

Article

A Systematic Review of Islamic Terrorist Webpages and Keywords

Hassan Awad Hassan Al-Sukhni¹, Madihah Mohd Saudi^{2,a}, Azuan Ahmad^{2,b}

¹Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia.
Email: hassan@usim.edu.my

²CyberSecurity and Systems Research Unit, Universiti Sains Islam Malaysia (USIM), Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia.
Email: ^amadihah@usim.edu.my, ^bazuan@usim.edu.my

Abstract— As the internet is considerably expanding and on-line information is increasing, the identification and detection of the large amounts of different web information become vitally important, particularly those of dark web or Islamic extremists. Webpages with extremist and terrorist content are believed to be the main factors in the radicalisation and recruitment of disaffected individuals who might be involved in terrorist activities at home or those who fight alongside terrorist groups abroad. In fact, the sheer volume of online data makes it practically impossible for authorities to carry out the individual examination for every webpage, post or conversational thread that might or might not be relevant to terrorism or contain terrorist sympathies. As terrorists exist within every nation and every religion, hence this paper presents a review and systematic analysis of existing webpages on Islamic terrorists. This include existing database of Islamic extremist words and existing techniques of web classifier for keywords. Based on this systematic review analysis, it will be the input to the formation of a new Islamic extremists WorldNet.

Keywords— Islamic terrorists; WorldNet; Web classifier; Systematic review analysis.

I. INTRODUCTION

Webpages displaying the contents of extremist and terrorist activities are suspected to be the essential factors in the recruitment of innocent civilians all over the world, who may be influenced and radicalized until they are apprehended to engage themselves in terrorism activities either in their homeland or even decide to fight together with terrorist groups abroad. It is alarming that the sheer volume of online data makes it virtually impossible for the authorities to scrutinize every webpage, conversational threads or social media posts in order to classify whether they are relevant to terrorism or those which contain elements of terrorist activities. Nowadays, the loose security of the internet has enabled easy access by the terrorist groups to vulnerable intelligence information to control their terrorism operations [1]. Of late, there are various researches that showed how terrorists abuse the usage of the internet to expand their commands to their networks in the whole world, control their activities, communicate with each other and develop their own intelligence system (C3I). For instance, Jenkins claims that terrorists use the Internet as a platform to broadcast their “terrorist news network” [2] that are proven as one of the influential techniques to reach a broader range of audience without being detected. Although one of the infamous site known as the Dark Web have recently gained attention from the government all around the world as well as getting media attraction, the general knowledge on how the Internet is used to campaign terrorism is somewhat limited. Therefore, it is high time that these issues pertaining to internet usage to promote terrorism to be investigated in detail as highlighted by the Institute for Security and Technology Studies (ISTS),

Dartmouth College [3] and Anderson [4]. Based on this situation, it is paramount to conduct computer guided data collection and analysis pertaining to this issue. Recent studies in this area have also shown that the rapid technological advancement opens up additional opportunities for authorities such as law enforcement and security agencies to use web classifier so as to accomplish effective and efficient solutions to solve the contemporary threats and other problems.

To date, web classifier is a commonly used method to identify topics, activities and group types of the Islamic terrorists all over the world [5]. This method has three essential processes [6-10]. The first one is called preprocessing of web content; which is a step to remove less important words and symbols in the web content. Secondly, the representation of webs in extracting the most vital information from the web is conducted. Later, the similarity among these web is compared. The final process of web classifier is to apply the classifier algorithm to track web detection of those that have Islamic extremists content on the basis of the similarities among deep learning. The current study aims to provide the restraints of web classifier in two stages, namely, the extraction of web and supervised feature selection related to the domain of Islamic terrorisms’ site. Although over the recent years, the area of web classifier has been extensively studied, it is still a popular domain among IT researchers as more improvements are required and called for it in that particular area [5-10]. Section II discusses two out of the three essential processes of web classifier. In Section III, Islamic extremists’ web classifier system is described. Lastly, Section IV presents the conclusion of this study.

