

Feature Selection from Colon Cancer Dataset for Cancer Classification using Artificial Neural Network

Md. Akizur Rahman[#], Ravie Chandren Muniyandi[#]

[#] *Research Center for Cyber Security, Faculty of Information Science and Technology, University Kebangsaan Malaysia, Bangi, 43600, Selangor, Malaysia*
E-mail: akiz.akizur@gmail.com; ravie@ukm.edu.my

Abstract— In the fast-growing field of medicine and its dynamic demand in research, a study that proves significant improvement to healthcare seems imperative especially when it is on cancer research. This research paved the way for such significant findings by the inclusion of feature selection as one of its major components. Feature selection has become a vital task to apply data mining algorithms effectively in the real-world problems for classification. The Feature selection has been the focus of interest for quite some time and much-completed work related to it. This study used feature selection for improving classification accuracy on the cancerous dataset. This study proposed Artificial Neural Network (ANN) for cancer classification by the feature selection on colon cancer dataset. The study used the best first search method in Weka tools for feature selection. The result of the experiment achieved 98.4 %, accuracy for cancer classification after feature selection by using the proposed algorithm. The result indicated that feature selection improves the classification accuracy based on the experiment conducted on the colon cancer dataset.

Keywords—Artificial Neural Network; cancer classification; feature selection; colon cancer.

I. INTRODUCTION

The tradition of the bioinformatics on the condition and illness of patients with the advent of microarray techniques lead to intake of adequate information. Beyond accurate diagnosis to the patients is extremely important for the appropriate treatment [1], [2]. This study focuses on the condition of patients. The majority of real-world classification problems required monitored learning in which the underlying of unfamiliar, class probabilities and class-conditional probabilities connected to a class label. The machine learning methods with the massive number of input features have dealt with difficulty, which is posing a remarkable challenge for researchers. The circumstances of relevant features are frequently unfamiliar as prior knowledge of real-world problems.

Consequently, the domain introduced several candidate features for better representation. Unfortunately, numerous fundamental ideas were either incompletely or entirely irrelevant or redundant. However, a relevant feature to the critical idea is neither irrelevant or redundant since it does not influence the irrelevant feature in any way to the key idea whereas, in the redundant feature, the critical idea cannot include anything new [3], [4]. A dataset has a huge size in many applications indicating that learning might not

be suitable before these useless features eliminate. A significance in the decreasing number of irrelevant features has the potential to decrease the time duration of a learning algorithm and produce a universal idea expansion. It is consequently beneficial for real-world classification problem to receive a superior connection as a fundamental idea [5]–[7]. Feature selection techniques attempt to select a subset of features, which may contribute to or affect the critical idea.

The basis for such statistical analysis and classification of multivariate data investigates two segments, namely feature selection and classification [1], [8]. The relevant features extract and reduce the dimensions of feature selection from the primary dataset. The selected features lead to the classification procedure using the classification algorithms. Feature selection and classification encompasses independent techniques. Both interactions have revealed and remained undiscovered among procedures. Specifically, the improvement of classification performance is the primary objective of this study.

Feature selection techniques, which depend on the utilization of the objective information, have two types: (1) unsupervised learning method and (2) supervised learning method. In unsupervised learning, either a principal component analysis (PCA) or an independent component analysis (ICA) regards correlations among variables or properties and changes the unique illustration for the

extraction of features with no significant loss of information. However, the fundamental values have no existing information, which can be deemed unsuitable for implementation in classification [9]. Contrastingly, supervised learning utilizes class information known earlier. Filter techniques evaluate the relevance of features with consideration of inherent characteristics of the data, whether they are comparatively significant or not in the current feature. Boundaries of unsupervised learning and fast computation aside from existing problems between selection and classification algorithms for disregarding feature dependencies and interaction distinguish this method [10].

The connection between the feature subset and classification algorithms involved the wrapper method and supervised learning method leading to augmentation. This selection method demonstrates the space of particular feature subset with several subsets of features that are generating systematically as their performance in a forward or backward direction; this seems comparable with defends of the former. The space of feature subsets grows exponentially with the number of features. It required heuristic methods for the searching of an optimal subset with the computationally exquisite approach. Additionally, it shows an over-fitting problem [11]. The wrapper method showed the accuracy improves of best-first search [12]. However, expanding the search seems to overflight the accuracy estimation guides. The search toward feature subsets benefits for the specific cross-validation folds. In the study of the classification, the statistical techniques of a large number designed for the classification of the original data or selected features for appropriate targets. Support Vector Machines (SVMs) have become a popular classification technique developed by Cortes et al. [13].

