IJWR

# *Expert Detection In Question Answer Communities*

Hamed Salimian* , MohammadAmin Fazli , Jafar Habibi
Computer Engineering Dept., Sharif University of Technology, Tehran, Iran; hamed@ce.sharif.edu*, fazli@sharif.edu, jhabibi@sharif.edu

***ABSTRACT***

**Community Question Answering has a crucial role in almost all societies nowadays. It is important for the owners of a community to be able to make it better and more reliable. One way to achieve this, is to find the users who have more knowledge, expertise, experience and skill and can well share their knowledge with others (which we call experts and aim to encourage them to be more active in the website). One method to use is to identify expert users, and whenever a new question is asked, we suggest this question to them to check and answer if its in their area of expertise. One way to encourage users to post replies, is to use gameplay techniques such as assigning points and badges to users. But as we will discuss, this method does not always detect expert users well, because some users will try to have small and insignificant but numerous activities that will make them gain a lot of points, however they are not experts. In this study, we examine the methods by which experts in a question-and-answer system can be found, and try to evaluate and compare these methods, use their ideas and positive points, and add our own new ideas to a new way of finding them. We used some ideas such as profile making for users, categorize users' expertise, A-Priori algorithm and showed that neural networks method results the best for the purpose of expert detection.**

***Keywords:*** ***Community Question Answering, Experts, Recommending System, Neural Networks, Machine Learning.***

## 1. Introduction

In recent years, countless Q&A (CQA) sites and forums have become available to the public on the Internet. Some of these systems have general applications and others are dedicated to a specific subject(s). Question and answer websites play a prominent role in our lives today. Users refer to these websites to find the answers of their questions, to ask questions, to share their knowledge with others, or at least to help the forum grow by sending positive and negative comments and votes. In these websites, there are usually a large number of membes, but there are not many active users among them. Among these active users, not all users can be considered as an expert user. In addition, there are some users who do not have many activities but have high expertise.

An expert user can be considered as a user who has a lot of expertise and skill (in one or more fields) and can share his or her knowledge with others. All CQA websites are based on knowledge sharing, which requires the expertise of experts, so it is necessary to discover these experts and somehow encourage them to participate more and more in thewebsite, for example [25], [26] and [27] have worked directly on the idea of identifying experts in a specific system and assigning tasks to them in a precise manner. But participation in some forums like stackOverflow requires specific knowledge and skills in a specific field, which makes the job more complicated. For this purpose, some forums assign score to users, which increases with every activity they do on the website. Some other websites give badges to users, and the number of badges in a profile indicates more expertise of the owner of that profile.

But does this score or badge always indicate the user's expertise? The answer is no [24][1]. In the following, by reviewing some of the previous works, we will try to investigate the answer to this question, check why this criterion is not appropriate, and finally we will seek to find other and more effective methods to find expert users.

The issue of finding experts in other areas of the software engineering is also an attractive issue, and many actions are taken in these areas as well. For example, [8] and [29] have worked on the idea of identifying experts from users' activity on GitHub.

## 2. 2. Previous Studies

In this section, we will first take a brief look at the history of finding experts, then we will check the algorithms and methods of finding experts, and finally we will check profile-based methods, which are one of the main ideas implemented in this article. At the end, we will have a brief look at the available methods of neural networks, machine learning and graph analysis.

### 2.1 The history of Finding an Expert

In the first study [1], it is stated that most of the existing methods for finding experts can be classified into two groups:

1) Credibility-based methods, which are based on link analysis of expert subject activities in the past.

2) Topic-based methods that are based on topic modeling techniques and latent meaning.

One of the models that is examined in this article is the MF-based method, which pays attention to the similarity of users, and by using this similarity, it can find users who can possibly have the answer to a question.

doi http://dx.doi.org/10.22133/ijwr.2022.342312.1118
Citation H. Salimian, M.A. Fazli, J. Habibi, " Expert Detection In Question Answer Communities," *International Journal of Web Research*, vol.5, no.1, pp. 25-32, 2022.

**Coressponding Author*
Article History: Received: 7 June 2022; Revised: 2 Augest 2002 ; Accepted: 6 August 2022

Finding experts is also widely used in other fields, such as [8], [10] and [12], but in this research, we only focus on finding experts in the space of question and answer forums.

