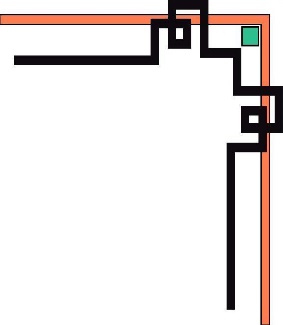
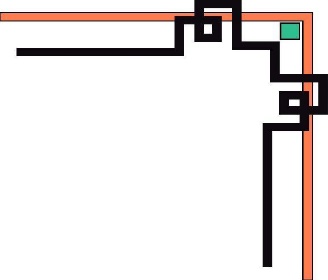
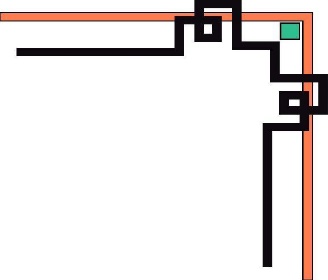
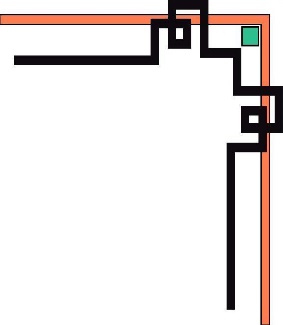
**BỘ GIÁO DỤC VÀ ĐÀO TẠO**



**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM**

**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**



**SỬ DỤNG MACHINE LEARNING TRONG ĐỀ TÀI DỰ BÁO THỜI TIẾT**

**Ngành: CÔNG NGHỆ THÔNG TIN**

**GVHD: Vũ Quang Huy**

**SVTH: Phạm Thành Phương\_15110102**

**Đàm Nhất Thống\_15110136**

**Nguyễn Huy Hoàng\_15110050**

**TP. HỒ CHÍ MINH – 11/2018**

Mục lục

[**I. Mục tiêu đề tài và cơ sở lí thuyết** 4](#_Toc528111337)

[**1.1** **Mục tiêu** 4](#_Toc528111338)

[**1.2** **Cơ sở lí thuyết** 4](#_Toc528111339)

[**II. Quá trình thực hiện (Plan and direction)** 5](#_Toc528111340)

[**III. Kết quả đạt được** 5](#_Toc528111341)

[**IV. Cơ sở lí thuyết** 5](#_Toc528111342)

[**1.Phương pháp** 5](#_Toc528111343)

[**2.Các giải thuật được áp dụng trong project** 5](#_Toc528111344)

[**V. Mô hình Flow Chart** 6](#_Toc528111345)

[**VI. Mô hình Block Diagram** 7](#_Toc528111346)

[**VII. Kết luận và hướng phát triển của đề tài** 7](#_Toc528111347)

**MỞ ĐẦU**

Ngày nay, với sự phát triển không ngừng nghỉ của công nghệ, đi kèm là hàng loạt các sản phẩm mang thiên hướng công nghệ cao được ra đời, để bắt kịp xu hướng của nhân loại thì chúng ta cần phải có sự hiểu biết nhất định về AI – Artificial Intelligence(Trí tuệ nhân tạo). Trí tuệ nhân tạo đang len lỏi hầu khắp mọi lĩnh vực của cộng đồng mà chúng ta có thể không để ý tới nó, ví dụ như Hệ thống nhận diện gắn thẻ khuôn mặt người của Facebook, hệ thống điểm danh bằng gương mặt,…

AI thể hiện một mục tiêu của con người, còn Machine Learning là một phương tiện giúp con người đạt được mục tiêu đó, và thực tế thì ML đã mang con người đi rất xa trong việc chinh phục AI nhưng vẫn còn một quãng đường dài ở phía trước. AI và ML có mối quan hệ chặt chẽ với nhau nhưng không hẳn là trung khớp bởi vì mục đích của nó là khác nhau và ML có những mục tiêu ngắn hạn như: làm cho máy tính có khả năng nhận thức cơ bản như con người như nghe, nhìn, hiểu được ngôn ngữ, giải toán, lập trình,… hỗ trợ con người xử lí một lượng lớn thông tin khổng lồ hay còn gọi là Big Data.

