

# Shahbaz Yousaf . Jarvis Consulting

I recently graduated with a Bachelor of Engineering, majoring in Computer Engineering from Toronto Metropolitan University, formerly Ryerson. Throughout my time at Ryerson, I honed my coding, problem-solving, and analytical skills through various courses and projects. I developed a keen interest in data engineering and devops as I delved into topics like algorithms, data structures, operating systems, data engineering and machine learning. This led me to my data science internship at Ansys, where I was working on big data processing. In my final year, I worked on a full-year capstone project in which I built a full-stack ML job recommender application. I have completed multiple certifications such as IBM Applied DevOps Engineering and IBM Data Science, reflecting my passion for continuous learning and skill enhancement. I consider myself disciplined, highly motivated, and adaptable, thriving both independently and within a team. I enjoy challenges that push my technical and leadership abilities, viewing them as opportunities for growth. Beyond tech, my interests include basketball, strength training, MMA, and cooking. Regularly engaging in these activities keeps me balanced and focused.

## Skills

**Proficient:** Python, Java, Linux/Bash, RDBMS/SQL, Docker, Agile/Scrum, Git

**Competent:** C, PySpark, Azure Databricks, HTML/CSS, ReactJS, Django, MATLAB, VHDL

**Familiar:** AWS, GCP, Jenkins, Terraform, Ansible

## Jarvis Projects

Project source code: [https://github.com/jarviscanada/jarvis\\_data\\_eng\\_ShahbazYousaf](https://github.com/jarviscanada/jarvis_data_eng_ShahbazYousaf)

**Linux Cluster Resource Monitoring App** [GitHub]: Built a monitoring system that records hardware specs and monitors real-time server usage and stores the data in a PostgreSQL database. The purpose of the collected data is to aid users in future resource planning, ie adding or removing servers, optimizing underused servers with more software and apps, or acquiring new servers to reduce the load on overused ones. Developed bash scripts to collect hardware and server data collection, used Docker for PostgreSQL provisioning and deployment, and leveraged Git and GitHub for version control. Configured crontab to run host\_usage.sh script every minute, automating data collection.

**Java Grep Application** [GitHub]: Developed a Java grep app mimicking Linux bash grep command, allowing users to search for matching strings in files. The application recursively searches a root directory based on a provided regex pattern, identifies matching lines, and saves them to a user-specified output file. Utilizing Docker, the app can be easily implemented by pulling the image from DockerHub. Technologies used in the project include Java, Maven, IntelliJ, Lambda & Stream API, and Docker.

**Python Data Analytics** [GitHub]: Collaborated with London Gift Shop, a UK-based retailer, to analyze customer behavior and diagnose revenue stagnation. Utilized Jupyter notebooks, Python, Numpy, matplotlib and pandas for analysis, generating insights on invoice distributions, monthly orders, and transaction patterns. These findings guided the marketing team in developing new marketing strategies to boost LGS's business.

## Highlighted Projects

**Machine Learning Driven Job Recommender System** [GitHub]: Developed a job recommender system achieving a 90% match rate. Used Python for NLP-based data preprocessing, reducing ML training time by 70%. Implemented a MySQL database for storing resumes and job descriptions. Created a user-friendly web app with Django for backend and React for frontend.

**NBA Data Model and Database:** Designed and implemented a 2021 NBA Season database using SQLite. Utilized ER Diagram for structure, inserted data, and implemented SQL queries. Developed a Java application using JDBC for database interaction. Created an XML version with an XML Schema and performed queries using XSLT.

**Gradient Descent for Univariate Linear Regression:** Implemented Gradient Descent algorithm for predicting final marks based on midterm marks using student data. Compared standardized and non-standardized values, conducted exploratory data analysis, and adjusted learning rates and iterations for improved accuracy. Validation with scikit-learn library showed similar outcomes.

## Professional Experiences

**Software Developer, Jarvis Consulting Group (2023-present):** Developed data engineering and DevOps projects at Jarvis for internal users using Java, Python, Docker, Git, PSQL, and Cloud (GCP). Worked in an agile environment with daily scrum meetings and bi-weekly sprint reviews/retrospectives to enhance development processes.

**Data Science Intern, Ansys (2022):** Developed Python scripts to support the Electronics Business Unit team in improving APIP analytics engine that processed software usage data, utilizing technologies like PySpark and Azure Databricks. Engineered nested sets tree algorithm, slashing hierarchal user data query time by around 80%.

## Education

**Toronto Metropolitan University (Formerly Ryerson University) (2016-2022),** Bachelor of Engineering, Electrical and Computer Engineering

## Miscellaneous

- IBM Applied DevOps Engineering Professional Certificate (2023)
- IBM Data Science Professional Certificate (2022)
- Basketball: Play organized basketball as part of a local league
- Strength Training: Current PRS are 275 lbs bench press, 275 lbs squat, 405 lbs deadlift
- MMA: Train muay thai and wrestling
- Cooking: Love making smash burgers and steaks