Journal of Radiation Research and Applied Sciences 15 (2022) 104–110

Contents lists available at ScienceDirect

Journal of Radiation Research and Applied Sciences

journal homepage: www.journals.elsevier.com/journal-of-radiation-research-and-applied-sciences

A wrapper-based feature selection approach to investigate potential biomarkers for early detection of breast cancer

Majdi R. Alnowami [a,*], Fouad A. Abolaban [a], Eslam Taha [b]

[a] Department of Nuclear Engineering,Faculty of Engineering, King Abdulaziz University, PO Box 80204, Jeddah, 21589, Saudi Arabia
[b] Center for Training and Radiation Prevention, King Abdulaziz University, P.O. Box 80204, Jeddah, 21589, Saudi Arabia

ARTICLE INFO

Keywords:
Breast cancer
Biomarkers
Feature ranking
Classification

ABSTRACT

Breast cancer (BC) biomarkers can radically improve the early detection in patients and, as a result, reduce mortality rate, whether for detecting individuals at increased risk of developing cancer or in the screening process. Finding a successful biomarker for breast cancer would be a fast and low-cost first solution to predicting BC, and it could potentially lead to a decline in the global BC mortality rate. However, biomarker exploration translates into the role of feature ranking and selection in machine learning terminology. This study explores the influence of using a particular biomarker or combinations of different biomarkers as predictors for breast cancer. Three different classification algorithms were integrated with a sequential backward selection model: support vector machine (SVM), random forests (RF), and Decision Trees (DTs). The result shows that the optimal set of biomarkers comprises Glucose, Resistin, homo, BMI, and Age using the SVM model. The sensitivity and specificity were 0.94 and 0.90, respectively and the 95% confidence interval for the AUC was [0.89, 0.98]. The result indicates that Glucose, Resistin, homo, BMI, and Age combined can serve as a crucial BC biomarker in BC screening and detection.

1. Introduction

Despite recent technological developments in diagnostic radiology, breast cancer detection remains a persistent challenge. According to recent statistical data, breast cancer is the most common cancer among females and the second-highest contributor to cancer mortality among humans after lung cancer (Siegel et al., 2018; Stewart and Wild-Organization, 2014). The reduction of mortality rate is associated with early detection, which can be achieved through screening programs (Lauby-Secretan et al., 2015; Rue et al., 2009). Early detection of breast cancer relies heavily on screening programs. Mammography is by far the most effective way to detect breast cancer at an early stage. However, socioeconomic limitations, and the geographical distribution and quality of mammography machines, can restrict access to mammographic screening in any given area. These factors are not directly under control, and sometimes they are tough to overcome. These practical challenges motivate researchers and health care providers to start searching for an alternative solution. That search leads to the exploration of Breast cancer biomarkers. Breast cancer biomarkers have been studied to improve early detection and reduce mortality. Biomarkers not only can be indicative of cancer presence (Kuppusamy et al., 2017; Levenson, 2007; Singh, 2019), but also of patient responsiveness to proposed treatment (Nicolini et al., 2018; Al-Khater et al., 2021). Biomarkers have been used to develop predictive models (Assiri & Kamel, 2016; Patrício et al., 2018; Santillán-Benítez et al., 2013). These models are developed by evaluating different biomarkers to find which single or combination of biomarkers is more indicative of breast cancer.

Researchers evaluated various biomarkers, and several predictive models were developed. Hwa et al. used logistic regression models to evaluate serum biomarkers' effectiveness in the early detection of breast cancer (Carneiro et al., 2020). Samples obtained showed Tissue polypeptide specific antigen (TPS) to provide the highest predictive value with 80% sensitivity. In comparison, insulin-like growth factor binding protein-3 (IGFBP-3) and breast cancer-specific cancer antigen 15.3 (CA15-3) sensitivity were reported to be 65% and 56%, respectively. Incorporating CA15-3 and IGFBP-3 with TPS showed the best multivariate logistic regression model with a sensitivity of 85%. Several studies evaluated adipokines for breast cancer, with many reporting higher leptin, resistin, and visfatin among confirmed cancer cases than patients with benign lesions and healthy subjects (Cust et al., 2009; Wu

Peer review under responsibility of The Egyptian Society of Radiation Sciences and Applications.
* Corresponding author.
E-mail address: malnowaimi@kau.edu.sa (M.R. Alnowami).

https://doi.org/10.1016/j.jrras.2022.01.003
Received 14 October 2021; Received in revised form 12 December 2021; Accepted 13 January 2022
Available online 3 March 2022

et al., 2009). Pham et al. evaluated nine biomarker predictors and proposed a novel generalized logistic dependent model (Pham & Pham, 2020). The novel model outperformed other models considered in his study, such as random forest and multiple logistic regression. It showed an accuracy ranging from 88.7% (trained with 100% of the dataset) to 100% (trained with 70% of the dataset) when used with Glucose, Age, Resistin, BMI, and MCP-1 as predictors.

Some researchers used imaging Biomarkers such as thermal imaging combined with various feature selection and classification methods for breast cancer detection. Zarei et al. evaluated a hybrid of Gaussian Mean Shift (GMS) and roulette wheel selection approach-based segmentation method for breast cancer detection using infrared thermal images. The suggested method showed to improve the performance of the Computer-Aided-Detection (CAD) system in comparison with other CAD systems using Mean Shift (MS), and Fuzzy C-Means (FCM) segmentation algorithms (Zarei et al., 2021). Darabi et al. presented a thermogram-based CAD system to detect breast cancer. Their CAD system used SVM and kNN algorithms for classification and the Random Subset Feature Selection (RSFS), the hybrid of minimum Redundancy Maximum Relevance (mRMR), and Genetic (GA) with RSFS algorithms for feature selection. Results showed RSFS with kNN and have accuracy and sensitivity of 85.36% and 94.11%, respectively (Darabi et al., 2021). However, this study focus on selecting Non-Imaging Biomarker.

