



International Journal of Emerging Technology and Advanced Engineering

Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue No., Current Month 2017)

Novel based question and answering for programming solution

S. Nandini¹, N. Hemalatha², C. Janani³, Komal Roy⁴

¹Assistant Professor

^{2,3,4}UG Scholar

¹sn.nandini21@gmail.com

Department of CSE

Dhanalakshmi College of Engineering, Chennai.

Abstract— Finding similar questions from historical archives has been applied to question answering, with well theoretical underpinnings 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. To alleviate such problem, we present a novel scheme to rank answer candidates via pairwise comparisons. In particular, it consists of one offline learning component and one online search component. In the offline learning component, we first automatically establish the positive, negative, and neutral training samples in terms of preference pairs guided by our data-driven observations. We then present a novel model to jointly incorporate these three types of training samples. The closed-form solution of this model is derived. In the online search component, we first collect a pool of answer candidates for the given question via finding its similar questions. We then sort the answer candidates by leveraging the offline trained model to judge the preference orders. Extensive experiments on the real-world vertical and general community-based question answering datasets have comparatively demonstrated its robustness and promising performance. Also, we have released the codes and data to facilitate other researchers.

Keywords— Community-based Question Answering, Answer Selection, Observation-guided Training Set Construction.

I. INTRODUCTION

The Information Retrieval (IR) systems have been applied successfully to a large scale of search area in which indexing and searching speed is important. Unfortunately, they return a large amount of documents that include indexing terms in a user's query. Hence, the user should carefully look over the whole text in order to find a short phrase that precisely answers his/her question. Question-answering (QA), an area of IR, is attracting more attention, as shown in the proceedings of AAAI and. A QA system

searches a large collection of texts, and filters out inadequate phrases or sentences within the texts. By using the QA system, a user can promptly approach to his/her answer phrases without troublesome tasks. However, most of the current QA systems have two problems as follows: It cannot correctly respond to all of the users' questions. It can answer the questions that are included in the pre-defined categories such as person, date, time, and etc. It requires more indexing or searching time than traditional IR systems do because it needs a deep linguistic knowledge such as syntactic or semantic roles of words. To solve the problems, we propose a QA system using a predictive answer indexer - MAYA. We can easily add new categories to MAYA by only supplementing domain dictionaries and rules. We do not have to revise the searching engine of MAYA because the indexer is designed as a separate component that extracts candidate answers. In addition, a user can promptly obtain answer phrases on retrieval time because MAYA indexes answer candidates in advance.

II. SYSTEM ARCHITECTURE

The following system architecture shows that the user browses his programming question for seeking exact answers about programming solutions. The user's question is get analysed from several websites which includes related answer but not the exact answer, and so the query gets generated accordingly in order to report users with exact answer along with semantic analysis and question classification for spotting and rectifying errors. The generated query retrieves web where answers from different website is been selected based on the ranking and review of users rated, among those ranking the most toppest and suggested ranking answer is been selected and provides the exact defined answer to the user.



International Journal of Emerging Technology and Advanced Engineering

Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue No., Current Month 2017)

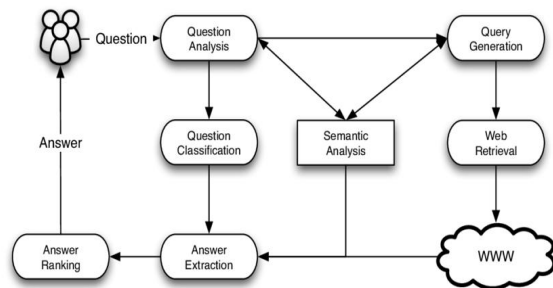


Figure 1 Architecture Diagram

III. BLOCK DIAGRAM

All The User must access the database. The user searches question in user search engine. The search engine analyses the users questions sequentially. The analysed question is sent to the database, where exact answer of the given question can be sought. Retrieves the best answer or exact answer from the database based on ranking provides best result to user.

IV. CONCLUSION

In this work, we present a novel scheme for answer selection in cQA settings. It comprises of an offline learning and an online search component. In the offline learning component, instead of time consuming and labor-intensive annotation, we automatically construct the positive, neutral, and negative training samples in the forms of preference pairs guided by our data-driven observations. We then propose a robust pairwise learning to rank model to incorporate these three types of training samples. In the online search component, for a given question, we first collect a pool of answer candidates via finding its similar questions. We then employ the offline learned model to rank the answer candidates via pairwise comparison. We have conducted extensive experiments to justify the effectiveness of our model on one general cQA dataset and one vertical cQA dataset. We can conclude the following points:

- 1) Our model can achieve better performance than several state-of-the-art answer selection baselines;
- 2) Our model is non-sensitive to its parameters;

3) Our model is robust to the noises caused by enlarging the number of returned similar questions;

4) The pairwise learning to rank models including our proposed PLANE are very sensitive to the error training samples.

V. FUTURE ENHANCEMENTS

Beyond the traditional pairwise learning to rank models, our model is able to incorporate the neutral training samples and select the discriminative features. It, however, also has the inherent disadvantages of the pairwise learning to rank family, such as noise-sensitive, large-scale preference pairs, and loss of information about the finer granularity in the relevance judgment. In the future, we plan to address such disadvantages in the field of cQA.

VI. REFERENCES

- [1] M. A. M. L. Wei Ding, He Jiang, Modern Advances in Intelligent Systems and Tools, ser. SCL. Springer, 2012, vol. 431. 1, 3
- [2] L. Nie, M. Wang, L. Zhang, S. Yan, B. Zhang, and T. S. Chua, "Disease inference from health-related questions via sparse deep learning," TKDE, vol. 27, no. 8, pp. 2107–2119, 2015. 1
- [3] A. Shtok, G. Dror, Y. Maarek, and I. Szpektor, "Learning from the past: Answering new questions with past answers," in Proceedings WWW'12. ACM, 2012, pp. 759–768. 1, 3
- [4] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne, "Finding high-quality content in social media," in Proceedings of WSDM'08. ACM, 2008, pp. 183–194. 2
- [5] J. Jeon, W. B. Croft, J. H. Lee, and S. Park, "A framework to predict the quality of answers with non-textual features," in Proceedings of SIGIR'06. ACM, 2006, pp. 228–235. 2
- [6] 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. 2

