

**MACHINE LEARNING FILE**

**SUBMITTED BY-**

**BHAVNOOR SINGH**

**01813202717**

**CSE-1**

**Github link for code :**

<https://github.com/Bhavnoor-Singh98/ML_LabWork/blob/main/Decision%20tree.ipynb>

**Question:** Study and implement the decision tree classifier on a breast cancer dataset using Scikit learn in python.

**Breast-Cancer-Prediction-using-decision-tree-algorithm**

The proposed data mining model is used to help medical researchers to analyse and predict the breast cancer cells in human body by analysing the breast cancer dataset and thus determine if the resultant cells are cancerous or not.

**Introduction:**

Breast cancer has become a very common phenomenon among women in these years and is the main reason of death among women in different age groups. Several reasons that may cause breast cancer includes age, genetics, dense breast tissues, alcohol consumption, radiation exposure and many as such. Breast cancer can be viewed in many stages for which numerous treatments can be prescribed depending upon the level and stages of breast cancer, but these treatments may vary depending upon the stage, cause and type of the cancer. When cells divide it causes cancer which then changes from a normal cell to what’s called a cancerous cell that requires gene alterations. Therefore, predicting whether the cells are cancerous has now become the need of the hour.

**DATASET DESCRIPTION:**

Dataset name: Wisconsin Diagnostic Breast Cancer (WDBC)database from the UC Irvine Machine Learning Repository. No of records(tuples): 699 records No of attributes:11 attributes

**Problem Statement:**

The proposed data mining model is used to help medical researchers to analyse and predict the breast cancer cells in human body by analysing the breast cancer dataset and thus determine if the resultant cells are cancerous or not. The proposed model tends to exhibit the use of Decision tree classification algorithm implementation using Jupyter Notebook in Python to visualize the data and find out the optimal accuracy of the model.

**Output**

The decision tree algorithms employed in python resulted into an efficient model with accuracy 94.16% which can very evidently state whether a particular cell is benign or malignant.