A Prognosis and Prediction of Breast Cancer using Machine Learning Techniques

Ms.Poonam R.Dholi [#1], Dr Varsha H. Patil [*2]

*#1 Asst. Prof. Department of Computer Engineering Matoshri College of Engineering & Research Center Nashik (M.S.) dholipoonam03@gmail.com, *2 Vice-Principal & HOD Department of Computer Engineering Matoshri College of Engineering & Research Center Nashik (M.S.) varsha.patil@gmail.com*

Abstract

Breast cancer is one of the major cause of cancers occurring in women compared to all other cancers. It is the most common type of all cancers and the main cause of women's death worldwide. Early prediction of breast cancer plays an important role in successful treatment and saving lives of number of patient every year. Today, it has become a hot issue in cancer research to make precise prognostic prediction for breast cancer patients. Research and prevention on breast cancer have more attention of research in recent years. Classification and machine learning methods are an effective way to classify data. Particularly in medical field, where those methods are extensively used in diagnosis and analysis to make decisions. In this paper, a performance comparison between different machine learning algorithms: Support Vector Machine (SVM), Decision Tree (J48), Naive Bayes (NB), K- Nearest Neighbor (KNN) and Logistic Regression on the Wisconsin Breast Cancer (original) datasets is conducted. The main intention is to calculate the correctness in classifying data with respect to efficiency of each algorithm in terms of accuracy, sensitivity, specificity, error rate and Time(seconds). Experimental results show that SVM gives the highest accuracy (97.59%) with lowest error rate. WEKA simulation tool used for all experiments.

Keywords: Diabetes Breast cancer, data mining ,Support vector Machine, Decision Tree, Naive Bays, K- Nearest Neighbor, Logistic Regression

1. Introduction

Breast cancer is one of the dangerous diseases in the most common of all cancers and is the leading cause of cancer deaths in women worldwide. Uncontrolled growth of breast cells causes breast cancer. The cause of breast cancer are not yet fully known, although a number of risk factors have been identified such a genetic risk, family history of breast cancer, age, overweight, late menopause, late age at first childbirth, age [3][21][22].

The classification of Breast cancer data can be useful to predict the outcomes of some diseases to discover the genetic behavior of tumors. The main role of breast cancer diagnosis is to distinguish between the malignant and benign breast masses, while the prognosis estimates recurrence of disease, predicts survival of patient and helps in establishing a treatment plan predicting the outcome of a disease. The machine learning and datamining techniques have create a new ways for healthcare diagnostic and prediction. Datamining is a powerful technology that can be used in all domains in order to detect hidden pattern from a large volume of data. In this paper, we utilize this advancement for predicting breast cancer risk based on a labeled datasets. Particularly, we utilize the clinical data as well as associated symptoms of patients to construct predictive model which can classify patient into different breast cancer categories, i.e. benign or maligent [4].

The organization of the paper is as follows. In Section 2, overview of related work on the use of DM techniques for breast cancer. In Section 3, describe our dataset, methods and background on data mining techniques. Presentation and discussion of our results in Section 4. We conclude our paper in Section 5.

International Conference on Contents, Computing & Communication (ICCCC-2022)

(**Available on SSRN. SSRN is an open-access online preprint community, owned by Elsevier**.)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

2. Related Work

A number of studies have been accounted for that have zeroed in on breast cancer malignant growth. These methods have applied various ways to deal with the given issue and accomplished high characterization exactness. Details of some of the preceding research works are given in the following.

Vikas Chaurasia et al. represents a report on available advance technologies to develop prediction model for breast cancer. In this paper to develop the prediction model no.of data mining algorism are applied like Naive Bayes, RBF Network, J48 using a large dataset (683 breast cancer cases). Also apply 10-fold cross-validation methods to measure the impartial estimate of the three prediction models for performance evaluation. The results based on average accuracy Breast Cancer dataset shows that the Naive Bayes is the best predictor with 97.36% accuracy on the given sample, RBF Network gives 96.77% accuracy, J48 gives 93.41% accuracy [3].

S.Kharya et al. The main objective of this study is to show the efficiency of these classifiers and their ensemble in the prediction of breast mass severity. This paper mainly focuses on Tree augmented Naïve Bayes and Markov blanket estimation learning algorithms techniques were applied on mammographic mass dataset. Bayesian network classifiers outperformed the multilayer perceptron neural network on the prediction of the severity of breast masses and they provide an proper way to rank the attributes that most extensively indicate the probability of default. On the basis of analysis concluded that the Bayesian network classifiers give remarkable results [6].

