



Applications of Machine Learning in Cancer Detection and Classification

¹Manasvi Kapoor, ²Vibhor Sharma

¹Student, ²Assistant Professor

^{1,2}Information Technology

^{1,2}Maharaja Agrasen Institute of Technology, New Delhi, India

Abstract—Breast cancer (BC) is one of the most common cancers among women worldwide, representing the majority of new cancer cases and cancer-related deaths according to global statistics, making it a significant public health problem in today's society. The early diagnosis of BC can improve the prognosis and chance of survival significantly, as it can promote timely clinical treatment to patients. Further accurate classification of benign tumours can prevent patients undergoing unnecessary treatments. Thus, the correct diagnosis of BC and classification of patients into malignant or benign groups is the subject of much research. Because of its unique advantages in critical features detection from complex BC datasets, machine learning (ML) is widely recognized as the methodology of choice in BC pattern classification and forecast modelling. Classification and data mining methods are an effective way to classify data. Especially in medical field, where those methods are widely used in diagnosis and analysis to make decisions.

Index Terms—Breast Cancer; Breast Cancer Detection; Medical Inputs; Machine Learning.

I. INTRODUCTION

Breast cancer is the second leading cause of death among women worldwide. In 2019, 268,600 new cases of invasive breast cancer were expected to be diagnosed in women in the U.S., along with 62,930 new cases of non-invasive breast cancer. Early detection is the best way to increase the chance of treatment and survivability. Data mining has become a popular tool for knowledge discovery which shows good results in marketing, social science, finance and medicine. Recently, multiple classifiers algorithms are applied on medical datasets to perform predictive analysis about patients and their medical diagnosis. For example, using machine learning techniques to assess tumour behaviour for breast cancer patients. One problem is that there is a class imbalance in the training data, since the probability of not having this disease is higher than the one of having it.

This paper introduces a comparison between six different classifiers: NB, SVM, KNN, LR, RF and DT with respect to accuracy in detection of breast cancer.

Breast cancer is the most found disease in the women, worldwide, where abnormal growth of a mass of tissue, cause the expansion of malignant cells leads to acute breast cancer. These malignant cells are originally created from milk glands of the breast. These malignant cells which are the main reason for breast cancer can be classified into different groups according to their unusual progress and capability affecting other normal cells. The capability of affecting means whether these malignant cells affect only the local cells or can spread throughout the full body. The effect of spreading these malignant cells throughout the whole body of the patient is called as metastasis.

It is very important to prevent this spreading effect by a diagnosis of cancer in the early stages using advanced techniques and equipment. In recent decades, there are many efforts to employ artificial intelligence and other related methods to assist in the detection of cancer in earlier stages.

Early detection of cancer boosts the increase of survival chance to 98%.

II. MACHINE LEARNING METHODS

Machine Learning is a process that machines (computers) are trained with data to make the decision for similar cases. ML is employed in various applications, such as object recognition, network, security, and healthcare. There are two ML types i.e. single and hybrid methods like ANN, SVM, Gaussian Mixture Model (GMM), K-Nearest Neighbor (KNN), Linear Regressive Classification (LRC), Weighted Hierarchical Adaptive Voting Ensemble (WHAVE), etc.

Before beginning with a detailed analysis of what machine learning methods work best for which kinds of situations, it is important to have a good understanding of what machine learning is – and what it isn't. Machine learning is a branch of artificial intelligence research that employs a variety of statistical, probabilistic and optimization tools to “learn” from past examples and to then use that prior training to classify new data, identify new patterns or predict novel trends (Mitchell 1997). Machine learning, like statistics, is used to analyze and interpret data. Unlike statistics, though, machine learning methods can employ Boolean logic (AND, OR, NOT), absolute conditionality (IF, THEN, ELSE), conditional probabilities (the probability of X given Y) and unconventional optimization strategies to model data or classify patterns. These latter methods actually resemble the approaches humans typically use to learn and classify. Machine learning still draws heavily from statistics and probability, but it is fundamentally more powerful because it allows inferences or decisions to be made that could not otherwise be made using conventional statistical methodologies (Mitchell 1997; Duda et al. 2001). For instance, many statistical methods are based on multivariate regression or correlation analysis. While generally very powerful, these approaches assume that the variables are independent and that data can be modeled using linear combinations of these variables. When the relationships are nonlinear and the variables are interdependent (or conditionally dependent) conventional statistics usually flounders. It is in these situations where machine learning tends to shine. Many biological systems are fundamentally nonlinear and their parameters conditionally dependent. Many simple physical systems are linear and their parameters are essentially independent.

