# Revolutionizing Healthcare: Developing an AI Model for Breast Cancer Prediction

# Introduction to AI in Healthcare

**Artificial Intelligence** is transforming healthcare by providing innovative solutions for early diagnosis. This presentation explores the development of an AI model specifically designed for **breast cancer prediction**, aiming to enhance patient outcomes and streamline diagnostic processes.

# Understanding Breast Cancer

Breast cancer is one of the most prevalent cancers among women. Early detection is crucial for effective treatment. Understanding the risk factors and symptoms can lead to timely medical intervention and significantly improve survival rates.

# Technological Innovations in Predictions

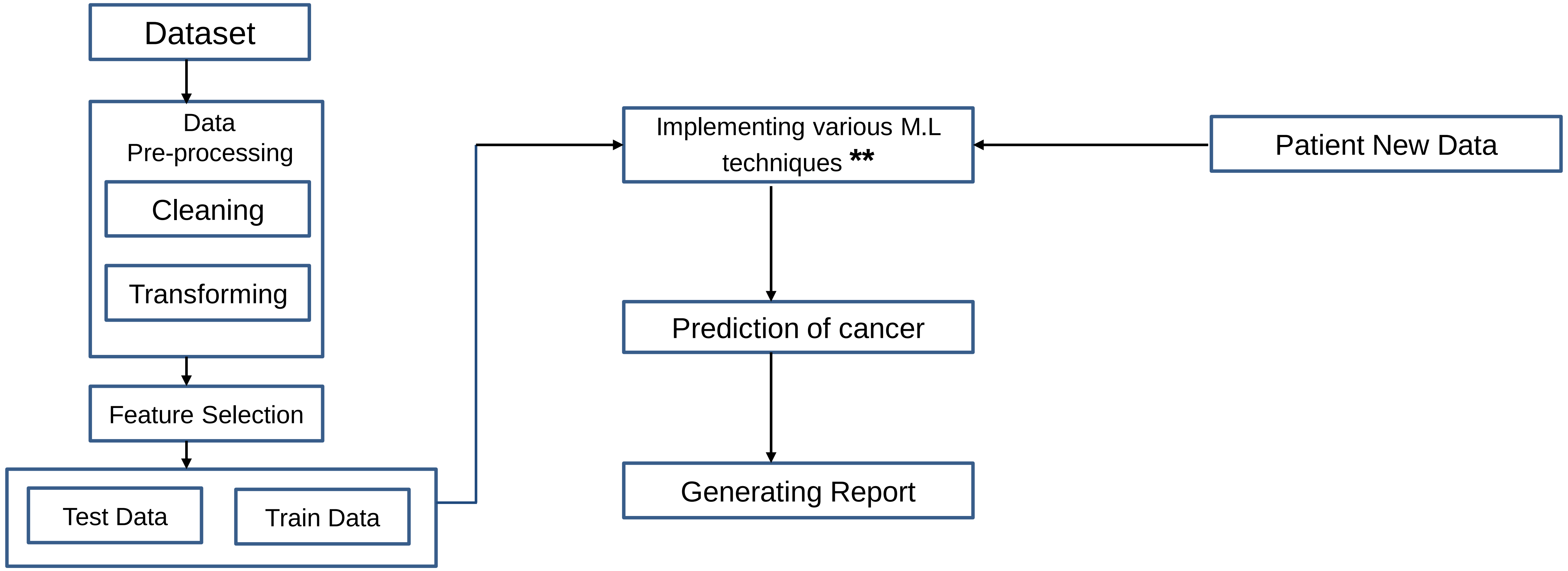
The integration of **Artificial Intelligence** and **Machine Learning** in breast cancer prediction models has revolutionized the way we analyze data. These technologies enable more accurate **risk assessments** and personalized treatment plans, ultimately leading to better patient outcomes.

# AI Model Development Process

Developing an effective AI model involves several steps: data collection, preprocessing, model selection, and training. The model is designed to analyze medical imaging and patient data to predict breast cancer risk with high accuracy.

