Propose Hybrid ACO and NB to Enhance Spam Filtering System

Dr. Soukaena Hassan Hashem
Computer Science Department, University of Technology/Baghdad.
Huda Adil Abd Algafore
Computer Science Department, University of Technology/Baghdad.
Email:Asola_adel95@yahoo.com

Received on: 7/6/2015 & Accepted on:20/1/2016

ABSTRACT

Unwanted e-mails became one of the most risk experienced by e-mail users, which may be either harmless or e-mails that represent a threat to the internet. Filtering systems are used to filter e-mail messages from spam. This paper introduces a proposed hybrid system to filter the spam; the proposal hybrid Ant Colony System (ACS) and Naive Bayesian (NB) classifier. Where, ACS will depend on the Information Gain (IG) as a heuristic measure to guide the ants search to select the optimal worst features then omitting these features. The remind features will be the subset which is used to train and test NB classifier to classify whether the mail message spam or not. The proposal is experimented on spambase dataset, and the results show that; the accuracy, precision and recall with NB which use a subset of features extracted by proposing IG-based ACS is higher than the traditional NB with all set of features.
Keywords: ACS, IG, NB, spam filtering

INTRODUCTION

Email has been becoming these days one of the most important methods of communication because it is inexpensive and faster than other communication methods. But the misusing of e-mail leads to create variety threats to the internet such as spam e-mail. Spam e-mail is "the practice of sending unwanted e-mail messages frequently with commercial content, in large quantities to a random group of recipients"[1].In comparison, managing cost of a spam is noticeably more than sending cost which is considered negligible. This leads to waste network resources, and storage. Meanwhile, traffic congestion cost, and leads to waste employees' productivity [2].Spam filtering system has been used to solve the spam problem. Naive Bayesian classifier is an area of interest of data mining contributes on manual classification of useful messages and spam, also reporting impressive precision and recall on the unseen messages. It may be surprising that the classification of the text can be effective in spam filtering systems [3].The benefits of using feature selection is the improvement in predication accuracy in some cases. Distinguishing the appropriate feature in a text message [4].Information gain is a statistical filter mechanism. It measures the text pureness in order to use the training sample. IG mechanism depends on the probabilities dimensions. These dimensions are between two probabilities of dispensers category, i.e., two categories "spam and non-spam"[5]. ACO algorithm is especially appealing for feature selection because there is no heuristic that is able to lead search for the most optimum feature subset each time. Also, it may be the status in which ants detect the optimum group of features during the process the search space [6, 7].

This paper introduces proposal to build spam filtering system that automatically classifies the mail message whether spam or not. The proposal system aims to train Naive Bayesian as a

https://doi.org/10.30684/etj.34.2B.3
2412-0758/University of Technology-Iraq, Baghdad, Iraq

spam filter system by hybrid it with preprocessing step include computing information gain for each feature to be used as part of the heuristic measure in ACS, that to guide the ant search for solving feature selection.

Related Works

Androutsopoulos I., et al., 2000, debated Naive Bayesian classifier could be implemented as spam e-mail filter, they implemented overall evaluation on their proposed work to contribute benchmark, while they keep in mind several factors to find out the effect of them on the filter performance (attribute set size, training data set size, process of grouping together the different inflected forms of a word so they can be analyzed as a single item and stop list) [3].Jensen R., 2006, showed that the main goal of feature selection for general applications is the reduction of the subset size of the features scope problem, while maintaining the accuracy high, can be reached by removing unnecessary features using two mechanisms; Ant Colony Optimization and fuzzy-rough data reduction process. He applied his method to two systems: classification of web and monitoring complex systems [7].

Guzella T.S.,et al.,2009, presented review of developed machine learning algorithm for spam filtering. They discussed two important aspects that are not extensively known in the literature; difficulties of updating a classifier depending on group of word representation, and the great difference between two models of NB. Generally, they concluded that although important achievements have been made in the recent years, some aspects still need to be explored, especially under more realistic settings of evaluation [8]. Chakraborty N., et al., 2012, discussed how e-mail becoming one of the widely used method of communication between individuals because of it is cheapest and fastest way of communication. They propose using hybrid system Neural Network and Naive Bayesian to enhance the spam filter. By training NN to recognize different forms of often used words in spam mails [1].

Theoretical Background

This section explains the theoretical back ground of the proposed system that includes Ant Colony System, Information Gain, and Naive Bayesian.

