International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 7 64 - 68

# Multimedia Answering and Retrieval System based on CQA with Media Query Generation

Priyanka N.Patil
Computer Science and Engineering
Walchand Institute and Technology
Solapur, India
*Pnpatil775@gmail.com*

Prof.Dr R.B Kulkarni
Computer Science and Engineering
Walchand Institute and Technology
Solapur, India
*Raj_joy@yahoo.com*

***Abstract*—**the question answering system which has recently received an attention from the various information retrieval systems, machine learning, information extraction and the natural language processing the goal of the QAS is to retrieve the answer to the question than full documents. This question answering system which works on the various modules related only to the question processing, the document processing, and the answer processing. This QAS which doesn't work properly with the main module which is questioning processing this system fails to categorize properly the questions. So to overcome the QAS the Community question answering (CQA) has gained popularity. As compare to QAS and automated QA sites the CQA sites are more effective. In this drawback available for community question answering system is that it only provides the textual answer. Here in this paper, we propose a scheme that enhances the textual answer with the multimedia data. The outline of Community question answering which mainly consists of three components: the selection of answer medium, the query generation for multimedia search and the selection and presentation of multimedia data. This approach automatically defines which type of media information should be added for the textual answer. Then it automatically collects the data from the web to supplement the answer.by handling an available dataset of QA pairs and adding them to a pool, in this, our approach is to allow a new multimedia question answering (MMQA) approach so as the users can find the answer in multimedia matching the questions pair those in the pool. Therefore, the users can approach MMQA from Web information will answer the questions in different media formats (text, video, and image) as particularly selected by the users.

***Keywords-:*** *CQA; Re-ranking; Selection Medium; MMQA; Query Generation; Intent Research*

*****

## I. INTRODUCTION

Question Answering (QA) which is dependent on the information retrieval for building this system that automatically generates the answer for a given question which is posed by the users. This question answering System which deals with closed domain and Open domain question. The closed domain deals with the question under a specific domain (such as medicine, automotive, maintains etc.). And the question on general ontologies and the world knowledge related deals with the open domain [2]. But this QAS attempts to deal with a wide range of question types including the What, Why, how question, the factoid, list, definition, imaginary question, semantically constrained and cross-lingual question. In this, the Question Processing part may fail to categorize properly the question or the information needed for extracting and generating of answers is not easily retrieved. However, most of the fully automated QA search engines that depend on the keyword matching that usually return too many low-quality matches and also not easy to tackle complex questions semantic and contextual processing to generate the answer.

The community question answering (CQA) allows the people with different backgrounds to share their knowledge and experiences [1]. As the number of CQA are based on the online services like Yahoo! The Answer, Metafilter, Answer bag, Wiki answer as shown in figure [1]. This cQA system comes up with an alternative for the users where the users can get the answers provided by other participants. This system not only provides the answer to the user but also acts as a platform for users where they can share their answer, discuss their opinion and rate the answer. Moreover, it develops the flexibility to the user in order to Get the best answer and also allow the users to gain more question-answer pairs for preservation and recovery of answered questions in cQA repositories. For example, one of the most well-known CQA systems hosts more than 15 million answered questions distributed in 8,000 categories.

Therefore, the existing system has the disadvantage that the community question answering (CQA) medium only provides the textual answer which is not more informative for many questions Textual answer itself is not acceptable enough to the user for understanding and memorize the content for various questions. However, the users generally post their URL (Uniform Resource Locator) for linking the textual answers with the equivalent matching images and videos to enrich the community –contributed answer with multimedia content. This confirms the importance of multimedia content, On the other hand, the increasing growth of multimedia content such has image and video over the web has been observed the status of the multimedia search. For example, YouTube Works on more than 100 million distinct videos and 65,000 uploads daily, and the traffic flow of this site balances for than 20% of all web traffic and 10% of the whole internet, including 60% of video watched online.

The photo-sharing containing site Flickr contained supplementary 5 billion images. This clearly shows the importance of multimedia content .In this paper we propose a model that enriches the textual answers with the corresponding media data in cQA. It mainly compromises three Components.

a) The selection of answer medium - For a given pair of question answer it predicts which form of media data should be added. Here we will categorize mainly into four types they are: text, text+ image, text+ video, text+ image+ video. It means that this approach will decide which type of data should be added for enriching the Textual answers with the combination of images and video.

b) The query generation for multimedia search- Query generation is for examining the search for the multimedia content. In order to find the multimedia content, we need to find the informative word which helps to recover the corresponding content.

c) The selection and presentation of multimedia data- Based on the generated informative and useful word which we will search and vertically collect corresponding media data i.e. video and image with multimedia search engines. This Multimedia as a search engine will provide only media data. So, to obtain a set of accurate data to enrich textual answer with representative of image or video the Re-ranking and duplicate removal data method is used.


