

# Shivam Saraf

+1-647-569-5470 | [sasaraf@uwaterloo.ca](mailto:sasaraf@uwaterloo.ca) | [LinkedIn](#) | [GitHub](#)

## EDUCATION

- University of Waterloo** Waterloo, CA  
*Bachelors in Computer Science; GPA: 3.98* *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.  
**Relevant Coursework:** Object-Oriented Software Development, Algorithms, Data Structures and Data Management, Introduction to Combinatorics, Computer Organization and Design, Foundation of Sequential Programs.

## SKILLS SUMMARY

- Languages:** Python, Javascript, C++, C, SQL, Golang, Scala, Schema, HTML, CSS3.
- Tools:** PostgreSQL, MySQL, SQLite, Docker, Git, Gitlab, Kubernetes, Jenkins, Jira, LaTeX.
- Frameworks:** NodeJS, Hadoop, Mongoose, Django, Flask, React, Angular, Spark, Cassandra, Airflow, AJAX, Jest.
- Libraries:** MongoDB, Numpy, Pandas, Scikit, NLTK, TensorFlow, Keras, Socket.io.

## EXPERIENCE

- 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** to develop 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

- Connec-Ted** Windsor, CA  
*Node JS, Twillio.* *March 2020*
  - Developed a **bot** that interfaces SMS sent from a client and provides desired directions and news **without internet access**.
  - Won the **People's Choice Award** at Winhacks, University of Windsor with over **400** applicants.
- Safe Crowds** Waterloo, CA  
*Tensorflow, 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.