International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-6, March 2020

Measuring Success of Heterogeneous Ensemble Filter Feature Selection Models

Noureldien A. Noureldien, Einas A. Mohammed

Abstract: One problem in utilizing ensemble feature selection models is machine learning is the fact that there is no guarantee that an ensemble model will improve machine learning classification performance. This implies that different ensemble models have different success probability, i.e. have different probability in improving the performance of machine learning. This paper introduces the concept of success probability for heterogeneous ensemble models and stated the definitions, notations, and algorithms necessary to the mathematical formulation and computation of the success probability. To show how the theory applied, we create an ensemble filter feature selection model that uses four filter feature selection algorithms (Correlation, Gain Ratio, Info Gain, and One R) as base filters and the Max as a combination method. The experimental results showed that the success probability of the developed ensemble filter model using a set of 9 machine learning algorithms is found to be 0.58.

Keywords: Filter Feature Selection, Ensemble Feature Selection Model, Combination Method, Success Probability, Classification Accuracy, Measuring Success Probability.

I. INTRODUCTION

Ensemble feature selection evolved from the field of ensemble learning, where multiple classifiers are combined to yield a more stable, and better-performing ensemble classifier [1].Similar to the construction of ensemble models for supervised learning, there are two essential steps in creating an ensemble feature selection model. The first steps the selection of a set of feature selection methods, called base feature selection methods, and generate their output. The second step is the aggregation or combination of the outputs results of the base methods into a single output.

This paper deals only with filter feature selection methods as base filters, and accordingly, with ensemble filter feature selection models. Filter feature selection methods output ranked list of dataset features. Thus, the ensemble model combines the multiple feature ranking lists obtained from base filters and generates a single ranking list.

Ensemble feature selection is used to improve the robustness of feature selection techniques, and therefore, it reduces the risk of choosing an unstable subset. Furthermore, an ensemble feature selection might give a better estimation to the optimum subset or ranking of features [1]. Hence an ensemble filter feature selection model is expected to improve the classification performance of classifiers and of any data mining method that based on feature selection [2].

There are several approaches or ways to design an ensemble [3]. The most common two particular ways are homogeneous and heterogeneous [4]: In the homogeneous ensemble, N rankers/models are generated using the same feature selection method but with different training data sets. In the heterogeneous ensemble, N rankers/models are created using different feature selection methods, but with the same training data (Figure 1). This approach takes account of the strengths and weaknesses of the individual methods. The several different methods are trained using the same training data, and the output is then combined using a combination

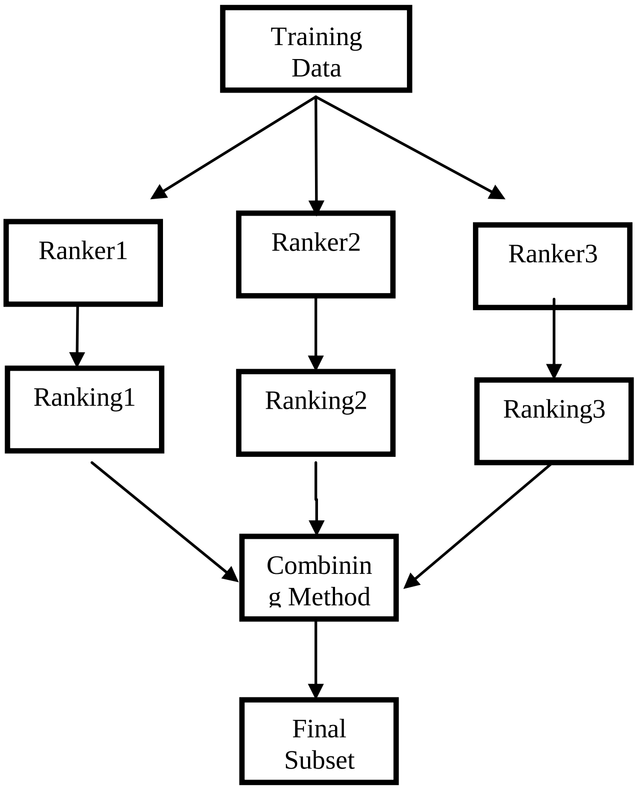

Fig. 1. Heterogeneous Ensemble Approach

The combination methods, also known as aggregation methods, are used in building ensemble models to produce a unique final output. Several different combination methods are available such as Intersection, Union, Max …etc. [5]. In this paper, we design our experimental ensemble model using heterogeneous approach and Max as a combination method. Since the output of each of the base filter methods is a ranked list that contains all dataset features, it is necessary to set a threshold that specifies the number of top-ranked elements that must be considered as a subset. Most works in the

Revised Manuscript Received on February 01, 2020.