Screening programs and biomarkers' effectiveness and accuracy in early cancer detection were sought to be improved by incorporating artificial intelligence techniques (McKinney et al., 2020; Shen et al., 2019). The successful application of artificial intelligence techniques on predictive modeling, object detection, and classification (Krizhevsky et al., 2017; Zhang et al., 2015) drew attention to the possibility of implementing such methods in medical diagnostic applications to improve diagnosis accuracy and reduce recall rate. The introduction of deep learning methods into mammography showed promising results in detecting and classifying breast lesions (Carneiro et al., 2015; Dhungel et al., 2015; Peng et al., 2016). Furthermore, breast cancer biomarkers have been used to develop predictive models(Assiri & Kamel, 2016; Patrício et al., 2018; Santillán-Benítez et al., 2013).

Features identified in the literature for breast cancer varied from one work to another, even when the same data is used(Aslan et al., 2018; Ghani et al., 2019; Li & Chen, 2018; Patrício et al., 2018; Rahman et al., 2020; Santillán-Benítez et al., 2013; Silva Araújo et al., 2019). This variation motivates us to investigate the factors that affect biomarkers ranking and selection techniques. This study focuses on age, BMI, Glucose, Insulin, HOMA, Leptin, Adiponectin, L/A ratio, resistin, and MCP-1. These biomarkers have been used in several studies (Aslan et al., 2018; Ghani et al., 2019; Li & Chen, 2018; Patrício et al., 2018; Rahman et al., 2020; Santillán-Benítez et al., 2013; Silva Araújo et al., 2019). As mentioned above, these studies reflect some level of disagreement on which single or combination of biomarkers is the best indicator for breast cancer.

Table 1
The quantitative features of patients and healthy controls in terms of their (mean ± standard deviation) and medians(interquartile) ranges.

Biomarkers	Median (interquartile range)		Mean (Standard deviation)	
	Controls	Patients	Controls	Patients
Age (years)	65 (34.5)	53 (23)	58.08 ± 18.96	56.67 ± 13.49
BMI (kg/m2)	27.69 (9.32)	27.41 (8.07)	28.32 ± 5.43	26.98 ± 4.62
Glucose (mg/dL)	87 (11)	98.5 (18)	88.23 ± 10.19	105.56 ± 26.56
Insulin (μU/mL)	5.48 (2.92)	7.58 (11.84)	6.93 ± 4.86	12.51 ± 12.32
HOMA	1.14 (0.93)	2.05 (3.44)	1.55 ± 1.22	3.62 ± 4.59
Leptin (ng/mL)	21.49 (26.19)	18.88 (25.18)	26.64 ± 19.33	26.6 ± 19.21
Adiponectin (μg/mL)	8.13 (5.63)	8.45 (6.93)	10.33 ± 7.63	10.06 ± 6.19
L/A	2.36 (4.75)	2.33 (3.18)	4.05 ± 4.06	3.91 ± 4.77
Resistin (ng/mL)	8.93 (6.36)	14.37 (14.97)	11.61 ± 11.45	17.25 ± 12.64
MCP-1(pg/dL)	471.32 (393.29)	465.37 (440.69)	499.73 ± 292.24	563.02 ± 384

2. Feature ranking

Biomarker exploration translates into the role of feature ranking and selection in machine learning terminology. The objective of using feature ranking and selection is to analyze the impact of using a specific feature or combinations of a different feature as predictors in the classification model's accuracy. In some cases, all features may not be helpful for some classification problems. Some features may be irrelevant, redundant, and noisy. Moreover, even when all features are relevant and contain information about the response variable, using a high number of features may negatively impact prediction accuracy(Iguyon & Elisseeff, 2003). Thus, the feature used to construct a specific classifier considerably impacts the classifier's accuracy, sensitivity, and computational cost. Therefore, feature ranking and selection aim to eliminate redundant or irrelevant features and reduce the feature dimension. Although both analysis processes, feature ranking and feature selection, are related, they target different outcomes. Feature ranking techniques rank each feature individually according to some decisive factors such as feature variance and feature relevance to the response. There are several types of feature ranking techniques, such as ranking the feature using the p-values of either the chi-square test statistics or the F-test statistic. The Mann–Whitney *U* test, a non-parametric test for equality of population medians of two separate samples, is another ranking technique. It is a dependence test such as a *t*-test, F-test. However, unlike the *t*-test and F-test, the Mann-Whitney *U* test is non-parametric. Considerable disagreements were highlighted in the rank of the features using different ranking methods(Wang et al., 2001; Yang & Mao, 2010). Feature in top positions in some ranks may appear at the bottom positions in the other ranks.

