

# DEVVRAT RAJESH MUNGEKAR

Calgary, AB | (249)-876-6843 | [devvratmungekar53@gmail.com](mailto:devvratmungekar53@gmail.com) | [LinkedIn](#) | [GitHub](#)

## SKILLS

Programming: **Python, PySpark, R, Shell Scripting/Bash**  
Cloud & DevOps: **Microsoft Fabric, Azure Data Factory, Azure Synapse Analytics, Azure Databricks, Azure Machine Learning, Azure AI Foundry, ADLS Gen 2, Azure OpenAI, Docker, Kubernetes, Git**  
Databases & Tools: **Azure SQL, Azure PostgreSQL, MySQL, Tableau**

## WORK EXPERIENCE

**Associate Data Scientist, LTIMindtree Canada @Microsoft, Calgary, AB** January 2024 – Present

- Designed and optimized **end-to-end data pipelines** using Azure Data Factory, Azure Databricks, Synapse Analytics, and ADLS Gen 2, improving throughput by 60%
- Built and managed scalable, production-grade **ETL pipelines** using Python and PySpark across structured and unstructured data
- Enabled **ML-based prediction workflows** by integrating Azure ML pipelines with enterprise data sources and deploying models to real-time and batch endpoints
- Deployed LLMs including Azure OpenAI, Hugging Face, and Llama to support **audience segmentation and prediction** use cases, increasing model accuracy by 20%
- Applied **CI/CD practices** using Azure DevOps and GitHub for automated deployment of data pipelines and models
- Worked on developing and improvising the flow rate by 20% using the Azure Prompt Flow and the Kubernetes cluster embedded in Azure Machine Learning
- Oversaw compute instances, clusters, and Azure Kubernetes deployments, supporting model registration, Prompt Flow jobs, and custom Docker scripts for environment configuration
- Worked cross-functionally with data scientists and business stakeholders to gather requirements and translate them into **robust data engineering solutions**
- Implemented monitoring and performance logging to enable **data quality validation** and root-cause troubleshooting

**Teaching Assistant, Trent University, Peterborough, ON** January 2023 – January 2024

- Assisted the course director in teaching Introduction to Databases, Introduction to Data Science, and Data Visualization courses, fostering engagement among 160 students
- Mentored students in SQL (MySQL and PostgreSQL), NoSQL (MongoDB), and machine learning pipelines, achieving a 99% student success rate
- Supported faculty in research projects, utilizing R for data analysis, and presented findings to stakeholders

**Software Engineer, Capgemini Technology Services India Ltd, Mumbai** August 2021 - January 2022

- Migrated 85% of custom SQL databases from Java architecture for optimized data storage in product lifecycle development.
- Improved query execution speed, enabling faster data validation and enhancing storage optimization.

## PROJECTS

### Backorder Inventory Management

[View in GitHub](#)

- Created an event-driven ML pipeline on GCP with BigQuery storage and used Streamlit for interactive geospatial clustering; achieved 97.3% accuracy
- Built a front-end web application on AWS Elastic Beanstalk, predicting inventory backorders with a 99.32% accuracy rate
- Developed a CI/CD pipeline using AWS Code Pipeline for seamless code integration and deployment

### Crime Analysis Application

[View in GitHub](#)

- Processed 7M+ records to create a machine learning pipeline on Google Cloud Platform, storing outcomes in BigQuery

- Designed an interactive Python web application using Streamlit for SQL-based data transformation and spatial clustering with 97.3% accuracy
- Published research paper “Crime Analysis using DBSCAN Algorithm” in IEEE doi: [10.1109/ICIRCA51532.2021.9544719](https://doi.org/10.1109/ICIRCA51532.2021.9544719)

## EDUCATION

**M.Sc. in Applied Modelling and Quantitative Analysis: Big Data Analytics**

May 2022 – September 2023

Trent University, Peterborough, ON

**B.E. in Computer Engineering**

August 2017 – July 2021

University of Mumbai, Mumbai, India

## CERTIFICATION

[AWS Certified Cloud Practitioner](#), Validation Number: C2BPCB227BEQQEC3

**AWS**