INTERNSHIP REPORT

***A report submitted in partial fulfillment of the requirements for 3rd Year***

#### BACHELOR OF ENGINEERING

**in**

#### ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

**By**

**Mr. Devrat Yashraj Deepak**

#### Exam Seat. No.: T190352011

#### Under Supervision of Digambar Padulkar,

**At Lets Grow More.**

#### (Duration: Start date [01/03/2023] to End date [1/04/2023])

****

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE VIDYA PRATISHTHAN’S KAMALNAYAN BAJAJ INSTITUTE OF ENGINEERING**

#### AND TECHNOLOGY,

#### BARAMATI.

**ACADEMIC YEAR: 2022-23**

#### DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

#### VIDYA PRATISHTHAN’S KAMALNAYAN BAJAJ INSTITUTE

**OF ENGINEERING AND TECHNOLOGY**

#### BARAMATI.

****

**CERTIFICATE**

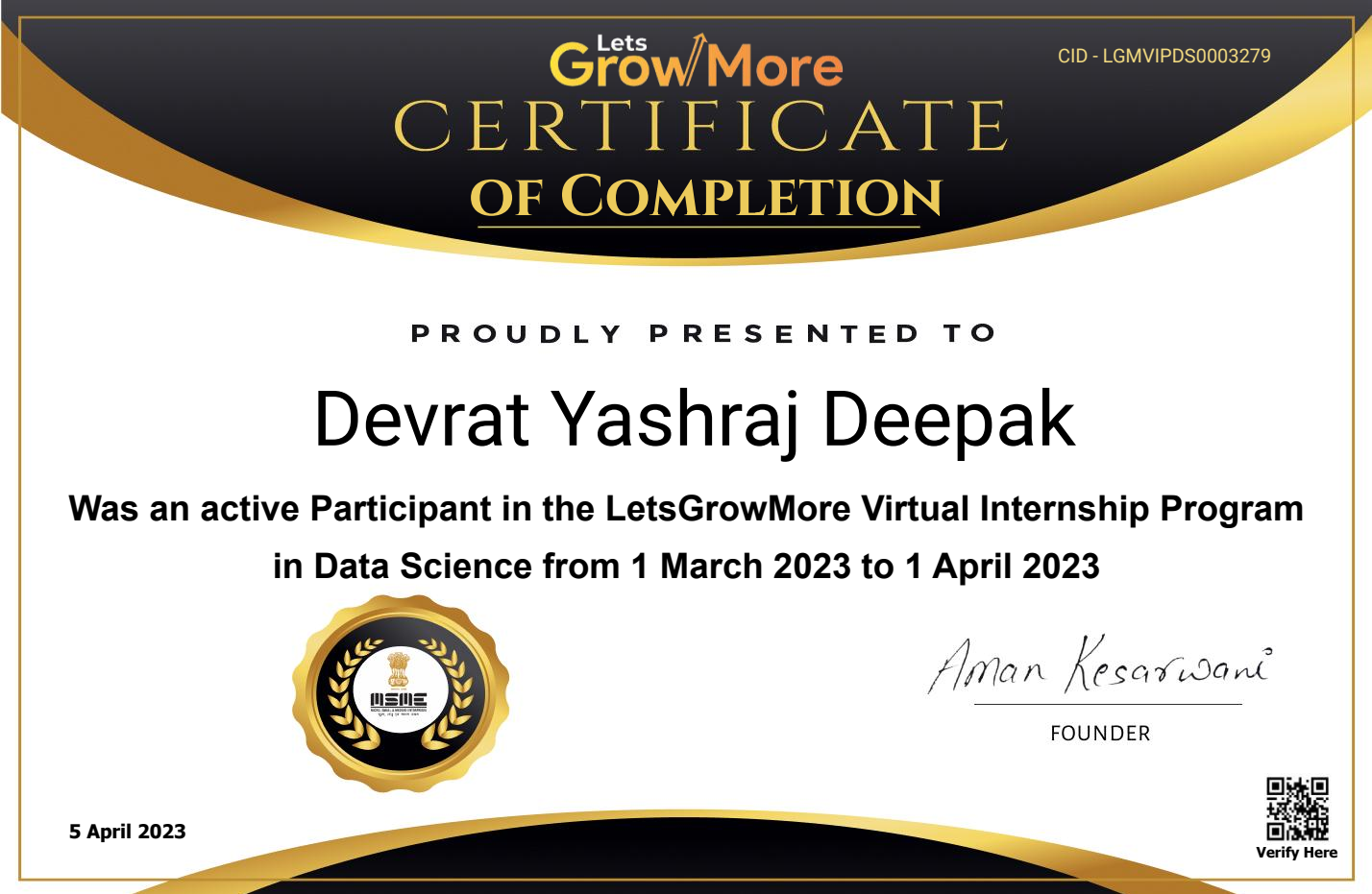
This is to certify that the “**Internship report”** submitted by **Mr. Devrat Yashraj Deepak ( Exam Seat**