D.Lavanya et al. paper focuses on the decision tree classifier, which is based on with or without feature selection technique in terms of accuracy, size of tree and time to build a model on various breast cancer datasets. According to the result CART has improved accuracy of feature selection of a particular dataset [7]. Ilias Maglogiannis et al. pointed in on the execution of SVM based classification algorithms nearly with elective techniques (Naïve Bayes, Bayesian Nets and Artificial Neural Nets) for the analysis and visualization of breast cancer disease. The primary reason for the paper is to test the capacity of our proposed SVM and ANN models for the Wisconsin anticipation and determination breast cancer disease issue in contrast with different classifiers. The optimized SVM algorithm performed remarkably, exhibiting high values of accuracy up to 96.91%, specificity up 97,67% and sensitivity up to 97,84%, superior than the alternative approaches for all classes. Similarly at a second stage, another optimized SVM classifier, was implemented for automated diagnosis of the Wisconsin Diagnostic Breast Cancer datasets and the accuracy was approximately 97%, while sensitivity and specificity indices were also satisfactory [11]. Leyi Wei et al. proposed a novel random-forest-based predictor called MePred-RF, MePred-RF is the sequence based information rather than techniques based on multiple and complex information as a input. It improve feature representation capability using powerful feature selection techniques [12]. Enli Zhang et al. proposed a novel regularized multi-view subspace clustering (rMV-spc) algorithm to obtain a representation matrix for each stage that balances concurrence across various stages. To avoid a heterogeneity of data the protein interaction network is incorporated into the rMV-spc through regularization. To solve the optimization problem using the interior point algorithm to obtain common modules. The proposed algorithm is based on artificial intelligence which discovers common modules in breast cancer dataset. These modules serve as biomarkers to predict stages of cancer.   The proposed model and algorithm effectively integrate heterogeneous data for dynamic modules [13]. SungHwan Kim proposed the new algorithm for the weighted KM-SVM to improve prediction accuracy. Typically, the KM-SVM has higher error rate than that it appears in the SVM, due to data reduction. To avoid this issue, suggested that the weighted KM-SVM (and SVM) and evaluated performance of each of classifiers through various experimental scenarios. Considering both issue, concluded that the proposed weighted KM-SVM (and SVM) is effective to diminish its

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

error rates. Particularly, after application of KM-SVM on TCGA cancer methylation data, results improved disease prediction performance [14]. Ujjwal Maulik at al. attempt a prediction scheme that combines fuzzy preference based rough set (FPRS) method for feature (gene) selection with semisupervised SVMs. To show the effectiveness of the proposed approach, compare the performance of this technique with the signal-to-noise ratio (SNR) and consistency based feature selection (CBFS) methods. Using six benchmark gene microarray datasets (including both binary and multi-class classification problems), demonstrate experimentally that proposed method can reach up to important success and is biologically applicable for cancer diagnosis and drug discovery [15]. Lin Zhang et al. proposed a Gene Interaction Regularized Elastic Net (GIREN) model based on integrating multiple data types. It can easily adopt gene measurement and gene to gene interaction information using elastic net formulation, enforcing structure sparsity and the "grouping effect" in solution to select the different features with prognostic value. An iterative gradient descent algorithm was also developed to solve the model with regularized optimization. GIREN was applied to human ovarian cancer and breast cancer datasets obtained from The Cancer Genome Atlas, respectively. By comparing the results of proposed system with competing algorithms (LASSO, Elastic Net and superPC, with or without average pathway expression features), GRIEN obtained more accurate results [17]. B.Padmapriya et al. analyzes the current research being carried out using the machine algorithms and datamining techniques to enhance the breast cancer diagnosis. This survey particularly focused on the classification algorithm ID3 and C4.5 in breast cancer. This research work concluded that the performance of C4.5 is better than the other algorithms [18]. Guoyin Zhang et al. propose a mobile medical health system based on cloud computing. Initially, head part examination was utilized to acquire agent highlights. Then, at that point, a worked on highlight subset was applied to help vector machine (SVM) in light of Sigmoid portion work. The informational collection was sorted by SVM as malignant growth patient and ordinary article trial results show that the strategy is worked on in accuracy, sensitivity, and specificity[19]. Deepika Verma et al. classified the five classification algorithms on WEKA Explorer and WEKA Experimenter interface. WEKA device is a decent grouping apparatus utilized in this paper. Naive bayes, SMO, REP Tree, J48 and MLP algorithms are utilized to order breast cancer malignant growth and diabetes dataset on WEKA interface. The exhibitions of these five calculations have been broke down on breast cancer disease and diabetes dataset utilizing preparing information testing mode. In the wake of breaking down the exhibitions of all algorithm, observed that Naive bayes gives 72.70%accuracy on breast cancer malignant growth dataset and SMO gives 76.80%accuracy on diabetes dataset[20]. D. A. Omondiagbe et al. applied features selection and features extraction techniques on Artificial Neural Network (ANN), Support Vector Machine (SVM) and Naive Bayes (NB) for the prediction of breast cancer using Wisconsin Diagnostic Breast Cancer (WDBC) Dataset. Coloration based feature selection (CFS), Linear Discriminant Analysis and Recursive Feature Elimination (RFE) were used as a feature selection techniques for comparative analysis. After the comparative analysis with different features selection methods, authors found that the accuracy rate of Artificial Neural Network was higher with 97% than the other algorithms. The accuracy of Support Vector Machine was 96.4% and Naive Bayes accuracy was 91% [24].

