**Table of Content**

1. Tumor
2. Types of Tumors
3. Breast Cancer Classification
4. Dataset
5. Workflow
6. Calculations by Algorithm

• Accuracy

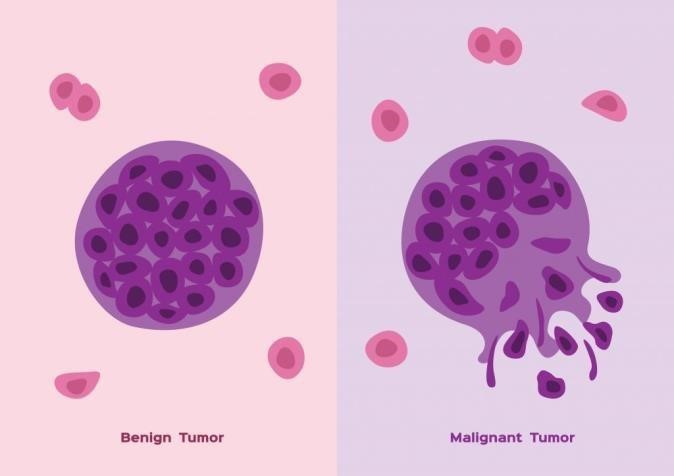
1. Libraries used

# 1. TUMOR:

Human body is made up of cells, tissues and organs etc. All the cells in our body divides and grows and sometimes what happens is some of cells in our body may divides repeatedly without some control. In that case it forms an abnormal mask and abnormal tissues those abnormal tissues are referred as tumor.

# 2. TYPES OF TUMOR:

1. Benign
2. Malignant



**BENIGN TUMORS:**

* Benign tumors are those which do not move to the other parts of body
* They are not as much harmful called as non-cancerous tumors  Slow growing

**MALIGNANT TUMORS:**

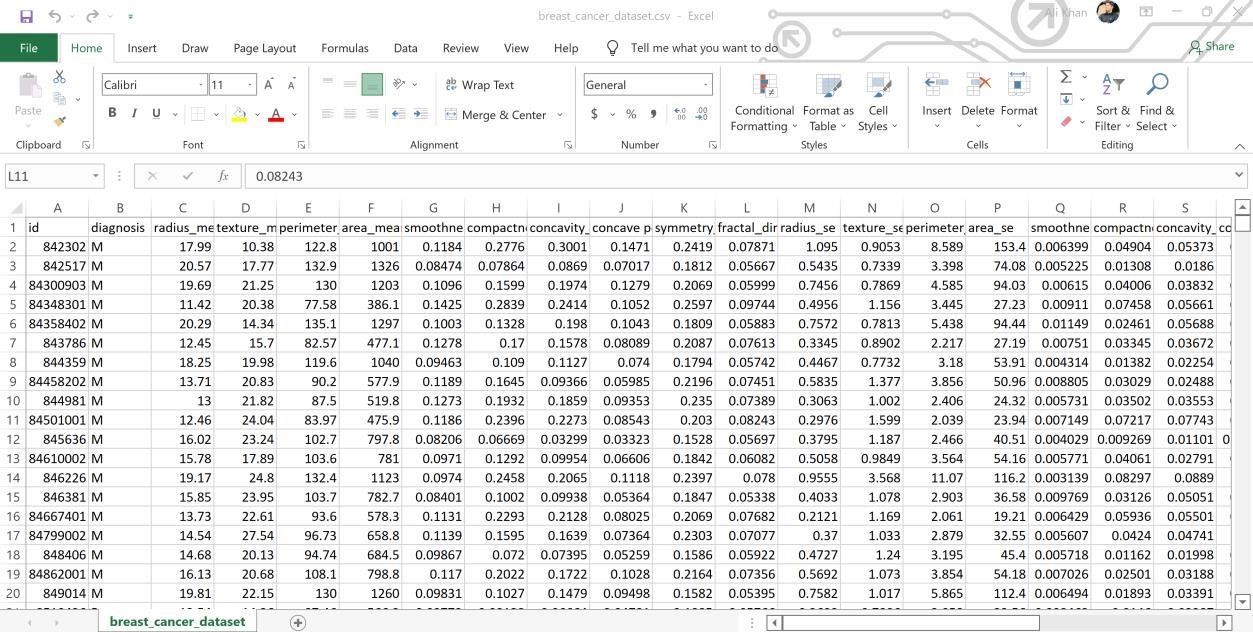
* Malignant tumors are those which have capability to move to the other parts of the body
* They are dangerous and called as cancerous tumors
* Fast growing

