**Giai đoạn 1:**  từ 25-8 đến 15-9

**Bài toán và phân tích**

**Bài toán:** Dự báo có mưa hay không có mưa sau 2-3 giờ tính từ thời điểm dự đoán.

* **Input**: có **8** input đầu vào

+ **Tốc độ gió** , **hướng gió** : Căn cứ vào kiếm thức địa lý sách giáo khoa, hương gió thổi từ biển đông mang theo hơi nước, mây mưa,…

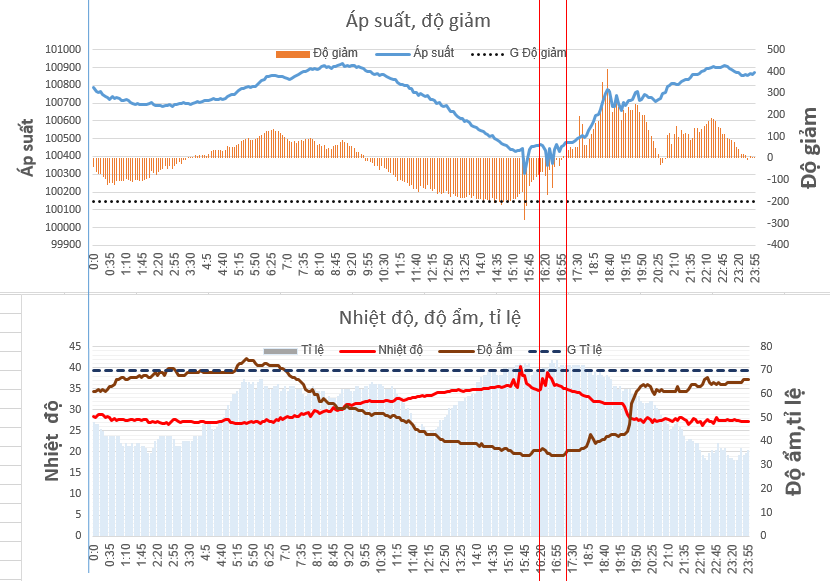
+ **Giờ trong ngày:** môi trường ở thời điểm khác nhau là khác nhau. Tùy thuộc mỗi mùa, tháng thời điểm mưa trong ngày là khác nhau. VD: tháng 10 mưa thương tập trong vào khoảng 6 giờ đến 10 giờ sáng

**+ Nhiệt độ thấp, cao nhất trong 1 giờ** : Khoảng nhiệt độ (thấp nhất, cao nhất), mà tại đây ta có thể xác định mưa hay không có mưa. Do không thể nào xác định chính xác được là tại 1 điểm nhiệt độ chính xác mà có thể dự đoán được nên ta lấy khoảng.

**+ Độ ẩm thấp, cao nhấp trong 1 giờ** : Khoảng độ ẩm (thấp nhất, cao nhất) , mà tại đây ta có thể xác định mưa hay không có mưa. Do không thể nào xác định chính xác được là tại 1 điểm độ ẩm chính xác mà có thể dự đoán được nên ta lấy khoảng.

**+ Tổng độ giảm áp suất trong 2 giờ:**

*Căn cứ vào ý kiến chuyên gia trong tài liệu Giáo trình Khí tượng Nông nghiệp (Đoàn Văn Điếm chủ biên) của Đại học Nông nghiệp 1, tại chương XI phần 1.3b.*



*Ta thấy rằng khi áp suất giảm cao với nhiệt độ thấp và độ ẩm cao thì sẽ là dấu hiệu cho khả năng mưa*

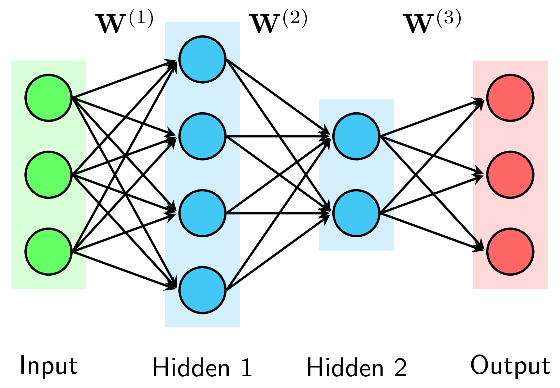
* **Ouput**:

**+ 1 output** : lượng mưa dự đoán được.

Dữ liệu được thu thập 5 phút 1 lần.

