Peter Zukerman

(+1) 216-538-6721 | pzukerman17@gmail.com | linkedin.com/in/peterzukerman | github.com/peterzukerman

Education

University of Washington — Seattle, WA
Master of Science in Computational Linguistics — *Sep. 2021 – Dec. 2022*

University of Illinois at Urbana-Champaign — Champaign, IL
Bachelor of Science in Computer Science and Linguistics — *Aug. 2017 – May 2021*

Experience

Natural Language Processing Intern — June 2020 - Aug. 2020
Raytheon BBN — *Cambridge, MA*

- Developed an AI visualization metric, enhancing detection by 83%
- Pioneered a sorting algorithm, boosting accuracy by a factor of 2.6 with no efficiency loss
- Designed logic for data models, increasing data collation and improving visualization

Speech Processing Intern — May 2019 - Sept. 2019
VoiceItt LTD. — *Tel Aviv, Israel*

- Integrated VoiceItt backend infrastructure with Amazon Alexa
- Introduced a low-resource data augmentation paradigm, improving speech recognition by 24%
- Expanded parselmouth library, optimizing Praat's usage for machine learning in Python

Research

Speech Perception (JP/US) | *University of Tokyo* — Since Oct 2019

- Analyzing instantaneous perception of Japanese English by listeners with various language profiles
- Working with Prof. Nobuaki Minematsu - published paper to Phonetic Society of Japan
- Concluded English speech utterances have 70% comprehension rates among Japanese native speakers

Natural Language Processing | *University of Illinois at Urbana-Champaign* — Since Jan 2019

- Co-authored ACL paper on adversarial speech synthesis, currently submitted
- Improved a morphological analyzer for the Siberian Yupik language with Prof. Lane Schwartz
- Wrote a Yupik E-Book parser, now in active use in St. Lawrence Island community
- IYIL 2019 Paper on initiating a tool-building infrastructure the Yupik community - Fort Wayne, IN
- Presented at CodeMash 2019 on Unicode, UTF-8, and foreign languages

Project Highlights

Twitter Sarcasm Detection | *Python, Jupyter, Keras, Pandas, BERT* — Dec. 2020

- Data augmentation and preprocessing increased usable dataset from 5000 to 20000 entries
- Implemented using BERT and improvements over various state of the art language models
- Ranked #1 on online text classification leaderboard with 1000+ competitors

Coursework

Linguistics: Syntax for CompLing, Semantics and Pragmatics
Computer Science: Data Structures, Algorithms, Systems Programming, Applied Machine Learning
Computational Linguistics: Shallow Techniques in NLP, Statistical NLP, NLP in Polysynthetic Languages
Planned Coursework: Deep Processing Techniques in NLP, Generative Grammar, Applied ML in NLP, Phonetics

Technical Skills

Languages: Python, Java, C/C++, JavaScript, HTML/CSS, Ruby, R
Developer Tools: Git, Docker, Linux, Google Cloud Natural Language, AWS
Libraries: pandas, Sci-Py, NumPy, Matplotlib, NLTK, carmel, Tensorflow, PyTorch

Foreign Languages: Russian (Fluent), Japanese (Business Proficiency), Korean and Spanish (Conversant)