Soham Gore

Mumbai, India | acad.soham@gmail.com | linkedin.com/in/sohamgore | github.com/debug-soham

Education

KJ Somaiya School of Engineering, Mumbai **2024 - Expected 2028**
Bachelor of Technology in Electronics and Computer Engineering *CGPA: **8.98** / 10 (After 1st Year)*
Honours in Data Science and Analytics

Technical Skills

Programming & Tools: Python, NumPy, Pandas, Scikit-learn, SQLite3, Java, C, Git, Docker

Data Analysis: EDA, Statistics, Probability, Hypothesis Testing, Data Visualization, Probability Distributions

Machine Learning: Linear & Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, XGBoost, SVM, Naïve Bayes, KNN, Clustering, PCA, Anomaly Detection

Feature Engineering & Optimization: Feature Selection, Data Scaling, Encoding, Regularization, Hyperparameter Tuning, Cross-Validation

Deployment: Flask, Streamlit, Docker (Basics)

Experience/Internships

Open Source Contributor **October 2025 - October 2025**
Hacktoberfest
- Enhanced the functionality of PaperScope, an AI-powered research assistant, demonstrating skills in Streamlit and AI API integration.
- Submitted improvements to PyDeepFlow, a Python package for deep learning, reinforcing knowledge of ML frameworks and package structures.
- Helped refine GNews (a Google News scraping API) and spewer (a debugging library), showcasing versatility in data extraction and software tooling.

Open Source Contributor **July 2025 - October 2025**
GSSoC - GirlScript Summer of Code
- Enhanced etsi-watchdog, a real-time data drift detection library, and etsi-failprint, an MLOps tool for root cause analysis of model failures.
- Implemented features for an AI-powered Driver Drowsiness Detection System, applying computer vision (CV) skills to improve road safety.

Technical Team Member **July 2025 - Present**
Datazen Somaiya *Mumbai - On-site*
- Collaborating on data-centric community projects focused on AI literacy and ethical ML.
- Leading events and hackathons as part of the technical team.

AI & ML Intern **June 2025 - July 2025**
Elevate Labs *Virtual*
- Engineered a Convolutional Neural Network (CNN) in Keras/TensorFlow that achieved 75% accuracy in classifying 10 distinct music genres from the 1,000-track GTZAN dataset by analyzing Mel-Frequency Cepstral Coefficients.

Projects

Music Genre Classification | *Deep Learning - TensorFlow, Keras, Librosa, NumPy*
- Developed a Convolutional Neural Network (CNN) achieving approx 75% accuracy on the GTZAN dataset for classifying 10 music genres.
- Extracted Mel-Frequency Cepstral Coefficients (MFCCs) as key audio features to train and validate the model.
- Implemented efficient data preprocessing and model optimization for balanced genre recognition.

Personal Finance Risk Classifier | *Machine Learning - Python, Scikit-learn, Pandas, Seaborn, CLI*
- Built a Gradient Boosting model to classify users as Conservative or Aggressive investors based on financial data.
- Engineered risk profiles using demographic and behavioural metrics, integrated into an interactive CLI tool.

Relevant Courses & Certifications

Machine Learning Specialization **June 2025**
DeepLearning.AI, Stanford Online