

Analysis of Deep Learning Algorithms for Breast Cancer Diagnosis on WBC

¹ Prachi Jain, ² Syeda Zoya Kulsum,

³ Rohan, Srivastava, ⁴ Aryan Singh

Final year students,

Department of Information and Science Engineering,
JSS Academy of Technical Education(JSSATE),
Visvesvaraya Technological University (VTU)
Bengaluru ,Karnataka,India

Dr. Rekha PM

Department of Information and Science Engineering
JSS Academy of Technical Education(JSSATE),
Visvesvaraya Technological University (VTU)
Bengaluru,Karnataka, India

Abstract - In clinical conclusion, the expectation of an illness plays a significant role in breaking down the medical pictures. An undesirable sudden cell development in any piece of the organ is known as a tumor. Tumor might be benign or malignant. Harmful/malignant tumors are viewed to be very risky. Along these lines, the early detection of even the Architectural Distortion (AD) of the tissue cells helps forestalling cancer. In ladies, breast cancer is treated as the most critical issue. There are different specialist studies about the prediction of breast cancer malignant growth. This paper plots to survey and analyze different deep learning procedures that are explicitly considered on breast cancer prediction.

The approach is to actualize ML algorithms with the necessary parameters for training and testing for better execution. Right now, careful overview on AI algorithms (SVM, DT-C4.5, Naïve Bayes, k-NN and ANN) is finished considering the ongoing exploration papers on expectation of breast cancer growth tumors on the Wisconsin Database from the UCI Repository.

Keywords-ML algorithms, classification, accuracy

I. INTRODUCTION

Breast cancer malignancy has the second most elevated death rate in ladies followed by lung disease. According to clinical insights, 1 in each 8 ladies is determined to have breast cancer in the course of their life. Nonetheless, intermittent clinical exams and individual tests help in early recognition and in this manner altogether increment the odds of endurance. Intrusive recognition systems cause break of the tumor, quickening the spread of malignant growth to bordering territories.

Henceforth, there emerges the requirement for a progressively powerful, quick, precise, and effective noninvasive disease recognition framework. Early recognition can give patients greater treatment alternatives. So as to identify indications of malignant growth, breast tissue from biopsies is stained to segregate the nuclei and cytoplasm for minute assessment. At that point, pathologists assess the degree of any irregular basic variety to decide if there are tumors. Architectural Distortion (AD) is an inconspicuous constriction of the breast tissue and

may speak to the soonest indication of malignant growth. Since it is probably going to be unnoticed by radiologists, a few methodologies have been proposed throughout the years yet none utilizing profound learning procedures.

In 2011, National Breast Cancer Coalition (NBCC) gave an advancement report on Breast Cancer malignant growth named "Breast Cancer Deadline 2020". As per the report, Breast Cancer will be completely removed by 2020 and in 90 % cases the passing of cancer disease tolerant kicked the bucket because of spreading of Breast Cancer in different pieces of body. All the tumors can't be a similar sort and their division should be possible through science of the tumor. So the sorts can be analyzed distinctively for better outcomes. It implies the results of breast cancer growth are not the equivalent in all patients, taking into account this reality it is exceptionally important to segregate them and give separate treatment. So there is the need of grouping and characterization systems.

A doctor may use one or more approaches to diagnose cancer:

1. Physical Exam
2. Laboratory tests
3. Imaging tests
4. Biopsy

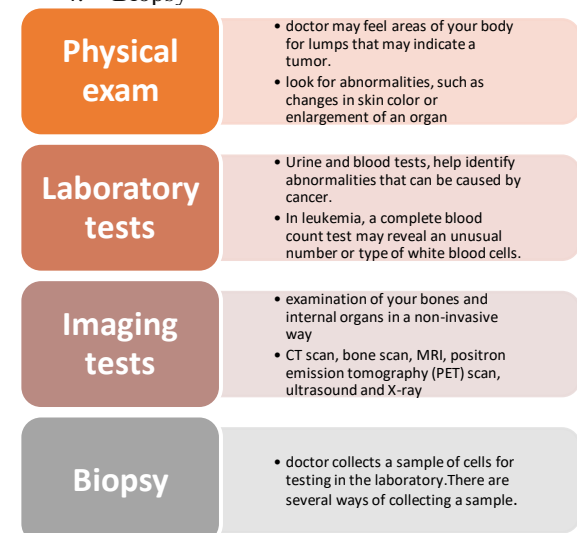


Fig 1. Approaches to diagnose the cancer

In the laboratory, pathologists observe various cell samples under the microscope. Normal cells seem uniform, with similar sizes and orderly arrangement. Cancer cells look haphazard, with varying sizes and without uniform organization.

