I. J. Education and Management Engineering, **2021, 4, 35-42**
Published Online August 2021 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijeme.2021.04.04

Evaluation of Machine Learning Techniques for Email Spam Classification

Mahmoud Jazzar
Palestine Technical University – Kadoorie, Faculty of Graduate Studies, Tulkarem, P.O. Box 7, Palestine.
Email: mjazzar@ptuk.edu.ps

Rasheed F. Yousef
Palestine Technical University – Kadoorie, Faculty of Graduate Studies, Tulkarem, P.O. Box 7, Palestine.
Email: r.f.yousef1@students.ptuk.edu.ps

Derar Eleyan
Palestine Technical University – Kadoorie, Faculty of Graduate Studies, Tulkarem, P.O. Box 7, Palestine.
Email: d.eleyan@ptuk.edu.ps

Received: 03 February 2021; Accepted: 07 March 2021; Published: 08 August 2021

Abstract: Electronic mail (Email) is one of the official and very common way of exchanging data and information over digital and electronic devices. Millions of users worldwide use email to exchange data and information between email servers. On the other hand, unwanted emails or spam became phenomenon challenging major companies and organizations due to the volume of spam which is increasing dramatically every year. Spam is annoying and may contain harmful contents. In addition, spam consume computers, servers, and network resources, causes harmful bottleneck, effect on computing memory and speed of digital devices. Moreover, the time consumed by the users to remove unwanted emails is huge. There are many methods developed to filter spam like keyword matching blacklist/whitelist and header information processing. Though, classical methods like blocking the source to prevent the spam are not effective. This study demonstrates and reviews the performance evaluation of the most popular and effective machine learning techniques and algorithms such as Support Vector Machine, ANN, J48, and Naïve Bayes for email spam classification and filtering. In con conclusion, support vector machine performs better than any individual algorithm in term of accuracy. This research contributes on the for the development of methods and techniques for better detection and prevention of spam.

Index Terms: Spam, spam filtering, machine learning algorithms, email classification.

1. Introduction

In this era of intelligent computing, user environment of software systems transformed farther complex. According to recent recorded spams from the world traffic, unwanted commercial bulk emails became huge challenge for email service providers. As such, the volume of spam reached around 53.95% of the world traffic during March 2020 [1]. The spammer collects valid email address using various ways such as website, chatrooms, and using computer viruses [2]. Spams are absolute and effective way to send millions of advertisements with no cost. Although spams are really annoying for digital users, it may cause serious hard effect on network performance, storage, mail servers, CPU, and computing memory. In addition, spams cause huge effect on computer network performance and consume digital users time to delete and get rid of the junk mails which on other hand affects the productive time [3]. On other hand, spams may contain infected links to disclose sensitive data like passwords, bank accounts and credit card information [2]. Further, spam may contain pornography contents in which should be not exposed to underage users [4]. As such, spam email may help cybercriminal for phishing, denial of service (DoS) attacks and more [5].

There are many forms of spam such as search engine spam(spamdexing), blog spam, social network spam, chatroom spam and e-mail spam [5]. Many methods and techniques were considered for stopping and fighting such digital harm. As such, governments enacted laws against the spam such as the Law of anti-spam which introduced in US, likewise awareness campaigns similar to do not reply, do not provide your email address on general websites and servers and never forward chain-letters [5].

Technically, there are many methods used so far to prevent and mitigate spam and junk mails. For example, blocking the spammers IP address and e-mail filtering. Email filtering techniques usually works based on the contents of the message, searching for specific phrases, words, and particular expressions. However, spammers started to avoid such methods by sending text messages contain no html codes and download images [5]. The volume and severity of spam is still and issue, Figure 1 below demonstrate the continuous severity of spam and the need for efficient spam filtering and detection techniques.