Vào vấn đề chính được nói tới trong đề tài này, Machine learning là một tập con của AI, với khái niệm cơ bản là một lĩnh vực nhỏ của Khoa học máy tính, có khả năng tự học hỏi dựa trên dữ liệu đưa vào mà không cần phải được lập trình cụ thể, giải các bài toán phức tạp với thời gian tính bằng phần trăm giây. Những năm gần đây, khi mà khả năng tính toán của các siêu máy tính được nâng lên một tầm cao mới và lượng dữ liệu khổng lồ được thu thập bởi các hãng công nghệ lớn, Machine Learning được tiến thêm một bước dài lên một tầm cao mới được gọi là Deep Learning giúp máy tính thực thi những bài toán mà ta có thể tưởng như không thể thực hiện được với các công nghệ cũ.

Áp dụng những cơ sở lí thuyết đã học ở trên lớp cũng như tự học ở nhà, nhóm em quyết định chọn đề tài Dự Báo Thời Tiết để làm sản phẩm cuối kì cho môn học. Trong quá trình thực hiện đề tài, chúng em đã được sự giúp đỡ và hướng dẫn tận tình của thầy Vũ Quang Huy để sản phẩm được hoàn thiện nhất.

# **I. Mục tiêu đề tài và cơ sở lí thuyết**

* 1. **Mục tiêu**

Vấn đề dự báo thời tiết hiện nay rất quan trọng trong cuộc sống chúng ta, đáp ứng nhu cầu cần biết được gần như chính xác thời tiết của ngày mai, ngày kia và trong vài ngày tới như thế nào để lên kế hoạch cho một chuyến đi, một lịch công tác hay một hoạt động ngoài trời. Nhóm đã tìm hiểu và cho ra đời sản phẩm Dự báo thời tiết sử dụng các kiến thức đã học trong ML, sử dụng các thuật toán phân tích tối ưu nhất để biết được gần như chính xác điều gì sẽ xảy ra trong một vài ngày tới.

* 1. **Cơ sở lí thuyết**

Để thực hiện được đề tài này, nhóm em quy nó về Time Series Analysis để có thể dễ hiểu và dễ áp dụng hơn. Sử dụng các thông số thời tiết cơ bản như: nhiệt độ, độ ẩm, lượng mưa, tốc độ gió,… đầu vào của thuật toán là thông số cơ bản của thời tiết ngày hôm trước, đầu ra là thông số cơ bản của ngày hôm sau.

Ở đề tài này, nhóm chúng em sử dụng thuật toán Linear Regression(Hồi qui tuyến tính) hiểu nôm na là thẳng, phẳng. Trong một không gian hai chiều, một hàm số được gọi là tuyến tính nếu đồ thị của nó có dạng đường thẳng. Trong không gian ba chiều, một hàm số được gọi là tuyến tính nếu đồ thị của nó có dạng một mặt phẳng. Trong không gian nhiều chiều, khái niệm mặt phẳng không còn phù hợp nữa, thay vào đó, một khái niệm khác ra đời được gọi là siêu mặt phẳng. Các hàm số tuyến tính là các hàm đơn giản nhất, vì chúng thuận tiện cho việc hình dung và tính toán. Tuyến tính rất quan trọng và hữu ích trong các bài toán ML nói chung và Dự báo thời tiết nói riêng.

