Annals of Mathematical Modeling
Vol. 3, No. 1, June 2023, pp. 1-13
ISSN 2715-7822

Lung cancer classification based on support vector machine-recursive feature elimination and artificial bee colony

Alhadi Bustamam [a,1,*], Zuherman Rustam[a,2], Selly A. A. K [a,3], Nyoman A. Wibawa[a,4], Devvi Sarwinda[a,5], NadyaAsanul Husna[a,6]

[a] Laboratory of Bioinformatics and Advanced Computing, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok 16424, Indonesia
[1] alhadi@sci.ui.ac.id*
* corresponding author

ARTICLE INFO

Keywords
Artificial bee colony
Lung cancer classification
Support vector machine

ABSTRACT

Early detection of cancerous cells can increase survival rates for patients by more than 97%. Microarray data, used for cancer classification, are comp osed of many thousands of features and from tens to hundreds of instances. Handling these huge datasets is the most imp ortant challenge in data classification. Feature selection or reduction is therefore an essential task in data classification. We prop ose a cancer diagnostic to ol using a supp ort vector machine for classifier and feature selection. First, we use supp ort vector machine-recursive feature elimination to prefilter the genes. This was enhanced with the artificial b ee colony algorithm. We ran four simulations using Ontario and Michigan lung cancer datasets. This approach provides higher classification accuracy than those without feature selection, supp ort vector machine-recursive feature elimination, or the artificial b ee colony algorithm. The accuracy of a supp ort vector machine using a feature selection-based recursive feature elimination metho d combined with the artificial b ee colony algorithm reached 98% with 100 b est features for the Michigan lung cancer dataset and 97% with 70 b est features for the Ontario lung cancer dataset. SVM with RFE-ABC as the feature selection metho d gives us an accurate result to diagnose Lung cancer using microarray data.

1. Introduction

Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells, which can result in death if untreated. According to the International Agency for Research on Cancer, there were 14.1 million new cases in 2012. By 2030, there will be 21.7 million more new cases, resulting in 13 million deaths (American Cancer Society). The World Health Organization notes that during 2015, lung cancer alone caused the death of 1.69 million people. Lung cancer may be classified into one of two types according to its growth pattern and the treatment required: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC); the former accounts for 80% of all lung cancer cases [1]. As the number of new cases of lung cancer is rising, and with it the number of deaths, a method to accelerate the classification process is needed [2], [3]. According to the

American Cancer Society, early treatment of cancerous cells can increase survival rates for patients by more than 97%. Thus, time plays a crucial role in detecting the disease. Although imaging techniques are a primary method of detection and diagnosis, they can only detect malignant cells in late stages of lung cancer, resulting in a low survival rate; 16% for NSCLC and 6% for SCLC [4].

Today, lung cancer can be detected by DNA microarray analysis. The main problem with this is one of classification problems in machine learning. Medical microarray data are characterized by its many features (i.e., genes), meaning that it requires a long computational time, and not all features are significant in classification. The situation is more complicated the larger the sample size. The high dimensionality of the features and the low population size cause overfitting of the classifier, necessitating a feature selection procedure to overcome this problem. Feature selection techniques can be categorized into three different methods: filter, wrapper, and embedded. The filter method is independent of classification algorithms and only the highest-ranking features are used, whereas the remaining low- ranking features are eliminated. Wrapper methods utilize a predetermined learning algorithm for feature subset evaluation, with the usefulness of a feature directly judged by the estimated accuracy of the learning method [5]. In the embedded method, feature selection is the search mechanism built into the classifier model [6].

2. Dataset

The data used in the present paper are from microarray analysis of lung cancer patients and were obtained from the Kent Ridge Biomedical Dataset Repository (http://datam.i2r.astar.edu.sg/datasets/krbd/). Two datasets were used: the Michigan lung cancer dataset and the Ontario lung cancer dataset (Table 1). The data are represented in a matrix in which each column represents genes as features, each feature has a value (gene expression value), and each row represents the characteristic of each sample along with its class label.

