A Preference-oriented Diversity Model Based on Mutual-information in Re-ranking for E-commerce Search

Huimu Wang*
Mingming Li*
wanghuimu1@jd.com
limingming65@jd.com
JD.com
Beijing, China

Dadong Miao†
Songlin Wang
Guoyu Tang
miaodadong@jd.com
wangsonglin3@jd.com
tangguoyu@jd.com
JD.com
Beijing, China

Lin Liu
Sulong Xu
Jinghe Hu
liulin1@jd.com
xusulong@jd.com
hujinghe@jd.com
JD.com
Beijing, China

ABSTRACT

Re-ranking is a process of rearranging ranking list to more effectively meet user demands by accounting for the interrelationships between items. Existing methods predominantly enhance the precision of search results, often at the expense of diversity, leading to outcomes that may not fulfill the varied needs of users. Conversely, methods designed to promote diversity might compromise the precision of the results, failing to satisfy the users' requirements for accuracy. To alleviate the above problems, this paper proposes a Preference-oriented Diversity Model Based on Mutual-information (PODM-MI), which consider both accuracy and diversity in the re-ranking process. Specifically, PODM-MI adopts Multidimensional Gaussian distributions based on variational inference to capture users' diversity preferences with uncertainty. Then we maximize the mutual information between the diversity preferences of the users and the candidate items using the maximum variational inference lower bound to enhance their correlations. Subsequently, we derive a utility matrix based on the correlations, enabling the adaptive ranking of items in line with user preferences and establishing a balance between the aforementioned objectives. Experimental results on real-world online e-commerce systems demonstrate the significant improvements of PODM-MI, and we have successfully deployed PODM-MI on an e-commerce search platform.

CCS CONCEPTS

• **Information systems** → **Learning to rank**; **Information retrieval diversity**.

KEYWORDS

Re-ranking, Mutual-information, Diversity, E-commerce Search

---

*Both authors contributed equally to this research.
†Corresponding author.

---

1 INTRODUCTION

Product Search Engine is a crucial component of E-commerce platforms. To improve the efficiency of consumer decision-making during searches, a slate containing a limited number of ranked items is often presented in the search engine. In this case, most existing ranking methods [1, 3–6, 6–8, 10, 12, 18, 19] often concentrate on predicting the click-through rate (CTR) for individual items, neglecting the critical mutual influence among contextual items and the user's demands. As a result, a re-ranking stage is introduced to rearrange the initial list from the ranking stage to meet user demands.

Nevertheless, User demands are inherently dynamic, shifting based on their search patterns and interactions. Searches for broad terms typically elicit a desire for a diverse range of options, while specific queries necessitate more targeted and precise results. This behavior highlights the importance of considering both accuracy and diversity in re-ranking algorithms.

However, the existing methods often demonstrate divergent priorities. In terms of precision, certain methods tend to prioritize precision over other factors, such as RNN-based models [2, 13, 19] . These methods suffer from feature information degradation with encoding distance and non-parallelizable characteristics caused by RNN [16]. Conversely, methods like context-wise re-ranking strategy [1, 15] leverage a context-wise evaluation model to capture the mutual influence among items and re-predict the CTR of each item. However, they do not directly model the diversity of user intent, resulting in a ranking that prioritizes precision over diversity and neglects the user's diverse needs. Regarding diversity, there exist effective diversity-promoting methods [3, 4]. However, these methods rely on point-wise rank scores to measure the relevance between items and user interest, without considering the interplay between contents. Although methods such as DPP extension [4, 9, 11], have addressed this issue, balancing relevance and diversity still depends on hyper-parameter tuning and may not satisfy the

diverse needs of different users. Moreover, methods based on greedy strategies may not obtain the global optimal solution.

To track these aforementioned issues, a novel Preference-oriented Diversity Model Based on Mutual-information (PODM-MI) is proposed in this paper. Specifically, it consists of two components including a preference-oriented network (PON) and a self-adaptive model(SAM). The PON is designed to model the representation of user diversity preference and the diversity within the candidate items. In contrast, the SAM dynamically adjusts the order of candidate items to align with the user's diversity preferences.

