

ANIKET JHA

+91 8210561852

aniketj26@iitk.ac.in

LinkedIn

GitHub

Education

Degree/Certificate	Institute	CGPA/%	Year
M.Tech (Department Of Management Sciences)	IIT Kanpur		2026-2028
B.Tech (Mechanical Engineering)	Panipat Institute of Engineering and Technology, Haryana	7.8	2019–2023
Higher Secondary Education(CBSE)	DAV, Kathmandu, Nepal	67.0	2018
Secondary Education (CBSE)	Damodar Academy, Nepal	9.6	2016

Projects

Marketing strategy to predict the most probable buyers from loyalty program participants

Machine Learning — Classification — Self Project

Objective

- Developing a predictive model using data analytics and machine learning to **identify the most probable buyers** for the new organic product line
- Enabling organisation to target a specific set of customers with personalised marketing messages and other sales techniques, thereby **optimising marketing strategy** and **increasing product penetration**

Approach

- Extracted and cleaned customer data** using spreadsheets and MySQL; handled nulls, missing data and encoded categorical features for ML use
- Built and validated a **logistic regression model (Scikit-learn)** to classify buying behaviour; used 80:20 train-test split
- Applied the trained model** to **predict purchase probabilities** for the remaining customer base and ranked them based on buying likelihood
- Shortlisted high-probability customers (>65%)** for focused marketing, enabling efficient targeting and resource optimisation
- Toolset:** Spreadsheets, MySQL, Python (NumPy and Pandas), Machine learning (Scikit-Learn, Joblib)

Result

- The machine learning model had an accuracy score of 81%
- Compared three scenarios: **(i) no model-based targeting, (ii) top 40% predicted buyers, and (iii) top 30% predicted buyers**
- 3rd scenario resulted in the highest estimated profit, demonstrating the effectiveness of model-based selective targeting

CO2 Emission Prediction

Machine Learning — Regression — Self Project

Objective

- The objective of the project is to **analyse the influence of various variables** on CO2 emissions in vehicles and identify the **most significant features** affecting CO2 emissions .

Approach

- Conduct **Exploratory Data Analysis (EDA), Correlation Analysis** on the provided dataset to investigate relationships between different vehicle features and CO2 emissions.
- Utilise **statistical techniques and visualisations** (Prediction Error visualisation, 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: **R² Score 0.874, RMSE 20.516.**
- The visualization helps to identify **patterns** in the **prediction errors** made by Linear Regression model.

Coursework & Skills

**In Progress*

Relevant Coursework	Probability and Statistics — Operation Research for Management — Data Mining and Knowledge Discovery — Introduction to Computing
Technical Skills	Programming: Python Machine Learning: TensorFlow, NumPy, Pandas, Matplotlib, Scikit-learn Tools: Tableau*, Power BI*, VS Code, Jupyter Notebook, Excel*, Word, Powerpoint
Soft Skills	Leadership — Communication Skills — Analytical Reasoning — Problem Solving — Curiosity