Table 1. Data Overview

Microarray Dataset	Total Samples	Total Features	Total Class
Michigan	96	7129	2
Ontario	39	2880	2

Table 2. Michigan lung cancer dataset

	Features				
Data	***A28102_at (nucleotide)***	***AB000114_at (osteomodulin)***	**…**	***Z97074_at (RABEPK)***	**Class**
No	1	2	…	7129	
X1	170	69.4	…	276	Cancer
X2	59.7	18.1	…	134.7	
…	…	…	…	…	…
X87	108	93	…	180	Non Cancer
…	…	…	…	…	…
X96	106.2	164	…	140.5	

Table 3. Ontario lung cancer dataset

Data	Features		…		Class
	A588029 (BRCA1)	***A417978 (DKK1)***		***A243359 (APOE)***	
No	1	2	…	2880	
X1	0.195	0.2	…	0	Cancer
X2	0.375	0.19	…	0.25	
…	…	…	…	…	…
X25	0.695	1.315	…	0.975	Non Cancer
…	…	…	…	…	…
X39	0.61	0.625	…	1.27	

The first column represents the sample involved in the research, and the last column represents the class label of each sample, which is either cancer (1) or non-cancer (-1) (Table 2, Table 3).

3. Method

This study Our proposed approach uses a combination of SVM-RFE and the ABC algorithm [7], since SVM-RFE does not take into account gene redundancy and its effectiveness becomes unstable at some gene values [8][9]; meanwhile, the ABC algorithm is very computational. The approach proposed herein seeks to improve the result obtained by SVM while simultaneously reducing the computational time of the ABC algorithm.

3.1. First Phase

In the first phase, SVM-RFE is used to select a subset of candidate genes from the data. The SVM-RFE algorithm was introduced for feature selection [8], [10], [11]. The SVM-RFE method is one application of recursive feature elimination that uses SVM weight as a ranking criterion and is classified as an embedded method [12]. The main principle of SVM-RFE is to eliminate features that have the least squares value of weights on each iteration. The procedure of recursive feature elimination in general is

- Perform classifier training to find the weight vector (w)
- Calculate ranking criteria for all features
- Elimination of features with the smallest ranking criterion values

SVM-RFE is now widely used for gene selection, and several improvements have been recently suggested [13], [14]. In this algorithm, the genes are removed recursively on the basis of the SVM classifier weights and later the samples are classified with SVM. The algorithm trained by SVM with a linear kernel and features eliminated recursively using the smallest ranking criterion. The weight vector is calculated according to equation (1) to generate a rank of features:

$$w = \sum_k a_k y_k x_k \quad (1)$$

When the algorithm ends, it will have produced a ranked feature list. The algorithm for SVM-RFE is shown as Algorithm 1 below: SVM-RFE, starting with all genes, recursively

removes the gene that is least significant for classification in a backward elimination [15]. The ranking score is given by the components of the weight vector w of the SVM as follows: is the class label of the sample is the Lagrange multipliers involved in maximizing the margin of separation of the classes [16].

When the algorithm ends, it will have produced a ranked feature list. The algorithm for SVM-RFE is shown as Algorithm 1 below: SVM-RFE, starting with all genes, recursively removes the gene that is least significant for classification in a backward elimination. The ranking score is given by the components of the weight vector w of the SVM as follows: $y_k \in$ is the class label of the sample α_κ is the Lagrange multipliers involved in maximizing the margin of separation of the classes [17].

3.2. Second Phase

After SVM-RFE feature selection, the selected features (genes) are chosen again using ABC, an optimization method proposed based on the social behavior of honeybees as they search for high-quality food sources. ABC has three main components: food sources, employed foragers, and unemployed foragers. Employed foragers are bees assigned to the task of conveying information about the distance and direction of the hive as well as the quality of the food source as a definite probability. Unemployed foragers are bees tasked with finding food sources [18].

