Machine Intelligence Research
www.mi-research.net

21(1), February 2024, 169-183
DOI: 10.1007/s11633-023-1422-z

A Simple yet Effective Framework for Active Learning to Rank

Qingzhong Wang Haifang Li Haoyi Xiong Wen Wang Jiang Bian
Yu Lu Shuaiqiang Wang Zhicong Cheng Dejing Dou Dawei Yin

Baidu Incorporated, Beijing 100085, China

Abstract: While China has become the largest online market in the world with approximately 1 billion internet users, Baidu runs the world′s largest Chinese search engine serving more than hundreds of millions of daily active users and responding to billions of queries per day. To handle the diverse query requests from users at the web-scale, Baidu has made tremendous efforts in understanding users′ queries, retrieving relevant content from a pool of trillions of webpages, and ranking the most relevant webpages on the top of the results. Among the components used in Baidu search, learning to rank (LTR) plays a critical role and we need to timely label an extremely large number of queries together with relevant webpages to train and update the online LTR models. To reduce the costs and time consumption of query/webpage labelling, we study the problem of active learning to rank (active LTR) that selects unlabeled queries for annotation and training in this work. Specifically, we first investigate the criterion – Ranking entropy (RE) characterizing the entropy of relevant webpages under a query produced by a sequence of online LTR models updated by different checkpoints, using a query-by-committee (QBC) method. Then, we explore a new criterion namely prediction variances (PV) that measures the variance of prediction results for all relevant webpages under a query. Our empirical studies find that RE may favor low-frequency queries from the pool for labelling while PV prioritizes high-frequency queries more. Finally, we combine these two complementary criteria as the sample selection strategies for active learning. Extensive experiments with comparisons to baseline algorithms show that the proposed approach could train LTR models to achieve higher discounted cumulative gain (i.e., the relative improvement $\Delta DCG_4 = 1.38\%$) with the same budgeted labelling efforts.

Keywords: Search, information retrieval, learning to rank, active learning, query by committee.

Citation: Q. Wang, H. Li, H. Xiong, W. Wang, J. Bian, Y. Lu, S. Wang, Z. Cheng, D. Dou, D. Yin. A simple yet effective framework for active learning to rank. *Machine Intelligence Research*, vol.21, no.1, pp.169–183, 2024. http://doi.org/10.1007/s11633-023-1422-z

1 Introduction

Baidu has established herself as the world′s largest Chinese search engine, serving hundreds of millions of daily active users and handling billions of queries per day. To date, Baidu has archived trillions of webpages for search. In addition to webpages and data resources, Baidu has invented a number of the most advanced search technologies, ranging from language models for content understanding[1, 2], domain-specific recommendation[3–6], online query–Ads matching for sponsored search[7, 8], and software/hardware co-designed infrastructures[9–11] for handling web-scale traffics of online search.

Generally, ranking the retrieved contents plays a critical role in a search engine, where learning to rank (LTR) is a standard workhorse. To achieve better ranking performance, we need to use a large amount of annotated data to train an LTR model. However, it is extremely expensive and time-consuming to label the ranks of relevant webpages for every query[12]. To address this issue, active learning[13, 14] to select a small number of most informative queries and relevant webpages for labelling is requested.

In this paper, inspired by uncertainty-based active learning methods, we present a simple yet effective approach to active learning for ranking. First, we investigate ranking entropy (RE), which characterizes the uncertainty of the ranking for every relevant webpage under a query using a query-by-committee (QBC) method[15]. Intuitively, RE can discover queries with ranking uncertainty, i.e., the predicted ranks of webpages in a query are indistinguishable using the LTR model. However, RE is also biased in favor of the low-frequency queries, i.e., the queries are less searched by users, as there are no sufficient supervisory signals (e.g., click-throughs) to train LTR models for fine predictions. The bias to the low-frequency queries would not bring sufficient information gain to LTR training. To alleviate this problem, we study

Research Article
Manuscript received on November 15, 2022; accepted on February 3, 2023
Recommended by Associate Editor Deng-Ping Fan
Colored figures are available in the online version at https://link.springer.com/journal/11633

yet another criterion – prediction variance (PV), which refers to the variances of rank prediction results among all relevant webpages for a query. Intuitively, we assume a query pairing to multiple webpages that have clearly distinguished orders of ranking as a query with high diversity. We further assume that the variance of rank prediction results would faithfully characterize the variance of ground-truth rank labels, i.e., the diversity of webpages in a query. We thus propose to use PV as a surrogate for the diversity of webpages in a query. Please refer to Section 3 for detailed comparisons and empirical analysis with real data.

More specifically, we report our practices in using the above two criteria to design query selection strategies for active learning. We conducted comprehensive empirical studies on these two criteria using realistic search data. The empirical studies show that the use of RE results in bias to the low-frequency queries, while the use of PV leads to the potential overfittings to the high-frequency queries. When incorporating low-frequency queries in labelling, the active learner might not be able to train LTR models well, due to the lack of supervisory signals (e.g., click-throughs) to distinguish the webpages for the queries. In contrast, when using high-frequency queries (hot queries) in labeling, the active learner might not be able to adapt the out-of-distribution queries (which is critical for ranking webpages at web-scale). Please see also in Sections 3.2 and 3.3 for the details of the criterion and empirical observations.

Finally, extensive experiments with comparisons to baseline algorithms show that the proposed approach (i.e., the combination of RE and PV) can train LTR models to achieve higher accuracy with fewer webpages labelled. Specifically, we have made the following contributions.

1) We study the problem of active learning for ranking in the Baidu search engine, where we focus on selecting queries together with relevant webpages for annotations to facilitate LTR model training and updates. We deploy the system in the Baidu search engine.

2) In the context of the Baidu search, we first consider commonly-used uncertainty metrics for active learning of LTR, namely ranking entropy (RE). We find that the use of RE could be biased by the frequency of queries, i.e., low-frequency queries normally have higher RE scores, as LTR models usually have not been well trained to rank webpages in such queries due to the lack of supervisory signals. To debias RE, we propose to study yet another diversity-based criterion – prediction variance (PV) that may favor high-frequency queries and are highly correlated to the true label variance of webpages under the query. In this way, we combine the two criteria for additional performance improvements.

3) We conduct extensive experiments, showing that our proposed approach is able to significantly improve the performance of LTR in the context of Baidu search. Specifically, we compare our proposals (the combination of RE and PV) with a wide range of sample selection criteria for active learning, including random selection, and expected loss prediction (aka ELO-DCG)[16]. The comparisons show that our proposals outperform other criteria, which discovers 43% more validated training pairs and improves DCG (e.g., $\Delta DCG_4 = 0.35\%–1.38\%$ in offline experiments and $\Delta DCG_4 = 0.05\%–0.35\%$ in online experiments) using the same budgeted labelling efforts under fair comparisons.

Note that in this work, we focus on the low-complexity criteria of sample selection in active learning for LTR. There also exists some sample set selection algorithms[17, 18] for active learning in the high-order polynomial or even combinational complexity over the number of unlabelled samples, which is out of the scope of this paper as we intend to scale-up active learning of LTR with large-scale unlabelled queries and webpages.

2 Related works

The goal of active learning (AL) is to select the most informative samples in the unlabelled data pool for annotation to train a model[19]. Generally, AL models are able to achieve similar performance but use fewer annotated data points. To select the most informative samples for labelling, two categories of methods, i.e., diversity-aware criteria and uncertainty-aware criteria, for sample selection have been studied. The diversity-aware methods[20, 21] measure the diversity of every subset of unlabelled samples and select the sample set with top diversity for labelling, where the core-set selection[22] leveraging the core-set distance of intermediate features is a representative method here. While diversity-aware methods work well on small datasets, they might fail to scale up over large datasets due to the need for subset comparisons and selections.

The uncertainty-aware methods[23–30] screen the pool of unlabelled samples and select samples with top uncertainty in the context of the training model (e.g., LTR models here) for labelling. While uncertainty-aware methods can easily be scaled up over large datasets due to their low complexity, a wide variety of uncertainty criteria have been proposed, such as Monte Carlo estimation of expected error reduction[31], distance to the decision boundary[32, 33], the margin between posterior probabilities[34], and entropy of posterior probabilities[35–37]. Cai et al.[38] proposed an active learning method based on the maximum model change (MMC) and applied it to the classification task. Similarly, expected model change maximization (EMCM)[39] employs the expected model change maximization to estimate the uncertainty, and EMCM is applied to the regression task. Settles et al.[40] introduced active learning into multiple-instance learning and applied active learning to multiple-instance logistic regression. The selection criterion also depends on uncertainty,

i.e., the model calculates multiple-instance uncertainty using the derivative of bag output with respect to instance output. Both MMC and EMCM use gradients to estimate the model change, however, computing gradients is normally more difficult for an LTR model than using QBC to estimate uncertainty. In addition, MMC and EMCM are designed for classification and regression tasks, while we focus on a more challenging task − webpage ranking.

Discussion. The most relevant works to this study are [16, 41–43]. As early as 2010, Long et al.[16, 41] proposed the expected loss optimization (ELO) framework, which selects and labels the most informative unlabelled samples for LTR and incorporates a predictor for discounted cumulative gain (ELO-DCG) to estimate the expected loss of given queries and documents. The work[42] further confirmed that ELO with DCG could work well with any ranker at scale and deliver robust performance. Cai et al.[43] followed the settings of ELO and extended DCG by incorporating the kernel density of queries, so as to balance the sample distribution and the model-agnostic uncertainty for sample selection. Compared to the above studies, this work revisits the problem of active learning for LTR at the web-scale in the 2020s, and we study new metrics of uncertainty for query selection with online LTR performances reported and data analyzed in the context of Baidu search.

