

Research Article

A Supervised Learning Identification System for Prognosis of Breast Cancer

Vandana Rawat,¹ Kamal Gulati²,^{ORCID} Upinder Kaur,³ Jitendra Kumar Seth,⁴ Vikas Solanki,⁵ A. Narasima Venkatesh,⁶ Devesh Pratap Singh,⁷ Neelam Singh,⁷ and Muralidaran Loganathan⁸^{ORCID}

¹Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

²Amity School of Insurance, Banking and Actuarial Science, Amity University, Noida, Uttar Pradesh 201 304, India

³Department of Computer Science & Engineering, Akal University, Bathinda, Punjab, India

⁴Department of IT, KIET Group of Institutions, Ghaziabad, India

⁵Institute of Engineering and Technology, Chitkara University, Punjab, Chandigarh, India

⁶Professor and Head of The Department, HRM and General Management, ISBR Business School, Bangalore 560 100, India

⁷Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

⁸St. Joseph University, Dar es Salaam, Tanzania

Correspondence should be addressed to Muralidaran Loganathan; muralinew1986@gmail.com

Received 14 June 2022; Accepted 15 July 2022; Published 28 August 2022

Academic Editor: Amandeep Kaur

Copyright © 2022 Vandana Rawat et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Breast cancer is one of the most dangerous cancers, accounting for a large number of fatalities each year. It is the leading cause of mortality among women globally. It is getting a lot of interest in the scientific community because of its possible life-threatening danger. As a consequence, many machine learning methods (MLMs) have been modified to provide the best results for early diagnosis of this malignancy. Machine learning methods (MLMs) offer several beneficial implications in breast cancer, including early prognosis, detection, and diagnosis. Compared to traditional statistical analysis, machine learning methods (MLMs) have the capacity to improve the analysis of various health data, such as unstructured, complicated, and noisy data. With the demanding prevalence of breast cancer and the arrival of “data reformation,” it is thus imperative to mention the ethical consequences of machine learning (ML) on society and cancer care. It offers conclusively strong tools, smart methods, and efficient algorithms that can help in the prognosis of breast cancer. The focus of this review is on supervised techniques such as classification and regression that may be implemented and used for breast cancer data analysis. Some supervised learning methods like Naive Bayes, AdaBoost, and support vector machine are presented in this work in the early identification of breast cancer. These algorithms have been analyzed for their accuracy and efficiency using various assessment metrics and methods.

1. Introduction

Breast cancer is among the most significant threats to women worldwide. Among females, this is the second biggest cause of death [1]. Breast cancer is the second largest cause of death, according to the World Health Organization (WHO) [2]. Breast cancer claimed the lives of 70,218 Indian women in 2012. It is a stepwise procedure offering a range of cell types, but prevention remains a significant challenge for

it, all over the world [2]. Huge methodological revolutions have occurred in this sector throughout the years; however, there are a variety of challenges.

In timely identification, several image processing methodologies have been applied to minimize the probability of death per year. As a result, a variety of machine learning techniques are applied to boost precision in the detection of breast cancer. In recent decades, various machine learning techniques have been developed and they are

divided into 3 stages: (a) preprocessing; (b) feature extraction; and (c) classification.

Preprocessing mammography pictures improves visibility of outlying regions and brightness, making data interpretation easier. Breast cancer survivors in certain industrialized nations have an overall 5-year survival rate of more than 80% due to early identification and care. Breast cancer stem cells have shown the genesis and methods of tumor-resistant strains, and various breast cancer genes have been found. People now have a plethora of pharmaceutical options for breast cancer prevention and treatment, and biological prevention has since been created to provide even greater advantages in terms of living standards [3–5].

2. Breast Cancer

Breast cancer has been defined as abnormal tissue development in the breast that results in malignant tumors. Non-invasive breast cancer refers to malignant cells that remain in the breast. Invasive malignancies cannot heal quickly and might spread to other regions of the body. These are deadly malignancies that can sprawl to other organs of the human body which results in metastatic cancer [6–9].

2.1. Categories of Breast Cancer. There are 2 categories of breast cancer: intrusive and non-intrusive. Ductal carcinoma in situ is a non-intrusive variety, whereas the rest are intrusive [10–13].

In Figure 1, various types of breast cancer are given, and they are defined in Table 1.

2.2. Breast Cancer Diagnosis. Multiple investigations are required for the diagnosis of breast cancer, such as a clinical examination, which is probing of the breast to discover cancer indicators, indicating the initial stage. Following the clinical examination, the next stage is medical imaging, which allows for the identification of tumor masses as well as the collection of data from the clinical examination. Many kinds of test include mammography, ultrasonography, and magnetic resonance imaging (MRI) [14–17]. One of these strategies is chosen based on the patient's circumstances as shown in Figure 2.

Machine learning is now in such high demand that it is being offered as a service. Machine learning is a high-barrier area that frequently necessitates specialist knowledge. A set of abilities and knowledge is required to design an efficient machine learning model, which includes the steps of pre-processing, feature selection, and classification processes. So, in Figure 2, some steps of breast cancer diagnosis are depicted.