Figure 1. Examples of QA from several popular cQA forums.

## II. RELATED WORKS

### *2.1 Text QA to multimedia QA*

In 1960"s the QA system has been started to investigate and it mainly focused on an expert system in particular domains. Consequently, the text-based QA has attained popularity in the year 1990 "s. depending upon the type of questions and expected answers, QA can be summarized as following classes: open domain- Based QA [2], Restricted-domain QA [3], Definitional QA and List QA [1]. The automated QA system still faces some difficulties in answering complex questions to overcome this cQA is an alternate approach to solving it. The existing cQA like Yahoo! Answers, Wiki Answers, answer-bag, and meta- filter supports pure text-based answers only which are not sufficient for users. To overcome this problem multimedia search has been introduced which adds images and videos along with the text. Multimedia QA system depends on video optical character recognition (VOCR) and automatic speech recognition (ASR).

### *2.2 Text QA to Multimedia CQA*

Further to overcome QA to CQA .CQA system has gained more popularity, in this system the more contribution is of users where multiple question can be asked and the askers need to wait for the post answer .so this textual answer can be enriched by using text +image+video. In this, varying as it depends upon on answer contributed by users and the Best answer selected by the asker positive and negative rating given by voters which is carried out in CQA in the generation of the Multimedia technique[1] in 2013"s .

### *2.3 The Multimedia Search*

A large amount of digital information stored on the web has been grown in enormously nowadays; hence extracting the desired information is an important task for the user to get the proper information. Therefore to overcome this problem multimedia search work is introduced which is classified into two categories: text-based search and content-based search [11]. Text-based search is based on the text queries and term based specification, where it matches the text with media on the web to restore the media data and in order to improve the performance, some machine learning methods are used which automatically explain the entities for gathering information. Various social websites like Flicker, Facebook uses the manually annotated media entities along with text-based search method which faces some technical issues. As a result to overcome this problem content-based search is used to filters the information by analyzing the content in the media instead of the metadata repository. The content-based retrieval has some restrictions such as high computational cost, difficulty in finding the visual queries and the large gap between the visual description and user's semantic Development. Multimedia search re-ranking algorithm is used for improving the search relevance by extracting the visual information of images and videos [4].

### *2.4 The Multimedia Re-ranking*

The Re-ranking algorithm has been considered into two techniques: pseudo relevance feedback [13] and graph based re-ranking [12].Pseudo-relevance feedback is used to bring together the relevant samples and that samples are expected to be irrelevant. The ranking model or classification is used for ranking the samples of data and it is responsible for response by labelling the results as relevant or irrelevant. The Graph classification technique is based on the two principles. First one is modification between the initial ranking list and processed ranking list should be small. The second one is visually similar samples should be ranked very closer. This algorithm builds a graph where the images or videos are representing the vertices and pairwise matches represent the edges. These two methods depend on the visual matches between two media entities. It should extend the similarity like color, texture, shape and so on; the almost query-adaptive system is used for guessing the person. Here we need to recognize theirs similarities of facial characteristics. Therefore Based on this identification of datasets, queries are classified into two classes namely person related query or non-person related query.

## III. SYSTEM ARCHITECTURE

The community question answering system is shown in following figure [2]. As cQA is a web-based system in this we provide the users ability to get the textual answer with appropriate media data. We provide the user to give a different multimedia question answering (MMQA) approach as the user can find the multimedia answer by matching their question those in the pool by matching their question in the pool in the form of image and video. For any complex question, we have provided the expert community to question the answer for the users. This cQA system gives the opportunities to Request information that we do not recognize, also, check information that we are not sure of, interact with people around us and see things from their viewpoint and they also help us to and fulfill our needs at the same time. In this system, the user needs to register first and then access the system by using the username and password.

1. Then the next users post their specific question on any topic and find answer provided by other participants. By leveraging these community efforts, are able to get the better answer than simply using the search engine.
2. Answer with better quality as generated based on human intelligence by the community contributed Experts
3. A large number of QA pair need to collect data in the repositories, and it facilities the preservation and search of answered question in MMQA datasets.

