Vishal Shukla

vishalshukla034@gmail.com
+91-8800768034
Faridabad,HR

Github **LinkedIn** **Portfolio**

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

CHANDIGARH UNIVERSITY MOHALI, PUNJAB
B.Tech in CSE 2019 - 2023

Skills

Languages and Tools	: C, Python, SQL, Git, GitHub, Heroku
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,ANN,Neural-Networks.
Mathematics for ML & DL	: Statistics, Probability, Linear algebra, Matrices
Tools for Data Analysis	: Tableau, Excel
Databases	: MySQL

Virtual Internships

Data Visualiation Internship | TCS Oct 2023 - Nov 2023

- Completed a simulation involving creating data visualizations for Tata Consultancy Services.
- Prepare dashboards on the data after cleaning the data.
- Created visuals for data analysis to help executives with effective decision making.

Data Analyst Internship | MeriSKILL Dec 2023 - Jan 2024

- 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.

Data Scientist Internship | PYSLIQ Jan 2024 - Mar 2024

- Performed web scraping using Beautiful Soup to generated dataset of 9980 companies from Ambition box and 5000 movies data from TMDB.
- Recommender System- Content Based - data preprocessing , feature engineering, model training and deployment at HEROKU.
- Developed a machine learning-based credit card fraud detection model with a 93% accuracy rate and f1 score -0.91.

Project

MOVIE RECOMMENDER SYSTEM || **Libraries:** Numpy, Pandas, Sk-Learn, NLTK, Streamlit 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.

CREDIT RISK MODELLING || **Libraries:-**Pandas, Scikit-learn, and XGB,Feature Engineering. Source Code
In my Credit Risk Modelling project, I developed a robust predictive model to assess credit risk using Python, leveraging libraries such as Pandas, NumPy, Scikit-learn, and XGBoost. I performed comprehensive data cleaning, feature engineering, and applied techniques like SMOTE for imbalance correction. The model's performance was evaluated using metrics such as AUC-ROC and precision-recall. This project demonstrated my ability to handle complex datasets, implement machine learning algorithms, and provide actionable insights for financial risk management.

VRINDA CLOTHING STORE ANALYSIS || **Tools:-** Excel, Pivot Tables & Pivot Charts , Data Cleaning. Source Code
The analysis of Vrinda Store's annual and monthly performance, conducted through a dynamic dashboard with pivot tables and slicers, revealed key insights: 64% of customers are women, primarily adults. Maharashtra, Karnataka, and Uttar Pradesh contribute 35% of total revenue. Peak sales occur from January to March, with March having the highest orders and 92% delivery success. Amazon, Flipkart, and Myntra are the top sales channels. To boost revenue, targeted promotions for women through these channels are recommended.