

Kirubaraj M

kmanimat@cs.cmu.edu | [linkedin.com/in/kirubarajm](https://www.linkedin.com/in/kirubarajm) | kirubarajm.me | github.com/kirubarajm | +1-415-417-8782

EXPERIENCE

Troutwood

January 2024 – December 2024

AI Engineer and Team Lead – Internship (Contract)

- **Leading** 4-member team building **Langchain LLM Agents** reducing research time for Financial Advisors by **50%**.
- Leveraged open source multimodal LLMs, embedding models, Vector DB to build AI Agent supported by LlamaIndex framework for **RAG** and deployed them on AWS EKS enabling enterprise level **security** and **privacy** for clients.
- Built near real time data aggregation pipeline for user data using AWS **SNS** and **SQS** to extract multi-user insights.
- Developed **API** based Tools and **Parallel Tool Calling** for Langchain Bedrock increasing chat **performance** by **8X**.

Oracle

November 2022 – July 2023

Senior Software Engineer

- **Led** a cross-functional team of 3 in **technical design, development, deployment** (CI/CD Pipeline) of multiple full-stack enterprise software projects for Oracle ERP, facilitating onboarding of **3** enterprise customers.
- Spearheaded **architecture design** and end-to-end development of **ELK Stack** with real time data ingestion, search and auto-fill features of account combinations increasing search result click rate by **75%** for thousands of end users.
- Transformed Monolith/JSF architecture of Flexfield component to independently deployable **ReactJS** component with backend **NodeJS microservices** deployed on Kubernetes, enabling portability and **scalability** across **12** teams.

Software Engineer

August 2020 – November 2022

- Unified multiple custom **PL/SQL** Packages and **Data Models** into seeded parametrized offerings avoiding code redundancy and increasing **maintainability**, resulting in decrease of inbound bugs by **27%** in the reporting product.
- Reconstructed the report visualization engine by parallelizing service calls, redesigning **ReactJS** renderer and tuning the underlying **SQL queries** by optimizing their joins, indexes and query hints reducing report rendering time by **70%**.
- Resolved **77** customer issues, emphasizing **performance** and high **availability** of business-critical reports, averaging 2 issues per week with 60% **Severity-1** escalations and an average Turnaround Time of **8 hours**.
- Optimized and added new **JUnit** automated regression tests decreasing failure rate by **18%** across **8** products.

Eat From Home (TOVO Technologies Private Limited) [[GitHub](#)]

August 2018 – August 2020

Co-Founder and CTO

- Designed, and developed an **Android** app for on-demand e-commerce with **45,000+** installs and **3,000+** DAUs.
- Independently built and managed the **NodeJS REST API** server for seamless integration of user, rider, and chef apps leveraging Shopify **GraphQL** APIS, and Firebase for order management, and real-time order tracking.
- Built a streaming big data pipeline of concurrent order locations and rider locations. Optimized order allocation system by implementing order batching algorithm leveraging **Kafka** and **Samza** bringing down delivery cost by **35%**.
- Built and deployed a Multi-Input Multi-Output Machine Learning model for last-mile prediction. Utilized model with real-time order data to enable delivery time **prediction**, eliminating Google Maps cost and reducing app load time by **40%**.
- **Led** and **mentored** 5 engineers/designers in development of rider/chef apps and ensured on-time delivery.

EDUCATION

Carnegie Mellon University

Pittsburgh, PA

Master of Software Engineering

December 2024

Coursework - Deep Learning, Cloud Computing, Distributed Systems, Data Structures, Algorithms, Database Systems

National Institute of Technology, Tiruchirappalli (NIT Trichy)

Tiruchirappalli, India

Bachelor of Technology – Electronics and Communication Engineering

July 2020

SKILLS

Backend/Data – Java, NodeJS, Golang, C | MySQL, MongoDB, Redis, Elasticsearch, Spark, Kafka, BigQuery, JUnit
Web/Mobile Dev – Javascript, Typescript, ReactJS, NextJS | REST API, Web Sockets, gRPC, GraphQL | CSS3, HTML5
AI/Infra – Python | Pytorch, LangChain, RAG, NLP | AWS, Azure, GCP, Terraform | Docker, Kubernetes, Linux, Git

PROJECTS

Integrating Short-Term News Sentiment To Stock Prediction [[Code](#), [Thesis](#), [Demo](#)]

January 2024 – May 2024

- A short-term stock prediction model using **attention techniques** to integrate news sentiment to historical stock data.
- Implemented **3** variations of **Transformer** Architecture on baseline **CNN-LSTM** model, improving MSE loss by **94.8%**.

Daily Locally [[Code](#), [Design](#)]

August 2019 – August 2020

- An e-commerce **Android** platform with **10k+** installs with a scalable **NodeJS** backend and **ReactJS** web dashboard.

Determining Human Body Measurements Using Deep Learning [[Code](#), [Thesis](#), [Demo](#)]

December 2019 – May 2020

- A **full-stack** web application powered by **ResNet** based deep learning model to measure human body with 2 images.