JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Comprehensive Analysis for Quality Answer Selection in Community Question Answering

[1]Dr. Kamal Sutaria, [2] Ms. Kruti Khalpada,
[1,2] Assistant Professor,
[1,2] Department of Computer Engineering,
[1] V.V.P. Engineering College, Rajkot, India
[2] Atmiya University, Rajkot, India

Abstract : Answer selection is one of the popular tools and has gained an increased research interest in extracting the quality answers in Community Question Answers (CQA) community. The aim of the QA is to retrieve the best answers for the user's query from the large collection of documents from the database. The CQA services allow the users to post questions about their interest and other users share their opinions with their knowledge. Traditional approaches on deep learning allow the network to capture a part of the semantic relationship but ignore the rich feature learning which limits the quality answer prediction. To overcome the certain limitation in those models, we propose a new approach named multi-scale stacked dilated convolution (Ms-SDC) with the attention-based deep bidirectional long short term memory (Att-deep BiDLSTM) to estimate the quality answers for the input query. This model tackles the problems of: i) high-level or rich feature representation learning, ii) quality answer prediction from the set of candidate answers, iii) modeling the complicate relationship among question-answer and answer-answer.

***IndexTerms* - Answer selection, deep neural network, community question answering, answer ranking.**

I. Introduction

In recent years, Community Question Answering (CQA) services have gained much attention due to the rapid growth of question-answering websites such as Quora, Stack Exchange, Yahoo! Answers, etc... These sites offer a user to ask questions, share their problems, which are answered by other users or experts and discuss their opinions with the entire world without any restrictions [Liu et al. 2011; Palanisamy and Foshay 2011]. The main objective of the answer selection in CQA technique is to obtain the suitable and an accurate answer for the input question of a user from the large collection of cQA threads which save the user's time [3]. The answer selection task is helpful for many knowledge-oriented intelligent systems like chatbot [4] and automatic question answering [5]. In the healthcare applications, health expert QA services is an efficient and an on-demand online approach to attain the medical services from home. In these services, the answers are given by the certified physician and provide numerous health information in the healthcare sites. These platforms provide guidance and thereby meet the general needs of the health user without visiting a hospital. However, the answer quality is not guaranteed and may be irrelevant to the question posted by the health consumer [6-8].

Due to the massive amount of content posted by a large number of users, the selection of accurate answers is quite challenging in the CQA community. The stack exchange website allows the users to thumps up or down to vote the answer to estimate whether the answers are of good or bad quality. This voting system plays a key role in determining the value of the answer as well as the answerer. In this, the recent answers are placed in the bottom of the list and would not be voted even if the answer is of good quality [9, 10]. Hence, a new tab named promising answers is introduced where the answers are recorded based on their usefulness which is predicted using the machine learning algorithms. However, the best answers are extraction from the list of answers but the high-quality answer prediction is equally important. There are many traditional approaches followed to extract the best answers, but still suffers from several setbacks due to the following reasons: a) the response posted follows the informal writing and incomplete sentence, b) complex relationship among the posted answers among themselves.

For the QA selection, the approaches followed are categorized under two schemes such as the traditional and the deep learning models. The traditional methods include the search engine, handcrafted rules [11] and information retrieval (IR) [12, 13] approach. The IR techniques use the keywords to perform matching but it does not guarantee in the correct answer section while the handcrafted features are impossible to cover all the patterns result in limited performances [14]. Recently, deep learning models gained popularity and applied in many Natural Language Processing (NLP) task such as sentence modeling and classification [15, 16], sentence summarization [17] and paraphrase detection [18]. The main task is to learn the semantic connection between the question and answers and finally answer is selected based on the matching similarity score. The deep learning model extracts the high level semantic and temporal features based on the collaborative decision strategy and the pre-trained model [19, 20]. Then the quality of the answer is predicted by the factorization machine (FM) that encloses the combination of the support vector machine and factorization model [21]. Convolutional neural network (CNN) captures the semantic information from the input sentences and converted into high-dimensional feature vector via pooling and convolution layers. In [22], uses the CNN and Bag of Word (BOW) representation that computes input and related question and score is identified by cosine similarity. However, the CNN models perform the semantic matching between the QA pair and do not consider the correlation criterion among the answer sequences [23].

