

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

**KHOA ĐIỆN – ĐIỆN TỬ**

**BỘ MÔN KỸ THUẬT MÁY TÍNH - VIỄN THÔNG**

-----🙞🙜🕮🙞🙜-----

**ĐỒ ÁN TỐT NGHIỆP  
ĐỀ TÀI: THIẾT KẾ HỆ THỐNG DỰ BÁO THỜI TIẾT**

**SỬ DỤNG THUẬT TOÁN MÁY HỌC**

**NGÀNH CÔNG NGHỆ KỸ THUẬT MÁY TÍNH**

**GVHD: TS. Võ Minh Huân**

Sinh viên: **Phạm Anh Tú**

MSSV: 15119156

**Nguyễn Văn Trung**

MSSV: 15119148

**Hoàng Anh**

MSSV: 15146124

TP.Hồ Chí Minh – tháng 6 năm 2019

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

**KHOA ĐIỆN – ĐIỆN TỬ**

**BỘ MÔN KỸ THUẬT MÁY TÍNH - VIỄN THÔNG**

-----🙞🙜🕮🙞🙜-----

**ĐỒ ÁN TỐT NGHIỆP  
ĐỀ TÀI: THIẾT KẾ HỆ THỐNG DỰ BÁO THỜI TIẾT**

**SỬ DỤNG THUẬT TOÁN MÁY HỌC**

**NGÀNH CÔNG NGHỆ KỸ THUẬT MÁY TÍNH**

**GVHD: TS. Võ Minh Huân**

Sinh viên: **Phạm Anh Tú**

MSSV: 15119156

**Nguyễn Văn Trung**

MSSV: 15119148

**Hoàng Anh**

MSSV: 15146124

TP.Hồ Chí Minh – tháng 6 năm 2019

**THÔNG TIN KHÓA LUẬN TỐT NGHIỆP**

1. **Thông tin sinh viên**

|  |  |
| --- | --- |
| Họ và tên sinh viên: Phạm Anh Tú  Email: 15119156@student.hcmute.edu.vn | MSSV: 15119156  Điện thoại: 0849918756 |
| Họ và tên sinh viên: Nguyễn Văn Trung  Email: 15119148@student.hcmute.edu.vn  Họ và tên sinh viên: Hoàng Anh  Email: 15146124@student.hcmute.edu.vn | MSSV: 15119148  Điện thoại: 0369884981  MSSV: 15146124  Điện thoại: 0346643034 |

1. **Thông tin đề tài**
   * Tên của đề tài: **Thiết kế hệ thống dự báo thời tiết sử dụng thuật toán máy học**
   * Đơn vị quản lý: Bộ môn Kỹ Thuật Máy Tính - Viễn Thông, Khoa Điện Điện Tử, Trường Đại Học Sư Phạm Kỹ Thuật Tp. Hồ Chí Minh.
   * Thời gian thực hiện: Từ ngày 18 / 02 / 2019 đến ngày 10 / 06 / 2019
   * Thời gian bảo vệ trước hội đồng: Ngày 19 / 06 / 2019
2. **Lời cam đoan của sinh viên**

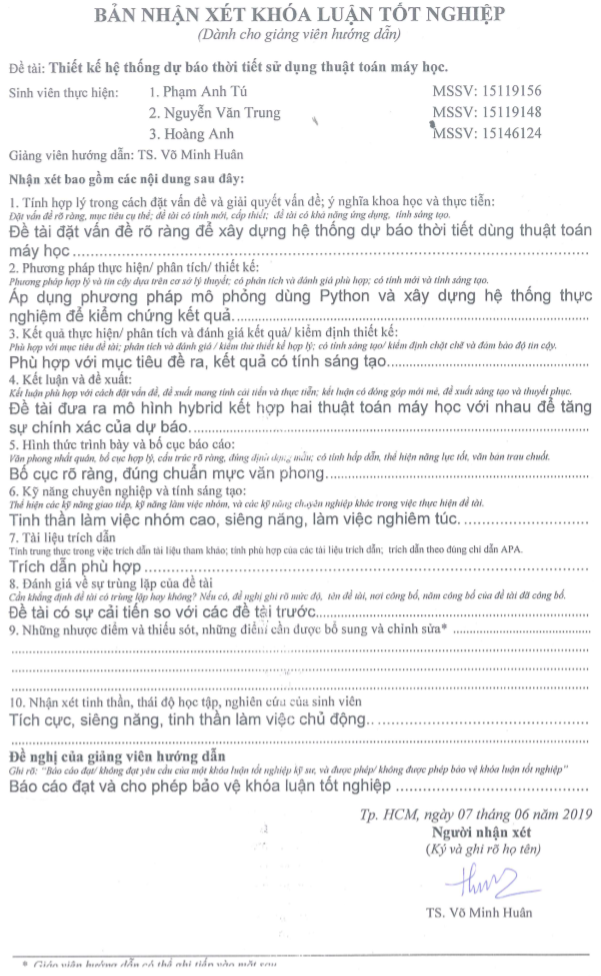
Chúng tôi – Phạm Anh Tú, Nguyễn Văn Trung, Hoàng Anh cam đoan KLTN là công trình nghiên cứu của chúng tôi dưới sự hướng dẫn của tiến sĩ Võ Minh Huân. Kết quả công bố trong KLTN là trung thực và không sao chép từ bất kỳ công trình nào khác.

|  |  |
| --- | --- |
|  | *Tp.HCM, ngày … tháng … năm 20…*  SV thực hiện đồ án  (*Ký và ghi rõ họ tên*) |

Giảng viên hướng dẫn xác nhận quyển báo cáo đã được chỉnh sửa theo đề nghị được ghi trong biên bản của Hội đồng đánh giá Khóa luận tốt nghiệp.

………………………………………………………………………………………………

|  |  |
| --- | --- |
| Xác nhận của Bộ Môn | *Tp.HCM, ngày … tháng … năm 20…*  Giáo viên hướng dẫn  *(Ký, ghi rõ họ tên và học hàm - học vị)* |

**

**BẢN NHẬN XÉT KHÓA LUẬN TỐT NGHIỆP**

*(Dùng cho giảng viên phản biện)*

Đề tài: **Thiết kế hệ thống dự báo thời tiết sử dụng thuật toán máy học**

Sinh viên thực hiện: 1. Phạm Anh Tú MSSV: 15119156

2. Nguyễn Văn Trung MSSV: 15119148

3. Hoàng Anh MSSV: 15146124

Giảng viên hướng dẫn: Ts. Võ Minh Huân

**Nhận xét bao gồm các nội dung sau đây:**

1. Tính hợp lý trong cách đặt vấn đề và giải quyết vấn đề; ý nghĩa khoa học và thực tiễn [15/100]:

*Đặt vấn đề rõ ràng, mục tiêu cụ thể [5]; đề tài có tính mới, cấp thiết [5]; đề tài có khả năng ứng dụng, tính sáng tạo [5].*

2. Phương pháp thực hiện/ phân tích/ thiết kế [25/100]:

*Phương pháp hợp lý và tin cậy dựa trên cơ sở lý thuyết [10]; có phân tích và đánh giá phù hợp [10]; có tính mới và tính sáng tạo [5].*

3. Kết quả thực hiện/ phân tích và đánh giá kết quả/ kiểm định thiết kế [25/100]:

*Phù hợp với mục tiêu [10]; phân tích và đánh giá / kiểm thử thiết kế hợp lý [10]; có tính sáng tạo/ kiểm định chặt chẽ và đảm bảo độ tin cậy [5].*

4. Kết luận và đề xuất [10/100]:

*Kết luận phù hợp với cách đặt vấn đề, đề xuất mang tính cải tiến và thực tiễn [5]; kết luận có đóng góp mới mẻ, đề xuất sáng tạo và thuyết phục[5].*

5. Hình thức trình bày, bố cục và chất lượng báo cáo [15/100]:

*Văn phong nhất quán, bố cục hợp lý, cấu trúc rõ ràng, đúng định dạng mẫu [5]; có tính hấp dẫn, thể hiện năng lực tốt, văn bản trau chuốt [15].*

6. Tài liệu trích dẫn [10/100]

Tính trung thực trong việc trích dẫn tài liệu tham khảo; tính phù hợp của các tài liệu trích dẫn; trích dẫn theo đúng chỉ dẫn APA.

7. Đánh giá về sự trùng lặp của đề tài

*Cần khẳng định đề tài có trùng lặp hay không? Nếu có, đề nghị ghi rõ mức độ, tên đề tài, nơi công bố, năm công bố của đề tài đã công bố.*

8. Những nhược điểm và thiếu sót, những điểm cần được bổ sung và chỉnh sửa\*

**Câu hỏi sinh viên phải trả lời trước hội đồng\*** *(ít nhất 02 câu)*

**Đánh giá chung**

* Điểm (Quy về thang điểm 10 không làm tròn): …………./10.
* Xếp loại chung (Xuất sắc, Giỏi, Khá, Trung bình, Yếu, Kém):…………………………………….

**Đề nghị của giảng viên phản biện**

*Ghi rõ: “Báo cáo đạt/ không đạt yêu cầu của một khóa luận tốt nghiệp kỹ sư, và được phép/ không được phép bảo vệ khóa luận tốt nghiệp”*

*Tp. HCM, ngày … tháng …. năm 20…*

**Người nhận xét**

(*Ký và ghi rõ họ tên)*

**LỜI CẢM ƠN**

Lời đầu tiên, nhóm thực hiện đề tài xin trân trọng cảm ơn GVHD TS. Võ Minh Huân đã tận tình chỉ dẫn cho nhóm các bước thực hiện đồ án này đúng tiến độ và đạt được thành quả. Trong quá trình làm việc với thầy nhóm không những tiếp thu những kiến thức bổ ích, mà còn rèn luyện được thêm nhiều kĩ năng hay, thái độ nghiêm túc trong quá trình nghiên cứu khoa học, điều này là rất cần thiết trong quá trình học tập và đặc biệt hơn nó sẽ là hành trang lớn cho con đường sự nghiệp sau này.

Đồng thời nhóm xin chân thành cảm ơn các giảng viên khoa đã truyền dạy cho chúng em những kiến thức cần thiết cho ngành.

Cuối cùng, nhóm xin chúc Quý Thầy, Cô của trường Đại học Sư Phạm Kỹ Thuật TPHCM dồi dào sức khỏe và gặt hái được nhiều thành công trong công việc.

Nhóm xin chân thành cảm ơn!

*Tp. Hồ chí Minh, tháng 6 năm 2019*

Sinh viên thực hiện

Phạm Anh Tú

Nguyễn Văn Trung

Hoàng Anh

**TÓM TẮT**

Dự báo thời tiết đã có từ xa xưa, từ khi có sự xuất hiện của con người và ngày càng gắn bó với cuộc sống của chúng ta. Từ thời cha ông của chúng ta, trong quá trình làm việc, sinh hoạt họ đã có thể dự đoán trước được các hiện tượng tự nhiên xảy ra trong thời gian gần. Dần dần kinh nghiệm quan sát được tích lũy, cộng thêm những tư duy sâu xa đã có thể giúp họ có được các dự báo chính xác hơn cho các hiện tượng khí tượng trong tương lai. Những phân tích, kinh nghiệm tích lũy dần được đúc kết thành các bài toán dự báo, cho phép tính toán khá chính xác hiện tượng thời tiết sắp xảy ra.

Hiện nay trên thế giới các hệ thống dự báo thời tiết được thực hiện bởi nhiều phương pháp khác nhau như: Phương pháp Synôp, phương pháp thống kê, phương pháp số trị,.. Trong mỗi phương pháp đó lại có rất nhiều phương pháp và mô hình dự báo cụ thể. Mỗi phương pháp đều có những ưu điểm và nhược điểm riêng. Việc ứng dụng phương pháp nào là tùy thuộc vào tài nguyên của hệ thống, trình độ của người dự báo, yêu cầu và mục đích của người xây dựng hệ thống,..

Ứng dụng các thuật toán máy học là một trong những phương pháp mới, có nhiều ưu thế vượt trội trong việc giải quyết các vấn đề phi tuyến. Các thuật toán máy học này nó mang tính chất khách quan và đáp ứng được nhiều yêu cầu của thực tiễn mà các dự báo bằng các phương pháp khác không đáp ứng được. Với những ưu thế vượt trội ở trên như tính linh hoạt, giải quyết được các vấn đề phi tuyến thích hợp trong việc xử lý dữ liệu có tính biến động lớn,… Đó là những lý do để nhóm chọn đề tài: Thiết kế hệ thống dự báo thời tiết sử dụng thuật toán máy học.

**MỤC LỤC**

[MỤC LỤC ix](#_Toc11608669)

[DANH MỤC CÁC BẢNG xii](#_Toc11608670)

[DANH MỤC CÁC HÌNH ẢNH xiii](#_Toc11608671)

[DANH MỤC CÁC TỪ VIẾT TẮT xv](#_Toc11608672)

[CHƯƠNG 1 1](#_Toc11608673)

[TỔNG QUAN 1](#_Toc11608674)

[1.1. GIỚI THIỆU ĐỀ TÀI 1](#_Toc11608675)

[1.2. TÍNH CẤP THIẾT CỦA ĐỀ TÀI 1](#_Toc11608676)

[1.3. MỤC TIÊU CỦA ĐỀ TÀI 2](#_Toc11608677)

[1.4. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU 2](#_Toc11608678)

[1.4.1. Đối tượng nghiên cứu 3](#_Toc11608679)

[1.4.2. Phạm vi nghiên cứu 3](#_Toc11608680)

[1.5. PHƯƠNG PHÁP NGHIÊN CỨU 3](#_Toc11608681)

[1.6. BỐ CỤC CỦA ĐỀ TÀI 3](#_Toc11608682)

[1.7. THỜI GIAN DỰ KIẾN THỰC HIỆN 4](#_Toc11608683)

[CHƯƠNG 2 5](#_Toc11608684)

[CƠ SỞ LÝ THUYẾT 5](#_Toc11608685)

[2.1. PHẦN CỨNG 5](#_Toc11608686)

[2.1.1. Vi điều khiển STM32F103C6T8 5](#_Toc11608687)

[2.1.2. NodeMCU ESP8266 6](#_Toc11608688)

[2.1.3. Module Zigbee 7](#_Toc11608689)

[2.1.4. Cảm biến DHT22 8](#_Toc11608690)

[2.1.5. Relay và Opto 9](#_Toc11608691)

[2.2. CÔNG NGHỆ TRUYỀN THÔNG KHÔNG DÂY ZIGBEE 10](#_Toc11608692)

[2.3. MÔ HÌNH THUẬT TOÁN ARIMA 13](#_Toc11608693)

[2.4. MÔ HÌNH THUẬT TOÁN SVMR 17](#_Toc11608694)

[2.4.1. Sơ lược về SVM (Support Vector Machine) 17](#_Toc11608695)

[2.4.2. Sơ lược về SVMr (Support Vector Machine-Regression) 18](#_Toc11608696)

[2.4.3. Xây dựng bài toán cho SVMr 18](#_Toc11608697)

[2.5. MÔ HÌNH KẾT HỢP (HYBRID) 21](#_Toc11608698)

[CHƯƠNG 3 22](#_Toc11608699)

[THIẾT KẾ HỆ THỐNG 22](#_Toc11608700)

[3.1. HÌNH THÀNH Ý TƯỞNG 22](#_Toc11608701)

[3.2. MÔ HÌNH HỆ THỐNG 26](#_Toc11608702)

[3.3. THIẾT KẾ PHẦN CỨNG 27](#_Toc11608703)

[3.3.1. Chức năng phần cứng 27](#_Toc11608704)

[3.3.2. Sơ dồ khối phần cứng 28](#_Toc11608705)

[3.3.3. Thiết kế từng khối 30](#_Toc11608706)

[3.3.3.1 Khối trung tâm 30](#_Toc11608707)

[3.3.3.2 Khối Node Sensor 31](#_Toc11608708)

[3.4. XÂY DỰNG DATABASE ONLINE 33](#_Toc11608709)

[3.4.1. Cách tạo tài khoản trên ThingSpeak 34](#_Toc11608710)

[3.4.2. Hướng dẫn gửi dữ liệu từ ESP8266 36](#_Toc11608711)

[3.5. THIẾT KẾ PHẦN MỀM 38](#_Toc11608712)

[3.6. THIẾT KẾ THUẬT TOÁN 41](#_Toc11608713)

[CHƯƠNG 4 46](#_Toc11608714)

[KẾT QUẢ THỰC NGHIỆM VÀ ĐÁNH GIÁ 46](#_Toc11608715)

[4.1. MÔ HÌNH THỰC NGHIỆM 46](#_Toc11608716)

[4.2. THỰC NGHIỆM 47](#_Toc11608717)

[4.2.1. Kiểm tra tính năng hoạt động của phần cứng 47](#_Toc11608718)

[4.2.2. Thực nghiệm thuật toán 48](#_Toc11608719)

[4.3. ƯU ĐIỂM 54](#_Toc11608720)

[4.4. NHƯỢC ĐIỂM 54](#_Toc11608721)

[CHƯƠNG 5 55](#_Toc11608722)

[KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 55](#_Toc11608723)

[5.1. KẾT LUẬN 55](#_Toc11608724)

[5.2.1. Những vấn đề nghiên cứu 55](#_Toc11608725)

[5.2.2. Những vấn đề hoàn thành 55](#_Toc11608726)

[5.2. HẠN CHẾ 55](#_Toc11608727)

[5.3. HƯỚNG PHÁT TRIỂN 55](#_Toc11608728)

[TÀI LIỆU THAM KHẢO 57](#_Toc11608729)

[PHỤ LỤC 58](#_Toc11608730)

# DANH MỤC CÁC BẢNG

[*Bảng 4.1:* *Phân tích dữ liệu thành 3 phần*. 49](#_Toc11608731)

[*Bảng 4.2:* *Bảng dưới đây liệt kê các thông số tốt nhất của 3 mô hình*. 52](#_Toc11608732)

[*Bảng 4.3:* *Bảng so sánh độ chính xác giữa 3 mô hình.* 54](#_Toc11608733)

# DANH MỤC CÁC HÌNH ẢNH

[*Hình 2.1:* *Vi điều khiển STM32F103C8T6 đen.* 5](#_Toc11769428)

[*Hình 2.2:* *NodeMCU ESP8266.* 6](#_Toc11769429)

[*Hình 2.3:* *Module Zigbee.* 7](#_Toc11769430)

[*Hình 2.1:* *Cảm biến DHT22.* 8](#_Toc11769431)

