Machine learning-Assignment3

Student id:700674089

Student Name: Jahnavi Danda

CRN: CS571013466

**GitHub Repository**

https://github.com/jxd40890/Machine-learning-Assignment3.git

**Assignment1 video URL**

https://drive.google.com/drive/folders/1e8A2F0D4Kp\_OPEdVWr4JazHXO208oTiS?usp=sharing

Question1

1. Find the correlation between ‘survived’ (target column) and ‘sex’ column for the Titanic use case in class.



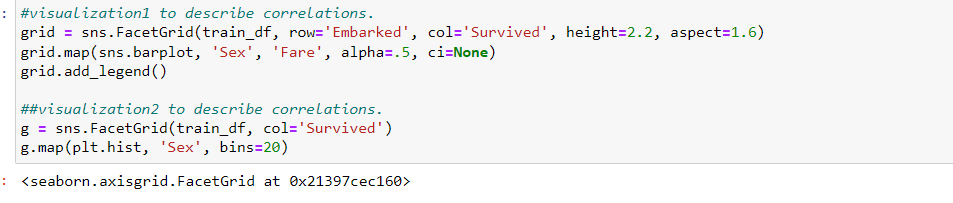
1. Do you think we should keep this feature?

**Answer**

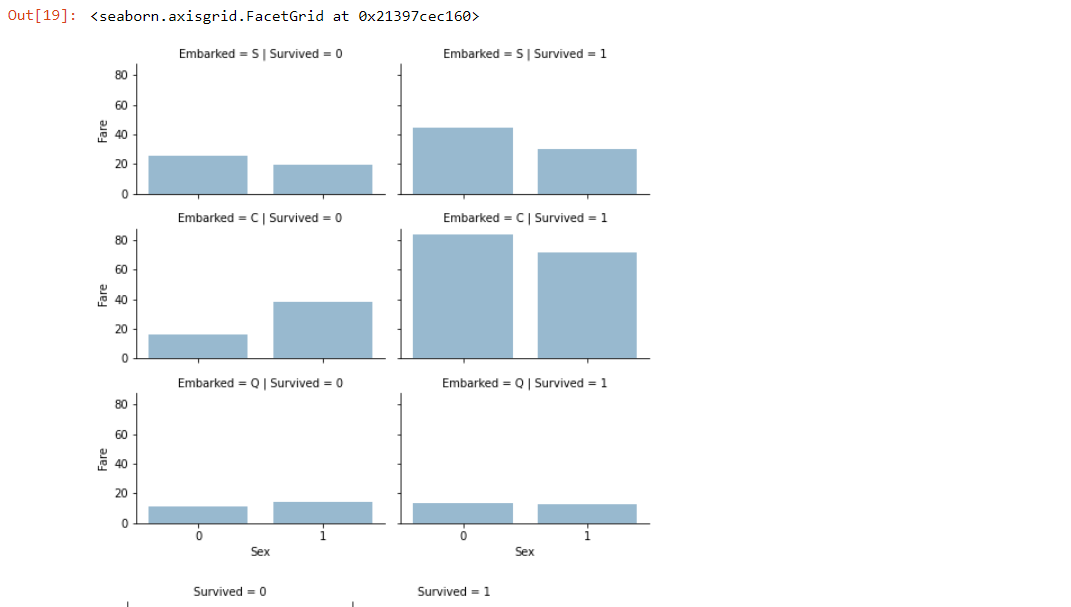
We don’t have to keep this feature since the correlation is Negative

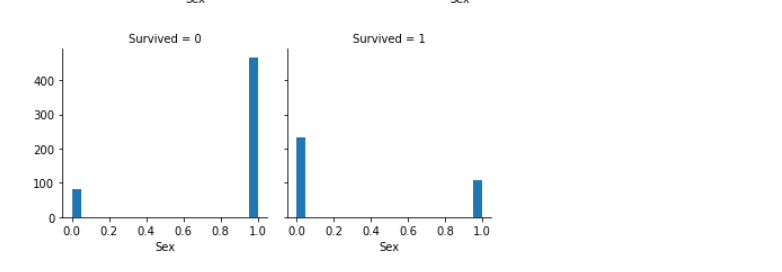
1. Do at least two visualizations to describe or show correlations.