### 2.2 Expert Finding Algorithms

In the first research of this section [28], gamification methods such as assigning credit to users and its effects, categorizing users, and finally comparing these categories are discussed. In this research, users are divided into two groups: sparrow and owl. Sparrow users are those who seek to increase their score and often cannot be counted on as a full-fledged expert. But the users who are in the group of owls are the users who have high knowledge and try to solve the important and basic problems of other users with their skills, expertise and experience. In the following, it will be checked that users with higher credit are not necessarily experts.

In another study [5], they analyzed the behavior of users with high and low credit in stackOverflow. In this research, graph analysis methods have been used to examine the relationships between users and the impact that each user has on the overall website, so that finally the users with the most influence can be identified. Using pagerank and Z-score criteria, many users were found who had high pagerank but not high score.

In the research [6], they have solved the problem of finding an expert by combining the learning of the question and answer community network structure and the semantic representation of the question with a ranking criterion called RMNL. Then, using a deep RNN network developed based on random walk, they generated rankings for users and questions in a Q&A forum network. In this research, which was published in 2020, Nobari and others tried to translate each skill area into a number of related words in order to improve the compatibility of these translated words with the questions asked and the answers given. For example, a word like java-ee is translated into the words http, session, request, controller and ejb. Their first translation model was called MI and their second model was called WE. The first model focused on common information and the second model on word embedding methods. Both of these models eventually proved to be helpful in finding experts.

In another study [7], published by Sumanth et al in 2018, they created a graph using data published by stackOverflow and using Python's SNAP library. The vertices of this graph which are directed are users, and each directed peer is from the user who sent a question to the user who sent the accepted answer. The important result that this research has for us is that some users have a high level of expertise in a specific field, such as the Python programming language, but they do not have a high level of score in the entire StackOverflow system. In other researches, the same idea of considering experts in each field separately has been implemented, like this research [9] that Zhao and others have worked on a model that can be used to find experts based on each topic. For this purpose they have invented a method called TEL that first creates the required data on the topics and Then it performs modeling to find experts.

Another study [29] checks the answers of users to more difficult questions and it is stated that users who answer more difficult questions have higher expertise and provide some criteria to find more difficult questions. There are specific works in the area of identifying difficult questions.

In [30], Lin uses the KGD-rank approach to identify difficult questions. In contrast to this probabilistic approach that requires extensive analysis of the user-user network, they follow a simple semi-supervised machine learning approach based on features, questions, and answers directly present in the data. In addition, the basis of their classification depends on the general characteristics of difficult questions. Liu proposed an approach to determine the effect of question quality on answer quality in Q&A community services [31]. However, apart from question quality, it also considers other characteristics such as the number of comments in labeling difficult questions.

Liu proposed an approach to determine the effect of question quality in determining answer quality in Q&A communities[11]. However, apart from question quality, it also considers other characteristics such as the number of comments in labeling difficult questions.

### 2.3 Expert Finding Methods

There are many researches such as [3] and [4] that have worked on the idea of creating a profile. Another research [13] has also focused on creating profiles for users and questions and then identifying experts using a recommender system based on those profiles. The input of these recommender systems is user characteristics and content created by users. The questions’ profile includes four categories of data:

1) Characteristics of textual data (such as the length of texts sent as questions and answers)

2) Non-textual features (metadata of questions and answers such as feedback received from other community members) and temporal data (when a question is posted and how long it takes for a user to respond to it)

3) Thread properties (such as information about question and answer threads and how the answers relate to each other and to the question)

4) Subject characteristics (such as some subject statistics)

In another research [13], they have worked on finding experts using the semantic matching of users' profiles. First, the difficulties of finding the similarities of the profiles of two users are pointed out, and then they provide a solution for this task. They used the semantic similarity between the two indexes to better discover the relationships between the concepts and words in the word bundles they used. In this article, they used a process called broadcast, which refers to a process that includes terms that are related to the main term in a user's profile.

In another research [32], social network analysis of Q&A forums and weighted HITS algorithm were used to identify the credibility of each user and finally identify experts. Yang et al. [18] present a topic-based machine learning expertise model that links common topics and expertise by integrating a textual content model and structure analysis.

There are also some methods based on completing the matrix for expert recommendation. In another study [19], user expertise is represented by labels. Users' expertise is translated into (user, tag) scores, which are highly dependent on the quality of the tags.

In addition, emerging deep learning models are integrated with the aforementioned methods to further improve the performance of expert finding methods[2]. These methods are capable of effective learning in high dimensions of specialized information, subject information and the interaction of specialized subjects.

