VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

FEATURE ANALYSIS FOR WEB FORUM QUESTION POST DETECTION

Adekunle Isiaka Obasa, Naomie Salim and Atif Khan
Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia
E-Mail: iaobasa@yahoo.com

ABSTRACT

A web forum which is also known as discussion board or Internet forum is an online community of users with a common interest. It is a problem-solving platform that engages experts across the globe. Both technical and non-technical problems are resolved on a daily basis within web forums. Research activities in this domain have been concentrated on answer detection with the assumption that the initial post of a thread is a question post. The quality of web forum question posts varies from excellent to mediocre or even spam. Detecting good question posts require utilization of salient features. In this paper, we implement a bag-of-words (BoW) model to mine web forum question posts. We empirically address the following questions in the paper. Can BoW model effectively detect web forum question post? What feature selection method is most appropriate for BoW model in this domain? Is choice of classifier influenced by web forum genre? We used three publicly available datasets of varying technical degrees for the experiments. The experimental results revealed that BoW can perform better than complex techniques that implement higher N-gram with part-of-speech tagging.

Keywords: web forum, bag-of-words, feature selection, question detection, dimensionality reduction.

INTRODUCTION

The question-answering (QA) paradigm, i.e. the process of getting precise answers to natural language (NL) questions, was started in late 1960-ies and early 1970-ies within the framework of NL understanding. The dawn of WWW has introduced the need for user friendly querying techniques that reduce information overflow, and poses new challenges to the research in automated QA. A large natural QA setting, which is community oriented is discussion board.

People mostly use the discussion boards (i.e. web forum) as problem-solving platforms. Web forum members post questions relating to some specific problem, and expect others to provide potential answers. This scenario is depicted in Figure-1. A number of commercial organizations such as Microsoft, Dell and IBM directly use online forums as problem-solving domain for answering questions and discussing needs raised by customers. [1] found that 90% of 40 discussion boards they studied contain question-answering knowledge. By using speech acts investigation on several sampled forums, [2, 3] discovered that question answering content is usually the largest type of content on forums.

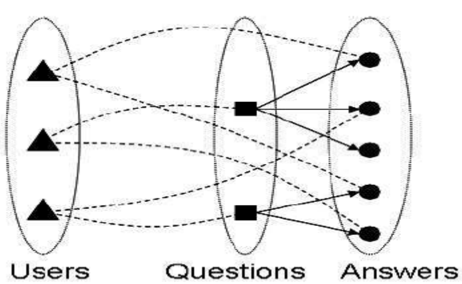


Figure-1. Network of interactions in forum connecting users, questions and answers ([4].

The collaborative activities within the forum offer a lot of benefits. In technical forums such as hardware or software forum, a lot of issues such as installing software or hardware, troubleshooting codes, fixing bugs, implementing tools, etc. are being discussed on a daily basis. For non-technical forum like travel, members share their travel experience with others. Good opinions are generated by members for the benefit of other members. It will be highly desirable to mine human knowledge being generated in the forum for the benefit of mankind. The aim of this paper therefore, is to mine standard initial posts as web forum question posts using bag-of-words with dimensionality reduction.

The rest of the paper is organized as follows. Section 2 gives description of the problem. Section 3 discusses related work while Section 4 presents the proposed approach. Experimental design is done in Section 5. Section 6 concludes the paper.

PROBLEMS FORMULATION

We consider web forum question post detection as classification problem. The problem is about getting salient features that can effectively classify web forum initial post as a question or not question. The initial post of a thread is considered as a whole as a question post if it contains a specific problem that needs to be solved otherwise it is non-question post. This problem definition is similar to that of [5] and [6]. For example, the following statements constitute a question post from Photography on the net, a digital camera forum.

"I have found that when I take pictures of scenery or landscape with no particular focus, that the camera has a difficult time focusing. I have tried the landscape mode but that does not work very well. I am mainly trying to do manual focus, however it is so difficult to tell by just

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

looking through the viewfinder. Are there any techniques or tips that anyone would recommend?"

The last sentence in the post is a question sentence; it gives little information about the real problem. The problem is the entire scenario that the author described using several sentences as a whole. It is therefore practical to treat the whole post as a question post.

