

Arabic Text Categorization Using Support vector machine, Naïve Bayes and Neural Network

Adel Hamdan Mohammad,
Computer Science Department,
The world Islamic Sciences and
Education University,
Amman-Jordan,

Tariq Alwada'n,
Computer Science
Department,
The world Islamic Sciences
and Education University,
Amman-Jordan

Omar Al-Momani
Network Department
The world Islamic Sciences and
Education University,
Amman-Jordan

Abstract

Text classification is a very important area in information retrieval. Text classification techniques used to classify documents into a set of predefined categories. There are several techniques and methods used to classify data and in fact there are many researches talks about English text classification. Unfortunately, few researches talks about Arabic text classification. This paper talks about three well-known techniques used to classify data. These three well-known techniques are applied on Arabic data set. A comparative study is made between these three techniques. Also this study used fixed number of documents for all categories of documents in training and testing phase. The result shows that the Support Vector machine gives the best results.

Keywords: Text Categorization, Naïve Bayesian, Text Classification, Support Vector Machine, Multilayer Perceptron Neural Network

1. INTRODUCTION

The number of documents available online has increased because of the rapid growth of internet. Also no doubt that number of documents available on line is countless. Regularly text documents include newspaper, research article, technical reports, blogs, journal paper...etc. This huge number of documents could be useful and valuable.[1]. Also this enormous number of documents required effective retrieval methods. Text Classification (TC) is a very important area in information retrieval technology and also it is an active research area. TC aims to categorize

this huge number of documents into a set of predefined categories based on its content. [2, 3]. Online textual documents and their categories are huge, numerous and diverge. TC is a method in data mining to give valuable information from large amount of data. [4]. TC has been used in many applications such as email filtering (spam or legitimate) [5, 6], news monitoring and searching for useful information on the web. [7]. The aim of TC methods and techniques is to extract any valuable information from unstructured textual resource, these valuable information then used to classify the document into a set of predefined categories.[2]. Text mining is not an easy process due to the massive availability of information in text documents and its diversity. Beside that the process of text mining includes derived linguistic features from the text. So far and because of this rapid development of web data, algorithms that can handle and develop text classification and increase its efficiency are still highly required. [8]

Most of systems used are designed to handle English language. Also many researches talks about English text classification and there are many algorithms designed to handle English documents. Few researchers talks about Arabic text classification. Arabic language considered one of the widely spoken languages. Arabic language is the main language in the Arab world and the secondary language in many other countries. [9]. Arabic language has rich morphology and a complex orthography [10]. Creating text classification system for Arabic

language is a not an easy task at all due to nature and difficulty of Arabic language.

This research will use Support vector machine (SVM), Naïve Bayesian (NB) and Multilayer Perceptron Neural Network (MLP-NN) to categorize Arabic data set.

The rest of this paper organized as follow: section 2 talks about Arabic language, section 3 talks about document pre-processing steps, Section 4 talks about term selection and weighting, section 5 talks about SVM, section 6 talks about Naïve Bayesian, section 7 talks about Multilayer Perceptron Neural Network, Section 8 talks about related studies, section 9 talks about data set used in experiments and section 10 talks about experimental and results, and finally section 11 demonstrate our conclusions.

2. ARABIC LANGUAGE.

Arabic language is one of six official language of the United Nations. Arabic language is the official language of 25 countries, also Arabic language is spoken from over 250 million .Arabic language consists from 28 letters plus hamza (ء) which is considered a letter by some Arabic linguistics. Arabic language is written from right to left. The letters of Arabic language have different shapes when appearing in a word depending on the position of the letter (at the beginning, at the end, at the middle). Different shapes and diacritics of Arabic language make parsing difficult task. [11]. Majority of Arabic words has its root, beside that over 80% of Arabic words can be mapped into 3-letter root. Fortunately, Arabic language has its built it filtering mechanism which means different words can be mapped into their root. Representing words to their root is very important and definitely will reduce the number of words. [12]

