SUMUKH ACHARYA

Bengaluru, India | +91 9972454072 | sumukh.acharya@gmail.com
sumukh-acharya.vercel.app | linkedin.com/in/sumukh-acharya-6859ab312/ | github.com/sumukhacharya03

EDUCATION

B.Tech in Computer Science and Engineering, PES University — Expected: **May 2026**

SKILLS

Languages:	Python, Java
AI/ML:	LLMs, PEFT, LoRA, QLoRA, Supervised Fine-Tuning
Data Science:	Pandas, NumPy, scikit-learn, PyTorch, TensorFlow, Librosa
Frameworks and Tools:	MySQL, Apache Kafka, Apache Spark, Docker, Git

EXPERIENCE

AI Intern — Sep 2025 - Present
Tvam Technologies — Bengaluru, India

- Fine-tuned large language models using PEFT techniques (LoRA, QLoRA) to improve training efficiency under constrained compute.
- Built data preprocessing and prompt engineering pipelines to enhance model performance and stability.
- Evaluated model variants using research literature and benchmarks to inform tuning and architectural decisions.
- Documented experimental results and insights to support iterative model improvement.

Python and Databases Teaching Assistant — Aug 2025 - Dec 2025
PES University — Bengaluru, India

- Mentored 150+ undergraduate students in Python programming and Database Management Systems (DBMS).
- Led hands-on lab sessions involving code debugging, SQL queries, and schema-level reasoning.
- Assisted faculty with grading, assessment design, and preparation of instructional materials.

Data Science Intern — Jun 2024 - Aug 2024
CODMAV Research Centre (PES University) — Bengaluru, India

- Led end-to-end development of a Lung Cancer Risk Prediction System using healthcare data.
- Implemented robust data preprocessing pipelines, including Random Forest based missing value imputation.
- Applied feature selection and dimensionality reduction techniques (PCA, BSO, RFE, SelectKBest).
- Trained and optimized ML models (XGBoost, CatBoost, SVM, KNN) and built an ensemble classifier.
- Achieved 98.75% accuracy and 96.25% recall on RFE-selected features; results validated by a medical professional.

PROJECTS

DFOS | Python, Sockets, Threading

- Designed a secure, multi-client file server with authenticated, role-isolated file access.
- Implemented concurrent upload, download, preview, deletion, and listing of user-specific files using multithreaded socket handling.

Grid Pulse | Python, Apache Kafka

- Engineered a real-time, event-driven data pipeline using Apache Kafka within a decoupled microservice architecture.
- Developed a live Streamlit dashboard consuming Kafka topics to visualize real-time driver availability trends.

Dynamic Fare Engine | Python, Time Series Analysis

- Modeled fare dynamics across bikes, autos, and cars using statistical and ML based forecasting techniques.
- Compared Holt-Winters, SARIMAX, XGBoost, and VAR models and combined them via ensemble learning.
- Assessed forecasting accuracy using SMAPE to inform dynamic pricing strategies.

PUBLICATIONS

- Acharya, S., et al. "Predictive Analytics for Early Lung Cancer Risk using Machine Learning" IEEE Xplore, 2024

CERTIFICATIONS

- IBM Data Science Professional Course by Coursera

ACHIEVEMENTS

- Ranked 7th out of 364 teams in Kaggle Data Analytics Hackathon-1