Chihsheng Jin

He/him/his, 23

Email: jinzhisheng2099@outlook.com
Address:
514 Lattimore Hall University of Rochester 500 Joseph C. Wilson Blvd. Rochester, NY 14627 United States

EDUCATION

§ University of Rochester | Master | Computational Linguistics | GPA: 4.0/4.0 09/2023-05/2025(expected)
§ Fudan University | Bachelor | Chinese language | GPA: 3.52/4.0 09/2019-07/2023
§ Shanghai International Studies University | Intercollegiate Minor | French | GPA:3.98/4.0 09/2020-07/2022
§ Beijing Language and Culture University | Summer School | Seminars on phonetics and phonology 07/2021

HONORS & AWARDS

Excellent Student Award | Fudan University | 2020
3rd –class Student Scholarship | Fudan University | 2020, 2019
2nd place in Fudan Original Music Competition | Fudan University | 2022

PROJECTS & PROFESSIONAL EXPERIENCE

§ Agenda-based Parser for Multiple Context-free Grammars

Project Description: For this course project, I developed an agenda-based parsing Python module for multiple context-free grammars a full test suite implemented in Pytest. Multiple context-free grammar is a class of grammar that is strictly more expressive than context-free grammar.

§ Music Classification from superficial features using DNN

Project Description: This was a project for my data mining class. I used various models including kNN, XGBoost, SVM, and DNN to utilize the superficial features in the *30000 Spotify songs dataset* from Kaggle for developing a music genre classifier. The classifier using DNN yielded the best performance with 71% percent accuracy which is on par with some of the classifiers that were trained on heavily engineered audio features.

§ Dynamics in the phonological encoding of bilingual speech production

Project Description: This is the thesis for my bachelor's degree. I used a picture-naming paradigm to test whether the phonological mapping status of cognates in Mandarin and Shanghai Dialect will take an effect on the speech production process and found a cognate facilitation effect when the phonological forms of the cognates match the basic mapping rules regardless of the discrepancy of their phonetic forms.

§ Late Qing Dynasty Dictionary Digitalization and Phonological Analysis

Project Description: I contributed to the digitalization of a late Qing Dynasty dictionary (*Petit Dictionnaire Français-Chinois*). After transcribing contents, I developed a Python program to analyze the phonological structure and map character pronunciations to Middle Chinese rime tables. This facilitated the study of phonological shifts in the Shanghai dialect over the past century.

WORK EXPERIENCE

§ Research Assistant in the FACTS.lab

Duration: 05/2024-05/2025 (expected)

Description: I'm currently working on event key summarization with Professor Aaron White. The objective of the series of studies is to improve machine summarization by training models to extract the information of a certain event from multiple source files. For now, I'm collecting annotation following FrameNet's ontology and training models to summarize events in a report text with keyed event arguments.

§ Project Intern at Shanghai Artificial Intelligence Laboratory

Duration: 07/2022-09/2022

Description: I managed student data using Python and Excel, organized key events such as the World Artificial Intelligence Conference 2022 and academic lectures, and assisted in recruitment for principal investigators' research groups by conducting interviews and evaluating candidate CVs.

§ Responsible Editor of *Yayan* (a school publication on language and culture)

Duration: 10/2019-06/2022

Description: I contributed to various aspects of the publication process as the responsible editor for *Yayan*. My responsibilities included developing new topics, curating articles from professors and peers, overseeing layout design, performing proofreading tasks, and managing the publication process.

PROFESSIONAL SKILLS

Skills

Machine learning, Statistics, Natural language processing, French (intermediate)

Tools

Python: good (Numpy, Pandas, Scikit-learn, Pytorch);

C/C++: moderate;

Praat: good (experienced in collecting phonetic data and moderate scripting experience);

SPSS: moderate (descriptive statistics, hypothesis testing, linear regression, factor analysis);

Ableton Live: advanced (music production and mixing);

Photoshop: advanced;

InDesign: good.