Journal of Computer Science 10 (11): 2232-2239, 2014
ISSN: 1549-3636

doi:10.3844/jcssp.2014.2232.2239 Published Online 10 (11) 2014 (http://www.thescipub.com/jcs.toc)

A COMPARATIVE STUDY OF COMBINED FEATURE SELECTION METHODS FOR ARABIC TEXT CLASSIFICATION

Aisha Adel, Nazlia Omar and Adel Al-Shabi

Knowledge Technology Group, Centre for AI Technology, Faculty of Information Science and Technology, University Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Received 2014-04-09; Revised 2014-10-15; Accepted 2014-11-11

ABSTRACT

Text classification is a very important task due to the huge amount of electronic documents. One of the problems of text classification is the high dimensionality of feature space. Researchers proposed many algorithms to select related features from text. These algorithms have been studied extensively for English text, while studies for Arabic are still limited. This study introduces an investigation on the performance of five widely used feature selection methods namely Chi-square, Correlation, GSS Coefficient, Information Gain and Relief F. In addition, this study also introduces an approach of combination of feature selection methods based on the average weight of the features. The experiments are conducted using Naïve Bayes and Support Vector Machine classifiers to classify a published Arabic corpus. The results show that the best results were obtained when using Information Gain method. The results also show that the combination of multiple feature selection methods outperforms the best results obtain by the individual methods.

Keywords: Feature Selection, Combination Method, Arabic Text Classification

1. INTRODUCTION

With the rapid growth of the Internet, the volume of the news and information available on the web is growing exponentially. Since there has been an explosion of information available on the Internet, this makes the process of analyzing and processing them manually a very difficult task. As a consequence, text classification has gained importance in hierarchical organization of these documents. The fundamental goal of the text classification is to classify texts into appropriate classes.

One of the problems of text classification is the huge number of features which reduce the performance of text classification and consume the time. Feature selection method is used to reduce the feature space by selecting the most relevant features (Maldonado and L'Huillier, 2013). Many feature selection methods have been proposed and investigated to improve the performance of English text classification. However, the work on feature selection for Arabic language are limited and most of studies in text classification for Arabic language are concerned with investigating the efficiency of text classification algorithms without enough attention to how the feature selection task can improve the accuracy of classification (Al-Salemi and Ab Aziz, 2010; Hawashin *et al*. 2013; Saad, 2011).

Our motivation to do this research is to enhance the robustness of the finally selected feature subsets of the class and get rid of the noisy and redundant features because there is another subset which supplies the same information

Corresponding Author: Nazlia Omar, Knowledge Technology Group, Centre for AI Technology, Faculty of Information Science and Technology, University Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

about the class. We need to combine two methods or more together to get rid of redundant and noisy features which degrade the performance of most classifiers.

This study introduces an investigation on the performance of five widely used feature selection methods and a combination approach of feature selection methods including Chi-square, Correlation, GSS Coefficient, Information Gain and Relief F. The main concern is to investigate the effectiveness of combining the individual feature selection methods on the performances of Arabic text classification.

The rest of the paper is organized as follows: Section two for related work, section three is methodology, section four is experimental work, section five gives the results and discussion of the experiments and section six concludes the paper.

2. RELATED WORK

Many feature selection and other preprocessing techniques have been applied for text classification. The bulk of feature selection work has been devoted for English and other Latin language.

Soares (2010) has proposed an algorithm based on a wrapper method to build an ensemble of models with a specific base classifier. The Class-Specific Ensemble Feature Selection (CEFS) algorithm applied in the above test with Naïve-Bayesian Classifier. The results showed an enhancements in the accuracy of prediction. Ren and Sohrab (2013) have introduced class-indexing-based term-weighting approaches. The proposed class-based indexing is incorporated with term, document and class index. They have investigated the efficiency of proposed class-indexing-based approaches, with other term weighing approaches to address the automatic text classification task. The results of the experiments have revealed that, the proposed term weighting approaches improved the classification task. Chen *et al*. (2009) have proposed two feature evaluation metrics for the Naïve Bayesian classifier, applied on multiclass text datasets: Multi-class Odds Ratio (MOR) and Class Discriminating Measure (CDM). Experiments of text classification have been carried out with Naïve Bayesian classifier. They have compared CDM and MOR with Information Gain and three variations of Odds Ratio. The results have indicated that, CDM and MOR gained better selection efficiency, compared to other feature selection methods.