3. Materials and Methods

In machine learning [18], there are various types of automatic models that are proposed to build from previous records in this type of learning. The machine learning model assesses and may generate future predictions based on that training model [19].

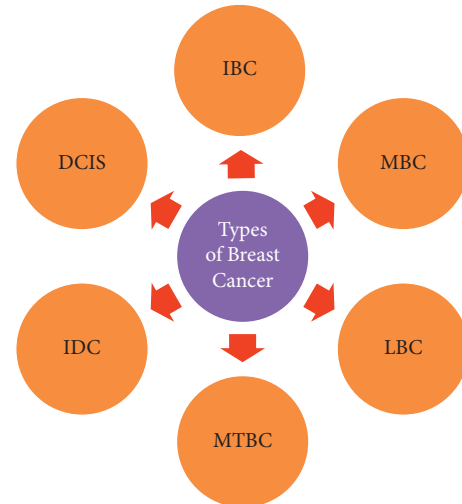


FIGURE 1: Multiple sorts of breast cancer.

To begin this investigation, the Wisconsin Breast Cancer Dataset was obtained from the UCI Machine Learning Repository. The datasets are classified into two types: malignant and benign. The points are derived from needle sputum imaging data from a breast tumor and describe the nucleus of the present picture. The WDBC database has been utilized to detect 212 cancer and 357 benign cases among 569 Wisconsin hospital patients [20–22].

The following subcategories will outline the process flow:

- (i) Dataset definition
- (ii) Dataset assessment
- (iii) Training and testing

3.1. Dataset Definition. The Wisconsin Breast Cancer Dataset (WBCD) is shown in this work with separation of benign and malignant status in Figure 3. This dataset contains 357 benign and 212 malignant breast cancer patients, respectively.

All the breast cancer attributes are listed in Figure 3 and Table 2 with real-value characteristics.

3.2. Dataset Assessment. The entire dataset has been examined for the data patterns. The mean radius feature of the dataset is counter-plotted and found that expected cancer-free patients have a mean radius of 1, whereas probable cancer-bearing patients have a mean radius greater than 1.

3.3. Training and Testing. The dataset numbers are examined based on their properties before being used in the following stage. The dataset was then divided into two halves at random: training and testing, 75 and 25, respectively.

Patients are confronted with a variety of interventions on a regular basis, including ultrasound, biopsy, and mammography, depending on the severity of their breast cancer signs. The most common of these operations is biopsy, which entails the removal of a piece of tissue or cell for

TABLE 1: Types of breast cancer.

Type	Summary
Ductal carcinoma in situ	It comes under the non-intrusive cancer category, which is created inside the breast lactation, as the name implies.
Ductal carcinoma	Caused by cancerous cells penetrating the lactation ducts and through canal wall.
Lobular carcinoma	Cancerous cells grow in the lobules, which are gathered in the lobes, in this type of cancer. They subsequently cross the lobules' walls and disseminate throughout the surrounding tissues.
Inflammatory carcinoma	It manifests as a non-healing tiny nipple wound.
Paget's disease	It manifests as a non-healing tiny nipple wound.
Other carcinomas	Rarest type.



FIGURE 2: The steps of breast cancer diagnosis.

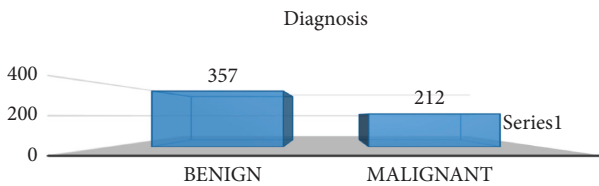


FIGURE 3: Breast cancer diagnosis.

investigation. A fine-needle aspiration (FNA) operation is used to remove a tissue sample from a breast and transmit it to a pathology laboratory for analysis under microscopy. Numerical qualities can be ascertained through an investigation of tissues at a microscopic level. The evidence received from fine-needle aspiration is evaluated in conjunction with other imaging analyses to find the possibility of the patient having a cancerous lump.

Six classification models are supposed to implement including ANN, logistic regression, AdaBoost, Naive Bayes, support vector machine, and K-nearest neighbor. The findings obtained are then assessed by the algorithms in order to find an efficient model. The workflow for the given dataset is illustrated in Figure 4.

4. Machine Learning Algorithms for Breast Cancer Prediction

4.1. *K-Nearest Neighbor (KNN)*. KNN [23] (K-nearest neighbor) is a supervised method that categorizes data depending on how its neighbors are categorized. KNN maintains track of all existing instances and classifies new ones using a matching score. The KNN method looks through a database for data that are similar to what are already available. These newly found data are known as the current data's nearest neighbors. With the test station in this article, KNN is used to choose generally relevant adjacent stations. Assume the road network has a M station [24, 25].

In this algorithm, a random sample size dataset has been taken with various data fields. It is the m^{th} station's historic

traffic flow data, which differ from the substation. The Euclidean distance is used to determine the degree of correlation between the substation and other locations.