Web forum initial post is often being considered as question post when mining answers from web forum [7-9] without due consideration for what the post is all about. Initial post can be an announcement, a report or an acknowledgement which does not require any answer from the members. Furthermore, some initial posts are trivial questions that cannot be mined for any knowledge discovery. An example of such that can be found in forum due to its less restrictiveness is a question post like *"Hi guys, check out my pictures on facebook. Can anybody say that I'm not handsome?"* In view of all these issues, it is desirable to first identify the web forum question post before looking for its answer.

RELATED WORKS

In this section, we will review works that are closely related to our mining approach i.e. bag-of-words (BoW). Notable research activities that involve the use of bag-of-words combined with some other approaches based on news articles, community-based question answering (CQA) and web forum corpus are shown in Table-1.

In Table-1, [10] used news article of Wall Street Journal corpus to determine opinion questions using BoW combined with n-gram. Their BoW was simply collection of opinion words which are positive and negative adjectives, nouns, verbs and adverbs. This in a way is a form of filtering out some word identities on a larger scale compared to the works of [6, 11]. This influenced the performance of the BoW and n-gram's higher result compared to others. [12] and [13] used BoW with 2 and 3-grams without feature selection to achieve similar results using different classifiers.

Table-1. Review of bag-of-words combination with other approaches.

Author/Year	What is combined with BoW	Feature Selection used for BoW	Learning method	Motivation	Result accuracy (*F-1* measure) (%)
[10]	BoW + 2-gram + 3-gram	Considered mainly opinion words	Naïve Bayes	Answering opinion questions by separating opinions from facts	87
[12]	BoW + POS + 2-gram + 3-gram	None	LibSVM	Determining whether CQA question has Objective or subjective orientation	72
[14]	BoW + 2-gram	Chi-square	LibSVM	Community QA question classification	75.3
[13]	BoW + 2-gram + 3-gram	None	Multinomial Naïve Bayes	Evaluation of subjectivity analysis in web forum.	72.3
[15]	BoW only	None	Multinomial Naïve Bayes	Classification of web forum posts	57.7

[14] combined BoW with 2-gram and applied chi-square as feature selection to obtain a slightly higher result. A very low result was realized by [15] that used only BoW. This confirms that BoW needs to be combined with some other approaches to enhance its performance. It is also worth noting that feature selection enhances BoW performance. A question one could ask here is what dimensionality reduction will be most suitable for enhancing BoW performance? This question, to a large extent is addressed in this study.

EXPERIMENTAL DESIGN

In this section, we show how our proposed approach is actualised. The section begins with a discussion about the proposed approach followed by datasets and their annotations. Thereafter, experimental setting is discussed.

Proposed approach

In question post detection, initial posts of threads are modelled as unordered collection of words picked from one of two probability distributions: one stands for question (Q) and the other non-question (NQ). This can be viewed as two literal bags full of words. One bag is filled with words found in question posts, and the other bag is filled with words found in non-question posts. While any given word is likely to be found somewhere in both bags, the "question" bag will contain question-related words such as "how", "can", and "where" which are much more frequent in question posts, while the "non-question" bag will contain words that have nothing to do with question.

To classify a post, the classifier assumes that the post is a pile of words that has been poured out randomly from one of the two bags, and uses an algorithm to determine which bag it is more likely to be. In summary, the steps are:

a. Detect and extract keywords,
b. Build a keyword dictionary, and
c. Use keyword dictionary to build term-document matrix
d. Use machine learning to train a classifier for the classification.

The above procedure will generate a set of keywords known as bag-of-words. As explained above, these keywords are the features that will be used to mine the questions post. Most of the values of the term-document matrix will be zeros since for a given document, a small fraction of it will be found in keyword dictionary. In view of this, bag-of-words are said to be typically high-dimensional sparse datasets that require a lot of memory. In addition, some of the non-zero features could be redundant or less effective for the task of question detection. In order to overcome the problem outlined above, we experiment with both filter and wrapper feature selections to obtain the most salient features.

Dataset and dataset annotation

Three different datasets were used for the experiments conducted in this research. We collected 16, 853 threads of Photography On The Net, a digital camera forum (CAM dataset) and 41,078 threads of Ubuntu Fora, an Ubuntu Linux community forum (Ubuntu dataset). In addition, we also collected 31, 998 threads of Trip Advisor-New York that contains travel related discussions on New York City (NYC dataset). All the datasets are made available publicly by [5, 13, 16]. These three fora are considered so as to evaluate the implemented methods on different domains of online fora. The Ubuntu dataset that contains a lot of configuration parameters and codes represents highly technical domain, CAM dataset that contains more of technical terms and some settings but no codes represents less technical domain while NYC dataset that does not contain codes, configuration settings and more of technical terms represents non-technical domain. Details of the datasets are shown in Table-2.

