Danqing Wang

Computer Science - University of California, Santa Barbara - CA
✉ dqwang122@gmail.com • dqwang122.github.io

Research Interest: Interested in all kinds of generation tasks, including natural language and biological sequence. Devoted to helping the machine generate new content that benefits human everyday life. Mainly focus on (but not limited to):

- Natural Language: Summarization, Multilingual, Controllable Generation, etc.
- Biological Sequence: Peptide Design, Antibody Discovery, etc.

Github: `https://github.com/dqwang122`

Education & Experience

- **PhD in Computer Science** — **University of California, Santa Barbara**
  Advisor: Lei Li — *2022.9 – Current*
- **Algorithm Researcher** — **AI-Lab, ByteDance**
  Advisor: Jiaze Chen, Hao Zhou and Lei Li — *2020.4 – 2022.9*
- **Master in Computer Science** — **Fudan Univerisity**
  Advisor: Prof. Xipeng Qiu and Prof. Xuanjing Huang — *2018.9 – 2021.1*
  GPA: 3.72/4.0 Ranking: 15/225
- **Bachelor in Computer Science and Technology** — **Fudan Univerisity**
  GPA: 3.62/4.0 Ranking: 10/74 — *2014.9 – 2018.6*

Main Publication

- **Accelerating Antimicrobial Peptide Discovery with Latent Sequence-Structure Model**
  Danqing Wang, *Zeyu Wen, Fei Ye, Hao Zhou, Lei Li* — ***arXiv 2022***
  - Sample peptides from the latent secondary structure space to control the peptide properties.
  - The generated peptides have a high AMP probability (93.62%) and 2/21 show high activity in wet laboratory experiments.
- **Enhancing Scientific Papers Summarization with Citation Graph**
  Chenxin An, Ming Zhong, Yiran Chen, ***Danqing Wang***, *Xipeng Qiu, Xuanjing Huang* — ***AAAI 2021***
  - The graph-based model based on citation graphs for scientific paper summarization (CGSUM).
  - The scientific papers summarization dataset Semantic Scholar Network (SSN).
- **Contrastive Aligned Joint Learning for Multilingual Summarization**
  Danqing Wang, *Jiaze Chen, Hao Zhou, Xipeng Qiu, Lei Li* — ***ACL 2021 Finding***
  - A large-scale multilingual summarization corpus MLGSum with 1.1 million articles and summaries in 12 languages.
  - Propose two tasks, contrastive sentence ranking and sentence aligned substitution, for multilingual summarization.
- **Heterogeneous Graph Neural Networks for Extractive Document Summarization**
  Danqing Wang*, *Pengfei Liu**, *Yining Zheng, Xipeng Qiu and Xuanjing Huang* — ***ACL 2020***
  - Introduce word nodes to model the cross-sentence relationship for extractive summarization.
  - Easily adapt the graph model from single to multiple document summarization.
- **Extractive Summarization as Text Matching**
  *Ming Zhong**, *Pengfei Liu**, *Yiran Chen,* ***Danqing Wang***, *Xipeng Qiu and Xuanjing Huang* — ***ACL 2020***
  - Formulate extractive summarization as a semantic text matching problem and select sentences in summary-level.
  - Achieve superior performance on six benchmark datasets, including state-of-the-art extractive result on CNN/DailyMail.
- **Searching for Effective Neural Extractive Summarization: What Works and What's Next**
  *Ming Zhong**, *Pengfei Liu**, ***Danqing Wang***, *Xipeng Qiu, Xuanjing Huang* — ***ACL 2019***
  - Models with autoregressive decoder are prone to achieving better performance against non auto-regressive ones.
  - LSTM is more likely to suffer from the architecture overfitting problem while Transformer is more robust.
- **A Closer Look at Data Bias in Neural Extractive Summarization Models**
  *Ming Zhong**, ***Danqing Wang****, *Pengfei Liu**, *Xipeng Qiu, Xuanjing Huang* — ***EMNLP 2019 Workshop***
  - Workshop on New Frontiers in Summarization
  - Define four measures in *constituent factor* and *style factors* to quantify the characteristics of summarization datasets.

Honor

- **May 2021** Shanghai Outstanding Graduates (5% of graduates)
- **Nov. 2020** Venustech Scholarship (1% of Fudan students)
- **Sept. 2019** Scholarship for Outstanding Students (First Prize)
- **Dec. 2017** Fudan's Undergraduate Research Opportunities Program