II. LITERATURE SURVEY

Paper[1] compares different machine learning algorithms on the basis of their performance on WBC datasets. The aim was to estimate exactitude in segregating the data with reference to their efficacy on basis of accuracy, susceptibility of all algorithms. The investigation questions proposed on these experiments are: which method attains finer accuracy? Which algorithm is more well planned? Classifiers reported in the paper were carried out with the help of pre-defined collection of methods from WEKA environment. Machine Learning methods executed and tried on actual challenges. It provides reliable framework for developers for assessing their work. The 10-fold cross validation test has been tested on each foretelling models. It divides the model into train set and test set. The investigatory outcomes of these experiments conclude that Support vector machine produces the best accuracy of 97.13%. Hence SVM proves to be best suitable for breast cancer risk prediction.

In the era of big data, data mining methods endure lot of summons. The outcome of this research provides perspective of implementing these techniques on health care methodology. Here four data mining techniques and eight hybrid models are checked on two datasets. PCA the dimension depletion method gives some benefits with regard to prognosis effectiveness and accuracy. Paper [2] this paper looks into on different information digging strategies for bosom malignant growth expectation to pick an exact method to foresee the disease. The intension of the paper is to separate and point out the exact model of bosom malignancy event on understanding records. SVM, ANN, NB classifier, Adaptive Boost tree are the four diverse KDD strategies are tried right now. WBC database (1991) and WBC (1995) datasets are utilized to check and look at the exactness of these systems.

Profound learning dummies can set up the highlights on which the outcome will depend on. Profound learning techniques are applied so as to vanquish the restrictions of AI calculations. To make a profound learning models that mirror the working of cerebrum. This is cultivated by neural systems. In paper [3] CNN a profound learning model is advanced for investigation of pictures are characterize them into harmful and considerate. The assessment utilizing quality datasets for the CNN model and VGGNet engineering are utilized to evaluate execution. The proposed work right now these two engineering on bosom disease pictures and predicts the sort of cell present. Highlights are drawn out from convolutional and pooling layers. At that point these properties are embedded into completely associated layer for classification. VGGNet was imagined before 2014. ImageNet includes high goals pictures and the point is to show the model accurately with the goal that it can separate information pictures into 1000 item classifications. As the

aftereffect of execution of these two design it is been noticed that CNN creates preferred precision of 86.32% over VGGNet. In the resulting work the presentation of VGGNet will be improved.

In paper [4] right now they have set up computational utilization of profound neural systems for order of bosom malignancy pictures. H&E stained images are used. This work applies different designs of profound learning. Histopathology focuses to separate among tumor cells and to execute prognostic estimation. Structures like ResNet, initiation comprises of enormous number of parameters to accomplish results in different PC errands. They have utilized LightGBM, the quick, dispersed, elite usage of angle helped trees, for regulated classification. For highlight extraction they have utilized ResNet-50 and initiation V3 systems. For preparing information, the information has been part into 10 stratified folds to keep up class dispersal. The anticipated class is chosen by most noteworthy likelihood score. For 4-class classification task, it produces 87.2% precision. For 2-class classification undertaking to identify carcinomas it is been accounted for 93.8% exactness, AUC 97.3%, and affectability/specificity 96.5/88.0% at the high-affectability working point. To our understanding this methodology surpasses other regular strategies in mechanized histopathological picture order.