[*Hình 2.2:* Sơ đồ nguyên lý của Relay và Opto 9](#_Toc11769432)

[*Hình 2.3:* *Cấu trúc các tầng của Zigbee.* 11](#_Toc11769433)

[*Hình 2.4:* *Các loại cấu trúc mạng.* 13](#_Toc11769434)

[*Hình 2.5:* *Chuỗi nhiệt độ đo tại TP. HCM (24/4/2019- 30/4/2019).* 14](#_Toc11769435)

[*Hình 2.6:* *Phân tích thuật toán SVM.* 18](#_Toc11769436)

[*Hình 2.7:* *Phân tích thuật toán SVMr* [3]*.* 19](#_Toc11769437)

[*Hình 2.8:* *Phân tích thuật toán SVMr với tập dữ liệu có điểm gây nhiễu.* 20](#_Toc11769438)

[*Hình 2.9:* *Kernel Trick.* 20](#_Toc11769439)

[*Hình 3.1:* *Sơ đồ tư duy của hệ thống.* 23](#_Toc11769440)

[*Hình 3.2:* *Minh họa dữ liệu trên Thingspeak.* 25](#_Toc11769441)

[*Hình 3.3:* *Mô hình hệ thống.* 26](#_Toc11769442)

[*Hình 3.4:* *khối chính của phần cứng hệ thống.* 27](#_Toc11769443)

[*Hình 3.5:* *Sơ đồ khối chi tiết của hệ thống.* 28](#_Toc11769444)

[*Hình 3.6:* *Mô hình đa truy cập.* 29](#_Toc11769445)

[*Hình 3.7:* *Sơ đồ khối của khối trung tâm.* 30](#_Toc11769446)

[*Hình 3.8:* *Khối kết nối phần cứng của khối xử lý trung tâm.* 31](#_Toc11769447)

[*Hình 3.9:* *Sơ đồ khối của khối Node Sensor.* 31](#_Toc11769448)

[*Hình 3.10:* *Khối kết nối phần cứng của khối Node Sensor.* 32](#_Toc11769449)

[*Hình 3.11:* *Minh họa quá trình truyền dữ liệu lên ThingSpeak.* 33](#_Toc11769450)

[*Hình 3.12:* *Giao diện trang chủ ThingSpeak.* 34](#_Toc11769451)

[*Hình 3.13:* *Tạo NewChannel.* 35](#_Toc11769452)

[*Hình 3.14:* *API Keys.* 35](#_Toc11769453)

[*Hình 3.15:* *Mở cửa sổ Manage Libraries.* 36](#_Toc11769454)

[*Hình 3.16:* *Tìm kiếm thư viện.* 36](#_Toc11769455)

[*Hình 3.18:* *Khởi động ThingSpeak.* 37](#_Toc11769456)

[*Hình 3.19:* *Gửi nhận dữ liệu lên ThingSpeak.* 38](#_Toc11769457)

[*Hình 3.20:* *Lưu đồ chính của bộ xử lý trung tâm.* 39](#_Toc11769458)

[*Hình 3.21:* *Lưu đồ khối xử lý trung tâm.* 40](#_Toc11769459)

[*Hình 3.22:* *Lưu đồ thuật toán SVMr.* 41](#_Toc11769460)

[*Hình 3.23:* *Lưu đồ thuật toán ARIMA.* 42](#_Toc11769461)

[*Hình 3.24:* *Lưu đồ mô hình Hybrid.* 43](#_Toc11769462)

[*Hình 3.25:* *Lưu đồ thuật toán.* 44](#_Toc11769463)

[*Hình 4.1:* *Khối điều khiển trung tâm.* 46](#_Toc11769464)

[*Hình 4.2:* *Khối điểm cảm biến* 46](#_Toc11769465)

[*Hình 4.3:* *Khối điểm cảm biến.* 46](#_Toc11769466)

[*Hình 4.4:* *Biểu đồ thể hiện khả năng đáp ứng của cảm biến đối với sự thay đổi nhiệt độ, độ ẩm.* 47](#_Toc11769467)

[*Hình 4.5:* *Hệ thống dự báo thời tiết thực tế* 48](#_Toc11769468)

[*Hình 4.6:* *MAE của SVMr.* 50](#_Toc11769469)

[*Hình 4.7:* *MAE của ARIMA.* 51](#_Toc11769470)

[*Hình 4.8:* *MAE residual của Hybrid model.* 51](#_Toc11769471)

[*Hình 4.9:* *Kết quả của SVMr model.* 52](#_Toc11769472)

[*Hình 4.10:* *Kết quả của ARIMA model.* 53](#_Toc11769473)

[*Hình 4.11:* *Kết quả của hybrid model.* 53](#_Toc11769474)

# DANH MỤC CÁC TỪ VIẾT TẮT

|  |  |
| --- | --- |
| **Từ viết tắt** | **Nghĩa đầy đủ** |
| IoTs | Internet of Things |
| SoC | System on a Chip |
| IEEE | Institute of Electrical and Electronics Engineers |
| MCU | Microcontroller Unit |
| CSMA | Carrier Sense Multiple Access |
| UART | Universal Asynchronous Receiver-Transmitter |
| API | Application programming interface |
| MAE | Mean Absolute Error |
| MSE | Mean Square Error |
| MAPE | Mean Absolute Percent Error |
| RMSE | Root Mean Square Error |
| ARM | Advanced RISC Machine |
| SVM | Support Vector Machine |
| SVMr | Support Vector Machine – Regression |
| RBF | Radial Basis Function |
| AR | Auto-Regressive |
| I | Integrated |
| MA | Moving Average |
| ARMA | Auto-Regressive Moving Average |
| ARIMA | Auto-Regressive Integrated Moving Average |

# CHƯƠNG 1

# TỔNG QUAN

## GIỚI THIỆU ĐỀ TÀI

Trên thế giới cũng như ở Việt Nam hiện nay, việc dự báo thời tiết là một vấn đề rất quan trọng trong cuộc sống của mọi người, các dự báo viên chủ yếu áp dụng các phương pháp: Synop, thống kê, phân tích ảnh mây vệ tinh, sản phẩm Radar thời tiết, dự báo thời tiết bằng mô hình số trị,… Tuy nhiên trong đề tài này nhóm đã nghiên cứu một mô hình thuật toán máy học để dự đoán thời tiết theo kiểu riêng của mình. Vậy tại sao lại áp dụng thuật toán máy học vào dự đoán thời tiết trong nông nghiệp? Dự báo thời tiết bằng máy học đang trở thành xu thế mới trên thế giới, vì tính chính xác và hiệu quả của nó trong các ứng dụng đời sống. Do đó, có rất nhiều thuật toán dùng để dự đoán thời tiết như: SVMr, Neural Networks, ARIMA,... Trong phạm vi đề tài này sẽ dùng hai thuật toán chính là SVMr (Support Vector Machine – Regression), ARIMA. Tất cả thuật toán đó đều có thể dùng để dự báo thời tiết theo chuỗi thời gian (time series). Bên cạnh đó, đề tài cũng sẽ đánh giá hiệu quả của các phương pháp dự đoán.

## TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Trong đời sống kinh tế- xã hội, để đạt được năng suất cũng như hiệu quả cao trong công việc thì vấn đề dự đoán của các yếu tố ở trong tương lai thì rất quan trọng. Một trong những yếu tố quan trọng mà chúng ta không thể không nhắc đến đó là yếu tố thời tiết. Vậy tại sao dự báo thời tiết là một vấn đề cần thiết của mọi người trong cuộc sống hằng ngày? Trong những năm gần đây, dự báo thời tiết đã thu hút nhiều sự quan tâm trong nhiều cộng đồng nghiên cứu vì nó có thể bảo vệ và giúp ích cho cuộc sống của con người. Đối với nước ta thì nông nghiệp luôn chiếm tỷ trọng cao trong cơ cấu kinh tế, vì vậy việc áp dụng các kỹ thuật tiên tiến vào sẽ hỗ trợ được một phần thời gian, sức lực của người nông dân, đồng thời cũng sẽ làm tăng năng suất trong nông nghiệp. Muốn cho năng suất cao thì yếu tố quan trọng đó chính là thời tiết, các thông tin thời tiết bao gồm nhiệt độ, tốc độ gió, bầu trời, độ ẩm,… Do sự thay đổi thất thường của các yếu tố thời tiết đã ảnh hưởng trực tiếp đến hoạt động sản xuất nông nghiệp, câu hỏi được đặt ra ở đây là liệu những thay đổi thời tiết có thể được dự báo với độ chắc chắn đáng tin cậy hay không, để từ đó có thể áp dụng những dự đoán đó vào trong hoạt động nông nghiệp đáp ứng một cách hiệu quả nhất. Với mong muốn sẽ được giải quyết được câu hỏi ở trên, nhóm quyết định tìm hiểu và thực hiện hệ thống cung cấp dữ liệu môi trường theo thời gian thực tại các khu vực nông thôn để người làm nông có thể phát hiện và quản lý các vấn đề cây bệnh để từ đó làm cho nền nông nghiệp phát triển theo hướng công nghiệp hóa-hiện đại hóa, nâng cao hiệu quả, tạo bước đột phá về năng suất, chất lượng nông sản, thỏa mãn nhu cầu ngày càng cao của xã hội.

## MỤC TIÊU CỦA ĐỀ TÀI

Từ các điểm cấp thiết của vấn đề dự báo thời tiết hiện nay, việc xây dựng một hệ thống dự báo như thế nào, nó sẽ hoạt động như thế nào? Để trả lời cho những câu hỏi này cũng chính là xây dựng mục tiêu của đề tài một cách bao quát và đầy đủ của toàn hệ thống. Nhóm thiết kế hệ thống IoTs bao gồm một thiết bị có thể gửi dữ liệu môi trường theo thời gian thực và một thuật toán máy học để dự đoán dữ liệu môi trường trong tương lai. Các dữ liệu môi trường bao gồm nhiệt độ, độ ẩm môi trường, hướng gió, mực nước sẽ được xử lý từ xa trên máy tính. Thuật toán máy học sử dụng Support Vector Machine – Regression (SVMr), ARIMA và kết hợp cả hai thuật toán để xử lý dữ liệu môi trường, từ đó dự đoán dữ liệu vào ngày tiếp theo. Các dữ liệu đã được dự đoán sẽ được gửi lên Internet để lưu trữ, người nông dân sẽ theo dõi trực tiếp và điều khiển tưới tiêu, bón phân,… một cách thuận tiện.

## ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

Với mục tiêu đề tài ở trên, nhóm cần phải xác định đối tượng nghiên cứu là gì, và phạm vi nghiên cứu đến đâu để từ đó có thể thực hiện nghiên cứu đề tài một cách có hệ thống và dễ dàng trình bày một cách hiệu quả nhất.

### Đối tượng nghiên cứu

* Vi điều khiển STM32F103C8T6, module Wifi ESP8266, các IC cho đọc cảm biến nhiệt độ.
* Database online, thuật toán SVMr, ARIMA.

### Phạm vi nghiên cứu

* Hoạt động của STM32F103C8T6, module wifi ESP8266, các IC cho đọc cảm biến nhiệt độ.
* Nghiên cứu database online.
* Nghiên cứu thuật toán SVMr.
* Nghiên cứu thuật toán ARIMA.

## PHƯƠNG PHÁP NGHIÊN CỨU

Để hoàn thành đề tài, việc thực hiện đề tài theo đúng phương pháp sẽ rất thuận lợi cho việc định hướng đề tài của mình cần làm gì và nó sẽ có những vấn đề gì cần tập trung nghiên cứu, chính vì những vấn đề như vậy nhóm phải áp dụng những phương pháp nào cho phù hợp? Dưới đây là một số phương pháp nhóm sẽ thực hiện xuyên suốt trong quá trình thực hiện đề tài.

* Nghiên cứu tài liệu.
* Phân tích, tổng hợp để trình bày các vấn đề.
* Liệt kê, sưu tầm tài liệu.

## BỐ CỤC CỦA ĐỀ TÀI

Có tổng cộng 5 chương trong quyển báo cáo để làm rõ đề tài này:

* *Chương 1: Giới thiệu* bao gồm các nội dung giới thiệu đề tài, tính cấp thiết đề tài, hình thành ý tưởng, mục tiêu đề tài, phương pháp nghiên cứu, đối tượng, phạm vi nghiên cứu.
* *Chương 2: Cơ sở lý thuyết* bao gồm cơ sở lý thuyết các vi điều khiển, cảm biến, các giao thức truyền nhận, thuật toán SVMr.
* *Chương 3: Thiết kế hệ thống* bao gồm các nội dung hình thành ý tưởng, thiết kế phần cứng và thiết kế phần mềm.
* *Chương 4: Kết quả thực nghiệm và đánh giá* bao gồm kết quả mô hình sản phầm thi công, hoạt động của hệ thống và đánh giá kết quả hoạt động của hệ thống.
* *Chương 5: Kết luận và hướng phát triển*.

# CHƯƠNG 2

# CƠ SỞ LÝ THUYẾT

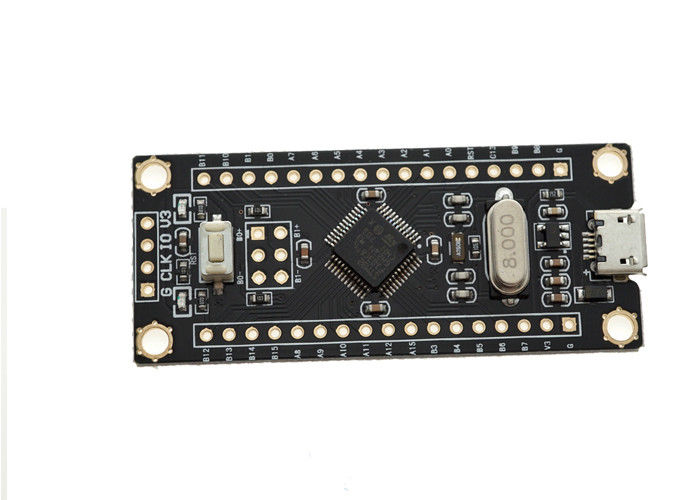
Với sự lựa chọn phần cứng cũng như thuật toán để thực thi đề tài này là rất quan trọng, nó gần như quyết định thành công của hệ thống này. Vậy chúng ta cần phải tìm hiểu, liệt kê phần cứng và giải thuật dự đoán đã được sử dụng trong hệ thống này là gì?

## PHẦN CỨNG

1. **Vi điều khiển STM32F103C6T8**

* *Giới thiệu:*

STM32 là một trong những dòng chip phổ biến của STMicroelectronics với nhiều họ thông dụng như F0, F1, F2, F3, F4,… STM32F103 thuộc họ F1 với lõi là ARM COTEX M3. STM32F103 là vi điều khiển 32 bit, tốc độ tối đa là 72MHz. Giá thành cũng khá rẻ so với các loại vi điều khiển có chức năng tương tự. Mạch nạp cũng như công cụ lập trình khá đa dạng và dễ sử dụng.



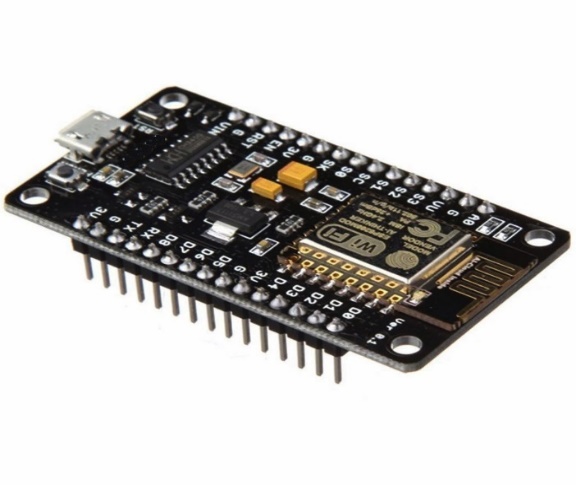
1. *Vi điều khiển STM32F103C8T6 đen.*

* *Thông số kĩ thuật*
* Điện áp hoạt động: 2V-3.3V DC.
* Chế độ Debug: SWD.
* Tần số làm việc: 72MHz.
* Bộ nhớ flash: 64K.
* SRAM: 20K.
* Hổ trợ các chuẩn giao tiếp: CAN, I2C, SPI UART/USART, USB.
* Tích hợp USB để cấp nguồn và giao tiếp.
* Có led báo nguồn.
* Nút nhấn Reset.
* Kích thước: 5.3cm x 2.2cm.

1. **NodeMCU ESP8266**

* *Giới thiệu****:***

Kit RF thu phát Wifi ESP8266 NodeMCU là kit phát triển dựa trên nền chip Wifi SoC ESP8266 với thiết kế dễ sử dụng và đặc biệt là có thể sử dụng trực tiếp trình biên dịch của Arduino để lập trình và nạp code, điều này khiến việc sử dụng và lập trình các ứng dụng trên ESP8266 trở nên đơn giản hơn. Kit RF thu phát Wifi ESP8266 NodeMCU được dùng cho các ứng dụng cần kết nối, thu thập dữ liệu và điều khiển qua sóng Wifi, đặc biệt là các ứng dụng liên quan đến IoTs.



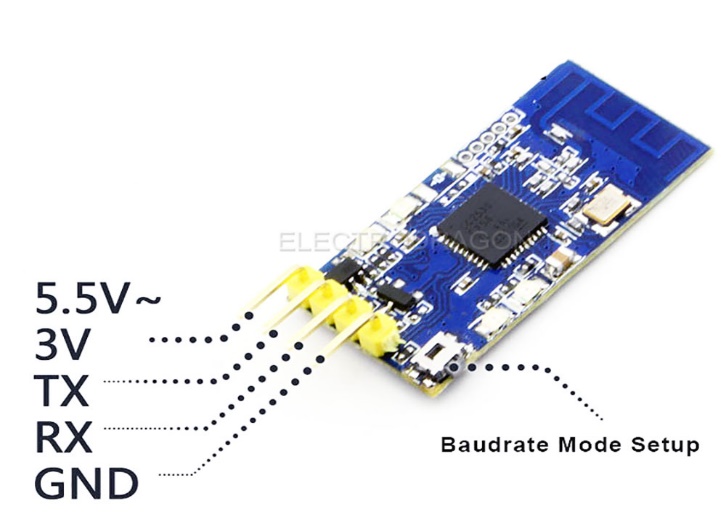
1. *NodeMCU ESP8266.*

* *Thông số kĩ thuật*
* IC chính: ESP8266 Wifi SoC.
* Phiên bản firmware: NodeMCU Lua
* WiFi: 2.4 GHz hỗ trợ chuẩn 802.11 b/g/n
* Tích hợp giao thức TCP/IP.
* Chip nạp và giao tiếp UART: CP2102.
* Chuẩn giao tiếp: UART, SPI.
* Cấp nguồn: 5VDC thông qua MicroUSB hoặc chân Vin.
* Dòng tiêu thụ: 10uA-170mA.
* GIPO giao tiếp mức 3.3VDC.
* Tương thích hoàn toàn với trình biên dịch Arduino.