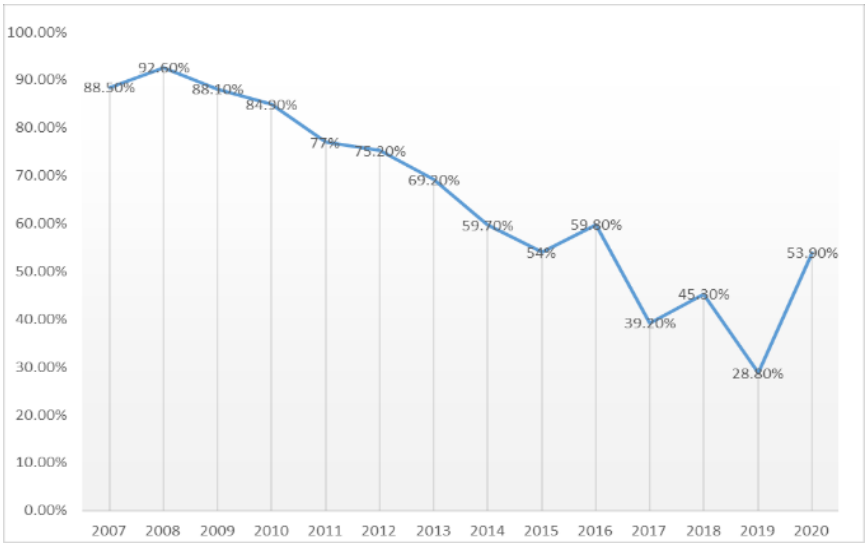

Fig.1. Graphical representation of global spam volume.

In general, there are major approaches used so far for spam filtering such as knowledge engineering and machine learning (ML). In knowledge engineering-based spam filtering, rules are developed and provided. As such, the detection is technically based on special keywords to categorize spam from ordinary email. This approach is considered as peer-to-peer approach in which requires regular rule update. The concern of this method strikes if the spammer can edit the rule to give spam emails special exception. On the other hand, peer to user approaches such as machine learning based filtering techniques is totally based on learning algorithm which provide more user convenient space. As such, no predefined rules are required. This study focusses on machine learning methods for email spam filtering and classifications. The most common machine learning algorithms types shall be reviewed for performance evaluation and for the capability of reduction of the annoying volume of spam.

2. Background

There has been lots of work devoted to spam detection and classification. As such, studies on email spam include spam detection using machine learning techniques [6,7]. In addition, machine learning algorithms competently used for classification and behavior detection of verity of data sets such as Coronary Artery Disease and more [8]. Related works on spam detection and classification includes ensemble methods, particle swarm optimization, and different machine learning algorithms and neural networks, however, the classification and accuracy of these algorithms is still an issue [7,9].

Spam is those unwanted electronic mails. The biggest challenge for spam filtering is that spam filters may categorize genuine email as spam or the filter actual spam email as real and genuine email. Spam and email classification is a common problem today for mail servers, network performance, and for the security of data and information. The volume of spam is frequently increasing inconstantly [10,15]. The challenge is to develop various methods to catch and label spam emails before it reaches to the user inbox.

Many techniques present short-term solutions to this problem such as black list/white list, keyword matching and header information analysis and the Bayesian classification. In black and white list methods, user act as the administrator of the filter, detected users or domains can be added as white or black list. However, such methods may not be very effective for the long term. Rule based methods have major disadvantages such as rules should be prepared by end users. In addition, the spam behaviors change over the time, therefore, using multiple spam filtering methods might be more effective than using a single spam filter tool [11].

There many techniques and methods found in the literature used for spam filtering, these methods can be categorized under the following five main categories [2].

- Content based filtering technique such as Support Vector Machine (SVM), K-Nearest Neighbour, Neural Networks and Na ïve Bayesian. This approach classifies emails by generating an automatic filtering rule based on the phrases, distribution of the words and the contents of the email [2].
- Case base spam filtering method is one of most popular approaches for spam filtering, known as sample base. Using modest collection method, all spam and ligament emails extracted for all the users. Passing through grouping process and evaluation, all data will be classified into two vectors for detection of spam and or for non-spam emails [2].
- Adaptive spam filtering technique: this approach detects and investigate the spam emails by grouping emails into groups, the incoming emails will be compared with each group. The detection whether the e-mail is spam or not spam anticipated by the percentage of similarity with the group which the email belongs to.
- Previous likeness-based spam filtering technique: this approach can be treated as instance and memory base filtering method. The used machine learning method such as k-nearest neighbour (kNN) filter and classify incoming emails with regarding to the training of particular instance [2].
- Rule based spam filtering technique such as spam assassin, this approach uses pre-defined rules or heuristics to assess massive number of patterns which are usually regular expressions against chosen message. A few similar patterns increase the rank of a message. Some ranking rules do not change over time, but other rules need constant updating to ensure effective filtering [2].