3. DOCUMENT PRE-PROCESSING

Document pre-processing is very essential to reduce the complexity of documents and to handle it in an easy form. [13]. Document pre-processing include several steps such as First, document cleaning, in this step document is cleaned from unnecessary words such as stop words, tags etc. Second, Document representation, in this step, Mostly, document is transformed to vector space model (VSM). Third, dimension reduction, in this step, the most useful and valuable features for classification are selected. Most commonly selection methods are Chi square and

document frequency, information gain and mutual information [14, 15, 16, 17]

Usually text documents are represented as a vector of term weights (word features). One major problem in classification for Arabic and English language is the high dimensionality of text documents. Data pre-processing reduce the complexity of text documents and make data handling and representation easier. Without data pre-processing several major problems will be encountered. Documents pre-processing can be classified into Feature Extraction (FE) and Feature Selection (FS). [18].

Feature Extraction (FE) is the first step of text pre-processing which includes cleaning text documents and stemming. [19]. FS used to eliminate text and to present it in a clear format. FS in very important and get much attention since it's effective, efficient and save storage space. [20].

Feature Selection (FS) is used after Feature Extraction to construct vector space. The main goal of FS is to select suitable features of the original document. FS aim to keep words with highest score according to a set of predefined measures.[21,22]

Representing all document words into their root is critical and could result into reducing the number of words. [12]. Root extraction (stemming) for Arabic language fall into two types. [23, 24, 25, 26]. First, getting the root based on the idea of removing prefixes, infixes, suffixes. Second, getting the root based on the weight of letters embedded within the text. [27].

4. TERM SELECTION AND WEIGHTING

Naturally, thousands of features can be found in document classification. So term selection techniques are used to reduce high dimensionality. Usually term selection based on term evaluation function. There are many term evaluation function used for English text categorization such as Chi square, information gain, Document Frequency Threshold [3,14,15,16,17].

After selecting the most important terms, each document is represented as weighted vector based on the words found in the text. Actually there are many weighting techniques such Term Frequency (TF), Inverse Document Frequency, Term Frequency Inverse Document

Frequency (TFIDF), and Normalized-TFIDF weighting. [3,14,15,16,17]

5. SUPPORT VECTOR MACHINE.

Support vector machines (SVMs) are one of the best well-known machine learning algorithms. SVMs are binary classifier initially proposed by Vladimir. [28]. Also, SVMs are supervised learning model used to analyse and categorize text data. Furthermore, SVM typically used for text categorization and regression analysis. SVM has been introduced and used in text classification and categorization by Thorsten Joachims in 1998 [29]. SVM training algorithm mainly used to builds a model that assigns new documents into a set of predefined categories. Beside that SVMs can be used as linear and non-linear classifier. [29, 30]

5.1 SUPPORT VECTOR MACHINE (THE SEPARABLE CASE, NON-LINEAR CLASSIFIER)

Consider the data shown in figure 1; simply data can be linearly separated. For defining a linear classifier the dot product between two vectors is required, defined as $w^T x = \sum_i w_i x_i$. A linear classifier is based on a linear discriminant function of the form $f(x) = w^T x + b$. Vector w is known as the weight vector, and b is called the bias. First consider the case $b = 0$. The set of points x such that $w^T x = 0$ are all points that are perpendicular to w and go through the origin — a line in two dimensions, a plane in three dimensions, and more generally, a hyperplane. [10, 29] The bias b translates the hyperplane away from the origin. The hyperplane $\{x: f(x) = w^T x + b = 0\}$ divides the space into two: the sign of the discriminant function $f(x)$ denotes the side of the hyperplane a point is on. The boundary between regions classified as positive and negative is called the decision boundary of the classifier. There are many cases in which the data is not linearly separated. But, Linear SVMs can be generalised for non-linear problems. Doing this required set up an objective function that trades off misclassifications against minimizing $\|w\|^2$ to find an optimal compromise. Thus we introduce a non-negative “slack” variable $\xi_i > 0$ for each training example.

