

EDUCATION				
Degree/Certificate		Institute/Board	CGPA / %	Year
M. Tech (Department of Management Sciences)		Indian Institute of Technology, Kanpur	–	2025 - Present
B. Tech (Industrial and Production Engineering)		NIT Jalandhar	7.73 CGPA	2019-2023
Higher Secondary Education		Board of Secondary Education, Rajasthan	85.60 %	2018
Secondary Education (CBSE)		Central Board of Secondary Education	9.2 CGPA	2016
PROJECTS				
CO2 Emission Prediction Machine Learning Regression (GitHub Link) (Self Project)				June 2025
Objective	The objective of the project is to analyze the influence of various variables on CO2 emissions in vehicles and identify the most significant features affecting CO2 emission.			
Approach	Conduct Exploratory Data Analysis (EDA), Correlation Analysis on the provided dataset to Investigate relationships between different vehicle features and CO2 emissions. Utilize statistical techniques, and visualizations (Prediction Error visualization, for a Linear Regression model which helps to understand how well Linear Regression model is performing) to assess the influence of variables. Toolset: NumPy, Pandas, Matplotlib, Seaborn			
Result	Performance achieved with Linear Regression: R ² Score 0.874, RMSE 20.516 The visualization helps to identify patterns in the prediction errors made by Linear Regression model.			
Customer Churn Prediction Machine Learning Classification (GitHub Link) (Self Project)				July 2025
Objective	To develop a binary classification model to accurately predict which customer is likely to leave			
Approach	This machine learning project used 20 independent variables across 7043 observations. For data preprocessing , the data was cleaned, and both One-Hot Encoding and feature scaling were performed. The analysis phase involved an Exploratory Data Analysis (EDA) , including Univariate and Bivariate Analysis , to gain insights. The models used were a Decision Tree Classifier and a Random Forest Classifier . To optimize performance, the project handled imbalanced data using SMOTE and performed PCA . Toolset: Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn.			
Result	Random Forest Classifier with SMOTE gave balanced results: Recall = 0.96, Precision = 0.94, F1 Score = 0.95.			
Customer Segmentation Machine Learning Clustering (GitHub Link) (Self Project)				August 2025
Objective	Segmented the customers into different clusters using demographic data, marketing-specific metrics and spending behavior data.			
Approach	Data Exploration & Visualization: Verified data integrity (no missing values), created histograms, count/scatter/violin plots to analyze distributions and relationships (Age, Income, Spending Score). K-Means Clustering: Applied Elbow Method (inertia) to determine optimal clusters across feature subsets (Age-Spending Score, Income-Spending Score, and combined). Visualized boundaries in 2D and employed Plotly for 3D cluster analysis. Toolset: Python (NumPy, Pandas), Matplotlib, Seaborn, Plotly, scikit-learn.			
Result	Successfully segmented customers into five distinct groups using K-Means clustering, allowing for targeted marketing strategies based on demographic and behavioral data.			
COURSEWORK & SKILLS				*in progress
Relevant Courses	Statistical Modelling for Business Analytics Probability & Statistics Operations Research for Management Introduction to computing			
Skills/Tools	Python ML Libraries: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn SQL MS Excel			
Certifications	- Scaler Academy Data Science content creation <i>internship</i> certification - SQL Basic certification – Hacker Rank - Python certification – Hacker Rank			
ACHIEVEMENTS & EXTRACURRICULAR				
- GATE DA 2025 (Data Science and AI): All India Rank 663, 98.84th Percentile - Actively participated in organizing the annual cultural fest activities as a Member of Team Cultural Affairs at NIT Jalandhar.				