Noureldien A. Noureldien, Faculty, Computer science and information technology, University of Science and Technology, Omdurman, Sudan.

Einas A. Mohammed, Master degree in Computer science at University of Science and Technology, Omdurman, Sudan.

literature use thresholds that are percentages of features [6, 7]. Several attempts have been made to derive the value of the threshold automatically [8, 9].

This will ease our selection of an ensemble feature selection model with a high probability that it will improve the classification and prediction of a classifier. Hereafter, we will refer to the ensemble model probability to improve classification accuracy as ensemble model success probability.

This paper is about how to measure an ensemble model success probability. The paper focuses only on heterogeneous ensembles of filter feature selection methods. The paper stated the necessary definitions, notations and the mathematical formulation, and computation algorithms. The developed theory is applied by building an ensemble filter feature selection model using four filter feature selection algorithms (Correlation, Gain Ratio, Info Gain, and One R) as base filters and the Max as a combination method.

The remainder of the paper is organized as follows: Section 2 provides a literature review to related work. Section3 presents the theoretical definitions, notations, and mathematical formulation and computational algorithms of ensemble success probability. In Section 4, we present the design of experiments and results and discussion is presented in Section 5. Finally, in Section 6, we draw the conclusion and recommendations for future work.

II. RELATED WORK

The concept of ensemble feature selection was introduced in 2008 [1]. Since then, many research work on ensemble feature selection is published. We can categorize most of the published work into four classes; (i) Work that focus on proofing that using ensemble filter feature selection models can improve the performance of machine learning rather than using single feature selection methods [1][10][11][12][13]. (ii) Research that compares the performance of several ensemble filter feature selection models to determine the best model [14][15]. (iii) Work that focuses on techniques of developing ensemble feature selection models [16][17][18], and (iv) Research work that focuses on providing reviews and surveys for existing approaches and applications of ensemble filter feature selection models [19].

In proofing the efficiency of ensemble models, researchers in [1] evaluate the performance of ensemble model on classification performance, using the Symmetrical Uncertainty, RELIEF, Random Forests, and SVM as base filters. The classifiers Random Forests, linear SVM, and the distance based k-nearest neighbor algorithm (KNN) were used as evaluators. From the results of ensemble feature selection model to a classifier it is observed that in most cases performance is increased. Similarly, an ensemble model of filters is proposed and its performance is compared with the base filter method in terms of classification prediction for each of the pair of filter and classifier [10]. An ensemble-based multi-filter feature selection model that combines the output of filter methods; information gain, gain ratio, chi-squared and Relief F is developed [11]., and the performance evaluation with NSL-KDD dataset demonstrated that the ensemble model, with 13 features, achieves better performance than individual filter feature selection methods using J48 classifier.

The model is [12] is validated using the classifiers, decision trees, random forests, KNN, and SVM on fourteen UCI, five gene expression, and two network datasets, and the experimental results show that the ensemble model achieves the best accuracy performance. The ensemble model in [13] combines base filters using SVM Rank. The model is tested using SVM as classifier. The obtained results on five UCI datasets showed that the use of the proposed ensemble gives better or comparable performance than the feature selection methods individually.

For research work that compares the performance of multiple ensemble filter feature selection models to determine the best model, the authors in [14] compare the performance of 17 ensembles of feature ranking techniques (rankers) that use a different number of base filters selected out of six commonly-used feature ranking methods. They utilize 16 real-world software data sets of different sizes and built 13,600 classification models. Their experimental results indicate that ensembles of very few rankers are more effective than ensembles of many or all rankers.

