Journal of Theoretical and Applied Information Technology
15th March 2021. Vol.99. No 5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A MACHINE LEARNING APPROACH FOR BREAST CANCER EARLY DETECTION

MUAWIA A. ELSADIG

Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam, Saudi Arabia

E-mail: muawiasadig66@gmail.com

ABSTRACT

The rapid increase in incidence of breast cancer is clearly noticeable. The cause of the disease is not clear and the reasons behind the increase of incidence are not well identified. In addition, a method for preventing its occurrence has yet to be discovered. Therefore, its early detection plays a major role in the treatment process and assists in achieving an acceptable survival rate. As a matter of fact, there are many methods, based on machine learning or statistical approaches, for distinguishing between benign and malignant images. However, most of them do not achieve the desired accuracy, due to the use of inaccurate features, the absence of proper classifiers, or inefficient datasets. Therefore, this study introduced an effective classifier approach based on a support vector machine (SVM) with an adequate features selection method that considers only the features with high influence and neglects the others. This scenario potentially enhances the accuracy of classification and reduces the computational overheads. In addition, the use of a trusted dataset and the application of proper validation methods were reflected in reliable and trusted results. Selecting of SVM is done after conducting real experimental scenarios for seven reputable classifiers in the field of breast cancer diagnosis. The experimental results reflect that the classifier approach, based on SVM, outperformed the other classifiers by obtaining the highest accuracy, reaching 97.4%. The contribution of this paper includes introducing an efficient SVM approach to predict breast cancer and presenting a comparison study for seven popular classifiers in this field. Our results have been thoroughly validated to nominate SVM as the best classifier for breast cancer detection.

Keywords: *Data mining, Machine Learning, Deep learning, Medical Images Classification, Breast Cancer Detection, Breast Cancer Diagnosis.*

1. INTRODUCTION

Nowadays, breast cancer is considered one of the main diseases in the world. Many surveys have shed light on the number of people who died because of breast cancer [1]. Mortality from breast cancer is very high compared to other types of cancer [2]. In 2012, a US government survey showed that 40 thousand people had died from this disease [1]. Moreover, 2 million new cases were diagnosed in 2018 [3]. The main reason for the low survival rate is due to the late discovery of this disease and the complications of its diagnosing process. Therefore, early detection has an important impact on reducing the risk by preventing the progression of the disease and reducing its morbidity rates [4, 5]. It allows patients to get appropriate treatment [6]. Dubey et al. have clearly stated that early detection of breast cancer can enhance and boost the survival rate up to 98% for small tumour cases and 73% for large tumour cases [3]. However, early detection is often a hard task due to the absence of symptoms in the beginning [7].

Kumari et al. [1] stated that breast cancer is dominant compared to other types, as illustrated in Figure 1, which gives the percentage of breast cancer compared with the other types of cancer [1]. This reflects the wide spread of the disease.

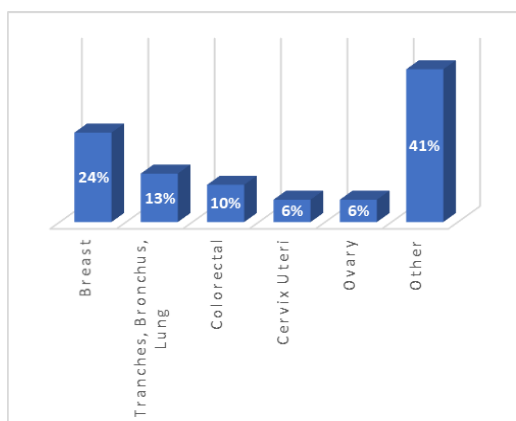

Figure 1: Cancer Types [1]

On the other hand, a recent report [8] shows the number of new cases and deaths in 2018 for different types of cancer disease. According to this

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

report, the numbers of new cases and deaths from breast cancer were 2,088,849 and 626,679, respectively. In respect to the number of new cases, breast cancer comes second, while in respect to the number of deaths, breast cancer is the fifth leading cause of death worldwide for both sexes, while for females breast cancer is the leading cause of death and has the highest number of new cases compared with other types of cancer [8]. Figures 2 and 3 show the number of new cases and deaths for cancer types over the year of 2018, respectively. In the United States, the number of new cases of breast cancer in 2019 was 271,270, while the number of deaths was 42,260 [9]. This reflects the importance of developing an efficient method that supports the early detection of this dangerous disease. This can be achieved by using advanced computing techniques to enhance diagnostic capability.

