

Tóm tắt bảng kế hoạch đồ án môn Khoa học dữ liệu CLC2018

Nhóm 15: (Đề tài: Dự đoán kết quả trận bóng rổ)

18127154 – Võ Ngọc Minh

18127155 – Vũ Công Minh

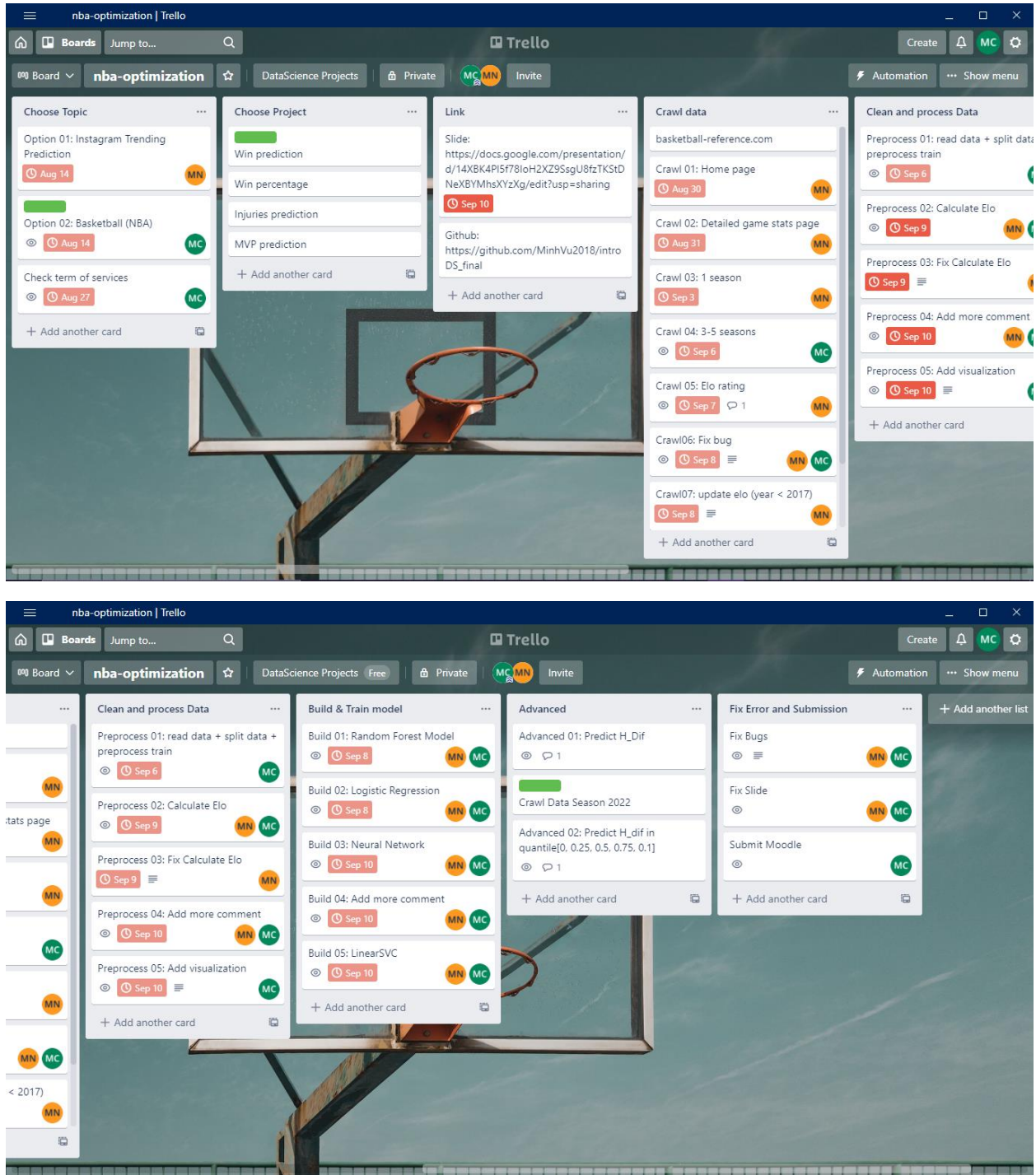
1. Phần mềm quản lý công việc: Trello

Lý do chọn: Giao diện đẹp, dễ thực hiện, có thông báo đến gmail (ngoài ra, còn có thể đổi giao diện thành timeline, calendar, map, ... nhưng chỉ dành cho Trello Premium)

