MAITRI MISTRY

mmistry3@umbc.edu | Mobile: +1 (443)-320-3273 | United States
Linkedin | Github

PROFESSIONAL SUMMARY

Data Analyst with 2+ years of experience in leveraging Python, SQL, and Power BI to drive data-driven decisions and optimize business processes. Proficient in advanced analytics tools (R, Apache Spark, Hadoop) and skilled in building ETL pipelines, predictive models, and interactive dashboards. Holds a Master's in Computer Science and a proven track record of transforming complex data into actionable insights to improve operational efficiency and business performance.

TECHNICAL SKILLS

Programming Languages: Python (Numpy, Pandas, NLTK, Scikit-learn, Keras, Tensorflow), SQL, PL/SQL, R, Java
Big Data & Cloud: Apache Spark, Hadoop, Microsoft SQL Server, Oracle, PostgreSQL, AWS (S3, EC2, Lambda)
Data Visualization: Tableau, Power BI (DAX, Power Query), Matplotlib, Seaborn, ggplot2
Data Pipeline: Informatica, Microsoft SSIS
Industry tools: Excel (VLOOKUP, PivotTables, Macros), Git, Jira, Jupyter notebook
Industry Knowledge: Data Warehousing, ETL, Agile, Data Structures and Algorithms, OOP

WORK EXPERIENCE

Data Analyst, Bank of America, *Gandhinagar* **June 2021 - June 2023**

- Analyzed historical data using **Excel** to identify trends and patterns, reducing default rates by 15%.
- Performed advanced data analyses using **SAS** and **Python** to identify risks and control deficiencies, increasing credit scoring model accuracy by 30%.
- Built and optimized **ETL** pipelines using **Informatica**, integrating data from 7+ sources, resulting in better data quality and availability for critical financial reporting.
- Enhanced data retrieval efficiency and reduced **SQL** query processing times by 25% through indexing and query optimization on **Oracle** database.
- Created **data visualizations** and dashboards using **Power BI**, enabling stakeholders to monitor key metrics and generate actionable insights to drive business strategies.
- Collaborated with cross-functional teams in **Agile environments** to align database management with business needs.

Data Analyst Intern, Nirma University, *Ahmedabad* **January 2021 - May 2021**

- Analyzed over 200,000 records using **Python** and **regression analysis** to identify critical parameters affecting manufacturing processes, leading to process improvements.
- Ensured high data quality for analysis by performing rigorous data cleaning and preparation using **Python**.
- Implemented machine learning models (**Random Forest, SVM, GBM, Neural Network**) to predict and reduce defects, achieving a 15% reduction in PCB defect rates and improving product quality.
- Designed dashboards and reports in **Tableau** to monitor assembly quality, achieving a 30% reduction in waste and costs through effective defect management.
- Published research on **"Parameter Optimization for Surface Mounter using a Self-alignment Prediction Model"**.

PROJECTS

Inventory Management Optimization | Python, Tableau

- Developed **statistical models** including **regression, ARIMA, and Exponential Smoothing** to forecast medical inventory quantities, achieving a MAPE below 10% and significantly enhancing inventory stocking decisions.
- Created interactive dashboards in **Tableau** to monitor inventory trends, sales performance, and profitability.

Twitter Sentiment Analysis Using Apache Spark | Python, NLTK, Pandas, Apache Spark, Tableau

- Processed 100,000+ tweets using **Apache Spark**, focusing on data cleaning and application of **NLP** techniques.
- Employed data visualization, generating word clouds and utilizing Latent Dirichlet Allocation (**LDA**) to extract and analyze public opinions.
- Implemented **Logistic Regression** and **Naive Bayes** to classify sentiments, achieving 90% accuracy.

Job Scam Detection | Python, R, SQL, Scikit-learn, Matplotlib

- Conducted exploratory data analysis **(EDA)** to identify attributes of fraudulent job postings, guiding feature engineering.
- Queried data using **SQL**, uncovering correlations between fraudulent behavior and job posting features.
- Fine-tuned predictive models to detect fake job descriptions, achieving 93% accuracy with **RFR**.

EDUCATION

University of Maryland - Baltimore County, Baltimore, MD **Expected: May 2025**
Master of Science in Computer Science **CGPA: 3.8/4.0**
Coursework: Machine Learning, NLP, Introduction to Data Science

Nirma University, Ahmedabad, Gujarat **2017 - 2021**
B.Tech. in Computer Engineering **CGPA: 3.23/4.0**