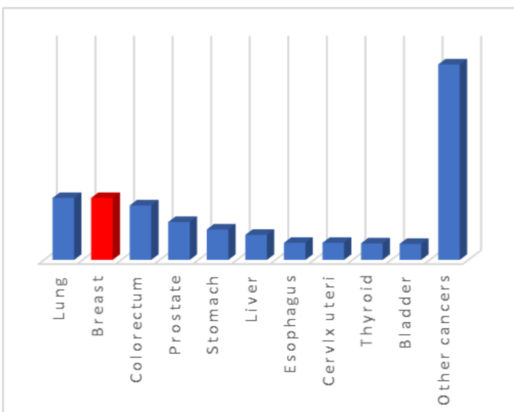

Figure 2: The chart presents the percentages of new cases for the ten common cancers worldwide in 2018 [8].

Microscopic analysis of a biopsy is one of the most important methods that is used for breast cancer diagnosis. However, this method needs a pathologist to perform specialized analysis. This analysis is costly and time-consuming and often leads to non-consensual results [10]. Therefore, the need for computer-aided diagnosis (CAD), such as classification algorithms, is greater.

Computer-aided diagnosis (CAD) assists medical experts in early diagnosis and thus increase the recovery rate. CAD uses machine learning approaches to predict whether a tumour is a malignant or benign cancer [11]. It is a computer-based system that supports specialists in taking decisions quickly [12, 13]. Interested readers in using computer techniques to extract information from different sources (images, text, etc.) are referred to [14,15]

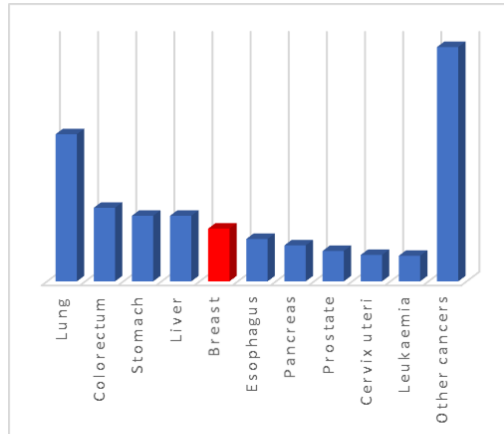

Figure 3: The chart presents the percentages of death cases for the ten common cancers worldwide in 2018 [8].

It is worth mentioning that machine learning techniques ensure their usefulness in discovering and defining patterns of huge medical image datasets and thus lead to successful classification and sorting out of these images. However, selecting an adequate machine learning approach, a trusted dataset, an efficient features selection method and an accurate validation approach is of great value when studying the time of survivability of breast cancer.

This paper proposes a machine learning approach for early detection of breast cancer. The approach uses an adequate classification algorithm and well common dataset after applying an efficient feature selection method that leads to the attainment of high classification accuracy. The selection of our classification algorithm is done through a practical comparison that is based on real experiments, including applying our feature selection method with appropriate training, testing and evaluation methods. These experiments are implemented on seven classifiers that are frequently used in the detection of breast cancer and are used to attain competent classification accuracy.

This paper gives a comprehensive review of breast cancer spread and the artificial intelligence methods that are used to detect such diseases, concentrating on some classifier approaches that are commonly used in this area. The core of the paper is to introduce a proper classifier that is capable of attaining high accuracy levels for early detection of this disease. The paper is organised as follows: the first section gives a brief introduction of breast cancer and its spread worldwide. It sheds light on some computer-aided diagnostic methods and their effective contribution to the diagnosis process for this disease. The next section gives a comprehensive review of the relative methods of detecting breast cancer, their achievements and limitations, with more focus on seven common classifiers that have a good reputation for detecting breast cancer. Section 3 demonstrates our proposed method and Section 4 shows details of our experimental work, its results and the methods that were applied to validate the results. In addition, this section gives a thorough discussion based on our experimental results. Finally, the paper is concluded in section 5, while section 6 sums up our future work.

2. RELATED WORK

The use of machine learning classifiers to support the diagnostic process is increasing rapidly in the medical field. As a matter of fact, evaluations of patients' data and experts' decisions are the most important influences in the diagnostic process. However, expert systems and artificial intelligence approaches for classification also support experts a great deal. Machine learning classifiers can help to minimise the possible errors caused by inexperienced experts, and also provide detailed medical data in a short time to be examined.