///////

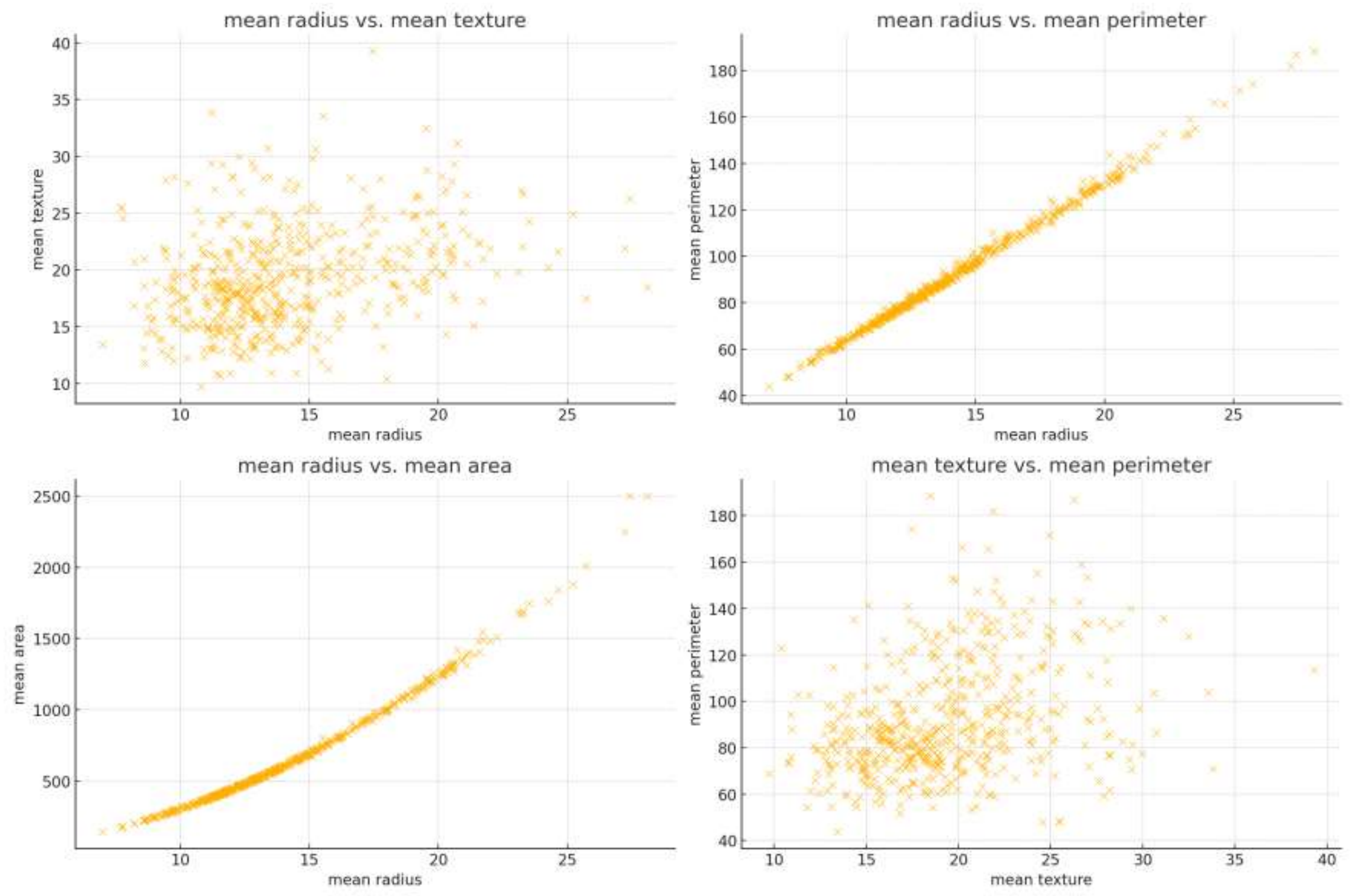
# Steps to Make the Model



# Visualization of the Features

- *Radius (mean of distances from the center to points on the perimeter)*
- *Texture (standard deviation of gray-scale values)*
- *Perimeter*
- *Area*
- *Smoothness (local variation in radius lengths)*
- *Compactness (perimeter^2 / area - 1.0)*
- *Concavity (severity of concave portions of the contour)*
- *Concave points (number of concave portions of the contour)*
- *Symmetry*
- *Fractal dimension ("coastline approximation" - 1)*

| | id | diagnosis | radius_mean | texture_mean | perimeter_mean | area_mean | smoothness_mean | compactness_ |
|---|---|---|---|---|---|---|---|---|
| 1 | 842302 | M | 17.99 | 10.38 | 122.8 | 1001 | 0.1184 | 0. |
| 2 | 842517 | M | 20.57 | 17.77 | 132.9 | 1326 | 0.08474 | 0.0 |
| 3 | 84300903 | M | 19.69 | 21.25 | 130 | 1203 | 0.1096 | 0. |
| 4 | 84348301 | M | 11.42 | 20.38 | 77.58 | 386.1 | 0.1425 | 0. |
| 5 | 84358402 | M | 20.29 | 14.34 | 135.1 | 1297 | 0.1003 | 0. |
| 6 | 843786 | M | 12.45 | 15.7 | 82.57 | 477.1 | 0.1278 | |
| 7 | 844359 | M | 18.25 | 19.98 | 119.6 | 1040 | 0.09463 | 0 |
| 8 | 84458202 | M | 13.71 | 20.83 | 90.2 | 577.9 | 0.1189 | 0. |
| 9 | 844981 | M | 13 | 21.82 | 87.5 | 519.8 | 0.1273 | 0. |
| 10 | 84501001 | M | 12.46 | 24.04 | 83.97 | 475.9 | 0.1186 | 0. |
| 11 | 845636 | M | 16.02 | 23.24 | 102.7 | 797.8 | 0.08206 | 0.0 |
| 12 | 84610002 | M | 15.78 | 17.89 | 103.6 | 781 | 0.0971 | 0. |
| 13 | 846226 | M | 19.17 | 24.8 | 132.4 | 1123 | 0.0974 | 0. |
| 14 | 846381 | M | 15.85 | 23.95 | 103.7 | 782.7 | 0.08401 | 0. |
| 15 | 84667401 | M | 13.73 | 22.61 | 93.6 | 578.3 | 0.1131 | 0. |
| 16 | 84799002 | M | 14.54 | 27.54 | 96.73 | 658.8 | 0.1139 | 0. |
| 17 | 848406 | M | 14.68 | 20.13 | 94.74 | 684.5 | 0.09867 | 0 |
| 18 | 84862001 | M | 16.13 | 20.68 | 108.1 | 798.8 | 0.117 | 0. |
| 19 | 849014 | M | 19.81 | 22.15 | 130 | 1260 | 0.09831 | 0. |
| 20 | 8510426 | B | 13.54 | 14.36 | 87.46 | 566.3 | 0.09779 | 0.0 |
| 21 | 8510653 | B | 13.08 | 15.71 | 85.63 | 520 | 0.1075 | 0 |