The research in paper [5] shows structured novel profound learning system for discovery of tumor cells in breast cytology pictures using transfer learning. Right now properties from pictures drawn out from recently prepared models of CNN i.e. GoogleNet, VGGNet, ResNet. These pictures are placed into completely associated layers that further predict whether the cell is favorable or dangerous. In the pre-handling stage highlights are removed from various designs and took care of into completely associated layers. Profound learning structures utilized here are prepared before for highlight extraction. The exchange learning procedure is then applied for characterization. For this arrangement two kinds of datasets are utilized: the standard benchmark dataset and privately advanced at LRH emergency clinic Peshawar, Pakistan. As the consequence of this work, it has been noticed that the precision correlation between singular designs and consolidated highlights of these structures produces distinctive exactness rates according to the dataset size given. As noted individual structures i.e. GoogleNet, VGGNet, ResNet gives 93.5%,94.15%,94.35% and joined system produces 97.525% of exactness. As indicated by the results, the proposed work gives more significant level of precision as contrast with single CNN design.

A proposition for the automatic diagnoses of breast cancer using histopathological images and concept of transfer and deep learning is made in paper [6]. Right now have proposed the programmed analyze of bosom disease utilizing histopathological pictures and idea of move and profound learning. The examination between convolutional neural system VGGNet design and shallower custom engineering. The dataset utilized for this work is benchmark BreakHis dataset. The designs made utilizing VGGNet parts and comprise convolutional layers with parameters. 16-layers sort of VGGNet is utilized, from

which uniquely crafted contains six convolutional layers. Ensembling is the method for creating various models and collaborating them to get the normal result. There are four diverse sort of picture goals in particular 40X,100X,200X,400X. The end troupe model created comprises of three custom structures of CNN. These designs are prepared with the ideas of move learning and increases more significant level of exactness. It has been noticed that shallower designs performs better than off-the-rack engineering and yields better execution. At certain goals, the presentation surpasses the best in class results, on the picked benchmark, by 10%.

The chief target paper [7] is to separate between the sores as generous or threatening by utilizing MMLP based classifier. This strategy helps in reenacting the natural property of the metaplasticity on MLP with the Back propagation(BPG). This MMLP calculation has been contrasted and a Classical BPG assumed a significant job to characterize WBCD database and with the assistance of as of late proposed calculations by different specialists that works on a similar database. The Multilayer Perception Neural Network (MLP) has been utilized for the arrangement of numerous grouping issues in design acknowledgment applications. The MMLP classifiers convey an extraordinary presentation obtained the accompanying outcomes on a normal for 100 systems. Our MMLP exhibit to be equivalent or here and there better than the beforehand cutting edge calculations applied to the WBCD database.

In paper [8] at first, eight diverse AI calculations are applied to the information first without applying any component determination technique and afterward by utilizing two of those. The results of the courses of action are differentiated and each other and with the delayed consequences of the primary case. The systems applied are SVM, KNN, MLP, Decision Trees, Random Forest, Logistic Regression, Adaboost and Gradient Boosting Machines. RFE (Recursive Feature Elimination) and RLR(Randomized Logistic Regression)feature removal systems are applied. These two methods start with the plan everything being equivalent and crash the most silly characteristics. In both datasets, SVM procedure is generally significant after segment decision systems, decision trees were organized with the least precision. MLP system gave close results with other AI methodologies on the primary data. In the ensuing data, it gave basically lower results appeared differently in relation to various methodologies. The precision in first educational record was high before the segment removal procedures so there was no colossal addition, yet there was a very enormous augmentation in the second dataset.

A point by point audit of the strengths, limitations, and performance of the latest CNNs applications in breaking down MG pictures is presented in paper [9]. It condenses in excess of 80 research reads for applying CNNs on different errands in mammography. This overview records the accepted procedures that overhaul the exhibition of CNNs including the pre-preparing of pictures and the utilization of multi-see pictures. Moreover, rest of the recorded procedures like transfer learning, data augmentation (DA),

batch normalization (BN), and dropout are attractive solutions to reduce over fitting and expand the speculation of the CNN models. Finally, this review distinguishes the difficulties in the examination and headings that require further examinations by the network.

