Journal of Theoretical and Applied Information Technology
29th February 2016. Vol.84. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

BINARY RELEVANCE (BR) METHOD CLASSIFIER OF MULTI-LABEL CLASSIFICATION FOR ARABIC TEXT

[1] ADIL YASEEN TAHA, [2] SABRINA TIUN

Faculty of Information Science and Technology

E-mail: [1]daloon.yaser@yahoo.com, [2]sabrinatiun@ukm.edu.my

ABSTRACT

Multi-label text classification has become progressively more important in recent years, where each document can be given multiple labels concurrently. Multi-label text classification is a main challenging task because of the large space of all potential label sets, which is exponential to the number of candidate labels. Among the disadvantages of the earlier multi-label classification methods is that they typically do not scale up with the number of specific labels and the number of training examples. A large amount of computational time for classification is required for a large amount of text documents with high dimensionality, especially, the Arabic language which has a very complex morphology and rich in nature. Furthermore, current researches have paid a little attention to the multi-label classification for Arabic text. Hence, this study aims to design and develop a new method for multi-label text classification for Arabic texts based on a binary relevance method. This binary relevance is made up from a different set of machine learning classifiers. The four multi-label classification approaches, namely: the set of SVM classifiers, the set of KNN classifiers, the set of NB classifiers and the set of the different type of classifiers were empirically evaluated in this research. Moreover, three feature selection methods (Odd ratio, Chi-square and Mutual information) were studied and their performances were investigated to enhance the performance of the Arabic multi-label text classification. The objective is to efficiently incorporate classification algorithms and feature selection to create a more accurate multi-label classification process. To evaluate the model, a manually standard interpreted data is used. The results show that the machine learning binary relevance classifiers which consists from a different set of machine learning classifiers attains the best result. It has achieved a good performance, with an overall F-measure of 86.8% for the multi-label classification of Arabic text. Besides, the results show an important effect from the used feature selection methods on the classification. Distinctly, the set of the different set of algorithms proves to be an efficient and suitable method for the Arabic multi-label text classification.

Keywords: *Arabic Text Classification, Multi-label Classification, Feature Selection, Statistical Methods*

1. INTRODUCTION

Text categorization (TC) is regarded as decision-making criteria which is very useful in content analysis and other such operation, but it holds if a particular text piece belongs to specific prescribed category [1]. On the other hand, the text categorization faces some crucial problems during data mining and machine learning emerge from union of bulk information available on Internet which amounts the information libraries available and electronic documentation. The concept of text categorization is to assign one document into different categories depending on its contents.

Supervised learning and unsupervised learning are the two machine learning methods. The former method is used to categorize documents by allotting predefined categories to address new documents which are usually used before text categorization systems with a high performance as an added advantage. However, unfortunately there are certain disadvantages of this method such as overlapping among the categories which occurs due to the large volume of labeled training documents essential for label allotment [2]. Contrary, unsupervised learning saves human efforts as compared to document's class label. However, supervised learning method will be used in the current study to scrutinize Arabic electronic documents. Supervised learning is based on two sorts of labels, i.e. a single label (first one) and multi-label (the other one).

The single-label text categorization (TC) problem assigns only one predefined category to every invisible Natural language, in case when there are two or more than two categories in category space. Due to the overlapping nature of text with one another in category space, the categorization of every single document becomes impossible. For

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

instance, *politics*, the field frequently overlaps with *economic*s. It hinders the single-label TC mission, on the other hand, in multi-label "TC any number of classifier from zero to N "(N means at least one) may be specified in the document [3]. More than one predefined class is assigned by the multi-label TC to an "unseen" document and it is called as overlapping TC tasks due to the target of assigning an object concomitantly to 1 or multiple categories.

