Comparative Analysis of Decision Trees and Neural Networks for Breast Cancer Detection in Medical Imaging Data

Harsheta Sharma

Sheridan College

MATH22500

Professor Farah Hussein

April 7th, 2025

Video Presentation Accompanying Report:

Comparative Analysis of Decision Trees and Neural Networks for Breast Cancer Detection in Medical Imaging Data Video - Harsheta Sharma

Introduction

Before we can describe the importance of detecting cancer, one must understand what cancer is. Cancer is a set of diseases which have affected multicellular living beings for more than 200 million years (Hausman, 2019). Cancer is different from other diseases in that the root causes of problems are not external but rather our own cells which have transformed into the building blocks of tumors. There are two types of cancers: (1) pre-cancers and (2) cancers, where one is less harmful and aggressive than the other. Cancer mainly involves genes and mutations within those genes. A growing cancer will redirect the processes and resources needed to coordinate healthy cellular growth toward tumors.

Cancer as a disease can have many impacts on an individual's life. It can affect an individual socially and economically in terms of how that individual can work and make money. It obviously affects the individual's health, and all these factors contribute to one's mental health as well.

Living with cancer can be a tough experience due to the amount of social isolation, as friends and family may not know how to support a person with cancer. Additionally, the constant hospital visits and treatments can make an individual feel drained.

A study looking at children and adolescents' perspectives on living with advanced cancer concluded that suffering and emotional distress were significant components of their lives. Many of these adolescents worked very hard to re-establish a sense of normalcy in their daily lives since cancer had affected them far beyond their health. Losing the ability to do the hobbies and activities they once enjoyed made the adolescents feel isolated. As one adolescent stated, "How much I would like to do everything easily, be comfortable like my peers around me now. How I wish I could be like a normal young girl without any difference from them… I miss my life so much" (Johnson, Ali, An, et al., 2024). Many of the adolescents described experiencing a deep sadness at times and feeling the need to emotionally escape or distract themselves (Johnson et al., 2024).

These effects magnify and translate into adult patients living with cancer, as their quality of life is significantly affected. A study researching the quality of life and psychological distress of patients with advanced cancer in the Philippines found that rates of anxiety and depression among cancer patients ranged from 22% to 58% (Bacorro, Que, Sy Ortin, et al., 2015, p. 98). Physical concerns such as fatigue, general pain, weight loss, and shortness of breath are also present during and after treatment(Manalo, Ng, Ozdemir, Malhotra, Finkelstein, Ong, & Teo, 2023).

Cancer also affects individuals economically due to the drastic lifestyle changes required following diagnosis. Many individuals express economic distress as they live with an imbalance between daily income, hospital bills, and other expenses (Meng, 2022). Additionally, their ability to survive becomes directly correlated to costs, as poor physical health and medical bills lead many to deprive their families of living expenses to receive treatment (Meng, 2022).

Fortunately with emerging research, almost 66% of adults survive more than five years after their diagnosis. However, long-term effects after treatments such as nausea, neuropathy, anxiety, depression, and reduced cognitive and physical capabilities linger on for a while (Bentley, Teckle, McQuarrie, et al., 2022).

A lot of research has been done on the impacts of cancer on income and economic outcomes for adult cancer survivors. Many report significant income loss due to cancer. A study done in 2018 for all cancers found that the annual income for such individuals was reduced by 65% in the first five years post-diagnosis for women and men combined (Bentley et al., 2022). Additionally, cancer treatments often lead to long-term physical and mental health issues. This includes bone fragility, effects on bone cells and metabolism, and a weakened immune system (Teissonnière, Point, Biver, et al., 2025).

Transitioning into breast cancer, “Breast cancer is the most common global malignancy and the lead cause of cancer deaths” (Katsura, Ogunmwonyi, Kankam, & Saha, 2022). Breast cancer comprises a range of malignancies which occur in the mammary gland. This cancer primarily affects women, and the chances of getting it increase with age. Breast cancer may present as a hard, immobile lump, swelling, or skin changes, among other symptoms (Katsura et al., 2022). Although diagnostic tests such as the triple assessment test exist, breast cancer continues to be a leading cause of cancer-related deaths, with 685,000 deaths worldwide in 2020 (Katsura et al., 2022). Life with breast cancer has many side effects similar to those of other cancers. However, additional issues include changes in body image due to hair loss and skin changes, as well as menopausal symptoms and infertility (Cohen, Anderson, Jensik, Xiang, Pruszynski, & Walker, 2012).

From the above analysis of cancer, cancer treatment, and specifically breast cancer, one can gain insight into the importance of detecting cancer. Detecting cancer is vital because it allows doctors to recognize signs of cancer and intervene early which increases the chance of successful intervention and treatment. Early detection can help prevent individuals from suffering full effects of both the disease and its treatment which is why advancements in breast cancer detection represent significant, life-saving milestones.

The use of machine learning algorithms for detecting breast cancer using medical image data has become a relevant topic in recent years. With advancements in AI expanding into medicine, machine learning algorithms such as decision trees and neural networks have proven to be excellent solutions for detecting breast cancers.

This report aims to explore the use of these two machine learning algorithms—Decision Trees and Neural Networks—to detect breast cancer in medical imaging data.

