THE UNIVERSITY OF HONG KONG

MSc in E-Commerce and Internet Computing

ECOM7126 Machine Learning for Business and E-Commerce
(2024-25)

Assignment 3 – Wine Analysis (Unsupervised Learning)

Wattson Wine is a major wine importer in Hong Kong and is the distributor of many suppliers of quality wine from around the world. Wine quality traditionally has been judged subjectively by wine experts (called connoisseur). The company management wishes to provide more objective assessment of wine quality to supplement the subjective judgment of the experts. The company requires the suppliers to provide laboratory certified information on the characteristics of the 1,500 brands of wine that it represents. The following is the list of the properties/attributes of each of the 1,500 brands (labelled with a sample number):

(1) acidity
(2) citric acid
(3) sugar
(4) chlorides
(5) density
(6) pH level
(7) sulphates
(8) alcohol

The dataset consists of the properties data of the 1,500 different brands in .csv format.

You are asked to analysis the data provided by the laboratories and give insights to the management of Wattson Wine how such data can be used to assist quality control and wine pricing for their wine.

Deliverable:
1. A report to the management in PDF format.
2. The Colab notebook (Python programs with comments and notes) that you use to produce your results in `.ipynb` (Colab notebook) format.

You should include the following in your report:

1. How you tackle the dataset and what you want to find out from the data.
2. The analyze of the data, results and how you achieve your objectives.
3. What you have learned from the data and suggestions/recommendations, if any, to the management as based on the results of your analysis.

Dataset provided: Wine_Dataset.csv

Deadline: 13 April 2025