Table-2. Dataset analysis

Dataset	No. of threads	No. of posts	Source
TripAdivisor	32,000	420,657	http://www.tripadvisor.com/
Ubuntu	41,078	198,828	http://ubuntuforums.org/
Photography	16,853	190,953	http://photography-on-the.net/

In order to obtain class labels for the question posts, we recruited three annotators. One worked on both NYC and Ubuntu datasets while the other two worked on the Ubuntu and NYC datasets separately. All the annotators were senior graduate students. Two of them were in computer science faculty and the other one in civil engineering. The two annotators from computer science were familiar with Ubuntu operating system and they were asked to annotate Ubuntu. The annotator from civil engineering was a member of TripAdvisor travel forum and was asked to do the second annotation for NYC dataset. 500 threads of photography were already annotated by [5]. The summary of question detection instances (i.e. initial posts) used for both training and testing in this research are shown in Table-3.

Table-3. Question detection dataset summary

Instances	CAM	Ubuntu	NYC
Total No. of Positive Instances (i.e. Questions)	204	223	225
Total No. of Negative Instances (i.e. Non-Questions)	204	223	225
Total No. of Initial Posts	408	445	450

Experimental setting

We used different supervised learning algorithms for our classification task. These algorithms include Multinomial Naïve Bayes (MNB), Support Vector Machines (LibSVM), Decision tree (J48), Sequential Minimal Optimisation (SMO) and Multilayer Perceptron (MP). In order to aid the experimentation carried out in this research, a freely available machine learning toolkit called weka is used. Weka is a pool of machine learning algorithm for data mining activities. The version of weka implemented in this study is weka 3.7.12 and can be downloaded at http://www.cs.waikato.ac.nz/ml/weka/.

Classification results are obtained using 10-fold cross-validation and 80% split (i.e. 80% training, 20%

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

testing). The performances of our classifiers were evaluated using precision, recall and F-1 measure metrics. Basic pre-processing such as removal of HTML tags and lower casing all words were performed on the corpus of initial posts used for the experiments. Dimensionality reduction is performed using both filters and wrapper. The filters considered in the study are: Chi-square, Information gain (Info. Gain), Gain ratio, Symmetrical uncertainty (Sym. Uncert.). These filters are experimented using three thresholds of 0, 5 and 10. The wrapper method is based on SMO classifier. SMO was determined empirically for the wrapper.

RESULTS AND DISCUSSIONS

In Table-4, screening of the three datasets (CAM, Ubuntu and NYC) using different reduction methods confirm chi-square, information gain, gain ratio and symmetrical uncertainty exhibiting the same feature reduction with only chi-square giving discriminative features for thresholds of 5 and 10. In the table, the 1775 features of CAM dataset were reduced to 253 features for all the four filters using threshold of 0. Chi-square gave 93 and 15 features for thresholds of 5 and 10 respectively. Classification results of the four for threshold of 0 are the same. In view of this, our empirical analyses are based on chi-square, wrapper and non-filtering.

Table-4. Datasets feature reduction analyses.

Dataset	Filter /Wrapper	Thresholds		
		0	5	10
CAM	Chi-square	253	93	15
	Info. Gain	253	0	0
	Gain Ratio	253	0	0
	Sym. Uncert.	253	0	0
	Wrapper(SMO)	63		
	No. Filter	1775		
Ubuntu	Chi-square	139	74	10
	Info. Gain	139	0	0
	Gain Ratio	139	0	0
	Sym. Uncert.	139	0	0
	Wrapper(SMO)	44		
	No. Filter	1626		
NYC	Chi-square	99	98	33
	Info. Gain	99	0	0
	Gain Ratio	99	0	0
	Sym. Uncert.	99	0	0
	Wrapper(SMO)	22		
	No. Filter	124		

Tables-5 through -9 show the results of the five different classifiers considered in this study. A total of 144 experiments were performed for the three datasets discussed in Section 4 above using bag-of-words with different dimensionality reductions outlined in Table-4. Both cross validation and 80% split were used to validate the results. As expected, the BoW without dimensionality reduction performed poorly with all the classifiers. The use of chi-square with different thresholds gives some improvements. An amazing observation with the use of chi-square thresholds is that higher thresholds with fewer feature space does not guarantee better performance. This reveals that higher threshold of chi-square does not optimize feature selection. The wrapper method with higher number of features often performs better than the higher threshold of chi-square with lesser number of features. Out of the five classifiers, multinomial Naïve Bayes works much better with chi-square using lower threshold especially on less technical datasets.