Other researches such as [20] and [21] have also worked on the idea of using machine learning and have obtained good results. Some other researches such as [15], [16] and [17] have also used the idea of using graph analysis. In these researches, people are considered as graph nodes and their relationships are edges. In these researches, they have used clustering methods, hypergraphs, or their innovative ideas, all of which have had important results.

Another group of methods that can be used to determine the experts, is gamification methods such as giving scores to users. In the following, we will have a look at the researches that try to find experts by considering the score.

In the first study [13] that we examine in this category, it is stated that the methods that are based on score cannot distinguish experts well. Because one of the ways users get credits is by answering questions. Now, if a user is familiar with many concepts (even if he or she is not an expert in any of them) because he or she sends a lot of answers, he or she will probably have high scores, if a person is an expert in only a few limited fields and because he or she cannot send many answers, So he or she will probably have a lowe rscore.

### 2.4 Expert Recommendation Methods

In this study [22] they developed a binary classifier that combined profile-based features (such as time since user registration or number of followers, etc.) with features of posts. Finally, by applying this model to the data, they concluded that users who have joined the website recently and users who are active are more interested in answering the questions that have been proposed.

In another article [23], Dike and others have studied the results of classification algorithms. In this research, it was found that among the algorithms, random forest performed better than Guassian Naïve Bayes in terms of F1-score criteria. Also, this algorithm has performed better than Linear Support Vector Classification.

To evaluate, for each question whose expert answerers are known, they run the desired algorithms and methods and derive a list of N suggested users who might answer the question. Then, for each of the users who answered and were in this list, they added one unit to a variable and finally took the average of this variable. The result of this research was that the LDA method performed better than the TF-IDF method and the linguistic model in finding the best experts for a question.

## 3. Problem Explanation

As discussed in the introduction section, finding experts in question and answer websites is very important. For this purpose and considering that most of the previous research studies have not considered all the aspects, we decided to present several methods and compare them, to find an optimal and more efficient method by considering many aspects resulting from a compilation of previous more important works.

In the previous sections, we have seen that each of the previous works have strengths and weaknesses, and we will try to use the combination of these positive points. We also try to improve the logic by creating our own personal scoring method and using some new ideas such as checking pairs of frequent tags. In the next section we will discuss the new ideas we have used.

The final goal of this research will be to find expert users who may be able to answer that question with a high probability for new questions or questions that have been sent for a while but have not received a suitable answer.

## 4. Proposed Solution

We have used the data set of the StackOverflo, which is available on Kaggle. As seen in the previous works section, one of the successful approaches to find experts is to create a profile for users and questions. In this research, we will use the same successful approach.

On the other hand, in some other reliable articles, we saw that expert users are not necessarily considered experts in the entire website and are considered expert in only few areas and not in the entire website. For this purpose, in our implementations, we have considered the labels placed on the questions as a measure of expertise for users. Of course, due to the large number of tags in the StackOverflow website, it is necessary to perform the processes mentioned further.

The time since the last activity of each user is also one of the criteria that has not been focused on in most of the previous works, and only some articles that happened to have relatively acceptable outputs have focused on the time of the user's last activity. In order to check how much expertise each user has for each specific question, we have considered the last time the user was active in the system. The reason for this is that a user may be an expert based on all the criteria, but because he or she has not been active in the system for months, so we cannot count on his or her possible answer to the question, and for this purpose, he or she should be removed from the list of possible experts.

The time elapsed since the user registered in the system is another item that should be considered. As seen in previous works, users who have recently joined the system are more willing to answer questions and create interaction. So, taking into account a number of restrictions (for example, the minimum number of interactions or the minimum number of correct answers sent), new users who do not have much activity in the system can also be considered in the possible list of experts.

Not fully relying on the scores set by the website is another point that we have considered. We have already seen that there are users who answer many simple questions or ask simple questions that only lead to an increase in their interaction, which makes the person get a high score in the website. This means that users who have a higher score in the website are probably more active users, but they cannot necessarily be considered experts.