Table 4. SVM-RFE Algorithm

	Training data $\{x_i, y_i\}_{i=1}^{N}$
Input Initialize	S= 1,2,…,D R = Ø
While S is not empty **do**	
	1. Restrict the features of X_i to the remaining S 2. Train SVM to get w 3. Compute $c_k = w_k^2, k = 1, \ldots, \lvert S \rvert$ 4. Find the feature with smallest value of c_k 5. Add p into R $(R - \{P\} \cup R)$ 6. Remove p from S $(S = S \backslash p)$
Output	Ranked feature list

There are two types of unemployed foragers: scout bees, which search for food sources randomly, and onlooker bees, which wait in the nest and determine the source of quality food using information provided by the employed foragers [19].

Initially, the scout bee finds a food source by searching randomly in the search area. The employed bee is sent out to exploit this food source and determine its quality while also exploring new food sources around old food sources. If a better source of food is found, the employed bee will leave the previous food source to exploit better food sources. All employed bees eventually return to the nest and share information about their food source with onlooker bees. Onlooker bees will choose to follow the instructions of the employed bee according to the specified probability, which is affected by the information given about the quality of the food source; higher-quality food sources have a greater chance of being

selected by the onlooker bee. The onlooker bee will then repeat the exploitation and exploration of selected food sources. An employed bee whose food source has run out will become a scout bee and find a new food source[20].

When implemented as a feature selection method, the food source (SM) represents a set of feature indexes to be optimized. The location of bees looking for food in the ABC algorithm represents the search location of the solution. Food sources that have the highest values are considered the best solutions. The value of the food source is represented as the fitness value, which, in this case, is the average value of the accuracy of the food sources obtained from the classification process using SVM.

There are four control parameters used in the ABC algorithm: the number of food source (SM), feature (D), the value of limit, and the maximum iteration. The stages of finding high-quality food source are as follows:

1) Scout Bee Phase I

Initialize the population of solutions $X_i^j, i = 1,2, ..., SM, j = 1,2,...,D$ using equation:

$$X_i^j = X_{min}^j + w(X_{max}^j - X_{min}^j), 0 \leq \omega \leq 1, \quad (2)$$

X_i^j : Food source in i order, $i = 1,2,\ldots, SM$

X_{min}^j : Minimal feature index, $j = 1,2,\ldots,D$

X_{max}^j : Minimal feature index, $j = 1,2,\ldots,D$

	Index Feature 1	Index Feature 2	Index Feature 3	...	Index Feature D
Food Source 1					
Food Source 2					
Food Source 3					
...					
...					
Food Source SM					

Fig. 1. Food source population in the ABC method

Each line in Figure 1 denotes the f^h food source, where the food source X_i contains the selected feature indexes with dimension D. With the ABC method as the feature selection method, the control parameter for $X_{min} = 1$ and X_{max} =maximum feature index. After initialization, the fitness value of the food source X_i is calculated, $i = 1, \ldots, SM$. The bee colony is divided into three parts: employed bees, onlooker bees, and scout bees.

2) Employed Bee Phase

After initialization, the worker bees look for new high-quality food sources around the original food source. New solution V_i^j are produced using the following equation:

$$v_i^j = X_i^j + \phi(X_i^j - X_i^j - X_k^j), k \neq i, -1 \leq \phi \leq 1 \quad (3)$$

and evaluated with equations (4) and (5). If $v_i^j < X_{min}$,, then

$$v_i^j = (v_i^j) mod(TotalFitur) + 1 \quad (4)$$

If $v_i^j < X_{max}$,,then

$$v_i^j = (v_i^j) mod (TotalFitur) + 1 \quad (5)$$

The fitness value is calculated by SVM accuracy, and the greedy selection process is applied using equation (6):

$$X_i = \begin{cases} V_i, & fit(V_i) > fit(X_i) \\ X_i, & fit(V_i) \leq fit(X_i) \end{cases} \quad (6)$$

and the probability values p_i for the selection X_i^j are calculated using equation (7):

$$p_i = \frac{fitness_i}{\sum_{n=1}^{SM} fitness_i} \quad (7)$$

3) Onlooker Bee Phase

After the food source is exhausted, the employed bee distributes information about food source quality and probability value with the onlooker bee. The onlooker bee selects a new food source based on the probability value. In this process, the onlooker bee generates a new food source according to equation (3) and calculates the fitness value using SVM accuracy and then applies greedy selection

4) Scout Bee Phase II

The scout bee regenerates a new food source using equation (2) when the fitness value of a food source does not increase during the iteration and the abandoned food counter exceeds a predetermined limit. Detailed pseudo-code of the ABC algorithm is given in table 3.