**No. : T190352011 )** is work done by his **Online** Mode and submitted during the **2022–2023** academic year, in partial fulfillment of the requirements for the **3rd Year** of **BACHELOR OF ENGINEERING in ARTIFICIAL INTELLIGENCE AND DATA SCIENCE,** at LETS GROW MORE**.**

|  |  |  |
| --- | --- | --- |
| **Digambar Padulkar** | **Mr. P. P. Ghorpade** | **Dr. P. M.Paithane** |
| **Internship Mentor** | **Dept Internship Coordinator** | **HOD, AI & DS** |

**Dr. D. B. Hanchate Dr. R.S. Bichkar Dean IIIC Principal**

# CERTIFICATE OF INTERNSHIP

****

iii

## ACKNOWLEDGEMENT

First, I would like to thank **LETS GROW MORE** for giving me the opportunity to do an internship within the organization.

I also would like to thank all the people that worked along with me **LETS GROW MORE** with their patience and openness they created an enjoyable working environment.

It is indeed with a great sense of pleasure and immense sense of gratitude that I acknowledge the help of these individuals.

I am highly indebted to Principal **Dr. R. S. Bichkar**, for the facilities provided to accomplish

this internship.

I would like to thank the Head of the Department **Dr. P. M. Paithane** for his constructive encouragement throughout my internship.

I would like to thank **Dr. D. B. Hanchate,** Dean IIIC, and **Mr. P. P. Ghorpade,** internship coordinator Department of **Artificial Intelligence and Data Science** for their support and advice to get and complete the internship in above said organization.

I am extremely great full to my mentor **Digambar Padulkar** and department staff members and friends who helped me in the successful completion of this internship.

**Mr. DEVRAT YASHRAJ DEEPAK**

**(Exam Seat No: T190352011)**

## ABSTRACT

Data science is an interdisciplinary field that involves the extraction of insights and knowledge from large and complex data sets using various techniques and tools. It combines elements of computer science, mathematics, and statistics, as well as domain-specific knowledge from various industries.

In summary, the data science internship at Lets Grow More offers an excellent opportunity for aspiring data scientists to gain hands-on experience, develop practical skills, and learn best practices in the field. The internship program provides a platform for the intern to contribute to the success of the company while also enhancing their professional development.

## Keywords:

Data science, Data visualization, Data analysis, Machine learning.

# Organization Information

Lets Grow More is an registered start-up which develops personalized software Throughout the software development, Lets Grow More is committed to producing exceptional software for each of the clients to ensure that they are aligned with the end goal. The team leverages state-of-the-art tools and techniques to ensure that the solutions they develop are accurate, reliable, and scalable.

An Lets Grow More believes in making our youth especially the students self-aware and exploring the untouched world of technology and tremendous growth-making fields. Lets Grow More believes in bridging the gap between students and their knowledge in the industrial field & bringing them a step closer to their dream. Lets Grow More is a registered start-up, always open to connecting new people with us over this journey of endless learning and great joy!

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S.N.** | **Contents** | **Page No.** |
| **1.** | Introduction | 9 |
| **2.** | Learning Objectives/Internship Objectives | 10 |
| **3.** | Motivation/Scope and Rationale of The Study | 11 |
| **4.** | Weekly Overview of Internship Activities | 12 |
| **5.** | About Industry/ Organization | 14 |
| **6.** | Software Requirements Specifications | 16 |
| **7.** | Technology | 17 |
| **8.** | Screenshots | 19 |
| **9.** | Results/Analysis/Inferences | 22 |
| **10.** | Conclusion And Future Scope | 24 |
| **11.** | Attendance Record | 25 |
| **12.** | Documents Related to Internship | 26 |
| **13.** | Bibliography | 29 |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page No.** |
| Fig. 1.1 | Confusion Matrix | 17 |
| Fig 1.2 | Task1\_Iris Flower Classification | 19 |
| Fig 1.3 | Task2\_Stock Market Analysis using LSTM Network | 20 |
| Fig 1.4 | Task3\_Exploratory Data Analysis on Terrorism dataset | 21 |
| Fig 1.5 | Offer Letter | 27 |
| Fig 1.6 | Gmail Offer Letter | 28 |
| Fig 1.7 | VIP Badge Selection Gmail | 29 |
| Fig 1.8 | Completion Certificate Gmail | 30 |

## INTRODUCTION

Lets Grow More INTERNSHIP

This internship was all about data science and analytics. I applied for this internship

from the website of Lets Grow More. I completed 3 tasks out of 5 tasks. The description of all the tasks is given below:

## Task-1:

IRIS FLOWER CLASSIFICATION - I predicted the species of IRIS Flower using the

attributes given such as petal length, sepal length, petal width and sepal width. I used Machine Learning algorithm for the classification of species.