Recurrent Neural Network (RNN), which uses the LSTM network that reads the input sentences and determines the long and short term dependencies among the sequences [24]. The feed-forward RN network captures the non-textual semantic features in the sequential data and learns the fixed dimensional feature representation. Also, the RNN based learning model captures the features at the sequence-to-sequence level and further learns the paragraph representation vectors. The sequential dependencies are learned by the bidirectional LSTM model and the relatedness between the QA pair is calculated by the cosine similarity score estimation [25]. The 2D-CNN based sentence representation models followed by LSTM network is utilized that identify the answer content from QA pairs [26, 27]. Despite these advances, still there certain problems remain unsolved in the community question answering task. However, the existing models approaches achieve a limited performance due to the embedding representation is integrated with the single neural structure which fails to capture the complex and semantic relationship between the QA pairs. Thus, to achieve improved performance in the community question answering, the complex patters should be learned effectively as possible. In this paper, an effective text representation obtains the semantic information and syntactic relations and presented a Hybrid Ms-SDCNN with the Att-deep BiDLSTM network models to capture rich features for QA selection.

The contribution of this research paper includes:

- We proposed a multi-scale stacked dilated convolution neural network (Ms-SDCNN) that performs dilated convolutions which combines more contextual information and extract more features between the QA without the loss of important features.
- For the retrieval purpose of the high-quality answers, we use the attention-based deep bidirectional LSTM model that captures deeper features and estimate the correlation among the answers.
- The proposed model is more flexible than other networks in which it learns with more rich features and allows to focus on obtaining top-ranked answers from the candidate answer set.
- Experimental results conducted in terms of the metrics such as precision, recall, accuracy, and F1-score samples show that the proposed method performs well than several existing approaches.

The remainder of this work is modeled as: In section 2, we first reviewed the related works of the answer selection task in CQA services. In section 3, we presented our proposed methodology of our (Ms-SDC+ Att-deep BiDLSTM) network model in the retrieval of high-quality answer selection. In section 4, the experimental and analysis performed under the two dataset result and its comparison with other methods are discussed. Finally, the paper is concluded based on the evaluated results in section 5.

II. RELATED STUDY

Several techniques related to the community question answering task are performed earlier with the data collected from the social sites to predict the answer for the QA pair. The deep neural network system is modeled with the word pre-trained embedding to learn the semantic representation features and further estimate the suitable answer for the relevant query. Thus final result for the given QA is to determine the similarity or relatedness of the question-answer and recognize the relevant answer with high quality. This will help the user to gain knowledge of the query with highly ranked answers with a considerable amount of time.

Zhou et al. [28], proposed a Recurrent CNN integrates both the convolutional and the RNN that capture both the matching and the correlations between the question-answer pair. Initially, the matching is performed by the CNN network between question and answer into two distributional sentence vectors. In this, the two parallel convolutional networks are employed to determine the representation of the question-answer matching patterns. Then, the hidden layer generates the fixed-length feature representation for the input QA set. Next, the semantic correlation among answers is estimated by the RNN network. Finally, the softmax layer predicts the answer quality for the input questions.

Elalfy et al. [29], discussed a hybrid model for predicting the quality answer for the input query. The first module uses the content features that include three types such as question–answer, answer content and answer to answer features. In the content module, preprocessing is carried out to extract the tokens from the QA, feature extraction is processed where an additional content features are added, and then classification is performed to extract the best answer for the question. The second module estimates the score of user expertise level score and user confidence level score to compute the non- content reputation score. Finally, both the content and non- content features are integrated together for the prediction.

Hu et al. [30], described a collaborative decision CNN (CDCNN) that learns the semantic knowledge of the non-linear features based on the pre-trained word embedding. This model is classified under four stages which include the modules such as CNN, dependency sensitive CNN (DSCNN), collaborative decision, multimodal learning, and a prediction component. Initially, the semantic features are learned by CNN and the DSCNN for the given input. Then, the collaborative decision integrates the external knowledge to a mixed representation to extract the precision information. Next, the multi-model layer combines the temporal and the mixed semantic features to the joint representation. At the last stage, the quality score is obtained using Factorization machine (FM) that determines the non-independent relationship of various features under sparsity. However, the richer features are not captured for the QA pair which causes low quality answers prediction.

