VISHNU VARDHAN REDDY MUKKARA

St. Louis, MO | +1(314) 359-5914 | vishred06@gmail.com | LinkedIn | Portfolio

PROFESSIONAL SUMMARY

Generative AI professional with expertise in LLM fine-tuning, prompt engineering, RAG pipelines, and AI/ML model deployment. Skilled in GPT, Llama, BERT, Falcon, and transformer-based models using Lora, PEFT, and RLHF, with experience in Python, SQL, Power BI, Porch, TensorFlow, and Azure Cognitive Search. Proficient in building end-to-end AI/ML pipelines, synthetic data generation, multi-modal AI, and model explainability, delivering scalable enterprise solutions in healthcare, manufacturing, and product analytics. Experienced in predictive modelling, anomaly detection, and AI-driven automation, ensuring accuracy, interpretability, and measurable business impact.

TECHNICAL SKILLS

Generative AI & Large Language Models (LLMs): GPT, Llama, Falcon, Mistral, Claude, Palm, T5, OPT, Hugging Face Transformers, Lang Chain, RAG (Retrieval-Augmented Generation), Lora, PEFT, Prompt Engineering, RLHF (Reinforcement Learning with Human Feedback), Fine-Tuning, Instruction Tuning, Synthetic Data Generation, Vector Databases (Pinecone, FAISS, Weaviate, Milvus)

Machine Learning & NLP: Python (Pandas, NumPy, Scikit-learn, XGBoost, LightGBM), TensorFlow, PyTorch, Keras, spaCy, NLTK, Transformers (BERT, RoBERTa, DeBERTa, DistilBERT, GPT-NeoX), Transfer Learning, Named Entity Recognition (NER), Text Classification, Summarization, Sentiment Analysis, Speech-to-Text, Translation Models

Deep Learning & Computer Vision: CNNs, RNNs, LSTMs, GANs, Autoencoders, Diffusion Models, Vision Transformers (ViT), OpenCV, YOLOv8, Hugging Face Diffusers, Image Segmentation, Object Detection, Multi-modal AI (CLIP, BLIP, LLaVA)

MLOps & Data Engineering: MLflow, Kubeflow, Databricks, Airflow, Prefect, Docker, Kubernetes, FastAPI, Flask, Streamlit, Ray, Model Deployment & Monitoring, CI/CD (GitHub Actions, GitLab CI/CD, Jenkins), Feature Stores (Feast, Tecton), Experiment Tracking

Cloud & Databases: AWS (SageMaker, Bedrock, S3, Redshift, Athena, Lambda), Azure (AI Studio, Synapse, Data Factory, Cognitive Services), GCP (Vertex AI, BigQuery, Dataflow, AutoML), SQL Server, PostgreSQL, MySQL, Snowflake, MongoDB, Neo4j, Graph Databases

Data Visualization & BI: Power BI, Tableau, Looker, Plotly, Matplotlib, Seaborn, Dash

Methodologies: Agile/Scrum, SDLC, CRISP-DM

PROFESSIONAL EXPERIENCE

Generative AI Intern **Aug 2024 - Sep 2024**

Excelerate | St. Louis, MO

- Engineered RAG pipelines using Lang Chain, FAISS, and Azure Cognitive Search, enabling contextualized responses across 50,000+ documents and reducing LLM hallucinations by 35%.
- Fine-tuned LLaMA-2 and GPT-3.5 models with domain-specific datasets using Lora and PEFT, improving response accuracy by 28% and strengthening stakeholder confidence in AI outputs.
- Designed dynamic prompt templates for summarization, classification, and reporting, standardizing outputs across workflows and cutting manual rework by 22%.
- Automated compliance and audit reporting with GPT-4 integrated into Azure Cognitive Search, reducing preparation time by 40% and ensuring regulatory adherence.
- Developed interactive dashboards in Power BI and Streamlet, visualizing LLM outputs, confidence scores, and retrieval context, which accelerated stakeholder validation of AI-generated recommendations.
- Pre-processed and validated structured and unstructured data in SQL Server and Azure Synapse, ensuring 98% data integrity for downstream AI/ML pipelines and improving model reliability.

Machine Learning Intern **May 2022 - Jun 2022**

Bharat Heavy Electricals Limited |Hyderabad, India

- Transformed turbine IoT telemetry data using Python (Pandas, NumPy) and SQL, producing structured datasets that enabled predictive maintenance pipelines and increased turbine uptime by 20%.
- Designed and implemented time-series forecasting and anomaly detection models with Scikit-learn and MATLAB, detecting efficiency deviations and allowing proactive maintenance scheduling.
- Integrated ML outputs into Power BI dashboards, providing plant managers with real-time KPI monitoring and actionable insights to optimize turbine performance.
- Automated ETL workflows using Python and SQL, accelerating dataset availability for model training and reducing manual preparation effort by 40%.
- Created synthetic datasets through augmentation to address missing sensor readings, enhancing model robustness and improving prediction consistency under varied operating conditions.
- Reconciled SCADA and monitoring system data using SQL validation queries, ensuring 100% compliance with regulatory audits and eliminating reporting discrepancies across departments.

PROJECTS

Automated Headlight Intensity Controller for Vehicles

- Collected and processed real-time LDR sensor data using Python, engineered features, and trained regression models to adjust headlight intensity dynamically, reducing glare incidents by 35%.
- Implemented predictive ML logic in Python to adapt brightness under varying light conditions, improving visibility and safety for oncoming drivers.
- Built simulation pipelines and integrated Python-based ML models with embedded firmware, enabling real-time AI-driven control on vehicle hardware.

Cardiovascular Disease Risk Prediction Using ML

- Pre-processed and merged clinical and Framingham datasets using Python, Pandas, and NumPy, applied feature engineering, imputation, and encoding, increasing model accuracy by 22%.
- Trained and evaluated logistic regression, random forest, and decision tree models with Scikit-learn, producing reliable risk predictions and generating actionable insights via Power BI dashboards.
- Documented end-to-end ML workflows, including data pipelines, training scripts, and evaluation metrics, ensuring reproducibility and facilitating integration into healthcare AI systems.

EDUCATION

M.S. in Computer & Information Systems **Jan 2024 - Dec 2025**

Saint Louis University | St. Louis, MO

B.Tech. in Electronics & Instrumentation Engineering **Aug 2019 - Jun 2023**

VNR Vignon Jyothi Institute of Engineering and Technology | Hyderabad, India

CERTIFICATIONS

- Certified in Fundamentals of Artificial Intelligence
- Certified Internet of Things
- Certified Machine Learning Techniques
- Certified Deep Learning
- IBM AI Practitioner - **Coursera**
- Generative AI with Large Language Models - **Coursera**
- Generative AI for Data Scientists - **Coursera**
- AWS Generative AI Applications Professional Certificate **- Coursera**