MAJOR PROJECT REPORT
MID-TERM

Department of Artificial Intelligence and Machine Learning

Submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF TECHNOLOGY (B. Tech)

in

Computer Science and Engineering (Al and ML)

by

UTKARSH PRATAP SINGH

Reg No: 219310098

submitted to

MANIPAL UNIVERSITY
JAIPUR
(University under Section 2(f) of the UGC Act)

DEPARTMENT OF AI and ML,
SCHOOL OF COMPUTER SCIENCE ENGINEERING,
MANIPAL UNIVERSITY JAIPUR,
RAJASTHAN, INDIA-303007

Acknowledgment of Participation

OFFER LETTER

DocuSign Envelope ID: AFF80195-BEC9-4528-BBA6-D679AF211723

GEMINI SOLUTIONS
VISUALIZE SUCCESS

11th January 2025

Dear **Utkarsh Pratap Singh,**

We are pleased to offer you the post of **Software Developer Trainee** with Gemini Solutions Private Limited as of **11th January 2025** and you will report to us on **20th January 2025** post which we shall make your final assignments. In addition to the duties that were outlined during your interviews, you shall perform such other duties as are customarily associated with such position and as the Board of Directors, or its authorized representative may from time to time require. You shall devote your full business efforts and time to the company and agree to perform your duties faithfully and to the best of your ability. **You agree not to engage in any other employment, occupation, or consulting activity for any direct or indirect remuneration without the prior approval of your supervisor.**

To address and appropriately deal with any instances or behavior that comes under following clauses, Zero Tolerance Policy (ZTP) (which includes immediate termination) is applicable to all employees of Gemini Solutions:

a) Demonstrate non-professional behavior/attitude towards customer/ clients.
b) Usage of profane, vulgar, or abusive language.
c) Offensive or inappropriate references to Race, Ethnicity, Religion, Gender, Lifestyle, Sexual Orientation, Disability and Age.
d) Consistent interruptions in a rude and sarcastic manner with client.
e) Unauthorized release of confidential information.
f) Inappropriate fraternization with clients.
g) Coming to work intoxicated or any disciplinary issues on the floor.
h) Extended unscheduled breaks, unapproved/ unscheduled leaves, reporting late to work etc.
i) Disparaging remarks about Gemini Solutions or Clients.

If the ZTP Committee decided to terminate the concerned employee immediately, then the standard full and final procedures are followed after the due consideration depending on seriousness of issue.

You understand and agree that neither your job performance nor promotions, commendations, bonuses or the like from the Company give rise to or in any way serve as the basis for modification, amendment, or extension, by implication or otherwise, of your employment with the Company.

You will be compensated **Rs.15,500 (Fifteen Thousand Five Hundred only)** per month for the work performed during the training. You will be paid in accordance with the company's normal payroll practices and be subject to the usual, required withholding.

While employed hereunder, you will not be entitled to participate in the employee benefit plans maintained by the Company, if any. You will be eligible for paid leave and paid holidays only as approved by your supervisor. The company will provide you with free Lunch and refreshments as additional benefit.

UPS

Registered Office: H-324, Gali No.-6, Raj Nagar-II, Palam Colony South West Delhi-110077
Principal Place of Business: Plot no-119, Udyog Vihar Phase-I, Sector-20, Gurugram, Haryana-122016
Branch Office: 1st Floor, Plot No 5, IT Park, Sector-22, Panchkula, Haryana 134112
CIN No: U72900DL2014PTC344614, **Phone No:** +91 124 6266300, **Email:** info@geminisolutions.in

DocuSign Envelope ID: AFF80195-BEC9-4528-BBA6-D679AF211723

GEMINI SOLUTIONS
VISUALIZE SUCCESS

The Company will also reimburse you for reasonable and documented travel, entertainment or other expenses incurred by you in the furtherance of or in connection with the performance of your duties hereunder, in accordance with the Company's expense reimbursement policy as in effect from time to time. The company will invest a lot of time, money, and energy in training you, so it is mandatory for you to complete the training.

You acknowledge and agree that you are executing this letter voluntarily and without any duress or undue influence by the Company or anyone else. You further acknowledge and agree that you have carefully read this letter and that you understand the terms, consequences, and binding effect of this letter.

This letter, together with any agreement you enter with the Company represents the entire agreement and understanding between the parties as to the subject matter herein and supersedes all prior or contemporaneous agreements, whether written or oral. If any provision hereof becomes or is declared by a court of competent jurisdiction to be illegal, unenforceable, or void, this letter will continue in full force and effect without said provision.

