

Đồ án cuối kỳ

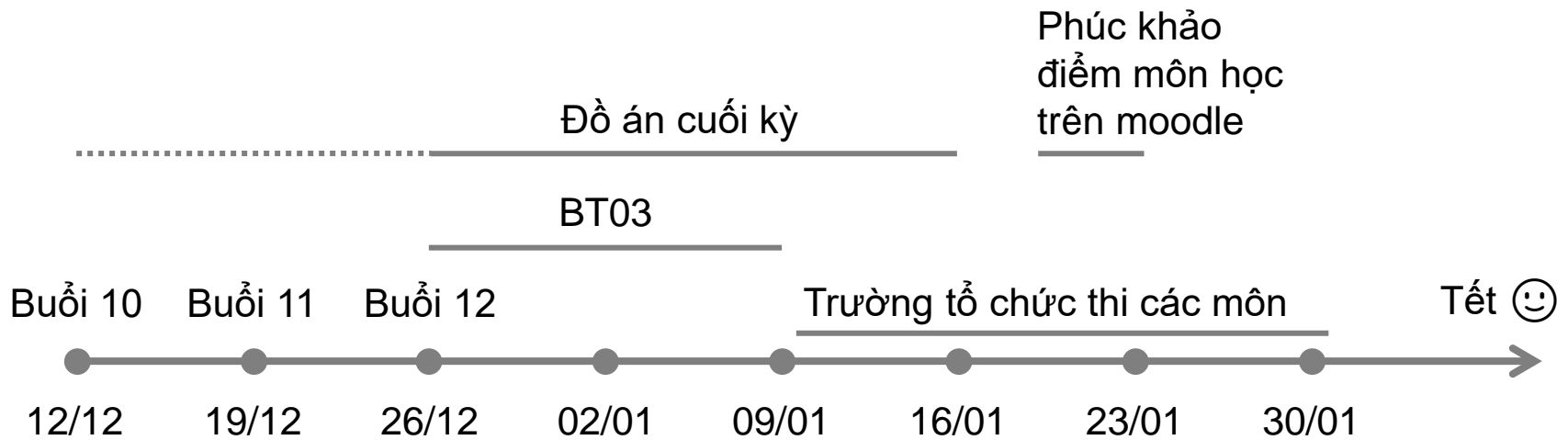
Trần Trung Kiên
ttkien@fit.hcmus.edu.vn
Cập nhật 11/12/2020



KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Kế hoạch hoàn thành môn học



Đồ án cuối kỳ

- ☐ Làm theo nhóm, 2 bạn / nhóm
 - ☐ Tất cả thành viên trong nhóm đều phải hiểu về đồ án của nhóm mình (tất nhiên là phải hiểu cả code)
 - ☐ Dùng Git & Github để quản lý phiên bản và hợp tác nhóm
- ☐ Áp dụng quy trình KHDL để trả lời cho một câu hỏi tự chọn
 - ☐ Câu hỏi phải có dạng: output được tính từ input theo công thức nào? (nghĩa là, phải mô hình hóa dữ liệu)
 - ☐ Phải tự thu thập dữ liệu (dùng web API, hoặc parse HTML)
- ☐ Các file cần nộp
 - ☐ File làm việc nhóm (viết file này khi nào?)
 - ☐ File dữ liệu (nếu nặng thì có thể up lên đâu đó và để link trong file txt)
 - ☐ File notebook
 - ☐ File slide (là bản tóm tắt của file notebook)

Đồ án cuối kỳ - các nội dung cần trình bày trong file notebook và file slide

- ☐ Thu thập dữ liệu
 - ☐ Khám phá dữ liệu
 - ☐ Đưa ra câu hỏi cần trả lời
 - ☐ Câu hỏi sẽ có dạng: output được tính từ input theo công thức nào?
 - ☐ Trả lời được câu hỏi thì có lợi ích gì?
 - ☐ Nguồn cảm hứng của câu hỏi: tự nghĩ hay tham khảo ở đâu?
 - ☐ Tránh chọn câu hỏi: giá nhà được tính từ các thông tin của nhà (diện tích, số phòng ngủ, ...) theo công thức nào?
- Vì đây là câu hỏi của BT03 😊

Đã có thể
bắt đầu
làm từ bây
giờ

Giả sử có nhóm làm như sau, bạn xem thử có vấn đề gì không?