Roy et al. [31], proposed a regression and classification model to rank the answers based on their effectiveness and then listed it under the promising answers tabs. In this, the ranking of the answer is estimated by considering the factors such as text quality, answers reputation and activeness of the answer. At first, the collected files are filtered and then features extraction identifies the textual and non-textual features to train the network. Then data labeling is performed that assign the labels for the answers in which the answers have 0 vote are labeled as low quality and the higher score are labeled as high-quality answers. Here, the random sampling technique is employed to eliminate the excess data from the classes. Finally, classification is done by Navies, Bayes, gradient boosting and random forest and then prediction performed by the regression model.

Adlouni et al. [32], proposed a supervised latent semantic analysis (LSA) model for the Arabic language in CQA services. This method is processed in both supervised and unsupervised approaches where the input sentences are first tokenized based on word punket and tokenizer and stop words are removed. The preprocessed text are then combined with the embedding model and forwarded to the neural networks that predict the similarity score. Here three types of embedding models are utilized such as

BiGRu, DotNet, and pyrmidNet that measure the similarity score of the input query and answers. At last, ranking is performed that minimizes the ranking based on the query sentences. However, this method consumes more time and the list-wise approach are difficult that affects the group structure.

Nguyen et al. [33], proposed a deep neural model that uses two models such as convolution and bidirectional LSTM network for estimating the similarity and ranking the related and input QA pairs. Initially, the input sentences are tokenized and converted into vectors using word2vec model. Then the vectors are converted into question matrix where each row defines word vector representation for the input token. The CNN model consists of the convolutional and the poling layers that generate the vector representation for the input question and related question. Also, the bidirectional architecture sequentially processes the data in both directions and generate a single vector representation. Next, the concatenation layer obtains the combined vector value and finally fed to the fully connected multi-layer perception (MLP) for prediction. This MLP includes two hidden layers that calculate the similarity score for the QA pair. Although the combined features are utilized, it suffers from high computational complexity and time-consuming.

Zhang et al. [34], described a hybrid encoder model based on CNN and GRU for the Chinese medical answer selection. The neural architecture such as multi-stack-CNN, Bidirectional Gated Recurrent Unit (BiGRU)-CNN is used to determine the relationship between the QA pair. Here, the encoder model is utilized to represent the input sentences into a low dimensional vector to estimate the similarity. This hybrid model is classified under three modules include embedding, encoder, and cosine similarity module. In this, the embedding module converts the input sentences into a dense vector and then the encoder structure captures and learns the relationship between the query and answer. Finally, the cosine measure is computed in the answer set to estimate the best answer for the question. However, these approaches fail in the complex patterns are not captured leads to low answer quality prediction.

Previous research focuses on the answer selection from the QA pair without determining the high level synthetic and semantic relationship between the QA pair. This leads to the low answer retrieval without the deeper representation among the question-answer set. Most of the techniques encode the sematic correlation features that characterize the sentence leads to dissimilarity with final representation. The lexical gap between the candidate answer sentences and the query and also for the complex feature representation may lead to obtaining the irrelevant answers. In addition, answer prediction based on the deep learning target on the single aspect of matching classification among the QA pair. Thus, we develop a model, with multi-scale stacked dilated convolution with the attention-based deep bidirectional long short term memory approach to obtain the rich features from the input sentences. Finally, we perform classification that determines the highly ranked answers and produces an accurate answer for the query related to the reviews.

III. PROPOSED ANSWER SELECTION APPROACH-MULTI-SCALE STACKED DILATED CONVOLUTION WITH ATTENTION-BASED DEEP BIDIRECTIONAL LSTM NETWORK

The proposed Ms-SDC with Att- based deep BiDLSTM approach for the quality answer prediction for the QA pairs is categorized under four phases: a) Embedding b) feature learning, c) prediction, and d) classification. In the first stage, the input query and the answers are modeled into the embedding matrix by representing each word in the sentences into the dense vector. In this, we train our embedding with the word2vec tool as the input module for the sentence representation. Then, the embedding matrix is fed into the multi-scale stacked dilated convolutional neural network structure for learn the semantic relationship between the QA sentence pair. Finally, the attention-based deep bidirectional network performs the answer-answer prediction to determine the high-quality answer from the selected set of answers from the convolution network and the highly ranked answers for the input question are classified at the final softmax layer. Figure.1 shows the overall architecture model of the proposed answer selection and ranking for the input query in question-answer community. Initially, the word embedding process is processed followed by the feature learning by the convolution model. Finally, the answer prediction is carried out by estimating the correlation between the answer-answer answer retrieval performed with the attention-based stack bidirectional network.