Feature ranking can be considered as a subset of the feature selection process. The feature selection method aims to eliminate irrelevant or redundant features. In other words, the aim is to find a subset of the feature set that can be used to train the classifier and improve accuracy and reduce execution time. However, all possible subset is computationally impracticable. Our study tested all possible subsets required to generate $(2^{10}) - 1 = 1024$ different sets of features and models. Current feature selection approaches fall into two general categories filter-based and wrapper-based algorithms. The filter-based assess the significance of features by their scores in various statistical tests for their correlation with the outcome variable. The feature subset selection is based on a user-specified threshold, which requires the user to choose an arbitrary cutoff on the number of features chosen. Moreover, Filter-based algorithms stand on the assumption that features with a higher variance may contain more useful information without considering the relationship between features (Iguyon & Elisseeff, 2003; Petković, Kocev, & Džeroski, 2019). Wrapper-based algorithms assess the significance of all possible feature subsets based on the performance of the classifying model. Based on each feature's impact on the classification model accuracy, the feature is added or removed from the feature subset. Wrapper-based approaches regularly reach better classification accuracy than the filters-based approach because the feature selection process is optimized for a particular classification model. However, wrappers are computationally expensive because they use a classifying algorithm to evaluate every subset of features. Turning the feature selection into a sequential decision process can significantly reduce the computational time and overcome the wrapper-based algorithm's drawback. Sequential feature selection is one of the most broadly used feature selection methods. It selects a feature subset by sequentially adding (forward search) or removing (backward search) until certain

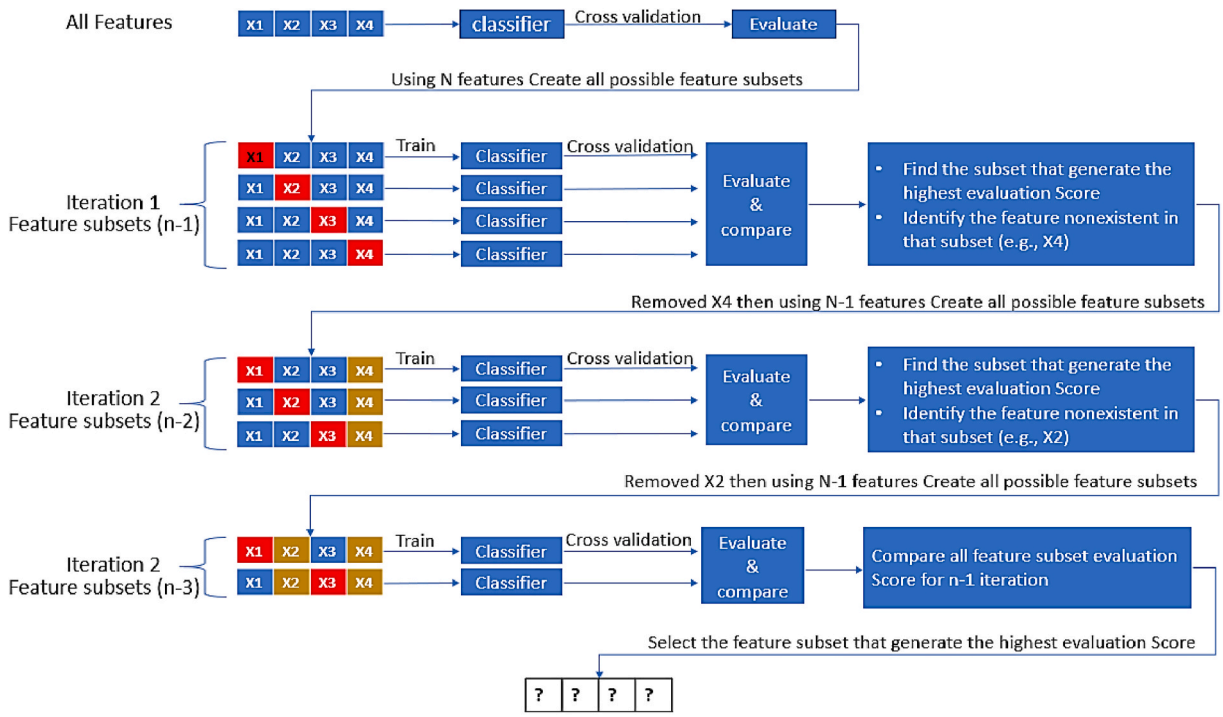

Fig. 1. Sequential backward selection (SBS) method: First, all features combined were used to train the classifier. Then, the classifier will be trained on each subset of n-1 features. The subset with the lowest predictive accuracy is removed. This process is repeated until the first local minimum of the cross-validation MCE is found or the model trained on all the possible candidate feature subset.

stopping conditions are reached. The objective of this study is to explore the influence of using a specific feature or combinations of different features as predictors for breast cancer, using the sequential feature selection method in the classification model's accuracy (Aggrawal & Pal, 2020).

3. Materials and methods

3.1. Data

The data was collected by the Gynecology Department of the University Hospital Centre of Coimbra (CHUC) between 2009 and 2013 (Crisostomo et al., 2016; Patrício et al., 2018). Patient data were for women diagnosed with breast cancer (BC) and before surgery and treatment. Healthy volunteers were nominated and participate in the study as controls. A total of 64 women with breast cancer (BC) and 52 healthy volunteers was included in the study. Collected data included age, weight, height, and Blood samples. Blood were taken to determine biochemical parameters in all participants. For each participant, multiple clinical features were examined or evaluated, including age, BMI, Glucose, Insulin, HOMA, Leptin, Adiponectin, L/A ratio, Resistin, and MCP-1. More details about the data collection process are available (Crisostomo et al., 2016; Patrício et al., 2018). Table 1 presents collected data for patients and controls in terms of their (mean± standard deviation) and medians(interquartile) ranges.