The function of SVMs is for maximum margin on the side of a hyperplane by separating into two data classes, which in turn create the highest potential space among hyperplanes separation. In multivariate classification, SVMs are responsible in identifying the support vectors (SVs) which are a dimension demote application to selected features. In case of the selected features for classification, it seems to indicate a negative response in any significant information, SVMs are unable to identify the correct directions of the hyperplanes, leading to a downgrade in the rendition of the SVMs [14]. The feature selection method is prevalent in various research fields from the literature investigation. Nowadays, most of the researchers used feature selection method for classification problem. That is why the feature selection method is also currently popular in the cancer research field.

Cancer is an alarming case for every human being that a thorough understanding of the classification of this disease is imperative. The traditional approaches to diagnosing cancer highly depend on the doctor's expertise and their visual inspections. Human beings naturally commit mistakes due to their limitations, yet humans can easily recognize patterns. There is a substantial quantity of data with low quality and redundant information that even for medical experts, may be tough to acquire the accurate classification. Computer-aided diagnostic tools are intended to help physicians for the accuracy of classification improvement [15]–[17].

Some methods and techniques have been used for cancer classification. Specifically, Onur Inan et al. [18] used breast cancer datasets for his experiment. This experiment used the Apriori algorithm (AP) with a Neural Network (NN) classifier for cancer. *N. N. Mohd Hasri* et al. [19] used support vector machine (SVM) with recursive feature elimination (RFE) method for cancer classification. Xin Sun et al. [20] suggested a dynamic weighting feature selection (DWFS) algorithm for classification of cancer. Maryam Yassi et al. [21] suggested a robust and stable feature selection through the integrating ranking methods (IRM) and the wrapper method in genetic data classification. The experiments of the proposed approach on five cancerous microarray datasets include colon cancer.

Thanh Nguyen et al. [22] introduced a new technique in the selection of features based on a modification of the analytic hierarchy process (MAHP). The author used different classifiers covering linear discriminant analysis, probabilistic neural network, k-nearest neighbors, multilayer perceptron, and support vector machine. Maolong Xi et al. [15] proposed Binary Quantum-behaved Particle Swarm Optimization (BQPSO) for feature selection from cancer datasets. The author used five microarray datasets including colon cancer. Aalaei et al. [17] used GA-Based classifier with the ANN.

This study also aimed to classify cancer by using feature selection method on cancerous datasets. Lingyun Gao et al. [23] used Fast Correlation-Based Feature Selection (FCBFS) technique with PA-SVM classifier for the filtering of irrelevant and redundant features leading to the improvement of the quality of cancer classification by. The proponent created a classification model based on PSO and ABC. Hanaa Salem et al. [24] presented a new method based on the gene expression profiles to classify human cancer diseases. The author proposed a method that combines both Information Gain (IG) and the Standard Genetic Algorithm (SGA). Information Gain (IG) and Genetic Algorithm (GA) used for feature selection and feature reduction in this experiment. Finally, this study used Genetic Programming (GP) for cancer type's classification. S.A. Ludwig et al. [25] presented a fuzzy decision tree algorithm in classifying gene expression data. The literature investigated provided information, which helped for the comparison of result with the proposed method.

Artificial Neural Networks (ANNs) is a machine learning approach. A prevalent type of ANNs is called feed-forward neural networks [26]–[30]. Artificial Neural Networks (ANNs) is similarly valued in estimating complex target functions with appropriate modeling and iterative learning. In nonlinear function, the target is determined according to the network architecture minus the linearity and conventions of limitation. These networks can apply easily due to the performance of improving the computer systems, with complexity and various areas to employ and evaluation models.

Consequently, neural networks utilized the regression and discriminated analysis for conventional statistical methods in the current period. The representation of a neural network formed as influential and adaptive nonlinear equation. They can provide information as regards to the multifaceted operational connections between the input and output data

[31]. Because of these attributes, once the output nodes accommodate real values, they are capable of becoming a regressor. Furthermore, a classifier can be defined if the outputs are an integer or absolute values. In overall, the feature selection methods produce actual values as outputs, and absolute values return by the classification algorithms. Hence, this study utilized neural networks for the estimation of the diagnostic of colon cancer classification.

This study contains an introductory section about the experiment with a review of the related study. The second section discusses the Feature Selection Method, Cancer, and Artificial Neural Network (ANN). The third section is a discussion on the experimental dataset while the fourth section presents the research experimental results and performance comparisons. Finally, in the last section, conclusion summarized the conducted research with the emphasis on the initial result of the experiment.

II. MATERIAL AND METHOD

In this section, the feature selection, feature selection model, and classification model described in details. Then, this manuscript describes Basic Artificial Neural Networks (ANNs) implementation with ANN Classifier. Next, the colon cancer microarray dataset defined; this, which used in this experiment.