3. Machine Learning and Modeling Approaches

3.1 Experiment Environment

Experiments describe in this paper are conducted using WEKA libraries. WEKA provides a machine learning environment for preprocessing, regression, association rules, classification and clustering. Many real times

International Conference on Contents, Computing & Communication (ICCCC-2022)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

based research are executed using WEKA tool .This tool offers a frame work for developer to build and evaluate their models.

3.2 Breast cancer dataset

The Wisconsin Breast Cancer (original) datasets from the UCI Machine Learning Repository is used in this study. Breast-cancer-Wisconsin has 699 instances (Benign: 499 Malignant: 200), 2 classes (70% malignant and 30% benign), and 11 integer-valued attributes.

Table 1: Sample of Dataset

ID	Clump thickness	Uniformity of cell size	Uniformity of cell shape	Marginal adhesion	Single epithelial cell size	Bare Nuclei	Bland Chromatin	Normal nucleoli	Mitoses	Class
1	5	1	1	1	2	1	3	1	1	-1
2	5	4	4	5	7	10	3	2	1	-1
3	3	1	1	1	2	2	3	1	1	-1
4	6	8	8	1	3	4	3	7	1	-1
5	4	1	1	3	2	1	3	1	1	-1
6	8	10	10	0	8	7	10	7	1	1
7	1	1	1	1	2	10	3	1	1	-1
8	2	1	2	1	2	1	3	1	1	-1
9	2	1	1	1	2	1	1	1	5	-1
10	4	2	1	1	2	1	2	1	1	-1

Table 2: Data Attributes

Sr. No.	Variable	Attribute	Domain
1	ID	Sample code number	ID number
2	Clump_Thickness	Clump thickness	1—10
3	Cell_Size	Uniformity of cell size	1—10
4	Cell_shape	Uniformity of cell shape	1—10
5	Adhesion	Marginal adhesion	1--10
6	Epi_Cell_Size	Single epithelial cell size	1--10
7	Nuclei	Bland nuclei	1--10
8	Chromatin	Bland chromatin	1—10
9	Nucleoli	Normal nucleoli	1--10
10	Mitoses	Mitoses	1--10
11	Class	Target variables	2-Benign,

International Conference on Contents, Computing & Communication (ICCCC-2022)

4

(**Available on SSRN. SSRN is an open-access online preprint community, owned by Elsevier**.)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

			4-Malignant

3.3 Data Mining Classification Methods

The data mining comprises of different strategies. Various strategies fill different needs, every strategy offering its advantages and disadvantages. The medical science uses classification and clustering techniques of data mining for prognosis and prediction. The classification category mainly used for prediction as patients to either a non- cancerous that is "benign" group or a cancerous that is "malignant" group and generate convention for the same. Hence the diagnostic of breast cancer are the broadly discussed classification problems. The classification is the most important step of data mapping regarding predefined targets in data mining.

  In data mining, classification is one of the most important tasks which maps data in to predefined targets means it's targets are supervised learning . The main aim of classification is to build the classifier based model which is based on attributes and object of some domain based cases. And then the classifier model is also used to predict the group of attributes related to domain based new cases. The commonly used data mining techniques for classification can be classified into the following groups.

