A picture containing font, text, logo, graphics

Description automatically generated

*``***Summer Internship Report**

**on**

***CHAT WITH PDF***

Submitted by

*Name: Harsh Kaushik*

*Roll No: 21CSU182*

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,

SCHOOL OF ENGINEERING AND TECHNOLOGY

THE NORTHCAP UNIVERSITY

GURUGRAM-122017

Internship Period: 14/06/2024 TO 31/07/2024



Table of Contents

|  |  |  |
| --- | --- | --- |
| S.NO. | TOPIC | PAGE NO. |
| 1 | Certificate Of Completion | 1 |
| 2 | Table Of Contents | 2 |
| 3 | Abstract | 3 |
| 4 | Introduction | 5 |
| 5 | Problem Statement | 6 |
| 6 | Social Relevance of The Project | 7 |
| 7 | Training Description | 9 |
| 8 | Analysis | 11 |
| 9 | Mentor Feedback | 13 |
| 10 | Conclusion | 14 |
| 11 | Bibliography | 16 |
| 11 | Stipend Proof | 17 |
| 12 | Appendix | 18 |

**ABSTRACT**

My internship was conducted at Vrinda Nano Technologies (VNT), located in Sector 7, IMT Manesar, Haryana. VNT is a forward-thinking company specializing in energy efficiency and conservation, with a strong focus on sustainability. The company aims to simplify and provide innovative and efficient energy solutions to make the world greener, connected, and protected. VNT is dedicated to achieving the highest quality standards transparently and ethically. VNT partners with major firms such as Jio, BSNL, and Ericsson, further underscoring its prominence in the industry.

During my internship, I was placed in the Research and Development (R&D) department, where I had the opportunity to immerse myself in cloud technologies. Over the course of my training, I focused on developing an AI chatbot capable of processing a PDF as input and answering questions based on the content of the PDF. This chatbot was designed to respond accurately to context-relevant questions and to refrain from answering when the questions had no context with the given PDF.

To achieve this, I utilized various AWS services including AWS Bedrock for natural language processing, S3 for storage, EC2 for computational needs, and IAM roles for secure access management. Additionally, I hosted this service on the cloud, making it accessible from anywhere, thus demonstrating the scalability and convenience of cloud-based solutions. The training goals were to understand and implement cloud-based AI solutions, enhance my technical proficiency in AI and ML, and to gain practical experience with AWS services.

In addition to technical skills, I gained valuable insights into the organizational structure and workflows at VNT. I observed how the R&D department collaborates internally and with other departments to drive innovation. I also noted the emphasis on teamwork and the collaborative approach to problem-solving. The interaction between technical employees and customers was another key element of my learning experience, providing a holistic view of how customer needs are integrated into the development process.

Overall, my internship at VNT not only strengthened my technical capabilities but also provided a comprehensive understanding of how a cutting-edge technology company operates, collaborates, and innovates. The experience of developing and deploying a cloud-hosted AI service highlighted the importance of accessibility, scalability, and efficient resource management in delivering effective technological solutions. This internship has been instrumental in shaping my career aspirations and providing a solid foundation for my future endeavors in the field of AI and ML.

VNT's commitment to energy efficiency was evident throughout my time there. The company’s services, ranging from electrical health safety studies to energy audits and fire safety services, aim to improve business sustenance and contribute to national goals of dependable maintenance and tailored solutions. VNT's mission to support the Bureau of Energy Efficiency's objectives and its vision to innovate and lead in energy solutions was a source of inspiration and a testament to the impact of dedicated efforts towards sustainability.

**1. INTRODUCTION**

My internship was with Vrinda Nano Technologies (VNT), an innovative company located in Sector 7, IMT Manesar, Haryana. VNT is at the forefront of energy efficiency and conservation, partnering with major firms like Jio, BSNL, and Ericsson. The company's mission is to deliver advanced energy solutions that contribute to a greener, more connected, and safer world.

VNT aligns with the Bureau of Energy Efficiency's goals, focusing on super high-efficiency SMPS systems, solar power initiatives, and a broad range of services including electrical health safety studies, earthing studies, energy audits, PQ and thermography studies, lightning protection adequacy studies, and fire safety services. VNT is committed to achieving the highest quality standards transparently and ethically.

