# Shiyi Shen

✉ shiys3447@yahoo.com

🔗 https://github.com/shiyishen3447

## Summary

Four-year experience in general linguistics plus one year in computational linguistics and natural language processing. Fluent in Python and conversant in Java; worked with Large data. Built mini projects and immersed in the state of the art NLP research toolkits, including NLTK, Scikit Learn, and UPenn Tree Bank. Conducted mini NLP projects like sentiment analysis with logistic regression algorithm, Chinese NER using Conditional Random Fields(CRF), building Youtube Comment Extractor, and detect folksiness in inauguration speeches using the appropriate textual analytic toolkits.

## Experience

### Research Assistant

Beth Israel Deaconess Medical Center

Oct 2017 - Oct 2017 (1 month)

Accomplished a mini project on bio text mining and bio-marker extraction. Understood the basic contribution of biomarkers to cancer prognosis and treatment predictions. Used Python, Excel, and a heat map to produce a small Bio-NLP report.

## Education

### Brandeis University

Master of Science - MS, Computational linguistics

2018 - 2022

Natural Language Processing; HMM; Naive Bayes; Finite State Morphology; NLTK; Chomsky's Transformational Grammar

### University of Pittsburgh

Bachelor of Arts - BA, Linguistics

2013 - 2017

Sentiment analysis; NER; Finite State Automata

## Skills

Computational Linguistics  •  Natural Language Processing (NLP)  •  Python (Programming Language)  •  Java  •  GitHub  •  NLTK  •  Sentiment Analysis  •  Information Extraction

## Honors & Awards

### SEALS scholarship - University of Pittsburgh

May 2014

Summer East Asian Language Studies (short for SEALS) is offered by the University of Pittsburgh Department of East Asian Languages & Literatures for a concentrated mix of intensive training in language and culture.