Unlike English language, a limited number of researches had been done for Arabic (Al-Salemi and Ab Aziz, 2010; Chantar and Corne 2011; Hawashin *et al*., 2013).

An investigation on three representation methods had been performed by Harrag *et al*. (2010) namely Document Frequency (DF), Latent Semantic Analyses (LSA) and Term Frequency Inverse Document Frequency (TFIDF). The results of experiments that performed on an Arabic dataset showed that TFIDF was the most effective method among the three feature reduction techniques. Duwairi *et al*. (2007) have compared and contrasted two feature selection techniques when applied to Arabic corpus. The dataset consisted of manually prepared Arabic text documents, collected from internet sites. They have employed stemming and light stemming as feature selection methods. The experiments have showed that the using of light stemming as a feature selection method obtained better results than using stemming.

Generally speaking, the work on feature selection for Arabic language used individual methods only, while using combination of feature selection methods may achieve better.

3. METHODOLOGY

Feature selection is an important preprocessing stage of text classification, which increases the performance of a predictive model. The main purpose of feature selection is to choose a subset of high discriminative features and eliminate the non-discriminative features.

In this study, we investigate the performance of five common feature selection methods with their combinations for Arabic text classification. We combine every two feature selection methods and we also combine the five feature selection methods. In both cases, two classifiers are used to conduct the experiments namely Naïve Bayes and Support Vector Machine. In literature, the studies that tried to combine the feature selection methods using different strategies, they combine either two or five feature selection methods like (Wang *et al*., 2010; Vege, 2012).

The key idea behind combining feature selection methods are that every individual method produces different types of errors and feature selection methods are combined to exploit their strengths. Combining feature selection methods are becoming more popular as they allow one to overcome the weaknesses of single methods. The combined feature selection methods always outperforms the best of its individuals in text classification task (Omar *et al*., 2013).

The following subsections describe briefly the classifiers and the feature selection methods used in this study. They also describe the used approach of combination of feature selection methods.

3.1. Naïve Bayes Classifier

Naïve Bayes has been one of the most popular machine learning methods since long ago. Its simplicity makes the framework attractive in different tasks and reasonable

performances are achieved in the tasks although this way of learning is based on an unrealistic independence assumption (Khalifa and Omar, 2014). The Naïve Bayes (NB) classifier generally uses Bayes' rule:

$$p(c_i \mid d) = \frac{p(c_i)\, p(d \mid c_i)}{p(d)}$$

where, $p(c_i \mid d)$ is the posterior probability of class C_i given a new document d, $p(C_i)$ is the probability of class C_i which can be calculated by:

$$p(c_i) = \frac{N_i}{N}$$

where, N_i is the number of documents assigned to class C_i and N is the number of classes, $p(d \mid c_i)$ is the probability of a document d given a class C_i and $p(d)$ is the probability of document d.

3.2. SVM Classifier

Support Vector Machine (SVM) classification algorithm is considered as one of the most robust and accurate machine learning algorithms (Ahmed, 2010). In simple words, given a group of training examples, each marked as belonging to one of two categories, *SVMs* training algorithm starts building a model that predicts whether a new example falls into which category. The method of SVM in its dual form is as follows:

$$\vec{\alpha} = arg\ min\{-\sum_{i=1}^{n}\alpha i + \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_i y_i y_i(\vec{x_i}, \vec{x_j})\}$$

$$Subjectto: \sum_{i=1}^{n}\alpha_i\, y_i = 0; 0 \le \alpha_i \le c$$

3.3. Feature Selection Methods

3.3.1. Chi-Square

Chi-square is a commonly used statistical test that determines the divergence from the distribution expected if one assumes the feature occurrence is obviously independent of the class value. As a statistical test, it is well known to act erratically for very minor expected counts, which are known in text classification both because of the rare occurring of word features and some other times because of having different positive training examples for a concept (Forman, 2008). The chi-square statistics is calculated by the following equation.

$$\chi^2(c,t) = \frac{N \times (AD\text{-}BC)}{(A+C)(B+C)(A+B)(C+D)}$$

Where:
A = The frequency of t and c occurrences;
B = The frequency of t occurrences without c,
C = The frequency of c without t,
D = The frequency of non-occurrence of both c and t and N is the quantity of document.