3.3. Support Vector Machine

Support vector machine (SVM) is a machine learning method developed by Boser, Guyon, and Vapnik in 1992. The main purpose of SVM is to find the optimal hyperplane, i.e., the hyperplane with maximum margin. The margin is the distance between the hyperplane with the closest data in each class. The hyperplane function is $w \,.\, x + b = 0$ with weight parameter (w) and bias parameter (b) . The primary optimization problem may be written as

$$min \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{N} \xi_i \quad (8)$$

$$s.t. \begin{cases} y_i (w.x_i + b) \geq 1 - \xi_i, & i = 1,2\ldots,N \\ \xi_i \geq 0 & i = 1,2\ldots,N \end{cases} \quad (9)$$

The parameter $C > 0$ controls the trade-off between the slack variable (ξ_i) penalty and margin. If at any point $\xi_i > 1$, the data the data are located on the wrong side and misclassified. If at any point $0 < \xi_i \geq 1$ the data are located inside the margin and still on the right side. If at any point $\xi_i = 0$, the data are located in a line parallel with the hyperplane and still on the right side. Equation (8) is complicated to solve, and this problem is more easily solved if it is converted into a dual form using Lagrange multipliers (α_x) so the optimization problem can be changed to

$$min \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_i y_i y_i (x_i . x_j) - \sum_{i=1}^{N} \alpha_i \quad (10)$$

$$s.t. \sum_{i=1}^{N} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1,2,\ldots,N$$

Table 5. Feature selection based on the artificial bee colony (ABC) algorithm

Input	Total Samples
	1. Initialize the population of solutions $X_i^J, i = 1,2, \dots, SM, i = 1,2, \dots, D$ using equation (2) 2. Calculate the fitness value of each food source 3. Iteration = 1 4. While Iteration < Maximal Iteration a. Produce new solutions V_i^J for the employed bee by using equation (3) b. Calculate the fitness value of each food source c. Apply the greedy selection process by equation (6) d. Calculate the probability values for the solution X_i^J by equation (7) e. Produce the new solution V_i^J for the onlookers from the solutions X_i^J selected depending on f. Calculate the fitness value of each food source g. Apply the greedy selection process by equation (6) h. Determine the abandoned solution for the scout, and replace it with a new randomly produced solution X_i^J by equation (3) i. Memorize the best solution achieved so far 5. Iteration = 1 + iteration 6. End While
Output	features with highest features

The weight vector is expressed as $w = \sum_{i=1}^{N} \alpha_i y_i x_i$, and the bias parameter is expressed as

$$b = \frac{1}{N_s} \sum_{i \in S} (t_i - \sum_{m \in S} \alpha_m t_m x_m^T . x_i),$$

Where S is a support vector index set and x_m is a support vector

4. Results and Discussion

The accuracy of lung cancer classification was calculated using the confusion matrix using different training and testing sets and is given as a percentage. The program were built on a computer with following specification processor Intel® Core™ i5-7200U CPU @2.50 GHz and RAM 4 GB in Windows 10. Running time is the time required to simulate a program, calculated using tic toc in Matlab R2017b, and is given in seconds. Four simulations were conducted: classification using SVM without feature selection, classification using SVM with SVM-RFE as a feature selection method, classification using SVM with ABC as a feature selection method, and classification using SVM with RFE-ABC as a feature selection method.

4.1. Simulation Results

First, classification was conducted using SVM without feature selection using cross validation. Table 4 shows the accuracy and running times of various k partitions in cross

validation for the Michigan lung cancer dataset. Table 5 shows the accuracy and running times of various k partitions in cross validation for the Ontario lung cancer dataset.