The internship allowed me to apply my academic knowledge in a practical setting. I worked in the Research and Development (R&D) department, which is known for its innovation and advanced research.

During my internship, I focused on cloud technologies and AI solutions. I developed an AI chatbot capable of processing PDFs and answering context-relevant questions. This project involved extensive use of AWS services, such as AWS Bedrock for natural language processing, S3 for storage, EC2 for computation, and IAM roles for secure access management. Hosting the chatbot on the cloud made it accessible from anywhere, showcasing the scalability and convenience of cloud-based solutions.

This experience required me to delve into cloud computing, artificial intelligence, machine learning, and natural language processing. It significantly enhanced my technical skills and provided practical insights into solving real-world problems. The internship not only improved my technical abilities but also offered valuable insights into VNT's collaborative and innovative work culture.

**2. PROBLEM STATEMENT**

The project aimed to address the challenge of extracting and retrieving specific information from extensive PDF documents efficiently and accurately. Traditional methods of searching through PDF documents are time-consuming and often yield irrelevant results due to the lack of context awareness. This inefficiency can hinder productivity, especially in environments where quick access to precise information is crucial.

The specific problem was to develop an AI-driven chatbot capable of understanding and processing the content of a PDF document, allowing users to ask questions and receive accurate, context-aware answers. The solution needed to ensure that the chatbot would only provide relevant information and refrain from answering questions when the context was not present in the given PDF. Additionally, the system had to be accessible from anywhere, highlighting the importance of a scalable and secure cloud-based implementation.

**3. SOCIAL RELEVANCE OF THE PROJECT**

The project holds significant social relevance by leveraging advanced AI technologies to enhance the accessibility and usability of information stored in PDF documents. In today's digital age, where vast amounts of data are generated and stored electronically, the ability to efficiently retrieve and utilize information is crucial for various societal sectors.

By developing an AI-powered chatbot capable of comprehensively analyzing and extracting information from PDF documents, the project addresses several key social aspects:

Firstly, it promotes **inclusivity** by making information more accessible to diverse user groups. Individuals with varying levels of technical proficiency can benefit from a user-friendly interface that allows natural language queries, reducing barriers to accessing and understanding complex information.

Secondly, the project **enhances productivity** across sectors by streamlining information retrieval processes. In fields such as education, research, and business, quick access to accurate information from documents can significantly improve decision-making and operational efficiency.

Moreover, the project contributes to **environmental sustainability** by reducing reliance on paper-based documentation. By facilitating digital access and analysis of PDF documents, it supports initiatives towards paperless offices and environmentally friendly practices.

Furthermore, the AI-driven approach promotes **data security** and **privacy**. By hosting the solution on a secure cloud platform with robust access controls, sensitive information within documents remains protected while enabling authorized users to extract insights efficiently.

Ultimately, the social relevance of the project lies in its ability to harness advanced technologies to solve practical challenges in information management. By making information retrieval more efficient and intuitive, the project empowers individuals and organizations to make informed decisions, fostering progress and innovation across various societal domains.

**4. TRAINING DESCRIPTION**

During my internship at Vrinda Nano Technologies (VNT), I was involved in a project focused on developing an AI-driven chatbot for extracting and analyzing information from PDF documents. This project was undertaken within the Research and Development (R&D) department, aiming to enhance document management and information retrieval processes using advanced technologies.

**Project Goals**

The primary goal of the project was to:

* Develop an AI-powered solution capable of parsing and understanding content within PDF documents.
* Enable users to interact with the chatbot through natural language queries to extract specific information efficiently.
* Ensure the scalability and accessibility of the solution by deploying it on cloud infrastructure.

**Methodology**

1. Research and Planning:

* Conducted a comprehensive review of AI and NLP techniques suitable for document analysis.
* Explored and selected Amazon Titan and Claude 2.1 models from Bedrock for integration into the project.

1. Technical Implementation:

* Utilized Python programming language and relevant libraries for PDF parsing and text extraction.
* Integrated Amazon Titan and Claude 2.1 models from Bedrock for enhanced AI capabilities in document analysis.
* Implemented NLP algorithms for entity recognition, keyword extraction, and sentiment analysis.
* Developed a user-friendly interface using Streamlit to allow users to interact with the chatbot seamlessly.
* Used DockerHub for containerization to ensure portability and scalability of the application.

