Question Classification in Social Media

تصنيف الأسئلة في الاعلام الاجتماعي

Mohan John Blooma, Dion Hoe-Lian Goh, Alton Yeow Kuan Chua
Division of Information Studies
Wee Kim Wee School of Communication & Information
Nanyang Technological University
Singapore 637718.
{bl0002hn, ashlgoh, altonchua}@ntu.edu.sg

خلاصة

يعتبر تصنيف الأسئلة جزءا هاما في أنظمة السؤال والجواب الحديثة. تعتمد أغلب الطرق في تصنيف الأسئلة على قواعد مبنية يدويا. أغلب الدراسات الحديثة قامت بتصنيف الاسئلة البسيطة باستخدام طرق التعلم الآلية وأغلبها أوصت باستخدام SVM كأحد أفضل طرق التصنيف أداءا. هذه الدراسة استخدمت التصنيف الشجري المعتمد على SVM كخوارزمية للتعلم الآلي للأسئلة المقدمة من المستخدمين وتم أخذ الأسئلة من موقع أجوبة ياهوو. (Yahoo!Answers). تكمن أهمية هذه الدراسة بمحاولتنا لتصنيف الأسئلة المعقدة والمتكونة من عدة أسطر والمكتوبة من أشخاص حقيقيين. في هذا البحث قدمنا دقة التصنيف والتي تم تحقيقها باستخدام كلا من مصنف قليل التفاصيل (coarse-grained classifier) ومصنف عالي التفاصيل أو الدقة (fine-grained classifier) لتوضيح فعالية الطريقة التي اتبعناها في هذا البحث على الأسئلة المعقدة. بالإضافة إلى ذلك، قدمنا مصفوفة الارتباك أو التشوش (confusion matrix) لتحليل النتائج المقدمة من المصنف الخاص بنا.

Question Classification in Social Media

Mohan John Blooma, Dion Hoe-Lian Goh, Alton Yeow Kuan Chua
Division of Information Studies
Wee Kim Wee School of Communication & Information
Nanyang Technological University
Singapore 637718
{bl0002hn, ashlgoh, altonchua}@ntu.edu.sg

ABSTRACT: *Question classification is an important part in modern Question Answering systems. Most approaches to question classification are based on handcrafted rules. Recent studies classify simple questions using machine learning techniques and recommends SVM as on of the best performing classifiers. This study applies a hierarchical classifier based on the SVM machine learning algorithm on questions posed by users, drawn from Yahoo! Answers. The significance of this study is that we attempted to directly classify complex questions with multiple sentence questions posed by real users. We report the accuracy achieved using both a coarse-grained classifier and fine-grained classifier to illustrate the effectiveness of our approach on complex questions. We also present a confusion matrix to analyze the results made by our classifier.*

Keywords: Question classification, question answering systems, support vector machines, hierarchical classifiers, social media

Received 12 June 2008; Revised and accepted 16 July 2008.

1. Introduction

Today, World Wide Web has moved from content provision towards an interactive information space where end users can actively create and share information. This phenomenon is facilitated by various platforms like wikis, blogs and community support systems. These platforms provide mechanisms that allow every user to access, or change the content of this information space with respect to access rights and correct authorization. These social information spaces are collectively termed as social media (Sebastian et al., 2007).

Social media is exponentially growing with various forms of content like opinions, arguments and solutions. These collective contributions of users result in harnessing burgeoning knowledge (Yang, et al., 2008; Gruber, 2008). Harvesting the knowledge generated in social media opens a new arena for researchers to probe. On one side, the absence of any editorial control in social media coupled with the ease in using these platforms lowers the barriers for users to contribute, leading to the proliferation of knowledge. On the other side, issues such as creation of content in diverse subjects, lack of structure, various styles of language, and high probability of spam, makes it difficult to separate the wheat from the chaff. (Bian et al, 2008; Heyman et al., 2008).

There are various strands of studies on reusing the knowledge built in social media. One of the dominant lines of investigation in mining and retrieval of information from knowledge built in social media are on improving retrieval using social features (Kirsch et al., 2006, Sebastian et al., 2007). Although studies have already explored mining and retrieval in social media, little research attention have been paid to the possibility of applying question answering (QA) techniques for harvesting the knowledge generated in social media (Jeon et al., 2006; Bian et al., 2008). This research is anchored upon the applying QA techniques for searching previously answered questions from social media.