Naive Baye's (NB)
The Naive Baye's is a quick technique for making of measurable prescient models. NB depends on the Bayesian hypothesis. This characterization method examinations the connection between each characteristic and the class for each occurrence to infer a contingent likelihood for the connections between the property estimations and the class. During preparing, the likelihood of each class is calculation depends on the no of events in the preparation dataset. This is known as the " prior probability " P(C=c). Not with standing earlier likelihood, the calculation likewise processes likelihood for the occurrence x given c with the presumption that the qualities are free. This likelihood turns into the result of the probabilities of each single characteristic. Then, at that point, the probabilities can be unsurprising from the frequencies of the examples in the training set.

Decision Trees (J48)
Decision tree is a tree where each non-terminal hub addresses a test or choice on the thought about information thing. Decision of a specific branch relies on the result of test. In decision tree to classify particular data item, traversing starts from root node and follow the assertions down up to reach to the leaf node or terminal node. Decision trees are nothing but the interpretation of special form of rules which is organized in a hierarchical order.
Decision tree can be utilized to order a model by beginning at the base of the tree and traveling through it until a leaf hub is reached, which gives the classification of the instance.

Logistic Regression(LR)
Logistic Regression is a supervised learning algorithm that includes more dependent variables. This algorithm is responses in the binary form. Logistic regression [23] can provide continuous outcome of a specific data. This algorithm is considered as the standard statistical model with binary variables. It is better option for linear regression which appoints a direct model to every one of the class and predicts concealed examples base on

(**Available on SSRN. SSRN is an open-access online preprint community, owned by Elsevier**.)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

larger part vote of the models.

K-Nearest Neighbor(KNN)

K-Nearest Neighbor algorithm is used in pattern recognition. It is a good approach for breast cancer prediction. In order to recognize the pattern, each class has given an equal importance. An item arranged by greater part of its neighbors. K is generally a positive number. The neighbors are chosen from a bunch of objects for which the right classification is known.

Support vector machine(SVM)

support vector machine are supervised learning models which can be used for classification or regression. Mainly the strategy is utilized to observe a hyper plane that isolates the instances of various results. Generally it is intended for two-class issues, SVMs track down a hyper plane with a greatest distance to the nearest point of the two classes; such a hyper plane is known as the optimal hyper plane. A bunch of occurrences that is nearest to the optimal hyper plane is known as a support vector [6].

4.   Results and Discussion

In predictive analytics a table of confusion matrix with two rows and two columns that reports the number of False Positives (FP), False Negatives (FN), True Positives (TP), and True Negatives (TN). This allows more detailed analysis than simple proportion of correct classifications (accuracy). Accuracy is not a reliable metric for the real performance of a classifier, because it will give confusing results if the dataset is unbalanced. Using WEKA tool, on the basis of the confusion matrix calculate the accuracy is predicted in terms of Sensitivity and Specificity for the methods Naive Baye's, Decision Tree, K- Nearest Neighbor, Logistic Regression and Support vector machine.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Table 3: Confusion Matrix of Training and Testing data

Algorithm	Training Data (499)			Testing Data (200)		
	Desired Result	Output Result		Desired Result	Output Result	
		Benign	Malignant		Benign	Malignant
J48	Benign	308	13	Benign	129	8
	Malignant	9	169	Malignant	8	55
Naïve Bayes	Benign	310	11	Benign	128	9
	Malignant	5	173	Malignant	2	61
Logistic	Benign	314	7	Benign	9	54
	Malignant	9	169	Malignant	131	6

(**Available on SSRN. SSRN is an open-access online preprint community, owned by Elsevier**.)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

SVM	Benign	315	6	Benign	131	6
	Malignant	6	172	Malignant	5	58
KNN	Benign	314	7	Benign	130	7
	Malignant	17	161	Malignant	5	58

Table 4: Performance analysis of Training and Testing data

Algorithms	Training Data					Testing Data				
	Accuracy	Sensitivity	Specificity	Error rate	Time	Accuracy	Sensitivity	Specificity	Error rate	Time
J48	95.59	0.96	0.949	4.41	0.09	92	0.942	0.873	8	0.08
Naïve Bayes	96.79	0.96	0.972	3.21	0.05	94.5	0.934	0.968	5.5	0.05
Logistic	96.79	0.978	0.949	3.21	0.23	92.5	0.956	0.857	7.5	0.23
SVM	97.59	0.981	0.966	2.41	0.73	94.5	0.956	0.921	5.5	0.69
KNN	95.19	0.978	0.904	4.81	0	94	0.949	0.921	6	0