In fact, many problems with using machine learning approaches are connected to the lack of precise and efficient validation. It is true that machine learning approaches improve survival prediction accuracy. However, selecting an adequate validation approach is of great value in studying the time of survivability of breast cancer [16].

Gao et al. [17] compared traditional methods vs. the machine learning–based methods that are used in breast imaging (mammography). Their aim was to emphasise the limitations of the traditional methods and to highlight the potential solutions of computer-aided detection (CAD) approaches. They concluded that CAD development is experiencing a worldview move based on the endless advancement of computing power and the rapid emergence of deep learning approaches. Therefore, we are witnessing today a promising area for developing machine learning-based algorithms that play an important and effective role in enhancing clinical care systems. In the same context, Araújo et al. [4] added that deep learning approaches represent efficient alternatives to conventional classification approaches, as they are capable of overcoming the obstacles of feature-based approaches.

Abdar et al. [2] pointed out that most of the research work on breast cancer histopathology image analysis are conducted using small datasets. In addition, the key obstacle to developing new histopathology image analysis methods is the absence of public, large and annotated datasets [18]. As a matter of fact, annotation is an important key to developing and validating any machine learning-based approach [2].

Numerous studies have highlighted the need to combine many classifiers together, rather than using a single classifier. This trend has led to efficient classification approaches that are capable of attaining acceptable classification accuracy. Interested readers are referred to [19-21].

2.1 Why these classifiers?

In this subsection, the paper introduces some relevant works on seven common classifiers that have been used frequently in breast cancer diagnosis. These classifiers reflect high performance in terms of classification accuracy and have made a good contribution to this field. What we have stated here in this subsection will fully support our idea of selecting these classifiers as the most common classifiers. So, our selection of the best classifier (through real experimental scenarios) will be based on a solid, accurate and concreate assumption.

Recently, Hosni et al. [22] reported that the classifiers most frequently used to build up ensemble classifiers are support vector machines, artificial neural networks and trees. The author in [23] also stated that it has been increasingly reported that the SVM classifier has superior accurate diagnosis capability [23].

Basing their work on the WBCD (Wisconsin Breast Cancer Database) dataset, Medjahed et al. applied a K-nearest neighbours classifier using different classification rules and distance types. Accordingly, they came up with the result that the two types of distance, Euclidean and Manhattan, are more effective in regards to classification accuracy and performance compared to other types of distance that were examined in their study. These two types of distance are time-consuming, but they still give better results [7].

Meenakshi et al. [24] conducted a comparison between two types of neural network classifier, radial basis function (RBF) and multilayer perceptron (MLP). They reported that MLP outperformed RBF by generating more accurate and specific results. In addition, Daniele Soria et al. [25] outlined that MLP was the most highly effective classifier in their study, which was conducted to compare three different classifiers: the C4.5 tree classifier, MLP and naïve Bayes (NB). Kathija et al. recommended the use of MLP to predict the survivability of breast cancer. This was based on their claim that MLP brings better accuracy and performance [26].

The naïve Bayes classifier is a simple classifier that is built on Bayes' theorem with independent assumptions. These assume that (i) the predictive attributes are conditionally independent and (ii) the numeric attribute values are normally distributed within each class [25]. In comparison with the MLP classifier, Daniele et al. [25] stated that MLP is a highly effective classifier with poor interpretability, while NB attains good performance and interpretability.

Moyano et al. [27] conducted a comparison study on artificial neural networks and logistic regression classifiers. They stated that these classifiers are the two most frequently used in clinical risk estimation. The two approaches were used to estimate breast cancer risk and showed similar performance. However, both approaches have their strengths and limitations, which should be taken into consideration. ANN is useful when there are complex relationships and implicit interactions in the data, whereas logistic regression is recommended when statistical inferences from the output need to be drawn. Both approaches can potentially assist physicians in understanding cancer risk factors, risk estimation and diagnosis [27]. Moyano et al. clearly stated that the two approaches may be used complementarily to help in decision making. This supports those who prefer using an ensemble classifier.

Random Forest is an ensemble machine learning technique that runs efficiently on large databases. It consists of a collection of tree-structured classifiers [28].

Nguyen et al. [29] developed a classifier model based on a random forest classifier and feature selection technique to help in breast cancer diagnosis. Their model, which is applied to two different datasets, has gained high classification accuracy and is considered a very promising result in this field. This promising result motivates us to include the random forest classifier as one of the classifiers under investigation in this paper.