Specifically, PON employs a multi-dimensional Gaussian distribution to model both the diversity preferences of users and the representation of diversity among candidate items, deviating from conventional embedding techniques, which yields two significant benefits: enhanced robustness compared to traditional embeddings and a more comprehensive characterization of diversity uncertainty with a higher-dimensional perspective. In the search scenario, history queries and the corresponding items of users are employed to represent the diversity preferences. Upon modeling the diversity representation of users and the candidate items, the SAM module adjusts the ranking of items to align with the user's diversity preferences. More precisely, SAM enhances the consistency between the diversity preferences of users and the diversity present in the candidate items. This consistency is accomplished by the maximization of mutual information between the user's diversity preferences and the candidate items' diversity. We introduce a variational posterior estimator to derive a tractable lower bound for the mutual information objective. This is also the first application of mutual information theory in the re-ranking model. After enhancing the consistency, a utility matrix predicated on the reinforced consistency is derived, which enables the adaptive re-ranking of items to reflect user preferences, thus maintaining equilibrium between diversity and accuracy. Moreover, the model is designed for straightforward implementation as an add-on module to any existing re-ranking algorithm.

The contributions of this paper can be summarized as follows:

- To the best of our knowledge, this is the first work that introduces mutual information to balance diversity and accuracy for E-commerce re-ranking.
- We propose a novel model that utilizes multi-dimensional Gaussian distribution to model diversity preference and adaptively matches the ranking results with user demands by maximizing the lower bound of mutual information.
- We conduct extensive experiments on a real-world dataset. Experimental results show that our model achieves significant improvement over the state-of-the-art models.
- PODM-MI has been successfully deployed on the home search platform in JD and brought substantial economic benefits.

2 PROBLEM FORMULATION

Typically, The goal of the re-ranking is to elaborately select candidate items from the input ranking list and rearrange them into the final ranking list, followed by the exhibition for users. Mathematically, with a certain user U and a ranking list of candidate items $C = \{c_i\}_{i=1}^{N_C}$, where c_i is the $i-th$ item and N_C is the total number
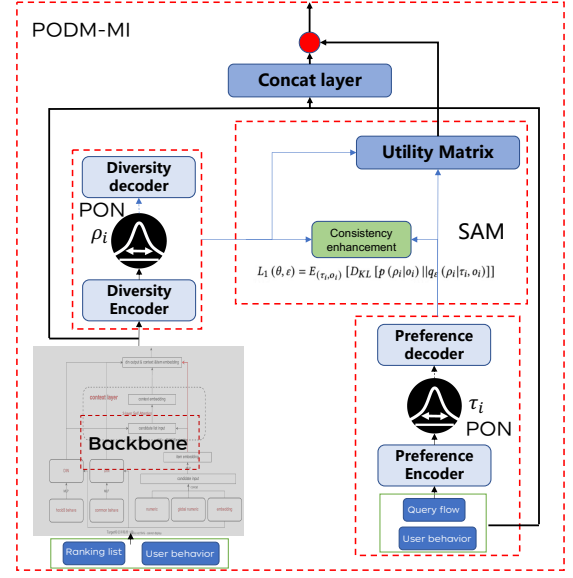

Figure 1: The framework of PODM-MI

of items in ranking list. The task of re-ranking stage is to learn a strategy π^*, $C \xrightarrow{\pi^*} F$ which selects and rearranges items from C, and present a final ranking list $(F = \{f_j\}_{j=1}^{N_F}, N_F < N_C)$ with the aim of meeting users demands such as diversity and accuracy.

3 METHODS

3.1 Overview

We present the overview structure of PODM-MI consisting of preference-oriented network (PON) and a self-adaptive model(SAM) in Figure 1. Specifically, we take ranking list $C = \{c_i\}_{i=1}^{N_C}$ and user's permutation-level behavior (such as click behavior flow, add-to-cart behavior flow) as the input of PODM-MI. Then PON is utilized to model both the diversity preferences of users and the representation of diversity among candidate items. Next, SAM enhances the consistency between the diversity preferences of users and the diversity present in the candidate items. Subsequently, a utility matrix is then derived from this enhanced correlation, influencing the scores produced by the backbone network. This matrix is dynamically tuned to user preferences, facilitating the reordering of final ranking outcomes to better satisfy user preferences. The detailed structure of PON and SAM will be introduced separately in the following sections.