II. PROCESS OF WEB CLASSIFIER

This section provides a comprehensive review of the prior studies in respect of extraction, as well as, the classifier algorithms of web classifiers. The first and second sub-sections provide detailed reviews of previous research on the subject of extraction terms and classifier algorithms on web classifiers. Apart from that, the later part of this section provides discussions on limitations in two out of the three processes of web classifiers.

A. Term Extractions

To extract the terms from indexed words is indeed a challenging task due to the fact that several terms or combinations of terms are substantial to be considered at this stage. To extract the information from the terms, various studies were conducted based on different focuses such as; employing named entity recognition [11], bag-of-words (BOW) [12-13], n-grams [14], as well as lemmatization algorithms [15-16]. Further to that, in improving the extraction step, previous researchers utilized ontology-based extraction [16], semantic extraction [16], Arabic word sense disambiguation [17-18], semantic word embedding [19], and semantic relationships [20]. Unfortunately, text data usually present drawbacks due to their high dimensionality, as well as, being ambiguous or having overlapping word senses. As a part of the solution in regards to this issue, previous studies proposed a variety of term extraction methods such as BOW,

N-grams, and named entity recognition to solve the high dimensionality of terms. The approaches mentioned above are referred to as syntactic extraction, while the overlapping word meanings like the semantic words found in Word Net are known as semantic extraction. This study initially discusses syntactic extraction and highlights the new proposal to improve the extraction method. As mentioned by Khoja [21], the verbs extraction method first deletes the suffixes and prefixes of terms, to which their roots are later checked in a dictionary. On occasion where the term is on the list, it comes back to the root; or it comes back to the original word without making any modification. A previous review of Arabic IR indicated that Arabic IR is possible to be improved, especially when the stems or roots are used in searching and indexing tasks. Al-Salemi and Aziz [14] evaluated the extraction method using BOW method and three levels of N-grams (3, 4 and 5). The results found that when using Naïve Bayes as a classifier, the BOW outperforms all tested N-grams. Other than that, a study by Al-Shamari [15] extracted nouns and verbs with a newly designed method known as lemmatization as a stemmer for Arabic words' root. The purpose of the algorithm is to select verbs and nouns from documents in Arabic based on prepositions and some related rules in other linguistic elements like the definite article "the". The result of that study showed that their proposed lemmatization stemmer outperforms other stemmers such as those of Khoja [21] and Larkey [22].

TABLE I
SUMMARY OF LITERATURE ON EXTRACTION

Authors	Dataset Used	Baseline Extraction	Outperform Extraction	Domain
[21]	Dataset manual	Khoja and Larkey stemmers	Noun with verbs for stemmer	Arabic clustering
[14]	TREC-2002	N-gram 3,4 and 5 level	BOW	Arabic classifiers
[11]	Arabic Wikipedia corpus	Name entity and their our system	Name entity and their our system are equivalent	Arabic classification
[12]	collection of news articles	BOW	BOW with support vector machines as classifiers better than K nearest neighbor	Arabic text categorization
[19]	Arabic TREC collection	BOW	word embedding similarities	Information Retrieval
[20]	Dataset manual	Ontology 'synonym, antonym, hypernym' and complex extraction	BOW	marker learning algorithm

As discussed in the review of literature on the extraction process shown in Table 1, previous researchers have highlighted in their experiments that syntactic extraction are divided into two methods' categorization, where one method extracts the verb words, while the other extracts noun words. However, these were not evaluated before as the extraction was directly conducted using stemmer process as mentioned previously in Section I. As highlighted by Al-Shamari [15], the verb words are known as important terms in the web. One of the concepts that support this approach is the notion of crime used by the Romanian nation in which they describe

the crime by two key aspects of 'nouns and verbs'. However, as shown in the previous studies, they used semantic word to replace the original word with their sense when the sense of the word is similar to it. The database is not created by using it to increase the training related to the Islamic terrorism and their patterns. After extracting the huge number of terms with their sense, the number of terms is not exactly important as features, which is the reason why they are needed to be reduced based on a technique known as feature selection, which is discussed in the next section.