A. Feature Selection

Feature selection is a method to reduce the number of attributes and selects a subset from the original features. In data pre-processing the feature selection frequently utilized to classify relevant features that are frequently unknown before and take away the noise and irrelevant or redundant features, which have zero significance in the classification task. The progress of classification accuracy is the main goals of feature selection [32]–[35]. This section illustrates the feature selection model and classification model as shown in Fig. 1 (a) and (b).

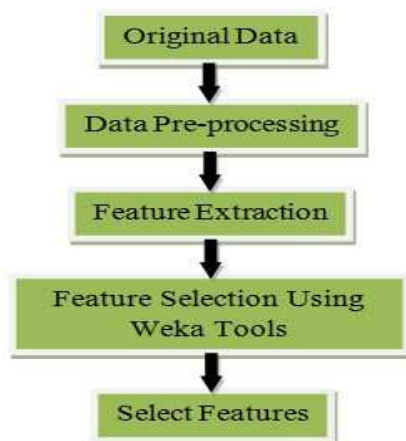


Fig. 1(a): Feature Selection Model

In Figure 1(a) the feature selection model depicts the functional block diagram. It consists of four steps; original data, data pre-processing, feature extraction, and feature selection using weka tools. Whereas in Figure 1(b) the classification model consists of three steps; selected features,

selected features pre-processing and ANN classifier with MATLAB tools.

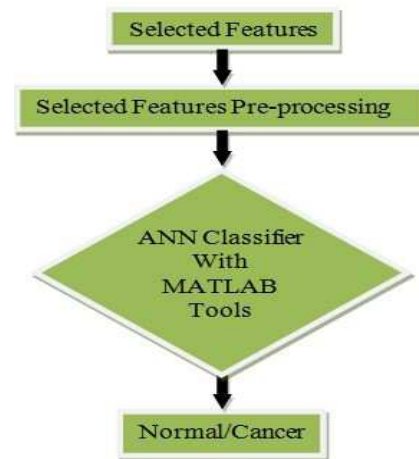


Fig. 1(b): Classification Model

The beginning part of the experiment in the original data step is subject to the extraction of feature and process of selection in identifying the input vector for the subsequent classifier. This leads to the decision on the class associated with vector pattern. Dimensionality reduction completed based on either feature selection or feature extraction. In the pre-processing step, the feature prepared and filtered for clearing noise and improving the features quality. Conversely, feature extraction encompasses the complete information content and plots the useful detail content into a lower dimensional feature. Feature selection based on neglecting those features from the available measurements that do not lead to class separable. In this case, redundant and irrelevant features disregarded. In the classification steps, ANN classifier utilized so that the ideal result of cancer classification could be obtained.

B. Artificial Neural Networks

Artificial Neural Networks (ANNs) are computing systems, which is a branch of computational intelligence that utilizes a variety of optimization tool with layers of computing nodes that have incredible processing information features [30], [36]–[38]. Nonlinearities can be detected but are not formulated as inputs; hence, they can learn and adapt. It possesses the following features according to the high parallelism, robustness, noise tolerance, and generalization, which are having the ability of clustering, function approximation, forecasting, and association, and lastly, performed massive parallel multi-factorial analyses for modeling complex patterns with small prior knowledge.

An artificial neural network implemented by the use of Neural Network Pattern Recognition apps in MATLAB Tool. Sigmoid hidden and Softmax output neurons, a two-layer feed-forward network can be classified as vectors randomly with enough neurons given in its hidden layer. In this study, the hidden layer used 18 neurons and 2 neurons in the output layer. ANNs used multiple layers for the training with an efficient method applied and with back propagation-learning algorithm to the network. Those outputs are probably

standard or cancer. A feed-forward neural network necessary architecture representation presented below in Fig. 2.

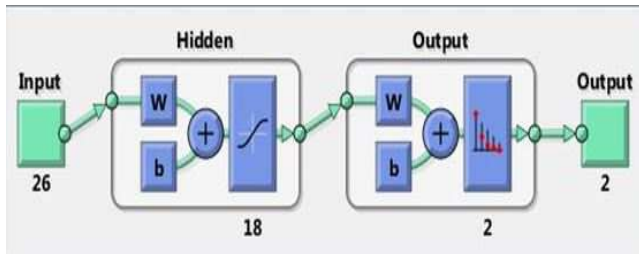


Fig. 2 Basic Architecture of Artificial Neural Network

The feed-forward network has links extending to a single direction only. There is no backward connection in the feed-forward network. All connections proceed from the input node toward the output node. The classifier needs to use for classification. This study used ANNs classifier for colon cancer classification. A typical Artificial Neural Networks (ANNs) classifier of graphical representation is presented in Figure 3.