Decision trees are non-parametric supervised machine learning algorithms which structure data like a tree and are commonly used for classification and regression. They are a good choice for this problem since they allow for easy interpretability of results, flexibility with data types and relationships, and speedy computation. Efficient computation and easy-to-interpret results are vital in a clinical setting where accuracy and speed are priorities. Decision trees allow us to trace the algorithm's path from root to leaf to determine which features led to the classification. This transparency allows us to improve and verify the model's accuracy. Decision trees can capture complex and non-linear relationships between different features, enhancing the accuracy of breast cancer classification by properly representing the underlying data (Ghiasi & Zendehboudi, 2021).

Versions of decision trees such as Random Forest and J48 are known to be computationally efficient which allows for very quick prediction which is especially important in clinical settings where timely and accurate decisions are needed(Al-Salihy & Ibrikci, 2017) . Multiple studies have also been done on looking into the effectiveness of using decision trees for detecting breast cancer where one found that ". The decision tree classification forecasts breast tumours with lower error average and higher precision of correctly classified cases 97.7%." (Al-Salihy & Ibrikci, 2017). Additionally, flexibility with data types and the ability to represent different types of relationships makes decision trees versatile and well suited for this problem.

Neural networks are deep learning models used to recognise patterns within data. Inspired by the way humans process information using neurons, they are made up of layers of tiny computing units also called neurons which work together to process information. Specific neural networks called Convolutional Neural Networks (CNNs) are excellent choices for this problem as they can learn intricate patterns in image data by extracting features we provide to the neural network such as cell shape(Currie, Hawk, Rohren, Vial, & Klein, 2019). The sheer accuracy and ability to constantly learn makes neural networks excellent choices for this problem. Additionally, the neural networks ability to process large amounts of information at once by expanding on previous learning adds to the accuracy and speed. CNNs specifically are excellent at automatic feature extraction from medical images which eliminates the need for extensive feature engineering by medical experts(Al Tawil, Shaban, & Almazaydeh, 2024). Moreover, these neural networks are able to generalise meaning they can process different datasets and adapt to variations; excellent for real world clinical applications such as the detection of breast cancer(Al Tawil, Shaban, & Almazaydeh, 2024). Lastly, many previous studies have shown promising results of breast cancer using CNN's. One study presented a CNN that effectively automates the detection of breast cancer with a remarkable accuracy of prediction of 99.86%.(Dabeer, Khan, & Islam, 2019).

For this experiment, $\mu 1$ represents the mean accuracy for decision trees while $\mu 2$ represents the mean accuracy for neural networks. A difference between the two would indicate a difference in the performance of the two machine learning algorithms

Null Hypothesis: the difference between $\mu 1$ and $\mu 2$ is not statistically different.

$$\mu 1 = \mu 2$$

Alternative Null Hypothesis: There is a statistical difference between $\mu 1$ and $\mu 2$.

$$\mu 1 \neq \mu 2$$

There will likely be a difference in the performance between the decision tree and the neural network due to a couple of factors making the difference between the two statistically different, which is the claim (The alternate null hypothesis). Firstly, neural networks can learn complex non linear relationships between features thus allowing for better accuracy while decision trees are limited in their representation power since the data is split based on simple decisions rules. These decision rules may restrict reducing performance for decision trees. Additionally, neural networks will probably do a better job at generalizing the predictions since they pick out their own features. Plus, since decision trees are supervised learning, the combinations in features will not automatically combine unless we reflect that in the splits which take place. Neural networks on the other hand do this automatically allowing for quicker learning and thus more accuracy. For those reasons, I am claiming that the performance between the two will be statistically different.

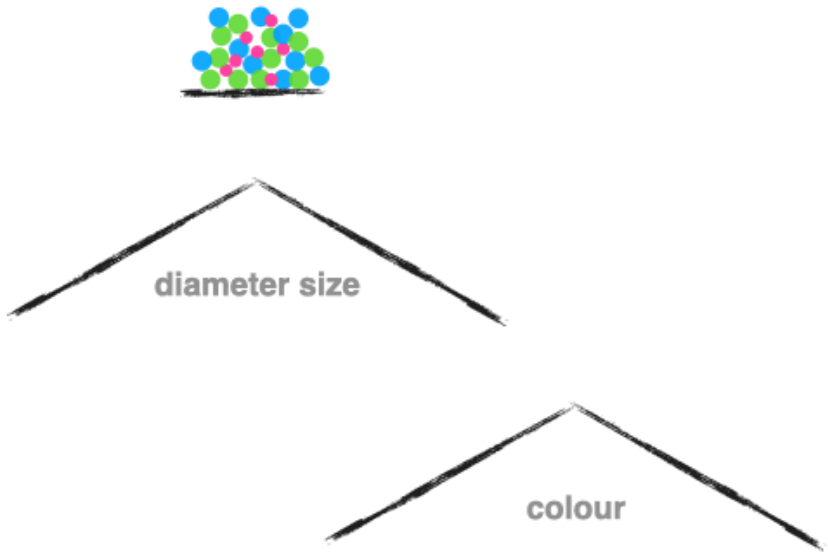
Literature Review

Decision trees are a type of non- parametric supervised machine learning algorithm which is used for classification and also regression(Vashist, Sagar, & Goyal, 2023). By structuring data into a tree-like structure, they consist of nodes, edges and leaves. The nodes represent the features or attributes, the edges represent the decision rules while the leaves represent the outcomes. Using the nodes/feature values and recursion, the algorithm will split the data to create branches (Song & Lu, 2015). The root node in a decision tree has the entire dataset and is where the splitting begins. The internal nodes will then correspond to a decision rule which leads to more splits until a leaf node is reached which has the final output.