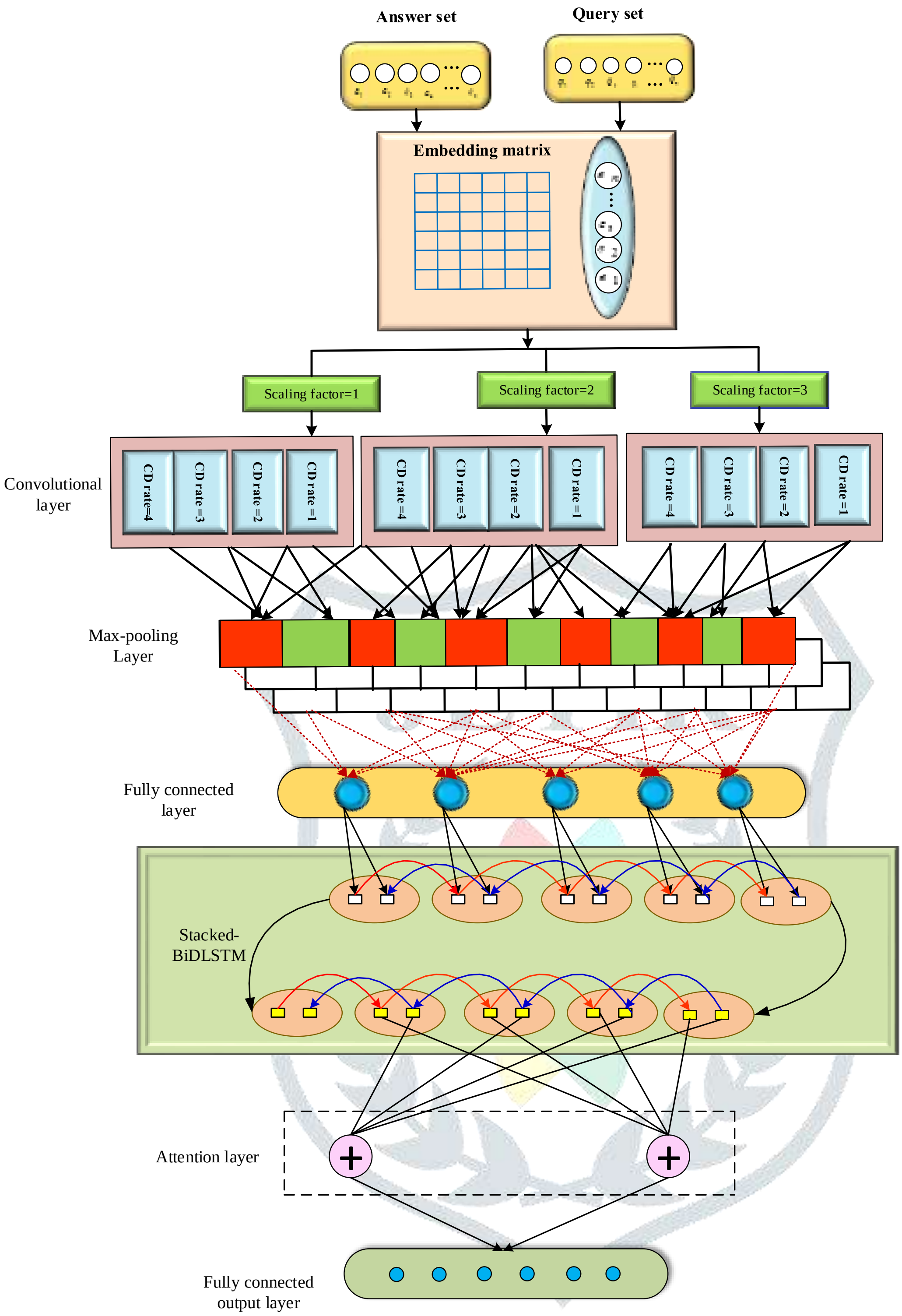

FIGURE. 1. PROCESS FLOW OF THE PROPOSED APPROACH

IV. CONCLUSION

In this research work, we presented Ms-SDCNN with Att-deep BiDLSTM task for retrieving an answer for the input query. This technique learns the complex and the high-level syntactic and semantic features among the QA pair that improves the performance of the answer selection. At first, word embedding applied over the input sentences is converted into the embedding matrix for representing each word into the dense vector. Then, the embedding vectors are fed into the convolutional architecture employs with the multi-scale module that converts the vector representation into three different scales. We will evaluate our proposed approach with the other methods in terms of accuracy, precision, recall, and F1-measures.