B. Supervised Feature Selections

The feature selection and the data distribution have a high influence on the performance of web classifier processes and classifier algorithms such as extraction [23]. The latter tends to achieve the local minima rather than the global minimum. The results obtained are highly appreciated in most occasions. This is particularly when the selected initial features selection are approximately far apart in that it can frequently differentiate the key category or class in a given data. Furthermore, the extraction methods quality and the classifiers algorithms key processing are both influenced by the initializing process of features selection [23]. Therefore, the preliminary points play a significant role in the result quality. For example, the failure of the classifiers algorithm in identifying the main categories features in certain data is possible if they are similar or close. This failure can also take place mostly if the feature selection algorithm is left uncontrolled. Moreover, associating the features with some procedures of optimization is absolutely essential in order to be less reliant on a certain initialization and data. More significantly, it enhances the process of the performance of web classifiers algorithm. Besides, it leads to yield good preliminary feature selection and better performance in refining the features to find the optimal feature selections [23].

The utilization of nature-inspired meta-heuristic algorithms has been common in many sectors including computer science [24], data mining [25], industry [26], agriculture [27], computer vision [28], forecasting [29], medicine and biology [30], scheduling [31], economy [32] and engineering [33]. The current study suggests its use in web classifiers. However, meta-heuristics are commonly still unclear in web classifiers. This is primarily for web mining where similar measurement between webs is used but a lot of 'words' or terms are specified for every web. On the other hand, the measurement of distance among points on the space is employed by web mining. The difference between web mining and data mining was discussed by Hearst (1999) [34]. In particular, data mining utilizes non-textual data as opposed to web mining which uses web textual data [34].

This section comprehensively outlines the most famous feature selection techniques based on meta-heuristic algorithms such as the Genetic Algorithms feature selection (GAs) [35], the Ant Colony Optimization feature selection (ACO) [36], the Particle Swarm Optimization feature selection [37] and the Harmony Search feature selection [38]. However, the GA has shortcomings [35, 39]. Therefore, further investigation is necessitated. The GA shortcomings are the uses of complex operators for selection and cross over, trapping into local optima, taking long runtime and weak local search. Behjat et al. [40] employed genetic algorithm as feature selection with Multi-Layer Perceptron as a classifier utilizing bag of word as extraction,

which they compared with support vector machine, Neural Network and Naïve Bayes, their proposed of genetic algorithm as feature selection with Multi-Layer Perceptron as classifier outperform other classifiers.

The ACO algorithm for feature selection was proposed by Rungsawang et al. [41], where they used Ant Colony as feature selection to decrease the number of features utilizing BOW as extraction and two classifiers, namely; C.45 and support vector machine. Based on their result, the Colony as feature selection is better than C.45 and support vector machine without feature selection. Other work conducted by El-Alfy [42] employed Ant Colony as a feature selection utilizing BOW and three classifiers, namely, Multi-Layer Perceptron, Ripper classifiers and Naïve Bayes; his results revealed that the Ant Colony as feature selection is better than three classifiers without feature selection.