A Question Answering (QA) system attempts to directly answer natural language questions posed by the user. It is a promising research area which combines natural language processing with information retrieval. The Text Retrieval Conference, TREC (http://trec.nist.gov/), has launched a QA track to support the competitive research on question answering, from 1999 (TREC8). The focus of the TREC QA track is to build a fully automatic open-domain QA system.

Current QA systems mainly deal with fact-oriented questions that ask for facts about a target, such as a person or an organization. To answer a question, a QA system has to identify definitions about the target from the corpus and summarize them to form an answer. One of the most important aspects of question answering is, understanding the question. Therefore, classifying questions has become a crucial step in building modern QA systems. This study aims to classify question posed in social media.

We define Question Classification to be the task that, given a natural language question, puts a question into one or more semantic categories. For example, for the question "Who wrote the book The Prince?" we can put into the category of "People" or "Author". "What do you all think of Machiavelli's The Prince?" can be put into the category "Books Reading". Question Classification not only enforces constraints on the plausible answers but also recommend different processing strategies.

Most approaches to Question Classification are based on handcrafted rules Sasaki et al., 2002). The manual construction of classifier rules requires much time spent in analyzing a large number of questions. Only recently, machine learning techniques have been applied to address the problem of Question Classification (Zhang and Lee, 2003; Suzuki et al., 2003). The machine learning approach is relatively well understood and has shown good performance given much training data. Moreover, a trained classifier can be easily adapted to a new domain. This research uses Support Vector Machines (SVM) a commonly used machine learning approach to classify questions (Zhang and Lee, 2003; Solorio and Coutino, 2004; Tamura et al., 2005).

Existing studies on Question Classifications are based on TREC style questions, i.e., open domain factual questions. For example, "Who invented the lamp?" They are simple, single sentence questions. This research is different from existing work in this domain as it classifies questions extracted from a user-oriented question answering system in social media -Yahoo! Answers. A typical question asked by a Yahoo! Answers user is usually complex, containing multiple sentences and detailed descriptions on the scenario which lead to the query. For example: How to set up a network printer in Windows XP on both computers? "I have two computers running on Windows XP. On a computer (desktop) it is the residential gateway (the one with the main internet connection), and on a laptop it uses Wi-Fi access to connect to internet. Conventional QA systems are typically not designed for such questions, and existing techniques may not be applicable or may work poorly for these questions. Therefore, constructing QA systems that can handle multiple-sentence questions is desirable. Classifying such multiple-sentence questions have not been attempted under the TREC QA track. This study therefore attempts to address this gap in the QA system literature.

This paper is organized as follows: First, we describe some previous approaches for question classification in Section 2. In Section 3, we present the Yahoo! Answers dataset we are using and describe the preprocessing method and machine learning algorithms used. In Section 4, we describe our experimental results. We conclude our study in Section 5.

2. Related Work

Traditionally, hand crafted rules were used to classify questions, while machine learning algorithms have gradually advanced into the current trend in the question classification in question answering domain.

Li and Roth (2002) solved the question classification problem based on the SNoW learning architecture. The SNoW is a multi-class classifier that is specifically tailored for learning in the presence of a very large number of features. They presented a hierarchical approach to classifying questions into 5 coarse classes and 50 more specific classes. They used lexical and syntactic features such as part-of-speech tags, chunks and head chunks together with two semantic features: named entities and semantically related words to represent the question. They reported 91% accuracy using 5,500 training data points on semantically related words as features.

Other machine learning approaches used in question classification are Hierarchical Directed Acyclic Graph (HDAG) kernel proposed by Suzuki and Taira (2003) and log-linear models presented by Blunsom and Kocik (2006). Shen (2005) also used automatic question classification for KDD 2005. Shen et. al (2006) used the Internet search engines to enrich the question key words and mapped into the intermediate objects and then into the target categories.

