

AAKASH PATEL

Calgary, Alberta, Canada | aakash98.pt@gmail.com | linkedin.com/in/aakashpatel05

SUMMARY

Data Engineer with 4+ years of experience building AI-powered automation systems and scalable data pipelines across Azure and AWS. Specializes in agentic AI solutions, LLM integration, and end-to-end workflow automation. Proven track record of reducing manual effort and operational costs through intelligent system design. Master's in Big Data Analytics. Builder of systems that work while you sleep.

TECHNICAL SKILLS

AI & Generative AI: Agentic AI Design, LLM Integration (OpenAI, Claude), Azure OpenAI, Azure AI Foundry, Copilot Studio, OpenAI SDK, Bing Grounding Search, RAG Architectures, Prompt Engineering, Model Evaluation

Cloud Platforms: Microsoft Azure (ADF, Databricks, Synapse, Functions, Event Hubs, Azure DevOps), AWS (Glue, EMR, Redshift, S3, Lambda)

Data Engineering: Apache Spark, PySpark, Kafka, Snowflake, ETL/ELT, Data Warehousing, Star/Snowflake Schema Design

Machine Learning & NLP: Scikit-learn, NLP Pipelines, Classification & Regression Models, TF-IDF, Feature Engineering, Model Evaluation

Programming: Python, SQL, Pandas, REST API Development

Infrastructure & DevOps: ARM Templates, CI/CD, Git

Analytics & Visualization: Power BI, Qlik

Databases: PostgreSQL, MySQL, Neo4j

EXPERIENCE

Data Engineer — Avenue Living

Calgary, Alberta, Canada | July 2025 – Present

- Built an agentic AI pipeline using Bing Grounding Search to autonomously find company websites, extract advisor data, and push into HubSpot CRM — zero human intervention
- Integrated Azure AI Foundry with Claude LLM to auto-generate intelligent summaries of ADF and Python project changes, auto-updating Confluence on every release — eliminated all manual documentation
- Developed web scraping agent using Sequentum for Canadian rental market data — reduced daily operational costs by 73% (from \$150+/day to under \$40/day)
- Redesigned event-triggered pipelines to batch-process file uploads, significantly reducing redundant Azure trigger costs
- Built centralized pipeline orchestration system for scheduling and managing multiple pipelines; manage CI/CD and DevOps responsibilities

Data Engineer — Citi

May 2022 – June 2025

- Designed and deployed end-to-end ETL pipelines using Azure Data Factory, handling ingestion, transformation, and loading of structured and semi-structured data from multiple internal and external sources
- Developed PySpark-based data processing frameworks on Azure Databricks and Azure Batch, optimizing performance and scalability for large-volume datasets
- Migrated data validation and monitoring to Azure Functions for serverless execution, enabling automated data quality checks and pipeline health tracking
- Built real-time and batch pipelines integrating Azure Event Hubs and Azure Stream Analytics for near-real-time insights and downstream reporting
- Implemented data warehouse solutions using Azure Synapse Analytics and Azure SQL MI, including schema design, partitioning, and query performance optimization

- Built Star and Snowflake schema models in Azure Synapse to support regulatory reporting, KPI dashboards, and BI workloads; created Power BI dashboards connected to Synapse datasets
- Automated Azure resource provisioning and infrastructure deployment via Azure DevOps CI/CD pipelines and ARM templates

AI & AUTOMATION PROJECTS

Kolena AI — Automated PDF Data Extraction Pipeline

- Built fully automated document intelligence pipeline — handles printed and handwritten PDFs dropped into SharePoint, auto-triggers, extracts structured data via AI prompting, runs ETL into data warehouse, serves live dashboards

Automated LinkedIn Content Generator

- End-to-end content automation: searches trending AI topics, generates posts via Claude API, refines with OpenAI, auto-publishes to LinkedIn via API — zero manual effort

Fake News Classifier (Academic)

- NLP-based ML classifier using TF-IDF, Logistic Regression, Random Forest, and Naive Bayes to detect fabricated news articles with high accuracy

EDUCATION

Master's in Big Data Analytics — Trent University

ML Projects: Fake news detection, credit card fraud detection, flight fare prediction, text classification

Bachelor's in Computer Engineering — A.D. Patel Institute of Technology

Publications: Indian Paper Currency Detection using OpenCV3 | Data Mining Algorithms

CERTIFICATIONS

Machine Learning and Deep Learning Masters • Apache PySpark by Example • Neo4j Fundamentals • Analyzing Big Data with Hive