3.3.2. Correlation

Correlation-based Feature Selection (CFS) is one of commonly known techniques to evaluate and rank the relevance of features by measuring correlation between features and classes and between some features and others (Suganya and Rajaram, 2012).

Given number of features k and classes C, CFS defines relevance of features subset by the use of Pearson's correlation equation:

$$Merit_s = \frac{kr_{kc}}{\sqrt{k + (k-1)r_{kk}}}$$

where, Merits is considered as the relevance of feature subset, r_{kc} is the average linear correlation coefficient between these features and classes and r_{kk} is the average linear correlation coefficient between different features.

3.3.3. Galavotti-Sebastiani-Simi (GSS) Coefficient

GSS method have been proposed as a simplified Chi Square statistic. The P N factor and the denominator have completely removed. The denominator have also removed, because the denominator gives high correlation coefficient score to rare words and rare categories (Uchyigit and Ma, 2008). The GSS CC value can be computed as follows:

$$GSS(t,c_i) = P(t,c_i).P(\bar{t},\bar{c_i}) - P(t,\bar{c_i}).P(\bar{t},c_i) \approx \frac{AD - CB}{N^2}$$

where,
N = The total number of training document,
A = The number of documents in class c_i that contain t.
B = The number of documents that contain the term t in other classes.
C = The number of documents in class c_i that does not contain the term t.
D = The number of documents that does not contain the term t in other classes.

3.3.4. Information Gain

Information gain represents the entropy reduction given a certain feature, that is, the number of bits of information gained about the category by knowing the

presence or absence of a term in a document (Ramalingam and Zheng, 2013):

$$IG(t) = -\sum_{i=1}^{|c|} p(c_i) log p(c_i)$$

$$+p(t)\sum_{i=1}^{|c|} p(c_i \mid t) log\ p(c_i \mid t)$$

$$+p(\bar{t})\sum_{i=1}^{|c|} p(c_i \mid \bar{t}) log\ p(c_i \mid \bar{t})$$

where, $p(c_i)$ represents the likelihood of the occurrence of *ci* class; $p(t)$ represents the likelihood of the occurrence of *t*; $p(\bar{t})$ represents the likelihood of the non-occurrence of $\bar{t}$.

3.3.5. Relief

Relief-f (Bolón-Canedo *et al*., 2013; Zhang and Sawchuk, 2011) is a commonly used metric for feature ranking that estimate the relevance of features according to how well its values distinguish the sampled instance from its nearest hit (instance of the same class) and nearest miss (opposite class). The Relief feature selection algorithm selects feature-instances randomly from the training data. For each sampled instance, the nearest hit and the nearest miss is found. A high weight is assigned to a feature if it differentiates between instances from different classes and has the same value for instances of the same class. Specifically, it tries to find a best estimate from the following probabilities to allocate as the weight for each term feature *f* (Sharma and Dey, 2012):

W_x = (different value of *f*/nearest instances from different classes)-(different value of *f*/nearest instances from same class)

4. COMBINATION OF FEATURE SELECTION METHODS

The main reason for the combination of feature selection methods is to compensate for the shortcomings of individual methods. Combining methods is a common technique in machine learning. The method used in this study is based on combination of feature selection techniques including Chi-square, Information Gain, Relief, Correlation and GSS Coefficient. First, it combines the top ranked n features resulted from k feature selection methods. Second, it calculates the average weight of every term obtained from every feature selection method using the following formula.

$$avg_{weight} = \frac{\sum_{i=1}^{k} weight_i}{k}$$

where, *k* is the number of ranking lists. Then, the method sorts the features according to the new weights and selects top m ranked features to form the final list of features. **Figure 1** shows the steps of the combination method.