Artificial Neural Networks is a classifier due to a single node that imitates biological neurons by acquiring the input data and executing data on a simple operation. The selection of the results transfers in other neurons is called ‘activation’ through every output of each node. The vector and node on the network connected with weight (w) values and those values constrain how inputs to output data connected. The connectivity of weight values with single nodes named as biases (b). Weight values through the network identified by the iterative flow of training data. The weight values proved during a training phase since the network discovers in identifying a particular cluster by the characteristics of typical input data.

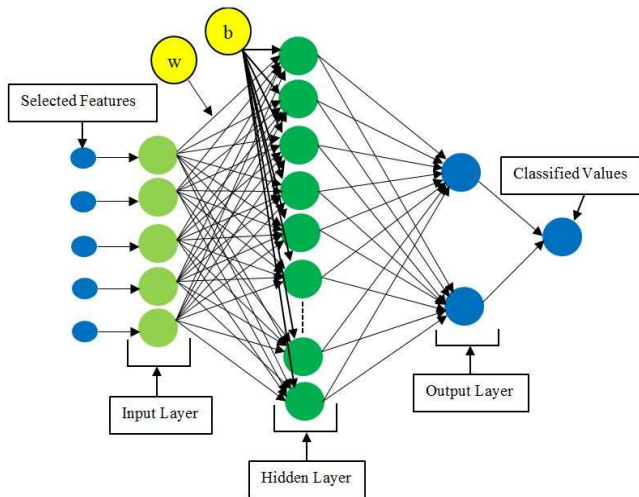


Fig. 3 ANN Classifier

C. The Dataset

The cancer is the most dangerous disease concerning all over in the world. Plenty of research about this issue have done and continually eager to find the fastest and accurate classification. Thus, the medical science community still is working to achieve the accuracy of classification. One kind of cancer diseases is the colon cancer. In the United States,

colon cancer is considered to rank as second and a major leading cause of mortality among men and women [39]. Besides that colon cancer ranks as third among the most common cancer; it also ranks fourth as the primary cause of cancer mortality global [40]. Colon cancer screening is essential for decreasing incidence, morbidity, and death from the disease. This study used colon cancer microarray dataset to distinguish cancer from normal samples for getting an initial result. The dataset is publicly available at <http://csse.szu.edu.cn/staff/zhuzx/Datasets.html>. The dataset consists of 2000 attributes and 62 instances with 2 classes, which are 22-normal, and 40- cancer. After feature selection, this study obtained the 26 most valuable features. At these 26 features and 62 instances with 2 classes used for colon cancer classification.

III. RESULT AND DISCUSSION

This experiment used an Intel(R) Core-(TM) i5-6400 CPU with the 2.10GHz processor; Installed memory 8.00 GB; Operating system 64-bit Windows 7. MATLAB and Neural Network toolbox utilized to implement the proposed system. The proposed algorithm tested by using colon cancer microarray datasets to achieve the initial result with feature selection. Table 1 shows the experiment results. The experiment has five results with Cross-Entropy, Percent of Error, and Classification Accuracy Percent. Cross-Entropy defines the results of proper classification in minimizing Cross-Entropy: the lower value, the better; zero means no error. Percent of error indicates that the fraction of samples is misclassified. Zero result indicates no misclassification; hence, 100 indicate maximum misclassification. Classification Accuracy Percent indicates an accurate classification. In this study after feature selection, the proposed algorithm ANNs achieved the initial experiment result with Classification Accuracy of 98.4%.

TABLE I
THE EXPERIMENTAL RESULTS

No of Experiments	Cross-Entropy (CE)	Percent of Error (%E)	Classification Accuracy Percent (%)
1	8.9772	3.2258	96.80
2	5.2208	3.2258	96.80
3	3.2004	1.6129	98.40
4	3.4364	4.8387	95.20
5	7.6149	6.4516	93.50

This study experiment conducted with the ANNs training performance, ANNs training state performance, ANNs error histogram result, Confusion matrixes performance and Receiver operating characteristic (ROC) curve result. The training performance of ANN is shown in Fig. 4.

The plot performance at the start of the training of cross-entropy shows a maximum error. This research used 62 samples dataset. ANNs performed with training, validation, and testing on the dataset. The network presented during training and accordingly adjusted the network by its error. Validation utilized to compute network simplification and to halt training through improving simplification. Testing ineffective on the training; therefore, the performance of the

network provided with independent measures during and after training.