Ant Colony System (Acs)

ACS is an ingredient of Ant Colony Optimization (ACO), that is metaheuristic inspired by real ant life. The follow of ACS:

1. Initially, "m" ants will be randomly positioned on "n" cities that are chosen according to certain initialization rule.
2. The State Transition Rule of ACS: When a tour is built in ACS, an ant "m" at the present position of node "i" selects the next move to the next node "j" by employing the rule of state transition by applying the following equation:
3.

$$j=\begin{cases} \arg\max_{u\in S^k(i)}\{[T(i,u)][[\eta(i,u)]^{\beta}\} & \text{if } q \le q_0 \\ J & \text{if } q \ge q_0 \end{cases} \quad \ldots (1)$$

Where $T(i,u)$ is the pheromone trace at edge (i,u). The heuristic desirability $\eta(i,u)=1/\delta(i,u)$is the inverse length from node "i" to node "u". $S^k(i)$ is a set of nodes that ant "k" keeps visiting and "i" represents the node where the ant is placed. Also, "β" is the parameter, which decides the importance of the relative pheromone against distance($\beta > 0$). "q" is a random number that is uniformly distributed in $[0,1]$, and q_0 is a parameter $(0 \le q_0 \le 1)$ which decides the relative importance of exploitation against exploration.The probability is known as random-proportional rule, and it is given in Eq.(2):

$$p^k(i,j) = \begin{cases} \frac{[T(i,j)][\eta(i,j)]^{\beta}}{\sum_{u\in S^k(i)}[T(i,j)][\eta(i,j)]^{\beta}} & \text{if } j \in S^k(i) \\ 0 & \text{otherwise} \end{cases} \quad .....(2)$$

4. Local Updating Rule of ACS: while an ant is building a tour, it changes the pheromone concentration on the visited edges by Eq.(3):

$$T(i,j) = (1-\rho)\,T(i,j) + \rho T_0 \quad ...(3)$$

Where, T_0 is the initial level of pheromone and ρ is the evaporating parameter of pheromone where $0 < \rho < 1$.

5. Global Updating Rule of ACS: updating will be performed after every ant that has completed its tours. In order to get a more directed search, by Eq.(4):
6.

$$T(i,j) = (1-\alpha)\,T(i,j) + \alpha\,\Delta T(i,j) \quad .....(4)$$

$$\text{Where } \Delta T(i,j) = \begin{cases} L_{gb}^{-1} & \text{if } (i,j) \in \text{globle best tour} \\ 0 & \text{otherwise} \end{cases} \quad ...(5)$$

Eq.(4), α $(0 < \alpha < 1)$ represents parameter of pheromone evaporating, and L_{gb} represents the global best tour length that is found up to the present iteration [9 and10].

Information Gain (Ig)

Being one of the ranking feature selection methods, IG has been proposed in most machine learning literatures. This technique purpose is to exclude redundant or irrelevant features from a certain feature vector. Entropy measure is commonly used to measure the purity of examples in an arbitrary collection. It is the basis of the ranking methods of IG attribute. It is considered system's unpredictability measure. However, the entropy of Y is:

$$H(Y) = -\sum_{y\in Y} p(Y)\log_2 p(Y) \quad ...(6)$$

Where $p(y)$ is the density function of marginal probability for the variable "Y" which is a random number [11 and 12].

Naive Bayesian (Nb)

Bayesian classifiers are considered as statistical classifiers. They are able to predict the probabilities of a class membership like the probability of a given tuple which belong to a particular class. The probability is calculated by Eq.(7).

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad ...(7)$$

In Bayesian expression, "X" is "evidence". Let "H" be certain hypothesis like the data tuple "X" that belongs to certain class "C". The probability of the hypothesis "H" holds the "evidence" or observed data tuple "X". P(H|X) is the posterior probability of "H" conditioned on "X" where, $P(h) = \frac{|h|}{N}$ is the estimated h priori probability and they are the patterns number in class "h" and "N" is the total patterns number and it is assumed that all hypotheses are likely equal to P(X|h) which is "X" conditional probability that is conditional on h, and P(X) is "X" prior probability, which is a constant. The hypothesis of maximum posteriori (MAP) used to mean that class h has maximum P(h|X)[13 and 14].

Description Of The Proposed System

The proposed system is offline NB classifier for spam filtering. Figure (1) will present the proposal; which is NB classifier use the IG-based ACS as feature selection, but ACS will be worked somehow inverse of the traditional concept. Stages of the proposal are:

1. Prepare and preprocess (normalization) spambase dataset to be used in training and testing.

2. The proposed algorithm which is called IG-based ACS feature selection will be hybrid with NB classifier. NB classifier is trained and tested on critical subset of features that is extracted by IG-based ACS.

