

Satwik Medipalli

AI Engineer / LLM Engineer | Multiagent Systems, RAG & Production Software
Richardson, TX | 945-544-2418 | sathwik.medipalli@gmail.com | LinkedIn | GitHub | Portfolio

SUMMARY

UT Dallas M.S. Business Analytics (May 2026) builder with three shipped internships and four live AI products: multiagent due diligence over SEC filings, calibrated LLM judge evaluation, a grounded S&P 500 research terminal, and fleet reliability ML with a claim verifying AI analyst. Every project is public and production deployed.

DEPLOYED PROJECTS

Diligence AI: Multiagent Due Diligence Platform

Spring 2026

LangGraph, CrewAI, Pinecone, RAG | Live Demo

- Built a six agent system that reads SEC 10K filings, extracts financials, classifies risk factors, and assembles citation backed analyst reports in minutes, with Pinecone RAG Q&A and RAGAS faithfulness gating every release

VoiceIQ: Evaluation Platform for Voice AI Agents

Fall 2025

Python, LLM as judge, Streamlit | Live Demo

- Shipped a testing harness that simulates conversations across scenario types, scores agents on five dimensions with self consistency checks, and validates the LLM judge with Bland Altman analysis plus precision/recall against reference scores

Lumen: S&P 500 Research Terminal with Grounded AI

Summer 2026

React, TypeScript, FastAPI, Python | Live Demo

- Built a full stack market terminal on live Yahoo quotes, SEC XBRL, and FRED data with a risk lab, portfolio backtests, and an AI layer constrained to explain only backend computed evidence with no invented numbers or causes

SiliconSentinel: Fleet Reliability with AI Analyst

Fall 2025

Python, XGBoost, Pydantic AI | Live Demo

- Delivered a reliability platform scoring 30 day device failure risk at 83% recall and 80% precision with survival models, SHAP deep dives, an ops RMA queue, and a Pydantic AI analyst that verifies model claims before operators act

EXPERIENCE

AI Research Intern

Jan 2024 - Apr 2024

Boltzmann Labs

- Built a format aware Python preprocessing engine that auto detects three biological data types across about 500K records at 85%+ accuracy and routes each file through the correct pipeline, replacing a manual prep workflow for the research team
- Turned a local ML prototype into a production cloud service by containerizing the backend with Docker, deploying on Kubernetes with AWS S3 storage, and exposing REST APIs consumed daily by the product and frontend teams

Data Analytics Intern

Aug 2023 - Dec 2023

Bostonlogix

- Shipped an Azure OpenAI NLP pipeline that extracts structured fields from about 100K unstructured text records and fully automates a previously manual review process for operations
- Migrated about 1M records into a SQL Server analytics store and audited and restored roughly 1000 failing API endpoints using Postman test suites across the platform

BI Developer Intern

Apr 2023 - Jul 2023

Ideabytes

- Built a real time ingestion pipeline streaming about 500K JSON sensor datapoints through APIs into Power BI, then presented the live operations dashboard directly to the CEO for review
- Developed regression and ARIMA models forecasting sensor performance at 92-95% one month accuracy and 80 to 90% at twelve months to schedule proactive maintenance before failures

EDUCATION

University of Texas at Dallas

Richardson, TX

M.S. Business Analytics

May 2026

GITAM University

Hyderabad, India

B.Tech Computer Science with Data Science

May 2024

TECHNICAL SKILLS

LLM & Agents: LangChain, LangGraph, CrewAI, OpenAI APIs, Pydantic AI, RAG, Pinecone, RAGAS, LLM as judge, prompt workflows, citation grounding

Software Engineering: Python, JavaScript/TypeScript, React, FastAPI, Flask, REST APIs, Docker, Kubernetes, Git, Postman

ML & Evaluation: Scikit learn, XGBoost, SHAP, survival analysis, model validation, Bland Altman, precision/recall

Data & Cloud: SQL, PostgreSQL, SQL Server, MongoDB, AWS S3, Azure, Pandas, NumPy, Streamlit