Figure 1 below can be used to visualise a simple decision tree for classifying colored balls.

Figure 1

Basic decision tree sorting colored balls using diameter and color decision rules (ML 2 GIFs, 2021)



In this example, the diameter size and color are decision rules and the data is split based on the decision rules leading to a classified output where all the balls have been sorted with their respective colors.

The performance of decision trees is evaluated using metrics such as accuracy , precision, specificity and F1 score and values which are derived from a confusion matrix are used to analyse their performance. A confusion matrix is just a method to visualise the results from a classification algorithm (Murel, 2024).

When constructing decision trees, it is vital to select the most informative features to split on because this allows us to ensure our results are properly classified. Some common algorithms for building decision trees include CART (classification and Regression Trees) , C4.5 and QUEST(Quick, Unbiased, Efficient Statistics Tree) (Song & Lu, 2015). Multiple decision trees such as the Random Forest which uses the CART algorithm and the J48 which is the Java implementation of the C 4.5 Algorithm are commonly used since they are known to be computationally efficient..

A decision tree can be used for cancer detection by looking into the clinical characteristics specifically classifying the lesions or the patients into different risk categories (Sripodok et al., 2024). One study developed a clinical diagnostic guide using a decision tree model which was used to predict localized Gingival Enlargement which is a condition where there is abnormal growth or swelling of the gingival tissue (our gums) in our mouths. The decision tree model aimed to find the difference between malignant and non-malignant LGE's through the recursive decision trees process. In this case, they mainly looked at the lesion size with the cutoff of 3cm and also looked at consistency and color thereafter(Sripodok et al., 2024). With breast cancer, the decision tree model would look at specific factors such as clump thickness, homogeneity of cell size, marginal adhesion or homogeneity of the cell shape as examples. The exact factors used would depend on the dataset being used(Vashist, Sagar, & Goyal, 2023). Usually a study will examine many phases of breast cancer and look at multiple datasets to

figure out an optimal algorithm for detection. But in general, tissue augmentation, shape abnormalities, roughness, and radius are important factors which would be considered(Vashist, Sagar, & Goyal, 2023).

Deep learning is a subset of machine learning that uses these neural networks. Neural networks, similar to a human brain's information processing process, are made of layers of artificial neurons which work together to pick important features from data(Yousif et al., 2022). Features are a property or character of the data we are looking at such as the color. This deep learning model has three layers including the input layer which is where the raw data is received, one or more hidden layers which is where the feature transformations and computations for patterns take place and an output layer which has the final classification.

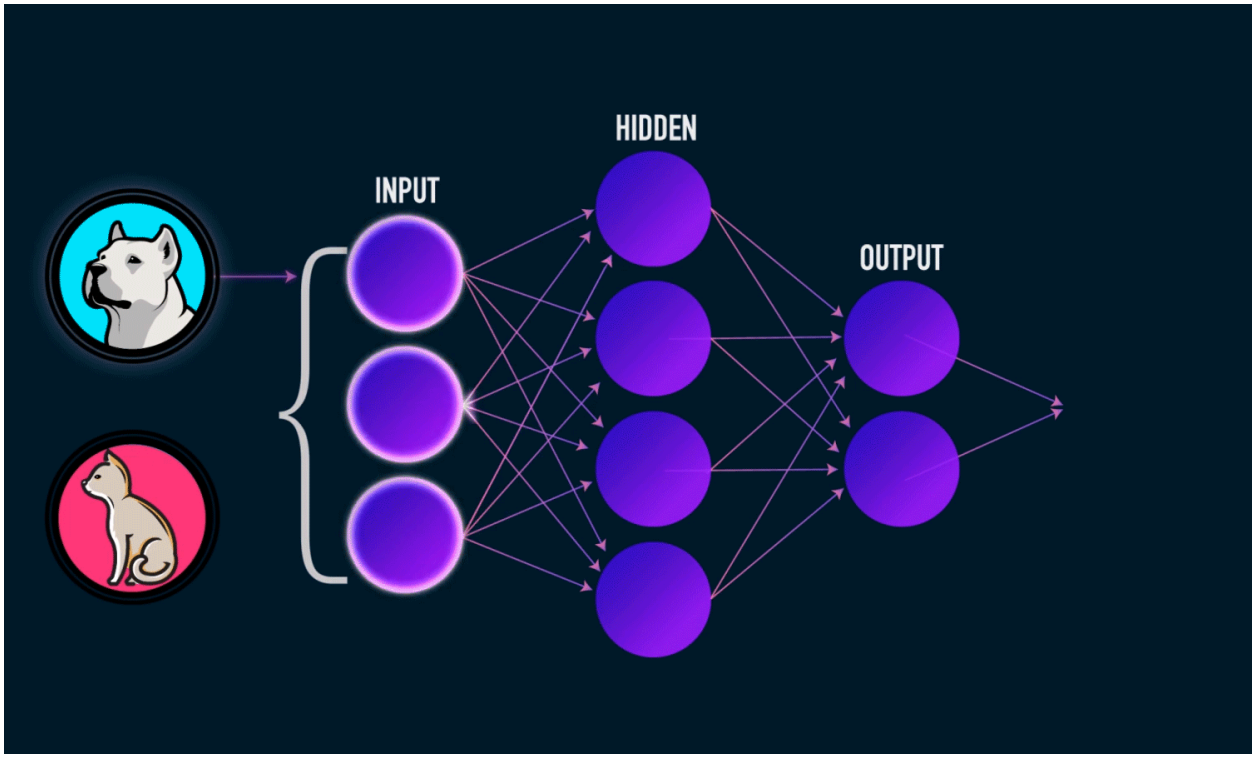
Through a process called forward propagation, the input layer data gets transferred across the layers using weights and activation functions until an output is classified. These neural networks will learn from generating hidden patterns from features from the data within these layers. The parameters used for the models can be changed and optimised by comparing the model's predictions to real values and updating the parameters accordingly to ensure accuracy allowing the model to learn. This process is known as backpropagation and uses functions like the mean squared error(Heenaye-Mamode Khan et al., 2021).

