Answer Ranking by Analyzing Characteristic of Tags and Behaviors of Users

Qian Wang[1], Lei Su[1](✉), Yiyang Li[1], and Junhui Liu[2]

[1] School of Information Engineering and Automation,
Kunming University of Science and Technology, Kunming 650093, China
m18288793132@163.com, s28341@hotmail.com
[2] School of Software, Yunnan University, Kunming 650091, China

Abstract. The quality of the ranking answer is good or bad, directly affects the high quality answers for users in the community question answering system. Learning method by sorting, establish the answer ranking model, is a research hotspot in community question answering system. The characteristics of tags and behavior of users, often have a direct relationship with the answer to the users' expectations. In this paper, ListNet is used as the ranking method which selects Neural Networks as the model and Gradient Descent as the optimization method to structure ListNet ranking model which blends in characteristics of tags and behaviors of user. Then, the ranking mode is utilized to finish experiment combining the answers feature space, and the result of experiment shows that the ListNet ranking model can improve effect of answers ranking obviously which blends in the characteristics of tags and behaviors of users.

Keywords: Community question answering system · User tags · Behavioral characteristics · ListNet · Gradient descent · Feature space

1 Introduction

With the rapid development of the question and answer services which are based on community, the Community-based Question Answering service has become a new knowledge-sharing model [1], social network. It is inconvenient for users to choose web pages from a large number of returned web pages through keyword matching by traditional search engines. However, the community question answering is an open, interactive network platform, which uses the collective wisdom of the network users, through the participation of users and provides a direct answer to the question, it provides a new way and platform for the sharing of the internet knowledge, and also brings new vitality life for answering technology. The community question answering system [2] develops rapidly with the mainstream of Baidu know and Yahoo! in recent years.

The answer ranking is an important issue to be solved in community question answering [3]. The effect of the answer ranking directly affects the quality of community question answering system and the users' experience. In order to return the best answer directly to users, you must choose the best answer from a number of answers. In answer ranking model, the answer of the most consistent with the users' needs will be put in the front row, so as to locate the target knowledge directly for users. So, the

J. Wan et al. (Eds.): CloudComp 2016, SPNCE 2016, LNICST 197, pp. 56–65, 2018.
https://doi.org/10.1007/978-3-319-69605-8_6

performance of the answer ranking model is very important, its accuracy directly determines the performance of the entire community question answering system. Therefore, it is worthy of study the sort of the answer ranking in the community question answering.

At present, ranking learning has attracted the interest of many scholars and has become a research hotspot of scholars. There have been many methods have achieved very good results in practice, such as the LambdaRank method proposed by Burges in 2006 [4]; list rank method based on the ListNet proposed by Cao in 2007, directly sequencing the whole return list [5]; the RankCosine method proposed by Qin in 2008, which is based on the level of query to construct the loss function, with the method of Boosting to optimize [6]. As for the answer ranking in community question answering system, the characteristics of the community that is the tags and behavior of users, which makes great effect on the answering ranking. For example, the level and experience value of respondents, areas of expertise, respondent's adoption rate and approval rate, the keywords to answer question and other user tags features that the respondents concerned, reflects the senior level of the person who answers the question, Further to say, it reflects the credibility of the answer provided. Generally speaking, the higher level of the respondents, the higher value of experience, the higher rate of adoption and the higher approval rate, the higher quality of the questions answered. The category of the question often being answered, the score of the respondents answered questions and the tourists together with other users' behavior have great impact on the credibility of the answers, directly responses to the answers with respect to issues related to the degree of matching, and further reflects the answer can be taken. Therefore, the characteristic of tags and behavior of users and other community features blend in the answers feature space to improve the effect of the answer ranking is worthy of study and realization.

In this paper, the characteristic of tags and behaviors of users blend in the answers feature space, combined with ListNet rank learning method to construct ranking model so as to improve the effectiveness of answer system in community question answering. Finally, the ranking mode is utilized to finish experiment combining the answers feature space, and the result of experiment show that the ListNet ranking model can improve effect of answers ranking obviously which blend in the characteristics of tags and behaviors of user.