A recent study conducted an experiment by applying four different selection methods and four common classifiers. This experiment was done on four different datasets. When applying feature selection, the results showed that an artificial neural networks approach outperformed the other approaches (Naïve Bayes, Support Vector Machines and Decision Trees) by attaining a noteworthy increase in breast cancer classification accuracy, while there was no increase in accuracy when applying a feature section with the other classifiers [11].

Therefore, and based on the aforementioned relevant studies, seven classifiers which have successfully ensured their efficiency in attaining good classification accuracy have been picked to undergo investigation in this paper. The seven classifiers are ANN, SVM, Random Forest, Logistic Regression, NB, KNN and DT.

3. THE PROPOSED METHOD

This paper has selected seven of the most popular classifiers that are used in breast cancer detection. These classifiers have been trained and tested using popular and trusted datasets after applying an efficient feature selection method to these datasets (picking the important features that have the most influence will greatly enhance the detection accuracy). Then the classifier that outperforms the others will be considered and proposed as a classifier approach for this paper. The selected classifiers are: ANN, SVM, Random Forest, Logistic Regression, NB, KNN and DT. The defence for selecting these classifiers is clearly stated in the previous section. Each classifier is built separately, and accordingly the results obtained are recorded and reported.

3.1 Dataset

This paper uses a worldwide and common dataset for the training and testing of all the classifiers used in this work. It uses the Wisconsin Breast Cancer datasets from the UCI Machine

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Learning Repository, which has 699 numerical instances and 30 features. This dataset has been enhanced by using a feature selection method that assists in reducing the number of features by considering only those that have a high influence and ignoring redundant and unweighted features. The subsection below gives more details in this regard. In this paper, the dataset has been partitioned into three sets for training and testing purposes, as follows: (i) set 1: 70% for training and 30% for testing (ii) set 2: 80% for training and 20% for testing (iii) set 3: 90% for training and 10% for testing. Each classifier is built, trained and tested by applying these training-testing sets. Applying these three sets of different training sizes is intended to reflect accurate and reliable results.

	A	B	C	D	E
1	texture_worst	radius_worst	perimeter_worst	perimeter_mean	radius_
2	17.33	25.38	184.6	122.8	
3	23.41	24.99	158.8	132.9	
4	25.53	23.57	152.5	130	
5	26.5	14.91	98.87	77.58	
6	16.67	22.54	152.2	135.1	
7	23.75	15.47	103.4	82.57	
8	27.66	22.88	153.2	119.6	
9	28.14	17.06	110.6	90.2	

Figure 4: Part of the revised dataset with ten features.

3.2 Features Selection:

Feature selection is a crucial part of building up an effective classification approach. The adequate decreasing of the number of features can definitely enhance classifier predictability and produce a less computationally intensive classification system. In the medical image classification field, limiting the number of features reduces the diagnostic costs and time. In the relevant literature, there are many methods and algorithms that are used for features selection, to support and increase classification accuracy and reduce systems cost. In our approach, the relief feature selection algorithm is used. This algorithm is designed for binary classification systems. For each feature, the relief algorithm calculates a feature score. Accordingly, the features are ranked to select the top features that have the most effective impact on the classification process. Based on the aforementioned feature selection algorithm, only ten features have been selected. Figure 4 shows a part of the enhanced dataset with the selected features.

3.3 Validation Methods:

In this study, the random sampling validation technique is applied to each experiment. This technique is repeated 20 times in order to attain reliable and realistic results. The classifier performance is evaluated using a confusion matrix, which is used to evaluate classification errors, false positive and false negative.

The accuracy is calculated by using the equation shown in (1).

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \qquad (1)$$

Where:

True positives (TP): refer to positive images that are classified correctly.

True negatives (TN): refer to negative images that are classified correctly.

False positives (FP): refer to negative images that are classified incorrectly.

False negatives (FN): refer to positive images that are classified incorrectly [30, 32].

Eventually, the seven classifiers are ranked according to their accuracy achievements.

3.4 Classifiers Implementation Tool

In this study, all experiments have been conducted using Orange data mining software. It is an open source and very useful machine learning tool to implement different classification algorithms and to visualize your data efficiently. Figure 5 shows one scenario of our experiments using Orange software.