1. **Module Zigbee**

* *Giới thiệu*

Module ZIGBEE CC2530 UART TTL 2.4G sử dụng IC CC2530 từ TI, mạch được lập trình sẵn firmware để có thể dễ dàng sử dụng như một module truyền nhận dữ liệu không dây chuẩn Zigbee với giao tiếp UART rất dễ kết nối với vi điều khiển hoặc máy tính (thông qua cáp chuyển USB-UART).



1. *Module Zigbee.*

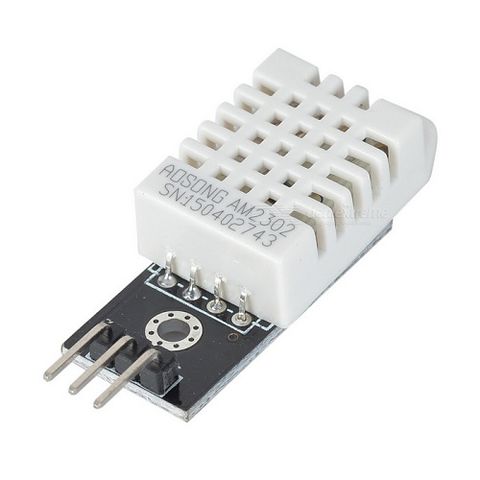
Module ZIGBEE CC2530 UART TTL 2.4G có khoảng cách truyền nhận xa, chuẩn truyền sóng Zigbee 2.4Ghz chuẩn công nghiệp rất ổn định và có khả năng cấu hình tạo thành mạng truyền nhận không dây với nhiều nút, điểm mạng khác khau qua giao thức Zigbee.

* *Thông số kỹ thuật*
* IC chính RF Zigbee SoC CC2530 từ TI.
* Điện áp sử dụng: 3 - 5.5VDC
* Dòng tiêu thụ khi nhận dữ liệu: tối đa 26.3mA
* Dòng tiêu thụ khi truyền dữ liệu: tối đa 37.5mA
* Dòng tiêu thụ ở trạng thái hoạt động: tối đa 2.6mA
* Chuẩn truyền sóng Zigbee 2.4Ghz.
* Tốc độ truyền sóng tối đa 3300bps.
* Công suất truyền: 4.5dbm
* Khoảng cách truyền lý tưởng: 250m (không có vật cản).
* Giao thức kết nối UART TTL (3.3VDC hoặc 5VDC), Baudrate tối đa 115200.

1. **Cảm biến DHT22**

* *Giới thiệu*

Cảm biến độ ẩm nhiệt độ DHT22 là cảm biến thông dụng tích hợp vừa đo được nhiệt độ và độ ẩm, độ chính xác khá cao. Giao tiếp với vi điều khiển qua chuẩn giao

tiếp 1 dây. Cảm biến độ ẩm nhiệt độ DHT22ra chân được tích hợp sẵn điện trở 5.1 KOhm. DHT22 có độ chính xác cao và khoảng đo hoạt động rộng. Module được thiết kế hoạt động ở mức điện áp 5VDC.

1. *Cảm biến DHT22.*

* *Thông số kỹ thuật*
* Điện áp hoạt động: 5VDC.
* Dòng tiêu thụ: 2.5mA.
* Khoảng đo độ ẩm: 0% - 100% RH sai số 2% RH.
* Khoảng đo nhiệt độ: -40 ~ -80 độ C sai số 0.5% độ C.
* Tần số lấy mẫu tối đa 0.5Hz (2 giây / lần).

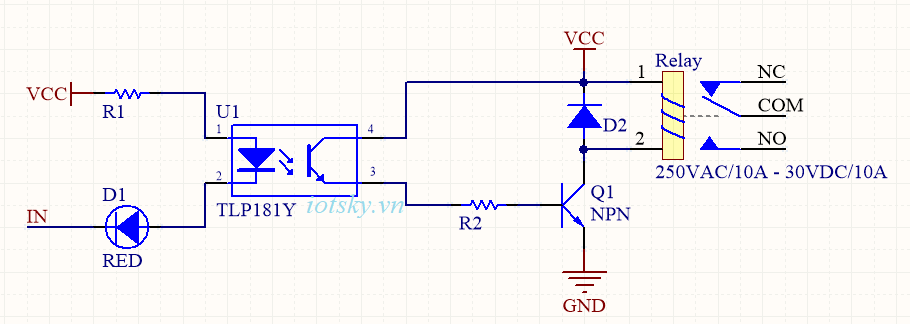
1. **Relay và Opto**

* *Giới thiệu :*

Opto là IC thường dùng để cách ly nguồn cao áp với vi điều khiển. Khi nguồn bị chập hoặc gặp sự cố sẽ không ảnh hưởng đến vi điều khiển.

Khi LED của Opto phát sáng với cường độ sáng nào đó, thì vùng base của transistor cảm quang tiếp nhận ánh sáng và giảm mức điện trở thuần tương đương của transistor, làm dòng qua transistor tăng.

Khi đó dòng điện chạy qua rơ le, dòng điện này sẽ chạy qua cuộn dây bên trong và tạo ra một từ trường hút. Từ trường hút này tác động lên một đòn bẩy bên trong làm đóng hoặc mở các tiếp điểm điện và như thế sẽ làm thay đổi trạng thái của rơ le.

Rơ le có 2 mạch độc lập nhau họạt động. Một mạch là để điều khiển cuộn dây của rơ le: Cho dòng chạy qua cuộn dây hay không, hay có nghĩa là điều khiển rơ le ở trạng thái ON hay OFF.

1. Sơ đồ nguyên lý của Relay và Opto

* *Thông số kĩ thuật*
* Sử dụng điện áp nuôi DC 5V.
* Relay tiêu thụ dòng khoảng 80mA.
* Điện thế đóng ngắt tối đa: AC250V~10A hoặc DC30V~10A.
* Có đèn báo đóng ngắt trên mỗi Relay.
* Có thể chọn mức tín hiệu kích 0 hoặc 1 qua jumper.

## CÔNG NGHỆ TRUYỀN THÔNG KHÔNG DÂY ZIGBEE

Công nghệ ZigBee được xây dựng dựa trên tiêu chuẩn 802.15.4 của tổ chức IEEE. Tiêu chuẩn 802.15.4 này sử dụng tín hiệu radio có tần sóng ngắn.

Nhờ chức năng điều khiển từ xa không dây, truyền dữ liệu ổn định, tiêu thụ năng lượng cực thấp, công nghệ mở đã giúp công nghệ ZigBee trở nên hấp dẫn sử dụng cho các ứng dụng, đặc biệt là ứng dụng trong nhà thông minh hiện nay.

Thực tế cho thấy hệ thống có thể hoạt động trong môi trường có dữ liệu dày đặc, hay trong vùng mà có nhiều đường truyền khác làm nhiễu thì hệ thống vẫn đảm bảo hoạt động liên tục đó là nhờ sự đánh giá chất lượng, sự phát hiện năng lượng tiếp nhận và đánh giá kênh rõ ràng.

Công nghệ đa truy cập nhận biết sóng mang CSMA được sử dụng để xác định thời điểm truyền, và tránh được những va chạm trong đường truyền.

* ***Truyền dữ liệu***

Tín hiệu công nghệ ZigBee có thể truyền xa đến 75m tính từ trạm phát, và khoảng cách có thể xa hơn rất nhiều nếu được tiếp tục phát từ nút liên kết tiếp theo trong cùng hệ thống. Các dữ liệu được truyền theo gói, gói tối đa là 128bytes cho phép tải xuống tối đa 104 bytes.

Tiêu chuẩn này hỗ trợ địa chỉ 64bit cũng như địa chỉ ngắn 16bit. Loại địa chỉ 64bit chỉ xác đinh  được mỗi thiết bị có cùng 1 địa chỉ IP duy nhất. Khi mạng được thiết lập, những địa chỉ ngắn có thể được sử dụng và cho phép hơn 65000 nút được liên kết.

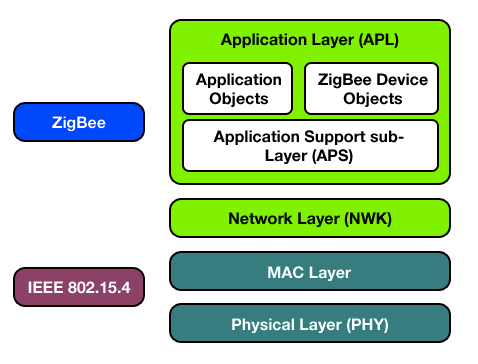
* ***Cấu trúc các tầng của Zigbee***

Ngoài 2 tầng vật lý và tầng MAC xác định bởi tiêu chuẩn 802.15.4, tiêu chuẩn ZigBee còn có thêm các tầng trên của hệ thống bao gồm: tầng mạng, tầng hỗ trợ ứng dụng, tầng đối tượng thiết bị và các đối tượng ứng dụng.

**Tầng vật lý:** có trách nhiệm điều biến, hoàn điều biết và gói tín hiệu vào không gian đồng thời giữ cho việc truyền tín hiệu được mạnh trong môi trường nhiễu.

**Tầng MAC:**  sử dụng như công nghệ đa truy cập nhận biết song mang CSMA để xác định hình dạng đường truyền để tránh va chạm xác định và xác định hình dạng mạng, giúp hệ thống mạnh và vững chắc.

**Tầng mạng – NWK** là 1 tầng phức tạp của ZigBee, giúp tìm kiếm, kết nối mạng và mở rộng hình dạng từ chuẩn 802.15.4 lên dạng lưới. Tầng này xác định đường truyền lên ZigBee, xác định địa chỉ ZigBee thay vì địa chỉ tầng MAC bên dưới.

1. *Cấu trúc các tầng của Zigbee.*

**Tầng hỗ trợ ứng dụng – APS** là tầng kết nối với tầng mạng và là nơi cài đặt những ứng dụng cần cho ZigBee, giúp lọc bớt các gói dữ liệu trùng lắp từ tầng mạng

**Tầng đối tượng thiết bị – ZDO** có trách nhiệm quản lý các thiết bị, định hình tầng hỗ trợ ứng dụng và tầng mạng, cho phép thiết bị tìm kiếm, quản lý các yêu cầu và xác định trạng thái của thiết bị.

**Tầng các đối tượng ứng dụng người dùng – APO:** là tầng mà ở đây người dùng tiếp xúc với thiết bị, tầng này cho phép người dùng có thể tuỳ biến thêm ứng dụng vào hệ thống.

* ***Các loại điểm trong mạng***

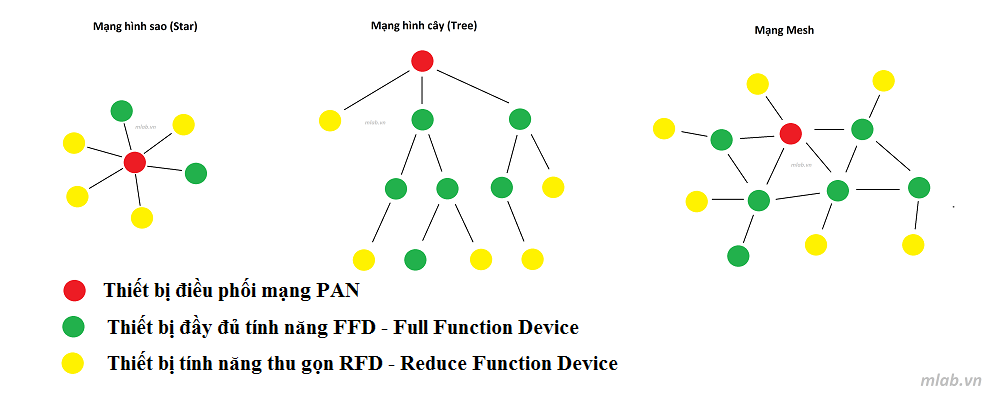
Thiết bị điều phối trung tâm – Zigbee Coordinator (ZC)

* Trong 1 mạng Zigbee, chỉ có một và duy nhất 1 Coordinator.
* Thiết lập mạng, địa chỉ mạng.
* Định tuyến đường truyền nếu trong mạng lưới (Mesh).
* Là gốc của Tree nếu trong mạng hình cây (Tree).
* Luôn hoạt động.
* Là thiết bị FFD.

Thiết bị Router – Zigbee Router (ZR)

* Tìm và tham gia mạng.
* Duy trì các gói tin thông qua mạng.
* Tham gia việc tìm đường: khám phá và duy trì đường đi.
* Cho phép các thiết bị khác tham gia mạng.
* Lưu các gói thông tin thay cho các Children đang ngủ.
* Luôn loạt động.
* Là thiết bị FFD.

Thiết bị End Device – Zigbee EndDevice (ZED)

* Tìm và tham gia mạng.
* Polling Parent của nó để kiểm tra có dữ liệu nào truyền đến nó hay không.
* Tìm một Parent mới nếu kết nối tới Parent cũ bị mất.
* Ngủ hầu hết thời gian để tiết kiệm năng lượng.
* Các thiết bị ZEDs không giao tiếp trực tiếp với nhau.
* Có thể coi là RFD.
* ***Có 3 loại cấu trục mạng***

1. *Các loại cấu trúc mạng.*

Mạng Zigbee chia làm 3 loại chính:

* **Mạng hình sao** (Star Topology): Gồm 1 nút trung tâm ZC, tất cả các nút khác đều được kết nối với nút trung tâm này, mạng hình sao bị hạn chế khoảng cách và sự mở rộng.
* **Mạng hình cây** (Cluster Tree Topology): Gồm một nút trung tâm ZC, các nút khác được liên kết với nhau theo mô hình giống một cái dễ cây, mạng này có khả năng mở rộng cao, tăng khoảng cách và quy mô của hệ thống.
* **Mạng hình lưới** (Mesh Topology): Gồm 1 nút trung tâm ZC, các nút trong mạng đều có thể kết nối với nhau (trừ ZED chỉ có thể kết nối với parent của nó). Khi một đường truyền bị lỗi, sẽ tự động tìm ra một đường truyền khác, tăng tính tin cây và kết nối trong mạng.

## MÔ HÌNH THUẬT TOÁN ARIMA

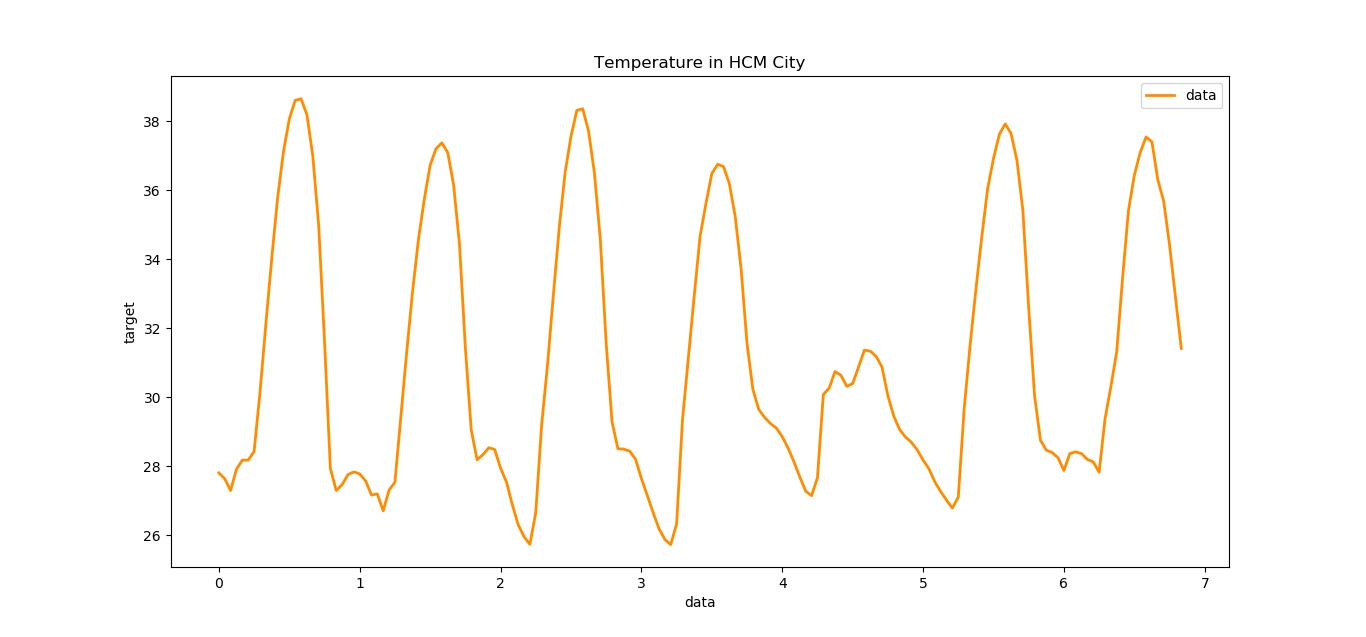
1. **Định nghĩa về chuỗi thời gian (Time Series)**

Chuỗi thời gian là một chuỗi các điểm dữ liệu, được đo theo từng khoảng khắc thời gian liền nhau theo một tần suất thời gian thống nhất. Một chuỗi thời gian có thể chia thành 4 phần chính: theo xu hướng (Trend), theo chu kỳ (Cycle), theo mùa (Seasonal), không theo quy tắc (Irregular).

Một chuỗi thời gian theo xu hướng (Trend) có thể tăng, giảm hoặc không tăng giảm trong một khoảng thời gian dài. Ví dụ: Đợt gia tăng dân số, số lượng nhà trong thành phố, dịch bệnh,…

Một chuỗi thời gian theo mùa (Seasonal) là một sự dao động trong vòng 1 năm, 1 tháng hoặc 1 ngày. Các yếu tố thay đổi theo mùa như: Khí hậu, thời tiết, phong tục, truyền thống,… Ví dụ: lượt nghe bài hát Last Christmas của Wham! tăng vào dịp Noel.

Một chuỗi thời gian thay đổi theo chu kỳ (Cycle) có thể bị tác động do hoàn cảnh, lặp lại theo chu kỳ. Thời gian của một chu kỳ kéo dài trong khoảng thời gian rất dài, thường là 2 năm hoặc hơn. Hầu hết các chuỗi thời gian về kính tế tài chính thường thay đổi theo chu kỳ.