The rest of this paper is organized as follows. Section 2 focuses on the characteristic of tags and behaviors of users blend in the answers feature space and the feature extraction; Sect. 3 presents the sorting method of ListNet in the community question answering ranking method; Sect. 4 reports on the classification experimental and results analysis based on the domain of "Baidu know"; Finally, Sect. 5 gives a summarize of the main study of the paper.

2 Factors Affecting the Ranking of Answers

There are many factors that affect the performance of the answer ranking in community question answering system, such as the similarity, the density and frequency of the query and the candidate answer.

In the community question answering, the level of respondents, the field of the respondents interest, the rate of adoption and the approval of the respondents, the questions and answers the experience value, answer the questions the focus of keywords, the score of the respondents and tourists, which attribute to the tags and users' behavior, they are important supporting elements for answer ranking and important factors to affect the answer ranking in community question answering system. Therefore, it is necessary to consider the effect of the tag and behavior of the user in the order of the ranking answer. The following details are the content of the two aspects.

2.1 User Tags

User tags include the level of the respondents, the professional field of the respondents, the rate of adoption and the approval of residents, the experience value and problem on keywords for answer. When the user answers the questions in the community question answering, there will be other relevant users, ask questions or tourists who have related knowledge needs will give their votes, scores or adopt or not. When a user answers a lot of questions, some of the answers will be adopted, so the rate of approval and adoption reflect the degree of the authority of the user. When a user takes some activities in a certain period of time in CQA, there will be relevant experience value, the more frequent activities, the more questions answered, the more experience value will be. So user tags reflect the community attitudes and the quality of answers, which is an important factor affecting the rank of the answer, it is very necessary to integrate users tags into the answer ranking model.

2.2 User Behavior

User behavior includes the score of the questions, as well as score of questions by tourists and the category of the questions and so on. In CQA System, for one particular issue, there will be a lot of other answers which provided by other users, the quality of these answers or credibility may be good or bad, the questioner will give the score or vote according to their own needs and the professional degree of the answer, and users with the same or similar knowledge needs will give the score or vote in the same way, the scoring or voting reflects the credibility of the answers to the corresponding problems. If a user often answers a question or a question of a particular field, then the user is likely to be good at this area or field in CQA and his answer to this question is of relatively high reliability. Therefore, the user behavior also reflects the quality of the relevant answers, which is an important factor that affects the rank of answer, so, the characteristic of tags and behaviors of users and other community features blend in the answers feature space to improve the effect of the answer ranking is worthy of study and implementation.

2.3 Method of Feature Extraction

The tags and behavior of users have an important effect for the effect of answer ranking in CQA. Therefore, in order to improve the ranking accuracy of the process, it is worth studying the importance of community characteristics. This thesis relies on the platform of the Baidu know to collect the answers to the problems by hand, while the characteristics of tags and behaviors of users were extracted so as to be blended in the answers feature space.

Word similarity computing is a basic research topic of natural language and widely used in natural language processing, information retrieval, text classification, automatic response, the meaning of the word row discrimination and machine translation field and other areas. It has attracted more and more researchers attention [10, 11] called "hit the extended version of Tongyicicilin [7, 8]".

In order to solve the problem of sparse matrix, this thesis introduces the method of calculating the semantic extension of words based on the synonyms Clilin method proposed by Liu and Wang [8, 9]. This method analyzed Clilin hierarchical structure, and combined the semantic with lexical chain extension and proposed a relatively novel text keyword extraction method based on the semantic relation between the word [9].

According to the hierarchical structure of the tree, all the words in the dictionary are divided into 3 levels, including 12 larger categories, 97 middle categories, and 1400 small classes [10]. All kinds of the small classes in the word forest contain a lot of words and each of them is divided into a number of words according to the meaning and the relevance of the word [10]. The words and expressions in each word group are divided into many lines according to the distance and the relevance of the word meaning. The same line of words is not the same word meaning, that is, the word meaning has a strong correlation [10]. The thesaurus is classified by hierarchical system, and the whole dictionary has 5 layers of structure. With the delicate classification of meaning step by step, the number of words in each category is very small to fifth layers, many of the words in the classification can't be classified again, that is, the atomic word group, atomic class or atomic node [10].

