Follow the below instructions and perform the tasks.  
  
Data for the assignment can be fetched using below code.  
  
*import pandas as pd*

*data = pd.read\_csv("https://raw.githubusercontent.com/AIP-BITS/BITS-DATA/main/cancer.csv")*

1) Basic Modelling [Marks:3]  
  
Do basic data preprocessing, EDA and train any default classifier as a baseline.  
  
2) Train a Decision Tree and fine tune the model using following   
[Marks:8]

* PCA to orient the data better
* Regularize the model using various hyperparameters.
* Serch for best hyperparameters using grid/random search

3) Train a Naive Bayes classifier and fine tune the model using following   
[Marks:8]

* Try different variations of naive bayes available in scikit learn and pick the default (without any hyper-parameter tuning) one that perform best on evaluation metrics.
* Search for best hyperparameters using grid/random search

4) Train a SVM classifier and fine tune the model using following   
[Marks:8]

* Train SVM classifier for various tunable parameters and store the   
  results (Accuracy, precison, recall,training time and testing time) for   
  each combinations in a dataframe.
* Search for best hyperparameters using grid/random search

[Marks:3]  
  
A model needs to be selected that gives the best recall and reasonable precision and F1 Score and the fastest prediction time.  
  
Select the parameters and retrain the final model from all the options available above and plot its confusion matrix  
Convert the notebook to html format and upload.   
Submit ONLY HTML format of the notebook.