TEAM ID : NM2023TMID14694

PROJECT TOPIC : ADVANCED BREAST CANCER PREDICTION WITH DEEP LEARNING

Introduction

Breast cancer is one of the major causes of death in women around the world. According to the American cancer society, 41,760 women and more than 500 men died from breast cancer recently1. Breast cancer occurs in four main types: normal, benign, in-situ carcinoma and invasive carcinoma [1]. A benign tumor involves a minor change in the breast structure. It is not harmful and does not classify as a harmful cancer. In cases of in-situ carcinoma, the cancer is only in the mammary duct lobule system and does not affect other organs. This type is not dangerous and can be treated if diagnosed early. Invasive carcinoma is considered to be the most dangerous type of breast cancer, as it can spread to all other organs. According to the authors in [1], breast cancer can be detected using several methods including X-ray mammography, ultrasound (US), Computed Tomography (CT), Portion Emission Tomography (PET), Magnetic Resonance Imaging (MRI) and breast temperature measurement. Usually, the golden standard is a pathological diagnosis for detecting breast cancer. This involves an image analysis of the removed tissue, which is stained in the lab to increase visibility. Hematoxylin and Eosin (H&E) are commonly used for the staining process. Breast cancer can be diagnosed using one of two approaches: histopathological image analysis or genomics. Histopathological images are microscopic images of breast tissue that are extremely useful in early treatment of the cancer.

Radio-genomics provide both radiological and genetic features that can enhance diagnosis. It can analyze tissues at the molecular level, helping with prediction and early detection of cancer. The main difference between imaging information and radio-genomics is the critical knowledge gap between imaging at the tissue level and analyzing the underlying molecular and genetic disease biomarkers. As imaging is less precise, it may lead to over- or under-treatment. While radio-genomics is much more effective than histopathological imaging, it is rarely used because the process involves datasets that are very expensive and require high computational power. As a result, a limited number of labs conduct radio-genomics experiments [2]. This research paper addresses the following research questions and highlights the deep learning models, looking at their performance, the datasets used and possibilities for breast cancer classification and detection.

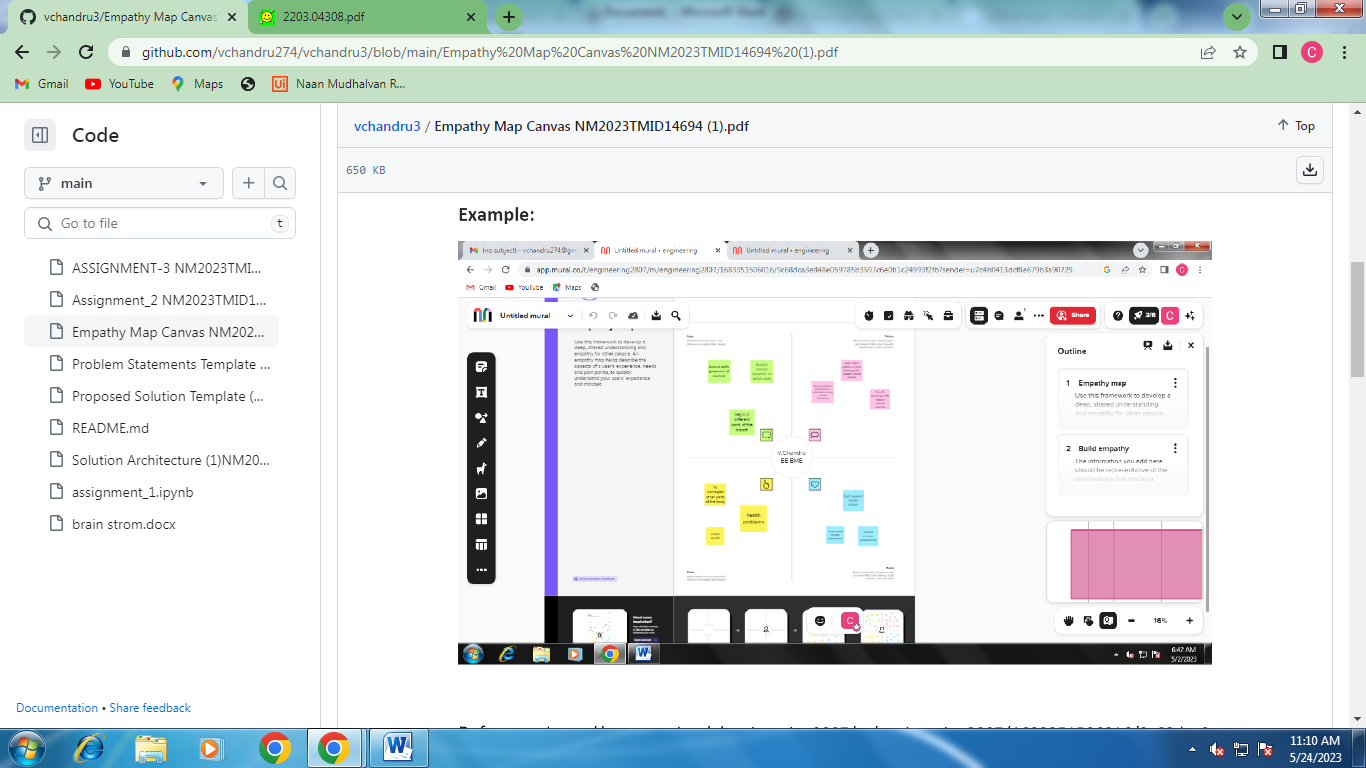
Abstract

Cancer is one of the most dangerous diseases to humans, and yet no permanent cure has been developed for it. Breast cancer is one of the most common cancer types. According to the National Breast Cancer foundation, in 2020 alone, more than 276,000 new cases of invasive breast cancer and more than 48,000 non-invasive cases were diagnosed in the US. To put these figures in perspective, 64% of these cases are diagnosed early in the disease’s cycle, giving patients a 99% chance of survival. Artificial intelligence and machine learning have been used effectively in detection and treatment of several dangerous diseases, helping in early diagnosis and treatment, and thus increasing the patient’s chance of survival. Deep learning has been designed to analyze the most important features affecting detection and treatment of serious diseases. For example, breast cancer can be detected using genes or histopathological imaging. Analysis at the genetic level is very expensive, so histopathological imaging is the most common approach used to detect breast cancer. In this research work, we systematically reviewed previous work done on detection and treatment of breast cancer using genetic sequencing or histopathological imaging with the help of deep learning and machine learning. We also provide recommendations to researchers who will work in this field. Key Words: Breast Cancer; Machine Learning; Deep Learning

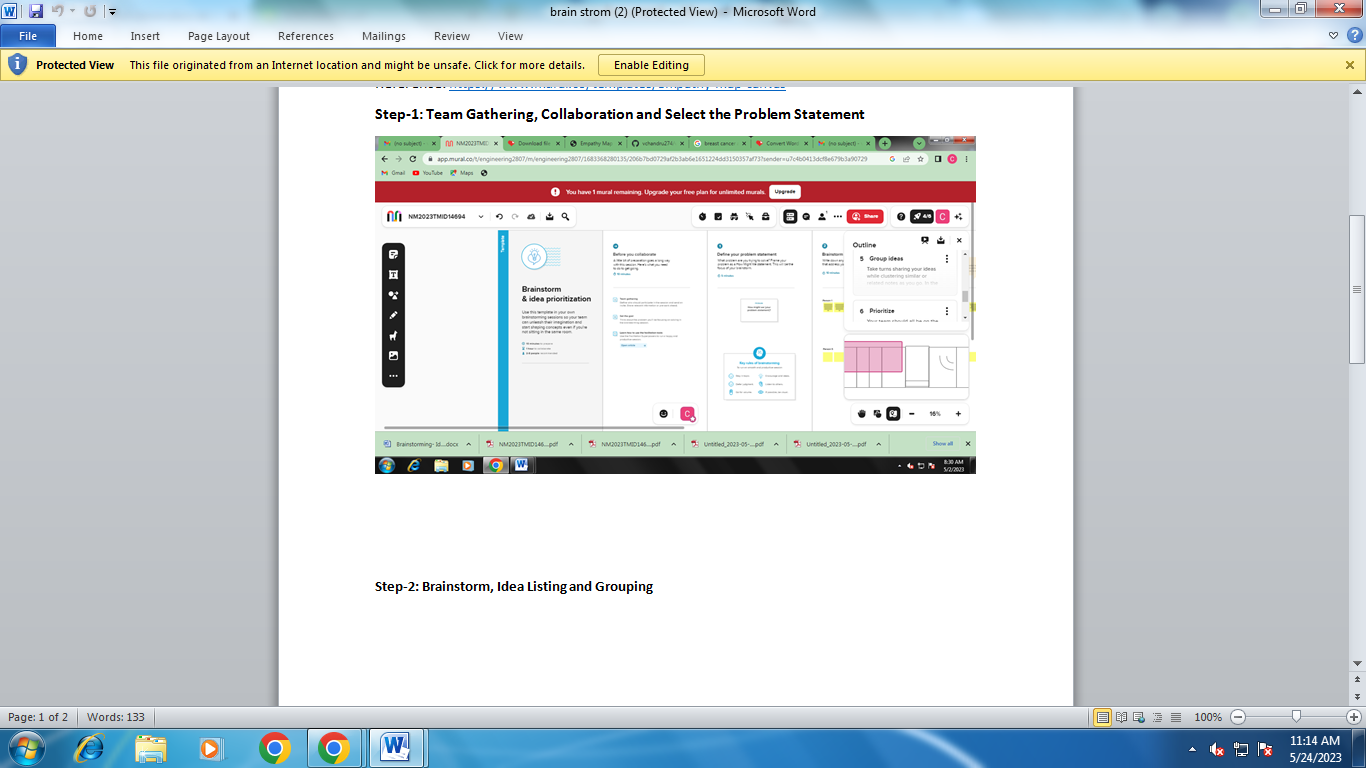
Conclusions and Future Research Directions

