

YASH SANJAYKUMAR PATEL

contact@yashpatel.work | +1 (647) 446-6971 | **LinkedIn:** [yashsanjaykumarpatel](https://www.linkedin.com/in/yashsanjaykumarpatel)

Portfolio: www.yashpatel.work

Experience

Mann Des Codes Inc. | Data Engineer | Toronto, ON Sept 2023 –Present

- Developed and optimized over 5 data pipelines, improving data processing efficiency by 20%.
- Managed databases with more than 2 million records, reducing query times by 15% through performance tuning.
- Implemented data warehousing solutions that supported real-time analytics, reducing reporting delays by 30%.
- Automated data workflows, cutting processing times by 25% and increasing team productivity.

Enkiube Technologies Inc. | Data Scientist (Intern) | Toronto, ON May 2023 –Sept 2023

- Managed 5000+ image annotations via Label Studio, increasing data quality for model training by 60% through streamlined data labeling processes.
- Conducted thorough reviews and validation of annotated data to ensure accuracy and compliance with project requirements.
- Labeled data for training machine learning models, enhancing model accuracy by 40%.
- Built and trained machine learning models using advanced statistics, augmenting business decisions by 30%.
- Developed 20+ Python features, optimizing app efficiency by 50% through strategic programming based on project needs.
- Incorporated LLMs into projects, enhancing AI capabilities and text generation accuracy by 40%.
- Contributed to critical project discussions, improving workflow planning through feasibility analysis and innovative problem-solving.
- Documented and communicated test results, edge cases, and critical issues to the development team.

Lambton College | Data Engineer (Intern) | Mississauga, ON May 2023 - Aug 2023

- Developed end-to-end AWS Pipeline as part of a collaborative team effort, extracting, processing, and loading data from a sample company dataset.
- Applied advanced data cleaning and transformation techniques to ensure the integrity and quality of the data within the pipeline.
- Leveraged Power BI to design and implement a visually compelling dashboard, providing actionable insights to stakeholders.
- Successfully presented project findings and dashboard functionality to project stakeholders, highlighting the impact of data engineering solutions on decision-making processes.
- Collaborated with the development team to understand data requirements and ensure they were met.

Amkay Software Solutions | Machine Learning Engineer (Intern) | Ahmedabad, India Sept 2020 - Mar 2021

- Undertook a leadership role in CRM development, strategically optimizing project management efficiency through a smart task delegation strategy.
- Built the CRM's robust back end, precisely aligning database design with intricate requirements of the project.
- Designed a four-tier user accessibility system embedded in the CRM, markedly improving task distribution and tracking efficiency by about 30%.
- Led a key project using machine learning to automate invoice data extraction, enhancing efficiency by 45% and reducing errors by 30%.
- Fostered a collaborative culture in a cross-functional team, promoting shared problem-solving and leading to early project completion.
- Monitored efforts to resolve issues and re-test bug fixes to ensure final product satisfaction.

Skills

- **Programming Languages:** Python, R, SQL, DAX, Scala

- **Tools & Software:** MySQL, PostgreSQL, NoSQL, Gemini, Chat-GPT, Tableau, MS Power BI, Google BigQuery, Hadoop Big Data, Spark, AWS, Azure, PowerPoint, Word, Advanced MS Excel (VLOOKUP, Pivot, Power Query), Google Analytics, Snowflake Data Warehousing, Jupyter Notebook, Google Cloud Platform, Bubble
- **Proficient Skills:** Data Annotation, Data Review, Data Visualization, RDBMS, Machine Learning, NLP, Project Management, ETL, Data Pipeline, Data Cleaning, Data Modeling, Data Analysis, BI Tools, Data Warehousing, Prompt Engineering
- **Soft Skills:** Analytical, Problem Solving, Time Management, Decision Making, Organizational Skills
- **Certificate:** Google Data Analytics Professional (Certified)

Personal Projects

Reddit Sentiment Tracker (ETL): Data Engineering Project (AWS, Power BI)

- Orchestrated real-time Reddit Sentiment Tracker for Canadian Politics using continuous data flow, AWS Lambda, ETL, and NLTK.
- Leveraged advanced technologies such as AWS S3, Athena, and Power BI for efficient data storage, analysis, and visualization.
- **GitHub Link:** <https://github.com/yashspatel/canadian-politics-sentiment-analysis>

Global COVID-19 Impact Visualizer: PowerBI Project

- Utilized global data and Power BI to present key insights, including total confirmed cases, deaths, and recoveries.
- Made complex epidemiological data accessible and understandable for diverse users.
- **GitHub Link:** <https://github.com/yashspatel/Global-Covid-19-Impact-Visualizer>

Netflix Trend Dashboard: Tableau Project

- Analyzed 6,000+ rows of data using Tableau, visualizing movie and TV show trends by country, release year, genre, and rating.
- Enabled tailored viewing experiences by allowing users to explore specific movie and show characteristics such as release date, duration, and cast.
- **GitHub Link:** <https://github.com/yashspatel/Netflix-Trends-Dashboard-in-Tableau>

Education

Lambton College, Mississauga

January 2022 - August 2023

- *Post Graduation Diploma of Big Data Analytics* | GPA: 3.65 | Mississauga, ON

Gujarat Technological University

June 2017 - May 2021

- *Bachelor of Engineering in Information Technology* | GPA: 3.9 | Gujarat, India