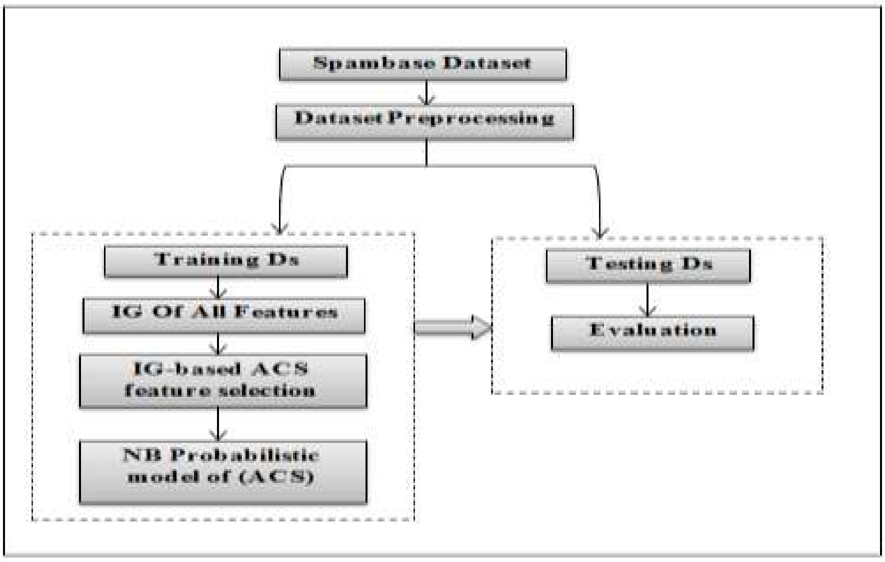

Figure (1) Block diagram of NB classifier with IG-based ACS feature selection

Dataset Description (Spambase Dataset)

The used dataset in the proposed system is a universal and standard spambase dataset consists of 4051 records (e-mails) and (57) features in addition to class type which were taken out of 4601 records, the rest of records are eliminated because they a redundant, spambase dataset is founded for scientific researches of spam filtering [15]. This dataset is prepared to be used with Weka machine learning software. The last column is titular represents the class type, which indicates whether the e-mail is considered a spam "1" or not "0". Most of the features indicate whether a particular word or character was frequently appearing in the e-mail. The run-length features (55-57) measure the length of sequences of consecutive capital letters.

Normalization of Spambase Dataset

The values of spambase dataset features have different ranges. Therefore the normalization process is applied on the values of these features to set them in a uniform range between [0, 1]. Normalization can be done using Eq.(8).

$$attx_{new} = \frac{attx_{old} - minold}{max_{old} - min_{old}} \qquad \dots(8)$$

Where

1. $attx_{old}$: old value of feature.
2. min_{old} : Minimum value that the feature $attx_{old}$ can get.
3. max_{old} : Maximum value that the feature $attx_{old}$ can get.

The normalized spambase dataset will be divided into two datasets these are: 1. Training dataset for training; which consist of (3007) records that represent the e-mails, and (58) columns that represent the features of the e-mails and the class type (ham and spam).

2. Testing dataset for testing the model which is consists of (1044) records that represent the e-mails to be tested and (57) columns present features of the e-mails.

Information Gain (Ig)

In order to improve the accuracy of the classification by calculating the entropy of each element in each feature with each class and the entropy of the entire class, to calculate the information gain of each feature as total sum of the subtraction of the total class entropy from the elements entropy of the features. The information gain will be obtained from the training dataset. It will provide (57) information gain. For more explanation see Figure (2).

1. Calculate the total entropy by the following , Total Entropy=

$$\left[-\frac{Class\ Spam}{The\ frequncy\ of\ Total\ Emails} log_2 \frac{\text{Class Spam}}{The\ frequncy\ of\ \text{Total E−mails}} - \frac{\text{Class Not Spam}}{The\ frequncy\ of\ \text{Total E−mails}} log_2 \frac{class\ Not\ Spam}{The\ frequncy\ of\ Total\ Emails}\right] \quad \text{... (9)}$$

2. Calculate T which is the summation of the frequency of a feature with the two classes spam and not spam.
3. Calculate the entropy of the feature with two classes using Eq.(10),Feature Entropy=$\frac{T}{3007}$[Entropy(feature(class|spam),feature(class|non-spam))] ..(10)

Information gain (feature) = Total Entropy – Feature Entropy ..(11)