3 Practical active learning to rank for web search: Sample selection criteria and empirical studies

In this section, we first review the system design of active learning to rank (active LTR) for web search and then present our proposed selection criteria for active learning with empirical observations.

3.1 Active learning to rank for web search at Baidu

As shown in Fig. 1, given a search query, denoted as q, from a user, the search engine frequently first retrieves all relevant webpages, denoted as $\{w_1, w_2, \cdots\}$, from the dataset and sorts the top-K relevant webpages for the best user reading experience through ranking. To rank every webpage under the query, the search engine pairs every webpage with the query to form a query-webpage pair, e.g., (q, w), and then extracts features from (q, w), denoted as the feature vector (x_q, x_w), where x_q denotes query-relevant features and x_w denotes webpage-relevant features and adopts the LTR model to predict the ranking score, e.g., {bad, fair, good, excellent, perfect}[1] at Baidu Search using (x_q, x_w).

[1] For human annotations, labels 0, 1, 2, 3, 4 and 5 denote bad, fair, good, excellent and perfect, respectively.

To train the LTR model, the search engine usually collects the historical search queries $\mathcal{Q} = \{q_1, q_2, \cdots\}$ and archives relevant webpages $\mathcal{W} = \{w_{1,1}, w_{1,2}, \cdots\}$, where $w_{i,j}$ denotes the j-th webpage associated with q_i. To scale up the active LTR on trillions of webpages/queries while ensuring the timeliness of a search engine, our active learning system (red path in Fig. 1) periodically picks up NEW queries appeared within the last ONE month i.e., $\mathrm{S} \subset \mathcal{Q}$, pairs every selected query in S with retrieved webpages and extracts feature vectors to form the unlabelled datasets denoted as $\mathcal{T} = \{(x_{q_1}, x_{w_{1,1}}), (x_{q_1}, x_{w_{1,2}}), \cdots\}$. Finally, the search engine recruits annotators to label $\mathcal{T}$ and retrains the LTR model with annotated data.

3.2 Sample selection criteria for active LTR

In this section, we present the two criteria proposed for active learning to rank webpages.

3.2.1 Ranking entropy

Uncertainty is one of the most popular criteria in active learning and QBC[15] approach has been widely applied to estimate the uncertainty scores of the unlabelled data. In this paper, we apply QBC to compute the RE of each webpage. Normally, there are M models $\{h_m(x_q, x_w); m = 1, \cdots, M\}$ to constitute a committee. Given the representation of a query-webpage pair $(x_{q_i}, x_{w_{i,j}}) \in \mathcal{T}$, the committee would provide a set of scores $\mathcal{S}_{i,j} = \{h_m(x_{q_i}, x_{w_{i,j}}); m = 1, \cdots, M\}$. Then, for any two webpages $\{w_{i,u}, w_{i,v}\}$ associated with q_i, we can easily calculate the probability that webpage $w_{i,u}$ is ranked higher than $w_{i,v}$ under query q_i, denoted as the probability of $w_{i,u} \succ w_{i,v}$, i.e.,

$$\pi_i^m(w_{i,u} \succ w_{i,v}) = \frac{1}{1 + \exp\left(\frac{-h_m(x_{q_i}, x_{w_{i,u}}) + h_m(x_{q_i}, x_{w_{i,v}})}{T}\right)} \quad (1)$$

where T denotes the temperature and $w_{i,u} \succ w_{i,v}$ denotes that $w_{i,u}$ is more relevant than $w_{i,v}$ under query q_i.

Similar to SoftRank[44], we can obtain the distribution over ranks based on the probabilities calculated by (1). We define the initial rank distribution of the v-th webpage $w_{i,v}$ as $p_{i,v}^{1,m}(r) = \delta(r)$, where $\delta(r) = 1$ if $r = 0$ and 0 otherwise. Then, we can calculate the distribution in the k-th step as follows:

$$p_{i,v}^{k,m}(r) = p_{i,v}^{k-1,m}(r-1)\pi_i^m(w_{i,u} \succ w_{i,v}) + p_{i,v}^{k-1,m}(r)\left(1 - \pi_i^m(w_{i,u} \succ w_{i,v})\right) \quad (2)$$

and the distribution in the last step will be the final ranking distribution. The computing procedure is shown in Algorithm 1, where one webpage is added to the list

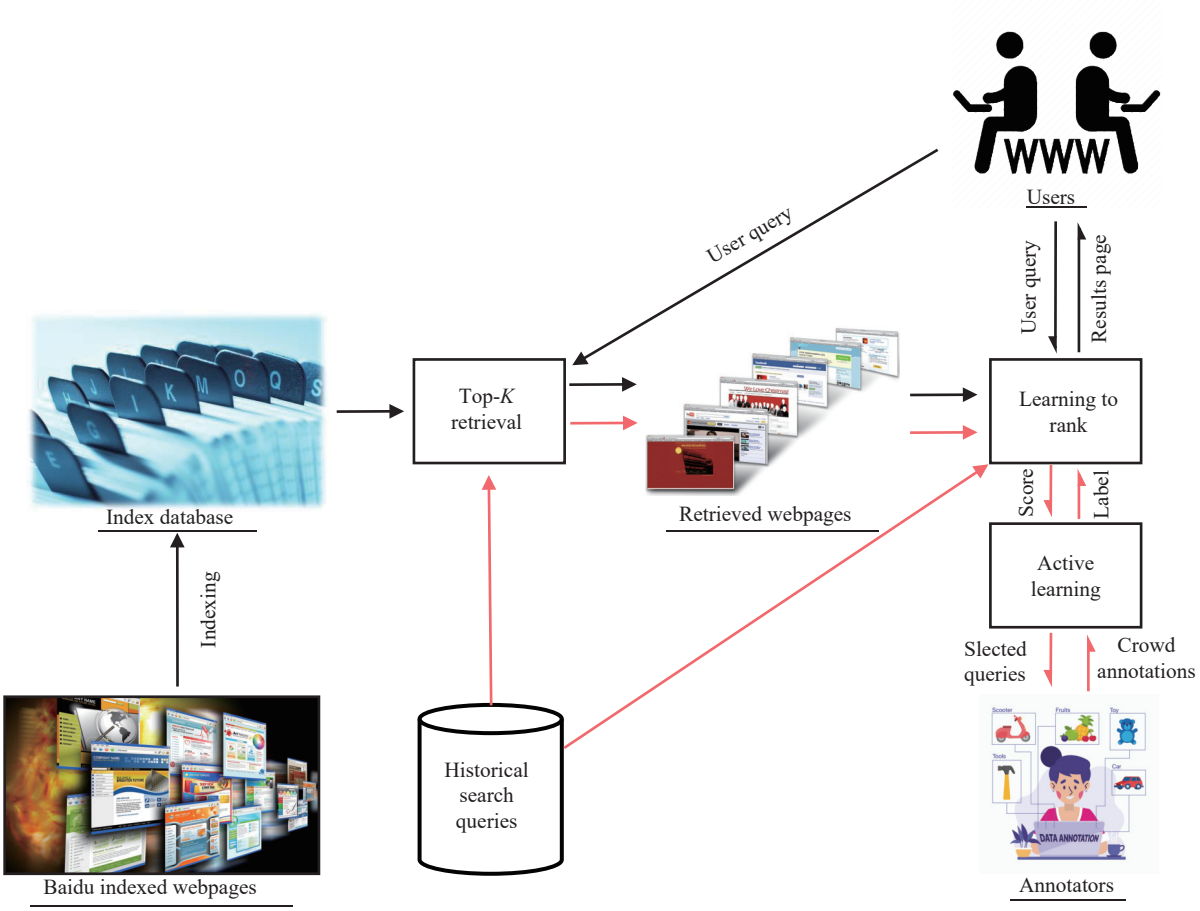


Fig. 1 An overview of the Baidu search system with the proposed active learning process. While the search engine records every search query from users and stores them in historical search queries, it periodically picks up the NEW queries that appeared within the last one month for annotation and retrains LTR models with annotated data.

for comparison in each step and the ranking distribution is updated using (1).

For each webpage $w_{i,j}$, we can obtain a set of ranking distributions $\{p_{i,j}^m(r); m = 1, \cdots, M; r = 1, \cdots, N_i\}$ using Algorithm 1, where N_i denotes the number of webpages associated with q_i. Finally, for every query-webpage pair with the feature vector $(x_{q_i}, x_{w_{i,j}}) \in \mathcal{T}$, we use the average distribution over the committee to compute its entropy score as follow:

$$p_{i,j}(r) = \frac{1}{M} \sum_{m=1}^{M} p_{i,j}^m(r) \tag{3}$$

$$E_{i,j} = -\sum_{r=1}^{N_i} p_{i,j}(r) \log_2 p_{i,j}(r). \tag{4}$$

Algorithm 1. Ranking distribution

Require: $\{\pi_i^m(w_{i,u} \succ w_{i,v}); u, v = 1, \cdots, N_i; u \neq v\}$

Ensure: $\{p_{i,v}^m; v = 1, \cdots, N_i\}$

1) **for** v in range(N_i) **do**
2) $p_{i,v}^m = [1, 0, \cdots, 0]$ /* N_i elements */
3) $p_{tmp} = [0, \cdots, 0]$ /* N_i elements */
4) $\pi_v = [\pi_i^m(w_{i,v} \succ w_{i,1}), \cdots, \pi_i^m(w_{i,v} \succ w_{i,N_i})]$
5) **for** u in range(1, N_i) **do**
6) **for** r in range($v + 1$) **do**
7) **if** r==0 **then**
8) $\alpha = 0$
9) **else**
10) $\alpha = p_{i,v}^m[r-1]$
11) **end if**
12) $p_{tmp}[r] = p_{i,v}^m[r] \times \pi_v[u-1] + \alpha \times (1 - \pi_v \times [u-1])$
13) **end for** $p_{i,v}^m = p_{tmp}$
14) **end for**
15) **end for**

Note that the goal of this paper is to select queries, hence, for a query q_i, we employ the average entropy of the webpages $\{w_{i,j}; j = 1, \cdots, N_i\}$ that are associated with q_i, i.e.,

$$RE(q_i) = \frac{1}{N_i} \sum_{j=1}^{N_i} E_{i,j}. \tag{5}$$

Higher $RE(q_i)$ refers to larger uncertainty in ranking results across the LTR models in the committee. Active

learners are expected to select queries with large ranking entropy, i.e., higher $RE(q_i)$ for $q_i \in \mathcal{Q}$, for annotation and training.

