environment Understand the basics of Jupyter Notebook and Anaconda
Day-3	03-05-2024	Python Fundamentals - Conditional Statements & Data Types	Learn and apply Python's basic data structures Understand and use conditional statements for decision-making
Day-4	04-05-2024	Python Loops & Iterations	Use loops efficiently to automate repetitive tasks Understand different looping structures and their use cases
Day-5	05-05-2024	Functions and Packages	Write reusable and modular Python functions Learn how to use and install external Python packages
Day -6	06-05-2024	Working with Various Python Libraries	Get familiar with essential Python libraries for ML Understand their role in data manipulation, visualization.

<h1 style="text-align:center">WEEKLY REPORT</h1>

<p style="text-align:center">WEEK – 1 (From Date:01-05-2024 to Date:06-05-2024)</p>

Objective of the Activity Done:

Detailed Report:

Day 1: the internship begins with an introduction to Machine Learning (ML). Participants will understand the definition of ML, its types (Supervised, Unsupervised, and Reinforcement Learning), and its working principles. They will explore how ML differs from traditional programming and learn about real-world use cases such as self-driving cars, recommendation systems, and fraud detection.

Day 2: the focus shifts to Python, the most widely used language for ML. Participants will set up their Python development environment using Anaconda and Jupyter Notebook. They will learn the basics of Python programming, including syntax, data types, and working with variables.

Day 3: participants will explore fundamental Python concepts such as conditional statements (if, elif, else) and different data structures like lists, tuples, and dictionaries. Understanding these concepts will help in structuring data and making logical decisions in programs.

Day 4: loops and iterations will be covered. Participants will work with for and while loops, as well as loop control statements (break, continue, pass). These concepts are essential for automating repetitive tasks in data processing.

Day 5: the session will introduce functions and packages in Python. Participants will learn how to create functions, pass arguments, and return values. Additionally, they will explore Python packages and learn how to install and use external libraries essential for ML.

Day 6: various Python libraries like NumPy, Pandas, Matplotlib, and Scikit-learn will be introduced. These libraries play a crucial role in data handling, analysis, visualization, and ML model building.

CHAPTER 2: ACTIVITY LOG FOR WEEK-2

Day	Date	Brief description of the daily activity	Learning Outcome
Day-7	07-05-2024	Introduction to NumPy	Understanding NumPy arrays Creating 1D, 2D, and 3D arrays Indexing and slicing arrays
Day-8	08-05-2024	Data Processing with NumPy	Perform mathematical computations on data using NumPy Handle missing or incorrect data using NumPy functions
Day-9	09-05-2024	Introduction to Pandas	Learn how to structure and analyze tabular data using Pandas Work with datasets for data processing
Day-10	10-05-2024	Data Analysis using Pandas	Performing basic operations on data (filtering, sorting, merging) Grouping and aggregating data Handling missing data
Day-11	11-05-2024	Data Visualization using Matplotlib	Plotting graphs using Matplotlib Line charts, bar charts, scatter plots Customizing plots (labels, titles, legends)
Day-12	12-05-2024	Advanced Visualization Techniques	Using different styles in Matplotlib Subplots and multiple plots in one figure Styling and annotating graphs

WEEKLY REPORT

WEEK – 2 (From Date:07-05-2024 to Date:12-05-2024)

Objective of the Activity Done:

Detailed Report:

Day 7: participants will engage in a hands-on practice session where they will apply Python concepts learned throughout the week. Small coding exercises and real-world tasks will help reinforce their understanding.

Day 8: NumPy, a fundamental library for numerical computing, will be introduced. Participants will learn how to create NumPy arrays, manipulate data structures, and perform mathematical operations on arrays.

Day 9: deeper insights into data processing using NumPy will be covered. Concepts such as reshaping arrays, handling missing values, and performing operations on large datasets will be discussed.

Day 10: participants will start working with Pandas, a powerful library for data manipulation and analysis. They will learn about Pandas Series and DataFrames, how to create and modify them, and how to import datasets for analysis.

Day 11: data analysis using Pandas will be explored in-depth. Participants will perform operations such as filtering, sorting, merging, and grouping data. They will also learn how to handle missing data, an essential step in preprocessing datasets for ML models

Day 12: data visualization using Matplotlib will be introduced. Participants will learn how to create basic plots like line graphs, bar charts, and scatter plots to visualize data trends.