Creating our own personal validation approach alongside using StackOverflow's validation system is our ultimate solution for user validation. It is true that we said that the validation procedure of the system itself may not directly lead to the discovery of expert users, but it cannot be ignored anyway. But to overcome the weaknesses of this procedure, we have also added our own validation procedure. In this section, after implementing some ideas, finally the best idea we came up with was to combine four topics: 1. Create a profile for users 2. Tag-centric the overall idea 3. Create a new scoring procedure 4. In Considering the difficulty of the questions. For this purpose, every user who creates an interaction on a question (including adding a comment, adding an answer and choosing an answer as the selected answer) will add a numerical weight to the desired tags on the question on his or her profile. This weight is the product of the importance of the activity (view less than the answer and answer less than the chosen answer) in the difficulty of the question. As we said in the second part, the difficulty of the question is considered with criteria such as the number of answers sent, the number of views sent, the ratio of the question score to the number of question views, the time between the time of sending the question and sending the selected answer, etc.

In some previous works, it was observed that some researchers tried to sample a part of the data for various reasons. In this research, we also implemented a sampling on the questions we have chosen. There were about 20 million questions in the data set, and we considered only the questions from 2018 onwards for two reasons:

This data includes about five million rows of data, which is a good fraction of the total data.

Some programming languages or general topics about which questions are asked have been discussed more in recent years (like Python vs. Java) and also some tags have been used more in more distant years and are now in great decline (such as assembly questions). Given that we wanted to recognize expert on new data, we found it necessary to remove older data.

Another sampling we did was on labels. In the dataset we worked on, there were more than 58,000 distinct tags that were used of about 58 million times. In the data of 2018 and later, this number was about 14.4 million. In order to reduce the size of the problem and reduce the processing, we separated the 100 most used tags from other tags and focused only on them. These 100 tags were used together about seven million times, which is about half of all uses, which isn't a bad approximation. It also shows that the other tags had very few uses (that is, exactly 57,653 tags had about seven million uses against only one hundred tags that had the same number of uses). In the selected tags, python, javascript, and java tags were the most repeated, respectively, with about 600 thousand, 520 thousand, and 355 thousand uses. The least used tags in this list is unit-testing with about 17 thousand uses. The distribution chart of the number of labels can be seen in Figure 1.

Modifying the profiles is also the last action we performed on the profiles. For this purpose, we checked which pairs of tags have many repetitions. This was done with the A-Priori algorithm. This idea was adopted in order that, for example, someone who is an expert in the field of

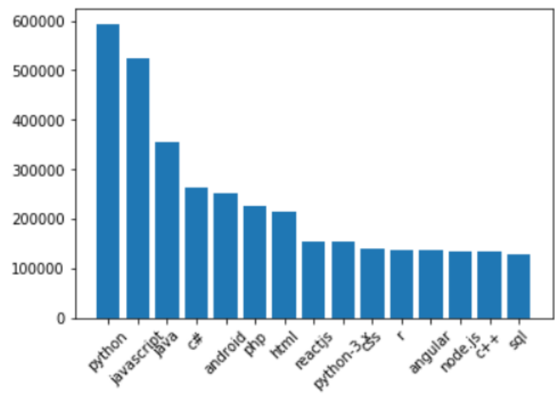


Figure. 1. Number of frequently used tags

nodeJS should undoubtedly be an expert in the field of javascript. Now, a question may be asked that has a nodeJS tag but not a javascript tag. At this stage, we have gone one step further and, in addition to directly using the used tags, we have also addressed the possibility of mastering the remaining tags. This case was another new idea that we used in this research.

Due to the combination of methods, one of the articles cannot be used as a criterion. For this purpose, we present several methods and compare these methods. One of these methods is the simplest possible method, which means that it is completely random, logically, all the presented methods should have better performance than this method.

The way of doing the work is that we try to include the information of the question and each user in our model and expect a number between zero and one from the model. The closer this number is to one, the more likely it is that the user will answer that question, and on the contrary, the closer this number is to zero, the more likely the user will not be able to answer that question.

We do not consider the entire user profile to find expert users. Since we have seen in previous studies that a user cannot be an expert in a large number of fields, for each user, we consider five tags that include the most points and ignore the rest of the tags.

We also considered the method of choosing the optimal answer in such a way that we know that the answer that is chosen as the selected answer is not necessarily the correct or the best answer and there may be a better answer. It is also possible that an unselected answer is the correct answer. For this purpose, if an answer has more upvotes than the question, there are more than five views on it, it is selected as the correct answer, the ratio of upvotes to the number of views is high, it is considered as an answer from an expert. So several answers may be considered expert answers in one question.