3.2 Preference-oriented Network

In the e-commerce search scenario, historical queries and their associated items provide an effective representation of user intent. Thus, our methodology not only employ typical behavioral patterns (click behavior flow, add-to-cart behavior flow,etc.) but also includes query trajectories to enhance the depiction of user preferences.

After determining the features of the user preferences modeling, the next step is to determine the modeling method for user preferences representation. Traditional models tend to treat the dynamism of the user preferences as deterministic, yielding static user

embeddings within a latent space. This approach is inadequate for capturing the complex nature of user preferences in dynamic real-world contexts. Note that distribution representations introduce uncertainties and provide more flexibility compared with one single fixed embedding. Previous works [20, 29] have also demonstrated the superiority of representing user preferences as a distribution rather than an embedding.

Therefore, we adopt a multi-dimensional Gaussian distribution, as delineated in Equation 1, to model the evolving trends of the user preferences. The multi-dimensional Gaussian distribution governed by a mean vector and one diagonal covariance matrix. To be specific, user intent has two embedding representations, which are for mean μ and covariance $\sum$ that obtained from preference encoder.

$$\tau(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_1|^{\frac{1}{2}}} exp\left(-\frac{1}{2}(x-\mu_1)^T \left(\sum\right)_1^{-1} (x-\mu_1)\right) \quad (1)$$

where x is a n-dimensional vector representing the user preferences, μ_1 is the mean vector , and $\sum_1$ is the covariance matrix.

We also model the candidate items with multi-dimensional Gaussian distribution.

$$\rho(y) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_2|^{\frac{1}{2}}} exp\left(-\frac{1}{2}(y-\mu_2)^T \left(\sum\right)_2^{-1} (y-\mu_2)\right) \quad (2)$$

where y is a n-dimensional vector representing ranking list, μ_2 is the mean vector , and $\sum_2$ is the covariance matrix.

3.3 Self-adaptive Model

Upon modeling both the user preferences and the diversity inherent in the candidate items, the next step is to consider how to make the ranking results highly correlated with user intent. Mutual information, which gauges the shared information between two variables, can be utilized to quantify the correlation between candidate items and user preferences.

By maximizing mutual information between the aforementioned two (as shown in (3)), the distribution of the candidate items is highly consistent with the distribution of user intent.

$$maxI(\rho_i; \tau_i | o_i) \quad (3)$$

where o_i represents the given ranking list.

However, estimating and maximizing mutual information is often intractable. Drawing inspiration from the literature of variational inference, we introduce a variational posterior estimator to derive a tractable lower bound for the mutual information objective:

$$I(\rho_i; \tau_i | o_i) \geq E_{(\rho_i, \tau_i, o_i)} \left[\log\left(\frac{q_\varepsilon(\rho_i | \tau_i, o_i)}{p(\rho_i | o_i)}\right)\right] \quad (4)$$

where $q(x)$ is the marginal distribution of $q(x|y)$ is the conditional distribution of X given Y, and $p(y|x)$ is the conditional distribution of Y given X.

The objective function for maximizing the lower bound of mutual information can be formulated as:

$$L_1(\theta, \varepsilon) = E_{(\tau_i, o_i)} \left[D_{KL}\left[p(\rho_i | o_i) || q_\varepsilon(\rho_i | \tau_i, o_i)\right]\right] \quad (5)$$

After enhancing their consistency, we design a learnable utility matrix to further align the final ranking results with user preferences. The learnable matrix is obtained by taking the dot product between a learnable weight matrix and the features that have been

aligned. Then we multiply the utility matrix with scores computed from backbone network.