2. Tóm tắt các mốc thời gian

- **Trước 14/08/2021:** Chọn đề tài Basketball, dự định lấy dữ liệu của giải đấu NBA (option 2 trong hình dưới)
- **14/08/2021 – 27/08/2021:** Chọn chủ đề (“Dự đoán Thắng/Thua”) + tìm các nguồn dữ liệu + kiểm tra điều khoản sử dụng
- **30/08/2021 – 31/08/2021:** Lấy dữ liệu của 1 trang HomePage và 1 trang chi tiết trận bóng
- **03/09/2021 – 10/09/2021:**
 - o Lấy dữ liệu 5 mùa bóng gần nhất (2016-2017 -> 2020-2021) và lấy elo từ mùa trước mùa bắt đầu lấy dữ liệu (trường hợp này là lấy elo các đội từ mùa 2015-2016)
 - o Thực hiện xử lý dữ liệu (bỏ các dữ liệu thừa => tính toán giá trị trung bình của các thuộc tính cần thiết => lưu lại data để thực hiện việc train model)
 - o Thực hiện huấn luyện model và dự đoán Thắng/Thua với 4 thuật toán (RandomForest, LogisticRegression, Neural Network, LinearSVC)
 - o Hoàn thành Slide thuyết trình
 - o Thêm mô tả trong các file ipynb
- **11/09/2021-12/09/2021:** Cập nhật từ các góp ý và nộp bài

3. Một số hình ảnh từ Trello



Crawl 05: Elo rating on nba-optimization | Trello

Board: nba-optimization

Choose Project

- Win prediction
- Win percentage
- Injuries prediction
- MVP prediction
- + Add another card

Crawl 05: Elo rating
in list [Crawl data](#)

MEMBERS: MN + DUE DATE: Sep 7 at 7:00 PM **OVERDUE**

Description

Add a more detailed description...

Activity Show details

Write a comment...

Minh Vũ Công Sep 7 at 5:19 PM

1. Term
2. Crawl data + datetime (1 break)
3. Elo(2018).txt
4. def cal_elo(): return int
5. Edit description.txt
6. tao folder data/
7. crawl_data.ipynb => data/
8. push git

✎ - Edit - Delete

SUGGESTED

- Join

ADD TO CARD

- Members
- Labels
- Checklist
- Dates
- Attachment
- Cover
- Custom Fields

POWER-UPS

- + Add Power-Ups

AUTOMATION

- + Add button

ACTIONS

- Move

Crawl06: Fix bug on nba-optimization | Trello

Board: nba-optimization

Choose Project

- Win prediction
- Win percentage
- Injuries prediction
- MVP prediction
- + Add another card

Crawl06: Fix bug
in list [Crawl data](#)

MEMBERS: MN MC + DUE DATE: Sep 8 at 7:00 PM **OVERDUE**

Description Edit

1. Description from summary
2. Preseason_data.csv from
3. ".Name" -> ".Team"
4. Reverse df
5. "data.csv" -> "raw_data.csv"
6. Check tên đội (*)

Activity Show details

Write a comment...

ADD TO CARD

- Members
- Labels
- Checklist
- Dates
- Attachment
- Cover
- Custom Fields

POWER-UPS

- + Add Power-Ups

AUTOMATION

- + Add button

ACTIONS

- Move
- Copy
- Make template

