Rahul Kashyap

Data Scientist

rk.rahulkashyap2@gmail.com • (902) 440 0921

LinkedIn • Halifax, NS, CA B3H 3Z5

Innovative **Data Scientist** with **6+ years** of experience in real-time audio transcription, LLM/RAG, and production-grade ML systems. Designed and deployed low-latency transcription and alerting pipelines, ICD code normalization services, and observable microservice architectures on Azure + OpenAI. Skilled in Python, PySpark, SQL, Power BI, and Excel, with a strong record of improving accuracy, reliability, and cost-efficiency through data-driven collaboration.

Technical Certifications

The Complete Agentic AI Engineering Course. Udemy	**2025**
IBM: Building AI Solutions Using Advanced Algorithms and Open-Source Frameworks	**2024**
IBM: Generative AI: Introduction and Applications, Coursera	**2024**
IBM: Data Science Specialization, Coursera	**2023**
Microsoft Power BI Desktop for Business Analytics, Udemy	**2021**
Certified in Data Science using Python, Analytix Labs	**2018**

Technical Proficiencies

Languages | Python, R, Java, C/C++, SQL, PySpark

Tools | VS Code, Cursor, Jupyter Notebook, Databricks, Tableau, OpenAI, Azure Storage/Key Vault/Monitor, FastAPI, Flask, Advanced Excel, Power BI, MS-Office

Process | Agile, Scrum, JIRA, Git

Professional Experience

ResolveHD, Nova Scotia, CA   **Feb 2025 – Oct 2025**
Data Scientist

- Built a real-time audio transcription platform for mic input, audio files, and live broadcast stream URLs using OpenAI API models, achieving under 8 seconds transcription latency and ~90% transcription accuracy through regex and prompt-engineering post-processing.
- Automated operations checklists (SOP) and emergency alerts triggered from recognized transcription text in real time, generating structured actions (timestamps, assignee, SOP step) and pushing them to backend services to reduce manual documentation and audit effort.
- Developed modular microservices (transcribe → classify → alert → export/notify) using an observable architecture pattern to ensure reliability, scalability, and easy monitoring in production.
- Shipped ICD condition code resolver microservices using a hybrid matcher (regex direct match, fuzzy distance, embeddings, RAG, BioBERT) to normalize free-text notes to ICD codes with approximately 75% accuracy.
- Delivered multiple end-to-end audio transcription demo products with integrated frontend and backend for go-to-market showcases.

Sobey School of Business, SMU, Nova Scotia, CA   **Mar 2024 – Jan 2025**
Data Scientist

- Developed machine learning models achieving 70% accuracy rate in distinguishing hateful from non-hateful image and text posts by implementing a sophisticated multi-modal machine learning & Image processing pipeline.
- Developed and tested 20 large language models (LLMs) to enhance content classification and hate speech detection capabilities.
- Improved model performance by addressing class imbalance with weighted random samplers and advanced data augmentation.

- Created an intensity scale for hate speech using the CLIP model, aiding in the assessment and moderation of harmful content.
- Devised and implemented sentiment analysis tools and emotion recognition models to classify content by hate intensity.
- Led efforts to proactively moderate content through data interpretation and sentiment metrics (NLP).

The Citco Group Limited, Nova Scotia, CA   **Sep 2023 – Mar 2024**
Data Analyst (Internship)
- Reduced database size by 10%, enhancing system efficiency and performance by optimizing data storage and removing redundancy.
- Predicted fund closure dates using machine learning models deployed on Databricks cloud platforms, improving forecast accuracy.
- Developed a comprehensive Power BI dashboard delivering critical business insights on data usability and infrastructure costs.
- Cataloged 377 tables across 33 schemas to streamline data management and analysis processes.
- Integrated SQL with Python/Pyspark to enhance data retrieval efficiency, streamlining the analysis process.

Sobey School of Business, SMU, Nova Scotia, CA   **Jan 2023 – May 2023**
Data Analyst (Internship)
- Extracted and wrangled Nova Scotia Open Data, Analyzed air quality data and its impact on asthma prevalence, identifying key correlations using statistical methods.
- Leveraged difference-in-differences analysis to validate findings, informing health strategies and policies.
- Utilized machine learning and descriptive analytics to provide actionable insights into health-related data.
- Applied Descriptive, and Prescriptive analytics to interpret complex datasets and provide insights.

Flipkart (Walmart Group), Ecommerce Giant, India   **Sep 2018 – July 2022**
Operations Business Analyst (Manager) – Last Mile
- Managed end-to-end project management of Same-Day Delivery and Delivery & Installation initiatives within 48 hours, significantly enhancing customer satisfaction and operational efficiency.
- Led comprehensive data science analyses on supply chain datasets, ensuring data integrity and optimizing processes to guide executive decision-making.
- Drove an 8% increase in quarterly revenue through strategic project management of delivery services.
- Improved reverse logistics efficiency by 45%, streamlining operations and reducing costs.
- Achieved a 25% cost reduction with vendors through effective implementation of a Vendor Scorecard system.
- Demonstrated proficiency in business intelligence tools by crafting sophisticated reports and dashboards.
- Developed and implemented a Vendor Scorecard system based on KPIs, resulting in cost savings and improvements.

Education

Master of Business Analytics | Sobey School of Business, Saint Mary's University (SMU), Halifax, NS, Canada	**2024**
PGP in Data Science using Python | Analytix Labs, Delhi, India	**2018**
Bachelor of Technology in Computer Science | Amity University, UP, India	**2017**

Awards & Activities

CEO's Award for Customer Excellence, Flipkart	**2022**
Business Optimization Award, [Received R&R Award], Flipkart	**2021**
Supply Chain Optimization, [Received R&R Award], Flipkart	**2021**
Business Innovation Award [Received R&R Award], Flipkart	**2020**