Vasantha Lakshmi Eda

+1 4807917472 | vasanthalakshmieda1@gmail.com | LinkedIn

SUMMARY

Full-stack **AI/ML professional** with **4+ years** of experience designing, deploying, and operating production AI systems across **LLMOps, NLP, machine learning, and data engineering**. Proficient across **the complete AI/ML lifecycle** from research, model development, and evaluation frameworks to scalable pipeline architecture and production delivery with hands-on expertise in LLMs (Claude 3, GPT, Azure OpenAI), RAG pipelines, BERTScore evaluation, AWS Step Functions, SageMaker, LangChain, PyTorch, TensorFlow, and MLOps tooling. Experienced across cloud platforms (AWS, Azure, GCP) and comfortable owning full-stack AI systems independently from architecture through frontend delivery. **AWS Certified Machine Learning** – Specialty Certificate Holder.

TECHNICAL SKILLS

Languages: Python, SQL, R (basic), Go, Java, Bash.

ML & DL: Regression, Classification, Clustering, XGBoost, LightGBM, RNN, CNN, LSTM, Transformers (BERT, GPT)

LLMs & Gen AI: Amazon Bedrock (Claude 3), Azure OpenAI, GPT, LangChain, LangGraph, RAG pipelines, prompt engineering, BERT Score evaluation, hallucination mitigation, Feature Engineering.

Frameworks: TensorFlow, PyTorch, Keras, Hugging Face Transformers, Scikit-learn, Streamlit, Matplotlib, Seaborn, Plotly

Libraries: Pandas, NumPy, NLTK, spaCy, OpenCV, Sentence-Transformers, Regex, FAISS, Boto3, JSON / YAML parsers, LangSmith

MLOps : MLFlow, Kubeflow, Airflow,Step Functions, SageMaker Pipelines, GitHub Actions, Azure DevOps, Jenkins, CI/CD Pipelines

Deployment: Docker, Kubernetes, Fast API, Flask, GitHub Actions

Data & Big Data: PySpark, Apache Spark, Kafka, Hive, Dask, Snowflake, Athena, Tableau, Power BI

Math & Statistics: Probability, Statistics, Linear Algebra, Calculus, Hypothesis Testing, A/B Testing.

Cloud: AWS (S3, EC2, lambda, SageMaker, Step Functions, Bedrock), Azure (Azure ML, Azure AI Search, Azure Open AI), GCP (BigQuery, Vertex AI)

Frontend & Automation: HTML, CSS, JavaScript, Microsoft Power Automate, n8n

PROFESSIONAL EXPERIENCE

Amazon – Data Scientist II — Oct 2025 – Apr 2026

- Architected an end-to-end **LLM-Powered Insights Automation System** delivering **monthly, weekly, daily** issue lifecycle reports, **reducing manual** reporting effort by **~80%** and enabling faster, data-driven decisions across teams.
- Designed and orchestrated scalable data pipelines using a **dual-package architecture** (**CDK**-based infrastructure for DAGs, S3, IAM, monitoring + Python science package), improving pipeline reliability and reducing failures.
- Orchestrated the end-to-end **inference pipeline** via an **AWS Step Functions** state machine sequencing data availability checks (**Lambda**), **SageMaker Processing Jobs** (data validation, transformation, top-contributor generation, Bedrock LLM inference, evaluation, and S3 writes) using **SQL and modular Python packages**, ensuring reliable execution across all pipeline stages.
- Developed AI-driven, **persona-based actionable recommendations** for builders, tool teams, and leadership by passing a rich **knowledge base** (business context, domain knowledge, and issue metadata) into Amazon Bedrock (Claude), improving decision consistency and operational efficiency.
- Implemented **evaluation frameworks** including ground-truth alignment (BERT-based similarity), field completeness, and data quality checks, achieving **93% AI output reliability** across all inference runs.
- Deployed a **Slack bot** that automatically delivers reports for **20+ MBR metrics** to targeted personas, improving visibility, reducing follow-ups, and streamlining communication.
- Automated an end-to-end **customer survey analysis pipeline** for **7+ tools** using **Python, NLTK, and spaCy** handling **tool-specific survey variations**, data ingestion, cleaning, transformation, **NLP-based sentiment analysis, and reason-grouping** on customer comments to generate structured CSAT insights, **reducing manual effort by ~60%.**
- **Built data ingestion, cleaning, and transformation** workflows handling tool-specific survey variations, ensuring scalable and accurate processing of engagement and feedback data.
- Delivered **structured CSAT insights** to product managers and leadership via **Amazon Quick Suite** dashboards including trend visualizations (line and donut charts), **reason groups, likes/dislikes** and satisfaction scores across tools, **accelerating product roadmap prioritization** and **improving tool satisfaction visibility** across teams.