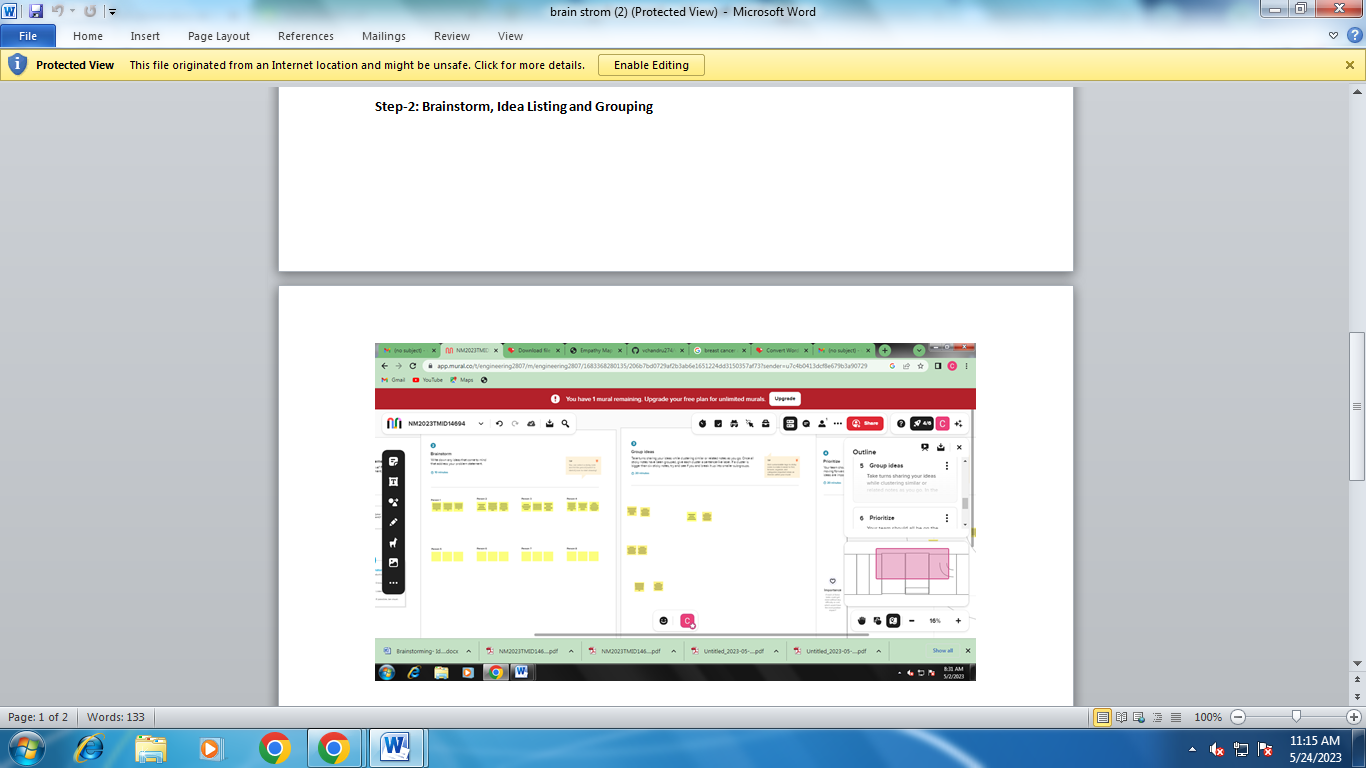
Most papers published in the field of breast cancer detection and subtype classification use machine learning techniques. However, deep learning models have not been heavily investigated in this domain. This presents researchers with opportunities to use various deep learning mechanisms to predict patient status such as LSTM, GAN and RNN, as these types of research have not yet been conducted in the field. Moreover, most papers focus only on the Accuracy metric to evaluate their performance and ignore confusion matrix parameters and AUC. This is insufficient since the Accuracy metric does not distinguish between false positive and false negative classifications. Future studies should include at least AUC and F-scores to assess the performance of each model. An important finding has been noticed which is the wide use of the CNN algorithm for both gene expression and MRI images data. Such models often obtain good results in comparison to other algorithms. Researchers might be interested to do further searching and implement more hybrid algorithms with CNN.