CHAPTER 3: ACTIVITY LOG FOR WEEK-3

Day	Date	Brief description of the daily activity	Learning Outcome
Day-13	13-05-2024	Hands-on Practice & Mini Project	Strengthen data analysis and visualization skills Apply Pandas and Matplotlib in real-world scenarios
Day-14	14-05-2024	Introduction to Scikit-learn	Understand the importance of Scikit-learn for ML Learn how to split data for model training
Day-15	15-05-2024	Introduction to Machine Learning Models	Learn how ML models are trained and tested Implement a simple regression model using Scikit-learn
Day-16	16-05-2024	Data Analysis using Pandas	Performing basic operations on data (filtering, sorting, merging) Grouping and aggregating data Handling missing data
Day-17	17-05-2024	Classification Models in ML	Understanding classification problems Implementing logistic regression Evaluating model accuracy)
Day-18	18-05-2024	Advanced Visualization Working with Real Datasets	Using different styles in Matplotlib Subplots and multiple plots in one figure Styling and annotating graphs

WEEKLY REPORT

WEEK – 3 (From Date:13-05-2024 to Date:18-05-2024)

Objective of the Activity Done:

Detailed Report:

Day 13: participants will engage in a hands-on practice session where they will apply Python concepts learned throughout the week. Small coding exercises and real-world tasks will help reinforce their understanding.

Day 14: NumPy, a fundamental library for numerical computing, will be introduced. Participants will learn how to create NumPy arrays, manipulate data structures, and perform mathematical operations on arrays.

Day 15: deeper insights into data processing using NumPy will be covered. Concepts such as reshaping arrays, handling missing values, and performing operations on large datasets will be discussed.

Day 16: participants will start working with Pandas, a powerful library for data manipulation and analysis. They will learn about Pandas Series and DataFrames, how to create and modify them, and how to import datasets for analysis.

Day 17: data analysis using Pandas will be explored in-depth. Participants will perform operations such as filtering, sorting, merging, and grouping data. They will also learn how to handle missing data, an essential step in preprocessing datasets for ML models

Day 18: data visualization using Matplotlib will be introduced. Participants will learn how to create basic plots like line graphs, bar charts, and scatter plots to visualize data trends.

CHAPTER 4: ACTIVITY LOG FOR WEEK-4

Day	Date	Brief description of the daily activity	Learning Outcome
Day-19	19-05-2024	Implementing Machine Learning Algorithms	Learn how different ML algorithms work Compare model accuracy using Scikit-learn
Day-20	20-05-2024	Hyperparameter Tuning and Model Optimization	Perform mathematical computations on data using NumPy Handle missing or incorrect data using NumPy functions
Day-21	21-05-2024	Hands-on ML Project	Applying ML techniques on a dataset Training, testing, and evaluating models Recap of the week's topics
Day-22	22-05-2024	Introduction to AI and Deep Learning	Performing basic operations on data (filtering, sorting, merging) Grouping and aggregating data Handling missing data
Day-23	23-05-2024	Working on ML Model Deployment	Saving and loading ML models Introduction to Flask for deploying models
Day-24	24-05-2024	Working on a Capstone Project	Evaluating final model performance Writing reports and documentation

WEEKLY REPORT

WEEK – 4 (From Date:19-05-2024 to Date:24-05-2024)

Objective of the Activity Done:

Detailed Report:

Day 19: participants will engage in a hands-on practice session where they will apply Python concepts learned throughout the week. Small coding exercises and real-world tasks will help reinforce their understanding.

Day 20: NumPy, a fundamental library for numerical computing, will be introduced. Participants will learn how to create NumPy arrays, manipulate data structures, and perform mathematical operations on arrays.

Day 21: deeper insights into data processing using NumPy will be covered. Concepts such as reshaping arrays, handling missing values, and performing operations on large datasets will be discussed.

Day 22: participants will start working with Pandas, a powerful library for data manipulation and analysis. They will learn about Pandas Series and DataFrames, how to create and modify them, and how to import datasets for analysis.

Day 23: data analysis using Pandas will be explored in-depth. Participants will perform operations such a filtering, sorting, merging, and grouping data. They will also learn how to handle missing data, an essential step in preprocessing datasets for ML models

Day 24: data visualization using Matplotlib will be introduced. Participants will learn how to create basic plots like line graphs, bar charts, and scatter plots to visualize data trends.

CHAPTER 5: OUTCOMES DESCRIPTION

Describe the work environment you have experienced:
During my virtual internship, I worked remotely with flexible hours. Communication was mainly through video calls and messaging platforms. I managed tasks independently while collaborating with my team online

Describe the real time technical skills you have acquired:
I have acquired technical skills in web development, including proficiency in HTML, CSS, and JavaScript. I'm experienced in using DOM manipulation, handling events, and working with arrays and objects. Additionally, I've learned to use advanced array methods, built-in JavaScript functions, and manage dynamic web content through event-driven programming.