REFERENCES

[1] Liu, Q., Agichtein, E., Dror, G., Gabrilovich, E., Maarek, Y., Pelleg, D., Szpektor, I.: Predicting web searcher satisfaction with existing community-based answers. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, ACM, 415-424 (2011).

[2] Palanisamy, R., Foshay, N.: Impact of user's internal flexibility and participation on usage and information systems flexibility. Global Journal of Flexible Systems Management, 14(4), 195-209 (2013).

[3] Yu, L., Hermann, K. M., Blunsom, P., Pulman, S.: Deep learning for answer sentence selection. arXiv preprint arXiv:1412.1632 (2014).
[4] Huang, J., Zhou, M., Yang, D.: Extracting Chatbot Knowledge from Online Discussion Forums. In IJCAI, 7, 423-428 (2007).
[5] Cong, G., Wang, L., Lin, C. Y., Song, Y. I., Sun, Y.: Finding question-answer pairs from online forums. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, ACM, 467-474 (2008).
[6] Hu, Z., Zhang, Z., Yang, H., Chen, Q., Zuo, D.: A deep learning approach for predicting the quality of online health expert question-answering services. Journal of biomedical informatics, 71, 241-253 (2017).
[7] Beloborodov, A., Braslavski, P., Driker, M.: Towards automatic evaluation of health-related CQA data. In International Conference of the Cross-Language Evaluation Forum for European Languages. Springer, Cham, 7-18 (2014).
[8] Beloborodov, A., Kuznetsov, A., Braslavski, P.: Characterizing health-related community question answering. In European Conference on Information Retrieval. Springer, Berlin, Heidelberg, 680-683 (2013).
[9] Molino, P., Aiello, L. M., Lops, P.: Social question answering: Textual, user, and network features for best answer prediction. ACM Transactions on Information Systems (TOIS), 35(1), 4 (2016).
[10] Jain, S., Dodiya, T.: Rule based architecture for medical question answering system. In Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012, Springer, New Delhi, 1225-1233 (2014).
[11] Xu, X., He, L., Shimada, A., Taniguchi, R. I., Lu, H.: Learning unified binary codes for cross-modal retrieval via latent semantic hashing. Neurocomputing, 213, 191-203 (2016).
[12] Xu, X., Shen, F., Yang, Y., Shen, H. T., Li, X.: Learning discriminative binary codes for large-scale cross-modal retrieval. IEEE Transactions on Image Processing, 26(5), 2494-2507 (2017).
[13] Lu, W., Wu, H., Jian, P., Huang, Y., Huang, H.: An empirical study of classifier combination based word sense disambiguation. IEICE Transactions on Information and Systems, 101(1), 225-233 (2018).
[14] Wang, S., Cao, L.: Inferring implicit rules by learning explicit and hidden item dependency. IEEE Transactions on Systems, Man, and Cybernetics: Systems (2017).
[15] Zhang, Y., Wallace, B.: A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. arXiv preprint arXiv:1510.03820 (2015).
[16] Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188 (2014).
[17] Rush, A. M., Chopra, S., Weston, J.: A neural attention model for abstractive sentence summarization. arXiv preprint arXiv:1509.00685 (2015).
[18] Iyyer, M., Boyd-Graber, J., Claudino, L., Socher, R., Daumé III, H.: A neural network for factoid question answering over paragraphs. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 633-644 (2014).
[19] Iyyer, M., Enns, P., Boyd-Graber, J., Resnik, P.: Political ideology detection using recursive neural networks. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 1113-1122 (2014).
[20] Wang, P., Xu, B., Xu, J., Tian, G., Liu, C. L., Hao, H.: Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification. Neurocomputing, 174, 806-814 (2016).
[21] Rendle, S.: Factorization machines with libfm. ACM Transactions on Intelligent Systems and Technology (TIST), 3(3), 57 (2012).
[22] Dos Santos, C., Barbosa, L., Bogdanova, D., Zadrozny, B.: Learning hybrid representations to retrieve semantically equivalent questions. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), 694-699 (2015).
[23] Hu, B., Lu, Z., Li, H., Chen, Q.: Convolutional neural network architectures for matching natural language sentences. In Advances in neural information processing systems, 2042-2050 (2014).
[24] Wang, D., Nyberg, E.: A long short-term memory model for answer sentence selection in question answering. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers, 707-712 (2015).
[25] Tan, M., Santos, C. D., Xiang, B., Zhou, B.: Lstm-based deep learning models for non-factoid answer selection. arXiv preprint arXiv:1511.04108 (2015).
[26] Mohtarami, M., Belinkov, Y., Hsu, W. N., Zhang, Y., Lei, T., Bar, K., Glass, J.: SLS at SemEval-2016 Task 3: Neural-based approaches for ranking in community question answering. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 828-835 (2016).
[27] Zhou, X., Hu, B., Chen, Q., Tang, B., Wang, X.: Answer sequence learning with neural networks for answer selection in community question answering. arXiv preprint arXiv:1506.06490 (2015).
[28] Zhou, X., Hu, B., Chen, Q., Wang, X.: Recurrent convolutional neural network for answer selection in community question answering. Neurocomputing, 274, 8-18 (2018).
[29] Elalfy, D., Gad, W., Ismail, R.: A hybrid model to predict best answers in question answering communities. Egyptian informatics journal, 19(1), 21-31 (2018).
[30] Hu, Z., Zhang, Z., Yang, H., Chen, Q., Zhu, R., Zuo, D.: Predicting the quality of online health expert question-answering services with temporal features in a deep learning framework. Neurocomputing, 275, 2769-2782 (2018).
[31] Roy, P. K., Ahmad, Z., Singh, J. P., Alryalat, M. A. A., Rana, N. P., Dwivedi, Y. K.: Finding and ranking high-quality answers in community question answering sites. Global Journal of Flexible Systems Management, 19(1), 53-68 (2018).
[32] El Adlouni, Y., Rodríguez, H., Meknassi, M., El Alaoui, S. O., En-nahnahi, N.: A multi-approach to community question answering. Expert Systems with Applications, 137, 432-442 (2019).
[33] Nguyen, V. T., Le, A. C.: Deep Neural Network-Based Models for Ranking Question-Answering Pairs in Community Question Answering Systems. In International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making. Springer, Cham, 179-190 (2018).