The results of MNB and SMO are the best of the 5 classifiers. SMO gave best result for CAM dataset (a less technical dataset) while MNB gave best results for Ubuntu (a highly technical dataset) and NYC (a non-technical dataset). A comparative analysis of the MNB and SMO is shown in Figure-2. SMO works better with the wrapper method while MNB favours chi-square with lower

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

threshold. Cross validation favours CAM dataset (a less technical dataset) and 80% split favours both Ubuntu and NYC. The MP classifier takes much longer time to generate results. Its computation for thousands of features was ignored in this study since such results cannot be better than the filters method.

Table-5. Empirical results using SMO classifier

Dataset	Feature selection method	No. of feature	Validation method	P	R	*F-1*
CAM	No Filter	1775	Cross	62.7	62.7	62.7
			80% Split	65	64.6	64.6
	Chi-square	253	Cross	81.7	79.9	80.2
			80% Split	74.8	72	71.5
		93	Cross	74	72.8	72.5
			80% Split	66.5	63.4	62.3
		15	Cross	72.2	71.1	70.7
			80% Split	74.8	72	71.5
	Wrapper	63	Cross	85	84.8	84.8
			80% Split	73	69.5	68.8
Ubuntu	No Filter	1626	Cross	59.3	59.3	59.3
			80% Split	66.3	66.3	66.1
	Chi-square	139	Cross	74.6	73.3	72.9
			80% Split	75.2	71.9	71.4
		74	Cross	69.1	68.1	67.6
			80% Split	71	67.4	66.6
		10	Cross	66.1	66.1	66.1
			80% Split	63.6	62.9	62.9
	Wrapper	44	Cross	75.5	75.2	75.2
			80% Split	76.6	76.4	76.3
NYC	No Filter	1224	Cross	70.5	70.5	70.5
			80% Split	69.1	68.9	68.9
	Chi-square	99	Cross	76.3	76.3	76.3
			80% Split	73.8	73.3	73.3
		98	Cross	76.5	76.5	76.5
			80% Split	73.8	73.3	73.3
		33	Cross	79.7	79.4	79.3
			80% Split	83.3	82.2	82.1
	Wrapper	22	Cross	83.2	82.9	82.9
			80% Split	85.1	84.4	84.4

ARPN Journal of Engineering and Applied Sciences

Table-6. Empirical results using MNB classifier

Dataset	Feature selection method	No. of feature	Validation method	P	R	*F-1*
CAM	No Filter	1775	Cross	73	73	73
			80% Split	66.4	65.9	65.8
	Chi-square	253	Cross	81.7	81.4	81.3
			80% Split	79.4	78	77.9
		93	Cross	69.6	68.9	68.6
			80% Split	61.5	59.8	59
		15	Cross	56.9	56.9	56.9
			80% Split	58.6	56.1	54.4
	Wrapper	63	Cross	65.2	64.7	64.4
			80% Split	75.7	73.2	72.8
Ubuntu	No Filter	1626	Cross	64.1	63.8	63.6
			80% Split	65.7	64	63.7
	Chi-square	139	Cross	74.7	73.9	73.7
			80% Split	80.6	75.3	74.5
		74	Cross	70.3	69	68.5
			80% Split	76.6	70.8	69.6
		10	Cross	62.9	62.9	62.9
			80% Split	70.8	64	61.9
	Wrapper	44	Cross	77.9	76.9	76.6
			80% Split	**81**	**80.9**	**80.8**
NYC	No Filter	1224	Cross	70.3	70.1	70
			80% Split	76.5	75.6	75.4
	Chi-square	99	Cross	82.5	81.8	81.7
			80% Split	**86.5**	**85.6**	**85.5**
		98	Cross	82.7	82	82
			80% Split	86.5	85.6	85.5
		33	Cross	82.3	81.8	81.8
			80% Split	84	82.2	82
	Wrapper	22	Cross	84.6	84.3	84.2
			80% Split	84	81.1	80.8

