Summer Internship Report

Submitted

by

MOHAMMED KHIZAR BAIG - AP19110010167

BOYAPATI SAI VENKAT – AP19110010174

ISHITA AGRAWAL – AP19110010511

G.B.N.V.S. ABHINAV – AP19110010166

N. SREE ARAVIND BHASKAR – AP19110010196

Department of Computer Science and Engineering

SRM University-AP, Andhra Pradesh, India

July - 2021

DATASHEET

Roll Numbers	:	1. AP19110010167 2. AP19110010174 3. AP19110010511 4. AP19110010166 5. AP19110010196	
Names of the student	:	1. MOHAMMED KHIZAR BAIG 2. BOYAPATI SAI VENKAT 3. ISHITA AGRAWAL 4. G.B.N.V.S. ABHINAV **5.** N. SREE ARAVIND BHASKAR	
Branch & Section	:	CSE B, E	
Batch	:	2019-2023	
Type of internship	:	Industry internship - Project with Faculty	
Company Name/Institute Name	:	Andhra Pradesh State Skill Development Corporation (APSSDC)	
Company/Institute Website	:	https://www.apssdc.in/home/homepage	
Start Date (MM/DD/YYYY)	:	06/07/2021	
End Date (MM/DD/YYYY)	:	07/14/2021	
Duration (No. of days)	:	45 Days	
Status of the internship	:	Completed	
Name of internship mentor (SRM Faculty)	:	Dr. Saleti Sumalatha	
Names & Signatures of the student	:	Mohammed Khizar Baig	
		Boyapati Sai Venkat	B. Sai Venkat.

		Ishita Agrawal	
		G.B.N.V.S. Abhinav	
		N. Sree Aravind Bhaskar	

ACKNOWLEDGEMENT

The internship opportunity we had with APSSDC along with Dr. Saleti Sumalatha was a great chance for learning and professional development. We extend our gratitude to all the stakeholders, especially the Department of Corporate Relations and Career Services of SRM University AP, Andhra Pradesh to enlighten us regarding this opportunity.

The Data Science Using Python internship was indeed very informative and helped in developing the knowledge of the practical and real-world scenarios.

We would also like to acknowledge the Industry Trainers Ms. Niharika, Ms. Lavanya, and Ms. Mouni to encourage and introduce several key concepts which are beneficial for the corporate ecosystem and the streamlined completion of the program.

Also a special thanks to Dr. Saleti Sumalatha, our internship mentor and faculty support for her overall support and guidance throughout this initiative.

We perceive this opportunity as a big milestone in our career enhancement. We will strive to use gained skills and knowledge in the best possible way, to improvise and attain the desired career objectives.

TABLE OF CONTENTS

I. INTRODUCTION

The internship was titled "Data Science Using Python" and was administered by the AP Skill Development Corporation for a duration of 1 month and 8 days. There were three separate instructors who managed different aspects of the entire program.

After a certain point, we had collaborated with Dr. Saleti Sumalatha to pursue our project titled "E-Commerce Review Analysis".

It is a project which we intend to make using Sentimental Analysis. It starts with the notion of an assessment of sentiments that have emerged as a method for understanding clients' emotions and an efficient knowledge representation for recognizing opinions on the products of E-Commerce companies and segregating them into categories.

II. OBJECTIVE OF THE INTERNSHIP

The primary objective of the internship program was to familiarise us with the concept of Data Science. It was done in a phase-wise manner starting from Python basics to machine learning introduction to finally the concept of data analysis.

Various key concepts, libraries, functions, classifiers, datasets, machine learning models, and techniques were introduced.

Along with technical knowledge, task management and peer learning were also significantly enhanced during the period. Practical experience of live scenarios in the domain of data science and machine learning was also a key byproduct of this program.

III. SKILLS ACQUIRED THROUGH THE INTERNSHIP

- ❖ Python Fundamentals
  - ➢ Strings
  - ➢ Iterators
  - ➢ Data Types
  - ➢ Control Statements
  - ➢ List
  - ➢ Sets
  - ➢ Functions
  - ➢ File Handling
  - ➢ Pandas
  - ➢ Object-Oriented Programming

- ❖ Machine Learning
  - ➢ KNN Classifier
  - ➢ Support Vector Classifier
  - ➢ Logistic regression
  - ➢ Polynomial regression
 - ➢ Dimensionality Reduction
  - ➢ K Means Clustering

- ❖ Data Analysis
  - ➢ Matplotlib
  - ➢ Data Cleansing
  - ➢ Data Preprocessing
  - ➢ Seaborn
    - ■ Joint map
    - ■ Heat map
    - ■ Pair plot
    - ■ Line Plot
    - ■ Distribution Plot
  - ➢ Scikit
  - ➢ Preprocessing
  - ➢ Encoding Techniques
  - ➢ Scaling Techniques

IV.   OVERVIEW OF THE PROJECT/WORK CARRIED OUT DURING INTERNSHIP

A huge amount of data is being generated every second in today's digital world. Such large data becomes very difficult to process using manual checking and traditional methods to sort/understand. So, to handle this, we are developing a Sentiment Analysis. Understanding the sentiment of textual data is a very crucial component in data science.