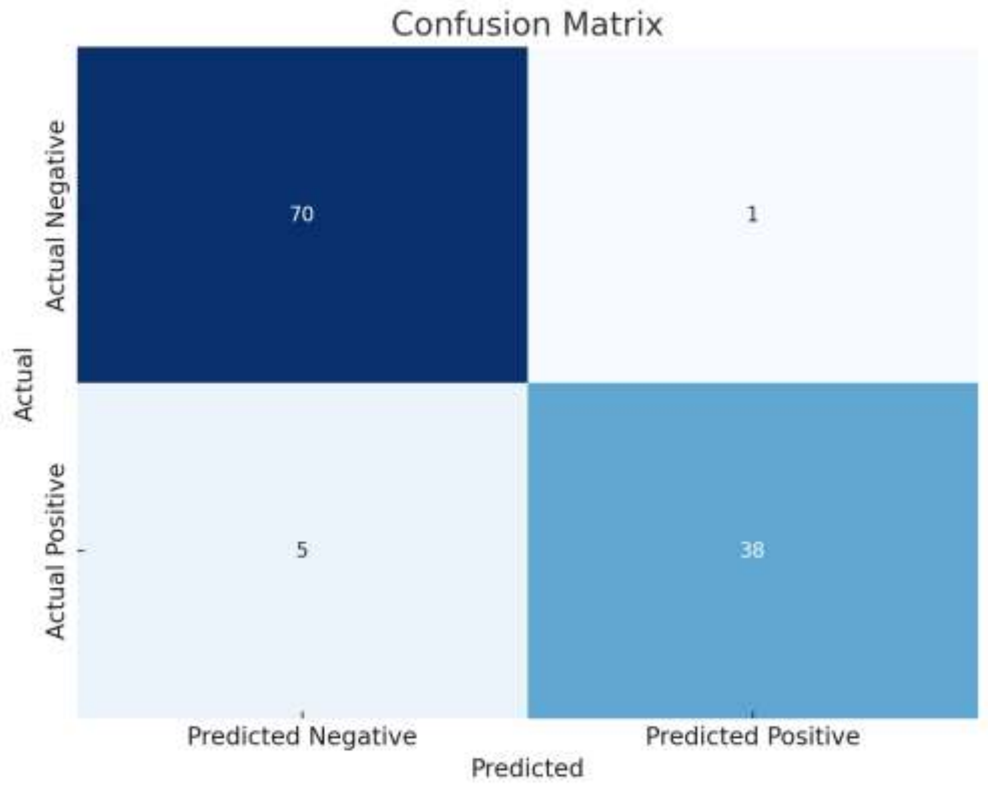

1. **Mean Radius vs. Mean Texture**
   Observation:The plot reveals a fairly scattered distribution, indicating that there is no strong linear correlation between the mean radius and the mean texture. Tumors with a similar radius can have a wide range of textures.
2. **Mean Radius vs. Mean Perimeter**
   Observation: There is a strong positive linear correlation between the mean radius and the mean perimeter. As the radius of the tumor increases, the perimeter also increases proportionally. This makes sense geometrically since the perimeter of a circle is directly proportional to its radius.
3. **Mean Radius vs. Mean Area**
   Observation: There is a strong positive non-linear (quadratic) correlation between the mean radius and the mean area. As the radius increases, the area increases more rapidly.
4. **Mean Texture vs. Mean Perimeter**
   Observation: There is a scattered distribution similar to the first plot, indicating no strong linear correlation between mean texture and mean perimeter. Tumors with a similar perimeter can have a wide range of textures.

