

Arunim Garg

Vancouver, Canada ✉ arunim27@gmail.com ☎ 604-655-7811 📧 in/arunimgarg 🌐 http://arunimgarg.com/

SUMMARY

A dynamic and responsible individual with a strong passion for leveraging data and machine learning to make a positive impact on people's lives. Proficient in R and experienced in Scala and python, I possess two years of hands-on industry experience in data analysis and a deep curiosity to develop ML models.

PROFESSIONAL EXPERIENCE

Data Research Analyst

Lakehead University & TBRHSC - Ontario, Canada - Part-time

Jan 2020 - Oct 2021 & Jan 2023 - Present

- Investigated family physicians' (FPs) billing codes to improve resource planning in Ontario, addressing the disparity between rural and urban populations.
- Utilized R to analyze, predict, and visualize the ICES physician and OHIP physician billing databases, handling 30K and 99M observations, respectively.
- Developed methodologies to analyze over 200M records of FPs annually, resulting in more accurate estimates of physician demand and supply in Ontario.
- Conducted hypothesis testing to evaluate our findings about physicians in rural areas, utilizing statistical analysis in R to derive actionable insights.

Data Engineer

Insurance Corporation of British Columbia (ICBC) - BC, Canada - Full-time Contract

May 2022 - November 2022

- Built 5 scalable data pipelines, processing billions of records from diverse sources to generate comprehensive summary tables for advanced analytics.
- Developed efficient ETL data pipelines using Scala, Spark, SQL, and the Hadoop ecosystem, optimizing processing time by up to 80%.
- Utilized tools such as HDFS, Apache Hive, DBeaver, and RDBMS (Oracle and MySQL) to analyze data, test queries, and support ad hoc requests.
- Collaborated with data engineers and QA analysts practicing agile development, actively participating in the new-hire buddy system and social events.

Data Analyst

ZLC Financial - BC, Canada - Full-time Contract

January 2022 - April 2022

- Led CRM migration, optimizing client relations for 20+ financial advisors through data extraction, auditing, and consolidation of financial data.
- Accelerated migration timeline by a month by implementing a new pipeline for auditing large datasets and automating data integrity review in MS Excel.
- Effectively communicated technical results to non-technical stakeholders, facilitating informed decision-making and driving business growth through data analysis and actionable recommendations.
- Reported directly to the CIO, providing daily touch-point meetings, developing weekly progress reports, and documenting procedures.

ML Developer

Communica - Alberta, Canada - Part-time Contract

November 2020 - July 2021

- Automated text-analyzing processes for a public affairs company by preparing an end-to-end NLP solution for text extraction, summarization, and text proofreading using a scalable ETL pipeline with Microsoft Azure services.
- Trained and tested a Microsoft Azure LUIS model on 1600 email documents to extract text fields with 98% accuracy.
- Developed scalable Python pipeline that reduced email review time by 200x. Used SQL queries in MS SQL Server to support ad hoc requests and extract data.
- Gathered client requirements to prepare and deliver quality product designs on time and within budget.

EDUCATION

Master of Computer Science, Artificial Intelligence Specialization

Lakehead University (LU) • Thunder Bay, ON • May 2021 • 90%

Bachelor of Science in Information Technology

Minor in Business Administration • Fairleigh Dickinson University (FDU) • Vancouver, BC • December 2018 • 3.65 (4.0 Scale)

RESEARCH PAPERS

To investigate the role of machine learning in medical research

Published - Computer Science Review Volume 40 • May 2021

Predicting family physicians based on their practice using machine learning

Published - 2021 IEEE International Conference on Big Data (Big Data) • December 2021

EXTRACURRICULAR

'Flatten the curve, Then Recover the Economy' Virtual Hackathon • Alberta Innovates • May 2020

- Developed an AI chatbot that can answer any COVID-19 related question to help increase awareness and help flatten the curve.
- Used AWS services including Sagemaker, Kendra, Lambda, DynamoDB, and Cognito and extended the Chatbot to be used with Amazon Alexa.

SKILLS & NOTABLE COURSES

Python (Tensorflow, Keras, PyTorch), R, SQL, Scala, Git, PowerBI, HDFS, Apache Spark, Apache Hive, Java, DBeaver, Hue, ELT/ETL data pipelines. Big Data, Machine Learning, Deep Learning, NLP, ML Specialization and Prompt Engineering for LLMs by DeepLearning.ai