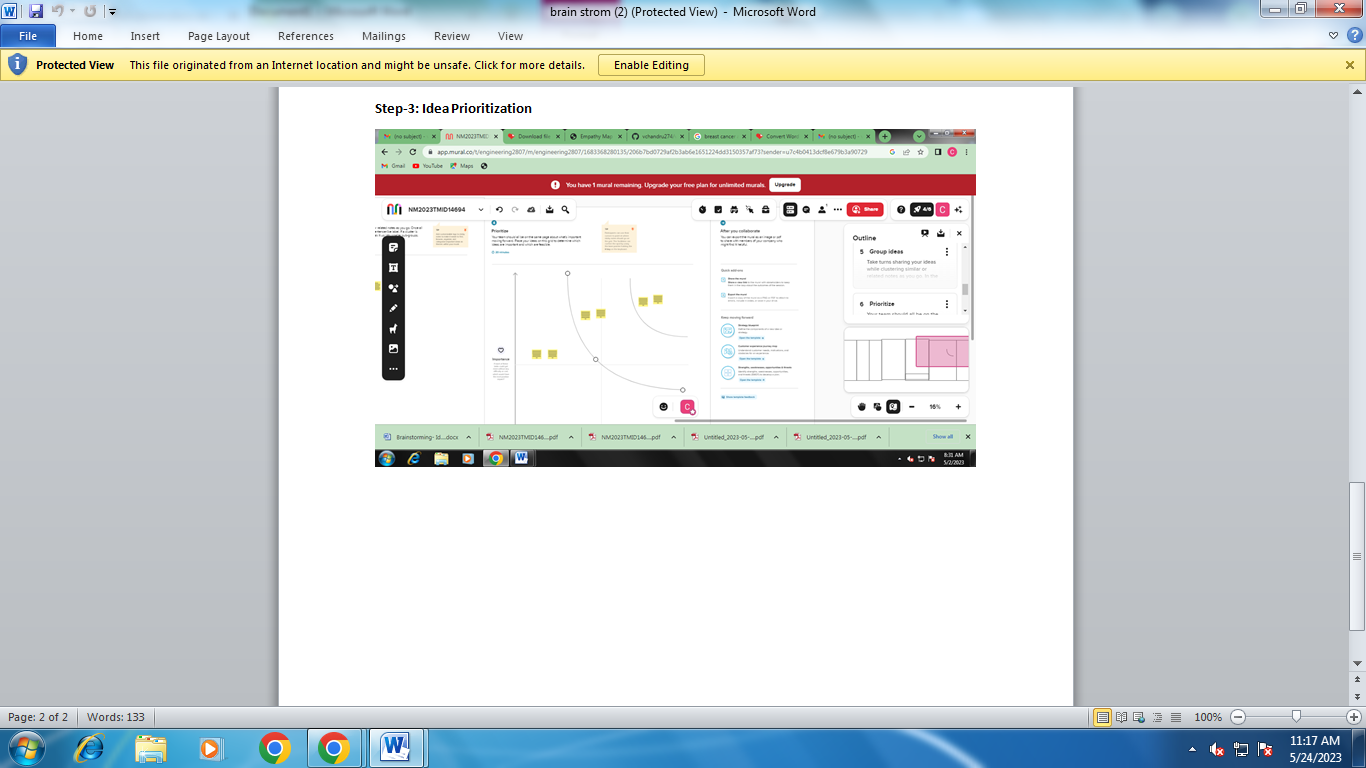
Empathy map



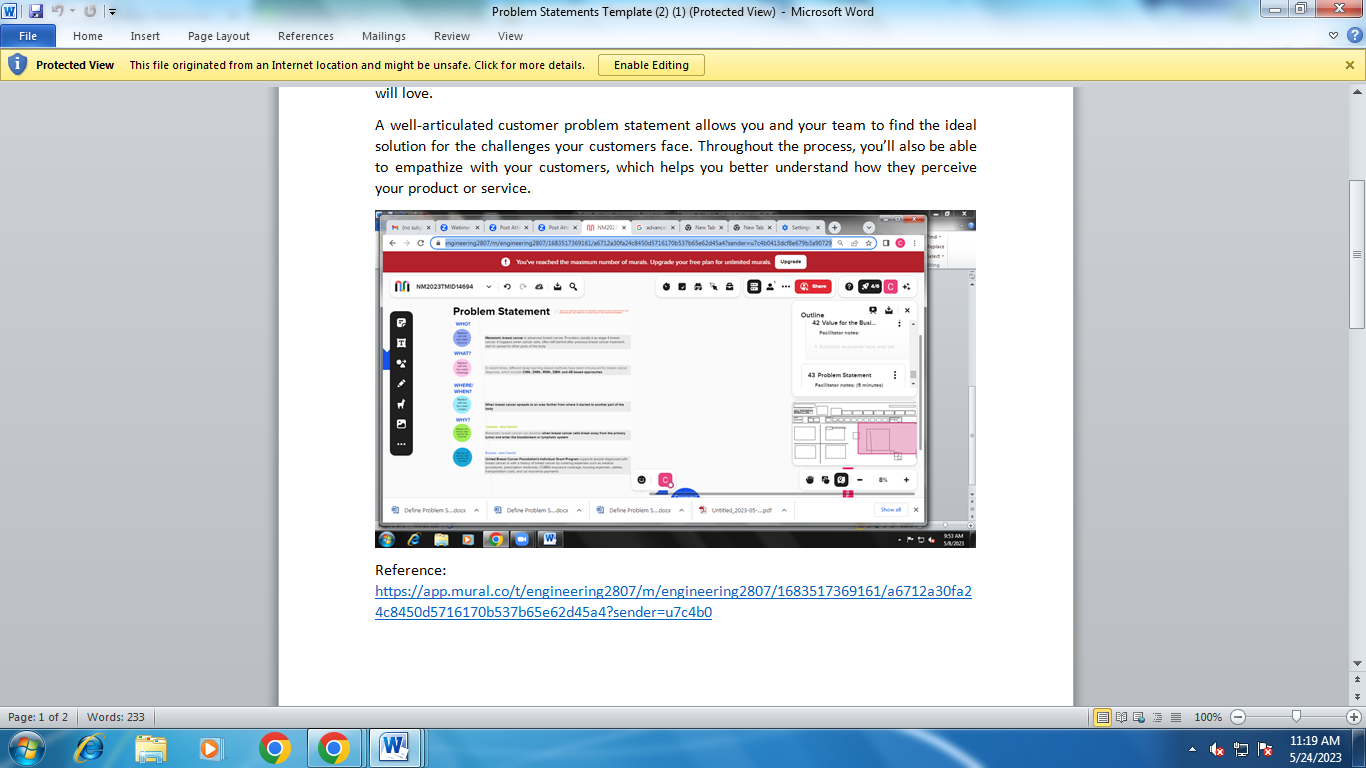
Brain storming

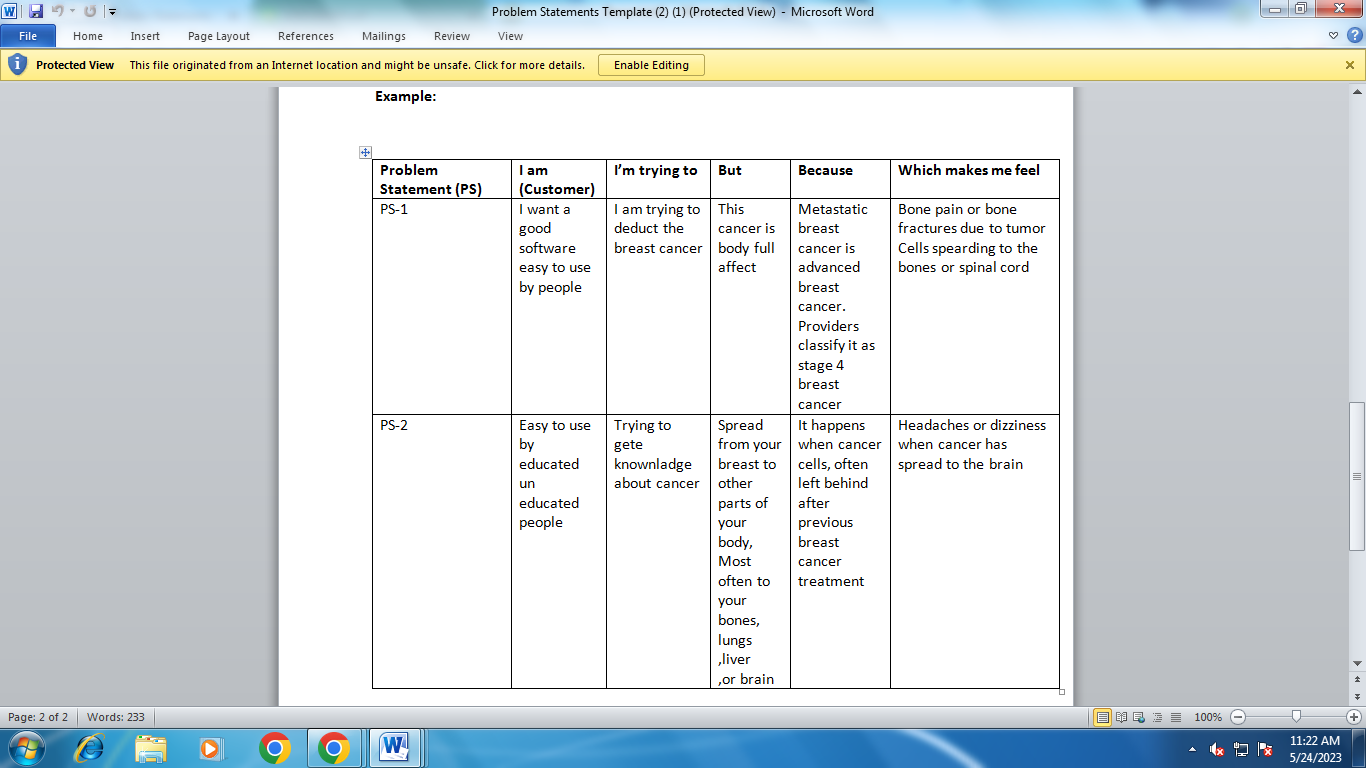






Problem statemant





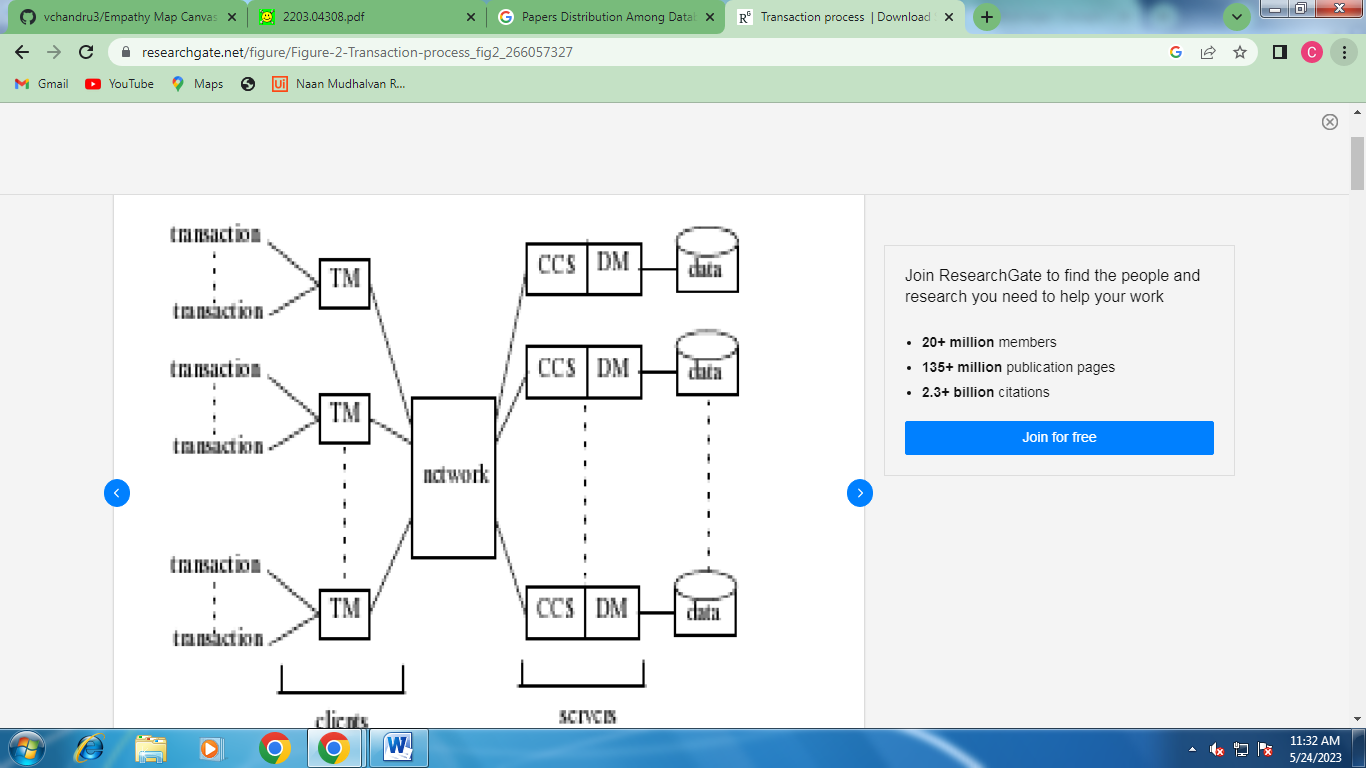
Related Work

Many studies have been conducted about breast cancer detection through imaging or through genomics. However, to the best of our knowledge, no research has been conducted including both techniques. The authors in [1] summarized the various techniques used to classify breast cancer using histopathological image analysis (HIA) based on different architectures of Artificial Neural Networks (ANN). The authors

