LITERATURE REVIEW

Cancer subtype classification has the potential to significantly improve disease prognosis and develop individualized patient management. Existing methods are limited by their ability to handle extremely high-dimensional data and by the influence of misleading, irrelevant factors, resulting in ambiguous and overlapping subtypes.[1]

Gene expression profiling based on microarray to figure out breast cancer classification provided vague insights for clinical study. [2]Further, development of a risk model that incorporates the gene expression–based “intrinsic” subtypes (luminal A, luminal B, HER2-enriched, and basal-like ) model predicted neoadjuvant chemotherapy efficacy with a negative predictive value for pCR of 97%.[3]

An integrated analysis of how copy number affects gene expression proved that Inherited variants (copy number variants and single nucleotide polymorphisms) and acquired somatic copy number aberrations (CNAs) were associated with expression in ∼40% of genes, with the landscape dominated by *cis*- and *trans*-acting CNAs.[4]

Microarray-based expression studies have demonstrated that breast cancer is both a clinically diverse and molecularly heterogeneous disease comprising subtypes with distinct gene expression patterns that are associated with outcome.[5]Several research groups have developed prognostic gene signatures by analyzing gene expression together with clinical outcome data[6].

The relevance of these subtypes for basic and translational research has led to their use in prognostic assessments [7], prediction of therapeutic efficacy [8], and retrospective analysis of clinical trials [9].

Today, Machine Learning (ML) techniques are being broadly used in the breast cancer classification problem. They provide high classification accuracy and effective diagnostic capabilities. In this paper, we present two different classifiers: Naive Bayes (NB) classifier and knearest neighbor (KNN) for breast cancer classification. We propose a comparison between the two new implementations and evaluate their accuracy using cross validation. Results show that KNN gives the highest accuracy (97.51%) with lowest error rate then NB classifier (96.19 %).[10]

The breast cancer is classified with serval techniques such as Softmax Discriminant Classifier (SDC), Linear Discriminant Analysis (LDA) [11], and Fuzzy C Means Clustering [12]. The knearest neighbors algorithm is one of the most used algorithms in machine learning [13]. Before classifying a new element, we must compare it to other elements using a similarity measure [14].

 In this paper, breast cancer diagnosis was conducted using least square support vector machine (LS-SVM) classifier algorithm. The robustness of the LS-SVM is examined using classification accuracy, analysis of sensitivity and specificity, k-fold cross-validation method and confusion matrix. The obtained classification accuracy is 98.53% and it is very promising compared to the previously reported classification techniques. Consequently, by LS-SVM, the obtained results show that the used method can make an effective interpretation and point out the ability of design of a new intelligent assistance diagnosis system.[15]

In this paper, a CAD scheme for detection of breast cancer has been developed using deep belief network unsupervised path followed by back propagation supervised path. The construction is back-propagation neural network with Liebenberg Marquardt learning function while weights are initialized from the deep belief network path (DBN-NN). Our technique was tested on the Wisconsin Breast Cancer Dataset (WBCD). The classifier complex gives an accuracy of 99.68% indicating promising results over previously-published studies. The proposed system provides an effective classification model for breast cancer. In addition, we examined the architecture at several train-test partitions [16]

1)Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data Runpu Chen, Le Yang, Steve Goodison, Yijun Sun

2) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.

*Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lønning PE, Børresen-Dale AL*

*Proc Natl Acad Sci U S A. 2001 Sep 11; 98(19):10869-74.*

3) Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes[J Clin Oncol](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667820/). 2009 Mar 10; 27(8): 1160–1167.

Published online 2009 Feb 9. doi: [10.1200/JCO.2008.18.1370](https://dx.doi.org/10.1200%2FJCO.2008.18.1370)

[Joel S. Parker](https://www.ncbi.nlm.nih.gov/pubmed/?term=Parker%20JS%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Michael Mullins](https://www.ncbi.nlm.nih.gov/pubmed/?term=Mullins%20M%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Maggie C.U. Cheang](https://www.ncbi.nlm.nih.gov/pubmed/?term=Cheang%20MC%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Samuel Leung](https://www.ncbi.nlm.nih.gov/pubmed/?term=Leung%20S%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [David Voduc](https://www.ncbi.nlm.nih.gov/pubmed/?term=Voduc%20D%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Tammi Vickery](https://www.ncbi.nlm.nih.gov/pubmed/?term=Vickery%20T%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Sherri Davies](https://www.ncbi.nlm.nih.gov/pubmed/?term=Davies%20S%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Christiane Fauron](https://www.ncbi.nlm.nih.gov/pubmed/?term=Fauron%20C%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Xiaping He](https://www.ncbi.nlm.nih.gov/pubmed/?term=He%20X%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Zhiyuan Hu](https://www.ncbi.nlm.nih.gov/pubmed/?term=Hu%20Z%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [John F. Quackenbush](https://www.ncbi.nlm.nih.gov/pubmed/?term=Quackenbush%20JF%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Inge J. Stijleman](https://www.ncbi.nlm.nih.gov/pubmed/?term=Stijleman%20IJ%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Juan Palazzo](https://www.ncbi.nlm.nih.gov/pubmed/?term=Palazzo%20J%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [J.S. Marron](https://www.ncbi.nlm.nih.gov/pubmed/?term=Marron%20J%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Andrew B. Nobel](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nobel%20AB%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Elaine Mardis](https://www.ncbi.nlm.nih.gov/pubmed/?term=Mardis%20E%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Torsten O. Nielsen](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nielsen%20TO%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Matthew J. Ellis](https://www.ncbi.nlm.nih.gov/pubmed/?term=Ellis%20MJ%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), [Charles M. Perou](https://www.ncbi.nlm.nih.gov/pubmed/?term=Perou%20CM%5BAuthor%5D&cauthor=true&cauthor_uid=19204204), and [Philip S. Bernard](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bernard%20PS%5BAuthor%5D&cauthor=true&cauthor_uid=19204204)