According to Addis [4], the single-label is more general than the multi-label classification due to the algorithm for single-label can be used for the multi-label, by converting a problem of multi-label with categories,$\{c1,c2.....,cn\}$ into n independent problems of single-label categorization with classes, for $i = 1,2,......,m$. The single-label is more general. This can be achieved if the classifications are randomly free of one another. However, generally the converse is not true. On the off chance that there is an algorithm to perform multi-label, it does not imply that it can be used for single-label order of text document and regularly categorized as non-covering [5]. In this label, for a given whole number k every component of C must be doled out to precisely k *(or < k, or >k)* components D.

Multi-label text categorization is a key stimulating task. The challenge comes from the large space of all possible label sets, which is exponential to the number of candidate labels. In the machine learning context, most of the research has been done in traditional single-label text categorization [6]. Multi-label categorization is a harder task than single-label categorization by nature, because an instance may be linked with multiple labels. Consequently, the single-label classifiers are easier to handle than the previous multi-label methods. One of the most common setbacks of the previous multi-label categorization methods is that they usually do not scale up with the number of distinct labels and the number of training examples.

A large amount of computational power for categorization is required for large amount of text documents with high dimensionality and in particular in the Arabic language which has a rich nature and very complex morphology. Moreover, accuracy of the multi-label text categorization based on machine learning algorithms in classifying multi-labeled Arabic documents written in different epoch, region and with a different style is the main problem here. Therefore, this study aims to propose a binary relevance classifier that consists of different classifiers (not necessary the same) to solve multi-label categorization problem.

2. RELATED WORK

Several approaches have been proposed in terms of Arabic text categorization. For instance, a smart Arabic text categorization system was proposed by Syaim et al. [7] based on Machine learning algorithms; the proposed system was employed for stemming and selection. They have employed normalized-TFIDF schema for Arabic text classification. They have utilized the Arabic text classification on a repository comprising of over 1K documents from three Egyptian newspapers (El Ahram, El Gomhoria and El Akhbar). They have dealt with 6 groups documents: (i) 233 documents from *Arts*; (ii) 233 documents from *Economics*; (iii) 280 documents from *Politic*s;(iv) 102 documents from *Information Technology*; (v) 121 documents related to *Women*; and (vi) 231 documents related to *Sports*. They have proposed a hybrid approach by incorporating Document Frequency and Information Gain, and established it to be ideal stemming for Arabic text; they have yielded a precision of 98%.

Hmeidi I, et al. [8] have studied Arabic text classification by making use of two machine learning approaches, such as, K nearest neighbour (KNN) and support vector machines (SVM). They have developed a unique compilation with the help of assorted news articles for training and testing. They have demonstrated that, their approach was effective, and however, revealed that, SVM was superior in terms of forecasting.

Hussien et al [9] have integrated the Sequential Minimal Optimization (SMO), Naïve Bayesian (NB) and J48 (C4.5) Algorithms and make use of weka program, and evaluated the algorithms in terms of reliability and time to achieve the outcomes. A significant amounts of characteristics or keywords in the documents resulted in an inadequate functionality concerning both, precision and time. Consequently, they have claimed that, it is quite crucial to pre-process the text prior to classification of documents, this is vital to gain knowledge from substantial data and decrease the time consumed by processing procedures. There are two methods of pre-processing: (i) eradication of stop-words; and (ii) normalization approach. The outcomes of the above study were evident enough to reveal that, the (SMO) classifier accomplishes the maximum reliability and smallest error rate, followed by J48 (C4.5), and (NB) classifier. However, the outcomes related to time consumption

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

reveals that, the SMO model was quicker, followed by NB model, however, J48 classifier consumed much higher time to deliver the results.

Alsaleem [10] has reviewed Naïve Bayesian method (NB) and Support Vector Machine algorithm (SVM) on various Arabic data sets. His comparisons were based on the most well-known text evaluation methods. The outcomes of distinct Arabic text categorization data sets have exposed that, SVM algorithm has outplayed the NB concerning all methods. It is worth to mention here that, there are just two works related to Arabic Multi-label text classification.

