ALEN K AJI

Email: alenkaji12@gmail.com | GitHub: https://github.com/AlenKAji | LinkedIn: https://www.linkedin.com/in/alenkaji/
Phone: +91 8594009559 | HackerRank : https://www.hackerrank.com/profile/b22cs0821

PROFILE SUMMARY

Machine Learning Engineer with hands-on experience building and deploying ML models in healthcare and environmental forecasting domains. Skilled in Python, TensorFlow, PyTorch, and model optimization. Experienced in full ML pipelines from data acquisition to production deployment.

ACADEMIC QUALIFICATION

IIT Guwahati   *Mar 2025 – Present*
- Artificial Intelligence and Machine Learning

Mar Athanasius College of Engineering   *Oct 2022 – Present*
- Computer Science Engineering
- CGPA: 9.3

WORK EXPERIENCE

ReverTech IT Solutions – Machine Learning Intern   *Jan 2025*

- Developed and deployed a Gradient Boosting Machine (GBM) model using scikit-learn for flood prediction, achieving 94% test accuracy.
- Engineered end-to-end ML workflows: data cleaning, exploratory data analysis, and model tuning.
- Collaborated with 3+ cross-functional teams to deliver AI-powered features in production.

SKILLS

Machine Learning: Scikit-learn, TensorFlow, PyTorch, Pandas, NumPy, Matplotlib

Programming Languages: Python, Java, SQL, C

Data Handling & Modelling: Data preprocessing, Feature engineering, Hyperparameter tuning, Cross-validation, Model evaluation metrics

Tools & Platforms: Git, GitHub, VS Code, MySQL, Excel, Google Workspace

Soft Skills: Stakeholder Management, Teamwork, Problem-Solving, Communication

POSITION AND RESPONSIBILITIES

- Sponsorship Team Lead, ENCIDE Club
- Volunteer, Training & Placement Cell

PROJECTS

Neural Network for Biomarker Prediction in Cancer Immunotherapy (GitHub)
- Developed a deep learning model using TensorFlow and PyTorch to classify RNA-seq data; achieved an AUC-ROC of 0.71 on unseen validation data.
- Performed dimensionality reduction and feature selection, reducing input features by 99% and enhancing model interpretability.

Flood Prediction System (GitHub)
- Built a predictive AI model leveraging Gradient Boosting and Random Forest to forecast floods based on precipitation and terrain data, achieving 94% classification accuracy.
- Implemented a robust data processing pipeline analysing precipitation patterns and terrain characteristics to deliver actionable insights for emergency response teams and disaster management authorities

Movie Genre Classification Model (GitHub)
- Implemented multiple supervised learning algorithms (Logistic Regression, Random Forest, Neural Networks) to classify movies into 20+ genres using textual metadata; achieved up to 67% accuracy.
- Engineered a preprocessing pipeline involving TF-IDF vectorization and multilabel binarization, improving feature quality and reducing dimensionality by 40%.

CERTIFICATIONS

- NPTEL Certificate – Programming, Data Structures & Algorithms using Python

AREAS OF INTEREST

Machine Learning, Bioinformatics, and AI-driven software development with a focus on deployment and optimization. Passionate about cybersecurity, full-stack development, scalable system architecture, and movies.