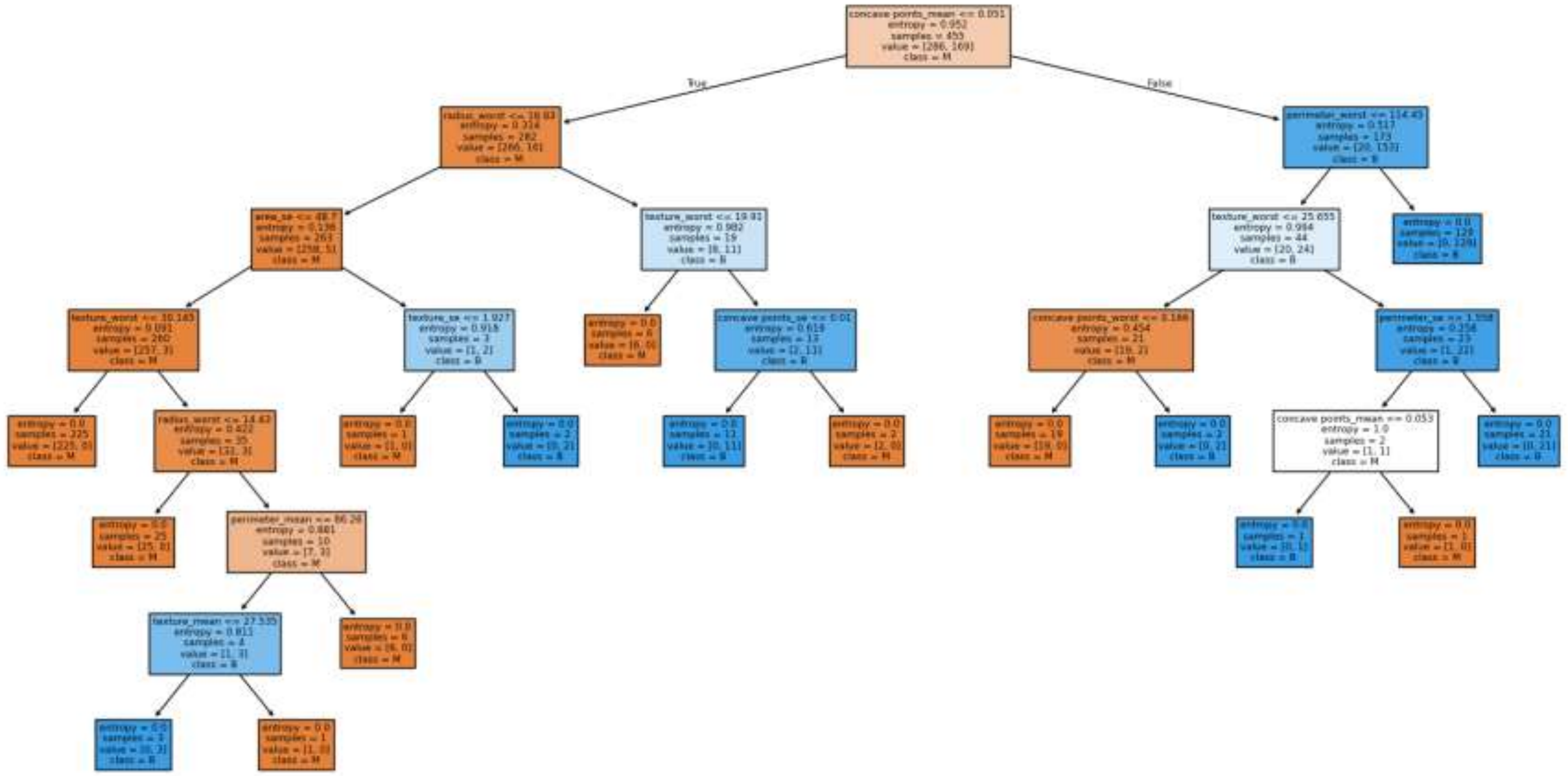
# Techniques Used for the Project

1. *Decision Tree*
2. *Random Forest Classification*
3. *Support Vector Classifier*
4. *Neural Network*

# Decision Tree

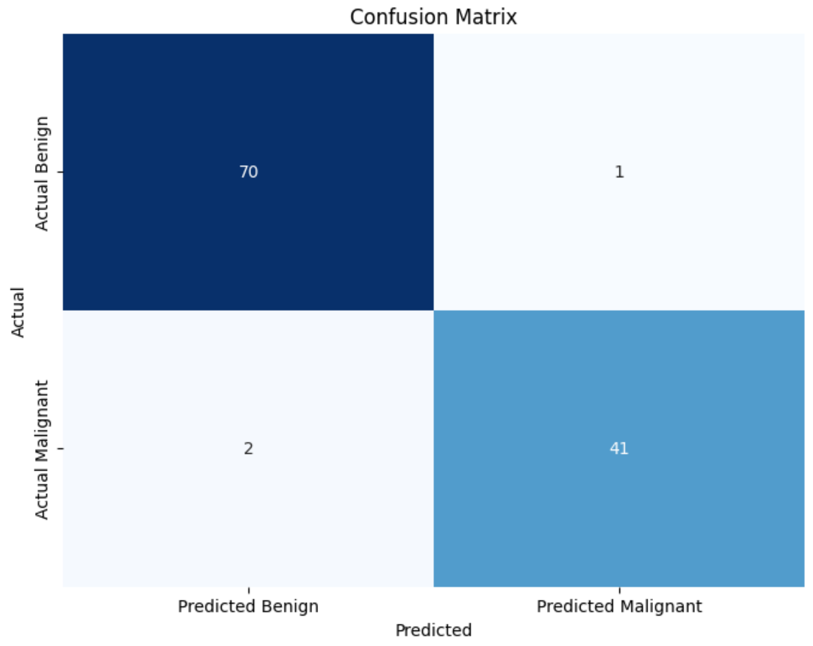
**This Model yielded 94.73% Accuracy for the prediction after taking the training size to 80% of the Data.
This Model had 6 cases where the data was predicted to be wrong from which there are 5 False negative cases.**