The performance of two ensemble feature selection ensemble models is compared in [15]. Two different ensemble techniques, Bagging and Boosting, are used to build the ensemble models. Begging ensemble model and Boosting ensemble model. These models were trained and tested respectively with the two collected subset of NSL-KDD dataset, 25-feature, and 35-feature datasets. The training and testing phase were conducted on two separate datasets (KDD Train+ and KDD Test+). The experimental results showed the best performance was produced by the bagging model that used J48 as the base classifier and worked on a 35-feature subset.

Example of research work that focuses on developing ensemble feature selection models includes; an ensemble feature selection approach based on feature selectors' reliability assessment [16]. The approach aims at providing a unique and stable, high predictive accuracy feature selection. The comparison of the proposed approach to several existing techniques and to individual feature selection algorithms is performed. The results obtained show that their approach often improves classification performance and feature selection stability for high dimensional data sets.

An ensemble approach for feature selection that aggregates the output of several different feature selection methods so that a more robust and efficient feature subset can be obtained is proposed [17]. Authors develop a genetic algorithm for aggregation. Experimental evaluations indicated that their proposed ensemble model is an efficient method, and it outperforms individual filter-based feature selection methods on sentiment classification. Similarly, an empirical exploration of the effectiveness of homogenous ensemble approach is studied [18]. To construct the model, a single feature selection algorithm is applied to several diversified datasets derived from the original set of records. The results of the experiments which are conducted on high-dimensional data from different domains show that the ensemble setting can lead to a significant improvement in stability without any degradation of the predictive performance.

Example of review and survey study is [19], where authors provide a review that covers many issues related to ensemble learning based feature selection, which includes the development of ensemble models, the stability measurement, etc.

This paper is different from all previous work on ensemble feature selection. This paper defines the concept of success probability of ensemble feature selection and shows how the success probability can be measured.

III. ENSEMBLE SUCCESS PROBABILITY: THEORETICAL FORMULATION AND COMPUTATION ALGORITHMS

Ensemble model success probability is the probability that an ensemble mode can improve the classification accuracy of a randomly selected classifier in comparison to the accuracy achieved by the base ensemble filters.

How to calculate the success probability of an ensemble filter feature selection model? To answer this question, we state the following definitions and notations.

Definition 1

If E is an ensemble filter model that uses the set of base filter selection methods F = {F1, F2… Fn}, and the combination method is &then we denote E as: E = f {F, &}. We denote the subset of features produced by the ensemble E as {E}, and the ranked subsets produced by the filter selection method Fi as {Fi}

Definition 2

Let C be a randomly selected classifier, then we denote the classification accuracy of C on the dataset D, using S as the feature subset selected by a filter feature selection method as: CAcc {DS}.

Definition 3

We say that E is successful when for a randomly selected classifier C, the classification accuracy of C using {E} is highest than the accuracy achieved by C using each of the ensemble base filters F1, F2, …Fn. i.e. E is successful when: CAcc{D{E}} >CAcc{D{Fi}}, for alli=1, 2,…,n.

Similarly, we say that E is partially successful when the classification accuracy of C using {E} share the highest accuracy with one or more of the ensemble base filters. i.e. E is partially successful when:

CAcc {D{E}} = CAcc{D{Fi}}, for some i,i=1,2,…,n.

And E is failed when the classification accuracy of C using {E} is not the highest accuracy compared to the accuracy of the ensemble base filters, i.e. E is failed when:

CAcc {D{E}} < CAcc{D{Fi}}, for some i; i=1,2,…,n.

Definition 4

Based on whether E is successful, partially successful, or failed when the classifier C uses it, we assign a weight to E, denoted as W (Ec) using the following equation:

$$W(\mathrm{Ec}) = \begin{cases} 1, if\ E\ is\ successful \\ \frac{1}{k}, if\ E\ is\ Part.\ Succ \\ 0, if\ E\ is\ failed \end{cases}$$

Definition 5

Let C = {C1, C2… Cm}, denote set of classifiers, and
E =f {F, &}, where F= {F1, F2… Fn} is a set of n base filters. Then the probability of success of E is defined as:

$$\mathrm{P(E)} = \sum_{i=1}^{m} W(Eci)/m$$

……………………………………..Equation (2)

