

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

Survey on Advanced Approach to Construct E-Learning QA Systems

Shubhangi Pawar, Rajesh Kulkarni.

Department of Computer Engineering, JSPM Narhe Technical campus, Pune, India

ABSTRACT: The data-driven approach can yield high levels of performance and nicely complements traditional question answering techniques driven by information extraction. In order for question answering systems to benefit from this vast store of useful knowledge, they must cope with large volumes of useless data. Question Answering systems (QA) uses similarity in questions and ranking the relevant answer to user. The web gives large data and that require more time as well as no relevancy in answers. To solve this problem proposed system proposed novel Pair-wise Learning to rank model i.e PLANE which can quantitatively rank answer candidates from the relevant question pool. Specifically, it uses two components i.e one offline learning component and one online search component. Our model is effective as well as achieves better performance than several existing questions answer selection system. User gets recommendation based on his profile. User recommend the new question to his friend and this is trust analysis so user can get top recommendation of newly arrived question of languages.

KEYWORDS: Community-based Question Answer, Answer Selection, Pair-wise learning technique

I. INTRODUCTION

In the web user, often, the hunger for questions is probably due to several reasons: 1) the questions are poorly written, ambiguous or not at all interesting; 2) QA systems can hardly address the newly published questions to the appropriate respondents; and 3) potential respondents have the corresponding experience, but are not available or are overwhelmed by the large volume of incoming questions. This case often occurs in vertical QA forums, whereby only authorized experts can answer these questions. Regarding the first case, the quality model of the application has been well studied, which can assess the quality of the application and serve to remind the respondents to reformulate their questions. Routing applications work by exploring the resources of the current system, in particular human resources. Beyond that, we can reuse the resolved questions from the past to answer new questions. In fact, a large number of historical QA pairs, over time, have been archived in the QA databases. Therefore, information seekers have a good chance of getting direct answers looking for repositories, instead of waiting long[12]. Inspired by this,[11] Wang et al. They have transformed the quality control task into the task of finding relevant and similar questions. However, candidates returned from the main application are generally associated with multiple answers and research on how to choose the correct answers from the relevant question group is relatively poor. When a question is asked, instead of naively choosing the best answer to the most pertinent question, In this paper, we present a new Pair-wise Learning to Rank model, dubbed PLANE, which can quantitatively classify candidates from the relevant question group. Figure 1 shows the workflow of the PLANE model, which consists of two components: offline learning and online research.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

II. RELATED WORK OR LITERATURE SURVEY

No	Paper Name and Year	Author Name	Proposed System Advantages and disadvantages	Referred Point
1.	Exploring heterogeneous features for query-focused summarization of categorized community answers (2016)	W. Wei, Z. Ming, L. Nie, G. Li, J. Li, F. Zhu, T. Shang, and C. Luo	Advantages: <ol style="list-style-type: none"> Proposes a three level scheme Generate answer in terms of two factors, i.e., novelty and redundancy. Disadvantages: <ol style="list-style-type: none"> More redundancy Less fine grained access control. 	We referred the term novelty.
2.	Multivcrank with applications to image retrieval (2016)	X. Li, Y. Ye, and M. K. Ng	Advantages: <ol style="list-style-type: none"> Develop a multivisual- concept ranking (MultiVCRank) scheme. Use a hyper graph concept. Disadvantages: Sometime generate wrong a hyper graph.	We referred the ranking scheme.
3.	Learning to find topic experts in twitter via different relations (2016)	W. Wei, G. Cong, C. Miao, F. Zhu, and G. Li.	Advantages: <ol style="list-style-type: none"> Develop a probabilistic method to jointly exploit three types of relations (i.e., follower relation, user-list relation, and list-list relation) for finding experts. Propose a Semi-Supervised Graph-based Ranking approach (SSGR) to offline calculate the global authority of users Disadvantages: Improper local relevance between users.	We referred the Semi-Supervised Graph-based Ranking approach.
4.	A ranking approach on large-scale graph with multidimensional heterogeneous information (2016)	W. Wei, B. Gao, T. Liu, T. Wang, G. Li, and H. Li.	Advantages: <ol style="list-style-type: none"> Reduce large-scale graph-based ranking problem. Propose an innovative and effective semi-supervised Page Rank (SSP) approach to parameterize the derived information within a unified semi-supervised learning framework (SSLF-GR), Disadvantages: <ol style="list-style-type: none"> Lack in data confidentiality Less security. 	We have referred the SSP approach.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