To break down the example of stroma encompassing ductal carcinoma in situ sores and entire slide pictures containing ductal carcinoma in situ with simultaneous intrusive malignancy were commented on by a pathologist (MES) in paper [10]. For each case, a subset of conduits containing ductal carcinoma in situ sores was named on WSIs with point markings in the focal point of the sore and assessed utilizing top notch rules dependent on atomic size and appearance, mitoses and discovery of rot this entire slide picture characterization framework depends on various profound CNNs. To engage assessment of reasonable execution of our figuring, the dataset was aimlessly part into a readiness set including 62% of the entire whole slide pictures and a testing set. With the remainder of the slides three neural system layers were made: Network 1 is , a convolutional neural system model meant was prepared using the way to deal with arrange fat, stroma, and epithelium. System 2 was prepared working on stromal districts perceived by Network 1. System 2 created a likelihood that a picture spoke to disease related stroma. System 3 was intended to give a score to the whole entire slide picture showing the likelihood that the slide contained intrusive disease subdividing entire slide pictures into areas comprising of epithelium, stroma, and fat accomplished a pixellevel 3-class arrangement precision of 95.5% contrasted with reference standard.

In paper [11] the general structure incorporates 4 stages: introductory one is the making sure about of picture, second separating highlights from the mammograms, picking logically compensating features, classifier to perceive right class of mammogram. Dim Level Co-event Matrix highlights are decide along 0° for all mammograms. In the proposed framework, 10 surface highlights characterized by Haralick et al are worked upon. Highlights little part choice is utilized to diminish include space that assists with diminishing the calculation time. This is practiced by emptying boisterous, dull and unimportant features i.e., it picks the convincing highlights to 10 highlights from the GLCM were evaluated along 0° . Highlights space is also decreased to six features by using the rank highlights strategy. Results show accuracy of 100% for endorsement and test data, and for the most part precision achieved by using the proposed procedure is 99.4% to get the hankering yield. The images used for Breast cancer diagnosis are found to be more useful when obtained with the CEDM technique rather than FFDM.

In paper [12] a shallow CNN is applied onto the Low Energy (LE) images first to derive "virtual" recombined and better images and then a deep CNN is applied to those images to extract the conspicuous features from them. The process is first tested on a small dataset of 49 images using both deep and shallow networks and then a larger dataset of 89 images from INbreast database is taken. A couple of low and high-vitality pictures is created after the organization of a difference medium specialist. The two

pictures are converged to upgrade contrast uptake territories and the recombined picture is then created. Paper [12][16] examines the upsides of recombined pictures from CEDM in helping the finding of breast lesions using a Deep-CNN method. The result obtained is compared with previous standards and it is found that there is improved accuracy observed with using shallow-deep CNN with FFD obtained images. Addition of recombined imaging features increases model performance accuracy of 0.89 with AUC of 0.91.

Mammograms are the most commonly used screening techniques for cancer diagnosis. A computer aided detection system using MGA (Modified Genetic Algorithm) tuned ANN is used for detection of tumor cases in mammograms here. 322 mammograms are used for evaluating the performance of the algorithm from the MAIS database in paper [13]. Processed images are obtained by extracting the ROI. MGI is applied in the end of the process which follows principles from natural evolution, containing 3 steps, namely: selection, crossover and mutation. ANN having 3 layers with the neurons equal to the number of features extracted is then tuned into MGA for classification of the expected tumors into either of the two classes- benign or malignant. The framework used in paper [12] is designed in MATLAB 7 and a CAD is developed. Using the above technique an accuracy of 97.8% is obtained on the taken database.

III. METHODOLOGY

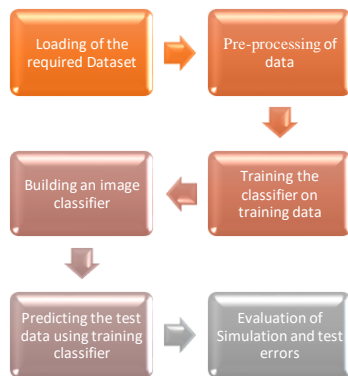


Fig 1. steps to implement algorithm

A. Loading of Dataset:

Loading the data before starting your machine learning project is the first step in the process and the most common format for machine learning data is CSV files. There are a number of ways to load a CSV file in Python. Loaded using python standard lib, pandas, numpy or directly from URL. Keras, Pillow etc are libraries used for loading datasets of images. The dataset used here is the Wisconsin Database from the UCI Repository.