4.2. *Support Vector Machine (SVM)*. Vapnik was the first one to discover the support vector machine in 1979 [26]. Vapnik suggested it for classification and regression problems again in 1995 [27]. Support vectors with multilayer perceptron and radial-basis function networks can be utilized for pattern classification. The SVM is a cutting-edge technique that incorporates maximum classification algorithms into statistical learning [28]. Linear and non-linear data are categorized using SVM techniques. It involves converting training dimensional space using non-linear projection. The classification task is carried out by SVM, which maximizes the margin and classifies both classes while reducing classification errors. Although the SVM is used to solve a range of problems, such as regression, data categorization is the most popular one. Figure 2 depicts the main concept. The goal is to identify a hyperplane that separates them by the greatest possible margin with the categorization of positive and negative data points.

4.3. *Artificial Neural Network (ANN)*. This algorithm is capable of performing a variety of tasks like prediction, curve fitting, and regression. It formulates the output using a transfer function [29, 30]. Each intake is multiplied by a weight, which acts as a link between the input, the neuron, and its layers. To acquire the result, the neuron uses a transfer function in the last stage. ANN approaches offer the advantage of having no prior understanding of mathematical computations across attributes, although they do demand less processing work and can handle multivariable problems [31].

4.4. *Logistic Regression (LR)*. The only machine learning method that is not a black-box approach is logistic regression. In general, black-box techniques are difficult to comprehend, but logistic regression demonstrates how they work. Logistic regressions are classified into three types: binary, multinomial, and ordinal [32]. Logistic regression originated as a classification technique rather than a linear regression. Scikit-learn solves multiclass classification issues with a more appropriate logistic regression implementation technique [33–35].

TABLE 2: Dataset attributes of breast cancer patients.

Attribute no.	Attribute description
Radius	The average of both the distances between the center and circumferential endpoints
Texture	Gray-scale rating standard error of the mean (SD)
Perimeter	The length between the snake's endpoints
Area	The whole pixels on the snake's inside, along with one-half of the pixels around its perimeter
Smoothness	Diameter length varies locally
Compactness	Perimeter ² /area
Concavity	The amplitude of the contour's sloping sections
Concave points	The frequency of concavities from the inside of the contour
Symmetry	The length disparity in both ways to the cell edge between lines perpendicular to the major axis
Fractal dimension	Output of large value is less regular contour

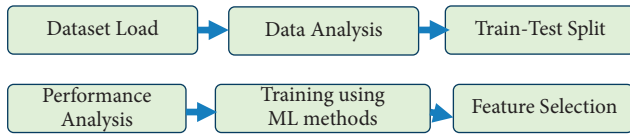


FIGURE 4: Data mechanism/workflow.

4.5. Naive Bayes. A probabilistic learner is called a Naive Bayes classifier. It is most commonly utilized in clinical conditions. Some of the significant benefits of this classifier are that it just requires a minimal scale of training dataset for computation. According to the Naive Bayes method, the occurrence of a particular feature in a class has no bearing on the inclusion or exclusion of any other feature [36]. Only the semantic content of the messages is classified by the Naive Bayes classifier. Operational filters take into account details like the existence of suspicious header, which provides extra properties to the message interpretation [37, 38].

Consider a set of experiments in which each statement is eventually represented as a vector VY_1, VY_2, \dots, VY_n , where Y_1, Y_2, \dots, Y_n are given properties. Each characteristic contains information about a single message token. In the simple case, all the attributes are from Boolean [39]:

$$\begin{aligned} X_i &= 1 \text{ (if the token is in the message);} \\ &\text{otherwise, } X_i = 0. \end{aligned} \quad (1)$$

Bayes' theorem represents the likelihood that a message with vector $(Y) = Y_1, Y_2, \dots, Y_n$ belongs to the following category c :

$$P(C|\vec{Y}) = \frac{P(C) \cdot P(\vec{Y}|C)}{P(Y)}. \quad (2)$$

4.6. AdaBoost. AdaBoost is a sort of classification boosting technique. It focuses on data points, and this technique combines the prognosis efficiency of numerous weak classifiers into a single strong classifier [40–44]. This method focuses on classifier “training” and “weighting.” It applies iteration to the provided method after assigning a weight value to every occurrence in the training dataset. Iteration raises the number of incorrect instances while lowering the number of correct ones. This iteration is carried out in order

to improve the weak classifiers' accuracy [45–47]. The AdaBoost algorithm reduces the exponential loss as follows.

Step 1: Let us check an example as follows:

$$(x_1, y_1), \dots, (x_n, y_n): x_i \in X, y_i \in (-1, +1). \quad (3)$$

Step 2: Assign the weight as $(W_1(i) = 1/N, i = 1 \dots \dots N, \text{ for } n = 1, \dots, N.$

Train the weak learner.

Step 3: Examine weak hypothesis $h_n: X \rightarrow R$ with its inaccuracy:

$$\epsilon_n = \sum_{i=h_m(x_i) \neq y_i} W_{n(i)}. \quad (4)$$

Step 4: Find

$$\epsilon_{m=R}. \quad (5)$$

Step 5: Update

$$K_{m+1}(i) = \frac{K_{m(i) \exp(-x_{m,y_m} M_m(x_m))}{c_m}. \quad (6)$$

D_{m+1} means distribution.

