Amruth Pai Thukaram

469-465-4829 | amruthpaiuni@gmail.com | linkedin.com/in/amruthpai | github.com/Immortal-Pi

Education

University of Texas at Dallas **Aug 2024 – May 2026**
Master of Science — Business Analytics & Artificial Intelligence *GPA 3.91*

Nitte Meenakshi Institute of Technology, India **Jun 2015 – Aug 2019**
Bachelor of Engineering — Computer Science Engineering *GPA 3.65*

Professional Experience

Data Analytics and Software Developer Intern | *Van Brunt & Associates, Texas* **May 2025 – Aug 2025**

- Designed machine learning pipeline to forecast daily and weekly peak loads in the ERCOT grid improving accuracy by 18%
- Engineered feature sets by integrating ERCOT system forecasts, weather data and tariff-based indicators like 4CP demand
- Built a dual-objective model to forecast load and 4CP events with 87% precision, minimizing commercial billing risk
- Integrated price signals, battery profiles, and weather anomalies to enhance forecasting accuracy by 14%
- Deployed modular forecasting pipeline using Python, SQL, and MLflow for reproducibility and scalability

Senior Software Engineer | *Infosys Limited, Bangalore, India* **Sep 2019 – Mar 2024**

- Migrated legacy reports from Excel macros to automated SQL jobs, improving reporting efficiency & reducing support tickets by 45%
- Built an ML-powered EOD tracking dashboard, improving workflow efficiency by 10% and enhancing visibility for teams
- Implemented data validation protocols for financial operations, reducing data inconsistency issues by 25%
- Led the integration of Java APIs with IBM iSeries systems to automate retrieval of financial close prices from upstream applications, streamlining inter-system communication and reducing client report delivery time by 60%
- Integrated SQL-based ETL pipelines with upstream APIs and downstream dashboards for seamless data flow

Software Engineer Intern | *Infosys Limited, Mysuru, India* **Jan 2019 – May 2019**

- Designed a full-stack inventory management system (Java, Spring Boot, MySQL), ensuring scalability & reliability
- Implemented RESTful APIs, facilitating efficient data exchange between frontend and backend components
- Optimized data retrieval, reducing query execution time by 20% which improved inventory tracking accuracy

Projects

Energy Demand Forecasting & Automation **Mar 2025 – Apr 2025**

- Forecasted daily campus energy demand to support smart load distribution strategies
- ETL workflow using ApacheAirflow to schedule tasks that pull energy data and weather data from OpenWeatherMap API
- Transformed and merged multi-source datasets using pandas, and stored the cleaned data in AWS RDS PostgreSQL instance
- Built and optimized an LSTM model in TensorFlow for time series forecasting, evaluated using RMSE, MAE, and R^2
- Achieved 93% R^2 score, reducing peak load inefficiencies by 30% through accurate forecasting

Surge Fare Estimation: An ML-Based Forecasting System **Jan 2025 – Mar 2025**

- Automated an end-to-end ML pipeline for surge price prediction using weather & geolocation data
- Evaluated regression models (RandomForest, Gradient Boosting and XGBoost) based on RMSE, MAE and R^2 using MLFlow, identifying XGBoost as the best-performing model with low RMSE and R^2 score of 96%
- Utilized Hyperopt for hyperparameter tuning, optimizing model performance and better prediction accuracy
- Developed and deployed a Flask-based web interface for real-time fare predictions on AWS EC2 using Docker

GraphRAG: Knowledge Graph Retrieval-Augmented Generation with LLMs **Dec 2024 – Jan 2025**

- Enhanced a RAG pipeline by integrating a knowledge graph to improve contextual relevance and entity-level understanding
- Converted unstructured data into knowledge graph via Open AI API enabling entity linking & relationship extraction
- Improved answer relevance and factual accuracy by enabling concept-driven responses over traditional vector RAG models

Llama-Law: Legal Assistant Chatbot (RAG) **Oct 2024 – Dec 2024**

- Built a multilingual personalized legal chatbot to assist public navigate govt websites
- Improved response relevancy using Crawl4AI for targeted web crawling of legal and procedural content from Govt. website
- Enhanced accessibility to government services by providing clear, up-to-date legal assistance
- Fine-tuned the LLM model on legal advice conversation dataset from Hugging Face to deliver accurate and context-aware responses

Blind Assist: A Navigation System for the Visually Impaired **May 2019 – Aug 2019**

- Constructed an IoT device for visually impaired users, enabling real-time object detection and text recognition
- Integrated YOLOv2 CNN model with custom object detection, Tesseract for text recognition, and OpenCV for path detection
- Implemented Google Text-to-Speech API to facilitate seamless user interaction with live feedback

Skills

- **Tools and Programming Languages:** Python, R, SQL, NoSQL, MongoDB, Tableau, Power BI, Neo4j, Gephi, MLFlow, DagsHub, DVC, Apache Airflow, Weights & Biases (W&B), GitHub Actions, JIRA, Service Now
- **Key Skills and Competencies:** Statistical Data Analysis, Data Visualization (Matplotlib, Seaborn, ggplot2, Altair), Database Management, Predictive Modeling, Machine Learning (NumPy, Pandas, Scikit-Learn, PySpark, Decision Tree, SVM, Naïve Bayes), Deep Learning (Keras, TensorFlow, PyTorch, Computer Vision), Time Series Analysis (LSTM), Natural Language Processing (NLP), Large Language Models (Prompt Engineering, RAG, Fine-Tuning, Azure OpenAI), A/B Testing
- **Cloud Technologies and DevOps:** AWS (EC2, S3), GCP, Azure, Docker, CICD, Linux, Flask, Streamlit, REST APIs, ETL
- **Certifications:** IBM Professional Data Science, DeepLearning.AI TensorFlow Developer