1. *Chuỗi nhiệt độ đo tại TP. HCM (24/4/2019- 30/4/2019).*

Các biến động ngẫu nhiên (Irregular) trong một chuỗi thời gian được gây ra bởi những ảnh hưởng không thể đoán trước được, không thường xuyên xuất hiện và cũng không lặp lại nhiều lần trong một mô hình cụ thể. Những biến động này có thể là chiến tranh, động đất, thiên tai,… Không có kỹ thuật thống kê nào có thể đo lường các dao động ngẫu nhiên trong một chuỗi thời gian.

1. **Mô hình thuật toán ARMA**

Mô hình thuật toán ARMA (p, q) là sự kết hợp của mô hình AR (p) và MA (q), phù hợp với mô hình chuỗi thời gian đơn biến. Trong mô hình AR (p), giá trị tương lai của một biến được kết hợp giữa các biến dữ liệu trong quá khứ và sai số ngẫu nhiên (random error). Về mặt toán học, mô hình AR (p) có thể được biểu thị dưới dạng [1]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.1) |
|  |  |  |

Với  và tương ứng với giá trị thực và sai số ngẫu nhiên tại thời điểm .

là các thông số của mô hình.

là hằng số.

là cấp bậc của mô hình.

Giống như mô hình AR (p) đối với các dữ liệu trong quá khứ, mô hình MA (q) sử dụng các sai số trong quá khứ để làm rõ các biến. Mô hình MA (q) được biểu diễn dưới dạng [1]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.2) |
|  |  |  |

Với và tương ứng với giá trị thực và sai số ngẫu nhiên tại thời điểm .

là giá trị trung bình của dãy dữ liệu.

là các thông số của mô hình.

là cấp bậc của mô hình.

Mô hình Auto-Regressive (AR) và mô hình Moving-Average (MA) có thể kết hợp hiệu quả với nhau để tạo thành một mô hình chung và tốt hơn, được gọi là mô hình ARMA. Về mặt toán học, một mô hình ARMA (p, q) được biểu diễn dưới dạng [1]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.3) |
|  |  |  |

Với , lần lượt tương ứng với cấp bậc của mô hình AR, MA.

Thông thường mô hình ARMA được thao tác bằng cách sử dụng ký hiệu lag (). Toán tử lag được định nghĩa là . Đa thức lag được sử dụng để biểu diễn mô hình ARMA như sau [1]:

Mô hình AR (p): .

Mô hình MA (q): .

Mô hình ARMA (p, q): .

Với:

|  |  |  |
| --- | --- | --- |
|  |  | (2.4) |
|  |  | (2.5) |
|  |  |  |

1. **Mô hình thuật toán ARIMA**

Mô hình ARMA chỉ có thể được sử dụng cho chuỗi thời gian cố định (stationary data). Tuy nhiên, trong thực tế nhiều chuỗi thời gian như những thứ liên quan đến kinh tế xã hội và kinh doanh cho thấy sự biểu diễn không cố định (non-stationary). Chuỗi thời gian kết hợp Trend và Seasonal cũng không cố định. Do đó, mô hình ARMA không đủ để định hình chính xác chuỗi thời gian không cố định và cũng thường gặp trong thực tế. Vì lý do đó, mô hình ARIMA (Auto-Regressive Integrated Moving Average) được đề xuất, đó là sự tổng quát hóa của mô hình ARMA, bao gồm cả trường hợp không cố định. Trong mô hình ARIMA, một chuỗi thời gian không cố định được chuyển thành cố định bằng cách áp dụng phương pháp sai phân hữu hạn giữa các điểm dữ liệu. Công thức toán học của mô hình ARIMA (p, d, q) sử dụng đa thức lag được biểu diễn như sau [1]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.6) |

Thay: (2.4) và (2.5). Ta được:

|  |  |  |
| --- | --- | --- |
|  |  | (2.7) |

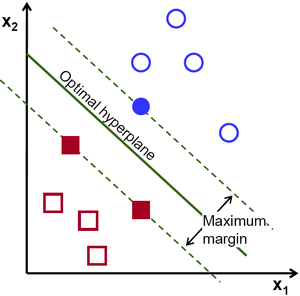
* Với p, d và q là các số nguyên lớn hơn hoặc bằng 0 và tương ứng với cấp bậc của các phần AR, I và MA.
* d là cấp bậc của vi phân. Với d = 1 có thể dùng trong hầu hết các trường hợp. Khi d = 0, thì ARIMA (p, 0, q) giảm xuống thành mô hình ARMA (p, q).
* Mô hình ARIMA (p, 0, 0) không khác gì ngoài mô hình AR (p) và ARIMA (0, 0, q) là mô hình MA (q).
* Mô hình ARIMA (0, 1, 0) là một loại mô hình đặc biệt, được gọi là mô hình Random Walk. Nó được sử dụng rộng rãi cho các dữ liệu không cố định.

## MÔ HÌNH THUẬT TOÁN SVMR

### Sơ lược về SVM (Support Vector Machine)

Support Vector Machine (SVM) là một loại thuật toán máy học có giám sát (supervised), có thể giải quyết được cả vấn đề phân loại (Classification) và hồi quy (Regression). Nhưng chủ yếu dùng để giải quyết vấn đề Classification.

SVM sẽ phân chia chính xác hai lớp (class) khác nhau, được mô tả bởi các điểm trong không gian nhiều chiều, bằng cách tìm một siêu phẳng (hyperplane) phân chia hai class đó, tức là các điểm thuộc một class sẽ nằm cùng một phía của siêu phẳng đó và các điểm thuộc class còn lại sẽ nằm trên phía còn lại của siêu phẳng. Đồng thời, thuật toán sẽ đi tìm siêu phẳng sao cho khoảng cách giữa các điểm gần siêu phẳng với siêu phẳng là lớn nhất. Khoảng cách đó gọi là margin.



1. *Phân tích thuật toán SVM.[[1]](#footnote-1)*

### Sơ lược về SVMr (Support Vector Machine-Regression)

SVM có thể được sử dụng để giải quyết vấn đề Regression, bao gồm các đặc trưng của thuật toán (margin lớn nhất). SVMr sử dụng đặc điểm tương tự SVM, chỉ với một vài khác biệt. Đầu tiên, ngõ ra là một con số rõ ràng nên việc dự đoán sẽ trở nên rất khó khăn vì có vô số kết quả. Trong trường hợp của Regression, giới hạn margin được đặt sấp xỉ thay vì phải đi tìm theo yêu cầu của vấn đề như SVM. Bên cạnh đó, thuật toán còn có một số lý do phức tạp hơn do đó cần phải được xem xét. Tuy nhiên, ý tưởng chính vẫn giống nhau: giảm thiểu lỗi, đi tìm siêu phẳng có margin lớn nhất.

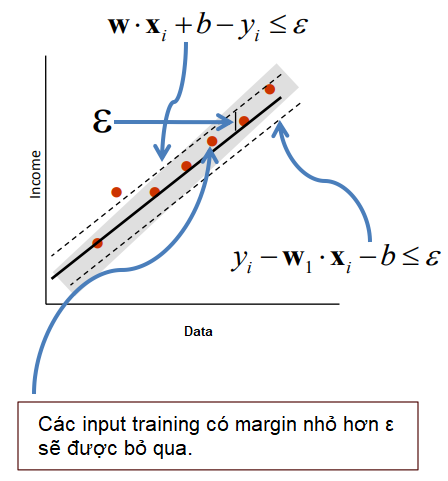
### Xây dựng bài toán cho SVMr

Ý tưởng của thuật toán là tìm một siêu phẳng , với độ lệch nhất định (ε) từ các input training y, có dạng phương trình [2]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.8) |

Bài toán tối ưu trong SVMr chính là bài toán tìm và b sao cho margin tại input training y tới siêu phẳng đạt giá trị lớn nhất [2]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.9) |

Phương trình trên có thể đưa về bài toán tối ưu có ràng buộc sau [2]:

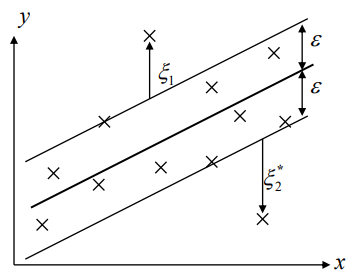
(2.10)

Với điều kiện:

1. *Phân tích thuật toán SVMr[[2]](#footnote-2).*

Việc xác định siêu phẳng ở trên được giả sử trong điều kiện lý tưởng khi các input training đều có margin nhỏ hơn hoặc bằng . Đối với những trường hợp đó, chúng ta cần sử dụng các biến slack .

(2.11)

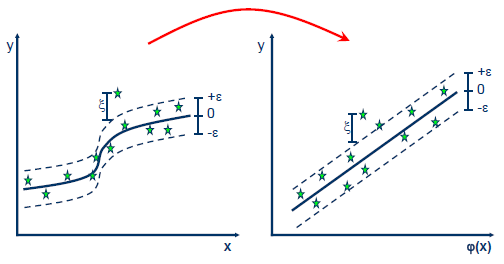
Với điều kiện:

Trong đó C là tham số xác định penalty degree.

1. *Phân tích thuật toán SVMr với tập dữ liệu có điểm gây nhiễu[[3]](#footnote-3).*

Trong trường hợp dữ liệu phi tuyến tính:

Khi dữ liệu bài toán là phi tuyến tính (non-linear), chúng ta phải sử dụng kernel ánh xạ dữ liệu sang không gian F nhiều chiều hơn để dữ liệu có thể biểu diễn sang dạng có tuyến tính (linear) [3].



1. *Kernel Trick[[4]](#footnote-4).*

Trong không gian F nhiều chiều hơn sẽ dẫn đến việc tính toán từng điểm dữ liệu tốn nhiều bộ nhớ và thời gian. Để có thể xử lý tính toán việc này dễ hơn, chúng ta sử dụng các hàm kernel function [2]:

|  |  |
| --- | --- |
| Polynomial: | (2.12) |
| Radial Basic Function: . | (2.13) |

## MÔ HÌNH KẾT HỢP (HYBRID)

Phương pháp kết hợp nhiều mô hình dự đoán khác nhau đã được nghiên cứu rộng rãi trong nhiều bài báo khác nhau. Một mô hình Hybrid có khả năng của cả linear và non-linear là một lựa chọn tốt cho việc dự đoán thời tiết. Mô hình Hybrid (Zt) có thể được biểu diễn như sau [3]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.14) |

Với là phần linear, là phần non-linear. Cả đều được đự đoán từ tập dữ liệu. Với là dữ liệu được dự đoán từ mô hình linear tại thời điểm t. biểu diễn cho sai số tại thời điểm t được thu thập từ mô hình linear. Khi đó [3]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.15) |

Các sai số đó sẽ được dự đoán từ mô hình non-linear và có thể biểu diễn như sau [3]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.16) |

Với f là hàm non-linear được tạo bởi mô hình non-linear và là sai số ngẫu nhiên. Từ đó, mô hình được kết hợp [3]:

|  |  |  |
| --- | --- | --- |
|  |  | (2.17) |

Với là kết quả dự đoán từ mô hình non-linear.

# CHƯƠNG 3

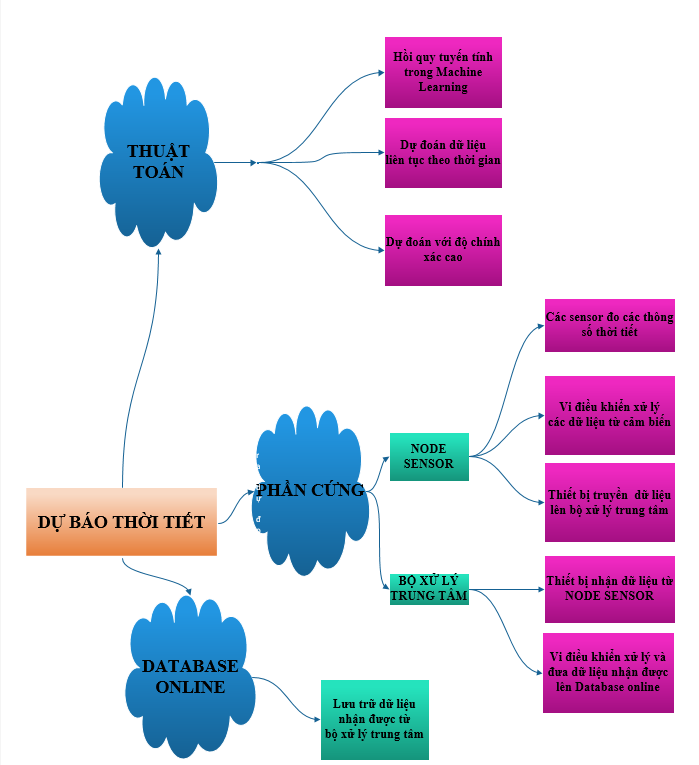
# THIẾT KẾ HỆ THỐNG

## HÌNH THÀNH Ý TƯỞNG

Việc dự báo thời tiết ở Việt Nam hiện nay vẫn chủ yếu dùng các thiết bị, kỹ thuật chưa hiện đại, dẫn đến nhiều sai số, khiến kết quả sai so với thực tế. Hiện nay công nghệ dự báo trên thế giới hoàn toàn có thể đưa ra các cảnh báo về thời tiết rất chính xác, thậm chí nhiều nơi có thể dự báo hằng giờ và một số bản tin có độ chính xác kéo dài từ ba đến năm ngày. Vậy làm thế nào để có thể dự đoán thời tiết nhanh và hiệu quả nhất? Trong đề tài này nhóm đã hình thành ra một ý tưởng để có thể khắc phục tình trạng dự báo thời tiết hiện nay bằng cách nghiên cứu và áp dụng một thuật toán sử dụng những tập dữ liệu thời tiết có sẵn và huấn luyện nó để nó có thể căn cứ vào đó dự đoán được thời tiết ở thời gian tương lai mà không phải đo đạc nhiều, giảm sai số ở mức thấp nhất có thể. Thuật toán nhóm sẽ áp dụng vào trong dự án này đó là SVMr, ARIMA dự báo thời tiết thuộc lớp bài toán dự báo hồi quy.

Để có được các dữ liệu thời tiết được đo trực tiếp bằng các cảm biến nhiệt độ, độ ẩm của không khí. Tuy nhiên, không phải lúc nào ở các vị trí khác nhau thì thông số thời tiết sẽ không đổi, chính vì vậy trong dự án này nhóm sẽ thiết kế nhiều Node Sensor đặt ở các vị trí khác nhau. Các thông số nhiệt độ, độ ẩm của từng Node sẽ được gửi lên bộ xử lý trung tâm, việc truyền nhận dữ liệu sẽ thông qua một thiết bị có thể giao tiếp theo chuẩn không dây với khoảng cách phù hợp. Bộ xử lý trung tâm sẽ đưa dữ liệu lên Database online thông qua một số giao thức. Dữ liệu trên Database online sẽ được huấn luyện (training) và từ đó đưa ra dự báo thời tiết để tưới tiêu một cách hợp lý.

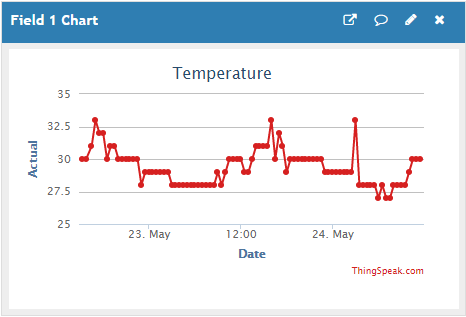
Vậy một hệ thống dự báo cần quan tâm những phần nào và quá trình hoạt động của chúng ra sao? Việc hình thành ý tưởng được coi như là bước đầu tiên và cũng là quá trình giúp chúng ta dễ dàng tổng quát hóa đề tài.



1. *Sơ đồ tư duy của hệ thống.*

Trong hệ thống này sẽ được chia làm ba phần chính kết hợp với nhau đó là: Phần cứng, Database Online và thuật toán. Việc lựa chọn phần cứng và thuật toán nhóm đã tìm hiểu, nghiên cứu các ưu điểm cũng như nhược điểm để đưa vào trong dự án này một cách hợp lý.

* **Phần cứng:** Bao gồm các Node Sensor và bộ xử lý trung tâm. Đối với Node Sensor sẽ có các cảm biến đo nhiệt độ và độ ẩm không khí. Việc đo nhiệt độ - độ ẩm của không khí nhóm sẽ chọn cảm biến DHT22, cảm biến này đo đạc với độ ổn định cao, có khả năng hoạt động liên tục trong thời gian dài, đồng thời dải nhiệt độ, độ ẩm có thể đo được rộng hơn so với những cảm biến khác. Việc truyền dữ liệu lên bộ xử lý trung tâm cần một Zigbee để có thể truyền dữ liệu đi, Zigbee là một giao thức được xây dựng theo chuẩn IEEE 802.15.4. Giao thức này được tạo ra nhằm phục vụ cho những ứng dụng yêu cầu giá thành và công suất thấp nhưng phải có khả năng linh động trong phạm vi rộng. Zigbee có đặc điểm là phạm vi hoạt động hẹp, tốc độ truyền Zigbee thích hợp cho các sensor không dây và chuyên dùng cho các ứng dụng giám sát, điều khiển. Có thể thấy rằng trong hệ thống dự báo thời tiết này, yêu cầu độ linh hoạt cao, giá thành thấp, tiêu thụ công suất nhỏ thì dùng chuẩn Zigbee là rất phù hợp. Đối với bộ xử lý trung tâm sẽ là một MCU có thể kết nối Wifi đó là ESP8266 và một Zigbee để nhận dữ liệu đo được từ Node Sensor. Bộ xử lý trung tâm sẽ thu thập dữ liệu và đưa lên DataBase online để lưu trữ.
* **DataBase Online:** Mọi thông tin về nhiệt độ - độ ẩm không khí sẽ được Zigbee thu thập, sau đó gửi lên Internet thông qua Wifi ở đây là Module ESP8266, những thông tin được gửi lên sẽ được lưu lại một server, server có sẵn được chọn là ThingSpeak. Từ điện thoại và máy tính ở bất kỳ đâu thì ta cũng có thể truy cập vào ThinkSpeak để quan sát được các thông số mà Zigbee đã gửi lên. ThingSpeak cho phép các cảm biến, các công cụ, và các Website để gửi dữ liệu lên đám mây (nơi mà dữ liệu được lưu trên một kênh riêng hoặc kênh công cộng). ThingSpeak lưu trữ dữ liệu trên các kênh riêng tư theo mặc định, nhưng những kênh công cộng có thể được dùng để chia sẻ dữ liệu với những người khác. Khi dữ liệu được lưu trữ trên ThingSpeak, chúng ta có thể phân tích và hình dung nó, tính toán dữ liệu mới, hoặc tương tác với phương tiện truyền thông, dịch vụ Web, và những thiết bị khác. Việc phân tích và tính toán dữ liệu trên ThingSpeak sẽ dễ dàng và trực quan hơn so với các Database online khác. ThingSpeak cung cấp quyền truy cập vào MATLAB giúp chúng ta có thể chuyển đổi, kết hợp và tính toán dữ liệu mới, đưa ra lịch trình tính toán để chạy vào những thời điểm nhất định, hiểu tương quan các mối quan hệ trong dữ liệu bằng các hàm vẽ có sẵn, kết hợp dữu liệu từ nhiều kênh khác nhau để xây dựng một phân tích tinh vi hơn.



