# FINAL PROJECT CHÂT LƯỢNG KHÔNG KHÍ

NHẬP MÔN KHOA HỌC DỮ LIỆU

Đỗ LIÊN KHOA-18120413

#### MỞ ĐẦU

Trong thời đại công nghệ 4.0 hiện nay, việc sử dụng máy móc đã trở thành chuyện không thể tránh khỏi, ngành công nghiệp máy móc đang trở nên phát triển hơn bao giờ hết.

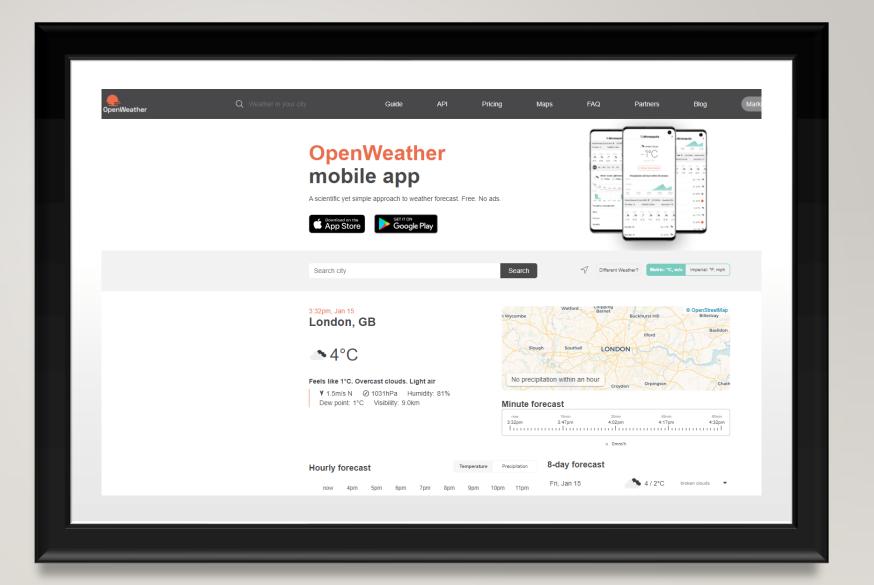
Nhưng với sự hoạt động liên tục và mạnh mẽ các loại máy như vậy dẫn đến ô nhiễm không khí rất nghiêm trọng. Ở Việt Nam, chỉ với khói bụi từ xe máy cũng đã ảnh hưởng rất nhiều đến chất lượng không khí. Vậy hiện tại chất lượng không khí ở Thành phố Hồ Chí Minh như thế nào?



### THU THẬP DỮ LIỆU

DỮ LIỆU ĐƯỢC THU THẬP TỪ TRANG

HTTPS://OPENWEATHERMA
P.ORG BĂNG API CỦA
TRANG WEB.



#### THU THẬP DỮ LIỆU

- aqi: chất lượng không khí, được sắp xếp theo thứ tự từ 1->5 với 1 = Good, 2 = Fair, 3 = Moderate, 4 = Poor, 5 = Very Poor
- co: lượng khí co trong không khí, đơn vị: μg/m3
- no: lượng khí no trong không khí, đơn vị: μg/m3
- no2: lượng khí no2 trong không khí, đơn vị: μg/m3
- o3: lượng khí o3 trong không khí, đơn vị: μg/m3
- so2: lượng khí so2 trong không khí, đơn vị: μg/m3

3	4	500.68	0.00	23.65	62.23	28.13	33.06	3
4	4	520.71	0.00	24.33	61.51	28.37	32.21	3

0.01 32.56 51.50 33.85

so2 pm2 5 pm10 nh3

dt

1609369200

8.11 1609365600 8.36 1609362000

8.61 1609358400

- pm2\_5: lượng bụi mịn có kích thước đường kính từ 2.5 tới 10 μm trong không khí, đơn vị: μg/m3
- pm10: lượng bụi mịn có kích thước đường kính nhỏ hơn hoặc bằng 2,5 μm trong không khí, đơn vị: μg/m3
- nh3: lượng khí nh3 trong không khí, đơn vị: μg/m3
- dt: thời gian đo, được tính theo UNIX time

#### TIỀN XỬ LÝ

Thực hiện các thao tác:

- Tách tập huấn luyện và tập validation.
- Xoá cột dt.

### HUẨN LUYỆN MÔ HÌNH

Sử dụng mô hình Neural Network để huấn luyện dữ liệu:

- Thay đổi tham số alpha để tìm được mô hình có độ lỗi thấp nhất với alpha = [0.01, 0.1, 1, 10, 100, 1000]
- => Độ lỗi thấp nhất là 3.195%

### TÀI LIỆU THAM KHẢO

- Slide bài giảng.
- Các bài tập BT01, BT02, BT03.
- https://scikit-learn.org/

## THANKS FOR WATCHING