

BA305 A1

May 2023

Breast Cancer Diagnosis Report



Prepared By:

Yilan Ruan, Brian Park, Shunnosuke
Kanematsu

Introduction

Breast cancer is undoubtedly one of the most aggressive cancers in the United States, affecting 1 in 8 women alone.¹ It accounts for 30% of all cancers diagnosed in females, and symptoms of breast cancer may include thickening or swelling in the breast, the presence of a lump in the breast or the underarm, redness and flaky skin near the area, and general pain in any area of the breast.² However, not all breast lumps are cancerous. Breast lumps can be classified as benign or malignant. Benign breast lumps are not cancerous and rarely spread to other body parts, while malignant breast lumps can spread to other body parts and be terminal.

A breast cancer diagnosis can be tricky as oftentimes, many patients with the illness do not experience any symptoms.³ Physicians that suspect breast cancer refer patients for a biopsy to confirm. During this procedure, a surgeon removes small breast tissue lesions for further diagnosis in the lab. Biopsies can be done with needles or open incisions and often range from \$1,929 to \$7,535 without insurance, depending on the physician's recommendation.⁴ Given the expensive treatments and the severity of the illness, it is critical for physicians to be able to make an accurate diagnosis of breast cancer in patients.

As part of our project, we created several classification models with the goal of accurately predicting the number of patients with malignant tumors. The classification models can support physicians with future diagnoses of patients with similar symptoms and characteristics. Python code for this project was inspired by individual Python data analysis knowledge, labs from the class, as well as use of ChatGPT.

Dataset Description

We found our data from Kaggle. It was uploaded by Nancy Alward and initially produced in a case study by the University of Wisconsin. The data comprises medical scans from a group of patients who underwent a fine needle aspirate (FNA) biopsy.

The original unreleased dataset had an arbitrary number of variables depicting the cell nuclei in the scans. A decision tree utilizing logistic regression was conducted to reduce the number of variables to 33 (Appendix 1). This can also explain the high accuracy scores in every model we ran.

Data Cleaning

To clean our data, we dropped the 'Unnamed: 32' column because all of its values are NaN, and we could not confidently decipher the meaning of the variable. We also dropped all of the standard error columns because it is not relevant in determining the predictive power of our models. Total columns dropped:

¹ "Breast Cancer Statistics: How Common Is Breast Cancer?" Breast Cancer Statistics | How Common Is Breast Cancer?, 12 Jan. 2023, www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html.

² Ibid.

³ Ibid.

⁴ "Breast Biopsy Procedures." Mdsave.Com, www.mdsave.com/procedures/breast-biopsy/d78af5cf. Accessed 8 May 2023.

```
'id,' 'radius_se,' 'Texture_se','Perimeter_se','Area_se','Smoothness_se','Compactness_se,'  
'concavity_se,' 'concave points_se','Symmetry_se','fractal_dimension_se' "Unnamed:32"
```

From the presentation, we split our models into 3 separate data sets (mean, worst, mean + worst). For the sake of simplicity, fitting the report into deliverable guidelines, and the scale of this project, we decided to only use the mean + worst in our final deliverable, as it gives us a comprehensive, aggregated view of all of our variables.

Exploratory Data Analysis

Our first step is to see the distribution of patients who were diagnosed as benign vs. malignant. We found that there are a total of 357 benign patients (63%) and 211 (37%) patients diagnosed with malignant breast cancer. There is a slight imbalance in our target variable, and we have decided to adjust the cost benefit matrix with higher consequences for false negatives (where patients who have breast cancer are incorrectly diagnosed as benign). With this split, the naive model predicted a baseline accuracy of 63%.

Our second step is to see the different characteristics of both groups. We analyzed values in the 'radius_mean', 'perimeter_mean', and 'area_mean' columns because physicians give significant importance to lesion size in breast cancer diagnosis.

We analyzed the distribution lesion sizes with 'area_mean' between the two groups of patients; we found that patients with malignant tumors had a lesion with a median area_mean of 932mm, which is 473.6mm larger than the lesions in patients with benign tumors with a reported value of 458.4mm (Appendix 2). Lesions in malignant patients are roughly twice as large as lesions in patients with benign diagnoses. We analyzed the distribution for the 'area_worst' columns, which stands for the most severe estimate of the size of the lesion (Appendix 3 and 4). We found that the median area worst is 1303.0mm for malignant cases while the median area worst is 547.4mm for benign cases. The median perimeter mean is 114.2mm for malignant patients while the median perimeter means for benign patients is 78.18mm (Appendix 5 and 6). The size of the tumor differs greatly between the two groups with a larger perimeter indicating a malignant diagnosis. This also matches the results we found when we ran the same histogram for the more severe estimate with the perimeter-worst variable.

Cost-Benefit Matrix

In order to create the cost-benefit matrix, we first needed to understand the progression of events from when healthcare providers conduct their initial evaluation up to when they may potentially conduct treatment options. This would ensure that we account for as many of the different procedures involved as possible. The process typically starts with an initial evaluation,

which may include a physical exam and medical history review. Based on the patient's symptoms and the results of the initial evaluation, the healthcare provider may recommend further diagnostic tests, including imaging tests such as mammography, ultrasound, or MRI.

If imaging tests suggest the presence of a suspicious breast mass, the healthcare provider may recommend an FNA biopsy to obtain a tissue sample for review, which, as mentioned earlier, costs around \$522.

After the biopsy, the tissue sample is sent to a laboratory for analysis to a pathologist who examines the cells under a microscope. The pathologist looks for any signs of abnormal or cancerous cells and determines whether the cells are malignant or benign. If the FNA biopsy shows no signs of cancer or abnormal cells (benign), the patient is in the clear, with the healthcare provider potentially recommending follow-up imaging tests or regular monitoring to ensure the condition does not progress. However, if the biopsy shows signs of cancer (malignant), the healthcare provider will recommend additional diagnostic tests, such as a core needle biopsy or a surgical biopsy, to confirm the diagnosis and assess the extent of the cancer. These additional diagnostic tests usually come around to cost \$5,946.⁵

After the completion of different biopsies, if cancer is confirmed, the healthcare provider will discuss treatment options with the patient. With the availability of myriads of different treatment options, it was impossible for us to accurately quantify the average costs. Nevertheless, given data on the costs related to breast cancer surgery, chemotherapy, and radiation therapy, we computed the average and determined that treatments cost approximately \$11,783.

With an understanding of typical medical procedures, we began to determine values for our cost-benefit matrix. For true negatives, we decided to give it a benefit value of approximately \$18K. The reason for this was that patients who were predicted to be benign as well as ending up as actually benign wouldn't need to be administered the core needle or surgical biopsies as well as the various treatment options mentioned earlier. These cost-savings yielded a total of \$17,729 which we rounded up to \$18K. For the true positives, we decided to give it a benefit value of about \$12K. Patients in the true positive category would have had their cells being identified as malignant at an early stage, indicating that they wouldn't need to undergo the various treatment options such as surgery, chemotherapy, and radiation therapy. Nevertheless, they would still need to undergo an additional round of biopsies (core needle or surgical), resulting in \$5,946 less of a benefit compared to the true negative patients, hence yielding a total savings of approximately \$12K.

