Shan Chen

Waltham, MA, USA 02453 | +1 (507) 210-1913 | shaaan@brandeis.edu | Web: shanchen.dev | Linkedin

EDUCATION

Brandeis University, Waltham, Massachusetts   **Expected: May 2022**
Master of Science, Computational Linguistics (STEM Designated, NLP + Data Science + Linguistics)

St. Olaf College, Northfield, Minnesota   Sep 2016 - May 2020
Bachelor of Arts, Mathematics, Japanese, Linguistics Concentration   Majors/Overall GPA: 3.7/4.0
- Honors/Awards: Cum Laude with distinctions, Japanese National Honor Society, Pi Mu Epsilon, JASSO Scholarship, National Japanese Exam Silver Prize, Henry Luce Research Grant
- Semester abroad: Nagoya University School of Engineering, Nagoya, Japan

RELATED COURSES

Algorithms & Data Structures, Applied Algorithms, Graph Theory, Neural Computation, Natural Language Annotation for Machine Learning, Object-Oriented Software Design (C++&Java), Information Retrieval, Statistical Machine Learning & NLP, Natural Language Processing, NLP System, Syntactic Theory, Semantics

LANGUAGES AND FRAMEWORKS

Python, Pytorch, C++, Java, R Studio, Flask, Flutter, Firebase, Docker, HTML & CSS, NoSQL
Chinese (Native), English (Native), Japanese (Professional working proficiency), German (Elementary)

WORK EXPERIENCE

NLP Intern, Boston Children's Hospital & Harvard Medical School**,** Longwood, MA   June 2021 - **Present**
- Computational Health Informatics Program NLP Research mentor: Timothy Miller, PhD
- Clinical NLP adaptation and pipeline development on electronic health records.
- Working with a postdoc on a #SMM4H challenge, developing a drug name normalization system for social media texts.

Graduate Research & Teaching Assistant, Brandeis Xue Lab, Waltham, MA   Aug 2020 - **Present**
- Mathematical Methods in Linguistics TA: received 100% positive feedback at the end of the semester.
- Nianwen Xue's Lab RA: building millions level Chinese COVID news corpus and adapting Chinese news documents temporal & event relations extraction models within the BBN system.

Data Engineering Intern, Elphi, Cambridge, MA   Dec 2020 - Feb 2021
- Worked on Real-time Dynamic data with Tableau and Machine learning loan prediction algorithms.
- Integrated the algorithm into the main app using GCP Pipeline to achieve systematic automation.

Software Developer, Unismart・ユニスマ, Tokyo, Japan   June 2020 - Sept 2020
- Developed a cross-platform app using Flutter and Firebase framework designed to help to connect 10k+ Waseda university students during the COVID-19 pandemic with an international team.
- Supported real-time push notification, messaging, course calendar system with Firebase API.

Linguistics Research Assistant, St. Olaf College, Northfield, MN   June 2018 - July 2019
- Conducted data analysis and synthesized data into a published paper (publication listed below). The Paper was published as a linguistics book chapter in 2019 by Lexington Press as: "*The Development of Willingness to Communicate in L2 Chinese Writing*".
- Generated and analyzed Japanese web-sourced linguistics data for Ito, Ph.D.'s linguistics research.

SELECTED PROJECTS

Washington Post Information Retrieval System   Apr 2021 - June 2021
- Web app retrieving information from Washington post corpus with boolean and vector space search.
- Fine-tuning FastText and BERT embeddings, adding spell checks and query expansion to improve TREC retrieving results by 20% using NDCG metrics.
- Leveraged frameworks: Flask + Python + Huggingface for front & backend implantations.