Ezzat et al. [11] have explored a method to classify huge quantities of data in the absence of training data and classification system. Their study was inspired by a real life problem. They have proposed a system known as "TopicAnalyzer", which fuses distinctive feature extraction, selection and classification approaches to be extremely versatile in accepting all textual data. The outcomes of assessing the TopicAnalyzer have demonstrated that, its precision is identical to current supervised classification systems. The authors have used annotated datasets to assess their system. They have stated three dataset, two of which are evidently single-labeled. However, the study neither has given any information about the remaining one, nor has outlined the precision measures appropriate for MTC.

Alwedyan et al. [12] have examined three Multi-class classification algorithms based upon association rule (MCAR), NB, and SVM. A dataset comprising over 5K Arabic documents was segregated into seven groups. The study has determined that, MCAR is more precise in terms of automatically classifying Arabic documents as opposed to the other two classifiers.

Ahmed et al. [13] have studied the transformation approach in an effort to take advantage of conventional TC algorithms. They have experimented with various base classifiers, such as, SVM (referred to as SMO in MEKA), NB, KNN2 (known as IBK in MEKA) and Decision tree (identified as J48in MEKA). These steps were executed utilizing the MEKA tool. However, it is crucial to have a huge volume of multi-labeled dataset.

3. PROPOSED METHOD

The framework of the proposed method as shown in Fig 1 consists of five main phases; *corpus*, *preprocessing*, *feature selection*, *multi-label categorization* and *evaluation*. Corpus phase discuss the dataset that has been used for classification. Preprocessing phase focuses on the tasks that have been performed regarding to turning the data into an appropriate format that is process-able. Feature selection phase aims to opt for the best discerning terms for training and testing. Multi-label categorization phase aims to carry out the collaborative machine learning binary classifier for Arabic text classification. Finally, evaluation phase aims to evaluate the propose method. However, the following sub-sections illustrates such phases in further details.

3.1. Multi-label Arabic Document

In order to evaluate the Arabic multi-label text categorization system, this study has used a standard corpus which is used by Ahmed et al. [13] which consist of about 10,000 articles written in Modern Standard Arabic (MSA). The articles belong to five general classes: *Arts, Sports, Politics, Economy*, and *Science*. Each domain is consisting of 2000 documents ((*Arts* (2000), *Sports* (2000), *Politics* (2000), *Economy* (2000) and *Science* (2000)). Each article has 1 to 5 labels and the total number of different labels is 32. Table 1 shows the description of such corpus.

Table 1. Corpus Description

Language	**Modern Standard Arabic (MSA)**
Categories	*Arts,* 2000 *Sports,* 2000 *Politics* 2000*, Economy* 2000*, and Science* 2000
No. of documents	10000, each domain is 2000 document
No. of possible multi-label categories	Total number of different labels is 32.

Journal of Theoretical and Applied Information Technology
29th February 2016. Vol.84. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