Solorio and Coutino (2004) proposed a language independent classification method based on SVM machine learning approach to deal with QC problems under different language. When training SVM with words as features, they reported 81.7% accuracy on English, 88.03% on Italian and 79.9% on Spanish. Tamura et al. (2005) used SVM to classify multiple sentence questions. Zhang and Lee (2003) present a QC method using Support Vector Machines method for question classification. They used bag-of-words and n-gram features and compared accuracy of SVM against Nearest Neighbors, Naive Bayes, Decision Trees and Sparse Network of Winnows (SNoW), with SVM producing the best results. They achieved 85.8% accuracy on 5500 training data points on bag-of-words features.

The accuracy of QC is very important to the overall performance of a QA system. Results of the error analysis of an open-domain QA system showed that 36.4% of the errors were generated by the question classification module (2003). The good performance of a QC algorithm thus helps greatly in finding suitable answers to a question from a set of documents in digital libraries and information retrieval systems. As most of the literature recommends SVM for obtaining higher accuracy in question classification, we used SVM in this study. However, a crucial difference is that we have used data from a user-oriented QA system where the questions are complex, containing multiple sentences. Hence, our results can augment existing research in question classification.

3. Methodology

In this section, we give our detailed experimental procedures and explain how we used the SVM machine learning algorithm to classify the questions we obtained from the Yahoo! Answers[1] dataset.

The Yahoo! Answers Dataset

Yahoo! Answers is a community-driven knowledge market Web site launched by Yahoo! that allows users to ask and answer questions posed by other users. It is becoming a new way to find and share information. Google Answers[2], Naver[3], and Wondir[4], are other community based user oriented QA services available online. In these portals, users can express specific information needs by posting questions, and get direct responses authored by other web users, rather than browsing results of search engines. It is possible to obtain answers to a wide variety of questions using these services in social media. Hence, these services rapidly built up a large collection questions and answers transforming into a valuable linguistic resource (Adamic et al., 2008). The importance of such CQA services is magnified as they create archives of millions of questions and hundreds of millions of answers, many of which are invaluable for the information needs of information seekers.

Recently there have been a number of researches on Yahoo! Answers. While (Blooma et al., 2007; Blooma et al., 2008; Adamic et al., 2008; Bian et al., 2008) developed a framework to predict the best answer, (Murata and Moriyasu, 2007) used the data from Yahoo! Answers to conduct a study on social network analysis. Work done by (Tamura et al., 2005) is similar to question classification except that they first extracted core sentences from the multiple sentences and converted it into a single sentence before applying the classifier. In Yahoo! Answers, users can ask questions on any topic, receive answers from other people, and share user's insights and experience. Because every question proposed in Yahoo! Answers by the user might be proposed by the user of a QA system, the questions and answers in Yahoo! Answers provide a good source for QA research. It provides a collection of real user questions that are very different from those used in the TREC series.

Question classification techniques can help users of Yahoo! Answers to find similar questions proposed in different ways and detect the similarity between questions and to provide related questions and answers. This work can also impose some constraints on the plausible answers and suggest different answering strategies.

Yahoo! Answers puts the natural langrage questions into 26 categories like "Arts & Humanities", "Beauty & Style" and into detailed 285 categories such as "Arts & Humanities/Philosophy", "Beauty & Style /Skin & Body". In our work, we randomly downloaded 2057 Questions from five categories of Yahoo! Answers. The categories were randomly selected but restricted to academic disciplines, so that the questions asked by the users would be more formal and classifiable.

Support Vector Machines

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. Given the data points of the form: {(x1,c1), (x2, c2),…,(xn,cn)} where ci is either 1 or 1 and each xi is a p-dimensional real vector, usually [0,1] or [-1,1] values. Support vector machines map input vectors to a higher dimensional space where a maximal separating hyper plane is constructed. More details on the SVM machine learning algorithm can be found in (Burges et al., 1998).