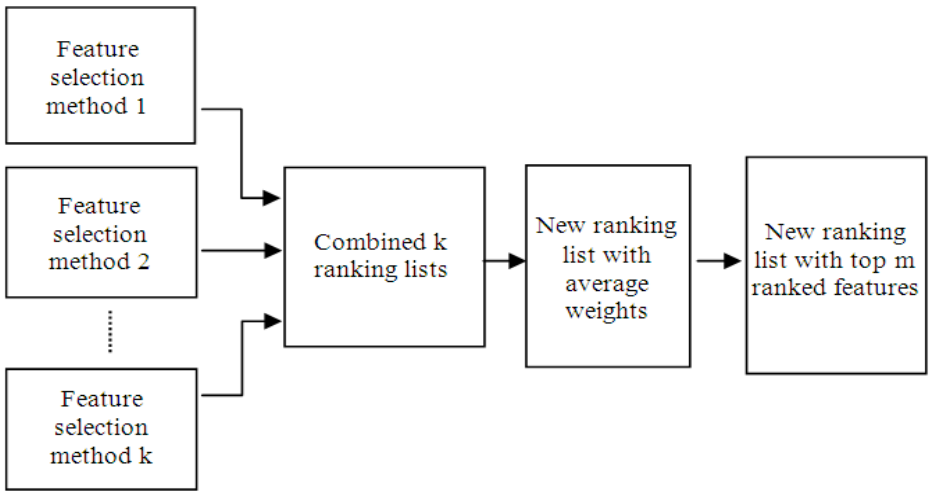

Fig. 1. Combination of multiple feature selection methods

5. EXPERIMENTAL WORK

Three different types of experiments have been performed to investigate the performance of five widely used feature selection methods including Chi-square, Correlation, GSS Coefficient, Information Gain and Relief F and their combinations. Every type of these experiments is conducted using two classifiers namely Naïve Bayes and Support Vector Machine. From other hand, two types of combination are performed. The first one combines the ranking lists resulting from two feature selection methods. The second one combines the ranking lists resulting from the five methods used in this study.To do so, the evaluation is performed on CNN Arabic published dataset. The dataset and performance measures used in this study will be described briefly.

5.1. CNN Arabic Corpus

This study uses CNN Arabic corpus which is collected by Saad (2011). This dataset contains 5,070 text documents. Each of them belongs to one of the six classes as shown in **Table 1**.

5.2. Performance Measures

In order to evaluate the feature selection methods the F1-measure is used which combines precision and recall. For ease of comparison, the Macro-averaged (Macro-F1) is used. Precision, recall, F_1 measure and macro F_1 are calculated using the following formulas consequently:

$$precision = \frac{true\ \ positive}{true\ \ positive + false\ \ positive}$$

$$recall = \frac{true\ \ positive}{true\ \ positive + false\ \ negative}$$

$$F1measure = \frac{2 * precision * recall}{precision + recall}$$

$$F_1^{macro} = \frac{1}{m}\sum_{i=1}^{m} F_1(c_i)$$

Table 1. Classes in CNN Arabic dataset

Class name	Documents no
Middle East news	1462
Entertainment	474
World news	1010
Sport news	762
Science news	526
Business	836

6. RESULTS

In order to compare the performance of the previously mentioned feature selection methods and to investigate the performance of the combination method, Naïve Bayes and SVM classifiers are used. After applying each feature selection method, the classification is performed with varying number of selected features. The experiments were carried out on published Arabic corpus namely CNN dataset. **Table 2** shows the results obtained when using every feature selection method individually with Naïve Bayes classifier.

Table 2 shows that the best result obtained is 87.3% Macro-F1 by using Information Gain method when number of features is 3000. The lowest result is 66.9% Macro-F1 obtained by Correlation feature selection method when the number of selected features is 500.

Table 3 shows the results obtained when using every feature selection method individually with SVM classifier.

