

ỦY BAN NHÂN DÂN THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC SÀI GÒN
KHOA TOÁN - ỨNG DỤNG



BÁO CÁO
HỌC PHẦN: KHAI PHÁ DỮ LIỆU (858016)

Giảng viên hướng dẫn: Đỗ Như Tài

Sinh viên thực hiện: Trần Quốc Hoàng (3123580015)

Trần Chí Vỹ (3123580065)

Phan Trần Hữu Tấn (3123580042)

Lớp: DDU1231

Thành phố Hồ Chí Minh, tháng 10 năm 2025

TRƯỜNG ĐẠI HỌC SÀI GÒN
NHẬN XÉT VÀ CHẤM ĐIỂM CỦA GIẢNG VIÊN

Họ và tên giảng viên: Đỗ Nhu Tài

Challenge 1: Titanic - Machine Learning from Disaster

Nội dung nhận xét:

Điểm:

Bằng số:

Bằng chữ:

GIẢNG VIÊN CHẤM

(Ký, ghi rõ họ tên)

QUÁ TRÌNH THAM GIA

Tên - MSSV	Phân công	Tham gia (%)
Trần Quốc Hoàng - 3123580015 (nhóm trưởng)	Viết code, hỗ trợ làm slide	100%
Trần Chí Vỹ - 3123580065	Viết paper	100%
Phan Trần Hữu Tân - 3123580042	Làm slide, poster	100%

LỜI CẢM ƠN

Lời nói đầu tiên cho chúng em xin trân trọng gửi lời cảm ơn chân thành tới Giảng viên Đỗ Như Tài đã tận tình chỉ bảo chúng em trong suốt quá trình học tập trong lớp. Chúng em xin cảm ơn Trường Đại học Sài Gòn và Khoa Toán - Ứng dụng đã tạo điều kiện và cơ hội cho chúng em để có thể được tiếp cận và học hỏi thêm kiến thức mới mẻ này.

Dù đã rất cố gắng trong quá trình nghiên cứu và thực hành, nhưng do kiến thức và kinh nghiệm còn hạn chế, chắc chắn bài làm vẫn không tránh khỏi những sai sót. Em mong Thầy có thể thông cảm và xem xét bỏ qua cho chúng em.

BÁO CÁO

Yêu cầu: tham gia challenge, viết report báo cáo (báo cáo dạng Paper, Slides, Poster, code liên quan + minh chứng), hạn nộp là 02/11.

The screenshot shows a competition page on Kaggle. At the top, there's a search bar and a user profile icon. Below the header, the competition title is 'Titanic - Machine Learning from Disaster'. To the right of the title are buttons for 'Submit Prediction' and more options. The main content area has tabs for 'Submission and Description' and 'Public Score'. Under 'Submission and Description', a file named 'submission.csv' is listed with a green checkmark, indicating it's complete, and was submitted 3 minutes ago. The public score for this submission is 0.75837.

Nhiệm vụ:

- Tuần 1: tìm hiểu về thử thách, tham khảo code trên mạng, viết code.
- Tuần 2: tham khảo mẫu của thầy và viết Paper, làm Slides và Poster.

Ý nghĩa của các thư mục và file:

- Trong thư mục chính sẽ chứa 3 thư mục con là Data, Graphs, Reports, 1 file csv kết quả dự đoán và 2 file code (ipynb và html). Phần chính cần phải được coi là file code ipynb (khuyến khích) và thư mục Reports.
 - Thư mục Data: là nơi lưu trữ các data được sử dụng trong code.
 - Thư mục Graphs: là các hình trực quan có trong file code hoặc liên quan tới chủ đề.
 - Thư mục Reports: là nơi chứa các file pdf chính như Paper, Slides, Poster liên quan tới báo cáo.
 - File submission.csv: là file kết quả dự đoán được lưu từ code.
 - File code ipynb: file code chính, hiển thị được hết các hình trực quan.
 - File code html: file code dạng web, không hiển thị được hình 3d.