## Task-2:

Stock Market Analysis using LSTM network - I explored the dataset by creating various charts and graphs using matplotlib and seaborn libraries of Python. Also predicted output after setting target and found accuracy of the model.

## Task-3:

Exploratory Data Analysis on Terrorism Dataset - I predicted the region which was affected

due to terrorist attacks and also danger zones . Plotting using seaborn library of Python.

## 2.LEARNING OBJECTIVES / INTERNSHIP OBJECTIVES

#### Gaining practical experience:

The internship provides an opportunity to gain practical experience in the field of data science by working on real-world projects and data sets.

#### Applying theoretical knowledge:

Interns can apply the theoretical knowledge gained through academic coursework to real- world scenarios and learn how to solve complex problems using data analysis techniques.

#### Building a professional network:

Internships offer a chance to meet and work with professionals in the field, which can help build a professional network and potentially lead to future job opportunities.

#### Developing new skills:

Interns can learn new skills and techniques that are not typically taught in academic coursework, such as how to work with large data sets or how to use specific tools or software.

#### Increasing employability:

Completing a data science internship can enhance one's employability by demonstrating practical experience and skills in the field, which can make one stand out to potential employers.

## 3. MOTIVATION / SCOPE AND RATIONALE OF THE STUDY

**Motivation/Scope:**

Data science is emerging as a field that is revolutionizing science and industries alike. Work across nearly all domains is becoming more data-driven, affecting both the jobs that are available and the skills that are required. As more data and ways of analyzing them become available, more aspects of the economy, society, and daily life will become dependent on data.

* 1. Internships are helpful for gaining hands-on experience in the domain which we like.
  2. Passionate to learn about data science
  3. Dive deeply to gain more knowledge of the domain
  4. To discover the patterns in data.
  5. To process the collecting, storing, segregating and analyzing data.
  6. Exploring and visualizing data, analysing data and/or applying machine learning to data decisions based on acquired insight.

## Rationale:

The rationale behind my participation in a data science internship is to bridge the gap between academic coursework and practical experience. While academic coursework provides a theoretical foundation in data science, it often lacks the practical application of working with real-world data sets or collaborating with other professionals in the field. Therefore, this internship provides me with the opportunity to apply the theoretical knowledge

I have learned in real-world scenarios, gain hands-on experience, and develop a deeper understanding of the field of data science. Additionally, this internship will enhance my employability by providing me with relevant experience and skills that are in high demand in the job market. Overall, this data science internship will be an excellent opportunity for me to grow both professionally and personally.

## WEEKLY OVERVIEW OF INTERNSHIP ACTIVITIES

**Week 1: Task Name: Study about the domain**

|  |  |  |  |
| --- | --- | --- | --- |
| **1st WEEK** | **DATE** | **DAY** | **NAME OF THE TOPIC/MODULE COMPLETED** |
| 01/03/23 | Wednesday | Learning statistics |
| 02/03/23 | Thursday | Collecting datasets |
| 03/03/23 | Friday | Data cleaning methods |
| 04/03/23 | Saturday | Data analyzing methods |
| 05/03/23 | Monday | Learning machine learning algorithms |
| 06/03/23 | Tuesday | Review |

## Week 2: Task Name: Iris flower classification

|  |  |  |  |
| --- | --- | --- | --- |
| **2nd WEEK** | **DATE** | **DAY** | **NAME OF THE TOPIC/MODULE COMPLETED** |
| 08/03/23 | Wednesday | Govt. Holiday |
| 09/03/23 | Thursday | Data preprocessing and visualization |
| 10/03/23 | Friday | Learning about various ML Algorithms |
| 11/03/23 | Saturday | Implementing ML Algorithm |
| 12/03/23 | Monday | Confusion Matrix |
| 13/03/23 | Tuesday | Uploading task on GitHub |

**Week 3: Task Name: Stock Market Analysis using LSTM Network**

|  |  |  |  |
| --- | --- | --- | --- |
| **3rd WEEK** | **DATE** | **DAY** | **Stock Market Analysis using LSTM Network** |
| 15/03/23 | Wednesday | Load the data , Train and Test dataset |
| 16/03/23 | Thursday | Analyze the dataset |
| 17/03/23 | Friday | Visualize the dataset |
| 19/03/23 | Monday | Prediction using LSTM Network |
| 20/03/23 | Tuesday | Uploading task on GitHub |

