

YATISH DURGA APPANAPALLI

Open to relocate • yatishdurga14@gmail.com • (660) 864-4977 • LinkedIn • Porfolio • Github

WORK EXPERIENCE

Data Engineer & AI Intern

May 2025 – present | Boston, USA

IAPP

- Designed AI frameworks with automated drift detection using **AWS CloudFormation** and **SageMaker Model Monitor**, cutting model debugging time by 30% and enhancing financial risk monitoring.
- Analyzed 80K+ e-commerce transactions with **Python**, **SQL**, and **AWS Glue** built **Power BI** dashboards using **Amazon Redshift** to support revenue forecasting and performance tracking for 15 stakeholders.
- Optimized AIGP use cases with **AWS Lambda**, improving operational efficiency and boosting feature adoption by 18%.
- Led cross-functional reviews leveraging **Amazon SageMaker** to improve ML model accuracy by 20%, accelerating deployment and enhancing financial decision-making.

Research Assistant(Data Science)

Aug 2024 – May 2025 | Lee Summit, USA

University of Central Missouri

- Processed and cleaned ECG R-peak data from MIT-BIH, preparing high-quality model inputs for heartbeat classification.
- Built CNN, XGBoost, and Random Forest models, achieving 98% accuracy, with results under peer-reviewed publication.

Data Engineer

Aug 2021 – Dec 2023 | Chennai, India

GAVS Technology

- Designed and maintained **ETL pipelines** using Cloud Composer (**Airflow**), **BigQuery** and Cloud Storage, optimizing workflows and reducing annual cloud costs by **\$45K**.
- Developed and tuned Snowflake SQL (**stored procedures, data partitioning, modeling**) transformation scripts, performing EDA and building **scalable pipelines** for ML models, increasing team productivity by 25%.
- Built Cloud Functions and Python scripts for PDF validation using Snowflake Tasks and serverless architecture, processing **1M+ records** for 3 clients and cutting manual review time by **45%**.
- Engineered data ingestion, cleansing, and **error logging** with **Snowpipe**, Teradata, and **GCP**, improving analytics reliability and supporting cross-business reporting.

Software Intern

May 2021 – Aug 2021 | Bengaluru, India

Tensile Tech

- Improved frontend efficiency by 20% using ReactJS, HTML/CSS, Bootstrap, and **GitHub**, integrating **REST APIs** with **Spark based ETL pipelines** for seamless ML model deployment.
- Implemented **Dockerized microservices** architecture with Docker, GitHub Actions, and **CI/CD**, incorporating data ingestion, transformation, and storage workflows with MLflow monitoring, reducing delivery time by **18%**.

EDUCATION

Masters in Data Science with AI, University Of Central Missouri

Jan 2024 – Dec 2025

Coursework : Machine Learning, Statistics, Python, ,Neural Networks, NLP, Computer Vision, Data Analysis, Deployments.

SKILLS

Cloud Platforms: AWS(S3, EC2, Lamda, S3, Glue, Sagemaker, Bedrock, Kiro)

Data Engineering: Databricks, Redshift, Postgres, Unity Catalog, Teradata, Oracle SQL Server, NoSQL, Terraform, Snowflake

Programming: Python, SQL, Spark, Pandas, NumPy, NLTK, Keras, TensorFlow, PyTorch, Docker, Kubernetes, MLflow

Visualization & Analytics: Power BI, Tableau, Alteryx, Matplotlib, Seaborn,

ACADEMIC PROJECTS

RAG-based Q&A Agent with LangChain

- Enhanced a RAG Q&A agent using LangChain, ChromaDB, and LLM, boosting query accuracy on unstructured PDFs by 40%.
- Optimized prompt formats and retrieval workflows through A/B testing, resulting in a 30% boost in answers

GenAI-Powered Mental Health Assistant

- Deployed a full-stack GenAI mental health assistant, using Streamlit, FastAPI, AWS Lambda, LangChain, OpenAI GPT, and FAISS, achieving a 45% boost in user query resolution rate through scalable RAG-powered support

Certifications & Achievements:

- Certified **AWS** cloud platform, Certified in **Artificial Intelligence Governance** Profession
- Databricks** Accredited Fundamentals, **Artificial Intelligence Governance Professional** Certified .