No waiver, alteration, or modification of any of the provisions of this Agreement will be binding unless in writing and signed by duly authorized representatives of the parties hereto. The internal substantive laws, but not the choice of law rules, of the State of India, shall govern this letter.

If you choose to accept this offer, please sign a copy of this letter in the space. We hope that this offer will be favorably received, and we look forward to working with you at Gemini Solutions Private Limited.

Sincerely,

Priyanka Gubrele
Vice President - Human Resource
Gemini Solutions Private Limited

ACCEPTED AND AGREED:

Utkarsh Pratap Singh

Signature: _____

Date: 11/01/2025

Registered Office: H-324, Gali No.-6, Raj Nagar-II, Palam Colony South West Delhi-110077
Principal Place of Business: Plot no-119, Udyog Vihar Phase-I, Sector-20, Gurugram, Haryana-122016
Branch Office: 1st Floor, Plot No 5, IT Park, Sector-22, Panchkula, Haryana 134112
CIN No: U72900DL2014PTC344614, **Phone No:** +91 124 6266300, **Email:** info@geminisolutions.in

DECLARATION

I hereby declare that the Major Project Report in Gemini Solutions Pvt Ltd as a Data Science Intern is an authentic record of my own work, during the period from 11th January 2025 to 30th June 2025, for the award of the degree of B.Tech in Computer Science and Engineering (Artificial Intelligence and Machine Learning), Manipal University Jaipur, Rajasthan.

Utkarsh

219310098

Date:
25/03
/2025

Certified that the above statement made by the student is correct to the best of our knowledge and belief.

Examined by:

Mr. Head of Department

ACKNOWLEDGEMENT

First and foremost, I express my sincere gratitude to Gemini Solutions Pvt Ltd for providing me with the incredible opportunity to undergo industrial training in Data Science. I thank the entire Gemini Solutions team for organizing such an enriching and insightful program.

I am deeply grateful to Manipal University Jaipur for facilitating my enrollment in this program. I would also like to sincerely thank Dr. Puneet Mittal (Head of the Department of Artificial Intelligence Machine Learning) for her constant support and guidance throughout my training. I am immensely thankful to all faculty members of our department for their cooperation and encouragement.

Finally, I would like to acknowledge the invaluable input from my peers and mentors during this training. Their thoughtful insights and suggestions greatly contributed to my successful completion of this program.

Utkarsh Pratap
Singh
219310098

COMPANY PROFILE

0.1 Company Profile: Gemini Solutions

Gemini Solutions is a cutting-edge technology company specializing in data analytics and business intelligence solutions. Founded with the mission to empower organizations through data-driven decision-making, Gemini Solutions provides tools and services that help businesses transform in the world of automation and artificial intelligence. By combining advanced analytics, machine learning, and user-friendly platforms, Gemini Solutions enables clients to optimize their operations, enhance customer experiences, and drive growth.

Core Services:

- **Data Analytics & Insights**: Delivering comprehensive data analysis to provide businesses with valuable insights for better decision-making.
- **Business Intelligence Solutions**: Developing customizable BI platform that help organizations visualize, interpret, and act on their data.
- **Software Development**: Offering software solutions tailored to the specific needs of clients, including automation tools, dashboards, and reporting systems.
- **Machine Learning & AI**: Implementing advanced AI algorithms to enable predictive analytics.

Gemini Solutions is a dynamic company that specializes in delivering data-driven solutions aimed at improving business decision-making through the power of analytics, business intelligence, and data science. The company's core focus lies in helping organizations unlock the potential of their data by providing cutting-edge tools and services for data visualization, data processing, and predictive analytics. By leveraging platforms such as Power BI and other advanced data analytics tools, Gemini Solutions enables clients to transform raw data into actionable insights, driving more informed strategic decisions.

Contents

ACKNOWLEDGEMENT iv

COMPANY PROFILE v

0.1 Company Profile: Gemini Solutions v
. .

LIST OF FIGURES vii

1 Introduction **1**

2 Formal Training Provided **3**

2.1 Formal Training Provided 3

2.1.1 Key Training Features 3

2.1.2 Key Training Modules 4

3 Industrial Training **5**

3.1 Objectives 5

3.2 Tools and Technology Used 5

3.3 Projects Completed. 6

3.4 Highlights of Training Exposure 7

**4 Future Scope. . . . .

List of Figures

1 Received offer letter from Gemini Solutions i

Chapter 1

Introduction

This report provides an overview of the projects I worked on during my internship at **Gemini Solutions**, where I had the opportunity to apply data analytics, business intelligence, and architectural design techniques. The projects not only enhanced my technical proficiency but also offered valuable experience in solving real-world business challenges. My work primarily involved analyzing datasets, designing data architectures, and visualizing insights to support business decisions.