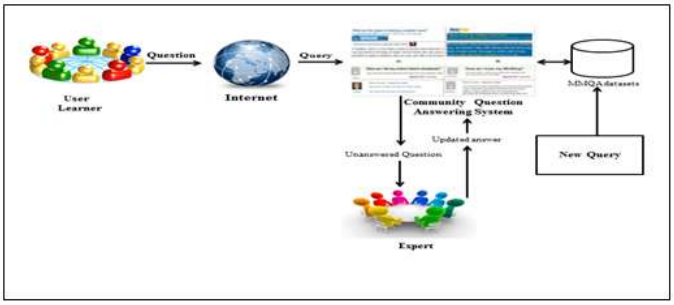


Figure 2.System Architecture.

## IV. PROPOSED FRAMEWORK

Figure [4] represents the framework of proposed system. For testing the collection of datasets files are required for the query generation .the query generation is carried with question-based classification, answer based classification and media analysis. The N-gram algorithm is used when the pre-processing of the given query selection from the database. Then the classification of a query is done by using KNN algorithm and Re-ranking algorithm to get the best answer from the history of the database. Then the medium of selection is carried in the form of text + image, text + video, text + image +video as shown in the figure [5]. For training datasets, Question and Answer separated from the datasets files and the text-based classifier algorithm are used for extraction the text feature data retrieval and the stored in the pre-processed dataset

### *4.1 Question based Classification*

Question-based classifications the question are classified into two steps first to identify the interrogative words (yes/no) and directly answer the question. Second classify the interrogative words using naive Bayes classifier. To extract text feature like the bigram text feature head word and class specific related words. Table [1] shows the four classes classification model based on class-specific related words.

### *4.2 Answer based classification*

Answer based classification the verb and bigram words are extracted from the Question. These verbs will help to identify how to enrich the textual answers i.e. either by means of image or video.

Table 1.Representation and classification of QA four models based on classes

| Categories | Class with the specific related words |
|---|---|
| TEXT | Country, city, population, distance, speed, age, birthday, websites, name, period, time, date, religions, schools, college, height etc. |
| TEXT + IMAGE | Who, image, whom, logo, pet, place, appearance, colour, clothes, looks like, appearance, figure, what is, symbol, smallest, Largest, band, photo, place etc. |
| TEXT+VIDEO | Songs, music, How to, How do, How can, dance, steps, recipe, differences, ways, story, films, story, said, first etc. |
| TEXT+IMAGE+VIDEO | War, battle, king, prime, kill, issues, minister, bombast, event, president, king, earthquake, nuclear etc. |

### *4.3 Query Generation*

This system is used for generating the queries before performing the search in multimedia search engine. These Queries will help to retrieve the most related images and videos from the web. The first step is query extraction to extract informative keywords from questions and answers [6]. The Second step is the query selection either from the questions or the answers or maybe the combination of both the question and answers. This query selection system includes the features of POS histogram and search performance calculation [1]. This Query generation for Multimedia search which supports only the 42-dimensional search estimate for each QA pairs


Figure 3 System generating the queries before performing the multimedia search Engine

### *4.4 Media Analysis*

Support Vector Machine (SVM) algorithm is used for training and classification it used in media resources and analysis. The training time of Support vector machine is slow but they are highly accurate appropriate media data. Media analysis for determining an appropriate selection of answer medium may be the image, video based on the query generation with different medium type. This search performance is based on the detail information that predicts the good top most results that are quite coherent [15].