5.	Reranking voting-based answers by discarding user behaviour biases .(2015)	X. Wei, H. Huang, C. Lin, X. Xin, X. Mao, and S. Wang.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Analyze two kinds of biases; position bias and appearance bias . 2. To identify the existence of these biases and propose a joint click model for dealing with both of them. <p>Disadvantages: The top answers obtained from this mechanism are not reliable, if the number of votes for the active question is not sufficient.</p>	We have referred kinds of biases; position bias and appearance bias
6.	Jaist: Combining multiple features for answer selection in community question answering (2015)	Q. H. Tran, V. Duc, Tran, T. T. Vu, M. L. Nguyen, and S. B. Pham.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Designed Answer Selection in Community Question Answering . 2. System identify the good or potentially good answers from the answer thread in Community Question Answering collections. <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Delegation problem. 2. Less fine grained access control. 	We referred the Answer Selection in Community Question Answering System.
7.	Ranking answers and web passages for non-factoid question answering: Emory university at TREC liveQA. (2015)	Savenkov	<p>Advantages:</p> <ol style="list-style-type: none"> 1. Represents how to automatically answer questions posted to Yahoo! Answers community question answering website in real-time. 2. Answers archive using Pairwise ranking criterion. <p>Disadvantages:</p> <ol style="list-style-type: none"> 1. Less security. 2. Does not find proper answer. 	We referred the Pairwise ranking criterion
8.	Learning to recommend descriptive tags for questions in social forums (2014)	L. Nie, Y. Zhao, X. Wang, J. Shen, and T. Chua	<p>Advantages: Presents a novel scheme that is able to comprehensively learn descriptive tags for each question.</p> <p>Disadvantages: In social-oriented question answering forums have at most one manually labelled tag, which is caused by incomprehensive question understanding or informal tagging behaviours.</p>	We referred the a novel scheme that is able to comprehensively learn descriptive tags for each question.
9.	Learning to rank for question routing in community question answering (2013)	Z. Ji and B. Wang.	<p>Advantages:</p> <ol style="list-style-type: none"> 1. proposes a general framework based on the learning to rank concepts for QR . 2. Uses SVM-based and Rankings-based methods to 	We referred the SVM classifier.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

			rank answers. Disadvantages: Ranking problem.	
10.	Exploiting user feedback to learn to rank answers in qa forums: A case study with stack overflow .(2013)	H. Dalip, M. A. Goncalves, M. Cristo, and P. Calado.	Advantages: <ol style="list-style-type: none"> propose a learning to rank (L2R) approach for ranking answers in Q&A forums . The L2R method was trained to learn the answer rating, based on the feedback users give to answers in Q&A forums. Disadvantages: Less integrity, data confidentiality	We referred the solution the cloud servers can offer various data services and outsourced delegation computation.

III. EXISTING SYSTEM AND DISADVANTAGES

Finding similar questions from historical archives has been applied to question answering, with well theoretical and great practical success. Nevertheless, each question in the returned candidate pool often associates with multiple answers, and hence users have to painstakingly browse a lot before finding the correct one[12].

Disadvantages:

- Information seekers usually have to wait a long time before getting answers to their questions.
- A large proportion of questions do not get any response even within a relatively long period.

IV. PROPOSED SYSTEM

System presents a novel scheme to rank answer candidates via pair wise comparisons. The system consists of one offline learning component and one online search component. In the offline learning component, we first automatically establish the positive, negative training samples in terms of preference pairs guided by our data-driven Approach. Present a novel model to jointly incorporate these types of training samples. In the online search component, we first collect a set of answer candidates for the given question via finding its similar questions. We then sort the answer candidates by working on the offline trained model to judge the preference orders. Proposed systems that allow a user to ask a question for computer language and receive an answer quickly and relevantly, with sufficient context to validate the answer. Current search engines can return ranked lists of answer with relevancy. System will check voting and review and according to that it will arrange ascending order. To match the feature we using lavenstine distance algorithm. User recommends the new question to his friend and this is trust analysis so user can get top recommendation of newly arrived question of language's.