grouped their work according to the applied dataset. They arranged it in ascending chronological order. This work found that ANNs were first used in the field of HIA around 2012. ANNs and PNNs were the most frequently applied algorithms. However, in feature extraction, most of the work used textural and morphological features. It was clear that Deep CNNs were quite effective for early detection and diagnosis of breast cancer, leading to more successful treatment. Prediction of Non-Communicable Diseases (NCDs) was conducted using many algorithms. In [2], the authors compared the performance of various classification algorithms. The classification algorithms were performed on eight NCD datasets using eight classification algorithms and a 10-fold crossvalidation method. These were evaluated using AUC as an indicator of accuracy. The authors stated that the NCD datasets have noisy data and irrelevant attributes. KNN, SVM and NN proved to be robust to this noise. In addition, they stated that the irrelevant attribute problem can be handled with some pre-processing techniques to improve the accuracy rate. Natural inspired computing (NIC) algorithms have been designed and applied to diagnose various human disorders. The authors in [3] introduced five insect-based NIC algorithms used for diagnosing diabetes and cancer. The authors found that it achieved a high level of performance in detecting different types of cancer (breast, lung, prostate and ovarian). To be more specific, breast cancer was detected using a hybridization of the guided ABC and neural networks. The authors also developed a highly effective methodology of detecting diabetes and leukemia. They concluded that the hybridization of NICs with other classification algorithms produces more precise and promising results. They mentioned that more work is required to detect different stages of diabetes and cancer. In [4], the authors demonstrated the effectiveness of NNs in the classification of cancer diagnoses, especially in the initial stages. According to their study, the majority of NNs have shown promise in detecting tumor cells. However, the imaging approach requires high computational capacity to preprocess the images. In [5], the authors reviewed different machine learning, deep learning and data mining algorithms related to breast cancer prediction. Several research papers on breast cancer were reviewed, with a total of 27 papers in machine learning, 4 papers in ensemble techniques and 8 papers in deep learning techniques. The authors mentioned that most of the papers used imaging, while only a few papers used genetics. The main algorithms used to predict breast cancer using genetics were SVM, decision tree and random forest. However, imaging techniques used several algorithms such as CNNs and Naïve Bayes.

Methodology

Our target topic is breast cancer detection using deep learning. We ended up using around 80 of the most recent papers related to breast cancer treatment and diagnosis. Some of the papers examined only deep learning, while others used a combination of machine learning and deep learning. In our search process, we mainly used the Scopus database to obtain the articles. This is to exclude nonrefereed publications However, in Figure 1, we state the distribution of selected papers among the existing databases. The top five databases are PubMed, ScienceDirect (Elsevier), IEEE, Springer and Nature.

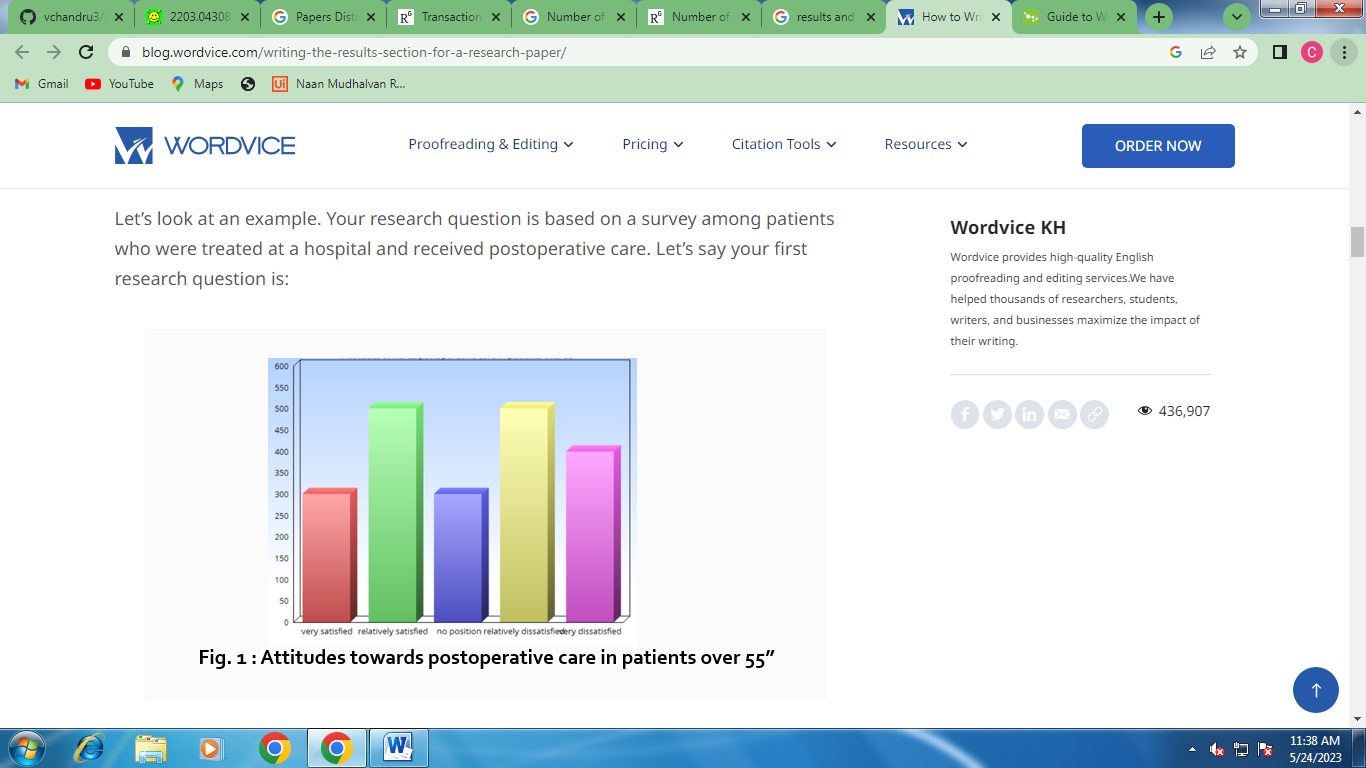


Number of publications between 2010 and 2020



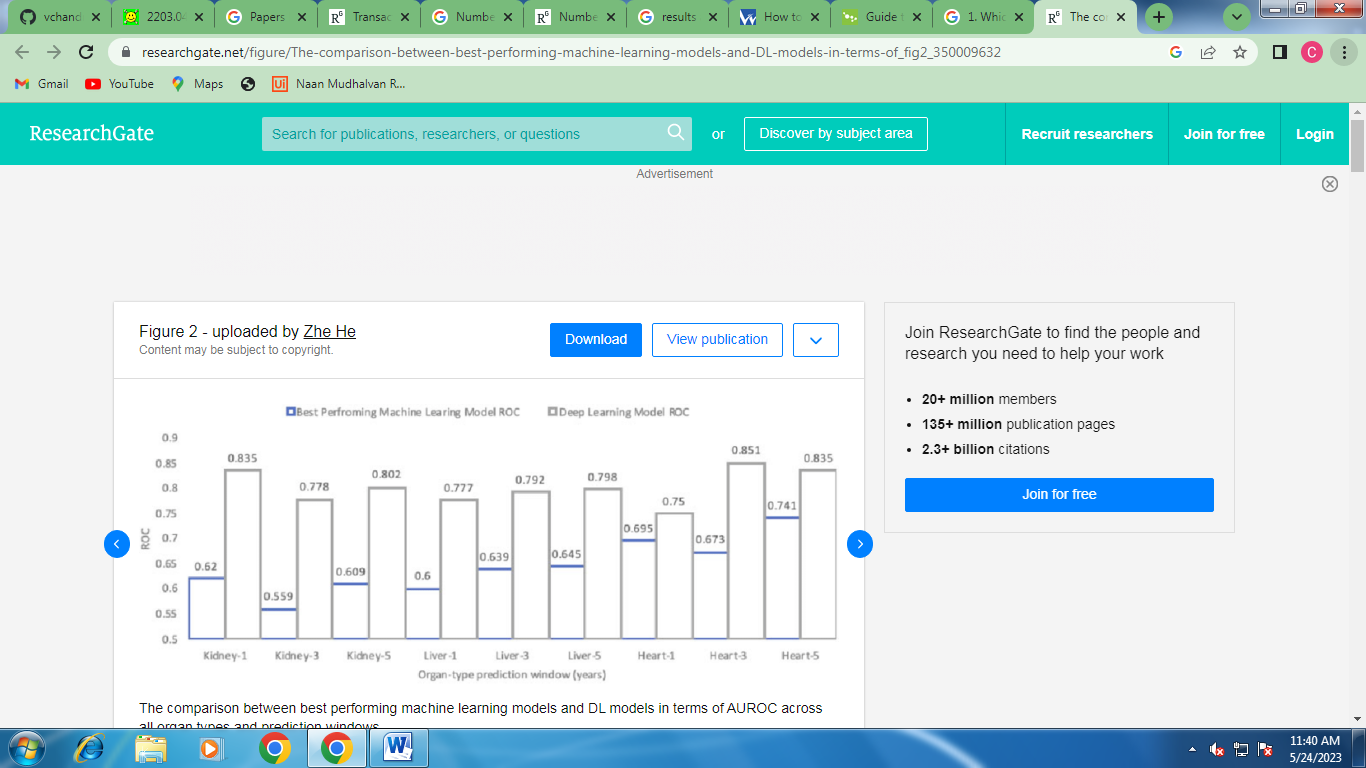
We can see that breast cancer study publications peaked around 2019 and 2020. We reduced the number of papers to 80 by including only papers using genetic expression and imaging and by focusing on journal and conference articles only. The imaging modalities we considered were ultrasound, radiography, mammography and magnetic resonance imaging (MRI), as well as various types of gene expression and gene sequencing. In this research, we focused on papers that implement the breast cancer detection using the techniques of AI, as well as papers that predict breast cancer using both gene data and image data. We applied the following eligibility criteria on each paper: (1) The language is English; (2) The topic is related to breast cancer detection and treatment; (3) The paper discusses hybrid models of machine learning and deep learning; (4) The paper purely discusses deep learning; (5) The paper discusses genetic expression data; (6) The paper discusses imaging data; (7) Only journal and conference publications are retained; (8) Only medical or biomedical engineering publications are kept which are related to the topic. Please note that non-refereed publications were excluded from the study. Firstly, we recorded the main information such as the name of the paper, year of publication, the list of authors and the publisher. Then, we included some information to conduct the systematic review, such as the algorithm used and whether the paper discusses only deep learning or a hybrid between DL and ML, the recorded accuracy and other performance evaluation parameters, the dataset, the features, and many other columns. We answered our research questions using these criteria.