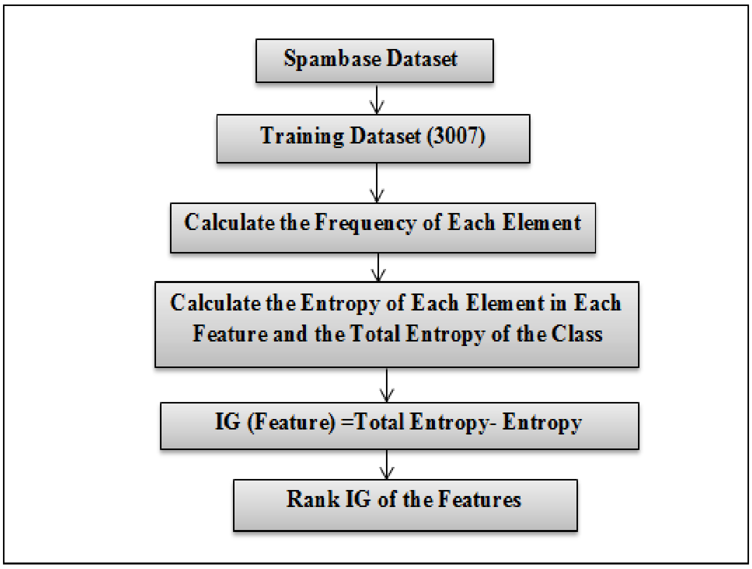

Figure (2) Block diagram of the customized information gain

Ig-Based Acs Feature Selection

The proposal will use ant colony system. ACS uses the obtained (57) IG as a heuristic measure in the guidance of ACS search. The proposal assumes each feature as a node and uses the ant colony system as a feature selection. ACS selects the worst best features that will increase accuracy of the Naive Bayesian classifier, precision and recall in classifying the test dataset e-mail messages.

The proposed IG-based ACS feature selection assumes the following parameters of ACS and for more explanation see Figure (3):

1. The initial pheromone(Ph_0) value is 1/No. of dataset features =1/57=0.01.
2. No. of nodes = No. of features=57, No. of ants = 5, No. of iterations =10.
3. Heuristic measure η_{ij} is the av, av = $(\frac{1}{IG_1} + \frac{1}{IG_2})$ / 2 ..(12)
4. The ants select the first node (feature) randomly.
5. Q is a random value between (0, 1), Q_0 is assumed to be 0.5.
6. If Q ≤ Q_0 the Ant will exploit :

   Transition rule = Select arg$max\ _{U\in N_i^K} Ph_{iu}\ (t)\eta_{iu}^{\beta}$(t) ..(13)

   Else if Q > Q_0 the Ant explores

$$\text{Transition rule} = \rho_{ij}^{k} = max \frac{ph_{ij}(t)\eta_{ij}{}^{\beta}(t)}{\sum_{u\epsilon N_i^K} ph_{iu}(t)\eta_{iu}{}^{\beta}(t)} \quad ..(14)$$

Where β: it is assumed =1, N_i^K (t): a set of nodes to be selected.

7. Local update pheromone rule $Ph_{ij}(t) = (1-\rho)Ph_{ij}(t)+\rho \frac{1}{n+Lnd}$ ..(15)

Where Lnd represents the length of the route between two nodes,ρ represents evaporate value is assumed 0.1

8. Global pheromone is updating at the end of each one iteration the according to Eq.(16):$Ph_{ij}(t+1)= (1-\rho_1)Ph_{ij}(t)+\rho_1 Dt$ ..(16)

Whereρ_1: represents evaporate value is assumed 0.1, Dt: 1/St, St: represents the length of the selected subset.

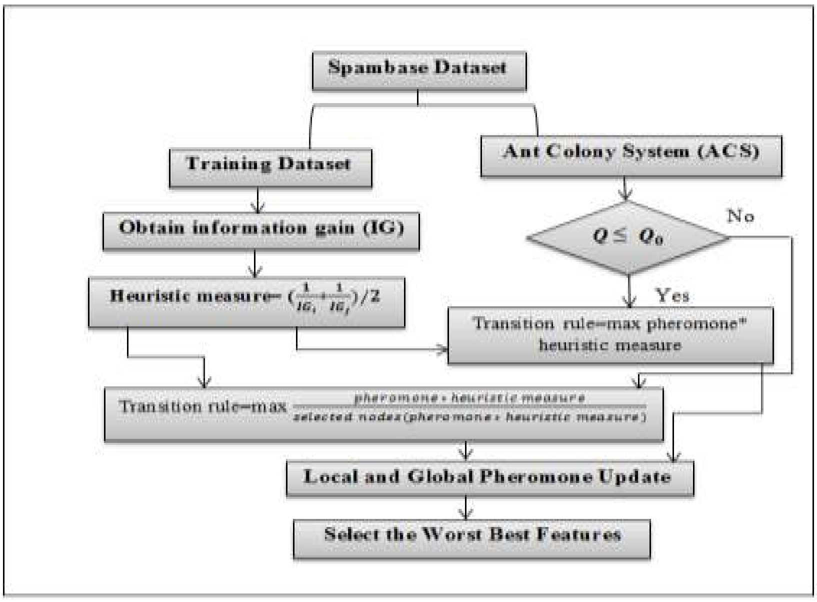

Figure (3) Block diagram of IG- based ACS feature selection

The Naive Bayesian Classifier

The work of NB will be consisted of two stages these are; training stage (the resulted NB probabilities will be used in classifying "test dataset e-mails" whether the incoming e-mail message is spam or not) and the classification stage both of them are based on the message content, see Figure (4).