$$\tilde{w} \cdot x_i + w_0 \geq 1 - \xi_i \quad \text{for } y_i = +1$$

$$\tilde{w} \cdot x_i + w_0 \leq -1 + \xi_i \quad \text{for } y_i = -1.$$

More details about SVMs equations in [29, 30, 31, 32, 33]

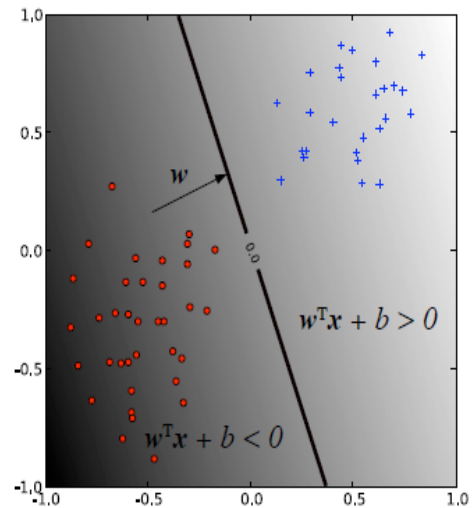


Figure 1: Linear classifier.

6. NAÏVE BAYESIAN

Naïve Bayes classifiers are simple probabilistic classifiers based on applying Bayes' theorem. Naïve Bayes are simple and powerful method for text classification. When Naïve Bayesian classifier is used for text classification problem we use the equation:

$$P(\text{class} | \text{document}) = P(\text{class}) \cdot P(\text{document} | \text{class}) / P(\text{document})$$

$P(\text{class} | \text{document})$: The probability that a given document D belongs to a given class C .
 $P(\text{document})$: The probability of a document.
 $P(\text{class})$: The probability of a class (or category), we can compute it from the number of documents in the category divided by documents number in all categories.
 $P(\text{document} | \text{class})$ represents the probability of document given class. [34, 35, 36].

This means that under the above independence assumptions we can rewrite it as:

$$P(\text{document} | \text{class}) = \prod p(\text{word } i | \text{class}) \text{ and } P(\text{class} | \text{document}) = p(\text{class}) \prod p(\text{word } i | \text{class}).$$

Where $P(\text{word } i | \text{class})$: The probability that the i -th word of a given document occurs in a document from class C , and this can be computed as follows:

$$P(\text{word } i | \text{class}) = (T_{ct} + \lambda) / (N_c + \lambda V) \text{ where } T_{ct}: \text{The number of times the word occurs in that category } C. N_c: \text{The number of words in category } C. V: \text{The size of the vocabulary table. } \lambda: \text{The positive constant, usually 1, or 0.5 to avoid zero probability. [35]}$$

7. MULTILAYER PERCEPTRON NEURAL NETWORK

Artificial neural network are a family of statistical learning models inspired by biological neural network (nervous system). Artificial neural network is available into different shapes such as supervised and unsupervised learning.

Modern neural network are based on the perceptron model and the Least Mean Square (LMS) (Figure 2). Perceptron model was introduced by Rosenblatt (1958). [37] The simplest perceptron model is a network model consists from an input layer and an output.

Single Layer Perceptron

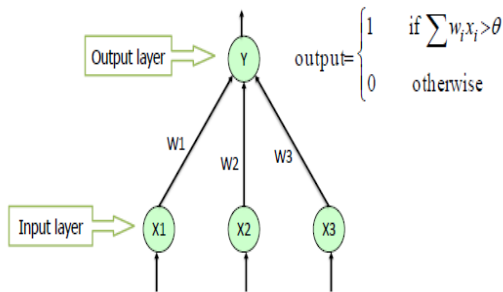


Figure 2: Single layer perceptron

A multilayer perceptron (MLP) (Figure 3) is a feed forward artificial neural network model which can be used to map set of input data into a set of appropriate outputs. MLP passing input to output through one or more connected hidden layers. There are several parameters affecting the results such as: number of hidden layers, weights and transfer functions. [38,39,40]