Results and Discussion



In this systematic study, our initial search turned up 1,000 conference and journal papers. After eliminating duplicated papers and unrelated studies that were “purely medical or about cancer in general”, we ended up with 80 papers related to both ML and DL. We wanted to focus on DL approaches or DL-ML hybrid models, so only papers related to DL were selected. Figure 3 explains our search methodology

1. Which deep learning models perform most effectively?

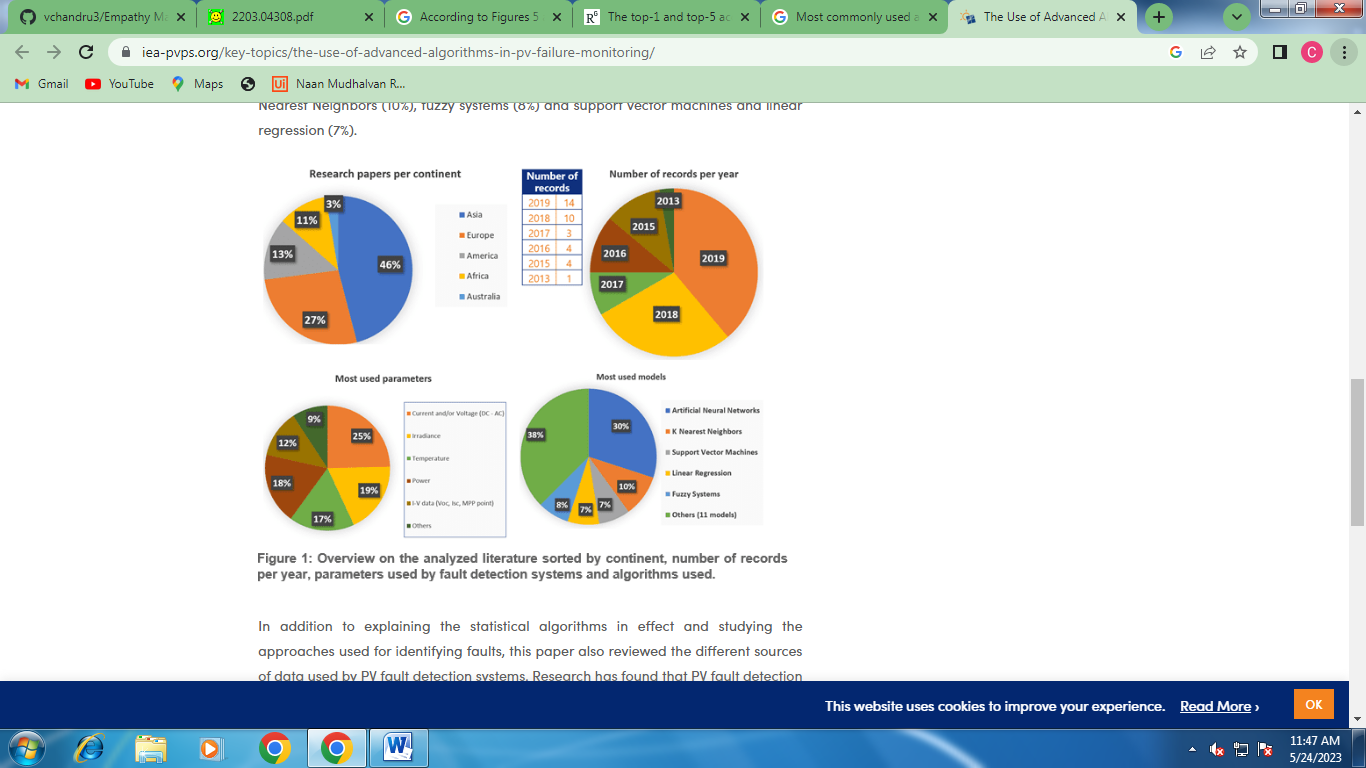


In the previous table, generally, the work in this area is divided into two main groups, the first involving binary classification (whether or not breast cancer is present) and the second classifying breast cancer types. We noticed that the use of binary classification results in the highest accuracy, and that it is generally more accurate than multiclass classification. In paper [11], the author used the BPNN algorithm in a highly effective way. We also see that most selected papers with binary classifications performed well in terms of most evaluation parameters. [18] However, the best performance for multiclass categorization or breast cancer subtypes classification obtained 97% accuracy. The author of the paper basically compared machine learning and deep learning on the task of breast cancer subtype classification. Most of the papers were strictly focused on accuracy and did not mention other parameters such as precision, recall, AUC, etc. This is an important limitation when we are talking about medical projects. Many models have been used, including CNN, DNN with attention mechanism, Feed Forward Neural Network (FFNN) and many other DL mechanisms. Because not all papers took into consideration the confusion matrix parameters, we will include only accuracy in our graphs. To state the findings from the above table, CNN model seems to be the number one used model among the papers for both binary and multiclass classification. A hybrid between machine learning and deep learning also seems to be affective for such models. We can see a hybrid between MLP and LR. As well as SVM algorithm from ML has been used a lot in these papers. Basically, when it comes to breast cancer detection using imaging data, we see high performance on binary classification. However, when it comes to classification of types, imaging did not produce results as accurately as gene expression data. Paper [77] obtained very high accuracy using a hybrid model between DL and ML for breast cancer detection. For multiclass differentiation or breast cancer subtype classification, the highest accuracy obtained using imaging data was 90% in a paper that uses ML. We can therefore see in the above table that in general

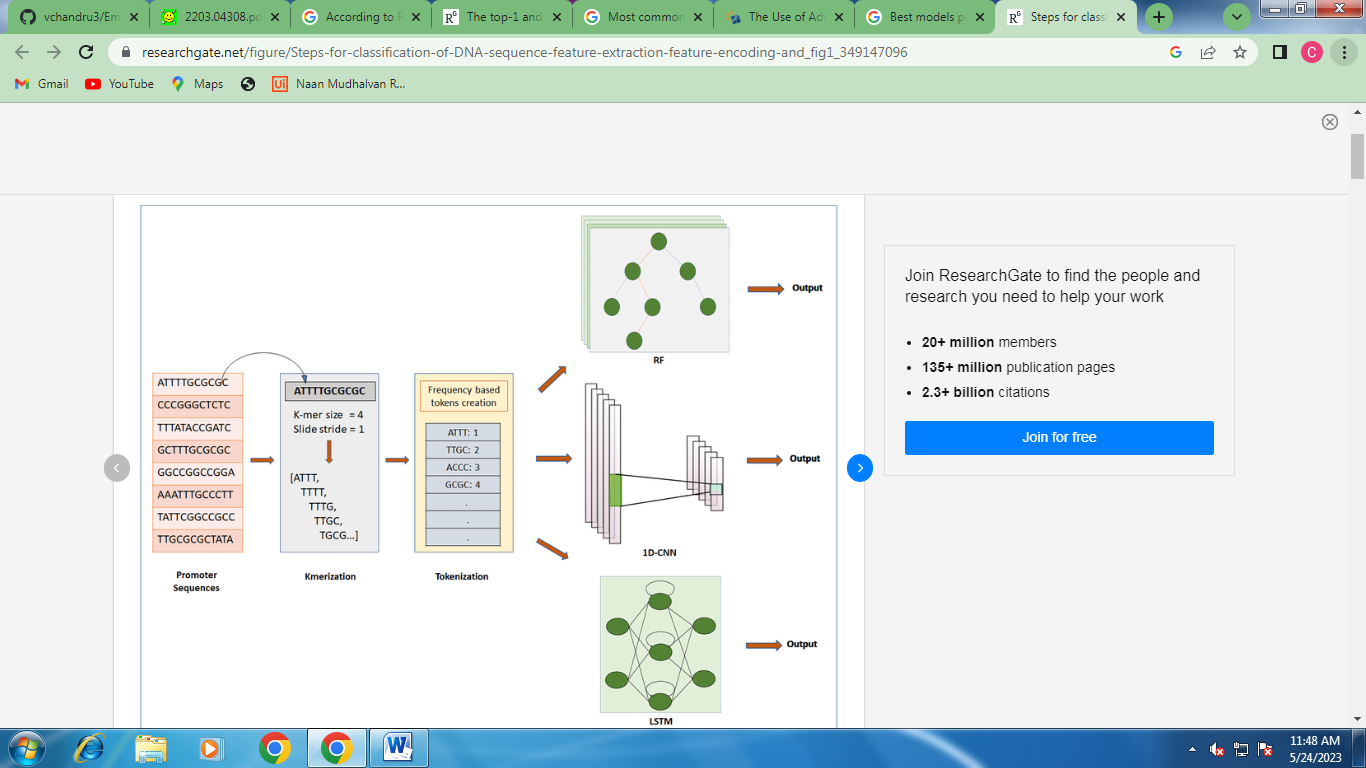
The use of CNN results in excellent performance for both gene expression and imaging. An example of this is paper [75], which produced 97% accuracy for binary classification. The models used include both deep learning standalone models and hybrid models consisting of both machine learning and deep learning algorithms. Comparing between hybrid and deep learning standalone models in gene expression, we can see that the standalone deep learning models obtain consistently higher accuracy. The hybrid technique in paper [14] obtained 87% accuracy, which was the lowest accuracy in all the models. The CNN in paper [12] obtained an use accuracy of 95%, while the algorithm BPNN achieved 99.8% accuracy. Many papers did not mention confusion matrix parameters. In fact, very few papers mentioned these parameters in which the authors mentioned sensitivity and specificity. Among the imaging papers, the highest accuracy was 99.7% for binary classification in paper [80], which used a hybrid model with both ML and DL.