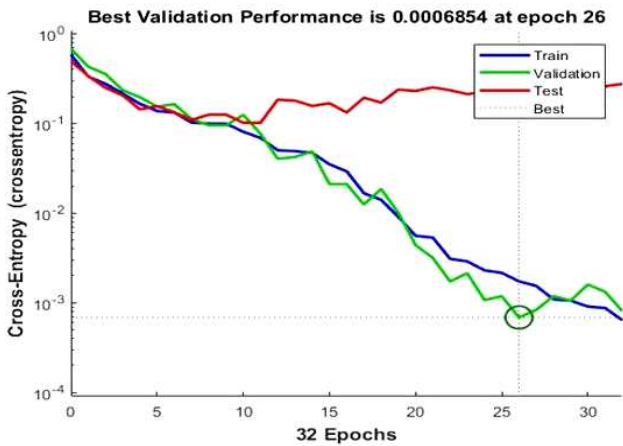


Fig. 4 ANN training performance

The dataset consisting of 62 samples, which divided randomly into three groups: 44 (70%) samples for training and 9 samples (15%) for both validation and testing. In this proposed method, the exact validation performance is 0.0006854 at epoch 26, and at this point of the cross-entropy, the error was almost zero. Fig. 5 shows the training state performance.

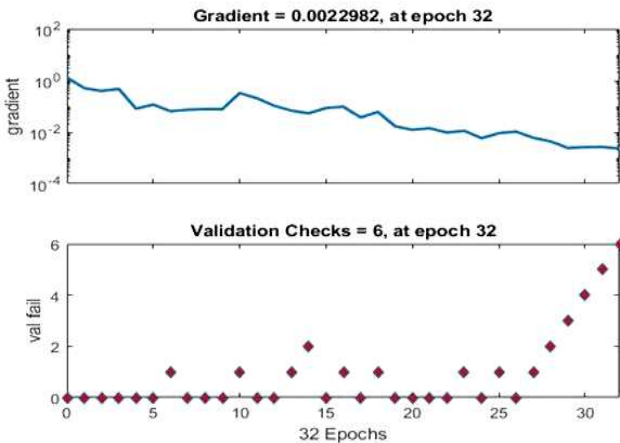


Fig. 5 ANN (Artificial Neural Network) training state performance

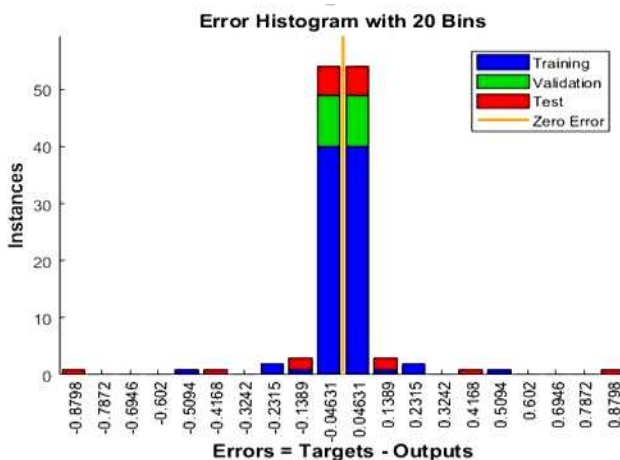


Fig. 6 ANN error histogram

The performance plot of training state showed the gradient is 0.0022982 at epoch 32. Within this stage, the cross-entropy error almost zero. The training continues until 6 at epoch 32. Fig. 6 shows the error histogram.

The Error Histogram plot represents the error values: the difference between critical values and predicted values. This study shows the error histogram with 20 bins. The proposed system shows the error of almost zero.

Fig. 7 shows the Confusion Matrix. There are three sets of overall confusion matrix of collectively confusion matrices namely: ANN training confusion matrix, ANNs validation confusion matrix and ANNs testing confusion matrix. All confusion matrices with first two diagonal cells demonstrate the number and percentage of correct classifications by the trained network. The Normal biopsy results are 22, which correctly classified with the corresponding 35.5% of the total 62 biopsies. On the other hand, the Cancer biopsy results are 39, which correctly classified with the corresponding 62.9% of total biopsies.



Fig. 7 ANN confusion matrix

The experiment also revealed that one of the results of cancer biopsies incorrectly classified as standard with the corresponding 1.6% of the total biopsy data. None of the results of normal biopsies is incorrectly classified as cancer with the corresponding 0.0% of the total biopsy data. The total result of 23 normal revealed that 95.7% correct and that 4.3% incorrect. The total result of 39 cancer was 100% correct and 0.0% incorrect. The total result of 22 typical cases 100% correctly classified as normal and 0.0% of the result of cancer. Out of the 40 cancer cases, 97.5% were correctly classified as cancer and 2.5% classified as normal. The total confusion matrix plots 98.40% of classification are correct, and 1.6% is misclassification for the proposed system. Fig. 8 shows the ROC Curve.