3. Impact of the Spam

Spam is not only annoying for end users, spam can be identified as cybercrime as they may harm individuals, corporate and governments. As such, spam became the most common way to distribute phishing emails and ransomware attacks. Therefore, anti-spam and email classifications methods become essential for the end user such that end users have no option to stop receiving emails from unknown sources. The following demonstrate some negative impact of the spam targeting the end user and the computing resources:

- Waste of the network performance and computing resources: receiving huge volume of spam consume the network resources and interrupt the bandwidth; which on the other hand; cause delay and bother on the functionality of the corporates.
- Cybercrime facilitation: spam is very effective method to send bulk messages, worms, and viruses to unlimited destinations with no cost.
- Spam consumes the time of the end user and delay ongoing operations flow.

4. Datasets and Machine Learning Techniques

One of the common and precession evaluation approach to examine the machine learning techniques for spam email filtering and for classification is to use well defined and clean dataset records. For this research, we used a dataset records from UCI machine learning repository [12]. The dataset contains 1367 spam e-mail and 4361as legitimate. In order to evaluated machine learning method we need to calculate the accuracy, precision, and recall. The general performance of the machine learning methods can be evaluated by calculating the positive and negative class values predictions or the overall accuracy as per the following.

$$Accuracy = (TP + TN) \,/\, (TP + FP + FN + TN) \tag{1}$$

Precision can be defined as the relation between correctly classified samples and those that are misclassified as positives. Precision can be calculated as per the following.

$$Precision = TP/(TP + FP) \tag{2}$$

Recall is the relation between properly classified instances and misclassified instances. Recall can be calculated as per the following.

$$Recall = TP/(TP + FN) \tag{3}$$

F-measure can be calculated as per the following to convey balance between the precision and the recall:

$$F - Measure = (2 * Precision * Recall)/(Precision + Recall) \tag{4}$$

Where:

$$True\ Positive\ Rate = TP/(TP + FN) \tag{5}$$

$$False\ Positive\ Rate = FP/(FP + TN) \tag{6}$$

$$True\ Negative\ Rate = TN/(TN + FP) \tag{7}$$

$$False\ Negative\ Rate = FN/(FN + TP) \tag{8}$$

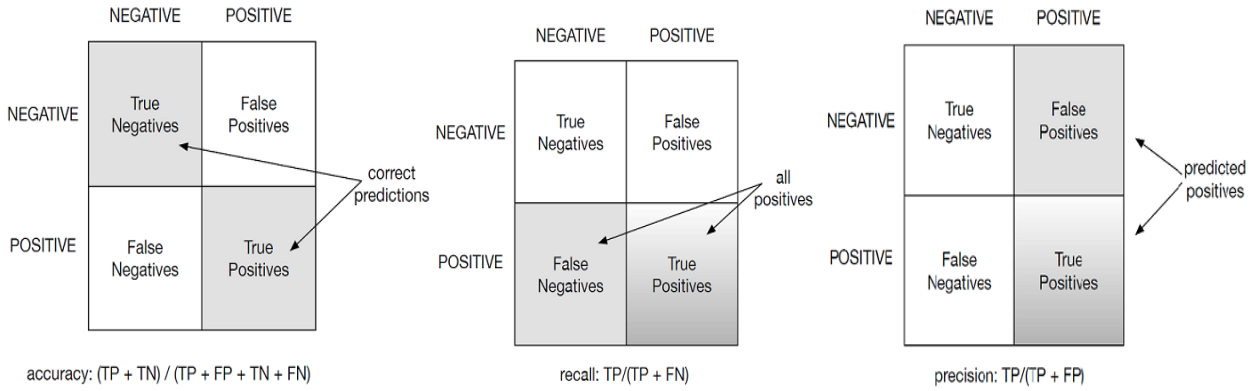

Fig.2. The label of accuracy, precision, and recall.