Step 6: Final hypothesis will be

$$H_{(a)} = \text{sign} \left(\sum_{m=1}^M \alpha_m h_m(a) \right). \quad (7)$$

5. Measurement and Experimental Result

5.1. Confusion Matrix for Different ML Algorithms. In the given breast cancer dataset, various machine learning methods [48, 49] are used in the given breast cancer dataset. The performance of a classification technique is shown and summarized using a confusion matrix [50]. The anticipated and actual classification are shown in a confusion matrix of dimension $n \times n$ connected with a classifier, where n is the

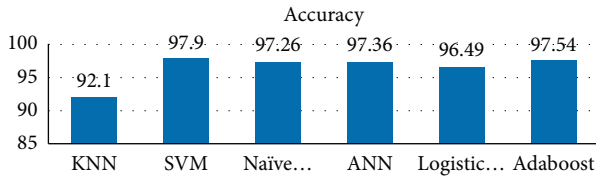


FIGURE 5: Accuracy.

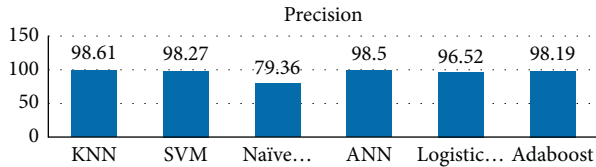


FIGURE 6: Precision.

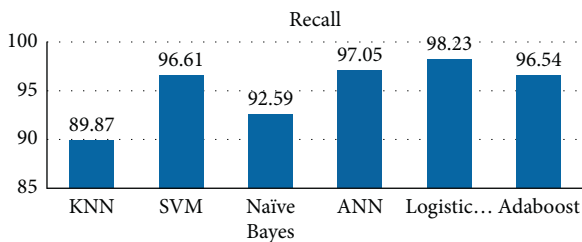


FIGURE 7: Recall.

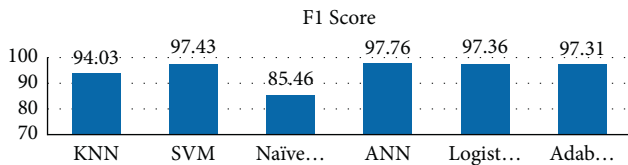


FIGURE 8: F1 score.

number of possible classes. A confusion matrix is used in this work [51] with the following interpretations for the entries.

A confusion matrix is frequently used to show the performance of a classifier based on the four values listed as TP, FP, TN, and FN [52–54]. The prediction accuracy and classification error [55] can be obtained from confusion matrix.

5.2. Results. The confusion matrix is used to assess method performance in classification. The results are then compared in the confusion matrix. In this matrix, 4 terms are used for evaluating performance, i.e., TP (true positive), FP (false positive), TN (true negative), and FN (false negative) [56–59].

Accuracy [60, 61] denotes correct figure of malignant patients whether it is positive or negative.

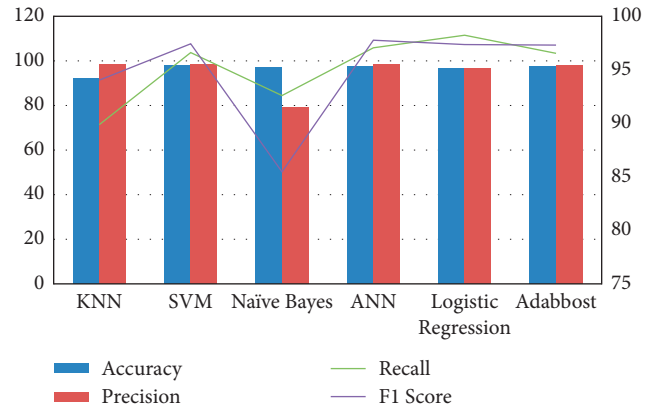


FIGURE 9: Comparative graph for breast cancer using supervised learning methods.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (8)$$

Precision is a metric used to determine how many of the positive forecasts are correct (true positives) [62, 63]. The formula is as follows.

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (9)$$

Among all the positive examples in the data, recall means the positive cases, which the classifier correctly predicted [64]. Sensitivity is a term that is occasionally used to describe it [65]. The formula is as follows.

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (10)$$

The F1 score [66] is a statistic that incorporates precision and recall. The harmonic mean of the two is a term used frequently [67]. The harmonic mean is a method of determining “average” of integers that is supposed to be superior than the arithmetic mean for ratios (such as precision and recall) [68]. In this situation, the F1 score [69] formula is as follows.

$$F1 \text{ score} = 2 * \left(\frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})} \right) \quad (11)$$

Six supervised learning algorithms were examined for their prognosis and detection capacities in the diagnosis of breast cancer [70, 71]. In this dataset, a maximum of 569 instances were utilized [72]. There are eight input parameters and one output parameter for each item. Figures 5–9 and Table 3 show the anticipated outcomes for the six suggested supervised learning algorithms (KNN, SVM, Naïve Bayes, ANN, logistic regression, and AdaBoost) based on the confusion matrix. Figure shows how the various approaches compare and depict all of the expected findings [73–75]. The anticipated results clearly show that, as compared to other algorithms, the bagging algorithm predicts breast cancer with the highest degree of accuracy. [56, 76].