3.5 Factors of Success

In this paper, there is more than one factor that greatly enhances either the classifiers' accuracy or the applied validation methods and thus supports our findings. These factors include:

- Selecting the most common and popular classifiers to place under investigation.
- Using a trusted international dataset from a trusted source.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Applying an adequate features selection method to pick only the features with greatest influence, which enhances the accuracy and reduces the processing time.
- Applying different sets of training-testing.
- Using a random sampling validation technique.
- This random sampling technique is repeated 20 times for each experiment in order to attain reliable and realistic results.

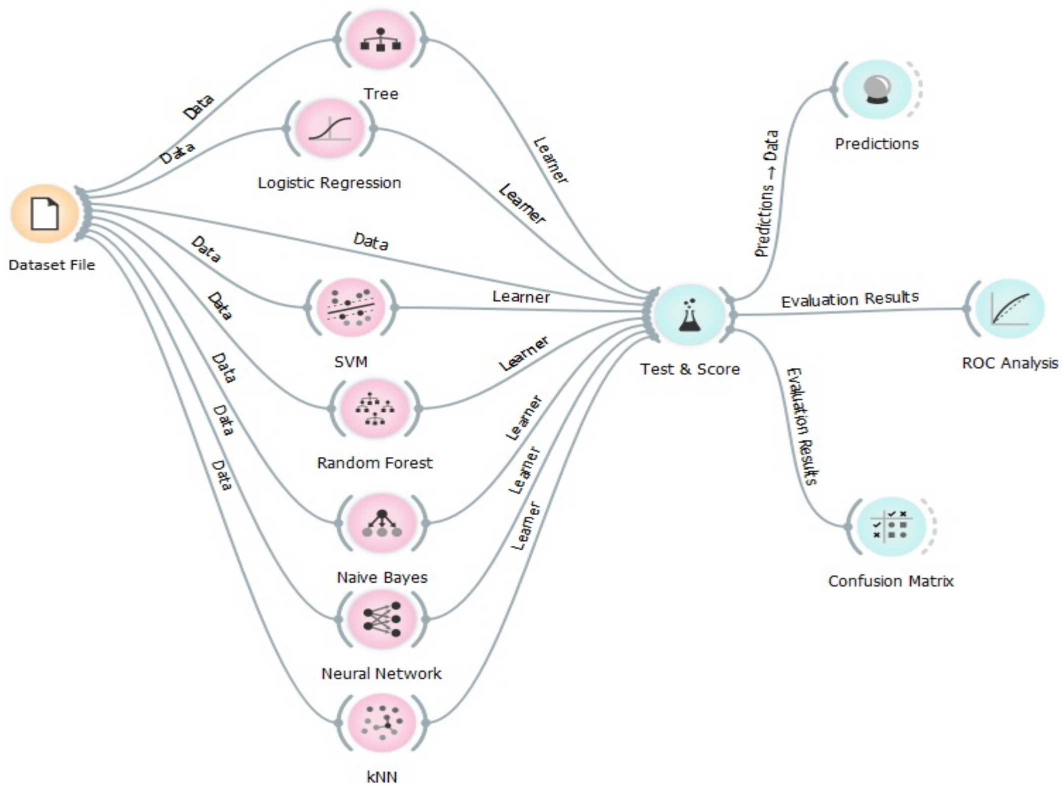

Figure 5: Classifiers Implementation using Orange software

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this paper, seven common classifiers for breast cancer have been selected to pick the one which attains the highest performance in terms of classification accuracy compared to the others, after applying different factors and techniques to enrich the classification accuracy. The selection of these classifiers is based on their good reputation in the classification domain, and especially in breast cancer diagnosis.

As per our method stated in the previous section - which includes data acquisition and pre-processing, features selection, building classifiers, classifier training, classifier testing and validation - all the experiments have been implemented accordingly. Table 1 shows the accuracy results for the seven classifiers over three different training-test sets (70%:30%, 80%:20% and 90%:10%). It is clearly evident that the SVM has outperformed the other classifiers by attaining the highest accuracy results: 97.3%, 97.4% and 96.8% over the three different experiments of training-testing sets, 70%-30%, 80%-20%, and 90%-10%, respectively.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1 Comparison of The Seven Classifiers' Accuracy

Training size 70%		Training size 80%		Training size 90%	
Classifier	Accuracy	Classifier	Accuracy	Classifier	Accuracy
SVM	0.973	SVM	0.974	SVM	0.968
MLP	0.972	MLP	0.970	MLP	0.964
RF	0.962	RF	0.958	RF	0.958
LR	0.944	LR	0.948	LR	0.950
KNN	0.933	KNN	0.934	DT	0.934
DT	0.931	DT	0.934	KNN	0.932
NB	0.930	NB	0.930	NB	0.932

A graphic presentation of our findings is also given in Fig. 6, which clearly reflects the accuracy percentages of the investigated classifiers which were achieved under three different benchmarks of training-test.