1. *Minh họa dữ liệu trên Thingspeak.*

* **Thuật toán:** Với thuật toán SVMr, nhóm không sử dụng kernel Linear vì kernel đó không phù hợp với tập dữ liệu, kernel Poly thực hiện tính toán quá lâu, nhóm sẽ chọn kernel RBF vì kernel được sử dụng phổ biến cho mô hình SVMr và có thể xử lý tính toán nhanh. Với thuật toán ARIMA, tập dữ liệu cần phải được xử lý làm mịn trước khi xử lý thuật toán, sau khi làm xử lý thuật toán, tập dữ liệu cần phải chuyển đổi ngược về tập dữ liệu cũ. Với mô hình Hybrid, nhóm sẽ cho ARIMA thực hiện phần linear, còn SVMr sẽ thực hiện phần non-linear.

## MÔ HÌNH HỆ THỐNG

Trong phần này sẽ được giới thiệu khái quát về mô hình hệ thống. Hệ thống sẽ hoạt động hiệu quả, dễ nâng cấp, bảo trì nếu được lựa chọn mô hình phù hợp, ngược lại, hệ thống sẽ kém hiệu quả và khó nâng cấp, bảo trì nếu mô hình hệ thống không phù hợp. Do vậy việc xác định mô hình hệ thống phù hợp là rất quan trọng.



1. *Mô hình hệ thống.*

* Hệ thống các điểm cảm biến sẽ đo thông tin nhiệt độ, độ ẩm không khí và gửi về cho điểm trung tâm 60 phút mỗi lần.
* Điểm trung tâm sẽ tính trung bình các thông số đó và gửi về cơ sở dữ liệu ở máy chủ Thingspeak.com.
* Khối thuật toán sẽ lấy dữ liệu từ cơ sở dữ liệu trên máy chủ đó, để tính toán và dự đoán được các thông số nhiệt độ, độ ẩm không khí tiếp theo. Khi có các thông số dự đoán, hệ thống sẽ dựa vào đó để đưa ra quyết định có cần tưới nước cho cây hay không?
* Khi có hệ thống đưa ra quyết định cần tưới nước, sẽ gửi yêu cầu lên máy chủ Thingspeak.com.
* Điểm trung tâm nhận được yêu cầu sẽ mở van để tưới nước theo yêu cầu của hệ thống.

## THIẾT KẾ PHẦN CỨNG

## Chức năng phần cứng

Phần cứng nhóm thiết kế sẽ được chia làm 2 khối chính: khối trung tâm và khối Node Sensor.



1. *khối chính của phần cứng hệ thống.*

* **Khối trung tâm:** nhận các dữ liệu của các điểm cảm biến bằng thiết bị Zigbee và gửi về cơ sở dữ liệu thông qua ESP8266. Điểm trung tâm nhận yêu cầu tưới nước từ hệ thống, điểm trung tâm sẽ điều khiển mở van điện từ để tưới nước cho khu vườn.
* **Khối Node Sensor:** mỗi 60 phút điểm cảm biến sẽ thu thập các thông tin nhiệt độ, độ ẩm không khí từ cảm biến DHT22, sau đó sẽ gửi các thông số đó về cho điểm trung tâm bằng thiết bị Zigbee.

## Sơ dồ khối phần cứng

Các khối chi tiết của phần cứng hệ thống:

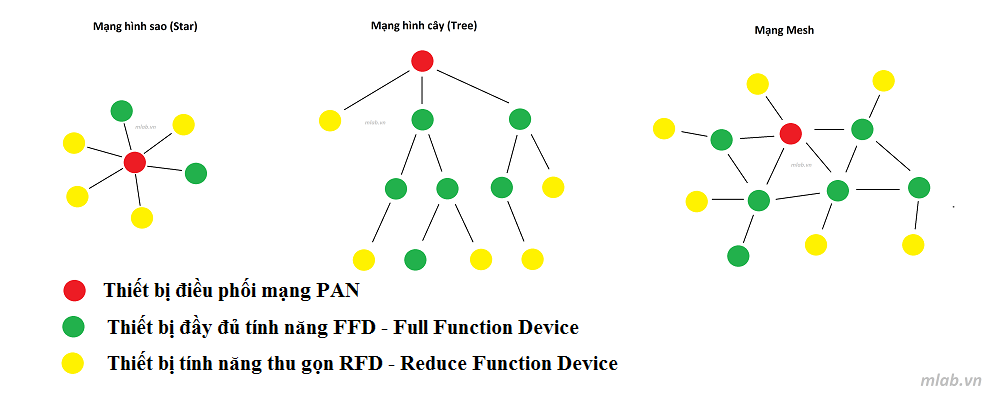


1. *Sơ đồ khối chi tiết của hệ thống.*

Từng khối trong hệ thống này sẽ được giải quyết các vấn đề thiết kế như sau:

* *Khối nguồn* có chức năng cung cấp nguồn cho tất cả các mạch trong hệ thống.
* *Khối điều khiển trung tâm* sẽ liên lạc giữa khối truyền nhận dữ liệu thông qua chuẩn truyền thông UART với khối Data thông qua wifi, nhận dữ liệu và đưa lên DataBase. Đồng thời khối trung tâm sẽ truyền tín hiệu cho khối điều khiển tưới nước của hệ thống.
* *Khối truyền nhận nhận dữ liệu* là một trạm trung gian để nhận dữ liệu từ khối Node Sensor cũng như truyền dữ liệu qua khối điều khiển trung tâm.
* *Khối Node Sensor* là một vi điều khiển xử lý dữ liệu thu được từ khối cảm biến. Vi điều khiển mà nhóm đã chọn trong đề tài này là STM32103C8.
* *Khối Sensor* bao gồm cảm biến nhiệt độ - độ ẩm của không khí DHT22. Các cảm biến này sẽ đọc tín hiệu Analog và đưa về cho khối Node Sensor xử lý.
* *Khối điều khiển tưới nước* ở đây sẽ là một cái van nước sẽ nhận tín hiệu từ khối điều khiển trung tâm, từ đó sẽ điều khiển van nước đóng mở phù hợp ở những thời điểm khác nhau.
* *Khối Data:* Để có thể dễ dàng quản lý và tính toán dữ liệu, nhóm đã chọn ThingSpeak làm DataBase Online chính trong hệ thống này.

Giải quyết vấn đề đa truy cập (tất cả các điểm cảm biến cùng gửi dữ liệu về điểm trung tâm cùng một lúc) như thế nào?



1. *Mô hình đa truy cập.*

Mô hình mạng nhóm đề xuất là mô hình hình sao, khi đó dữ liệu các điểm cảm biến sẽ gửi dữ liệu về điểm trung tâm. Mô hình này sẽ giúp hệ thống đơn giản và phù hợp cho hệ thống có ít điểm cảm biến (khoảng 2 đến 50 điểm cảm biến).

Để giải quyết vấn đề đa truy cập khi tất cả điểm cảm biến cùng gửi dữ liệu cùng lúc cho điểm trung tâm, nhóm thực hiện biện pháp hỏi từng điểm cảm biến để nhận dữ liệu. Tức khi điểm trung tâm cần dữ liệu sẽ gửi yêu cầu đến lần lượt từng điểm cảm biến, sau đó điểm cảm biến sẽ gửi trả dữ liệu về điểm trung tâm.

Nếu dữ liệu bị sai thì điểm trung tâm sẽ yêu cầu gửi lại. Khi điểm trung tâm nhận biết được dữ liệu của điểm cảm biến nào bị sai (nhận đúng địa chỉ điểm cảm biến, nhưng giá trị nhiệt độ, độ ẩm bị sai) thì điểm trung tâm sẽ gửi lại yêu cầu điểm cảm biến đó gửi lại dữ liệu cho tới khi đúng. Khi điểm trung tâm nhận được dữ liệu sai (địa chỉ điểm cảm biến sai, hoặc dữ liệu không đúng chuẩn) thì sẽ yêu cầu tất cả điểm cảm biến gửi lại (trừ những điểm cảm biến đã gửi đúng cả địa chỉ và giá trị cảm biến).

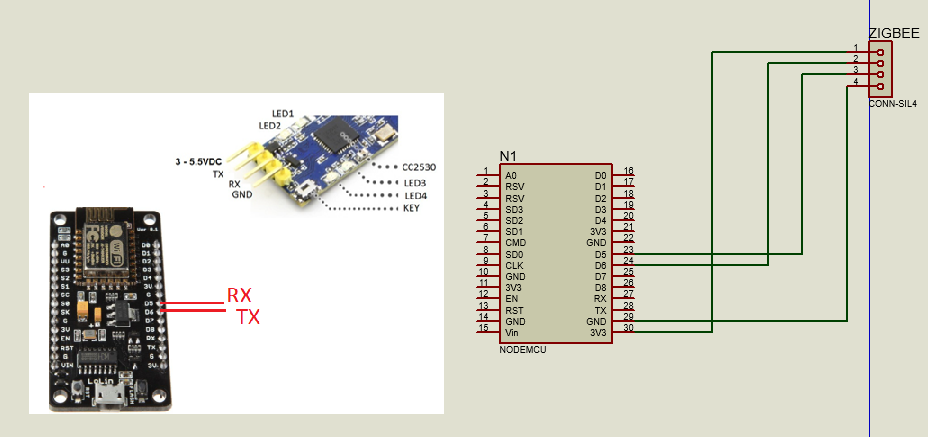
## Thiết kế từng khối

## Khối trung tâm



1. *Sơ đồ khối của khối trung tâm.*

Khối xử lý trung tâm là khối xử lý các tín hiệu nhận được từ các Node Sensor, từ đó nó sẽ tính toán và đưa những dữ liệu phù hợp lên DataBase Online để lưu trữ. Bộ xử lý trung tâm sử dụng NodeMCU ESP8266 được kết nối với module Zigbee theo chuẩn UART. Module Zigbee nhận dữ liệu đo được từ các các node cũng thông qua một module Zigbee nữa nhưng ở đây nó đóng vai trò là truyền dữ liệu đọc được từ các cảm biến. Việc giao tiếp giữa bộ điều khiển trung tâm và các Node Sensor đều được thông qua tín hiệu không dây theo chuẩn cấu hình ban đầu của Zigbee. Khoảng cách trung bình mà các Module Zigbee có thể truyền nhận được là 10-75m nên việc sử dụng nó trong một hệ thống dự báo thời tiết ở từng vùng khác nhau là hoàn toàn phù hợp.



1. *Khối kết nối phần cứng của khối xử lý trung tâm.*

Để thực hiện ý tưởng tưới nước cho cây trồng một cách hợp lý ở những vị trí và thời điểm khác nhau, bộ xử lý trung tâm được kết nối thêm với một thiết bị điều khiển tưới nước, nó hoạt động như một cái van nước bình thường nhưng ở đây nó sẽ đóng mở theo tín hiệu điều khiển của bộ điều khiển trung tâm. Từ các thông số của thời tiết như nhiệt độ, độ ẩm ở các thời gian và địa điểm khác nhau, bộ điều khiển trung tâm sẽ tính toán và xử lý để đưa ra tín hiệu số (digital) để điều khiển van đóng mở một cách hợp lý.

## Khối Node Sensor

Khối Node Sensor sẽ xử lý các tín hiệu nhận được từ các cảm biến để đưa ra các thông số thời tiết chính xác. Khối Node Sensor sử dụng một bộ xử lý riêng khác với bộ xử lý trung tâm, STM32F103C8T6 là một sự lựa chọn có thể nói là tối ưu nhất của nhóm thực hiện bởi giá thành cũng như giao tiếp dễ dàng với các module khác.



1. *Sơ đồ khối của khối Node Sensor.*

Bộ xử lý này giao tiếp với khối trung tâm thông qua sóng radio với tần số ngắn nên nó cũng được gắn với một module Zigbee. Có thể thấy việc sử dụng sóng radio để giao tiếp giữa các module với khoảng cách xa là rất hợp lý, dễ dàng trong việc lắp đặt hệ thống, cho dù các Node bất cứ ở đâu mà nằm trong vùng làm việc hiệu quả của Zigbee thì cũng đều nhận được tín hiệu tốt.



1. *Khối kết nối phần cứng của khối Node Sensor.*

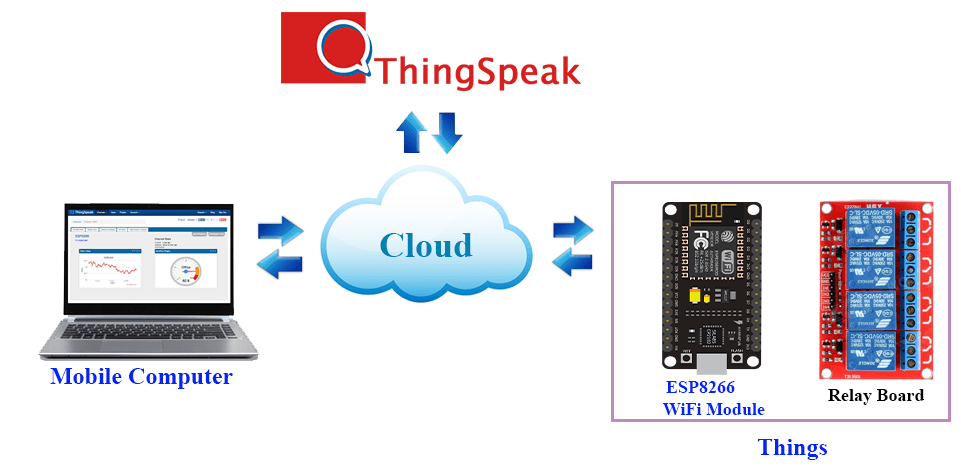
Để cho việc đo các thông số độ ẩm, nhiệt độ của không khí nhóm đã lựa chọn cảm biến DHT22. Việc đọc các thông số nhiệt độ và độ ẩm sẽ diễn ra liên tục trong 60 phút sẽ đọc các thông số này 5 phút một lần và sau đó sẽ tính trung bình gửi về cho STM32F103C8T6 xử lý và đưa lên bộ xử lý trung tâm. Trong quá trình các cảm biến gửi tín hiệu về cho vi điều khiển, đôi lúc sẽ bị lỗi thì ngay lúc đó vi điều khiển sẽ gửi lệnh yêu cầu cho đọc tín hiệu và gửi lại một lần nữa. Sau khi đã có được dữ liệu nhận được từ các cảm biến thì vi điều khiển ở khối Node Sensor này sẽ gửi lên bộ xử lý trung tâm để đưa lên server.

## XÂY DỰNG DATABASE ONLINE

Với dữ liệu đã lấy được từ phần cứng thì làm thế nào để lưu trữ và xử lý một cách dễ dàng và không tốn kém, ai cũng có thể quan sát được? Từ những câu hỏi trên, nhóm đã tìm hiểu và đưa ra quyết định sử dụng ThinkSpeak làm DataBase Online cho dự án này.

Thingspeak được đề xuất bởi ioBrigde vào năng 2010 như một dịch vụ hổ trợ các ứng dụng IoT.

Dữ liệu từ các cảm biến sẽ được MCU ESP8266 thu thập và gửi lên mạng Internet, công việc tiếp theo là nhận dữ liệu được gửi lên, xử lý và lưu trữ lại. Để thực hiện quá trình này cần có một không gian lưu trữ trên mạng Internet để lưu dữ liệu và một tên miền để tham chiếu tới không gian đang sử dụng.



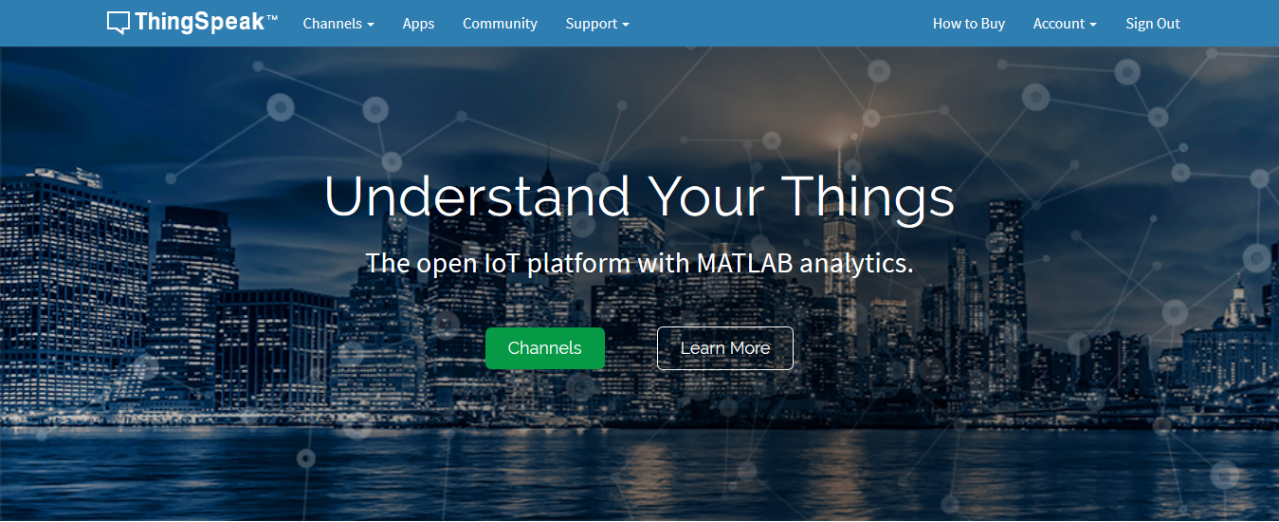
1. *Minh họa quá trình truyền dữ liệu lên ThingSpeak[[5]](#footnote-5).*

Thingspeak được tích hợp công cụ tính toán MATLAB từ Mathworks. Cho phép người dùng hình dung và phân tích dữ liệu được tải lên bằng matlab. Công cụ matlab trực tiếp trên Thingspeak server, do đó người dùng không phải tốn chi phí cho việc mua bản quyền matlab từ Mathworks cũng như phần cứng để chạy ứng dụng

Ta có thể thấy tất cả mọi thông tin về nhiệt độ, độ ẩm, ánh sáng,..sẽ được thu thập, sau đó gửi lên internet thông qua Wifi ở đây là module ESP83266, những thông tin được gửi lên sẽ được lưu lại trên một server, server có sẵn được chọn là ThinkSpeak. Từ điện thoại và máy tính bất kỳ đâu thì ta cũng có thể truy cập vào ThinkSpeak để quan sát được các thông số đã gửi lên.

