

Sanjay Chhaba

✉ sanjaychhaba20@gmail.com | 📞 8000310803
🌐 github.com/Sanjay-Chhaba | [linkedin.com/in/sanjay-chhaba/](https://www.linkedin.com/in/sanjay-chhaba/)

Skills

Languages: C/C++, kotlin, Python, JavaScript, SQL

Technologies & Tools: AWS, PowerBI, Tableau, Pandas, numpy, sklearn, Machine Learning, Deep Learning, Langchain, Tensorflow, NLP, GenAI, prompt engineering

Work Experience

Simyog Sept 2024 - Dec 2024

Analyst

- Designed interactive dashboards with SQL and **Apache Superset**, providing stakeholders with actionable insights to support data-driven decisions.
- Engineered data transformation pipelines, creating optimized models for seamless data processing and platform-wide usage.

Turing, USA remote May 2024- July 2024

Data Scientist

- worked on fastAPI integration of AI in their projects using langchain to call LLMs.
- Enhanced the llms by giving the RLHF training and few shot learning.

Hanyang University, South Korea May 2023 - July 2023

Data Scientist

- Formulated a Clustering algorithm which is based on Fuzzy clustering problem where the algorithm does the image segmentation with the accuracy upto 98 %. Mainly involve into developing of mathematical intuition of algorithm.
- Attained accurate classification into **3 distinct classes**, yielding a visual figure that showcased the algorithm's report. It can be used into cc-tv cameras or can be use for face detection.

Education

Indian Institute of Technology, Guwahati Sep 2020 - May 2024

B.Tech. in Biosciences and Bioengineering

CGPA: 7.68/10

Relevant Coursework: Structural Analysis ,Databases, DSA and Cloud Services to Deploy projects, Machine Learning and Deep learning for drug Discovery, Data Mining, Image Processing.

Project Work

Predictive Sales Forecasting using Power BI Aug 2023 - Nov 2023

- Built an interactive sales forecasting dashboard using historical sales data (2019–2021), applying DAX for time-series analysis, rolling averages, and predicting next 15 days of sales.
- Visualized trends by product category, region, and customer segment, uncovering 10 % YoY growth, seasonal peaks (Q4), and key drivers like laptops and office supplies.

NYC Taxi Time Prediction Nov 2022 - Dec 2022

- Built a predictive model that predicts the total ride duration of taxi trips in New York. Dataset is based on the 2016 NYC yellow cab trip record data which contains 15,00,000 trip records and 11 features.
- Univariate and Bivariate analysis was done to get more insights into the dataset, followed by PCA for feature extraction. Multiple Linear Regression and XGBoost Regression models were used.

Building a Happy Workforce April 2023

- Presented case studies exemplify multifaceted strategies to elevate employee **satisfaction and productivity**. These initiatives target the root causes behind a **3% market share reduction** and a **20% decline** in service quality.
- Formulated strategy by dissecting the problem chart, articulating issues, and setting a **timeline with KPIs**. Aims: **15% revenue boost, 30% product quality increase** for enhanced customer satisfaction.
- Tools used such as **Porter's Five Forces, SWOT analysis** and cycle of problems.