The following key projects were the focus of my internship:

Automated Working Status:

Developing individual python scripts for 9 POC's to automatically test their working functionalities on a daily basis.

Central Repository:

Developed a StreamLit application to host the details of all the POC's in the Data Science Department.

Email Notification:

A python script to check the working status of the links of all the deployed POC's.

Set a cron job to schedule this script twice daily.

These projects provided me with hands-on experience in applying data science techniques, architecture design principles, and python development.

Chapter 2

Formal Training Provided

2.1 Formal Training Provided

The training program at **Gemini Solutions** was designed to provide a balanced mix of theoretical knowledge and hands-on experience. Below are the key features of the training:

The training program at Gemini Solutions was designed to provide a balanced mix of theoretical knowledge and hands-on experience, integrating real-world project work to reinforce learning. Below are the key features of the training:

Key Features of the Training

- **Quizzes and Hands-on Labs:** Each training module included conceptual quizzes and practical labs. The quizzes reinforced theoretical knowledge, while the labs focused on real-world applications. For instance, I developed Python scripts for automating the working status of 9 Proof of Concepts (POCs) within the Data Science department. These scripts were designed to automatically test POC functionalities daily, applying the concepts I learned in automation and Python programming.

- **Weekly Masterclasses:** Regular masterclasses led by industry experts helped clarify complex concepts. These sessions focused on practical applications in data science, automation, and architecture design. One of the sessions directly supported my work on creating a **StreamLit application** to host the details of all POCs in the department, giving me insights into building user-friendly data dashboards.

- **Project-based Learning:** The training was project-oriented, with each module tied to real tasks. This hands-on approach helped bridge the gap between theory and practice. For example, I applied my knowledge of Python for automating daily checks on POCs by creating a script that would send **email notifications** about the working status of deployed POCs. A **cron job** was scheduled to run the script twice daily, enhancing operational efficiency.

- **Progress Tracking:** Progress was monitored using a dashboard that tracked completion rates, quiz scores, and performance in practical labs. This helped me stay on top of my development and focus on areas needing improvement, such as refining my understanding of automation workflows and system scheduling.

Key Training Modules

The formal training was divided into several focused modules, equipping me with both theoretical

knowledge and practical skills. Below are the key modules relevant to my work:

1. **Python for Data Engineering and Automation**

This module focused on Python programming with a particular emphasis on automation and workflow management. I learned to work with libraries like Pandas, NumPy, and scheduling tools like **cron jobs**. This module was crucial for the **Automated Working Status** project, where I developed Python scripts to automatically test the working functionalities of 9 POCs. Additionally, I created an **email notification system** to inform stakeholders about the working status of these POCs, making the process more efficient and hands-free.

2. **StreamLit for Application Development**

This module introduced me to **StreamLit**, an open-source framework for building data applications. I used the skills from this module to create a **central repository** using StreamLit, which hosted and displayed the details of all POCs within the Data Science department. This application provided a centralized interface for the team to access and monitor all POCs, improving visibility and tracking.

3. **Excel and Power BI for Data Analysis and Visualization**

Although my primary focus was on automation and application development, I also learned to use **Excel** and **Power BI** for data visualization and analysis. The practical skills gained here were directly applied to the **Form Response Analysis** project, where I created dashboards to visualize response trends. These tools also helped in tracking the daily performance of POCs and generating reports to inform stakeholders about their operational status.

Conclusion

The training program at Gemini Solutions not only enhanced my theoretical knowledge but also provided me with practical experience through projects like the **Automated Working Status**, **StreamLit Application for Central Repository.These projects helped me develop a comprehensive understanding of data science, automation, and application development, preparing me for more complex tasks in the field.

Chapter 3

Industrial Training

1.1 Objectives

The objective of this project was to leverage advanced data analytics, business intelligence, and data architecture design techniques to address real-world business challenges faced by Gemini Solutions. The goal was to analyze diverse datasets, implement scalable data architectures such as the bus matrix with star schema, and create insightful visualizations to support strategic decision-making. By doing so, the project aimed to enhance Gemini Solutions's operational efficiency, improve data-driven business strategies, and provide actionable insights that would lead to better user engagement and optimized service offerings.