The semantic extension of words include two parts: word similarity and word correlation calculation. Word similarity calculation is based on the synonyms word Lin encoding distance to the two words semantic similarity calculation. Its main idea is to determine the two words in the word forest belong to which layer of branches. Then according to the semantic distance of the two words to calculate the similarity between the two words, which, the closer of semantic distance between two words, the higher word similarity they are.

The formula for calculating the similarity of words is as follows:

$$sim(w_1, w_2) = d.\left(\frac{n-k+1}{n}\right).\cos\left(n.\frac{\pi}{180}\right) \tag{2.1}$$

In this formula, $sim(w_1, w_2)$ is semantic similarity $(0 < sim < 1).d$ is coefficient. The two word similarity calculation decided by the needs of the encoding branch. n is the total number of nodes in the branch layer. k is the distance between branches.

Then calculate the words in the semantic relevancy. Make use of the semantic relation between the words in the "a synonym in the word", and calculate the relevance degree of the two words by means of statistical methods. Firstly, find out the correlation calculation of the word w_1 and w_2 and the corresponding encoding code1 and code2 in "a synonym in the word forest", if the coding code1 is equal to code2, and two codes' bit 8 is marked as "#", then these two words correlation degree is 1; if the coding code1 is equal to code2, but the two code s' bit 8 is marked as "=", then the two words correlation degree is 0.85; otherwise, we must calculate out the times of the two words appear at the same time and the times of them appear alone, then statistical information is substituted into the formula to calculated correlation degree between the two words. The formula for calculating the correlation of words is as follows:

$$\mathrm{rel}(w_1, w_2) = \frac{count(w_1, w_2)}{\min(count(w_1), count(w_2))} \tag{2.2}$$

In this formula, $count(w_1, w_2)$ is the number of the w_1 and w_2 both appears in the question. $count(w_1)$ and $count(w_2)$ is the number of the w_1 and w_2 appears alone in the question. $\min(count(w_1), count(w_2))$ is the w_1 and w_2 minimum number of occurrences alone.

3 Attribute Reduction Based on List Net Sort Method

The rank learning has three category methods: based on PointWise, PairWise and ListWise. And ListNet is a rank method based on ListWise. Cao came up with the method of the feedback corresponding to the entire list of search rank [11]. ListNet rank model is used Neural Network as a model ω, based on the probability of the entire arrangement of the feedback list $p_s(\pi)$ and Partition Function $f\left(x_j^{(i)}\right)$, and use Gradient Descent as an optimization method. Through continuous training, so that the loss of function is the best and then output the sort model. Arranged probability formula is as follows:

$$p_s(\pi) = \prod_{j=1}^{n} \frac{\varphi(s_{\pi(j)})}{\sum_{k=j}^{n} \varphi(s_{\pi(k)})} \tag{3.1}$$

$\phi()$ is an increasing function and constant greater than 0. For example, linear function $\varphi(x) = \alpha x, a > 0$, Exponential function $\varphi(x) = exp(x)$, this thesis chooses the Exponential function. π is a list of retrieve and feedback query. $S_\pi(j)$ is the score of the goal scoring function of the arrangement for the first chapter of the document score. Because of the low efficiency of a ranked list of the entire sorting list, we only calculate the first k article document at present; that is, $Top(k)$ is the Probability Model. Here is the formula:

$$p_s(\wp_k(j_1, j_2, \cdots, j_k)) = \sum_{\pi \in \wp_k(j_1, j_2, \cdots, j_k)} p_s(\pi) = \prod_{t=1}^{k} \frac{exp(s_{j_t})}{\sum_{l=t}^{n} exp(s_{j_l})} \tag{3.2}$$