Table 3 shows that the best result obtained is 93.1% Macro-F1 by using Information Gain method when number of features is 4000. The lowest result is 71.4% Macro-F1 obtained by Correlation feature selection method when the number of selected features is 500.

The results in **Table 2 and 3** show that the best performance among the used feature selection methods is obtained by Information Gain method. The lowest performance is achieved by Correlation method in most cases because it depends on the correlation between features and may not take into account the correlation with class.

Table 4 shows the results obtained when using all possible combinations of two feature selection methods with Naïve Bayes classifier.

Table 5 shows the results obtained when using all possible combinations of two feature selection methods with SVM classifier.

The results in **Table 4 and 5** show that the binary combination of the feature selection methods outperform the results in **Table 2 and 3** consequently which obtained using individual methods.

Table 6 shows the results obtained when combining the ranking lists obtained from the five feature selection methods with Naïve Bayes classifier.

Table 7 shows the results obtained when combining the ranking lists obtained from the five feature selection methods with SVM classifier.

Table 6 and 7 show the performance of the combination of multiple feature selection methods namely Chi-square, Information Gain, Relief, Correlation and GSS Coefficient with Naïve Bayes and SVM classifiers consequently.

Table 2. The results of individual FS methods with NB

Macro-F						
Feature size FS method	500.0	1000.0	2000.0	3000.0	4000.0	5000.0
Chi-square	79.4	84.6	86.8	86.8	86.5	85.6
Info Gain	79.1	85.1	87.2	87.3	86.4	86.1
Relief-F	76.4	82.3	84.6	86.6	86.2	86.2
GSS	78.3	83.2	85.9	86.5	86.6	85.1
Correlation	66.9	77.7	85.7	87.1	86.4	85.3

Table 3. The results of individual FS methods with SVM

Macro-F						
Feature size FS method	500	1000	2000	3000	4000	5000
Chi-square	86.1	90.2	92.1	92.5	92.8	92.7
Info Gain	86.6	90.3	92.2	92.8	93.1	92.8
Relief-F	84.7	89.6	92.1	92.7	92.9	92.5
GSS	86.6	90.4	92.4	92.4	92.9	92.8
Correlation	71.4	83.9	91.3	91.9	92.4	92.6

Table 4. The results of binary combination of FS methods with NB classifier

Macro-F						
Feature size FS method	500.0	1000.0	2000.0	3000.0	4000.0	5000.0
Chi and I. Gain	79.5	84.9	88.7	87.8	87.7	86.1
Chi and Relief F	79.2	86.4	87.3	88.1	87.3	86.9
Chi and GSS	79.3	84.4	86.7	87.5	87.2	86.8
Chi and Corr.	70.8	79.7	86.6	88.9	87.5	85.8
I. Gain and Relif F	79.3	87.3	88.7	89.4	88.4	87.8
I. Gain and GSS	78.4	83.6	86.8	87.4	87.5	86.2
I. Gain and Corr.	71.1	81.2	87.7	89.1	88.6	87.3
Relif-F and GSS	79.8	83.6	86.5	86.7	87.9	87.1
Relif-F and Corr.	68.3	75.4	79.2	84.2	85.8	86.3
GSS and Corr.	69.7	73.9	80.6	85.3	86.7	87.1

Table 5. The results of binary combination of FS methods with SVM classifier

Macro-F						
Feature size FS method	500.0	1000.0	2000.0	3000.0	4000.0	5000.0
Chi and I. Gain	87.4	91.6	94.1	94.7	94.2	93.8
Chi and Relief F	86.5	91.1	93.8	94.5	94.3	93.7
Chi and GSS	87.2	91.5	93.8	94.2	94.3	93.9
Chi and Corr.	85.8	90.7	93.4	93.8	93.7	93.4
I. Gain and Relif F	87.3	91.5	93.8	94.5	94.8	93.7
I. Gain and GSS	87.4	91.7	94.1	94.5	94.2	93.9
I. Gain and Corr.	84.7	90.5	93.5	94.3	94.3	93.8
Relif-F and GSS	87.2	91.5	93.8	94.6	94.4	93.7
Relif-F and Corr.	81.9	89.7	93.6	94.4	94.7	93.9
GSS and Corr.	84.1	90.7	93.2	93.6	93.8	93.7