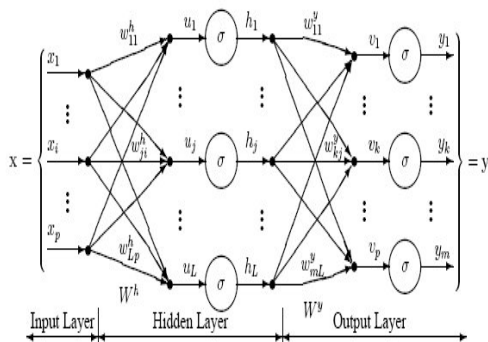


Figure 3: Multilayer Perceptron Neural Network Model

Authors in this paper use a three layer feed-forward neural network with hyperbolic tangent (tanh) activation function in the hidden layer, followed by a linear output layer. Number of input layers used in this paper is

400 and 600. Also the number of hidden layers is 30. Training is done by using back propagation learning rule based on supervised learning. Trial and error approach has been used to find the most suitable number of hidden layers that provide the most suitable classification accuracy. Number of output layers is 8 (number of categories in corpus). Also learning rate parameter is 0.01 and Mean Square Error is 0.001.

8. RELATED STUDIES

Tarek Fouad (2009) [31] apply SVM model in classifying Arabic text documents. Results are compared with Bayes, K-NN and Rocchio classifier. Experiments are done on 1132 documents. Experiments shows that Rocchio classifier gives better results when the size of feature set is small. Also results shows that SVM outperform the other classifiers when the size of the feature set is large enough.

El-Kourdi (2004) [41] applied Naïve Bayesian (NB) to classify Arabic web data. The results show that the average accuracy was 68%. Experiments are done on 300 web document per category.

Saleh Alsaleem (2011) [35] uses NB and SVMs on Arabic data set. Experiments are done on several Arabic data set categories. Recall and precision are used as measures. Results show that SVM are better than NB.

Abdelwadood Moh'd A MESLEH (2007) [42] implement SVM for Arabic language. Chi square classifier is used as feature selection. Collected corpus contains 1445 documents. Documents fall into nine categories. Results show that F-measure (F=88.11).

Fouzi Harrag (2010) [39] present a model based on the neural network for classifying Arabic texts. Authors propose singular value decomposition (SVD). Experiments done on in-house consist from 5734 tokens (14 major categories). Number of words after pre-processing step is 1065 words. Results show that MLP NN model using SVD method is better for capturing the non-linear relationships between the input document vectors and the document categories than that of basic MLP NN model.

Al-Harbi (2008) [43] uses SVM and C5.0 to classify Arabic text documents on seven different Arabic corpora by using a recognized statistics technique. C5.0 classifier, in general, gives better accuracy.

9. DATA SET

Dataset set used in our experiments collected from Aljazeera news web site (<http://www.aljazeera.net>), Saudi Press Agency (SPA) (<http://www.spa.gov.sa/index.php>) and Al-hayat (<http://www.alhayat.com/>). Dataset consists of 1400 Arabic documents belong to different categories (see table 1).

Category	Total Number of documents	Number of documents used for training	Number of documents used in testing.
Computer	175	110	65
Economics	175	110	65
Education	175	110	65
Law	175	110	65
Medicine	175	110	65
Politics	175	110	65
Religion	175	110	65
Sports	175	110	65
Total	1400	880	520

Table 1: data set

Approximately, dataset is divided into two parts. Dataset used for training is 880 documents (63%) and 520 documents is used for testing (37%).

10. EXPERIMENTS AND RESULTS

In this paper three evaluation measures are used (recall, precision and F1). Precision is the

fraction of retrieved documents that are relevant. Recall is the fraction of relevant documents that are retrieved and F1 is a measure of a test's accuracy. Consider table 2 and suppose that a: relevant document retrieved, b: irrelevant document retrieved, c: relevant document not retrieved, d: irrelevant document not retrieved. (See table 2).

Iteration	Relevant	Irrelevant
Document retrieved	a	b
Document not retrieved	c	d

Table 2: Document possible sets based on a query

$$F1 = (2 * Precision * Recall) / (Recall + Precision)$$

$$Precision = a / (a + b)$$

$$Recall = a / (a + c)$$

In other words, high precision means that an algorithm returned significantly more relevant results than irrelevant, also high recall means that an algorithm returned most of the relevant results. So we look to get high values of precision and recall.