Table 6. SVM without features selection for the Michigan lung cancer dataset

k	Accuracy (%)	Running time (s)
3	85.878	4.251
4	85.720	5.705
5	86.863	7.400
6	86.760	8.831
7	85.953	10.352
8	88.661	12.069
9	88.360	13.433
10	89.022	15.201

Table 7. SVM without features selection for the Ontario lung cancer dataset

k	Accuracy (%)	Running time (s)
3	66.154	0.899
4	68.000	0.916
5	67.286	1.072
6	65.079	1.103
7	68.993	1.112
8	69.125	1.149
9	69.111	1.158
10	69.867	1.230

For the Michigan lung cancer dataset, the highest accuracy (89.022%) was obtained using k = 10 with running time of 15.201 s, whereas the fastest running time (4.251 s) was obtained using k = 3. For the Ontario lung cancer dataset, the highest accuracy for classification without feature selection (69.867%) was obtained using k = 10 with a running time of 1.230 s. The fastest running time for classification without feature selection (0.899 s) was obtained using k = 3. According to these results, k = 10 was used for cross validation in all further feature selection.

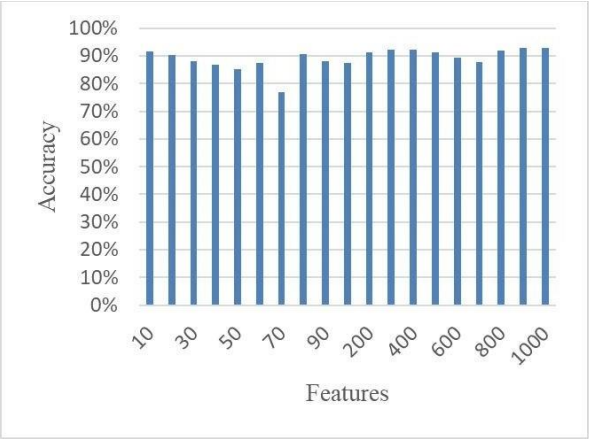

Fig. 2. Average accuracy of cancer classification for the Michigan lung cancer dataset with SVM-RFE feature selection

The second simulation was classification using SVM with SVM-RFE as the feature selection method. Table 6 shows the accuracy and running times of classification in this simulation for the Michigan lung cancer dataset. Table 7 shows the accuracy and running times of classification in this simulation for the Ontario lung cancer dataset. Figure 2 and Figure 3 shows the average accuracy of classification in this simulation for the Michigan lung cancer dataset and Ontario lung cancer dataset.

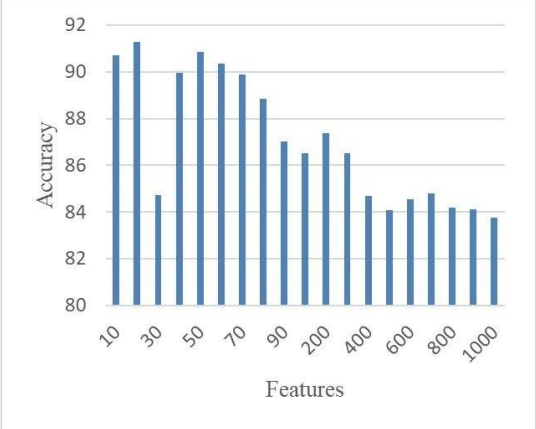

Fig. 3. Average accuracy of cancer classification for the Ontario lung cancer dataset with SVM-RFE feature selection

Table 8. SVM with SVM-RFE as feature selectio for The Michigan lung cancer dataset with linear kernel and 10- fold cross validation

Features	Accuracy (%)	Running time (s)
10	91.467	42.591
20	90.200	55.156
30	87.978	66.515
40	86.689	82.997
50	85.044	97.491
60	87.533	109.117
70	76.756	120.209
80	90.511	140.238
90	87.911	154.055
100	87.267	176.420
200	91.311	339.327
300	92.178	478.984
400	92.244	650.929
500	91.336	774.999
600	89.222	936.834
700	87.800	1098.192
800	91.800	1256.116
900	92.711	1431.470
1000	92.911	1558.052

For the Michigan lung cancer dataset, the highest accuracy of classification using SVMRFE (2.911%) was obtained using k = 10 and 1000 features with a running time of

1558.98 s. For the Ontario lung cancer dataset, the highest accuracy of classification with SVM-RFE as the feature selection method (91.267%) was obtained using 20 features with a running time 1.428 s.