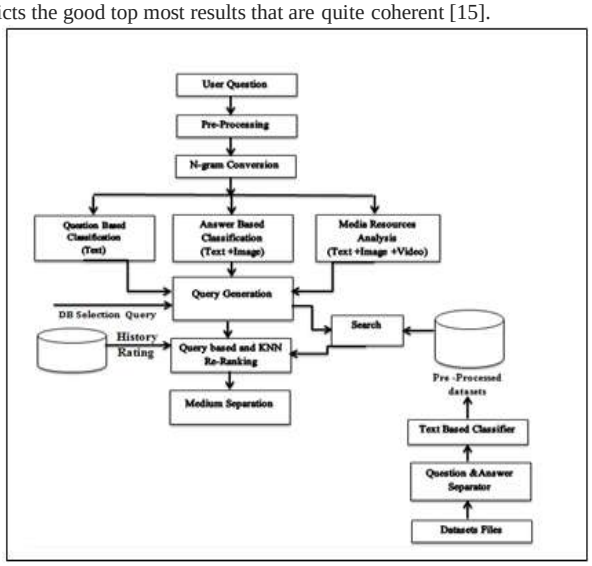


Figure 4.Proposed System Framework

### *4.6 Medium Selection*

The medium selection performs by learning a various four-class of classification model based on the outcomes of the question-based classification, the answer-based classification, and the media resource analysis. For question-based classification, we have the following four scores, i.e., the confidence scores that the question should be answered by text, text+ image, text+ image+ video, and "text+ video. In the same way, for the answer-based classification, we similarly have four scores. Therefore For media resource analysis, we have three scores SVM with linear kernel according to the figure [5].


Figure 5 Selection of medium in form of MMQA Text, Text+Image, Text+Image+Video


Figure 6. MMQA using Text, Text +Images by using Re-ranking and KNN algorithm.

### *4.7 N-Gram Conversion*

N-grams can also be used for effective approximate matching a set of similar items and by exchanging the sequence of items to a set of n-grams [11]. The N-gram is the sequence occurring at the character or the token level which is used for information Retrieval for finding the similar documents N-gram has the fixed access with opinion mining, bags-of –words (BOWs) includes part-of-speech(POS) and also extract the features of syntactic and semantic for extractions of patterns. Legomena n-gram is comparisons of words and replacing twice occurring words .N-gram is used for efficient approximate matching and has one type of probabilistic model used to predict the next items in the sequence form of (n-1) order as shown in figure [6]. In this converting of the sequence of Items to the sets of n-gram embedded in the vector space and one sequence items is compared with the other sequence in an efficient manner. An n-gram is used for information Retrieval for finding the similar document and this model predicts ai based on - [ai-(n-1)… ai-1], in the Probability it is defined as P (ai | ai-(n-1)….. ai-1).

Algorithm for N-gram

Input:
A= $(a_1, a_2 \ldots\ldots a_n)$ denote the set of words n-grams
Where $(a_1, a_2 \ldots\ldots a_n)$ = features of n gram set.
For each $a_x = (a_{x1,\ldots\ldots} a_{xd})$ denotes the tuple in A
$wt\ (a_x)$ is the weight for the feature $a_x$ in the training data

Output: accurate text Retrieval
1. Compute n-gram features of text data.
2. Calculate weight for n-gram of each word.
3. If one n-gram is the sub sums of other.
then
4. Calculate semantic orientation of each feature.
5. If one set is parallel to other.
then
6. Compute the co-relation between features of two sets

### *4.7 Re-ranking and KNN algorithms*

The image search reranking suffers from the infidelity of the expectations under which the text-based search result. However these resultant images contain more irrelevant images.so to overcome their ranking concept rises to re-rank images those are retrieved images which are based on the text and nearly by the image, data and data of image also the visual various types of Image. Reranking of these images is carried out by differentiating between various methods number of times and then the high ranked images used as noisy data. Further, the KNN algorithm is used for classification to learn to

resolve to rank. The main invocation of the overall method is in gathering text and metadata of images, visual features in order to succeed an automatic image ranking as shown in the above figure [6].

Algorithm for Re-Ranking

Input: Retrieved Answer
Output: Re-Ranked
1. For each retrieved answer compute answer distance vector with source query.
2. For each retrieved answer calculate the rating from history log.
3. Sort the retrieved answer according to the distance vector*
4. Rearrange the sorted answer question according to the ranking
5. Provide ranking according to the distance from source query.

### *4.7 Intent Search*

Intent search technique will enrich the user intention to find query question generated by the users. For this medium selection4.4 perform the dynamic role in various four-class classification model based on the question-based, answer based classification and media analysis. And the users can acquire the best answer from the set of relevant answers for the text-based query. This search performance will predict the users to propose the scheme that enriches the textual answer in large cQA corpses with images and video information from the relevant data from the database. Then for the user-provided question, in this, we can perform question search to find multimedia answer in the cQA corpuses. The intent for audio very few users likes to answer the question using these kinds of the medium in QA, so the speech content can be presented in the form of text. Also, the technique of OCR(optical character recognition) and ASR(Automatic Speech Recognition) to recognize the textual data from the low pixels image to high pixels image and to convert the audio to text data format by using the ASR technique.

