**B. MÔ TẢ NGHIÊN CỨU**

**B1. Giới thiệu về đề tài**

Đo pH đóng vai trò quan trọng trong nhiều lĩnh vực, từ khoa học và kỹ thuật đến đời sống sinh hoạt hàng ngày. Trong nông nghiệp, đo pH giúp chọn loại cây trồng, kiểm tra dung dịch dinh dưỡng và hệ thống tưới tiêu. Trong sản xuất thực phẩm, nó đảm bảo chất lượng và an toàn, cũng như trong lên men và nước uống. Trong hóa chất và mỹ phẩm, việc đo pH giúp kiểm soát quá trình sản xuất và đảm bảo an toàn cho người dùng. Hơn nữa, trong lĩnh vực y tế, đo pH chẩn đoán và điều trị nhiều bệnh, và trong môi trường, giúp đánh giá mức độ ô nhiễm và quản lý xử lý nước thải. Ngoài ra, đo pH còn được ứng dụng trong nhiều lĩnh vực khác như: công nghiệp dệt may, sản xuất giấy, luyện kim, ... Việc đo pH giúp chúng ta kiểm soát và điều chỉnh các yếu tố quan trọng trong nhiều lĩnh vực, từ đó đảm bảo chất lượng sản phẩm, an toàn cho sức khỏe và bảo vệ môi trường. Hiện nay, việc sử dụng bút đo pH để xác định độ pH trong các dung dịch là rất phổ biến. Tuy nhiên, với phương pháp đo pH bằng quỳ tím, ta chỉ sử dụng được 1 mẫu giấy cho 1 lần đo, và với phương pháp đo bằng bút đo pH, đầu bút thường bị hư hỏng khi liên tục tiếp xúc với nước, dẫn đến chi phí thay thế cao và ảnh hưởng đến độ tin cậy của kết quả đo.

Do đó, đề tài sẽ tập trung nghiên cứu, tìm hiểu và thiết kế bộ thí nghiệm đo pH trong nước bằng quang phổ hồng ngoại gần và áp dụng các mô hình máy học để đưa ra các dự đoán về độ pH với độ chính xác cao nhất có thể và phát triển thêm các hướng phát triển khác.

**B2. Mục tiêu, nội dung, kế hoạch nghiên cứu**

**B2.1. Mục tiêu**

Mục tiêu tổng quan: Đề tài tập trung nghiên cứu và thiết kế bộ thí nghiệm đo độ pH trong nước áp dụng những tác động của quang phổ cận hồng ngoại đến phân tử nước và khả năng đâm xuyên của quang phổ. Thiết kế và xây dựng bộ thí nghiệm mang tính chất thử nghiệm có khả năng đo pH với độ bền rất cao, từ đó tiết kiệm chi phí cho người sử dụng.

Mục tiêu cụ thể:

* Tìm hiểu, nghiên cứu các phương pháp đo pH hiện nay.
* Xác định các bước sóng có phản ứng mạnh nhất với phân tử nước.
* Xây dựng bộ dữ liệu với 200 mẫu, từ đó áp dụng các phương pháp tiền xử lý dữ liệu, mô hình máy học để đưa ra mối tương quan giữa các bước sóng thu được khi đã đi qua nước với độ pH thực tế.
* Nghiên cứu các cảm biến, nguồn thu, phát phù hợp với đề tài.
* Thiết kế phần cứng với giá thành rẻ, có thể tái sử dụng nhiều lần.
* So sánh và đưa ra các điều chỉnh phù hợp với độ sai số so với độ pH \_\_\_ %.

**B2.2. Nội dung và phương pháp nghiên cứu**

**2.2.1. Nội dung thực hiện**

* Nội dung 1: Tìm hiểu tổng quan đề tài.
* Nội dung 2: Thiết kế cấu trúc phần cứng của hệ thống.
* Nội dung 3: Lập trình vi điều khiển, lập trình ứng dụng.
* Nội dung 4: Thu thập bộ dữ liệu và xây dựng mô hình máy học. Chọn ra bước sóng và mô hình có kết quả cao, phù hợp.
* Nội dung 5: Đóng gói, thử nghiệm, đánh giá sản phẩm.

**2.2.2. Xây dựng hệ thống phần cứng**

Phần cứng cơ bản sẽ gồm 4 thành phần chính:

* Vi điều khiển: là bộ điều khiển trung tâm, nơi tiếp nhận xử lý các tín hiệu từ cảm biến quang phổ và áp dụng các mô hình máy học, xử lý dữ liệu để đưa ra dự đoán.
* Cảm biến quang phổ: dùng để thu thập các bước sóng được phát ra từ nguồn phát sau khi đi qua nước; cảm biến quang phổ phải thu thập các bước sóng nằm trong dải NIR và giá thành dễ tiếp cận.
* Đèn halogen: loại đèn phổ biến nhất trên thị trường hiện nay là đèn led với độ tiêu thụ điện năng thấp và tuổi thọ sử dụng dài. Tuy nhiên, đèn led chỉ có thể phát ra ánh sáng với 1 bước sóng cố định, vì thế đèn led sẽ không phù hợp với đề tài. Đèn halogen có khả năng phát ra ánh sáng với dải bước sóng trải dài từ 350nm – 1000nm và tuổi thọ tương đối cao, vì thế sẽ phù hợp với đề tài hơn.