Patients in the false positive category were predicted to have malignant tumors, but actually carried benign tumors. This is clearly an error healthcare providers seek to avoid as it could prove costly not just from a monetary perspective, but from a reputation standpoint as

⁵ "Breast Biopsy Procedures." Mdsave.Com, www.mdsave.com/procedures/breast-biopsy/d78af5cf. Accessed 8 May 2023.

well. False positive patients will unnecessarily be administered the second round of biopsies (core needle/surgical), resulting in costs that could've been avoided if the patient was correctly classified. We delegated a cost of about \$6K for the \$5,946 that hospitals would incur when administering the core needle or surgical biopsy. In terms of the false negative patients, this is also an error that healthcare providers would not want to incur. This error is much worse than the false positive error. Not only is treatment of cancer more expensive, but it also puts the patient at a greater risk. False negative patients will initially not receive the second round of biopsies because of the FNA biopsy classifying them as benign. However, as they incur more symptoms and check in with healthcare providers, it will be clear that they do indeed have cancerous cells. They will need to undergo the second round of biopsies as well as immediate treatment. For this reason, we delegated a total cost of \$18K which could've been avoided, had the tumor been correctly classified from the start. That being said, the \$18K is probably the least accurate of the other estimates we have provided because it is impossible to quantify the costs affiliated with emergency treatments such as those administered to patients who were falsely classified as benign. The cost-benefit matrix that we created indicates that it is much more important for healthcare providers to lower the number of false negatives they classify rather than false positives (Appendix 7).

KNN

The first classification algorithm we created was the K-Nearest Neighbors model. Creating our KNN model was relatively simple, however, we needed to standardize our training data due to the fact that our KNN model would be computing Euclidean distances. After determining the best k-value, the model for all variables which had an accuracy of 96.5%.

After optimizing for k using a for loop, we noticed that the k-value with the highest accuracy and net benefit was a k-value of 19 (Appendix 8). A caveat to our model having a high k-value is that it wouldn't be able to capture the local structure within the data. That being said, it makes the tradeoff of smoothing out the data, and reducing the level of noise. Another caveat to the model is its lack of interpretability. Other than taking the majority vote of its nearest neighbors, there is essentially no statistical techniques used in making the classifications. In one of the few cases where a relatively small dataset proved advantageous to us, KNN was much easier to run and capture any relationships within the data.

From the 60/40 train/test split, the model provided a total of 228 predictions, out of which 141 correctly predicted malignant tumors (true positives), and 77 correctly predicted benign tumors (true negatives). The model also produced 8 false negative predictions, implying misclassification of malignant tumors as benign, and 2 false positive predictions, representing benign tumors misclassified as malignant (Appendix 9). As previously mentioned, reducing false negative predictions is crucial since they signify the possibility of missing malignant tumors that may lead to insufficient or delayed treatment. Conversely, false positive predictions may lead to

unnecessary medical interventions and psychological distress for patients. Despite that, our model's exceptional performance was evident since it had only 2 false positive classifications, approaching 0, which is quite remarkable. Overall, the model performed reasonably well, with a high number of true positives and true negatives, but further optimization could be done to minimize false negative and false positive rates. Based on our cost-benefit analysis, the net benefit value was \$3,366,000.

To conduct a thorough evaluation of the model, we employed several evaluation metrics in this study, including precision, true positive rate (TPR), false positive rate (FPR), F1 score, and Matthews Correlation Coefficient (MCC). The precision score of 0.99 represents near-perfect predictive power in detecting malignancy. Moreover, the TPR of 0.85 indicates that the model correctly identified 94% of all actual malignant cases, a crucial metric for medical applications. The FPR of 0.007 signifies that the model misclassified only a marginal proportion of benign cases as malignant, demonstrating near-perfect specificity. The F1 score of 0.91 suggests an excellent balance between precision and TPR. Lastly, the MCC of 0.87 indicates a strong correlation between the benign and malignant cases' predictions, suggesting that both were predicted well.

Overall, our KNN model exhibited exceptional proficiency in accurately categorizing the breast cancer data as either malignant or benign, as it achieved near-perfect scores across a range of performance metrics.

Naive Bayes

Naive Bayes (NB) was another classification algorithm that we utilized to create an accurate predictive model. The reason we imported GaussianNB rather than MultinomialNB was that our data's inputs were all numerical. Of the 3 GNB models that we created, the model on only the worst attributes of the cell nuclei returned an accuracy score of 92.5%. Compared to our KNN model, our GNB model was a lot easier and quicker to make because all we needed to do was plug and play the code from the scikit-learn package.

With regards to the confusion matrix, our model made a total of 228 predictions, of which 138 were true positives (correctly predicted malignant tumors) and 76 were true negatives (correctly predicted benign tumors). However, the model also made 6 false negative predictions (misclassified malignant tumors as benign) and 8 false positive predictions (misclassified benign tumors as malignant) (Appendix 10). As mentioned earlier, the false negatives are particularly important to reduce as they represent cases where a malignant tumor was missed, potentially leading to delayed or inadequate treatment, while false positives can lead to unnecessary medical interventions and stress for patients. Overall, the model appears to perform reasonably well with a high number of true positives and true negatives, but could be further optimized to reduce the false negative and false positive rates. Based on our cost-benefit matrix, the net benefit value was \$3,240,000.

In order to further evaluate the model, we computed various evaluation metrics. The model evaluation metrics used in this study included precision, true positive rate (TPR), false positive rate (FPR), F1 score, and Matthews Correlation Coefficient (MCC). The precision score of 0.9 indicates a strong predictive power in detecting malignancy as 90% of the models predictions that were malignant were correct. Additionally, the TPR of 0.93 indicates that out of all the actual malignant cases the data had, the model correctly identified 93% of the malignant cases, which is a crucial metric for medical applications. The FPR of 0.05 suggests that the model misclassified only a small proportion of benign cases as malignant, indicating that the model is highly specific. The F1 score of 0.92 indicates a good balance between precision and TPR. Finally, the MCC of 0.87 suggests that both the benign and malignant cases were predicted well via the strong correlation.

Overall, our GNB model performed well in accurately classifying the data as malignant or benign, achieving high scores across various evaluation metrics.

Logistic regression

Logistic Regression (LR) was another classification algorithm that we utilized to create an accurate predictive model. In order to create the model, we imported the LogisticRegression package from scikit-learn's Linear Model module. This model returned an accuracy score of 93.9%. . Creating our LR model had a somewhat similar process to how we created our KNN model because of how we added the step of determining the optimal cutoff value for the highest possible accuracy score. However, there was no need for us to standardize our training data.

