

# Yifu Hou

Chicago, IL 60615 | (872) 303-8538 | yifu@uchicago.edu

## EDUCATION

---

### University of Chicago

Chicago, IL

MS *Computational Analysis and Public Policy*

Expected June 2023

Coursework: Machine Learning, Statistics, Databases, Data Visualization, Big Data and High Performance

### Communication University of China

Beijing, China

BA *Communications, Media and Journalism* (GPA: 3.8/4.0)

June 2021

## SKILLS

---

**Python** (NumPy, Pandas, PyTorch, Scikit-Learn) | **SQL** (MySQL, PostgreSQL) | **R** | **Microsoft Excel**

**Cloud Storage & Computing** (Spark, AWS) | **Data Visualization** (D3, matplotlib)

**Statistical Analysis** (Linear Regression, Hypothesis Testing, Causal Inference, Predictive Model, Program Evaluation)

**Machine Learning** (CNN, RNN, BERT & SentenceTransformer, Support Vector Machine)

## EXPERIENCE

---

### Computer Science Department, The University of Chicago

Chicago, IL

*Teaching Assistant*

Sept 2022 – Present

- Assisted teaching Python language, Hash Table, Web Scraping, Record Linkage, Decision Tree, etc.

### United Nations Women

Remote

*Research and Data Intern*

June 2022 - Aug 2022

- Classified over 600 NGO commitment survey responses with **Natural Language Processing**, identified need of support and engagement in migrant and LGBT groups, improved internal efficiency on commitment evaluation
- Improved internal efficiency by building a **Python** framework for UN Women outreach contact spreadsheet, merging databases with **Pandas** and cleaning duplicates with record linkage methods
- Evaluated energy access by analyzing World Bank and UNDESA energy data with **SQL**, contributed research result for 2022 UN Gender Snapshot

### The Nielsen Company Ltd.

Beijing

*Digital Ads Rating Intern*

Oct 2020 - Jan 2021

- Evaluated brand health for LinkedIn, identified a 22% gap between brand recognition rate and conversion rate by analyzing consumer survey data with **SQL** and **Excel**
- Analyzed brand equity index (BEI) by user's city, gender, and career status, created **data visualization** for brand funnels across multiple competing products
- Conducted **A/B Testing** and pre-processed consumer data in collaboration with the Data Science team

## PROJECTS

---

### Backtrack: 8K Loan Information Detection with S-BERT

Jan 2023 - Mar 2023

- Improved project internal efficiency by building **NLP** pipeline with **Semantic Similarity model** and **FAISS sentence embedding method**, correctly identified 82.5% of 8-K forms containing loan/bond information
- Developed **Logistic Regression** and **Neural Network** models, reduced false negative rate to 0.11 and achieved 0.938 recall score for the classification of loan/bond information

### Fake News Detection Application with BERT

- Attained 87% testing accuracy in fake news detection by training **BERT Sentence Transformer** on 40,000+ news

## COMMUNITY INVOLVEMENT

---

Board member of OUTPolitik – UChicago Harris LGBTQ student organization

June 2022 - Present

Red Cross Volunteer at Köln, Germany

Summer 2018 & 2019