

## BÀI TẬP KIỂM TRA BUỔI 2 – KNN & Bayes thơ ngây

- a. Đọc dữ liệu từ tập dữ liệu đánh giá chất lượng rượu vang **đỏ** trên trang UCI  
<https://archive.ics.uci.edu/ml/datasets/wine+quality>
- b. Tập dữ liệu có bao nhiêu phần tử, Có bao nhiêu nhãn? Ghi chú kết quả trong file code sau câu lệnh.  
len(tên biến) : số lượng phần tử  
import numpy as np  
np.unique(tên biến) : liệt kê các giá trị khác nhau của biến  
tên biến.value\_counts() : liệt kê số lượng và giá trị khác nhau của biến
- c. Sử dụng **6 phần** để xây dựng mô hình và sử dụng **4 phần** để thực hiện đánh giá mô hình. Anh/chị ghi chú lại số lượng phần tử trong tập test và nhãn của các phần tử thuộc tập test trong file code
- d. Xây dựng mô hình KNN với 7 láng giềng dựa trên tập dữ liệu học tạo ra ở bước c.
  - i. Đánh giá độ chính xác tổng thể và độ chính xác của từng lớp cho toàn bộ dữ liệu trong tập test
  - ii. Đánh giá độ chính xác tổng thể và độ chính xác của từng lớp cho **8** phần tử đầu tiên trong tập test
- e. Xây dựng mô hình Bayes thơ ngây, đánh giá độ chính xác tổng thể và độ chính xác cho từng phân lớp
- f. Với nghi thức hold-out (2/3 để học, 1/3 để kiểm tra), so sánh độ chính xác tổng thể của mô hình KNN và Bayes thơ ngây