In a neural network, first data is fed into the input layer and each of these inputs will match a feature such as age, pixel value etc.

The input layer is connected to the hidden layer through channels as seen by the arrows in the image below. Each of these channels has a weight which is a numerical value. The inputs are multiplied to these weights and the sum of these is sent as sum to the neurons of the next layer. In the hidden layer, each neuron has a value called a bias which is added to the previous input sum. This total sum is passed through a function called an activation function. The activation function indicates if the neuron will get activated which means it will transfer data to the next channel. This is the essence of forward propagation. In the output layer, the neuron with the highest value determines the output(IBM, n.d.).

Figure 2

Example of neural network with input, hidden and output layer classifying animals (Priyal, 2023)



Specifically Convolutional neural networks (CNNs) are designed for image data and have shown excellent accuracy in identifying breast cancer from medical data. A study found that a deep learning model called BCCNN was able to classify breast cancers into different classes showing excellent precision within classification(Abunasser et al., 2023). CNNs have convolutional layers which automatically extract hierarchical features from the images such as textures or abstract patterns which allows for a very accurate analysis of medical scans/images.

In general, neural networks for cancer detection require a large dataset with images labeled as cancerous or absence of cancerous tissue or maybe with a specific type of cancer. The neural network model will then get trained as described, adjusting its internal parameters(the weights) to ensure correct classification(Currie, Hawk, Rohren, Vial, & Klein, 2019).Once the model is trained, it will be used to analyze new images and provide a prediction about the likelihood of cancer. This method is especially useful for detecting breast cancer as ANN's and CNNs mentioned earlier have shown great accuracy in mammography interpretation. Specifically, CNNs have a unique capacity to automatically extract complex patterns, features and pertinent information from very sophisticated medical images as mentioned which allows for quicker learning, enhanced accuracy and less workload for the medical experts. Different studies have attempted to compare the performance of different CNN architectures for breast cancer prediction. One study used the VGG19, AlexNet and ResNet50 to compare their performances and found that ResNet50 was the most accurate , achieving an accuracy of 93.27% while AlexNet was the most computationally efficient(Al Tawil, Shaban, & Almazaydeh, 2024)..ANNs on the other hand rely on manual extracted features from medical images or clinical data e.g. patient history.

These manual features need to be designed by experts based on what is known and relevant to cancer (Ayer, Chen, & Burnside, 2013). After training using ANN or CNNs the neural network is ready to analyse new unseen data based on what it has learned. Many factors can be looked at for breast cancer detection. ANNs mainly use patient demographic risk factors, mammographic findings e.g. edge gradient or construct, BI- RADS descriptions which are standardized lexicon of Breast Imaging Reporting and Data System which have descriptors to look for (Ayer, Chen, & Burnside, 2013). CNNs on the other hand will mainly use raw pixel data of medical images since they are able to learn to identify shapes, boundaries and extract complicated patterns and features themselves (Ayer, Chen, & Burnside, 2013). CNNs are able to perform object detection(locating potential tu,ours), object segmentation and object classification which all helps in the detection of cancer (Currie, Hawk, Rohren, Vial, & Klein, 2019).

Methods

This experiment was conducted using the Weka machine learning platform specifically using its breast cancer dataset which consists of images from patients diagnosed with cancer and healthy individuals. This dataset contains 286 patient records with a mixture of nominal and categorical attributes, such as age, menopause status, tumor size, and breast density. The dataset was divided into two sets, a training set which consisted of 66% of the data and a testing set which consisted of 34% of the data. Both models were run 30 times to collect a sample of 30 values and the model accuracy was collected for each element in the sample.

For the decision tree model, a J48 decision tree which is Weka's version of the C4.5 decision tree was used. The split was made as aforementioned with 66% of the data for training and 34% for testing. An excel table was made for this sample consisting of Trail Number, Correctly Classified Instances, Correctly Classified Instances in Percentage, Incorrectly Classified Instances and Incorrectly Classified Instances in Percentage columns were made. Then the first run was conducted and the data was collected for the columns in excel. This was continued for a total of 30 runs ensuring the Random seed was changed each time to ensure we got a different subset of the data. Failing to change the Random seed would produce the same result since the exact same subset of data was used.

