Ryan Timbrook

Description:

Design a simple, lowcost sensor that can distinguish between red wine and white wine.
The sensor must correctly distinguish between red and white wine for at least 95% of the samples in a set of 6497 test samples of red and white wine.
Your technology is capable of sensing the following wine attributes: Fixed acidity Free sulphur dioxide Volatile acidity Total sulphur dioxide Citric acid Sulphates Residual sugar pH Chlorides Alcohol Density

Tasks:
1.Read WineQuality.pdf.
2.Use the RedWhiteWine.csv/arff file that is provided.
Note: If needed, remove the quality attribute, which you will not need for this assignment.
3.Build an experiment using Decision Trees.

Questions:
1-What is the percentage of correct classification results, using all attributes?
        Answer: Using the entire dataset's 11 attributes yields an accuracy score of %97.8
2-What is the percentage of correct classification results, using a subset of attributes?
        Answer 2.1: Limiting the dataset to 6 attributes yields an accuracy score of 96.1%
        Attribute List: fixed.acidity,volatile.acidity,citric.acid,residual.sugar,chlorides,density
        Answer 2.2: Limiting the dataset to 5 attributes yields an accuracy score of 96.2%
        Attribute List: fixed.acidity,citric.acid,residual.sugar,chlorides,density
        Answer 2.3: Limiting the dataset to 4 attributes yields an accuracy score of 95.5%
        Attribute List: fixed.acidity,citric.acid,residual.sugar,chlorides
        Answer 2.4: Limiting the dataset to 3 attributes yields an accuracy score of 95.7%
        Attribute List: fixed.acidity,residual.sugar,chlorides

        *BELOW ACCEPTANCE CRITERIA of 95% Accuracy - >
        Answer 2.5: Limiting the dataset to 2 attributes yields an accuracy score of 92%
        Attribute List: fixed.acidity,residual.sugar

3-What is the AUC of your model?, What is the best AUC achieved?
        Answer: The AUC of the heighest accuracy score test was 98%

4-Which are the minimum number of attributes, why?

Answer: The minimum number of attributes that yielded a greater than 95% accuracy score was 3. They were (fixed.acidity,residual.sugar,chlorides)