Success in machine learning is not always guaranteed. As with any method, a good understanding of the problem and an appreciation of the limitations of the data is important. So too is an understanding of the assumptions and limitations of the algorithms being applied. If a machine learning experiment is properly designed, the learners correctly implemented and the results robustly validated, then one usually has a good chance at success. Obviously if the data is of poor quality, the result will be of poor quality (garbage in = garbage out). Likewise if there are more variables than events to predict then it is also possible to create a series of redundant learners. This is a set of learning algorithms that seems to perform at the same (low) level regardless of the choice of input data. The problem of too many variables and too few examples is called the “curse of dimensionality” (Bellman 1961). This curse is not restricted to machine learning. It also affects many statistical methods as well. The only solution is to reduce the number of variables (features) or increase the number of training examples. As a general rule, the sample-per-feature ratio should always exceed 5:1 (Somorjai et al. 2003). Not only is the size of the training set important, so too is the variety of the training set. Training examples should be selected to span a representative portion of the data the learner expects to encounter. Training too many times on too few examples with too little variety leads to the phenomenon of over-training or simply training on noise (Rodvold et al. 2001). An over-trained learner, just like an overtired student, will generally perform poorly when it tries to process or classify novel data.

Sometimes conventional statistics proves to be more powerful or more accurate than machine learning. In these cases the user's initial determinations about the interdependence and nonlinearity of the data would have been wrong. This is not necessarily a weakness to machine learning, it is just a matter of choosing the right tool for the right job. Likewise, not all machine learning methods are created equal. Some are better for certain kinds of problems while others are better for other kinds of problems. For instance some machine learning algorithms scale nicely to the size of the biological domains, others do not. Likewise some methods may have assumptions or data requirements that render them inapplicable to the problem at hand. Knowing which method is best for a given problem is not inherently obvious.

This is why it is critically important to try more than one machine learning method on any given training set. Another common misunderstanding about machine learning is that the patterns a machine learning tool finds or the trends it detects are non-obvious or not intrinsically detectable. On the contrary, many patterns or trends could be detected by a human expert – if they looked hard enough at the data. Machine learning simply saves on the time and effort needed to discover the pattern or to develop the classification scheme. Recall that with any interesting discovery, it is frequently obvious to the casual observer – particularly after the discovery has been made.

There are three general types of machine learning algorithms: 1) supervised learning; 2) unsupervised learning and 3) reinforcement learning. They are essentially classified on the basis of desired outcome of the algorithm (Mitchell, 1997; Duda et al. 2001). In supervised learning algorithms a “prescient provider” or teacher gives the learning algorithm a labeled set of training data or examples. These labeled examples are the training set that the program tries to learn about or to learn how to map the input data to the desired output. For instance a labeled training set might be a

set of corrupted images of the number “8” (Figure 1). Since all the images are labeled as being the number “8” and the desired output is the uncorrupted “8”, the learner is able to train under the supervision of a teacher telling it what it is supposed to find. This is the process by which most school children learn. In unsupervised learning, a set of examples are given, but no labels are provided. Instead it is up to the learner to find the pattern or discover the groups. This is somewhat analogous to the process by which most graduate students learn. Unsupervised learning algorithms include such methods as self-organizing feature maps (SOMs), hierarchical clustering and K-means clustering algorithms. These approaches create clusters from raw, unlabeled or unclassified data. These clusters can be used later to develop classification schemes or classifiers.

