Wenyan Li Email:wenyanl62@gmail.com
https://www.linkedin.com/in/wenyanli/

Education

- **University of Maryland** College Park, MD
  Master of Science in Electrical Engineering *Aug 2018*
  - Relevant Coursework: Computational Linguistics, Machine Learning, Database Design, Convex Optimization, Computer Processing of Pictorial Information
- **Northwestern Polytechnical University** Xi'an, China
  Bachelor of Engineering in Electrical Engineering and Automation *June 2016*
  - Ranked 1/97 (top 1%); Outstanding Graduate Award

Employment

- **Comcast Applied AI** — *Senior Research Engineer, Machine Learning* Washington, D.C.
  NLP & Data Science *Jan 2019 – Present*
  - Designed an **unsupervised auto-annotation pipeline** which used user behavioral modeling to automatically identify errors in speech recognition and NLP systems and suggest corrections; summarized the work into a conference paper as the first-author and submitted a patent as main contributor
  - Developed a context-based approach that discovered misclassified user queries in question answering systems by performing semantic search with Sentence-BERT and clustering
  - Leveraged **subword-level query representation** and adversarial training in customer care dialogue system for misspelled user queries, which improved classification accuracy by 18% and increased user experience stability
- **JD Digits AI Lab** — *Research Intern* Mountain View, CA
  Customer Service Chatbot *Oct 2018 – Dec 2018*
  - Implemented attention-based CNN and RNN models for user query classification in the online question answering system

Research Experience

- **CLIP Lab, University of Maryland** — *Master's Thesis Research* College Park, MD
  Deep Learning for Verb Prediction; Advisor: Jordan Boyd-Graber *Sep 2017 – Aug 2018*
  - Developed an end-to-end and incremental verb prediction model for translation latency reduction in simultaneous machine translation, and significantly improved prediction accuracy in both German and Japanese
  - Implemented synonym-aware verb prediction for German and provided interpretable visualization of the prediction process
- **Computational Biology Group, University of Maryland** College Park, MD
  Predicting Phenotype from Genomic Sequences; Advisor: Max Leiserson *Sep 2017 – Dec 2017*
  - Experimented with random forest and an attention-based LSTM model for genotype-phenotype reasoning which predicts genetic interactions directly from DNA/amino-acid sequences

Publications & Patents

- **W. Li**, A. Grissom II, J. Boyd-Graber, "ANVIIL: An Attentive Recurrent Model for Incremental Prediction of Sentence-final Verbs", *Findings of EMNLP*, 2020
- **W. Li** and F. Ture, "Auto-annotation for voice-enabled entertainment systems", in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR, July 2020*
- "Systems and Methods for Training Voice Query Models". U.S. Application Serial No.: 63/056,361. filed July 24, 2020. Patent Pending.

Programming Skills

- **Languages**: Python, MATLAB, SQL
- **Frameworks and Tools**: PyTorch, Tensorflow, Keras, Scikit-Learn, PySpark, Git, Docker, Snorkel, Latex