

KEERTHANA VANGALA

vangala.keerthana20@gmail.com | +1 929 662 1511 | LinkedIn

Education

State University of New York at Buffalo, USA

M.S. in Computer Science

August 2024 – December 2025

GPA: 3.54/4.00

Courses: Machine Learning, Data-Intensive Computing, Statistical Data Mining, Database Systems (Capstone), Algorithms Analysis

G. Narayanamma Institute of Technology and Science, India

B.Tech. in Electronics and Communication Engineering

July 2018 – June 2022

GPA: 8.5/10.0

Courses: Python, Cloud Computing, Database Management Systems, Data Structures and Algorithms, Operating Systems

Skills

Languages:	Python, R, C, C++, SQL, Java, Bash, React, PySpark, Mainframe (COBOL), HTML, CSS
Data Engineering:	PyTorch, Apache Spark, Hadoop, Kafka, Hive, Airflow, Databricks, ETL Pipelines
Cloud & Infrastructure:	AWS (EC2, S3, Lambda, QuickSight, Redshift, SageMaker), Docker
Databases & Storage:	MySQL, MongoDB, Redis, Data Warehousing (Redshift, Snowflake), Supabase
Data Visualization	Tableau, Power BI, QuickSight, Plotly

Experience

Resilience

Software Development Intern

Remote, US

May 2025 – August 2025

- Designed and deployed an **AI-enabled Hazard & Security Risk Assessment (HVA/SRA)** web app using **React**, enabling real-time risk evaluation across **10+ teams** and improving assessment efficiency by **35%**.
- Integrated **Supabase (Auth, Postgres)** with dynamic relational models for secure storage and **real-time analytics**, boosting data retrieval speed by **40%**.
- Implemented RESTful APIs in Python and automated CI/CD pipelines via AWS Amplify, reducing deployment time by 30% in collaboration with a 4-member team. **Tech Stack:** React, Supabase, Python, AWS Amplify, CI/CD, REST APIs

Deloitte

Analyst

Hyderabad, India

September 2022 – August 2024

- Vanguard:** Migrated legacy **COBOL+DB2** mainframe applications to the **AWS Cloud** using **PySpark ETL**, reducing infrastructure cost 25%.
- Automated dependency mapping for hundreds of mainframe jobs via metadata parsing, cutting manual tracing 85 %.
- Developed scalable AWS Glue ETL workflows integrating DB2 outputs with S3, Redshift, QuickSight, processing 500+ jobs daily.
- Tech Stack:** Python, PySpark, AWS Glue, DB2, COBOL, S3, Redshift, QuickSight, Step Functions, ETL Pipelines
- Elevance Health:** Engineered the "**Timber**" AI-driven claim-verification platform automating validation of **ICD-10** and **CPT/HCPCS** codes.
- Integrated COBOL + DB2 backend with an AI-based rule engine detecting 2000+ billing and claim inconsistencies monthly.
- Implemented and tested **edit-trigger logic** for claim-pricing & compliance across payer networks.
- Created automated **Postman API** collections to simulate batch claims and ensure end-to-end data consistency.
- Optimized workflows, cutting review time 30% and saving **\$16.8 K/month**; visualized analytics via **QuickSight, React, Power BI**.
- Tech Stack:** Python, COBOL, DB2, React, AWS QuickSight, Power BI, Postman, ICD-10/CPT/HCPCS Validation, AI Rule Engine, Claim Edit Automation

Projects

Maternal Infant Insight – LLM-Augmented Risk Prediction

October 2024 – January 2025

- Programmed an **end-to-end ML pipeline** using **scikit-learn, XGBoost, pandas** with integrated **EDA** and statistical validation for maternal health data.
- Benchmarked **Logistic Regression, Random Forest, XGBoost** models ($F1 = 0.94$, $Recall = 0.93$); applied **SHAP** for interpretability.
- Integrated an **LLM-based explainer** using **GPT-3.5/T5** to summarize risk drivers for clinicians.
- Tech Stack:** Python, scikit-learn, XGBoost, SHAP, GPT-3.5/T5, NumPy, Matplotlib, EDA

Noti Research Mini – Agentic Multi-Agent Research Assistant

August 2025– October 2025

- Designed an **agentic AI system** with planner, summarizer, and citation agents orchestrated via **LangChain**.
- Architected a **FastAPI backend** supporting **JWT authentication, SQLAlchemy + SQLite** persistence, and modular orchestration.
- Developed a lightweight **HTML/JS** front-end for interactive research querying and visualization.
- Tech Stack:** Python, FastAPI, SQLAlchemy, SQLite, JWT, bcrypt, LangChain, OpenAI API, Docker

Predictive Loan Advisor – Machine Learning Capstone

October 2022 – November 2022

- Orchestrated an **ETL pipeline on Databricks (PySpark)** processing 50K+ loan records, including feature encoding, scaling, and correlation analysis.
- Identified key drivers: **Debt-to-Income Ratio, Credit History, LTV**, improving prediction accuracy.
- Trained and tuned **XGBoost, Random Forest, Logistic Regression** on **AWS SageMaker** ($AUC = 0.96$).
- Built dashboards using **AWS QuickSight, Power BI** for loan portfolio insights.
- Tech Stack:** Databricks, PySpark, SageMaker, S3, Aurora, XGBoost, QuickSight, Power BI, Python

Certifications

- AWS Certified Cloud Practitioner (CLF-C01) & AWS Certified Developer – Associate (DVA-C02)**
- Completed **Postgraduate Program in Software Engineering and Data Science** from Great Learning, in association with **IIIT Hyderabad**