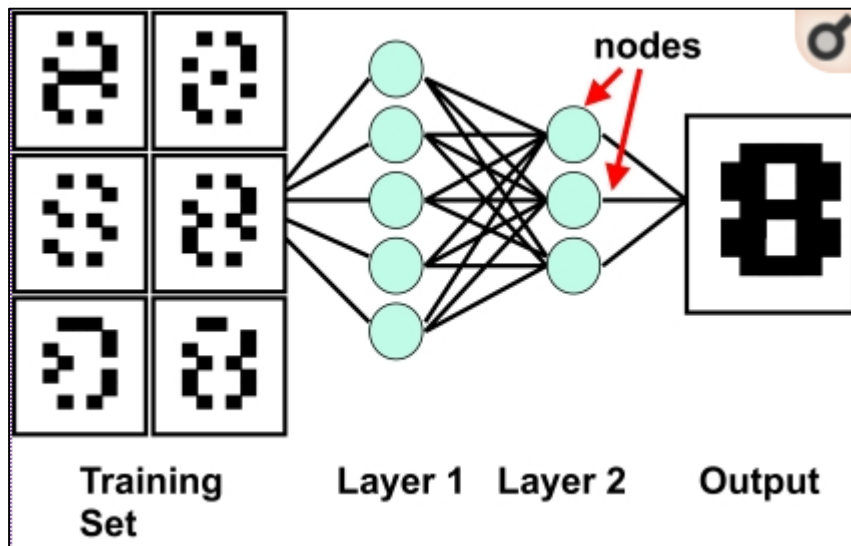


Figure 1.

An example of how a machine learner is trained to recognize images using training set (a corrupted image of the number “8”) which is labelled or identified as the number “8”.

Following are the used ML algorithms:

2.1. NAÏVE BAYES (NB) CLASSIFIER

Bayes refers to a probabilistic classifier that applies Bayes’ theorem with robust independence assumptions. In this model, all properties are considered separately to detect any existing relationship between them. It assumes that predictive attributes are conditionally independent given a class. Moreover, the values of the numeric attributes are distributed within each class. NB is fast and performs well even with a small dataset. However, it is difficult to find independent properties in real life have deployed NB classifier for breast cancer detection and it gave the maximum accuracy with only five dominant.

2.2. DECISION TREES (DT)

DT is a data mining technique used for early detection of breast cancer. It is a model that presents classifications or regressions as a tree. In this model, the data set is broken to small sub-data, then to smaller ones. As a result, the tree is developed and at the last level, the result is revealed. In a tree structure, the leaves characterize the class labels whereby the branches characterize conjunctions of feature leading to the class labels. Hence, DT is not sensitive to noise.

2.3. K-NEAREST NEIGHBOURS (KNN)

KNN is a supervised learning method which is used for diagnosing and classifying cancer. In this method, the computer is trained in a specific field and new data is given to it. Additionally, similar data is used by the machine for detecting (K) hence, the machine starts finding KNN for the unknown data. It is recommended to choose a large dataset for training also K value must be an odd number.

2.4. SUPPORT VECTOR MACHINES (SVM)

SVM is a supervised pattern classification model which is used as a training algorithm for learning classification and regression rule from gathered data. The purpose of this method is to separate data until a hyperplane with high minimum distance is found. SVM is used to classify two or more data types. SVM include single or hybrid models such as Standard SVM (St-SVM), Proximal Support Vector Machine (PSVM), Newton Support Vector Machine (NSVM), Lagrangian Support Vector Machines (LSVM), Linear Programming Support Vector Machines (LPSVM) and Smooth Support Vector Machine (SSVM).

2.5. RANDOM FOREST (RF) CLASSIFIER

Random forests or random decision forests are an ensemble learning for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance. Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration in packages such as scikit-learn.

2.6. LOGISTIC REGRESSION (LR) CLASSIFIER

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. Mathematically, a binary logistic model has a dependent variable with two possible values, such as pass/fail which is represented by an indicator variable, where the two values are labelled "0" and "1". In the logistic model, the log-odds (the logarithm of the odds) for the value labelled "1" is a linear combination of one or more independent variables ("predictors"); the independent variables can each be a binary variable (two classes, coded by an indicator variable) or a continuous variable (any real value). The corresponding probability of the value labelled "1" can vary between 0 (certainly the value "0") and 1 (certainly the value "1"), hence the labelling; the function that converts log-odds to probability is the logistic function, hence the name. The unit of measurement for the log-odds scale is called a logit, from logistic unit, hence the alternative names.