Given that SVM has shown good performance for many natural language related applications, such as text classification (Joachims, 2002) and named entity recognition (Solorio and L'opez; 2004), we used it as classification algorithm in our present work. We used SVMmulticlass software to process our Yahoo! Answers data. SVMmulticlass is an implementation of the multi-class SVM described in (Crammer and Singer, 2001). While the optimization problem is the same as in Crammer and Singer, (2001), this implementation uses a different algorithm which is described in Tsochantaridis et al., (2004). However, we are attempting to classify complex question when compared to Crammer and Singer, (2001) and Tsochantaridis et al., (2004) in which simple questions were effectively classified.

---

[1] http://answers.yahoo.com/
[2] http://answers.google.com/answers/
[3] http://www.naver.com/
[4] http://www.wondir.com/wondir/jsp/index.jsp

Dataset Preprocessing

The raw Yahoo! Answers dataset is in the format of Web pages. It consists of the question, a sub-question and all the answers given by different users. It also contains user's profiles. The profile of the asker of the question, and the profile of the answerer whose answer was chosen as the best answer are available.

Often, users of Yahoo! Answers provide some explanations of the questions they ask. For example, in the category "Business & Finance/Corporations", there is a question "Why do companies have a share buyback program?" At the same time, there is an explanation "I have noticed companies buying back shares. BP is one. What is the reason for this strategy? Is it good for investors?" Here, we call this part as a sub-question.

We extracted the question and its sub-question, if it exists, and its category and sub-category from the downloaded web pages. In our dataset, 1024 questions had sub-questions. For every sub-category, we divided the questions into the training set and testing set. We then merged all the training sets part and testing sets and obtained a total of 1534 training questions and 523 testing questions.

Feature Space

In our work, we only extracted syntactic features of the questions. For each question, we extract unigram and bigram features following the algorithm proposed in (Mladeni and Grobelink, 1998). Then we converted the features into a vector space to be used by the SVM classifier. In this study, we only consider whether the unigram or bigram features appears or not in one question. We used 1 or 0 to represent the value of the feature in the vector space. Note that there is often very short questions proposed by the user like "Where is Jantarmantar?" This question has only one word "Jantarmantar" that is not the stop-word. So, we are not able to extract the bigram feature of this question. As a result, we do not use this question in our experiments to calculate bigram feature classification accuracy.

Hierarchical Classifier

Question classification is a multi-class classification problem. From the viewpoint of the target intention, a hierarchical structure of taxonomy should firstly be given. Many question taxonomies have been defined so far but these are not standardized. For example, the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005) [4] gave 67 predefined categories to categorize 800,000 queries. These categories include: Shopping/Stores & Products, Living/Family & Kids, and Entertainment/Games & Toys, and so on. The hierarchy defined by (Li and Roth, 2002) contains 6 coarse classes (ABBREVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION and NUMERIC VALUE) and 50 fine classes (abbreviation, definition, animal, body, event, food, technique, term, city, country, speed, temperature, etc.).

In the Yahoo! Answers taxonomy, a question can be mapped to one of the 5 possible coarse classes and then one of the 23 possible fine classes. The question classifier uses a sequence of two simple classifiers, each utilizing the SVM algorithm. The