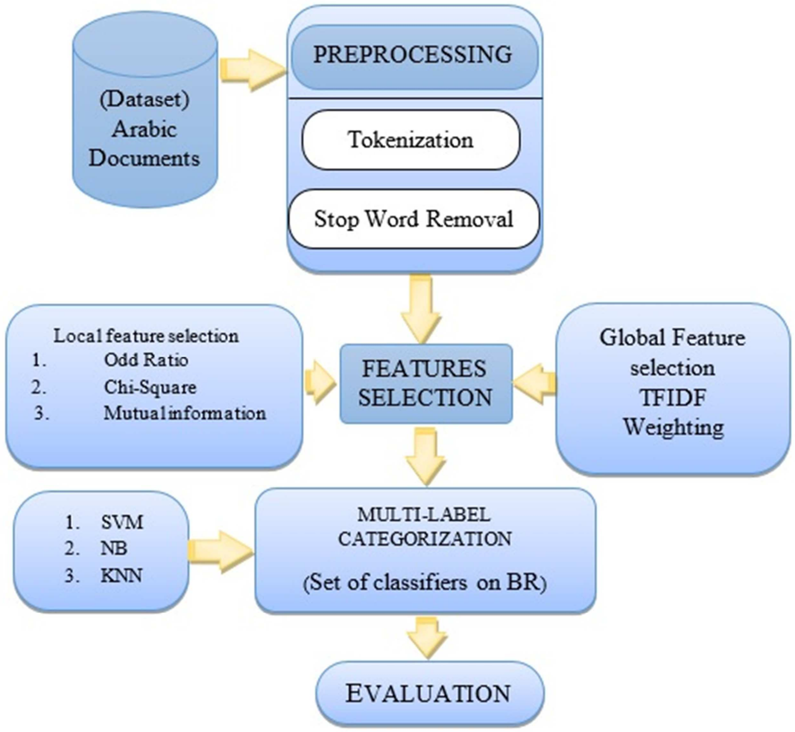

Figure 1. The Proposed Research Method Framework

3.2. Preprocessing

The pre-processing phase is an important task in designing a data set and it is crucial to pre-process the data with machine learning approaches. In fact, it consists of two sub-tasks; (i) tokenization and (ii) stop-words removal. The objective of tokenization is the survey of the words in a sentence. Written data is just a bulk of fonts at the beginning. All subsequent procedures in information recovery need the words of the data set. The foremost usage of tokenization is classifying the meaningful keywords. Whereas, stop-words removal task aims to eliminate the unnecessary words, such as, pronouns, prepositions, conjunctions and others. Such words are usually used by authors for linguistically enhancing the structures of the sentences. These words, which are so frequently found in the texts and which do not provide more valuable information about the text content, are called the stop words.

3.3. Feature Selection

In text categorization, a document is described by a vector of features (terms) and feature values, also called attributes and attribute values. A common text representation is to assign for each word (feature), the TFIDF (term frequency inverse document frequency). Obviously, this representation could lead to very high number of features for vast document collections. While feature selection is also necessary in single label text categorization task due to the high dimensionality of text features and the existence of irrelevant (noisy) features, it is especially important in multi-label text categorization as it includes many single label text categorization tasks.

The proposed feature selection mechanism is composed of two parts; First, in the global feature selection based on the word weighting. Second is the local feature selection based on statistical methods that collects the essential content from a document. Three statistical feature selection methods have been implemented are; Chi Squared ($\chi 2$), Mutual Information (MI) and odd ratio (OR) for feature selection measures. These statistical measurement achieved the best performance to tackle the categorization problem [14]. The main aim of global and local feature selection is to reduce

Journal of Theoretical and Applied Information Technology
29th February 2016. Vol.84. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the high dimensionality terms by getting rid of irrelevant (noisy) features (terms) in order to increase the performance and efficiency of classification task. In the following, both mechanism of feature selection will be explained with more details.

3.3.1 Global Feature Selection

The global dimensionality reduction focuses in classifying a reduced set of terms from all the categories. Before any binary classification take place, each document from all training data set is represented as a set of terms. Then, each term is assigned a TFIDF (term frequency inverse document frequency): A term is assigned a weight based on two measures: (1) tf , the frequency of a term in a single document, and (2) df , the number of documents in the corpus that contain the given term. N is the total number of documents. Such weighting task has been performed based on the following equation:

$$\text{TFIDF}_i = tf_i \cdot \log(\frac{N}{n}) \quad (1)$$

3.3.2 Local Feature Selection

Unlike, global feature selection, in the local feature selection only terms form documents of a single domain training set (not all training dataset) are weighted. In the binary classification, the training set of single domain consists of documents that belong to this domain and documents which are not belong to this domain. One of feature selection methods, i.e., Chi-square (χ2), Odd Ratio (OR) and Mutual Information (MI) is used in order to selects the best terms that represent this domain (class) and select the best terms that not represent this domain. Such statistical methods are illustrated as follows:

i. Chi-square (χ2)