## Week 4: Task Name: Exploratory Data Analysis on Terrorism dataset

|  |  |  |  |
| --- | --- | --- | --- |
| **4th WEEK** | **DATE** | **DAY** | **Exploratory Data Analysis on Terrorism dataset** |
| 23/03/23 | Wednesday | Load the data , Train and Test dataset |
| 24/03/23 | Thursday | Plotting using Seaborn library |
| 25/03/23 | Friday | Visualize the dataset |
| 26/03/23 | Saturday | Prediction of models accuracy |
| 27/03/23 | Monday | Plotting again the predicted results |
| 28/03/23 | Tuesday | Uploading task on GitHub |

* + 1. **ABOUT INDUSTRY/ ORGANIZATION**

## Company background-organization and activities

At Lets Grow More believes in making our youth especially the students self-aware and exploring the untouched world of technology and tremendous growth-making fields. Lets Grow More believes in bridging the gap between students and their knowledge in the industrial field & bringing them a step closer to their dream. Lets Grow More is a registered start-up, always open to connecting new people with us over this journey of endless learning and great joy!

The organization turns the digital dreams of the clients into a reality. The organization works closely with the users throughout development to ensure that organization is still aligned with the end goal. organization is committed to producing exceptional software for each of its clients.

## Services :

### Website Development

Our website developers provide expert web application development and web design services to our clients. Lets Grow More offers a variety of website design and development services, We specialize in developing interactive, scalable, brand-oriented, and business-ready custom web solutions. Our aim is to build profitable digital products engineered to spark creativity and boost ROI.

### Cloud Computing

Whatever type of cloud computing services you use, one thing is certain: large quantities of data will move back and forth between your end users and the cloud provider's data centers, over the internet. That's why whatever type of cloud services you use, Lets Grow More can help your organization have a better cloud experience and reap greater benefits from your cloud services investment.

### Graphic Design

Our creative designers offer outstanding and attention-grabbing visuals for brochures, banners, flyers , logos, business cards, e-books, and many more. Our graphic designs can develop a great impression on your prospect’s mind. With a complete mixture of professionalism and stunning creativity skills, our expert designers can establish a bright profile for your business in the online world.

### Animation Design

Animations make your brand more engaging and interactive. This makes them great assets to share on social media, distribute through email, feature internally and host on your website— they’re that versatile. Our passionate team of motion graphics designers empower you to tell great stories for your business. We believe it's only when you tell great stories you become one.

### Database Management

### Your databases are at the core of your business. Whether you’re operating on-premises or in the cloud, using a traditional or open-source platform, our global experts are available around the clock to help you turn your databases into a business differentiator. We offer a full range of solutions to support your needs—from assuming total ownership of your database environment to augmenting your internal team to helping you refactor for the cloud.

### Machine Learning Solutions

Artificial Intelligence & Machine Learning offerings help organizations build highly- customized solutions running on advanced machine learning algorithms. Our machine learning services offer advanced algorithms to help organizations in solving key business challenges, enabling data-driven decision-making and creating innovative business models.

## Scope and Object of the Study

### To discover the patterns in data.

* To process the collecting, storing, segregating and analyzing data.
* Exploring and visualizing data, analysing data and/or applying machine learning to data decisions based on acquired insight.

## Internship Place Details-

Internship Mode: Online

Address: Allahabad – 211001 , Uttar Pradesh

Contact : [Contact@letsgrowmore.in](mailto:Contact@letsgrowmore.in)

## SOFTWARE REQUIREMENTS SPECIFICATIONS

**Software Requirement: -**

|  |  |
| --- | --- |
| Operating system | Linux |
| Python interpreter | IDLE python |
| Application | Jupyter Note- book |
| Programming Language | Python |

## Hardware Requirement: -

|  |  |
| --- | --- |
| Processor | 2.0GHz |
| RAM | 4 GB(minimum) |
| Hard Disk | 2 GB |

**Dataset Used: -**

|  |  |
| --- | --- |
| TASK-1 | Iris.csv |
| TASK-2 | stockmarket.csv |
| TASK-3 | Terrorism.csv |

## TECHNOLOGY

#### LOGISTIC REGRESSION:-

Logistic regression is a data analysis technique that uses mathematics to find the relationships between two data factors. It then uses this relationship to predict the value of one of those factors based on the other. The prediction usually has a finite number of outcomes, like yes or no.

#### DECISION TREE ALGORITHM :-

#### Decision Tree Algorithm is a supervised algorithm technique that can be used for both classification and regression purpose as well. It is a hierarchical , tree , structure , which consist of a root node, branches , internal nodes and leaf nodes.