3.2.2 Prediction variance

In our work, we assume a query, pairing to multiple webpages that have clearly distinguished orders of ranking, as a query with high diversity. While the true orders of ranking could be obtained through human annotations (i.e., labelling every webpage under the query using scores of five levels 0, 1, 2, 3 and 4 in this work), we propose to use the rank prediction results of a trained LTR model to measure such diversity. Given a checkpoint of the online GBRank[45] model, we propose adopting the variance of predicted ranking scores (namely prediction variance, PV) to measure the diversity of webpages for the query.

Similar to RE, we also use the predictions of the committee to compute the prediction variance. Given the outputs of the committee $\mathcal{S}_{i,j} = \{h_m(x_{q_i}, x_{w_{i,j}});\ m = 1, \cdots, M\}$, the prediction variance of a query q_i with N_i retrieved webpages can be computed using the following equations:

$$\mu_m(q_i) = \frac{1}{N_i} \sum_{j=1}^{N_i} h_m\left(x_{q_i}, x_{w_{i,j}}\right) \tag{6}$$

$$STD_m(q_i) = \sqrt{\frac{1}{N_i} \sum_{j \in 1}^{N_i} \left(h_m\left(x_{q_i}, x_{w_{i,j}}\right) - \mu_m(q_i)\right)^2}. \tag{7}$$

Finally, we calculate the prediction variance $PV(q_i)$ as follows:

$$PV(q_i) = \frac{1}{M} \sum_{m=1}^{M} STD_m(q_i). \tag{8}$$

For active learning, we assume queries with large prediction variances, i.e., higher $PV(q_i)$ for $q_i \in \mathcal{Q}$, as the candidates for annotation and training.

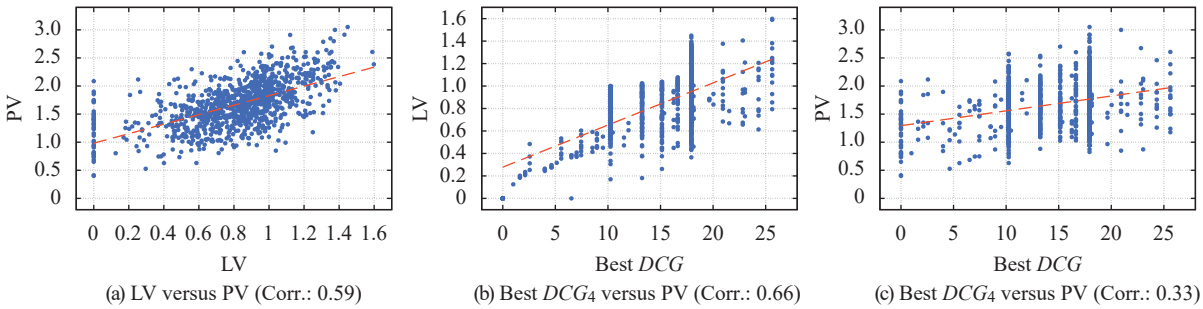
3.3 Empirical studies on proposed criteria for active LTR

While the first criterion RE directly measures the uncertainty of ranking results under a query (either due to the defects of learned models or simply the difficulty to rank), we now hope to validate whether the second proposed criterion PV can characterize the diversity of webpages in a query and how PV improves LTR. We conduct empirical studies based on 1 000 realistic query data points drawn from the validation set and hope to test two hypotheses as follows.

Does PV characterize the variance of human-annotated ground truth ranking scores for LTR? Here to test our hypothesis, we fetch a past checkpoint of the online LTR model in Baidu search and use the model to predict the ranking scores for every webpage in the 1 000 queries. Note that, such 1 000 queries are obtained from the validation set and have no overlap with the queries used for training the LTR model. In Fig. 2(a), we plot the scatter points of LV (label variance) versus PV with Pearson correlation $Corr = 0.59$ and p-value < 0.05. In this way, we can conclude that PV significantly correlates to the variance of human-annotated ground truth ranking scores (LV) and faithfully characterizes the difficulty of ranking every query.

Does PV correlate with the information gain of query selection for LTR? Another hypothesis in our mind is that selecting queries with a high diversity of webpages for annotations could bring more information gain to LTR. We thus need to correlate LV and PV with certain information measures of queries. In this study, we use a measure namely Best DCG_4 which refers to the estimate of the upper bound of DCG_4 that uses the human-annotated ground truth ranking scores as the prediction results of the ranking. Intuitively, the best DCG_4 reflects the optimal DCG_4 that can be achieved by any algorithm. The correlation studies have been performed and illustrated in Figs. 2(b) and 2(c). The significance could be found in the correlations between LV and Best DCG_4 and the correlations between PV and Best DCG_4. The observations suggest that queries with higher diversity in webpages are usually more informative for LTR, regardless of whether the diversity was measured by LV or PV.

Based on the above two observations, we could conclude that 1) PV could faithfully characterize LV (label



(a) LV versus PV (Corr.: 0.59) (b) Best DCG_4 versus PV (Corr.: 0.66) (c) Best DCG_4 versus PV (Corr.: 0.33)

Fig. 2 Correlation studies and empirical observations on criteria based on 1 000 queries ($N = 1\,000$). Best DCG_4 refers to the estimate of the upper bound of DCG_4, where we use the human-annotated ground-truth labels to compute the DCG_4 score for every query. Corr. denotes Pearson correlation coefficient.

variance − the diversity of webpages), although PV was estimated using the prediction results of a model, and 2) a query with higher LV or PV is usually more informative for LTR, since LTR models are normally trained using pairwise or listwise loss and a small LV leads to small loss values and gradients.

3.4 Combined criteria

To be simple, we use the weighted sum of RE and PV as the acquisition function in active learning. For each query $q_i \in \mathcal{Q}$, the acquisition function is

$$f(q_i) = RE(q_i) + \alpha \times PV(q_i) \tag{9}$$

where α is a hyperparameter to balance ranking uncertainty and webpage variance. We select the queries that have the largest values of $f(q_i)$ in each cycle of active learning.

4 Experiments

In this section, we present the results of experiments, where we first introduce the results of offline experiments and then figure out the online performance of our proposals, both in comparison with baseline algorithms.

4.1 Offline experiments and results

In this section, we present the details of offline experiments with introductions to the setups and results.

4.1.1 Setups

To conduct offline experiments, we construct a dataset for LTR. We classify the queries in the last month into 10 categories based on the frequency, and then we filter out the erotic and illegal queries in each category. Finally, we randomly sample 1 500 queries from each category and for each query, we select 60 retrieved documents for human annotation, resulting in a dataset composed of 15 000 queries and 900 000 documents. Note that the dataset with 15 000 queries is relatively large and we present the comparison between our dataset and existing LTR datasets in Table 1. In the dataset, the label of each query-document pair is in 5 levels: bad, fair, good, excellent and perfect and the corresponding relevant scores are {0, 1, 2, 3, 4, 5}, respectively.

To train the model, first, we split the dataset into a training set (14 000 queries) and a validation set (1 000 queries). In the beginning of active learning, we randomly select N_0 queries from the training set as the base and in each cycle of active learning, we set the batch size $bs = 100$, i.e., we select 100 queries from the pool (the rest of the training set) using the acquisition function[2]. The quota is 2 000 queries, i.e., we run active learning for 20 cycles. We also conduct ablation studies on the value of α in (9), where $\alpha = \{0.5, 1.0, 1.5\}$. We set the number of committees M to 9, i.e., nine variants of GBRank with different numbers of trees (100, 300, 500) and maximum depth (1, 3, 5).

Table 1 Comparison between our dataset and existing datasets for learning to rank

Dataset	Train		Validation & Test	
	# Queries	# Documents	# Queries	# Documents
Yahoo set1[46]	19 944	473 134	9 976	236 743
Yahoo set2[46]	1 266	34 815	5 064	138 005
Microsoft[47]	18 900	2 261 000	12 600	1 509 000
Tiangong[48]	3 449	333 813	100	10 000
Ours	14 000	840 000	1 000	60 000

To evaluate the performance of an LTR model, we use discounted cumulative gain (DCG), calculated as follows:

$$DCG_K = \sum_{k=1}^{K} \frac{G_k}{\log_2(k+1)} \tag{10}$$

where G_k denotes the weight assigned to the webpage's label at position k. A higher G_k indicates that the webpage is more relevant to the query. Additionally, a higher DCG_K indicates a better LTR model. In this paper, we consider the DCG of the top 4 ranking results, i.e., DCG_4. In addition, we consider another important metric – the percentage of the irrelevant webpages in top K, which is computed as follows:

$$R_{01} = \frac{N_{01}}{K} \tag{11}$$

where N_{01} denotes the number of the irrelevant webpages[3]. Obviously, a lower R_{01} indicates a better LTR model. Additionally, we consider R_{01} in the top 4 in this paper.

In addition to DCG and R_{01}, we also compare the distribution of the selected queries and the number of valid training pairs obtained by using different methods, which is able to reflect the label diversity of webpages. A large label diversity means that webpages are uniformly distributed on each label and a small diversity indicates that most webpages have the same label.