1. Training stage; NB classifier will be trained using known features (words) that appear in either spam or non-spam as in the following:
   - Calculate the probability of each class, using Eq. $P(C_i) = \frac{|C_i|}{N}$ ..(17)
   - Calculate the conditional probability of each element occurring in each feature to each class using Eq. $P(X|C_i)= \prod_{j=1}^{n} P(E_j|C_i)$ ..(18)
2. Classification stage; NB filter classifies e-mail in the testing dataset using the NB probabilities of training dataset.
   - Apply NB classifier on each record (e-mail) in testing data set.
   - Take the probabilities of the current record elements from training stage.
   - If the probability of an element is not found the classifier take the average probability of the nearest two elements.
   - Calculate the posterior probabilities of the current record (e-mail) using Eq.$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$ … (19)

- If the result of the posterior probabilities of class spam is > the result of posterior probabilities of class non-spam, the record (e-mail) is classified as spam, and vice versa.

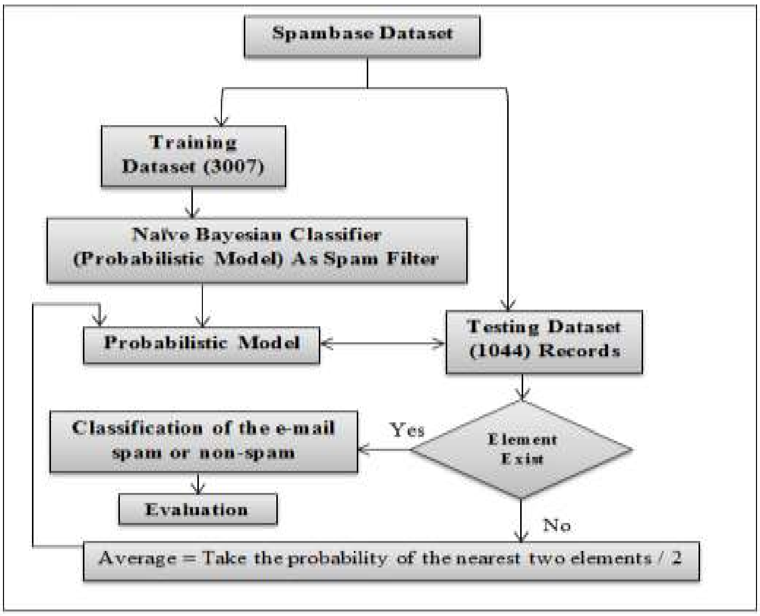

Figure (4) Flowchart explain Naive Bayesian Classifier

Experimental Work And Results

To experiments the proposed models spam filtering system will depend on spambase dataset. The proposal train three classifiers (NB, IG-based NB, and IG-Based ACS NB) on training dataset, then the constructed classifiers are tested on testing dataset. The proposed system has been experimented (i.e., trained and tested) for many times to estimate the accuracy of the classifiers, finally highlight which of them rate higher accuracy results. This section explains results according to standard evaluation measures of classifications.

Three classification models have been experimented, where these models have been trained and tested on the same training and testing datasets. That is to estimate the validation and the accuracy of these constructed models on the same testing dataset. The classification measures are:

1. True Positive (TP): infected e-mail that correctly categorized as spam.
2. False Positive (FP): e-mail that incorrectly categorized as spam.
3. True Negative (TN): e-mail that correctly categorized as e-mail.
4. False Negative (FN): infected e-mail that incorrectly categorized as e-mail.

Table (1), presents the performance of the three classifiers NB, IG-based NB, IG-Based ACS NB with set of features along 3007 e-mails (training dataset) as in Figure (5); In the Table (1) these samples obtained during classification process have the highest accuracy result. These samples are taken from Naive Bayesian model, Information Gain based Naive Bayesian model and Information Gain Based Ant Colony System Naive Bayesian model. For instance in the first record of this Table is Naive Bayesian classifier which have all features and classification results are TN= 780, TP= 157, FN= 23, FP= 84.

Table (1): Classification results of NB, IG-based NB, and IG-Based ACS NB (testing of 1044 e-mails)

Classifier	Selected Features	TN	TP	FN	FP	Accuracy	Precision	Recall
NB	57	780	157	23	84	89.75%	65.15%	87.22%
IG-based NB	10	798	165	15	66	92.24%	71.43%	91.67%
	15	799	162	18	65	92.04%	71.37%	90.00%
	17	803	161	19	61	92.33%	72.52%	89.44%
	20	771	161	19	93	89.27%	63.39%	89.44%
	30	786	156	24	78	90.22%	66.67%	86.67%
	40	785	155	25	79	90.03%	66.34%	86.11%
	45	780	156	24	84	89.65%	65.00%	86.67%
	47	777	156	24	87	89.36%	60.20%	86.67%
	50	776	156	24	88	89.27%	63.93%	86.67%
IG-Based ACS NB	47	820	153	27	44	93.19%	77.66%	85%
	37	810	160	20	54	92.91%	74.76%	88.88%
	29	826	161	19	38	94.54%	80.90%	89.44%
	27	828	164	16	36	95.01%	82%	91.11%
	24	834	161	19	30	95.30%	84.29%	89.44%

Figure (5) contains representation of the highest and most accurate (TN, TP, FN, FP) results from each one of the proposed classification models as shown in Table (1) of the (NB, IG-based NB and IG-Based ACS NB).

