



Registration Number – 22BCY10154
Email: sauryadavaligarh@gmail.com
Phone: +91 7983322687
[Leet Code](#)
[LinkedIn](#)
[GitHub](#)

Saurabh Yadav

Technical Skills:

- Python : ★★★★★
- C++ : ★★★★★
- Machine Learning : ★★★★★
- NumPy : ★★★★★
- Pandas : ★★★★★
- Matplotlib : ★★★★★
- Scikit-Learn : ★★★★★

EDUCATION

Board	Tenure	Educational institution	CGPA/Percentage
B. Tech (CSE)Cyber	2022 –2026	VIT Bhopal University	8.14/10
Class XII	2021 - 2022	St. Fidelis School, Aligarh	64.4%
Class X	2019 - 2020	St. Fidelis School, Aligarh	80.3%

PROJECTS

Credit Card Fraud Detection using PySpark and Machine Learning	<i>Duration: April 2025</i> <ul style="list-style-type: none">• Designed and implemented a scalable fraud detection system using PySpark to process large-scale credit card transaction datasets (143 MB), achieving a Test AUC of 0.9286, showcasing proficiency in handling big data for real-world applications.• Applied machine learning techniques with a Random Forest Classifier and SMOTE to address class imbalance, improving recall for fraud detection by 15% compared to baseline models.• Optimized data preprocessing pipelines with Pandas and PySpark, reducing processing time by 20% through efficient data cleaning and transformation, aligning with data science workflow needs.• Evaluated model performance using 3-fold cross-validation, achieving precision, recall, and F1-score of 0.9993, with ongoing efforts to refine metrics for production readiness.• Developed visualizations (confusion matrix, feature importance) using Matplotlib and Seaborn to support data-driven insights, enhancing collaboration with cross-functional teams.• Initiated a CI/CD pipeline with GitHub Actions for automated linting (flake8) and testing (pytest), demonstrating early experience in deployment workflows.• Managed project with Git for version control, ensuring reproducibility and code quality. Key Skills: PySpark, Machine Learning, Random Forest, SMOTE, Data Preprocessing, Big Data, Python, Pandas, Matplotlib, Seaborn, Scikit-Learn, SQL (Basic), GitHub Actions, CI/CD, Data Visualization, Cross-Validation, AUC, Precision, Recall, F1-Score

EXPERIENCE

Unified Mentor	Remote Machine Learning Intern (September 24 – October 24) <ul style="list-style-type: none">- Collaborated on innovative projects in the field of Machine Learning, contributing to the development and optimization of algorithms for ML-applications.- Designed and implemented an Animal Image Classification system and Forest Cover Type Prediction using different ML models in Google Colab notebook.- Conducted data analysis and visualization to derive insights from complex datasets, aiding in decision-making processes for project strategies.- Gained hands-on experience with various ML frameworks and tools, enhancing skills in Python, SVM, and image classification techniques.

CO-CURRICULAR

Coding	▪ Leet Code – 40+ Questions solved Hacker Rank – 4 Star in C++
Patent	▪ Indian Patent granted for an innovative evaporative air conditioner design.

ADDITIONAL INFORMATION

Languages	▪ English, Hindi
-----------	------------------