Sumukh Acharya

Bengaluru | sumukh.acharya@gmail.com | 9972454072 | sumukh-acharya.vercel.app | GitHub | LinkedIn

Skills

- **Languages:** Python, R
- **Machine Learning and Deep Learning:** pandas, numpy, tensorflow, scikit-learn, Matplotlib, Seaborn, Pytorch, keras, Librosa, Statsmodels, LightGBM, Joblib, OpenCV, Random Forest, PCA, RFE, SelectK, BSO, XGBoost, SVM, CatBoost, KNN, Ensemble, CNN, LSTM, Siamese Networks, Autoencoders
- **Database and Big Data:** MySQL, Hadoop, Kafka, Spark
- **Operating System:** Windows, Linux
- **Web Dev:** HTML, CSS, TypeScript, JavaScript, React, Astro, Tailwind
- **Others:** Git, GitHub, Blender, Docker, Kubernetes, VSCode, MSOffice, Vercel, Streamlit, Figma, Canva

Experience

CODMAV, Research Intern | June 2024 - July 2024 **GitHub Link**

- **Tools Used:** pandas, numpy, scikit-learn, Matplotlib, seaborn, XGBoost, CatBoost, SVM, KNN, PCA, RFE, SelectK, BSO, Random Forest
- Led the development of a **lung cancer risk prediction system** using advanced feature selection (**PCA, BSO, RFE, SelectKBest**) and machine learning models (**XGBoost, CatBoost, SVM, KNN**), achieving **98.75% accuracy**; co-authored on a research paper published on **IEEE Xplore**.

Projects

DFOS – Distributed File Orchestration and Synchronization **GitHub Link**

- **Tools Used:** socket, TCP Protocol, ThreadPoolExecutor, Multi-Threading, Client-Server Architecture
- **Engineered** a secure, **multi-client file server** with user authentication, enabling upload, download, preview, deletion, and listing of user-specific files with **concurrent handling** and graceful shutdown support.

My Portfolio – Portfolio Website using Astro and TailwindCSS **GitHub Link**

- **Tools Used:** Astro, TailwindCSS, TypeScript, Vanilla JavaScript, Responsive Design
- Built a **responsive** portfolio website showcasing my experience, projects and achievements, incorporating custom **animations** and **smooth** user experience; deployed using **Vercel**.

RideWave – Fare Forecasting in Quahog City **GitHub Link**

- **Tools Used:** numpy, pandas, seaborn, matplotlib, scikit-learn, statsmodel, SARIMAX, XGBoost, VAR, Ensemble
- **Developed** a comprehensive **time series forecasting** and **pricing strategy** by performing data visualisation, outlier detection, and feature engineering; implemented **SARIMAX**, **Holt-Winters**, **XGBoost**, and **VAR** models with **ensemble** techniques across all the vehicle types (bike, auto, car) for RideWave in Quahog City.

Achievments

- IBM Data Science Professional Course by Coursera
- Learn Photorealism with Blender Course by PESU I/O
- Secured 7th place out of 364 teams in a Kaggle Data Analytics Hackathon
- Participated in HackNight organised by ACM and contributed to Open-Source

Education

PES University, B-Tech in Computer Science and Engineering 2022-2026

Coursework: Machine Learning, Databases, Deep Learning, Software Engineering, Cloud Computing, Big Data, Data Analysis, Statistics for Data Science, C/Java, Data Structures and Algorithms, Operating Systems, Computer Networks