After creating a for loop, we noticed that the cutoff value with the highest accuracy and net benefit was a cutoff value of 0.15, which is the probability threshold that was used to classify the predicted outcomes of the logistic regression model (Appendix 11). If the predicted probability of the diagnosis was greater than 0.15, the predicted outcome would have been classified as malignant, and if it is less than or equal to 0.15, it would have been classified as benign. Since the cutoff value was lower than the default probability cutoff of 0.5, it indicated that our model would be less selective in classifying malignancy, making the trade off of classifying more false positives (benign cases predicted as malignant), but less false negatives (malignant cases predicted as benign).

The confusion matrix generated returned 140 accurately predicted malignant tumors and 79 correctly identified benign tumors. The model produced 6 false negative predictions, indicating the misclassification of malignant tumors as benign, as well as 3 false positive predictions, signifying benign tumors erroneously identified as malignant. As previously stated, minimizing false negative predictions is critical since they indicate the possibility of missing malignant tumors, potentially leading to inadequate or delayed treatment. Conversely, false positive predictions may result in unnecessary medical interventions and emotional distress for

patients. Nevertheless, our model's impressive performance is evident since it only had 3 false positive classifications, which is remarkable as it approaches zero (Appendix 12). Overall, the model performed quite well, with a high number of true positives and true negatives, though further optimization is possible to reduce false negative and false positive rates. Based on our cost-benefit analysis, the net benefit value was estimated at \$3,366,000, highlighting the potential economic savings of our model in clinical practice.

For this model, we once again utilized multiple evaluation metrics to comprehensively assess the model's performance, including precision, true positive rate (TPR), false positive rate (FPR), F1 score, and Matthews Correlation Coefficient (MCC). The precision score of 0.92 demonstrates the model's ability to accurately predict malignancy, as 92% of the model's malignant predictions were correct. Additionally, the TPR of 0.92 highlights that the model accurately identified 92% of all actual malignant cases, which is a crucial metric for medical applications. The FPR of 0.048 indicates that the model misclassified only a small percentage of benign cases as malignant, demonstrating a high level of specificity. Moreover, the F1 score of 0.92 suggests an outstanding balance between precision and TPR. Finally, the MCC of 0.872 reveals a robust correlation between the predictions of benign and malignant cases, indicating that both were predicted accurately.

With our LR model, we also wrote some code that outputted the original and exponentiated coefficients for the regression (Appendix 13). This would enable us to determine which of the columns are most important in classifying the cell nuclei in breast tissues. We were heavily interested in the exponentiated coefficients because they were the most useful in determining the magnitude of the change in odds per one-unit increase in the corresponding predictor variable. From our coefficient estimates, we noticed that there were 3 variables that played a key role in determining the classification of a patient. First was the symmetry of the cell nuclei. An increase in one unit of symmetry was associated with an increase in the odds of having a malignant tumor by a factor of 21,279. Next was the compactness of the cell nuclei. An increase in one unit of compactness was associated with an increase in the odds of having a malignant tumor by a factor of 7,710. Third was the concave points of the cell nuclei. An increase in one unit of the nuclei's concave points is associated with an increase in the odds of having a malignant tumor by a factor of 392. The results of our LR model analysis highlight that symmetry, compactness, and concave points are the most important attributes that heavily determine how a patient will be classified.

Overall, the performance of our LR model in accurately distinguishing between malignant and benign breast cancer data was very good. The exponentiated coefficients in the LR model also indicate the relevance of each feature in predicting the likelihood of malignancy.

Decision Trees

The next model we trained was a decision tree classification model. Similar to any other decision tree, we first took a look at the full tree with no limitations to get a general idea of where we stood with its accuracy score as well as determining the number of dimensions we were working with (nodes, leaves, maximum depth, features, classes). As seen in Appendix 14, running the basic decision tree gave us 2 classes, 20 features, 23 nodes, 12 leaves, with a maximum depth of 5. This tree gave us an accuracy score of 90.8%. We attributed this relatively low accuracy score to possibly result from our small dataset. The confusion matrix for this model returned 0 False Positives and 0 False Negatives on the training data (100% accuracy score), which strongly suggested an overfit decision tree, which further explained the lower accuracy rate on the test set. For the test set, this model predicted 7 false negatives and 14 false positives, and a net benefit of \$3,030,000.

In order to prune the tree, we used a combination of KFold and GridSearch cross validation methods to optimize our parameters. Kfold was used with 5 splits, which in turn fed into the number of folds for GridSearchCV. The pruned decision tree can be seen in Appendix 15, with 2 classes, 20 features, 15 nodes, 8 leaves, and a maximum depth of 3. This pruned tree returned a far greater accuracy score of 98.23%, with 2 false negatives and 0 false positives. It must be noted that this reduced tree also used a reduced test set, and thus could have inflated the accuracy score. The reduced testing sample also caused our calculations for net benefit, which depends on values of the confusion matrix, to be lower than the rest of the models despite its high accuracy score. Thus, calculating the net benefit proved to be difficult, and was instead approximated ~\$3,400,000. To investigate such a drastic increase in accuracy scores, we used a feature importances graph to visualize which features were the most influential to our model. Looking into our feature importances (Appendix 16) we discovered that `concavepoints_mean` outweighed the other features in our decision tree model by approximately tenfold. Concave points represent the number of concave portions or indentations of the cell nucleus, indicating an irregular shape and presence of malignant cancer cells. Another interesting finding was that our other models such as logistic regression have weighed other features to be important as well. We were also surprised that a model that considered one feature to be far more important than the rest performed so well with the testing set.

Random Forest

As an extension from the decision tree model, we also looked into running a random forest model. As an ensemble method, there would be less concerns that the model would overfit. This model returned 9 false negatives and 5 false positives, with an accuracy score of 93.86% and net benefit of \$3,186,000. This model, despite a GridSearch cross validation with 5 folds, did not produce a significant improvement. The model, shown in Appendix 17, shows the results of our random forest model and its feature importances graph.

As can be seen from the graph, this model features a much more balanced assortment of features and their rankings, with “area_worst” and “concavepoints_worst” showing to be the most important features of this model.

Neural Networks

In this analysis, we explored neural networks as a model for binary classification. Before we developed our model, we preprocessed our data. Data transformation was applied to normalize the values into 0-1, and as per the other models, we split our data into a training set and test set on a 60-40 split. We decided to run the neural network model only on the dataset with both mean and worst values because we had limited processing power and time.

When we created the model for the mean+worst dataset, we selected 20 nodes for our input layer because it matched the number of columns we were evaluating. We selected 1 node for our output layer and ‘sigmoid’ as our activation function because we are conducting a binary classification where the end result could only be two values. We selected the Root Mean Square Propagation for our optimization algorithm because it helps accelerate convergence and is suited for high dimensional data. We experimented with a number of different nodes and hidden layers to find the model that maximized accuracy while minimizing the binary cross-entropy as our loss function.