ARPN Journal of Engineering and Applied Sciences

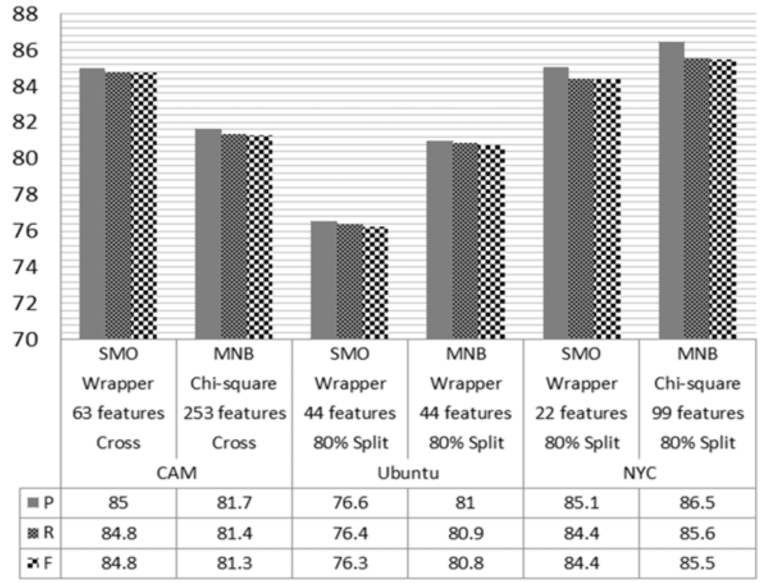


Figure-2. Comparative analysis of the two best classifiers (MNB and SMO) on the three datasets.

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Table-7. Empirical results using LibSVM classifier

Dataset	Feature selection method	No. of feature	Validation method	P	R	*F-1*
CAM	No Filter	1775	Cross	72.8	55.6	45.4
			80% Split	75.3	48.8	33.3
	Chi-square	253	Cross	71.5	66.4	64.3
			80% Split	70.8	59.8	55
		93	Cross	68.9	67.2	66.4
			80% Split	68.8	62.2	59.6
		15	Cross	70.1	69.4	69.1
			80% Split	71.2	68.3	67.6
	Wrapper	63	Cross	61	59.3	57.7
			80% Split	62.3	53.7	46.7
Ubuntu	No Filter	1626	Cross	59	52.6	42.1
			80% Split	22.3	47.2	30.3
	Chi-square	139	Cross	65.9	64	62.9
			80% Split	66.9	60.7	58.1
		74	Cross	66.2	65.2	64.6
			80% Split	65.9	62.9	61.9
		10	Cross	64.7	64.7	64.7
			80% Split	66.6	66.3	66.3
	Wrapper	44	Cross	62.9	60.4	58.5
			80% Split	68.1	67.4	66.7
NYC	No Filter	1224	Cross	69.4	69.4	69.4
			80% Split	59.7	58.9	58.4
	Chi-square	99	Cross	80.5	80.3	80.2
			80% Split	81.5	81.1	81.1
		98	Cross	80.5	80.3	80.2
			80% Split	82.9	82.2	82.2
		33	Cross	80.6	80.5	80.5
			80% Split	81.1	81.1	81.1
	Wrapper	22	Cross	84.1	83.6	83.5
			80% Split	86	85.6	85.5

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-8. Empirical results using J48 classifier

