

# SUKRUTH KESHAVARAM MOTHAKAPALLY

Dublin, CA | 603-688-4686 | [sukruth0402@gmail.com](mailto:sukruth0402@gmail.com) | [GitHub](#) | [LinkedIn](#) | [Portfolio](#)

## EDUCATION

Northeastern University - Boston, MA

Sept 2022 - Dec 2024

Master of Science in Computer Software Engineering

GPA: 3.8

Relevant Courses: Data Structures and Algorithms in Python and Java, Advances in LLMs, Cloud, Big Data, Parallel Computing.

Visvesvaraya Technological University - Bangalore, India

Aug 2016 - June 2020

Bachelor of Engineering in Computer Science

GPA: 3.0

Relevant Courses: Java and Python App Programming, Machine Learning, Object-Oriented Concepts, Big Data Analytics

## WORK EXPERIENCE (~ 3 years)

Graduate Teaching Assistant – Northeastern University (Python, ML, Object Oriented)

Jan 2023 – Dec 2024

- Designed and graded assessments for 40+ students in the courses - Data Structures and Algorithms, and Data Science.

AI Software Engineer Intern – [PipeIQ.ai](#), Palo Alto, CA (AI, Python, AWS, SQL, React, TS, Node.js)

Jan 2024 - Aug 2024

- Member of a startup's founding engineering team, taking ownership of multiple microservices using Python and Node.js, leveraging RESTful APIs and AWS services like Lambda, DynamoDB, and S3, accelerating MVP launch timeline by 30%.
- Developed a dynamic and interactive user interface with React, Next.js, and TypeScript, enabling cross-stack features like authentication, content delivery, and payment processing, resulting in a 40% boost in platform performance and response time.
- Engineered a multi-agent AI interface, 'AI Copilot', leveraging OpenAI, scikit-learn, and LangChain to automate and personalize tasks such as email generation, intent classification, and campaign creation, boosting revenue growth by 70%.
- Developed RAG based AI chatbot using OpenAI embedding and FAISS, into marketing platform, increasing customer productivity.

Systems Engineer – Infosys (client - Intel), Bangalore, India (Python, C++, SQL, CICD, Testing)

Dec 2020 - June 2022

- Developed Intel driver deployment solutions in C++ on Linux, leveraging automation tools and configuration management techniques.
- Built automated testing frameworks in Python for SSD firmware on Linux, enhancing system reliability and performance by 15%.
- Performed root cause analysis of driver performance metrics and logs using SQL, identifying and mitigating production issues.
- Re-engineered driver deployment process with Python and Shell scripting, reducing driver update time by 30% for 1 million+ users.
- Optimized CI/CD pipelines with Jenkins and Git, improving code quality and reliability in driver deployment workflows.

Full-Stack Developer Intern – Infosys, Mysore, India (JavaScript, MERN, NoSQL, Agile)

Feb 2020 - Mar 2020

- Collaborated on backend development using Node.js, Express.js, and MongoDB for a dynamic employee skills analysis platform.
- Developed interactive React components and integrated JavaScript to enable real-time data visualization on the platform.

Software Developer Intern – Bosch, Bangalore, India (Python, Node.js, REST, SQL, AWS)

July 2018 - Aug 2018

- Contributed to Python and Node.js backend services for shipment tracking app, integrating RESTful APIs and supporting 30+ endpoints.
- Optimized SQL / NoSQL queries, improving shipment data processing time by 30% and deployed a dockerized app on AWS.

## ACADEMIC PROJECTS

[Custom LLM & GPT Fine-Tuning with Multi-Agent AI](#) (Python, PyTorch, LLM, RLHF, NLP, Deep Learning)

- Built scalable GPT-based LLM integrating multi-head attention, positional encodings, and tokenization for instruction-tuned text generation.
- Fine-tuned model with LoRA, leveraging instruction backtranslation, self-augmentation, and datasets to align responses with intent.
- Applied DPO, RLHF, and multi-agent AI orchestration to refine dialogue quality, adherence, and overall model performance.

[Personalized AI Job Tracker - Full-Stack App](#) (Python, LLM, FastAPI, JavaScript, Postgres, AWS, CI/CD, Docker)

- Built AI-powered Chrome extension with JavaScript, automating job tracking and improving productivity through advanced data processing.
- Developed and deployed REST APIs with FastAPI, leveraging multi-agent AI using OpenAI LLMs to perform job data analytics.
- Orchestrated end-to-end deployment with Docker, CI/CD pipelines via GitHub Actions, and hosted on AWS, ensuring high availability.

[AI-based Stack Overflow Full-Stack App](#) (Python, LLM, Airflow, Docker, FastAPI, CICD, GCP)

- Led a team of 3 to develop Dockerized Python app, integrating GenAI features (RAG, text generation, summarization) into Stack Overflow.
- Built data pipeline with Airflow, deployed app on GCP, using FastAPI for backend and Streamlit for UI, ensuring efficient workflows.

[Audio Emotion Classifier with Deep Learning and Parallel Computing](#) (Python, PyTorch, CNN, NLP)

- Spearheaded development of an Audio Emotion Classifier leveraging advanced Deep Learning and NLP methodologies.
- Optimized model's efficiency by integrating parallel computing, effectively detecting complex emotion patterns from audio data.

## TECHNICAL SKILLS

- Programming** - Python, Java, JavaScript, TypeScript
- AI / LLM** - AI agents, Function calling, LangChain, LlamaIndex, RAG, fine-tuning, RLHF, DPO, Prompt Engineering
- ML | NLP** - TensorFlow, Keras, Sklearn, PyTorch | Sbert, Cohere, Hugging Face, NLTK, Embeddings
- Backend** - Flask, FastAPI, Django, Node.js, Express.js, Spring Boot
- Frontend** - React, Redux, Next.js, Angular.js, HTML, CSS, Bootstrap, SASS/SCSS
- Databases** - SQL (PostgreSQL, MySQL, SQL Server), NoSQL (MongoDB, Redis), Elasticsearch
- Cloud** - AWS (EC2, S3, RDS, IAM, SQS, Lambda, ECS), GCP (CloudSQL, GCS, BigQuery, Pub/Sub)
- Devops** - CI/CD, Git, GitHub, GitHub Actions, Jenkins, Docker, Terraform, packer, Kubernetes