

The Question Answer System Using Deep Learning

Vaishali Fulmali¹, Prof. K. P. Moholkar²

JSPM's Rajarshi Shahu College of Engineering, Pune, India

¹vaishali.fulmali@gmail.com, ²kpmoholkar_comp@jspmrscoe.edu.in

Abstract

Question-answer systems are intelligent systems that can be used to give answers to the user's questions. A question answering system can be designed that can automatically answer a question, in a similar way as a human can find answers to questions. Community question answering (CQA) services were becoming popular over the past few years. It permits the members of the community to post a question as well as answer the questions. CQA texts are noisy, therefore, the application of the traditional text mining methods does not give good results. In the proposed system, a deep learning approach is used for automatic question- answering. The questions and answers are embedded and the deep neural network is trained to find the similarity between questions. The most relevant answer to the user's question is provided based on the highest matching score of the user's question with the questions available in the dataset. The purpose of the proposed system is to design an automated question answering system. The proposed system uses hierarchical clustering to organize the dataset and Bi-LSTM algorithm to find the relevant question.

Keywords: Social networking, Community based question answering, Recurrent neural networks, LSTM, Deep Learning

1. Introduction

In the field of natural language processing, one of the key problems is to decide whether or not two sentences have nearly the same meaning. The main implementation of question similarity is the Question Answering system [1]. After receiving a question from the user, a QA system finds out a list of predictable questions from a huge dataset, leading to a few hundred ones. Then, it chooses almost similar questions set from the list with the help of a deep learning (BiLSTM) algorithm. The answers to existing questions are most likely the right answers to the real question. In such a system, the key matter is to decide the similarity between the questions. Answer selection is the task of giving an answer to the existing question which is most similar to the user's question. Recently, machine learning algorithms can resolve various complicated pieces of work in different fields of science and engineering. Deep learning is part of machine learning. Deep learning methods produce a good performance which is not relying on any feature engineering or expensive external resources. In the proposed system, a new technique for detecting similar questions to the question of a user using BiLSTM is used.

2. Literature Review

Ziye Zhu [11] proposes a community question answering system in which problems like how to

Make choice of finest answer out of multiple candidates for a sole question are considered.

The system is capable of searching and giving a satisfactory answer to users by similarity calculation. Experimental results show that the answer selection and the candidate question re-ranking methods provide huge accuracy as well as reliability. Data analysis shows that authors use the power of the user attribute information given by the response providers to extract the best answer for each question in the original Q-R database.

Xiang Cheng[5] proposes a question answering system in which the user who raised the question on CQA sometimes waits for several days. However, several answers are of low quality, so the user who asks the question cannot get satisfied. Question routing is used to solve such problems. It helps to route the new questions to the appropriate users who give answers. The question routing can enhance the possibility of answer and the quality of answer also of the routed users.

M. Breja[6] proposes the question answering system which can act just like social networking sites. Social networking analysis helps to investigate the features of structures of social circles.

G. Zhou[7] proposes the question-answer system which can provide a statistical machine translation technique to retrieve the questions in the system. In order to find the most relevant question and a new question to the users who wants to solve it, word ambiguity or mismatch are the main issues. The potentially rich semantic information drawn from other languages is used to find these issues. This work provides good results

Fei Wu [2] uses the temporal interactions between answers (how previously posted answers convey the later answers). Usually, any user first reads other posted answers, and then based on his or her knowledge, he or she gives a more accurate answer. In this way, the answer quality is improved, which is called temporal interactions. The causal influence between question-answer and answer is how suitable an answer is for a given question. A long short-term memory (LSTM) is used to find the question-answer influence and the answers-answer interactions to find the best answer. The experimental results show the advantage of the model comparing with other traditional methods.

N. Viriyadamrongkij [3] proposes the system to find the complexity of the question using the content of the question. Authors want to distinguish hard questions from easy ones. Concept hierarchy is used which is extracted from online as well as from books. Question Scope of the javascript related question measures the difficulty levels.

J. Liu [4] proposes a question-answer system in which, in order to find the quality of the question in community question answering systems, four factors are evaluated. These factors are asker, category, answer related, and question. Then by applying a CSMRLP algorithm, the prediction of question quality is done and classification is performed.

B. Ojokoh [10] proposes a standard method for recognition as well as classification of questions. At first, the identification of the questions is done using its semantic occurrence of a part of speech. Then classification is done using the Naive Bayes classification.

J. Wang [8] proposes the question-answer system which plays a vital role in the field of traditional medical systems. It can provide guidance to diagnose various diseases. This system uses an algorithm that provides an answer to any medical question present in the system. The dataset is formed from the online medical CQA systems. The algorithm tries to perform cases retrieval which are of a similar type, answer quality estimation and finally suggest the answer.

