**ICP6 – Machine Learning Algorithms**

**Algorithm** – Classification/Naive Bayes

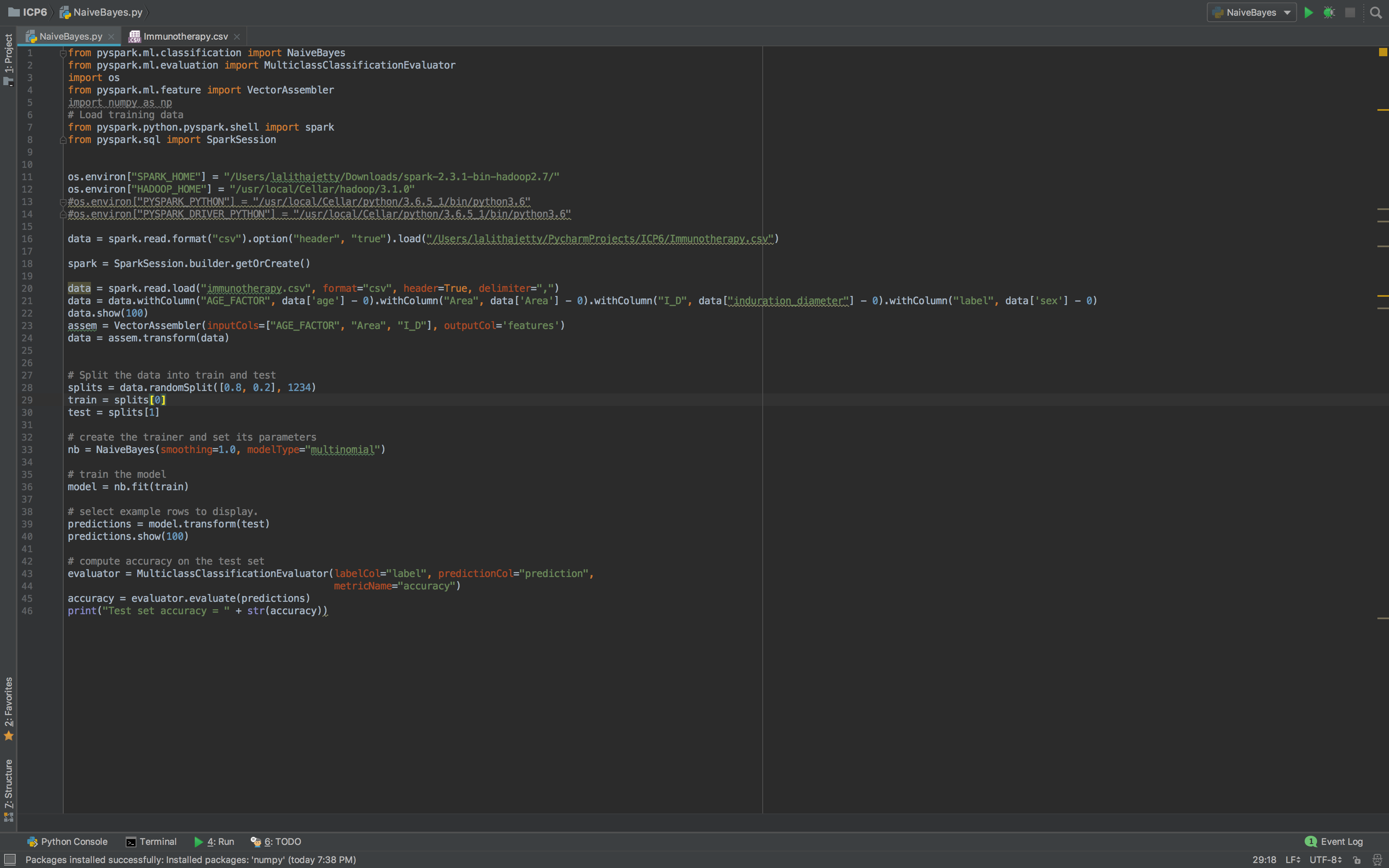
**Data Set** - Immunotherapy dataset

**Columns** - sex, age, Time, Number\_of\_Warts, Type, Area, induration\_diameter, (7 - features) Result\_of\_Treatment(label)