According to the above table, the mostly used algorithm for breast cancer detecting and subtype classification is CNN, and that does make sense, since our data is MRI images and CNN considered the best for computer vision problems. To conclude, many algorithms were used in the studies. Some papers used several models in series, while others used only one model. According to Figure 4, ANN and CNN were the most widely used algorithms in both gene sequence and images data. Many other algorithms were used, such as DNN and SVM, but most of the papers used CNN and ANN with various parameters and properties.

Most commonly used algorithms in the papers

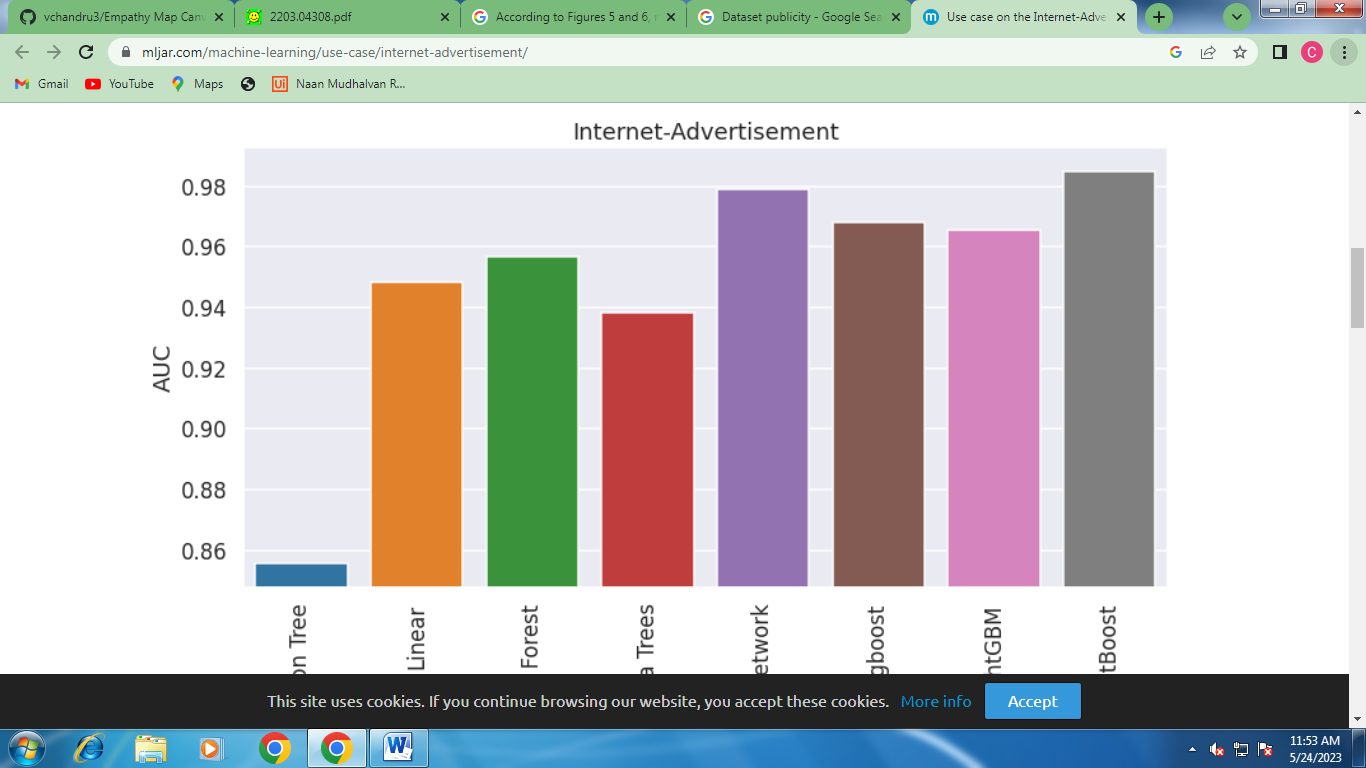


Best models performance for gene sequence classifiers

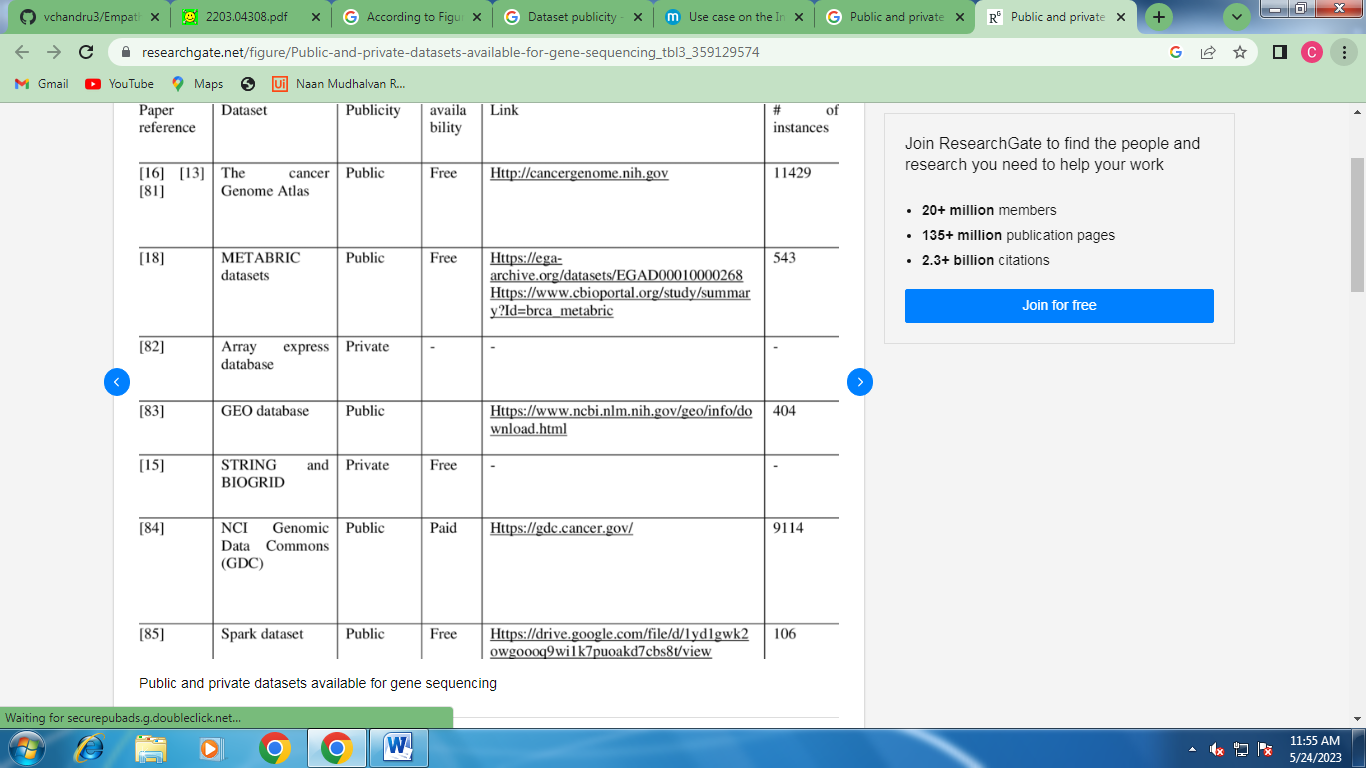


2. What datasets are available for gene sequencing and MRI?

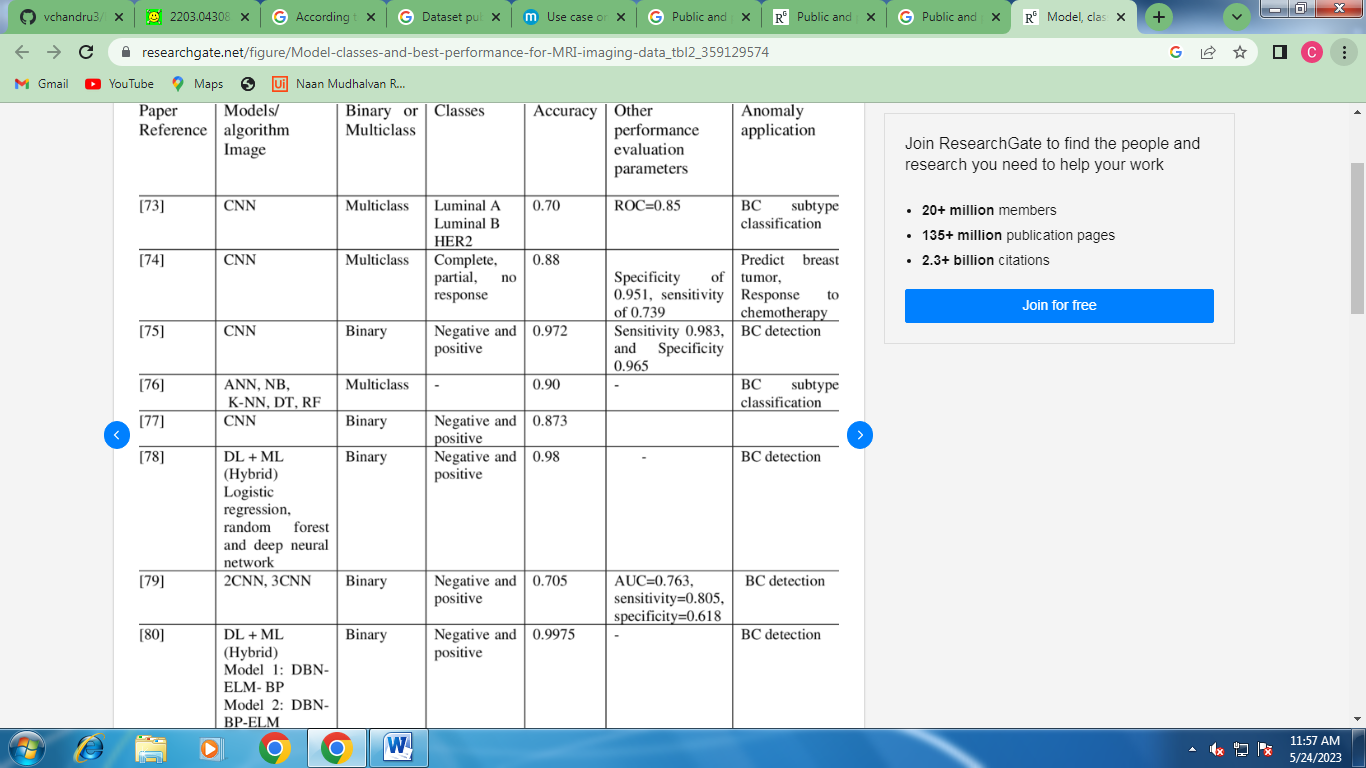
According to the figure below, many public and private datasets were used. The private datasets were sourced from various universities in the US and EU. There were very few publicly available datasets that were not for free. Table 3 shows the public and private datasets available for gene sequencing

****

Public and private datasets available for gene sequencing



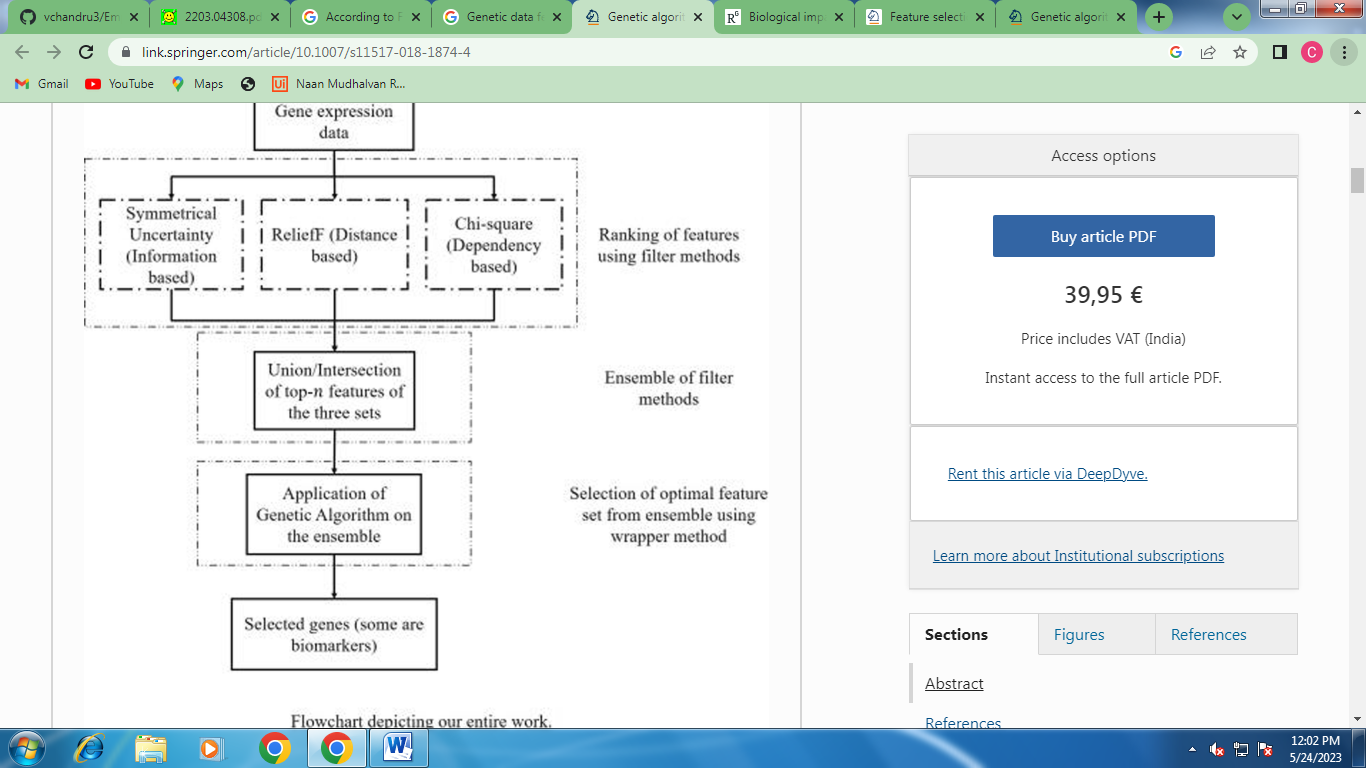
Although gene expression data is not as common as imaging data, we found many public and private datasets containing gene expression data for both healthy and sick people. The most well-known of these is the cancer Genome Atlas [86], which is a project aiming to identify the complete set of DNA changes in many different types of cancer. Studying these changes may help researchers understand how different types of cancer form. The dataset contains gene data for many cancer types, including breast cancer. The second most commonly used dataset is the METABRIC dataset, which contains clinical traits, expression, CNV profiles and SNP genotypes derived from breast tumors collected from participants of the METABRIC trial. The GEO database contains genome data and DNA sequencing from cancer detection research. NCI Genomic Data Commons (GDC) also provides researchers with a large number of gene related data for cancer research and analysis. The most used dataset in the above table is the cancer Genome Atlas dataset. It provides researchers with a big number of instances. It provides a clinical data for each participants including some general information.