Behjat et al. [43] used Binary Particle Swarm Optimization as feature selection and Multi-Layer Perceptron as classification and bag of word as extraction; they compare with information gain as feature selection and three classifiers, namely, BP Neural Network, Linear Discriminant and support vector machine. Their proposed output from genetic algorithm and other feature selection of information gain are the feature selection and other classifiers like BP Neural Network, Linear Discriminant and support vector machine. On the other hand, Moradi and Gholampour [44] used particle swarm optimization for the subset selection of feature, utilizing k-nearest neighbor as classifier, and BOW as extraction and on the other side they utilized filter-based methods among which term variance, information gain, fisher score and mRMR and five familiar wrapper-based approaches including particle swarm optimization, genetic algorithm, simulated annealing and ant colony optimization. Wang et al. [10] utilized fresh document frequency and term frequency combined feature selection method (DTFS) in order to improve the text classification performance by the proposed optimal document frequency based feature selection approach (ODFFS) and a predetermined threshold were applied for selecting the most discriminative features. The other proposed approach was the optimal term frequency based feature selection (OTFFS) approach and another predetermined threshold was applied to choose more discriminative features, then they combined ODFFS and OTFFS to select the remaining features. The feature selection of optimization utilized global best harmony oriented harmony search (GBHS) with fuzzy Support Vector Machine (FSVM) and Naïve Bayesian (NB) classifiers; their experimental results revealed that DTFS outperforms other approaches like Chi-square, t-test based feature selection, comprehensively measure feature selection, term frequency based information gain, two step based hybrid feature selection approach and improved term frequency inverse document frequency approach on six corpuses.

TABLE II
ADVANTAGES AND DISADVANTAGES OF THE MAIN OPTIMIZATION FEATURE SELECTION

		Advantages	Disadvantages
Meta-heuristic feature selection	GA PSO ACO HS	1. Tend to move relatively quickly toward high-quality solutions and thus offer an effective way of dealing with large complicated issues. 2. Helpful in cases in which traditional feature selection are stuck at the local optimum.	1. Have no guarantee to find optimal feature selection solution. 2. Many of the heuristic approaches needed for parameter tuning are based on the used dataset. 3. Initial population and the number of iterations can impact the quality of the final solution. 4. Some heuristic approaches suffer from slow convergence rate while others suffer from premature convergence. 5. Some heuristic algorithms have weak local search while others have weak global search.

C. Limitations of Feature Extraction and Term Extraction

The key limitations of literature summarized in Table 1 and 2 are presented as follows:

- The researchers presumed that the syntactic "word-level" or semantic "sense-level" are good for extracting information from the Islamic terrorists' web without overcoming the lack of words related to Islamic terrorists such as creating rich training of sense words extracting, and simultaneously avoiding the extraction of the insignificant features [11-22].

- The researchers assumed that feature selection can decrease the number of features [23-40] in carrying out a comparison between "Harmony Search, genetic algorithm and others mentioned before" of feature selection without detecting the weaknesses of these algorithms for Islamic terrorists' web. They offered a solution for such weaknesses by hybrid [41-44] with other algorithms and then compared the applied algorithms with another one so as to increase their performance, without addressing the effect with other processes in the web classifiers such as classifiers algorithms and extraction method.

- All the proposed works generally dealt with the weakness of web classifiers in respect of promoting a part of the process, where the output of every stage was influenced by the next process accuracy [6, 8, 9, and 12]. In short, the problems need to be detected one by one for each of the effects to take place on the other processes.

- As mentioned above, there are three processes of web classifiers, from the bottom of "classifiers algorithms" up to "extracting the terms from the Islamic terrorists' web". Thus, the notion of working from bottom to up guides the user to identify or describe the problem of extracting information from Islamic terrorists when sorting out the weakness of feature selection algorithms. However, it can be said that it is so straightforward to detect the extracting information problem and recognize the information which should be avoided.

D. Need

It is a known fact that all the current approaches of identifying and detecting Islamic terrorists' web have presumed that all the Islamic terrorists do exist. However, there are two issues which are relevant to feature selection algorithms, in which the performance of extraction methods and classifier algorithms heavily relies on selecting the number of features to avoid misclassification. Thus, it is expected that the result of this approach is often suboptimal [27] as well as an issue relevant to the features of extraction [6, 8, 9, and 12]. Each of these two problems affects other processes [6, 8, 9, and 12]. Consequently, to overcome such challenges, this study provides high performance and effective alternatives to assist in detecting and identifying the Islamic terrorists' web or types of Islamic extremists in webs by using the proposed approach as will be discussed in the next section.