3.2. Feature selection and machine learning classifiers

In this study, we used the Sequential backward selection (SBS) method (Theodoridis & Koutroumbas, 2009; Wang et al., 2001; Aggrawal & Pal, 2020). As illustrated in Fig. 1, First, all features combined were used to train the classifier. Then, the classifier will be trained on each subset of n-1 features. The subset with the lowest predictive accuracy, i.e., higher misclassification rate, is removed. In another way the feature nonexistent in the feature subset that shows the highest performance is removed. The process will be repeated for the remaining subset of feature (n-1), the classifier will be trained on each subset of n-2 and then one feature will be eliminated based on the predictive accuracy. This process is repeated until the first local minimum of the cross-validation MCE is found or the model trained on all the possible candidate feature subset based on the criteria mentioned above. 10-fold cross-validation was applied to the training set to evaluate and compare the performance of each classifier along with the anticipated feature subset.

Three different classification algorithms were integrated with the Sequential backward selection model: support vector machine (SVM) (Ma & Guo, 2014), random forests (RF) (Cutler et al., 2012), and Decision Trees (DTs) (Rokach & Maimon, 2008; Safavian & Landgrebe, 1991). The development of this method employs 10-fold Monte Carlo Cross-Validation (MCCV) to ensure stability. Moreover, to avoid the resubstitution error, the data was randomly partitioned into two groups: 70% training set and 30% testing set. The training and test dataset is created with equal distribution.

3.3. Performance metrics

Several widely used statistical measures to evaluate classifiers' performances were computed. The list below provides insights into a few basic concepts and measurement procedures that summarize these metrics.

- Confusion matrix (CM) confusion matrix (CM) is such that CM_i^j is equal to the number of observations in group I and classified to be in group j.
  - True Positive (TP): CM_1^1, True Positive (TP): when the model predicted as Positive, and they were Positive.
  - True Negative (TN): CM_2^2, True Negative (TN): when the model predicted as Negative, and they were Negative.
  - False Positive (FP):CM_2^1, False Positive (FP): when the model predicted as Positive, but they were Negative.

Table 2
Presents the selected feature based on performance in each iteration for each classification model (SVM, RF, DTs).

Classifier		SVM										RF										Dts									
Itration		0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
Feature	Glucose (mg/dL)	X	X	X	X	X	**X**	X	X	X	X	X	X	X	X	**X**	X	X	X	X	X	X	X	X	X	X	X	**X**	X	X	X
	Resistin (ng/mL)	X	X	X	X	X	**X**	X	X	X	-	X	X	X	X	**X**	X	X	X	X	-	X	X	X	X	X	X	**X**	X	X	-
	HOMA	X	X	X	X	X	**X**	-	-	-	-	X	X	X	X	-	-	-	-	-	-	X	X	X	X	-	-	-	-	-	-
	Insulin (μU/mL)	X	X	X	-	-	-	-	-	-	-	X	X	X	-	-	-	-	-	-	-	X	X	X	-	-	-	-	-	-	-
	BMI (kg/m2)	X	X	X	X	X	**X**	X	X	-	-	X	X	X	X	**X**	X	X	X	-	-	X	X	X	X	X	X	**X**	-	-	-
	Age (years)	X	X	X	X	X	**X**	X	-	-	-	X	X	X	X	**X**	X	X	-	-	-	X	X	X	X	X	X	**X**	-	-	-
	MCP-1(pg/dL)	X	X	X	X	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-
	Adiponectin (μg/mL)	X	X	X	X	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-	-	X	X	-	-	-	-	-	-	-	-
	L/A	X	X	-	-	-	-	-	-	-	-	X	X	X	X	**X**	X	-	-	-	-	X	X	X	X	X	X	-	-	-	-
	Leptin (ng/mL)	X	-	-	-	-	-	-	-	-	-	X	X	X	X	**X**	-	-	-	-	-	X	X	X	X	X	-	-	-	-	-

– False Negative: CM_1^2, False Negative: when the model predicted as Negative, but they were Positive.

- Misclassification rate (%) (MCE): The number of misclassified observations divided by the number of observations; $MCE = \frac{FP+FN}{TP+TN+FP+FN}$
- Sensitivity (Se) also be referred to as the recall, hit rate, or TPR (True Positive Rate): The proportion of actual negatives that are correctly identified as positives; $Se = \frac{TP}{TP+FN}$
- Specificity (Sp) also be referred to asTrue Negative Rate (TNR): The proportion of actual negatives that are correctly identified as negatives; $Sp = \frac{TN}{TN+FP}$
- Receiver Operating Characteristic (ROC) curve and Area Under the ROC Curve (AUC): ROC is a probability curve and AUC represents degree or measure of separability. The ROC curve is plotted with TPR (True Positive Rate) against the FPR (False Positive Rate) where FPR = 1-Sp. TPR is on y-axis and FPR is on the x-axis.
- Precision (Pr) or Positive predictive value (PPV): The ratio of true positives to all predicted positives; $Pr = \frac{TP}{TP+FP}$
- F-score or F measure: The harmonic mean of the precision and recall; $Pr = 2 \cdot \frac{Pr \cdot Se}{Pr+Se}$

4. Results

As mentioned earlier, the data was randomly partitioned into two groups: 70% training set and 30% testing set. The training set and the Sequential backward selection (SBS) method were used to select the features and train the classifier. The test set was used to evaluate the selected feature subset and classifiers' performances. 10-fold cross-validation was applied to the training set to evaluate and compare each classifier's performance along with the anticipated feature subset. Table 2 presents the selected feature based on performance in each iteration for each classification model (SVM, RF, DTs). The x sign represents the features selected at that specific iteration, where the **X** represents the selected features that achieved minimum cross-validation (MCE).

