Aditi Gupta

+91-9667664778 | guptaaditi.0825@gmail.com | linkedin.com/in/guptaaditi8/ | github.com/aditi-gupta-git

Education

Vellore Institute of Technology (VIT) — **Bhopal, India**
B.Tech in Computer Science and Engineering (AI & ML Specialization) — Aug 2023 - Jun 2027
Current CGPA: 8.81/10

Dayanand Anglo Vedic (DAV) Public School — **Gurugram, India**
Grade XII: 93.2% — Apr 2022 - Mar 2023

Delhi Public School — **Gurugram, India**
Grade X: 92.6% — Apr 2020 - Mar 2021

Skills

- **Languages:** Python, Java, C++, SQL
- **Data Science/ML:** Pandas, NumPy, Scikit-learn, Matplotlib, Power BI
- **Tools:** Git, Jupyter Notebook, MATLAB, VS Code
- **Concepts:** Data Structures, Algorithms, Operating Systems, Probability & Statistics, Computer Networks

Projects

Student Performance Predictor — Sep 2025
End-to-End Machine Learning Pipeline and Deployment

- Designed and executed an end-to-end regression project to predict student math scores, managing the full lifecycle from data ingestion and model development to deployment as an interactive web application.
- Engineered a modular and reusable data processing pipeline using Scikit-learn to automate data cleaning, feature scaling, and transformation, ensuring consistent data handling for both training and real-time inference.
- Developed and evaluated multiple regression models, including CatBoost and XGBoost, to identify the most accurate predictor based on R-squared and MAE metrics.
- Implemented a highly modular source code structure with dedicated components for data ingestion, transformation, and model training, promoting reusability, reproducibility, and scalability.
- Productionalized the final regression model by building and deploying a lightweight Flask web app.
- Tools: **Python, Scikit-learn, Pandas, CatBoost, XGBoost, Flask**

Surrogate Modelling for Binary Distillation — Sep 2025
Surrogate Modeling and Process Simulation with DWSIM

- Developed machine learning surrogates to predict **distillate purity** (x_D) and reboiler duty (Q_R), enabling rapid alternatives to computationally expensive process simulations.
- Generated and curated a dataset of 373 samples through DWSIM simulations; applied rigorous data cleaning, validation, and unit consistency checks to ensure reliability.
- Employed a block-based validation strategy (holding out unseen operating regions) to evaluate model extrapolation beyond standard train–test splits.
- Designed reproducible ML pipelines with automated preprocessing, feature scaling, categorical encoding, and hyperparameter optimization via `RandomizedSearchCV`.
- Benchmarked multiple algorithms including Polynomial Regression, Random Forest, AdaBoost, SVR, and XGBoost. XGBoost delivered the best performance, achieving $R^2 \approx 0.95$ for x_D and 0.93 for Q_R with the lowest error.
- Conducted diagnostic analyses (parity plots, error distributions, residuals) and feature importance studies. Identified Reflux Ratio and Feed Mole Fraction as key drivers of separation efficiency, and Boilup Ratio and Reflux Ratio as primary determinants of energy demand.
- Tools: **Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, XGBoost, DWSIM**

Leadership & Extracurricular Activities

- Served as **Head Girl** at Delhi Public School; represented student body and coordinated school-wide initiatives.
- Served as **Mother Teresa House Captain**; led team to win the House Cup by organizing and motivating peers in competitions.
- Anchored multiple school events, improving **public speaking** and **communication skills**.