The machine learning repository (dataset) used to examine the following methods:

- J48 decision tree: Each branch node of the tree represents a choice between a number of alternatives and each leaf node represents classification or decision. The decision tree generated by J48 can be used for various classification problems. J48 algorithm uses greedy technique to induce decision tree for classification [13].
- Support Vector Machine (SVM): The concept of SVM is to separate two or more classes using a hyperplane that maximizes the margin between the classes.
- Artificial Neural Networks (ANN) Classifier: An artificial neural network is an adaptive system that changes its structure based on information and data flows through the artificial network during a learning phase. The ANN is based on the principle of learning by example. The attributes stream into the input layer, go through the hidden layers, and produce an output at the output layer [14,15]. Except for the input layer, every neuron gets signals from the neurons of the past layer straight weighted by the interconnect values between neurons. The neuron then creates its output by passing the summed signal through sigmoid or other types of activation function [11].
- Naive Bayes-classifier (NB): A Naive Bayes classifier is a probabilistic machine learning model that is used for classification tasks. The core of the classifier is based on the Bayes theorem. The NB classifier for spam can be defined as per the following.

$$NB = \arg\max P(ci)\ \ \prod_{a=1}^{b} P(wk|ci) \tag{9}$$

Where, $P(ci)$ is the set of target classes (spam or non-spam), and $P(wk|ci)$ is the probability that word wk occurs in the email, given that the email belongs to class ci [10]. The following diagram in Figure 3 illustrate the evaluation process using the dataset

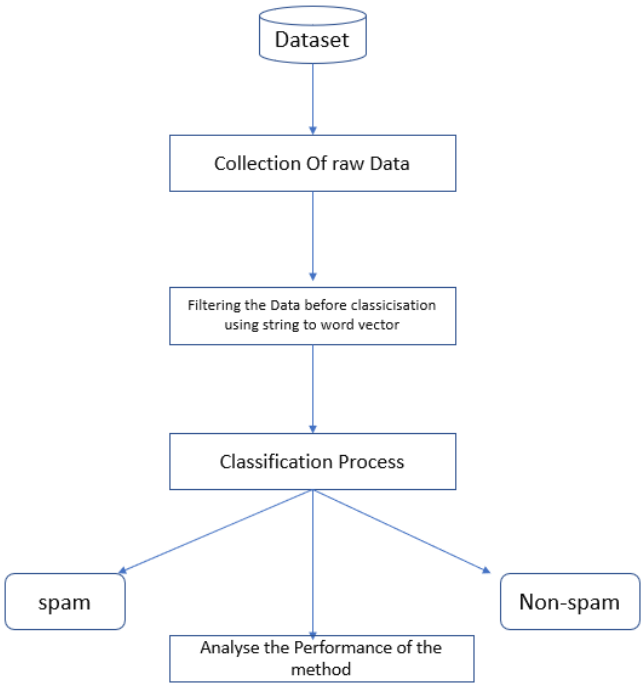

Fig.3. Evaluation Process.

5. Experimentation Result and Discussion

Table 1 below illustrate the precision, accuracy, and recall of the discussed methods. Experimental results demonstrate that Na ïve Bayes, J48, Support vector machine and ANN methods has an accuracy of 92.8%, 91.8%, 93.91% and 91.05% Consecutively. The results also demonstrate that the precision of the methods Na ïve Bayes, J48, Support vector machine and ANN as 93.98%, 91.21%, 92.98% and 89.09% Successively. In addition, recall of 90.22%, 87.9%, 90.23% and 88.19% uninterruptedly.

Table 1. Experimental Results for ML Methods Accuracy, Precision, and Recall

ML method	Accuracy	Precision	Recall
Naive Bayes	92.8	93.98	90.22
J48	91.8	91.21	87.9
Support Vector Machine	93.91	92.98	90.23
ANN	91.05	89.09	88.19

In spite of the training time required for SVM as compared with Naive Bayes and J48, the False Positive Rate is less for SVM. Experimentation results shows that SVM is the best methods in terms of accuracy and False Positive Rate (FPR). Technically, machine learning techniques used for algorithmic development require accuracy testing and training. Therefore, the developer must evaluate the machine learning method in order to precisely deploy the most optimal technique for spam filtering and classification. As example, false positive rate measure may determine and answer the question on the capability of spam filters to determine authentic from spam emails.