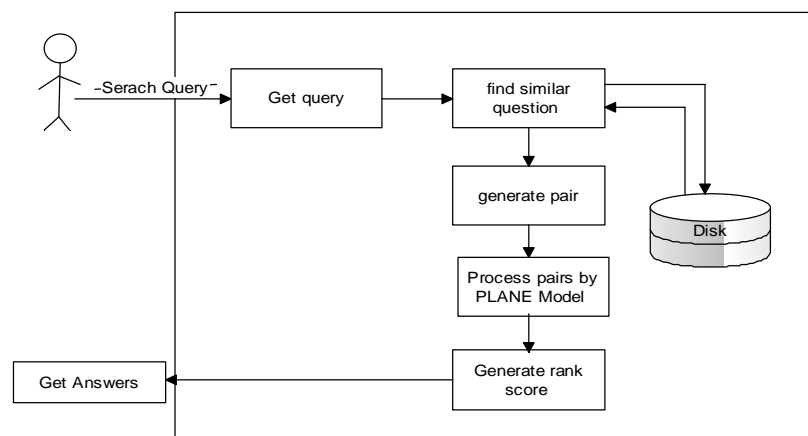


Figure 01: Proposed System i.e PLANE model

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 1, January 2018

Advantages:

- Our model can achieve better performance than several state-of-the-art answer selection Techniques.
- Our model is robust to the noises caused by enlarging the number of returned similar questions.
- The pair wise learning to rank models including our proposed PLANE are very sensitive to the error training samples.
- Provide relevancy

V. CONCLUSION

This paper presents a system for selecting answers in the QA system. It consists of an offline search and online search component. In learning offline components, instead of taking a lot of time and effort, it automatically creates positive, neutral, pairs. In the online search component, for particular question first system gather all a group of answers to find candidates through their similar questions. So we use the offline model to classify candidate responses through pair-wise comparison. Proposed system will give top recommended question to user according to trust based analysis.

REFERENCES

- [1] W. Wei, Z. Ming, L. Nie, G. Li, J. Li, F. Zhu, T. Shang, and C. Luo, "Exploring heterogeneous features for query-focused summarization of categorized community answers," *Inf. Sci.*, vol. 330, pp. 403–423, 2016.
- [2] X. Li, Y. Ye, and M. K. Ng, "Multivcrank with applications to image retrieval," *TIP*, vol. 25, no. 3, pp. 1396–1409, 2016.
- [3] W. Wei, G. Cong, C. Miao, F. Zhu, and G. Li, "Learning to find topic experts in twitter via different relations," *TKDE*, vol. 28, no. 7, pp. 1764–1778, 2016.
- [4] W. Wei, B. Gao, T. Liu, T. Wang, G. Li, and H. Li, "A ranking approach on large-scale graph with multidimensional heterogeneous information," *TOC*, vol. 46, no. 4, pp. 930–944, 2016.
- [5] X. Wei, H. Huang, C. Lin, X. Xin, X. Mao, and S. Wang, "Reranking voting-based answers by discarding user behavior biases," in *Proceedings of IJCAI'15*, 2015, pp. 2380–2386.
- [6] Q. H. Tran, V. Duc, Tran, T. T. Vu, M. L. Nguyen, and S. B. Pham, "Jaist: Combining multiple features for answer selection in community question answering," in *Proceedings of SemEval'15. ACL*, 2015, pp. 215C–219.
- [7] Savenkov, "Ranking answers and web passages for non-factoid question answering: Emory university at TREC liveqa," in *Proceedings of TREC'15*, 2015.
- [8] L. Nie, Y. Zhao, X. Wang, J. Shen, and T. Chua, "Learning to recommend descriptive tags for questions in social forums," *TOIS*, vol. 32, no. 1, p. 5, 2014.
- [9] Z. Ji and B. Wang, "Learning to rank for question routing in community question answering," in *Proceedings of CIKM'13. ACM*, 2013, pp. 2363–2368.
- [10] D. H. Dalip, M. A. Goncalves, M. Cristo, and P. Calado, "Exploiting user feedback to learn to rank answers in qa forums: A case study with stack overflow," in *Proceedings of SIGIR'13. ACM*, 2013, pp. 543–552.
- [11] K. Wang, Z. Ming, and T.-S. Chua, "A syntactic tree matching approach to finding similar questions in community-based qa services," in *Proceedings of SIGIR'09. ACM*, 2009, pp. 187–194.
- [12] M. A. M. L. Wei Ding, He Jiang, *Modern Advances in Intelligent Systems and Tools*, ser. SCI. Springer, 2012, vol. 431.
- [13] M. Surdeanu, M. Ciaramita, and H. Zaragoza, "Learning to rank answers on large online qa collections," in *Proceedings of ACL'08. ACL*, 2008, pp. 719–727.
- [14] C. Shah and J. Pomerantz, "Evaluating and predicting answer quality in community qa," in *Proceedings of SIGIR'10. ACM*, 2010, pp. 411–418.
- [15] Hieber and S. Riezler, "Improved answer ranking in social question-answering portals," in *Proceedings of SMUC'11. ACM*, 2011, pp. 19–26.