ROHAN DODIYA

Jersey City, NJ | P: +1 201-993-1446 | dodiyarohan.019@gmail.com | https://www.linkedin.com/in/rohandodiya/|
https://rohan19072.github.io/rohandodiya.github.io/

EDUCATION

NEW JERSEY INSTITUTE OF TECHNOLOGY **GPA: 3.81/4**
Masters in Data Science **09/2023 – 05/2025**

GUJARAT TECHNOLOGICAL UNIVERSITY **GPA: 8.75/10**
Bachelors in Computer Engineering **07/2019 – 05/2023**

WORK EXPERIENCE

DATA SCIENTIST July 2022-May 2023
DHRIMA SOLUTIONS LLP

- Developed predictive maintenance algorithms using NLP and achieved up to 90% accuracy by leveraging various machine learning techniques.
- Utilized SQL (PostgreSQL/MySQL) for query optimization, validation, and analytics in support of real-time dashboards and insights.
- Built and automated scalable ETL workflows to clean, integrate, and validate structured and unstructured datasets from multiple sources, improving data integrity for business reporting and analytics.
- Created dashboards in Tableau and Python for business stakeholders, increasing accessibility to key KPIs and performance metrics.

PROJECTS

SENTIMENT EXTRACTION MODEL

- Developed a user-friendly web interface using Flask, enabling seamless interaction with the sentiment extraction tool and providing real-time analysis from customer reviews.
- Refined and prepared textual data for sentiment analysis by implementing advanced preprocessing techniques, ensuring high-quality input for analysis.
- Utilized selenium for web scrapping, automating data extraction from online source to feed the model.
- Achieved an accuracy of 85% by training the model and further developed measures for recall, precision, and F1 score to conduct a thorough analysis.

PATIENT DEMOGRAPHIC& MEDICAL ANALYSIS USING SQL

- Developed complex SQL queries to analyze patient demographics, identifying age and gender distribution trends for improved population health insights.
- Investigated prevalence of medical diagnoses across demographic groups, uncovering patterns based on age and gender to inform targeted healthcare strategies.
- Analyzed visit patterns and appointment times to determine peak hours, aiding in resource planning and operational efficiency for healthcare providers.
- Designed risk stratification logic to classify patients based on smoking status and chronic conditions, enabling early identification of high-risk individuals.

SKILLS

- **Languages & Tools:** Python, R, SQL, Git, Excel, Power BI, Tableau, Jupyter, VS Code
- **Libraries & Frameworks:** Scikit-learn, Pandas, NumPy, Flask, Selenium, BeautifulSoup
- **Techniques:** Data Cleaning, EDA, Machine Learning, A/B Testing, NLP, , Web Scraping, Forecasting

CERTIFICATIONS:

- Excel Advanced Functions and Formulas (LinkedIn Learning)
- SQL-Intermediate (Hackerrank)
- SAP S/4 HANA Essential Training