

# LE ZHANG

📞 (+86)17317812540    ✉ z118300720002@gmail.com    👤 <https://zhangle.netlify.app/>

## EDUCATION

### Fudan University

*Bachelor of Engineering in Biomedical Engineering*

Sep. 2018 – June 2022(expected)

Shanghai, China

- GPA: **Major: 3.81 / 4.0**, Overall: 3.75 / 4.0 (freshman: 3.57, sophomore: 3.9, junior: 3.82 )
- Rank: 5<sup>th</sup> / 227

### Core Courses

- **Mathematics:** Mathematical Analysis A (4.0/4.0), Linear Algebra, Probability, Statics and Stochastic Process (4.0/4.0), Elements of Information Theory (4.0/4.0), Engineering Mathematics (4.0/4.0)
- **EE&CS:** Programming, Python Programming (4.0/4.0), Data Structure (4.0/4.0), Machine Learning (4.0/4.0), Computer Architecture (4.0/4.0), Database System (4.0/4.0), Digital Image Processing (4.0/4.0), Medical Imaging Technology (4.0/4.0), Digital Logic (4.0/4.0), Signal and Communication System (4.0/4.0)

**Research Interest:** Natural Language Processing, Multi-modal Machine Learning

## EXPERIENCE

### Georgia Tech's Social and Language Technologies (SALT) lab

May 2021 – Sep 2021

*Summer Research Intern, Advisor: Prof. Diyi Yang*

Remote

- Proposed **TreeMix**, a novel compositional data augmentation method for natural language understanding
- Evaluated on a wide range of benchmarks including GLUE and empirically validated that TreeMix outperforms current state-of-the-art data augmentation methods

### Shanghai QiZhi Research Institute

Feb 2021 – Apr 2021

*Research Intern, PI: Prof. Hang Zhao*

Shanghai, China

- Expanded the task formulation and datasets introduced in *Visual Indicates Sound* by reprocessing original datasets and adding multi-modal features
- Designed a framework based on patch-wise contrastive learning and introduced a sound embedding module to utilize sound signals as an indicator for the transformation
- Designed a crowd-source labeling HTML page for Amazon Mechanical Turk

### Fudan Data Intelligence and Social Computing Lab (DISC)

Sep 2020 – present

*Research assistant, Advisor: Prof. Zhongyu Wei*

Shanghai, China

- Built a multi-modal Chinese event dataset: Collected multi-modal documents (image or video & text) from official and commercial press accounts on 微博 Sina-weibo using crawler. Predefined categories and sub-types for social events e.g. {Military: [conflict, weapon operation, transportation]}
- Designed a crowd-source labeling pipeline as well as a crowd-source labeling HTML page

## PUBLICATION & MANUSCRIPTS

1. **Le Zhang**, Zichao Yang, Diyi Yang. TreeMix: Compositional Constituency-based Data Augmentation for Natural Language Understanding. **Submitted to** *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*. [page][code]

## SELECTED COURSE PROJECTS

### Machine Learning [page][code] | Python, PyQt5

November 2020

- Developed a face mask detection & classification framework based on YOLO5 and ResNet50, and a GUI implemented with PyQt5
- Applied various techniques such as Mixup and BatchNorm

### Database System [page][code] | Python, PyQt5, MySQL

December 2020

- Implemented an administration system connected with PolarDB, providing services such as adding courses and enrolling students
- Designed a GUI for the system based on PyQt5 and replaced SQL commands with simple click-operations

## SELECTED AWARDS

- Fudan Huatai Securities Technology Scholarship for 2020-2021 ( Highest scholarship sponsored by Huatai Corp., top 1% )
- Fudan HUAWEI Student Scholarship for 2019-2020 ( Highest scholarship sponsored by HUAWEI Corp., top 1% )
- Fudan University Outstanding Undergraduate Student Scholarship Second Prize for 2018-2019 (top 5%)

## SKILLS

**Programming Languages:** Python, C/C++, JavaScript, HTML/CSS, SQL, Bash, MATLAB

**Tools and Frameworks:** Git, L<sup>A</sup>T<sub>E</sub>X, Pytorch, Huggingface, Docker