**Code Screenshot**



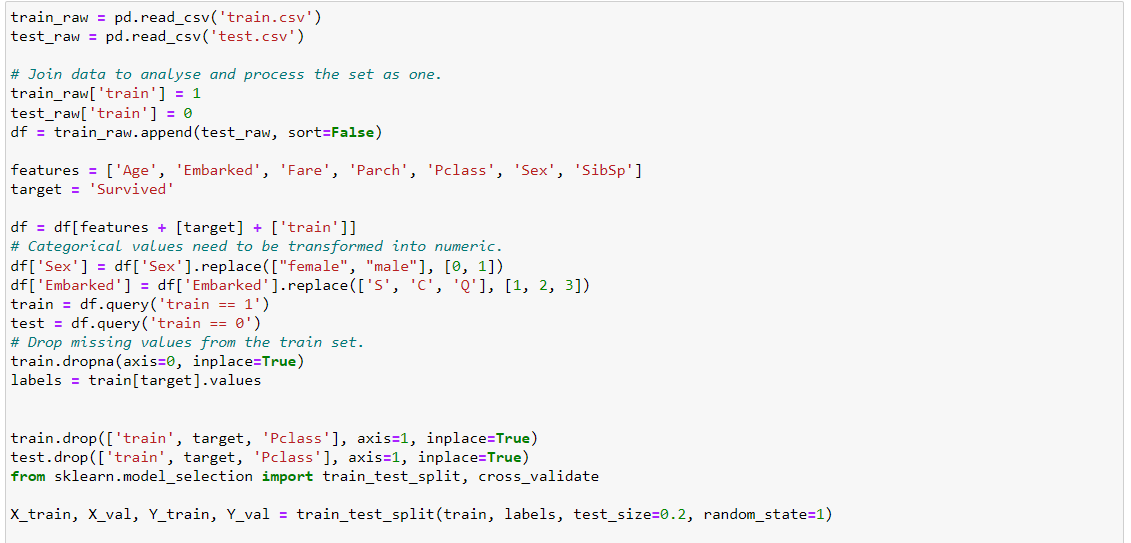
**Output screenshot**

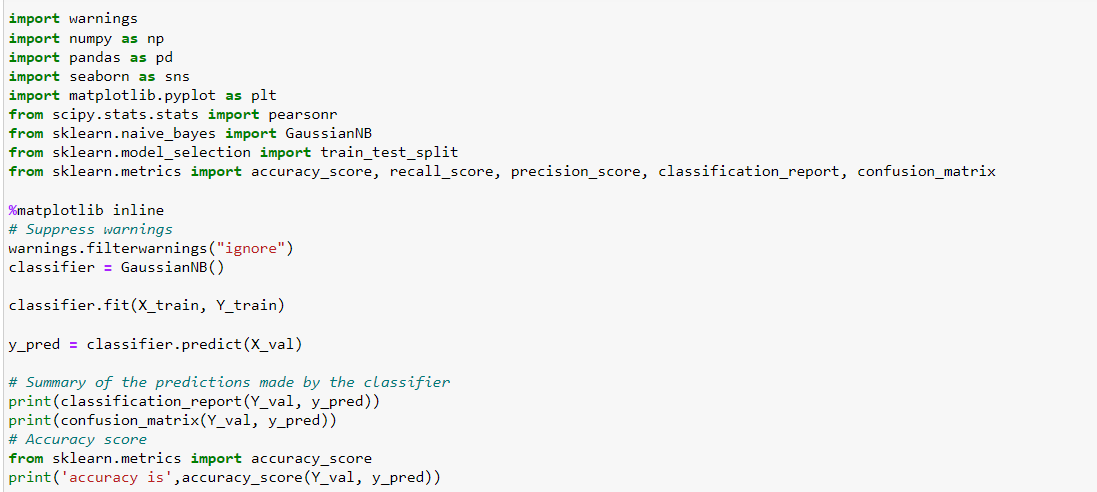




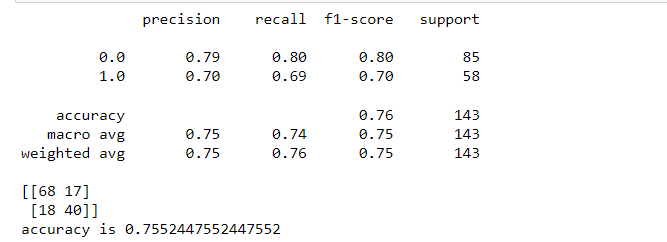
1. Implement Naïve Bayes method using the scikit-learn library and report the accuracy

Code Screenshot





Output Screenshot