More datasets are available for imaging data than for genetic data. The most used imaging dataset is the Wisconsin breast cancer dataset, which is obtained from the UCI repository. It contains features that are computed from a digitized image of a Fine Needle Aspirate (FNA) of a breast mass. Many other breast cancer imaging datasets could be found, most of them public and free. If the dataset is very large, like the DDSM dataset, it can be used on its own. If the dataset is not large enough to produce high performance results, it can be merged with another dataset to improve the data pool. Researchers mainly used Wisconsin breast cancer dataset and MRI dataset, as we can notice both of them are public and contain a lot of samples.

A. What are the most commonly used features for breast cancer classification

We can see that the features are divided into two categories: tumor information and protein types/status. The tumor information includes the size of the tumor and the grade, which depends on cancer type. The ER and PR status represent the proteins that exist in the area. Number of positive lymph nodes represents how many lymph nodes contain the cancer and is recorded as a real number. Metastasis records the number of sites the cancer has spread into, and may include the lungs, bones, etc. HER2 is another protein that indicates the growth of the cancer inside the breast area. PAM50 is an index or test that shows if breast cancers are likely to metastasize (spread to other organs). The following tables (Tables 5 & 6) show the features, their description and types for gene as well as imaging



When looking at imaging data, many general features are considered such as symmetry, compactness and concave points. Another type of feature is related to the breast image specifically. For example, marginal adhesion quantifies the degree to which cells on the outside of the epithelial wall tend to stick together. Another example is uniformity of cell size, which is represented by 0 if lacking uniformity and 1 if uniform.

Clump thickness describes if cells are mono- or multi-layered. Uniformity of cell size evaluates consistency in cell size in the sample.