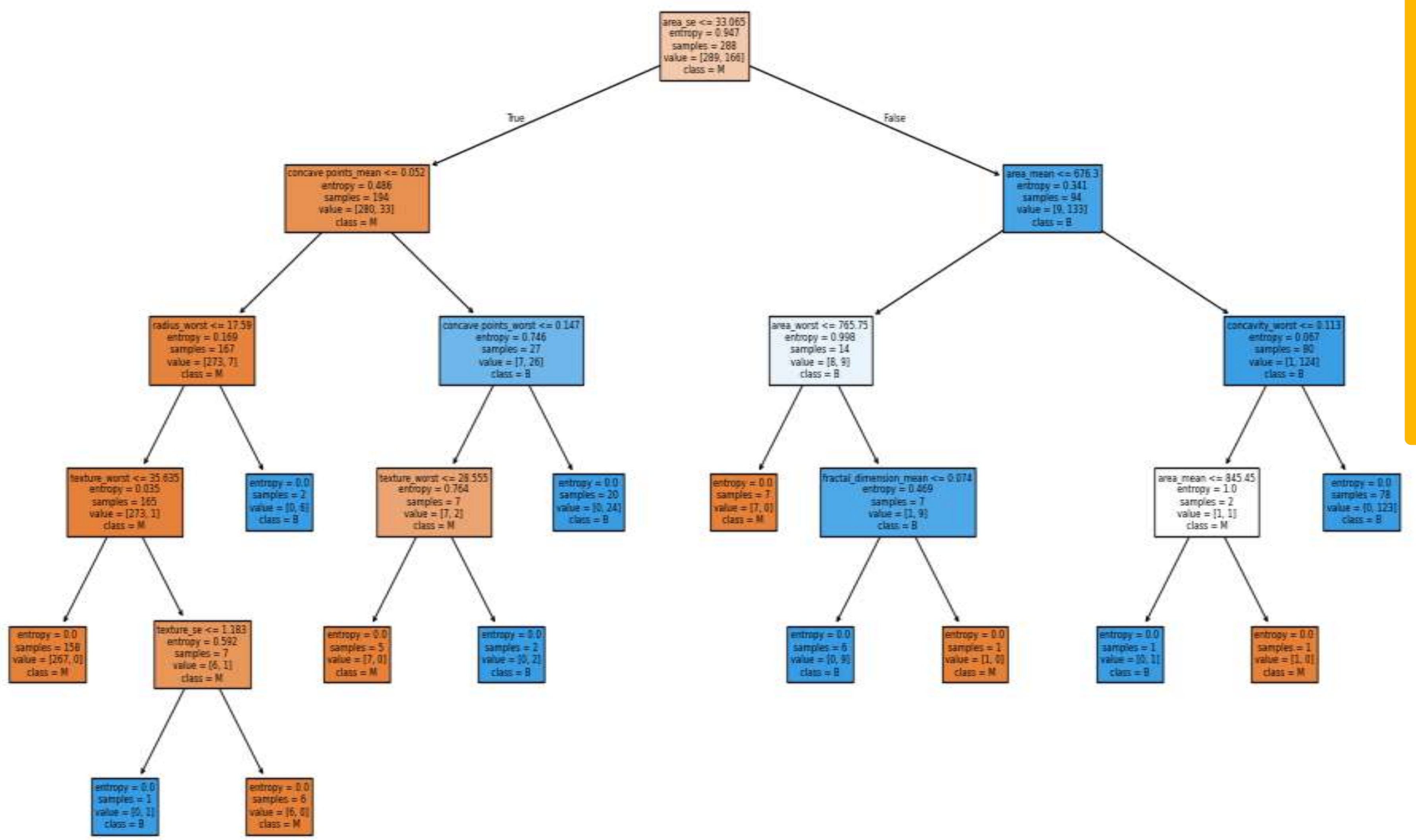


# Random Forest Classifier

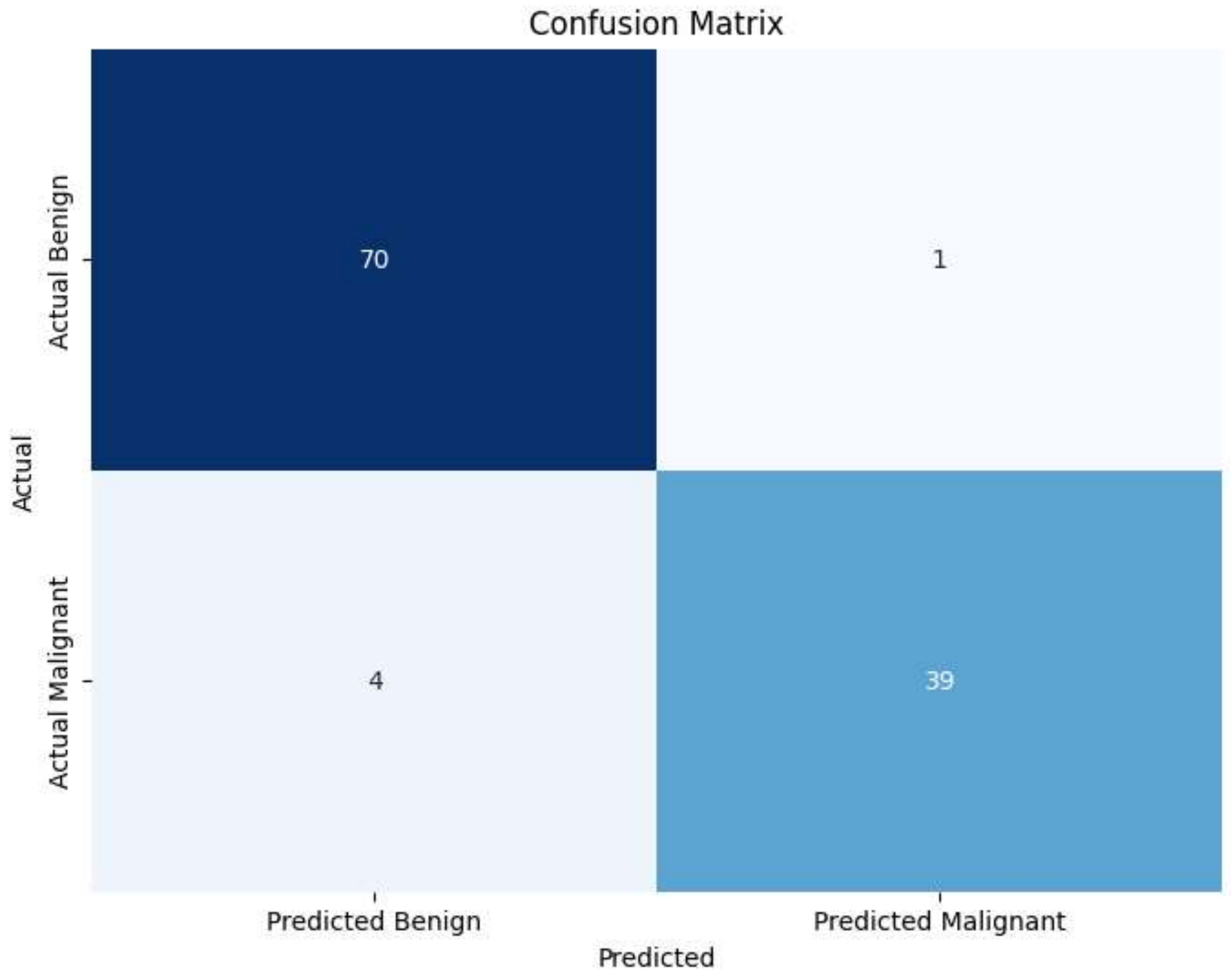
This model is the better version of the decision tree while doing large and complex databases, and this model yielded 97.36% of accuracy.
This model had 3 cases where the data was predicted to be wrong from which there are 2 false negative cases



