ABSTRACT:

This project will detect weather a person has breast cancer or not as Breast cancer is a significant health concern worldwide, especially among women. Timely and accurate detection of breast cancer plays a crucial role in successful treatment and improved patient outcomes. Advances in healthcare data collection have resulted in large datasets containing information about breast cancer patients, including clinical data, imaging data (e.g., mammograms), and histopathological information. Leveraging these datasets along with machine learning techniques offers an opportunity to enhance the accuracy and efficiency of breast cancer detection.

PROBLEM STATEMENT:

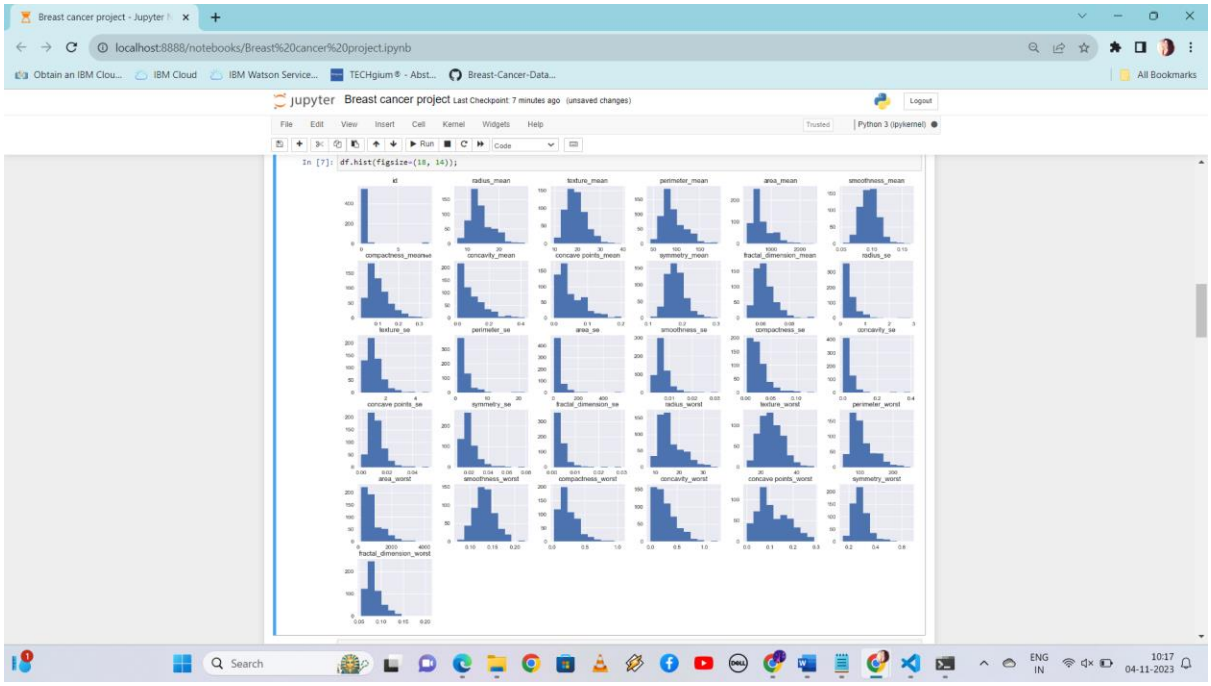
Breast cancer is a leading cause of cancer death in women. Early detection is essential for improving patient outcomes. Current methods for breast cancer detection are not always accurate. Machine learning can be used to develop more accurate methods for breast cancer detection. The goal of this project is to develop a machine learning model to detect breast cancer from medical dataset.

The primary objective of this project is to harness the power of machine learning to address these healthcare challenges. This process is inherently time-consuming, prone to human error, and costly due to the expertise required and the extensive workforce needed. Moreover, even with highly trained professionals, there is always a risk of missed or misdiagnosed cases.