In this formula $\wp_k(j_1, j_2, \ldots, j_k)$ is a permutation of the previous K document. The paper chooses $k = 1$. Then the probability model becomes:

$$p_s(\wp_1(j_1)) = \frac{\exp(s_j)}{\sum_1^n \exp(s_{j_1})} \tag{3.3}$$

The Loss Function of ListNet ranking method is as follows:

$$L(y^{(i)}, z^{(i)}(f_\omega)) = -\sum_{\forall g \in \wp_k} p_{y^{(i)}}(g) \log(p_{z^{(i)}(f_\omega)}(g)) \tag{3.4}$$

In this formula, Permutation probability is a neural network model of the rank function or scoring function. Its formula is as follows:

$$f_\omega(x_j^{(i)}) = \langle \omega, x_j^{(i)} \rangle \tag{3.5}$$

Here $< . >$ is Inner product. ListNet is used as the ranking method which selects Gradient Descent as the optimization method and Gradient descent calculation formula is:

$$\Delta\omega = \frac{\partial L(y^{(i)}, z^{(i)}(f_\omega))}{\partial \omega} = -\sum_{\forall g \in \wp_k} \frac{\partial p_{z^{(i)}(f_\omega)}(g)}{\partial \omega} \frac{p_{y^{(i)}}(g)}{p_{z^{(i)}(f_\omega)}(g)} \tag{3.6}$$

In the process of ranking model repeated training, using Gradient Descent method to continuously optimize the loss function, until the order of the model's loss function is optimal. The pseudo-code of ListNet learning method is shown in Fig. 1.

ListNet learning method

Input: training set $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})\}$
Parameters settings: steps T, learning rate η
Initialization parameters ω
For $t = 1$ to T do
For $t = 1$ to m do
Put list of documents query $x^{(i)}$ of i $q^{(i)}$ to the neural network combining gradient descent ω calculation corresponding scores list $z^{(i)}(f_\omega)$
Use (3.6) calculation Gradient $\Delta\omega$
Update gradient $\omega = \omega - \eta \times \Delta\omega$
End for
End for
Output Neural Network Model

Fig. 1. The pseudo-code of ListNet

4 Experiment Results and Analysis

This section is to verify the rank results of a list ranking method that blend in the characteristics of tags and behaviors of users. The experimental data is from Baidu know, collected a total of 150 questions and 1499 answers, the question category covers 10 small classes, and marked annotation of the degree of correlation between the answers and questions. These 10 sub-classes are: the use of mobile phones, health care, the common sense of life, the employment, fitness, outdoor sports, holiday tourism, flowers, birds, fish and insects, and pediatric traumatology.

Experiments adopted 10-fold cross-validation method, divide the 150 questions and 1499 answers corpus according to the proportion into test set T_1 and training set T_2. Here, the Proportion of T_1 is 20, the Proportion of T_2 is 80. That is to say, take 30 questions and 300 answers for the test set, the remaining 120 questions and 1199 answers for the training set. In order to verify the effect of the combination of user tags and behavior feature, we will put each relevant answer and the characteristics of tags and behavior of users blended in answer feature space, in the training and testing phases of answer ranking model, the steps of List Net is set to 1500 and learning rate is 0.00001.

4.1 The Comparison of Different Ranking Methods

This thesis chooses the rank of ListNet method as the answer method in the CQA system, with the extended version of tongyicicilin based on hit word similarity computing as a method of data processing [11]. In order to make the experiment more sufficient, the experiments were carried out 50 and 136 dimensional feature space respectively, and used five kinds of ring methods including ListNet and NDCG evaluation, MAP evaluation and P@1 evaluation of the 3 evaluation methods to compare. The result of answering sorting in different sorting methods, different dimensions of the feature space and different evaluation indicators of the answers in community question answering system are shown in Table 1.