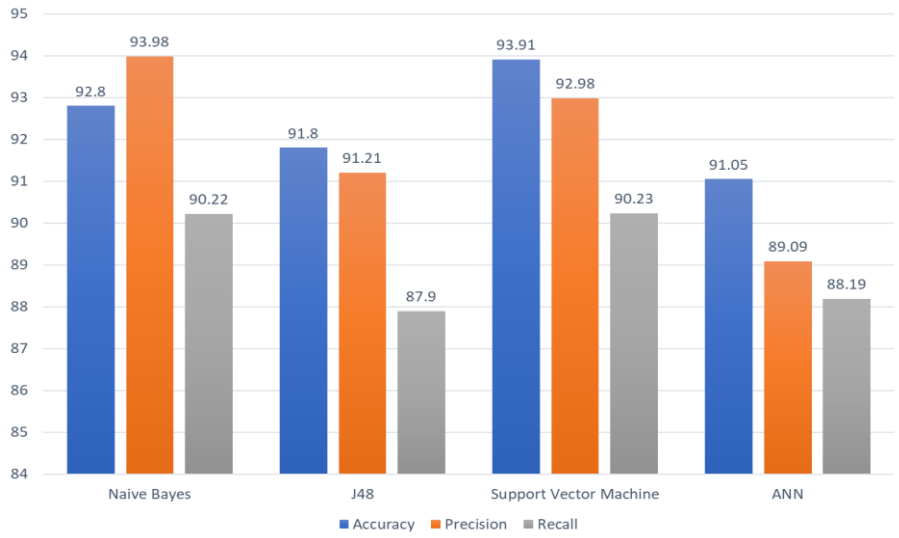

Fig.4. Graphical representation of experimental results.

Table 2 below demonstrate very close experimental results from the literature [13,14]. Therefore, we can conclude that the high value of recall and precision gives the SVM priority to be considered during the development of the filtering and classification mechanizes. SVM demonstrate the highest accuracy over the other machine learning methods which means the efficiency of email spam classification.

Table 2. Experimental Results for ML Methods Accuracy, Precision, and Recall from Previous Works

ML method	Accuracy	Precision	Recall
Naive Bayes	92.8	92.15	90.06
J48	92.07	91.61	88.09
Support Vector Machine	94.06	93.21	90.18
ANN	87.10	93.70	91.05

The following diagram in Figure 5 represent graphical representation and comparison based on experimental outputs with other studies such as in [13,14]. However, the following factors need to be considered when referring to such experimentation which include differences in the training data, stochastic learning algorithm, evaluation procedures, and the platform used during the evaluation.

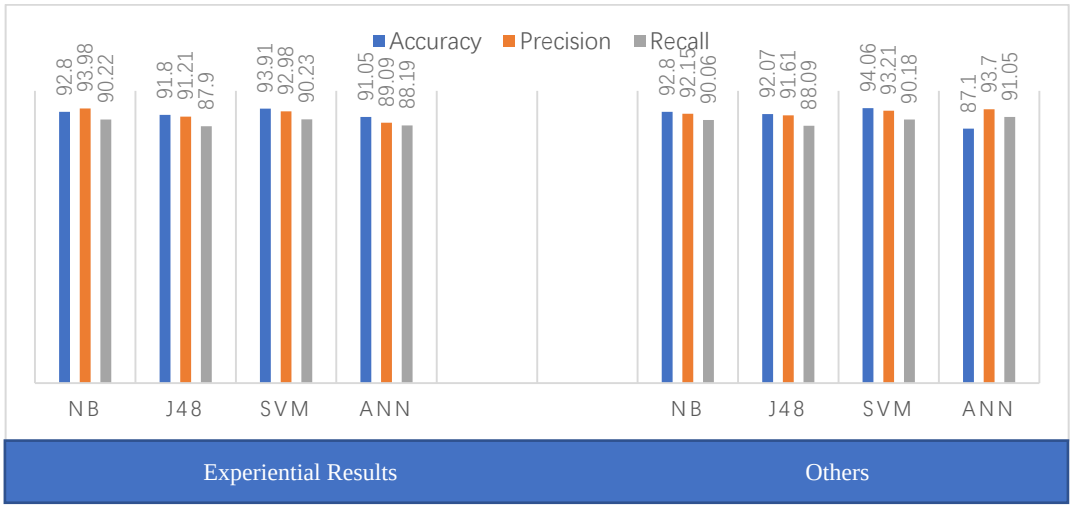

Fig.5. Graphical representation of experimental results vs. results from other studies.