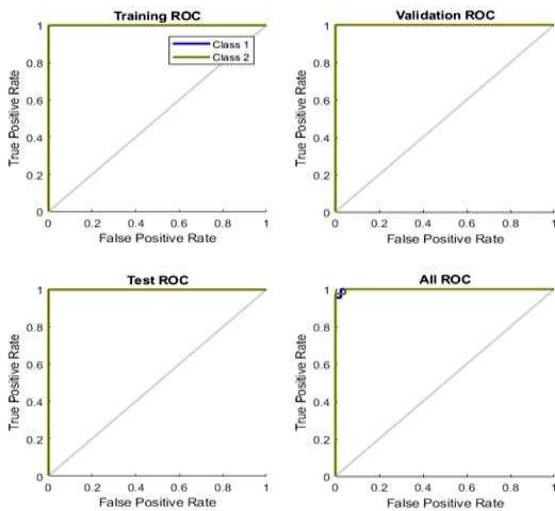


Fig. 8 ANN ROC Curve

ROC curve indicates the correct positive rate and false positive rate in different edge setting of the network, which illustrates a premium percentage result for the system. The Neural Network after training, validation, and testing performed 98.40% correct classification for the two classes of normal and cancer.

TABLE II
PERFORMANCE COMPARISON WITH OTHER RESEARCH WORKS

Author Name & Reference	Method	Accuracy %	Year
Xin Sun et al.[20]	DWFS	96.67%	2013
Maryam Yassi et al.[21]	IRM with Wrapper	96.00%	2014
Thanh Nguyen et al. [22]	MAPH+PNN	86.45 %	2015
Maolong Xi et al.[15]	BQPSO	93.55%	2016
Lingyun Gao et al. [23]	FCBFS	93.55 %	2017
Simone A. Ludwig et al. [25]	FDT	80.28	2018
This Study	ANN	98.40%	2018

Table 2 indicates that this study outperformed the results of other experiments (authors). It should note that they used a different method for classification. All results and works of the authors compared here based on the classification of accuracy, which primarily concerned about cancer classification. The table shows the different experimental results and methods of various authors. The table compared based on colon cancer dataset. All authors experimented to achieve accuracy in the classification of cancer.

As compared, the table reveals the feature selection by S. A. Ludwig et al. [25] indicating 80.28% lower classification accuracy. The author-selected 26-features, which were the same number of features selected in this study. Hence, the performance was not similar and achieved different accuracy, which was much lower; perhaps the author used Fuzzy Decision Tree (FDT) Algorithms for the classification of colon cancer dataset. Moreover, the comparison table shows that Xin Sun et al.[20] selected 23-features from colon cancer dataset and achieved the upper classification accuracy

of 96.67% by using dynamic weighting feature selection (DWFS) algorithm with classifier 1-NN (1-Nearest Neighbor); however, there was low accuracy compared to the result of this study.

Conversely, the author Maryam Yassi et al. [21] performed feature selection by using IRM (integrating ranking method) with Wrapper method. The author selected 50-features from colon cancer dataset, and the experiment achieved 96.00% of the classification accuracy that was still low accuracy from our study. Maolong Xi et al.[15] selected 5-features and achieved 93.55 % classification accuracy from the colon cancer dataset by using binary quantum-behaved particle swarm optimization (BQPSO) algorithm. Lingyun Gao et al.[23] revealed the same classification accuracy of 93.55% from 14-selected features on colon cancer dataset by using FCBFS method for feature selection with PA-SVM as a classifier for cancer classification. Both researchers showed similar results, but our study has a higher result. Thanh Nguyen et al. [22], used the MAPH method for feature selection from cancer dataset and the classifier PNN (probabilistic neural network) used for cancer classification. The author gave an unclear explanation about how many features selected from colon cancer dataset. The result of this experiment achieved 86.45% classification accuracy, which is low accuracy than our study.

The result of our study was different compared with the previous works; thus, our study selected 26 most valuable features from colon cancer dataset, used different feature selection method, and used different classifier. This study of cancer classification achieved **98.40%** accuracy using the ANN with 3-layers namely; Input layer, Hidden layer, and an Output layer. Moreover, there were 26-inputs, 18-hidden neurons, and 2-outputs to indicate either normal or cancer. The initial weights and biases selected randomly. Activation Functions is Tangent-Sigmoid for hidden layer while Log-Sigmoid for the output layer. Based on the comparison table, this study revealed a significant increase in the classification accuracy of cancer.