Table 3 and figures 4, 5 give precision; recall and F1 measures generated used SVM and NB. After analysing results we found that SVM results are little bit better than NB except in law and sports category.

Category Name	SVM			NB		
	Precision	Recall	F1	Precision	Recall	F1
Computer	0.696	0.647	0.671	0.665	0.651	0.658
Economics	0.791	0.87	0.830	0.781	0.862	0.819
Education	0.615	0.651	0.632	0.715	0.691	0.703
Law	0.879	0.881	0.878	0.899	0.891	0.895
Medicine	0.877	0.881	0.879	0.812	0.881	0.845
Politics	0.673	0.698	0.685	0.563	0.598	0.580
Religion	0.891	0.851	0.870	0.791	0.721	0.754
Sports	0.791	0.711	0.749	0.811	0.785	0.796
Average	0.778	0.774	0.775	0.754	0.759	0.756

Table 3: SVM, NB Results

Table 4 and figures 5, 6 give precision; recall and F1 measures generated used MLP_NN using 400 and 600 input layers. 600 input layers gives a little bit better results than 400 input layers.

400 Input layers	MLP-NN			600 Input layers	MLP-NN		
	Precision	Recall	F1		Precision	Recall	F1
Computer	0.596	0.584	0.589939	Computer	0.619	0.684	0.649879
Economics	0.688	0.645	0.665806	Economics	0.61	0.667	0.637228
Education	0.569	0.599	0.583615	Education	0.699	0.611	0.652044
Law	0.778	0.811	0.794157	Law	0.791	0.711	0.74887
Medicine	0.798	0.811	0.804447	Medicine	0.811	0.799	0.804955
Politics	0.771	0.511	0.614635	Politics	0.799	0.611	0.692467
Religion	0.741	0.691	0.715127	Religion	0.788	0.701	0.741958
Sports	0.521	0.661	0.582709	Sports	0.621	0.699	0.657695
Average	0.68275	0.664125	0.668804	Average	0.71725	0.685375	0.698137

Table 4: MLP-NN Results

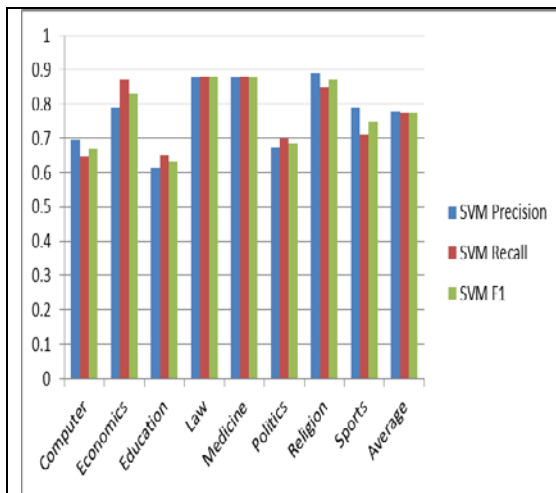


Figure 4: SVM (Precision, Recall, F1)

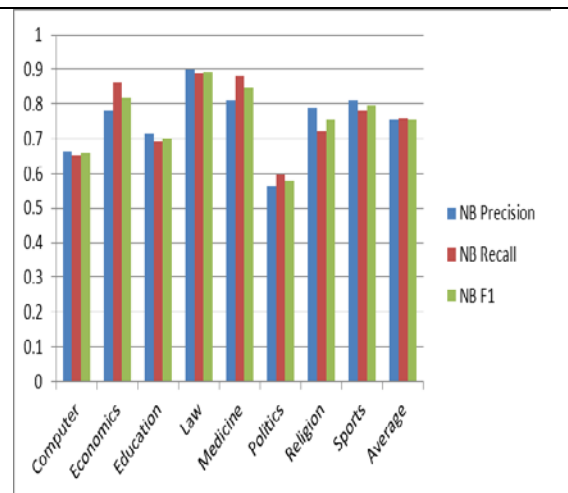


Figure 5: NB (Precision, Recall, F1)

