

GUANLIN LI

◊ {H}omepage & {K}nowledge Vault
Algorithm Engineer
JD Inc. Beijing, China

WORK EXPERIENCE

Intelligent Algorithm Group, JD Industry. Jul. 2024 - Current

- **JDipedia project:** Management and construction of various knowledge-related assets, namely, lexicons, knowledge graph, and task-specific rule bases.
 1. **Lexicon curation:** LLM-based filtering and crowd-sourcing for brand-pattern lexicons.
 2. **Industrial product kg construction:** LLM-based product-attribute bipartite knowledge graph construction; Neo4J for kg loading and visualization, cypher for common graph queries.
 3. **Rule base curation:** attribution for rationale extraction, rationales cleaning with ambiguity score, rule base extension using a human-in-the-loop process with rationales as *constitution*.
- **Product matching/Category recognition related**
 1. **LLM-based zero-shot filtering of customer materials:** Qwen-72b for filtering of explicit query-conflicting candidates after the recall stage, 4.8/5.6pp improvements on top-1/3 acc.
 2. **GNN-based category recognition:** GraphSAINT-based GCN model training via the above kg data, about 0.8pp improvement over MLP baseline with the same node embedding feature.

Intelligent Model Group, Risk Control Center, JD Technology. Jan. 2023 - Jun. 2024

- **Anti-fraud related**
 1. **Order-level multimodal antifraud model:** contribution to legoflow (a tensorflow framework for config-based model construction); fusion of text, sequence, and gathering point features using one neural model, about 15pp of top1/10 recall improvements over online baseline.
- **Interpretability toolkit construction**
 1. **Attribution-based risky subsequence mining:** Integrated gradient attribution and frequent-pattern mining to find meaningful and high risky subsequence of browse pages/search words.
 2. **Swap set analysis based on SHAP:** SHAP-based salient feature discovery of the groups of swap-in and swap-out applicants, given two A-card models over different time windows.

EDUCATION

Harbin Institute of Technologh, Harbin, China Sep. 2015 - Apr. 2022

Ph.D. in Computer Technology

Supervisor: Tiejun Zhao

Thesis: *Understanding neural machine translation: representation, learning dynamics and behavior.*

Central China Normal University, Wuhan, China Sep. 2011 - Jun. 2015

BA. in Computer Science and Technology

Department of Computer Science

GPA: 90.6, 3/143

Central China Normal University, Wuhan, China Sep. 2011 - Jun. 2015

Minor in English and Linguistics

PUBLICATIONS AND PREPRINTS

1. *Detecting Source Contextual Barriers for Understanding Neural Machine Translation* [J], **Guanlin Li**, Lemao Liu, Conghui Zhu, Rui Wang, Tiejun Zhao, Shuming Shi, TASLP 2021.
2. *Understanding Data Augmentation for Neural Machine Translation: Two Perspective towards Generalization* [C], **Guanlin Li**, Lemao Liu, Guoping Huang, Conghui Zhu, Tiejun Zhao, Shuming Shi, EMNLP 2019.
3. *Understanding and Improving Hidden Representations for Neural Machine Translation* [C], **Guanlin Li**, Lemao Liu, Xintong Li, Conghui Zhu, Tiejun Zhao, Shuming Shi, NAACL 2019.
4. *On the Word Alignment from Neural Machine Translation* [C], Xintong Li, **Guanlin Li**, Lemao Liu, Max Meng, Shuming Shi, ACL 2019.
5. *Evaluating Explanation Methods for Neural Machine Translation* [C], Jierui Li, Lemao Liu, Huayang Li, **Guanlin Li**, Guoping Huang, Shuming Shi, ACL 2020.
6. *Generative Bridging Network for Neural Sequence Prediction* [C], Wenhui Chen, **Guanlin Li**, Shuo Ren, Shujie Liu, Zhirui Zhang, Mu Li, Ming Zhou, NAACL 2018.
7. *Attending From Foresight: A Novel Attention Mechanism for Neural Machine Translation* [J], Xintong Li, Lemao Liu, Zhaopeng Tu, **Guanlin Li**, Shuming Shi, Max Q.-H. Meng, TASLP 2021.
8. *Demystify The Learning of Unsupervised Neural Machine Translation*, **Guanlin Li**, Lemao Liu, Taro Watanabe, Conghui Zhu, Tiejun Zhao, ICLR 2021 rejected, 5 4 6 5.
9. *Understanding Learning Dynamics for Neural Machine Translation*, Conghui Zhu, **Guanlin Li**, Lemao Liu, Tiejun Zhao, Shuming Shi, arXiv 2020 (**2.5, 3.5, 4.5** rejected from ACL 2020).
10. *Approximate Distribution Matching for Sequence-to-Sequence Learning*, Wenhui Chen, **Guanlin Li**, Shuo Ren, Shujie Liu, Zhirui Zhang, Mu Li, Ming Zhou, arXiv 2018.

INTERNSHIP

- Research Intern at Tencent AI Lab** Jan. 2018 - Mar. 2021
- Mentor: **Lemao Liu**
 - Topics: Researched about interpretability problems in neural machine translation, such as attention and word alignment, learning dynamics and input intervention.
- Research Intern at Microsoft Research Asia** Mar. 2017 - Aug. 2017
- Mentor: **Shujie Liu**
 - Topics: Researched about reinforcement learning approach to neural machine translation, such as applying GAN as reward model or coaching mechanism for further improve translation generation.
- Student at Summer School on NLG, Summarisation, and Dialogue Systems.** July. 2015
- Mentor: **Kees van Deemter**
 - Topics: Studied the **data-to-text** approach of natural language generation, and used **SimpleNLG** for constructing a weather forecasting nlg system.

COMMUNITY CONTRIBUTIONS

- Series Seminars on Practicals in NLP** Sep. 2017 - Jan. 2018
- About **9 seminars** covering basic NLP techs to sota papers.
 - **30 audiences** from students in HIT.

Series Lectures on Deep Learning Basics and Its Application in NLP Oct. 2016 - Dec. 2016

- **5 lectures** based on the Deep Learning book, ranging from nn basics to structured prediction.
- Up to **40 audiences** in my department for every lecture.

STUDENT PROJECTS

Gaokao History QA System (863 Program)

Aug. 2016 - Oct. 2016

- **Task:** building a QA system which generates long-text answers for Gaokao history exam.
- **Outcome:** achieving close-to 60 points out of 100 points under human evaluation.

3D Scene Description Generation with Geometric Objects

Nov. 2014 - Apr. 2015

- **Task:** building an NLG system for describing virtual 3D scenes.
- **Outcome:** achieving Excellent Undergraduate Thesis award.

SOFTWARD SKILLS

Good Level Python, PyTorch, SQL, Linux, Tmux, Vim, Java

Basic Level C(++) , Git, Matlab

LANGUAGE SKILLS

Chinese (native speaker), English (TEM-8, Test for English Majors Grade Eight)

MISCS

Level 10 Er-hu (a traditional Chinese music instrument)

Others Finger-style guitar