We made the decision to experiment with a different number of hidden layers because deep learning (3+ hidden layers) is associated with image recognition, and our data points came from a digitalized scan of a cell nucleus.

The results of our neural network models are shown in appendix 18. In our neural network model, there is not a significant increase in accuracy as we tweaked our parameters. We saw slight increases in accuracy when we increased the number of nodes used and a subtle decline in our loss function of about 4-5% as we increased our layers. Slight discrepancies can exist due to random initialization and difference in hardware.

For our model, we selected [8,8] for the hidden layer sizes because it is a top 3 accuracy score and the lowest loss score. We decided to use 2 hidden layers instead of 1 because our data points come from a digitized image and we wanted to capture the dimensionality of the data. We refrained from using 3 layers to prevent overfitting as feature selection was done on this dataset prior.

Next, we created the confusion matrix from this new model in appendix 19. The model incorrectly classified 5 people who have breast cancer as benign and incorrectly classified 4 people who do not have breast cancer as malignant. The misclassification is slightly higher than our KNN model and signals that this neural network may be underperforming due to parameter selection. With this, we performed the cost-benefit analysis and received a net value of \$2,964,000.

Just like prior models, we utilized various evaluation metrics to evaluate. The model evaluation metrics used in this model included precision, true positive rate (TPR), false positive rate (FPR), F1 score, and Matthews Correlation Coefficient (MCC). For neural networks, we calculated a precision score of 0.9375 which indicates a strong predictive power in detecting malignancy. Additionally, the True Positive Rate of 0.949 indicates that out of all the actual malignant cases the data had, the model correctly identified 94.9% of the malignant cases, which is a critical metric for medical diagnosis. There is a False Positive Rate of 0.033 suggests that the model misclassified only a small proportion of benign cases as malignant. A F1 score of 0.9423 indicates a good balance between precision and TPR. Finally, the MCC of 0.887 suggests that both the benign and malignant cases were predicted well via the strong correlation. The MCC score is similar to the Naive Bayes model.

Overall, our neural network model achieved high scores across various evaluation metrics, indicating high predictive power. With only 2 hidden layers and 8 nodes each, our NN model displayed the potential to achieve a higher accuracy level and playing around with the parameters could increase the model's capability to accurately predict malignant cases.

Conclusion

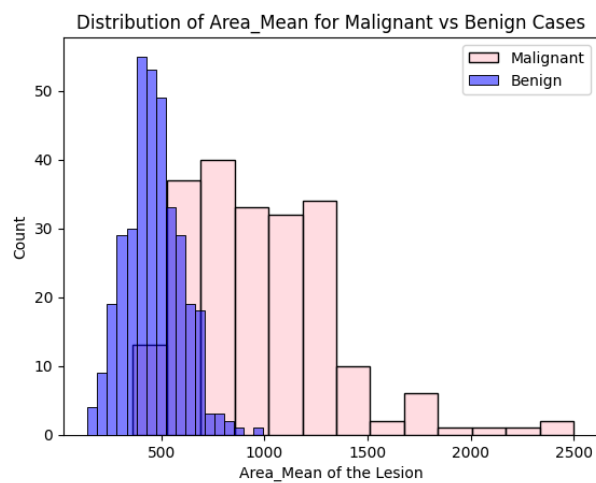
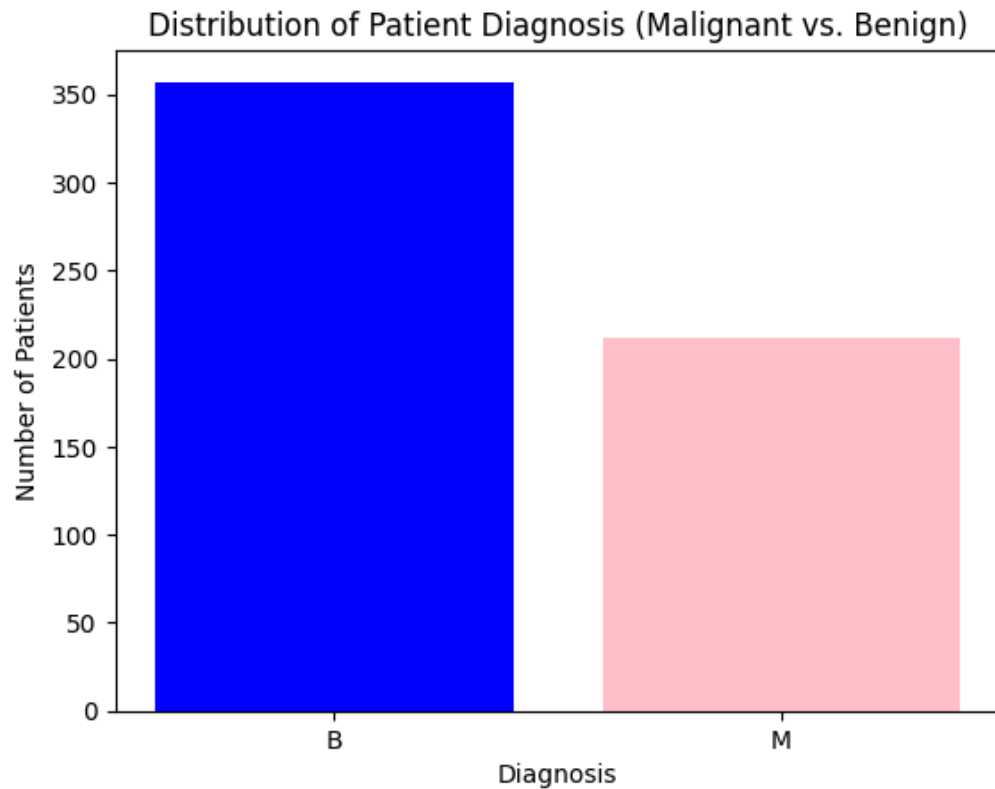
Comparing our models side by side, we reached the highest accuracy score with the reduced decision tree model, which has an accuracy score of 98.23%. This model has high performance across all of the evolution metrics and can be run quickly. For future analysis, we can gather datasets of patient scans from other types of biopsies. With Fine Needle Aspirate Biopsy being the most affordable and least invasive option, it is oftentimes the least accurate. Using a combination of these models, there is potentially an opportunity to incorporate these models to compare the accuracy of other biopsies which will help distinguish which combination of method and models is the most efficient in diagnosing breast cancer.