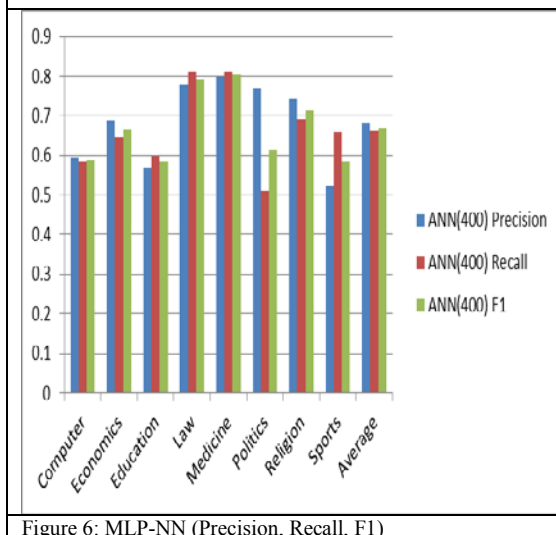


Figure 6: MLP-NN (Precision, Recall, F1)

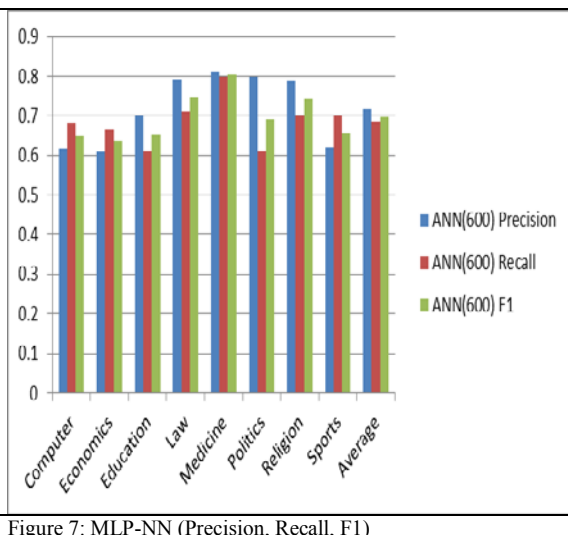


Figure 7: MLP-NN (Precision, Recall, F1)

11. CONCLUSIONS AND FUTURE WORKS.

In this paper authors discuss the problem of text classification in general and Arabic text classification in specific. Authors use three important text classification algorithms and apply experiments and tests on in-house developed data set. The three algorithms used are SVM, NB and MLP-NN algorithms. The importance of this study based on applying those three algorithms on large Arabic data set and then applying a comparative study. The average measures obtained from applying mentioned algorithms indicate that SVM algorithm outperformed NB and MLP-NN. Average precision results of SVM, NB and MLP-NN using 600 input layers respectively is 0.778, 0.754, and 0.717. This small fraction of difference forced authors to think that features extraction and selection is critical and affect results in many ways. Beside that authors thinks that MLP-NN could give better results if feature extraction and selection in pre-processing steps is developed enhanced since 600 input layer gives promising results. Future works of authors is applying different feature extraction and selection methods using MLP-NN and make a comparative study on feature extraction and selection.

