

# Alvin Hou

Stanford, CA 94305

[alvinhou@cs.stanford.edu](mailto:alvinhou@cs.stanford.edu) | Github: [alvinbhou](https://github.com/alvinbhou) | Website: [alvinhou.com](https://alvinhou.com) | LinkedIn: [linkedin.com/in/alvinbhou/](https://linkedin.com/in/alvinbhou/)

## EDUCATION

---

### Stanford University

Stanford, CA

*M.S. in Computer Science | Artificial Intelligence track*

*Sep 2019 - Mar 2021*

**GPA:** 4.0/4.0 **Awards:** Citadel AI Coding Competition 2020 2<sup>nd</sup> place **Teaching:** TA of CS224n Natural Language Processing

**Coursework:** Machine Learning, Natural Language Processing, Machine Learning with Graphs, Mining Massive Data Sets, Principles of Data-Intensive Systems, ML System Design, Parallel Computing, Data Management and Data Systems1

### National Taiwan University

Taipei, Taiwan

*B.B.A. in Information Management*

*Sep 2014 - Jun 2018*

**Rank:** 1/39 **GPA:** Overall 3.95/4.3; Major 4.05/4.3 **Teaching:** TA of Calculus

**Honors:** NTU Presidential Award, Phi Tau Phi Honor Society, Get Fresh System Development Contest National Championship

## WORK EXPERIENCE

---

### Twitter

San Francisco, CA

*Machine Learning Engineer | Recommendations Platform*

*May 2021 - Present*

### Apple

Cupertino, CA

*Software Engineer Intern | PySpark · Spark ML*

*June 2020 - Sep 2020*

- Designed and built a **new scalable machine learning pipeline**; implemented **anomaly detection** algorithms with PySpark which achieved **1500%** runtime performance boost compared to previous works (151 → 10 minutes for 155 million of data)
- Run dimensionality reduction and unsupervised clustering algorithms with **SparkML** and **Scikit-Learn** to provide data insights
- Built **Tableau Dashboards** for visualizing large-scale ML analysis reports to support ~100 engineers in building new products

### ASUS Intelligent Cloud Services

Taipei, Taiwan

*Software Engineer (Machine Learning) Intern | NLP · PyTorch · Java · DevOps*

*May 2019 - Jul 2019*

- Implemented Multi-Task Learning BERT models for **Natural Language Understanding** with Pytorch; fine tuned the models on 100k+ training data and achieved **92.8 F1** score with 22 class on 15,000 testing data
- Built a data annotation tool for displaying incorrect parsing result; constructed an **autonomous CI/CD workflow** for a team of 10 Machine Learning and Linguistic engineers (**Azure Pipeline, Docker, Python, Java**)

### IBM

Taipei, Taiwan

*Software Engineer Intern | Fullstack · Salesforce*

*Jul 2017 - Jun 2018*

- **Fullstack developer** of IBM Impact Grants program, developed the NGO Teach For Taiwan's website in production
- Reduced the NGO's operating cost by  $\frac{2}{3}$  through integrating **Salesforce Sales Cloud** into its donation and membership system

### Industrial Technology Research Institute

Hsinchu, Taiwan

*Software Engineer Intern | Deep Learning · Web*

*Jul 2016 - Feb 2018*

- Researched on **Plants Recognition Project** to classify and detect 240 Taiwan endemic plant species images with **Keras**; improved top-5 accuracy from 60.0% to 81.6% using ResNet-50; integrated the model with Telegram chatbot to support end users
- Developed and deployed a photo map feature using **Google Map API**, on Taiwan's largest education media website Cloudplay

## SKILLS & ACTIVITIES

---

- **Programming Language:** Python, JavaScript, Java, C/C++, SQL **Web Development:** Node.js, React, Flask, PostgreSQL
- **Machine Learning:** Pytorch, Spark, Scikit-Learn, Numpy, Pandas, FastAPI **Others:** Linux, Git, Azure, Docker, Bash
- **Organizations:** NTU Open Source Community: Open sourced **campus support bot** and web extensions with **20,000+ users**

## SELECTED PROJECTS

---

**Recommendation System for CS Graduate Programs** [FastAPI, PostgreSQL, Docker, Heroku] [[Website](#)] [[Code](#)] [[Docs](#)]

- Built APIs with **FastAPI** and **Postgres** that returns personalized program recommendations based on a student's background
- Deployed the website and Swagger docs on Heroku with a CI/CD workflow using **Docker Compose** and **CircleCI**