By adjusting the values of the matrix, we can control the relative importance of different items and trends in the ranking process. For example, if certain items are more suitable for user preferences, we can adjust the values of the matrix to give them more weight in the ranking process. This approach allows the ranking results to be adaptively adjusted according to user intent.

3.4 Model Training

The proposed model is easily deployed as a follow-up module after any ranking algorithm by directly using the existing ranking feature vectors. Therefore, The loss of mutual information can be considered as an auxiliaries loss. For example, we use the order data as label and minimize the loss function for a backbone (such as PRM[15]):

$$L_2 = -\sum\sum z_i log(P(z_i | Y; \theta)) \quad (6)$$

where z_i is the label (the label is positive if the item is purchased)

Then the total loss for the training process is :

$$L_{total} = L_1 + \chi L_2 \quad (7)$$

where L_1 is the loss function of mutual information and χ is the hyper-parameter.

4 EXPERIMENTS

4.1 Datasets and Metrics

We collect search logs of the user for over two months from JD.com, one of China's largest B2C e-commerce websites, with the dataset size exceeding one billion and encompassing hundreds millions of items and users.

For accuracy estimation, we choose Area Under ROC (Auc) to evaluate offline performance. $Auc_{ord}@k$ ($K \in \{1, 3, 5, 10\}$) denote Auc for the top-k items with respect to actual purchasing behaviors. For diversity evaluation, we use entropy to evaluate the richness of top-k items. $Brand/Shop@k$ ($K \in \{10, 20\}$) denote the entropy related to brand and store of items. Moreover, we use UCVR (User Conversion Rate), UCTR (User Click Rate) and UV_add_cart (User Add-to-cart Rate) to evaluate online performance.

4.2 Baselines

In the industrial field, there are many classical methods for re-ranking. The most widely used work could be divided into several backbones: list-wise, evaluator-generator,etc. Without loss of generality, we select PRM [15] (considering the personalized information) as baseline with the backbone of list-wise, and EGR [14] as the representative of the backbone of evaluator-generator. Our method is a general model, which could be adapted in various versions of re-ranking framework. In this paper, we chose the PRM with PODM-MI and EGR with PODM-MI for a fair comparison.

4.3 Implementation Details

To ensure a fair comparison among different methods, we keep the feature columns, vocabulary size, the dimension of query/item unchanged. Specifically, we set the batch size as 1024. The learning

Table 1: Overall performance comparison on the different methods in terms of $auc_{ord}@k$ and $entropy@k$.

Methods	Auc_ord@1	Auc_ord@3	Auc_ord@5	Auc_ord@10	NDCG	Brand@10	Brand@20	Shop@10	Shop@20
PRM	0.4236	0.6022	0.7176	0.8386	0.6854	1.2926	1.6299336	1.6389594	2.1226752
PRM+TODMI	0.4271	0.6052	0.7206	0.8412	0.6882	1.2953	1.63303047	1.64076226	2.12479788
Boost	**+0.83%**	**+0.50%**	**+0.42%**	**+0.31%**	**+0.40%**	**+0.21%**	**+0.19%**	**+0.11%**	**+0.10%**
EGR	0.4279	0.6052	0.7192	0.8428	0.6875	1.2965	1.6332	1.6406	2.1248
PODM-MI	0.4326	0.609	0.7217	0.845	0.6905	1.2997	1.6364	1.6427	2.1267
Boost	**+1.10%**	**+0.63%**	**+0.35%**	**+0.26%**	**+0.44%**	**+0.25%**	**+0.20%**	**+0.13%**	**+0.09%**

Table 2: Online performance of A/B tests in JD. p-value is obtained by t-test.

	UCVR	UCTR	UV_add_cart	Brand@10	Shop@10
Gain	+0.66%	+0.24%	+0.35%	+0.26%	+0.20%
p-value	0	0.01	0.01	0	0

rate is searched from 10^{-4} to 10^{-2}. All models use Adagrad as the optimizer. The hyper-parameter in loss function is set to 1.