**CONFUSION MATRIX:-**

The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known. The matrix itself can be easily understood, but the related terminologies may be confusing. Since it shows the errors in the model performance in the form of a matrix, hence also known as an error matrix.

|  |  |  |
| --- | --- | --- |
| n = total predictions | Actual: No | Actual: Yes |
| Predicate: No | True Negative | False Positive |
| Predicate: Yes | False Negative | True Positive |

Fig1.1 :- Confusion Matrix

**LINEAR REGRESSION: -**

Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression. Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc.

#### **LOAD THE DATA , TRAIN AND TEST DATASET**: -

Loading dataset properly is also necessary so that it should not give problem later . After loading dataset I trained the dataset using Long-Short Term Memory network . At last I tested it by splitting the dataset into training and testing and predicted the accuracy score.

**ANALYZE THE DATASET** : -

Data analysis refers to the process of manipulating raw data to uncover useful

insights and draw conclusions. So analyzation of dataset and understanding the insights from it.

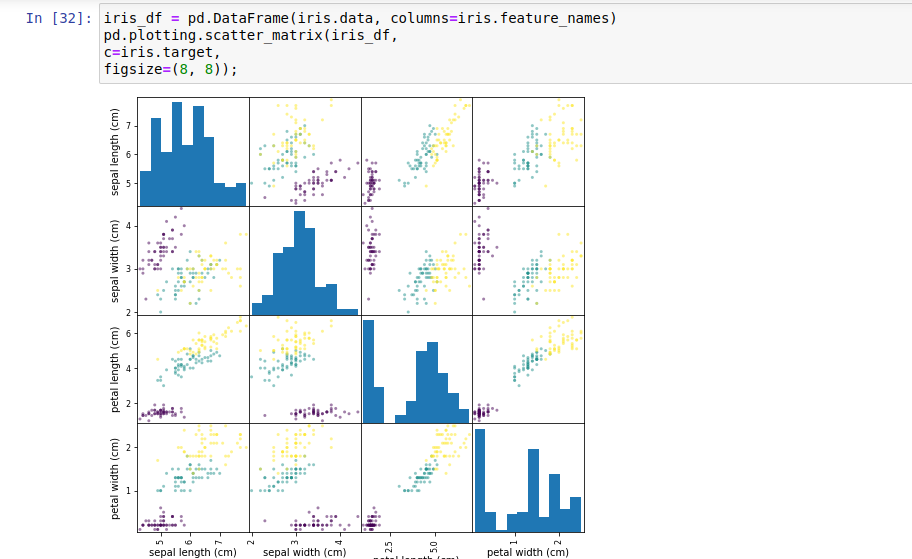
#### VISUALIZE THE DATASET: -

After analysing the dataset finding of correlation and trends in the dataset . And then representing it pictographically using seaborn , matplotlib library.

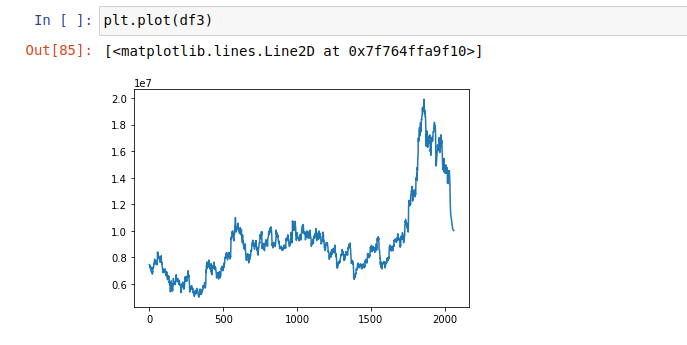
#### PREDICTION USING LSTM NETWORK: -

At the end we need to do train the model and predict our results. Use of LSTM network in this model to train the network .

