

# ARJUN MOHAMMED

✉ ai3moham@uwaterloo.ca — ☎ 2269788822 — 🌐 arjun.mohammed.io — in junjun99 — 📷 junjun99

## SUMMARY

---

- **Languages:** Python, C++, PHP, Bash, Java, C#
- **ML Tools:** sklearn, numpy, pandas, nltk, spaCy, gensim, textblob, jupyter, matplotlib, stanford coreNLP
- **Technologies:** Git, Linux command line, Mongo, SQL, Flask, Laravel, Jira, Google Firebase, L<sup>A</sup>T<sub>E</sub>X
- **Experience:** text mining, information extraction, AWS (S3 & SageMaker), data pipelines and storage

## EDUCATION

---

### University of Waterloo

Sept, '17 – Pres

*Candidate for Bachelor of Mathematics in Statistics with Minor in Computer Science, Honours Co-op Program*

- GPA: 3.8/4.0, Expected Graduation: Apr, '22
- Relevant Courses: Data Structures and Data Management, Applied Linear Models, Algorithms, Object-Oriented Software Development, Computational Statistics and Data Management, Applied Probability, Combinatorics

## EXPERIENCE

---

### Data Science Intern

May – Aug, '19 & May, '20 – Pres

*Loom Analytics* [↗](#)

*Toronto, ON*

- Developed and prototyped a new import feature that automatically generates database schemas based on user input to prevent data redundancy, reducing storage space up to 40% for insurance claims, while facilitating faster queries
- Created python modules to extract data points from legal cases in pdf to cut search and discovery time by 80%
- Implemented a full stack, multi-processing text classification solution leveraging Facebook's fastText that allows users to annotate legal cases, split data for model creation and predict incoming files through a web interface

### Data Analyst Intern

Sep – Dec, '18

*HelloGbye* [↗](#)

*Toronto, ON*

- Collected and parsed flight requests in email thread data using the most recent spell-check algorithms to increase accuracy of tokenization by 30% in data collected from non-native English speaking regions
- Utilized Stanford CoreNLP to extract names and location data from flight requests to streamline accurate passenger tagging and limit expensive API calls by reducing location data ambiguity

### Technical Operations Intern

Jan – Apr, '18

*Interset AI and Cybersecurity (now Micro Focus)* [↗](#)

*Ottawa, ON*

- Automated the installation of the insider threat platform through bash scripting and introduced a single sign-on feature to meet the enhanced security needs of new clients

## RESEARCH

---

### Local Differential Privacy Frameworks

Sep, '19 – Pres

- Research on the extension of central privacy frameworks to the local and decentralised settings to eliminate the need for trust between users and a central server under Prof. Xi He

## PROJECTS

---

### LabelLearn (Grand Prize Winner of HackMIT 2019) [↗](#)

Sep, '19

- Data labelling tool that facilitates consistent decisions from users by showing the history of each data-point and live class distribution, while also predicting new data-points using AWS ML solutions
- Utilises AWS BlazingText, Google Firebase, ExpressJS, NodeJS, Python, sklearn, pandas, spaCy, gensim

### Data Collection, Probability and Statistics [↗](#)

Aug, '17 – Pres

- Co-author of a free textbook and solutions manual for Caribbean A-level Probability and Statistics course
- Currently 1000+ downloads from high-school students all over the Caribbean with another book in progress

## HOBBIES

---

- Travel, Table Tennis, Rocket League, Ukulele, Pool, Gym, Squash, Reading, Guitar, Watching stand-up comedy