There are two baselines for comparison, the first one is random selection and the second one is ELO-DCG[16] – an uncertainty-based active learning method for ranking.

4.1.2 Offline results

Here, we first present the statistical characteristics of the selected queries (with webpages retrieved) for annota-

[2] The reason for using $bs = 100$ is that annotating the relevant scores are expensive and time-consuming. We can only annotate 500 queries per day and in the following offline experiments, we also consider $bs = 500$.

[3] We consider the webpages with labels of 0 and 1 as irrelevant webpages.

tions and then introduce the details about the valid query-webpage pairs formed from annotation results for training. Finally, we present the performance improvements of the proposed criterion in comparison to baseline criteria, such as ELO-DCG[16, 41].

Distribution of selected queries

Fig. 3 shows the distribution over categories of 1 000 selected queries. Category 0 is composed of the most frequent queries, while category 9 contains the least frequent (only one time in the one-month search log) queries. For random selection, we randomly select 100 queries from each category. Compared with random selection, LV prefers relatively frequent queries, such as Categories 2–4, but selects fewer low-frequency queries in Categories 7–9, indicating that the webpages associated with low-frequency queries have similar human annotations. PV performs similarly in high-frequency queries but selects much fewer low-frequency queries. In contrast, RE selects the most low-frequency queries. However, it is difficult to construct enough training pairs if there are too many low-frequency queries since irrelevant webpages dominate these queries. In Fig. 4(c), we demonstrate the distribution of labels, where we use 1 000 randomly selected queries to obtain the statistics. Obviously, label 0 dominates low-frequency queries leading to difficulties in constructing training pairs for GBRank, hence, we need to balance the number of selected queries in each category and RE+PV is able to achieve the goal (see Fig. 4(f)). Interestingly, the existing work – ELO-DCG[16, 41] favors high-frequency queries. The reason is that for low-frequency queries, the best DCG and the DCG based on the average relevant scores are very small, leading to a relatively small ELO-DCG. Generally, the Baidu search engine can well handle high-frequency queries to satisfy users′ demands and selecting more high-frequency queries cannot benefit the gain of Baidu search.

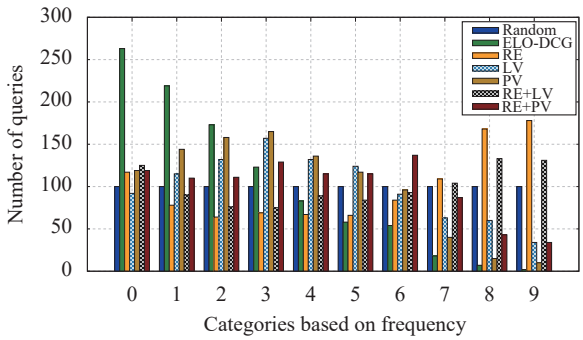


Fig. 3 Distribution of 1 000 selected queries using different criteria. LV stands for label variance, PV for prediction variance and RE for ranking entropy.

Looking at Fig. 4, where we present the distribution of labels over categories. For random selection, one observation is that low-frequency queries have more webpages with the label 0. In addition, the distributions of labels for each category are different and unbalanced. Using more webpages with label 0 is able to provide more training positive-negative pairs, for example, suppose we have N relevant webpages and M irrelevant webpages. Then we can construct $N \times M$ training pairs and a larger M enlarges the number of pairs, which could reduce the percentage of irrelevant webpages in the top K ranking results. However, too many webpages with similar labels could reduce the number of valid training pairs, hurting ranking quality. By contrast with random selection, ELO-DCG[16, 41] selects more webpages with label 2 since it prefers high-frequency queries that have more relevant (Labels 2, 3, 4) webpages. Although ELO-DCG[16, 41] can select more relevant webpages, the diversity among webpages for each query is relatively low. While RE performs in the opposite way, selecting more webpages with label 0, which also lacks diversity among webpages. Intuitively, if the webpages of a query have the same label, then it is difficult to rank them, i.e., higher uncertainty. Moving on to Figs. 4 (d)–4(f) (LV, PV and RE+PV), the distribution of labels is more balanced, indicating that the webpages are diverse. On the one hand, a higher diversity score could result in more training pairs, on the other hand only considering diversity selects more queries that the trained model can easily rank the associated webpages and these queries cannot further improve the ranking model. In contrast, RE+PV is able to balance diversity and uncertainty, selecting more useful and informative queries.

Number of training pairs

For GBRank[45] which uses pairwise loss, the number of training pairs is crucial. With more training pairs fed to the training procedure, LTR models are expected to deliver better performance. Table 2 presents the number of pairs obtained using different approaches. We can easily conclude that the proposed approach is able to obtain more training pairs compared with random selection and the existing work – ELO-DCG[16, 41]. In terms of the number of valid pairs composed of two webpages with different human annotated labels, random selection obtains 764 527 pairs for 1 000 queries, while the number increases by 25% when using ELO-DCG[16, 41]. Using the criteria LV and PV that improve the diversity among webpages, the number of valid pairs can be improved by 36% and 28%, respectively. Only using RE can also achieve an 8% increase compared with random selection, but it is inferior to ELO-DCG, LV and PV since RE selects more low-frequency queries and most webpages associated with these queries are labelled 0 (see Fig. 3 and Fig. 4(c)). Although selecting more low-frequency queries could benefit in solving the problem of unusual queries and attracting more users, it is difficult to retrieve relevant webpages for these queries and irrelevant webpages are less useful to train GBRank. Combining RE and PV is able to alleviate the problem. The number of valid pairs surges by 43% and 50% by using RE+PV and

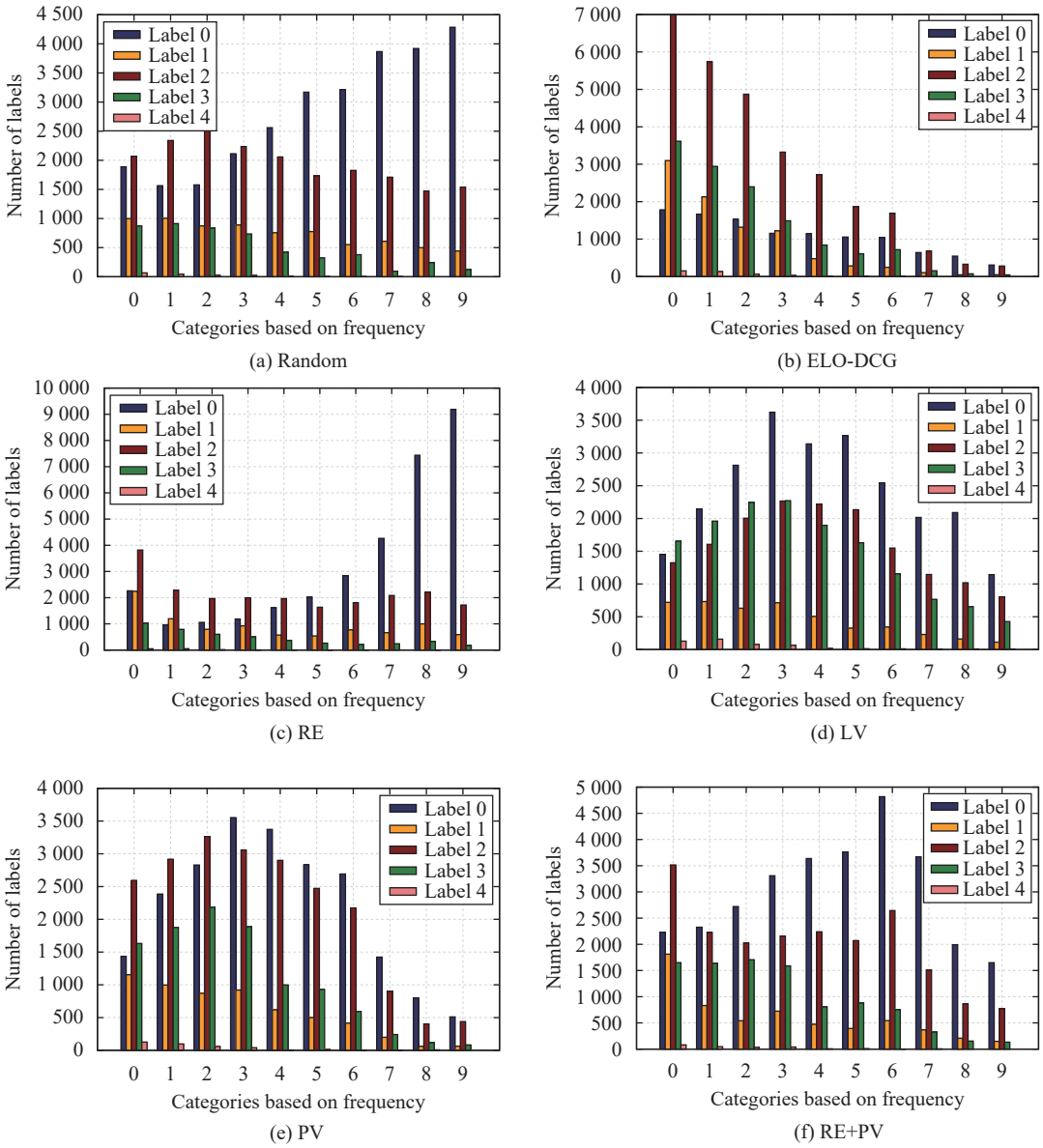


Fig. 4 Distribution of labels over categories. We use 1 000 selected queries and the corresponding webpages to obtain the statistics.