III. PROPOSED WEB CLASSIFIER

A. Data Collection and Performance Measure

In this study, the datasets have been gathered manually from the news whereby half of it is test datasets and the remaining are training datasets. From an approximate number of 2000 websites evaluated for this study purpose, at least half of them are found to be related to the Islamic extremists. In order to search the related websites, the researchers generated a lot of vocabulary using Arabic WorldNet to gather as many important words related to the Islamic extremists while the other dataset is suggested to be used as a testing dataset of dark web to evaluate this study's proposal. Furthermore, the current study uses a general accuracy measure and a general F-measure so as to measure the external quality. Such measures are popular web classifier measures [45]. The existence of higher general accuracy and F-measures provides the best classifier detection.

B. Proposed of Extraction Features

As mentioned in Section II, regarding the weakness of extraction, syntactic extraction such as lemmatization is proposed to be employed [15] to extract information based on three questions, namely, where, who, and when, or regarding semantic "replace the original words" [16-20]. These methods however do not collect rich terms of

"information related to Islamic extremists". To solve that, this study suggests utilizing Arabic Word Net to generate all related words "vocabularies" to increase the number of words related to the Islamic extremists. These words play a significant role to be extracted in that most of such features are often utilized to describe Islamic extremists in general (syntactic \cup semantic).

TABLE III
SOME OF PREVIOUS WORK ON CLASSIFIERS USING OPTIMIZATION FEATURE SELECTION

Authors	Year	Data set	Extraction	With or without feature selection	Feature selection	Optimization feature selection	Classifiers
[40]	2012	LingSpam benchmark corpora	BOW	GA	GA	Multi-Layer Perceptron with GA	Multi-Layer Perceptron, SVM and Naïve Bayes.
[43]	2013	LingSpam benchmark corpora	BOW	BPSO	BPSO	Multi-Layer Perceptron with BPSO	BP Neural Network, support vector machine and Linear Discriminant
[10]	2015	PU2, CSDMC2010, PU3, Lingspam, Enron-spam and Trec2007	BOW	with	Chi-square, Information gain, Improved Gini index, Comprehensively measure feature selection, Ambiguity measure, t-test based feature selection, Term frequency based information gain and Improved term frequency inverse document frequency	global best harmony oriented harmony search	fuzzy Support Vector Machine and Naïve Bayesian
[41]	2011	Not Provided	BOW	with	ant colony	ant colony	C.45 and SVM
[42]	2009	spambase	BOW	with	ant colony optimization	ant colony optimization	Multi-Layer Perceptron, Naïve Bayes and Ripper classifiers
[44]	2016	12 datasets taken from the UCI machine learning repository	BOW	with			

C. Proposed Krill Herd Algorithm as Feature Selection

In order to overcome the unimportant features problems as mentioned in Sections II, Krill Herd is used as feature selection whereby they are reflected in three classifiers namely; Nave Bayesian, Support Vector Machine and K-nearest Neighbor to track and detect as illustrated in Fig.2. It is also to reflect the proposed extraction on feature selection and three classifiers algorithms. For that, the characteristics generated from the Krill Herd as feature selection will cover the problems of huge features and extract the core features from the Islamic terrorists' web. In addition, the proposed krill herd as feature selection will help avoid misclassification and increase the performance of web classifier.