Dataset	Feature selection method	No. of feature	Validation method	P	R	*F-1*
CAM	No Filter	1775	Cross	56.9	56.9	56.9
			80% Split	56.8	56.1	55.9
	Chi-square	253	Cross	64.4	64	63.7
			80% Split	67.5	64.6	63.8
		93	Cross	64.2	63.7	63.4
			80% Split	67.5	64.6	63.8
		15	Cross	69.1	68.1	67.7
			80% Split	70.3	67.1	66.3
	Wrapper	63	Cross	63.5	63.5	63.5
			80% Split	65.4	64.6	64.5
Ubuntu	No Filter	1626	Cross	57.8	57.8	57.7
			80% Split	59.8	59.6	59.6
	Chi-square	139	Cross	58.1	58	57.8
			80% Split	64.3	62.9	62.6
		74	Cross	58.6	58.4	58.2
			80% Split	62.3	60.7	60.2
		10	Cross	60.7	60.7	60.7
			80% Split	64.6	64	64
	Wrapper	44	Cross	64.1	63.6	63.3
			80% Split	66.7	66.3	65.6
NYC	No Filter	1224	Cross	65	65	65
			80% Split	65.6	65.6	65.6
	Chi-square	99	Cross	74.4	74.3	74.2
			80% Split	72.9	72.2	72.1
		98	Cross	74.4	74.3	74.2
			80% Split	72.9	72.2	72.1
		33	Cross	75	74.9	74.9
			80% Split	80.2	78.9	78.7
	Wrapper	22	Cross	81.6	81.2	81.1
			80% Split	81.6	80	79.8

VOL. 11, NO. 3, FEBRUARY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

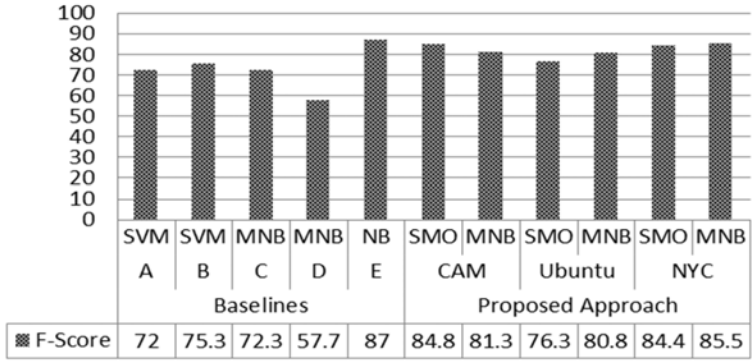
Table-9. Empirical results using MP classifier

Dataset	Feature selection method	No. of feature	Validation method	P	R	*F-1*
CAM	No Filter	1775	Cross	-	-	-
			80% Split	-	-	-
	Chi-square	253	Cross	72.1	71.3	71.1
			80% Split	66.4	65.9	65.8
		93	Cross	62.2	61.8	61.4
			80% Split	57	56.1	55.7
		15	Cross	65	64.5	64.1
			80% Split	65.5	62.2	60.9
	Wrapper	63	Cross	81.9	81.9	81.9
			80% Split	**83**	**82.9**	**82.9**
Ubuntu	No Filter	1626	Cross	-	-	-
			80% Split	-	-	-
	Chi-square	139	Cross	68.7	68.3	68.2
			80% Split	69.1	68.5	68.5
		74	Cross	65.8	64.9	64.4
			80% Split	73.7	71.9	71.7
		10	Cross	61.9	61.3	60.8
			80% Split	63.3	62.9	62.9
	Wrapper	44	Cross	77.6	77.5	**77.5**
			80% Split	**77.9**	**77.5**	77.3
NYC	No Filter	1224	Cross	-	-	-
			80% Split	-	-	-
	Chi-square	99	Cross	72	71.8	71.8
			80% Split	74.2	73.3	73.2
		98	Cross	73.5	73.4	73.4
			80% Split	75.8	75.6	75.5
		33	Cross	71.6	71.6	71.6
			80% Split	76.8	76.7	76.7
	Wrapper	22	Cross	**83.2**	**82.9**	**82.9**
			80% Split	82.9	82.2	82.2

Baselines

We consider the five works of Table-1 as our baselines. The comparative analysis of the baselines and our proposed approach is shown in Figure-3. The F-measure metric is used for the comparison. Our proposed approach outperformed four out of the five baselines with the two classifiers and the three datasets. The work of [10] that slightly outperformed our approach was actually based on a set of selected opinion words. This in a way is similar to feature selection method proposed in this study. Our approach generally selects viable words which have the potential of revealing important latent words.

ARPN Journal of Engineering and Applied Sciences



A= Li et al., 2008, B = Aikawa et al., 2011, C = Biyani et al., 2014
D = Bhatia *et al*., 2015, E = Yu and Hatzivassiloglou, 2003

Figure-3. Comparing proposed approach with baselines

CONCLUSIONS

In this paper, we addressed bag-of-words feature analysis for detecting web forum question post. Web forum question post detection is treated as a classification problem. The contributions of the paper are:

a) We evaluate the performance of different feature selection approaches on web forum data. We confirm that filter method favours less technical dataset while wrapper method performs better on highly technical and non-technical datasets.
b) We confirm that higher thresholds for filter method will reduce feature space but may not enhance performance. This confirms that filter method irrespective of thresholds will not optimize feature selection.