Chi-square is one of the most commonly used feature selection algorithms. The feature selections focus on measuring lack the flexibility between the term and the category [14]. In the text categorization tasks, it measure the independence of two random variables (1) the occurrence of a term t (2) the occurrence of a class c. It is also used widely in the text categorization research and it performs well in comparison with other feature selection algorithms [15]. The χ2 value for each term t in a category c is calculated by the following equation:

$$X^2\,(t,c) = \frac{N(\,AD - CB\,)^2}{(\,A + C\,)(\,B + D\,)(A + B\,)(\,C + D\,)} \quad (2)$$

Where N: is the total number of Arabic training documents. A is the number of Arabic documents in class c and contain term t. B is the number of Arabic documents that do not belong to class c but contain term t. C is the number of Arabic documents that do not belong to class c and do not contain term t. D is the number of Arabic documents that do not belong to class c and not contain term t [16].

ii. Mutual Information (MI)

The mutual information measure is derived from information theory [17]. It provides a formal way to model the mutual information between the terms and the classes. The mutual information MI (t, c) between the term t and the class c is defined based on the level of co-occurrence between the category c and term t [18]. Such measure is calculated using the following equation:

$$MI(x,y) = log_2 \frac{p(x,y)}{p(x)p(y)} \quad (3)$$

iii. Odd Ratio (OR)

Odds ratio feature selection measure is originally developed for selecting terms for relevance feedback in text categorization. The main idea is that the distribution of features (terms) on the related texts is different from the distribution of features on the non-related texts. Odd Ratio reflects the odds of the word occurring in the positive class normalized by that of the negative class [19]. Such measure is calculated using the following equation:

$$Odd\ Ratio\ (f) = \frac{A * D}{C * B} \quad (4)$$

3.4. Multi-label Categorization

Generally, it is impossible to categorize each document under a single label, because of the natural overlapping of the category spaces. As an example, the economics field often dovetails the politic one. This fact brings to enforce different constraints on the categorization task, depending on the application. As mentioned before with multi-label text categorization a document can belong to multiple classes. Not only does the training data have documents with multiple labels, the classifier has to be able to map a single document into multiple classes. The training algorithm has to be adapted to be able to handle multiple labels. Multi-Label text categorization can be defined as the classification task where a classifier x or a set of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classifiers assigns each document d to zero or more predefined class labels.

- **Set of Classifiers based on Binary Relevance (BR) Method**

This study proposes a set of *n* classifiers that are combined in order to classify multi-label text. For this purpose, three classifiers have been used; Support Vector Machine (SVM), Naïve Bayes (NB) and K-Nearest Neighbor (KNN). The set classification has two tasks; first, a set of n classifiers of the same type (for example a set of n NB classifiers) are used. Each classifier is trained independently using a data set from one domain from the set of domains {*Arts, Sports, Politics, Economy,* and *Science*}. Each classifier can only predict if the document belongs to one class or not. The second task contains a set of n classifiers of different types (for example a set of SVM, NB, SVM, KNN, NB classifiers) are used. As in set of n binary classifiers from the same type, each classifier is trained independently using a data set from one domain from the set of domains {*Arts, Sports, Politics, Economy,* and *Science*}. Each classifier can only predict if the document belongs to one class or not. The criteria for the selection of the type of the classifiers are based on evaluating the three types (NB, SVMN and KNN) on each domain. For example, if we evaluated (trained and tested) the three classifiers (NB, SVM and KNN) on data from *Sport* domain and the results show that NB is outperforms others in *Sport* domain. Then, the NB is selected for *Sport* field. Table 2 shows the pseudo code of the set binary classifier algorithm.

Table 2. Pseudo Code Of The Set Binary Classifier Algorithm

A binary relevance method on set of classifiers.

