YISONG MIAO

+65 94609940 | miaoyisong@gmail.com | www.yisong.me

Natural Language Processing, Information Retrieval, Conversational Recommender Systems

EDUCATION

National University of Singapore (School of Computing) Singapore
Master of Computer Science Candidate (GPA: 4.3/5) Aug 2018 – Now
Undergraduate Exchange Student (GPA: 4.2/5) Jan 2017 – May 2017

University of Chinese Academy of Sciences Beijing, China
Founding Class of Bachelor of Science in Computer Science (GPA: 3.5/4) Sept 2014 – June 2018

PUBLICATIONS

[1] *Estimation-Action-Reflection: Towards Deep Interaction Between Conversational and Recommender Systems*
Wenqiang Lei, Xiangnan He, **Yisong Miao**, Qingyun Wu, Richang Hong, Min-Yen Kan & Tat-Seng Chua ACM WSDM 2020

- This work is my first research experience. I was closely advised by Dr. Lei, Dr. He, Dr. Kan and Dr. Chua.
- We introduce a novel conversational recommender system, which is able to recommend item to user through conversation. Our work emphasise on the strategy of the conversation, which is solved by Reinforcement Learning. We also innovate classic Factorization Machine model to achieve both item prediction and attribute prediction through multi-task learning framework.
- My contributions in this work include the whole research pipeline: idea innovation, literature review, model design, system implementation, experiments and evaluation, paper writing.
- Project Website: https://ear-conv-rec.github.io/

[2] *Towards Computing Contextual Lexical Contrast: Cont2Lex Corpus, Recognition benchmarks, and Preliminary Analyses*
Wenqiang Lei*, **Yisong Miao***, Runpeng Xie, Bonnie Webber, Meichun Liu, Tat-Seng Chua, Nancy Chen, (In submission to Transaction of ACL, * equal contribution)

- In this work, I gain more independent research training under the guidance of Dr. Lei, and Dr. Chua.
- We systematically investigate a fundamental NLP phenomenon: the lexical relation should be dependent on the context where they are shown. We pay special attention to lexical contrast, which is always under investigated. We first built a large and high-quality corpus to benchmark this task. Later we scrutinise the performance of state-of-the-art contextual word embeddings like BERT and ELMo, GPT, etc.
- My contribution in this work is also the whole research pipeline. My progress compared with [1] is that I gain more independent research skills.

ACHIEVEMENTS

- Oral presentation at ACM Web Search and Data Mining Conference Houston, USA, Feb. 2020
- Reviewer / Secondary Reviewer Services: NLPCC'2019, AAAI'2020
- First Runner-up in Kaggle In-class Competition of News Article Classification. [Code]: https://github.com/YisongMiao/CS5228-project, [Leaderboard]: https://www.kaggle.com/c/cs5228-article-category-prediction

PATENT

An Advanced Conversational Recommender System Emphasising on the Interaction Between Conversational Components and Recommender Components, SG Non-Provisional Application No. 10202000482X
Wenqiang Lei, Xiangnan He, **Yisong Miao,** Min-Yen Kan & Tat-Seng Chua

RESEARCH EXPERIENCE

National University of Singapore (School of Computing) Singapore
Master Student at Web Information Retrieval and Natural Language Processing Group Aug 2018 - Now

- **Master thesis** is on the topic of conversational recommendation, accepted to WSDM'2020, Houston, USA.
  - Conducted comprehensive literature survey and formalise this interesting task as a research problem, and distilled three key research questions: (1) what to ask and what to recommend? (2) what strategy to ask and recommend? (3) how to adapt to user's online feedback?

- Designed and implemented the three phase workflow of our conversational recommendation system. The *Estimation* stage for item and attribute prediction using BPR; *Action* stage for deciding asking attributes or making recommendation; *Reflection* stage for updating user prediction when rejecting a list of items.
- Designed and conducted extensive experiment to evaluate our system.

- **Ongoing research projects** on Implicit Discourse Relation Detection
  - As is widely known in the academic community, existing system's performance on Implicit Discourse Relation Detection is far from usable. In this work, we strive to take discourse relation as instance, to answer what information can neural network models capture.
  - We first capture complex linguistic features, then designed a series of probing task to validates if neural network models can capture them.

University of Chinese Academy of Science, National University of Singapore Beijing, Singapore
Bachelor Thesis, jointly advised by Prof Min-Yen Kan, and Prof Yanyan Lan Jan 2018 - May 2018
Webpage Classification and Metadata Extraction using URL only, the motivation of such project is: traditional webpage classification methods require fetching full webpage, which can be quite time-consuming.
- Designed and implemented a deep learning workflow for char-level and word-level LSTM model, the latter one has achieved the strongest performance so far.

Activities And Services

Special Interest Group in Math/Interdisciplinary Modeling at UCAS Beijing, China
Founder, Head of the Group Feb 2015 - Sept 2015
- Founded the Special Interest Group in Math/Interdisciplinary Modeling in my university(UCAS). The group grew from 0 to 50+ members during my service.
- My responsibilities included: Organizing Seminars, Inviting Lectures, Giving Tutorials.

University Student Union at UCAS Beijing, China
Deputy Director of Social Practice Department Sept 2014 – August 2015
- Led the department to hold Thanksgiving Day series activities(for 200) and Music Band Festival(for 300).

Selected Awards

Meritorious Winners in the Interdisciplinary Contest in Modeling Beijing, China, Jan 2016
- Designed and created a System-Dynamics model to predict the future water shortage based on real-world data in multiple areas as the leader of team, the same model can give solutions to the shortage.

Second Prize of National Olympiad in Physics Competition (50 / ~5000) China, 2013

Additional Information

- Programming Languages: python, c/c++, java, sql
- Natural Language: Mandarin Chinese(Native), English(Fluent, TOEFL-ibt 99), Spanish(Beginner)
- Tools and Frameworks: PyTorch, Stanford CoreNLP, spaCy, scikit-learn, numpy, scipy, postgresql
- Interests: 10th Grade in Erhu (A Traditional Chinese Musical Instrument), swimming, jogging