<u>Algorithm1. Pseudo-code for constructing an ensemble filter feature selection model</u>

Input: D = {attr1, attr2, attr3… attrm}, a dataset D with m features (attributes)
Input: F = {F1, F2… Fn}, the set of n ranker/filter methods
Input: T: A threshold of the number of top ranked features to be selected as a subset
Output: {E}, a subset of features produced by the ensemble model E
1 for each Fi from i =1 to ndo
2 Obtain {D {Fi}}, where D {{Fi}} is a ranked list of T top features of D ranked by Fi
3 end
4 Obtain {E}, where {E} is obtained by combining the T top features in each D {{Fi}} using a combination method.

<u>Algorithm 2. Pseudo-code of calculating the success probability of an ensemble model</u>

Input: S= {{E}, {F_1}, {F_2}… {F_n}}, a set of subsets obtained by the ensemble model and the N base filter methods. Each of the subsets {F_1}, {F_2}… {F_n} contains features.
Input: D = {$attr_1$, $attr_2$, $attr_3$… $attr_m$}, a dataset D with m features (attributes)
Input: F = {F_1, F_2… F_n}, the set of n ranker/filter methods
Input: C = {C_1, C_2… C_m}, a set of m classifiers
Output: P (E): the probability of success of E
1) set W(E) = 0, initially the weight of E is 0
2) for each C_i from i =1 to m do
3) for each from j=1 1 to n do
4) Obtain C_{iAcc} {D {Fj}}; the classification accuracy for Ci using each of the bases filters subsets.
5) end
6) Obtain C_{iAcc} {D {E}}; the classification accuracy of C using the ensemble model subset.
7) calculate W(E_{ci}); W(E_{ci}) is calculated for each classifier using equation 1
8) W (E) = W (E) + W (E_{ci}); accumulate the ensemble gained weights as stated in equation 2.
9) end
10) P (E) = W (E)/m, calculate the success probability as stated in equation 2.
11) Output P(E)

IV. EXPERIMENTAL DESIGN

To implement algorithm 1, we select randomly four base filter feature selection methods, namely Correlation Attribute Eval, Gain Ratio Attribute Eval, Info Gain Attribute Eval and One R Attribute Eval, and we select the NSL-KDD 20 % as a dataset. The NSL-KDD data set is suggested to solve some of the inherent problems of the KDDCUP'99 data set. KDDCUP'99 is the most widely used data set for anomaly detection [9].

To construct the ensemble model, the four base feature selection methods are applied individually to the NLS-KDD 20% dataset. The output is an ordered set of the dataset features (41 features). From the output of the ranked features, the subset of each method is obtained using a threshold value. Subsets will be combined together using the Max method to generate the ensemble subset.

To apply the Max combination method, we need to determine how much features to pick up of the top-ranked features from each ranked set, i.e. to determine the value of T, the threshold of the number of top ranked features to be selected as a subset.

For our experiments, we used three different threshold values to delimit data dimensionality.

- 25%. This threshold selects the top 25% of the ranked features list (top 10 features)
- 38%. This threshold selects the top 38% of the ranked features list (top 15 features)
- 50%. This threshold selects the top 50% of the ranked features list (top 20 features)

Table (I) shows the ensemble model output generated using the 10, 15, and 20 top ranked features.

To calculate the success probability of the developed ensemble model, we apply algorithm 2. To apply algorithm 2, we make the following selections:

1) 1A set of 9 classifiers from three different classes, shown in table (II) below.
2) The top 10, 15, and 20 ranked features as threshold.

The classification accuracy is calculated for each classifier using the base filter feature selection methods subsets and the ensemble subset with the three thresholds. Tables (III) to (V)

Shows the accuracy of each classifier per base filter selection methods and the ensemble model using 10, 15, and 20 top ranked elements.