Table 6. The results of combination of five FS methods with NB classifier

Macro-F						
Feature size	500.0	1000.0	2000.0	3000.0	4000.0	5000.0
Combination of 5FS methods	80.3	86.6	89.5	90.1	88.3	87.4

Table 7. The results of combination of five FS methods with SVM classifier

Macro-F						
Feature size	500.0	1000.0	2000.0	3000.0	4000.0	5000.0
Combination of 5FS methods	87.9	91.8	94.5	95.1	94.7	94.2

Tables 6 and 7 show that the Macro F1 results after using the combination of multiple feature selection methods outperform the results obtained by using every method individually.

7. DISCUSSION

The performance obtained of combination of feature selection methods in both scenario of combination, binary and multiple indicates that the combination of feature selection methods are indeed informative for text classification tasks especially when the number of features is extremely large (1000 and above).

Table 6 and 7 shows that the performance became more stable especially with increase size of selected features when applying the combination of feature selection methods. The average improvement rate after applying binary and multiple combination ranges between (1-2.5%) in macro-average *F*1.

The results in **Table 4 to 7** show that the combination of feature selection methods performs better than individual methods since the success of the methods depends on various variables. It is more likely that the combination of different feature selection methods obtains more effective performance in text classification as they allow one to overcome the weaknesses of single approaches. The combined methods always outperform the best of its individuals in feature selection task (Omar *et al*., 2013). The previous studies (Harrag *et al*., 2010; Omar *et al*., 2013; Soares, 2010; Vege, 2012) pointed out that using combined methods improve the results, in spite of the different techniques of combination and the different purposes of using combined methods, this point deals with our findings.

In this study, the focus was on combination of feature selection methods to select a subset of features that most representative of the class. The interesting opportunity for future research in this area is towards the investigation of combining the features using different representation techniques (i.e., bigram, trigram or n-gram) to propose a new method that can select the most discriminative features of the class which can be more useful for text classification.

8. CONCLUSION

This study introduced a combination of feature selection methods to improve the performance of text classification. First, we evaluate the performance of five common feature selection methods on a published Arabic dataset. Then, we evaluate the performance of all possible binary combinations of these five methods. Finally, we evaluate the performance of the combination of the five methods in order to determine the most appropriate features for classification. Comparing the performance of the individual methods with the performance of the combination methods shows that combining two feature selection methods outperforms the individual methods, while combining the five methods significantly improves the classification performance. Although many feature selection methods exist in text categorization, it is hard to state one is generally superior to others since the success of the methods depends on various variables. It is more likely that combining different feature selection methods obtains more effective performance in text categorization.

9. REFERENCES

Ahmed, I.A., 2010. Myocardial-infraction based on intelligent techniques. Am. J. Applied Sci., 7: 349-349. DOI: 10.3844/ajassp.2010.349.351

Al-Salemi, B. and M.J. Ab Aziz, 2010. Statistical bayesian learning for automatic Arabic text categorization. J. Comput. Sci., 7: 39-39. DOI: 10.3844/jcssp.2011.39.45

Bolón-Canedo, V., N. Sánchez-Maroño and A. Alonso-Betanzos, 2013. A review of feature selection methods on synthetic data. Know. Inform. Syst., 34: 483-519. DOI: 10.1007/s10115-012-0487-8

Chantar, H.K. and D.W. Corne, 2011. Feature subset selection for Arabic document categorization using BPSO-KNN. Proceedings of 3rd World Congress on Nature and Biologically Inspired Computing, Oct. 19-21, IEEE Xplore Press, Salamanca, pp: 546-551. DOI: 10.1109/NaBIC.2011.6089647

Chen, J., H. Huang, S. Tian and Y. Qu, 2009. Feature selection for text classification with naïve Bayes. Expert Syst. Applic., 36: 5432-5435. DOI: 10.1016/j.eswa.2008.06.054

Duwairi, R., M. Al-Refai and N. Khasawneh, 2007. Stemming versus light stemming as feature selection techniques for Arabic text categorization. Proceedings of the 4th International Conference on Innovations in Information Technology, Nov. 18-20, IEEE Xplore Press, Dubai, pp: 446-450. DOI: 10.1109/IIT.2007.4430403

Forman, G., 2008. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res., 3: 1289-1305.