TABLE 3: Comparison among various supervised learning methods.

ML method	Accuracy	Precision	Recall	F1 score
KNN	92.10	98.61	89.87	94.03
SVM	97.90	98.27	96.61	97.43
Naive Bayes	97.26	79.36	92.59	85.46
ANN	97.36	98.50	97.05	97.76
Logistic regression	96.49	96.52	98.23	97.36
AdaBoost	97.54	98.19	96.54	97.31

The support vector machine (SVM) gives the highest accuracy on the test dataset, which is 97.9%, as seen in this comparison table and graph. SVM has a higher accuracy percentage than other supervised learning algorithms. Support vector machine is the optimal algorithm for prediction, according to this research.

6. Conclusion

Breast cancer patients with a good prognosis and early discovery have a better chance of being cured. Finally, in the current study, several of the most prominent machine learning methods were utilized to classify “breast cancer” based on ten criteria. Some of the machine learning methods which are applied on this dataset are KNN, SVM, Naive Bayes, ANN, logistic regression, and AdaBoost with respective accuracy of 92.10%, 97.90%, 97.26%, 97.36%, 96.49%, and 97.54%. The defined methods have been applied on the prognosis of breast cancer. As a result, support vector machine (SVM) has been the most proficient algorithm among all supervised learning methods.

In future, some models can be implemented for supervised learning methods with some other techniques like genetic algorithm, ACO (ant colony optimization), PSO (particle swarm optimization), and so on so that early detection can be performed with efficient accuracy.

Data Availability

The data used to support the findings of this study are available from the author Vandana Rawat upon request (vandanarawat@geu.ac.in).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] T. V. Swathi, S. Krishna, and M. V. Ramesh, “A survey on breast cancer diagnosis methods and modalities,” In *Proceedings of the 2019 International Conference on Wireless Communications Signal Processing And Networking (WiSPNET)*, pp. 287–292, IEEE, Chennai, India, March 2019.
- [2] H. Dhahri, E. Al Maghayreh, A. Mahmood, W. Elkilani, and M. Faisal Nagi, “Automated breast cancer diagnosis based on machine learning algorithms,” *Journal of healthcare engineering*, pp. 1–11, 2019.
- [3] A. Bhoi, R. C. Balabantaray, D. Sahoo et al., “Mining Social media Text for Disaster Resource Management Using a Feature Selection Based on forest Optimization,” *Computers & Industrial Engineering*, vol. 169, Article ID 108280, 2022.
- [4] M. S. Mekala, G. Dhiman, G. Srivastava et al., “A DRL-based service offloading approach using DAG for edge computational orchestration,” *IEEE Transactions on Computational Social Systems*, pp. 1–9, 2022.
- [5] K. Yadav, A. Jain, N. M. Osman Sid Ahmed, S. A. Saad Hamad, G. Dhiman, and S. D. Alotaibi, “Internet of thing based koch fractal curve fractal antennas for wireless applications,” *IETE Journal of Research*, pp. 1–10, 2022.
- [6] T. Thomas, N. Pradhan, and V. S. Dhaka, “Comparative analysis to predict breast cancer using machine learning algorithms: a survey,” In *Proceedings of the 2020 International Conference on Inventive Computation Technologies (ICICT)*, pp. 192–196, IEEE, Coimbatore, India, February 2020.
- [7] B. Sumathy, A. Chakrabarty, S. Gupta et al., “Prediction of diabetic retinopathy using health records with machine learning classifiers and data science,” *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 2, pp. 1–16, 2022.
- [8] R. Nair, M. Soni, B. Bajpai, G. Dhiman, and K. M. Sagayam, “Predicting the death rate around the world due to COVID-19 using regression analysis,” *International Journal of Swarm Intelligence Research*, vol. 13, no. 2, pp. 1–13, 2022.
- [9] A. Alferaidi, Y. Kusum, Y. Alharbi et al., “Distributed deep CNN-lstm model for intrusion detection method in IoT-based vehicles,” *Mathematical Problems in Engineering*, vol. 2022, p. 8, Article ID 3424819, 2022.
- [10] P. W. Dalrymple, “Data, information, knowledge: the emerging field of health informatics,” *Bulletin of the American Society for Information Science and Technology*, vol. 37, no. 5, pp. 41–44, 2011.
- [11] H. Ding, X. Cao, Z. Wang et al., “Velocity clamping-assisted adaptive salp swarm algorithm: balance analysis and case studies,” *Mathematical Biosciences and Engineering*, vol. 19, no. 8, pp. 7756–7804, 2022.
- [12] M. S. Mekala, G. Srivastava, J. C. W. Lin, G. Dhiman, J. H. Park, and H. Y. Jung, “An Efficient Quantum Based D2D Computation and Communication Approach for the Internet of Things,” *Opt Quant Electron*, vol. 54, 2021.
- [13] S. Swain, B. Bhushan, G. Dhiman, and W. Viriyasitavat, “Appositeness of optimized and reliable machine learning for healthcare: a survey,” *Archives of Computational Methods in Engineering: State of the Art Reviews*, pp. 1–23, 2022.
- [14] P. M. Griffin and H. B. Nembhard, *Health Data and Informatics*, John Wiley & Sons, Hoboken, NJ, USA, 2016.
- [15] W. Viriyasitavat, L. Da Xu, G. Dhiman, A. Sapsomboon, V. Pungpapong, and Z. Bi, “Service Workflow: State-Of-The-Art and Future Trends,” *IEEE Transactions on Services Computing*, 2021, In press.
- [16] G. Dhiman and A. R. Yildiz, “Nature-inspired algorithms for real-life complex engineering problems,” *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, vol. 14, no. 3, p. 251, 2021.
- [17] K. K. Singamaneni, K. Ramana, G. Dhiman, S. Singh, and B. Yoon, “A novel blockchain and Bi-linear polynomial-based QCP-ABE framework for privacy and security over the complex cloud data,” *Sensors*, vol. 21, no. 21, p. 7300, 2021.
- [18] D. Hanahan and R. A. Weinberg, “The hallmarks of cancer,” *Cell*, vol. 100, no. 1, pp. 57–70, 2000.
- [19] V. R. Brus, L. I. Voronova, and V. I. Voronov, “Neural network classification of cardiac activity based on cardiogram data for driver support system,” In *Proceedings of the 2020*