REFERENCE

- [1] Motaz K Saad, Wesam Ashour, "Arabic Text Classification Using Decision Trees"(2010) proceedings of the 12th international workshop on computer science and information technologies CSIT'2010, Moscow – Saint-Petersburg, Russia, 2010
- [2] Mofleh Al-diabat (2012), "Arabic Text Categorization Using Classification Rule Mining "Applied Mathematical Sciences, Vol. 6, 2012, no. 81,pp. 4033 – 4046
- [3] Sebastiani, F. (2002) 'Machine learning in automated text categorization' ACM Publication:. ACM Computing Surveys, Vol. 34, No. 1, March 2002, pp. 1–47
- [4] Rasha Elhassan, Mahmoud Ahmed (2015), "Arabic Text Classification review " International Journal of Computer Science and Software Engineering (IJCSSE), Volume 4, Issue 1, January 2015
- [5] Adel Hamdan, , Raed Abu-Zitar "Spam Detection Using Assisted Artificial Immune System", Volume: 25, Issue: 8(2011) pp. 1275-1295, International Journal of Pattern Recognition and Artificial Intelligence, .
- [6] Raed Abu-Zitar ,Adel Hamdan, , " Application of Genetic Optimized Artificial Immune System and Neural Networks in Spam Detection ,Applied Soft Computing, Volume 11, Issue 4, June 2011, Pages 3827-3845 ,Elsevier, 2011
- [7] Dharmadhikari, C.S., Ingle, M. and Kulkarni, P.(2011) "Empirical Studies on Machine Learning Based Text Classification Algorithms," Advanced Computing: An International Journal (ACIJ), Vol.2, 2011.
- [8] Shang, W., Huang, H., Zhu, H., Lin, Y., Qu, Y., and Wang Z., " A Novel Feature Selection Algorithm for text categorization." Elsevier, science Direct Expert system with application -2006, 33(1), pp.1-5, 2006.
- [9] Rehab Duwairi (2007), " Arabic c text categorization", the international Arab journal of information technology , Vol4, No2,April 2007.
- [10] Mesleh, A.M. (2008), "Support Vector Machines Based Arabic Language Text Classification System: Feature Selection Comparative Study," Advances in Computer and Information Sciences and Engineering, Springer Science + Business Media B.V., 2008
- [11] Rehab Duwairi (2005) ,Machine learning for Arabic text categorization ,” Journal of American society for information science and technology (JASIST), Vol57, No8,pp1005-1010, 2005
- [12]Eldos T . (2003), "Arabic Text Data Mining" A root Based Hierarchical Indexing Model”, International Journal of Modeling and Simulation, vol23, no3,pp158-166,2003
- [13] Aurangzeb Khan, Baharum Baharudin, Lam Hong Lee*, Khairullah khan, (2010)“ A Review of Machine Learning Algorithms for Text-Documents Classification, journal of advances in information technology, vol. 1, no. 1, February 2010.
- [14] Franca Debole et al., "Supervised Term Weighting for Automated Text Categorization", proceedings of SAC-03, 18th ACM Symposium on Applied Computing, Melbourne, 2003, USA
- [15] Al-Zaghoul F., Al-Dhaheiri S., "Arabic Text Classification Based on Features Reduction Using Artificial Neural Networks", UKSim, pp. 485-490. 2013.
- [16] Johannes Furnkranz., "A Study Using n Gram Features For Text Categorization", Technical Report OEFAl-TR-1998-30.
- [17] Abu-Errub A., "Arabic Text Classification Algorithm using TFIDF and Chi Square Measurements", International Journal of Computer Applications 93 (6), 40-45, 2014.
- [18] Liu, H. and Motoda,(1998) ., "Feature Extraction, construction and selection: A Data Mining Perspective.", Boston, Massachusetts(MA): Kluwer Academic Publishers.
- [19]Wang, Y., and Wang X.J.(2005), " A New Approach to feature selection in Text Classification", Proceedings of 4th International Conference on Machine Learning and Cybernetics, IEEE- 2005, Vol.6, pp. 3814-3819, 2005.
- [20] Yan, J., Liu, N., Zhang, B., Yan, S., Chen, Z., Cheng, Q., Fan, W., and Ma, W., "OCFS: Optimal Orthogonal centroid Feature selection for Text Categorization." 28 Annual International conference on Reserch and Informational reterival, ACM SIGIR, (2005).
- [21] Zi-Qiang Wang, Xia Sun, De-Xian Zhang, Xin Li "An Optimal Svm-Based Text Classification Algorithm" Fifth International Conference on Machine Learning and Cybernetics, Dalian,pp. 13-16 , 2006.Barizal, , pp.122-129, 2005.

