

# Aditya Ravi

adiravi1104@gmail.com | GitHub: [Link](#) | Leetcode: [Link](#) | LinkedIn: [Link](#)

Passionate AI and data science enthusiast with experience in machine learning, web development, and full-stack applications. Skilled in building scalable solutions and eager to apply knowledge to real-world challenges.

## Education

<b>KJ Somaiya School of Engineering</b> – BTech in Artificial Intelligence and Data Science (9.76 CGPA)	June 2023 - Present
<b>St. Gregorios High School</b> – 11th & 12th Grade ISC Science (91.4%)	June 2021 - June 2023
<b>St. Gregorios High School</b> – 10th Grade ICSE Science (94.3%)	March 2009 - May 2021

## Skills

**Languages:** Python, C, Java, R, JavaScript

**Database:** MongoDB, PostgreSQL

**Frameworks & Libraries:** React.js, Node.js, Flask, NumPy, Pandas, Scikit-learn, TensorFlow,

**Tools & Platforms:** Git, GitHub, Jupyter Notebook, Linux, Tableau

## Experience

**AI/ML Intern**, Godavari Biorefineries Ltd– Fort, Mumbai Jan 2025 – Present (on pause due to company side delays)

- Developing an AI-powered impurity detection system using YOLOv8 to identify and separate impurities in sugarcane at varying levels of granularity.
- Contributing to dataset creation, camera equipment selection, and model training pipeline to ensure accurate impurity detection.
- Collaborating with cross-functional teams to refine detection accuracy, enhance real-time processing, and integrate the system into the production workflow.

**Machine Learning Intern**, Suvidha Foundation– Nagpur, Maharashtra May 2025 – July 2025

- Fine-tuned 6 advanced news summarization models (BART, Pegasus, DeepSeek-v1, FLAN-T5, T5, BigBird-Pegasus) on real-world datasets, achieving top scores (e.g., BART ROUGE-1: 0.44, ROUGE-2: 0.20, ROUGE-L: 0.30).
- Engineered and optimized a Multi-Document Summarization (MDS) pipeline using PRIMERA, increasing ROUGE-1 from 0.36 to 0.44 and ROUGE-L from 0.18 to 0.20 through targeted fine-tuning—improving summary quality and coherence by over 22% and 8% respectively.
- Drove social impact by raising ₹20,000+ through foundation-led fundraising campaigns, benefiting over 1,500 community members via coordinated outreach and execution.

## Projects

**Churn Retention Analysis** Jan 2025 - May 2025

([GitHub Repo](#))

- Developed an end-to-end machine learning pipeline for customer churn prediction using a dataset of 780K+ customer records over a 90-day period.
- Achieved 88% AUC and 83% recall on test data using a Voting Ensemble model with RandomForest, LightGBM, and XGBoost, leveraging behavioral and trend-based engineered features.
- Designed and deployed an expected tenure regression model (MAE: 2.3,  $R^2$ : 0.939), improving churn segmentation granularity and enabling time-sensitive retention targeting.
- Delivered business-focused churn insights based on early tenure patterns and usage behaviors, helping prioritize interventions for high-risk user cohorts.

## Certifications

- Machine Learning Specialization (Andrew Ng, Stanford University) ([certificate 1](#), [certificate 2](#), [certificate 3](#))