To present the number of optimal features to train the classifier, we plotted the MCE on the test set as a function of the number of features illustrated in Fig. 2. The figure shows the cross-validation MCE as a function of the number of features in each iteration. The X-axis represents the number of features removed from the feature set. The Y-axis represents the best cross-validation MCE in each model's iteration. Fig. 2 illustrates that the SVM model reached the minimum cross-validation MCE value when five features were eliminated. The curve stays slightly flat for the RF model over the range from 4 features to 5 features. However, the DTs model reached the minimum cross-validation MCE value when six features were eliminated. The figure illustrated that the SVM classification model using five features achieved the minimum cross-validation MCE compared to other models and other feature subsets. In general, four features confirmed necessary: Glucose, Resistin, BMI, and Age, three features confirmed unimportant: Insulin, MCP-1, and Adiponectin, and three exploratory features: HOMA L/A and Leptin. The table shows that Glucose, Resistin, BMI, and Age were common optimal features across the three classifiers with an additional HOMA for SVM and L/A and Leptin for Dts.

To further investigate the impact of each feature subset on the overall performance of the classifier (SVM, RF, DTs) using the test set, a 95% confidence interval for the AUC was plotted as illustrated in Fig. 3. The figure shows that SVM produces a 95% confidence interval for the AUC [0.89, 0.98] when five features were eliminated. The best 95% confidence interval for the AUC for RF and Dts were [0.81, 0.94] and [0.88, 0.72] respectively.

Finally, to evaluate the performance of the finally selected optimal feature subset for each classifier, test data was used to train each model, and the ROC curve was plotted as illustrated in Fig. 4. The ROC curves show that using the selected optimal features, Glucose, Resistin, homo,

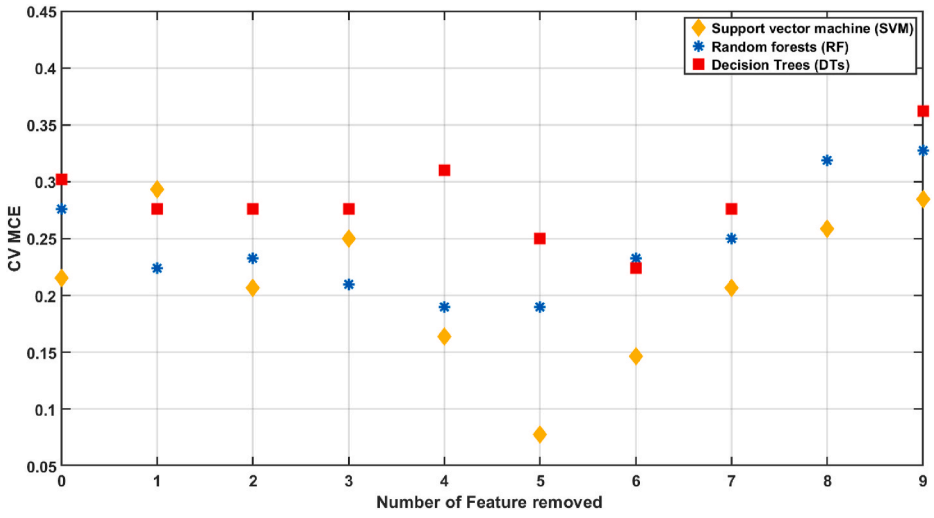

Fig. 2. The cross-validation MCE as a function of the number of features in each iteration. The X-axis represents the number of features removed from the feature set. The Y- axis represents best achieved the cross-validation MCE in each iteration for each models.

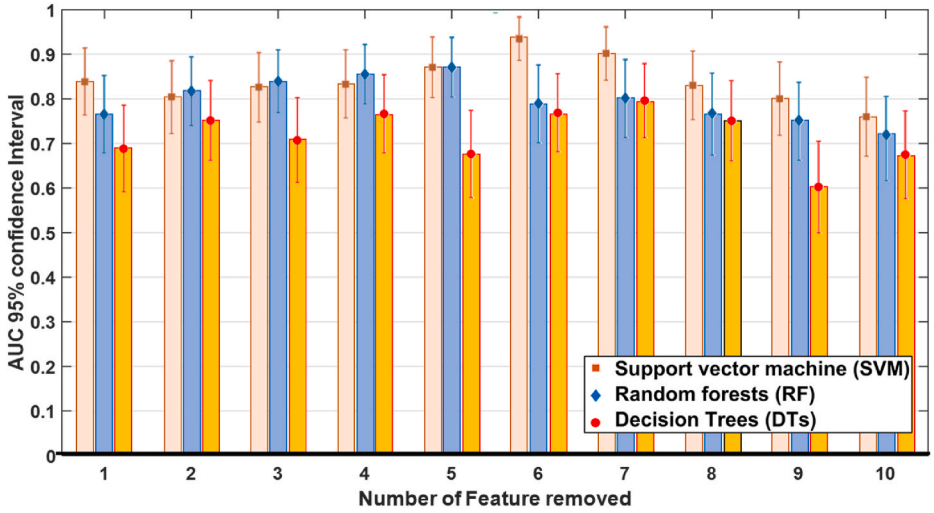

Fig. 3. The values of AUC and the respective 95% confidence intervals (CI) as a function of the number of features in each iteration. The X-axis represents the number of features removed from the feature set. The Y- axis AUC value.