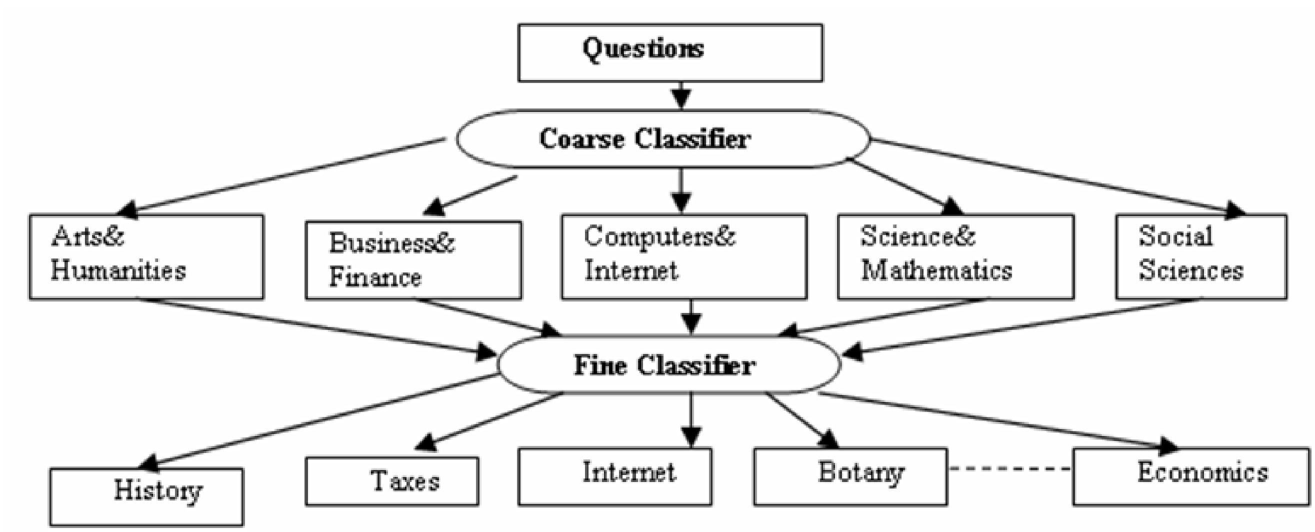

Figure 1. Hierarchical classifier

first classifies questions into coarse classes (Coarse Classifier) and the second into fine classes (Fine Classifier). The second classifier depends on the first in that its candidate labels are generated by expanding the set of retained coarse classes from the first into a set of fine classes; this set is then treated as the possible classes set for the second classifier. The basic structure of the hierarchical classifier is shown in Figure 1.

4. Results and Discussion

We classified the questions obtained from the Yahoo! Answers dataset using the SVM machine learning algorithm. In this section, we give our detailed experimental results. We also discuss the results obtained from our analysis.

4.1 Results

We compare the unigram and bigram features used in our classifier. The following tables show the results obtained from our experiments. The "training" and "testing" columns are the number of questions in our training dataset and testing dataset. The "correct" and "incorrect" columns are the number of questions correctly or incorrectly classified by the SVMmulticlass software. The classification accuracy, which is the proportion of the correctly classified questions among all test questions, is listed in "accuracy" column. In our work, we measure the question classification performance by the accuracy values.

It is evident from the results obtained that bigram features are better than unigram features to classify the questions by the course classifier without using sub-question. Without using sub-questions, we obtained an accuracy of 75% using bigram

	Training	**Testing**	**Correct**	**Incorrect**	**Accuracy**
unigram	1534	500	312	188	62.4%
bigram	1500	128	96	32	75%

Table 2. Yahoo! Answers question classification accuracy for coarse classes

	Training	**Testing**	**Correct**	**Incorrect**	**Accuracy**
unigram	1534	500	287	213	57.4%
bigram	1500	128	80	48	62.5%

Table 3. Yahoo! Answers question classification accuracy for fine classes

	Training	**Testing**	**Correct**	**Incorrect**	**Accuracy**
unigram	1534	513	343	170	66.9%
bigram	1525	269	162	107	60.2%

Table 4. Yahoo! Answers question classification (with sub-question) accuracy for coarse classes

	Training	**Testing**	**Correct**	**Incorrect**	**Accuracy**
unigram	1534	513	298	215	58.1%
bigram	1525	269	143	126	53.2%

Table 5. Yahoo! Answers question classification (with sub-question) accuracy for fine classes

features and 62.4% using unigram features by the course classifier. This result once again demonstrates that bigram features yield a higher accuracy for sentence question classification consistent with previous work in this area.

However, the results show that bigram features demonstrated a decrease in performance of the classifier when considering the sub-questions. On including sub-questions for classification, we obtained an accuracy of only 60.2% using bigram features and 66.9% using unigram features by the course classifier (see Table 4). This is much lower than the 75% of accuracy obtained without including the sub-questions (see Table 2). Hence we may conclude that bigrams are not useful when classifying complex, multiple-sentence questions. This is because sub-questions may contain unnecessary sentences which increase the number of noisy features for question classification. This is evident on comparing results from Table 2 and 4, and Table 3 and 5 where the corresponding accuracy decreased after including the sub-questions. Although, Tamura et. al, (2005) used multiple sentences, they had extracted core sentence to classify in order to obtain a higher accuracy for the bigram features. However, their approach adds an additional layer of complexity due to the need for preprocessing of questions. Other similar studies used simple sentence questions to classify and hence obtained higher accuracy for the bigram features.

