HARIKRISHNAN S

Data Analyst,AI/ML Engineer
harikrishnansivadasan@gmail.com | +918075665216 | Sivahari North Aryad P.O. Alappuzha
Linkedin | HackerRank | GitHub | LeetCode

Profile Summary

Aspiring Data Scientist with strong expertise in data analysis, statistical modeling, and machine learning. Handson experience in building data pipelines, feature engineering, and deploying AI/ML solutions including LLM-powered applications, predictive analytics, and deep learning models. Proficient in Python, TensorFlow/Keras, SQL, and modern AI/ML frameworks such as LangChain and Hugging Face, with practical exposure to MLOps (Docker, GitHub Actions CI/CD, AWS, FAISS). Passionate about applying data driven insights, scalable pipelines, and advanced AI techniques to solve complex challenges in healthcare, automation, and business analytics.

Education

Cochin University of Science and Technology — Kochi
Computer Science with specialization in Data Science Masters — July 2022 - April 2024
CGPA: 8.04

KMEA Engineering College — Kochi
Mechanical Engineering B.Tech — June 2016 - August 2020
CGPA: 6.87

Experience

Brototype | Data Science AI/ML Bootcamp — Aug 2024 | Ongoing
Attending a self-learning bootcamp focused on Data Science and AI/ML. Gain hands-on experience in Machine Learning, Deep Learning, and Large Language Models (LLMs), while learning industry practices through collaboration with peers and mentorship from top industry professionals.

Palm Fibre India Pvt Ltd | ETP Plant Operator — Alappuzha, India | February 2021 - July 2022
As an ETP Operator cum Maintenance professional, I managed efficient Effluent Treatment Plant operations, adhering to environmental standards. I oversaw chemical handling, optimized processes, and maintained equipment. Quality control and safety compliance were top priorities, along with meticulous documentation and collaboration with cross-functional teams. I ensured regulatory compliance and preparedness for emergencies while keeping up to date with industry trends. My role combined technical expertise, problem solving, and a strong commitment to environmental sustainability, safety, and compliance.

Skills

Programming Languages:	Python, C, HTML, CSS
Libraries/Frameworks:	NumPy, Pandas, Matplotlib, scikit-learn, TensorFlow, Keras, Hugging Face Transformers, LangChain
LLMs / GenAI:	Groq , OpenAI API, Hugging Face Models
Tools / Platforms:	Git, GitHub Actions (CI/CD), Docker, Streamlit, Tableau, VS Code
Cloud / Deployment:	AWS (EC2, deployment)
Databases / Vector DBs:	SQL, PostgreSQL, FAISS

Projects / Open-Source

LetterForge – AI-powered Cover Letter Generator | Live Demo | GitHub *Python, LLMs, LangChain, Hugging Face, Docker, AWS*

Developed an AI-powered application that analyzes CVs and job descriptions to generate customized cover letters, featuring an agentic module for fetching similar job postings and an interactive chatbot for CV–JD queries. Designed pipelines for skill extraction, gap analysis, and personalized recommendations, integrating Hugging Face models and LangChain for efficient LLM orchestration. Implemented an interactive Streamlit dashboard for visualization and user engagement, and ensured modular, production-ready code through containerization with Docker. Deployed the system on AWS with GitHub Actions–based CI/CD for automated builds, seamless updates, and scalable performance in real-world usage.

Real-Time Health Monitoring System | GitHub *Python, Machine Learning, Streamlit, Sensors, Docker, AWS*

Built a real-time health monitoring dashboard that ingests patient sensor data, applies ML models to predict health risk, and triggers alerts for high risk patients. Designed with modular ML pipeline components for ingestion, transformation, training, and prediction. Deployed on AWS using Docker, ensuring scalability and reliability.

Oral Cancer Detection | GitHub *Deep Learning, TensorFlow, Python, React-Native, Git*

Developed an oral cancer detection system using CNN (ResNet) and built a React Native Android app for real-time cancer risk analysis. Showcased proficiency in DL model development and mobile app integration for critical healthcare applications.

Forecasting mRNA Abundance from Genomic Sequence | GitHub *Python, Deep Learning, VGG16, matplotlib*

Implemented a VGG16-based model to predict mRNA abundance from genomic sequences. Demonstrated the use of CNN architectures in computational biology and showcased deep learning applications in genomics research.

Importance of Feature Selection in Machine Learning | GitHub *Machine Learning, Python, matplotlib, numpy, pandas*

Performed a feature selection analysis on the Iris dataset using KNN, Logistic Regression, and Decision Tree algorithms. Evaluated filter, wrapper, and embedded methods to identify optimal feature subsets that maintained classification accuracy while significantly reducing model training time. Demonstrated how strategic dimensionality reduction improves computational efficiency without compromising predictive performance across all three algorithms. This systematic comparison provided practical insights into feature importance and model optimization for classic classification tasks

Certifications

- Basic Image Classification with TensorFlow - **Coursera**
- Build a Data Science Web App with Streamlit and Python - **Coursera**