1.2 Tools and Technology Used

During my internship at Gemini Solutions, I utilized a range of tools and technologies that were essential for data science. These included:

- **- Streamlit:** Used Streamlit to create a Central Repository, to list all the Data Science POC's.
- **- Python:** Programming language used for automation and scripting.
- **- Selenium:** Used selenium to create automated test cases.
- .
- **- Requests:** Library for making HTTP requests in EMail notification script..
- **- Cron:** Job scheduler used for automating script execution.
- **- VS Code:** Code editor used for development.

- **- Git/GitHub:** For maintaining a repository for the team to access the projects.

1.3 Projects Completed: -

1. POC CENTRAL REPOSITORY:

Developed a StreamLit application to list the details of all the past and ongoing POC's in the Data Science department

- **Technology Stack Used:**

  Python

  StreamLit

  SMTP Lib

  Linux Cron Job

This stack enabled the creation of a fully functional application to list all the POC's in the Data Science, emphasizing a user friendly experience for all the stakeholders involved.

2. EMAIL NOTIFICATION SCHEDULER:

Created a python script to send the Link status of the POC's to all the respective POC leads.
The script was written in python using the following packages/libraries:
Ssl, smtplib, email.mime.multipart, email.mime.text

Scheduled a cron job on the instance server where the email script was deployed.

The cron job is set for 10 am and 6 pm IST, daily.

```
# 10 am IST
30 4 * * * /home/ubuntu/POC-Central-Repository/myenv/bin/python /home/ubuntu/POC-Central-Repository/email_notification.py
# 6 pm IST
30 12 * * *  /home/ubuntu/POC-Central-Repository/myenv/bin/python3 /home/ubuntu/POC-Central-Repository/email_notification.py
```

3. AUTOMATED WORKING TEST CASES

Created working test cases for all the deployed POC's in the data science department.

Objective: *To automate the verification procedure for checking the working status of the POC's.*

The test cases were made using individual python scripts for all the POC's. Each python script was executed using the selenium function in python to automate the process.

- **Technology Stack Used:** Python | Selenium

TRAINING CONCLUSION

Successfully developed and launched a comprehensive Streamlit app to enlist and manage all data science POC's, ensuring centralized tracking and accessibility.

- Automated email notifications by creating a Python script, deployed and scheduled via a cron job on a hosted instance, achieving consistent communication twice daily.

- Developed and implemented automated Python scripts using Selenium to test the functionality of all deployed POC's, ensuring reliability and performance.

- Fine-tuned a mental health chatbot leveraging Hugging Face datasets and Unsloth base models, enhancing the accuracy and relevance of user interactions.

- Utilized VS Code and Google Colab for efficient development, testing, and deployment of various projects, streamlining the coding and execution process.
- Managed version control and collaboration using Git and GitHub, maintaining a robust codebase and enabling seamless teamwork.

Chapter 4

Future Scope and Training Benifits

- Throughout this project, I have significantly enhanced my technical and project management skills by working on various critical components, such as developing a comprehensive Streamlit app to manage and enlist all data science POC's. This experience provided me with practical exposure to designing scalable web applications and creating intuitive interfaces that centralize project tracking.
- I also deepened my understanding of automation and scripting by creating a Python script for email notifications, which I deployed and scheduled with a cron job on a hosted instance. This automation ensures reliable and timely communication, reflecting my growing proficiency in backend processes and operational efficiency.
- Additionally, I developed automated Python scripts using Selenium to test the functionality of all deployed POC's, further honing my skills in quality assurance and testing automation. This project also allowed me to fine-tune a mental health chatbot using Hugging Face datasets and Unsloth base models, enhancing my capabilities in natural language processing and AI-driven solutions.
- The supportive working environment at Gemini Solutions Pvt Ltd has been pivotal in my growth, providing valuable insights into industry best practices, collaborative problem-solving, and agile project management. My project management abilities have been sharpened through Git/GitHub version control, ensuring seamless collaboration and codebase manageme

Gemini Solutions - This report presents a series of data-driven projects, including form response analysis, Bus Matrix Architecture integration, word cloud visualization, and gender prediction. These projects showcase the application of various analytical tools and techniques to derive actionable insights and enhance data understanding.