According to the performance of the machine learning techniques Support Vector Machine (SVM) performed the best with accuracy 97.9% and low error rate i.e. 2.41% in both training and testing data. Naïve Bayes and Logistic regression performed with accuracy 96.79% and error rate 3.21 for training data and for testing data NB performed with accuracy 94.5% with error rate 5.5 and for Logistic performed with accuracy 92.5% with error rate 7.5%, Decision Tree performed with accuracy 95.59% with error rate 4.41% for training data and for testing data accuracy 92% with error rate 8%, K- Nearest Neighbor performed with accuracy 95.19% and error rate 4.81% for training data and for testing data accuracy 94% with 0 error rate as shown in figure1 and figure 2.

International Conference on Contents, Computing & Communication (ICCCC-2022)
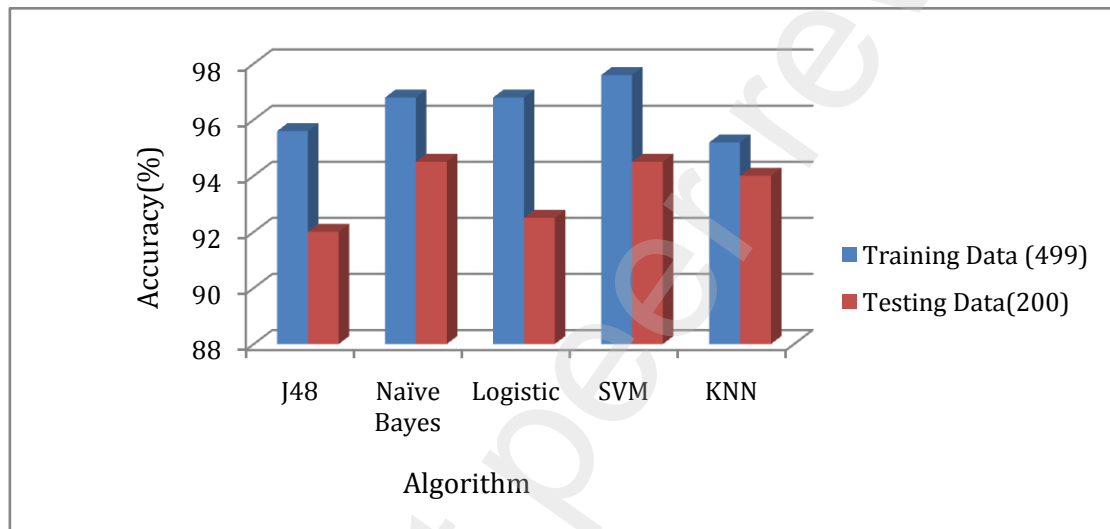
26th and 27th February,2022

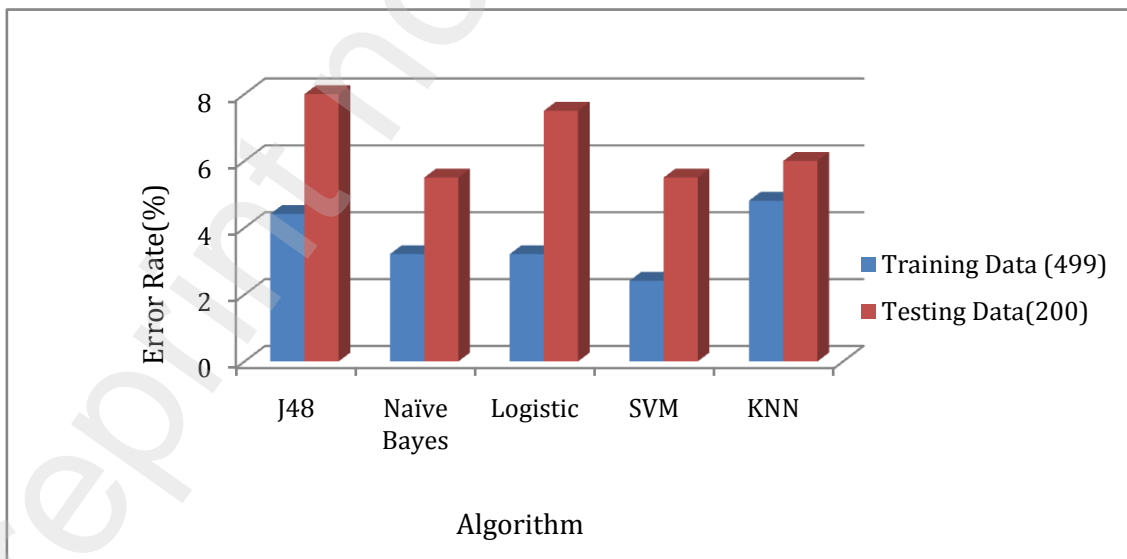Figure 1: Accuracy of Classification methods