III. PREVIOUS RELATED WORKS

3.1. Case Study 1 – Cancer Risk or Susceptibility Prediction

Of the 79 papers surveyed in this review, relatively few papers (just 3) employed machine learning to predict cancer risk susceptibility. One of the more interesting papers (Listgarten et al. 2004), used single nucleotide polymorphism (SNP) profiles of steroid metabolizing enzymes (CYP450s) to develop a method to retrospectively predict the occurrence of "spontaneous" breast cancer. Spontaneous or non-familial breast cancer accounts for about 90% of all breast cancers (Dumitrescu and Cotarla 2005). The hypothesis in this study was that certain combinations of steroid-metabolism gene SNPs would lead to the increased accumulation of environmental toxins or hormones in breast tissue leading to a higher risk for breast cancer. The authors collected SNP data (98 SNPs from 45 different cancer-associated genes) for 63 patients with breast cancer and 74 patients without breast cancer (control). Key to the success of this study was the fact that the authors employed several methods to reduce the sample-per-feature ratio and investigated multiple machine learning methods to find an optimal classifier. Specifically, from a starting set of 98 SNPs the authors quickly reduced this set to just 2–3 SNPs that seemed maximally informative. This reduced the sample-per-feature ratio to a respectable 45:1 (for 3 SNPs) and 68:1 (for 2 SNPs) instead of close to 3:2 (had all 98 SNPs been used). This allowed the study to avoid falling victim to the "curse of dimensionality" (Bellman 1961; Somorjai et al. 2003). Once the sample size was reduced, several machine learning techniques were employed including a naïve Bayes model, several decision tree models and a sophisticated support vector machine (SVM). The SVM and naïve Bayes classifiers attained the highest accuracy using only a set of 3 SNPs and the decision tree classifier attained the highest accuracy using a set of 2 SNPs. The SVM classifier performed the best with an accuracy of 69%, while the naïve Bayes and decision tree classifiers achieved accuracies of 67% and 68%, respectively. These results are approximately 23–25% better than chance. Another notable feature to this study was the extensive level of cross validation and confirmation performed. The predictive power of each model was validated in at least three ways. Firstly, the training of the models were assessed and monitored with 20-fold cross-validation. A bootstrap resampling method was employed by performing the cross-validation 5 times and averaging the results so as to minimize the stochastic element involved with partitioning of the samples. Secondly, to minimize the bias in feature selection (i.e. selecting the most informative subset of SNPs), the selection process was performed within each fold for a total of 100 times (5 times for each of the 20 folds). Finally, the results were compared against a random permutation test which at best, had a predictive accuracy of 50%. While the authors attempted to minimize the stochastic element involved with partitioning of the samples, a better method may have been to use leave-one-out cross-validation which would have removed this stochastic element completely. That being said, the multiple cross-validations resulted in a standard deviation that was not more than 4% for any of the reported accuracies and since all the methods performed close to 25% better than chance, this standard deviation is deemed negligible. While no external validation set was reported in this study, we

have recently learned that the results described in this paper have been duplicated with a similar follow-on study of another 200 individuals (S. Damaraju, personal communication). Overall, this study nicely illustrates how the proper design, careful implementation, appropriate data selection and thorough validation of multiple machine learners can produce a robust and accurate cancer-risk prediction tool. It also highlights how machine learning can reveal important insights into the biology and polygenic risk factors associated with spontaneous or non-familial breast cancer.

3.2. Case Study 2: Prediction of Cancer Survivability

Nearly half of all machine learning studies on cancer prediction were focused on predicting patient survivability (either 1 year or 5 year survival rates). One paper of particular interest (Futschik et al. 2003) used a hybrid machine learning approach to predict outcomes for patients with diffuse large B-cell lymphoma (DLBCL). Specifically, both clinical and genomic (microarray) data were combined to create a single classifier to predict survival of DLBCL patients. This approach differs somewhat from the study of Listgarten et al. (2004) which only employed genomic (SNP) data in its classifier schema. Futschik et al. hypothesized, correctly, that clinical information could enrich microarray data such that a combined predictor would perform better than a classifier based on either microarray data alone or clinical data alone. In assembling the test and training samples, the authors collected microarray expression data and clinical information for 56 DLBCL patients. The clinical information was obtained from the International Prediction Index (IPI) which consists of a set of risk factors, that when properly assessed, allows patients to be separated into groups ranging from low-risk to high-risk. The data from the patient's IPI classifications was then used to create a simple Bayesian classifier. This classifier achieved an accuracy of 73.2% in predicting the mortality of DLBCL patients. Separately from the Bayesian classifier, several different types of "evolving fuzzy neural network" (EFuNN) classifiers were also developed to handle the genomic data. The best EFuNN classifier used a subset of 17 genes from the microarray data. This optimal EFuNN had an accuracy of 78.5%. The EFuNN classifier and the Bayesian classifier were then combined into a hierarchical modular system to generate a consensus prediction. This hybrid classifier attained an accuracy of 87.5%, a clear improvement over the performance of either classifier alone. This was also 10% better than the best performing machine learning classifier (77.6% by SVMs).