IV. CONCLUSIONS

This research aimed at feature selection for the accurate classification algorithm. Without feature selection, this study achieved 95.2% accuracy from the colon cancer dataset, which has 62 instances, 2000 attributes, and 2 classes. After feature selection, the proposed algorithm implemented on the colon cancer dataset, which has 62 instances, 26 selected attributes, and two classes. For feature selection, this study used Best first search method and the Artificial Neural Networks (ANNs) for cancer classification. ANNs Algorithm achieved **98.40%** accurate classification on the initial result. In this regard, the ANNs have several advantages including the ability to process a large amount of data, reduced likelihood of overlooking relevant information, and reduction of classification time. ANNs have proven suitable for satisfactory classification of cancer. In future work analysis of more samples and further improving neural network method will be carried out to improve for greater accuracy and reduce computational discrepancy.

ACKNOWLEDGMENT

The Ministry of Higher Education of Fundamental Research Grant Scheme has supported this study (Malaysia; Grant code: FGSR/1/2015/ICT04/UKM/02/3).

REFERENCES

- [1] I. Sarkar et al., "Characteristic attributes in cancer microarrays," *J. Biomed. Inform.*, vol. 35, no. 2, pp. 111–122, Apr. 2002.
- [2] N. Elkhani and R. C. Muniyandi, "Membrane computing inspired feature selection model for microarray cancer data," *Intell. Data Anal.*, vol. 21, no. S1, pp. S137–S157, Apr. 2017.
- [3] G. H. John, R. Kohavi, and K. Pfleger, "Irrelevant Features and the Subset Selection Problem," in *Machine Learning Proceedings 1994*, Elsevier, 1994, pp. 121–129.
- [4] N. Elkhani and R. C. Muniyandi, "Membrane Computing to Model Feature Selection of Microarray Cancer Data," *Proc. ASE BigData Soc.* 2015, p. 13, 2015.
- [5] R. Kohavi and D. Sommerfield, "Feature subset selection using the wrapper method: overfitting and dynamic search space topology," *Proceedings of the First International Conference on Knowledge Discovery and Data Mining*. AAAI Press, pp. 192–197, 1995.
- [6] D. Koller and M. Sahami, "Toward optimal feature selection," *ICML'96 Proc. Thirteen. Int. Conf. Int. Conf. Mach. Learn.*, pp. 284–292, 1996.
- [7] K. Islam, G. Mujtaba, R. R.-... ICE2T), 2017 International, and undefined 2017, "Elevator button and floor number recognition through hybrid image classification approach for navigation of service robot in buildings," *researchgate.net*.
- [8] S. N. Das, M. Mathew, and P. K. Vijayaraghavan, "An Approach for Optimal Feature Subset Selection using a New Term Weighting Scheme and Mutual Information," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 1, no. 3, p. 273, 2011.
- [9] M. S. Park and J. Y. Choi, "Theoretical analysis of feature extraction capability of class-augmented PCA," *Pattern Recognit.*, vol. 42, no. 11, pp. 2353–2362, Nov. 2009.
- [10] Y. Saeys, I. Inza, and P. Larranaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, Oct. 2007.
- [11] M. Monirul Kabir, M. Monirul Islam, and K. Murase, "A new wrapper feature selection approach using the neural network," *Neurocomputing*, vol. 73, no. 16–18, pp. 3273–3283, Oct. 2010.
- [12] R. Kohavi, "Feature Subset Selection as Search with Probabilistic Estimates," *AAAI Fall Symp. Relv.*, pp. 122–126, 1994.
- [13] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
- [14] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.
- [15] M. Xi, J. Sun, L. Liu, F. Fan, and X. Wu, "Cancer Feature Selection and Classification Using a Binary Quantum-Behaved Particle Swarm Optimization and Support Vector Machine," *Comput. Math. Methods Med.*, vol. 2016, pp. 1–9, 2016.
- [16] M. A. Rahman and R. C. Muniyandi, "Review of GPU implementation to the process of RNA sequence on cancer," *Informatics Med. Unlocked*, vol. 10, pp. 17–26, Jan. 2018.
- [17] S. Aalaei, H. Shahraki, A. Rowhanimanesh, and S. Eslami, "Feature selection using genetic algorithm for breast cancer diagnosis: an experiment on three different datasets," *Iran. J. Basic Med. Sci.*, vol. 19, no. 5, pp. 476–82, May 2016.
