- **TECHNICAL REPORT -**

# ĐÁNH GIÁ CÁC MÔ HÌNH HỌC MÁY TRÊN TẬP DỮ LIỆU BOSTON HOUSING BẰNG CÁC CHỈ SỐ ACCURACY, PRECISON, RECALL, F1 SCORE

**1. Giới thiệu**

➢ Tập dữ liệu dự đoán thời tiết là một tập dữ liệu thực tế, được sử dụng để dự đoán các hiện tượng thời tiết như mưa, nắng, hoặc gió dựa trên các yếu tố như lượng mưa, nhiệt độ cao nhất, nhiệt độ thấp nhất, và tốc độ gió.  
➢ Xem tập dữ liệu tại đây: Seattle Weather Dataset  
➢ Báo cáo này đánh giá ba mô hình học máy: **Cây quyết định** (Decision Tree - ID3 Algorithm), **Mạng nơ-ron nhân tạo** (Neural Network - MLP Classifier), và **Perceptron**. Hiệu suất của mỗi mô hình được đánh giá dựa trên các thước đo chính gồm: **Độ chính xác** (Accuracy), Độ chính xác theo nhãn (Precision), **Độ hồi đáp** (Recall), và **Chỉ số F1** (F1 Score). Mục đích là xác định mô hình nào dự đoán thời tiết chính xác nhất.

**2. Các chỉ số đánh giá**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| **Tên chỉ số** | Độ chính xác | Độ chính xác theo nhãn | Độ hồi đáp | Chỉ số F1 |
| **Mục đích** | Đánh giá tỷ lệ dự đoán đúng trên tổng số dự đoán | Đánh giá khả năng mô hình chỉ ra nhãn đúng trong số các dự đoán đúng | Đánh giá khả năng mô hình nhận diện đúng các nhãn thực | Đánh giá sự cân bằng giữa Precision và Recall |
| **Giá trị** | [0,1] | [0,1] | [0,1] | [0,1] |
| **Ý nghĩa** | - Giá trị Accuracy = 1 nghĩa là mô hình dự đoán hoàn toàn chính xác (tất cả dự đoán đều đúng).  - Nếu Accuracy = 0, mô hình hoàn toàn sai. Giá trị thấp có thể chỉ ra sự mất cân bằng dữ liệu (bias) trong tập dự đoán. | - Precision = 1 cho thấy mô hình không có lỗi dương tính giả (FP = 0), tức là tất cả các mẫu được dự đoán là dương đều thực sự là dương.  - Precision thấp chỉ ra rằng mô hình hay gán nhầm các mẫu âm thành dương (nhiều FP). | - Recall = 1 có nghĩa là mô hình không bỏ sót mẫu dương nào (FN = 0), tức là tất cả các mẫu dương đều được nhận diện đúng.  - Recall thấp nghĩa là mô hình bỏ sót nhiều nhãn dương, không nhận diện đúng các mẫu dương thực sự. | - F1 Score = 1 nghĩa là mô hình hoàn hảo trong việc cân bằng giữa Precision và Recall.  - F1 Score càng gần 0 nghĩa là mô hình kém, có sự mất cân bằng lớn giữa Precision và Recall (một trong hai có thể rất thấp). |
| **Lưu ý** | * Accuracy cao không phải lúc nào cũng tốt nếu dữ liệu bị lệch * Precision cao nhưng Recall thấp nghĩa là mô hình rất chính xác khi dự đoán nhãn dương nhưng bỏ sót nhiều nhãn dương thực tế. * Recall cao nhưng Precision thấp có nghĩa là mô hình phát hiện được nhiều nhãn dương, nhưng dự đoán nhiều nhãn dương sai (nhiều FP) | | | |
| **Công thức tính** |  |  |  |  |

**Chú thích:**

* **TP (True Positive):** Dự đoán đúng nhãn dương.
* **TN (True Negative):** Dự đoán đúng nhãn âm.
* **FP (False Positive):** Dự đoán nhầm nhãn dương thành âm.
* **FN (False Negative):** Dự đoán nhầm nhãn âm thành dương.

**3. So sánh hiệu suất các mô hình**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Mô hình** | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| **Perceptron** | 0,74 | 0,74 | 0,74 | 0,74 |
| **Decision tree** | 0,81 | 0,81 | 0,81 | 0,81 |
| **Neural Network** | 0,82 | 0,82 | 0,82 | 0,82 |
| **Stacking** |  |  |  |  |

Từ kết quả đó, nhóm đã đánh giá được kết quả như phần trình bày dưới đây:

**3.2. Perceptron**

* **Accuracy** = 0.74: Mô hình Perceptron có độ chính xác đạt 74%, tức dự đoán chính xác 74% trường hợp trong tập dữ liệu.
* **Precision** = 0.74: Trong tất cả các dự đoán của Perceptron, 74% dự đoán là chính xác.
* **Recall** = 0.74: Mô hình tìm ra đúng 74% các trường hợp thực tế của lớp được dự đoán.
* **F1 Score** = 0.74: Đạt sự cân bằng giữa Precision và Recall, chỉ số F1 là 74%, thể hiện mô hình hoạt động ổn định.

**3.3. Decision tree**

* **Accuracy** = 0.81: Mô hình Decision Tree dự đoán chính xác 81% trường hợp trong tập dữ liệu.
* **Precision** = 0.81: 81% trong số các dự đoán của Decision Tree là đúng.
* **Recall** = 0.81: Mô hình này tìm ra đúng 81% các trường hợp thuộc lớp cần dự đoán.
* **F1 Score** = 0.81: Cho thấy sự cân bằng tốt giữa Precision và Recall với giá trị 81%, hiệu suất tổng thể cao và ổn định.

**3.4.** **Neural Network**

* **Accuracy** = 0.82: Neural Network dự đoán đúng 82% trường hợp trong tập dữ liệu.
* **Precision** = 0.82: 82% dự đoán của mô hình là chính xác.
* **Recall** = 0.82: Neural Network tìm ra đúng 82% các trường hợp thực tế của lớp dự đoán.
* **F1 Score** = 0.82: Với chỉ số F1 là 82%, mô hình này cho thấy sự cân bằng giữa Precision và Recall, cùng hiệu suất vượt trội.

**4. Nhận xét**

* **Neural Network** là mô hình có hiệu suất tốt nhất với các chỉ số Accuracy, Precision, Recall và F1 Score đều đạt 0.82, nhỉnh hơn so với các mô hình còn lại.
* **Decision Tree** cũng cho thấy hiệu suất rất tốt với các chỉ số đều đạt 0.81, chỉ kém nhẹ so với Neural Network**.**
* **Perceptron** tuy có hiệu suất ổn định nhưng thua kém hơn so với hai mô hình trên, với các chỉ số đều là 0.74.
* **Stacking** chưa có kết quả để đánh giá.

**5. Kết luận**

* Dựa trên phân tích hiệu suất, Neural Network là mô hình dự đoán tốt nhất cho tập dữ liệu này, nhờ vào các chỉ số vượt trội về độ chính xác và khả năng phân loại.
* Decision Tree cũng là một lựa chọn tốt với hiệu suất tương đối cao và ổn định.
* Perceptron tuy có hiệu suất thấp hơn nhưng vẫn đạt được mức độ ổn định và cân bằng giữa Precision và Recall.