RE+LV, respectively, which is a remarkable improvement. Considering the number of neg-pos pairs, our proposed approach is also superior to random selection and ELO-DCG[16, 41]. Generally, the number of neg-pos pairs is related to the percentage of the irrelevant webpages in top K since using more neg-pos pairs to train an LTR model, it would be easier to distinguish relevant webpages from irrelevant webpages. Using RE+PV, the number of neg-pos pairs drastically increases by 50% compared with random selection, and it is also boosted by 34% compared to the existing active learning approach − ELO-DCG[16, 41].

LTR performance comparisons

In this experiment, we use the valid query-webpage pairs obtained by various strategies to train LTR models (GBRank models with cross-entropy loss) and compare the ranking quality of these LTR models on our validation dataset of 1 000 queries. The ranking quality is measured using DCG. Fig. 5 shows the comparison among different approaches.

Let us first pay attention to base100 − the top two subfigures in Fig. 5. The proposed approach RE+PV achieves better performance than random selection and ELO-DCG[16, 41]. We can see that using RE+PV selected queries to train GBRanK, DCG_4 increases faster than its counterparts, such as random selection, ELO-DCG and RE. Compared with random selection, the relative improvement of DCG_4 ranges from 0.35% to 1.38% using RE+PV. Compared to the existing work ELO-DCG[16, 41], the proposed RE+PV boosts DCG_4 by at least 0.37%. Random selection outperforms ELO-DCG and RE when selecting more training data. The possible reason is that

Table 2 Number of training pairs obtained by using different criteria and the relative improvement compared with random selection. If two webpages associated with a query have different labels, then they constitute a valid pair. Neg-pos pair denotes that a pair is composed of irrelevant and relevant webpages. We use each approach to select 1 000 queries to obtain the statistics.

Criterion	# Valid pairs	# Neg-pos pairs
Random	764 527	534 500
ELO-DCG[16]	959 228 (25%↑)	598 555 (12%↑)
LV	1 039 474 (36%↑)	757 492 (42%↑)
PV	979 051 (28%↑)	704 621 (32%↑)
RE	823 411 (8%↑)	562 464 (5%↑)
Ours (RE+PV)	**1 091 176 (43%↑)**	**803 723 (50%↑)**
RE+LV (Upper bound)	**1 149 083** (50%↑)	**827 133** (55%↑)

ELO-DCG and RE are biased by query frequency, e.g., ELO-DCG prefers high-frequency queries, whereas RE is in favor of low-frequency queries. Interestingly, ELO-DCG and RE perform similarly to each other, although the distributions of the selected queries are different. Looking at PV and LV that are related to the diversity among webpages, both outperform random selection in most cycles. In regard to R_{01} (top-right subfigure in Fig. 5), the proposed RE+PV also achieves competitive performance compared with its counterparts, e.g., R_{01} drops by at most 6.31% compared to random selection. Although ELO-DCG is able to obtain higher DCG_4 scores at the beginning of AL cycles, it performs even worse considering the metric R_{01}. The reason is that ELO-DCG selects more webpages with label 2 but fewer webpages with label 0, hence, it is difficult for GBRank to distinguish irrelevant webpages.

Moving on to base500 (bottom subfigures in Fig. 5), our proposed approach − RE +PV also achieves higher DCG_4 than random selection and ELO-DCG[16, 41]. The relative improvement of DCG_4 is at most 0.84% compared to random selection. Moreover, R_{01} shows decreases in most cycles, e.g., in cycle 9, R_{01} drops by 3.76% using RE+PV.

We also present the relative improvement of the average DCG_4 over AL cycles for each category in Table 3. We can easily find that most AL approaches obtain better performance on high-frequency queries compared with random selection, e.g., for category 0, using RE, RE+0.5PV and RE+LV boost DCG_4 by 1.13% compared with random selection in the base100 scenario. DCG_4 also increases by 0.42% using ELO-DCG[16, 41] for category 0 since ELO-DCG selects more high-frequency queries (see Fig. 3). In contrast, for low-frequency queries,

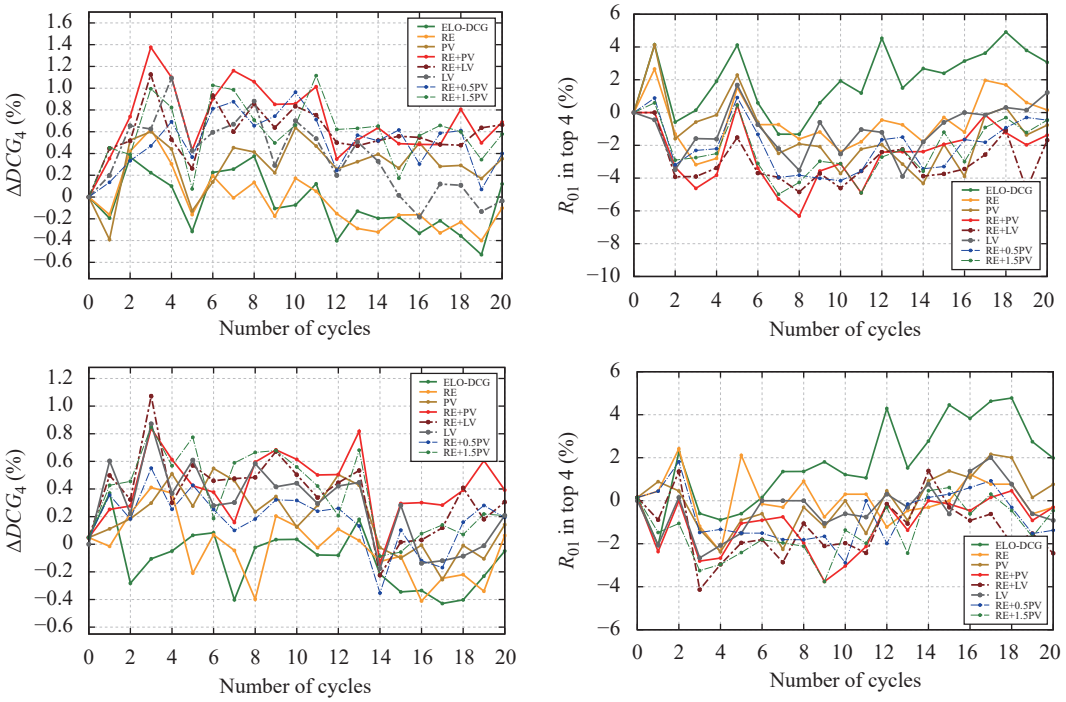


Fig. 5 Relative improvements of DCG_4 (i.e., ΔDCG_4) and R_{01} (i.e., ΔR_{01}) compared with using random selection in each active learning cycle with the same budget. Top: Base set is composed of 100 queries. Bottom: Base set is composed of 500 queries.

Table 3 Relative improvement of average DCG_4 (i.e., ΔDCG_4) over all active learning cycles compared with random selection. Red numbers represent the relative increases, blue numbers are the relative decreases and the bold numbers are the highest relative improvements.

Base	Method	Categories based on frequency										All
		0	1	2	3	4	5	6	7	8	9	
100	ELO-DCG	0.42%↑	0.14%↓	0.30%↓	0.31%↑	0.93%↑	0.18%↓	0.18%↓	0.00%↑	1.44%↓	0.74%↓	0.00%↑
	LV	0.28%↑	0.29%↓	0.22%↑	0.78%↑	0.51%↑	0.72%↑	0.27%↓	0.66%↑	0.24%↓	0.25%↑	0.43%↑
	PV	0.50%↑	0.07%↑	0.15%↑	0.08%↓	0.42%↑	0.27%↑	0.36%↑	0.66%↑	0.12%↑	0.62%↓	0.35%↑
	RE	**1.13%**↑	0.57%↓	0.22%↑	0.54%↓	0.08%↓	0.54%↑	0.00%↑	0.28%↓	1.08%↓	0.00%↑	0.00%↑
	RE+0.5PV	**1.13%**↑	0.29%↑	0.15%↑	0.86%↑	0.76%↑	0.45%↑	0.36%↑	0.85%↑	0.48%↓	0.62%↑	0.52%↑
	RE+PV	0.57%↑	**0.43%**↑	0.00%↑	**0.86%**↑	1.18%↑	0.36%↑	**0.82%**↑	1.14%↑	**0.96%**↑	**1.23%**↑	**0.78%**↑
	RE+1.5PV	0.35%↑	0.22%	0.30%↑	0.70%↑	**1.43%**↑	0.36%↑	0.45%↑	**1.42%**↑	0.60%↑	0.62%↑	0.61%↑
	RE+LV	**1.13%**↑	0.29%↑	**0.45%**↑	0.23%↑	1.35%↑	**1.08%**↑	0.64%↑	0.85%↑	0.60%↓	0.00%↑	0.61%↑
500	ELO-DCG	0.28%↑	0.00%↑	0.52%↑	0.39%↓	0.17%↓	0.36%↓	0.54%↓	0.28%↓	0.48%↓	0.00%↑	0.09%↓
	LV	0.08%↑	0.21%↓	**0.90%**↑	0.08%↓	0.17%↑	0.91%↑	0.18%↓	0.19%↑	0.12%↑	1.12%↑	0.26%↑
	PV	0.14%↓	0.00%↑	0.82%↑	0.77%↓	0.17%↑	0.63%↑	0.00%↑	0.19%↑	0.48%↑	0.99%↑	0.17%↑
	RE	0.28%↑	0.29%↓	0.37%↑	0.69%↓	0.75%↓	0.36%↑	0.27%↑	0.17%↓	0.84%↓	1.61%↑	0.09%↓
	RE+0.5PV	0.35%↑	0.29%↑	0.75%↑	0.46%↓	0.42%↓	0.63%↑	0.09%↑	0.09%↑	0.84%↓	1.24%↑	0.17%↑
	RE+PV	0.07%↑	**0.43%**↑	0.22%↑	**0.23%**↑	0.17%↑	0.27%↑	**0.63%**↑	**0.47%**↑	**1.20%**↑	1.12%↑	**0.43%**↑
	RE+1.5PV	**0.49%**↑	**0.43%**↑	0.52%↑	0.15%↓	**0.25%**↑	**0.72%**↑	0.36%↑	0.28%↑	0.24%↓	0.74%↑	0.34%↑
	RE+LV	0.35%↑	**0.43%**↑	0.67%↑	0.39%↓	0.00%↑	**0.72%**↑	0.36%↑	0.28%↑	0.12%↓	**1.36%**↑	0.34%↑

the performances of different approaches vary, e.g., DCG_4 for category 9 obtained by ELO-DCG drops by 0.74% compared with random selection, whereas DCG_4 obtained by our proposed approach − RE+PV is 1.23% higher than using random selection. Interestingly, the performances of using $bs = 100$ and $bs = 500$ in category 3 are different. We can see that the test models except for PV and RE outperform random selection when using $bs = 100$, while they are inferior to random selection when $bs = 500$. The possible reason is that for the queries belonging to Category 3, it is difficult to annotate the relevance scores[4], resulting in many noisy labels, and selecting more queries introduces more noisy labels; hence, the performance decreases when using $bs = 500$.

