# Tarun Chauhan

101 700E220020

tarunchauhan4272@gmail.com

+91 7895328830



Greater Noida, UP

Github LinkedIn Medium Leetcode

Data Science enthusiast with a strong academic background in mathematics and computer science, experienced in data analysis, Machine Learning, and proficient in Python, SQL. Committed to utilizing data-driven insights to solve complex real-world problems..

## Education

## GL Bajaj Institute of Technology and Management,

B.Tech in CSE, Specialization in Artificial Intelligence and Machine Learning

Greater Noida, UP

2020 - 2024

## Skills

Languages and Tools : C, C++, Python, SQL, Git, GitHub

Libraries & Frameworks : Numpy, Pandas, Matplotlib, Seaborn, Plotly, PySpark, Sk-Learn, Flask

Data Science & Machine Learning : Data gathering, Data cleaning, EDA, Feature engineering, Feature selection &

Extraction, Supervised and Unsupervised learning algos.

Mathematics for ML & DL : Statistics, Probability, Linear algebra, Matrices

Tools for Data Analysis : Tableau, Power bi, Excel

Databases : MySQL

# **Experience**

## **Data Scientist Intern | Innomatics Research Lab**

Jan 2023 - May 2023

- Performed web scraping using Beautiful Soup to generated dataset of 650 smartphones from smartprix website.
- Conducted data cleaning, assessment and exploratory data analysis.
- Apply machine learning algorithms to make predictions based on the data.

#### Data Analyst Internship | KPMG

Nov 2022 - Jan 2023

- Conduct data analysis and extract meaningful insights, including data quality assessment.
- Use data visualization techniques to represent complex information, in the form of charts, graphs, and tables.
- Work closely with the data analytics team.

#### Machine Learning Internship | Feynn Labs

July 2022 - Oct 2022

- Worked on Al & ML projects that help small businesses.
- Created a clothing classification deep learning model using CNN with accuracy of 92%.
- Performed exploratory data analysis on McDonald dataset using visualization techniques.
- Developed a machine learning-based credit card fraud detection model with a 91% accuracy rate.

# **Project**

## Movie Recommender System

Source code

This is a content based movie recommender system which trained on a dataset containing 5000 movies, This is an end to end machine learning project with GUI web application, which recommend the movies using cosine similarity. Library used: Numpy, Pandas, Sk-Learn, NLTK, Streamlit

#### **Big Mart Sales Forecasting**

Source code

An Al Product / Service Prototyping for Small Business, which can help the big marts for rapid growth by predicting the sales for products.

Library used: Numpy, Pandas, Matplotlib, Seaborn, Sk-Learn

## **Certifications**

- Python(basic) | SQL(basic) | SQL(intermediate) | SQL(advance) HackerRank
- · Python for Data Science IBM
- Machine Learning 1stop.ai

## **Extra Curricular**

- Worked with Nvidia Deep learning training program on DGX A100 server
- Badge of completion for the Open-Access Data Science & Advanced Analytics Virtual Experience Program Issued by **BCG**