4.4 Offline Experimental Results

The experimental results are shown in 1. From the results, we can conclude that PODM-MI achieves a significant improvement with different backbones. Specifically: 1) Regarding precision metrics, the PODM-MI model significantly outperforms the baseline model in terms of AUC, which demonstrates that the ranking results of the model can effectively meet users' needs and illustrates its superiority and effectiveness. 2) Regarding diversity, the PODM-MI model shows varying degrees of improvement in store and brand diversity compared to other methods, indicating that the model effectively improves the diversity of ranking results. It is noteworthy that while improving diversity, precision has also been significantly improved, indicating that our model has effectively balanced precision and diversity. 3) Regarding compatibility, the PODM-MI model shows significant improvements in both performance and diversity compared to the original backbone when deployed on different backbones, indicating that our model is modular and can be easily deployed as a follow-up after any ranking algorithm.

4.5 Online Experimental Results

To investigate the effectiveness of the PODM-MI, we conduct online A/B testing in the JD e-commerce search engine. The online performance is compared against the previous EGR model online and is averaged over four consecutive weeks. As shown in Table.2, the performance of PODM-MI increases by 0.66% in UCVR (p-value=0), 0.26% in terms of diversity (p-value=0), and 0.24% in UCTR (p-value=0.01). Small p-value means statistically significant, which demonstrates that PODM-MI not only enhances the likelihood of user purchases but also augments the diversity of items displayed in search results. Note that every 0.10% increase in UCVR or UCTR brings great revenues for the company, hence the improvement achieved by the PODM-MI was significant. Consequently, PODM-MI has been fully integrated into the JD e-commerce search engine.

4.6 Correlation Visualisation

Distinct entropy values for the ranking results correspond to varied user intents. In order to assess whether the ranking results are
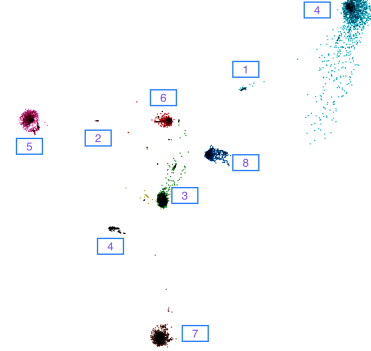
Figure 2: The correlation visualisation

high correlated with user intent, we visualized the distribution of entropy values of ranking results according to user intent using T-SNE [17] dimensionality reduction. For enhanced clustering, we categorized the entropy levels into 8 distinct groups.

The results presented in Figure 2 demonstrate that the clustering of user behavior flows with different levels of result diversity is highly noticeable, with well-defined boundaries for the distribution clusters of user intent. This indicates that the proposed model has successfully captured the underlying trends of user intentions and adjusted the ranking results accordingly.

Notably, the entropy values of the ranking results increase as the user's intentions become more diverse, reflecting a higher degree of diversity in the ranking outcomes. Conversely, as the user's intentions become more precise, the entropy values decrease, indicating a higher level of accuracy in the ranking results.

5 CONCLUSIONS

This paper proposes a preference-oriented diversity model based on mutual-information, which harmonizes accuracy and diversity in re-ranking. This is the first work that introduces mutual information to balance diversity and accuracy for E-commerce re-ranking. It consists of two components including PON and SAM. PON utilizes Multidimensional Gaussian distributions based on variational inference to model users' diversity preferences with uncertainty. SAM establishes the high correlation between the candidate items and user preferences by maximizing their mutual information. Then utility matrix allows the ranking results to be adaptively adjusted according to user preferences. Extensive experiments conducted on a real-world dataset and the successful deployment on an e-commerce search platform show that the proposed framework can significantly surpass other state-of-the-art models.

REFERENCES

[1] Qingyao Ai, Keping Bi, Jiafeng Guo, and W Bruce Croft. 2018. Learning a deep listwise context model for ranking refinement. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 135–144.

[2] Irwan Bello, Sayali Kulkarni, Sagar Jain, Craig Boutilier, Ed Chi, Elad Eban, Xiyang Luo, Alan Mackey, and Ofer Meshi. 2018. Seq2slate: Re-ranking and slate optimization with rnns. *arXiv preprint arXiv:1810.02019* (2018).