- Systems of Signals Generating and Processing in the Field of on Board Communications*, pp. 1–5, Moscow, Russia, 2020.
- [20] R. Kumar, V. Joshi, G. Dhiman, and W. Viriyasitavat, “An improved exponential metric space approach for C-mean clustering analysing,” *Expert Systems*, Article ID e12896, 2021.
- [21] S. P. Singh, G. Dhiman, P. Tiwari, and R. H. Jhaveri, “A soft computing based multi-objective optimization approach for automatic prediction of software cost models,” *Applied Soft Computing*, vol. 113, Article ID 107981, 2021.
- [22] M. Soni, G. Dhiman, B. S. Rajput, and K. N. Tejra, “Energy-Effective and Secure Data Transfer Scheme for Mobile Nodes in Smart City Applications,” *Wireless Persedictive analytics for chronic kidney disease using machine learning technique Commun*, 2021.
- [23] A. Holzinger, “Machine learning for health informatics,” in *Machine Learning for Health Informatics* Springer, New York, NY, USA, 2016.
- [24] A. Charleonnann, T. Fufaung, T. Niyomwong, W. Chokchueyattanakit, S. Suwannawach, and N. Ninchawee, “Predictive analytics for chronic kidney disease using machine learning techniques,” In *Proceedings of the 2016 Management and Innovation Technology International Conference (MITIcon)*, Bang-San, Thailand, 2016.
- [25] E. Boonchieng and K. Duangchaemkarn, “Digital disease detection: application of machine learning in community health informatics,” In *Proceedings of the 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 1–5, Khon Kaen, Thailand, July 2016.
- [26] R. Bhardwaj, A. R. Nambiar, and D. Dutta, “A study of machine learning in healthcare,” In *Proceedings of the 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, pp. 236–241, Turin, Italy, 2017.
- [27] A. P. Tafti, E. Larose, J. C. Badger, R. Kleiman, and P. Peissig, “Machine learning-as-a service and its application to medical informatics,” in *Machine Learning and Data Mining in Pattern Recognition*, Springer, Cham, Berlin, Germany, 2017.
- [28] B. Nithya and V. Ilango, “Predictive analytics in health care using machine learning tools and techniques,” In *Proceedings of the 2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 492–499, Madurai, Tamilnadu, 2017.
- [29] H. Sharma, C. Mao, Y. Zhang, and H. Vatani, “Portable Phenotyping system: A Portable Machine-Learning Approach to i2b2 Obesity Challenge,” In *Proceedings of the 2018 IEEE International Conference on Healthcare Informatics Workshop (ICHI-W)*, pp. 86–87, New York, NY, USA, 2018.
- [30] F. Ahamed and F. Farid, “Applying internet of things and machine-learning for personalized healthcare: issues and challenges,” In *Proceedings of the 2018 International Conference on Machine Learning and Data Engineering (iCMLDE)*, pp. 19–21, Sydney, Australia, 2018.
- [31] R. Sivakani and G. A. Ansari, “Machine learning framework for implementing alzheimer’s disease,” In *Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0588–0592, Chennai, India, 2020.
- [32] M. Moreb, T. A. Mohammed, and O. Bayat, “A novel software engineering approach toward using machine learning for improving the efficiency of health systems,” *IEEE Access*, vol. 8, Article ID 23169, 2020.
- [33] R. Katarya and P. Srinivas, “Predicting Heart Disease At Early Stages Using Machine Learning: A Survey,” In *Proceedings of the 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 302–305, Coimbatore, India, 2020.
