

Opposition Report on

Identification the quality of Wine

Bala Kiran M

h19baman@du.se

Data Science

Kiran Sai T

h19kirta@du.se

Data Science

Kailash Chowdary B

h19kaibo@du.se

Data Science

I. INTRODUCTION

The introduction was simple, and the main idea of the report has been explained. In Literature review there was no statements about the past studies which would explain about the dataset. The literature review just explains about the process which have been taken in the current report which is irrelevant in the Literature review. There was no clear introduction about the methods which have been used in the report.

II. DESCRIPTION OF DATA

The dataset has been explained in overview. There was an introduction about the features in the dataset. The explanation lacks with the clear overview on the features in the dataset. For a normal reader it would be difficult to understand the features from the explanation as there was a lack of description about the main features.

III. METHODS

The methods which have been used has been explained but there is a lack of information that how the data is been cleaned? How the null values have been handled? On what features the data is been analyzed to understand?

Selecting Relevant feature has been cleared explained but rather than selecting the features from the correlation plot they could choose PCA model for better selection of the models.

The implications of the model and the relevant things have a clear explanation for understanding the methods which have been taken in the report. The starting of the Methods says that there are four different steps to go through, but they have been explained only three.

IV. LEARNING CURVE

There is a clear explanation about the models which have been used for the analysis but there is no clear explanation about the training data and testing data how they have been split. There is no clear definition of the methods and terms which have been used and how there are been used in the report. In overall all the methods should be explained in the methods rather in the learning curve.

V. RESULTS

The correlation plot on the features is unable to understand and draw conclusions from the plot. It is very hard to understand the important features and dependencies of the features from the plot. There is no clear explanation about the features that how they are correlated with each other on the heatmap. There is no clear explanation about how the features would be affecting the quality.

The predictions and the confusion matrix say 0 and 1 but there is no state of explanation of the labels and how they are related.

In the results we have learning curves but there is no clear explanation about the axis labels and the measures. A normal reader would be having difficulties to understand the learning curves as there is no clear explanation about the labels and the intervals and how they have been measured.

The conclusion drawn from the learning curves are not so clear and unable to understand. We were not able to understand meaning about the training examples and the scores.

VI. DISCUSSION

The points mentioned in the discussion are not clearly explained and could not be related. The content mentioned is not been correctly related. There are conclusions mentioned in the discussion rather continuing the topic for the future.

VII. REFERENCES

The quality of the references is not been correctly mentioned. The links which have been given were not working and could not relate with the content.

VIII. OVERALL IMPRESSION

As a team we felt they had a very good idea of the dataset and the methods which needs to perform. Like using the SVM method for the better results. We felt it was a decent attempt overall.

They were some points where they were lacking. There is no clear explanation about the methods and the features which have been used. The IEEE format has not been correctly used and the references are not been correctly mentioned. There were grammatical errors and spelling mistakes in the report. There was no clear explanation about the data. The figures were not correctly explained, and irrelevant content has been filled at places.