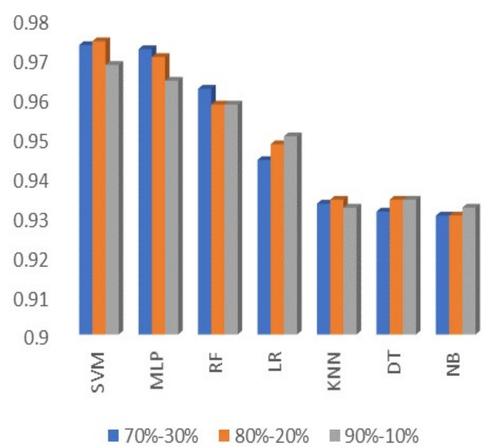

Figure 6: Graphic representation of the classification accuracy achieved by the seven classifiers under different training-testing sets

The rest of the figures - Figures 7, 8 and 9 - illustrate the Receiver Operating Characteristic (ROC) curve for the seven classifiers investigated under three different training-testing sets: 70%-30%, 80%-20%, and 90%-10%, respectively.

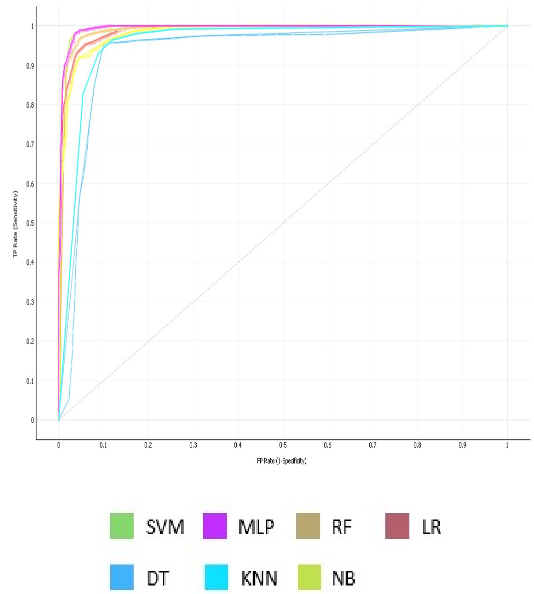

Figure 7: ROC curve: 70%-30% training-testing scenario

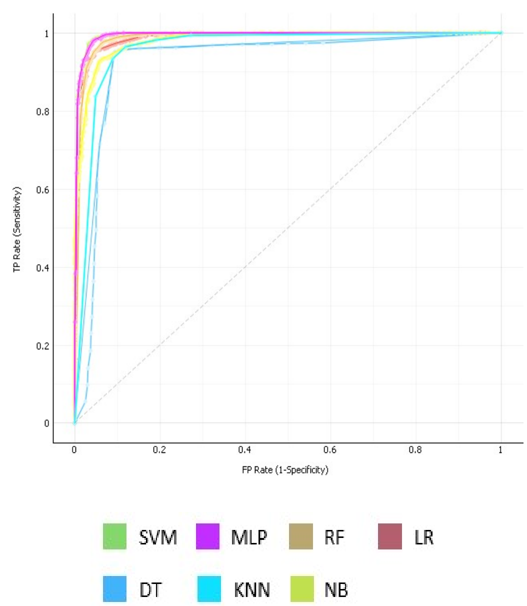

Figure 8: ROC curve: 80%-20% training-testing scenario

Journal of Theoretical and Applied Information Technology
15th March 2021. Vol.99. No 5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

These figures reflect that, in terms of specificity and sensitivity, the SVM is the best compared to other classifiers.

