

Harish Muthyala

Address: 16200 Space Center Blvd., Houston, TX | **Mobile:** +1 (281) 965-2335

✉ harishcmuthyala@gmail.com

in [linkedin.com/in/harish-muthyala](https://www.linkedin.com/in/harish-muthyala)

github.com/harishcmuthyala

EDUCATION

Master's of Science in Computer Science , <i>University of Houston</i>	Aug. 2024 – May 2026
<i>Machine Learning, Generative AI, Advanced Operating Systems GPA: 3.8/4.0</i>	<i>Houston, TX</i>
Bachelors in Computer Science , <i>Vellore Institute of Technology</i>	Jul. 2018 – Apr 2022
<i>Data Structures, Algorithms, Computer Networks, Artificial Intelligence</i>	<i>Vellore, India</i>

TECHNICAL SKILLS

Languages: Python, Java, C/C++, SQL (Postgres, DynamoDB), JavaScript, HTML/CSS, R
Frameworks: Langchain, RAG, MLOps, Tensorflow, pandas, Numpy, Agents, React, Node.js, Flask, FastAPI
Cloud: AWS S3, Lambda, Sagemaker, Bedrock, EC2, ECS, VPC, Codepipeline, Quicksight, Azure Foundations
Developer Tools: Git, Docker, Kubernetes, Postman, VS Code, Jupyter
Credentials: AWS Solutions Architect - Associate, Accenture Trailblazer Award, AWS Article, Hawks Scholarship

EXPERIENCE

Generative AI Engineer	May 2023 – July 2024
<i>Senior Analyst, Accenture AWS Business Group (AABG)</i>	<i>Hyderabad, India</i>
<ul style="list-style-type: none">Architected Retrieval-Augmented Generation (RAG) pipeline, reducing manual underwriting processes by 75% for a credit underwriting workflow.Implemented OpenSearch Serverless vector DB with Titan embeddings, enabling fast retrieval of structured client data from Excel and other financial documents.Designed few-shot prompt templates, improving accuracy for generating complex financial notations.Engineered preprocessing pipelines using Pandas for data cleaning, chunking, and conversion, enabling robust and scalable ingestion of tabular financial data.	
Machine Learning Operations Engineer	Aug. 2022 – Jan. 2024
<i>Application Engineering Analyst, Accenture AWS Business Group (AABG)</i>	<i>Hyderabad, India</i>
<ul style="list-style-type: none">Deployed SageMaker Autopilot pipelines in secure VPC environments for telecom customer churn prediction, supporting scalable API-based inference.Implemented real-time model drift detection via SageMaker Model Monitor, enabling proactive model retraining strategies.Automated migration of QuickSight Dashboards across AWS accounts, preserving dataset integrity and improving reporting for analytics stakeholders.	
Information Technology Project Analyst	Oct. 2024 – Present
<i>Office of Information Technology, University of Houston</i>	<i>Houston, TX</i>
<ul style="list-style-type: none">Collaborated with cross-functional teams to manage and deliver technology projects aligned with university strategic goals, ensuring project visibility and successful implementation.Served as internal SME and consultant for migrating IT service tools from FootPrints to TeamDynamix, affecting 80+ IT staff and significantly improving workflow visibility and tracking.	

PROJECTS

Model Context Protocol <i>Python, Claude, Research, Langchain</i>
<ul style="list-style-type: none">Conducted comprehensive research on MCP architecture for LLM communication and context managementAnalyzed protocol integration patterns with frameworks like LangChain and demonstrated real-world applicationsImplemented MCP client-server architecture enabling seamless LLM-application communicationCreated implementation examples showcasing file creation and Google Maps integration via MCP servers
Exploratory Data Analysis on Customer Churn Prediction <i>Random Forest, Python, Jupyter, Pandas, Git</i>
<ul style="list-style-type: none">Developed a customer churn prediction model for the telecom sector using RF algorithm with .92 accuracyConducted extensive exploratory data analysis to identify key factors influencing customer attritionImplemented machine learning techniques to calculate individual customer churn probability, enhancing retention strategies