Table 3 and 5 shows the decrease in accuracy on using fine classifiers when compared to the accuracy in Table 2 and 4 respectively obtained using course classifiers. However, Li and Roth, (2002), were able to obtain significant improvement in fine classifiers using semantically related words. The low performance of fine classifiers will have to be explored in our future work by increasing the data set and also by using semantic features to reduce the noisy features in real user questions.

4.2 Discussion

When referring to the performance of a classification model, we are interested in the model's ability to correctly predict or separate the classes. When looking at the errors made by a classification model, the confusion matrix gives a fuller picture. We only analyze the coarse classifier on unigram features here due to space constraints. The confusion matrix is given in Table 6. We use the label 1, 2, 3, 4, 5 to represent the categories: Arts & Humanities, Business & Finance, Computers & Internet, Science

		Predicted class					
		1	2	3	4	5	Total
Known class (class label in data)	1	61	4	5	10	20	100
	2	6	74	10	7	9	106
	3	4	10	73	4	9	100
	4	15	5	8	51	16	95
	5	17	9	4	15	53	98

Table 6. Confusion matrix by course classifier on unigram feature

& Mathematics, Social Science. The rows correspond to the known categories of our data. The columns correspond to the predictions made by our classifier. The value of each of element in the matrix is the number of predictions made with the class corresponding to the column for examples with the correct value as represented by the row.

In Table 6, of 100 label 1 (Arts & Humanities) test data points, the classifier misclassified 20 into label 5 (Social Science). At the same time, of 98 label 5 (Social Science) test data points, the classifier misclassified 17 into label 1 (Arts & Humanities). This indicates the close relationship of the two labels. In reality, it is often difficult to distinguish Social Science from Arts & Humanities.

We discuss some misclassified examples here:

"Who said, "a rose is a rose is a rose"?" The correct category is "Arts & Humanities", but the coarse classifier failed to relate the "who said" as an author, output as "Social Science" because of "rose".

Take another example: "Why do skilled workers resent close supervision?" The correct category is "Social Science/ Sociology", but the fine classifier failed to understand the question, output as "Business & Finance/ Corporations" because of "worker" and "supervision" terms.

Because the category of a Yahoo! Answers question is given by the user, the user mislabels the category in some cases. For example, a user categorizes the question “Where to get chalk brushes? Yes this is a repost but I didn’t get any answers to my 1st question?” into “Arts & Humanities”. Surely it is not correct. The classifier classified it into “Science & Mathematics” because there is a “questions” and “answer” feature in that question.

We summarize that some of the misclassifications are partly due to the limitations of our dataset. Machine learning algorithms perform better when a bigger training set is available. Since we only have 2057 questions in our research, we believe that given more training data, our results will be improved.

5. Conclusion

This paper discusses the problem of question classification for a questions posed in social media. A popular user-oriented QA system Yahoo! Answers is used in this study. The question classification is performed using support vector machines. We developed a hierarchical classifier and used it to classify questions into coarse classes and fine classes. Our experimental results showed that this approach can be useful for classifying complex, multi-sentence questions such as those found in Yahoo! Answers. We also compared the effectiveness of unigram and bigram features on question classification. In our work, we experimented with the use of sub-questions. However, the results showed nearly no improvement in classification accuracy. Although Tamura et al, (2005) classified multiple sentence questions, they extracted core sentence from the multiple sentences and applied the classification to the simple sentences. It is worth investigating the core sentence extraction approach in our future work.

Finally, we analyzed our results and presented a confusion matrix to illustrate some classification errors. In this paper, we only tested unigram and bigram features on our dataset. Work by Li and Roth, (2002) concluded that semantic features contribute to the accuracy of the classifier. Our future work is aims to exploit the syntactic structure of the question (Zhang et al., 2003) and to extract the semantic features like named entities, target words and semantically related words from the Yahoo! Answers data set and explore the influence of these features. Meanwhile, other machine learning algorithms can also be applied to our dataset including SNoW learning architecture (Li and Roth, 2002), HDAG kernel (Suzuki et al., 2003) and log-linear models (Blunsom et al, 2006).