For the neural network, the MultilayerPerceptron which is an ANN was used and found under functions on Weka. The exact same excel table was set up as for the decision tree. The same split was done on the entire dataset where 66% was for training and 35% was for testing. The first run was performed and the data was collected into the excel table. Like the decision tree, a total of 30 runs were performed for this neural network ensuring the Random seed was changed each run.

Once all the data was collected from the experiment, the analysis could be done on excel.

The data given by Weka for the two algorithms were medical images of patients with cancer and healthy individuals. Once the experiment was performed, we collected our own data on the accuracy of the model for each trail. The columns on the excel are the trail number which is the element number of the sample. We also have the Correctly Classified Instances column which has the number of images which were correctly classified from the subset. The Correctly Classified Instances in Percentage provides the accuracy of the model for each run since it is the ratio of correctly classified images over incorrectly classified images * 100. The

Incorrectly Classified Instances column has the number of images incorrectly classified while the Incorrectly Classified Instances in Percentage column just presents that as a ratio. From these columns, our analysis was performed on the Correctly Classified Instances in Percentage since our study aimed to look at the accuracy of each model. The same columns were used for both models to ensure consistency for comparison. The dataset collected has 30 rows for each model representing the 30 trails with the 4 columns as aforementioned.

The decision tree dataset was produced by running the J48 - Weka's version of the C4.5 Decision Tree model on the breast cancer dataset provided by Weka containing medical images of cancer diagnosed and healthy individuals. The neural network dataset was produced by running the MultilayerPerceptron neural network on that same breast cancer dataset. 30 runs were performed for each model to collect 30 elements in the sample using Random seed splits for each run.

This data is relevant since it allows us to test the accuracy of both models on detecting breast cancer. The data we used for this experiment is relevant since it is medical images of patients with and without a diagnosis. The data collected is relevant since it measures the accuracy of the model for each run allowing us to compare the performance of each model.

The column *Correctly Classified Instances in Percentage (Accuracy)* was a representation of the following formula automatically applied and included in each Weka result:

Accuracy $= (Correctly\ Classified\ Instances \div Total\ Instances) * 100$

The accuracy of each model was a representation of the number of correctly classified images over the total number of images for each subset in each run. This accuracy, represented as a percentage, indicated the effectiveness and performance of each model during each run. The higher the accuracy, the better for each run.

Traditionally, accuracy would be defined using the four outcomes of a confusion matrix which are True Positive, False Positive, True Negative and False Negative in the formula below:

Figure 3

General Formula for Accuracy from the Google Developers Crash Course on ML (Google, n.d.)

$$\text{Accuracy} = \frac{\text{correct classifications}}{\text{total classifications}} = \frac{TP + TN}{TP + TN + FP + FN}$$

A perfect model would mean zero false positives and zero false negatives providing 100% accuracy. But since our data already provides an accuracy column consisting of the correctly classified instances over the total instances, the use of these four outcomes is not needed.

Recall is the true positive rate since it is the proportion of true positives that were correctly classified as positives. In our case, this would be the proportion of correctly classified images/ actual number of images with a cancer diagnosis (all actual positives). This can be represented as a formula below

Recall = (Number of correctly identified cancer-positive images) / (Total number of actual cancer-positive images)

Traditionally, recall is represented using the formula below:

Figure 4

Recall formula sourced from the Google Developers Crash Course on ML (Google, n.d.)

$$\text{Recall (or TPR)} = \frac{\text{correctly classified actual positives}}{\text{all actual positives}} = \frac{TP}{TP + FN}$$

But since our data does not provide us the number of actual positives, we are unable to compute a recall.

Precision is the measure of how many positive predictions were truely correct. For example, the model predicted 10 positives, the precision would be a measure of how many of those 10 are actually positives. This is different from recall which measures the model's ability to find all positive cases compared to precision which checks how many of those positives are correct.

Precision can be represented using the formula below in our case:

Precision = (Number of correctly identified cancer -positive images) / (Total number of images model predicted to have cancer)

Figure 5

Precision formula for the Google Developers Crash Course on ML (Google, n.d.)

$$\text{Precision} = \frac{\text{correctly classified actual positives}}{\text{everything classified as positive}} = \frac{TP}{TP + FP}$$

But since we do not have access to the true positives and other outcomes of the confusion matrix, we are unable to compute the precision.

To test our hypothesis with the gathered data, a t - test was performed. Our hypothesis looked to test if there was a difference between the performance of the two models. The data for both samples was independent. This is because even though the underlying data used to train both models was the same (Breast Cancer Dataset from Weka), the percentage split which was done - 66% for training and 34% for testing, divided the dataset to make it unequal. Then, the Random seed for each trail created a subset of the data for each trail leading to even more change between the data. These two transformations made the data used for both models quite different thus independent. Thus, to test our hypothesis, we needed to use

a method involving independent samples. Next, sigma or the population standard deviation was unknown. Since we just ran the models on the dataset provided and collected samples, we had no access to sigma. Since sigma was unknown, we knew the Z distribution could not be used since its calculation required sigma leaving us to use the T distribution using the t-test.