Nghiệm của bài toán Linear Regression: Cách phổ biến nhất để tìm nghiệm của một bài toán tối ưu là giải phương trình đạo hàm bằng 0

# **II. Quá trình thực hiện (Plan and direction)**

Để hoàn thành được project thì nhóm chúng em trước hết tìm hiểu về thuật toán Linear Regression ,trong đó thuật toán Linear Regression sử dụng mô hình statsmodel và scikit learn. Nhóm chúng em trước hết học các thuật toán cơ bản về học máy.Tìm hiểu kĩ thuật toán Linear Regression.Nhóm chúng em chia ra 3 phần :

-Phần 1 : Nhóm chúng em sẽ lấy dữ liệu api để xem nhiệt độ tối đa ,nhiệt độ tối thiểu,độ ẩm,áp suất tối thiểu …

-Phần 2 Mô tả các quy trình và các bước cần thiết để xây dựng mô hình hồi quy tuyến tính nghiêm ngặt để dự đoán giá trị nhiệt độ trung bình hàng ngày trong tương lai dựa trên tập dữ liệu. Để xây dựng mô hình hồi quy tuyến tính thì nhóm em sẽ tìm hiểu về hai thư viện chính là Scikit Learn và StatModel.

-Phần 3 Mạng Nơ ron.

# **III. Kết quả đạt được**

1. Một ứng dụng hoàn chỉnh để ta có thể nhìn vào và dự báo chính xác được nhiệt độ ,độ ẩm ,lượng mưa,mây của ngày đó và những ngày khác.
2. Khi người dùng nhìn vào có thể xem được ngày đó có nắng hay mưa ,thời tiết như thế nào.

# **IV. Cơ sở lí thuyết**

## **1.Phương pháp**

Khi thu thập các tập dữ liệu để cung cấp cho các mô hình, có một số tham số được gọi là dữ liệu được phân loại bao gồm: giông bão, mưa, sương mù, u ám, nhiều mây, mây, mây rải rác và có thể nhìn mây rõ ràng. Đây có thể được phân loại thành bốn lớp.

1. Mưa, giông bão và tuyết rơi vào mưa

2. Hầu như có mây, sương mù và u ám vào rất nhiều mây

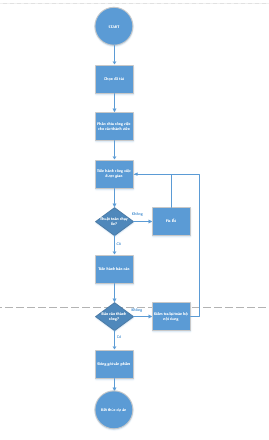
3. Những đám mây rải rác và một phần có mây vào đám mây vừa phải

4. Những đám mây rõ ràng

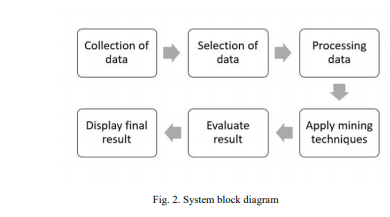
## **2.Các giải thuật được áp dụng trong project**

Linear Regression và 1 biến thể của Function Regression..

# **V. Mô hình Flow Chart**

****

# **VI. Mô hình Block Diagram**



# **VII. Kết luận và hướng phát triển của đề tài**

Nghiên cứu này đề xuất một hiệu quả là dự báo thời tiết chính xác và mô hình dự báo sử dụng các khái niệm hồi quy tuyến tính(linear regression) và phương trình Function Regression. Tất cả các những khái niệm này là một phần của việc học máy. Đơn giản phương trình là một mô hình dự đoán thời tiết rất hiệu quả và sử dụng các thực thể nhiệt độ, độ ẩm và điểm sương, có thể được sử dụng để dự đoán thời tiết đáng tin cậy. Điều này mô hình cũng tạo điều kiện cho việc ra quyết định trong cuộc sống hàng ngày. Nó có thể mang lại kết quả tốt hơn khi áp dụng cho sạch hơn và lớn hơn bộ dữ liệu. Việc xử lý trước các tập dữ liệu có thể có hiệu quả trong dự đoán là dữ liệu chưa xử lý cũng có thể ảnh hưởng đến hiệu quả của mô hình.