# Yu Fang

MASTER GRADUATE · UNIVERSITY OF MINNESOTA - TWIN CITIES, COMPUTER SCIENCE

+1 530-761-3439 | fangx174@umn.edu | goucejia.github.io | goucejia

## Education

**University of Minnesota, Department of Computer Science and Engineering**    *Minneapolis, Minnesota*
MASTER OF SCIENCE, COMPUTER SCIENCE    *Jan. 2020 - May 2022*
- Advisor: Prof. Dongyeop Kang
- GPA: 3.9 / 4.0

**University of Minnesota - Twin Cities, College of Science and Engineering**    *Minneapolis, Minnesota*
BACHELOR OF SCIENCE, COMPUTER SCIENCE    *Sep. 2014 - Dec. 2017*
- Technical GPA: 3.8 / 4.0

## Research Experience

**Yanqi Lake Beijing Institute of Mathematical Sciences and Applications (BIMSA)**    *Beijing, China*
ADVISOR: DR. HAIHUA XIE    *Oct. 2022 - Present*
- Project: Seq2Seq Grammatical Error Diagnosis for Chinese language with transformer-based language models. [P2]

**University of Minnesota, Twin Cities - Dept of Computer Science and Engineering**    *Minneapolis, MN*
ADVISOR: PROF. DONGYEOP KANG - MINNESOTA NLP GROUP    *Jun. 2021 - Present*
- Directed research: Data-centric approach analysis for understanding contrastive learning in language models & generative methods of high-quality contrastive sample pairs. [P1]
- Project: Interactive narrative analysis on open-domain fan-fiction data based on transportation theory. [P3]

**Peking University Founder Group - State Key Laboratory of Digital Publishing Technology**    *Beijing, China*
SUPERVISORS: PROF. LIANGCAI GAO, PROF. YONGTAO WANG    *Jun. 2018 - Jul. 2019*
- Project: Adaption of YOLO-based models on OCR page object detection tasks.
- Project: Target-format PDF-file + annotation generation tool for auto-generation of valid new training samples.
- Model performance measurements tool, track-B task datasets and the cTDaR website for the ICDAR 2019 Table Recognition Competition. [C1]

## Publications

### CONFERENCE

[C1] Liangcai Gao, Yilun Huang, Hervé Déjean, Jean-Luc Meunier, Qinqin Yan, **Yu Fang**, Florian Kleber, and Eva Maria Lang. 2019. ICDAR 2019 Competition on Table Detection and Recognition (cTDaR). *International Conference on Document Analysis and Recognition (ICDAR)*, pages 1510–1515, Sydney, Australia. DOI:10.1109/icdar.2019.00243

### PREPRINT

[P1] **Yu Fang**, Petros Karypis, Dongyeop Kang. (2022) Understanding Contrastive Learning via Meta-Informative Cube Analysis. (In submission)

[P2] Haihua Xie, Xiaoqing Lyu, Xuefei Chen, **Yu Fang**. (2022) String Editing Based Chinese Grammatical Error Diagnosis with a Unified Error Correction Mechanism. (In submission)

### IN PREP

[P3] **Yu Fang***, Kelsey Neis.* (2022) Understanding Narrative Transportation in Fantasy Fan-fiction.

## Professional Experience

| | |
|---|---|
| 2022-2023 | **Research Assistant (Remote)**, Yanqi Lake Beijing Institute of Mathematical Sciences and Applications (BIMSA) |
| 2021-2023 | **Research Assistant**, University of Minnesota - Twin Cities, Minnesota NLP group |
| 2018-2019 | **Research Assistant**, Peking University Founder Group - State Key Laboratory of Digital Publishing Technology |

## Teaching Experience

| | | |
|---|---|---|
| Spring 2022 | **CSCI 4041: Algorithms and Data Structures**, Graduate Teaching Assistant | *UMN* |
| Fall 2021 | **MSBA 6421 : Predictive Analytics**, Graduate Teaching Assistant | *UMN* |

## Awards

| | |
|---|---|
| Spring 2017 | **Dean's list**, College of Science & Engineering, UMN |
| Spring 2016 | **Dean's list**, College of Liberal Arts, UMN |
| Spring 2015 | **Dean's list**, College of Liberal Arts, UMN |

## Outreach & Professional Development

### SERVICE AND OUTREACH

| | |
|---|---|
| Reviewer | **EMNLP**, 2022 |

## Work & Internship Experience

**PKU Founder IT Group** — *Beijing, China*
SOFTWARE ENGINEER — *Jun. 2018 - Jul. 2019*
- Developed the online annotation API for our trained OCR model using python-Django & HTML-CSS-JS to display the model-predicted annotation results.
- Implemented User module, with modification and sharing function to update the annotation dataset.

**MaesInfo (Chengdu) Co. Ltd.** — *Chengdu, China*
SOFTWARE DEVELOPER INTERN — *Feb. 2018 - Jun. 2018*
- Developed a data management tool for reliability test.
- Developed the User module of Melinked.com.

## Language & Skills

- **Language**: Chinese(native), English(proficient)
- **Programming Language**: Python, Java, JS+HTML+CSS, C++
- **Machine Learning Framework**: Pytorch, Tensorflow