Using the t- test, we had two options to calculate the p value for the independent samples. We either use the pooled or unpooled method which are both tests present on excel. Using the pooled variance requires us to assume that sigma 1 and sigma 2 which are the standard deviations of sample 1 and 2 are equal. Unpooled method does not require them to be equal. Since we had no information on sigma, we needed to conduct an F test. The F test was used to look at the ratio of the two sample variances using the formula below:

$$F_{STAT} = \frac{S_1^2}{S_2^2} \quad df_1 = n_1 - 1; \quad df_2 = n_2 - 1.$$

From conducting we get an Fstat which is compared against a critical value to determine if we can assume equal or unequal variance for the population. The F test is right skewed thus finding the F critical requires us to divide alpha or the significance level by 2. The criteria for interpreting the F test was:

If F <F critical and p value is > 0.05 - assume equal variances
If F> F critical and p value is < 0.05 - assume unequal variances

Figure 6

F Test conducted on Excel to verify equal or unequal variances

F-Test Two-Sample for Variances		
	Correctly Classified Instances in Percent	*Correctly Classified Instances in Percentage*
Mean	0.7013745	0.652577333
Variance	0.001535464	0.001503695
Observations	30	30
df	29	29
F	1.021127702	
P(F<=f) one-tail	0.477745857	
F Critical one-tail	1.860811435	

Conducting the F-test using the data analysis tool found that F critical is around 1.86 while the F value is 1.02 making F < F Critical. Additionally, the p - value is around 0.48 which is bigger than 0.05. With all this information, we assumed equal variances.

Knowing we needed to perform the T-test, we needed to use independent samples and assume equal variances, I moved forward with performing the T- test : Two Sample Assuming Equal Variances with a significance level of 0.05 on excel. This translates to the independent samples using pooled variances

method. The significance level was chosen to be 0.05 since it was the default for this test and is the safest alpha to choose when alpha is not given. The hypothesised mean difference was set to 0 since that was being tested in the hypothesis. The independent samples using pooled variances test using the t-distribution were performed allowing us to get a p value to conclude is there significant difference in the performance of the two models.

Results, Analysis, Discussion and Conclusion

Upon performing the T-test: Two Sample Assuming Equal Variances with a significance level 0f 0.05, we produced the table below. This table included the mean of accuracy of both models, their individual variances, the pooled variance, the t-stat and also the P values for one and two tailed tests. Our number of interest , highlighted in pink, is the P(T<=t) two tails since our hypothesis required a two tailed test checking to see if there was a difference in performance.

Figure 7

T test chart conducted on Excel displaying accuracy means, pooled variance and p - value for both machine learning models

t-Test: Two-Sample Assuming Equal Variances using a significance level of 0.05

Statistic	Decision Tree (J48)	Neural Network
Mean	0.7013745	0.652577333
Variance	0.001535464	0.001503695
Observations	30	30
Pooled Variance	0.001519579	
Hypothesized Mean Difference	0	
df	58	
t Stat	4.848177667	
P(T<=t) one-tail	4.85368E-06	
t Critical one-tail	1.671552762	
P(T<=t) two-tail	9.70735E-06	
t Critical two-tail	2.001717484	

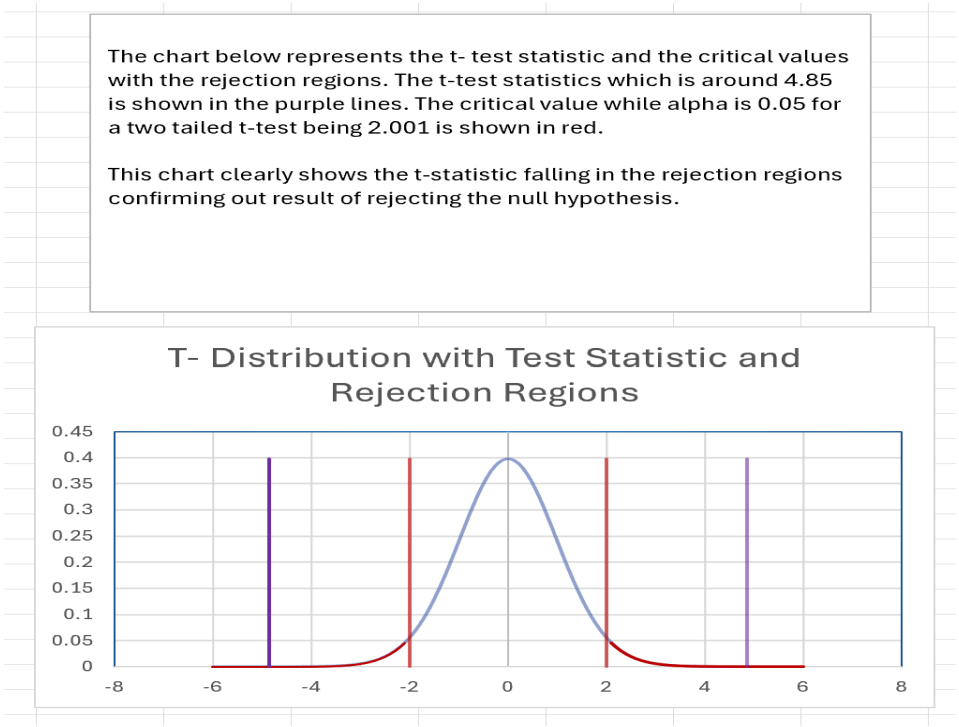
Our criteria for rejecting the null hypothesis was if the p-value was less than 0.05, the null hypothesis would be rejected and we could conclude that there is a significant difference in the performance of the machine learning algorithms. Conducting the t-test gave us a p-value of 9.70735E-06

which is much smaller than 0.05. Since our p-value<0.05, we can safely reject the null hypothesis and conclude that there is a significant difference in the performance of the machine learning algorithms.