6. Concluding Remarks

Due to the huge volume of spam increase, many techniques and methods have been developed and used. The aim is to reduce the volume of spam and hence reduce the volume of malware, phishing attacks and cybercrimes. However, many studies shows that it is very hard to completely prevent spam. On the other hand, the unexpected rise in volume of false positives and false negatives need huge attention. In order to not waste human and computing resources, the used method should be evaluated carefully taking into account the above mentioned. This study focused on evaluation of major and most popular machine learning methods and techniques. The evaluation was based on the accuracy of machine learning techniques and methods using machine learning data repository dataset. Experimentation concludes that SMV is one the best methods in term of accuracy and relevant false positive rate. The future work should focus on evaluation relevant to using such techniques and methods in social networks spam since the spam now not limited only for email, spammer similarly targeting social network sites and more.

Acknowledgment

The authors wish to thank Palestine Technical University-Kadoorie (PTUK) for supporting this research work as part of PTUK research fund.

References

[1] Statista, “Global Spam Volume,” https://www.statista.com/statistics/420391/spam-email-traffic-share/, Retrieved Dec 18, 2020.
[2] E. G. Dada, J. S. Bassi, H. Chiroma, S. M. Abdulhamid, A. O. Adetunmbi, & E. O. Ajibuwa, “Machine learning for email spam filtering: review, approaches and open research problems," *Heliyon*, *5*(6), e01802. https://doi.org/10.1016/j.heliyon.2019.e01802
[3] B. Yu, Z. Xu, “A Comparative Study for Content-Based Dynamic Spam Classification Using Four Machine Learning Algorithms,” *Knowledge-Based Systems*, 21(4), 355–362. https://doi.org/10.1016/j.knosys.2008.01.001
[4] K. Tretyakov, “Machine Learning Techniques in Spam Filtering,” *Data Mining Problem-oriented Seminar*, MTAT.03.177, May 2004, pp. 60-79.
[5] A. Bhowmick, S. M. Hazarika, “E-Mail Spam Filtering: A Review of Techniques and Trends,” *In Lecture Notes in Electrical Engineering, Springer Singapore*, 2017; pp 583–590.
[6] S. M. Abdulhamid, M. Shuaib, O. Osho, I. Ismaila, J. K. Alhassan,"Comparative Analysis of Classification Algorithms for Email Spam Detection," International Journal of Computer Network and Information Security (IJCNIS), Vol.10, No.1, pp.60-67, 2018.DOI: 10.5815/ijcnis.2018.01.07
[7] M. Zavvar, M. Rezaei, S. Garavand,"Email Spam Detection Using Combination of Particle Swarm Optimization and Artificial Neural Network and Support Vector Machine,"International Journal of Modern Education and Computer Science (IJMECS), Vol.8, No.7, pp.68-74, 2016.DOI: 10.5815/ijmecs.2016.07.08
[8] B. Nazlı, Y. Gültepe, H. Altural, " Classification of Coronary Artery Disease Using Different Machine Learning Algorithms, " International Journal of Education and Management Engineering (IJEME), Vol.10, No.4, pp.1-7, 2020. DOI: 10.5815/ijeme.2020.04.01
[9] O. Oluwatoyin, A. Bodunde, G. Titus, A. Ganiyu, "An Improved Machine Learning-Based Short Message Service Spam Detection System," International Journal of Computer Network and Information Security (IJCNIS), Vol.11, No.12, pp.40-48, 2019. DOI: 10.5815/ijcnis.2019.12.05
[10] D. K. Renuka, T. Hamsapriya, M. R. Chakkaravarthi and P. L. Surya, "Spam Classification Based on Supervised Learning Using Machine Learning Techniques," *2011 International Conference on Process Automation, Control and Computing*, Coimbatore, 2011, pp. 1-7, doi: 10.1109/PACC.2011.5979035.
[11] A. W. Awad, “Machine Learning Methods for Spam E-Mail Classification,” *International Journal of Computer Science and Information Technology*, *3*(1), 173–184. https://doi.org/10.5121/ijcsit.2011.3112
[12] UCI, “UCI Machine Learning Repository,” https://archive.ics.uci.edu/ml/index.php, Retrieved Dec 12, 2020.
[13] M. Shajideen and B. V., "Spam Filtering: A Comparison Between Different Machine Learning Classifiers," *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, 2018, pp. 1919-1922, doi: 10.1109/ICECA.2018.8474778.
[14] S. Jukic, J. Azemovic, D. keco, and J. Kevric, “Comparison if Machine Learning Techniques in Spam E-mail Classification,” *Southeast Europe Journal of Soft Computing," Vol. 4 No.1, March 2015.*
[15] A.S. Aski, and K. N. Sourati, “Proposed efficient algorithm to filter spam using machine learning techniques,” *Pacific Science Review A: Natural Science and Engineering*, *18*(2), 145–149. https://doi.org/10.1016/j.psra.2016.09.017

