

DEEKSHITH T N

☎ +91 9019636311 ✉ deekshithn89997@gmail.com 🌐 <https://www.linkedin.com/in/deekshith-t-n-44016b245/>

Education

SJB Institute of Technology Bengaluru
Bachelors of Electrical and Electronics Engineering

September 2020 – July 2023
Bengaluru, Karnataka

Experience

MSR Cosmos IT LLP

Generative AI - Software Engineer

September 2024 – Present

Hyderabad, Telangana

- **RAG Expertise:** Implemented RAG systems with advanced chunking strategies, improving retrieval accuracy by 40% and reducing latency by 60%. Processed over **300,000+ documents** for efficient and accurate retrieval workflows.
- **Social Media Sentiment Analysis:** Developed a sentiment analysis pipeline for the AP government using GPT-4, achieving 85% accuracy across **100,000+ social posts** Applied clustering techniques for actionable insights reporting.
- **LLMs Fine-Tuning:** Fine-tuned text-based, speech-based, image-based, and multimodal large language models (LLMs) to create innovative AI-driven solutions tailored for diverse applications.
- **Backend and DevOps** Engineered FastAPI solutions with Docker and Azure, reducing deployment time by 60% through CI/CD optimization for scalable, efficient, and high-performance backend systems.

Simpragma Solutions Pvt. Ltd.

AI/ML developer

August 2023 – September 2024

Bengaluru, Karnataka

- **Speech-to-Text Models:** Optimized Whisper Large, Medium, and Small models using CTranslate2, Faster-Whisper, and TensorRT pipelines, achieving sub-1-second response times with minimal resources for efficient speech recognition.
- **Text-to-Speech Models** Integrated scalable **Coqui TTS** and **Google TTS** for high-quality voice synthesis, using fine-tuned cloning techniques for enhanced, natural-sounding real-time interactions.
- **Chat Bot and Voice Bot Development:** Built a voice bot and chat bot using **GPT-3.5** and **Qdrant** for personalized advice and efficient data retrieval. Deployed with **Docker** created API with **FastAPI** for scalability and reliability.

Pie Infocomm Pvt. Ltd (Internship)

AI/ML developer

Febraury 2023 – July 2023

Bengaluru, Karnataka

- **Resume Analysis Tool Development:** Developed an NLP-based tool to extract and categorize key resume details, improving screening accuracy. Created an algorithm to match candidate skills and experience with job requirements for efficient recruitment.

Technical Skills

Languages: Python, SQL, MongoDB

Developer Tools: FastAPI, Azure Function Apps, Docker, Git, Gitlab, Linux, Vector DBs

Technologies/Libraries: Machine Learning, Deep Learning, Numpy, Pandas, Scikit-learn, Keras, Pytorch, Transformers, Generative AI, Langchain, Huggingface, RAG Evaluation and Optimizations

Algorithms: Linear and Logistic Regression, SVM, KNN, Decision trees, Random Forest, Bagging, Boosting, Clustering

Certification/Awards : Awarded Rising Star Award in 2023 and Awesome Card Award in 2024.

Projects

Digital Library BOT | *Python3, GPT-4, Qdrant, Multi-Model RAG, FastAPI, MongoDB, Azure,*

- Developed a RAG-based chatbot application with 96% retrieval accuracy, utilizing advanced data pipelines and language models for insightful responses and real-time user interactions. Deployed the solution with Azure for efficient, scalable integration.

SmartCommerce Search Engine | *Python3, GPT-4o, Pinecone, Multi-Model RAG, Azure, FastAPI*

- Developed a RAG-based chatbot application with 98% retrieval accuracy, utilizing advanced data pipelines and language models for insightful responses. Deployed the solution with Azure for efficient, scalable integration.

SQL-LLM Chatbot for Data Base Management | *Python3, GPT-4, Agents, RAG, FastAPI, Azure*

- Created a SQL-LLM chatbot for DB, allowing natural language commands to query and create databases from SQL scripts. Integrated CSV/XLSX import and RAG techniques to generate relevant responses from tabular data.

Andhra Pradesh Social Media Sentiment Analysis | *Python3, KNN, GPT-4, HDBSCAN, Zero-Shot Classification*

- Collected and preprocessed social media data for sentiment analysis using HDBSCAN and KNN. Integrated GPT-4.0 for cluster naming and MongoDB for efficient storage and reporting.

BFSI Chatbot and Voice Bot Development | *Python, STT, NLP, NLU, TTS, FastAPI, Flask, spaCy, Whisper*

- Developed chatbots and voice bots for BFSI, automating customer queries and financial guidance for Bajaj and DMI Finance, enhancing user experience with NLP-based dialogue management and personalized interactions.