# 4)The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups

* [Christina Curtis](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-1),
* [Sohrab P. Shah](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-2),
* [Suet-Feung Chin](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-3),
* [Gulisa Turashvili](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-4),
* [Oscar M. Rueda](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-5),
* [Mark J. Dunning](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-6),
* [Doug Speed](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-7),
* [Andy G. Lynch](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-8),
* [Shamith Samarajiwa](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-9),
* [Yinyin Yuan](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-10),
* [Stefan Gräf](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-11),
* [Gavin Ha](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-12),
* [Gholamreza Haffari](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-13),
* [Ali Bashashati](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-14),
* [Roslin Russell](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-15),
* [Steven McKinney](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-16),
* [METABRIC Group](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#group-1),
* [Anita Langerød](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-103),
* [Andrew Green](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-104),
* [Elena Provenzano](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-105),
* [Gordon Wishart](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-106),
* [Sarah Pinder](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-107),
* [Peter Watson](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-108),
* [Florian Markowetz](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-109),
* [Leigh Murphy](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-110),
* [Ian Ellis](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-111),
* [Arnie Purushotham](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-112),
* [Anne-Lise Børresen-Dale](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-113),
* [James D. Brenton](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-114),
* [Simon Tavaré](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-115),
* [Carlos Caldas](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-116) &
* [Samuel Aparicio](https://www.nature.com/articles/nature10983?__hstc=12316075.cde2cb5f07430159d50a3c91e72c280a.1492646400108.1492646400109.1492646400110.1&__hssc=12316075.1.1492646400111&__hsfp=1773666937#auth-117)
* [- Show fewer authors](javascript:;)

[*Nature*](https://www.nature.com/nature) **volume 486**, pages346–352(2012)

5) Desmedt C, Haibe-Kains B, Wirapati P, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes, Clin Cancer Res., 2008, vol. 14 16(pg. 5158-5165)

6) 12.SotiriouC,  Pusztai. Gene-expression signatures in breast cancer, *N Engl J Med.*, 2009, vol. 360(pg. 790-800)

7) Parker JS, Mullins M, Cheang MCU, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes, J Clin Oncol., 2009, vol. 27 8(pg. 1160-1167)

8) Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer, J Clin Oncol., 2008, vol. 26 8(pg. 1275-1281)

9) Pusztai L, Broglio K, Andre F, et al. Effect of molecular disease subsets on disease-free survival in randomized adjuvant chemotherapy trials for estrogen receptor-positive breast cancer, J Clin Oncol., 2008, vol. 26 28(pg. 4679-4683)

10) Breast cancer classification using machine learning[2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)](https://ieeexplore.ieee.org/xpl/conhome/8384822/proceeding) [10.1109/EBBT.2018.8391453](https://doi.org/10.1109/EBBT.2018.8391453)

**Publisher:** IEEE

11) A. Alarabeyyat, A.M., "Breast Cancer Detection Using K-Nearest Neighbor Machine Learning Algorithm" in 9th International Conference on, IEEE, v.i.e.E.(DeSE), pp. 35-39, 2016

12) MF. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis", Expert Syst Appl, vol. 36, no. 2, pp. 3240-3247, March 2009.

13) S.K. Prabhakar, H. Rajaguru, N. Maglaveras, I. Chouvarda, P. de Carvalho, "Performance Analysis of Breast Cancer Classification with Softmax Discriminant Classifier and Linear Discriminant Analysis" in Precision Medicine Powered by pHealth and Connected Health. IFMBE Proceedings, Singapore:Springer, vol. 66, 2018.

14) P. Baldi, S.R.B., S.r.B. Pierre Baldi, Bioinformatics: The machine learning approach, 2001.

15) Author links open overlay panel[Ahmed M.Abdel-Zaher](https://www.sciencedirect.com/science/article/pii/S0957417415007101" \l "!)[Ayman M.Eldeib](https://www.sciencedirect.com/science/article/pii/S0957417415007101" \l "!) <https://doi.org/10.1016/j.eswa.2015.10.015>

16) Breast cancer diagnosis using least square support vector machine https://doi.org/10.1016/j.dsp.2006.10.008 K. Polat,S. Güneş

17)[1] Sun, D., Wang, M., & Li, A. (2019). A Multimodal Deep Neural Network for Human Breast Cancer Prognosis Prediction by Integrating Multi-Dimensional Data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *16*(3), 841–850. doi: 10.1109/tcbb.2018.2806438

17) S, N. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 1–30. Retrieved from <http://jmlr.org/papers/volume15/srivastava14a.old/srivastava14a.pdf>