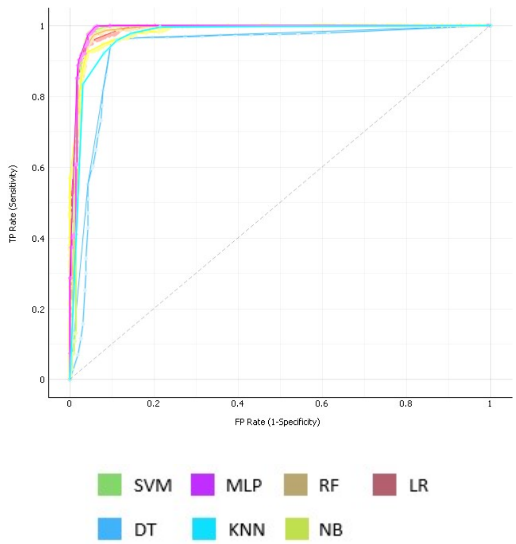

Figure 9: ROC curve: 90%-10% training-testing scenario

Based on these findings, the SVM would be suggested as the best classifier for early detection of this kind of disease. However, more work that considers different datasets, different features selection methods, different training-testing parameters, etc. is required to support this finding.

5. CONCLUSION

After a thorough review and efficient experiments that are based on applying an adequate features selection method, this paper has introduced an approach to predicting breast cancer using SVM. The proposed approach has been successfully validated to attain high accuracy, reaching 97.4%. It outperforms the other six common classifiers that have a good reputation in this field.

Also, this paper presented a comparison study and assessment of the top most common classifiers in terms of their classification accuracy and reported their achievements. The strength of the proposed approach relies on three phases of enhancement as follows:

Phase 1: Applying an adequate feature selection method is the most effective way to enhance classification accuracy. Applying this method reduces the number of dataset features. Reducing the number of features – by selecting the weighted features and ignoring the ones that have a light weight – enriches the classifier’s performance in terms of classification accuracy and also reduces the overheads by decreasing the computational cost and time. The most effective way to enhance classification accuracy is by selecting the proper classification features.

Phase 2: The proper selection of the top classifiers, which is based on their good reputation in this field and on their classification accuracy they achieved, which was calculated through experiments conducted using an enhanced version of a reliable and trusted dataset. Dataset without irrelevant, noisy, and redundant features.

Phase 3: The use of a random sampling technique which is repeated 20 times in order to attain reliable and realistic results.

Eventually, we can conclude that the study contributes by introducing an effective classifier approach based on a support vector machine (SVM) with an adequate features selection method that considers only the features with high influence and neglects the others. Reduction of irrelevant, noisy, and redundant features has potentially enhanced the accuracy of our classification model.

Moreover, the paper has presented a through experimental comparison of the top seven classifiers which is another important contribution for this work.

6. FUTURE WORK

Our future work would include different implementation scenarios under different datasets using SVM.

ACKNOWLEDGMENT

This work is a part of our project No. 2019-348-DSR that supported by the Deanship of Scientific Research of IAU. Therefore, the author is highly appreciating their encourage and support.

REFRENCES:

[1] N. Kumari and K. Verma, "A Survey On Various Machine Learning Approaches Used For Breast Cancer Detection," International Journal of Advanced Research in Computer Science, vol. 10, no. 3, p. 76, 2019.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455-1462, 2015.

[3] A. K. Dubey, U. Gupta, and S. Jain, "A Survey on Breast Cancer Scenario and Prediction Strategy," 2015: Springer, pp. 367-375.

[4] T. Araújo et al., "Classification of breast cancer histology images using convolutional neural networks," PloS one, vol. 12, no. 6, p. e0177544, 2017.

[5] R. A. Smith, V. Cokkinides, and H. J. Eyre, "American Cancer Society guidelines for the early detection of cancer, 2004," CA: a cancer journal for clinicians, vol. 54, no. 1, pp. 41-52, 2004.

[6] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. P. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," Pattern Recognition Letters, vol. 125, pp. 1-6, 2019.

[7] S. A. Medjahed, T. A. Saadi, and A. Benyettou, "Breast cancer diagnosis by using k-nearest neighbor with different distances and classification rules," International Journal of Computer Applications, vol. 62, no. 1, 2013.

[8] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: a cancer journal for clinicians, vol. 68, no. 6, pp. 394-424, 2018.

[9] M. D. Bacolod et al., "Prediction of blood-based biomarkers and subsequent design of bisulfite PCR-LDR-qPCR assay for breast cancer detection," BMC cancer, vol. 20, no. 1, p. 85, 2020.

[10] G. Aresta et al., "Bach: Grand challenge on breast cancer histology images," Medical image analysis, 2019.