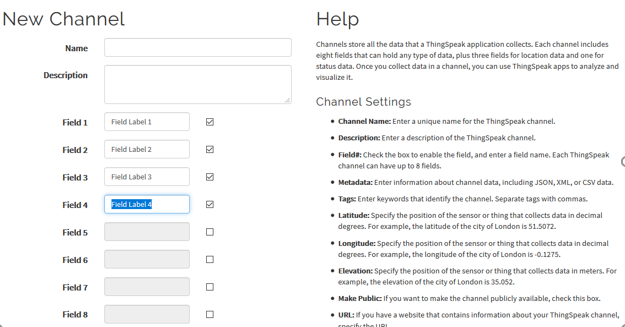
### Cách tạo tài khoản trên ThingSpeak

Thingspeak có mỗi quan hệ chặc chẽ với Mathworks, Inc. Trên thực tế, tất cả các tài liệu của Thingspeak tích hợp vào trang web tài liệu trên Mathworks và thậm chí người dùng có thể đăng nhập vào Thingspeak bằng tài khoản của Mathworks. Để đăng nhập trên ThingSpeak thì chúng ta phải có tài khoản MathWork, do đó mình phải tạo một tài khoản bên đó rồi đăng nhập bên ThingSpeak [5].



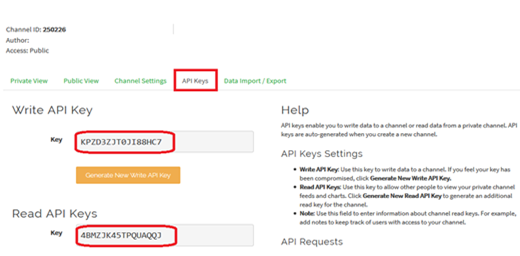
1. *Giao diện trang chủ ThingSpeak.*

Sau khi tạo tài khoản và đăng nhập thành công, chọn New Channel để tạo kênh mới nhận dữ liệu.



1. *Tạo NewChannel.*

Điền tên, mô tả, chọn số Field nhận dữ liệu. Thingspeak cung cấp tối đa 8 Field cho mỗi channel và có thể tạo nhiều channel trong một tài khoản. Sau khi điền đầy đủ thông tin, chọn Save Channel ở cuối trang.

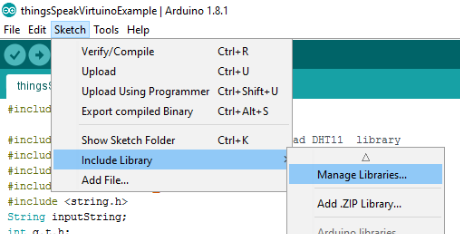


1. *API Keys.*

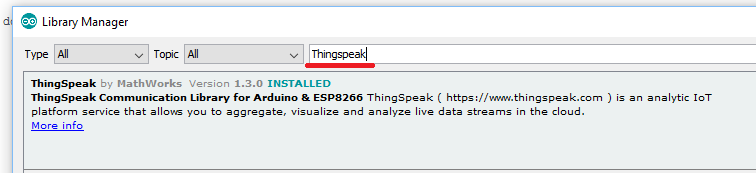
Sau khi tạo channel, phần thông tin quan trọng nhất của channel là API Keys. Với hai mã Read API Key và Write API Key cho phép mọi thiết bị IoTs điều có thể Read và Write dữ liệu lên.

### Hướng dẫn gửi dữ liệu từ ESP8266

**Bước 1:** Thêm thư viện ThngSpeak cho NodeMCU.



1. *Mở cửa sổ Manage Libraries.*



1. *Tìm kiếm thư viện.*

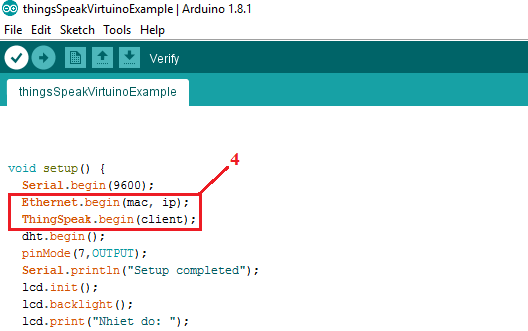
Ở cửa sổ Library Manager, gõ từ khóa “Thingspeak” vào thanh tìm kiếm để tiến tìm và tải thư viện.

**Bước 2:** Kết nối NodeMCU đến ThingSpeak

Để kết nối ESP8266 đến Thingspeak trước tiên phải thực hiện một số khai báo cần thiết:

1. Thêm thư viện Thingspeak vào đầu chương trình.
2. Khai báo địa chi IP vào địa chỉ mac. Địa chỉ mac ở đây chính là địa chỉ mac của Ethernet Shield( board mở rộng giúp Arduino Mega kết nối internet).
3. Khai báo thông tin channel và mã Write API key. Ở đây nhóm thực hiện đề tài sử dụng hai channel. Một là channel Sensor làm nhiệm vụ nhận và hiển thị các giá trị do cảm biến gửi lên. Hai là channel Command, Arduino Mega sẽ đọc giá trị ở Field 1 thuộc channel Commad để cập nhật chế độ hoạt động và gửi về cho khối cảm biến.
4. Khởi động Thingspeak cho chương trình.

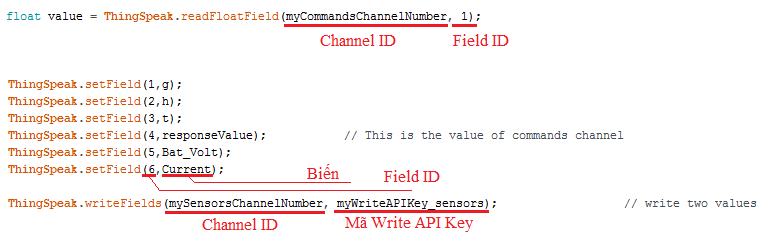


1. *Khai báo biến cho ThingSpeak.*
2. *Khởi động ThingSpeak.*

Sau khi khai báo các biến cần thiết cho Thingspeak và khởi động Thingspeak cho chương trình, lúc này chường trình đã sẵn sàng kết nối đến Thingspeak server.

**Bước 3:** Gửi nhận dữ liệu lên Thingspeak

* Đọc dữ liệu: sử dụng hàm readFloatField(ChannelID, FiealID). Với các tham số: Channel ID là sô Channel được Thingspeak cung cấp, Field ID là số thứ tự field trong channel.



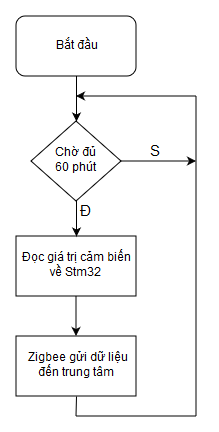
1. *Gửi nhận dữ liệu lên ThingSpeak.*

* Gửi dữ liệu: Sử dụng hàm SetField( FielID, Biến), và hàm writeFields( ChannelID, WriteAPIKey) để gửi dữ liệu lên channel. Trong đó:
* Biến: giá trị muốn gửi lên Thingspeak.
* WriteAPIKey: là mã Write API Key của channel Sensors.

## THIẾT KẾ PHẦN MỀM

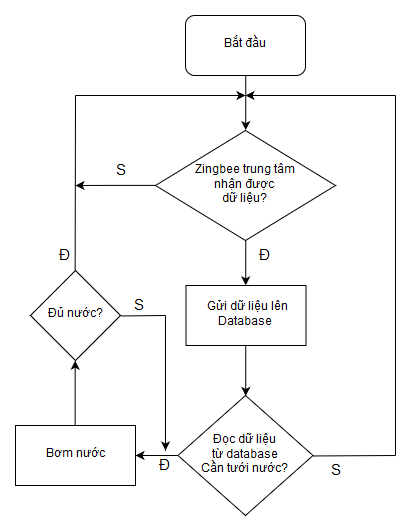
Thiết kế phần mềm là một quá trình giải quyết vấn đề và lập kế hoạch cho giải pháp của thuật toán. Sau khi các mục đích và đặc điểm kĩ thuật của phần mềm được quyết định. Mục tiêu của thiết kế này là một quá trình chuyển hóa các yêu cầu phần mềm một cách hiệu quả

Hệ thống phần cứng bao gồm 2 khối chính bao gồm khối Node Sensor và khối xử lý trung tâm như đã nói ở phần thiết kế phần cứng. Trong phần này sẽ mô tả cách hoạt động của phần cứng bằng phần mềm. Khối Node Sensor sẽ làm rõ vấn đề đọc dữ liệu từ các cảm biến để gửi lên bộ xử lý trung tâm. Đối với khối xử lý trung tâm sẽ làm rõ vấn đề cách thức nhận dữ liệu từ các điểm cảm biến và gửi dữ liệu lên Database để xử lý. Hoạt động chi tiết của 2 khối trong hệ thống này sẽ được biểu diễn qua các lưu đồ dưới đây:

*  Khối Node Sensor hoạt động theo lưu đồ sau:

1. *Lưu đồ chính của bộ xử lý trung tâm.*

Giải thích lưu đồ: Đây là một vòng lặp hoạt động của vi điều khiển (STM32F103C8T6). Đầu tiên kit sẽ delay một khoảng thời gian 60 phút để cho phần cứng ổn định, sau khi hết 60 phút vi điều khiển sẽ đọc các giá trị Analog của cảm biến đo được, rồi xử lý tính toán và xử lý đưa ra các thông số nhiệt độ, độ ẩm. Sau đó Zigbee được kết nối với vi điều khiển thông qua chuẩn UART, sau khi đã gửi các dữ liệu đo được thì lại tiếp tục chờ 60 phút sau đọc giá trị cảm biến và gửi dữ liệu lên thông qua Zigbee. Như vậy theo như ở trên lưu đồ, cứ mỗi 60 phút sẽ cập nhật giá trị nhiệt độ, độ ẩm.

* **Khối xử lý trung tâm hoạt động theo lưu đồ sau:

1. *Lưu đồ khối xử lý trung tâm.*

Giải thích lưu đồ: Khi Zigbee nhận dữ liệu nhiệt độ, độ ẩm từ một Zigbee ở bên khối Node Sensor, nếu việc nhận dữ liệu hay bên khối Node Sensor truyền dữ liệu bị lỗi thì sẽ được yêu cầu thực thi lại, dữ liệu nhận được không bị lỗi sẽ được gửi lên database online thông qua ESP8266. Thiết bị ngọai vi được kết nối với vi điều khiển của bộ xử lý trung tâm là máy bơm nước. Sau khi gửi lên dữ liệu lên Database online, vi điều khiển sẽ đọc dữ liệu từ DataBase online, từ đó đưa ra tín hiệu điều khiển máy bơm nước đóng mở van để tưới cây hợp lý. Nếu dữ liệu đưa ra với các thông số nhiệt độ, độ ẩm phù hợp với điều kiện cây trồng cần nước thì sẽ cho phép bơm nước với lượng nước vừa đủ, sau khi bơm đủ lượng nước sẽ quay lại nhận dữ liệu từ Zigbee và trường hợp các thông số không phù hợp với điều kiện cây trồng thì sẽ cũng quay lại chu kỳ đọc nhiệt độ, độ ẩm.

## THIẾT KẾ THUẬT TOÁN

Với 3 mô hình thuật toán khác nhau, thuật toán nào có khả năng dự đoán chính xác cao nhất? Làm thế nào thực hiện 3 mô hình thuật toán đó?

* Mô hình SVMr

1. *Lưu đồ thuật toán SVMr.*

Giải thích:

* Thuật toán SVMr sử dụng kernel RBF với nhiều thông số khác nhau: C = {1, 10, 100} và gamma = {0.1, 0.2,…, 10}.
* Tính toán độ chính xác giữa các dữ liệu dự đoán và test data theo nhiều thông số C, gamma khác nhau để tìm ra thông số tối ưu nhất.
* Khi áp dụng thuật toán vào thực nghiệm, nhóm sẽ chọn thông số C và gamma tốt nhất.
* Mô hình ARIMA



1. *Lưu đồ thuật toán ARIMA.*

Giải thích:

* Xử lý làm mịn training data làm ổn định tập dữ liệu, giúp đơn giản hóa dữ liệu để việc xử lý thuật toán tăng độ chính xác hơn. Việc xử lý có nhiều cách để thực hiện: vi phân dữ liệu, căn bậc 2 dữ liệu,…
* Sau khi có kết quả dự đoán cần phải chuyển kết quả dự đoán về dạng ban đầu, vì training data dùng để dự đoán đã được xử lý.
* Mô hình Hybrid

1. *Lưu đồ mô hình Hybrid.*

Giải thích:

* Trong mô hình Hybrid, phần linear sẽ được xử lỷ bởi thuật toán ARIMA, còn phần non-linear sẽ được xử lý bởi thuật toán SVMr.
* Cũng như ARIMA, training data cần được xử lý làm mịn.
* Trong lúc thực hiện thuật toán ARIMA, thuật toán sẽ tính toán các sai số ở các training data gọi là residual. Sau đó, Thuật toán sẽ tiếp tục dự đoán các dữ liệu tương lai.
* Thuật toán SVMr sẽ sử dụng các residual làm training data và thực hiện dự đoán các residual tương lai.
* Kết quả dự đoán tương lai cộng với các residual tương lai để cho kết quả dự đoán cuối cùng.
* Lưu đồ hoạt động nhận dữ liệu từ database để thực hiện thuật toán như sau:

1. *Lưu đồ thuật toán.*

Giải thích:

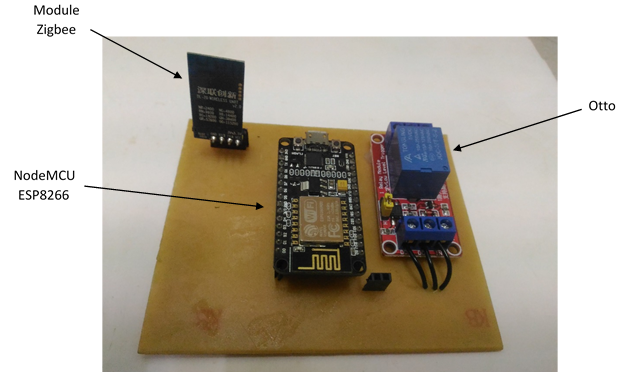
* Chương trình đọc dữ liệu từ database Thingspeak và lưu giá trị index.
* Chương trình kiểm tra có dữ liệu mới bằng cách so sánh giá trị index được lưu trước đó với giá trị index hiện tại. Nếu trước đó không có index (chưa từng chạy chương trình bao giờ) thì chương trình mặc định thực hiện thuật toán.
* Chương trình lọc các dữ liệu lỗi (nhiều dữ liệu cùng 1 giờ, dữ liệu bị gửi sai giờ,…) và tạo thành chuỗi dữ liệu để chuẩn bị chạy thuật toán.
* Sau khi tạo được chuỗi dữ liệu, nhóm chọn thuật toán tốt nhất để thực hiện dự đoán dựa vào các thông số đánh giá độ chính xác.
* Gửi dữ liệu dự đoán lên Thingspeak để phần cứng thực hiện công việc tưới cây nếu kết quả dữ liệu vượt ngưỡng.

# CHƯƠNG 4

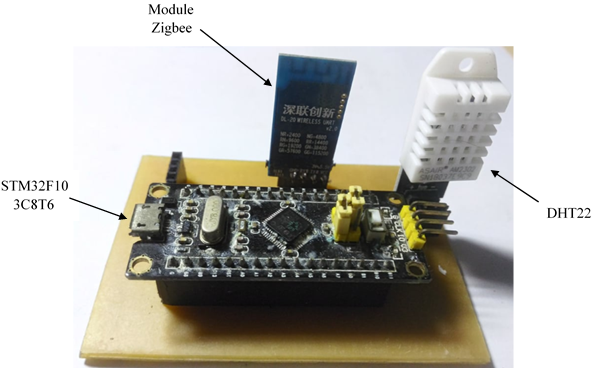
# KẾT QUẢ THỰC NGHIỆM VÀ ĐÁNH GIÁ

## MÔ HÌNH THỰC NGHIỆM

Khối điều khiển trung tâm:



Opto

1. *Khối điều khiển trung tâm.*
2. *Khối điểm cảm biến*
3. *Khối điểm cảm biến.*

## THỰC NGHIỆM

### Kiểm tra tính năng hoạt động của phần cứng

Nhóm đã tiến hành kiểm tra hoạt động của các điểm cảm biến và bộ xử lý trung

tâm có đúng như của nhà sản xuất cũng như phù hợp với môi trường thực tế. Nhóm đã đánh giá độ nhạy của cảm biến DHT22 với sự thay đổi của môi trường (hơi nước đang sôi) trong 20s đầu kể từ khi bắt dầu đo. Sau khi khảo sát đã cho chúng ta thấy độ chính xác +/-0.5 độ, còn đối với độ ẩm thì độ chính xác 2-5%.

Kiểm tra khoảng cách giữa các điểm cảm biến, các điểm cảm biến đặt ở những vị trí khác nhau có khoảng cách 20 m so với bộ xử lý trung tâm, nhóm đã đưa ra kết luận việc dùng Zigbee để truyền nhận dữ liệu trong khu vườn hay nông trại thì rất phù hợp, dữ liệu truyền nhận được ổn định và ít xảy ra vấn đề trong lúc hệ thống hoạt động.

|  |  |
| --- | --- |
|  |  |

1. *Biểu đồ thể hiện khả năng đáp ứng của cảm biến đối với sự thay đổi nhiệt độ, độ ẩm.*
2. *Hệ thống dự báo thời tiết thực tế*

Trong các điều kiện thời tiết nhóm đã tiến hành chạy thực tế sản phẩm để đánh giá chính xác.