As feedback plays a crucial role when it is related to e-commerce, products, digital marketing, and so on. Every host wants to improve their product/work, it'll more effective when he/she goes through what their customer wants, they receive the feedback through public and to go through what they want, we have a solution "Sentimental Analysis", that sorts the feedback through the sections (Positive, Negative) and helps them to improve.
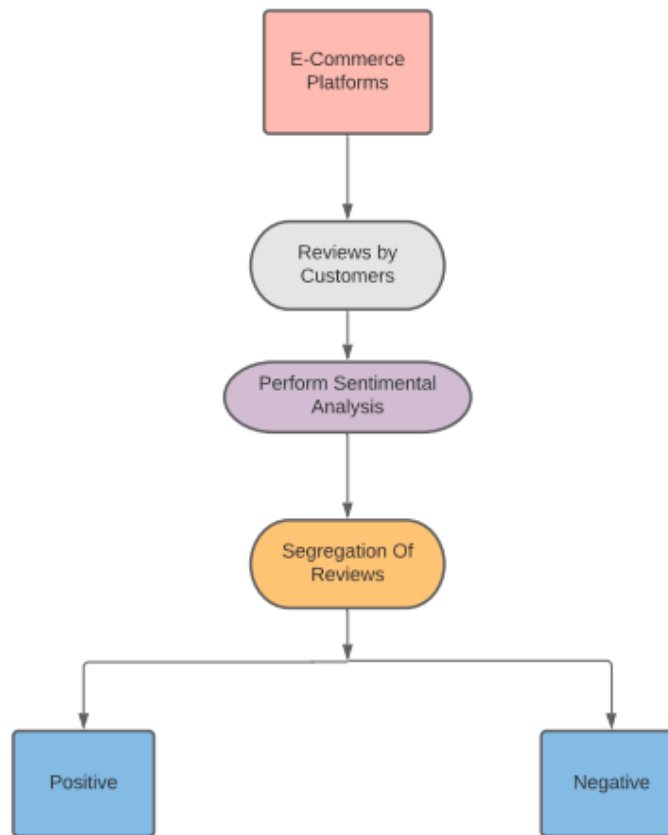
V. RESULTS/OUTPUT

The output of the project is basically the reviews data segregated into two categories namely positive and negative. This will determine the polarity of reviews so it helps for the product improvisation because we can't acknowledge the huge data we can drop it down into these categories so that by the count of these categories we can analyze the reviews. This helps in enhancing the business operation.

Also at the end of the internship period, a meticulous assessment was conducted by APSSDC. It comprised of 50 Questions which were categorized into three divisions namely:

- Python
- Data Analysis
- Machine Learning

The exam was a synthesis of all the key concepts learned during the program. The overall result of this internship was significant technical knowledge enlightenment along with work ethics.

B000LQOCH0	ABXLMWJIXXAIN	Natalia Corres "Natalia Corres"	1219017600	"Delight" says it all	confection around century light pillowy citrus...	Positive
B000UA0QIQ	A395BORC6FGVXV	Karl	1307923200	Cough Medicine	looking secret ingredient robitussin believe f...	Positive
B006K2ZZ7K	A1UQRSCLF8GW1T	Michael D. Bigham "M. Wassir"	1350777600	Great taffy	great taffy great price wide assortment yummy ...	Positive

VI.   CONCLUSION

It was indeed a very informative and enlightening experience. Many new concepts and skills were acquired during the span of this internship. The live problem-solving tasks. It was a great professional and training experience which undoubtedly piqued my interest in the domain of Data Science.

Will definitely strive to use gained skills and knowledge in the best possible way, to improvise and attain the desired career objectives.

VII.   REFERENCES (If Any)

- Singla, Z., Randhawa, S., & Jain, S. (2017). Sentiment analysis of customer product reviews using Machine Learning. 2017 International Conference on Intelligent Computing and Control (I2C2). https://doi.org/10.1109/i2c2.2017.8321910

- Sklearn naive Bayes classifier python: Gaussian naive Bayes Scikit-Learn tutorial. DataCamp Community. (n.d.). Retrieved September 20, 2021,from https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn.

- 1.4. support vector machines. scikit. (n.d.). Retrieved September 20, 2021, from https://scikit-learn.org/stable/modules/svm.html.

- Michael J. Greis CFA, Michael, J., Greis, & amp; Cfa. (2018, October 25). Mainstreaming Sustainable Investing. CFA Institute. Retrieved September 20, 2021,from https://www.cfainstitute.org/en/research/foundation/2018/mainstreaming-sustainable-investing?s_cid=ppc_RF_Google_Search_MainstreamingSustainableInvesting.