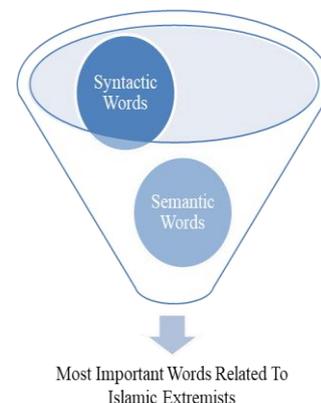


Fig.1 Proposal of Method to Extract Most Important Features of Islamic Extremists

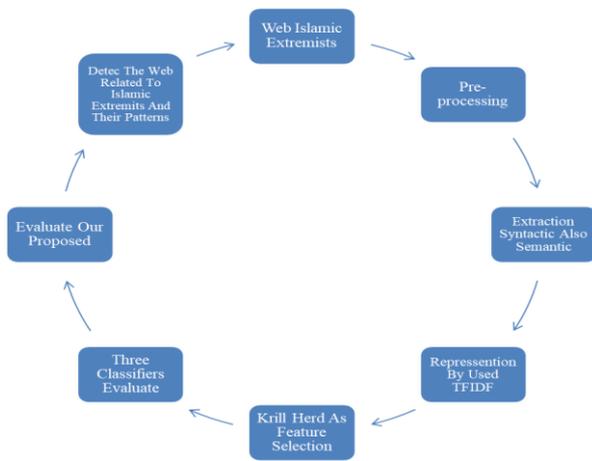


Fig. 2 Proposal of Enhanced Web Classifier for Islamic Extremists' Web to Avoid Unimportant Features

V. CONCLUSION

This study identifies and detects some limitations in Islamic terrorists' web. Firstly, it tackles the fault identification and detection in the extraction methods. Then, it examines the weakness of avoiding un-important features by using krill herd algorithm and their reflection of two processes on three classifier algorithms namely; Nave Bayesian, Support Vector Machine and K-nearest Neighbor. Consequently, this study aims at enhancing the reliability of Islamic terrorists' web detection by efficient extractions as well as the feature selection of web classifiers. Furthermore, it is used for Islamic terrorists' web classifiers where its results are the best testimony for its efficiency. This is because it aims to enhance the feature selection for web classifiers, as well as, the extraction of information, which tracks and detects the Islamic terrorists' web that can outperform the traditional web classifiers based on two criteria of time and performance. It is hoped that this study's suggestion will enhance the performance and the effectiveness of detecting Islamic Terrorists' web.

REFERENCES

- [1] S. M., Schneider, K., Foot, M., Kimpton, and G. Jones, "Building thematic web collections: challenges and experiences from the September 11 Web Archive and the Election 2002 Web Archive," 3rd ECDL Workshop on Web Archives. Trondheim, Norway: 2003.
- [2] (2004) Institute for Security Technology Studies: "Examining the Cyber Capabilities of Islamic Terrorist Groups". Report, ISTS. [Online]. Available: <http://www.ists.dartmouth.edu/>
- [3] (2004) Jenkins, B. M.: *World Becomes the Hostage of Media-Savvy Terrorists: Commentary*. USA Today. [Online]. Available: <http://www.rand.org/>.
- [4] A. Anderson, "Risk, Terrorism, and the Internet. Knowledge," *Technology & Policy*, vol. 16(2), pp. 24-33, 2003.
- [5] E. Reid, J. Qin, Y. Zhou, G. Lai, M. Sageman, G. Weimann, and H. Chen, "Collecting and analyzing the presence of terrorists on the web: A case study of jihad websites," in *International Conference on Intelligence and Security Informatics* (pp. 402-411). Berlin, Heidelberg: Springer, 2005.
- [6] N. Abe, and M. Kudo, "Non-parametric classifier-independent feature selection," *Pattern Recognition*, vol. 39 (1), pp. 737-746, 2006.
- [7] S. Hong, W. Lee, and M. Han, "The feature selection method based on genetic algorithm for efficient of text clustering and text