Khối điều khiển trung tâm hoạt động ổn định, dữ liệu sau khi nhận được được gửi lên ThingSpeak đúng với nhiệt độ, độ ẩm đọc được từ các cảm biến. Khối cảm biến đọc nhiệt độ tương đối ổn định, có đôi lúc bị lỗi thì STM32F103C8T6 sẽ yêu cầu đọc và gửi lại. Từ dữ liệu nhiệt độ và độ ẩm đọc được sẵn sàng chuẩn bị gửi lên bộ xử lý trung tâm.

### Thực nghiệm thuật toán

Với phần cứng và thuật toán được sử dụng, đánh giá hệ thống dự đoán trên đại lượng thước đo nào?

Dữ liệu nhiệt độ được sử dụng tính toán được đo tại Thành phố Hồ Chí Minh bao gồm dữ liệu trên Internet (từ 0h 14/5/2019 đến 6h 30/5/2019) và dữ liệu đo được lưu trữ trên database của nhóm (từ 7h 30/5/2019 đến 23h 3/6/2019). Với tập dữ liệu này, nhóm phân tích thành 3 phần như bảng 4.1. Với training data dùng để chạy thuật toán với nhiều thông số khác nhau, validation data dùng để theo dõi độ chính xác theo từng thông số của thuật toán và test data dùng để đánh giá độ chính xác của thuật toán.

|  |  |  |
| --- | --- | --- |
| Training data | Validation data | Test data |
| 0h 14/5/2019 –  11h 3/6/2019 | 12h 3/6/2019 –  17h 3/6/2019 | 18h 3/6/2019 –  23h 3/6/2019 |

1. *Phân tích dữ liệu thành 3 phần*.

Kết quả độ chính xác từ 3 phương pháp sẽ được tính thông qua công thức MAE (Mean Absolution Error), MSE (Mean Square Error), MAPE (Mean Absolution Percent Error), RMSE (Root Mean Square Error) [1]:

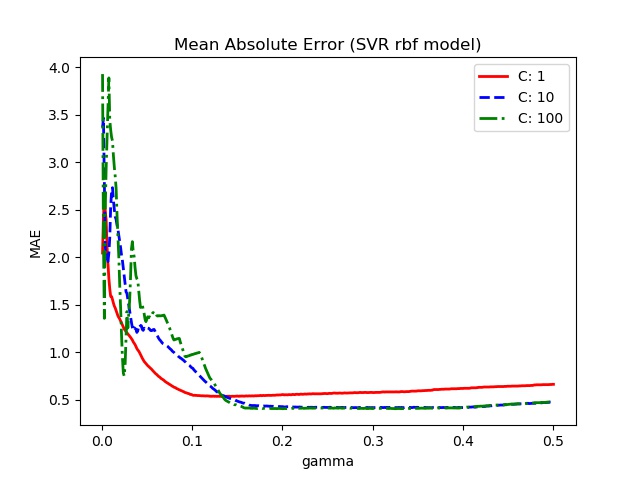
|  |  |  |
| --- | --- | --- |
|  |  | (4.1) |
|  |  | (4.2) |
|  |  | (4.3) |
|  |  | (4.4) |

Với N là tổng số các mẫu được dự đoán.

dt là giá trị nhiệt độ thực tại thời điểm t

zt là giá trị nhiệt độ dự đoán tại thời điểm t.

Với thuật toán SVMr, Nhóm sử dụng kernel RBF bao gồm các thông số C, gamma cần phải được cân chỉnh để thuật toán được tối ưu nhất. Độ chính xác MAE theo từng thông số được biểu diễn theo như hình dưới đây:



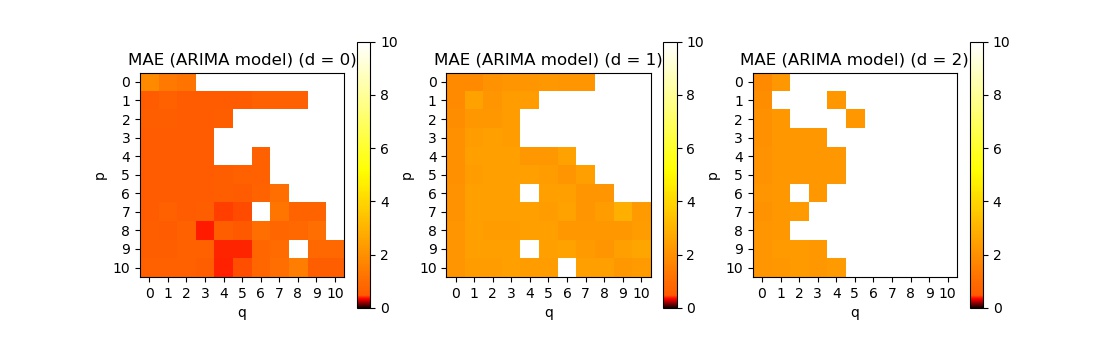
1. *MAE của SVMr.*

Với Trục Y là giá trị của MAE.

Trục X là giá trị gamma chạy từ 0.00 đến 0.50 với độ chia là 0.01.

Theo như hình 4.4, với C=1, đường màu đỏ, cho kết quả MAE ổn định trong khoảng gamma từ 0.1 đến 0.5 nhưng MAE không được nhỏ. Cũng như C=1, C=10 và C=100, lần lượt đường màu xanh dương và xanh lá, cũng cho kết quả MAE ổn định trong khoảng gamma từ 0.1 đến 0.5 và có MAE rất nhỏ. C=100 có MAE nhỏ nhất nhưng với gamma nhỏ có thể dẫn đến MAE rất cao.

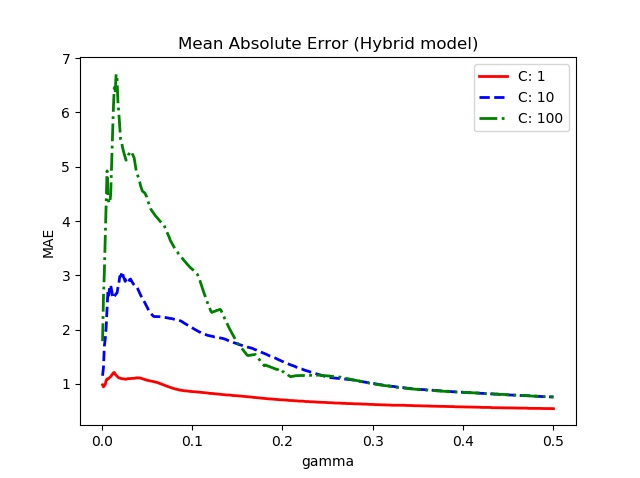
Với thuật toán ARIMA bao gồm các thông số p, d, q cần phải được cân chỉnh để phù hợp với thuật toán. Độ chính xác MAE theo từng thông số được biểu diễn theo như hình dưới đây:



1. *MAE của ARIMA.*

Vì thuật toán sử dụng 3 thông số nên nhóm tách thành 3 hình tương ứng với mỗi thông số d để dễ theo dõi. Kết quả MAE càng nhỏ thì ô màu tại thông số (p, d, q) càng đậm.

Với mô hình Hybrid, Nhóm sẽ sử dụng thông số (p, d, q) tốt nhất dựa trên hình 4.5 để chạy thuật toán ARIMA. Sau đó, sử dụng các residual cân chỉnh các thông số C, gamma cho thuật toán SVMr.



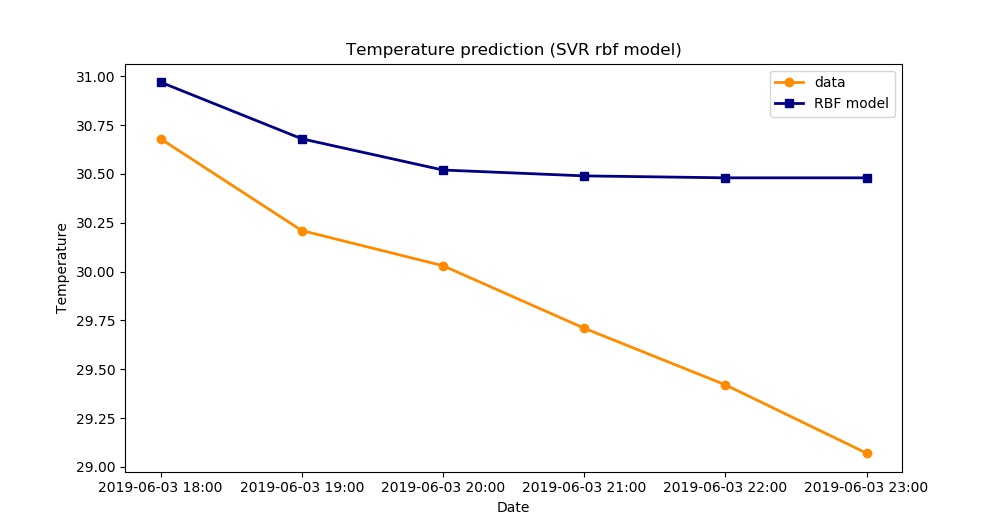
1. *MAE residual của Hybrid model.*

Với những thông số trên, nhóm sẽ bắt đầu tiến hành đánh giá độ chính xác trên tập test data.

1. *Bảng dưới đây liệt kê các thông số tốt nhất của 3 mô hình*.

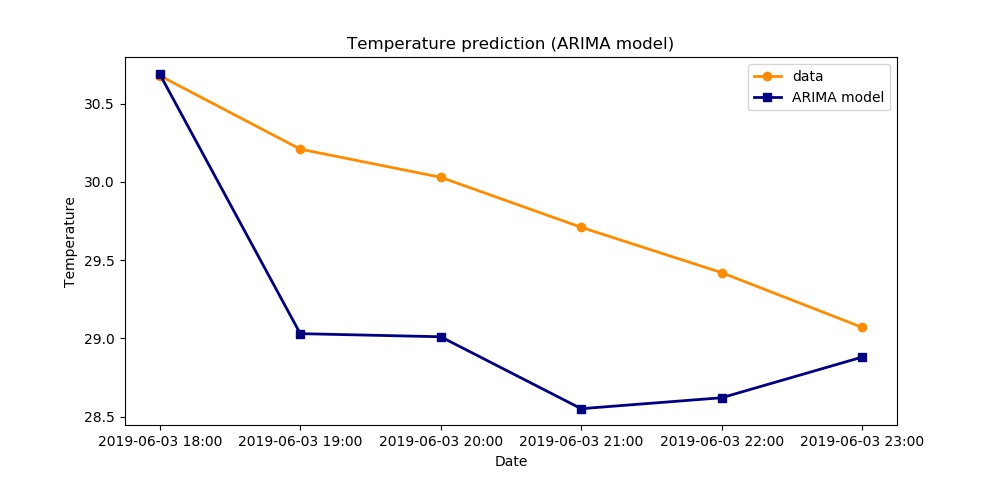
|  |  |  |
| --- | --- | --- |
| Mô hình SVMr | Mô hình ARIMA | Mô hình Hybrid |
| C = 100  gamma = 0.32 | (p, d, q) = (8, 0, 3) | (p, d, q) = (8 , 0, 3)  C = 1  gamma = 0.50 |

Kết quả dự đoán từ mô hình SVMr:



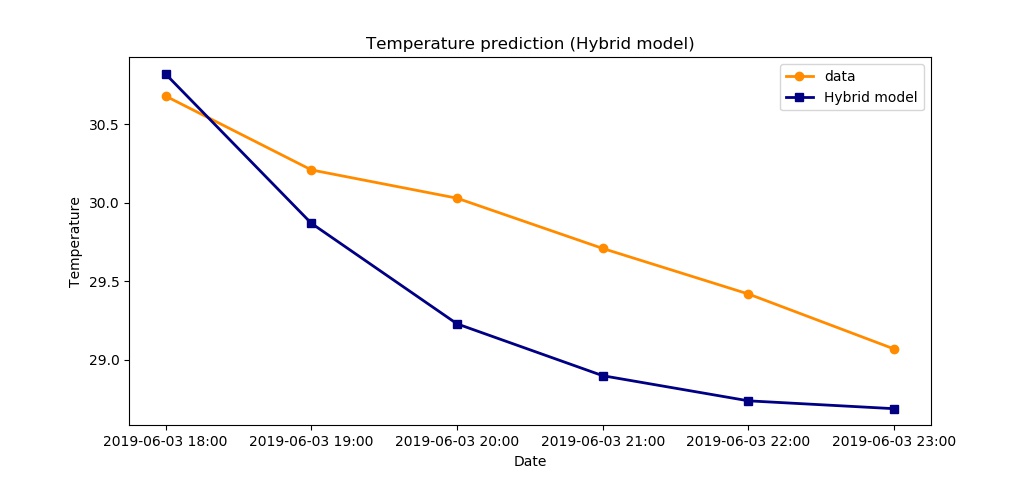
1. *Kết quả của SVMr model.*

Kết quả dự đoán từ mô hình ARIMA:



1. *Kết quả của ARIMA model.*

Kết quả dự đoán từ mô hình Hybrid:



1. *Kết quả của hybrid model.*
2. *Bảng so sánh độ chính xác giữa 3 mô hình.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mô hình | MAE | MSE | MAPE | RMSE |
| SVMr | 0.750 | 0.711 | 2.54% | 0.843 |
| ARIMA | 0.727 | 0.742 | 2.435% | 0.862 |
| Hybrid | 0.525 | 0.34 | 1.765% | 0.583 |

Với kết quả độ sai số từ bảng 4.3, cả 3 mô hình đều thực hiện dự đoán rất tốt. Với việc kết hợp 2 mô hình nhau, mô hình Hybrid thực hiện dự đoán tốt nhất. So với đề tài trước đó (Kết quả sai số nhiệt độ: ±2°C) [4], mô hình thuật toán có kết quả tốt hơn.

## ƯU ĐIỂM

* Có mức tiêu thụ năng lượng thấp.
* Xây dựng được hệ thống truyền nhận bằng công nghệ Zigbee với khoảng cách xa.
* Tìm hiểu được mô hình năng lượng của hệ thống, qua đó thiết lập được các chế độ hoạt động của khối cảm biến giúp tiết kiệm năng lượng.
* Khối cảm biến có khả năng mở rộng phạm vi và số cảm biến.
* Hệ thống có thể tự động cập nhận thông tin trên trang web.
* Có khả năng cập nhật thông tin tình hình thời tiết cho cộng động.

## NHƯỢC ĐIỂM

* Chất lượng và độ nhạy của các cảm biến không cao, nên đôi khi hoạt động không ổn định.
* Vì đây là hệ thống dự báo thời tiết nên luôn đặt ở ngoài trời, khi thời tiết rất xấu như mưa to, gió lớn có thể lam hỏng hệ thống làm gián đoạn việc lấy dữ liệu từ các cảm biến.
* Chưa đưa vi điều khiển điểm cảm biến vào trạng thái tiết kiệm năng lượng.
* Thông số thuật toán được chọn vẫn còn chưa chính xác, nên trong một số trường hợp sẽ dự báo sai.

# CHƯƠNG 5

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## KẾT LUẬN

### Những vấn đề nghiên cứu

* Nghiên cứu xây dựng hệ thống dự báo thời tiết bằng thuật toán máy học phù hợp với thực tế.
* Nghiên cứu về dữ liệu online.
* Nghiên cứu về mô hình năng lượng hệ thống.

### Những vấn đề hoàn thành

Xây dựng thành công hệ thống dự báo thời tiết bằng thuật toán máy học.

Các dữ liệu nhận được từ các cảm biến được vi điểu khiển của khối điểm cảm biến đưa lên bộ xử lý trung tâm rồi từ đó đưa lên ThingSpeak.

Thuật toán để dự đoán dữ liệu các thông số thời tiết đã đáp ứng được yêu cầu đặt ra, dự đoán với độ chính xác cao.

## HẠN CHẾ

* Vẫn còn sử dụng một số module mua ngoài.
* Chưa tích hợp được còi báo động trên khối điều khiển trung tâm khi các chỉ số vượt ngưỡng hoặc có sự cố.
* Hệ thống đôi lúc truyền nhận dữ liệu không ổn định.

## HƯỚNG PHÁT TRIỂN

Từ những vấn đề nhóm đã hoàn thành và các hạn chế của đề tài, nhóm đề xuất ra hướng phát triển như sau:

* Có thể tăng thêm số lượng khối điểm cảm biến.
* Gia tăng số cảm biến trong khối điểm cảm biến.
* Nâng cao khả năng dự đoán thời tiết bằng các loại cảm biến tốt hơn.
* Xây dựng ứng dụng giao tiếp người dùng cho nhiều nền tảng hệ điều hành khác nhau.
* Thêm tính năng cảnh báo khi có sự cố cháy xảy ra như còi báo động ở khối trung tâm hoặc chức năng gọi điện cho người nông dân.
* Nên sử dụng mạch pin để tiết kiệm năng lượng cho các điểm cảm biến.
* Phát triển thuật toán tối ưu hơn, hoặc kết hợp các thuật toán hỗ trợ khác, cải thiện hơn cho ARIMA và SVMr.

