ROHIT MANVAR

Linkedin: https://www.linkedin.com/in/rohit-manvar-141261216

Github: https://github.com/RohitManvar

Portfolio: https://rohitmanvar.github.io

Address: Vadodara, Gujarat, India

Email: rohitmanvar.dev@gmail.com

Mobile: +91 9662787336

EDUCATION

The Maharaja Sayajirao University, Baroda   Vadodara, India
Master of Computer Application (MCA)   **Aug 2024 - May 2026**

Atmiya University, Rajkot   Rajkot, India
Bachelor of Computer Applications (BCA), CGPA: 8.17   **Sep 2021 - Apr 2024**

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java
Data Engineering: Apache Spark, Pyspark, Snowflake
Cloud & Databases: AWS (EC2,S3,Elastic beanstalk), PostgreSQL
Frameworks & Libraries: Streamlit, Numpy, Pandas
Platforms: Visual Studio, Jupyter Notebook, VS Code
Tools & Concepts: OOP, Git/GitHub, Docker, ETL/ELT Pipelines

INTERNSHIP

Data Science Intern | Plasmid | Remote   **Nov 2024 - May 2025**

- Attended structured video lectures on data science tools and techniques before working on hands-on projects.
- Worked on a **Product Recommendation System** project using real-world datasets.
- Cleaned and transformed raw data using **Pandas**, **NumPy,** ensuring structured input for modeling.
- Built recommendation models using **FAISS** with C-contiguous and cosine similarity for normalized vector search.
- Visualized data insights and evaluation metrics using **Matplotlib** and **Seaborn**.

PROJECTS

Custom ELT Data Pipeline | Data Engineering Project   Link

Tech Stack: Python, Apache Airflow, dbt, PostgreSQL, Docker, Pandas

- Designed and implemented a custom **ELT pipeline** for automated data ingestion and transformation.
- Utilized Apache **Airflow DAGs** for scheduling, dependency management, and **workflow orchestration**.
- Built modular **dbt models** to perform SQL-based data transformation and analytics layer creation.
- Integrated **PostgreSQL** as the central data warehouse for structured storage and querying.
- Containerized the entire pipeline with **Docker** for portability and reproducibility across environments.

Movie Recommendation System | Web Application   Link

Tech Stack: Python, Streamlit, Pandas, NumPy, Scikit-learn, GCP

- Built a movie recommendation web app using **Streamlit** that suggests movies based on user preferences.
- Trained the model on the Kaggle Movie Dataset with preprocessing via Pandas and NumPy.
- Implemented **content-based filtering** using **Scikit-learn's TF-IDF Vectorizer** and cosine similarity.
- Integrated OMDb API for dynamic posters and details, and deployed on **Google Cloud Platform** (GCP).

CERTIFICATIONS

- Completed Data Science Internship Program at Plasmid (2024-2025)   Link
- Data for Business Analysts Using Microsoft Excel (Coursera by SAP-E2E)   Link