M.R.Sumalatha [9] proposes a system in which the prediction of an intelligent Contributor and their expertise is done. Based on this prediction the performance of question answer system is improved.

K.P.Moholkar [12] proposes a system that provides hybrid model that recognize the appropriate answer to the question asked. It takes advantage of convolutional neural network to extract features. Also the capability of LSTM can handle long term dependancies to retain the context of the question.

3. Proposed System

The questions are extracted from Quora dataset. Corresponding answers are extracted and customized dataset is prepared using best answers to each questions. The questions are clustered using hierarchical clustering. Bidirectional LSTM is used to retain the context of the questions. The output of BiLSTM is given to fully connected layer to classify similar type of questions. When a user ask a question, the system identifies the relevant domains and then check for the similar questions in the selected domain. The system extracts the answers which are most likely to the users question.

The dataset is clustered using hierarchical clustering. The training set is embedded in the system. The bidirectional LSTM is used to find the context of the question. The output of lstm is given to the dense layer for classification of similar type of questions. The figure 1 shows the training phase of the system.

The testing phase of system is shown in figure 2. The question is input to the cluster identification. After identification of cluster the trained model is loaded in the system. Similar questions are identified in the available questions and the answer is fetched and provided as output.

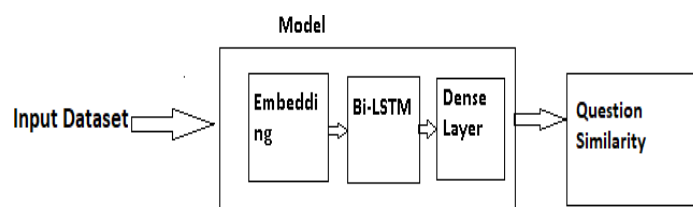


Fig. 1 Training the system

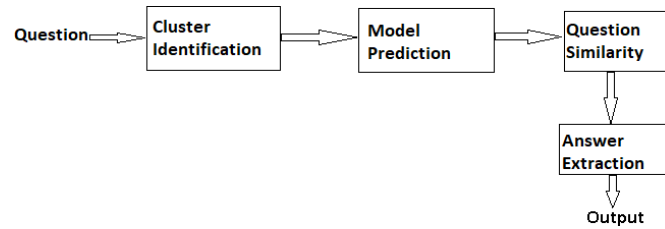


Fig. 2. Testing the system

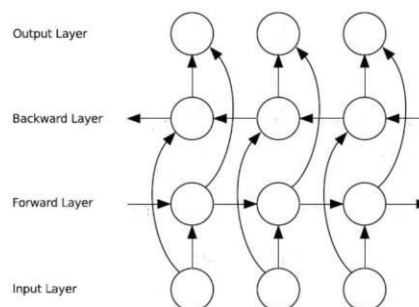
4. Motivation

The development of neural network which is trained to enhance the probability of prediction of a sentence. To provide a satisfactory answer to a user question.

5. Algorithm

RNN(recurrent neural network) reads the document from left to right and after every word updates the state. By the time we reach the end of the document the information obtained from first few words is completely lost. Ideally we want to forget the information added by stop words (a,the.etc). Selectively read the information added by previous sentiments bearing words (awesome ,amazing etc). Selectively write new information from current word to the state. This can be achieved in LSTM (Long short term memory) neural networks. LSTM are type of RNN but they can manage long term relationships as well. The Bi-LSTM are a type of LSTM in which the output is connected to previous state as well as next state. So we can predict based on future as well as past reference of data.

Fig 3. BiLSTM



The fig 3 shows the basic structure of Bi-LSTM network. The LSTM has three gates:

A. Forget gate layer:

This applied to the input at the current time step t and the hidden state at the previous step, i.e. $z(t) = [x_t, h_{t-1}]$. Since the output is a number between 0 and 1 for each element, it controls the amount of information to be retained from the previous time step $t-1$.

$$z(t) = (h_{t-1}, x_t) \quad (1)$$

$$f(t) = \sigma(W_f z_t) \quad (2)$$

B. Input gate layer:

It is similar to the forget gate. It controls which elements state vector C have to be updated.

$$I(t) = \sigma(W_i z_t) \quad (3)$$

With these functions, the state C is updated. In other words, the state at the time t depends on the state at the previous time $t-1$.

