



HARI SANKARA KOTESWARA LAL KELUTH

+1 (940)-843-7985 | harisankara.keluth@gmail.com | [LinkedIn](#) | [Github](#) | [Portfolio](#) ♦ WILLING TO RELOCATE

EDUCATION

University of North Texas,

MS in Computer and Information Science, GPA: 3.6/4.0

Aug 2023 – May 2025

VR Siddhartha Engineering College

BTech in Computer Science Engineering, GPA: 3.5/4.0

July 2019 – July 2023

PROFESSIONAL EXPERIENCE

AI/ML Intern | GrantAide | Hybrid (USA/ Remote)

Apr 2025 – Present

- Helped build AI-powered tools for document search and grant writing by working with frameworks like LangChain, FAISS, and OpenAI GPT-4, supporting smarter and faster content recommendations.
- Assisted in developing Flask-based APIs that connect to Firebase Firestore and Google Cloud Run, providing real-time features like text summarization, entity detection, and chat services for users.
- Contributed to automating document organization and improving search accuracy using PyTorch and Hugging Face Transformers, while learning to deploy these models for multiple users in a cloud environment.

Machine Learning Engineer | Cadential Technologies Pvt Ltd | India

Jan 2022 – Jul 2023

- Assisted in processing large datasets (over 6 million records) using PySpark and Elasticsearch, helping to improve data quality by supporting fuzzy matching and data management tasks.
- Helped develop and test machine learning models in Python with frameworks like TensorFlow, PyTorch, and Scikit-learn, supporting predictive analytics for business processes.
- Contributed to building NLP-based search and recommendation features using tools like BERT and Hugging Face Transformers, making it easier for users to find information in big datasets.
- Worked with deep learning models such as CNNs and RNNs for simple image and text classification tasks, helping improve automation in basic computer vision and language processing projects.
- Supported the deployment of machine learning models into production using Docker, Kubernetes, and CI/CD tools on AWS, and participated in routine testing and improvements for stable model performance.

AWS Cloud Intern | AICTE | India

Oct 2021 – Dec 2021

- Supported data processing and transfer tasks using AWS Glue, S3, and Lambda, helping the team improve how quickly and accurately data moved through our systems.
- Assisted in building and deploying machine learning models on AWS SageMaker and EC2, while learning to use tools like AutoML, Athena, and DynamoDB to help with model training and managing cloud data.

Projects

ChatPDF using Retrieval Augmented Generation (RAG) (GPT-3.5, FAISS, Hugging Face, OpenAI, NLP)

- Developed an AI-powered PDF chatbot using RAG architecture for contextual search and semantic content extraction, enabling accurate question-answering on unstructured documents.
- Leveraged FAISS for vector indexing and GPT-3.5 Turbo for response generation, enhancing retrieval precision and language understanding in document-centric queries.

Insurance Cross-Sell Prediction (FastAPI, Docker, AWS ECS, MLflow, CI/CD, DVC, Evidently AI)

- Built a production-ready MLOps pipeline for cross-sell prediction including data ingestion, transformation, DVC-based versioning, and MLflow-driven experiment tracking to ensure performance reproducibility.
- Deployed the model with FastAPI on Docker inside AWS ECS, integrated CI/CD via GitHub Actions, and used Evidently AI for real-time drift detection and monitoring in production.

SKILLS

- Machine Learning & AI:** Deep Learning, NLP, Computer Vision, LLMs (GPT-4, Gemini, Llama, Hugging Face Transformers), RAG, LangChain, FAISS, TensorFlow, PyTorch, Scikit-learn
- Data Engineering & Big Data:** SQL, MongoDB, Snowflake, ETL, Data Pipelines, PySpark
- Cloud & MLOps:** AWS (Glue, Redshift, EC2, SageMaker, Lambda), GCP, CI/CD, Kubernetes, Docker
- Software Development:** Python, Java, C, React.js, Flask, Node.js, REST APIs.
- Data Analytics & Optimization:** Tableau, Power BI, Feature Engineering, A/B Testing.

Certifications

[Google Professional ML Engineer](#) | [AWS Solutions Architect – Associate](#) | [AWS Cloud Foundations](#)