Appendix

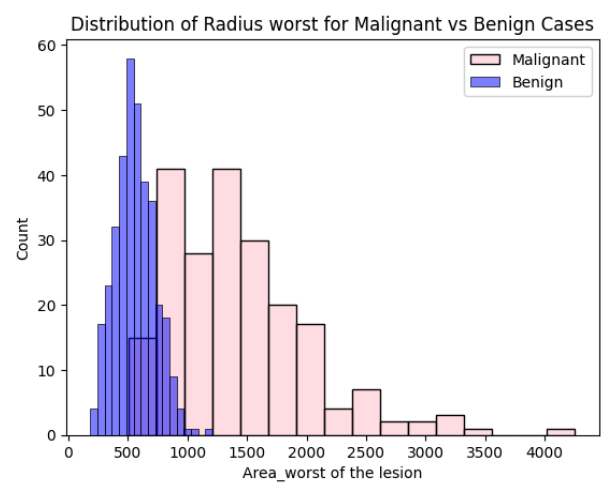
[Appendix 1: Variable Type Book]

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	id	569 non-null	int64
1	diagnosis	569 non-null	object
2	radius_mean	569 non-null	float64
3	texture_mean	569 non-null	float64
4	perimeter_mean	569 non-null	float64
5	area_mean	569 non-null	float64
6	smoothness_mean	569 non-null	float64
7	compactness_mean	569 non-null	float64
8	concavity_mean	569 non-null	float64
9	concave points_mean	569 non-null	float64
10	symmetry_mean	569 non-null	float64
11	fractal_dimension_mean	569 non-null	float64
12	radius_se	569 non-null	float64
13	texture_se	569 non-null	float64
14	perimeter_se	569 non-null	float64
15	area_se	569 non-null	float64
16	smoothness_se	569 non-null	float64
17	compactness_se	569 non-null	float64
18	concavity_se	569 non-null	float64
19	concave points_se	569 non-null	float64
20	symmetry_se	569 non-null	float64
21	fractal_dimension_se	569 non-null	float64
22	radius_worst	569 non-null	float64
23	texture_worst	569 non-null	float64
24	perimeter_worst	569 non-null	float64
25	area_worst	569 non-null	float64
26	smoothness_worst	569 non-null	float64
27	compactness_worst	569 non-null	float64
28	concavity_worst	569 non-null	float64
29	concave points_worst	569 non-null	float64
30	symmetry_worst	569 non-null	float64
31	fractal_dimension_worst	569 non-null	float64
32	Unnamed: 32	0 non null	float64

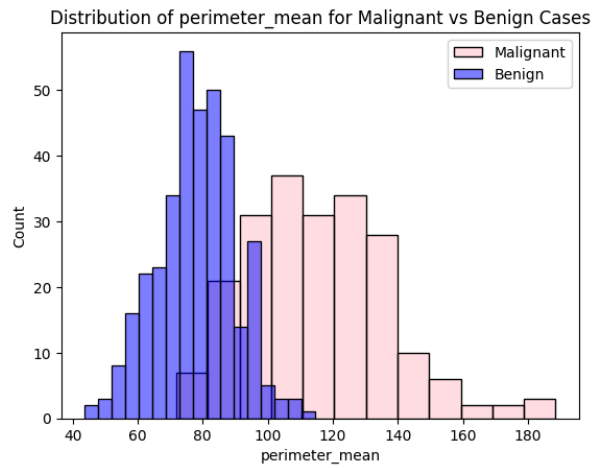
[Appendix 2: Distribution of Patient Diagnosis]



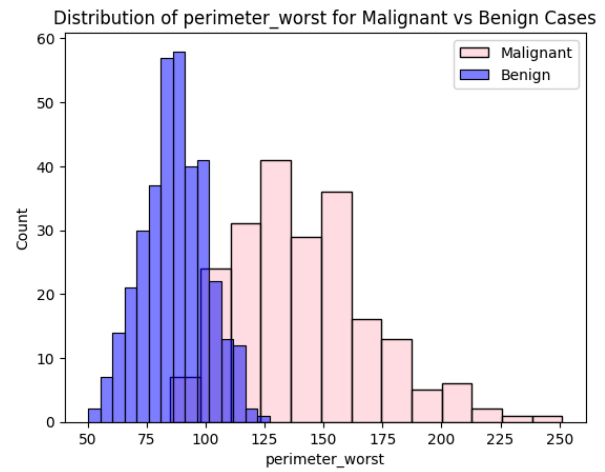
[Appendix 3: Guestimate Area of Lesion]



[Appendix 4: Most Severe Area of Lesion]



[Appendix 5: Perimeter of Lesion]

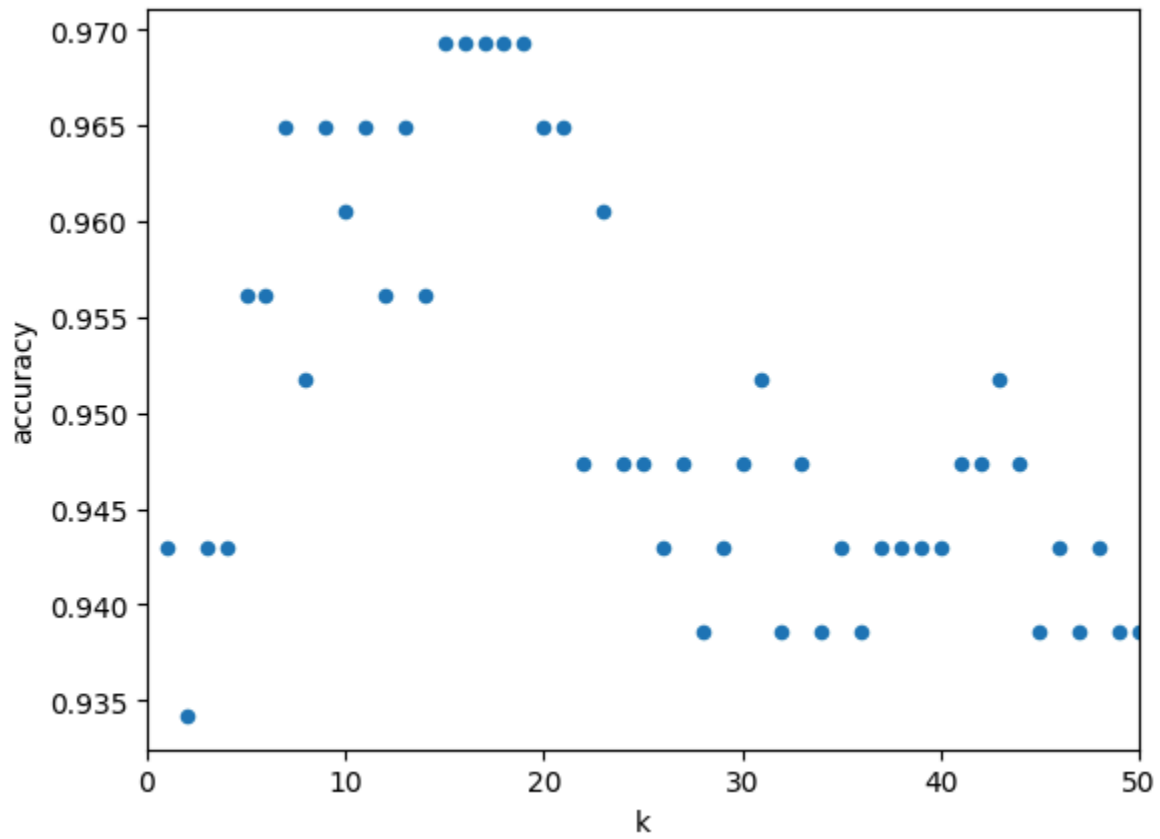


[Appendix 6: Most Severe Perimeter of Lesion]

