DONGSHENG WANG

PhD student at School of Electronic Engineering, Xidian University, Xi'an, Shanxi

Email: wds_dana@163.com · Phone: (+86) 15091626037 · Homepage: wds2014.github.io

❤ RESEARCH INTEREST

My long-term research goal is to build explainable machines that can learn from human prior knowledge and apply such common sense in their decision process under the Bayesian framework.

My general research lies at statistical machine learning and its real-world applications, such as Deep Bayesain learning, Representation learning, *etc*. My research directions involves:

- Deep generative model, *e.g.* Variational Autoencoder, Gamma belief network, *etc*.
- Theory and applications about Optimal transport.
- Knowledge-Aware machine learning, *e.g.* Knowledge embedding, Knowledge incorporation, *etc*.
- Corresponding issues in computer vision and natural language processing, *etc*.

🎓 EDUCATION

Xidian University, Xi'an, Shanxi, China   2018 – present

PhD student in Electronics Engineering (EE),   *advised by Bo Chen*

Xidian University, Xi'an, Shanxi, China   2014 – 2018

B.S. in Electronics Engineering (EE)

👤 PUBLICATIONS

[1] **Dongsheng Wang**, Yishi Xu, Miaoge Li, Bo Chen, Minayuan Zhou. Knowledge-Aware Bayesian deep topic model. Submission to Neurips2022.

[2] **Dongsheng Wang**, He Zhao, Dandan Guo, Bo Chen, Minayuan Zhou. Building Cross-Layer Transport to Learn Hierarchical Topic Embeddings. Submission to Neurips2022.

[3] Yishi Xu, **Dongsheng Wang**, Zhibin Duan, Bo Chen, Minayuan Zhou. HyperMiner: Topic Taxonomy Mining with Hyperbolic Embedding. Submission to Neurips2022.

[4] **Dongsheng Wang**, Dandan Guo, He Zhao, Bo Chen, Minayuan Zhou. Improving Neural Topic Models by Contrastive Learning with BERT Submission to TNNLS.

[5] **Dongsheng Wang**, Dandan Guo, He Zhao, Huangjie Zheng, Korawat Tanwisuth, Bo Chen, Mingyuan Zhou. [Representing Mixtures of Word Embeddings with Mixtures of Topic Embeddings](). The Tenth International Conference on Learning Representations. (ICLR 2022).

[6] Zhibin Duan, **Dongsheng Wang**, Bo Chen, Chaojie Wang, Wenchao Chen, Yewen Li, Jie Ren, Mingyuan Zhou. [Sawtooth Factorial Topic Embeddings Guided Gamma Belief Network](). Thirty-eighth International Conference on Machine Learning. (ICML 2021).

[7] Zhibin Duan, Yishi Xu, Bo Chen, **Dongsheng Wang**, Chaojie Wang, Mingyuan Zhou. TopicNet: Semantic Graph-Guided Topic Discovery. Thirty-fifth Conference on Neural Information Processing Systems. (NeurIPS 2021).

[8] **Dongsheng Wang**, Chaojie Wang, Bo Chen, Mingyuan Zhou. Ordinal Graph Gamma Belief Network for Social Recommender Systems. Arxiv.

[9] Chaojie Wang, Hao Zhang, Bo Chen†, **Dongsheng Wang**, Zhengjue Wang, Mingyuan Zhou. Deep Relational Topic Modeling via Graph Poisson Gamma Belief Network. Thirty-fourth Conference on Neural Information Processing Systems. (NeurIPS 2020).

[10] Ruiying Lu, Dandan Guo, **Dongsheng Wang**, Bo Chen, Mingyuan Zhou. Hierarchical Topic-Aware Contextualized Transformers. Submission to IEEE Trans ASLP2022.

👥 PROJECT AND COMPETITIONS

DeeCamp, Sinovation Ventures. Beijing, China   2019.7- 2019.8

Summer Camp,   Product description generation in E-commerce

We focus on high quality product description generation in E-commerce, and an efficient algorithm is developed to generate accurate and attractive description for items only according to the input keywords of the product title. The presented method is based on Transformer under the encoder-decoder framework. In order to make personalized recommendation, we also consider user attributes, which are beneficial for modeling user interests. The generated results achieve comparable performance with existing E-commerce platforms, such as *JD* and *taobao*.

Global AI Innovation Contest. Shanxi, China   2022.4- 2022.6

AI Competitions,   Key attributes matching in E-commerce

We explore multi-modality algorithm for image and text matching task. Given the image of a item, together with an attached description, the developed method need to 1) predict whether such two modalities are semantically consistent in the sentence level; 2) predict whether the key attributes in provided description accurately match the image in the fine-grained level.

- We propose a multi-modality framework based on BERT, and view this problem as a multi-label classification task.
- An efficient sampling strategy is designed to generate positive and negative samples.
- The results achieve 40/317.

Domain knowledge incorporation and inference methods in AI. Shanxi, China   2022.7- present

Advised by Bo Chen

We propose a general Bayesian framework that allows algorithms integrate domain prior knowledge (*e.g.* WordNet, human taxonomy) and improve itself via the side information. Specifically, the proposed framework jointly

models the node feature and the edges by sharing the latent variables. The former aims to learning semantic node embedding, while the latter incorporate relational information to enhance the representation ability.

- We explore the hierarchical node representation and provide interpretable tools by visualizing different level semantics in embedding spaces.

- In order to model such hierarchical structures more accurate, The hyperbolic embedding and contrastive learning are employed on the knowledge graph.

- We adopt the graph adaptive technique to allow the proposed model can add/reweight edges according to the current dataset, which provides an option to addresses the mismatch issue between the provided knowledge graph and tha target dataset.

⚙ ACADEMIC SERVICES

NeurIPS 2022 reviewer   2022

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) reviewer   2022

ICML 2022 reviewer   2022

ICLR 2022 reviewer   2021