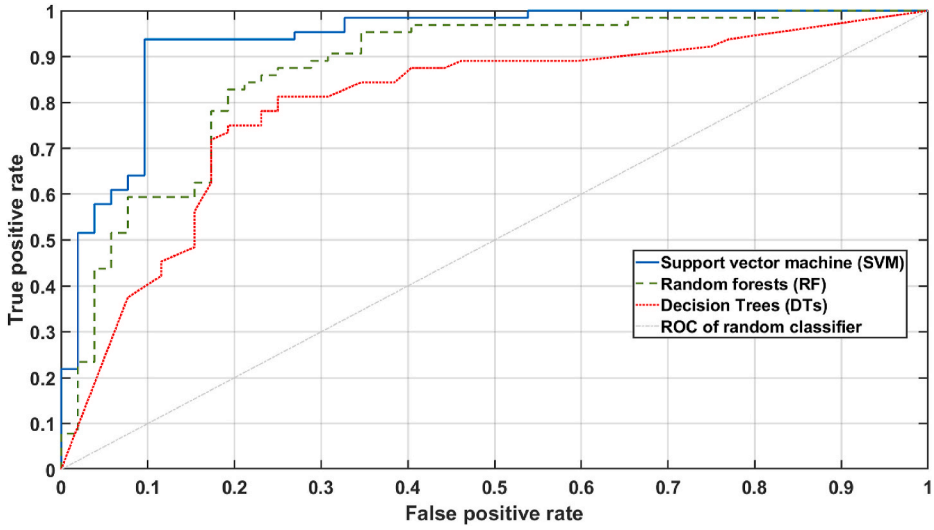

Fig. 4. The ROC curve of the finally selected optimal feature subset for each classifier.

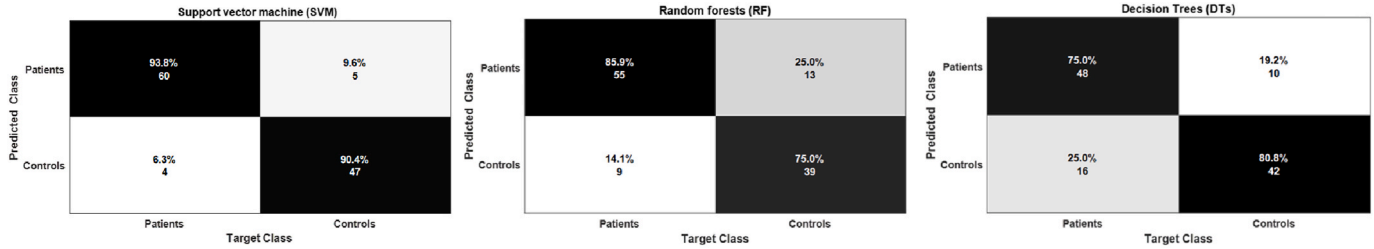

Fig. 5. The confusion matrix for all classifiers (SVM, RF, DTs) using Leave-One-Out Cross-Validation and trained with the associated optimal selected feature subset. According to the confusion matrix, 4 patients were classified incorrectly as control, whereas five control were classified as a patient using the SVM classifier.

Table 3
The performance metric for all classifiers (SVM, RF, DTs) using Leave-One-Out Cross-Validation and the optimal selected features associated with the classifier.

Evaluation metrics	Model		
	SVM	RF	Dts
AUC 95% CI	[0.89, 0.98]	[0.80,0.94]	[0.73, 0.88]
Accuracy	0.92	0.81	0.78
Sensitivity (Se)	0.94	0.86	0.75
Specificity	0.90	0.75	0.81
Precision	0.92	0.81	0.83
F_measure	0.93	0.83	0.79
Selected Features	Glucose, Resistin, HOMA, BMI, and Age	Glucose, Resistin, BMI, Age and Leptin	Glucose, Resistin, BMI and Age

BMI, and Age, the SVM model performance for classifying the data is superior. The true positive rate is higher, and the false positive rate is lower than for RF and Dts models at all cut-offs. The area under the curve for SVM is larger than the area under the curve for RF and DTs. Fig. 5 shows the confusion matrix for all classifiers using Leave-One-Out Cross-Validation. Regarding the confusion matrix, four patients were classified incorrectly as control, whereas five control were classified as a patient using the SVM classifier.

Table 3 summarizes the performance metric for all classifiers (SVM, RF, DTs) using Leave-One-Out Cross-Validation and the optimal selected features associated with the classifier. SVM with Glucose, Resistin, homo, BMI, and Age achieved sensitivity and specificity of 0.94% and 0.90% respectively and overall accuracy of 0.92%.

5. Discussion

The Sequential backward selection model was combined with three different classification algorithms in this study: support vector machine (SVM), random forests (RF), and Decision Trees (DTs). This method was evaluated using 10-fold Monte Carlo Cross-Validation (MCCV) to ensure stability. Furthermore, to avoid the resubstitution mistake, the data were divided into two groups at random: a 70% training set and a 30% testing set. Fig. 2 showed the MCE on the test set as a function of the number of features selected in each iteration. The minimum the cross-validation MCE, the higher accuracy in the model performance. The SVM model reached the minimum cross-validation MCE value when five features were eliminated. However, the DTs model required six features to be eliminated to reach the minimum cross-validation MCE value. The RF model achieved almost the same minimum cross-validation MCE when five or four features were eliminated. As illustrated in Table 2, the models disagreed on the optimal subset of features choices. However, some biomarker/features were common across all models.