- [34] C. H. Pan and G. Dhiman, “Study on automatic tracking method of marking points in sports image sequence,” *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, vol. 14, no. 7, pp. 708–717, 2021.
- [35] M. S. Mekala, G. Dhiman, R. Patan et al., “Deep learning-influenced joint vehicle-to-infrastructure and vehicle-to-vehicle communication approach for internet of vehicles,” *Expert Systems*, vol. 39, no. 5, Article ID e12815, 2021.
- [36] H. O. Alanazi, A. H. Abdullah, and K. N. Qureshi, “A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care,” *Journal of Medical Systems*, vol. 41, no. 4, p. 69, 2017.
- [37] R. Kumar and G. Dhiman, “A comparative study of fuzzy optimization through fuzzy number,” *International Journal of Modern Research*, vol. 1, no. 1, pp. 1–14, 2021.
- [38] P. K. Vaishnav, S. Sharma, and P. Sharma, “Analytical review analysis for screening COVID-19 disease,” *International Journal of Modern Research*, vol. 1, no. 1, pp. 22–29, 2021.
- [39] F. Ahamed and F. Farid, “Applying Internet of Things and machine-Learning for personalized healthcare: Issues and challenges,” In *Proceedings of the 2018 International Conference on Machine Learning and Data Engineering (iCMLDE)*, p. 19, Sydney, Australia, December 2018.
- [40] K. Govindan, H. Mina, and B. Alavi, “A Decision Support System for Demand Management in Healthcare Supply Chains Considering the Epidemic Outbreaks: A Case Study of Coronavirus Disease 2019(COVID-19),” *Transp Res E Logist Transp Rev*, vol. 138, 2019.
- [41] Z. Rayan, M. Alfonse, and A.-B. M. Salem, “Machine learning approaches in smart health,” *Procedia Computer Science*, vol. 154, pp. 361–368, 2019.
- [42] V. K. Gupta, S. K. Shukla, and R. S. Rawat, “Crime tracking system and people’s safety in India using machine learning approaches,” *International Journal of Modern Research*, vol. 2, no. 1, pp. 1–7, 2022.
- [43] T. Sharma, R. Nair, and S. Gomathi, “Breast cancer image classification using transfer learning and convolutional neural network,” *International Journal of Modern Research*, vol. 2, no. 1, pp. 8–16, 2022.
- [44] S. K. Shukla, V. K. Gupta, K. Joshi, A. Gupta, and M. K. Singh, “Self-aware execution environment model (SAE2) for the performance improvement of multicore systems,” *International Journal of Modern Research*, vol. 2, no. 1, pp. 17–27, 2022.
- [45] K. G. Sheela and R. V. Anu, “MachineLearningbasedHealthMonitoringSystem,” *Materials Today Proceedings*, vol. 24, pp. 1788–1794, 2020.
- [46] K. Munir, H. Elahi, A. Ayub, F. Frezza, and A. Rizzi, “Cancer diagnosis using deep learning: a bibliographic review,” *Cancers*, vol. 11, no. 9, p. 1235, 2019.
- [47] R. X. Zhang, G. B. Huang, N. Sundararajan, and P. Saratchandran, “Multicategory classification using extreme learning machine for microarray gene expression cancer diagnosis,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, pp. 485–495, 2007.
- [48] J. A. Cruz and D. S. Wishart, “Applications of Machine Learning in Cancer Prediction and Prognosis,” *Cancer Informatics*, vol. 2, pp. 59–77, 2006.
- [49] W. L. Bi, A. Hosny, M. B. Schabath et al., “Artificial intelligence in cancer imaging: clinical challenges and applications,” *CA: A Cancer Journal for Clinicians*, vol. 69, no. 2, pp. 127–157, 2019.