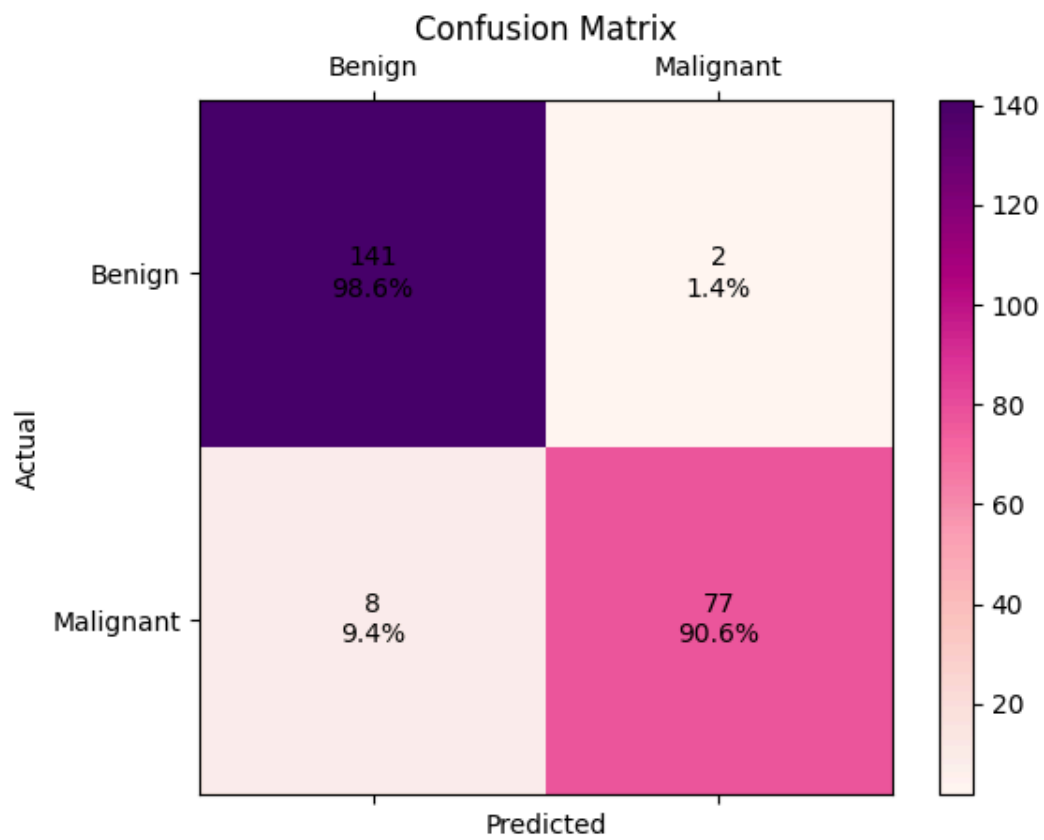
[Appendix 7: Cost Benefit Matrix]

	Pred. Benign (0)	Pred. Malignant (1)
Actual Benign (0)	Benefit: \$18K	Cost: -\$6K
Actual Malignant (1)	Cost: -\$18K	Benefit: \$12K

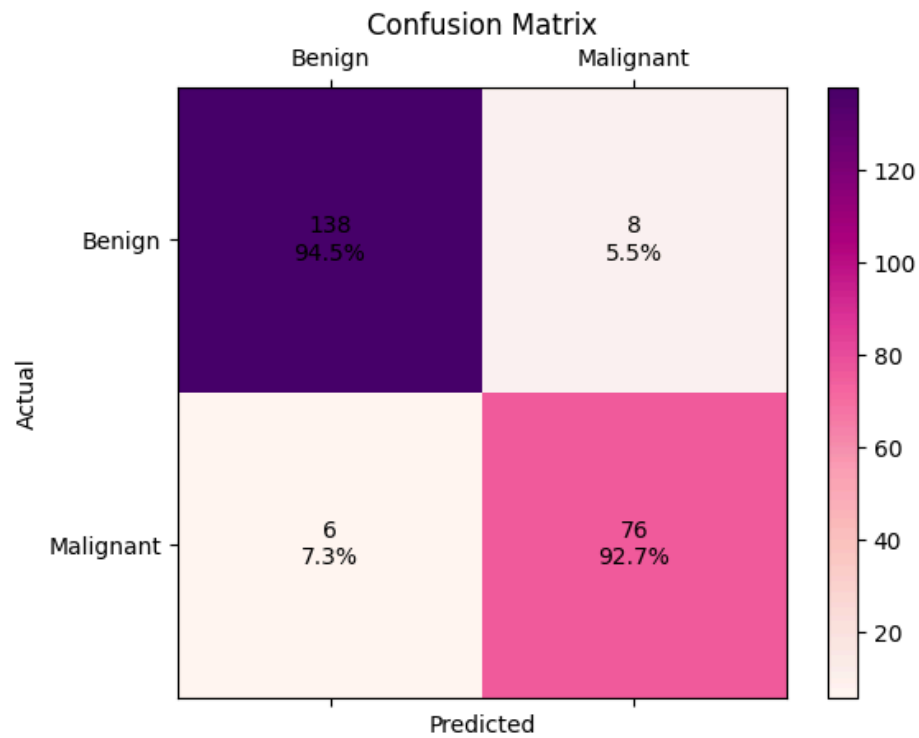
[Appendix 8: KNN K-value vs. Accuracy]