The third simulation involved classification of SVM with ABC as the feature selection method. For the Michigan lung cancer dataset, the highest accuracy of classification using ABC (94.778%) was obtained using 60 features. For the Ontario lung cancer dataset, the highest accuracy of classification using ABC (93%) was obtained using 60 features.

Figure 4 shows the average accuracy of cancer classification for the Michigan lung cancer dataset with ABC feature selection. Figure 4 shows the average accuracy of cancer classification for the Ontario lung cancer dataset with ABC feature selection.

The last simulation used SVM with RFE-ABC as the feature selection method. Figure 5 shows the average accuracy of cancer classification for the Michigan lung cancer dataset in this simulation and shows the average accuracy of cancer classification for the Ontario lung cancer dataset in this simulation.

Table 8 shows the results of research conducted using linear kernel, 10-fold cross validation, and SVM classification methods for the Michigan lung cancer dataset. Table 9 shows the results of research conducted using linear kernel, Table 10, Table 11 10-fold cross validation, and SVM classification methods for the Ontario lung cancer dataset

Table 9. SVM with SVM-RFE as feature selectio for The Ontario lung cancer dataset with linear kernel and 10-fold cross validation

Features	Accuracy (%)	Running time (s)
10	90.700	1.357
20	91.267	1.428
30	84.733	1.504
40	89.967	1.582
50	90.833	1.635
60	90.333	1.699
70	89.900	1.803
80	88.833	1.887
90	87.033	1.930
100	86.500	2.021
200	87.367	2.956
300	86.533	4.200
400	84.700	5.333
500	84.100	6.640
600	84.533	8.202
700	84.800	9.434
800	84.200	11.133
900	84.133	12.824
1000	83.767	14.876

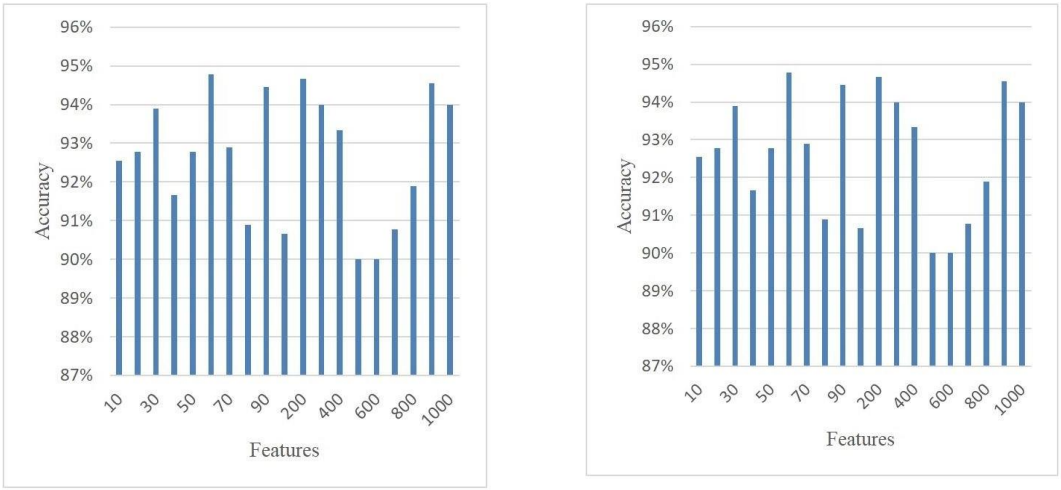

Fig. 4. Average accuracy for cancer classification for the Michigan lung and Ontario lung cancer dataset with ABC feature selection