Algorithm for intent search
Input: - Src Question
Output: - Set of most Relevant Answer

1. N-gram =compute n-gram (Src Question)
2. Set distance_vector = ← φ
3. for q for each dataset question
   Compute Histogram
   Set qg =compute ngram (q)
   Distance_vector[q] = compute distance (ngram (qg))
   End for
4. Order (sort) distance_vector for lowest to highest (i=1……n)
5. Select the k-Answer with the lowest distance vector.
6. Return to the k-nearest answer
7. End

## V. PERFORMANCE AND RESULT ANALYSIS

### *5.1 Result and Discussion*

In this Table [2] we are comparing the number of QA sets, Retrieval time in sec, Accuracy, recall and precision all this is calculated for the extraction of text retrieval media data. We are matching the performance of the proposed system with the existing system in terms of accuracy rate. From the end of this research section, we can say that the proposed system has higher effectiveness than the existing system. From the graph, we can see that the accuracy of the system is reduced somewhat in the existing given system than the proposed system. From this graph we can take the accuracy of proposed system is better which will be the best ones.

$$\text{Precision} = \frac{|\{\text{relevant answer}\} \cap \{\text{retrieved answer}\}|}{\{\text{Retrieved answers}\}}$$

$$\text{Recall} = \frac{|\{\text{relevant answer}\} \cap \{\text{retrieved answer}\}|}{\{\text{Relevant answers}\}}$$

Table 2.Calculation and Representation of performance and result analysis

| *No. Of QA* | *Retrieval time (Sec)* | *Precision* | *Recall* | *Accuracy* |
|---|---|---|---|---|
| 500 | 1.5 | 0.6 | 0.76 | 0.58 |
| 1000 | 1.8 | 0.75 | 0.65 | 0.69 |
| 1500 | 2.1 | 0.79 | 0.44 | 0.77 |
| 2000 | 2.5 | 0.88 | 0.34 | 0.91 |