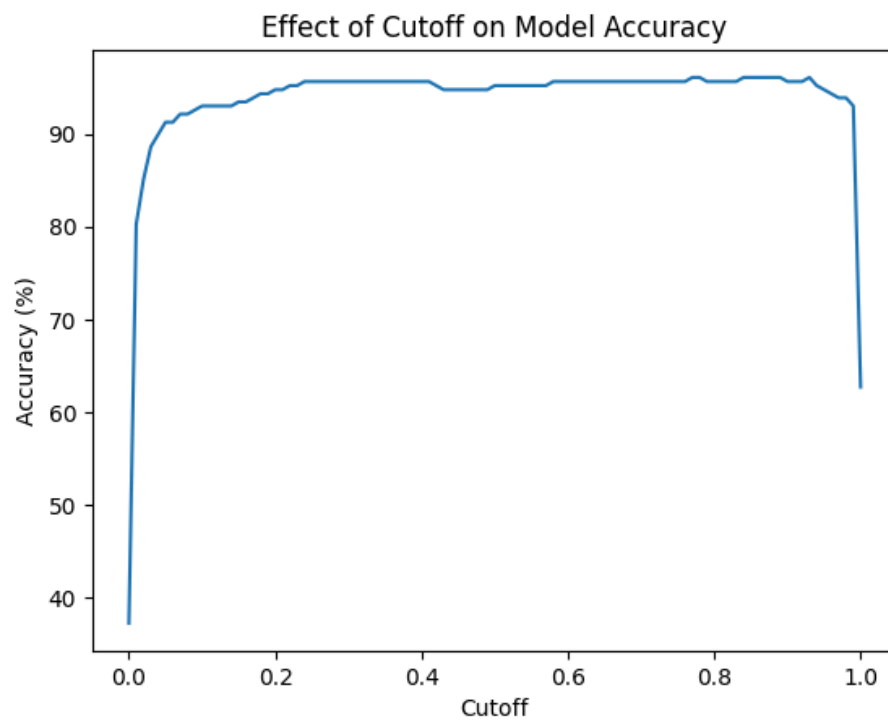
[Appendix 9: KNN Confusion Matrix]



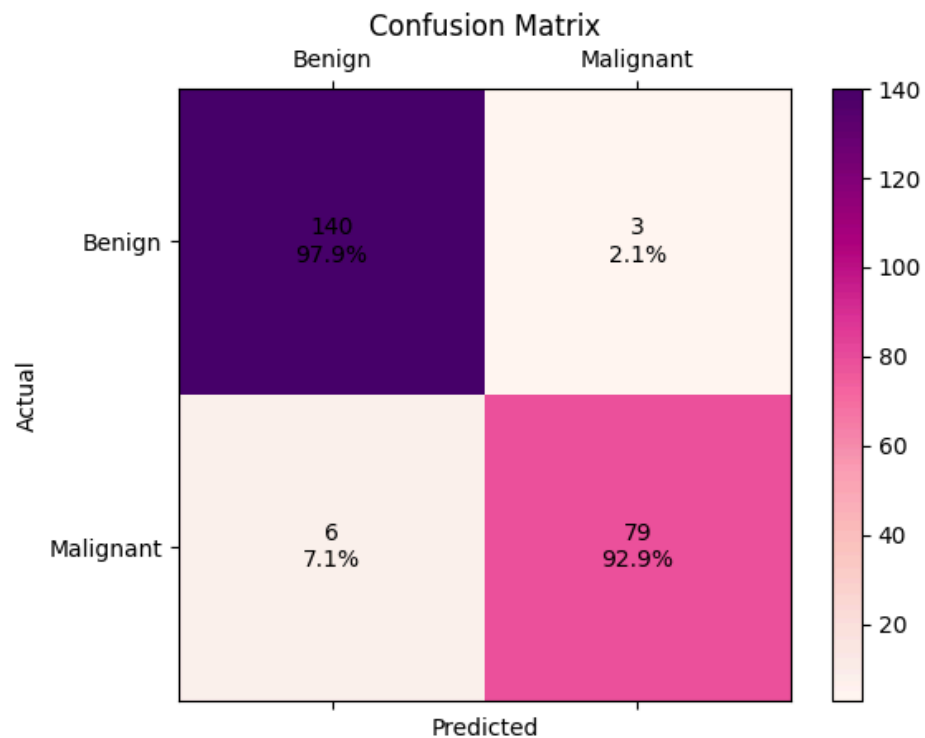
[Appendix 10: NB Confusion Matrix]



[Appendix 11: LR Cutoff vs. Accuracy]



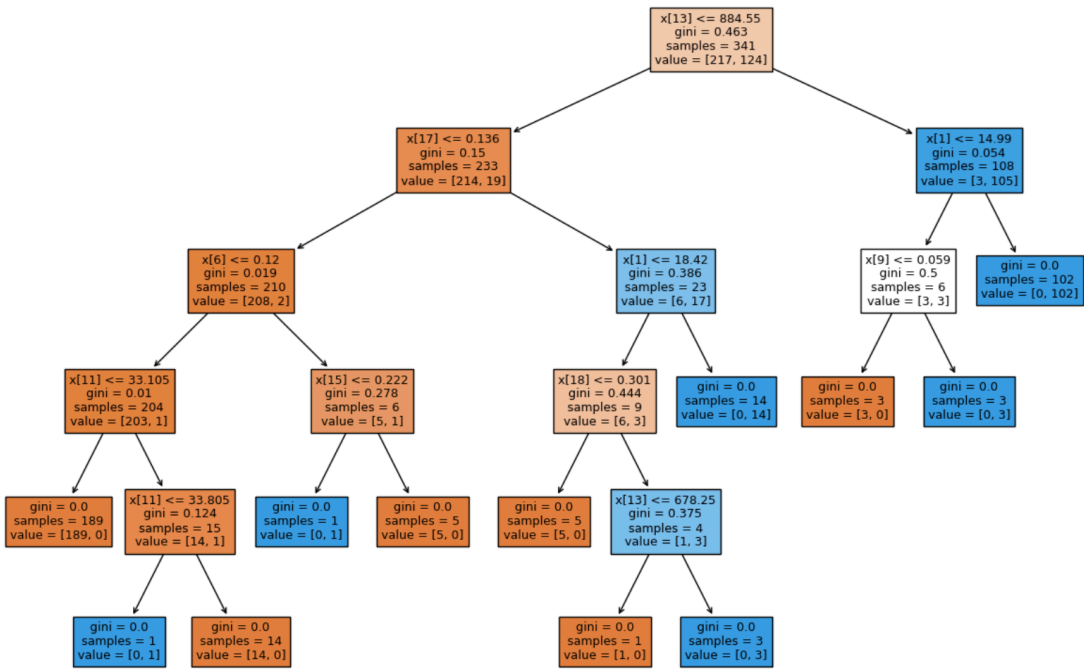
[Appendix 12: LR Confusion Matrix]