[3] Jaime Carbonell and Jade Goldstein. 1998. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. 335–336.

[4] Laming Chen, Guoxin Zhang, and Eric Zhou. 2018. Fast greedy map inference for determinantal point process to improve recommendation diversity. *Advances in Neural Information Processing Systems* 31 (2018).

[5] Qiwei Chen, Changhua Pei, Shanshan Lv, Chao Li, Junfeng Ge, and Wenwu Ou. 2021. End-to-End User Behavior Retrieval in Click-Through Rate Prediction Model. *arXiv preprint arXiv:2108.04468* (2021).

[6] Qiwei Chen, Yue Xu, Changhua Pei, Shanshan Lv, Tao Zhuang, and Junfeng Ge. 2022. Efficient Long Sequential User Data Modeling for Click-Through Rate Prediction. *arXiv preprint arXiv:2209.12212* (2022).

[7] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*. 7–10.

[8] Yu Gong, Ziwen Jiang, Yufei Feng, Binbin Hu, Kaiqi Zhao, Qingwen Liu, and Wenwu Ou. 2020. EdgeRec: recommender system on edge in Mobile Taobao. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2477–2484.

[9] Zhe Gong, Wei-Lun Chao, Fei Sha, and Chengxiang Zhang. 2014. Practical diversified recommendations on youtube with determinantal point processes. In *Proceedings of the 8th ACM Conference on Recommender systems*. ACM, 281–284.

[10] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: a factorization-machine based neural network for CTR prediction. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. 1725–1731.

[11] Yanhua Huang, Weikun Wang, Lei Zhang, and Ruiwen Xu. 2021. Sliding spectrum decomposition for diversified recommendation. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 3041–3049.

[12] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. 2018. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 1754–1763.

[13] Zihan Lin, Hui Wang, Jingshu Mao, Wayne Xin Zhao, Cheng Wang, Peng Jiang, and Ji-Rong Wen. 2022. Feature-aware diversified re-ranking with disentangled representations for relevant recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 3327–3335.

[14] Yue Lv, Yongfeng Zhang, Jiaxin Mao, Qingyao Ai, and Weidong Liu. 2020. Validation Set Evaluation can be Wrong: An Evaluator-Generator Approach for Maximizing Online Performance of Ranking in E-commerce. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Xi'an, China, July 25-30, 2020*. ACM, 1171–1180. https://doi.org/10.1145/3397271.3401093

[15] Changhua Pei, Yi Zhang, Yongfeng Zhang, Fei Sun, Xiao Lin, Hanxiao Sun, Jian Wu, Peng Jiang, Junfeng Ge, Wenwu Ou, et al. 2019. Personalized re-ranking for recommendation. In *Proceedings of the 13th ACM conference on recommender systems*. 3–11.

[16] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. 1986. Learning representations by back-propagating errors. *nature* 323, 6088 (1986), 533–536.

[17] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 11 (2008).

[18] Guorui Zhou, Na Mou, Ying Fan, Qi Pi, Weijie Bian, Chang Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Deep interest evolution network for click-through rate prediction. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 5941–5948.

[19] Tao Zhuang, Wenwu Ou, and Zhirong Wang. 2018. Globally optimized mutual influence aware ranking in e-commerce search. *arXiv preprint arXiv:1805.08524* (2018).

A COMPANY PORTRAIT

JD.com, Inc., also known as Jingdong, is a Chinese e-commerce company headquartered in Beijing. It is one of the two massive B2C online retailers in China by transaction volume and revenue, a member of the Fortune Global 500. When classified as a tech company, it is the largest in China by revenue and 7th in the world in 2021.

B PRESENTER PROFILES

Huimu Wang is a researcher in the Department of Search and Recommendation at JD.com Beijing. He received his doctor degree in Institute of Automation, Chinese Academy of Sciences. His research focuses on reinforcement learning and natural language processing.

Mingming Li is a researcher in the Department of Search and Recommendation at JD.com Beijing. He received his doctor degree in Institute of Information Engineering, Chinese Academy of Sciences. His research focuses on information retrieval.