Inputs:
1. Training datasets: $TD = \{td_1, td_2, td_3, td_4\ , td_5\ \}$
2. Classifiers set: $C = \{c_1, c_2, c_3, c_4\ , c_5 : c_i\ \in \{SVM, KNN, NB\}\}$
3. Labels : $L = \{\text{Arts, Sports, Politics, Economy, and Science}\ \}$
4. Test set

Outputs : MCR Multi-label classification result of test set

$For(\ i = 0; i <\ size\ of\ domains(5)\ ; i++)$
    $begin$
    $Train\_classifier(c_i, td_i) // train\ classifier\ c_i using\ dataset\ from\ domain\ td_i$
    end
$For\ each\ (document\ d\ in\ Test\ set)$
$Begin$
$Document_{label}$={};
    $For\ each\ (classifier\ c_i\ \ in\ \ C)$
    $Begin$
    $r_i = Classify(c_i, d)\ //result\ of\ \ c_i\ is\ 1\ (if\ the\ document\ d\ belong\ to\ td_i)\ or\ 0\ if\ not$
    $If(r_i == 1)$
    $Document_{label}.add(label_i)$
    End
$MCL[d] = Document_{label};$

End

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.5. Evaluation

Several metrics have been recently proposed to evaluate the performance of multi-label classifiers. They range from considering the performance of the multi-label classifier over each class independently of the rest, to considering the performance of all the classes at the same time. For the purpose of comparison, we used three different multi-label evaluation measures which are normally used in multi-label categorization [20-22]:

1. Average precision over the d class variables (precision per label) as in the following equation:

$$\mathrm{M\_PRECISION} = \sum_{i=1}^{d} \frac{\mathrm{TP_i}}{\mathrm{PT_i} + \mathrm{FP_i}} \tag{5}$$

2. Average recall over the d class variables (recall per label) as in the following equation:

$$\mathrm{M\_RECALL} = \sum_{i=1}^{d} \frac{\mathrm{TP_i}}{\mathrm{PT_i} + \mathrm{FP_i}} \tag{6}$$

3. Average F measure over the d class variables (F measure per label) as in the following equation:

$$\mathrm{M\_F\beta} = \sum_{i=1}^{d} \frac{(\beta^2 + 1)\mathrm{Pr} \times \mathrm{Re}}{\beta^2 \mathrm{Pr} + \mathrm{Re}} \tag{7}$$

4. RESULTS AND DISCUSSION

Basically, the results have been obtained from both tasks on the set of classifiers. First, the set of *n* classifiers from the same type have been evaluated with the three feature selection methods χ^2, MI and OR. Second, the set of *n* classifier from different types have been evaluated with the three feature selection methods χ^2, MI and OR. Table 3 depicts such results.

Table 3 Experimental Results

Set classifier	Average Precision	Average Recall	Average F-measure
set of NBs with χ2	0.733	0.763	0.737
set of NBs with OR	0.717	0.778	0.732
set of NBs with MI	0.767	0.746	0.753
set of SVMs with χ2	0.83	0.781	0.803
set of SVMs with OR	0.754	0.795	0.774
set of SVMs with MI	0.826	0.779	0.799
set of KNNs with χ2	0.973	0.793	0.858
set of KNNs with OR	0.88	0.779	0.807
set of KNNs with MI	0.859	0.779	0.797
set of diff. type classifiers with χ2	0.91	0.83	0.863
set of diff. type classifiers with OR	0.82	0.82	0.815
set of diff. type classifiers with MI	0.801	0.812	0.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

