ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

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

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

A Advanced Approach to Construct E-Learning QA System

Shubhangi Suresh Pawar, Prof.R.H.Kulkarni.

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

ABSTRACT: The Novel 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 copy 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 Pairwise 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: Answer Selection, Community-based Question Answering, Question-Answer pairs, Pairwise 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) cQA 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 cQA 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 cQA databases. Therefore, information seekers have a good chance of getting direct answers looking for repositories, instead of waiting long. Inspired by this, 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 are 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 Pairwise Learning to Run 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.

II. MOTIVATION

The main motivation is to overcome the problem of to find the similar questions, Because each question in the returned candidate pool consist with multiple answers, and hence users get trouble to browse a lot before finding the correct one. So we motivate to construct a novel approach a novel Pairwise Learning to rANk model i.e PLANE which can quantitatively rank answer candidates from the relevant question pool.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

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

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

III. REVIEW OF LITERATURE

1. Duyu Tang Author proposed a joint segmentation and classification framework for sentence-level sentiment classification. It is widely recognized that phrasal information is crucial for sentiment classification. However, existing sentiment classification algorithms typically split a sentence as a word sequence, which does not effectively handle the inconsistent sentiment polarity between a phrase and the words it contains, such as {"*not bad*," "*bad*"} and {"*a great deal of*," "*great*"}. We address this issue by developing a joint framework for sentence-level sentiment classification. It simultaneously generates useful segmentations and predicts sentence-level polarity based on the segmentation results. Specifically, we develop a candidate generation model to produce segmentation candidates of a sentence; a segmentation ranking model to score the usefulness of a segmentation candidate for sentiment classification; and a classification model for predicting the sentiment polarity of a segmentation [8].
2. Q. Le and T. Mikolov author proposed a three level scheme, which aims to generate a query-focused summary-style answer in terms of two factors, i.e., novelty and redundancy. Specifically, we first retrieve a set of Qas to the given query, and then develop a smoothed Naïve Baye's model to identify the topics of answers, by exploiting their associated category information[9].

3. X. Wei, H. Huang Author proposed in generating a vote, a user's attention is influenced by the answer position and appearance, in addition to real answer quality. Previously, these biases are ignored. As a result, the top answers obtained from this mechanism are not reliable, if the number of votes for the active question is not sufficient. The author solves this problem by analyzing two kinds of biases; position bias and appearance bias. To identify the existence of these biases and propose a joint click model for dealing with both of them [5].
4. Q. H. Tran ,Author designed Answer Selection in Community Question Answering. In this task, the systems are required to identify the good or potentially good answers from the answer thread in Community Question Answering collections. This system combines 16 features belong to 5 groups to predict answer quality. This final model achieves the best result in subtask A for English, both in accuracy and F1-score[6].
5. Savenkov ,Author represents how to automatically answer questions posted to Yahoo! Answers community question answering website in real-time. This system combines candidates extracted from answers to similar questions previously posted to Yahoo! Answers and web passages from documents retrieved using web search. The candidates are ranked by a trained linear model and the top candidate is returned as the final answer. The ranking model is trained on question and answer (Qn A) pairs from Yahoo! Answers archive using pair wise ranking criterion. Candidates are represented with a set of features, which includes statistics about candidate text, question term matches and retrieval scores, associations between question and candidate text terms and the score returned by a Long Short-Term Memory (LSTM) neural network model[7].
[6] W. Wei, Author developed a probabilistic method to jointly exploit three types of relations (i.e., follower relation, user-list relation, and list-list relation) for finding experts. Specifically, propose a Semi-Supervised Graph-based Ranking approach (SSGR) to offline calculate the global authority of users. In SSGR, employ a normalized Laplacian regularization method to jointly explore the three relations, which is subject to the supervised information derived from Twitter crowds. Then online compute the local relevant between users and the given query. By leveraging the global authority and local relevance of users, here rank all of users and find top-N users with highest ranking scores [3].
[7] W. Wei, B. Gao, Author addresses the large-scale graph-based ranking problem and focus on how to effectively exploit rich heterogeneous information of the graph to improve the ranking performance. Specifically, propose an technique and effective semi-supervised Page Rank (SSP) technique to parameterize the derived information within a unified semi-supervised learning framework (SSLF-GR), then simultaneously optimize the parameters and the ranking scores of graph nodes[4].

IV. SYSTEM OVERVIEW

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

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

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

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

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 gets recommendation based on his profile. 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.