References

[1] Adamic, L. A., Zhang, J., Bakshy, E., Ackerman, M. S., (2008) Knowledge sharing and Yahoo answers: Everybody knows something. *In the proceedings of WWW 2008. Beijing, China*.

[2] Bian, J., Liu, Y., Agichtein, E., Zha, H., (2008) Finding the right facts in the crowd : Factoid question answering over social media. *In the proceedings of WWW 2008, Beijing China.*

[3] Blooma, M.J., Chua, A., Goh, D., Ling, Z. Towards a hierarchical Framework for Predicting the Best Answer in a Question Answering System. *Proceedings of the 10th International Conference on Asian Digital Libraries*, December 10-13, Vietnam, 2007.

[4] Blooma, M.J., Chua, A., Goh, D. A Predictive Framework for Retrieving the Best Answer. *Proceedings of the 23rd Annual ACM Symposium on Applied Computing,* March 16-20, Brazil., 2008.

[5] Blunsom, P., Kocik, K., Curran, J. R. (2006). Question Classification with Log-Linear Models, *In the Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, Seattle, Washington, USA, 2006.

[6] Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and KnowledgeDiscovery 2.*, p. 121-167, 1998.

[7] Crammer, K Singer, Y. (2001). On the Algorithmic Implementation of Multi-class SVMs, *JMLR*, 2001.

[8] Joachims, T. (2002) Learning to Classify Text using Support Vector Machines: Methods Theory and Algorithms, *The Kluwer International Series in Engineering and Computer Science, 668,* Kluwer Academic Publishers, 2002.

[9] Murata, T. Moriyasu, S. Link prediction of social networks based on weighted proximity measures, 2007.

[10] Li, X., Roth, D. (2002). Learning question classifiers. *In the proceedings of the 19th international conference on Computational linguistics,* Taipei, Taiwan. 1-7.

[11] Mladeni, D., Grobelink, M. (1998). Word sequences as features in text learning. *17th Electro. & Comp. Science Conf.,* Slovenia, p. 145-148, 1998.

[12] Tsochantaridis, I, Hofmann, T., Joachims, T., Altun, Y. (2004) Support Vector Learning for Interdependent and Structured Output Spaces, ICML, 2004.

[13] Sasaki,Y., Isozaki, H., Hirao, T., Kokuryou, K., Maeda, E. (2002) : *NTT's QA Systems for NTCIR QAC-1*. Working Notes, NTCIR Workshop 3, Tokyo, p. 63-70, 2002.
[14] Solorio, T., Coutino, M. P. (2004). A language independent method for question classification, *In the Proceedings of the 20th international conference on Computational Linguistics,* Geneva, Switzerland.
[15] Solorio, T., L´opez, A. (2004). Learning named entity classifiers using support vector machines. In Alexander Gelbukh, editor, *Fifth International Conference on Intelligent Text Processing and Computational Linguistics*, CICLing 2004, 2945 of Lecture Notes in Computer Science, p. 158-167. Springer, 2004.
[16] Shen, D., Pan, R., Sun, J., Pan, J. J., Wu, K., Yin, J., Yang, Q. (2005). Q2C@UST: our winning solution to query classification in KDDCUP 2005, *In the proceedings of ACM SIGKDD,* 7 (2), 100-110.
[17] Shen, D., Sun, J., Yang, Q., Chen, Z. (2006). Building bridges for web query classification. *In the proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, Seattle, Washington, USA, 2006.
[18] Shen, D. (2005). An Ensemble Method for Query Classification, *In the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'05)*. Chicago, USA.
[19] Suzuki,J., Taira, H., Sasaki, Y., Maeda, E. (2003). Question Classification using HDAG Kernel. *In Workshop on Multilingual Summarization and Question Answering,* p. 61-68.
[20] Tamura, A., Takamura, H., Okumura, M. Classification of Multiple-Sentence Questions. *IJCNLP 2005,* LNAI 3651, p. 426-437, 2005.
[21] Zhang, W., Lee, S. (2003) Question classification using support vector machines, *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, Toronto, Canada.