B. Pre-Processing of Data:

Before raw data could be sent through a machine learning model it has to undergo preprocessing. And it's simply

because data in the real world are generally Incomplete, Noisy and Inconsistent.

- handling missing values: missing values can be handled by either deleting the entire row having a missing value or by imputing the missing values.
- fitting data to the dataset parameters is another step
- splitting data into train and test set constitutes the next step
- feature scaling is then done to standardize data

C. Building an image classifier:

Image categorization points to the labeling of images into one of a number of previously defined classes. Here, the survey focuses on comparing the binary classification (benign or malignant) of different techniques in ML applied on the taken dataset. This step involves importing the libraries from different frameworks (TensorFlow, Keras etc) and using them to fit and transform data for our model building phase. We can import data from Google drive and use Google Colab as a platform for building our model.

D. Training the Classifier Model:

This step will train the model on the training set images and validate it using, the validation set. Preparing a model requires defining a loss function, optimizer and metrics. It is difficult to ascertain the ideal loads for a neural network, there are numerous questions and rather, the issue of learning is given a role as a hunt or streamlining issue and a calculation is utilized to explore the space of potential arrangements of loads the model may use to make great forecasts.

Seeking to minimize the error caused by an optimization function and choose the one that best fits the model gives rise to what is called the loss function, which in brief minimizes error of the optimization function. The estimators used to estimate the error could be the Maximum Likelihood Estimator and maximize the optimization function.

E. Predicting the test data using the classifier:

After training the classifier model by running it for a certain number of epochs, we can save the model to use it for predicting risk in new images that aren't labeled. The testing data is divided into two parts, one containing all the features of the testing images and the other having all the target labels. When images are passed to the classifier, it resorts them into either of their binary classes and then checks against its labels and outputs accuracy for the rightly classifies test data. This helps in improving the model.

F. Evaluation of Simulation and test errors:

Strategies for assessing a model's presentation are isolated into 2 classes: in particular, holdout and Cross-approval. The two strategies utilize a test information not seen by the

model to assess model execution. It's not recommended to use the data we used to build the model to evaluate it as it may lead to overfitting. AUC, Logarithmic Loss, Confusion Matrix are other evaluation metrics used to validate a model and its performance.

IV. MACHINE LEARNING ALGORITHM

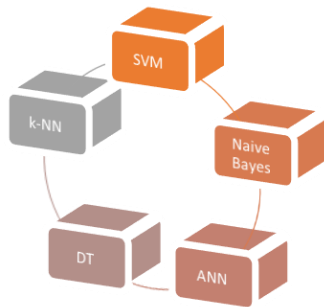


Fig. 2 Different ML algorithms

A. Support Vector Machine

Support Vector Machine is called as SVM. It is Supervised Machine learning calculation utilized for characterization and relapse. In SVM calculation the informational collection is plotted as n-dimensional (where n is no. of examples in dataset). It separates various highlights in the hyper plane and thinks about them quite well. It is utilized in written by hand computerized acknowledgment, picture acknowledgment, face discovery, Bioinformatics and some more.

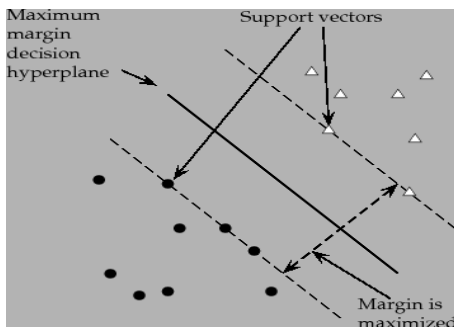


Fig 3 SVM Classification

B. Naïve-Bayes Classifier:

Naïve Bayes is a statically and feasibility classifiers, which is established on Bayes Theorem. Every feature of the attribute is independent to each other attribute. It is a classification technique which was designed to classify the high-dimensional datasets. The probability of events is calculated as:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability
Posterior Probability
Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

Fig 4. Probability calculation

Where y and X are different features. y is known as class variable and X is the evidence i.e. probability of event prior evidence.

