Shivam Saraf

+1-647-569-5470 | sasaraf@uwaterloo.ca | LinkedIn | GitHub

Education

- **University of Waterloo** — Waterloo, CA
  Bachelors in Computer Science; GPA: 3.98, Average: 90.40 % — *September 2019 - April 2024*

  Scholarships and Awards: *Computer Science International Upper-Year Scholarships, President's Research Award, President's Scholarship of Distinction, Upper-Year Term Distinction, Dean's Honours List.*

Skills Summary

- **Languages**: Python, R, Javascript, C++, C, SQL, Scala, Scheme, HTML, CSS3.
- **Frameworks**: NodeJS, Hadoop, Kafka, Django, Flask, React, Spark, Airflow, Jest.
- **Tools**: PostgreSQL, MySQL, Selenium, Ansible, Docker, Kubernetes, Tableau, Grafana, REST API's, JIRA.
- **Libraries**: MongoDB, Numpy, Kubeflow, Pandas, Scikit, NLTK, TensorFlow, Keras, OpenCV, Socket.io.

Experience

- **Index Exchange** — Toronto, CA
  Software Developer — *May 2021 - Present*
  - Designed and deployed a **High Availability Hashicorp Vault** cluster provisioned with a **Consul** backend for secret management within the company.
  - Currently developing and automating several **CI/CD** pipelines, with **Ansible**, to dockerise and deploy the Vault service on several **Kubernetes** or bare metal production clusters.
  - Transported metrics and logs, of the Vault service, using **Telegraf** and **Filebeat** to **Kafka** clusters for analysing and monitoring the system in **Grafana.**

- **University of Waterloo** — Waterloo, CA
  Research Assistant — *September 2020 - December 2020*
  - Worked under Professor **Srinivasan Keshav** to estimate the lighting load of the largest commercialised buildings in the world.
  - Integrated **Google's Streetview API** with **Python** to automate the process of extracting images with over **100** different **fov's (Field of Views)** and **radii.**
  - Utilised **OpenCV** for developing feature detection algorithms to estimate the dimensions of the windows in order to calculate the flux and the intensity of light entering the building.

- **Novantas Incorporated** — New York, USA
  Data Engineer — *May 2020 - August 2020*
  - Extensively used **Scala** and **Spark-SQL** to improve the performance of the existing queries by over **92%** and developed **MapReduce** jobs in **Java** for log analysis, analytics and data cleaning.
  - Assisted in the design and implementation of **Spark Pipelines** that ingest more than **500 GB** of data daily through multiple clients.
  - Worked on performance tuning in **Hive** and **Impala** using multiple methods such as **dynamic partitioning** and **cost based optimization** coupled with the creation of several dashboards and worksheets in **Tableau.**
  - Developed, maintained and deployed models in **PMML** by writing scripts in **Python** that acquire data from more than **10,000** sources. Developed **MapReduce** jobs in Python for Machine Learning and Predictive Analysis in **Hue** on **AWS**.

- **Coffee 'N Code** — Waterloo, CA
  Senior Project Lead — *January 2020 - April 2020*
  - Responsible for developing a **four month course** to educate High School and University students in **Python.**
  - **Course content** revolved around building strong foundational concepts in **Object Oriented Programming** in order for students to be comfortable developing **2-Dimensional games** in Python.
  - Reduced needed budget for future events by **10%** due to a record turnout of over **500** people coupled with a streamlined and an improved project process flow.

Projects

- **Safe Crowds** — Waterloo, CA
  Tenserflow, Pandas, OpenCV, Matplotlib, Scikit-learn. — *September 2020*
  - Developed a program using an Object Detection Model, **YOLOv3**, to verify if social distancing is being practiced between people.
  - Utilised **Scikit-Learn** and **Pandas** to integrate algorithms such as K-Means Clustering to classify red zones.