Nidhi Patel

Boston, MA, 02215 | nidhipatel066@gmail.com | +1 857 891 9378 | linkedin.com/in/nidhipatel
github.com/nidhi0805 | Personal Website

EDUCATION

Northeastern University, Boston, MA — Sept 2023 – June 2025
Master's in Data Analytics — GPA: 3.8/4.0
Related Courses: Communication and Visualization for Data, Database Management, Big Data

Vellore Institute of Technology, Vellore, India — July 2018 – May 2022
Bachelor of Technology in Electronics and Communication — GPA: 8.24/10.0

TECHNICAL SKILLS

Programming: Python, SQL (MS SQL Server, Teradata, Snowflake, PostgreSQL, MySQL, MongoDB), R
Data Engineering/Visualization: ETL, Stored Procedures, Triggers, Indexing, Query Optimization, PowerBI, Tableau
Big Data & Cloud: Hadoop, Spark, Azure Databricks, GCP, AWS
Machine Learning: Scikit-Learn, TensorFlow, NLP, LLMs, A/B Testing, Sentiment Analysis, Feature Engineering
Certifications: Azure Fundamentals by Microsoft AZ-900, Google Data Analytics by Google, Microsoft Power BI Desktop for Business Intelligence

WORK EXPERIENCE

Hewlett Packard Enterprise, Remote, US — July 2024 – Dec 2024
Data Science Intern

- Built scalable data pipelines handling millions of records using Python, reducing processing time by 20%.
- Developed **predictive models** to identify factors driving quote-to-order conversion by analyzing 10M+ quote-to-order conversions, and improving the understanding by 15%.
- Engineered **feature selection** techniques, leveraging domain insights to boost win rate prediction by 12%.
- Designed dashboards in **Power BI** and **Qlik** reducing shipment delays by 25% across 50+ regions.
- Applied **A/B testing** to assess pricing elasticity, improving revenue forecasting accuracy.
- Automated model training and **hyperparameter tuning** in **Dataiku**, reducing model deployment time by 30% and enhancing predictive performance.

PepsiCo Global Business Services, Hyderabad, India — July 2022 – Aug 2023
Graduate Engineer Trainee | Data Engineer

- **Optimized ETL pipelines** processing 100M+ records daily, improving data accuracy by 15%.
- Led migration of 5TB+ data from SharePoint to **Azure Data Lake**, implementing bronze, silver, and gold layers via **Databricks** & **PySpark**, reducing reporting time by 20%.
- Improved query performance by 30% through **indexing**, **partitioning**, and **parallel processing**.
- Automated **data validation and anomaly detection** with SQL triggers & stored procedures, ensuring integrity.
- Designed and optimized SQL queries, views, and stored procedures, improving data transformation efficiency.

PROJECTS

Sentiment Analysis for Opioid Harm Reduction | *Python, NLP, ML, Flask* 🐙 — Jan 2025 - Mar 2025

- Applied **NLP techniques** and **sentiment analysis** to Reddit discussions on opioid use to identify public perceptions and emerging trends. Utilized **BERT-based embeddings** to extract key discussion topics. Created a Flask-based dashboard with real-time visualizations, mapping sentiment shifts and high-risk substance mentions.

EuroSAT Classification | *Python, TensorFlow, Keras, Satellite Image Classification* 🐙 — Sept 2024 - Dec 2024

- Developed a deep learning model to classify satellite images from the **EuroSAT dataset**, leveraging **convolutional neural networks (CNNs)**, optimized model performance using data augmentation and **hyperparameter tuning**.

Civic Tech Hackathon 2025 - Winner | *LLM, Ollama, Deepseek, React, MongoDB* 🐙 — Feb 2025

- ThinkFlow is an AI-driven learning platform that simplifies complex topics into personalized mind maps and quizzes. Featuring in-built accessibility tools like text-to-speech and adaptive fonts for inclusive education.