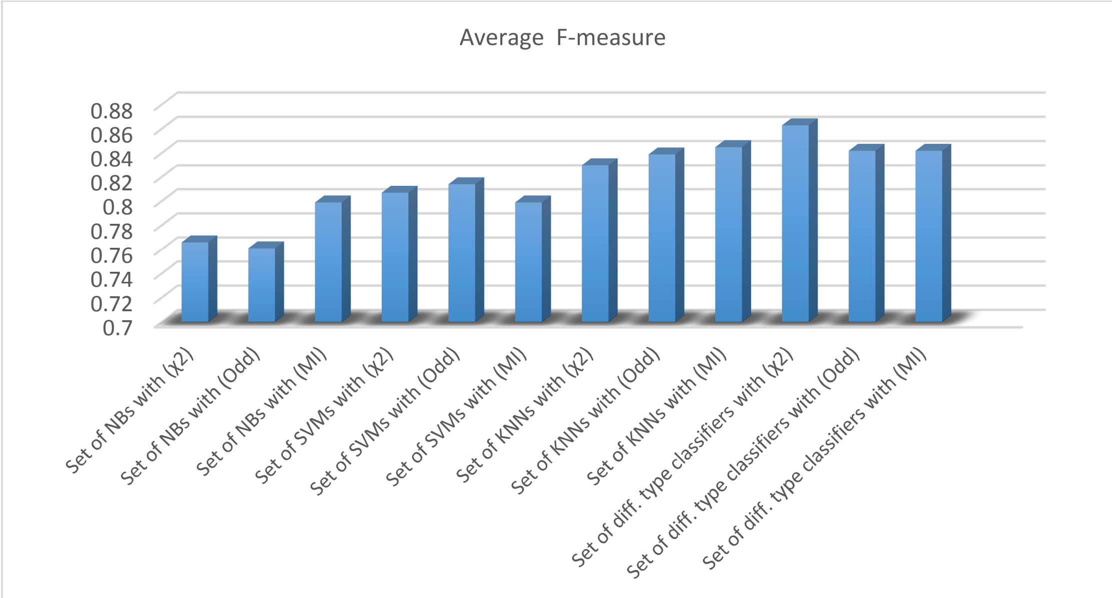

Figure 2. Experimental Results

As shown in Table 3 and Fig 2, the experiments of Arabic multi-label text categorization, the highest result yield by the set of the different set of algorithms with $\chi 2$ 86.8% average f-measure and the lowest result yield by the set of NB classifiers (OR) with 76.1% average f-measure. In addition, the results obtained using the set of the different set of algorithms method is outperformed that obtained using other methods. These results indicate that the set of the different set of algorithms method is the most suitable technique for Arabic multi-label text categorization. Finally, results show that each of the feature selection methods (Chi-square ($\chi 2$), Odd Ratio (OR) and Mutual information (MI)) has different effect on the quality of Arabic multi-label categorization depends on the categorization method used. In which, for the set NB classifiers, MI feature selection method outperformed than the other feature selection methods (OR and $\chi 2$). In the set of SVM classifiers, feature selection method of OR gives the best results, whereas for set of different classifiers (SVM and KNN), feature selection of $\chi 2$ yields to the best result.

5. CONCLUSION

The primary aim of study is to design a prototype tool and to describe a new methodology for multi-label text categorization for Arabic texts based on binary relevance (BR) approach. The set of classifiers are setup based on binary relevance (BR) approach where the set contains a different set of machine learning classifiers including Support vector machine (SVM), K-nearest neighbor classifier (KNN) and Naive Bayes (NB). The binary relevance (BR) classifier set is constructed based on the results of single binary classifiers on each single domain or class. One of the drawbacks of this study is the lack of well-annotated data sets. A new dataset which cover more labels and annotated manually is needed to evaluate used or new methods for Arabic multi-label text categorization.

6. ACKNOWLEDGMENT

This research project is funded by Malaysian Government under research grant ERGS /1/ 2013/ ICT07/UKM/03/1.

REFERENCES

[1] David D Lewis and William A Gale, "A sequential algorithm for training text classifiers," in *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, 1994, pp. 3-12.
[2] Youngjoong Ko, "Text Categorization using Unlabeled Data," Sogang University, 2003
[3] Fabrizio Sebastiani, "Machine learning in automated text categorization," *ACM computing surveys (CSUR),* vol. 34, pp. 1-47, 2002
[4] Andrea Addis, "Study and Development of Novel Techniques for Hierarchical Text Categorization," University of Cagliari, 2010
[5] Maria-Luiza Antonie and Osmar R Zaiane, "Text document categorization by term association," in *Data Mining, 2002. ICDM*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2003. Proceedings. 2002 IEEE International Conference on, 2002, pp. 19-26.

