

LALIT SINGH AHAKEY

+91 6265746174 <mailto:lalits26@iitk.ac.in> | [linkedin.com/in/lalit-singh-ahakey](https://www.linkedin.com/in/lalit-singh-ahakey) | github.com/Lalit-Ahakey

EDUCATION

Degree / Certificate	Institute	CGPA / %	Year
M.Tech (Department of Management Sciences)	Indian Institute of Technology, Kanpur	In Progress	2026 – Present
B.E. (Civil Engineering)	Jabalpur Engineering College	6.77 / 10	2016 – 2020
Higher Secondary Education	Jawahar Navodaya Vidyalaya, Singodi, Chhindwara	88.40%	2016
Secondary Education	Jawahar Navodaya Vidyalaya, Singodi, Chhindwara	9.6 / 10	2014

SELF PROJECTS

Employee Attrition Prediction | Machine Learning, Classification

July 2026

Objective: Develop a predictive classification model to identify employee turnover risk.

- **Preprocessing:** Engineered a robust dataset by handling missing values, extracting logical features, and addressing a severe class imbalance (84:16 ratio).
- **EDA:** Performed univariate and bivariate analysis; statistically validated findings using the Mann-Whitney U test, Chi-squared test, and Spearman correlation; checked features for multicollinearity.
- **Models Used:** Logistic Regression, Random Forest Classifier.
- **Result:** Logistic Regression outperformed Random Forest Classifier (ROC-AUC 0.789 vs 0.779, PR-AUC 0.552 vs 0.410, “Left”-class recall 0.64 vs 0.26).

Insurance Charges Prediction | Machine Learning, Regression

June 2026

Objective: Develop a regression model to forecast individual medical insurance charges based on demographic and lifestyle features.

- **Preprocessing:** Executed comprehensive data preprocessing, including categorical encoding and a strict 80/20 train-test split to prevent data leakage.
- **EDA:** Visualized multi-variable relationships and used correlation heatmaps to isolate the highest-impact predictors.
- **Model Used:** Linear Regression.
- **Result:** Achieved an R^2 score of 0.783 (explaining ~78.3% of variance), with an RMSE of 5796.28.

Car Price Prediction | Machine Learning, Regression

June 2026

Objective: Develop regression models to estimate the market value of used vehicles based on vehicle specifications, brand, and historical usage features.

- **Preprocessing:** Constructed an automated data pipeline to handle scaling, median imputation for numeric features, and one-hot encoding for categorical variables.
- **EDA:** Analysed the target variable (“selling price”) distribution and corrected heavy right-skewness using a log-transformation, optimizing the data shape for linear modelling.
- **Models Used:** Linear Regression, Random Forest Regressor.
- **Result:** Random Forest performed best, achieving a test R^2 of 0.901 (~90.1% variance explained) with a 5-fold cross-validated mean R^2 of 0.908.

COURSEWORK & SKILLS

Current Coursework (*in progress): Probability & Statistics* | Operations Research for Management* | Introduction to Computing* | Data Mining and Knowledge Discovery*

Programming: Python

ML Libraries: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn

Databases & Tools: MySQL (in progress), Excel

Soft Skills: Analytical Ability, Decision Making, Problem Solving, Communication