

Rahul Sharma

☎ 1-647-572-2414 • ✉ sharmarah494@gmail.com • in [sharmarah](https://www.linkedin.com/in/sharmarah) • 🌐 [rahul494](https://rahul494.com)

PROFESSIONAL EXPERIENCE

Software Developer

Oct. 2020 - Present

Canadian Institute for Health Information (CIHI) | Toronto, ON

Business Intelligence and Full Stack Development

- Developed SQL-based dashboards and reports under strict government data requirements (PHIPA), achieving a 4.8s P95 latency.
- Engineered custom Qlik frontend extensions using HTML, CSS, and JavaScript, expanding organizational data analysis capabilities.
- Currently building a full stack version of a Physician Resource Planning forecasting application, utilizing FastAPI Python backend and React.js frontend to predict the number of physicians Canada will need in the upcoming years.
- Collaborated with subject matter experts and R&D team to develop a PyTorch based recurrent neural network to detect sensitive personal health information, reducing manual review efforts with a 98% accuracy in classification.

ETL and Database Management

- Scheduled database backups, facilitated performance analytics, and provided user administration support for a Physician Resource Planning forecasting application, reducing administrative overhead by 80%.
- Developed and maintained in-memory BI and Oracle data models by debugging source tables and SQL view definitions.
- Orchestrated ETL data workflows for Canadian public health reports, aiding experts in data-driven decision-making and planning.

Technical Leadership

- Lead expert for several applications, leveraged knowledge to create markdown/confluence pages minimizing defect downtime and enhancing operational efficiency by 90%.
- Delegated requirements to consultants to migrate business intelligence applications between MicroStrategy to Qlik and underlying database management systems.
- Provided mentorship to new junior developers, fostering their professional development and integration into the team.

Business Intelligence Developer

Sep. 2018 - Aug. 2019

Regional Municipality of York | Newmarket, ON

Business Intelligence and Full Stack Development

- Engineered real-time, dynamic dashboard charts and reports utilizing HTML, CSS, JavaScript, and Power BI.
- Successfully delivered a full-stack accounting/capital planning software within a tight fiscal year-end deadline, collaborating with teammates to empower business analysts with robust asset management and expenditure tracking tools.

ETL and Database Management

- Devised ETL scripts to efficiently handle data from multiple databases, ranging from 10MB to 10GB in size.
- Designed and optimized relational database models, creating complex SQL queries, triggers, constraints, and stored procedures for enhanced performance.
- Developed Python scripts leveraged for web data collection, cleaning, and analytical capabilities.

EDUCATION

Georgia Institute of Technology | Online

Expected 2026

Master of Science in Analytics

York University – Lassonde School of Engineering | Toronto, ON

Sep. 2015 - May 2020

Bachelor of Science in Computer Science with Honors

Relevant Courses: Data Structures, Algorithms, Object Oriented Programming, Big Data Systems, Data Analytics, Regression Analysis

ACHIEVEMENTS & CERTIFICATIONS

AWS Certified Cloud Practitioner Certification

Expected 2024

Bayesian Statistics: Techniques and Models with Honors | Coursera

Awarded in May 2020

University of California Santa Cruz

Big Data Specialization | Coursera

Awarded in Apr. 2020

University of California San Diego

RESEARCH PROJECTS

Mental Health Analysis using Twitter [[report](#) / [code](#)]

- Utilized big data technologies such as Hadoop and Apache Spark to conduct descriptive analytics on user tweets.
- Implemented various MapReduce algorithms to pull data from Hadoop.
- Performed various natural language processing techniques to provide descriptive analytics.
- Conducted analytics using methods such as sentiment analysis, n-grams, skip-grams, and TF-IDF.

Image Recognition for Environmental Settings [[report](#) / [code](#)]

- Implemented various machine learning models to determine whether an image was taken outdoors.
- Applied L1 and L2 regularization methods for feature selection, resulting in a performance increase of 14%.
- Achieved a prediction accuracy of 90% utilizing machine learning algorithms such as penalized logistic regression, SVM, and a convolutional neural network.

PERSONAL PROJECTS

WAP Data Engineering Design Pattern and AWS Ecosystem for Music Subscription Service [[architecture](#)]

- Architected and implemented cloud ETL workflow to write, audit, and publish data from end user to production database.
- Designed robust AWS ecosystem leveraging services such as S3, Lambda, RDS, Step Functions, and Simple Email Service.
- Integrated multi-stage data processing by implementing Lambda driven auditing and Python code to ensure high data integrity and reliability.
- Enhanced user experience by automating notifications upon successful completion of data orchestration.

CI/CD Pipeline Solution for Python Development

- Designed and implemented complete CI/CD pipeline using AWS CodePipeline, integrating AWS CodeBuild for seamless automation.
- Configured web services such as S3, EC2 instances, and IAM roles to deploy code changes automatically to production for Python applications.

Apache Airflow Financial Data Aggregation for Banking Institutions

- Designed and orchestrated an Apache Airflow based pipeline to automatically aggregate financial metrics from multiple bank branches daily, ensuring timely and accurate consolidation of key transaction data.
- Implemented advanced data transformation logic within the Airflow DAGs to compute crucial financial metrics, followed by the automated generation of comprehensive daily financial reports for executive review.
- Automated the distribution of daily financial reports to key stakeholders and bank management via email, facilitating prompt financial insights and strategic decision-making based on fresh data.

TECHNICAL SKILLS

Programming Languages: Python, Java, R, SQL, HTML, CSS, JavaScript

Databases & Data Warehousing: Oracle, SQL Server, MySQL, PostgreSQL

Big Data & ETL Tools: Hadoop, Apache Spark, Apache Airflow

Cloud Computing & AWS Services: S3, EC2, RDS, Lambda, Simple Email Service, IAM, CodePipeline, CodeBuild

Business Intelligence & Data Visualization Tools: Qlik Sense, MicroStrategy, Power BI, Oracle APEX, Tableau