We use highly technical, less technical and non-technical datasets to establish that strong classification algorithms will be consistent with different forum genres.

In this study, we have been able to show that simple bag-of-words with dimensionality reduction can outperform highly expensive n-gram approaches.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education Malaysia, Kaduna Polytechnic, Kaduna, Nigeria and Soft Computing Research Group (SCRG) of Universiti Teknologi Malaysia (UTM). The work was also supported in part by grant from Vote 4F373.

REFERENCES

[1] G. Cong, L. Wang, C.-Y. Lin, Y.-I. Song and Y. Sun. 2008. Finding question-answer pairs from online forums, In: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. pp. 467-474.

[2] J. Kim, G. Chern, D. Feng, E. Shaw, and E. Hovy. 2006. Mining and assessing discussions on the web through speech act analysis, In: Proceedings of the Workshop on Web Content Mining with Human Language Technologies at the 5th International Semantic Web Conference.

[3] J. Kim, E. Shaw, D. Feng, C. Beal, and E. Hovy. 2006. Modeling and assessing student activities in on-line discussions, In: Proc. of the AAAI Workshop on Educational Data Mining.

[4] J. Bian Y., Liu D., Zhou E., Agichtein and H. Zha. 2009. Learning to recognize reliable users and content in social media with coupled mutual reinforcement. Paper presented at the Proceedings of the 18th international conference on World wide web, Madrid, Spain.

[5] L. Hong and B. D. Davison. 2009. A classification-based approach to question answering in discussion boards, In: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. pp. 171-178.

[6] L. Sun, B. Liu, B. Wang, D. Zhang and X. Wang. 2010. A study of features on Primary Question detection in Chinese online forums, in Fuzzy Systems and Knowledge Discovery (FSKD), Seventh International Conference. pp. 2422-2427.

[7] R. Catherine, A. Singh, R. Gangadharaiah, D. Raghu and K. Visweswariah. 2012. Does Similarity Matter? The Case of Answer Extraction from Technical Discussion Forums, in COLING (Posters). pp. 175-184.

[8] P. Deepak and K. Visweswariah. 2014. Unsupervised Solution Post Identification from Discussion Forums, In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, Maryland, USA. pp. pages 155–164.

[9] N. Kumar, K. Srinathan, and V. Varma. 2015. Unsupervised Deep Semantic and Logical Analysis for Identification of Solution Posts from Community Answer in International Journal of Information and Decision Sciences. Report No: IIIT/TR/2015/-1.

[10] H. Yu and V. Hatzivassiloglou. 2003. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences, In: Proceedings of the 2003 conference on Empirical methods in natural language processing. pp. 129-136.

[11] B. Wang, B. Liu, C. Sun, X. Wang and L. Sun. 2009. Extracting Chinese question-answer pairs from online forums, in Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on. pp. 1159-1164.

[12] B. Li, Y. Liu, and E. Agichtein. 2008. Cocqa: co-training over questions and answers with an application to predicting question subjectivity orientation, In: Proceedings of the conference on empirical methods in natural language processing. pp. 937-946.

[13] P. Biyani, S. Bhatia, C. Caragea and P. Mitra. 2014. Using non-lexical features for identifying factual and opinionative threads in online forums, Knowledge-Based Systems, 69 (October 2014), 170-178.

[14] N. Aikawa, T. Sakai, and H. Yamana. 2011. Community QA Question Classification: Is the Asker Looking for Subjective Answers or Not? IPSJ Online Transactions. 4: 160-168.

[15] S. Bhatia, P. Biyani, and P. Mitra. 2015. Identifying the role of individual user messages in an online discussion and its use in thread retrieval, Journal of the Association for Information Science and Technology, doi: 10.1002/asi.23373

[16] B. Sumit, B. Prakhar, and M. Prasenjit. 2012. Classifying User Messages for Managing Web Forum Data, Fifteenth International Workshop on the Web and Databases (WebDB 2012), Scottsdale, AZ, USA.