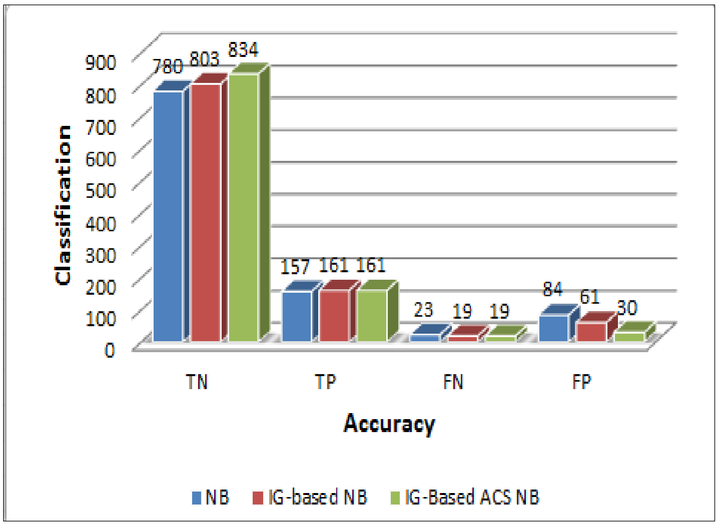

Figure (5): The highest accuracy results of testing dataset

Figure (6) below represents the highest and most accurate (Accuracy, Precision and Recall) results of each one of the proposed models during the classification process as shown in Table(1) of the (NB, IG-based NB and IG-Based ACS NB) .

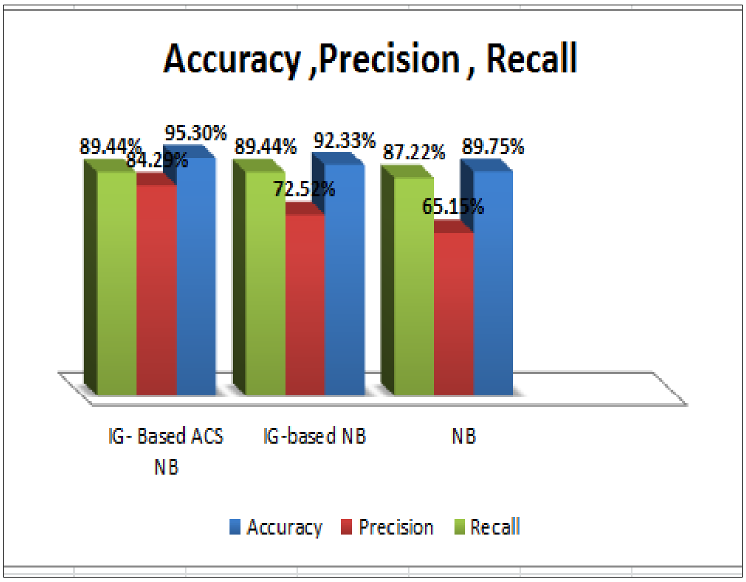

Figure (6): Represents accuracy, precision and recall

Table (2) represents samples of the highest accuracy results by using the ant colony system first transition rule to select the features to be turned off to increase accuracy of the classification process. For instance, in Table (2), the first record will be the IG-Based ACS NB classifier takes the highest accuracy result obtained during 8th iteration, which turns off 28 features that improves accuracy result from 89.75% to 94.54%.

Table (2) high accuracy by using IG-Based ACS NB Classifier first transition rule

Classifier	Iteration	No. of selected node to be turned off	Selected features to be turned off	Start node	Accuracy
IG-Based ACS NB	8	28	55 , 54 , 7 , 9 , 5 , 1 , 4 , 0 , 31 , 38 , 50 , 14 , 35 , 44 , 41 , 30 , 11 , 12 , 34 , 8 , 29 , 49 , 27 , 17 , 28 , 10 , 19 , 16	55	94.54%
	8	30	55 , 9 , 5 , 1 , 4 , 22 , 2 , 0 , 31 , 38 , 50 , 14 , 35 , 44 , 41 , 30 , 11 , 12 , 34 , 8 , 29 , 49 , 27 , 17 , 28 , 10 , 19 , 16 , 36 , 54	55	95.01%
	9	33	55 , 1 , 4 , 22 , 2 , 56 , 25 , 18 , 0 , 31 , 38 , 50 , 14 , 35 , 44 , 41 , 30 , 11 , 12 , 34 , 8 , 29 , 49 , 27 , 17 , 28 , 10 , 19 , 16 , 36 , 54 , 7 , 9	55	95.30%

Figure (7) shows the highest three accuracy results obtained by applying first transition rule of IG-Based ACS NB classification model. Where Figure (8) shows the different number of selected features and their accuracy.