- [22] Montanes,E., Ferandez, J., Diaz, I., Combarro, E.F and Ranilla, J., " Measures of Rule Quality for Feature Selection in Text Categorization", 5th international Symposium on Intelligent data analysis , Gerneny-2003, Springer- Verlag 2003, Vol2810, pp.589-598, 2003.
- [23] Al-Shalabi R, and Evans M., "A Computational Morphology System For Arabic", in proceedings of Computational approaches to semitic languages workshop (COLING'98), Montreal, Canada, pp.58-65,1998.
- [24] El-Sadany T. A. and Hashish M. A. , "An Arabic Morphological System ." IBM Systems Journal, vol. 28, no 4,pp 600-612,1989.
- [25] Gheith M. and El-Sadany T., "Arabic Morphological Analyzer on a personal Computer", in proceedings of the Arabic Morphology Workshop Stanford University, California, USA, pp55-65, 1987.
- [26] Hilal Y, "Automatic Processing of Arabic Language and application [in Arabic] ", in proceedings of the 1st Kuwaiti Computer conference, Kuwait, pp. 145-171, 1989.
- [27]Al-Shalabi, R., Kanaan, G., and Muaidi, H. (2003). New Approach for Extracting Arabic Roots. Proceeding of the International Arab Conference on Information Technology. Alexandria, Egypt.
- [28] Vladimir, N., Vapnik. 1995. The Nature of Statistical Learning Theory. Springer-Verlag Berlin.
- [29] Thorsten Joachims. "Text categorization with support vector machines: learning with many relevant features". In Proceedings of the 10th European Conference on Machine Learning ECML-98, Chemnitz, Germany. Pages 137–142. 1998.
- [30] Cristianini, N., and J. Shawe-Taylor. 2000 An Introduction to Support Vector Machines (and other kernel-based learning methods). Cambridge University Press
- [31] Tarek Fouad Gharib, Mena Badieh Habib , and Zaki Taha Fayed, "Arabic Text Classification Using Support Vector Machines", International Journal of Computers and Their Applications, Vol (16), Issue(4) ,(2009).
- [32] W.S. Noble. What is a support vector machine? Nature Biotechnology, vol (24) Number (12) 2006.
- [33] N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines. Cambridge UP, Cambridge, UK, 2000
- [34] Russell, Stuart; Norvig, Peter (2003) [1995]. Artificial Intelligence: A Modern Approach (2nd ed.). Prentice Hall. ISBN 978-0137903955
- [35] Saleh Alsaleem, Automated Arabic Text Categorization Using SVM and NB, International Arab Journal of e-Technology, Vol. 2, No. 2, June 2011
- [36] Rish, Irina (2001). An empirical study of the naive Bayes classifier . IJCAI Workshop on Empirical Methods in AI.
- [37] Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain". Psychological Review 65 (6): 386–408
- [38] M. Caudill, and C. Butler, "Understanding Neural Networks: Computer Explorations", Vols. 1 and 2, MIT Press, Cambridge MA, USA, 1992.
- [39] Fouzi Harrag, Eyas Al-Qawasmah," Improving Arabic Text Categorization Using Neural Network with SVD", Journal of Digital Information Management, Volume 8 Number 4 □August 2010
- [40] D. Rumelhart, G. Hinton, and R. Williams, "Learning Internal Representations by Error Propagation", Parallel Distributed Processing, MIT Press, Cambridge MA, USA 1986
- [41] El-Kourdi, M., Bensaïd, A., and Rachidi, T. "Automatic Arabic Document Categorization Based on the Naïve Bayes Algorithm," 20th International Conference on Computational Linguistics, 2004, Geneva
- [42] Abdelwadood Moh'd A Mesleh, "Chi Square Feature Extraction Based Svms Arabic Language Text Categorization System, Journal of Computer Science 3 (6): 430-435, 2007
- [43] Al-Harbi S., Almuhareb A., Al-Thubaity A., Khorsheed M. S., and Al-Rajeh A., "Automatic Arabic Text Classification," 9es Journéesinternationals, France, pp. 77-83, 2008.