Vraj Pandya

Ontario | Phone: +1 604-389-2372 | E-mail: vraj.pndya@gmail.com | Linkedin | Github

Education

Master of Science in Computer Science *Wilfrid Laurier University* — Waterloo, ON
GPA: 3.9/4.0 — Sep 2023 – Dec 2024

Bachelor of Engineering in Information and Communication Technology *Gujarat Technological University* — Ahemdabad, GJ
GPA: 3.87/4.0 — Jul 2018 – Jul 2022

Work Experience

Wilfrid Laurier University — Waterloo, ON
Research Analyst — Dec 2023 – Jan 2025
• Conducted geospatial analysis on LandSAT datasets, applying OpenCV for satellite imagery enhancement.
• Developed **deep learning models (TensorFlow, PyTorch, ArcPy)** achieving **95-96%** accuracy in road extraction.
• Applied statistical validation, **reducing error rates by 4-6%** and improving model robustness.
• **Published research at IEEE-ICMI 2025** and contributed to OpenStreetMaps, improving mapping accuracy by 30%.

Inclusivity Institute For Better Data — Toronto, ON
Product Data Analyst — Jan 2024 – May 2024
• Led a 5-member team in an Agile environment to analyze **10M+ StatsCan data records**, predicting rental prices using **Python & MySQL**.
• Built interactive **Tableau dashboards (Streamlit),** increasing user engagement on website by 22-30%.
• Streamlined ETL pipelines from **NoSQL** for rental trend analysis, improving processing efficiency by 20%.

Tatvic Analytics — Ahmedabad, India
Data Analyst — Jul 2022 − Aug 2023
• Designed and implemented KPIs to improve **website performance by 17-21%** across multiple industries.
• Built **ETL pipelines with Airflow** and real-time **Looker/Power BI dashboards**, cutting monitoring time from a **week to 4 hours**.
• Conducted cohort and quantitative analysis (Google Analytics 4, Python, BigQuery), increasing client **user acquisition by 25-30%.**
• Developed **predictive models** to analyze historical sales patterns, improving revenue projections by 50,000$ in annual revenue for clients.
• Analyzed sales and revenue data in **Excel/Google sheets**, to optimize pricing strategies and improve demand forecasting.
• Performed A/B testing with the design team on product designs, improving user engagement and conversion rates by 15%.
• Trained clients on GA4 analytics, enabling independent execution of **15+ analyses per quarter**.

Indian Space And Research Center — Ahmedabad, India
Data Analyst — Aug 2021 – Jun 2022
• Utilized **AWS S3 for efficient storage** and retrieval of INSAT-3D satellite data, boosting processing speed by 20%.
• Developed **statistical models using MySQL** for spatial data analysis, uncovering 5+ new research insights on cloud behavior.
• Created geographic **visualizations in Python and reports in PowerBI**, aiding future atmospheric research studies.

Projects

Market Basket Analysis (Instacart's Data) – Data Analysis and Machine Learning with Python and PowerBI [Link]
• Conducted EDA on **30M+ transactions**, uncovering insights that can improve cross-selling by **15%.**
• Applied **Apriori & FP-Growth algorithms**, statistically increasing bundling revenue by **10-20%.**

Sales and Customer Analysis – Data Analysis with MySQL [Link]
• Developed complex SQL queries using **stored procedures**, **window functions** and **CTEs** to analyse YoY growth and sales trends.
• Created **8+ data-driven** reports analysing product performance, market penetration, and customer segmentation.
• Generated actionable insights from **1M+ sales data records** to support strategic business decisions across multiple markets.

Movie Recommendation System – Data Analysis and Machine Learning with PySpark, Python [Link]
• Built a distributed movie recommender processing 20M+ records, reducing computation time by 93% from two hours to 8 minutes.
• Improved model accuracy and loading times by 15% by using collaborative filtering techniques on PySpark from Scikit-Learn and Pandas.

Technical Skills

Languages & Databases: Python, SQL (MySQL, PostgreSQL), NoSQL, BigQuery, ElasticSearch
Data Visualization: Power BI, Tableau, Looker, Streamlit, Matplotlib, DAX
Machine Learning & Artificial Intelligence: TensorFlow, PyTorch, Keras, NLP, RAG, LangChain, Embedding Optimization
Big Data & ETL: AWS S3, AWS EC2, PySpark, Airflow, Databricks, Azure, Batch Processing
Quantitative Methods: A/B Testing, Hypothesis Testing, Simulation & Optimization
Soft Skills: Problem-Solving, Teamwork, Cross-functional Collaboration, Analytical Thinking