- [18] O. Inan, M. S. Uzer, and N. Yilmaz, "A New Hybrid Feature Selection Method Based on Association Rules and Pca for Detection Of Breast Cancer," *Int. J. Innov. Comput.*, vol. 9, no. 2, pp. 727–739, 2013.
- [19] N. N. Mohd Hasri, N. H. Wen, C. W. Howe, M. S. Mohamad, S. Deris, and S. Kasim, "Improved Support Vector Machine Using Multiple SVM-RFE for Cancer Classification," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 4–2, p. 1589, Sep. 2017.
- [20] X. Sun, Y. Liu, M. Xu, H. Chen, J. Han, and K. Wang, "Feature selection using dynamic weights for classification," *Knowledge-Based Syst.*, vol. 37, pp. 541–549, Jan. 2013.
- [21] M. Yassi and M. H. Moattar, "Robust and stable feature selection by integrating ranking methods and wrapper technique in genetic data classification," *Biochem. Biophys. Res. Commun.*, vol. 446, no. 4, pp. 850–856, Apr. 2014.
- [22] T. Nguyen, A. Khosravi, D. Creighton, and S. Nahavandi, "A novel aggregate gene selection method for microarray data classification," *Pattern Recognit. Lett.*, vol. 60–61, pp. 16–23, Aug. 2015.
- [23] L. Gao, M. Ye, and C. Wu, "Cancer Classification Based on Support Vector Machine Optimized by Particle Swarm Optimization and Artificial Bee Colony," *Molecules*, vol. 22, no. 12, p. 2086, Nov. 2017.
- [24] H. Salem, G. Attiya, and N. El-Fishawy, "Classification of human cancer diseases by gene expression profiles," *Appl. Soft Comput.*, vol. 50, pp. 124–134, Jan. 2017.
- [25] S. A. Ludwig, S. Picek, and D. Jakobovic, "Classification of Cancer Data: Analyzing Gene Expression Data Using a Fuzzy Decision Tree Algorithm," *Springer, Cham*, 2018, pp. 327–347.
- [26] J. Hopfield, "Artificial neural networks," *IEEE Circuits Devices Mag.*, vol. 4, no. 5, 1988.
- [27] K. Islam, G. Mujtaba, ... R. R.-E. T., and undefined 2017, "Handwritten digits recognition with the artificial neural network," *ieeexplore.ieee.org*.
- [28] G. Bebis and M. Georgiopoulos, "Feed-forward neural networks," *IEEE Potentials*, vol. 13, no. 4, 1994.
- [29] K. T. Islam, R. G. Raj, and G. Mujtaba, "Recognition of Traffic Sign Based on Bag-of-Words and Artificial Neural Network," *Symmetry (Basel)*, vol. 9, no. 8, p. 138, Jul. 2017.
- [30] N. M. Nawli, N. A. Hamid, N. A. Samsudin, M. A. Mohd Yunus, and M. F. Ab Aziz, "Second Order Learning Algorithm for Back Propagation Neural Networks," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 4, pp. 1162–1171, 2017.
- [31] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert Syst. Appl.*, vol. 36, no. 1, pp. 2–17, Jan. 2009.
- [32] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997.
- [33] N. Elkhani and R. Muniyandi, "Review of the Effect of Feature Selection for Microarray Data on the Classification Accuracy for Cancer Data Sets," *Int. J. Soft Comput.*, vol. 11, no. 5, pp. 334–342, 2016.
- [34] S. S. Shreem, S. Abdullah, and M. Z. A. Nazri, "Hybridising harmony search with a Markov blanket for gene selection problems," *Inf. Sci. (Ny)*, vol. 258, pp. 108–121, Feb. 2014.
- [35] S. Sahran, D. Albashish, A. Abdullah, N. A. Shukor, and S. Hayati Md Pauzi, "Absolute cosine-based SVM-RFE feature selection method for prostate histopathological grading," *Artif. Intell. Med.*, Apr. 2018.
- [36] F. M. Selaru et al., "Artificial neural networks distinguish among subtypes of neoplastic colorectal lesions," *Gastroenterology*, vol. 122, no. 3, pp. 606–613, Mar. 2002.
- [37] K. Islam and R. Raj, "Real-time (vision-based) road sign recognition using an artificial neural network," *Sensors*, vol. 17, no. 4, 2017.
- [38] O. Alesawy and R. C. Muniyandi, "Elliptic Curve Diffie-Hellman Random Keys Using Artificial Neural Network and Genetic Algorithm for Secure Data over Private Cloud," *Inf. Technol. J.*, vol. 15, no. 3, pp. 77–83, Jun. 2016.
- [39] J. Redmond, R. Vanderpool, and R. McClung, "Effectively Communicating Colorectal Cancer Screening Information to Primary Care Providers," *Am. J. Heal. Educ.*, vol. 43, no. 4, pp. 194–201, Jul. 2012.
- [40] M. A. Pourhoseingholi, S. Kheirian, and M. R. Zali, "Comparison of Basic and Ensemble Data Mining Methods in Predicting 5-Year Survival of Colorectal Cancer Patients," *Acta Inform. Med.*, vol. 25, no. 4, pp. 254–258, Dec. 2017.