In general, four features confirmed necessary: Glucose, Resistin, BMI, and Age, three features confirmed unimportant: Insulin, MCP-1, and Adiponectin, and three exploratory features: HOMA L/A and Leptin. The table shows that Glucose, Resistin, BMI, and Age were common optimal features across the three classifiers with an additional HOMA for SVM and L/A and Leptin for Dts.

The result shows that the optimal set of biomarkers comprises Glucose, Resistin, homo, BMI, and Age using the SVM model. The sensitivity and specificity were 0.94 and 0.90, respectively and the 95% confidence interval for the AUC was [0.89, 0.98]. The result indicates that Glucose, Resistin, homo, BMI, and Age combined can serve as a crucial BC biomarker in BC screening and detection.

6. Conclusions

This study explores the influence of using a particular biomarker or combinations of different biomarkers as predictors in the accuracy of the classification model for breast cancer. It shows that the feature used to construct a specific classifier considerably impacts the classifier's accuracy, sensitivity, and computational cost. Therefore, it is crucial to understand that classification approaches vary in how they handle data. Thus, they respond differently toward biomarkers selection approaches. Although the current study is based on a small sample of participants, the findings suggest that the biomarkers selection process relies on the classification model used as much as the biomarkers.

It is essential to highlight that nowadays, the newly emerging area of artificial intelligence in deep learning has enabled machines to determine features-of features needed for data classification automatically. Deep learning has seen numerous advancements and has contributed to significant precision improvements in many Computer-Aided Systems. However, deep learning requires far more data than a conventional algorithm for machine learning. Collecting more data would be a fruitful area for further work. More data on breast cancer biomarkers would help us establish greater accuracy on this matter. Collecting more data and examining the impact of modern Machine learning approaches can be described as this work's natural progression.

Acknowledgement

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under grant no. RG-88-135-42. The authors, therefore, gratefully acknowledge DSR technical and financial support.

References

Aggrawal, R., & Pal, S. (2020). Sequential feature selection and machine learning algorithm-based patient's death events prediction and diagnosis in heart disease. *SN Computer Science, 1*(6).

Al-Khater, K. M., Almofty, S., Ravinayagam, V., Alrushaid, N., & Rehman, S. (2021). Role of a metastatic suppressor gene kai1/cd82 in the diagnosis and prognosis of breast cancer. *Saudi Journal of Biological Sciences, 28*(6), 3391–3398.

Aslan, M. F., Celik, Y., Sabanci, K., & Durdu, A. (2018). Breast cancer diagnosis by different machine learning methods using blood analysis data. *International Journal of Intelligent Systems and Applications in Engineering, 6*(4), 289–293.

Assiri, A. M., & Kamel, H. F. (2016). Evaluation of diagnostic and predictive value of serum adipokines: Leptin, resistin and visfatin in postmenopausal breast cancer. *Obesity Research & Clinical Practice, 10*(4), 442–453.

Carneiro, G., Nascimento, J., & Bradley, A. P. (2015). Unregistered multiview mammogram analysis with pre-trained deep learning models. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 9351, pp. 652–660). Springer.

Carneiro, G., Nascimento, J., & Bradley, e. a. (2020). Prediction of breast cancer and lymph node metastatic status with tumour markers using logistic regression models. *New England Journal of Medicine, 9*(6), 84–90. arXiv:1504.03293. URL http://www.iarc.fr/en/publications/books/wcr/%0Ahttps://books.google.dk/books/about/World_Cancer_Report_2014.html?id=OQHbngEACAAJ&pgis=1.

Crisostomo, J., Matafome, P., Santos-Silva, D., Gomes, A. L., Gomes, M., Patrício, M., Letra, L., Sarmento-Ribeiro, A. B., Santos, L., & Seiça, R. (2016). Hyperresistinemia and metabolic dysregulation: A risky crosstalk in obese breast cancer. *Endocrine, 53*(2), 433–442.

Cust, A. E., Stocks, T., Lukanova, A., Lundin, E., Hallmans, G., Kaaks, R., Jonsson, H., & Stattin, P. (2009). The influence of overweight and insulin resistance on breast cancer risk and tumour stage at diagnosis: A prospective study. *Breast Cancer Research and Treatment, 113*(3), 567–576.

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In *Ensemble machine learning* (pp. 157–175). Springer.

Darabi, N., Rezai, A., & Hamidpour, S. S. F. (2021). Breast cancer detection using rsfs-based feature selection algorithms in thermal images. *Biomedical Engineering: Applications, Basis and Communications, 33*, 2150020, 03.

Dhungel, N., Carneiro, G., & Bradley, A. P. (2015). Automated mass detection in mammograms using cascaded deep learning and random forests. In *2015 international conference on digital image computing: Techniques and applications, DICTA 2015* (pp. 1–8). IEEE.

Ghani, M. U., Alam, T. M., & Jaskani, F. H. (2019). Comparison of classification models for early prediction of breast cancer. In *2019 international conference on innovative computing (ICIC)* (pp. 1–6).

Iguyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*(Mar), 1157–1182.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90.

Kuppusamy, P., Govindan, N., Yusoff, M. M., & Ichwan, S. J. (2017). Proteins are potent biomarkers to detect colon cancer progression. *Saudi Journal of Biological Sciences, 24*(6), 1212–1221.