The result can also be represented in the form of a distribution below. The following distribution looks at the t values for the test statistics which was 4.85 and the critical value which was 2.001.

Figure 8

Chart Representation of the results from the t - test confirmation rejection of null hypothesis



According to this chart, since our test statistic's t-value (purple lines) falls far into the rejection regions (red lines), we can reject the null hypothesis stating that there is no difference in the performance of decision trees and neural networks for detecting breast cancer.

For this experiment, $\mu 1$ represented the mean accuracy for decision trees while $\mu 2$ represented the mean accuracy for neural networks. A difference between the two would indicate a difference in the performance of the two machine learning algorithms. Our null hypothesis was that $\mu 1$ and $\mu 2$ is not statistically different while our alternative null hypothesis was there is a statistical difference between $\mu 1$ and $\mu 2$. Based on the p-value found and the chart displaying the critical regions and test statistics, we concluded that we should reject the null hypothesis. Rejecting the null hypothesis means supporting the alternative null hypothesis which states that there is a statistical difference between the mean accuracy, an indicator of performance, between the two machine learning models.

Additionally, based on the t-test performed, it can be concluded that decision trees outperform neural networks in detecting breast cancer since the mean accuracy for the decision tree was 0.7013745 while the mean accuracy for the neural network was 0.652577333 making decision trees a more suitable choice for this approach.

This conclusion is supported by other studies performing a comparative analysis of different machine learning methods in the context of breast cancer. One study conducted in 2005 confirmed that looking into breast cancer survivability found " The aggregated results indicated that the decision tree induction method (C5) performed the best with a classification accuracy of 93.6% which is better than any reported in the published literature, the ANN model (with multi layered perceptron architecture) came out to be second best with a classification accuracy of 91.2%..."(Delen, Walker, & Kadam, 2005).

This concluded my comparative analysis of decision trees, specifically the J48 and neural networks, specifically multilayer perceptrons, which is an ANN in detecting breast cancer. My report has found that there is a significant difference in the performance of these two machine learning methods for detecting cancer and concluded that decision trees outperform an ANN in this context. The discussion opens to expand onto this hypothesis by performing a comparative analysis between the J48 decision tree and CNN's to determine if there is a difference in performance and the effectiveness of each in detecting breast cancer.

References

Abunasser, B. S., Al-Hiealy, M. R. J., Zaqout, I. S., & Abu-Naser, S. S. (2023). Convolution Neural Network for Breast Cancer Detection and Classification Using Deep Learning. *Asian Pacific Journal of Cancer Prevention: APJCP, 24*(2), 531–544. https://doi.org/10.31557/APJCP.2023.24.2.531

Al-Salihy, N., & Ibrikci, T. (2017). Classifying breast cancer by using decision tree algorithms. In *Proceedings of the 8th International Conference on Software Technology and Engineering* (pp. 144–148). https://doi.org/10.1145/3056662.3056716

Al Tawil, A., Shaban, A., & Almazaydeh, L. (2024). A comparative analysis of convolutional neural networks for breast cancer prediction (Tech. Rep.). Faculty of Information Technology, Applied Science Private University. https://www.researchgate.net/publication/381071759_A_comparative_analysis_of_convolutional_neural_networks_for_breast_cancer_prediction

Ayer, T., Chen, Q., & Burnside, E. S. (2013). Artificial neural networks in mammography interpretation and diagnostic decision making. *Computational and Mathematical Methods in Medicine, 2013*, 832509. https://doi.org/10.1155/2013/832509

Bacorro, W. R., Que, J. C., Sy Ortin, T. T., & Others. (2015). A cross-sectional analysis of symptom burden among adult cancer patients in a Filipino tertiary care cancer center. *Journal of Clinical Oncology, 33*(29_suppl), 98–98. https://doi.org/10.1200/jco.2015.33.29_suppl.98

Bentley, C., Teckle, P., McQuarrie, L., & Others. (2022). Impact of cancer on income, wealth and economic outcomes of adult cancer survivors: A scoping review. *BMJ Open, 12*, e064714. https://doi.org/10.1136/bmjopen-2022-064714

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Çinar, D., & Çetin, A. (2025). Quality of life in breast cancer patients. In *Managing side effects of breast cancer treatment*. Springer. https://doi.org/10.1007/978-3-031-75480-7_28

Cohen, M., Anderson, R., Jensik, K., Xiang, Q., Pruszynski, J., & Walker, A. (2012). Communication between breast cancer patients and their physicians about breast-related body image issues. *Plastic Surgical Nursing, 32*(3), 101–105. https://doi.org/10.1097/PSN.0b013e3182650994

Currie, G., Hawk, K. E., Rohren, E., Vial, A., & Klein, R. (2019). Machine learning and deep learning in medical imaging: Intelligent imaging. *Journal of Medical Imaging and Radiation Sciences, 50*(4), 477–487. https://doi.org/10.1016/j.jmir.2019.09.005

