Artificial Intelligence with Python

Lab Task 2

The Census Income Data Set (Income_data.txt) presents a challenge of predicting whether income exceeds $50K/yr based on census data. The data is stored in Income_data.txt file. There are two distinct classes namely '<=50K' and '>50K'. Apply necessary pre-processing steps on this data and state them clearly. Note that there should not be more than 70% data in train. The model has to be built which will accurately classify the test data. Use various classification techniques and indicate the results in tabular form. Perform a benchmarking analysis of the results.