RUTVIK PATEL

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

OBJECTIVE

A data science professional with over 1.5 years of experience in data analysis and machine learning, I am skilled in programming, designing innovative AI solutions, and advanced data reporting with Power BI. Seeking to leverage my expertise in Python and machine learning techniques, I aim to drive innovative solutions for impactful insights and efficient decision-making. My commitment lies in harnessing analytics and machine learning to benefit the greater good.

TECHNICAL SKILLS

- **Programming and Scripting:** Python, C, C++, Java, HTML, JavaScript, CSS, MATLAB, Shell
- **Development Tools:** Git, Docker, ETL (Power Query)
- **Database Management, Visualization & Cloud:** SQL, NoSQL, MongoDB, Hadoop, Azure ML Studio, Power BI, Looker Studio
- **Python Libraries:** Pandas, SciPy, NumPy, Scikit-Learn, PySpark, LangChain, Keras, PyTorch, GPT4All, Sentence Transformers
- **Machine Learning Techniques:** Regression, Deep Learning, Decision Trees, Naive Bayes, Clustering, ML and DL Libraries

WORK EXPERIENCE

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

- Developed a methodology for the archival of Encampment Data to facilitate data analysis and future reporting.
- Crafted an advanced Power BI dashboard integrating data from SQL Server, Dynamics CRM, and Network Folders, aiding the city's homelessness strategy.
- Established a dashboard to analyze the correlation between weather patterns, shelter availability, and CSR communications, leveraging CRM and historical weather data REST API, which improved reporting and decision-making, saving up to 20% of time.
- Created dashboards instrumental in supporting executive decision-making and public statistics reporting.
- Engineered Designed automation scripts that cut staff time by 60% for routine processes and monthly reporting.
- In the process of developing an integrated dashboard for executive decision-making, incorporating data from diverse sources, including Census data.
- Spearheaded the creation of process flowcharts, digital presentations, and public informative videos using Synthesia, enhancing organizational efficiency and public access to government program information.
- Standardized and documented various internal processes for consistency and future reference.

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

- Contributed to the development of a personalized AI chatbot, specializing in prompt engineering for tailored user interactions.
- Implemented an audio segmentation module within a voice cloning and translation framework, enabling its application in multi-speaker environments.
- Engaged in the initial conceptualization and design planning of the voice synthesis AI software architecture.
- Researched various tools and resources necessary for the development phases of the AI-based product, including open-source tools, architectures, and Large Language Models (LLMs).
- Developed a Retrieval-Augmented Generation (RAG) based model to customize chatbots using available user data and chat history.
- Assisted in the testing of a Large Language Model (LLM) based product for text-based interactions, employing tools such as Langchain, Hugging Face, Sentence Transformers, and GPT4All.

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 by standardizing data collection forms and transitioning to web-based versions, achieving a 40% reduction in employee processing time.
- Designed and deployed an automated script for daily financial reporting and alerts to the Accounting department regarding outstanding and receivable account balances, achieving a 25% improvement in operational efficiency.

RUTVIK PATEL

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

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***

PROJECTS

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

- Conducted an extensive analysis of over 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.

Reinforcement Learning based 2nd player for Pong Game

- Created a Pong game from the ground up, integrating a Deep Convolutional Neural Network (CNN) with Q-learning to facilitate real-time learning capabilities for AI-versus-human gameplay.
- Optimized the training process by initially pairing the machine learning-based AI with a predefined hard-wired AI opponent, enhancing the learning curve.
- Refined model efficacy through iterative adjustments to custom reward and loss functions, culminating in a notable score of 7.88 out of 20 (moving average over 25 games) following 250 training episodes, indicating substantial progress in model 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.

Database Management for a Product Manufacturing Company

- Developed a 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.

Twitter Data Analysis: FIFA 2022

- Scraped approximately 5 million tweets, totaling 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 preprocessing 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.

Stock Price Movement Prediction Using Deep CNN and Other Deep Learning Models

- Implemented a **CNN**-based **classification** strategy with automated technical analysis using **candlestick charts** to predict stock price trends, achieving over **81% accuracy** in the top 10 Indian stocks on the NSE, complemented by scripts for automatic chart generation.
- Utilized **LSTM**, **RNN**, and **GRU** models for forecasting stock prices in time-series data, outperforming others with a MER under 14%.

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)