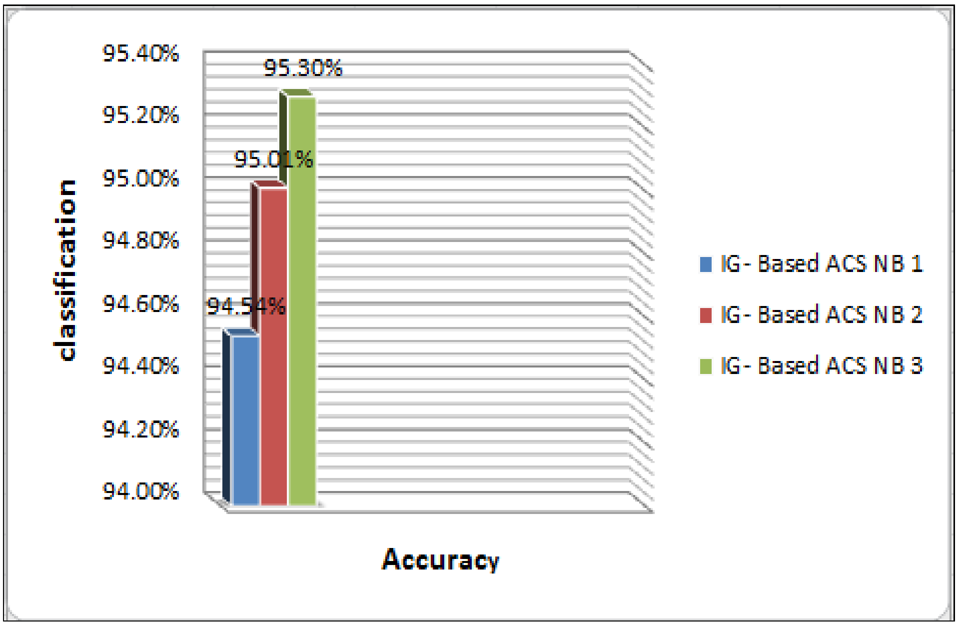

Figure (7): the highest accuracy of IG-Based ACS NB classifier

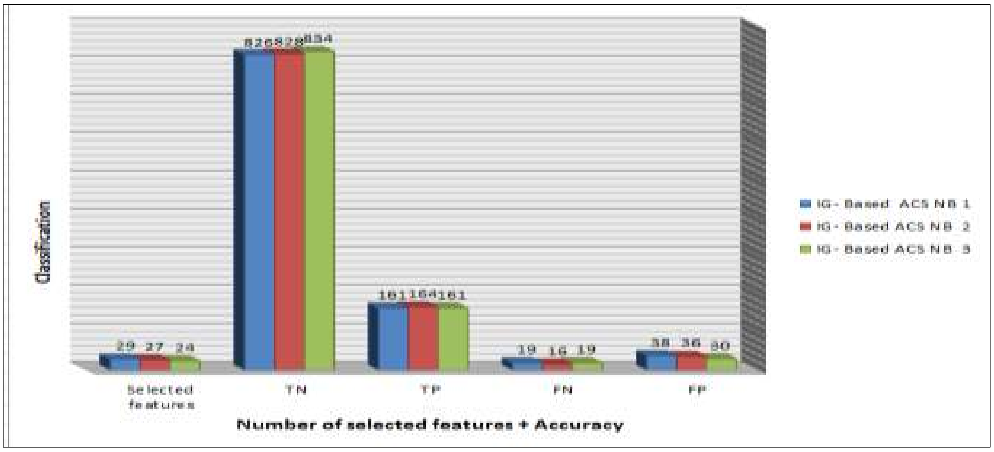

Figure (8): represent number of selected features with accuracy of the IG-Based ACS NB classifier

Table (3) below represents the samples of the highest accuracy results by using the ant colony system second transition rule to select the features to be turned off to increase accuracy of the classification process. For instance, in Table (3), the first record will be the IG-Based ACS NB classifier takes the highest accuracy result obtained during 2nd iteration, which is turn off10 features that improves accuracy result from 89.75% to 92.24%.

Table (3) high accuracy by using IG-Based ACS NB Classifier second transition rule

Classifier	Iteration	No. of selected node to be turned off	Selected features to be turned off	Start node	Accuracy
IG-Based ACS NB	2	10	54 , 43 , 45 , 32 , 33 , 42 , 53 , 3 , 48 , 13	54	92.24%
	5	20	54 , 52 , 20 , 24 , 15 , 45 , 32 , 33 , 42 , 53 , 3 , 48 , 13 , 47 , 43 , 40 , 46 , 39 , 37 , 21	54	89.75%
	9	30	54 , 4 , 45 , 32 , 33 , 42 , 53 , 3 , 48 , 13 , 47 , 43 , 40 , 46 , 39 , 37 , 21 , 51 , 52 , 20 , 24 , 15 , 23 , 6 , 55 , 26 , 18 , 25 , 56 , 2	54	85.82%

CONCLUSIONS

1. Through the use of the proposed system Naive Bayesian classifier accuracy achieved (89.75%) displayed in Table (1), a good results are obtained for the classification of spambase dataset e-mail messages by the likelihood an element occurring in either category more explanation in Figure (5) and Figure (6), also provide higher speed computational result when is applying to large database such spambase (4051) records as shown in flowchart (4).