What are the most effective feature selection and feature extraction methods:

Feature extraction methods:

The literature discusses various methods for feature extraction. Feature-based Strategy (FES) is a method used in some of the papers for finding image displacements. This strategy pinpoints features (for example image edges, corners and other structures that can be localized in two dimensions) and tracks these as they move from frame to frame. This involves two stages. Firstly, the features are found in two or more consecutive images. The act of feature extraction, if done correctly, will reduce the amount of information to be processed (and so reduce the workload). It will also contribute to obtaining a higher level of understanding in the field, by its very nature of eliminating the unimportant information. Secondly, these features are matched between the frames. In the simplest and commonest cases, two frames are used and two sets of features are matched to produce a single set of motion vectors. Alternatively, the features in one frame can be used as seed points from which to use other methods (i.e.: gradient-based methods -- see the following section) to find the motion [23]. Other authors used a multi-level wavelet transformation method, which is a wavelet-based discrete signal analysis method that can extract multilevel time-frequency features from a time series by decomposing the series level by level into low and high frequency sub-series. CNN was also used for feature selection and extraction. This is a very broad question, but you can look at general CNN architecture for two main types of imaging classification: “feature extractor” based on convolutional layers, and “classifier”, usually based on fully connected layers. Feature extraction usually refers to one of two options: either the last hidden layer – i.e.: the last layer before the output layer [see vgg16 example below, the 4096x1x1 layer], or the last convolutional layer after flattening [in vgg16 -the conv5 layer 7x7x512] [93].

Feature selection methods

Several feature selection methods were used, starting with XGboost. The authors in [94] used XGboost and a random forest multilayer network analysis of mRNA and protein expression profiles in breast cancer patients. Random forests are often used for feature selection in a data science workflow because the treebased strategies used by random forests naturally ranks features by how well they improve the purity of the node. This means that they assign each feature a status on whether they reduce impurity over all trees (called GINI impurity). Nodes producing the greatest decreases in impurity occur at the start of the trees, while nodes with the smallest reduction in impurity occur at the end of trees. Thus, by pruning trees below a particular node, we can create a subset of the most important features. On the other hand, authors in [83] used principal component analysis (PCA), a technique for reducing the dimensionality of datasets, increasing interpretability but at the same time minimizing information loss. They achieved this by creating new uncorrelated variables that successively maximize variance.

Comparing gene sequence data with image data, for breast cancer detection problem, what is the drawbacks, challenges, and advantages for each:

In this survey, we compare between genetic data and imaging data as they are related to breast cancer detection. Here, we summarize the findings and results.

The differences between genetic sequencing data and imaging data

We collected papers that use either genetic sequencing data or imaging data to detect breast cancer or to classify it into subtypes. Both strategies produced good results. Regarding imaging data, many features could be easily extracted, but not all of them were effective. Genetic expression data contains fewer features but may be more effective. The best performance using imaging data in our systematic review produced 99.3% accuracy, whereas the most accurate use of genetic data resulted in 99.8% accuracy.

In general, dealing with imaging data is much easier than dealing with genetic data. There are more ways to preprocess imaging data, and more techniques to extract features from it. On the other hand, in our survey, we found that genetic data is consistently more accurate, especially in the context of multiclass prediction.

Imaging data requires more data cleaning work after a feature’s extraction phase because most techniques involve CNNs extracting related and nonrelated features, possibly leading to poor performance.

On the other hand, dealing with genetic data poses more challenges. Genetic data processing is complex and expensive. It is much easier to find imaging data for breast cancer projects. When it comes to genetic data, it is hard to find datasets that are large enough with the right labels.

We cannot conclude that gene expression will give better prediction and more accurate results, based on the above comparison we can state that each datatype contains drawbacks and challenges.

Compliance with Ethical Standards:

The authors thank the University of Sharjah for supporting this work. Funding is through the competitive project “Applications of Machine Learning in Metastatic Breast Cancer Detection” Conflict of Interest: The authors declare that they have no conflict of interest. Informed Consent: This study does not involve any experiments on animals.

References:

[1] Zhou X, Li C, Rahaman MM, Yao Y, Ai S, Sun C, et al. A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks. IEEE Access 2020;8:90931–56. https://doi.org/10.1109/ACCESS.2020.2993788.

[2] Sutanto DH, Ghani MKA. A BENCHMARK OF CLASSIFICATION FRAMEWORK FOR NON-COMMUNICABLE DISEASE PREDICTION : A REVIEW 2015.

[3] Gautam R, Kaur P, Sharma M. A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings. Prog Artif Intell 2019;8:401–24. <https://doi.org/10.1007/s13748-019-00191-1>.

[4] Mahmood M, Al-Khateeb B, Alwash WM. A review on neural networks approach on classifying cancers. IAES Int J Artif Intell 2020;9:317–26. https://doi.org/10.11591/ijai.v9.i2.pp317-326

[5] Fatima N, Liu L, Hong S, Ahmed H. Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis. IEEE Access 2020;8:150360–76. https://doi.org/10.1109/ACCESS.2020.3016715

[6] Wisesty UN, Mengko T, Purwarianti A. Gene mutation detection for breast cancer disease: A review. IOP Conf Ser Mater Sci Eng 2020;830:32051. <https://doi.org/10.1088/1757-899X/830/3/032051>.

[7] Pang T, Wong JHD, Ng WL, Chan CS. Deep learning radiomics in breast cancer with different modalities: Overview and future. Expert Syst Appl 2020;158:113501. https://doi.org/https://doi.org/10.1016/j.eswa.2020.113501

[8] Punitha S, Amuthan A, Joseph KS. Enhanced Monarchy Butterfly Optimization Technique for effective breast cancer diagnosis. J Med Syst 2019;43:206. https://doi.org/10.1007/s10916-019-1348-8.

[9] Umamaheswari TS, Sumathi P. Enhanced firefly algorithm (EFA) based gene selection and adaptive neuro neutrosophic inference system (ANNIS) prediction model for detection of circulating tumor cells (CTCs) in breast cancer analysis. Cluster Comput 2019;22:14035–47. https://doi.org/10.1007/s10586-018-2183-2.

[10] Chen R, Yang L, Goodison S, Sun Y. Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data. Bioinformatics 2019;36:1476–83. <https://doi.org/10.1093/bioinformatics/btz769>

[11] Dhannoon BN. Predication and Classification of Cancer Using Sequence Alignment and Back Propagation Algorithms in Brca1 and Brca2 Genes. Int J Pharm Res 2019;11. <https://doi.org/10.31838/ijpr/2019.11.01.062>

[12] Mostavi M, Chiu Y-C, Huang Y, Chen Y. Convolutional neural network models for cancer type prediction based on gene expression. BMC Med Genomics 2020;13. <https://doi.org/10.1186/s12920-020-0677-2>.

[13] Liu Q, Hu P. Association Analysis of Deep Genomic Features Extracted by Denoising Autoencoders in Breast Cancer. Cancers (Basel)

[14] Chen H, Gao M, Zhang Y, Liang W, Zou X. Attention-Based Multi-NMF Deep Neural Network with Multimodality Data for Breast Cancer Prognosis Model. Biomed Res Int 2019;2019:9523719. <https://doi.org/10.1155/2019/9523719>.

[15] Jiang P, Huang S, Fu Z, Sun Z, Lakowski TM, Hu P. Deep graph embedding for prioritizing synergistic anticancer drug combinations. Comput Struct Biotechnol J 2020;18:427–38. <https://doi.org/https://doi.org/10.1016/j.csbj.2020.02.006>.

[16] Danaee P, Ghaeini R, Hendrix DA. A Deep Learning Approach for Cancer Detection and Relevant Gene Identification. Biocomput. 2017, n.d., p. 219–29. <https://doi.org/10.1142/9789813207813_0022>.

[17] Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. Cytometry A 2017;91:566–73. <https://doi.org/10.1002/cyto.a.23065>.

[18] El-Nabawy A, El-Bendary N, Belal NA. A feature-fusion framework of clinical, genomics, and histopathological data for METABRIC breast cancer subtype classification. Appl Soft Comput 2020;91:106238. <https://doi.org/https://doi.org/10.1016/j.asoc.2020.106238>.

[19] Karakış R, Tez M, Kılıç YA, Kuru Y, Güler İ. A genetic algorithm model based on artificial neural network for prediction of the axillary lymph node status in breastcancer. Eng Appl Artif Intell 2013;26:945–50. <https://doi.org/https://doi.org/10.1016/j.engappai.2012.10.013>.

[20] Zhang F, Chen J, Wang M, Drabier R. A neural network approach to multi-biomarker panel discovery by high-throughput plasma proteomics profiling of breast cancer. BMC Proc 2013;7:S10–S10. <https://doi.org/10.1186/1753-6561-7-S7-S10>.

[21] Alsaleem MA, Ball G, Toss MS, Raafat S, Aleskandarany M, Joseph C, et al. A novel prognostic two-gene signature for triple negative breast cancer. Mod Pathol 2020;33:2208–20. <https://doi.org/10.1038/s41379-020-0563-7>.

[22] Lancashire LJ, Powe DG, Reis-Filho JS, Rakha E, Lemetre C, Weigelt B, et al. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. Breast Cancer Res Treat 2010;120:83–93. https://doi.org/10.1007/s10549-009-0378-1.2019;11. <https://doi.org/10.3390/cancers11040494>

RESULT:

Advanced breast cancer prediction with deep learning:

Cancer is one of the most dangerous diseases to humans, and yet no permanent cure has  
been developed for it. Breast cancer is one of the most common cancer types. According to  
the National Breast Cancer Foundation, in 2020 alone, more than 276,000 new cases of  
invasive breast cancer and more than 48,000 non-invasive cases were diagnosed in the US.  
To put these figures in perspective, 64% of these cases are diagnosed early in the disease's  
cycle, giving patients a 99% chance of survival. Artificial intelligence and machine learning.

Github link:

https://github.com/vchandru274/vchandru3