[34] Zhang, Y., Lu, W., Ou, W., Zhang, G., Zhang, X., Cheng, J., Zhang, W.: Chinese medical question answer selection via hybrid models based on CNN and GRU. Multimedia Tools and Applications, 1-26 (2019).
[35] Wu, F., Duan, X., Xiao, J., Zhao, Z., Tang, S., Zhang, Y., Zhuang, Y.: Temporal interaction and causal influence in community-based question answering. IEEE Transactions on Knowledge and Data Engineering, 29(10), 2304-2317 (2017).
[36] Abujabal, A., Yahya, M., Riedewald, M., Weikum, G.: Automated template generation for question answering over knowledge graphs. In Proceedings of the 26th international conference on world wide web. International World Wide Web Conferences Steering Committee, 1191-1200 (2017).
[37] Liu, M., Fang, Y., Choulos, A. G., Park, D. H., Hu, X.: Product review summarization through question retrieval and diversification. Information Retrieval Journal, 20(6), 575-605 (2017).
[38] Limbasiya, N., Agrawal, P.: Semantic Textual Similarity and Factorization Machine Model for Retrieval of Question-Answering. In International Conference on Advances in Computing and Data Sciences. Springer, Singapore, 195-206 (2019).
[39] Filice, S., Croce, D., Moschitti, A., Basili, R.: Kelp at semeval-2016 task 3: Learning semantic relations between questions and answers. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 1116-1123 (2016).
[40] Barrón-Cedeno, A., Da San Martino, G., Joty, S., Moschitti, A., Al-Obaidli, F., Romeo, S., Uva, A.: Convkn at semeval-2016 task 3: Answer and question selection for question answering on arabic and english fora. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 896-903 (2016).
[41] Mohtarami, M., Belinkov, Y., Hsu, W. N., Zhang, Y., Lei, T., Bar, K., Glass, J.: SLS at SemEval-2016 Task 3: Neural-based approaches for ranking in community question answering. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 828-835 (2016).