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, bản đồ

Mô tả được tạo tự động

*Hình 1: Bộ thí nghiệm với những phần cứng chính*

**2.2.2.3. Xây dựng hệ thống phần mềm**

Hình 3 mô tả mô hình xây dựng dự kiến của hệ thống, thiết bị sẽ thực hiện việc đo lấy dữ liệu từ cảm biến, sau đó được đưa vào các phương pháp tiền xử lý để tối ưu hóa dữ liệu. Dữ liệu này sẽ được phục vụ cho quá trình huấn luyện các mô hình máy học. Sau khi hoàn tất huấn luyện, dựa vào các phương pháp đánh giá sẽ chọn ra một mô hình máy học phù hợp nhất. Từ đó thực hiện việc đo và dự đoán độ pH, cuối cùng sẽ kết quả dự đoán sẽ hiển thị trên màn hình. Vi điều khiển đóng vai trò quan trọng trong việc chuyển đổi quang phổ dữ liệu thành giá trị pH và điều khiển các cảm biến, đèn Halogen. Khi khởi động, vi điều khiển sẽ truyền tín hiệu để khởi tạo các cảm biến sau đó sẽ khởi động nguồn phát. Khi ánh sáng của đèn Halogen được chiếu xuyên qua nước thì cảm biến sẽ thu được giá trị của các bước sóng. Khi đã có dữ liệu từ cảm biến, vi điều khiển sẽ xử lý dữ liệu và áp dụng các mô hình máy học để huấn luyện dữ liệu.

(Chèn hình)

**2.2.2.4. Phương pháp chọn bước sóng**

Hệ thống sẽ ứng dụng những tính chất của quang phổ hồng ngoại gần cụ thể là ở bước sóng nằm trong khoảng 750nm – 970nm (SW-NIR) bởi sự hấp thụ của gốc -OH trong phân tử nước đến dải bước sóng này, bước sóng 970nm sẽ được quan tâm nhiều để nghiên cứu và thử nghiệm vì nó đã được sử dụng ở các nghiên cứu khác nhau. Đèn halogen chính là sự lựa chọn phù hợp để làm nguồn phát, từ đó ta sẽ phát hiện được chính xác các bước sóng tác động mạnh đến phân tử nước và tối ưu hóa số lượng bước sóng để có thể tăng tốc độ xử lý để đưa ra kết quả một cách nhanh nhất nhưng vẫn đảm bảo về độ chính xác.

**2.2.2.5. Xây dựng bộ dữ liệu**

Với mục tiêu đã đề ra, đề tài sẽ thu thập đa dạng dữ liệu hơn bằng cách tăng độ trải dài pH của các mẫu nước từ đó bộ dữ liệu sẽ trực quan và đáng tin cậy hơn. Bộ thí nghiệm sẽ áp dụng các mô hình máy học để tìm ra mối liên hệ giữa các bước sóng và độ pH thực tế, bút đo pH sẽ được dùng để thu thập dữ liệu về độ pH thực tế từ đó các mô hình máy học tìm ra được sự liên quan giữa bước sóng và pH thực tế.

**2.2.2.6. Phương pháp máy học**

Phân tích hồi quy là một phương pháp thống kê được sử dụng để nghiên cứu mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập. Mục tiêu của phân tích hồi quy là xác định mô hình toán học tốt nhất để dự đoán hoặc giải thích biến phụ thuộc dựa trên các biến độc lập. Công trình sẽ ứng dụng các mô hình hồi quy để tìm ra mối tương quan giữa độ pH và các bước sóng được thu bởi cảm biến. Sau khi chọn ra được mô hình hồi quy phù hợp, mô hình sẽ được nạp lên vi điều khiển và tiến hành dự đoán độ pH của nước.  
**2.2.2.7. Phương pháp đánh giá kết quả**

**B2.3. Kế hoạch nghiên cứu**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **STT** | **Nội dung** | **Thời gian** | | | | | |
| T11  2023 | T12  2023 | T1  2024 | T2  2024 | T3  2024 | T4  2024 |
| 1 | Nội dung 1 | X |  |  |  |  |  |
| 2 | Nội dung 2 |  | X | X |  |  |  |
| 3 | Nội dung 3 |  |  | X | X |  |  |
| 4 | Nội dung 4 |  |  |  | X | X |  |
| 5 | Nội dung 5 |  |  |  |  |  | X |

**B3. Kế hoạch dự kiến**

Các kết quả cần đạt được để làm tiêu chí đánh giá trong thời gian 6 tháng thực hiện đề tài như sau:

1. Phần cứng được thiết kế ổn định, chắc chắn, hạn chế bị nhiễu bởi điều kiện môi trường và sai số giữa các lần đo.
2. Xây dựng được dataset lớn hơn 200 mẫu với độ pH đa dạng của các mẫu nước.
3. Chọn ra 3 - 4 bước sóng ảnh hưởng cao nhất đến phân tử glucose nhằm tối ưu hóa bộ dữ liệu từ đó giảm độ phức tạp về dữ liệu và tăng tốc độ đối với các mô hình máy học và vi điều khiển.
4. Các phương pháp đánh giá mô hình máy học đạt hiệu quả tốt. Chọn ra một mô hình đạt hiệu quả cao nhất (R2 phải đạt trên 90%), phù hợp với vi điều khiển.
5. Hệ thống hoạt động ổn định có độ chính xác cao. Đạt kết quả sai số trung bình giữa thiết bị và độ pH chuẩn không quá 8%.

**B4. Tài liệu tham khảo**