# 3. BREAST CANCER CLASSIFICATION:

This algorithm will classify the tumors as malignant or benign. That’s why it is known as Breast Cancer Classification. Logistic Regression model will be trained on the dataset and then it will be tested on the new data.

# 4. DATASET:

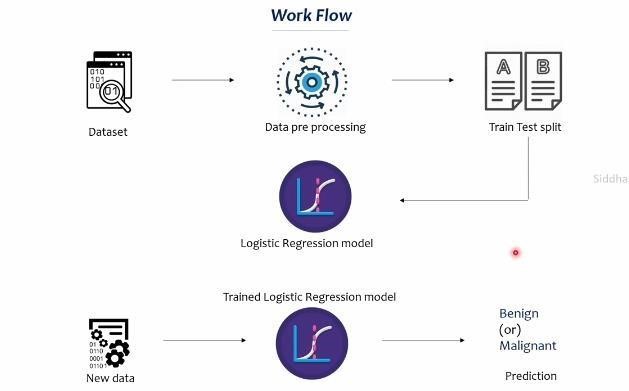
Fine needle aspiration: It is a type of biopsy procedure. In fine needle aspiration, a thin needle is inserted into an area of abnormal appearing tissue or body fluid. As with other types of biopsies, the sample collected during fine needle aspiration can help make a diagnosis or rule out conditions such as cancer. The data we use has been derived from this particular test called Fine needle aspiration. This is a standard procedure.



**5.**

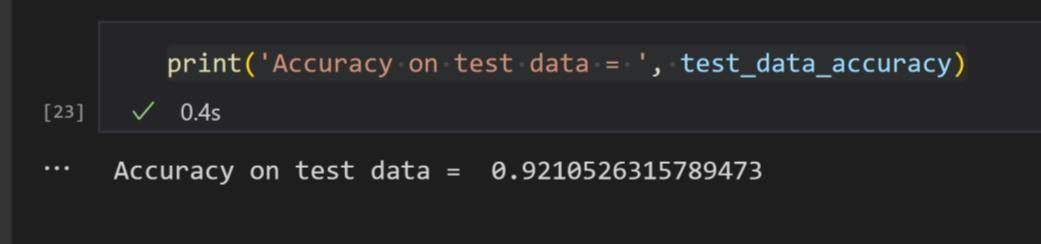
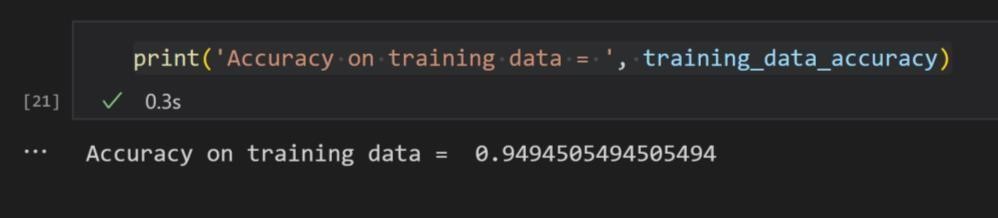


**WORKFLOW:**



1. **CALCULATION BY ALGORITHM**

• **Accuracy**



1. **ABOUT LIBRARIES:** 
   * **import numpy as np** (Used to make numpy arrays)
   * **import pandas as pd** (Used to create pandas dataframe, which are helpful to analyze the process data in more structured way)
   * **import sklearn.datasets**(Used to import the breast cancer data)
   * **from sklearn.model\_selection import train\_test\_split**(Splits the data into training and testing part)
   * **from sklearn.linear\_model import LogisticRegression**(Logistic regression is used because we have binary decision)
   * **from sklearn.metrics import accuracy\_score**(Used to evaluate our model i.e. how many correct predictions our model is making)