Describe how could you improve your communication skills:
To improve my communication skills, I could practice active listening, ensuring I fully understand others before responding. I could also focus on being more concise and clear in both written and verbal communication. Regularly seeking feedback and engaging in discussions or presentations would help me refine my ability to convey ideas effectively and confidently.
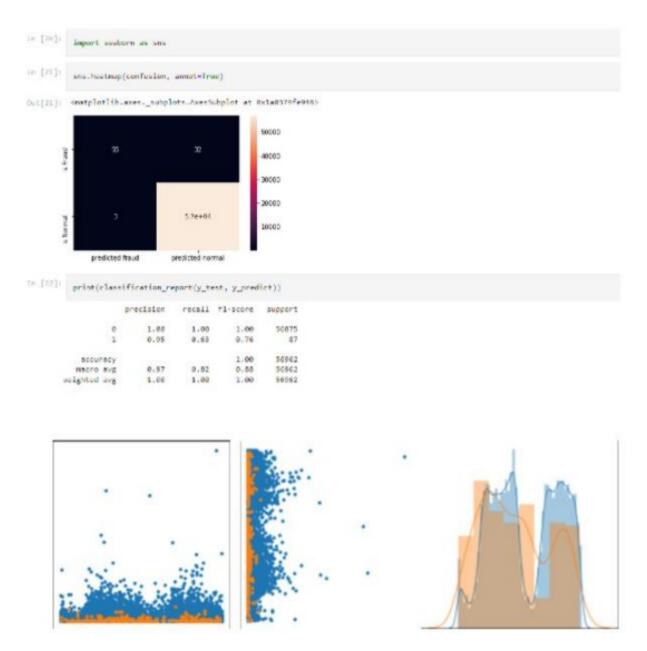
```
in [20]:   import seaborn as sns
```

```
in [21]:   sns.heatmap(confusion, annot=True)
```

```
Out[21]:   <matplotlib.axes._subplots.AxesSubplot at 0x1a8376fe968>
```



```
in [22]:   print(classification_report(y_test, y_predict))

                   precision    recall  f1-score   support

              0        1.00      1.00      1.00     50875
              1        0.05      0.63      0.76        87

       accuracy                           1.00     50962
      macro avg        0.57      0.82      0.88     50962
   weighted avg        1.00      1.00      1.00     50962
```
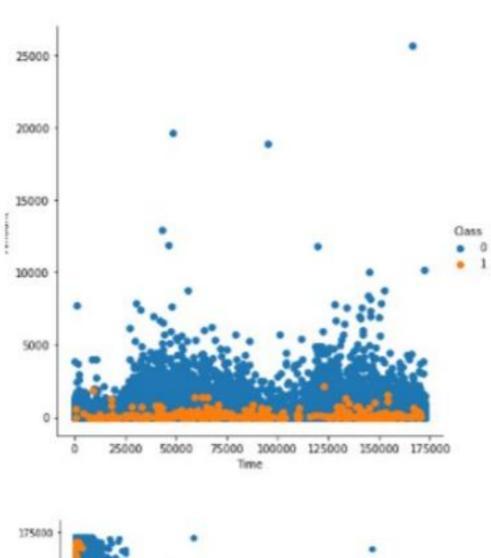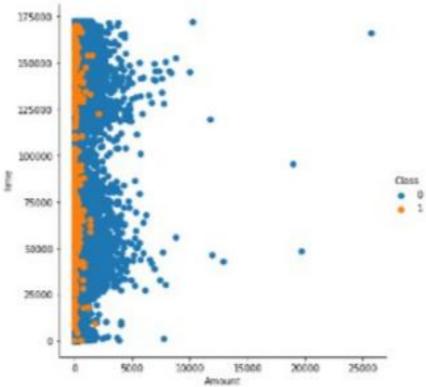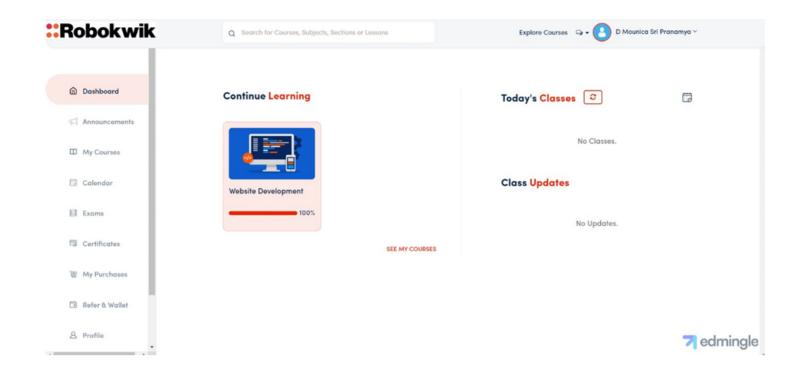

Fig:(Training dataset and visualizing)

Fig:(Dataset preprocessing live demo)

Fig:(completion of given material)