**FINAL EXAMINATION - DATA MINING**

**DEADLINE 15/08/2021**

**1. Yêu cầu.** Nộp bài cho lớp trưởng tổng hợp trước ngày Deadline

**2.** **Bài nộp**. PDF(Convert từ ipynb Code) + IPYNB file + Word ghi chú nhận xét đánh giá từng bước làm Mining Data (nếu bước có cần làm hay không thì nêu lý do) và ít nhất 3 thuật toán máy học dùng để so sánh kết quả phân loại hay dự đoán

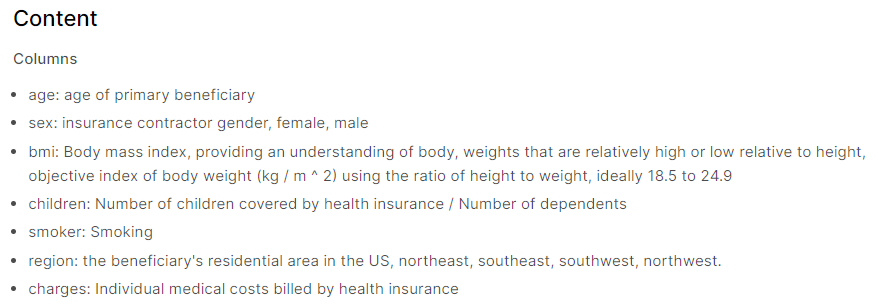
**3.** **Tối đa 2 người 1 nhóm.**

**4. Các bước Mining cần tiến hành**.

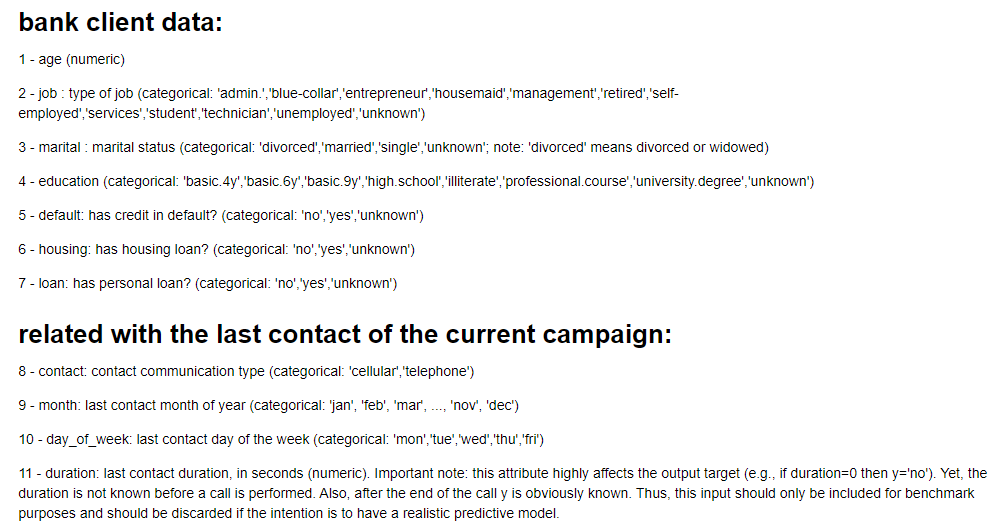
* 1. Data Integration
  2. Data Cleansing
  3. EDA
  4. Encoding
  5. Feature Scaling
  6. Feature Selection (Feature Importance)
  7. Model Building
  8. Model Evaluation

**5. Chọn 1 trong 2 bài sau để làm**

**5.1. Bài toán dự đoán.**  insurance.csv target : charges



**5.2. Bài toán phân loại**. bank-full.csv target : y - has the client subscribed a term deposit? (binary: 'yes','no')



---------------------------------------------------------------GOOD LUCK -------------------------------------------------------