Table 1. The result of different sorting method dimension and evaluation index

Dimensions	Evaluation indexes	Ranking method				
		RankNet	RankBoost	AdaRank	LambdaRank	ListNet
50	NDCG	0.7063	0.6808	0.6941	0.6998	0.7114
	MAP	0.8190	0.7956	0.7958	0.8488	0.8483
	P@1	0.7633	0.7000	0.5667	0.7867	0.7900
136	NDCG	0.6925	0.6977	0.6779	0.6977	0.7094
	MAP	0.8216	0.8224	0.7958	0.8417	0.8341
	P@1	0.7800	0.8333	0.5667	0.7500	0.8367

From Table 1 we can be see, as for the RankNet ranking method, RankBoost ranking method and AdaRank ranking method through NDCG evaluation methods, MAP evaluation methods and P@1 evaluation method for answer ranking results evaluation, ListNet ranking methods have better sorting effect in 50 and 136 dimensions feature space experiments. As for LambdaRank ranking method, in the case of MAP evaluation is slightly better than the ListNet method, but there is also an obvious gap between ListNet in the NDCG evaluation and P@1 evaluation method. In general, whether it is in the 50dimensional feature space or in the 136 dimensional feature space, with the NDCG evaluation methods, MAP evaluation methods and P@1 evaluation methods the evaluation results show that ListNet ranking method performed better and more valid than other ranking methods in community question answering system answers ranking task.

4.2 The Rank Method Blend in Tags and Behavior of User

In this section, the characteristics of tags and behavior of users are blended in the feature space to improve the effect of the answer to the question in CQA systems. In order to verify the validity of characteristics of the tags and behavior of the user in the CQA system, different ranking method are used. That is, RankNet ranking method, RankBoost ranking method, LambdaRank ranking method, AdaRank ranking method and ListNet ranking method and different feature space dimensions 50, 59,136 and 145 dimensional are used to do experiment. At the same time, with different evaluations such as NDCG evaluation methods, MAP evaluation methods, and P@1 evaluation methods to evaluate the results of the ranking. After the characteristics of the tags and behavior of user blended in 50 and 136 dimensional feature space turn to 59 and 145 dimensional feature space. The result of the experiment is shown in Tables 2 and 3.

Table 2. The result of 50 and 59 dimensional feature space

Rankting method	Evaluation					
	NDCG		MAP		P@1	
	Unfused feature	Fusion features	Unfused feature	Fusion features	Unfused feature	Fusion features
RankNet	0.7063	0.7103	0.8190	0.8629	0.7633	0.8100
RankBoost	0.6808	0.6808	0.7956	0.7956	0.7000	0.7000
AdaRank	0.6941	0.6941	0.7958	0.7958	0.5667	0.5667
LambdaRank	0.6998	0.7134	0.8488	0.8359	0.7867	0.7667
ListNet	0.7114	0.8041	0.8483	0.8889	0.7900	0.9000

Table 3. The result of 136 and 145 dimensional feature space

Ranking method	Evaluation					
	NDCG		MAP		P@1	
	Unfused feature	Fusion features	Unfused feature	Fusion features	Unfused feature	Fusion features
RankNet	0.6925	0.7315	0.8216	0.8300	0.7800	0.8133
RankBoost	0.6977	0.6977	0.8224	0.8224	0.8333	0.8333
AdaRank	0.6779	0.6779	0.7958	0.7958	0.5667	0.5667
LambdaRank	0.6977	0.7097	0.8417	0.8257	0.7500	0.7767
ListNet	0.7094	0.7503	0.8341	0.8723	0.8367	0.8867