For this purpose, each comment that the user sends adds one point, the answer he or she sends adds three point, and each correct answer he or she has sent adds five points to the profile of the user. Then we multiply this score by a coefficient, which is calculated as follows: the ratio of points to observations with a weight of 0.5, the number of views (up to 10 items) with a weight of 0.2, the length of time that has passed since the question (up to 72 hours) with a coefficient of 0.1 and answer points up to 2500 with a coefficient of 0.2.

On the other hand, for each question, we considered the level of difficulty as a number between zero and one. The time interval between sending the question and sending the answer that is later selected as the correct one, the number of answers (at least three), the number of comments on the question (at least three), the upvotes of the question alone (up to 2500, the highest upvote among all Questions are from 2018 onwards) and the ratio of question views to question upvotes are all our criteria for scoring question difficulty. Of course, in this method, we first assume that all the questions are completely simple (that is, with zero difficulty) and then, for each of the mentioned criteria, we add the level of difficulty if necessary. Then we multiply this difficulty coefficient by the score wementioned at the beginning of this paragraph and add this amount to all the question tags in the user profile. This creates our personalized profile for the user. For example, if a question has a difficulty of 0.75 and the usefulness of an answer is 0.5, we add 0.5×0.75 to the value of all tags in the sender profile of that answer. It should be noted that if the difficulty level of the question or the usefulness of the answer may sometimes become negative (for example, an answer or question that has a lot of negative points), then we consider it to be zero if these two numbers are always between zero and one. .

In the following, for example, we see a part of the profile of one of the users:

*{'python': 44445.2994120573,*

*'javascript': 40785.373228161494,*

*'java': 24850.98611114744,*

*'c#': 20163.131018360233,*

*'android': 17469.807392388288,*

*'php': 12274.59503363403,*

*'html': 13432.882015758008,*

*'reactjs': 12154.339674976654,*

*'python-3.x': 10858.816594901491,*

*'css': 9782.591122652624,*

*'r': 10856.262039384912,*

*'angular': 12430.032216431942,*

*'node.js': 7853.19516844874,*

*'c++': 18396.492872628103,*

*'sql': 9905.570381176232,*

*'jquery': 7010.051779561444}*

So far and with this profile creation method, we have considered the user's expertise in each field separately, and we have included the difficulty of the questions. These two cases are among the main principles that we have used in this research, and the second case was a new idea that is difficult to pay attention to with these details.

### 4.1 Cosine distance

The first method we have used is the similarity of the profile of the user who has asked a question with the user under investigation. We can hope that if two users have similar expertise, they will probably be able to answer each other's questions. For this purpose, we consider each user's profile as a 100-point vector of the points we have given him or her based on his or her performance so far, and to calculate the similarity of the profiles, we include the cosine distance of the 100-point vector of both profiles. This method managed to correctly identify 48% of the correct answers.

In the following, we didn't just check the accuracy of this method with correct answers, we checked this method on all the answers. In the previous method, we could only discuss True Positives and False Negatives, but to check the efficiency, we need to check False Positives and True Negatives as well. For this purpose, we added two columns to the data set, which were the difficulty of the question and the usefulness of the answer. Then, we considered any answer that had a usefulness greater than 0.5 as an expert answer with a label of one, and otherwise, we considered that answer as a non-expert answer with a label of zero. Then we again applied the cosine distance method of the profile vector of the user who owns the question and the user who owns the answer. In this case, about two million TN, 1.8 million FN, 600 thousand FP and 840 thousand TP were created.

### 4.2 machine learning methods

As the next method, we used machine learning methods. Three methods will be examined in this section, which we will be discussed further.

1) Linear regression

The first method we implemented in this section was the linear regression method. In this method, which we avoid detailing, we considered our data set as an input and rendered the output of the usefulness of the question and considered it as an output.

We did not consider some features of the dataset as input, for example, the answer score or the number of answer views. Because our final model is supposed to predict regardless of the answer and only based on the characteristics of the questioner and the respondent.

In this method, we considered the test set as 25% of the training set. In total, in this method, about 650 thousand TN, 46 thousand FN and 620 thousand TP were created on the test set, and no FP samples were observed. The ROC diagram of this method can be seen in Figure 2.

