Team Work Group 9

1. Thông tin thành viên.

|  |  |  |
| --- | --- | --- |
| STT | Họ tên | MSSV |
| 1 | Hứa Phú Thành |  |
| 2 | Trần Luật Vy |  |

1. Phân công công việc.

|  |  |
| --- | --- |
| Thời gian | Công việc |
| 2 / 1 / 2021 – 5 / 1 /2021 | Chọn câu hỏi và trang web chứa data để xử lí.  Thống nhất chọn nội dung là: dự đoán giá xe ô tô (dựa trên các thuộc tính như hãng sản xuất, số chỗ ngồi, chiều dài xe, tải trọng xe, nhiên liệu sử dụng, …).  Trang web để parse dữ liệu: [www.cars-data.com](http://www.cars-data.com) |
| 6 / 1 / 2021 – 8 / 1 / 2021 | Xem cấu trúc trang web và check file robots.txt  Thành: lấy link các page car và link thông tin chi tiết car, lưu dữ liệu lấy được xuống file.  Vy: từ link thông tin chi tiết car lấy thông tin chi tiết lưu xuống file. |
| 10 / 1 / 2021 – 13 / 1 /2021 | Tiền xử lý.  Thành: chuẩn hóa dữ liệu về đúng dạng, xử lí các cột dạng số và caterogical, phân tích loại bỏ các cột không cần thiết, xử lí multilable choice.  Vy: chia tập dữ liệu, tạo pipeline. |