CMG Financial – Agentic AI Prompt Engineer — July 2025 – Oct 2025

- Engineered a **RAG-based mortgage policy search** system using **LangChain, Azure OpenAI, and Azure AI Search** indexing 40+ policy documents to enable loan officers and customers to query guidelines via natural language, improving decision consistency and reducing manual policy lookup time.
- Launched an **AIO paydown simulator** as a chat interface using **Python, RAG, and n8n** serving **100+ end users** to model and analyze **loan repayment scenarios** conversationally, reducing dependency on manual financial calculations.
- Independently delivered an **AI-powered meeting transcript automation** system using an **event-driven architecture**, an event listener monitors drive recordings, triggers **a Python package to scrape transcripts**, and passes them to **OpenAI** for priority **extraction and summarization**, proactively shipped while the team awaited Microsoft Graph API permissions.
- Designed end-to-end **automation workflows** using **Microsoft Power Automate, n8n, and HTML/CSS/JavaScript**, integrating AI-driven applications **into frontend interfaces to ensure seamless user interaction** across all solutions.

GoTackle (Powered by GoLogic) – Data Engineer — Jan 2025 – June 2025

Volunteered with GoTackle through the Arizona State University Professional Training Program.

- Designed and implemented serverless data pipelines using **AWS Lambda, Redshift,** and **BigQuery** to process high-volume business data in real time.

- Built **AI-powered business insight** engines using **OpenAI GPT, Vertex AI**, and **LangChai**n to generate automated recommendations and reports.
- **Automated data transformations** using dbt and Airflow, ensuring timely and reliable analytics **for 10,000+ SMB users**.
- Engineered event-driven workflows with **AWS EventBridge and DynamoDB**, improving system scalability and reliability.
- Collaborated with cross-functional teams **to integrate AI-driven insights** into **user dashboards**, enhancing platform intelligence and usability.

Arizona State University - Research Assistant **Jan 2023- Dec 2024**

- Built a Retrieval-Augmented Generation **(RAG)** pipeline leveraging **knowledge graph triplets** to power semantic search and contextual recommendations, improving information retrieval accuracy **by 35%.**
- Fine-tuned pre-trained models (BERT, GPT-2) on custom research datasets using **Hugging Face Transformers**, enabling domain-specific question answering and text summarization.
- **NLP** pipelines using **SpaCy, Transformers**, and custom tokenizers to extract structured data from **10 M+** unstructured data from different sources, enhancing automated coding workflows and reducing manual review time by **40%**.
- Conducted comprehensive data cleaning and processing on datasets exceeding **500,000 records**, improving data quality **by 30%** and delivering actionable insights **through 10+ visualizations**.
- Developed **a ML model** for data matching, increasing data accuracy by 25% and reducing manual data reconciliation efforts by 60%.

Cardlytics – Data Science Analyst **Jul 2021- Nov 2022**

- Built a **Streamlit-based application** to aggregate and standardize merchant location data from multiple sources (Mastercard, AGG), improving processing efficiency by **4x** and reducing manual effort from **5 FTEs to 1**.
- Developed a **REST API for merchant location cleaning (MLC)** to normalize and extract phone numbers from bank transaction data, improving data quality for downstream analytics.
- Enhanced time-series forecasting of market trends using **XGBoost**, **LSTM**, and **LightGBM**, reducing operational risk by 25%. Deployed **ML models** via **FastAPI** and **Docker** with **RESTful APIs** for scalable integration.
- Collaborated with data scientists on a **Character-level CNN model** for merchant classification, achieving multi-class prediction across 14,000+ categories.
- Optimized complex **SQL queries** for large-scale financial data extraction, improving retrieval performance **by 25%.** Streamlined financial reporting using **SAP Business Objects**, reducing report generation time **by 30%.**
- Implemented **MLOps** for the full model lifecycle and established **CI/CD pipelines** with **Azure DevOps** for automated deployment.
- Developed and optimized batch and real-time **ETL pipelines** using **Python**, **PySpark**, **AWS**, and **Apache Kafka**, improving processing speeds by 15%.
- Automated batch processing with **Python scripts**, reducing manual workload by 50%. Conducted statistical analysis using **T-tests** and **ANOVA** to examine speaker traits in client datasets, supporting behavioral insights for financial advisory.

Cardlytics – Data Science Analyst Intern (Volunteer) **Feb 2021- Jun 2021**

- Designed an **unsupervised clustering model** to group millions of transaction records, leveraging fuzzy matching and n-gram analysis to map transactions to likely merchant brands using OpenStreetMap data.
- Took part in **documentation, code reviews, and knowledge sharing**, strengthening best practices across the AI/ML intern cohort.
- Built automated pipelines using **Python, web scraping, and Google search APIs** to create a location corpus mapping cities to zip codes, improving data coverage and accuracy.
- Pre-processed large financial datasets using **NumPy**, **Pandas**, and **Scikit-learn** for high-quality models.

EDUCATION

Master of Science in Information Technology (Arizona State University - AZ, USA) **(GPA: 4/4)**
Bachelor of Technology in Information Technology (Jawaharlal Nehru Tech University – India) **(GPA:8.3/10)**

EXTRA CIRCULAR ACTIVITES

Open Source Contributor – Shared ML notebooks and solutions on Kaggle, improving model performance and community learning.
Hackathon Winner – Won Smart India Hackathon 2021 for building an AI-driven NLP solution under real-world constraints.
Student Club Leader – Led the AI Club and organized technical sessions, project sprints, and college fest events.
Community Member – Engaged with MLOps. community and Papers with Code to exchange ideas and showcase ML projects.