

# VENKATA SAMYUKTA MALAPAKA

[✉](mailto:venkatasamyuktamalapaka@gmail.com) [linkedIn](https://www.linkedin.com/mvsamyukta/)

[github](https://github.com/vmalapaka)

[github](https://github.com/vmalapaka)

## EDUCATION

### University of Massachusetts, Amherst

Feb'23 – May'25

Master of Science in Computer Science (Data Science Honors)

### Jawaharlal Nehru Technological University, Hyderabad, India

Aug'16 – May'20

Bachelor of Technology in Computer Science and Engineering

## TECHNICAL SKILLS

**Programming Languages:** Java, Python, Scala, Javascript, Typescript, HTML/CSS, R, Linux, React

**Databases/Storage :** MongoDB, NoSQL, Oracle SQL, MySQL, PostgreSQL, Elasticsearch, Redis, Django

**Frameworks/Libraries:** SpringBoot, MVC, Confluence, Angular, Tensorflow, Pytorch, OpenCV, Pinecone, Terraform

**Softwares/Tools:** AWS, Docker, Kubernetes, Atlassian, Apache Spark, Apache NiFi, Kibana, Kafka, Logstash, Jenkins, Groovy, Firebase, Langchain, HuggingFace, GIT, Azure, GCP, NodeJS, OpenAI API

## EXPERIENCE

### Amazon - Software Development Engineer

May 2025 – Ongoing

- Engineering cryptographic solutions within Amazon Global Payments infrastructure, designing proprietary encryption protocols and implementing secure data transmission pipelines to mitigate DDOS attacks and ensure PCI-DSS compliance for financial transactions

### Umass BioNLP Lab - Research Scientist

Jan 2025 – Ongoing

- Researcher on a project build a prediction model to identify new social determinants of health(SDOH), and generate a Grapg-RAG for impact scores of SDOH on each other using MIMIC-III/IV datasets with 40,000+ patient data

### IgniteIQ - AI Engineer

Oct 2023 – Dec 2023

- Architected prompt optimization framework leveraging semantic embeddings, context compression algorithms, and distributed caching across 6 production chatbot applications
- Achieved 95% response accuracy while reducing token consumption and inference costs by 40% through hybrid RAG

### Verizon - Full Stack Software Development Engineer 2

Feb 2020 – Jan 2023

- Worked in a cross-functional 5G team to build a global web app for Verizon field tech scheduling, driving \$5M in profit
- Built RESTful APIs in Spring Boot and developed a TypeScript-based Angular front-end to enhance product features
- Eliminated 2 hours/week of downtime through AWS deployments, shadow rollouts, and cron jobs that auto-update the application database every 3 minutes.
- Built a high-performance report tool with Hive on HDFS, Spark and ELK for enabling real-time analytics on 18TB of data
- Optimized query performance with partitioning JSON data, indexing, and caching, cutting report generation time by 90% for 15+ non-BI clients

## PROJECTS

### YOLOv11-Based Crowd Flow Detection on Image Data

Sep 2024 – Dec 2024

- Addressed the challenge of predicting crowd flow patterns in resource-constrained environments using static images instead of bandwidth-intensive video data
- Developed a two-stage pipeline leveraging YOLOv11 for object detection and flow classification, supported by custom datasets with detailed annotations
- Achieved 85% precision and 90.1% mAP50 in detection, with 55% accuracy in flow classification, showcasing potential for real-world applications

### Personalization in Large Language Models

Sep 2023 – Dec 2023

- Fine-tuned Flan-T5-base architecture through engineered personalized prompts, utilizing AdamW optimizer with 5-step warmup scheduling, linear decay rate, and L2 regularization across 10 training epochs
- Engineered context-aware prompts incorporating user conversation history via BM25 information retrieval algorithm, achieving 80% accuracy on personalization benchmarks with contextually-relevant generation per user

### Dynamic Stock Trading Application

Feb 2023 – May 2023

- Developed web application implementing distributed systems patterns: Redis-backed caching layer, consensus-based leader election for replication, and fault-tolerant architecture supporting 1000+ concurrent transactional requests
- Designed RESTful APIs for client-facing communication and gRPC protocols for inter-service transactions, integrating AWS Lambda serverless compute for elastic scalability and reduced operational overhead
- Implemented Agile methodology with sprint-based iterations and utilized distributed storage architectures for horizontal scalability, optimizing throughput and reducing latency under concurrent load conditions