2.By taking number of the highest information gain calculated of each feature from the training dataset then applying Naive Bayesian classifier over the selected features to classify the testing dataset e-mail messages as display in Table (1) when take the 10 highest information gain make accuracy results increased from (89.75%) to (92.24%) even higher results when taking the 17 highest information gain of features (92.33%) more explanation in Figure (5) and (6), it provide higher accuracy results than applying Naive Bayesian classifier only.

3. As it is seen in the experimental work that displayed in previous sections, ant colony system depends on information gain as a heuristic measure which is static value in Eq.(12) and pheromone updating is a dynamic value in Eq.(15) and (16) that update on the selected nodes to be turned off, the ant colony system work in reverse of traditional method as shown in Block diagram (3).

4.In IG-Based ACS NB classifier when ants use the first transition rule in Eq.(13) to decide which is the next node, then applying Naive Bayesian, higher accuracy result is given compared with using a second transition rule in Eq.(14) as seen in Table (2). When turn off 28 selected features that are selected by the first transition rule in Eq.(13) accuracy result is improved from using only Naive Bayesian (89.75%) to (94.54%) and representation in Figure (7).

5.When the ants in IG-Based ACS NB classifier use the second transition rule in Eq.(14) to decide which node to go next then applying Naive Bayesian, the accuracy result decreases compared with using the first transition rule in Eq.(13) as shown in Table (3). When turn off 30 selected features by using the second transition rule in Eq.(14) leads to decrease accuracy result from using Naive Bayesian (89.75%) to (85.82%), these results is not so promising.

REFRENCES

[1]Chakraborty N. and Patel A.,"E-mail Spam Filter using Bayesian Neural Networks", International Journal of Advanced Computer Research ,Volume 2, Number 1, March 2012.

[2]Al-Jarrah O., Khater I., and Al-Duwairi B., " Identifying Potentially Useful E-mail Header Features for E-mail Spam Filtering", The Sixth International Conference on Digital Society, 141Copyright (c) IARIA, 2012.

[3]Androutsopoulos I., Koutsias J., Chandrinos K.V., Paliouras G.and Spyropoulos C.D., "An Evaluation of Naive Bayesian Anti-Spam Filtering", 11th European Conference on Machine Learning, pp. 9-17, 2000.

[4]El-Halees A., "Filtering Spam E-Mail from Mixed Arabic and English Messages: A Comparison of Machine Learning Techniques" The International Arab Journal of Information Technology, Vol. 6, No. 1, January 2009.

[5]Paswan M.K., Bala P.S. andAghill G.,"Spam Filtering: Comparative Analysis of Filtering Techniques",IEEE- International Conference On Advances In Engineering, Science And Management , March 30, 31, 2012.

[6]Abd-AlsabourN. and RandallM., "Feature selection for classification using an ant colony system", Sixth IEEE international conference on e-Science, Australia, December,2010.

[7]JensenR., "Performing Feature Selection with ACO", Department of Computer Science, The University of Wales, Aberystwyth, UK rkj@aber.ac.uk.

[8]Guzella T.S. and Caminhas W.M.,"A Review Of Machine Learning Approaches To Spam Filtering ",Expert Systems with Applications 36 (2009) 10206–10222.

[9]Yinga K.S.andLiaoa C.J., "An Ant Colony System For Permutation Flow-Shop Sequencing",Computers & Operations Research 31 (2004) 791–801 © Elsevier.

[10]Hasheem S.H., Saleh D.M., "Proposal to Enhance Network Routing by Using Ant Colony Optimization", Eng. & Tech. Journal .Vol31, Part (B), No. 6 , 2013.

[11]Kumar V. and Minz S.,"Feature Selection: A Literature Review "Smart Computing Review, vol.4, no.3 , June 2014.

[12]NovakovicJ., "Using Information Gain Attribute Evaluation To Classify Sonar Targets",17th Telecommunications forum TELFOR, Belgrade, November 24-26, 2009.

[13]Han J., Kamber M., and Pei J., "Data Mining Concepts And Techniques", 2012 by Elsevier Inc.

[14]Hasheem S.H.,Abdulmunem I.A.,"A Proposal to Detect Computer Worms (Malicious Codes) Using Data Mining Classification Algorithms", Eng. & Tech.Journal .Vol31,No.2, 2013.

[14]UCI Machine Learning Repository, archive.ics.uci.edu /ml/datasets/ Spambase.