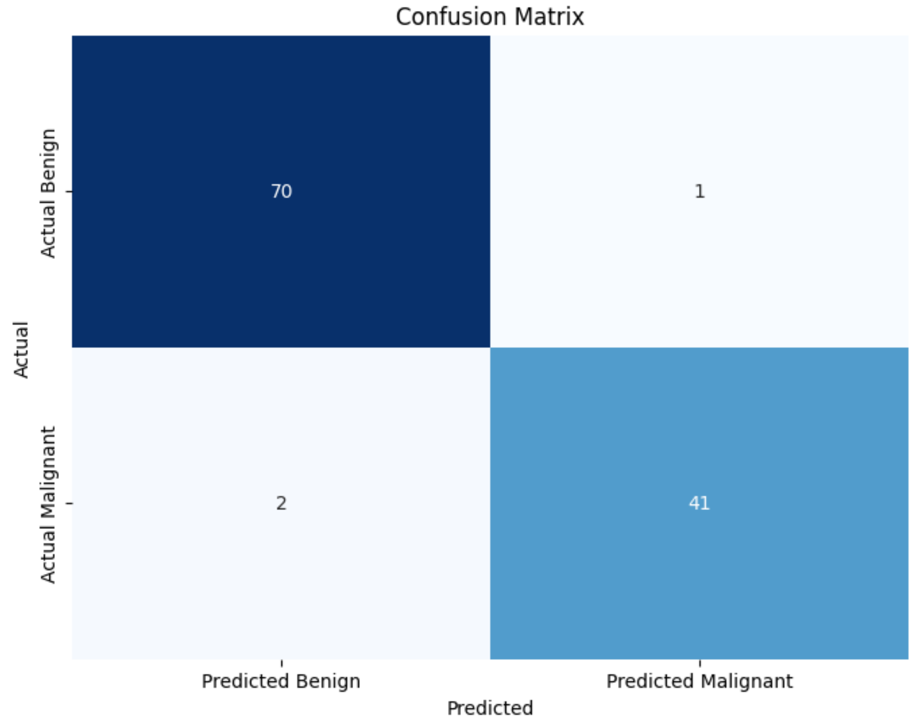
# Support Vector Classifier

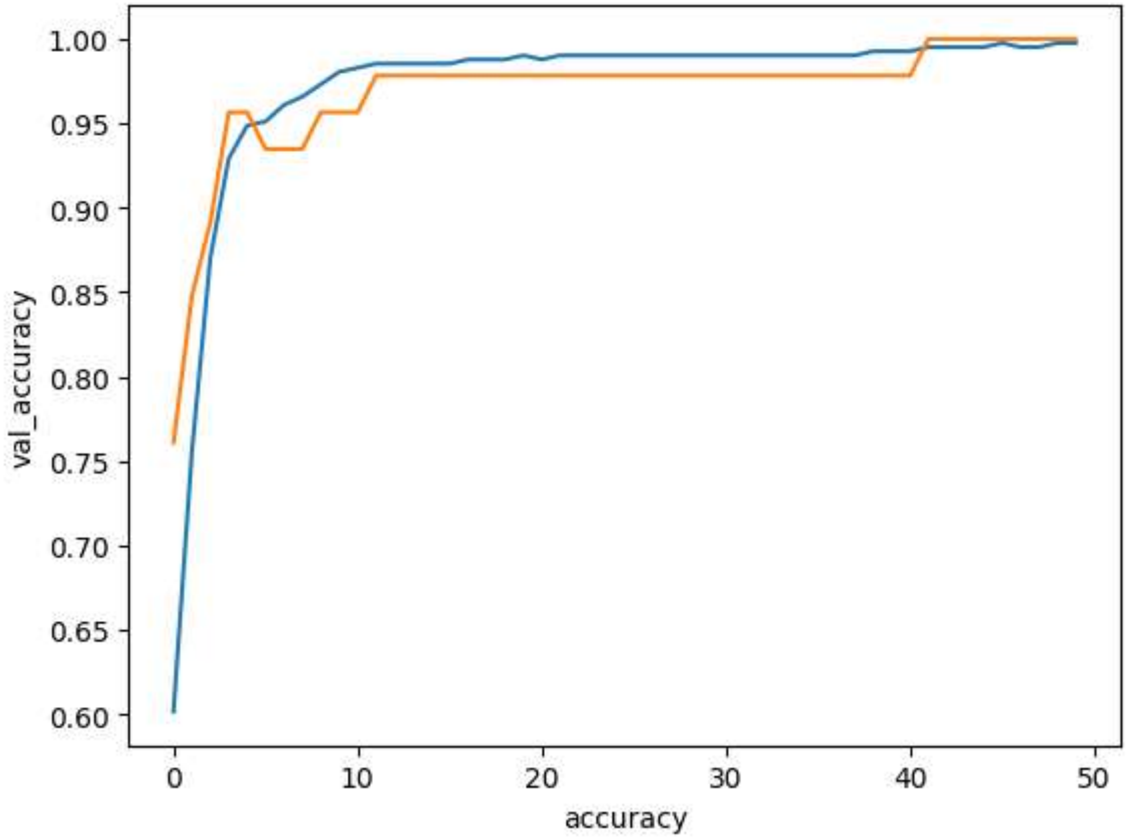
**This Model yielded 95.61% of the accuracy, which makes it a better model than the decision tree model.**
**This model had 5 cases of wrong predictions from there were 4 false negative cases.**



# Neural Network

**This model yielded 97.36% of accuracy which is the same as the random forest model. We used 50 epochs for the model.**
**This may suggest that the highest accuracy achieved by any technique is similar to random forest model.**
**This model had 3 wrong cases from which 2 were false negative cases.**



# Real time Usage of the Project

- *The last two decades a variety of different ML techniques and feature selection algorithms have been widely applied to disease prognosis and predictions.*
- *This model provided that it gets accurate data can be used in clinical practice as a medical tool*
- *This model will help in reducing the human efforts which will help doctors to take proper decisions and steps on time resulting in the benefit of patient and overall society.*

# Data Sources and Features

The AI model utilizes diverse data sources, including **medical records** imaging data, and genetic information. Key features such as age, family history, and **tumor characteristics** are crucial for improving prediction accuracy.

# Results and Impact

Preliminary results indicate that the AI model significantly improves early detection rates of breast cancer. The impact on patient care includes reduced diagnostic time and increased treatment options, ultimately leading to better patient outcomes.

# Conclusion and Future Directions

The development of an AI model for breast cancer prediction represents a significant advancement in healthcare. Future directions include refining the model, expanding data sources, and exploring **integration** into clinical practice for real-time decision support.