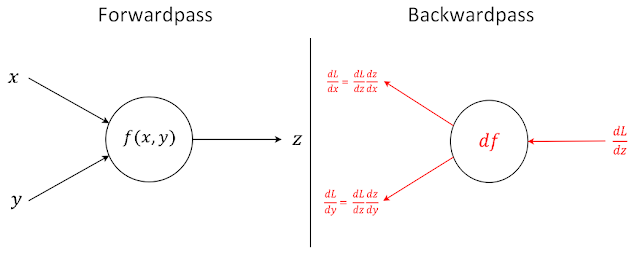
* **Mô hình**:

+ Multi-layer Perceptron ( MLP )



* + - * Hàm truyền ở hidden layer : Sigmoid
      * Hàm truyền ở output layer : Sigmoid

+ Backpropagation



**Kết quả**

* **Kết quả** : kết quả dự đoán được sau khi train trên mô hình neural không cho kết quả khả thi, những dự đoán cho kết quả rất không chính xác.
* **Nguyên nhân và cái tiên**:

+ Output mong muốn không thích hợp với mô đình, phương hướng cãi tiến là output: nhận kết quả là 0 và 1, chia bộ dữ liệu thành 2 class là có mưa (1), không có mưa (0).

+ Input nhiệt độ thấp nhất, nhiệt độ cao nhất, độ ẩm thấp nhất, độ ẩm cao nhất. Không dùng được, do không giúp bộ dữ liệu phân chia thành 2 lớp rõ ràng, không thích hợp với phương án mới. Cải tiến : nhiệt độ trung bình trong 1 giớ, độ ẩm trung bình trong 1 giờ.

**Giai đoạn 2:** từ 25-9 đến 15-10

* **Input:** có 6 input

**+ Tốc độ gió** : không thay đổi

**+ Hướng gió :** chia thành 12 hướng tương ứng ( mỗi hương 30 độ).

**+ Giờ trong ngày:** không thay đổi.

**+ Nhiệt độ trung bình trong 1 giờ:** đặc trung nhiệt độ

**+ Độ ẩm trung bình trong 1 giờ:** đặc trung độ ẩm

**+ Tổng độ giảm áp suất trong 2 giờ:** không thay đổi

* **Output:** có 1 output

**+** nhận giá trị 0 và 1

* **Mô hình:**

**+** Multi-layer Perceptron ( MLP )

* + - * Hàm truyền ở hidden layer : Sigmoid
      * Output layer : Softmax Regression

**+** Backpropagation

**Kết quả**

Kết quả dự đoán 1 số ngày đại diện

Ngày 28-09:

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 2 | Không mưa | Không mưa | Ok |
| 3 | Không mưa | Không mưa | Ok |
| 4 | Không mưa | Không mưa | Ok |
| 5 | Có mưa | Không mưa | No |
| 6 | Có mưa | Không mưa | No |
| 7 | Có mưa | Không mưa | No |
| 8 | Có mưa | Không mưa | No |

Độ chính xác: 42%

Ngày 01 -10:

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 12 | Không mưa | Có mưa | No |
| 13 | Không mưa | Có mưa | No |
| 14 | Không mưa | Có mưa | No |
| 15 | Không mưa | Không mưa | Ok |
| 16 | Không mưa | Không mưa | Ok |
| 17 | Không mưa | Không mưa | Ok |
| 18 | Không mưa | Không mưa | OK |
| 19 | Không mưa | Không mưa | Ok |
| 20 | Có mưa | Không mưa | No |
| 21 | Có mưa | Không mưa | No |
| 22 | Có mưa | Không mưa | No |
| 23 | Có mưa | Có mưa | Ok |

Độ chính xác: 50%

Ngày 03-10

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 10 | Có mưa | Có mưa | Ok |
| 11 | Có mưa | Có mưa | Ok |
| 12 | Có mưa | Có mưa | Ok |
| 13 | Có mưa | Không mưa | No |
| 14 | Có mưa | Không mưa | No |
| 15 | Có mưa | Không mưa | No |
| 16 | Không mưa | Không mưa | Ok |
| 17 | Không mưa | Không mưa | Ok |
| 18 | Không mưa | Không mưa | Ok |
| 19 | Không mưa | Có mưa | No |
| 20 | Không mưa | Có mưa | No |
| 21 | Không mưa | Có mưa | No |

Độ chính xác: 50%

Kết quả thực tế đo được có độ sai lệnh so với kết quả ngoài thực tế, cách khác phục là tăng gia trị khi thu thập được từ cảm biến mưa xác định mưa lên đê loại bỏ trường họp chỉ mưa nhẹ vài hạt nhưng vẫn nhận đó là đã mưa.

