Rishi Shah

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

Technical Skills

Languages: Python, C++, Java, R, SQL (PostgreSQL, sqlite3), JavaScript, HTML/CSS, Bash, Javascript
Frameworks: Flask, React, Node.js, Django, REST API, jinja2, Material-UI
Developer Tools: Docker, AWS, Git, LinuxCL, Jenkins, JIRA
Libraries: pandas, NumPy, Matplotlib, SciKitLearn, Tensorflow, Plotly, NLTK

Experience

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

- Updated and integrated a **REST API** in **node.js** with a **PostgreSQL** database for file/data management in the pipeline, increasing **efficiency** of data transfer by **300%**
- Created a **COVID-19 analysis** program 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**
- Designed and produced various **Python, R** and **bash** software to upgrade and **re-implement proprietary** legacy code, leading to **operation times cut** by **50%** and a **25% reduced codebase**
- Worked with **AGILE** methodology, created documentation, reports/tickets in **JIRA**, **Jenkins tests** and **deployed 7** projects to **production**

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

- Created and modified **R, Python** and **Bash** scripts to create metrics data from cancer files. Reported data to **4** new **dashboards** and **increased** the **rate of analysis** by **50%**
- Installed **12** cancer research **projects & software** in **docker** containers on **AWS**, using **Amazon ECS**
- Integrated software together to run in a **workflow** using **wdl**, **increasing** the **speed** of the pipeline by **70%**
- Effectively worked on a **cluster** of high performance **computing nodes** in a **LinuxCL** environment

Projects

IMG://REPO | *Python, Flask, Jinja2, Sqlite3, Image Recognition, 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 the most common colours in the image and used a **machine learning API** to get image characteristics

COVID-19 API & Visualizer | *Python, Flask, Sqlite3, Web-Scraping, Plotly, Pandas, MatPlotLib*

- **REST API** to get data from the **Ontario** and **Canadian government** websites and return **cases, deaths**, etc
- Integrated a backend to return **requests** for data from the API by **date** and **province** through connecting to a **database** and using **dataframe**
- Creates a **pi-chart** and **line graph** of overall cases in provinces and **trends** of new cases/deaths respectively

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**

Various Data Science & ML Projects | *Python, SciKitLearn, NumPy, Pandas, NLTK*

- Implemented a **Logistic Regression** algorithm with **11 features** and **2 engineered features** to predict success of a loan with a **90% accuracy**
- Created a **natural language processor** using **NLTK** and a **Naive-Bayes algorithm** to detect spam emails with a **success rate** of **97%**
- Used a **Random Forest Classification** algorithm to detect **benign/malignant** breast cancer with **98%** accuracy

Education

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