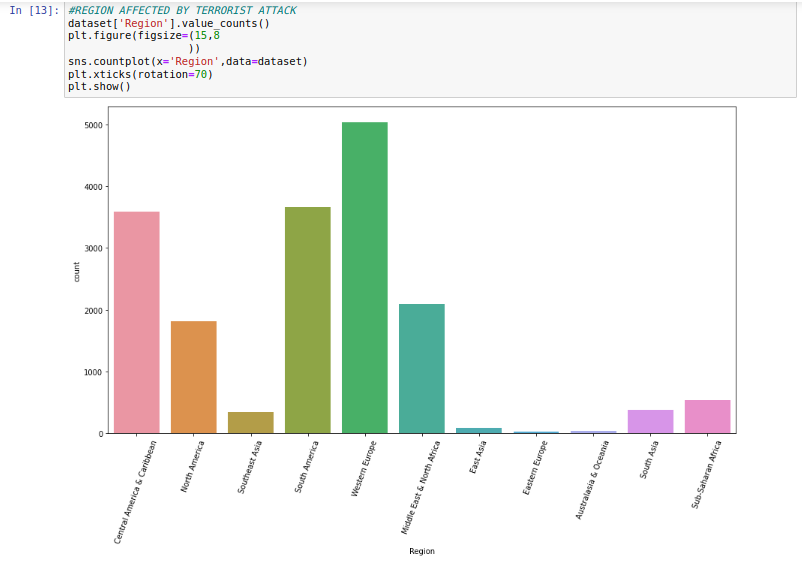
## SCREENSHOTS

****

#### Fig 1.2 :-Task1\_Iris Flower Classification



**Fig 1.3 :-Task2\_Stock market prediction using LSTM Network**



**Fig 1.4 :-Task3\_Exploratory Data Analysis on Terrorism dataset**

## RESULTS / ANALYSIS / INFERENCES

**Results :**

Accuracy: - Task 1: The accuracy of the model is 97.37 %. Accuracy: - Task 2: The accuracy of the model is 90 %.

Accuracy: - Task 3: The accuracy of the model is 82 %.

## Analysis :

* Data Preparation.
* Feature Selection.
* Model Selection.
* Model Training and Evaluation.

## Inferences :

**Task Name : Iris flower classification.**

1. Machine learning algorithms can accurately classify iris flower species based on their characteristics. The selected algorithm's accuracy depends on various factors such as the quality of the data, feature selection, model selection, and hyper parameter tuning.
2. The four features of the iris flower data set - sepal length, sepal width, petal length, and petal width - are the most significant factors in distinguishing between the three species. These features can help in accurately classifying the iris flowers, and similar features may be used in other classification problems.
3. The accuracy of the model can be improved by selecting the appropriate machine learning algorithm, optimizing the hyper parameters, and increasing the size of the data set.
4. The results of the analysis can be useful in various real-world applications, such as identifying plant species in agriculture, identifying disease types in medical diagnosis, and detecting fraud in finance.

5. The iris flower data set is a widely used data set in machine learning, and the analysis of this dataset can serve as a basis for understanding and analyzing other classification problems.

**Task Name : Stock market prediction using LSTM Network**

* 1. Machine learning algorithms LSTM Network can do effective analysis of stock market dataset. The selected algorithm's performance depends on the quality of the data, the features selected, and the model parameters.
  2. The analysis can help identify the key factors that contribute to unemployment rates, such as economic growth, population demographics, education levels, and industry trends.
  3. LSTM Network can be used to predict good accuracy , so that user can gain insights and understand variations of rise and fall in the market.
  4. The results of the analysis can be useful in various real-world applications, such as public policy planning, and investment decision-making.
  5. Machine learning can help to develop more accurate and reliable models fo r predicting results . Using graph one can easily understand the uprise and downfall of share market or stock market.

**Task Name : Exploratory Data Analysis on Terrorism dataset**

1. Most of the attacks were attacked through explosives and then through firearms. Attacks were more during 2014 and then in 2015. When compared to attacks from 1970 onwards , the last 6 years scored a maximum. But from 2014 onwards count started decreasing.
2. Almost Every day has the same contribution but attacks were low during 31st and high during 15th and 1st. Iraq dominates all the countries and it has the highest number of attacks and then Pakistan, Afghanistan, and India follow it.
3. The analysis can help identify the key factors that contribute to terrorist attacks, higher chances of terrorist attacks, Zones , external conditions.
4. Machine learning can be used to identify patterns and trends in dataset, providing insights into the long-term effects of attacks.

## 10.CONCLUSION AND FUTURE SCOPE

#### CONCLUSION :

During the internship, I worked on various projects such as iris flower classification, stock market analysis, and exploratory analysis on terrorism dataset. The analysis of these projects highlighted the effectiveness of machine learning algorithms in solving complex problems and generating insights. The internship helped in developing skills such as data wrangling, data Preprocessing , feature engineering, model selection, and evaluation. Additionally, the internship also emphasized the importance of data quality, data ethics, and data privacy in the field of data science. The internship provided an opportunity to work on real-world projects, allowing us to apply the concepts and techniques learned in the classroom to real-world problems.

#### FUTURE SCOPE :

1. Continued growth in the field of data science: The field of data science is continuously evolving, with new tools, techniques, and technologies being developed. The internship can help individuals stay up-to-date with the latest trends and technologies in the field.
2. Increased demand for data scientists: With the increasing amount of data being generated in various industries, the demand for data scientists is also increasing. The internship can help individuals develop the skills and experience needed to pursue a career in data science.
3. Opportunities for further learning: The internship can provide a foundation for further learning in data science, such as pursuing a master's degree or other advanced training programs.