Lauby-Secretan, B., Scoccianti, C., Loomis, D., Benbrahim-Tallaa, L., Bouvard, V., Bianchini, F., & Straif, K. (2015). Breast-cancer screening — viewpoint of the IARC working group. *New England Journal of Medicine, 372*(24), 2353–2358.

Levenson, V. V. (2007). Biomarkers for early detection of breast cancer: What, when, and where? *Biochimica et Biophysica Acta (BBA) - General Subjects, 1770*(6), 847–856. URL https://www.sciencedirect.com/science/article/pii/S0304416507000402.

Li, Y., & Chen, Z. (2018). Performance evaluation of machine learning methods for breast cancer prediction. *Applied and Computational Mathematics, 7*(4), 212–216.

Ma, Y., & Guo, G. (2014). *Support vector machines applications* (Vol. 649). Springer.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., et al. (2020). International evaluation of an ai system for breast cancer screening. *Nature, 577*(7788), 89–94.

Nicolini, A., Ferrari, P., & Duffy, M. J. (2018). Prognostic and predictive biomarkers in breast cancer: Past, present and future. *Seminars in Cancer Biology, 52*, 56–73. emerging trends in biomarker discovery: ease of prognosis and prediction in cancer. URL https://www.sciencedirect.com/science/article/pii/S1044579X17300524.

Patrício, M., Pereira, J., Crisóstomo, J., Matafome, P., Gomes, M., Seiça, R., & Caramelo, F. (2018). Using Resistin, glucose, age and BMI to predict the presence of breast cancer. *BMC Cancer, 18*(1), 29.

Peng, W., Mayorga, R. V., & Hussein, E. M. (2016). An automated confirmatory system for analysis of mammograms. *Computer Methods and Programs in Biomedicine, 125*, 134–144.

Petković, M., Kocev, D., & Džeroski, S. (2019). Feature ranking for multi-target regression. *Machine Learning, 109*(6), 1179–1204.

Pham, H., & Pham, D. H. (2020). A novel generalized logistic dependent model to predict the presence of breast cancer based on biomarkers. *Concurrency Computation, 32*(1), e5467.

Rahman, M., Ghasemi, Y., Suley, E., Zhou, Y., & Wang, S. (2020). Machine learning based computer aided diagnosis of breast cancer utilizing anthropometric and clinical features. *IRBM*.

Rokach, L., & Maimon, O. Z. (2008). *Data mining with decision trees: Theory and applications* (Vol. 69). World scientific.

Rue, M., Vilaprinyo, E., Lee, S., Martinez-Alonso, M., dia Carles, M., Marcos-Gragera, R., Pla, R., & Espinas, J. A. (2009). Effectiveness of early detection on breast cancer mortality reduction in Catalonia (Spain). *BMC Cancer, 9*, 326.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics, 21*(3), 660–674.

Santillán-Benítez, J. G., Mendieta-Zerón, H., Gómez-Oliván, L. M., Torres-Juárez, J. J., González-Bañales, J. M., Hernández-Peña, L. V., & Ordóñez-Quiroz, A. (2013). The tetrad BMI, leptin, leptin/adiponectin (L/A) ratio and CA 15-3 are reliable biomarkers of breast cancer. *Journal of Clinical Laboratory Analysis, 27*(1), 12–20.

Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep learning to improve breast cancer detection on screening mammography. *Scientific Reports, 9*(1), 1–12.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018, CA. *A Cancer Journal for Clinicians, 68*(1), 7–30.

Silva Araújo, V. J., Guimarães, A. J., de Campos Souza, P. V., Rezende, T. S., & Araújo, V. S. (2019). Using resistin, glucose, age and bmi and pruning fuzzy neural network for the construction of expert systems in the prediction of breast cancer. *Machine Learning and Knowledge Extraction, 1*(1), 466–482.

Singh, B. K. (2019). Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm. *Biocybernetics and Biomedical Engineering, 39*(2), 393–409.

B. W. Stewart, C. P. Wild, Organization WH. reportWorld Cancer Report 2014, IARC Nonserial Publication. URL http://www.iarc.fr/en/publications/books/wcr/%0Ahttps://books.google.dk/books/about/World_Cancer_Report_2014.html?id=OQHbngEACAAJ&pgis=1.

Theodoridis, S., & Koutroumbas, K. (2009). *Pattern recognition* (4th ed.).

Wang, W., Jones, P., & Partridge, D. (2001). A comparative study of feature-salience ranking techniques. *Neural Computation, 13*(7), 1603–1623.

Wu, M. H., Chou, Y. C., Chou, W. Y., Hsu, G. C., Chu, C. H., Yu, C. P., Yu, J. C., & Sun, C. A. (2009). Circulating levels of leptin, adiposity and breast cancer risk. *British Journal of Cancer, 100*(4), 578–582.

Yang, F., & Mao, K. (2010). Robust feature selection for microarray data based on multicriterion fusion. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8*(4), 1080–1092.

Zarei, M., Rezai, A., & Falahieh Hamidpour, S. S. (2021). Breast cancer segmentation based on modified Gaussian mean shift algorithm for infrared thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1–7.

Zhang, Y., Sohn, K., Villegas, R., Pan, G., & Lee, H. (June, 2015). Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction, 07-12. In *Proceedings of the IEEE computer society conference on computer vision and pattern recognition* (pp. 249–258). arXiv:1504.03293.