Authors' Profiles

Mahmoud Jazzar is currently working as an assistant professor in computer science and director of the academic quality department at Palestine Technical University – Kadoorie. He served as director of Kadoorie center for innovation in teaching and learning during 2017 – 2018. Prior working at Palestine Technical University - Kadoorie, Jazzar worked as Dean with Royal University for Women in the Kingdom of Bahrain and as assistant professor in computer science with Al-quds University, Curtin University of Technology-Sarawak, and Birzeit University. Jazzar is member of IEEE Computer Society, IAENG, MySEIG, and the Malaysian Information Technology Society (MITS). He joined many organizing and technical program committees and as reviewer of many international conferences and journals. His main research lies in the area of Computer and Network Securities, Intrusion Detection and Protection, Forensics, and Intelligent Systems. He has supervised several research projects, published one book and several scientific research papers in his research domain.

Rasheed. F Yousef is a senior information security engineer, awarded maser degree in computer since from Al-Quds university-Palestine and honor degree in Information and telecommunication technology from Al-Quds Open University-Palestine in 2012 and 2008 respectively. He had experience in banking as network security specialist and now working as senior information security engineer and field consultant in cyber security domain for the government and private sectors. In addition, he is currently pursuing graduate degree of science in cybercrimes & digital evidence analysis at Palestine Technical University – Kadoorie.

Derar Eleyan has a good relevant diversity experience in academic and industry. He is the manager of the Erasmus+ project "Pathway in forensic computing" and associate professor in information systems. Eleyan is currently working as the president assistant for international academic cooperation. He served five years as an assistant professor at Birzeit University in the Department of Computer Science teaching variety of courses at the undergraduate and postgraduate levels. He has worked also as lecturer and course team leader in computing at South East Essex college of Arts and Technology where he taught various modules as, information system, project management, database, web database, and website management, Computer and Business Ethics, Research methods. He has a good expertise as an Information Systems Consultant at BAS Computer Systems a private company since 2003 till 2006. On 2006, He was a visiting lecturer at the University of Manchester, teaching an MSc course of Enterprise System Modelling, collaborating with Prof. Loucopoulos. He has served as an external reviewer to some conferences in information science, information systems and business process modelling. His research interests focus in System dynamics, software quality, Information Systems, Business Process Modelling, Customer service and satisfaction and return on investment, information technology management, IT project management, Quality of Service, Academic quality and performance evaluation, Business and computer ethics, Software testing quality assurance, Usability and e-commerce. He is a member of some societies as British Computer Society (BCS), Institute for Learning (IFL), Systems Dynamics Society (SDS).

How to cite this paper: Mahmoud Jazzar, Rasheed F. Yousef, Derar Eleyan, " Evaluation of Machine Learning Techniques for Email Spam Classification", International Journal of Education and Management Engineering (IJEME), Vol.11, No.4, pp. 35-42, 2021. DOI: 10.5815/ijeme.2021.04.04