- classification," *Interbational Journal Advance Soft Compu. Appl*, vol. 7(1), pp. 2074-8523, 2015.
- [8] S. Manek, D. Shenoy, M.C. Mohan, and K.R. Venugopal, "Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier," *World wide web*, vol. 20(2), pp. 135-154, 2017.
- [9] A. Onan, S. Korukoğlu, & H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Systems with Applications*, pp. 232-247, 2016.
- [10] Y. Wang, Y. Liu, L. Feng, and X. Zhu, "Novel feature selection method based on harmony search for email classification," *Knowledge-Based Systems*, pp.311-323, 2016.
- [11] F.B. Mesmia, K. Haddar, N. Friburger, and D. Maurel, "CasANER: Arabic Named Entity Recognition Tool," in *Intelligent Natural Language Processing: Trends and Applications*. Springer, Cham., pp. 173-198, 2018.
- [12] I. Hmeidi, B. Hawashin, and E. El-Qawasmeh, "Performance of KNN and SVM classifiers on full word Arabic articles," *Advanced Engineering Informatics*, vol. 22(1), pp. 106-111, 2008.
- [13] K., Alsmearat, M. Al-Ayyoub, and R. Al-Shalabi, "An extensive study of the bag-of-words approach for gender identification of arabic articles," in *Computer Systems and Applications (AICCSA), IEEE/ACS 11th International Conference*, November 2014, pp. 601-608.
- [14] B. Al-Salemi, and M.J.A. Aziz, "Statistical bayesian learning for automatic arabic text categorization" *Journal of computer Science*, vol. 7(1), pp. 39, 2011.
- [15] Al-Shammari, "Lemmatizing, stemming, and query expansion method and system", patent application, Pub.no.:US 2010/0082333 A1, Apr.1, 2010.
- [16] A.M. Al-Zoghby, A. Elshawi, and A. Atwan, "Semantic Relations Extraction and Ontology Learning from Arabic Texts—A Survey," *Intelligent Natural Language Processing: Trends and Applications*. Springer, Cham, pp. 199-225, 2018.
- [17] B. Elayeb, "Arabic word sense disambiguation: a review," *Artificial Intelligence Review*, pp.1-58, 2018.
- [18] S.A. Salloum, A.Q. AlHamad, M. Al-Emran, and K. Shaalan, "A Survey of Arabic Text Mining," *Intelligent Natural Language Processing: Trends and Applications*. Springer, Cham, pp. 417-431, 2018.
- [19] A. El Mahdaouy, S.O. El Alaoui, and E. Gaussier, "Improving Arabic information retrieval using word embedding similarities," *International Journal of Speech Technology*, pp.1-16, 2018.
- [20] A. Benabdallah, M.A. Abderrahim, and M.E.A. Abderrahim, "Extraction of terms and semantic relationships from Arabic texts for automatic construction of an ontology," *International Journal of Speech Technology*, vol. 20(2), pp. 289-296, 2017.
- [21] S. Khoja, "APT: Arabic part-of-speech tagger proceedings of the student workshop," at the *second meeting of the north American chapter of the ocaion for computational linguistics (NAACL01)*, pp. 20-26, 2001.
- [22] S., Larkey, L., Ballesteros, and E. Connell, "Improving stemming for Arabic information retrieval: light stemming and co-occurrence analysis," in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. Tampere, Finland: ACM*, 2002, pp. 275-282.
- [23] M. Abualigah, T. Khader, and A. Al-Betar, "Unsupervised feature selection technique based on harmony search algorithm for improving the text clustering," *Computer Science and Information Technology (CSIT)*, 7th International Conference, pp. 1-6, 2016.
- [24] B. Akay, and D. Karaboga, "Parameter Tuning for the Artificial Bee Colony Algorithm," *Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems*, Berlin, Heidelberg: Springer, vol. 5796, pp. 608-619, 2009.
- [25] S. Rana, S. Jasola, and R. Kumar, "A Review on Particle Swarm Optimization Algorithms and Their Applications to Data Clustering," *Artificial Intelligence Review*, vol. 35(3), pp. 211-222, 2011.
- [26] W. Fox, H. Xiang and Lee., "Industrial applications of the ant colony optimization algorithm," *The International Journal of Advanced Manufacturing Technology*, vol. 31, pp. 805-814, 2007.
- [27] C. Geem, Y. Tseng, and B. Park. "Harmony search for generalized orienteering problem best touring in China," *Lecture Notes in Computer Science, Advances in Natural Computation*, vol. 3612, pp. 741-750, 2005.
- [28] E. Connolly, R. Granger and Sabourin., "An adaptive classification system for video-based face recognition," *Information Sciences*, vol. 192, pp. 50-70, 2012.