Figure 2: Error rate of Classification methods

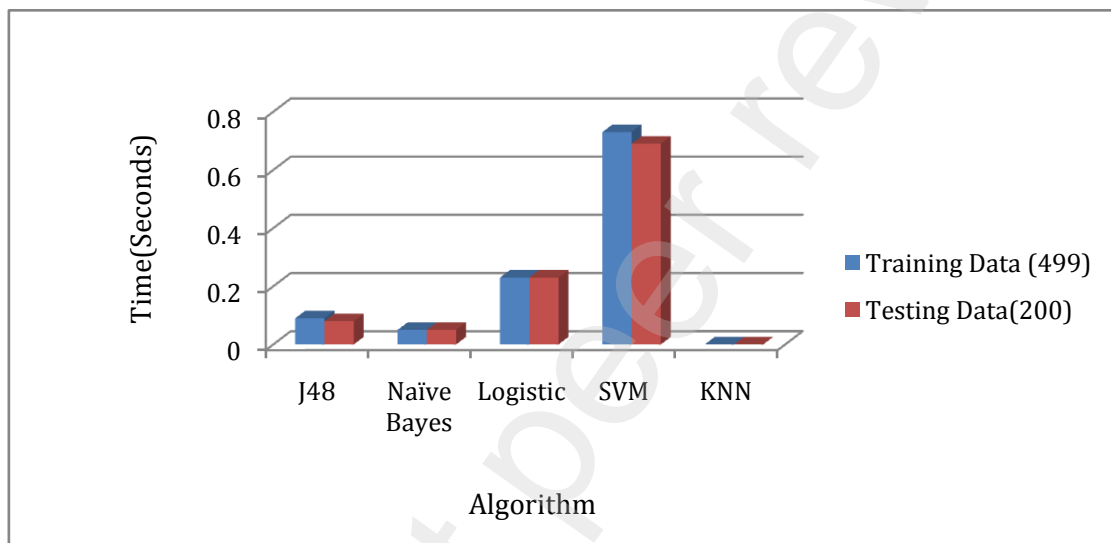**International Conference on Contents, Computing & Communication (ICCCC-2022)**

Figure 3: Execution Time of Classification methods

5. Conclusion

In this study, the technique for DT, NB, KNN, Logistic and SVM models are utilized as the order to foresee the idea of breast cancer malignant growth with different traits. The after effects of expectation will serve to determinations of sickness and the reasonable therapy for treatment. The dataset is utilized in this study is WBCD which contains 699 records and 11 attributes. After pre-processes the crude information of WBCD dataset and acquired the information that contains 683 records with 9 attributes and the file showing whether the volunteer has the malignant tumour. After comparing the accuracy of five classification models, the result has revealed that SVM is selected as the primary classification model in this study. Now a day, the factors of malignant breast cancer become more and more complex. The restricted crude information affects the exactness of results. Likewise, the SVM additionally can be joined with different information mining methods to acquire more precise and effective outcomes in the future work.

References

[1] Hamid Karim Khani Zand , "A Comparitive Survey On Data Mining Techniques for Breast Cancer Diagnosis and Prediction" Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 , 2015 Vol.5 (S1), pp. 4330-4339/Karim

[2] Rozita Jamili Oskouei1, Nasroallah Moradi Kor2,3, Saeid Abbasi Maleki,"Data mining and medical world: breast cancers' diagnosis, treatment, prognosis and challenges"

9

(**Available on SSRN. SSRN is an open-access online preprint community, owned by Elsevier**.)

