Prit Italiya

+1-431-990-2685 | prititaliya2244@gmail.com | www.linkedin.com/in/prititaliya/ | https://prititaliya.github.io/portfolio/

Summary

Aspiring Data Scientist with expertise in machine learning, data engineering, and software development. Skilled in building recommendation systems, analysing large datasets, and creating scalable pipelines using Scikit-learn, Pandas, and NumPy to drive data-driven decisions.

Skills

Data Analysis: Data Cleaning, Data Visualization, Statistical Analysis, ETL (Extract, Transform, Load), Data Integration
Data Tools: Pandas, NumPy, Matplotlib, Seaborn, R, Excel, SQL, Google Data Studio
Software Engineering: Object-Oriented Programming, Design Patterns, Algorithms, API Integration
Languages: Python, Java, JavaScript, C, C++, R, PHP, SQL
Version Control: Git, GitHub
Database Management: MySQL, MongoDB, PHPMyAdmin, Google BigQuery
Cloud Platforms: Google Cloud (GCP), AWS (basic knowledge), Zapier Automation
Machine Learning Techniques: Linear Regression, Logistic Regression, Classification, K-Nearest Neighbours (KNN)
Machine Learning Libraries/Tools: Scikit-learn, Pandas, NumPy
Machine Learning Concepts: Data Preprocessing, Model Evaluation (Accuracy, Precision, Recall), Feature Engineering
Machine Learning Applications: Predictive Modelling, Classification Problems, Data Analysis

Education

University of Manitoba
B.C.Sc., Computer Science   Jan 2023 - May 2026

Gujarat Technological University
Diploma in Engineering, Information & Technology   Aug 2019 - Jul 2022

Work Experience

Naos Blockchain Capital
Data Engineering Intern   Nov 2024 - Present
- Developed and optimized data-driven business intelligence dashboards for a live MVP environment in collaboration with co-founders. Leveraged SQL, Python, Google Data Studio, and Tableau to improve decision-making processes by transforming raw data into actionable insights
- Implemented scalable data pipelines and advanced analytics workflows using Google Cloud Products, ensuring efficient data capture, cleansing, and processing. Demonstrated expertise in predictive modeling, database management, and end-to-end solution design, which improved data accuracy and facilitated strategic insights through collaborative GitHub version control.

CSSA
Accessibility Advocate   Aug 2024 - Present

Projects

Socioeconomic Income Analysis in Mexico
- Analysed average daily salary and income distribution across socioeconomic levels in Mexico, focusing on disparities and correlations by region and occupation.
- Conducted correlation and distribution analyses to identify significant economic disparities across socioeconomic levels.
- Found that A/B level employees have the highest average income, while E level shows economic challenges.
- Discovered that administrative roles yield higher incomes than operational ones, indicating occupation's impact on income.

Personalized Movie Recommendation System
- Built a robust recommendation system by integrating movie metadata and user ratings datasets to predict user preferences.
- Pre-processed data using Pandas, merging and cleaning datasets to create a user-item matrix with over 1,800 movies.
- Computed item-based similarities using cosine similarity to dynamically recommend movies based on user ratings.
- Optimized the recommendation pipeline by aggregating duplicate entries, normalizing data with StandardScaler, and efficiently handling sparse data matrices.
- Designed a Python-based algorithm to suggest top 10 similar movies for any input movie, enhancing personalized user experience.

College Website (Full-Stack)
- Led development of a faculty management website with responsive design, real-time updates, document sharing, MySQL integration, and user authentication using phpMyAdmin.
- Contributed to feature enhancements through user feedback.