

# Shikang Wu

✉ wushikang@bytedance.com | 🌐 shikangwu.github.io | 🎓 Google Scholar | 🔗 LinkedIn

## RESEARCH INTERESTS

---

Recommender systems, information retrieval, large-scale ranking models, multimodal recommendation, representation learning, and generative recommendation.

## EMPLOYMENT HISTORY

---

ByteDance Beijing, China

*Tech Lead / Staff ML Scientist (3-1), Douyin-Search* Jan. 2026 – Present

- Leading the direction of the Douyin Next-Gen Search Engine. We are building a next-generation search engine based on a Generative Search Foundation Model, a unified model combining both retrieval and ranking capabilities, and a Large Recommendation Model, an extreme-scale ML architecture scaling towards 4B dense parameters, 100T sparse features, and 100K sequence lengths.

*Senior ML Scientist (2-2), Douyin-Search* 2025 – 2026

- Lead of the Large Recommendation Model (LRM) for Douyin Search, focusing on model scaling, sequential modeling, all-token design, multi-modal co-training, and next-generation search tools.

*Senior ML Scientist (2-1), Douyin-Search* 2024 – 2025

- Lead of the Large Recommendation Model (LRM) for Douyin Search.
- Focused on sparse activation scaling (30×, using MoE, memory networks, etc.), sequence scaling (20×), sparse scaling (10×), and a multi-distribution learning all-token design.

Baidu Beijing, China

*Senior ML Scientist (T5, cross-level promotion), Baidu Search* 2023 – 2023

*ML Scientist (T3), Baidu Search* 2021 – 2023

Meituan Beijing, China

*ML Scientist Intern, Recommendation* Jun. 2020 – Sep. 2020

## EDUCATION

---

Beijing Jiaotong University Sep. 2018 – Jun. 2021  
*M.E., School of Information Technology; advised by Zhihao Wu and Youfang Lin* Beijing, China

Beijing Jiaotong University Sep. 2014 – Jun. 2018  
*B.E., School of Information Technology* Beijing, China

## PUBLICATIONS

---

### MSN: A Memory-based Sparse Activation Scaling Framework for Large-scale Industrial Recommendation.

- **Shikang Wu\***, Hui Lu\*, Jinqiu Jin\*, Zheng Chai\*, Shiyong Hong, Junjie Zhang, Shanlei Mu, Kaiyuan Ma, Tianyi Liu, Yuchao Zheng, Zhe Wang<sup>†</sup>, Jingjian Lin. *KDD 2026*. [arXiv](#)

### MDL: A Unified Multi-Distribution Learner in Large-scale Industrial Recommendation through Tokenization.

- Shanlei Mu\*, Yuchen Jiang\*, **Shikang Wu\***, Shiyong Hong, Tianmu Sha, Junjie Zhang, Jie Zhu, Zhe Chen, Zhe Wang<sup>†</sup>, Jingjian Lin. *KDD 2026*. [arXiv](#)

## **TokenMixer-Large: Scaling Up Large Ranking Models in Industrial Recommenders.**

- Yuchen Jiang\*, Jie Zhu\*, Xintian Han\*, Hui Lu\*, Kunmin Bai\*, Mingyu Yang\*, **Shikang Wu\***, Ruihao Zhang\*, Wenlin Zhao\*, Shipeng Bai, Sijin Zhou, Huizhi Yang, Tianyi Liu, Wenda Liu, Ziyang Gong, Haoran Ding, Zheng Chai, Deping Xie, Zhe Chen, Yuchao Zheng<sup>†</sup>, Peng Xu. *KDD 2026*. [arXiv](#)

## **HyFormer: Revisiting the Roles of Sequence Modeling and Feature Interaction in CTR Prediction.**

- Yunwen Huang\*, Shiyong Hong\*, Xijun Xiao\*, Jinqiu Jin\*, Xuanyuan Luo, Zhe Wang, Zheng Chai<sup>†</sup>, **Shikang Wu<sup>†</sup>**, Yuchao Zheng, Jingjian Lin. *SIGIR 2026*. [arXiv](#)

## **LEMUR: Large scale End-to-end MULTimodal Recommendation.**

- Xintian Han\*, Honggang Chen\*, Quan Lin\*, Jingyue Gao\*, Xiangyuan Ren\*, Lifei Zhu, Zhisheng Ye, **Shikang Wu**, Xionghang Xie, Xiaochu Gan, Bingzheng Wei, Peng Xu, Zhe Wang, Yuchao Zheng<sup>†</sup>, Jingjian Lin, Di Wu, Junfeng Ge. *CIKM 2026*, under review. [arXiv](#)

## **NOTES**

---

\* indicates equal contribution; † indicates corresponding author.