## 11. ATTENDANCE RECORD

#### Name & Address of Organization: -

Vidya Pratishthan’s Kamalnayan Bajaj Institute of Engineering and Technology Baramati -413102

|  |  |  |
| --- | --- | --- |
| Name of Student | Mr. Devrat Yashraj Deepak | |
| Roll. No | 2237031(**T190352011**) | |
| Name of Course | TE - Artificial Intelligence & Data Science | |
| Date of Commencement of Training.: | | 01/03/2023 |
| Date of Completion of Training: | | 1/04/2023 |

#### Initials of the student:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Month & Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Mar- 2023 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**Note:**

1. The attendance Sheet should remain affixed in Daily Training Diary. **Do not remove or tear it off.**
2. Students should sign/initial in the attendance column. Do not mark ‘P’
3. Holidays should be marked in **red ink** in the attendance column. Absent should be marked as ‘**A’ in Red Ink.**

**Signature of Company internship supervisor /**

**Dept. Mentor with company stamp/ seal**

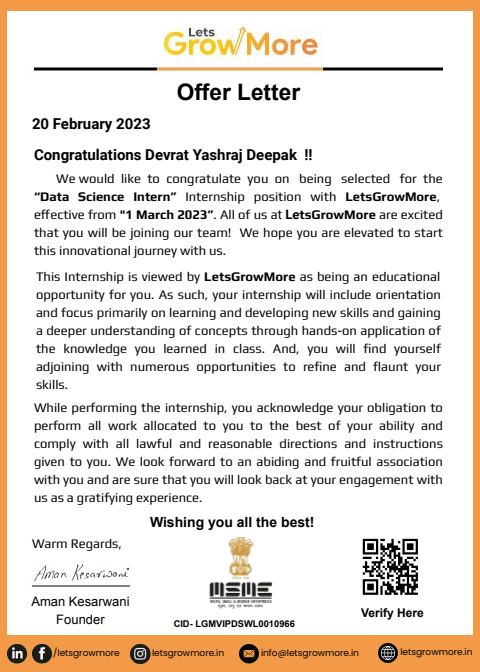
**(Name ) Contact No.**

## 12. DOCUMENTS RELATED TO INTERNSHIP

* 1. **Offer Letter:**

**Lets Grow More**

INTERNSHIP OFFER LETTER



**OIB/M1/IP8571**

**Dear Devrat Yashraj Deepak**

We are delighted to welcome you for the internship of **Data Science** in our fellowship. This internship is observed by **Lets Grow More** as being a learning opportunity for you.

In essence, your internship will embrace orientation and give emphasis on learning new skills with a deeper understanding of concepts through hands-on application of the knowledge you gained as an intern. Our team is confident that you will acknowledge your obligation to perform all work allocated to you to the best of your ability within lawful and reasonable direction given to you.

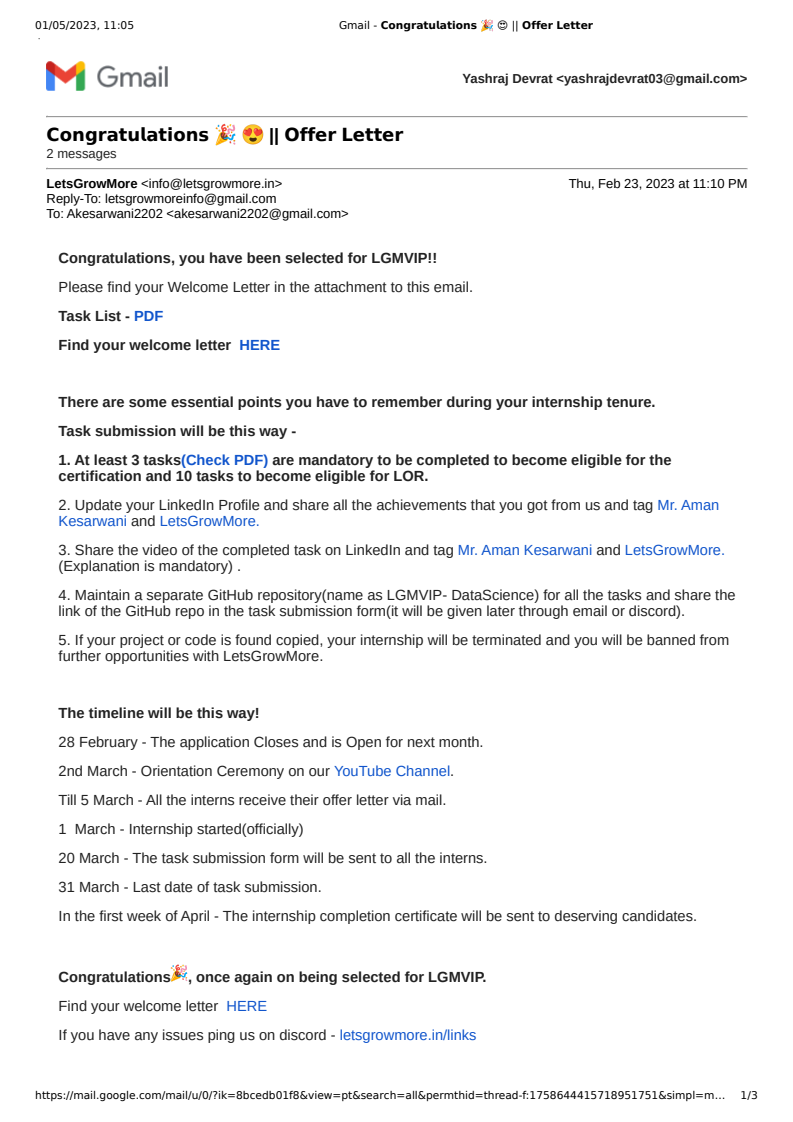
We look forward to a worthwhile and fruitful association which will make you equipped for future projects. wishing you the most enjoyable and truly meaningful internship program experience.