[11] N. Lindqvist and T. Price, "Evaluation of Feature Selection Methods for Machine Learning Classification of Breast Cancer," ed, 2018.

[12] Q. Li and R. M. Nishikawa, Computer-aided detection and diagnosis in medical imaging. Taylor & Francis, 2015.

[13] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," Computerized medical imaging and graphics, vol. 31, no. 4-5, pp. 198-211, 2007.

[14] M. Abdelmagid, A. Ahmed, and M. Himmat, "Information Extraction Methods and Extraction Techniques in the Chemical Document's Contents: Survey," ARPN Journal of Engineering and Applied Sciences, vol. 10, no. 3, 2015.

[15] M. Abdelmagid et al., "Survey on Information Extraction from Chemical Compound Literatures: Techniques and Challenges," Journal of Theoretical and Applied Information Technology, vol. 67, no. 2, 2014.

[16] I. Mihaylov, M. Nisheva, and D. Vassilev, "Application of Machine Learning Models for Survival Prognosis in Breast Cancer Studies," Information, vol. 10, no. 3, p. 93, 2019.

[17] Y. Gao, K. J. Geras, A. A. Lewin, and L. Moy, "New Frontiers: An Update on Computer-Aided Diagnosis for Breast Imaging in the Age of Artificial Intelligence," American Journal of Roentgenology, vol. 212, no. 2, pp. 300-307, 2019.

[18] M. Veta, J. P. W. Pluim, P. J. Van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: A review," IEEE Transactions on Biomedical Engineering, vol. 61, no. 5, pp. 1400-1411, 2014.

[19] M. Abdar et al., "A new nested ensemble technique for automated diagnosis of breast cancer," Pattern Recognition Letters, vol. 132, pp. 123-131, 2020.

[20] D. Saraswathi and E. Srinivasan, "An ensemble approach to diagnose breast cancer using fully complex-valued relaxation neural network classifier," International Journal of Biomedical Engineering and Technology, vol. 15, no. 3, pp. 243-260, 2014.

[21] S. Lee, M. Amgad, M. Masoud, R. Subramanian, D. Gutman, and L. Cooper, "An Ensemble-based Active Learning for Breast Cancer Classification," 2019: IEEE, pp. 2549-2553.

[22] M. Hosni, I. Abnane, A. Idri, J. M. C. de Gea, and J. L. F. Alemán, "Reviewing Ensemble Classification Methods in Breast Cancer," Computer methods and programs in biomedicine, 2019.

[23] M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," Expert systems with applications, vol. 36, no. 2, pp. 3240-3247, 2009.

[24] C. Meenakshi, M. Govindarajan and A. M. Sameeullah, "Detection of Breast Cancer Using

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MLP and RBF Classifiers," IMS Manthan, 5 (1), pp. 35 – 38, 2010.

[25] D. Soria, J. M. Garibaldi, E. Biganzoli, and I. O. Ellis, "A comparison of three different methods for classification of breast cancer data," 2008: IEEE, pp. 619-624.

[26] A. Kathija, S. S. Nisha, and M. M. Sathik, "Breast cancer data classification using neural network approach of MLP algorithm," International Journal of Trend in Research and Development, vol. 4, no. 3, pp. 275-279, 2017.

[27] T. Ayer, J. Chhatwal, O. Alagoz, C. E. Kahn Jr, R. W. Woods, and E. S. Burnside, "Comparison of logistic regression and artificial neural network models in breast cancer risk estimation," Radiographics, vol. 30, no. 1, pp. 13-22, 2010.

[28] K. Vy and P. K. Sinha, "Random forest classifiers: a survey and future research directions," Int. J. Adv. Comput, vol. 36, no. 1, pp. 1144-1153, 2013.

[29] C. Nguyen, Y. Wang, and H. N. Nguyen, "Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic," Journal of Biomedical Science and Engineering, vol. 6, no. 05, p. 551, 2013.

[30] M. A. Elsadig and Y. A. Fadlalla, "Packet Length Covert Channel: A Detection Scheme," 2018: IEEE, pp. 1-7.

[31] M. A. Elsadig, "Resolving Network Packet Length Covert Channels," Diss. Sudan University of Science & Technology, 2018.

[32] A. Ahmed, O. M. Barukab and M. A. Elsadig, " Heterogeneous Multi-Classifier Method based on Weighted Voting for Breast Cancer Detection," International Journal of Advances in Science Engineering and Technology, vol. 7, no. 4, 2019.