

Harsh Makwana

Los Angeles, CA | +1 (980) 327 9431 | harshmakwana181@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Experienced Data professional with a strong background in building scalable data pipelines, automating ETL workflows, and developing efficient data models. Skilled in performance optimization and cost reduction solutions. Seeking to contribute to a dynamic team by leveraging analytics to drive impactful business outcomes and innovation.

SKILLS

- **Languages:** Python, MySQL, JavaScript, R, C++, Scala
- **Databases:** CosmosDB, SQL Server RDBMS, PostgreSQL, Snowflake
- **Frameworks:** Sklearn, PyTorch, TensorFlow, Keras, NLTK, Pandas, NumPy, Seaborn, Matplotlib, Flask, Django
- **Tools and Services:** Microsoft Azure, AWS, Spark, Kafka, Hadoop, Docker, Kubernetes, Power BI, Tableau, Github

EXPERIENCE

Lugano

Nov 2024 - Jun 2025

Data Analytics Engineer

- Maintained real-time data integration pipelines across SQL RDBMS, Azure Storage, Dynamics 365, and NetSuite using Azure Logic Apps, Azure Functions, and Power Automate.
- Built data warehouse and data marts in Microsoft Fabric using medallion architecture and serverless features, reducing reporting turnaround time by 50%.
- Implemented CI/CD pipelines with GitHub and Shell scripting, reducing deployment time by 70%.
- Maintained and monitored 50+ reporting pipelines in Microsoft Fabric, ensuring consistent data accuracy and timely report generation.
- Optimized compute usage with dynamic scripting using UDFs, DAGs, and multithreading, reducing processing costs by 40%.
- Built monitoring dashboards for business operations and warehouse lifecycle using Azure HDInsight, Log Analytics, and Power BI, reducing issue detection time by 70%.
- Built semantic models using Fabric Notebooks and SQL Server, improving backend query performance by 60% for faster report loading.
- Developed executive KPIs in Power BI DAX, increasing performance management efficiency by 80%.
- Developed financial and ad-hoc reporting using SparkSQL and Power BI, reducing manual work by 90%.
- Built a custom GenAI chatbot using Azure OpenAI and Function Apps, reducing employee query response time by 80% and improving sales and operations visibility.

Outlier

Jan 2024 - Nov 2024

AI consultant - Data Scientist

- Optimized performance of LLMs and multimodal generative AI models by Unit testing, and prompt engineering.
- Specialized in Reinforcement Learning from Human Feedback (RLHF) to train AI models, successfully aligning their actions with human preferences to improve user satisfaction
- Developed models and algorithms in Scala, Swift, and Python for prototyping and production, leading to more efficient data processing and analysis
- Enhanced AI chatbots by integrating RAG techniques, improving response accuracy by 30%.

Numerator

Sep 2020 - Jul 2021

Data Analyst

- Automated data extraction using AWS SDK and integrated data from various sources by triggering Lambda functions, improving data processing efficiency
- Transformed data according to business needs using PySpark scripts in AWS Glue jobs, enhancing data accuracy and usability for stakeholders
- Implemented incremental data loading into AWS S3 using AWS Glue, reducing system load by 30%.
- Performed data analysis into AWS SageMaker and deployed models were integrated into data pipeline via Lambda functions and API Gateway.
- Collaborated with the data science team to develop predictive models leveraging techniques like deep learning and gradient boosting, leading to improved forecasting accuracy

New Portland Enterprise

Aug 2018 - Aug 2020

Data Analyst

- Implemented data pipelines in Azure Data Factory (ADF) using linked services to extract, transform, and load data from Legacy System.
- Developed real-time data extraction processes using Spark Streaming and Kafka by Batch Extraction with Micro-Batching.
- Executed data transformation using Azure Databricks by creating Pyspark notebooks and SQL scripts, ensuring data integrity.
- Utilized SSIS to load large datasets into Azure Synapse incrementally, improving data refresh efficiency by 30%.
- Documented Notebooks of the ETL pipelines for monitoring, auditing, and references.
- Utilized NLP for sentiment analysis and K-means for customer segmentation using Databrick Notebooks, enhancing customer understanding and targeting strategies
- Designed 50+ dynamic analytics dashboards using PowerBI, facilitating instant data-driven insights.

PROJECTS

Real-Time Stock Price Prediction Using LSTM

- Developed an LSTM-based model to predict stock prices using historical data from financial APIs, and deployed it via Django for real-time predictions.
- Created interactive visualizations to compare model predictions with actual stock prices.
- Tech: TensorFlow, Pandas, NumPy, Matplotlib, Django Lang: Python

Natural Language Processing (NLP) for Chatbot Development

- Developed and deployed an NLP-based chatbot using Flask, including intent recognition and entity extraction.
- Integrated the chatbot with a web interface for a seamless user experience.
- Tech: NLTK, PyTorch, Flask Lang: Python

Credit Risk Model

- Designed a credit risk model using ML to predict loan default likelihood based on various historical financial data.
- Visualized the insights for better engagement.
- Tech: AWS SageMaker, S3, Scikit-Learn, Seaborn Lang: Python

Predictive Analytics for Retail Sales

- Collected and cleaned datasets from various car dealerships and built ML models to predict future sales.
- Evaluated model performance using RMSE, MAE, and R-squared, and created visualizations to present results.
- Tech: Scikit-Learn, Pandas, NumPy, Seaborn, Jupyter Notebook Lang: Python

EDUCATION

University of Massachusetts - Dartmouth

Masters, Data Science

- **GPA:** 3.33

Sep 2021 - Dec 2023

Dartmouth, MA

Gujarat Technological University

Bachelor, Computer Engineering

- **GPA:** 3.26

Aug 2016 - Aug 2020

Ahmedabad, India