

# Siddharth Kumar

BTech CSE - PES University Bangalore

Ph:91 9606183715

[siddharth78980@gmail.com](mailto:siddharth78980@gmail.com)

Bangalore, KA



## SUMMARY

Task-driven individual with experience in MERN stack, NLP, C, and Python. Fond of problem-solving, and adept at leveraging superior communicative and interpersonal skills to interact with diverse individuals and groups at every level. Possesses excellent leadership qualities and loves working with people.

## PROJECTS

### Private Question Answering Model

Designed and Developed a secure **LLM** (Private Language Model) for confidential question answering on sensitive company documents, utilizing cutting-edge techniques such as **K-means clustering** and **Pinecone Vector DB** for enhanced long-term model memory. ([GitHub](#))

- Utilized HaystackDocumentStore for building Question Answering and Search pipelines. Key components: DocumentStore (Pinecone Vector DB), Retriever (EmbeddingRetriever), and Generative Model (Seq2SeqGenerator).
- Sentence Transformers: Used for document embeddings. Facilitated efficient similarity matching during retrieval.
- BART Model: Implemented for sequence-to-sequence generation. Capable of producing human-like text.

### Large Scale Server Performance Monitoring System (Python)

An application that simultaneously connects to multiple servers and executes specific commands in real-time using ansible. The results are displayed on a dashboard-like user interface created using Streamlit. ([GitHub](#))

- Used an open-source software called **ansible** which execute set commands on all the servers at once and return the output.
- The output is received by the frontend (Streamlit - A Python Library) which parse the data and displays the data in a dashboard like fashion.
- The software also a custom shell which can take in data and run that particular command on all the servers at once.
- Stack: Python, Ansible, Streamlit

### Multimodal Question Generation Model (Final Year Project)

The objective of this model is to process a page from a **textbook** that includes both **text and pictures** and automatically generate **questions** along with its corresponding answer key. The intended application for this model is in the field of **education**, where it can assist teachers in generating questions more efficiently, saving time and effort.

- Models:** T5 base Finetuned on SQuAD1.1 Dataset, BLIP Image Captioning
- Helper Libraries:** Wikipedia, Feature Extraction (Image Captioning).
- Datasets:** Sense2Vec, RACE Dataset, SQuAD1.1 Dataset.
- Stack:** Python, Flask API, Next.JS, Tailwind CSS.

### Tourism Specialists (C)

A software made in **C** language which works as a **Hotel booking and Vacation planning system**.

- This project aimed at providing users with full command of customizing their Vacations and Trips. We used concepts like **CSV manipulation** to maintain our database.
- At the end of the project, an all-rounded software was created to allow users to have a better experience in **Hotel booking and vacation planning**.

### Pricera (Python)

It's a software which uses **Python** and **bs4** to scrape the web and get data from amazon.

- Pricera is made to help customers with **E-commerce** shopping. This helps users compare multiple electronic products and choose the best-suited gadget at the best price.
- It uses concepts like **Web scraping to scrape Amazon** and get instant results of any new products in the market.

### Hotel Management Website (MERN)

A website which allows users to make hotel bookings as well as let owners add new hotels, with login and signup using JWT (Json web tokens) and cookies.

Made using MongoDB, Express, React.Js, Node.js.

## TECHNICAL SKILLS

**Languages:** C/C++, Python, JavaScript,

**HTML+CSS Libraries:** C++, STL, ReactJs,

Nodejs, Git, GitHub

**Cloud/Databases:** MongoDB, Firebase, MySQL

**Relevant Coursework:** Data Structures & Algorithms,

OS, Object Oriented Programming, Database

Management System, Software Engineering

**Areas of Interest:** Web Design and Development,

Software Development, Machine Learning.

**Soft Skills:** Problem Solving, Self-learning, Presentation, Adaptability

## PROFESSIONAL EXPERIENCE

### Summer Internship

[Rakuten Symphony](#)

06/2023 - 07/2023

Bangalore, KA

Rakuten Symphony is a **Rakuten Group company**, providing global B2B services for the mobile telco industry and enabling next-generation, cloud-based, international mobile services.

- Designed and Developed a secure **LLM** (Private Language Model) for confidential question answering on sensitive company documents, utilizing cutting-edge techniques such as **K-means clustering** and **Pinecone Vector DB** for enhanced long-term model memory.
- Built a robust Multi-Server Performance Monitoring System Dashboard using Ansible and Streamlit. Successfully deployed the system on internal company servers, enabling efficient monitoring and management of multiple servers simultaneously, supporting the DevSecOps team in ensuring optimal performance and security.
- Key Skills: Private Large Language Model, K-means Clustering, Pinecone Vector DB, Multi-Server Monitoring, Ansible, Streamlit, Server Management, File Management System.

### Internship

[CDSAML @ PESU](#)

06/2022 - 07/2022

Bangalore, KA

An Internship program hosted by PESU under the HOD of the college

- Floor Plan Generation Using CNN**
- Our research project aims to assist architects in planning floor plans for buildings by automating the process.
- By taking inputs and constraints from the client, our system generates multiple unique floor plan designs for them to choose from
- This streamlines the iterative process architects often go through, providing efficient and diverse options for the clients.

## KEY ACHIEVEMENTS

### Head and Founder of A college Club (Hallothon)

In the year 2022, I took the initiative to establish a college club and assumed the position of its head. This club had the proud tradition of hosting an annual hackathon, which proved to be a resounding success, attracting a massive participation of 300 enthusiastic individuals eager to showcase their skills and compete in the event.

### Won CDSAML hackathon conducted by PESU

CDSAML was a summer internship program organized by PESU in the month of June and July of 2022. Our team researched on the topic of floor plan generation using CNN

