

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-NEWARK, NJ
Masters in Data Science

GPA: 3.81/4
09/2023 – 05/2025

GUJARAT TECHNOLOGICAL UNIVERSITY-GUJARAT, INDIA
Bachelors in Computer Engineering

GPA: 8.75/10
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.
- Created educational dashboards and data visualizations using Tableau and Python, which enhanced understanding of complex data insights and supported strategic decision-making
- 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.
- Collaborated with cross-functional teams to identify KPIs and reporting needs.
- Leveraged SQL in PostgreSQL/MySQL to ensure efficient data extraction, validation, and delivery

PROJECTS

SENTIMENT EXTRACTION MODEL

- Built a web-based NLP model to extract sentiment from customer reviews in real-time.
- Employed Selenium for dynamic web scraping; implemented advanced text preprocessing.
- Achieved 85% model accuracy; evaluated using precision, recall, and F1-score.
- Designed front-end interface using Flask for live interaction and usability testing.
- Conducted performance testing and fine-tuning of model for production readiness.

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, Hypothesis Testing, Web Scraping, Forecasting
- **Databases:** MySQL, PostgreSQL

CERTIFICATIONS:

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