**Nhận xét**: Kết quả dự đoán cho kết quả không cao, chưa thỏa yêu cầu, cần cải thiện.

**Cải thiện**:

* Nhiệt độ trung bình trong 1 giờ -> Độ biến thiên nhiệt độ trong 1 giờ .
* Độ ẩm trung bình trong 1 giờ -> Độ biến thiên độ ẩm trong 1 giờ.

Thể hiện rõ hơn đặc trưng, sự thay đổi của nhiệt độ, độ ẩm. Giúp phân rõ hơn 2 class trong dữ liệu.

**Giai đoán 3**: tự 20-10 đến nay

* **Input:** có 6 input

**+ Tốc độ gió** : không thay đổi

**+ Hướng gió :** chia thành 12 hướng tương ứng ( mỗi hương 30 độ).

**+ Giờ trong ngày:** không thay đổi.

**+ Tổng độ giảm nhiệt độ trong 1 giờ:**

**+ Tôngt độ giảm độ ẩm trong 1 giờ:**

**+ Tổng độ giảm áp suất trong 2 giờ:** không thay đổi

* **Output:** có 1 output

**+** Nhận giá trị 0 và 1

Input được chuẩn hóa về thuộc khoảng [-1,1].

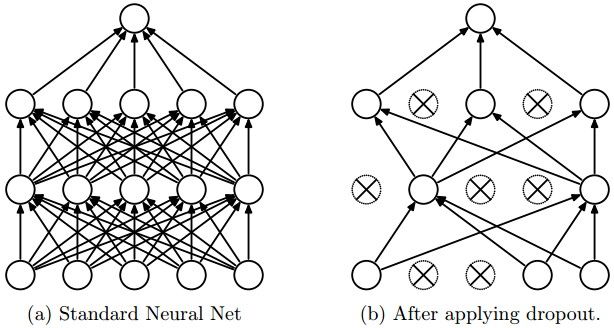
* **Mô hình:**

**+** Multi-layer Perceptron ( MLP )

* + - * Hàm truyền ở hidden layer : Sigmoid
      * Output layer : Softmax Regression

**+** Backpropagation

+ Dropout : Dropout là cách thức mà chúng ta giả định một phần các unit bị ẩn đi trong quá trình training, qua đó làm giảm tích hòa trộn (hay nói cách khác là 1 hidden unit không thể dựa vào 1 unit khác để sửa lỗi lầm của nó, dễ cho chúng ta thấy các hidden unit không đáng tin cậy). Tại mỗi step trong quá trình training, khi thực hiện Forward Propagation (Lan truyền xuôi) đến layer sử dụng Drop-Out, thay vì tính toán tất cả unit có trên layer, tại mỗi unit ta “gieo xúc xắc” xem unit đó có được tính hay không dựa trên xác suất p



**Kết quả**

Kết quả dự đoán 1 số ngày đại diện:

Ngày 03-10:

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 2 | Không mưa | Không mưa | Ok |
| 5 | Không mưa | Không mưa | Ok |
| 10 | Không mưa | Không mưa | Ok |
| 12 | Có mưa | Không mưa | No |
| 15 | Có mưa | Có mưa | Ok |
| 18 | Không mưa | Không mưa | Ok |
| 23 | Không mưa | Không mưa | Ok |

Độ chính xác: 87%

Ngày 07-10:

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 2 | Không mưa | Không mưa | Ok |
| 5 | Không mưa | Không mưa | Ok |
| 6 | Có mưa | Không mưa | No |
| 8 | Không mưa | Không mưa | Ok |
| 13 | Không mưa | Không mưa | Ok |
| 17 | Không mưa | Không mưa | Ok |
| 19 | Không mưa | Không mưa | Ok |
| 22 | Không mưa | Không mưa | Ok |

Độ chính xác : 90%

Ngày 18-10:

|  |  |  |  |
| --- | --- | --- | --- |
| Giờ trong ngày | Kết quả thực tế đo được | Kết quả dự đoán | Chính xác |
| 2 | Không mưa | Không mưa | Ok |
| 7 | Không mưa | Không mưa | Ok |
| 8 | Có mưa | Có mưa | Ok |
| 9 | Có mưa | Không mưa | No |
| 10 | Có mưa | Không mưa | No |
| 12 | Có mưa | Có mưa | Ok |
| 15 | Không mưa | Không mưa | Ok |
| 19 | Không mưa | Không mưa | Ok |
| 20 | Không mưa | Không mưa | Ok |

Độ chính xác : 78%