1. Deployment and Optimization:

* Deployed the solution on cloud infrastructure using AWS services such as EC2 for computing and S3 for storage.
* Conducted iterative testing and optimization to improve the chatbot's performance and responsiveness.
* Implemented feedback loops to refine algorithms and enhance user experience.

***Accomplishments***

* Successfully developed and deployed an AI-driven chatbot capable of processing PDF documents and providing relevant information to users.
* Enhanced document management efficiency within VNT, reducing manual search times and improving productivity.
* Documented project methodologies, algorithms used, and outcomes achieved for future reference and knowledge sharing.

1. **ANALYSIS**

**Relevance of Training Work**

My internship at Vrinda Nano Technologies (VNT) significantly enhanced my grasp of the engineering discipline, particularly in AI, machine learning, and natural language processing (NLP). Through developing an AI-powered chatbot for extracting information from PDF documents, I gained practical insights into applying theoretical knowledge to real-world engineering challenges.

1. Technical Learnings:

* Advanced AI Techniques: I mastered implementing and optimizing NLP algorithms for text parsing and sentiment analysis, crucial for extracting insights from unstructured data.
* Cloud Deployment: Employing STREAMLIT for the interface and leveraging Docker Hub for containerization, I streamlined deployment and enhanced application accessibility.

2. Organizational Insights:

* Company Structure: VNT's collaborative structure fostered agile practices, facilitating seamless cross-departmental communication and innovation.
* Team Collaboration: Engaging in effective teamwork underscored the importance of shared knowledge and skill development.
* Client Interaction: Aligning technical solutions with client needs highlighted the significance of client-centric service delivery.

**Performance Analysis**

Strengths:

* Technical Proficiency: Successfully integrating complex AI algorithms demonstrated strong technical acumen.
* Problem-Solving Skills: Overcoming challenges like data preprocessing showcased analytical prowess.
* Communication Skills: Active participation in team discussions and presentations highlighted effective communication abilities.

Areas for Growth:

* Time Management: Enhancing efficiency in task prioritization and project planning.
* Technical Mastery: Further refining skills in advanced AI techniques and staying updated with industry advancements.
* Client Engagement: Developing hands-on experience in client interaction to better address their needs and feedback.

**MENTOR FEEDBACK**

**A close-up of a letter

Description automatically generated**

**6. CONCLUSION**

My summer training at Vrinda Nano Technologies (VNT) has significantly advanced my understanding of the engineering profession, both technically and organizationally. This experience has been invaluable in applying theoretical knowledge to real-world challenges, particularly in AI, machine learning, and natural language processing (NLP).

**Technical Advancements:**

Developing an AI-driven chatbot for PDF document analysis allowed me to:

* Implement advanced AI algorithms for text parsing and sentiment analysis.
* Utilize cloud platforms such as AWS, STREAMLIT for interface development, and Docker Hub for efficient application deployment.

**Organizational Insights:**

Working within VNT's collaborative structure and agile methodologies provided insights into:

* Effective team dynamics and cross-functional collaboration.
* Client-centric service delivery and project management practices.

**Learnings and Benefits:**

This training has enriched my:

- Technical proficiency in AI and NLP applications, essential for data-driven decision-making.

- Problem-solving skills, particularly in optimizing algorithms and improving application performance.

- Communication abilities through presentations and collaborative interactions within a professional environment.

In conclusion, my internship at VNT has been pivotal in bridging academic knowledge with practical skills, preparing me for a successful career in engineering. The experience underscored the importance of continuous learning and adaptive problem-solving in navigating complex technological landscapes.

**BIBLIOGRAPHY**

1. Streamlit Documentation

<https://docs.streamlit.io/get-started/fundamentals/main-concepts>

1. Amazon AWS Documentation

<https://docs.aws.amazon.com/>

1. Docker Documentation. (n.d.). "Docker Overview

<https://docs.docker.com/get-started/overview/>

1. Vrinda Nano Technologies. (2024). Internal documents and presentations provided during internship at VNT.