C. k-NN:

k-NN is an apathetic model since it doesn't pick up anything during the preparation stage and learns in testing stage. It is occurrence based learning. It is a non-parametric realizing which retain the resultant of characterize off inconspicuous information. It is utilized for grouping and relapse calculations. The yield of the grouping are in type of 1, -1 and 0. This calculation is utilized for design acknowledgment and interruption identification. It requires some investment to register the outcome so it is less productive among the others.

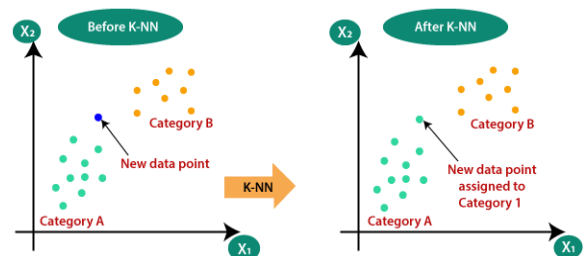


Fig 5. k-Nearest Neighbour

D. Decision Tree(C4.5):

Decision Table is a classification algorithm which build tree structure data format. The datasets are divided into sub nodes or sub leaf. Each sub nodes represent the instances of the datasets. The leaf node is called as class label. Some datasets include missing values are not calculated by the algorithms but this algorithm show accurate result still having missing values.

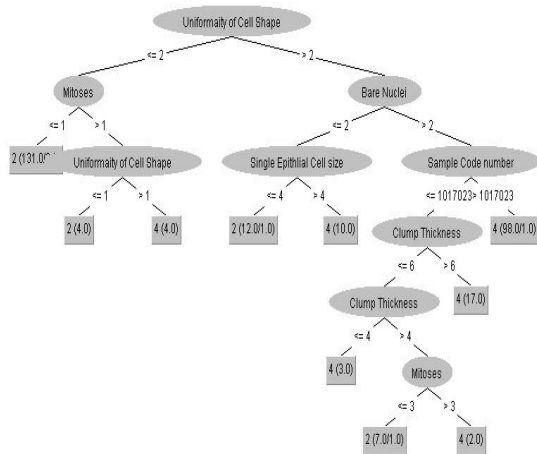


Fig 6. DT Example

E. ANN:

Artificial Neural Networks is supervised learning using Multi-Layer Perceptron (MLP) with back propagation by implementing a bunch of weights. The weights act as bridge to the input through the outcome units. This strategy is by all accounts mind boggling and reasonable structure. Be that as it may, the expectation precision is exceptionally high.

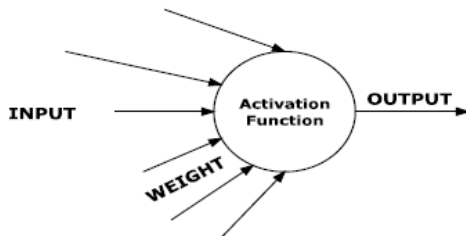


Fig 7. Artificial Neuron

V. DISCUSSION

The exhaustive study done by mulling over all the exploration papers referenced above demonstrates that the best strategy for bosom disease finding on WBC is Support Vector Machine. In outline, SVM had the option to show its capacity as far as adequacy and productivity dependent on precision and review. Contrasted with a decent measure of research on Breast-malignant growth Wisconsin found in writing that think about arrangement correctnesses of information mining calculations, our test results make the most noteworthy estimation of exactness (97.28 %) in characterizing bosom disease dataset. It tends to be seen that SVM outflanks different classifiers as for exactness, affectability, particularity and accuracy; in ordering bosom malignant growth dataset.

TABLE I

| Number | Model | Accuracy |
|--------|-------|----------|
| 1 | C 4.5 | 95.13 |
| 2 | SVM | 97.13 |
| 3 | NB | 95.99 |
| 4 | k-NN | 95.27 |
| 5 | ANN | 89.88 |

VI. CONCLUSION

This survey paper records the research of related papers that have been published and discusses the methodologies used in breast cancer detection previously on Wisconsin Breast Cancer Database. It is evident that age, sex, alcohol, urban/rural region and weight are the major factors that influence the occurrence of Breast Cancer tumors. The chances of breast cancer cases can be high when it is a hereditary disease. There is a need of large amounts of data for analyzing therefore data mining may be used. The symptoms of breast cancer aren't similar in any two patients, in view of this fact it is very necessary to categorize them and give separate treatment. For grouping of alike symptoms clustering can be used. The best results in different conditions may be discovered by optimizing the results.