[6] Clifford Tawiah and Victor Sheng, "Empirical comparison of multi-label classification algorithms," in *Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

[7] Mostafa M Syiam, Zaki T Fayed, and Mena B Habib, "An intelligent system for Arabic text categorization," *International Journal of Intelligent Computing and Information Sciences,* vol. 6, pp. 1-19, 2006.

[8] Ismail Hmeidi, Bilal Hawashin, and Eyas El-Qawasmeh, "Performance of KNN and SVM classifiers on full word Arabic articles," *Advanced Engineering Informatics,* vol. 22, pp. 106-111, 2008.

[9] FO Majed Ismail Hussien, "Minwer AL-dwan, Ahlam Shamsan," ARABIC TEXT CLASSIFICATION USING SMO, NAÏVE BAYESIAN, J48 ALGORITHMS,"" *International Journal of Research and Reviews in Applied Sciences,* vol. 9, p. 10, 2011.

[10] Saleh Alsaleem, "Automated Arabic Text Categorization Using SVM and NB," *Int. Arab J. e-Technol.,* vol. 2, pp. 124-128, 2011.

[11] Heba Ezzat, Souraya Ezzat, Samhaa El-Beltagy, and Moustafa Ghanem, "Topicanalyzer: A system for unsupervised multi-label arabic topic categorization," in *Innovations in Information Technology (IIT), 2012 International Conference on*, 2012, pp. 220-225..

[12] Jaber Alwedyan, Wa'el Musa Hadi, Ma'an Salam, and Hussein Y Mansour, "Categorize arabic data sets using multi-class classification based on association rule approach," in *Proceedings of the 2011 International Conference on Intelligent Semantic Web-Services and Applications*, 2011, p. 18.

[13] Nizar A Ahmed, Mohammed A Shehab, Mahmoud Al-Ayyoub, and Ismail Hmeidi, "Scalable multi-label arabic text classification," in *Information and Communication Systems (ICICS), 2015 6th International Conference on*, 2015, pp. 212-217.

[14] Luigi Galavotti, Fabrizio Sebastiani, and Maria Simi, "Experiments on the use of feature selection and negative evidence in automated text categorization," in *Research and Advanced Technology for Digital Libraries*, ed: Springer, 2000, pp. 59-68.

[15] Şerafettin Taşcı and Tunga Güngör, "Comparison of text feature selection policies and using an adaptive framework," *Expert Systems with Applications,* vol. 40, pp. 4871-4886, 2013Retrieved from.

[16] Fadi Thabtah, M Eljinini, Mannam Zamzeer, and W Hadi, "Naïve Bayesian based on Chi Square to categorize Arabic data," in *proceedings of The 11th International Business Information Management Association Conference (IBIMA) Conference on Innovation and Knowledge Management in Twin Track Economies, Cairo, Egypt*, 2009, pp. 4-6.

[17] Thomas M Cover and Joy A Thomas, *Elements of information theory*: John Wiley & Sons, 2012.

[18] Charu C Aggarwal and ChengXiang Zhai, "A survey of text classification algorithms," in *Mining text data*, ed: Springer, 2012, pp. 163-222.

[19] George Forman, "An extensive empirical study of feature selection metrics for text classification," *The Journal of machine learning research,* vol. 3, pp. 1289-1305, 2003

[20] Concha Bielza, Guangdi Li, and Pedro Larranaga, "Multi-dimensional classification with Bayesian networks," *International Journal of Approximate Reasoning,* vol. 52, pp. 705-727, 2011

[21] Jesse Read, Bernhard Pfahringer, Geoff Holmes, and Eibe Frank, "Classifier sets for multi-label classification," *Machine learning,* vol. 85, pp. 333-359, 2011.

[22] L Enrique Sucar, Concha Bielza, Eduardo F Morales, Pablo Hernandez-Leal, Julio H Zaragoza, and Pedro Larrañaga, "Multi-label classification with Bayesian network-based set classifiers," *Pattern Recognition Letters,* vol. 41, pp. 14-22, 2014.