Harrag, F., E. El-Qawasmeh and A.M.S. Al-Salman, 2010. A Comparative Study of Statistical Feature Reduction Methods for Arabic Text Categorization. In: Networked Digital Technologies, Zavoral, F., J. Yaghob, P. Pichappan and E. El-Qawasmeh (Eds.)., Springer Berlin Heidelberg, pp: 676-682. ISBN-10: 978-3-642-14305-2

Hawashin, B., A.M. Mansour and S. Aljawarneh, 2013. An efficient feature selection method for Arabic text classification. Int. J. Comput. Applic., 83: 1-6. DOI: 10.5120/14666-2588

Khalifa, K. and N. Omar, 2014. A hybrid method using lexicon-based approach and naive Bayes classifier for Arabic opinion question answering. J. Comput. Sci., 10: 1961-1968. DOI: 10.3844/jcssp.2014.1961.1968

Maldonado, S. and G. L'Huillier, 2013. SVM-based feature selection and classification for email filtering. Proceedings of the International Conference, Feb. 6-8, Springer Berlin Heidelberg, Algarve, pp: 135-148. DOI: 10.1007/978-3-642-36530-0_11

Omar, N., M. Albared, A.Q. Al-Shabi and T. Al-Moslmi, 2013. ensemble of classification algorithms for subjectivity and sentiment analysis of arabic customers' reviews. Int. J. Adv. Comput. Technol., 5: 77-77.

Ramalingam, P. and B. Zheng, 2013. Optimization of a computer-aided detection scheme using a logistic regression model and information gain feature selection method. Global J. Breast Cancer Res., 1: 1-7.

Ren, F. and M.G. Sohrab, 2013. Class-indexing-based term weighting for automatic text classification. Inform. Sci., 236: 109-125. DOI: 10.1016/j.ins.2013.02.029

Saad, G., 2011. Evolutionary psychology in the business sciences. 1st Edn., illustrated, Springer Science and Business Media, ISBN-10: 3540927840, pp: 397.

Sharma, A. and S. Dey, 2012. A comparative study of feature selection and machine learning techniques for sentiment analysis. Proceedings of the ACM Research in Applied Computation Symposium, Oct. 23-26, San Antonio, TX, USA, DOI: 10.1145/2401603.2401605

Soares, C., 2010. A class-specific ensemble feature selection approach for classification problems. Proceedings of the 48th Annual Southeast Regional Conference, Apr. 15-17, Oxford, MS, USA. DOI: 10.1145/1900008.1900054

Suganya, R. and S. Rajaram, 2012. Content based image retrieval of ultrasound liver diseases based on hybrid approach. Am. J. Applied Sci., 9: 93-938. DOI: 10.3844/ajassp.2012.938.945

Uchyigit, G. and M.Y. Ma, 2008. Personalization Techniques and Recommender Systems. 1st Edn., World Scientific, ISBN-10: 9812797025.

Vege, S.H., 2012. Ensemble of feature selection techniques for high dimensional data. MSc. Theses, Vege, Sri Harsha.

Wang, H., T.M. Khoshgoftaar and A. Napolitano, 2010. A comparative study of ensemble feature selection techniques for software defect prediction. Proceedings of the 9th International Conference on Machine Learning and Applications, Dec. 12-14, IEEE Xplore Press, Washington, DC, pp: 135-140. DOI: 10.1109/ICMLA.2010.27

Zhang, M. and A.A. Sawchuk, 2011. A feature selection-based framework for human activity recognition using wearable multimodal sensors. Proceedings of the 6th International Conference on Body Area Networks, Nov. 07-10, Beijng, China, pp: 92-98. PMID: 2318798