Note that we also conduct ablation studies on the hyperparameter α in (9), finding that $\alpha = 1$ is a better choice and in our online experiments we use this value.

4.2 Online experiments and results

To report the online performance of our proposals, we carried out an A/B test with real-world web traffics. Note that online testing is expensive and time-consuming, so we only compare our proposed model with the existing baseline that employs random selection in the Baidu search engine.

4.2.1 A/B test setups

We used 0.6% real-world web traffics on the Baidu search engine to conduct the A/B test, where the 0.6% traffics were randomly partitioned into two folders (0.3%) each to evaluate the performance of our proposals (RE+PV) and random selection respectively. This online A/B test lasted for 13 days/cycle. In each cycle, we use our proposed approach to select 500 queries from the historical query pool composed of hundreds of millions of queries, and each query has 100 retrieved webpages. Then we filter out the pornographic webpages and the webpages forbidden by the government, resulting in approximately 60 webpages for each query. After that, we hire people to annotate the relevance scores and each query-webpage pair has at least 6 scores annotated by different workers. Finally, our expert annotators evaluate the quality of annotations to ensure that the accuracy is higher than 85% and then we use the weighted sum of the scores as the final label to train our LTR model. Similar to the offline experiments, we also calculate ΔDCG_4 and ΔR_{01} between the two methods based on the ground-truth annotation results.

As we have mentioned, we can only label 500 queries per day, hence, in the online experiments we use the $bs = 500$. Additionally, note that our query pool is com-

[4] For the queries belonging to Category 0, it is easy to annotate since many webpages are highly relevant to the queries. Likewise, we can easily annotate queries belonging to Category 9, since the webpages are irrelevant to queries.

posed of hundreds of millions of historical queries, therefore, the selection criteria should be as simple as possible, otherwise, our cluster cannot handle the selection in a few hours. Moreover, since we need to update the LTR model every day, and to avoid some unexpected influences on the search engine, we only use the 0.6% traffic to test the proposed approach.

4.2.2 Online performance

The comparison is shown in Fig. 6, where we only present the relative improvements. Compared to random selection, the proposed RE+PV is able to boost DCG_4 in all cycles and the largest relative improvement is around 0.35% when considering all queries. We also compare RE+PV to random selection on low-frequency queries (Categories 7–9) and DCG_4 increases by at most 0.85%, indicating that the proposed approach benefit low-frequency query search. In terms of R_{01}, our proposed approach – RE+PV can also reduce the percentage of irrelevant webpages in the top K results, e.g., R_{01} decreases by at most 1.9% considering all queries, while it drops by at most 2.6% on low-frequency queries. Basically, the online performance is consistent with our offline results and the proposed active learning approach outperforms the baseline.

5 Conclusions and future work

In this work, we revisited the problem of active learning for ranking webpages in the context of Baidu search, where the key problem is to establish the training datasets for learning to rank (LTR) models. Given trillions of queries and relevant webpages retrieved for every query, the goal of active learning is to select a batch of queries for labelling and train the current LTR model with the newly labelled datasets incrementally, where the labels here refer to the ranking score of every webpage under the query. To achieve the goals, for every query, this work proposed two new criteria − RE and PV that could measure the uncertainty of the current LTR model to rank webpages in a query and the diversity of ranking scores for webpages in a query respectively. Specifically, RE estimates the entropy of relevant webpages under a query produced by a sequence of online LTR models updated by different checkpoints, using a QBC method, while PV estimates the variance of prediction results for all relevant webpages under a query. Our experimental observations find that RE may pop low-frequency queries from the pool for labelling while PV prioritizes high-frequency queries more. Furthermore, the estimate of PV significantly correlates to the diversity of true ranking scores of webpages (annotated by humans) under a query and correlates to the information gain of LTR. Finally, we combine these two complementary criteria as the sample selection strategies for active learning. Extensive experiments with comparisons to baseline algorithms show that the proposed approach could train LTR models, achieving higher DCG (i.e., $\Delta DCG_4 = 0.35\%–1.38\%$ in offline experiments, $\Delta\text{DCG}_4 = 0.05\%–0.35\%$ in online experiment) using the same budgeted labelling efforts, while the proposed strategies could discover 43% more valid training pairs for effective training. Note that the queries selected by active learning are more informative, and we can use fewer queries to train LTR models, achieving satisfactory performance, which saves millions of yuan per year.

Recently, deep learning approaches have been applied to web search and real-world products. For example, in the Baidu search engine, we have used pretrained large models[49]. Kaleido-BERT(bidirectional encoder representations from transformers)[50] and AliCoCo(Alibaba cognitive concept net)[51] employ large models for e-commerce search and both of them introduce specific knowledge into the search engine. One future direction should be cold-started active learning, i.e., using the pretrained models to select samples for annotation. In the field of web search, artificial intelligence-generated content (AIGC) should draw much attention. ChatGPT[52] has shown the strong ability of big models to answer questions. Additionally, some works on text-to-image and image-to-text translation[53–58] should provide content for search engines, therefore, another future direction could be combining both AIGC and current search engines to well satisfy users′ demands.

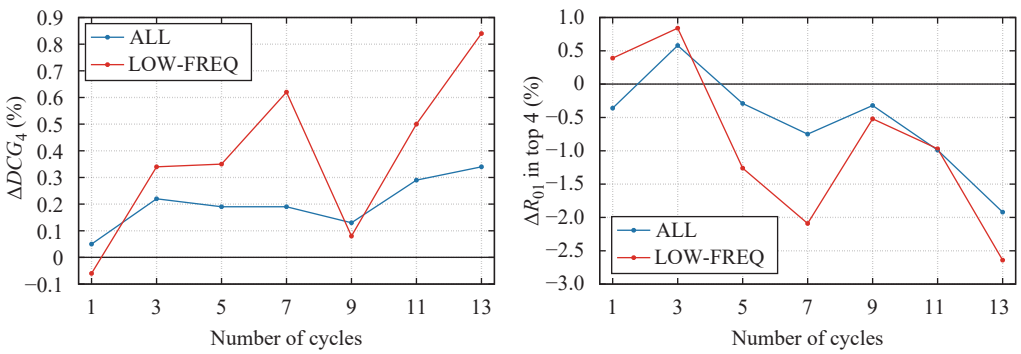


Fig. 6 Online performance. We only report the relative improvement with p-value < 0.05 over the baseline. Left: Performance based on DCG_4. Right: Performance based on R_{01}.

Acknowledgements

This work was supported in part by the National Key R&D Program of China (No. 2021ZD0110303).

Declarations of conflict of interest

The authors declared that they have no conflicts of interest to this work.

References

[1] Y. Sun, S. H. Wang, Y. K. Li, S. K. Feng, X. Y. Chen, H. Zhang, X. Tian, D. X. Zhu, H. Tian, H. Wu. ERNIE: Enhanced representation through knowledge integration, [Online], Available: https://arxiv.org/abs/1904.09223, 2019.

[2] Y. Sun, S. H. Wang, Y. K. Li, S. K. Feng, H. Tian, H. Wu, H. F. Wang. Ernie 2.0: A continual pre-training framework for language understanding. In *Proceedings of AAAI Conference on Artificial Intelligence*, Palo Alto, USA, pp. 8968–8975, 2020. DOI: 10.1609/aaai.v34i05.6428.

[3] J. Z. Huang, S. Q. Ding, H. F. Wang, T. Liu. Learning to recommend related entities with serendipity for web search users. *ACM Transactions on Asian and Low-resource Language Information Processing*, vol. 17, no. 3, Article number 25, 2018. DOI: 10.1145/3185663.

[4] J. Z. Huang, W. Zhang, Y. M. Sun, H. F. Wang, T. Liu. Improving entity recommendation with search log and multi-task learning. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, ACM, Stockholm, Sweden, pp. 4107–4114, 2018. DOI: 10.5555/3304222.3304341.

[5] J. Z. Huang, H. F. Wang, W. Zhang, T. Liu. Multi-task learning for entity recommendation and document ranking in web search. *ACM Transactions on Intelligent Systems and Technology*, vol. 11, no. 5, Article number 54, 2020. DOI: 10.1145/3396501.

[6] M. Fan, Y. B. Sun, J. Z. Huang, H. F. Wang, Y. Li. Meta-learned spatial-temporal POI auto-completion for the search engine at baidu maps. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, Singapore, pp. 2822–2830, 2021. DOI: 10.1145/3447548.3467058.

