**Lab02 – Data Mining –Missing Tutorials**

**Bài tập về nhà** :

* Đọc dữ liệu Titanic.csv

**-** Giải quyết vấn đề Missing Data trên tập dữ liệu này và chạy thử các mô hình máy học so sánh các cách khử missing value theo bài mẫu “Lab02 - Missing Tutorials.html”

**Mô tả : Titanic**

pclass: A proxy for socio-economic status (SES)  
1st = Upper  
2nd = Middle  
3rd = Lower  
  
age: Age is fractional if less than 1. If the age is estimated, is it in the form of xx.5  
  
sibsp: The dataset defines family relations in this way...  
Sibling = brother, sister, stepbrother, stepsister  
Spouse = husband, wife (mistresses and fiancés were ignored)  
  
parch: The dataset defines family relations in this way...  
Parent = mother, father  
Child = daughter, son, stepdaughter, stepson  
Some children travelled only with a nanny, therefore parch=0 for them.

**Ghi chú:**

Bài thực hành : DM Lab02

Tiêu đề Email : MSSV\_HoTenSinhVien\_ DM Lab02

Gửi mail : [tuantran261083course@gmail.com](mailto:tuantran261083course@gmail.com" \t "_blank)

Nội Dung Email :

- Ghi đã làm đầy đủ hay thiếu bài gì, chức năng gì

- Ghi có làm thêm được gì hay không (khuyến khích làm thêm để cộng điểm)

- Nếu nộp trễ ghi lý do nộp trễ

Dealine : 17/04/2022

File nộp : File MSSV\_HoTenSinhVien\_ DM Lab02.pdf chụp hình code và chụp hình kết quả chạy (không cần gửi code)