

**Xiaohan<sup>(Shaw Han)</sup> Yang**  
xiaohan\_yang@g.harvard.edu | (857) 498-1264

## EDUCATION

---

Harvard University Cambridge, MA  
MS, Data Science GPA: 4.00 / 4.00 Expected Dec 2022

- Relevant Coursework: AI, Systems Development for Computational Science, Deep Learning in NLP

New York University Shanghai New York, NY  
BS, Business and Finance, and Data Science GPA: 3.88 / 4.00 May 2019

- Relevant Coursework: Databases, Data Structures, Intro to Optimization, Intro to Machine Learning
- Honors: Magna Cum Laude, Outstanding Graduate of Shanghai, Dean's Undergraduate Research Fund Grant

## PROFESSIONAL EXPERIENCES

---

Tencent Holdings Ltd. Shenzhen, China  
*Recommender System Research Intern* May-July 2021

- Researched and experimented with nine treatment effect models to optimize the overall ROI in social media advertisements, leading to a 4.2% relative increase in conversion rate
- Developed a calculator to compute the minimum sample size required by A/B tests, and published a blog about corresponding theories of Z-tests and Hoeffding's inequality on the company's internal forum
- Explored and summarized the theories behind the Explain Data function in Tableau to help to automatize the data analysis pipeline for non-tech colleagues in the Financial Technology group

BioClou Technology Co., Ltd. Shanghai, China  
*Data Analyst Intern* Mar-Apr 2021

- Built an automatic pipeline that could generate customized reports regularly for the clients who used the conversational AI services, reducing labor costs by over 46,000 dollars annually for this startup
- Drafted a financial model and helped to prepare the pitch deck for the fund-raising purpose

Institution of New Economic Development Chengdu, China  
*Research Assistant* June 2019-June 2020

- Proposed and conducted a research project to quantify implicit gender differences in the Chinese labor market by building an LSTM model to classify 26 million anonymous résumés via Keras
- Contributed to the investigation of the Covid-19 outbreak in Sichuan by collecting data about confirmed cases and analyzing people's sentimental reaction on Micro Blog
- Translated The Doomsday Calculation by William Poundstone, which will be published in 2021

## RESEARCH & PROJECTS

---

Harvard University Cambridge, MA

- *Thesis in progress*: Designing active learning strategies to develop a semi-automatic annotation tool for coreference resolution, which will mitigate the time-intensive nature of annotation in the field
- *Covid prediction*: Applied a conditional RNN model to predict the increase of confirmed cases in each county in the US, reducing 20% prediction loss of a traditional epidemic model
- *AutoDiff package*: Designed and implemented a Python package with three teammates, which can be used easily to perform forward-mode automatic differentiation and to compute higher-order differentiation

## SKILLS & INTERESTS

---

- Technical: Python 3 (pandas, NumPy, scikit-learn, matplotlib, PyTorch), SQL, Markdown, Bash
- Language: Mandarin (native), English (fluent)
- Affiliations: Harvard CSSA (business development chair), Harvard iLab (venture program participant)