[7] M. Fan, J. C. Guo, S. Zhu, S. Miao, M. M. Sun, P. Li. MOBIUS: Towards the next generation of query-ad matching in baidu″s sponsored search. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, USA, pp. 2509–2517, 2019. DOI: 10.1145/3292500.3330651.

[8] T. Yu, Y. Yang, Y. Li, X. D. Chen, M. M. Sun, P. Li. Combo-attention network for baidu video advertising. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, USA, pp. 2474–2482, 2020. DOI: 10.1145/3394486.3403297.

[9] J. Ouyang, S. D. Lin, W. Qi, Y. Wang, B. Yu, S. Jiang. SDA: Software-defined accelerator for large-scale DNN systems. In *Proceedings of IEEE Hot Chips 26 Symposium*, Cupertino, USA, 2014. DOI: 10.1109/HOTCHIPS.2014.7478821.

[10] W. J. Zhao, J. Y. Zhang, D. P. Xie, Y. L. Qian, R. L. Jia, P. Li. AIBox: CTR prediction model training on a single node. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, Beijing, China, pp. 319–328, 2019. DOI: 10.1145/3357384.3358045.

[11] J. Ouyang, M. Noh, Y. Wang, W. Qi, Y. Ma, C. H. Gu, S. Kim, K. I. Hong, W. K. Bae, Z. B. Zhao, J. Wang, P. Wu, X. Z. Gong, J. X. Shi, H. F. Zhu, X. L. Du. Baidu kunlun an AI processor for diversified workloads. In *Proceedings of IEEE Hot Chips 32 Symposium*, Palo Alto, USA, 2020. DOI: 10.1109/HCS49909.2020.9220641.

[12] B. Settles. From theories to queries: Active learning in practice. In *Proceedings of the Active Learning and Experimental Design workshop in Conjunction with AISTATS*, Ft. Lauderdale, USA, 2011.

[13] D. Cohn, L. Atlas, R. Ladner. Improving generalization with active learning. *Machine Learning*, vol. 15, no. 2, pp. 201–221, 1994. DOI: 10.1007/BF00993277.

[14] S. Y. Huang, T. Y. Wang, H. Y. Xiong, J. Huan, D. J. Dou. Semi-supervised active learning with temporal output discrepancy. In *Proceedings of IEEE/CVF International Conference on Computer Vision*, IEEE, Montreal, Canada, pp. 3427–3436, 2021. DOI: 10.1109/ICCV48922.2021.00343.

[15] Y. Freund, H. S. Seung, E. Shamir, N. Tishby. Selective sampling using the query by committee algorithm. *Machine Learning*, vol. 28, no. 2, pp. 133–168, 1997. DOI: 10.1023/A:1007330508534.

[16] B. Long, J. Bian, O. Chapelle, Y. Zhang, Y. Inagaki, Y. Chang. Active learning for ranking through expected loss optimization. *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 5, pp. 1180–1191, 2015. DOI: 10.1109/TKDE.2014.2365785.

[17] K. Wei, R. Iyer, J. Bilmes. Submodularity in data subset selection and active learning. In *Proceedings of the 32nd International Conference on Machine Learning*, ACM, Lille, France, pp. 1954–1963, 2015. DOI: 10.5555/3045118.3045326.

[18] S. Hanneke, L. Yang. Minimax analysis of active learning. *Journal of Machine Learning Research*, vol. 16, no. 12, pp. 3487–3602, 2015.

[19] D. A. Cohn, Z. Ghahramani, M. I. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research*, vol. 4, pp. 129–145, 1996. DOI: 10.1613/jair.295.

[20] H. T. Nguyen, A. Smeulders. Active learning using pre-clustering. In *Proceedings of the 21st International Conference on Machine Learning*, ACM, Banff, Canada, pp. 79–86, 2004. DOI: 10.1145/1015330.1015349.

[21] Y. H. Guo. Active instance sampling via matrix partition. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, ACM, Vancouver, Canada, pp. 802–810, 2010. DOI: 10.5555/2997189.2997279.

[22] O. Sener, S. Savarese. Active learning for convolutional neural networks: A core-set approach. In *Proceedings of the 6th International Conference on Learning Representations*, Vancouver, Canada, 2018.

[23] A. Kapoor, K. Grauman, R. Urtasun, T. Darrell. Active learning with gaussian processes for object categorization. In *Proceedings of the 11th International Conference on Computer Vision*, IEEE, Rio de Janeiro, Brazil, pp. 1–8, 2007. DOI: 10.1109/ICCV.2007.4408844.

[24] Z. Wang, J. P. Ye. Querying discriminative and representative samples for batch mode active learning. *ACM Transactions on Knowledge Discovery from Data*, vol. 9, no. 3, Article number 17, 2015. DOI: 10.1145/2700408.

[25] Y. Gal, Z. Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning*, ACM, New York, USA, pp. 1050–1059, 2016. DOI: 10.5555/3045390.3045502.

[26] Y. Gal, R. Islam, Z. Ghahramani. Deep bayesian active learning with image data. In *Proceedings of the 34th International Conference on Machine Learning*, ACM, Sydney, Australia, pp. 1183–1192, 2017. DOI: 10.5555/3305381.3305504.

[27] S. Ebrahimi, M. Elhoseiny, T. Darrell, M. Rohrbach. Uncertainty-guided continual learning with bayesian neural networks. In *Proceedings of the 8th International Conference on Learning Representations*, Addis Ababa, Ethiopia, 2020.

[28] X. Y. Zhan, H. Liu, Q. Li, A. B. Chan. A comparative survey: Benchmarking for pool-based active learning. In *Proceedings of the 30 th International Joint Conference on Artificial Intelligence*, pp. 4679–4686, 2021.

[29] X. Y. Zhan, Q. Z. Wang, K. H. Huang, H. Y. Xiong, D. J. Dou, A. B. Chan. A comparative survey of deep active learning, [Online], Available: https://arxiv.org/abs/2203.13450, 2022.

[30] X. Y. Zhan, Z. Y. Dai, Q Z. Wang, Q. Li, H. Y. Xiong, D. J. Dou, A. B. Chan. Pareto optimization for active learning under out-of-distribution data scenarios, [Online], Available: https://arxiv.org/abs/2207.01190, 2022.

[31] N. Roy, A. McCallum. Toward optimal active learning through monte carlo estimation of error reduction. In *Proceedings of the 18th International Conference on Machine Learning,*, San Francisco, USA, pp. 441–448, 2001.

[32] S. Tong, D. Koller. Support vector machine active learning with applications to text classification. *Journal of Machine Learning Research*, vol. 2, no. 1, pp. 45–66, 2002. DOI: 10.1162/153244302760185243.

[33] K. Brinker. Incorporating diversity in active learning with support vector machines. In *Proceedings of the 20th International Conference on Machine Learning*, ACM, Washington, USA, pp. 59–66, 2003. DOI: 10.5555/3041838.3041846.

[34] D. Roth, K. Small. Margin-based active learning for structured output spaces. In *Proceedings of the 17th European Conference on Machine Learning*, Springer, Berlin, Germany, pp. 413–424, 2006.

[35] B. Settles, M. Craven. An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, ACM, Honolulu, USA, pp. 1070–1079, 2008. DOI: 10.5555/1613715.1613855.

[36] A. J. Joshi, F. Porikli, N. Papanikolopoulos. Multi-class active learning for image classification. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Miami, USA, pp. 2372–2379, 2009. DOI: 10.1109/CVPR.2009.5206627.

[37] W. J. Luo, A. G. Schwing, R. Urtasun. Latent structured active learning. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, ACM, Lake Tahoe, USA, pp. 728–736, 2013. DOI: 10.5555/2999611.2999693.

[38] W. B. Cai, Y. Zhang, S. Y. Zhou, W. Q. Wang, C. Ding, X. Gu. Active learning for support vector machines with maximum model change. In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, Nancy, France, pp. 211–226, 2014. DOI: 10.1007/978-3-662-44848-9_14.

[39] W. B. Cai, Y. Zhang, J. Zhou. Maximizing expected model change for active learning in regression. In *Proceedings of the 13th IEEE International Conference on Data Mining*, Dallas, USA, pp. 51–60, 2013. DOI: 10.1109/ICDM.2013.104.

[40] B. Settles, M. Craven, S. Ray. Multiple-instance active learning. In *Proceedings of the 20th International Conference on Neural Information Processing Systems*, ACM, Vancouver, Canada, pp. 1289–1296, 2007. DOI: 10.5555/2981562.2981724.

[41] B. Long, O. Chapelle, Y. Zhang, Y. Chang, Z. H. Zheng, B. Tseng. Active learning for ranking through expected loss optimization. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, Geneva, Switzerland, pp. 267–274, 2010. DOI: 10.1145/1835449.1835495.

[42] M. Bilgic, P. N. Bennett. Active query selection for learning rankers. In *Proceedings of the 35th international ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, Portland, USA, pp. 1033–1034, 2012. DOI: 10.1145/2348283.2348455.

[43] W. B. Cai, M. H. Zhang, Y. Zhang. Active learning for ranking with sample density. *Information Retrieval Journal*, vol. 18, no. 2, pp. 123–144, 2015. DOI: 10.1007/s10791-015-9250-6.

[44] M. Taylor, J. Guiver, S. Robertson, T. Minka. SoftRank: Optimizing non-smooth rank metrics. In *Proceedings of International Conference on Web Search and Data Mining*, ACM, Palo Alto, USA, pp. 77–86, 2008. DOI: 10.1145/1341531.1341544.

[45] Z. H. Zheng, K. K. Chen, G. Sun, H. Y. Zha. A regression framework for learning ranking functions using relative relevance judgments. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Amsterdam, The Netherlands, pp. 287–294, 2007. DOI: 10.1145/1277741.1277792.