The EFuNN classifier was validated using a leave-one-out cross-validation strategy. This was likely due to the small sample size. As with Case Study #1, no external validation set was available to test the generality of the model. With only 56 patients (samples) being classified via 17 gene features, the sample per feature ratio (SFR) is just over 3. As a rule, an SFR of less than 5 does not necessarily guarantee a robust classifier (Somorjai et al. 2003). However, it is quite evident that the authors were aware of this issue and went to considerable lengths to justify their approach by explaining, in detail, the inner workings of their classifier. This included a description of how the Bayesian classifier was built, how the EFuNN works, and how the two classifiers work together to give a single prediction. In addition, the authors also investigated, and subsequently confirmed, the independence of the microarray data from the clinical data. This attention to detail is particularly exemplary for a machine learning investigation of this kind. This study nicely demonstrates how the power of using both clinical and genomic data in cancer prognosis can substantially enhance prediction accuracy.

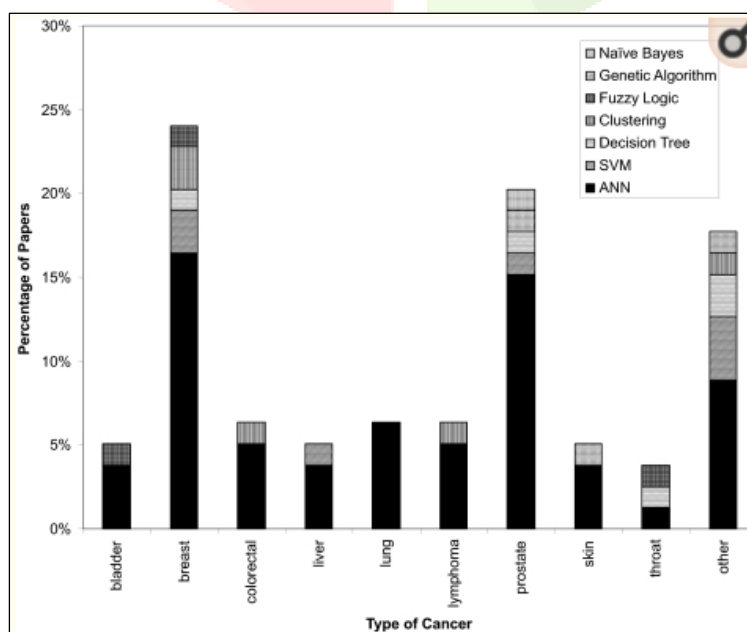


Figure 2.

A histogram showing the frequency with which different types of machine learning methods are used to predict different types of cancer. Breast and prostate cancer dominate, however a good range of cancers from different organs or tissues also appear to be compatible with machine learning prognoses. The "other" cancers include brain, cervical, esophageal, leukemia, head, neck, ocular, osteosarcoma, pleural mesothelioma, thoracic, thyroid, and trophoblastic (uterine) malignancies.

IV. CONCLUSION

In this review we have attempted to explain, compare and assess the performance of different machine learning that are being applied to cancer prediction and prognosis. Specifically we identified a number of trends with respect to the types of machine learning methods being used, the types of training data being integrated, the kinds of endpoint predictions being made, the types of cancers being studied and the overall performance of these methods in predicting cancer susceptibility or outcomes. While ANNs still predominate it is evident that a growing variety of alternate machine learning strategies are being used and that they are being applied to many types of cancers to predict at least three different kinds of outcomes. It is also clear that machine learning methods generally improve the performance or predictive accuracy of most prognoses, especially when compared to conventional statistical or expert-based systems. While most studies are generally well constructed and reasonably well validated, certainly greater attention to experimental design and implementation appears to be warranted, especially with respect to the quantity and quality of biological data. Improvements in experimental design along with improved biological validation would no doubt enhance the overall quality, generality and reproducibility of many machine-based classifiers. Overall, we believe that if the quality of studies continues to improve, it is likely that the use of machine learning classifier will become much more commonplace in many clinical and hospital settings.