So if the methodologies are combined in a single framework the chances of finding better solution will be increased. So a hybrid framework consists of classification, clustering, association and optimization will prove to be better in the above situation. It is concluded that out of the 5 techniques used on the Wisconsin Database, SVM proves to be a better classifier with higher accuracy and less loss. Further, the scope of this survey paper can be extended to cover other algorithms with added techniques like PCA or segmentation.

VII. REFERENCES

- [1] Karabatak, M., Cevde t-Ince, M.: An expert system for detection of breast cancer based on association rules and neural network. Expert Systems with Applications 36, 3465–3469 (2009)
- [2] Hiba Asria*,Hajar Mousannifb,Hassan Al Moatassimec,Thomas Noeld, Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. . Published by Elsevier B.V. doi: 10.1016/j.procs.2016.04.224, 2016
- [3] Asri H, Mousannif H, Al Moatassime H, Noel T. Big data in healthcare: Challenges and opportunities. 2015 Int Conf CloudTechnol Appl. 2015:1-7. doi:10.1109/CloudTech.2015.7337020.
- [4] V. Chaurasia and S. Pal, (2014) "Data Mining Techniques: To Predict and Resolve Breast Cancer Survivability," 10–22, 2014.
- [5] B.Krishnakumar, K.Kousalya, R.S.Mohana, K.Dinesh, S.Santhiya, Classification of Breast Cancer using Deep Learning Architecture, International Journal of Recent Technology and Engineering (IJRTE),Volume-8,DOI:10.35940/ijrte.D5317.118419,2019
- [6] J. Wang, L. Perez, The effectiveness of data augmentation in image,Classification using deep learning, in: IEEE Conference on Computer and its Vision and Pattern Recognition, 2017, pp. 1–8.
- [7] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations, 2015, pp. 1–14.

-
- [8] O. Penatti, K. Nogueira, J. Santos, Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 44–51.
- [9] A.S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: an astounding baseline for recognition, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2014, pp. 512–519.
- [10] E. Hoffer, N. Ailon, Deep metric learning using triplet network, in: International Workshop on Similarity-Based Pattern Recognition, (2015) 84–92.
- [11] Alireza Osareh, Bitashadgar, "A Computer Aided Diagnosis System for Breast Cancer", International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011
- [12] Orozco-Monteagudo, M., Taboada-Crispí, A., Del Toro-Almenares, A web based report A.: Training of multilayer perceptron neural networks by using cellular Genetic algorithms. In: Martínez-Trinidad, 2006. LNCS, vol. 4225, pp. 389–398. Springer, Heidelberg (2006)
- [13] Guijarro-Berdiñas, B., Fontenla-Romero, O., Perez-Sanchez, B., Fraguera, P.: A linear learning method for multilayer perceptrons using least-squares. LNCS, vol. 4225, pp. 365–374. Springer, Heidelberg (2007)
- [14] Andina, D., Jevtić, A., Marcano, A., Barrón-Adame, M.: Error weighting in artificial neural networks learning interpreted as a metaplasticity model. In: Mira, J., Álvarez, J.R. (eds.) IWINAC 2007. LNCS, vol. 4527, pp. 244–252. Springer Heidelberg (2007)
- [15] Weiming Zhi¹(B), Henry Wing Fung Yueng², Zhenghao Chen², Seid Miad Zandavi², Zhicheng Lu², and Yuk Ying Chung², Using Transfer Learning with Convolutional Neural Networks to Diagnose Breast Cancer from Histopathological Images, Springer International Publishing AG, pp. 669–676, 2017
- [16] Misra, B.B., Biswal, B.N., Dash, P.K., Panda, G.: Simplified polynomial neural network for classification task in data mining. In: Evolutionary Computation, CEC 2007, pp. 721–728 (2007)