[46] O. Chapelle, Y. Chang. Yahoo! Learning to rank challenge overview. In *Proceedings of International Conference on Yahoo! Learning to Rank Challenge*, ACM, Haifa, Israel, 2011.

[47] T. Qin, T. Y. Liu, J. Xu, H. Li. LETOR: A benchmark collection for research on learning to rank for information retrieval. *Information Retrieval*, vol. 13, no. 4, pp. 346–374, 2010. DOI: 10.1007/s10791-009-9123-y.

[48] Q. Y. Ai, K. P. Bi, C. Luo, J. F. Guo, W. B. Croft. Unbiased learning to rank with unbiased propensity estimation. In *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, Ann Arbor, USA, pp. 385–394, 2018. DOI: 10.1145/3209978.3209986.

[49] L. X. Zou, W. X. Lu, Y. D. Liu, H. Y. Cai, X. K. Chu, D. H. Ma, D. T. Shi, Y. Sun, Z. C. Cheng, S. M. Gu, S. Q. Wang, D. W. Yin. Pre-trained language model-based retrieval and ranking for web search. *ACM Transactions on the Web*, vol. 17, no. 1, Article number 4, 2022. DOI: 10.1145/3568681.

[50] M. C. Zhuge, D. H. Gao, D. P. Fan, L. B. Jin, B. Chen, H. M. Zhou, M. H. Qiu, L. Shao. Kaleido-BERT: Vision-language pre-training on fashion domain. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Nashville, USA, pp. 12642–12652, 2021. DOI: 10.1109/CVPR46437.2021.01246.

[51] X. S. Luo, L. X. Liu, Y. H. Yang, L. Bo, Y. P. Cao, J. H. Wu, Q. Li, K. P. Yang, K. Q. Zhu. AliCoCo: Alibaba E-

commerce cognitive concept net. In *Proceedings of ACM SIGMOD International Conference on Management of Data*, Portland, USA, pp. 313–327, 2020. DOI: 10.1145/3318464.3386132.

[52] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Gray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, R. Lowe. Training language models to follow instructions with human feedback. In *Proceedings of the 36th Conference on Neural Information Processing Systems*, 2022.

[53] Q. Z. Wang, A. B. Chan. Describing like humans: On diversity in image captioning. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Long Beach, USA, pp. 4190–4198, 2019. DOI: 10.1109/CVPR.2019.00432.

[54] Q. Z. Wang, J. N. Wang, A. B. Chan, S. Y. Huang, H. Y. Xiong, X. J. Li, D. J. Dou. Neighbours matter: Image captioning with similar images. In *Proceedings of the 31st British Machine Vision Conference*, 2020.

[55] J. N. Wang, W. J. Xu, Q. Z. Wang, A. B. Chan. Compare and reweight: Distinctive image captioning using similar images sets. In *Proceedings of the 16th European Conference on Computer Vision*, Springer, Glasgow, UK, pp. 370–386, 2020. DOI: 10.1007/978-3-030-58452-8_22.

[56] H. Zhang, W. C. Yin, Y. W. Fang, L. X. Li, B. Q. Duan, Z. H. Wu, Y. Sun, H. Tian, H. Wu, H. F. Wang. Ernie-vilg: Unified generative pre-training for bidirectional vision-language generation, [Online], Available: https://arxiv.org/abs/2112.15283, 2021.

[57] Q. Z. Wang, J. Wan, A. B. Chan. On diversity in image captioning: Metrics and methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 2, pp. 1035–1049, 2022. DOI: 10.1109/TPAMI.2020.3013834.

[58] J. N. Wang, W. J. Xu, Q. Z. Wang, A. B. Chan. On distinctive image captioning via comparing and reweighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 2, pp. 2088–2103, 2023. DOI: 10.1109/TPAMI.2022.3159811.

Qingzhong Wang received the B. Eng. degree in automation, the M. Eng. in control science and engineering from Harbin Engineering University, China in 2013 and 2016, respectively, and the Ph. D. degree in computer science from City University of Hong Kong, China in 2021. He is now a researcher in Big Data Laboratory, Baidu Research, China.

His research interests include computer vision and vision-language learning.

E-mail: qingzwang@outlook.com

ORCID iD: 0000-0003-1562-8098

Haifang Li received the B. Eng. degree in mathematics from Shandong University, China in 2011, and the Ph. D. degree in mathematics from University of Chinese Academy of Sciences, China in 2016. She is now a senior algorithm engineer at Baidu Inc., China. Before that, she was an assistant researcher at Institute of Automation, Chinese Academy of Sciences, China.

Her research interests include information retrieval and data mining.

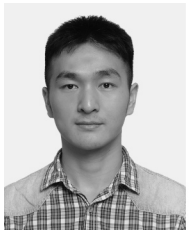
E-mail: lihaifang16@gmail.com

Haoyi Xiong received the Ph. D. degree in computer science from Telecom SudParis and Pierre and Marie Curie University, France in 2015. He is currently a principal architect at Big Data Laboratory, Baidu Inc., China. From 2016 to 2018, he was a Tenure-track assistant professor with Department of Computer Science, Missouri University of Science and Technology, USA. Before that, he was a postdoc at University of Virginia, USA from 2015 to 2016. He has published more than 70 papers in top computer science conferences and journals. He was a co-recipient of the 2020 IEEE TCSC Award for Excellence in Scalable Computing (Early Career Researcher) and the prestigious Science & Technology Advancement Award (First Prize) from Chinese Institute of Electronics in 2019.

His research interests include AutoDL and ubiquitous computing.

E-mail: haoyi.xiong.fr@ieee.org (Corresponding author)

ORCID iD: 0000-0002-5451-3253



Wen Wang received the Ph. D. degree in computer science from Department of Software Engineering, East China Normal University, China in 2021. He is now a senior algorithm engineer at Baidu Inc., China.

His research interests include information retrieval and recommendation systems.

E-mail: wangwen15@baidu.com

Jiang Bian received the B. Eng. degree in logistics systems engineering from Huazhong University of Science and Technology, China in 2014, the M. Sc. degree in industrial systems engineering from University of Florida, USA in 2020, and the Ph. D. degree in computer science from University of Central Florida, USA in 2020. He is a researcher in Baidu Research, China.

His research interests include internet of things, sports analytics and ubiquitous computing.

E-mail: jiangbian03@gmail.com

ORCID iD: 0000-0002-6337-9375

Yu Lu received the B. Eng. degree in computer science from Xidian University, China in 2010, and the M. Eng. degree in computer science from Xi′an Jiaotong University, China in 2014. He is now a senior algorithm engineer at Baidu Inc., China.

His research interests include information retrieval and data mining.

E-mail: luyu06@baidu.com

Shuaiqiang Wang received the B. Sc. and Ph. D. degrees in computer science from Shandong University, China in 2004 and 2009, respectively. He visited Hong Kong Baptist University, China, as an exchange doctoral student in 2009. He is currently a principal algorithm engineer at Baidu Inc., China, leading the Web Search Ranking Strategy Group that advances the document ranking for the Baidu Search Engine. Previously, he was a research scientist and senior algorithm engineer at JD inc.,

China, taking responsibility for the feed recommendation at JD.com. Before that, he worked as an assistant professor at University of Manchester, UK in 2017 and University of Jyvaskyla in Finland, from 2014 to 2017 respectively. Earlier, he served as an associate professor at Shandong University of Finance and Economics, China from 2011 to 2014, and a postdoctoral researcher at Texas State University, USA from 2010 to 2011. He served as Senior PC Member of IJCAI, and PC Member of WWW, SIGIR and WSDM in recent years. He published over 50 papers in leading journals and conferences.

His research interests include information retrieval, recommendation systems and data mining.

E-mail: wangshuaiqiang@baidu.com
ORCID iD: 0000-0002-9212-1947

Zhicong Cheng received M. Sc. degree in computer science from Peking University, China in 2011. He is now a distinguished architect at Baidu Incorporated, China.

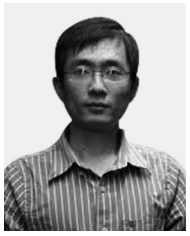
His research interests include learning to rank, machine learning, information retrieval and question answering.

E-mail: chengzhicong01@baidu.com

Dejing Dou received the B. Eng. degree in electronic engineering from Tsinghua University, China in 1996, and the Ph. D. degree in artificial intelligence from Yale University, USA in 2004. He is a professor with Computer and Information Science Department, University of Oregon, USA and the lead of the Advanced Integration and Mining Laboratory (AIM Laboratory). He is also the director of the NSF IUCRC Center for Big Learning (CBL).

His research interests include artificial intelligence, data mining, data integration, information extraction, biomedical and health informatics.

E-mail: doudejing@baidu.com
ORCID iD: 0000-0001-7561-1672



Dawei Yin received the B. Sc. degree in computer science from Shandong University, China in 2006, the M. Sc. and Ph. D. degrees in computer science from Lehigh University, USA in 2010 and 2013, respectively. From 2007 to 2008, he was an M.Phil. student in The University of Hong Kong, China. He is senior director of Engineering at Baidu Incorporated, China. He is managing the search science team at Baidu, leading Baidu′s science efforts of web search, question answering, video search, image search, news search, app search, etc. Previously, he was senior director, managing the recommendation engineering team at JD.com between 2016 and 2020. Prior to JD.com, he was senior research manager at Yahoo Labs, leading relevance science team and in charge of Core Search Relevance of Yahoo Search. He published more than 100 research papers in premium conferences and journals, and was the recipients of WSDM2016 Best Paper Award, KDD2016 Best Paper Award, WSDM2018 Best Student Paper Award.

His research interests include data mining, applied machine learning, information retrieval and recommender systems.

E-mail: yindawei@acm.org (Corresponding author)
ORCID iD: 0000-0002-8846-2001