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

Am J Cancer Res 2017;7(3):610-627

[3] Vikas Chaurasia1, Saurabh Pal1 and BB Tiwari,"Prediction of benign and malignant breast   cancer using data mining techniques" Journal of Algorithms & Computational Technology 2018, Vol. 12(2) 119–126

[4]   Tuan Tran and Uyen Le,"Predicting Breast Cancer Risk: A Data Mining Approach" 6th International Conference on the Development of Biomedical Engineering in   Vietnam (BME6), IFMBE Proceedings 63

[5] Yixuan Li1, Zixuan Chen2, "Performance Evaluation of Machine Learning Methods for Breast   Cancer Prediction" Applied and Computational Mathematics 2018; 7(4): 212-216

[6] S.Kharya , D. Dubey , and S. Soni," Predictive Machine Learning Techniques for Breast Cancer Detection", International Journal o Computer Science and Information   Technologies, Vol. 4 (6) , 2013, 1023-1028.

[7] D.Lavanya, Dr.K.Usha Rani,"Analysis of Feature Selection with Classfication: Breast Cancer   Datasets", Indian Journal of Computer Science and Engineering (IJCSE), Vol. 2 No. 5 Oct-Nov 2011

[8]   Hiba Asria, Hajar Mousannif, Hassan Al Moatassimec,Thomas Noel "Using Machine   Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis", ELSEVIER, FAMS 2016

[9]   Ravi Aavula ,Dr. R. Bhramaramba, "A Survey on Latest Academic Thinking of Breast Cancer Prognosis", International Journal of Applied Engineering Research ISSN 0973-4562   Volume 13, Number 7 (2018) pp. 5207-5215

[10] Elias Zafiropoulos, Ilias Maglogiannis, loannis Anagnostopoulos," A Support Vector Machine Approach to Breast Cancer Diagnosis and Prognosis", Artificial Intelligence Applications and Innova,2006, in IFIP International Federation for Information Processing, Volume 204

[11] Ilias Maglogiannis · Elias Zafiropoulos, Ioannis Anagnostopoulos" An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers",Springer, Published online: 12 July 2007

[12] Leyi Wei, Pengwei Xing, Gaotao Shi, Zhiliang Ji, and Quan Zou" Fast prediction of protein methylation sites using a sequence-based feature selection technique", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2016

[13] Enli Zhang and Xiaoke Ma" Regularized Multi-View Subspace Clustering for Common   Modules Across Cancer Stages", Molecules 2018

[14] SungHwan Kim" Weighted K‑means support vector machine for cancer prediction" Springer   2016

[15] Ujjwal Maulik, Debasis Chakraborty," Fuzzy Preference Based Feature Selection and Semisupervised SVM for Cancer Classification" IEEE Transactions on Nanobioscience, Vol. 13, No. 2, June 2014

[16] Majid Mahrooghy, Ahmed B. Ashraf,Dania Daye, Elizabeth S. McDonald, Mark Rosen, Carolyn Mies, Michael Feldman, and Despina Kontos, " Pharmacokinetic Tumor Heterogeneity as a Prognostic Biomarker for Classifying Breast Cancer Recurrence Risk", IEEE Transactions on Biomedical Engineering, 2015

[17] Lin Zhang, Hui Liu, Yufei Huang, Xuesong Wang, Yidong Chen and Jia Meng, " Cancer Progression prediction Using Gene Interaction Regularized Elastic Net", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2015

[18] B.Padmapriya, T.Velmurugan, " A Survey on Breast Cancer Analysis Using Data Mining Techniques", ResearchGate 2014

[19] Guoyin Zhang, Liang Kou, Ye Yuan, Jianguo Sun, Yun Lin, Qingan Da,Wenshan Wang,"An Intelligent

International Conference on Contents, Computing & Communication (ICCCC-2022)

26th and 27th February,2022

Method of Cancer Prediction Based on Mobile Cloud Computing", Springer Nature 2017

[20] Deepika Verma, Nidhi Mishra" Analysis And Prediction Of Breast Cancer And Diabetes Disease Datasets Using Data Mining Classification Techniques", Ieee International Conference On Intelligent Sustainable Systems (Iciss), 2017

[21] Breast Cancer Wisconsin. https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin

[22] World Health Organization. http://www.who.int.

[23] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, ``An introduction to logistic regression analysis and reporting,'' J. Educ. Res., vol. 96, no. 1, pp. 3_14, 2002.

[24] D. A. Omondiagbe, S. Veeramani, and A. S. Sidhu, ``Machine learning classi_cation techniques for breast cancer diagnosis,'' IOP Conf. Ser., Mater. Sci. Eng., vol. 495, Jun. 2019, Art. no. 012033.

International Conference on Contents, Computing & Communication (ICCCC-2022)