Thank you

Team Lets Grow More

**Fig 1.5: -Offer Letter**

## Gmail of Offer Letter :

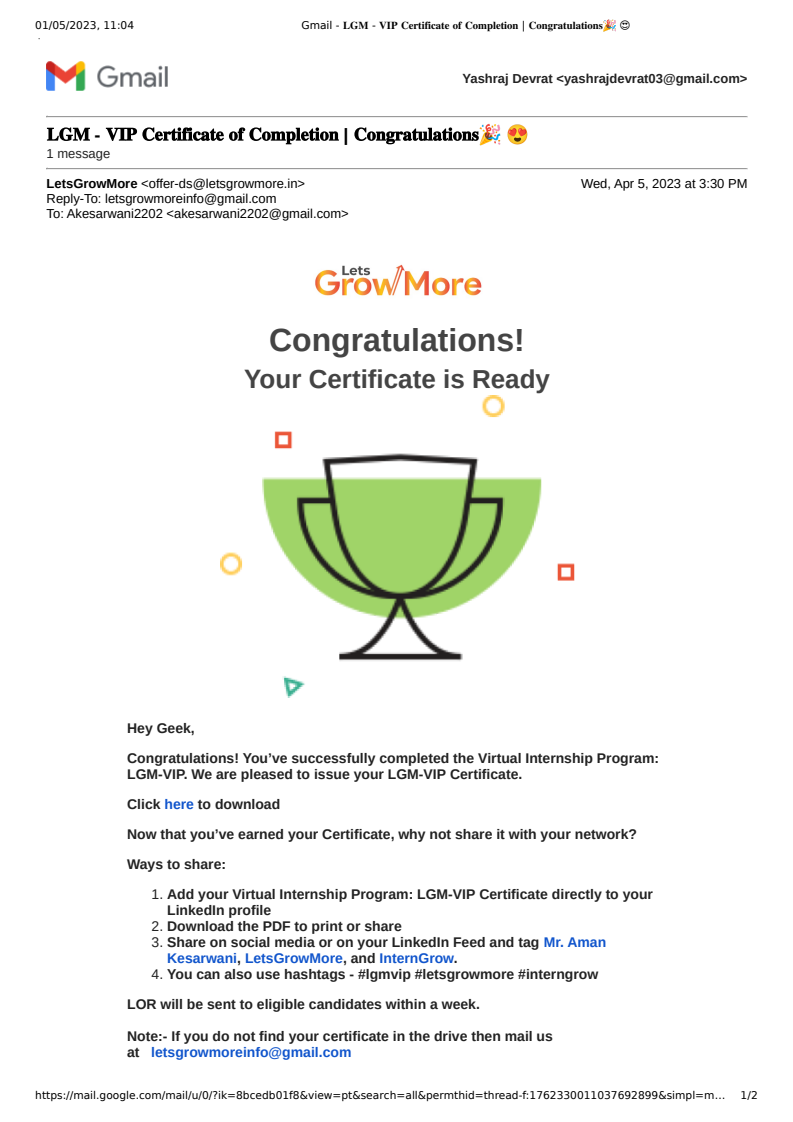


**Fig 1.6 : - Gmail Offer Letter**

## VIP Badge of Selection Gmail :

****

## **4.**Completion Certificate Gmail :

**Fig 1.8 : - Completion Certificate Gmail**

## 13. BIBLOGRAPHY

* + 1. https://data-flair.training/blogs/iris-flower-classification/
    2. https://[www.analyticsvidhya.com/blog/2020/08/building-sales-prediction-web-application-](http://www.analyticsvidhya.com/blog/2020/08/building-sales-prediction-web-application-) using-machine-learning-dataset/
    3. https://thecleverprogrammer.com/2021/07/12/unemployment-analysis-with-python/
    4. Python for Data Analysis-Book
    5. Data Science and Big Data Analytics-Book