**Báo cáo TeamWork**

**FINAL PROJECT : DATA SCIENCE**

**Trần Ngọc Tịnh Nguyễn Ngọc Năng Toàn  
 18120597 18120600**

**Link GitHub : https://github.com/TranNgocTinh0810/DS-Final**

**I Lịch trao đổi & phân công công việc.**

**Lần 1: Ngày 26/12/2020   
 Nội dung: Lựa chọn đề tài.**

**Lần 2: Ngày 01/01/2021  
 Nội dung: + Thay đổi đề tài.  
 +Phân công công việc lần 1 :**

|  |  |
| --- | --- |
| **Trần Ngọc Tịnh ( 18120597 )** | **Nguyễn Ngọc Năng Toàn ( 18120600 )** |
| **+ Tìm hiểu API. + Lấy dữ liệu.** | **+ Tìm hiểu dữ liệu vừa lấy về**  **+ Tiền xử lí dữ liệu** |

**Sau đó thành viên tìm hiểu vấn đề của đồng đội để hiểu & sửa lỗi.**

**Lần 3: Ngày 03/01/2021**

**Nội dung: + Tìm hiểu về các mô hình ở BT03 & Demo.  
 + Tìm hiểu các siêu tham số ở các mô hình.**

**Lần 4: Ngày 05/01/2021**

**Nội dung: + Thống nhất về các cách tiền xử lí cho các phương pháp.  
 + Thực hiện các phương pháp.**

**II Vấn đề cụ thể.**

**Lần 1: Sau khi tham khảo các trang web hỗ trợ API mà thầy đã hướng dẫn ở buổi 4 tại trang web** [**https://github.com/public-apis/public-apis**](https://github.com/public-apis/public-apis) **thì tụi em dự định sẽ làm về API của AirVisual để có thể tìm hiểu & dự đoán về chỉ số bụi mịn PM trong không khí của các thành phố trên thế giới( đặc biệt là Việt Nam ). Nhưng AirVisual chỉ hỗ trợ đối với những tài khoản miễn phí dữ liệu là trực tuyến( không có lịch sử ) nên chúng em phải chuyển đổi chủ đề. Chủ đề tụi em thống nhất lựa chọn tiếp theo là sự kiện rất đáng chú ý ở năm 2020 đó là về   
Covid-19.**

**Lần 2: Câu hỏi nhóm em đặt ra là dự đoán số lượng người chết của mỗi quốc gia,cột output là deaths. Nhưng vấn đề nhóm gặp phải đó là dữ liệu này quá ít. Chỉ có 147 quốc gia, bên cạnh đó dữ liệu này còn rời rạc. Dẫn đến việc dự đoán một con số chính xác có kết quả là khá tệ, nếu như không muốn nói là kết quả không đúng kết quả nào trong bộ test.**

**Nhóm đã quyết định thay vì dự đoán một con số chính xác thì sẽ dự đoán một khoảng, các khoảng thì sẽ được trình bày ở phần phân tích dữ liệu phía dưới.**