- [29] Chaturvedi, "Applications of genetic algorithms to load forecasting problem," *Soft Computing*, Berlin ,Heidelberg: Springer, 2008, pp. 383–402.
- [30] E. Christmas, T.M. Keedwell, J.R.B. Frayling and Perry, "Ant colony optimisation to identify genetic variant association with type 2 diabetes," *Information Sciences*, vol. 181, pp. 1609–1622, 2011.
- [31] W.D. Guo, A.R. Li, G.W. Mileham and Owen., "Applications of particle swarm optimisation in integrated process planning and scheduling," *Robotics and Computer-Integrated Manufacturing*, vol. 25, pp. 280–288, 2009.
- [32] K. Zhang, Y. Liu, X. Tana and He, "Random black hole particle swarm optimization and its application" *IEEE International Conference Neural Networks and Signal Processing*, ICNNSP, pp. 359–365, 2008.
- [33] Manoj & Elias, "Artificial bee colony algorithm for the design of multiplier-less nonuniform filter bank transmultiplexer," *Information Sciences*, vol. 192, pp. 193–203, 2012.
- [34] M.A. Hearts, "Untangling text data mining" in *Proceedings of ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics*, University of Maryland June 20-26 (invited paper).
- [35] D. E. Goldberg, "Genetic Algorithms in Search Optimization and Machine Learning," 1stEd. Boston, USA: Addison-Wesley Reading Menlo Park, 1989.
- [36] P. Shelokar, V.K. Jayaraman, and B.D. Kulkarni, "An ant colony approach for clustering," *Analytical Chimica Acta*, vol. 509(2), pp. 187-195, 2004.
- [37] H. Aghdam, and S. Heidari, "Feature selection using particle swarm optimization in text categorization" *Journal of Artificial Intelligence and Soft Computing Research*, vol. 5(4), pp. 231-238, 2015.
- [38] R. Forsati, M. Mahdavi, M. Shamsfard, and M. R. Meybodi "Efficient stochastic algorithms for document clustering," *Information Sciences*, vol. 220, pp. 269-291, 2013.
- [39] O. Hrstka, and A. Kučerová, "Improvements of real coded genetic algorithms based on differential operators preventing premature convergence," *Advances in Engineering Software*, vol. 35, pp. 237-246, 2004.
- [40] A.R. Behjat, A. Mustapha, H. Nezamabadi-pour, M.N. Sulaiman, and N. Mustapha, "GA-based feature subset selection in a spam/non-spam detection system," in *Computer and Communication Engineering (ICCCE)*, IEEE International Conference, pp. 675-679, 2012.
- [41] A. Rungsawang, A. Taweessiriwate, and B. Manaskasemsak, "Spam host detection using ant colony optimization," in *IT Convergence and Services*, Netherlands: Springer, pp. 13-21, 2011.
- [42] E.S.M. El-Alfy, "Discovering classification rules for email spam filtering with an ant colony optimization algorithm," in *Evolutionary Computation, 2009, CEC'09, IEEE Congress*, pp. 1778-1783.
- [43] A.R. Behjat, A. Mustapha, H. Nezamabadi-pour, M.N. Sulaiman, and N. Mustapha, "A PSO-Based Feature Subset Selection for Application of Spam/Non-spam Detection" in *Soft Computing Applications and Intelligent Systems*, pp. 183-193, 2013.
- [44] P. Moradi, and M. Gholampour, "A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy," *Applied Soft Computing*, vol. 43, pp. 117-130, 2016.