RUTVIK PATEL

London, Canada | [LinkedIn](#) | +1 519-992-3504 | rutvik05@gmail.com|[Portfolio](#)

TECHNICAL SKILLS

- **Programming and Scripting:** Python, C, C++, Java, HTML, JavaScript, CSS, MATLAB, Shell
- **Database Management, Visualization & Cloud:** SQL, MySQL, NoSQL, MongoDB, Hadoop, Power BI, Azure ML Studio
- **Development Tools:** Git, Docker, ETL (Power Query)
- **Python Libraries:** Pandas, NumPy, Scikit-Learn, PySpark, LangChain, Keras, PyTorch, TensorFlow
- **AI/ML Techniques & APIs:** Deep Learning, Classical ML Techniques, Generative AI, OpenAI & Gemini APIs

SOFT SKILLS

- **Creativity:** Developed creative solutions for data archiving, analysis, and access challenges to assist with London's homelessness strategy.
- **Leadership:** Led design of AI software architecture assignment at Vosyn and development of Dashboards at City of London.
- **Teamwork:** Collaborated on website for UN SDGs at Community of Guardians, and process flowcharts at City of London

EDUCATION

The University of Western Ontario | London, Ontario, CA   *Sep 2021-Feb 2023*
Master of Science in Computer Science (Specialization in Artificial Intelligence)   ***GPA: 3.90/4.00***

Pandit Deendayal Energy University | Gandhinagar, Gujarat, IN   *July 2017-June 2021*
Bachelor of Technology in Information and Communication Technology   ***GPA: 9.38/10.0***

WORK EXPERIENCE

City of London | London, ON, Canada   *Oct 2023-Present*
Training & Information Intern – Housing Stability Services

- Developed a methodology for the archival and collection of **encampment data** to facilitate **data analysis** and future **reporting**.
- Crafted an advanced **Power BI dashboard** integrating data from **SQL Server**, **Dynamics CRM**, **Weather Data API** and Network Folders, for executive use to assist with the city's **homelessness strategy**.
- Established a dashboard for daily client communication and housing services data reporting for strategic use, significantly improving reporting time and quality, saving more than **70%** of the time.
- Engineered automation scripts using **Python** that cut staff time by **90%** for routine processes.
- Spearheaded the creation of process **flowcharts** and **presentations** to assist employees.
- Produced public informative videos using **Synthesia** (AI video creator) by conducting cross-organizational research and surveys with community partners, enhancing efficiency and public access to information on government programs (some available on [this](#) page).

Vosyn Inc. | Toronto, ON, Canada   *Sept 2023-Mar 2024*
Machine Learning Engineer

- Implemented an audio segmentation module within a voice cloning and translation framework, enabling its application in multi-speaker environments, thereby enabling further development of the product. Utilized **AWS** for deploying and batch processing audio segmentation.
- Spearheaded the initial conceptualization and design planning of the voice synthesis AI software architecture.
- Implemented **Retrieval-Augmented Generation (RAG)** for customizing chatbots with user data and chat history, automating prompt generation through **OpenAI GPT-3.5 API**, removing the need for manual prompt generation.
- Assisted in the testing and review of a **Large Language Model (LLM)** based voice-synthesis and chat framework for text-based interactions, employing tools such as **LangChain**, Hugging Face, Sentence Transformers, and GPT4All, and suggested improvements.

Hiloni Stock Broking Pvt. Ltd. | Ahmedabad, GJ, India *Mar 2021-Aug 2021*
Data Science Intern

- Developed client-centric dashboards in **Power BI** for premium clients, integrating daily updates from **SQL Server** and **Market Data API** to boost engagement and investment tracking.
- Implemented **K-Means clustering** to efficiently segment 200 clients into six categories for targeted advisor and stockbroker assignments, enhancing service delivery and reducing operational costs.
- Innovated standardized data collection forms to web-based, achieving **40%** reduction in employee processing time.
- Designed and deployed an automated script using data from **MySQL** server for daily financial reporting and alerts to the Accounts department regarding account balances of various clients, achieving a **25%** improvement in operational efficiency.

PROJECTS

Comprehensive Retail Product Analytics using Power BI: Sales, Sizes, and Pricing Insights

- Conducted an extensive analysis of over a **SQL database** of 250,000 retail entries using **Power BI** over a six-month period, focusing on identifying top-performing SKUs and optimizing Manufacturer's Suggested Retail Prices (MSRPs).
- Developed and utilized advanced visuals, including market share **pie charts**, month-over-month (MoM) comparisons, and sales pattern **bubble charts**, to enhance inventory management and pricing strategies.
- Leveraged insights from a wide array of data, effectively pinpointing the top 10% of SKUs, determining competitively priced products by weight, and highlighting the potential of select product types for strategic decision-making.

Database Management for a Product Manufacturing Company

- Developed a **MySQL** database management system tailored for a consumer product manufacturer, focusing on schema design and seamless data integration to streamline operations.
- Achieved a significant improvement in operational efficiency, with the system reducing time spent on manual entries and paper-based order searches by more than **50%**, highlighting its effectiveness in enhancing productivity.
- Created a **Graphical User Interface** for the **SQL database** using **Java** Swing.

Twitter Data Analysis: FIFA 2022

- Scraped approximately 5 million tweets, totalling 14GB, that included 53 different hashtags related to FIFA 2022 using **SNScraper**, and stored the data efficiently in an **HDFS cluster**.
- Employed data cleaning, wrangling, and pre-processing techniques with the **NLTK** library, followed by sentiment analysis using **Textblob** to derive insights into the public perception of FIFA 2022.
- Applied **SparkMLlib** for the development of **Logistic Regression** and **Naive Bayes** models for **sentiment classification**, achieving **f1 scores of 0.9 and 0.73**, respectively, demonstrating high efficacy in sentiment analysis.

Microstrip Antenna Design

- Utilized **CST Studio Suite** to design and simulate a **Microstrip Antenna** based on provided parameters for lab demonstrations.
- Etched the designed antenna on copper substrate and used lab equipment to evaluate the antenna's performance.

Stellar Classification and Comparative Analysis of Classification Algorithms

- Performed a detailed analysis of six classification algorithms - **Decision Tree**, **Random Forest**, **XGBoost**, **AdaBoost**, **Naive Bayes**, and **Deep Convolutional Neural Network (CNN)** - to assess their performance in stellar classification.
- Utilized the Stellar Dataset from Kaggle to conduct thorough weight balancing and feature selection, employing both correlation analysis and univariate/multivariate techniques to refine the data preparation process.
- Identified **Deep CNN** as the most effective algorithm, achieving a notable **classification accuracy** of **96.5%**, underscoring its superiority in handling complex classification tasks.

CERTIFICATIONS AND ACHIEVEMENTS

- **Pursuing** Data Engineering Essentials using SQL, PySpark (Udemy), and Azure Data Scientist Certification
- **Ranked** in the 89th percentile among more than 55,000 participants in TCS Codevita Coding Competition
- **Completed** online courses on Machine Learning (deeplearning.ai), SQL (Stanford Online), Agile Software Development (University of Minnesota), and various Short courses on Generative AI (deeplearning.ai)