ADITI GUPTA

+91-9667664778 | guptaaditi.0825@gmail.com | LinkedIn: guptaaditi8 | GitHub: aditi-gupta-git

EDUCATION

Vellore Institute of Technology (VIT) Bhopal — Bhopal, India
B.Tech in Computer Science and Engineering (Specialization: AI & ML) — Aug 2023 – Jun 2027
CGPA: 8.81/10
Relevant Coursework: Data Structures & Algorithms, OOP, DBMS, OS, Linear Algebra

DAV Public School — Gurugram, India
Senior Secondary Education (Grade XII) – 93.2% — Apr 2022 – Mar 2023

TECHNICAL SKILLS

Programming Languages: Python, C++, Java, SQL
Machine Learning & AI: Machine Learning, Deep Learning, TensorFlow, Scikit-learn
Data Science & Analytics: Pandas, NumPy, Matplotlib, Seaborn, Feature Engineering, Exploratory Data Analysis (EDA), Statistical Analysis
Tools & Platforms: Flask, AWS, Render, Git, GitHub, Jupyter Notebook, VS Code, DWSIM, LaTeX

EXPERIENCE

Machine Learning Intern | *FOSSEE, IIT Bombay (Remote)* — Oct 2025 – Present

- Developing ANN-based predictive models for crude oil characterization using data from 50+ crude oil assays, estimating crude properties faster and more accurately than traditional mathematical correlations
- Contributing to open-source engineering modeling initiatives integrating machine learning with process simulation tools for chemical engineering applications

PROJECTS

Binary Distillation Surrogate Model | GitHub — Sep 2025
ML-Based Process Simulation Approximation using DWSIM — *Python, XGBoost, Random Forest, DWSIM, Scikit-learn*

- Developed machine learning surrogate models to replace computationally expensive DWSIM process simulations, predicting distillate purity (xD) and reboiler duty (QR) with XGBoost achieving R^2 of 0.95 and 0.93 respectively
- Generated and curated 373-sample dataset through systematic DWSIM simulations by varying 6 key operating parameters (Reflux Ratio, Boilup Ratio, Feed Mole Fraction, Feed Flowrate, Number of Stages, Feed Thermal Condition)
- Performed comprehensive EDA using univariate, bivariate, and multivariate analysis techniques; identified Reflux Ratio and Feed Mole Fraction as primary drivers of separation purity, and Feed Flowrate as key factor in energy demand
- Implemented and compared 6 regression algorithms (Linear, Polynomial, Random Forest, AdaBoost, SVR, XGBoost) with hyperparameter tuning via RandomizedSearchCV; validated models using residual and parity plots

Student Performance Predictor | GitHub | Live Demo — Sep 2025
End-to-End ML Pipeline with Web Deployment — *Python, CatBoost, XGBoost, Flask, Render*

- Built end-to-end ML system analyzing impact of demographic, socioeconomic, and academic factors on student mathematics performance; architected modular pipeline (Data Ingestion, Transformation, Model Training) ensuring scalability
- Engineered automated preprocessing pipeline using Scikit-Learn ColumnTransformer with StandardScaler for numerical features and OneHotEncoder for categorical variables, handling 1,000 student records across 8 features
- Trained and evaluated 10+ regression algorithms (Linear, Ridge, Lasso, Random Forest, XGBoost, CatBoost, AdaBoost, Gradient Boosting) with GridSearchCV optimization; Ridge Regression achieved best performance with R^2 of 0.88
- Deployed production-ready Flask web application on Render with custom exception handling and logging; created responsive UI enabling real-time predictions based on user-input student characteristics

CERTIFICATIONS

Applied Machine Learning in Python – University of Michigan
AWS Technical Essentials – Amazon Web Services (AWS)
Cloud Computing Certification (Elite) – NPTEL

LEADERSHIP & EXTRACURRICULAR

Head Girl – Represented student body and coordinated 15+ school-wide initiatives
House Captain – Led Mother Teresa House team to win House Cup through organizing inter-house competitions