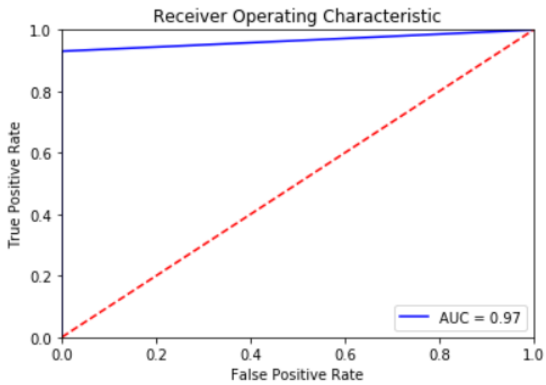


Figure. 2. ROC of linear regression

2) Logistic regression

The second method we implemented in machine learning methods, was the logistic regression. As we will see below, this method has worked better than the previous one. Here too, we considered the training data as 75% of the data and the test data as 25% of the data. In total, in this method, about 650 thousand TN and 667 thousand TP were created on the test set, and no FP and FN samples were observed.

Due to the absence of any FP or FN in this method, the area under the ROC diagram will be equal to one, and the display of the diagram is irrelevant, and for this purpose, we will not draw it.

3) Neural Networks

In this method, several different architectures of neural networks were investigated, and finally a deep neural network with the following architecture provided the best answer:

a) The first layer with eight neurons and ReLU activation function

b) The second layer with 16 neurons and ReLU activator function

c) The third layer with 16 neurons and ReLU activator function

d) The fourth layer with eight neurons and ReLU activator function

e) The fifth layer with single neuron for final prediction with sigmoid activation function

Also, dropout of 0.2 was considered after each layer to prevent overfitting.

Neural networks with about 614,000 TN, 38,000 FN, 39,000 FP and 629,000 TP provided relatively acceptable results. Of course, the area under the ROC diagram for this method shows the number 0.98, which is a very good number and can be seen in Figure 3.

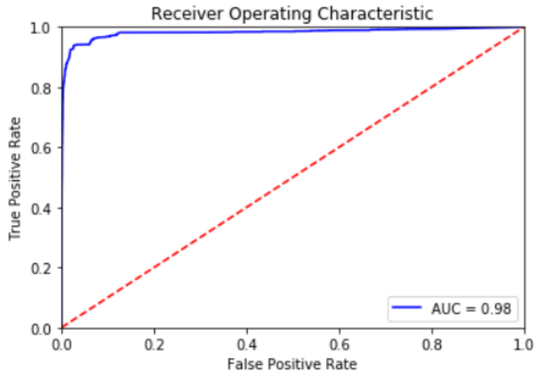


Figure. 3. ROC of Neural Networks

## 5. Conclusion

In this research, we first discussed the reason for the importance of expert users. Then we reviewed some of the previous activities that focused on finding experts or at least checking expert users and stated their strengths and weaknesses. Then we combined the previous ideas with new ideas of our own, such as creating a personal profile, creating new scores for each user, checking users' fields of interests, taking temporal characteristics into account, using methods such as A-priori to refine the profile, etc., and then we tried to create different models with four methods of cosine distance, linear regression, logistic regression and neural networks. Since our method had many differences with the basic methods from which we got ideas and even our final approach and main goal was not exactly aligned with many of them, we inevitably compared our methods with each other to validate that in the end the method Logistic regression gave the best result, followed by the neural network method.

## 6. Threats To Validity

Several cases in this research can be considered as a threat to validity. For example, we separated the data from 2018 onwards, which, of course, we had enough reasoning for this, but there are some programming languages or some areas discussed in StackOverflow that have been hot topics for many years (such as Java), although it is possible Their popularity may have been decreased a bit in recent years, but they are still worth checking out.

Also, we removed the tags that had few repetitions. But maybe users who are experts in those areas have more expertise than users who are experts in hotter areas. The next point is that if a question or answer gets a negative score in our procedures, we consider it zero, but a question that is negative is really different from a question that is zero, and this difference was not seen in our research.

## 7. Future Works

As one of the suggestions for future works, it is possible to research the data before 2018. Of course, we said that due to the many changes in technology and the trends of developers and other activists in the computer fields, we have removed the data from the years before 2018, but this review of how much the approaches and trends have changed since 2018 can be an interesting topic for further researches. Also, if some topics have been favored by everyone for a long time, our methods and ideas can be considered on that data set.