Table- I: Ensemble Model Output Using 10, 15, 20 Top Ranked Features

Feature Selection Method	T=10	T=15	T=20
Correlation Attribute Eval	29,33,34,12,39,4,38,25,26,23	29,33,34,12,39,4,38,25,26,23,32,3,41,27,40	29,33,34,12,39,4,38,25,26,23,32,3,41,27,40,28,35,30,31,8
Gain Ratio Attribute Eval	12,26,4,25,39,6,30,38,5,29	12,26,4,25,39,6,30,38,5,29,3,37,34,33,8	12,26,4,25,39,6,30,38,5,29,3,37,34,33,8,35,23,31,41,32
Info Gain Attribute Eval	5,3,6,4,30,29,33,34,35,38	5,3,6,4,30,29,33,34,35,38,12,39,25,23,26	5,3,6,4,30,29,33,34,35,38,12,39,25,23,26,37,32,36,31,24
One R Attribute Eval	5,3,6,4,29,30,34,33,35,12	5,3,6,4,29,30,34,33,35,12,23,25,38,39,26	5,3,6,4,29,30,34,33,35,12,23,25,38,39,26,32,36,37,24,31
Ensemble output	3,5,6,12,23,25,26,30,33,34,35,38,39	5,6,8,23,27,30,32,35,37,40,41	5,6,8,24,27,28,32,36,37,40,41

Table- II: The Selected Classifiers

Class Name	Selected Algorithms
Bayes	Naïve Bayes, Bayesnet, Naïve Bayes Multi NminalText
Rules	Decision Table, J Rip, One R
Trees	Decision Stump, J48, Random Tree

Table- III: Classification Accuracy Using 10 Top Ranked Features as a Threshold

Machine Learning Algorithms	Performance based on correlation	Performance based on Gain Ratio	Performance based on Info Gain	Performance based on One R	Performance based on ensemble
Bayes .Bayes Net	90.6 %	92.5 %	93.7 %	95.2 %	94.8 %
Bayes .Naive Bayes	88.1 %	86.1 %	87.7 %	88.2 %	86.3 %
Bayes. Naïve Bayes Multi Nominal Text	53.3 %	53.3 %	53.3 %	53.3 %	53.3 %
Rules .Decision Table	95.3 %	98.8 %	98.9 %	98.9 %	99.4 %
Rules. J Rip	96.4 %	98.5 %	99.6 %	99.5 %	99.8 %
Rules. One R	88.1 %	96.2 %	96.2 %	96.2 %	96.4 %
Tree .Decision Stump	87.7 %	92.2 %	92.2 %	92.2 %	92.2 %
Tree.J48	96.8 %	99.1 %	99.5 %	99.5%	99.8 %
Tree .Random Tree	96.3 %	99.2 %	99.5 %	99.5 %	99.9%

Table- IV: Classification Accuracy Using 15 Top Ranked Features as a Threshold

Machine Learning Algorithms	Performance based on correlation	Performance based on Gain Ratio	Performance based on Info Gain	Performance based on One R	Performance based on ensemble
Bayes .Bayes Net	91.9 %	96.4 %	93.5 %	93.5 %	97.4 %
Bayes .Naive Bayes	89.3 %	89.4 %	89.1 %	89.1 %	70.9 %
Bayes. Naïve Bayes Multi Nominal Text	53.3 %	53.3 %	53.3 %	53.3 %	53.3 %
Rules .Decision Table	97 %	99 %	99.1 %	99.1 %	99.1 %
Rules .J Rip	99 %	99.5 %	99.5 %	99.5 %	99.6 %
Rules. One R	91.4 %	96.2 %	96.2 %	96.2 %	96.4 %
Tree .Decision Stump	87.7 %	92.2 %	90 %	92.2 %	92.2 %
Tree.J48	99.1 %	99.5 %	99.1 %	99.6 %	99.8 %
Tree .Random Tree	98.9 %	99.5 %	99 %	99.4 %	99.9 %

Table- V: Classification Accuracy Using 20 Top Ranked Features as a Threshold

Machine Learning Algorithms	Performance based on correlation	Performance based on Gain Ratio	Performance based on Info Gain	Performance based on One R	Performance based on ensemble
Bayes .Bayes Net	92 %	94.4 %	95.7 %	95.7 %	96.9 %
Bayes .Naive Bayes	88.9 %	91 %	90.4 %	90.4 %	89.5 %
Bayes. Naïve Bayes Multi Nominal Text	53.4 %	53.4 %	53.4 %	53.4 %	53.4 %
Rules .Decision Table	97.2 %	99 %	99 %	99 %	99.1 %
Rules .J Rip	99 %	99.5 %	99.5 %	99.6 %	99.4 %
Rules .One R	91.4 %	96.3 %	96.3 %	96.3 %	96.4 %
Tree .Decision Stump	87.7 %	92.2 %	92.2 %	92.2 %	92.2%
Tree.J48	99 %	99.6 %	99.6 %	99.4 %	99.6 %
Tree .Random Tree	98.9 %	99.5 %	99.9 %	99.5 %	99.9 %

