JIAZHAO LI

1839 Shirley Ln, Ann Arbor, MI. 48105
(+1)734 604 1596, jiazhaol@umich.edu, https://jiazhaoli.github.io

EDUCATION

University of Michigan, Ann Arbor U.S.
M.S in Electrical Computer Engineering *Sept.2017 – May.2019*
Ph.D in Informatics *Sept.2020 – Apr.2025*
Nankai University China
B.S in Electrical Engineering *Sept.2013 – June.2017*

RESEARCH INTEREST

Natural Language Processing Health Informatics.
Question Answering, Knowledge Graph, Domain-adaptive Pre-train Model, Interpretable model.

RESEARCH EXPERIENCE

PharmMT: A Neural Machine Translation Approach to Simplify Prescription Directions.
In Findings of EMNLP'20 *Sept 2019 - Feb 2020*

- Built Nerual Network-based MT model between Prescription and Pharmacy directions corpus.
- Augmented model using MIMIC-III domain-specific pre-trained word embedding, external information from Drug/ Strength.
- Applied ensemble learning and numerical checking to improve accuracy and avoid fictitious generations.
- Applied BLEU score and SARI score to do automatic evaluation on MT performance and developed web app to do manual evaluation by pharmacists.

Re-ranking biomedical literature for precision medicine with pre-trained neural models.
ICHI'20 *Jan 2019 - May 2019*

- TREC precision medicine information retrieval challenge on ontology topics.
- calculating the relevant score using lexical-matching based iterate information retrieval method.
- calculating the relevant score using domain-adaptive contextual word embedding model BioBERT . Combining two relevant score using Rank Fusion.
- 6.2% improvement on inferred NDCG and 6.8% improvement on R-precision against SOTA models .

Identify Medication Relations from Clinical Narratives [Paper]

· Identifying medication relations between drugs and associated attributes automatically from clinical narratives to develop advanced tools for decision support. This is part of 2018 national clinical NLP challenge.

- Developed and shared python tokenization package for pre-processing MIMIC clinical notes for team.
- Feature engineering including part-of-speech tag, named-entities-recognition tag, pre-trained token embedding and bi-direction relative position of target entities pair.
- Developed Bi-LSTM models to extract 8 relations between drug names and adverse events associated concepts with F1 0.892 outperforming CNN and SVM model.

Baby blues: Analyzing Facebook and health forums on Pregnancy [Paper]

- Systematically analyzed questions posted on pregnancy forums by young mothers and contrasted it to a unique dataset of Facebook posts by expectant adolescent mothers.
- Implemented the topic model Latent Dirichlet Allocation (LDA) to extract the top 10 themes of questions in different trimesters during pregnancy.

- Concluded that Facebook is chosen as a self-expression place to seek emotional support while health forums served as professional information providers.

Movie Revenue prediction with Hierarchical Model [Poster] *Jan.2019 - Apr.2019*

- Observed movie revenue following two mixed Gaussian Distribution: High revenue and Low revenue.
- Based on observation, trained High-Low binary Random Forest Classifier based on labels resulted from Gaussion Mixture Model (GMM) clustering.
- Trained movie Gradient Boosting Regression (GBR performed best) separately on two revenue group and using Back-Off strategy to solve cold start problem.

Video Segments Retrieval System based on Attentive CNN [Report] *Sep.2018 - Nov.2018*

- Enhanced video clip embedding with attentive-weighted contextual video segments embedding.
- Generated cross latent feature between video clip embedding and corresponding video content description text embedding through outer product.
- Trained ACNN model on TACoS dataset with loss function on video-text similarity and offset of video clips achieved 0.347 (IoU=0.5) and 0.719 (IoU=0.1) in Top10.

PUBLICATIONS

Lester CA, Ding Y, **Li J**, Jiang Y, Rowell B, Vydiswaran VGV. Comparing human versus machine translation of electronic prescription directions. *Under review at JAPhA (Journal of the American Pharmacists Association).*

Chang T, DeJonckheere M, Vydiswaran VGV, **Li J**, Buis L, Guetterman T. Transforming Mixed Methods Research with Natural Language Processing. *Under review at the Journal of Mixed Methods Research.*

Jiazhao Li, Corey Lester, Xinyan Zhao, Yuting Ding, Yun Jiang, and V.G.Vinod Vydiswaran. PharmMT: A Neural Machine Translation Approach to Simplify Prescription Directions. *In Findings of EMNLP, the 2020 Conference on Empirical Methods in Natural Language Processing. Pages:2785–2796.*

Jiazhao Li, Adharsh Murali, Qiaozhu Mei, V.G.Vinod Vydiswaran. Re-ranking biomedical literature for precision medicine with pre-trained neural models. *Proceedings of the IEEE International Conference of Healthcare Informatics (ICHI), 2020.*

Zhao X, **Li J**, Lester C, Ding Y, Jiang Y, Vydiswaran VGV. *Focused representation with lexical constraints for parsing prescription instructions to decrease medication error.* (Under review)

Zhao X, **Li J**, Lester C, Jiang Y, Vydiswaran VGV *Focused representation models for transcribing prescription instructions.* (Poster MIDAS 2019 Symposium)

WORK EXPERIENCE

Graduate Student Research Assistant	Sep.2020 - present	Prof. VG Vinod Vydiswaran
Research Associate	Aug.2019 - Aug.2020	Prof. VG Vinod Vydiswaran
TA Grader	Jan.2019 - Apr.2019	Prof. Rada Mihalcea

HONORS AND AWARDS

- Sep. 2016. Third Prize Scholarship of the University (10%), NKU
- Sep. 2016. Merit Student of the University (5%), NKU