SYSTEM ARCHITECTURE

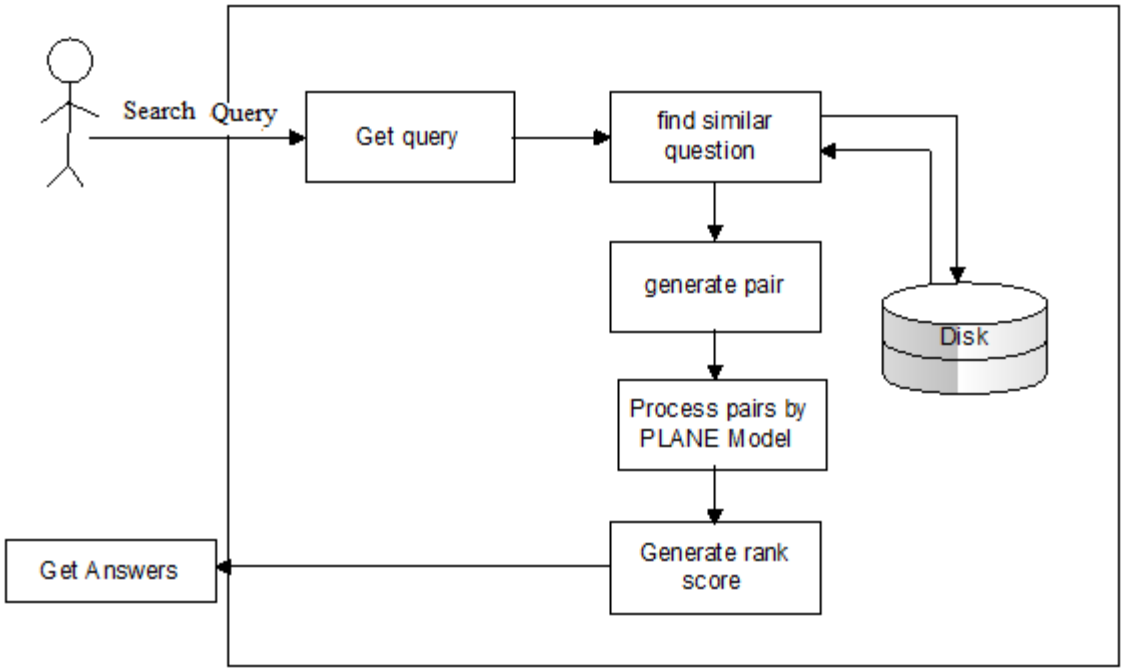

Fig. 01 System architecture

V. MATHEMATICAL MODEL

Notation

1.a_j^i =be the j th answer of i'th question q
2. a_i^0 =be the best answer
3.q=Entered question.
4.a1 be the votes of answer
5.C be the class of answer.
6. A_{11} =all similar question of q

Equation:

A_{11} =avg(feature (all matched question))---------- (1)

Eq.1 is obtained by synonym and Lavenstine distance algorithm. That match user entered question with all available database question. Finally it gives similar question of entered question .

a_i^0 =avg(a1,C)---------- (2)

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

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

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Eq.2 Here all review are passed to naïve baye's. Best answer by applying naïve bays'e and considering votes to that Answer.

VI. ALGORITHMS

Levenshtein distance algorithm.
This is used to get similar question for entered question
Input. : User entered question
Processing: 1. text string x = a1 for matching
2. Data From Available questions
3. Distance to match string x in available questions
System will asked question as string and will match that question available from database, according to specified distance.
Output: Get matched similar questions.

Algorithm 2: Naive Bayes
This algorithm is used to get review is positive or negative according to that we get relevancy in answer.
Input: Review
Processing:
Step 1: Take review
Step 2: Preprocess the review
Step 3: Pass to naive bayes class.
Step 4: Get positive and negative score according to specify its dictionary.
Step 5:Get max score and declare as positive or negative.
Output: Predicated class of review.

VIII.EXPERIMENT RESULT

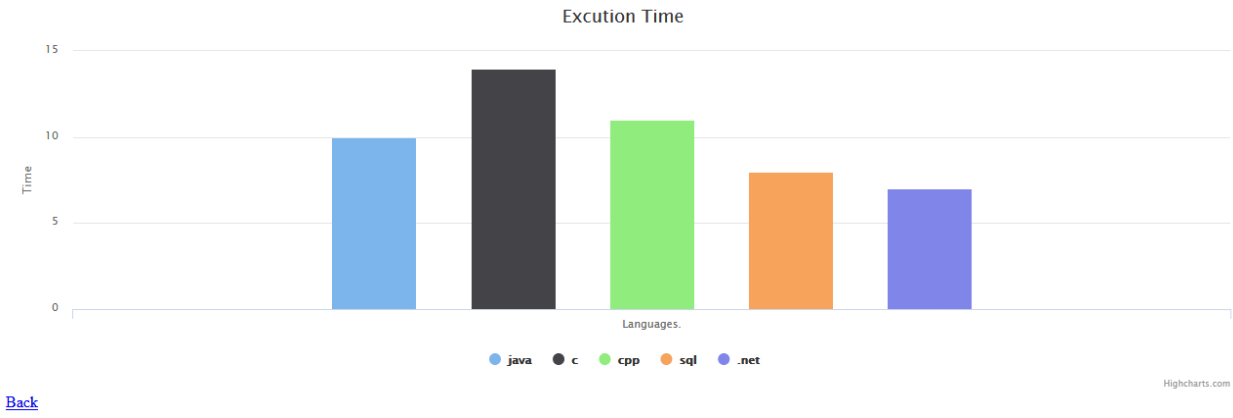

Fig2. Shows no. of search top recommended question of each language.

Explanation: The graph is trust based top recommendation analysis. That for each language top recommended question will display to user and that question will be new in that particular language.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

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

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Table: 1

language	Recommended question
java	10
c	14
cpp	11
sql	8
.net	7

Table1. shows no. of recommended que. for each language

IX. CONCLUSION

We present a system for selecting answers in the cqa system. It consists of an online search and online search component. In learning online components, instead of a job that takes a lot of time and effort, it automatically creates positive, neutral, pairs. In the online search component, a particular question is, first of all, gathering a group of answers to find candidates through their similar questions. So we use the offline model to classify candidate responses through pairs 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]A Joint Segmentation and Classification Framework for Sentence Level Sentiment Classification Duyu Tang, Bing Qin, Furu Wei, Li Dong, Ting Liu, and Ming Zhou. EEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 23, NO. 11, NOVEMBER 2015
[9] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proceedings of ICML'14*. Morgan Kaufmann Publishers Inc., 2014, pp. 1188–1196.
[10] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay, "Accurately interpreting clickthrough data as implicit feedback," in *Proceedings of SIGIR'05*. ACM, 2005, pp. 154–161.