The next action that can be considered is to pay attention to the votes issued by the user himself or herself on the questions and answers of others. The fact that the user has given a good vote to appropriate and difficult questions or has added useful answers with his or her positive vote is a sign of the user's expertise and vice versa.

Another thing to consider is question or answer edits.Considerable results can be achieved by being precise in the discussion of editing posts, their connection with the views sent, the connection of the edits with the responses sent, etc. Also, if a user edits another post, it can also contain important points for us.

## Declarations

***Funding***

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

***Authors' contributions***

HS: Study design, acquisition of data, interpretation of the results, modeling, analysis, drafting the manuscript; MF:

Study design, Supervision, interpretation of the results, drafting the manuscript, revision of the manuscript; JH: Supervision, revision of the manuscript;

### *Conflict of interest*

The authors declare that there is no conflict of interest.

## References


[1] S. Yuan, Y. Zhang, J. Tang, W. Hall, and J. B. Cabotà, "Expert finding in community question answering: a review", *Artificial Intelligence Review*, vol. 53, no. 2, pp. 843-874, 2020.

[2] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative filtering and deep learning based recommendation system for cold start items", *Expert Systems with Applications*, vol. 69, pp. 29-39, 2017.

[3] Y. Tian, W. Ng, J. Cao, and S. McIntosh, "Geek talents: Who are the top experts on github and stack overflow?", *CMC-COMPUTERS MATERIALS & CONTINUA*, vol. 61, no.2, pp.465-479, 2019.

[4] N. Ghasemi, R. Fatourechi, and S. Momtazi, "User embedding for expert finding in community question answering", *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 15, no. 4, pp.1-16, 2021.

[5] Z. Zhao, Q. Yang, D. Cai, X. He, and Y. Zhuang, "Expert Finding for Community-Based Question Answering via", in *Ijcai*, vol 15, 2016, pp. 3000-3006.

[6] A. D. Nobari, M. Neshati, and S. S. Gharebagh, "Quality-aware skill translation models for expert finding on StackOverflow", *Information Systems*, vol. 87, p. 101413, 2020.

[7] P. Sumanth, and K. Rajeshwari, "Discovering Top Experts for Trending Domains on Stack Overflow", *Procedia computer science*, vol. 143, pp. 333-340, 2018.

[8] S. Panichella, G. Canfora, and, A. Di Sorbo, "Won't We Fix this Issue? Qualitative characterization and automated identification of wontfix issues on GitHub", *Information and Software Technology*, vol. 139, p. 106665, 2021.

[9] M. Bhanu, and J. Chandra, "Exploiting response patterns for identifying topical experts in stackoverflow", in *Eleventh International Conference on Digital Information Management (ICDIM)*, IEEE, 2016, pp. 139-144.

[10] J. A. Oliveira, M. Viggiato, D. Pinheiro, and E. Figueiredo, "Mining Experts from Source Code Analysis: An Empirical Evaluation", *Journal of Software Engineering Research and Development*, vol. 9, pp. 1-16, 2021.

[11] J. Liu, H. Shen, and L. Yu, "Question quality analysis and prediction in community question answering services with coupled mutual reinforcement", *IEEE Transactions on Services Computing,*, vol. 10, no. 2, pp. 286-301, 2015.

[12] A. Bozzon, M. Brambilla, S. Ceri, M. Silvestri, and G. Vesci, 'March. Choosing the right crowd: expert finding in social networks', In *Proceedings of the 16th international conference on extending database technology*, 2013, pp. 637-648.

[13] I. Srba, M. Grznar, and M. Bielikova, "Utilizing non-qa data to improve questions routing for users with low qa activity in cqa", in *he 2015 IEEE/ACM international conference on advances in social networks analysis and mining 2015*, 2015, pp. 129-136.

[14] J. Wang, J. Sun, H. Lin, H. Dong, and S. Zhang, "Convolutional neural networks for expert recommendation in community question answering", *Science China Information Sciences*, vol. 60, no. 11, pp. 102-110, 2017.

[15] P. Cifariello, P. Ferragina, and M. Ponza, "Wiser: A semantic approach for expert finding in academia based on entity linking", *Information Systems*, vol. 82, pp.1-16, 2019.

[16] F. Amato, G. Cozzolino, and G. Sperlì, "A hypergraph data model for expert-finding in multimedia social networks", *Information*, vol. 10, no. 6, p.183, 2019.

