Nithil Eshwar Mani

7162925911 | nithiliisd@gmail.com | nithil3007.github.io/portfolio | linkedin.com/in/nithil-eshwar | github.com/Nithil3007

EDUCATION

University at Buffalo 2025
Master of Science (Data Science) CGPA: 3.7/4.0
Courses - Statistical Learning, Data Mining, Data Models and Query languages, Predictive Analysis, Analysis of Algorithms.

College of Engineering, Guindy (CEG), Anna University 2023
Bachelor of Engineering (Computer Science and Engineering) CGPA: 3.3/4.0
Courses - Machine Learning, NLP, Big Data Analytics, Database Management Systems, Object-Oriented Analysis and Design.

EXPERIENCE

Software Engineer Intern, **Green InfoTech (S-corp)** - Union City, CA July 2025 - Present

- Worked on the **MedScribe AI** project as part of a 5-person team, which focuses on automating paperwork by transcribing patient encounters, generating notes and filling medical forms.
- Built a **Clinical transcript RAG** to provide context-aware responses. Used **AWS Lambda** and **Amazon Bedrock's** models to generate responses and embeddings. Adopting AWS Lambda instead of server-based architectures **reduced costs** by **60%**.
- Developed an **AI Call Agent** to automate patient intake (appointment scheduling, insurance information verification). Used **FastAPI** for building **APIs** and **ElevenLabs** for **conversation handling**, **storage** and **retrieval**.
- Optimized a **Text-to-SQL agent** to provide SQL results for natural language queries. Used **Vanna AI** to train the model and wrote **70+ SQL query prompts**. Improved performance on complex queries by **60%**.
- Designed an **admin portal** to organize user profiles, analyse user statistics and activity, manage permissions and feature flags. Ensured reliable user **database connectivity, query performance** and enforced **HIPAA compliance**.
- Wrote **terraform** scripts to deploy applications in **AWS Amplify** and **AWS App Runner**. Created **Docker** files to build application images and implemented CI/CD pipelines for automated deployment.
- Utilized **AWS CloudWatch** to monitor application performance and **AWS Cognito** to manage user authentication.
- Improved code efficiency by **20%** by creating API endpoints that return smaller, more specific responses. Performed **unit testing**, **integration testing** and **API testing** using **pytest** to ensure reliability.

Research Assistant, **University at Buffalo** - Buffalo, NY February - May 2025

- Collaborated as part of a 4-member team on the development and testing of **UnionLabs**, a cloud-based distributed platform to share data, code, software and hardware resources for research in next-generation networks and wireless Internet of Things.
- Established **AWS EC2-local server** connectivity with API endpoints.
- Conducted thorough **testing** of **back-end**, **database**, and **API endpoints** to ensure robustness and reliability of the platform.
- Improved user **accessibility** and **responsiveness** by **70%** through optimized **React.js** code.
- The cloud-based remote-access solution enabled ease and efficiency and **reduced lab costs** by **60%**.

PROJECTS

Automated Data Analysis and Visualization (FastAPI, Pandas, React, Typescript) - Built a full-stack **web application** that performs intelligent **data analysis, visualization** and **provides AI insights** on CSV files. **Fine-tuned** the **gpt-oss-20b** model to automatically identify important features, determine data types, and recommend univariate and bivariate analyses.

Question-and-Answer Agent for Research Papers (Langchain, Cypher, Streamlit) - Built a **Q+A agent** for Research papers using **Graph RAGs** and **GPT-4o**. Deployed the app using **Streamlit**. Created knowledge graphs and wrote **data retrieval functions** in **cypher**. The Graph RAG gave better results in **identifying links** between multiple papers compared to traditional RAGs by **70%**.

HEP-TH (High Energy Physics Theory) Paper classification using Sci-BERT (Python, TensorFlow, SageMaker AI) - Classified over 30000 HEP-TH papers dating from 1991-2004 as influential and non-influential. Used **AWS SageMaker Blazing Text** for classification and **Sci-BERT** word embedding for tokenising. Performed hyperparameter optimization and obtained **75%** accuracy.

Elastic Net Attack and Inception-ResNet V1 for Retinal OCT images (Python, Pytorch, AWS SageMaker AI) - Optimized the **Inception-Resnet V1** model to detect Choroidal Neovascularization in over **10000 Retinal OCT images**. Obtained better accuracy (97.10%) compared to the **Amazon SageMaker Image Classification Model** (95.43%). Trained the model to overcome Elastic-Net attack. Improved the accuracy on adversarial examples from 7.5% to 51.39%.

TECHNICAL SKILLS

- **Programming and Querying**: Python, C, C++, R, Java, Golang, React.js, Node.js, Typescript, SQL, Cypher
- **Software Development and Containerization**: FastAPI, Flask, Django, Git, Docker, Kubernetes, Agile Methodologies
- **Databases and Analytics**: MySQL, PostgreSQL, Neo4j, DuckDB, MongoDB Tableau, PowerBI, Google Analytics, Metabase
- **AWS Suite**: Amplify, Apprunner, Amazon EKS, Amazon EC2, Glue, Athena, Redshift, Spark, SageMaker, Bedrock, Cognito
- **AI and Programming Frameworks**: Langchain, BigQueryML, Pytorch, Tensorflow, Scikit-learn, Matplotlib
- **Certifications**: **AWS Certified Machine Learning Engineer Associate, AWS Certified Data Engineer Associate**