

# YUAN ZHANG

✉ yuanz2299@gmail.com ☎ 650-924-6009 📄 yzhang0.github.io

## SKILLS

---

**Languages/Tools** PHP, Python (Django), JavaScript (React), C++, SQL, Hadoop (Spark, Hive)

## EXPERIENCE

---

### Facebook

*Instagram Ads* 2021 - Present

- Ensure ads remain compliant for EU privacy requirements by rolling out a **new data annotation system** across the Instagram Ads org.
- Fix rendering errors by resolving the ambiguity between different ad products, surfaces, and stacks, saving the company **millions in revenue**.
- Safeguard critical codepaths with exception handling and better testing guidance.
- Lead a tooling workstream and **mentored 4+ engineers** to improve dev efficiency.

*Site Integrity* 2019 - 2021

- Designed spam detection system for **new Messenger products**: cross app messaging, end-to-end encrypted messaging, and communities messenger.
- Implemented the reporting flow on E2EE messages to mitigate adversarial attacks affecting >100k users.
- Launched Instagram content spam classifiers and **doubled the amount of daily spam caught**.
- Led war rooms to mitigate **spam attacks flagged by VPs** and in the news that target **over 300k users**.

### Facebook

Jan 2019 - Apr 2019

*Software Engineering Intern*

*Menlo Park*

- Improved Fake News classifiers' precision by 25% using **gradient boosting** and **word embedding**.

### Salesforce

May 2018 - Aug 2018

*Software Engineering Intern*

*San Francisco*

- Optimized platform components in **Python** and **Java** with replication and scalability features to transport large scale data from key-value store to relational databases.
- Developed **ETL pipelines** and **UDF** to efficiently store, query and analyze data in **Hadoop (Oozie, Hive)**.

### Scotiabank

Sep 2017 - Dec 2017

*Data Science Intern*

*Toronto*

- Cleaned and feature engineered millions of customer browsing data points using **R** and **Hadoop Framework (HDFS, MapReduce, Spark)** to investigate the efficiency of third party data for marketing campaigns.
- Predicted digital banking likelihood of new customers using **xgboost** and random forest with **73%** accuracy.

### IBM

Jan 2017 - Apr 2017

*Data Analytics Developer*

*Ottawa*

- Built frontend data visualization using **D3.js**.
- Assessed data visualization libraries **Jinja2**, **Kibana**, **Plot.ly**, and **NVD3** built a new monitoring dashboard.

## EDUCATION

---

### University of Waterloo, Canada

Sep 2015 - Aug 2019

Honours B. Math., Joint Major in Computer Science and Statistics

Coursework: Operating Systems, Object-Oriented Programming, Algorithms, Data Structures, Statistical Learning, Forecasting, Applied Linear Models.