1.Accuracy

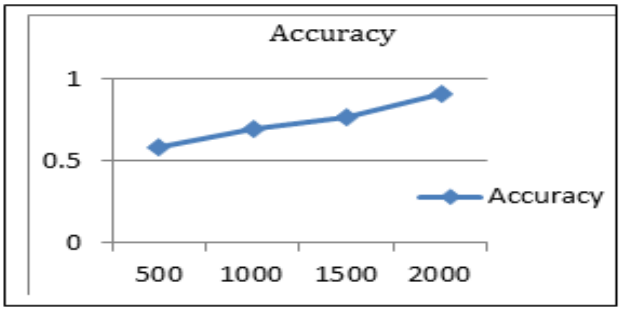


Figure 7. Graphical representation of Accuracy

2. Retrieval Time

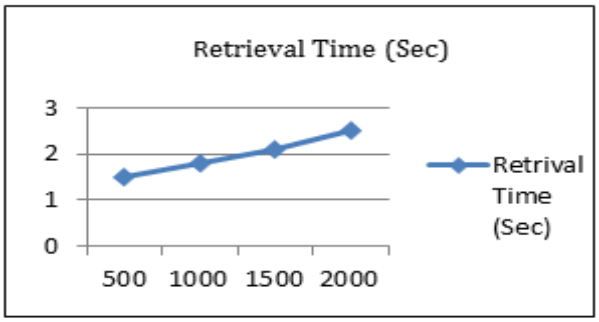


Figure 8. Graphical representation of retrival time

3. Precision

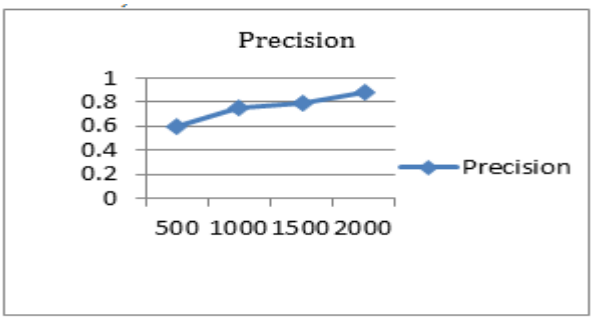


Figure 9 Graphical representation of Precision

4. Recall

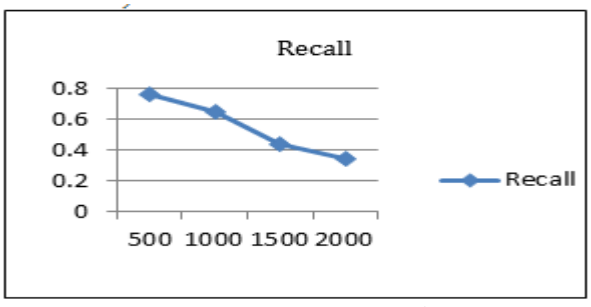


Figure 10 Graphical representation of Recall

## VI . CONCLUSION

Existing system uses a different scheme to answer questions using the media data by leveraging the textual answers in cQA. For a given QA pair, our scheme first determines which type of medium is suitable for enriching the novel textual answer. Following that, it impulsively creates a query based on the QA knowledge and then implements multimedia search with the query.so for proposed relevance ranking for social image search; at the same time take the relevance and the diversity into account. It leverages both visual information of the images and the semantic information of tags. Finally, re-ranking and duplicate removal are executed to obtain a set of images and videos for presentation along with the original textual answer. In our future work, we will further develop the scheme, such as developing for the better query generation method and for audio very few user like to answer the question using these kinds of medium in QA, for better speech content can be presented in text form and examining the relevant segments from a video.

## REFERENCES

[1] Liqiang Nie, Meng Wang, Member IEEE, Yue Gao, Zheng-Jun Zha, Member, IEEE,and Tat-Seng Chua, Senior Member, IEEE Beyond Text QA: Multimedia Answer Generation by Harvesting Web Information IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 2, FEBRUARY 2013.

[2] Quarteroni and S. Manandhar, Designing an interactive open domain question answering system, J.Natural Lang. Eng., vol. 15, no. 1, pp. 73–95, 200

[3] D. Mollá and J. L. Vicedo, Question answering in restricted domains: An overview, Commutate. Linguist. vol. 13, no. 1, pp. 41–61, 2007.

[4] T. Yeh, J. J. Lee, and T. Darrell, Photo-based question answering, in Proc. ACM Int. Conf. Multimedia, 2008.

[5] H.Yang, S.Neo, L.Chaisorn, Chua Available: Generation by Harvesting Web Information- IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 2, FEBRUARY 2013.

[6] H. Cui, M.-Y. Kan and T.-S. Chua, Soft pattern matching models for definitional question answering, ACM Trans. Inf.Syst., vol. 25, no.2, pp. 30–30, 2007.

[7] R. C. Wang, N. Schlaefer, W. W. Cohen, and E. Nyberg, Automatic set expansion for list question answering, in Proc. Int. Conf. Empirical Methods in Natural Language Processing, 2008.

[8] L. A. Adamic, J .Zhang, E. Bakshy, and M. S. Ackerman, Knowledge sharing and Yahoo answers: Everyone knows something, in Proc. Int. World Wide Web conf., 2008.

[9] G. Zoltan, K. Georgia, P. Jan, and G.-M. Hector, Questioning Yahoo! Answers, Stanford Info Lab, 2007, Tech. Rep.

[10] H. Yang, T.-S. Chua, S. Wang, and C.-K. Koh, Structured use of external knowledge for event-based open domain question answering, in Proc. ACM Int. SIGIR Conf., 2003.

[11] S. K. Shandilya and N. Singhai, Article: A survey on: Content based image retrieval systems, Int. J. Comput. Appl., vol. 4, no. 2, pp. 22–26, 2010.

[12] Selecting Attributes for Sentiment Classification Using Feature Relation Networks, Abbasi, A. ,IEEE Sheldon B. Lubar Sch. of Bus., Univ. of Wisconsin - Milwaukee, Milwaukee, WI, USA ; France, S. ; Zhu Zhang ; Hsinchun Chen,IEEE vol 23 march 2011

[13] Y.Liu, T. Mei, X.-S. Hua, J. Tang, X.Wu, and S. Li, Learning to video search rerank via pseudo preference feedback, in Proc. Int. Conf. Multimedia & Expo, 2008.

[14] X. Li and D. Roth, Learning question classifiers, in Proc. Int. Conf. Computational Linguistics, 2002.

[15] J. Zhang, R. Lee, and Y. J. Wang, Support vector machine classifications for microarray expression data set, in Proc. Int. Conf. Computational Intelligence and Multimedia Applications, 2003.

[16] A. Tamura, H. Takamura, and M. Okumura, Classification of multiple-sentence questions, in Proc. Int. Joint Conf. Natural Language Processing, 2005