Table 10. Accuracy and running times of classification for the Michigan lung cancer dataset with linear kernel and 10-fold cross validation

Feature selection methods	**Michigan lung cancer dataset**		
	Accuracy (%)	***Running time(s)***	***Total feature***
Without feature selection	89.022	15.201	7129
SVM-RFE	92.911	1558.98	1000
ABC	94.778	155101.357	60
RFE-ABC	98.000	21672.982	1100

Table 11. Accuracy and running times of classification for the Ontario lung cancer dataset with linear kernel and 10- fold cross validation

Feature selection methods	**Michigan lung cancer dataset**		
	Accuracy (%)	***Running time(s)***	***Total feature***
Without feature selection	69.867	1.230	2880
SVM-RFE	91.267	1.428	20
ABC	93.000	2273.98	60
RFE-ABC	97.000	1914.715	70

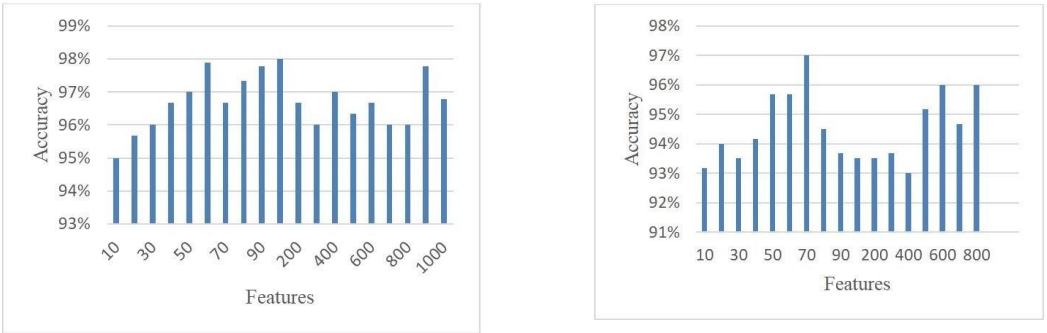

Fig. 5. Average accuracy of cancer classification for the Michigan lung and Ontario lung cancer dataset with RFE-ABC feature selection

For the Michigan lung cancer dataset, the highest accuracy using RFE-ABC (98.000%) was obtained using 100 features. For the Ontario lung cancer dataset, the highest accuracy using RFE-ABC (97.000%) was obtained using 70 features.

5. Conclusion

Based on the simulations, feature selection provides higher classification accuracy and SVM with RFE- ABC as the feature selection method gives us more accurate result to diagnose Lung cancer using microarray data than SFM-RFE and ABC.

Acknowledgment

This research was supported by a PDUPT 2018 research grant from Ministry of Research and Technology and Directorate of Higher Education.

References

[1] T. M. St John, "With Every Breath," *A Lung Cancer Guid. Vancouver*, 2003.

[2] J. M. Luna *et al.*, "Predicting radiation pneumonitis in locally advanced stage II–III non-small cell lung cancer using machine learning," *Radiother. Oncol.*, vol. 133, pp. 106–112, 2019, doi: https://doi.org/10.1016/j.radonc.2019.01.003.

[3] B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," *Eur. J. Oper. Res.*, vol. 265, no. 3, pp. 993–1004, 2018, doi: https://doi.org/10.1016/j.ejor.2017.08.040.

[4] J. A. Tsou, J. A. Hagen, C. L. Carpenter, and I. A. Laird-Offringa, "DNA methylation analysis: a powerful new tool for lung cancer diagnosis," *Oncogene*, vol. 21, no. 35, pp. 5450–5461, 2002.

[5] M. A. Khan *et al.*, "Lungs cancer classification from CT images: An integrated design of contrast based classical features fusion and selection," *Pattern Recognit. Lett.*, vol. 129, pp. 77–85, 2020, doi: https://doi.org/10.1016/j.patrec.2019.11.014.