[17] N. Zhao, J. Cheng, N. Chen, F. Xiong, and P. Cheng, "A Novel Expert Finding System for Community Question Answering", *Complexity*, 2020.

[18] B. Yang, and S. Manandhar, "Tag-based expert recommendation in community question answering", in *EEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, IEEE, 2014, pp. 960-963.

[19] Z. Zhao, L. Zhang, X. He, W. Ng, "Expert finding for question answering via graph regularized matrix completio", *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 4, pp. 993-1004, 2014.

[20] S. Yuan, Y. Zhang, J. Tang, W. Hall, and J. B. Cabotà, "Expert finding in community question answering: a review", *Artificial Intelligence Review*, vol. 53, no. 2, pp.843-874, 2020.

[21] M. Miri, and H. Beigy, "Expert Finding in Community Question Answering", $6^{th}$ *Iran Web conf*, 2020.

[22] Z. Liu, and B. J. Jansen, "Predicting potential responders in social Q&A based on non-QA features", *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, 2014, pp. 2131-2136.

[23] D. van Dijk, M. Tsagkias, and M. de Rijke, "Early detection of topical expertise in community", in *the 38th international ACM SIGIR conference on research and development in information retrieval*, 2015, pp. 995-998.

[24] S. Wang, D. M. German, T. H. Chen, Y. Tian, and A. E. Hassan, "Is reputation on Stack Overflow always a good indicator for users' expertise? No!", n *2021 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, IEEE, 2021, pp. 614-618. IEEE

[25] M. S. Faisal, A. Daud, A. U. Akram, and R. A. Abbasi, N. R. Aljohani, and I. Mehmood, "Expert ranking techniques for online rated forums", *Computers in Human Behavior*, vol. 100, p. 168-176, 2019.

[26] J. Lock, and P. Redmond, "Embedded experts in online collaborative learning: A case study", *The Internet and Higher Education*, vol. 48, p. 100773, 2021.

[27] P. K. Roy, J. P. Singh, and A. Nag, "Finding Active Expert Users for Question Routing in Community Question Answering Sites", *Machine Learning and Data Mining in Pattern Recognition*, Springer, 2018, 440-451.

[28] D. Movshovitz-Attias, Y. Movshovitz-Attias, P. Steenkiste, and C. Faloutsos, "Analysis of the reputation system and user contributions on a question answering website: Stackoverflow", in *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013*, IEEE, 2013, 886-893.

[29] C. L. Lin, Y. L. Chen, and H. Y. Kao, "Question difficulty evaluation by knowledge gap analysis in question answer communities", in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, IEEE, 2014, pp. 336-339.

[30] J. Liu, H. Shen, and L. Yu, "Question quality analysis and prediction in community question answering services with coupled mutual reinforcement", *IEEE Transactions on Services Computing,*, vol. 10, no. 2, pp. 286-301, 2015.

[31] R. Thiagarajan, G. Manjunath, and M. Stumptner, "Finding experts by semantic matching of user profiles", *Doctoral dissertation, CEUR-WS*, 2008.

[32] J. Shen, W. Shen, and X. Fan, "Recommending experts in Q&A communities by weighted HITS algorithm", *International Forum on Information Technology and Applications*, vol. 2, pp. 151-154, 2009.

[33] F. Riahi, Z. Zolaktaf, M. Shafiei, and E. Milios, "Finding expert users in community question answering", in *the 21st International Conference on World Wide Web*, 2012, pp. 791-798.


**Hamed Salimian** is a PhD student in computer engineering at Sharif University of Technology. He has received his M.Sc and B.Sc degree in computer engineering as well from Sharif and Kharazmi university. His area of interests include machine learning, data mining, financial markets analysis and software engineering.

**MohammadAmin Fazli** received his BSc in hardware engineering and MSc and PhD in software engineering from Sharif University of Technology, in 2009, 2011 and 2015 respectively. He is currently a faculty member at Sharif University of Technology and R&D head at Sharif's Intelligent Information Solutions Center. His research interests include Game Theory, Combinatorial Optimization, Computational Business and Economics, Graphs and Combinatorics, Complex networks and Dynamical Systems.

**Jafar Habibi**, Associate Professor at Computer Engineering Department of Sharif University of Technology, has received his PhD degree in computer science from Manchester University in 1998. His area of interests include: Software Engineering, Software Architecture and Evolution, Simulation and Performance Evaluation, Social and Complex Networks, Information Systems and Enterprise Architecture.