- [50] S. S. Mahdavi, M. Moradi, W. J. Morris, S. L. Goldenberg, and S. E. Salcudean, "Fusion of ultrasound B-mode and vibro-elastography images for automatic 3-D segmentation of the prostate," *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2073–2082, 2012.
- [51] L. Wang, F. Chu, and W. Xie, "Accurate cancer classification using expressions of very few genes," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 1, pp. 40–53, 2007.
- [52] M. Liang, Z. Li, T. Chen, and J. Zeng, "Integrative data analysis of multi-platform cancer data with a multimodal deep learning approach," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 12, no. 4, pp. 928–937, 2015.
- [53] A. P. Athreya, A. J. Gaglio, J. Cairns et al., "Machine learning helps identify new drug mechanisms in triple-negative breast cancer," *IEEE Transactions on Nanobioscience*, vol. 17, no. 3, pp. 251–259, 2018.
- [54] S. Hussein, P. Kandel, C. W. Bolan, M. B. Wallace, and U. Bagci, "Lung and pancreatic tumor characterization in the deep learning Era: novel supervised and unsupervised learning approaches," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1777–1787, 2019.
- [55] A. A. R. Magna, H. Allende-Cid, C. Taramasco, C. Becerra, and R. L. Figueroa, "Application of machine learning and word embeddings in the classification of cancer diagnosis using patient Anamnesis," *IEEE Access*, vol. 8, Article ID 106198, 2020.
- [56] N. Fatima, L. Liu, S. Hong, and H. Ahmed, "Prediction of breast cancer comparative review of machine learning techniques and their analysis," *IEEE Access*, vol. 8, pp. 150 360–150 376, 2020.
- [57] S. Iqbal, G. F. Siddiqui, A. Rehman et al., "Prostate cancer detection using deep Learning and traditional techniques," *IEEE Access*, vol. 9, Article ID 27085, 2021.
- [58] M. Marsden and B. Weyers, "Intraoperative margin assessment in head and neck cancer using label-free fluorescence lifetime imaging, machine learning and visualization," *Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems*, vol. 11631, Article ID 116310, N, D.G, 2021.
- [59] G. Xu, N. Yin, and Q. Yang, "Study on the feasibility to detect cancer tumors by combining nanotechnology with SQUID" applied superconductivity," *IEEE Transactions on*, vol. 20, no. 3, pp. 1956–1959, 2010.
- [60] D. C. Nguyen and F. Azadivar, "Application of computer simulation and genetic algorithms to gene interactive rules for early detection and prevention of cancer," *IEEE Systems Journal*, vol. 8, no. 3, pp. 1005–1013, 2014.
- [61] M. Al Ahmad, A. Najar, A. El Moutaouakil et al., "Label-free cancer cells detection using optical sensors," *IEEE Access*, vol. 6, Article ID 55807, 2018.
- [62] M. H. Waseem, M. S. A. Nadeem, A. Abbas et al., "On the feature selection methods and reject option classifiers for robust cancer prediction," *IEEE Access*, vol. 7, Article ID 141072, 2019.
- [63] C. Liu, Y. Huang, J. A. Ozolek, M. G. Hanna, R. Singh, and G. K. Rohde, "SetSVM: an approach to set classification in nuclei-based cancer detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 351–361, 2019.
- [64] Q. Zhou, B. Yong, Q. Lv, J. Shen, and X. Wang, "Deep autoencoder for mass spectrometry feature learning and cancer detection," *IEEE Access*, vol. 8, Article ID 45156, 2020.
- [65] T. K. BamunuMudiyanselage, X. Xiao, Y. Zhang, and Y. Pan, "Deep fuzzy neural networks for biomarker selection for accurate cancer detection," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 12, pp. 3219–3228, 2020.
- [66] Y. Rejani and S. T. Selvi, "Early Detection of Breast Cancer Using SVM Classifier Technique," 2009, <https://arxiv.org/ftp/arxiv/papers/0912/0912.2314.pdf#:~:text=In%20the%20field%20of%20medical,fit%20in%20a%20feature%20space>.
- [67] B. Y. Sun, Z. H. Zhu, J. Li, and B. Linghu, "Combined feature selection and cancer prognosis using support vector machine regression," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 6, pp. 1671–1677, 2011.
- [68] M. Varjonen, "Three-dimensional digital breast tomosynthesis in the early diagnosis and detection of breast cancer," in *International Workshop on Digital Mammography* Springer, Berlin, Germany, 2006.
- [69] H. Saoud, A. Ghadi, and M. Ghailani, "Proposed approach for breast cancer diagnosis using machine learning," In *Proceedings of the 4th International Conference On Smart City Applications*, pp. 1–5, Casablanca, Morocco, October 2019.
- [70] A. F. M. Agarap, "On breast cancer detection: an application of machine learning algorithms on the Wisconsin diagnostic dataset," In *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing*, pp. 5–9, New York, NY, USA, February 2018.
- [71] B. Al-Shargabi and F. A. Al-Shami, "An experimental study for breast cancer prediction algorithms," In *Proceedings of the Second International Conference on Data Science*, pp. 1–6, E-Learning And Information Systems, Dubai, UAE, December 2019.
- [72] N. Eyal, M. Last, and E. Rubin, "Comparison of three classifiers for breast cancer outcome prediction," In *Proceedings of the 16th International Conference on Engineering Applications of Neural Networks (INNS)*, pp. 1–6, September 2015.
- [73] R. Shen, Y. Yang, and F. Shao, "Intelligent breast cancer prediction model using data mining techniques," In *Proceedings of the 2014 Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*, vol. 1, pp. 384–387, IEEE, August 2014.
- [74] S. Ara, A. Das, and A. Dey, "Malignant and benign breast Cancer Classification using Machine Learning Algorithms," In *Proceedings of the 2021 International Conference on Artificial Intelligence (ICAI)*, pp. 97–101, IEEE, Slamabad, Pakistan, April 2021.
- [75] P. H. Abreu, M. S. Santos, M. H. Abreu, B. Andrade, and D. C. Silva, "Predicting breast cancer recurrence using machine learning techniques: a systematic review," *ACM Computing Surveys*, vol. 49, no. 3, pp. 1–40, 2016.
- [76] J. T. Wassan, H. Wang, F. Browne, and H. Zheng, "A comprehensive study on predicting functional role of metagenomes using machine learning methods," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 3, pp. 751–763, 2019.