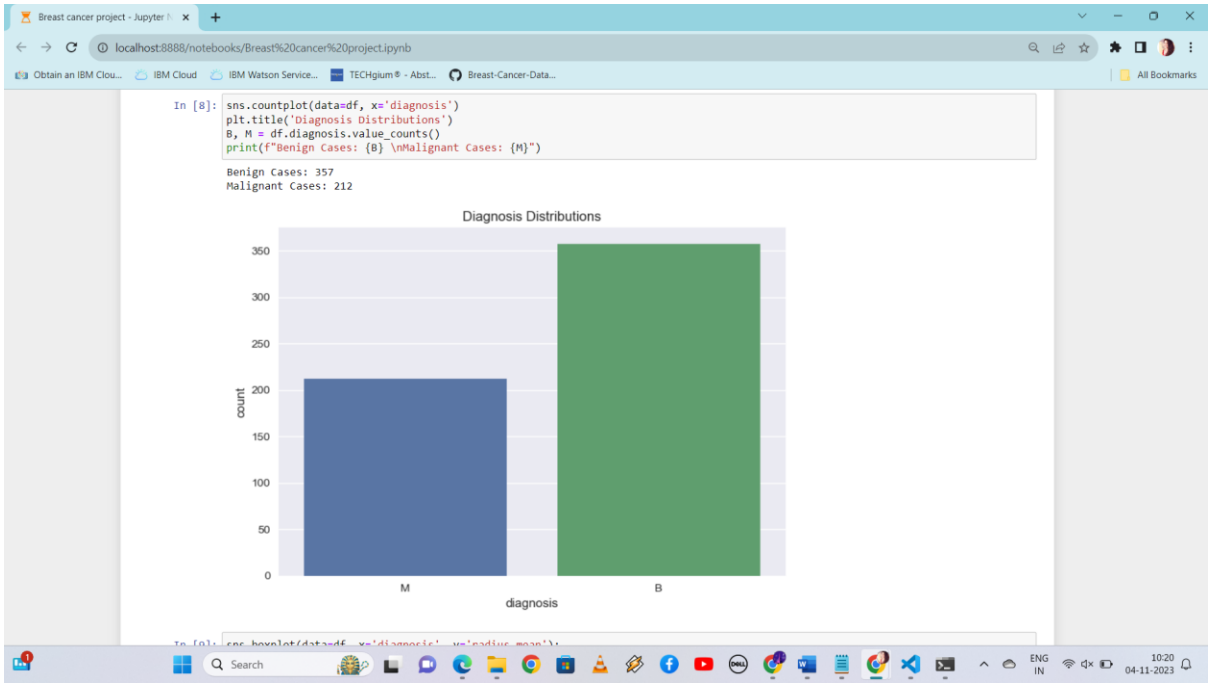
One of the key challenges this project faces is the collection and preprocessing of diverse medical data. Healthcare data comes in various formats and from numerous sources, including electronic health records, wearable devices, and medical imaging. This data must be aggregated, cleaned, and harmonized to create a consistent and reliable dataset for analysis. Moreover, ensuring data privacy and compliance with regulations, such as HIPAA, is paramount to protect patient information.

RESULT AND DICUSSION:

DATA PREPROCESSING

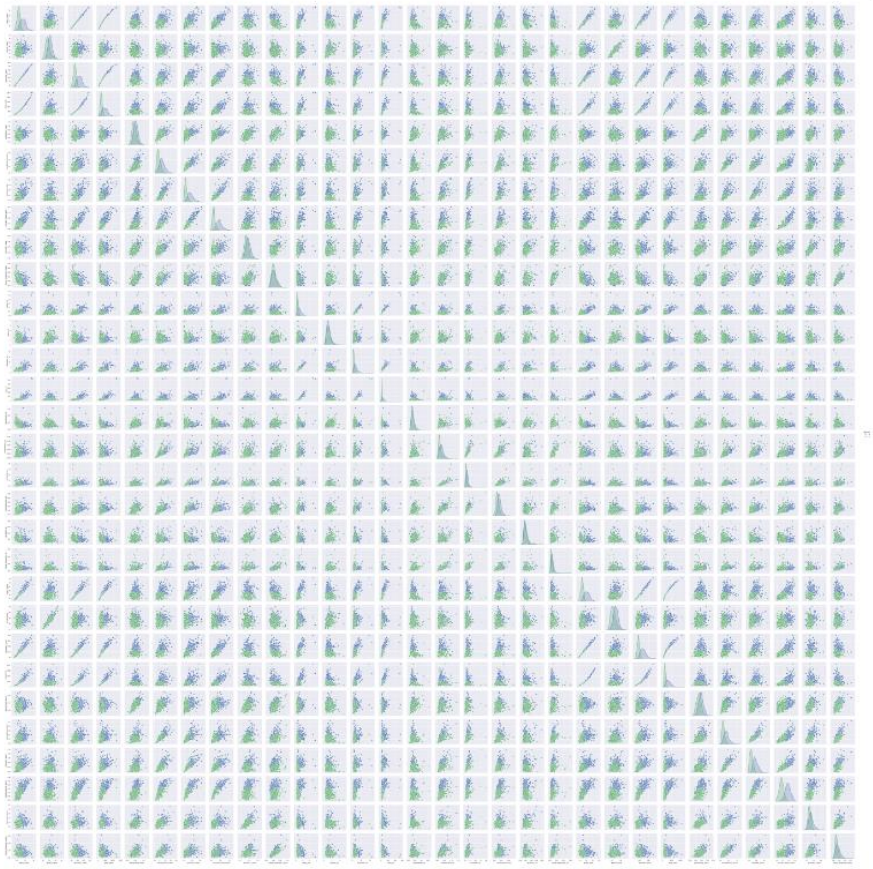


FEATURE EXTRACTION



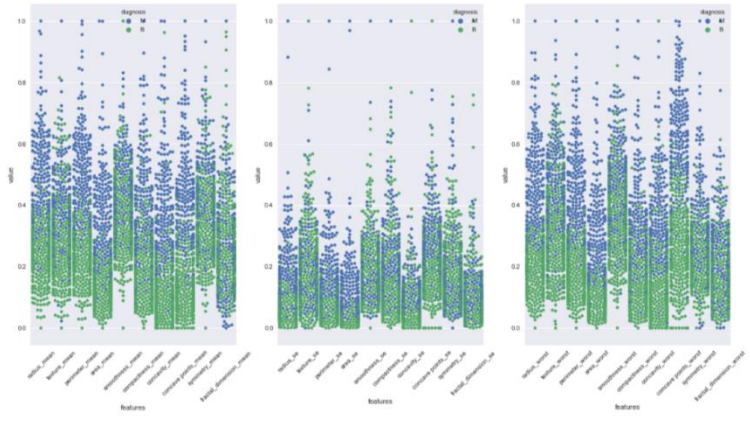
DATA SELECTION

In [13]:
```
sns.pairplot(data=df, hue='diagnosis');
```



COMPARING DIAGONISIS FEATURES

```
plt.figure(figsize=(20, 10))
plt.subplot(1, 3, 1)
sns.swarmplot(data=dfs[0], x='features', y='value', hue='diagnosis')
plt.xticks(rotation=45);
plt.subplot(1, 3, 2)
sns.swarmplot(data=dfs[1], x='features', y='value', hue='diagnosis')
plt.xticks(rotation=45);
plt.subplot(1, 3, 3)
sns.swarmplot(data=dfs[2], x='features', y='value', hue='diagnosis')
plt.xticks(rotation=45);
```



In [11]:

```
plt.figure(figsize=(20, 10))
plt.subplot(1, 3, 1)
sns.pointplot(data=dfs[0], x='features', y='value', hue='diagnosis', linestyles="", dodge=0.3)
plt.xticks(rotation=45);
plt.subplot(1, 3, 2)
sns.pointplot(data=dfs[1], x='features', y='value', hue='diagnosis', linestyles="", dodge=0.3)
plt.xticks(rotation=45);
plt.subplot(1, 3, 3)
sns.pointplot(data=dfs[2], x='features', y='value', hue='diagnosis', linestyles="", dodge=0.3)
plt.xticks(rotation=45);
```