- ☐ Thu thập dữ liệu về phim gồm các cột: đạo diễn, diễn viên, thể loại, doanh thu, điểm đánh giá
- ☐ Câu hỏi: điểm đánh giá được tính từ các thông tin còn lại theo công thức nào?
- ☐ Ý nghĩa: dùng công thức này để dự đoán điểm đánh giá của các phim mới ra để biết là có nên đi xem hay không

Khi dự đoán điểm đánh giá của một phim mới ra thì phim này đâu đã có thông tin về doanh thu

Giả sử có nhóm làm như sau, bạn xem thử có vấn đề gì không?

- ☐ Thu thập dữ liệu về thông tin nhà và giá trên một trang rao bán nhà (ai thích đăng sao cũng được)
- ☐ Câu hỏi: giá nhà được tính từ các thông tin nhà theo công thức nào?
- ☐ Ý nghĩa: với một ngôi nhà cần mua hoặc cần bán, dùng công thức này để có được giá tham khảo

Giá nhà trong dữ liệu thu thập không làm “correct output” được

vì ở trên trang rao bán nhà ai thích đăng giá sao thì đăng, những nhà này cũng không phải là những nhà đã được bán

Đồ án cuối kỳ - các nội dung cần trình bày trong file notebook và file slide

- ☐ Thu thập dữ liệu
- ☐ Khám phá dữ liệu
- ☐ Đưa ra câu hỏi cần trả lời
 - ☐ Câu hỏi sẽ có dạng: output được tính từ input theo công thức nào?
 - ☐ Trả lời được câu hỏi thì có lợi ích gì?
 - ☐ Nguồn cảm hứng của câu hỏi: tự nghĩ hay tham khảo ở đâu?
 - ☐ Tránh chọn câu hỏi: giá nhà được tính từ các thông tin của nhà (diện tích, số phòng ngủ, ...) theo công thức nào?
Vì đây là câu hỏi của BT03 😊
- ☐ Tiền xử lý dữ liệu
- ☐ Mô hình hóa dữ liệu
- ☐ Nhìn lại quá trình làm đồ án
 - ☐ Đã gặp những khó khăn gì? (hay mọi chuyện đều thuận lợi)
 - ☐ Có học được gì hữu ích? (hay không học được gì cả)
 - ☐ Nếu có thêm thời gian thì sẽ làm gì?
- ☐ Tài liệu tham khảo

Nên bắt đầu làm sớm, mỗi ngày làm một ít

Đồ án cuối kỳ - các mốc thời gian

- ☐ Gọi x là ngày vấn đáp (dự kiến: $x = 16/01/2021$)
- ☐ **Trước 9h sáng ngày $x-2$** : tất cả các nhóm sẽ upload đồ án lên Github, đồng thời mở một “issue” ở Github để thông báo là đã hoàn thành và nhờ các bạn góp ý
- ☐ **Trong ngày $x-2$** : mỗi nhóm (tất cả thành viên chứ không phải là một người đại diện) sẽ đọc và góp ý cho một nhóm khác (mình sẽ phân công sau) bằng chức năng “issue” của Github
- ☐ **Ngày $x-1$** : các nhóm hoàn thiện đồ án dựa trên góp ý của nhóm khác, cập nhật trên Github và nộp bài trên Moodle
- ☐ **Ngày x** : vấn đáp online, lịch cụ thể của từng nhóm sẽ được thông báo sau (mình sẽ để các bạn ký tên trước vào phiếu điểm, dự kiến là vào buổi 12, lúc đó chắc là học offline)