Mẫu: 2a\_ĐTTL

## TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HÒ CHÍ MINH

## KHOA/VIỆN CNNT

## 02-ĐỀ THI CUỐI KỲ

Môn thi: Phân Tích Dữ Liệu

Lớp/Lớp học phần: DHHTTT15B

Ngày thi: 19/06/2021

Thời gian làm bài: 60 phút (Không kể thời gian phát đề)

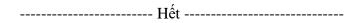
Câu 1: (LO1) (5 điểm) Hãy so sánh dữ liệu định tính và dữ liệu định lượng.

<u>Câu 2: (LO2)</u> (5 điểm) Sinh viên sử dụng dulieuxettuyendaihoc.csv để xây dựng mô hình dự báo kết quả điểm thi đại học môn 2 (**DH2**) dựa trên mô hình hồi quy tuyến tính, với tỉ lệ tập huấn luyện và kiểm tra là 80% và 20%. Biết rằng đề dự báo thì người ta dựa trên các thông tin: trung bình điểm số trong các học kì của học sinh các năm trung học phổ thông Hãy trả lời các câu hỏi và **trình bày ra giấy**:

- a. Hãy mô tả mô hình dự báo
- b. Hãy phân tích mô hình hồi quy tuyến tính

Họ và tên thí sinh .....

c. Hãy giải thích giá trị RMSE của mô hình trên tập huấn luyện và tập kiểm tra.



Lưu ý: - Giáo viên ra đề và trưởng bộ môn ký duyệt vào mặt sau của đề.

- Đề thi (được) sử dụng tài liệu.
- Sinh viên phải **nộp bài làm và source code**

Cán bộ coi thi không giải thích gì thêm.