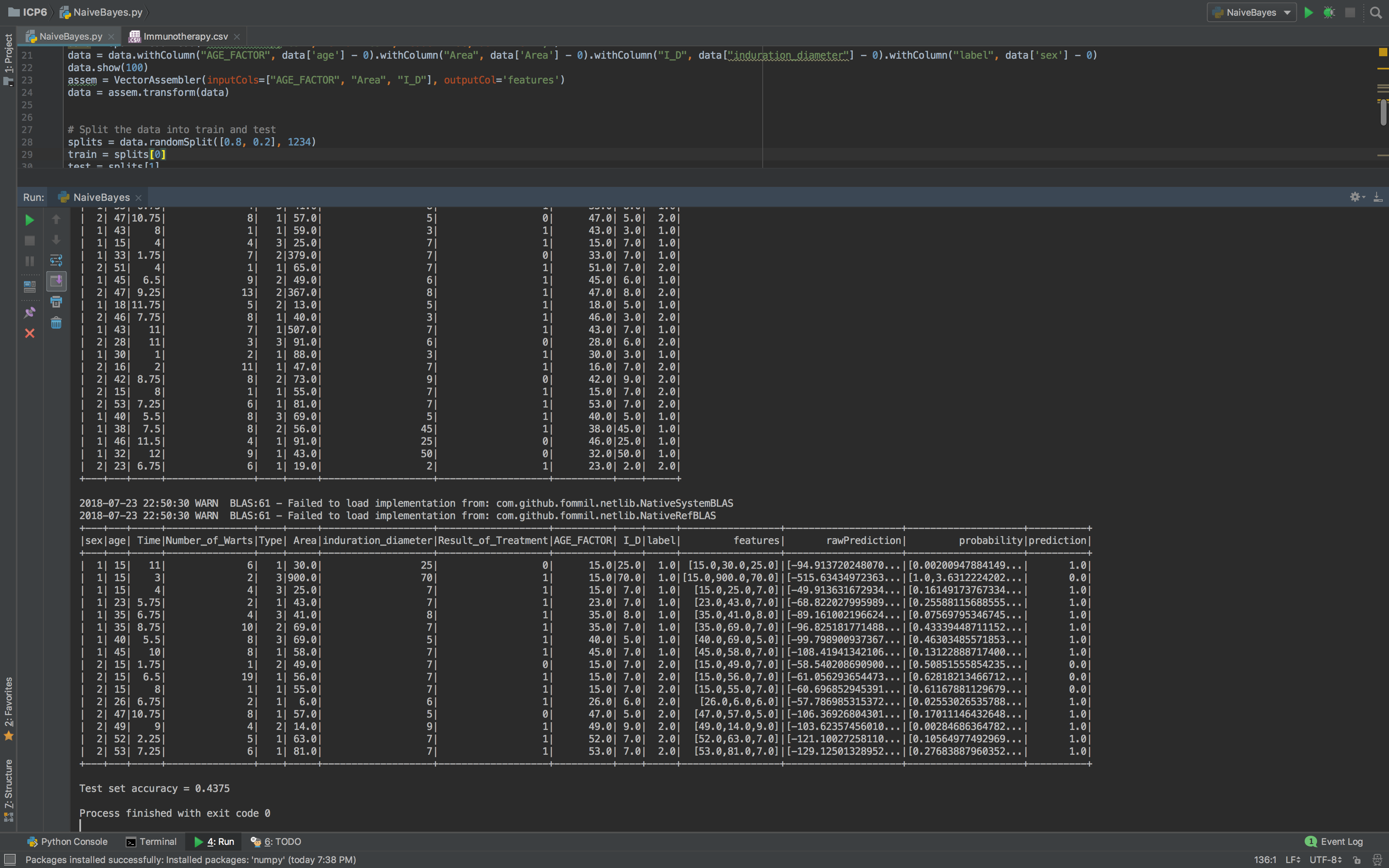
This dataset contains information about wart treatment results of 90 patients using immunotherapy. This dataset has 7 different features like sex, age, time etc., and one label which has two categories (wart treatment curing the disease = 1, wart treatment not curing the disease = 0)

In the below code we have only considered three features sex, age and induration\_diameter features to predict the Result\_of \_Treatment. And the cross-validation parameter is set to 80-20 which means 80% of the data = 72 samples are used to train the algorithm and 20% of the data = 18 samples are used to test the model. Random function is used to divide the samples to training and testing. Then the accuracy is calculated which is correct predictions in test set/total predictions in test set.

Code :



Output:



Test set accuracy = 0.4375