Tables 2 and 3 show that each dimension feature space, the characteristics of tags and behavior of user blended in the feature space, and there is no obvious improvement, and there is even a sign of decline for RankBoost ranking method, AdaRank ranking methods and AdaRank ranking methods. However, as for the RankNet ranking method and ListNet ranking method with the characteristics of the tags and behavior of user characteristics, the ranking effect is significantly improved. And it can also be seen that the ListNet ranking method is far better than the RankNet ranking method. For example, the characteristics of tags and behavior of user blend in the 50 dimensional feature space, the result of RankNet ranking method in NDCG evaluation method, MAP evaluation method and P@1 evaluation methods increase respectively for 0.004, 0.0439 and 0.0467; however, ListNet increase for respectively 0.0927, 0.0406 and 0.11; the characteristics of tags and behavior of user blend in the 136 dimensional feature space, the result of RankNet ranking method in NDCG evaluation method, MAP evaluation method and P@1 evaluation methods increase respectively 0.039, 0.0084 and 0.0333; but the characteristics of the tags and behavior of the user, the result of ListNet ranking method increase respectively 0.0409, 0.0382 and 0.05. After the characteristics of the tags and the behavior of the user blend in the answers feature space, the result of the ListNet ranking method in the community question answering system is still better than other methods, and compared with other methods, the ranking effect is more obviously than other methods. So, the experiment proves that the ListNet ranking method is effective in the community question answering system again, and it is quite obvious that the result of answer ranking with the characteristics of tags and behavior of user blend in the feature space.

5 Conclusion

This paper, mainly introduces the answer list ranking method of the characteristics of the tags and the behavior of the user in the community question answering system, with the characteristics of the tags and behavior of the user blended in the answering feature space to improve the answer ranking accuracy effectively. Experimental comparison of multiple angles from the ranking method of comparison, the dimensions of the features comparison and evaluation index of verify the list ranking algorithms ListNet in community question answering ranking task effectiveness, and the experiment proves that the ListNet ranking method is effective in the community question answering system, and it is obvious that the effect of the user tag and the user behavior characteristics on the answer ranking.

Acknowledgement. This work is supported by the National Science Foundation on of China under the Grant No. 61365010.

References

1. Zhang, Z.F., Li, Q.D.: Summary of community question answering system. Comput. Sci. **37** (11), 19–23 (2011)
2. Yan, X., Fan, S.X.: Coarse grain question classification method for question answering community. Comput. Appl. Softw. **30**(1), 220–223 (2013)
3. Voorhees, E.M.: Overview of the TREC 2007 Question Answering Tack. In: TREC 2007, vol. 10(8), pp. 115–123. Gaithersburg, New York (2007)
4. Quoc, C., Le, V.: Learning to rank with nonsmooth cost functions. In: Proceedings of the Advances in Neural Information Processing Systems, vol. 19, pp. 193–200. Vancouver, Canada (2007)

5. Cao, Z., Qin, T., Liu, T.Y., et al.: Learning to rank: from pairwise approach to listwise approach. In: Proceedings of the 24th International Conference on Machine Learning, pp. 129–136. ACM, Lisboa, Portugal (2007)
6. Qin, T., Zhang, X.D., Tsai, M.F., et al.: Query-level loss functions for information retrieval. Inf. Process. Manag. **44**(2), 838–855 (2008)
7. Yu, G., Pei, Y.J., Zhu, Z.Y.: Research on text similarity based on semantic similarity of words. Comput. Eng. Des. **27**(2), 241–244 (2006)
8. Tian, Y.L., Zhao, W.: Word similarity calculation method based on synonyms. J. Jilin Univ. (Inf. Sci. Ed.) **28**(6), 603–604 (2010)
9. Liu, R.Y., Wang, L.F.: Keyword extraction algorithm based on semantic extension and lexical chain. Comput. Sci. **40**(12), 265–266 (2013)
10. Seung, H.S., Opper, M., Sompolinsky, H.: Query by committee. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory, pp. 287–294. ACM, New York, America (1992)
11. Mei, J.J., Zhu, Y.M., Gao, Y.Q.: Synonymy Thesaurus. Shanghai Lexicographical Publishing House, Shanghai (1993). pp. 106–108
12. Li, B., Gao, W.J., Qiu, X.P.: An answer ranking model based on syntax analysis and statistical method. J. Chin. Inf. **23**(2), 23–27 (2009)
13. Lian, X.: Research on some questions of community question answering system, pp. 2–3. School of Computer and Control Engineering, Nankai University, Tianjin (2014)