# QNA Time

//////

# References


1. https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset



```
"id","diagnosis","radius_mean","texture_mean","perimeter_mean","area_mean","smoothness_mean","compactness_mean","concavity_mean","concave
points_mean","symmetry_mean","fractal_dimension_mean","radius_se","texture_se","perimeter_se","area_se","smoothness_se","compactness_se","concavity_se","concave
points_se","symmetry_se","fractal_dimension_se","radius_worst","texture_worst","perimeter_worst","area_worst","smoothness_worst","compactness_worst","concavity_worst","concave
points_worst","symmetry_worst","fractal_dimension_worst",
842302,M,17.99,10.38,122.8,1001,0.1184,0.2776,0.3001,0.1471,0.2419,0.07871,1.095,0.9053,8.589,153.4,0.006399,0.04904,0.05373,0.01587,0.03003,0.006193,25.38,17.33,184.6,2019,0.1622,0.6656,0
.7119,0.2654,0.4601,0.1189
842517,M,20.57,17.77,132.9,1326,0.08474,0.07864,0.0869,0.07017,0.1812,0.05667,0.5435,0.7339,3.398,74.08,0.005225,0.01308,0.0186,0.0134,0.01389,0.003532,24.99,23.41,158.8,1956,0.1238,0.1866
,0.2416,0.186,0.275,0.08902
84300903,M,19.69,21.25,130,1203,0.1096,0.1599,0.1974,0.1279,0.2069,0.05999,0.7456,0.7869,4.585,94.03,0.00615,0.04006,0.03832,0.02058,0.0225,0.004571,23.57,25.53,152.5,1709,0.1444,0.4245,0.
4504,0.243,0.3613,0.08758
84348301,M,11.42,20.38,77.58,386.1,0.1425,0.2839,0.2414,0.1052,0.2597,0.09744,0.4956,1.156,3.445,27.23,0.00911,0.07458,0.05661,0.01867,0.05963,0.009208,14.91,26.5,98.87,567.7,0.2098,0.8663
,0.6869,0.2575,0.6638,0.173
84358402,M,20.29,14.34,135.1,1297,0.1003,0.1328,0.198,0.1043,0.1809,0.05883,0.7572,0.7813,5.438,94.44,0.01149,0.02461,0.05688,0.01885,0.01756,0.005115,22.54,16.67,152.2,1575,0.1374,0.205,0
.4,0.1625,0.2364,0.07678
843786,M,12.45,15.7,82.57,477.1,0.1278,0.17,0.1578,0.08089,0.2087,0.07613,0.3345,0.8902,2.217,27.19,0.00751,0.03345,0.03672,0.01137,0.02165,0.005082,15.47,23.75,103.4,741.6,0.1791,0.5249,0
.5355,0.1741,0.3985,0.1244
844359,M,18.25,19.98,119.6,1040,0.09463,0.109,0.1127,0.074,0.1794,0.05742,0.4467,0.7732,3.18,53.91,0.004314,0.01382,0.02254,0.01039,0.01369,0.002179,22.88,27.66,153.2,1606,0.1442,0.2576,0.
3784,0.1932,0.3063,0.08368
84458202,M,13.71,20.83,90.2,577.9,0.1189,0.1645,0.09366,0.05985,0.2196,0.07451,0.5835,1.377,3.856,50.96,0.008805,0.03029,0.02488,0.01448,0.01486,0.005412,17.06,28.14,110.6,897,0.1654,0.368
2,0.2678,0.1556,0.3196,0.1151
844981,M,13,21.82,87.5,519.8,0.1273,0.1932,0.1859,0.09353,0.235,0.07389,0.3063,1.002,2.406,24.32,0.005731,0.03502,0.03553,0.01226,0.02143,0.003749,15.49,30.73,106.2,739.3,0.1703,0.5401,0.5
39,0.206,0.4378,0.1072
84501001,M,12.46,24.04,83.97,475.9,0.1186,0.2396,0.2273,0.08543,0.203,0.08243,0.2976,1.599,2.039,23.94,0.007149,0.07217,0.07743,0.01432,0.01789,0.01008,15.09,40.68,97.65,711.4,0.1853,1.058
,1.105,0.221,0.4366,0.2075
845636,M,16.02,23.24,102.7,797.8,0.08206,0.06669,0.03299,0.03323,0.1528,0.05697,0.3795,1.187,2.466,40.51,0.004029,0.009269,0.01101,0.007591,0.0146,0.003042,19.19,33.88,123.8,1150,0.1181,0.
1551,0.1459,0.09975,0.2948,0.08452
84610002,M,15.78,17.89,103.6,781,0.0971,0.1292,0.09954,0.06606,0.1842,0.06082,0.5058,0.9849,3.564,54.16,0.005771,0.04061,0.02791,0.01282,0.02008,0.004144,20.42,27.28,136.5,1299,0.1396,0.56
09,0.3965,0.181,0.3792,0.1048
```

# Thanks!

**Made by:**
**Arjav Jain**
**Saurabh Sharma**
**Swayam Bansal**
**Yash aggarwal**

***arjavjain_2494***
***aggarwal_yash***
***supersam9276***