V. RESULT AND DISCUSSION

Based on whether the ensemble model (E) is successful, partially successful or failed with each classifier, table (6) shows the success probability for the base filters and the ensemble model calculated with thresholds: T=10,15and 20, and the average success probability is shown in the last column.

Table- VI: Ensemble Model Success Probability

Filter Feature Method	Success Probability T=10	Success Probability T=15	Success Probability T=20	Success Probability T=10,15,20
Correlation	2%	2%	2%	2%
Gain Ratio	5%	17%	20%	14%
Info Gain	5%	6%	14%	8.3%
One R	27%	10%	16%	17.7%
Ensemble Model	61%	65%	48%	58%

Table (VI) shows that the ensemble filters achieved the highest success probability. This makes the ensemble model achieves a better accuracy with approximately 58% of the total performance comparisons, that is to say, the probability that the ensemble model can improve the accuracy of a random machine learning algorithm is 0.58.
Since this probability is p>0.5, this implies that using the developed ensemble filter models will improve the performance of machine learning algorithms, rather than using individual filter feature selection algorithms.

The ensemble model achieves the highest accuracy with 61%, 65%, and 48% of classifiers using top 10, 15, and 20 top ranked features respectively. This makes for this particular data set, NLS-KDD 20% data set, using the top 15 ranked features is the best choice to build the ensemble model.

VI. CONCLUSION AND RECOMMENDATIONS FOR FUTURE WORK

In the era of big data with multi-dimensional, the feature selection process becomes an important preprocess for machine learning. Having an optimal subset of features that attain robust stability and improve learning objectives is a primary research aim. Ensemble feature selection is evolved to satisfy this need. Selecting a suitable ensemble model for specific machine learning algorithm with high probability that it will improve learning objectives (classification/prediction) is an obvious need.

This paper stated definitions and algorithms for measuring the success probability of an ensemble model and it shows experimentally how this can be applied.

Our future work will focus on improving the stated theory, determining the effect of learning technique on the ensemble model performance, comparing the ensemble filter models that use different combination methods, and measuring the success probability of ensemble models that share the same base filters but having different combination methods. Also, it is crucial to check the efficiency of an ensemble model using different datasets and to investigate whether the ensemble of ensemble feature selection models worth or not.