V. REFERENCES

- [1] Aha D. Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms. *International Journal of Man-Machine Studies*. 1992;36:267–287. [[Google Scholar](#)]
- [2] Ando T, Suguro M, Hanai T, et al. Fuzzy neural network applied to gene expression profiling for predicting the prognosis of diffuse large B-cell lymphoma. *Jpn J Cancer Res*. 2002;93:1207–12. [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
- [3] Ando T, Suguro M, Kobayashi T, et al. Multiple fuzzy neural network system for outcome prediction and classification of 220 lymphoma patients on the basis of molecular profiling. *Cancer Sci*. 2003;94:906–13. [[PubMed](#)] [[Google Scholar](#)]
- [4] Atlas L, Cole R, Connor J, et al. Performance comparisons between backpropagation networks and classification trees on three real-world applications. *Advances in Neural Inf. Process. Systems*. 1990;2:622–629. [[Google Scholar](#)]
- [5] Bach PB, Kattan MW, Thornquist MD, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*. 2003;95:470–8. [[PubMed](#)] [[Google Scholar](#)]
- [6] Baldus SE, Engelmann K, Hanisch FG. MUC1 and the MUCs: a family of human mucins with impact in cancer biology. *Crit Rev Clin Lab Sci*. 2004;41:189–231. [[PubMed](#)] [[Google Scholar](#)]
- [7] Bellman R. *Adaptive Control Processes: A Guided Tour*. Princeton University Press; 1961. [[Google Scholar](#)]
- [8] Bocchi L, Coppini G, Nori J, Valli G. Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks. *Med Eng Phys*. 2004;26:303-12. [[PubMed](#)] [[Google Scholar](#)]
- [9] Bollschweiler EH, Monig SP, Hensler K, et al. Artificial neural network for prediction of lymph node metastases in gastric cancer: a phase II diagnostic study. *Ann Surg Oncol*. 2004;11:506–11. [[PubMed](#)] [[Google Scholar](#)]
- [10] Cancer country profiles 2014, WHO, <http://www.who.int/cancer/country-profiles/en/>
- [11] M. Stalin, and R. Kalaimagal, —Breast cancer diagnosis from low-intensity asymmetry thermogram breast images using fast support vector machine, Imanager's Journal on Image Processing, vol. 3, no. 3, pp. 17–26, 2016.
- [12] R. Kirubakaran, T. C. Jia, and N. M. Aris, —Awareness of Breast Cancer among Surgical Patients in a Tertiary Hospital in Malaysia, Asian Pacific Journal of Cancer Prevention, 2017, vol. 18, no. 1, pp. 115–120.
- [13] T. M. Khan, and S. A. Jacob, —Brief review of complementary and alternative medicine use among Malaysian women with breast cancer, Journal of Pharmacy Practice and Research, 2017, vol. 47, no. 2, pp. 147–152.