Dabeer, S., Khan, M. M., & Islam, S. (2019). Cancer diagnosis in histopathological image: CNN-based approach. *Informatics in Medicine Unlocked, 16*, 100231. https://doi.org/10.1016/j.imu.2019.100231

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial Intelligence in Medicine, 34*(2), 113–127. https://doi.org/10.1016/j.artmed.2004.07.002

Gadaleta, E., Thorn, G. J., Ross-Adams, H., Jones, L. J., & Chelala, C. (2022). Field cancerization in breast cancer. *The Journal of Pathology, 257*(5), 561–574. https://doi.org/10.1002/path.5902

Google. (n.d.). Accuracy, precision, and recall. Google Developers. Retrieved April 6, 2025, from https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall

Hausman, D. M. (2019). What is cancer? *Perspectives in Biology and Medicine, 62*(4), 778–784. https://doi.org/10.1353/pbm.2019.0046

Heenaye-Mamode Khan, M., Boodoo-Jahangeer, N., Dullull, W., Nathire, S., Gao, X., Sinha, G. R., & Nagwanshi, K. K. (2021). Multi-class classification of breast cancer abnormalities using deep convolutional neural network (CNN). *PLOS ONE, 16*(8), e0256500. https://doi.org/10.1371/journal.pone.0256500

IBM. (n.d.). Neural networks. IBM. Retrieved April 6, 2025, from https://www.ibm.com/think/topics/neural-networks

Johnson, A., Ali, R., An, E., & Others. (2024). Children's and adolescents' perspectives on living with advanced cancer: A meta-synthesis of qualitative research. *Quality of Life Research, 33*(8), 2095–2106. https://doi.org/10.1007/s11136-024-03688-y

Katsura, C., Ogunmwonyi, I., Kankam, H. K., & Saha, S. (2022). Breast cancer: presentation, investigation and management. *British Journal of Hospital Medicine (London, England : 2005), 83*(2), 1–7. https://doi.org/10.12968/hmed.2021.0459

Manalo, M. F., Ng, S., Ozdemir, S., Malhotra, C., Finkelstein, E. A., Ong, K. D., & Teo, I. (2023). Quality of life and psychological distress of patients with advanced cancer in the Philippines. *Quality of Life Research, 32*(8), 2271–2279. https://doi.org/10.1007/s11136-023-03389-y

Martínez-Vizcaíno, V., Cavero-Redondo, I., Reina-Gutiérrez, S., Gracia-Marco, L., Gil-Cosano, J. J., Bizzozero-Peroni, B., Rodriguez-Artalejo, F., & Ubago-Guisado, E. (2023). Comparative effects of different types of exercise on health-related quality of life during and after active cancer treatment: A systematic review and network meta-analysis. *Journal of Sport and Health Science, 12*(6), 726–738. https://doi.org/10.1016/j.jshs.2023.01.002

Meng, X., An, Z., Xu, Y., & Others. (2022). Living experiences of people with advanced cancer with low socioeconomic status: A systematic review of qualitative evidence. *Palliative Medicine, 37*(4), 444–459. https://doi.org/10.1177/02692163221137106

ML 2 GIFs. (2021, January 3). Classification task using decision trees [GIF]. ML 2 GIFs. https://ml2gifs.substack.com/p/decision-trees

Murel, J., & Kavlakoglu, E. (2024, January 19). What is a confusion matrix? IBM. https://www.ibm.com/think/topics/confusion-matrix

Priyal. (2023, August 17). Neural nets. Medium. https://medium.com/@p._./neural-nets-299072b6e45f

Song, Y. Y., & Lu, Y. (2015). Decision tree methods: Applications for classification and prediction. *Shanghai Archives of Psychiatry, 27*(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

Sripodok, P., Lapthanasupkul, P., Arayapisit, T., Kitkumthorn, N., Srimaneekarn, N., Neeranadpuree, V., Amornwatcharapong, W., Hempornwisarn, S., Amornwikaikul, S., & Rungraungrayabkul, D. (2024). Development of a decision tree model for predicting the malignancy of localized gingival enlargements based on clinical characteristics. *Scientific Reports, 14*(1), 22185. https://doi.org/10.1038/s41598-024-73013-7

Teissonnière, M., Point, M., Biver, E., Hadji, P., Bonnelye, E., Ebeling, P., Kendler, D., Villiers, T., Holzer, G., Body, J.-J., Fuleihan, G., Brandi, M., Rizzoli, R., & Confavreux, C. (2025). Bone effects of anti-cancer treatments in 2024. *Calcified Tissue International, 116*, 54. https://doi.org/10.1007/s00223-025-01362-0

Vashist, A., Sagar, A., & Goyal, A. (2023). Breast cancer detection by using decision tree. In *Proceedings of the International Conference on Intelligent Engineering and Management (ICIEM 2023)* (pp. 73–84). Springer. https://doi.org/10.1007/978-3-031-45121-8_7

Yousif, M., van Diest, P. J., Laurinavicius, A., Rimm, D., van der Laak, J., Madabhushi, A., Schnitt, S., & Pantanowitz, L. (2022). Artificial intelligence applied to breast pathology. *Virchows Archiv: An International Journal of Pathology, 480*(1), 191–209. https://doi.org/10.1007/s00428-021-03213-3