REFERENCES

1. Yvan Saeys, Thomas Abeel, and Yves Van de Peer, Robust Feature Selection Using Ensemble Feature Selection Techniques, ECML PKDD 2008, Part II, LNAI 5212, pp. 313–325, 2008
2. S. H. Vege, Ensemble of Feature Selection Techniques for High Dimensional Data, 2012.
3. M. Bramer, Data for data mining, in: Principles of Data Mining, Springer, 2013, pp. 9–19.
4. B. Seijo-Pardo, I. Porto-D´ıaz, V. Bol´on-Canedo, A. Alonso-Betanzos, Ensemble Feature Selection: Homogeneous and Heterogeneous Approaches, Knowledge-based Systems, 2017.
5. A. Tsymbal, M. Pechenizkiy, and P. Cunningham, Diversity in Ensemble Feature Selection, Technical Report TCD-CS-2003-44, 2003.
6. V. Bol´on-Canedo, N. S´anchez-Maro˜no, A. Alonso-Betanzos, Recent Advances and Emerging Challenges of Feature Selection in the Context of Big Data, Knowledge-Based Systems 86 2015, pp. 33–45.
7. V. Bol´on-Canedo, N. S´anchez-Maro˜no, A. Alonso-Betanzos, A Review of Feature Selection Methods on Synthetic Data, Knowledge and Information Systems 34 (3) (2013) 483–519.
8. T. M. Khoshgoftaar, M. Golawala, J. Van Hulse, An Empirical Study of Learning from Imbalanced Data Using Random Forest, in the 19th IEEE International Conference on Tools with Artificial Intelligence, ICTAI,Vol. 2, IEEE, 2007, pp. 310–317.
9. M. Mej´ıa-Lavalle, E. Sucar, G. Arroyo, Feature Selection with A Perceptron Neural Net, in: Proceedings of the international workshop on feature selection for data mining, 2006, pp. 131–135.
10. V. Bolo´n-Canedo,N.Sa´nchez-Maron˜o and A. Alonso-Betanzos, An Ensemble of Filters and Classifiers for Microarray Data Classification,Laboratory for Research and Development in Artificial Intelligence (LIDIA), Vol. 45, PP. 531-539, 2011.
11. OpeyemiOsanaiye,H.Cai, Kim-KwangR.Choo,AliDehghantanha,ZhengXuand M.Dlodlo, Ensemble-based Multi-filter Features Selection Method for DDoS Detection in Cloud Computing, EURASIP Journal on Wireless Communications and Networking, 2016.
12. Nazrul Hoque, Mihir Singh, Dhruba K., and EFS-MI: An Ensemble Feature Selection Method for Classification, Complex and Intelligent Systems, v 4, issue 2, pp 105-118, June 201.
13. Borja Seijo-Pardo(B), Ver´onica Bol´on-Canedo, Iago Porto-D´ıaz, and Amparo Alonso-Betanzos, Ensemble Feature Selection for Rankings of Features, Springer International Publishing Switzerland, IWANN 2015, Part II, LNCS 9095, pp. 29–42, 2015.
14. Huanjing Wang ; Taghi M. Khoshgoftaar ; Amri Napolitano, A Comparative Study of Ensemble Feature Selection Techniques for Software Defect Prediction, Proceedings of the Ninth International Conference on Machine Learning and Applications, 2010, Washington, DC, USA.
15. Ngoc Tu Pham, Ernest Foo, Suriadi Suriadi, Improving Performance of Intrusion Detection System Using Ensemble Methods and Feature Selection, In Proceedings of the Australasian Computer Science Week Multi conference, Bris band, Queensland, Australia — January 29 - February 02, 2018
16. Afef Ben Brahim & Mohamed Limam, Ensemble Feature Selection for High Dimensional Data: A new Method and a Comparative Study," Advances in Data Analysis and Classification, Springer; German Classification Society - vol. 12(4), pages 937-952, December 2018.
17. Aytug˘ Onan, Serdar Korukog˘ lu, A Feature Selection Model Based on Genetic Rank Aggregation for Text Sentiment Classification, Journal of Information Science, 2017, Vol. 43(1) 25–38.
18. Barbara Pes, Ensemble Feature Selection for High-dimensional Data: A Stability Analysis across Multiple Domains, Neural Computing and Applications, https://doi.org/10.1007/s00521-019-04082-3. , Springer, February 2019
19. Donghai Guan, Weiwei Yuan, Young-Koo Lee, Kamran Najeebullah & Mostofa Kamal Rasel (2014) A Review of Ensemble Learning Based FeatureSelection, IETETechnicalReview, and 31:3, 190-198, DOI: 10.1080/02564602.2014.906859.

AUTHORS PROFILE

Prof Noureldien is a Computer Science Professor and working as the Dean of the faculty of computer science and information technology, University of Science and Technology, Omdurman, Sudan. He has more than 20 years of teaching experience. Prof Noureldien received his B.Sc. and M.Sc. degrees from the University of Khartoum, Sudan, and received his Ph.D. in 2001 from Sudan University of Science and Technology. He has published over 20 research article in reputed journals. He has also presented more than 15 research papers in international conference. He has successfully supervised tens of M.Sc. students and three PhD dissertations. He is currently supervising M.Sc. and PhD students in the field of Data Mining and cloud security.

Ms .Enass has completed her Master degree in Computer science at University of Science and Technology, Omdurman, Sudan. She is now pursing her Ph.D. in computer science. Her areas of interest include Data Mining and Information Security.