C. Output gate layer:

Finally, the hidden state at the time t is calculated, and output is given if t is also the final time (i.e. the last element of the input vector)

$$O_t = \sigma(W_o z_t) \quad (4)$$

$$h_t = O_t * \tanh(C_t) \quad (5)$$

5. Usefulness

This work gives satisfactory answers to the input question. The customized dataset is used in this system. Using Hierarchical clustering the searching process can be improved. The system helps to find the most similar question to the user's question to provide appropriate answer. This system can be useful in educational field, medical field, industries etc.

6. Result and Discussion

In this work, we have finished with training the dataset. The table shows the training dataset results. We have used Quora dataset which is applied to both LSTM and BiLSTM and the graph in figure 5 shows the number of similar questions recognised.

Table 1. Training Dataset Results

LSTM

Phase	No. of questions	Domain Recognized	Similar Question Found
1	100	100	29
2	100	100	32
3	100	100	35

BiLSTM

Phase	No. of questions	Domain Recognized	Similar Question Found
1	100	100	35
2	100	100	41
3	100	100	45

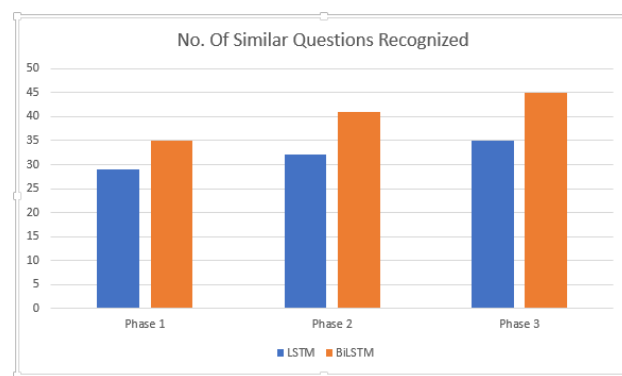


Fig 4. No. Of similar questions recognized

7. Conclusion

The proposed system can give appropriate and satisfactory results in finding the similarity of question. The system find most relevant question to the users question and thus helps in finding the appropriate answer to the users question.

Acknowledgments

I am deeply grateful to my project guide Prof.K.P. Moholkar for her guidance in the field of Deep Learning

REFERENCES

- [1] E. Karimi, B. Majidi and M. T. Manzuri, "Relevant Question Answering in Community Based Networks Using Deep LSTM Neural Networks," 2019 7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Bojnord, Iran, 2019, pp. 1-5.
- [2] F.Wu et al., "Temporal Interaction and Causal Influence in Community- Based Question

- Answering,” in IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 10, pp. 2304- 2317, 1 Oct. 2017.
- [3] N. Viriyadamrongkij and T. Senivongse, ”Measuring difficulty levels of JavaScript questions in Question-Answer Community based on concept hierarchy,” 2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhon Si Thammarat, 2017, pp. 1-6.
 - [4] J. Liu, H. Shen and L. Yu, ”Question Quality Analysis and Prediction in Community Question Answering Services with Coupled Mutual Reinforcement,” in IEEE Transactions on Services Computing, vol. 10, no. 2, pp. 286-301, 1 March-April 2017.
 - [5] X. Cheng, S. Zhu, S. Su and G. Chen, ”A Multi-Objective Optimization Approach for Question Routing in Community Question Answering Services (Extended Abstract),” 2018 IEEE 34th International Conference on Data Engineering (ICDE), Paris, 2018, pp. 1765-1766.
 - [6] M. Breja, ”Social network analysis in question answering community,” 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, 2017, pp. 314-318.
 - [7] G. Zhou, Z. Xie, T. He, J. Zhao and X. T. Hu, ”Learning the Multilingual Translation Representations for Question Retrieval in Community Question Answering via Non-Negative Matrix Factorization,” in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 7, pp. 1305-1314, July 2016.
 - [8] J. Wang, C. Man, Y. Zhao and F. Wang, ”An answer recommendation algorithm for medical community question answering systems,” 2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Beijing, 2016, pp. 139-144.
 - [9] M. R. Sumalatha and N. A. Priyanka, ”Analyzing and predicting knowledge of contributors in community question answering services,” 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, 2016, pp. 2473-2477.
 - [10] B. Ojokoh, T. Igbe, A. Araoye and F. Ameh, ”Question identification and classification on an academic question answering site,” 2016 IEEE/ACM Joint Conference on Digital Libraries (JCDL), Newark, NJ, 2016, pp. 223- 224.
 - [11] Z. Zhu, X. Liu, H. Li and T. Li, ”UB-CQA: A user attribute based community question answering system,” 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Nanjing, 2017.
 - [12] Moholkar, Kavita P. and Suhas Haribhau Patil. ”Hybrid CNN-LSTM Model for Answer Identification.” 2019.