# TÀI LIỆU THAM KHẢO

|  |  |
| --- | --- |
| [1] | R. Adhikari và R. K. Agrawal, trong *An Introductory Study on Time series Modeling and Forecasting*, LAP Lambert Academic Publishing, 2013. |
| [2] | A. . J. Smola và B. Schölkopf, “A tutorial on support vector regression,” *Statistics and Computing,* tập 14, số 3, pp. 199-222, 2004. |
| [3] | P.-F. Pai và C.-S. Lin, “A hybrid ARIMA and support vector machines model in stock price forecasting,” *Omega,* tập 33, số 6, pp. 497-505, 2004. |
| [4] | T. Truong, A. Dinh và K. Wahid, “An IoT environmental data collection system for fungal detection in crop fields,” trong *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, Windsor, Ontario, 2017. |
| [5] | TuanPM, “Serial - Hardware Serial,” [Trực tuyến]. Available: https://arduino.esp8266.vn/peripherals/serial.html. [Đã truy cập 23 3 2013]. |
| [6] | I. Bhattacharyya, “Support Vector Regression Or SVR,” Coinmonks, 29 6 2018. [Trực tuyến]. Available: https://medium.com/coinmonks/support-vector-regression-or-svr-8eb3acf6d0ff. [Đã truy cập 21 3 2019]. |
| [7] | TuanPM, “Đo nhiệt độ, độ ẩm và gởi lên Thingspeak,” [Trực tuyến]. Available: https://arduino.esp8266.vn/projects/dht11-thingspeak.html. [Đã truy cập 22 3 2019]. |
| [8] | P. V. Hoàn, Vi xử lý nâng cao, Thành phố Hồ Chí Minh: Trường ĐH Sư phạm Kỹ thuật TP.HCM, 2016. |

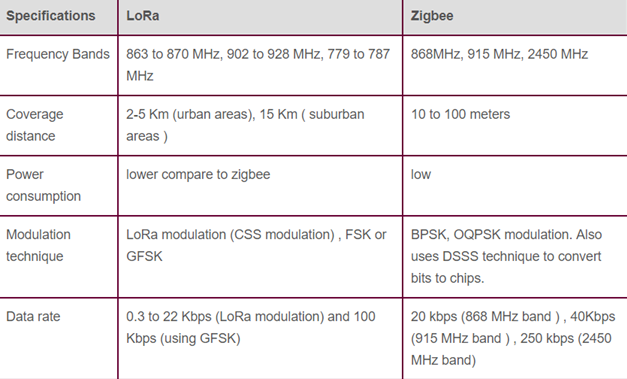
# PHỤ LỤC

**BIÊN BẢN HỌP NHÓM NGÀY 5/3/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Lựa chọn phần cứng.  Lựa chọn Database Online để lưu trữ dữ liệu.  Lựa chọn thuật toán dự đoán thời tiết. | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Lựa chọn phần cứng cho Node Sensor | | | | Nguyễn Văn Trung | | | 15h15 | |
| Lựa chọn phần cứng cho bộ xử lý trung tâm | | | | Hoàng Anh | | | 15h45 | |
| Lựa chọn DataBase Online | | | | Phạm Anh Tú | | | 16h15 | |
| Lựa chọn module truyền nhận dữ liệu giữa NodeSensor và bộ xử lý | | | | Hoàng Anh | | | 16h30 | |
| Lựa chọn thuật toán dự đoán thời tiết. | | | | Phạm Anh Tú | | | 16h45 | |

|  |  |  |
| --- | --- | --- |
| **4.** **Những lưu ý và các quyết định đưa ra trong cuộc họp** | | |
| **Chủ đề** | **Người đưa ra ý kiến** | **Thời gian** |
| Lựa chọn phần cứng trong Node Sensor cho hợp lý để có thể đo nhiệt độ, độ ẩm chính xác: Bao gồm vi điều khiển là STM32F103C8T6, cảm biến đo nhiệt độ, độ ẩm của không khí (DHT22), cảm biến độ ẩm đất  Lý do chọn vi điều khiển STM32F103C8T6: thuộc họ F1 với lõi là ARM COTEX-M3. STM32F103 là vi điều khiển 32-bit, tốc độ tối đa là 72Mhz. Giá thành cũng khá rẻ so với các loại vi điều khiển có chức năng tương tự. Mạch nạp cũng như công cụ lập trình khá đa dạng và dễ sử dụng.  Đưa ra lý do chọn các cảm biến nhiệt độ, độ ẩm của không khí và đất: Các cảm biến này đo đạc với dộ ổn định cao, có khả năng hoạt động liên tục trong thời gian dài, đồng thời dải nhiệt độ, độ ẩm có thể đo được rộng hơn so với những cảm biến khác. | Nguyễn Văn Trung | 15h15 |
| Lựa chọn phần cứng của Bộ xử lý trung tâm: ESP8266. Chọn ESP8266 làm vi điều khiển cho bộ xử lý trung tâm vì nó có khả năng kết nối Internet thông qua WiFi một cách nhanh chóng và sử dụng rất ít linh kiện.  Bộ xử lý gắn thêm vào máy bơm nước, ESP8266 điều khiển để có thể tưới cây dựa vào các thông số thời tiết. | Hoàng Anh | 15h45 |
| Lựa chọn Database Online là ThingSpeak cho phép các cảm biến, các công cụ và các Website để gửi dữ liệu lên đám mây (nơi mà dữ liệu được lưu trên một kênh riêng hoặc kênh công cộng). Khi dữ liệu được lưu trữ trên ThingSpeak, chúng ta có thể phân tích và hình dung nó, tính toán dữ liệu mới, hoặc tương tác với phương tiện truyền thông, dịch vụ Web, và những thiết bị khác. | Phạm Anh Tú | 16h15 |
| Lựa chọn module truyền nhận dữ liệu giữa NodeSensor và bộ xử lý. Có 2 lựa chọn là dùng LoRa hoặc Zigbee.  Sau khi tìm hiểu thì cuối cùng đã đưa ra quyết định chọn Zigbee. Lý do: Việc chọn zigbee sẽ ít tốn năng lượng hơn Lora. Zigbee có đặc điểm là phạm vi hoạt động hẹp, tốc độ truyền Zigbee thích hợp cho các sensor không dây và chuyên dùng cho các ứng dụng giám sát, điều khiển. (Hình PL.1) | Hoàng Anh  Nguyễn Văn Trung | 16h30 |
| Lựa chọn thuật toán dự đoán thời tiết: Dữ liệu liên tục theo thời gian nên nhóm phải sử dụng thuật toán hồi quy, với yêu cầu độ chính xác cao. Thuật toán ARIMA và SVR là 2 thuật toán được sử dụng phổ biến. | Phạm Anh Tú | 16h45 |

|  |  |  |
| --- | --- | --- |
| **5. Thực thi** | | |
| **Hoạt động** | **Người thưc hiện** | **Dự kiến ngày hoàn thành** |
| Mua linh kiện phần cứng. | Nguyễn Văn Trung  Hoàng Anh | 6/3/2019 |



*Hình PL.1: So sánh giữa LoRa và Zigbee.*

**BIÊN BẢN HỌP NHÓM NGÀY 21/3/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Thiết kế thuật toán dự đoán thời tiết. | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Thiết kế mô hình dự đoán thời tiết sử dụng thuật toán SVR. | | | | Hoàng Anh | | | 20h30 | |
| Thiết kế mô hình dự đoán thời tiết sử dụng thuật toán ARIMA. | | | | Nguyễn Văn Trung  Phạm Anh Tú | | | 21h30 | |
| Thiết kế mô hình dự đoán thời tiết sử dụng phương pháp kết hợp ARIMA và SVR. | | | | Phạm Anh Tú | | | 22h30 | |

|  |  |  |
| --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | |
| **Chủ đề** | **Người đưa ra ý kiến** | **Thời gian** |
| Training data là tập dữ liệu phi tuyến tính, và yêu cầu thuật toán có thể xử lý nhanh và chính xác nên sử dụng kernel RBF cho thuật toán SVR.  Đồng thời, kernel RBF được sử dụng phổ biến cho thuật toán SVR. | Hoàng Anh | 20h30 |
| Thuật toán ARIMA có thể dự đoán chính xác theo chu kỳ 1 ngày, 1 tháng hay 1 năm. Nên training data cần phải xử lý chuyển đổi tập dữ liệu đồng dạng với nhau theo chu kỳ. | Nguyễn Văn Trung  Phạm Anh Tú | 21h30 |
| Kết hợp 2 thuật toán với 1 thuật toán dự đoán dữ liệu tương lai, còn dữ liệu còn lại dự đoán dựa vào sai số của thuật toán trước.  Do SVR có thể thực hiện dự đoán dữ liệu phi tuyến tính rất tốt nên ARIMA sẽ dự đoán dữ liệu tương lai còn SVR sẽ dự đoán dữ liệu sai số tương lai. | Phạm Anh Tú | 22h30 |

**BIÊN BẢN HỌP NHÓM NGÀY 22/3/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Thiết kế Database Online | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Đăng ký ThinkSpeak, tạo các field và channel | | | | Hoàng Anh | | | 15h15 | |
| Cách đưa dữ liệu lên ThinkSpeak | | | | Nguyễn Văn Trung | | | 20h30 | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | | | |
| **Chủ đề** | | **Người đưa ra ý kiến** | | **Thời gian** |
| Tạo 2 channel: Dữ liệu thực và Dữ liệu dự đoán.  Mỗi channel có 2 field: Dữ liệu nhiệt độ, dữ liệu độ ẩm. | | Hoàng Anh | | 15h15 |
| Thêm thư viện ThingSpeak cho NodeMCU.  Khai báo địa chi IP vào địa chỉ mac.  Khai báo thông tin channel và mã Write API key.  Cách đọc dữ liệu và gửi dữ liệu. | | Nguyễn Văn Trung | | 20h30 |
| **5. Thực thi** | | | | |
| **Hoạt động** | **Người thưc hiện** | | **Ngày hoàn thành** | |
| Tiến hành đăng ký tài khoản trên ThinkSpeak | Hoàng Anh | | 22/3/2019 | |
| Tiến hành đọc và gửi dữ liệu về ThingSpeak | Nguyễn Văn Trung | | 27/3/2019 | |

**BIÊN BẢN HỌP NHÓM NGÀY 23/3/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Thiết kế phần cứng  Thiết kế ngôi nhà chứa điểm cảm biến | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Thiết kế điểm cảm biến | | | | Nguyễn Văn Trung | | | 15h15 | |
| Thiết kế giao tiếp bộ xử lý trung tâm với điểm cảm biến | | | | Hoàng Anh | | | 15h15 | |
| Thiết kế bộ xử lý trung tâm | | | | Phạm Anh Tú | | | 16h15 | |
| Thiết kế mô hình ngôi nhà cho điểm cảm biến | | | | Hoàng Anh | | | 17h00 | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | | | |
| **Chủ đề** | **Người đưa ra ý kiến** | | | **Thời gian** |
| Vi điều khiển cho node sensor là STM32F103C8T6 đọc nhiệt độ, độ ẩm từ DHT22 theo chuẩn truyền dữ liệu 1 dây.  Vi điều khiển đọc cảm biến độ ẩm đất thông qua điện áp bằng bộ chuyển đổi tương tự sang số ADC 12bit.  Vi điều khiển giao tiếp với module Zigbee bằng chuẩn UART. | Nguyễn Văn Trung | | | 15h15 |
| NodeMCU ESP8266 kết nối với Zigbee bằng chuẩn UART, để nhận và truyền dữ liệu giữa các node. | Hoàng Anh | | | 15h15 |
| NodeMCU dùng chuẩn Wifi để gửi và nhận dữ liệu với các server, database. | Phạm Anh Tú | | | 16h15 |
| Ngôi nhà chứ điểm cảm biến có kích thước 15x15x15 cm.  Có các lỗ thông gió và mica trong suốt cho ánh sáng vào được bên trong để đo nhiệt độ, độ ẩm chính xác. | Hoàng Anh | | | 17h00 |
| **5. Thực thi** | | | | |
| **Hoạt động** | | **Người thưc hiện** | **Ngày hoàn thành** | |
| Tiền hành vẽ PCB cho điểm cảm biến | | Nguyễn Văn Trung | 28/3/2019 | |
| Tiến hành vẽ PCB cho điểm trung tâm | | Hoàng Anh | 28/3/2019 | |
| Vẽ CAD ngôi nhà | | Hoàng Anh | 28/3/2019 | |

**BIÊN BẢN HỌP NHÓM NGÀY 29/3/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Lập trình cho phần cứng | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| H1oảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Lập trình cho NodeSensor, đọc nhiệt độ, độ ẩm từ các cảm biến | | | | Nguyễn Văn Trung | | | 8h | |
| Lập trình cho Bộ xử lý trung tâm, kết nối với ThingSpeak | | | | Hoàng Anh | | | 8h45 | |
| Đưa dữ liệu từ NodeSensor lên bộ xử lý trung tâm theo chuẩn UART thông qua ZigBee | | | | Nguyễn Văn Trung  Hoàng Anh | | | 9h45 | |

|  |  |  |
| --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | |
| **Chủ đề** | **Người đưa ra ý kiến** | **Thời gian** |
| Cần đọc nhiệt độ và dộ ẩm với độ phân giải ±0.01 | Nguyễn Văn Trung | 8h |
| Nên để thời gian cập nhật một lần là 60 phút, khi có lỗi thì yêu cầu gửi lại | Hoàng Anh | 8h45 |
| Nên viết code Debug ở trong Code để dễ dàng kiểm soát khi có lỗi | Phạm Anh Tú | 9h45 |

|  |  |  |
| --- | --- | --- |
| **5. Thực thi** | | |
| **Hoạt động** | **Người thưc hiện** | **Ngày hoàn thành** |
| Lập trình cho NodeSensor, đọc nhiệt độ, độ ẩm từ các cảm biến, và đưa dữ liệu đọc được lên bộ xử lý trung tâm thông qua Zigbee. | Nguyễn Văn Trung | 4/4/2019 |
| Tại bộ xử lý trung tâm nhận dữ liệu về thông qua Zigbee rồi đưa lên ThingSpeak. | Hoàng Anh | 4/4/2019 |

**BIÊN BẢN HỌP NHÓM NGÀY 5/4/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Lập trình thuật toán dự đoán thời tiết. | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Lập trình mô hình dự đoán thời tiết sử dụng thuật toán SVR. | | | | Hoàng Anh | | | 20h30 | |
| Lập trình mô hình dự đoán thời tiết sử dụng thuật toán ARIMA. | | | | Nguyễn Văn Trung | | | 21h30 | |
| Lập trình mô hình dự đoán thời tiết sử dụng phương pháp kết hợp ARIMA và SVR. | | | | Phạm Anh Tú | | | 22h30 | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | | | |
| **Chủ đề** | | **Người đưa ra ý kiến** | | **Thời gian** |
| Sử dụng dữ liệu thời tiết TP.HCM trong 2 tuần có sẵn. Chia dữ liệu thành 2 phần: training data, test data.  Lựa chọn kernel và thông số tối ưu cho thuật toán. | | Hoàng Anh  Phạm Anh Tú | | 20h30 |
| Sử dụng dữ liệu thời tiết TP.HCM trong 2 tuần có sẵn. Chia dữ liệu thành 2 phần: training data, test data.  Tiền xử lý chuyển đổi training data để loại bỏ dữ liệu nhiễu.  Lựa chọn thông số để tối ưu cho thuật toán. | | Nguyễn Văn Trung  Phạm Anh Tú | | 21h30 |
| Sử dụng dụng thuật toán ARIMA để dự đoán các dữ liệu tiếp theo và trả kết quả sai số từng điểm training data.  Dựa vào các dữ liệu kết quả sai số từ thuật toán ARIMA, SVR sẽ dự đoán các sai số từng điểm tiếp theo.  Cộng kết quả dự đoán từ ARIMA và kết quả dự đoán sai số từ SVR với nhau. | | Phạm Anh Tú | | 22h30 |
| **5. Thực thi** | | | | |
| **Hoạt động** | **Người thưc hiện** | | **Dự kiến ngày hoàn thành** | |
| So sánh kết quả 3 phương pháp trên. | Phạm Anh Tú | | 25/4/2019 | |

**BIÊN BẢN HỌP NHÓM NGÀY 10/5/2019**

| **Tên đề tài** | Thiết kế hệ thống dự báo thời tiết bằng thuật toán máy học | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Người hướng dẫn** | TS. Võ Minh Huân | | | | **Địa điểm** | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh | | |
| **1. Mục tiêu cuộc họp** | | | | | | | | |
| Thu thập số liệu, đánh giá hệ thống  Chỉnh sửa hệ thống | | | | | | | | |
| **2. Thành viên** | | | | | | | | |
| **Họ và tên** | | **MSSV** | **E-mail** | | | | | **Điện thoại** |
| Nguyễn Văn Trung | | 15119148 | 15119148@student.hcmute.edu.vn | | | | | 0369884981 |
| Hoảng Anh | | 15146124 | 15146124@student.hcmute.edu.vn | | | | | 0346643034 |
| Phạm Anh Tú | | 15119156 | 15119156@student.hcmute.edu.vn | | | | | 0849918756 |
| **3. Tóm tắt quá trình của cuộc họp** | | | | | | | | |
| **Chủ đề** | | | | **Người đưa ra ý kiến** | | | **Thời gian** | |
| Thu thập số liệu | | | | Nguyễn Văn Trung | | | 19h15 | |
| Chỉnh sửa hệ thống | | | | Hoàng Anh | | | 20h45 | |
| Đánh giá hệ thống | | | | Phạm Anh Tú | | | 21h15 | |

|  |  |  |
| --- | --- | --- |
| **4. Những lưu ý và các quyết định đưa ra trong cuộc họp** | | |
| **Chủ đề** | **Người đưa ra ý kiến** | **Thời gian** |
| Số liệu chưa chính xác, cần phải đo nhiệt độ ngoài trời. | Nguyễn Văn Trung | 19h15 |
| Hệ thống chưa ổn định, khoảng thời gian đo mỗi lần chưa đúng 1 giờ. | Hoàng Anh | 20h45 |
| Điểm trung tâm cần phát hiện dữ liệu sai và yêu cầu điểm cảm biến gửi lại dữ liệu | Phạm Anh Tú | 21h15 |
| **5. Thực thi** | | |
| **Hoạt động** | **Người thưc hiện** | **Dự kiến ngày hoàn thành** |
| Lập trình lại khoảng thời đo mỗi lần cho chính xác 1 giờ 1 lần. | Nguyễn Văn Trung | 5/6/2019 |
| Lập trình kiểm tra lỗi dữ liệu cho điểm trung tâm và di chuyển điểm cảm biến ra ngoài trời. | Hoàng Anh | 5/6/2019 |

1. G. Drakos, " Support Vector Machine vs Logistic Regression", Available: https://towardsdatascience.com/support-vector-machine-vs-logistic-regression-94cc2975433f. [Đã truy cập 23 3 2019]. [↑](#footnote-ref-1)
2. P. Paisitkriangkrai, " Linear Regression and Support Vector Regression", Available: https://cs.adelaide.edu.au/~chhshen/teaching/ML\_SVR.pdf. [Đã truy cập 23 3 2019]. [↑](#footnote-ref-2)
3. P. Paisitkriangkrai, " Linear Regression and Support Vector Regression", Available: https://cs.adelaide.edu.au/~chhshen/teaching/ML\_SVR.pdf. [Đã truy cập 23 3 2019]. [↑](#footnote-ref-3)
4. Dr. Saed Sayad, " Support Vector Machine - Regression (SVR)", Available: https://www.saedsayad.com/support\_vector\_machine\_reg.htm. [Đã truy cập 29 12 2018]. [↑](#footnote-ref-4)
5. U. Murliza, "ESP8266 IoT ThingSpeak Control Relay", Available: https://getosy.blogspot.com/2016/05/esp8266-iot-thingspeak-control-relay.html. [Đã truy cập 1 6 2019]. [↑](#footnote-ref-5)