Question2: -

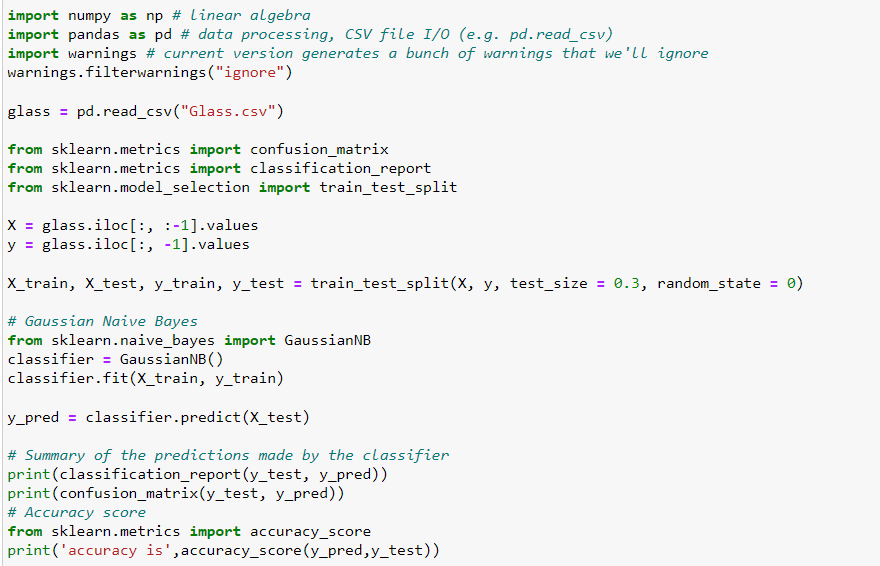
1. Implement linear SVM method using scikit library

a. Use the glass dataset available in Link also provided in your assignment.

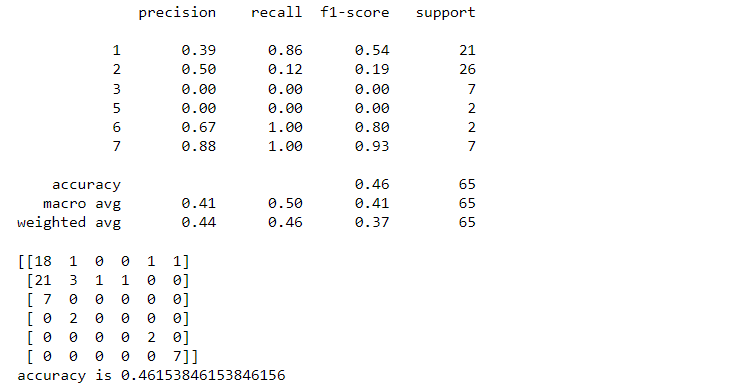
b. Use train\_test\_split to create training and testing part.