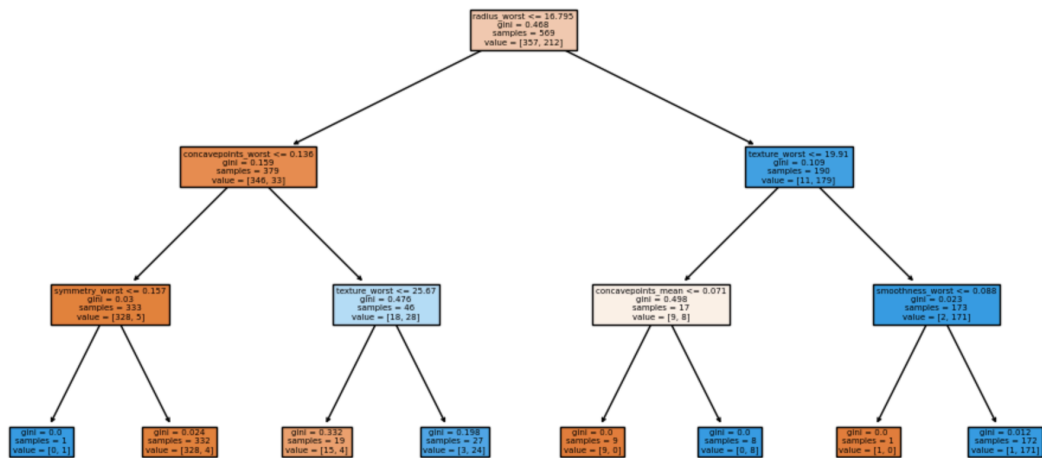
[Appendix 13: Exponentiated Coefficients]

```
radius_worst: 0.17481655886984745
texture_worst: 1.2472359653348204
perimeter_worst: 0.8664028878316166
area_worst: 1.0441315407065261
smoothness_worst: 36.575561025197445
compactness_worst: 7709.713789051001
concavity_worst: 10.612202477649118
concave_points_worst: 392.12382080379416
symmetry_worst: 21279.282102773253
fractal_dimension_worst: 2.532806356529277
```

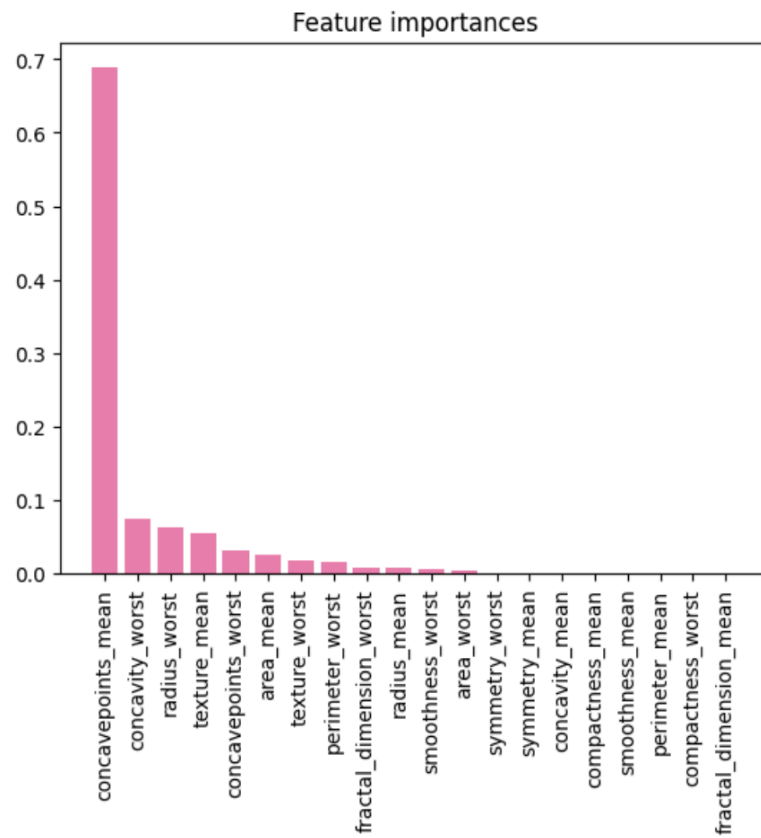
[Appendix 14: Full Decision Tree]



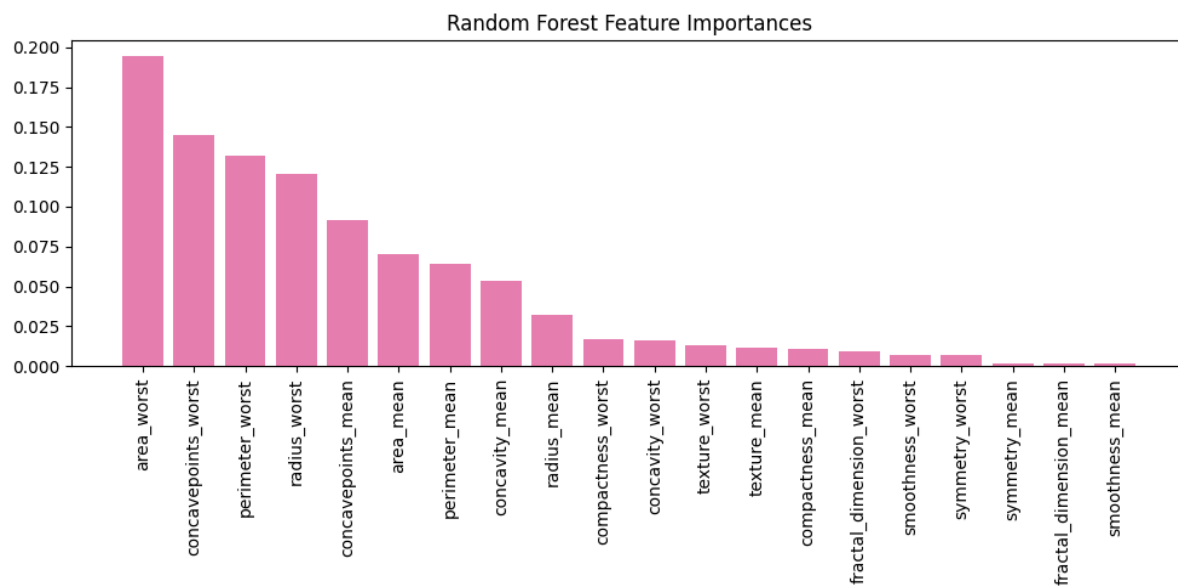
[Appendix 15: Pruned Decision Tree]



[Appendix 16: Decision Tree Feature Importances]



[Appendix 17: Random Forest Feature Importances]



[Appendix 18: Mean + Worst (ALL) Dataset Nodes]

Number of Hidden Layers	Number of Nodes	Epochs (Max:256)	Loss	Accuracy Train	Accuracy Valid
1	4	80	0.1101	0.9912	0.9561
	6	53	0.0748	0.9795	0.9605
	8	53	0.0807	0.9795	0.9518
	32	53	0.0651	0.9853	0.9518
2	4,4	74	0.0783	0.9824	0.9561
	6,6	60	0.0653	0.9824	0.9561
	8,8	53	0.0654	0.9824	0.9605
	32,16	41	0.0662	0.9795	0.9649
3	4,4,4	53	0.0663	0.9824	0.9518
	6,6,6	24	0.1073	0.9677	0.9342
	8,8,8	41	0.0665	0.9765	0.9649
	32,16,8	45	0.0645	0.9795	0.9561

[Appendix 19: Mean + Worst (ALL) Confusion Matrix]

