

# Yang Yang

☎ (313)327-9512

✉ hit.youngyang@gmail.com

🌐 linkedin.com/in/young-yang

## Skills

- **Programming Languages:** Python, R, JAVA, C/C++, JavaScript, HTML
- **Frameworks & Tools:** Scikit-learn, Numpy, Pandas, NLTK, Pytorch, JupyterNotebook, Git, Docker, SOLR, SQL
- **Machine Learning Knowledge:** Various models and methods for deep learning, regression, dimension reduction, topic modeling, NER, information retrieval, sentiment analysis, Random Forest, RNN, LSTM, Transformers, BERT, etc.

## Experience

### University of California San Diego | Postdoc Employee

San Diego, CA

THE ROBERTS LAB, HALICIOĞLU DATA SCIENCE INSTITUTE

09/2017 – Present

- Led independent researches on policy diffusion process in autocracy government using a corpus-based framework
- Applied state-of-the-art NLP models to various research projects (see below)
- Managed multiple research projects and delivered quarterly policy reports to CEOs from different industrial sectors
- Published interactive data visualizations on Chinese society, work was cited by popular media like Politico and NYT
- Developed a data analytics infrastructure that supports interactive topic spotting, tracing, and news searching for researchers

### Yale University | Research Scientist

New Haven, CT

YALE SCHOOL OF PUBLIC HEALTH

09/2016 – 08/2017

- Led a five-person team to investigate patients' health-seeking behavior change after using a smart healthcare system
- Proposed process redesign suggestions that reduce 20% of patients' waiting time
- Was invited to give a talk on this research at the 2017 iHEA world congress, a top conference in health economics

### Zhejiang University | Research Assistant

Hangzhou, China

CENTER FOR INFORMATION TECHNOLOGY AND ECONOMY & SOCIETY SYSTEMS

09/2012 – 07/2016

- Designed and conducted research on the sharing economy platform (DiDi) use and service supply influences
- Provided suggestions of driver incentives to the platform and proposals of industrial policy to the government
- Led research groups of 5-10 members to accomplish 3 national-award-winning research projects (including Taobao/Tmall inter-provincial trade analysis, Weibo use in public service, cloud computing impacts on enterprises)

## Selected Projects

### Congress Tweets @China

UC San Diego, 2019-2020

- Developed a live tracker that analyzes the aspect-level sentiment towards foreign countries in congressmen's tweets
- Fine-tuned a Transformer for aspect-level sentiment prediction, reached an accuracy of 85%

### Name Ethnicity Analysis

UC San Diego, Summer 2020

- Trained a GBM that use first and last name to make the ethnicity prediction among 187 nations
- Increased accuracy of the model by 32% when compared to the based model using the word embeddings

### AI Bias Research

UC San Diego, Spring 2020

- Created a sentence completion app that let users compare political rhetoric variations between AIs trained on different corpora

### CCP Elite Visualization Portal

UC San Diego, Fall 2020

- Developed an interactive dashboard which allow users to visualize some of the key characteristics of roughly 1700 CCP elite who were or are active government officials during 2012 and 2022.

### Newspaper Analysis in Real-Time (with San Diego Supercomputer Center)

UC San Diego, 2018-2019

- Developed a news analytics platform that supports interactive topic spotting, tracing and analyzing news articles.
- Reduced the search latency by 80% using text-specific search DB (SOLR) and increased ACU 5 times using Play Framework

### Policy Document Infrastructure

UC San Diego, 2017-2018

- Designed and implemented the data pipeline (from data scraping) of 200,000+ policy records and the data storage
- This solution enables 50+ researchers at UC San Diego to perform a fast and efficient search on Chinese policy documents

### Hackathon on Chinese Verdict Analysis (with Columbia Law School & Berkeley Law)

UC San Diego, Spring 2019

- Design the analysis strategy and pipeline of text data cleaning
- Draw hypotheses and tested on the pattern of long duration of the judicial hearing, delivered a high-quality report in 3 days

### Sharing Economy Platform Use and Service Supply Influences

Zhejiang University, 2015-2016

- Extracted patterns of taxi drivers' client searching and continuous working intention from 29 million trip trajectories.

## Education

### Zhejiang University

Hangzhou, China

PH.D., MANAGEMENT SCIENCE AND ENGINEERING

09/2011 – 06/2016

Thesis: *Impacts of Taxi Hailing App on Taxi Drivers' Drivers' behavioral Decision and Operation Performance*

### Harbin Institute of Technology

Harbin, China

BS WITH HONORS, INFORMATION MANAGEMENT AND INFORMATION SYSTEMS

09/2007 – 06/2011