Rishi Shah

416-995-6279 | rishi.shah@uwaterloo.ca | rishishah.ca | linkedin.com/in/rishishah2001 | github.com/rishi-shah12

Technical Skills

Languages: Python, C++, Java, R, SQL (PostgreSQL, sqlite3), NoSQL (Elasticsearch), HTML/CSS, MATLAB
Frameworks: Flask, React, Django, REST API, Jinja2, YAML, JSON, Material-UI
Developer Tools: AWS, Azure, Docker, Git, LinuxCL, Jenkins, JIRA, Excel
Libraries: Pandas, NumPy, Matplotlib, SciKitLearn, Tensorflow, Ray, Plotly, NLTK, Pickle

Experience

Software Engineer June. 2021 – Present
Biotech Square *Toronto, CA*

- Authored **Extract-Transfer-Load microservice pipelines** for drug approvals using **AWS Lambda** and a **Elasticsearch** database
- Designed **distributed computing** programs using **Ray** to run on **AWS EC2** instances, **improving performance speed** by **40%**
- Utilized **natural language processing algorithms** for **entity** recognition, **sentiment** analysis and **summarization**, **increasing** search-able fields by **1.25x**

Software Developer May. 2021 – Aug. 2021
Environment & Climate Change Canada *Ottawa, CA*

- Developed a **full-stack web app** for **OCR** which **extracts & transforms** data using **computer vision**
- Leveraged **Microsoft Azure** to create **supervised neural networks** for various layouts resulting in **2x quicker** extraction
- Created a **dashboard** for crucial **statistics & reports reducing prep-time** by **90%**

Bioinformatics Developer Aug. 2020 – Dec. 2020
Ontario Institute for Cancer Research *Toronto, CA*

- Created a **COVID-19 analysis ETL pipeline** to analyze sequenced data while creating a **JSON metrics file, graphs, charts** and a **pdf report** in **Python** and **R**. **Reduced time-to-completion** by **5 hours**
- Integrated a **REST API** for file/data management in the pipeline, increasing **efficiency** of data transfer by **300%**
- Designed **Python, R** software to **re-implement legacy** code, leading to **operation times cut** by **25%**
- Managed **Jenkins tests**, worked in **AGILE** methodology, and **deployed 7** projects to **production**

Bioinformatics Programmer Jan. 2020 – Apr. 2020
Ontario Institute for Cancer Research *Toronto, CA*

- Installed and managed **12 projects & software** in containers for use in **microservice pipelines**
- Created **R, Python** and **Bash** scripts for metrics data from **cancer pipelines**. Reported data to new **dashboards** and **increased** the **rate of analysis** by **50%**
- Combined software together to run in a **workflow** using **wdl**, **increasing** the **speed** of the pipeline by **70%**

Projects

IMG://REPO | *Python, Flask, Jinja2, Sqlite3, Image Classification, Colour Processing*

- Developed a **full-stack web application** using with **Flask** serving a **REST API** with **Jinja2** as the frontend
- Implemented a **CRUD database** and login system with **browser cookies** and **authentication tokens**
- Leveraged **image processing** to get most common image colour and used a **machine learning API** for labels

NOSQL Database Emulation | *C++, Object Serialization, JSON, Data Management, File Organization*

- Created a **C++** program to emulate a database like **MongoDB** where **objects** are **serialized** in **files** and delivered through a custom **JSON output function**
- Utilized the **CRUD** structure in creation and allows for **efficient delivery** of a single or all entries
- Leveraged **file organization** to keep the object information within files for **permanent storage**

Education

University of Waterloo Waterloo, CA
Bachelor of Applied Science in Computer Engineering, Honours *Sept. 2019 – April 2024*

Certifications:
Azure: AI Fundamentals (AI-900), JP Morgan: Software Engineering Program, Goldman Sachs: Engineering Program