- [14] L. Caplan, —Delay in breast cancer: implications for the stage at diagnosis and survival,|| *Frontiers in Public Health*, 2014, vol. 2, Article 87, pp. 1–6.
- [15] M.A. Richards, A.M. Westcombe, S.B. Love, P. Littlejohns, and A.J. Ramirez, —Influence of delay on survival in patients with breast cancer: a systematic review,|| *The Lancet*, 1999, vol. 353, no. 9159, pp. Systems (RACS '14), Towson, Maryland, 2014, pp. 5 1119-1126.
- [16] B. Stewart and C.P. Wild, *World Cancer Report 2014*, International Agency for Research on Cancer, WHO, 2014.
- [17] S. A. Korkmaz, and M. Poyraz, —A New Method Based for Diagnosis of Breast Cancer Cells from Microscopic Images: DWEE—JHT,|| *J. Med. Syst.*, vol. 38, no. 9, p. 92, 2014.
- [18] P. Louridas, and C. Ebert, —Machine Learning,|| *IEEE Softw.*, vol. 33, no. 5, pp. 110–115, 2016.
- [19] A. Simons, —Using artificial intelligence to improve early breast cancer detection, —2017. Retrieved on April 10, 2018, from <https://www.csail.mit.edu/news/using-artificialintelligence-improve-early-breast-cancer-detection>
- [20] E. Ali, and W. Feng, —Breast Cancer classification using Support Vector Machine and Neural Network,|| *International Journal of Science and Research*, pp. 2013, 2319-7064.
- [21] S. Medjahed, T. Saadi, and A. Benyettou, —Breast Cancer Diagnosis by using k-Nearest Neighbor with Different Distances and Classification Rules,|| *International Journal of Computer Applications*, 2013, vol. 62, no. 1, pp. 0975 – 8887.
- [22] R. Sumbaly, N. Vishnusri, and S. Jeyalatha, —Diagnosis of Breast Cancer using Decision Tree Data Mining Technique,|| *International Journal of Computer Applications*, 2014, vol. 98, no. 10, pp. 0975 – 8887.
- [23] M. Elgedawy, —Prediction of Breast Cancer using Random Forest, Support Vector Machines and Naïve Bayes,|| *International Journal of Engineering and Computer Science*, 2017, vol. 6, no. 1, pp. 19884- 19889.
- [24] R. Senkamalavalli, and T. Bhuvaneshwari,|| *Improved classification of breast cancer data using hybrid techniques*, —*International Journal of Advanced Research in Computer Science*. 2017, vol. 8, no. 8, pp. 454-457.
- [25] A. Hazra, S. Mandal, and A. Gupta|| *Study and Analysis of Breast Cancer Cell Detection using Naïve Bayes, SVM and Ensemble Algorithms*,|| *International Journal of Computer Applications*. 2016, vol. 145, no.2, pp. 0975 – 8887.
- [26] S. Gc, R. Kasaudhan, T. K. Heo, and H.D. Choi, —Variability Measurement for Breast Cancer Classification of Mammographic Masses,|| in *Proceedings of the 2015 Conference on research in adaptive and convergent systems (RACS)*, Prague, Czech Republic, 2015, pp. 177–182.
- [27] C. Wang, W. Wang, S. Shin, and S. I. Jeon, —Comparative Study of Microwave Tomography Segmentation Techniques Based on GMM and KNN in Breast Cancer Detection,|| in *Proceedings of the 2014 Conference on Research in Adaptive and Convergent* 303–308.
- [28] C. L. Chowdhary, and D. P. Acharjya, —Breast Cancer Detection using Intuitionistic Fuzzy Histogram Hyperbolization and Possibilitic Fuzzy c-mean Clustering algorithms with texture feature-based Classification on Mammography Images,|| in *Proceedings of the International Conference on Advances in Information Communication Technology & Computing*, Bikaner, India, 2016, pp. 1–6.
- [29] Bryce TJ, Dewhirst MW, Floyd CE, Jr, et al. Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys*. 1998;41:239–45. [[PubMed](#)] [[Google Scholar](#)]
- [30] Burke HB, Bostwick DG, Meiers I, et al. Prostate cancer outcome: epidemiology and biostatistics. *Anal Quant Cytol Histol*. 2005;27:211–7. [[PubMed](#)] [[Google Scholar](#)]
- [31] Catto JW, Linkens DA, Abbod MF, et al. Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modeling and artificial neural networks. *Clin Cancer Res*. 2003;9:4172–7. [[PubMed](#)] [[Google Scholar](#)]

- [32] Claus EB. Risk models used to counsel women for breast and ovarian cancer: a guide for clinicians. *Fam Cancer*. 2001;1:197–206. [[PubMed](#)] [[Google Scholar](#)]
- [33] Cochran AJ. Prediction of outcome for patients with cutaneous melanoma. *Pigment Cell Res*. 1997;10:162–7. [[PubMed](#)] [[Google Scholar](#)]
- [34] Cortes C, Vapnik V. Support-vector networks. *Machine Learning*. 1995;20:273–297. [[Google Scholar](#)]
- [35] Crawford ED, Batuello JT, Snow P, et al. The use of artificial intelligence technology to predict lymph node spread in men with clinically localized prostate carcinoma. *Cancer*. 2000;88:2105–9. [[PubMed](#)] [[Google Scholar](#)]
- [36] Dai H, van't Veer L, Lamb J, et al. A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients. *Cancer Res*. 2005;65:4059–66. [[PubMed](#)] [[Google Scholar](#)]
- [37] Dettling M. BagBoosting for tumor classification with gene expression data. *Bioinformatics*. 2004;20:3583–93. [[PubMed](#)] [[Google Scholar](#)]

