

# Zekun Wang

3111 Long Blvd – Nashville, TN – 37203, USA

✉ zekun.wang.us@gmail.com ☎ +1 (615) 972-0128

🔗 waittim.github.io in in/zekun-wang/

## Education

### Vanderbilt University

*Master of Science in Data Science*

GPA: 3.93/4.00

Core courses: Data Science Algorithms, Data management system, Modeling & Machine Learning, Big Data & Scaling

Nashville, TN

Aug.2019 – Apr.2021

### Central South University

*Bachelor of Science in Statistics*

Final semester ranking: 1/96

Changsha, China

Sep.2014 – Jun.2018

## Skills

**Programming Languages:** Python(Pandas, Scikit-Learn, Keras), R(ggplot2, Shiny, H2o.ai), SQL, MATLAB

**Software & others:** Tableau, SPSS, Eviews, Microsoft Office, Caffe, LaTeX, RDBMS(MySQL)

## Professional Experience

### Latin American Public Opinion Project (LAPOP), Vanderbilt University

*Research Assistant*

Nashville, TN

Feb.2020 – Now

- Designed user interface of dashboard by Sketch, and created an interactive website demo through Shiny R.
- Created Tableau Dashboard to help unskilled users complete autonomous exploratory data analysis.

### Hunan Fenghuang Taifeng Cultural Tourism Investment Ltd.

*Intern, Leader of Statistical Analysis Intern Group*

Fenghuang, China

Jan.2018 – Mar.2018

- Created ARIMA to anticipate the tourism conditions and analyzed 23 dimensions of survey data by SPSS.
- Led analysis and presented with commercial report for the \$300M "Nanhuali" tourism development project.

### Hunan Yuanfang Information Technology Ltd.

*Intern, Operation Analyst*

Changsha, China

May.2016 – Jul.2017

- Extracted elements from articles and typesetting feature as variables to perform multivariate statistical analysis.
- Applied findings to improve the UX of social media and led to an increase in advertising revenue by more than 150%.

## Projects

### H1B Visa Application Analysis

*Provided employment guidance for international students in DS-related majors.*

Nashville, TN

Oct.2019 – Dec.2019

- Explored the trends of data-jobs flow in different fields and the salary distribution of different roles from 3M observations.
- Packages & skills used: R(ggplot2, rMaps, gganimate), text processing, time series analysis, data visualization

### Trending Video Analysis and Prediction

*Discovered essentials of YouTube trending videos and made a trending-prediction model.*

Nashville, TN

Oct.2019 – Nov.2019

- Analyzed 41K trending videos on YouTube and findings about When/How/What questions of creating trending videos.
- Built a NLP model(Naive Bayes) to predict trending levels and did sentiment analysis based on the text data.
- Packages & skills used: R, Python(nltk, sklearn), feature engineering, natural language processing, sentiment analysis

### HR Consulting: Employee Attrition Prediction

*Developed the ML models that can potentially save up to \$30M for companies.*

Nashville, TN

Sep.2019 – Oct.2019

- Examined the dataset, collected insights from collaboration with 40 HRs from various industries.
- Dived into feature engineering and boosted the performance by building a random forest model.
- Packages & skills used: R(h2o), feature engineering, random forest, consulting

### Natural Background Leaf Family-Classification with Deep Learning

*Earned 1st Prize of Stats department and Undergraduate Excellent Graduation Thesis of CSU.*

Changsha, China

Mar.2018 – Jun.2018

- Conducted web crawler to form a dataset containing 18K leaf-images.
- Innovated watershed algorithms with maker-controlled feature extraction and increase accuracy by 50%.
- Finetuned deep learning models with convolutional neural networks to identify the leaf and have an accuracy of 94.31%.
- Packages & skills used: Python, Caffe, Matlab, dCNN, image segmentation, image classification