

THEJESH MALLIDI

+1 517-730-4554 | thejesh.mallidi18@gmail.com
github.com/thejesh-m | linkedin.com/in/thejeshm | www.thejesh.in | Irving, TX

EXPERIENCE

Quantiphi Analytics

Irving, TX, USA

Machine Learning Engineer

- Architected an **LLM-powered large document processing pipeline** for a Tier-1 telecom, transforming **2M+** complex contracts from unstructured legalese to structured graphical format preserving document hierarchy for superior data integrity.
- Empowered legal teams to surface risks and key contract terms across 2M+ documents and cut review cycles by **60%**, through a **GraphRAG-powered Agentic Contract Assist chatbot** with sub-second retrieval latency and **90%+ answer accuracy**.
- Automated deployment with **CI/CD** across **Red Hat OpenShift** and **GCP Cloud Run** to ensure seamless version upgrades and 99.9%+ uptime and integrated metadata storage solutions with **ElasticSearch**, **BigQuery** and **Google Cloud Storage** for fast, reliable contract analytics.

Michigan State University

East Lansing, MI, USA

Graduate Research Assistant - Applied AI

01/2024 – 05/2025

- Pretrained domain-adapted BERT (MLM) with domain-aware masking (boosted mask probability for curated terminology) on 1M abstracts, delivering **+15% average accuracy** on downstream tasks.
- Built a **domain-specific knowledge graph** with over **2 million entities** and **1.4 million relationships** to accelerate hypothesis generation in plant science, leveraging **Named Entity Recognition (NER)** and **Relation Extraction**.
- Fine-tuned LLMs, including BERT and LLaMA variants and created a **custom BERT-based multi-relation extraction model** for information extraction, achieving **94%** accuracy in entity extraction and **84%** in relation identification.
- Developed a **conversational research assistant** with **GraphRAG + agentic workflows (LangGraph)** to reason over the KG and literature.
- Led large-scale topic analysis of **700k evolutionary biology research abstracts**, uncovering temporal trends, research shifts, and literature gaps to inform future studies; leveraged **BERT-based topic modeling** and **Agentic AI** to extract ~150 high-quality topics.

Quantiphi Analytics

Bangalore, India

Machine Learning Engineer

05/2021 - 07/2023

- Automated document processing pipelines for Banking, Financial Services, and Insurance (**BFSI**) workflows using **computer vision** and **NLP-based models**, which reduced manual review time by **75%** and achieved **90%** data extraction accuracy.
- Delivered **10+ Proof of Concepts (PoCs)** on **document classification & extraction, object detection, image classification, and semantic segmentation** in collaboration with cross-functional teams of frontend developers and business stakeholders.
- Developed **insurance-focused predictive models** for risk assessment and claim outcome prediction using Logistic Regression, Decision Trees, Bagging, Random Forest, and Gradient Boosting Machines (GBM).
- Built and optimized real-time inference pipelines for Nvidia edge devices using **DeepStream** with **custom C++ plugins**, achieving 27 FPS for latency-sensitive applications.
- Orchestrated the deployment of **20+ machine learning models** into production environments, leveraging **CI/CD pipelines** across **AWS SageMaker, GCP Vertex AI, and Kubernetes**, decreasing model deployment time by **30%**.

SKILLS

- Core Expertise:** Data Science, Machine Learning, Deep Learning, Predictive Modeling, Natural Language Processing (NLP), Computer Vision, Large Language Models (LLMs), Generative AI, Agentic AI, RAG, Graph ML, MultiModal AI, LLM Orchestration
- Programming Languages:** Python, R, C++, C, Bash
- AI/ML Frameworks & Libraries:** PyTorch, TensorFlow, TensorRT, ONNX, LangGraph, LangChain, LlamaIndex, AutoGen, Vector databases, Hugging Face Transformers, OpenCV, NumPy, Pandas
- Big Data & Streaming:** Hadoop Spark, Apache Kafka, Apache Beam
- Databases:** MySQL, PostgreSQL, Neo4j Graph DB, Elastic Search
- Cloud & MLOps Tools:** GCP, AWS, Docker, Git, Flask, FastAPI, CI/CD Pipelines, Jenkins, MLFlow, KubeFlow

EDUCATION

Michigan State University, USA

08/2023 - 04/2025

Master of Science in Data Science

GPA: 3.95/4.0

Madanapalle Institute of Technology & Science, India

06/2017- 07/2021

Bachelor of Technology in Computer Science

GPA: 9.23/10.0