[6] M. Wiesweg *et al.*, "Machine learning reveals a PD-L1–independent prediction of response to immunotherapy of non-small cell lung cancer by gene expression context," *Eur. J. Cancer*, vol. 140, pp. 76–85, 2020, doi: https://doi.org/10.1016/j.ejca.2020.09.015.

[7] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020, doi: https://doi.org/10.1016/j.neucom.2019.10.118.

[8] J. A. ALzubi, B. Bharathikannan, S. Tanwar, R. Manikandan, A. Khanna, and C. Thaventhiran, "Boosted neural network ensemble classification for lung cancer disease diagnosis," *Appl. Soft Comput.*, vol. 80, pp. 579–591, 2019, doi: https://doi.org/10.1016/j.asoc.2019.04.031.

[9] P. A. Mundra and J. C. Rajapakse, "SVM-RFE with MRMR filter for gene selection," *IEEE Trans. Nanobioscience*, vol. 9, no. 1, pp. 31–37, 2009.

[10] G. Wang, G. Zhang, K.-S. Choi, K.-M. Lam, and J. Lu, "Output based transfer learning with least squares support vector machine and its application in bladder cancer prognosis," *Neurocomputing*, vol. 387, pp. 279–292, 2020, doi: https://doi.org/10.1016/j.neucom.2019.11.010.

[11] Q. Gu *et al.*, "Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer," *Eur. J. Radiol.*, vol. 118, pp. 32–37, 2019, doi: https://doi.org/10.1016/j.ejrad.2019.06.025.

[12] K. R. Kavitha, G. S. Rajendran, and J. Varsha, "A correlation based SVM-recursive multiple feature elimination classifier for breast cancer disease using microarray," in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2016, pp. 2677–2683.

[13] K.-B. Duan, J. C. Rajapakse, H. Wang, and F. Azuaje, "Multiple SVM-RFE for gene selection in cancer classification with expression data," *IEEE Trans. Nanobioscience*, vol. 4, no. 3, pp. 228–234, 2005.

[14] Y. Tang, Y.-Q. Zhang, Z. Huang, X. Hu, and Y. Zhao, "Recursive fuzzy granulation for gene subsets extraction and cancer classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 6, pp. 723–730, 2008.

[15] F. Yuan, L. Lu, and Q. Zou, "Analysis of gene expression profiles of lung cancer subtypes with machine learning algorithms," *Biochim. Biophys. Acta - Mol. Basis Dis.*, vol. 1866, no. 8, p. 165822, 2020, doi: https://doi.org/10.1016/j.bbadis.2020.165822.

[16] Y. Huo, L. Xin, C. Kang, M. Wang, Q. Ma, and B. Yu, "SGL-SVM: A novel method for tumor classification via support vector machine with sparse group Lasso," *J. Theor. Biol.*, vol. 486, p. 110098, 2020, doi: https://doi.org/10.1016/j.jtbi.2019.110098.

[17] A. Bustamam, A. Bachtiar, and D. Sarwinda, "Selecting Features Subsets Based on Support Vector Machine-Recursive Features Elimination and One Dimensional-Naïve Bayes Classifier using Support Vector Machines for Classification of Prostate and Breast Cancer," *Procedia Comput. Sci.*, vol. 157, pp. 450–458, 2019, doi: https://doi.org/10.1016/j.procs.2019.08.238.

[18] P. K R and N. N C, "Lung Cancer Survivability Prediction based on Performance Using Classification Techniques of Support Vector Machines, C4.5 and Naive Bayes Algorithms for Healthcare Analytics," *Procedia Comput. Sci.*, vol. 132, pp. 412–420, 2018, doi: https://doi.org/10.1016/j.procs.2018.05.162.

[19] Y. Xu, P. Fan, and L. Yuan, "A simple and efficient artificial bee colony algorithm," *Math. Probl. Eng.*, vol. 2013, 2013.

[20] X. Yu, J. Zhang, J. Fan, and T. Zhang, "A faster convergence artificial bee colony algorithm in sensor deployment for wireless sensor networks," *Int. J. Distrib. Sens. Networks*, vol. 9, no. 10, p. 497264, 2013.