2. Evaluate the model on testing part using score and

**Code screenshot**



**Output screenshot**

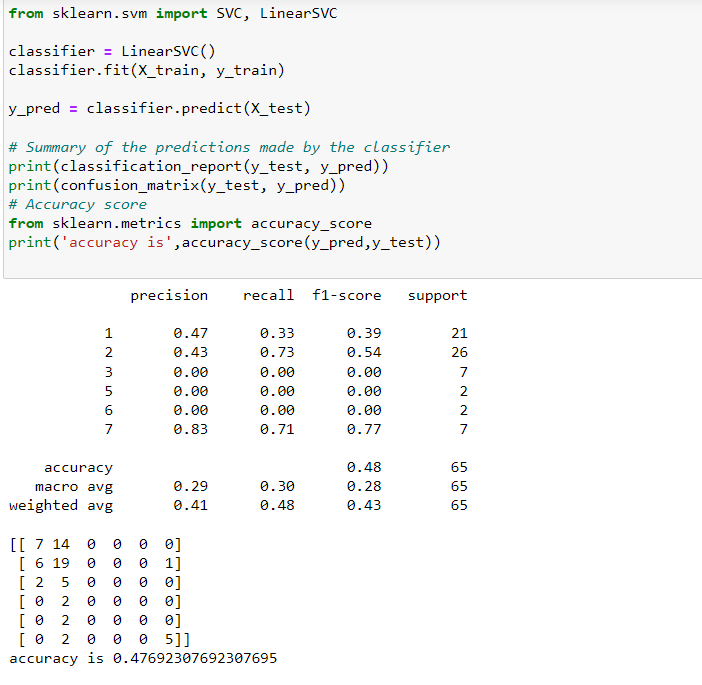


Question3: -

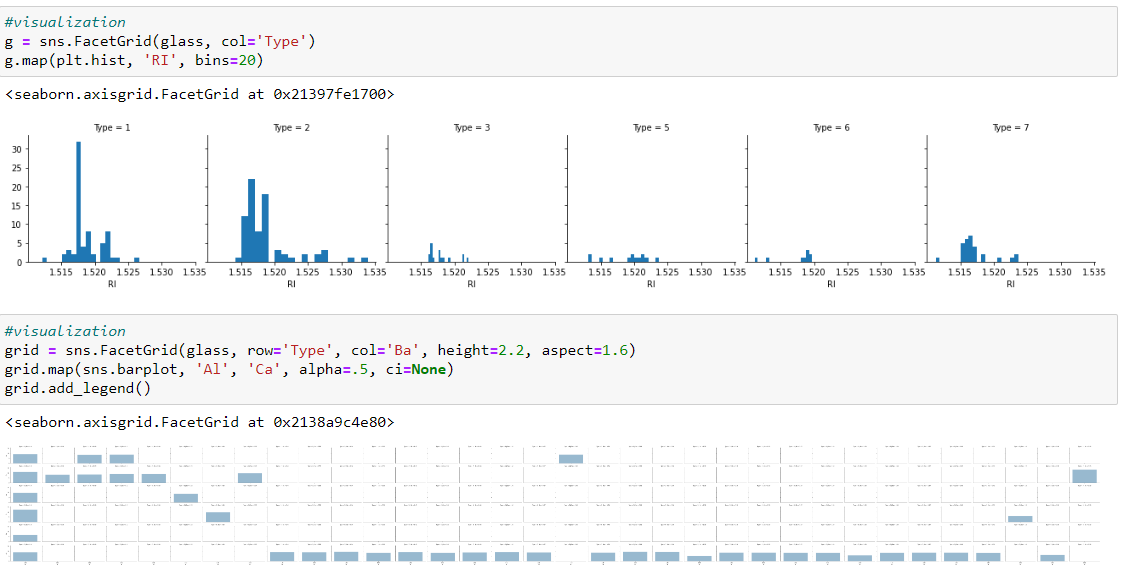
1. Implement linear SVM method using scikit library a. Use the glass dataset available in Link also provided in your assignment. b. Use train\_test\_split to create training and testing part.

2. Evaluate the model on testing part using score and classification\_report(y\_true, y\_pred) Do at least two visualizations to describe or show correlations in the Glass Dataset

**Code Screenshot**



**Output Screenshot:**



**Which algorithm you got better accuracy? Can you justify why?**

Naïve Bayes is giving more accuracy many times when compared to svm since the dataset is small.

NYB is better at snippets than long documents.

And also based on kfold cross validatation NYB is showing better mean than SVM.

