ĐẠI HỌC QUỐC GIA TP.HCM

**TRƯỜNG ĐH KHOA HỌC TỰ NHIÊN**

A picture containing logo

Description automatically generated

Nhóm 8 N2

**REFLECTIVE REPORT ĐỒ ÁN NHẬP MÔN CÔNG NGHỆ THÔNG TIN**

Ngành: Khoa học dữ liệu

Khóa học năm: 2022

Người hướng dẫn khoa học: Lê Nhựt Nam

Tp.HCM, ngày 13, tháng 1, năm 2023

Thông tin thành viên.

|  |  |  |  |
| --- | --- | --- | --- |
| **MSSV** | **Họ tên** | **Email** | **Vai trò** |
| **22280041** | Trảo An Huy (Trưởng nhóm) | [Traohuy098@gmail.com](mailto:Traohuy098@gmail.com) | Scrum Master,  DEV phụ |
| **22280040** | Trần Gia Huy |  | Developer, Designer |
| **22280042** | Đinh Xuân Khang |  | Developer |
| **22280047** | Nguyễn Lê Đăng Khoa | [khoadz2609@gmail.com](mailto:khoadz2609@gmail.com) | BA vs TESTER |

**MỤC LỤC**

[**I)** **Phương pháp thực hiện đồ án.** 3](#_Toc124445466)

[**1)** **Bài toán đang giải quyết là gì?** 3](#_Toc124445467)

[**2)** **Thông tin về dữ liệu?** 3](#_Toc124445468)

[**3)** **Tạo cơ sở dữ liệu như thế nào.** 5](#_Toc124445469)

[**4)** **Sử dụng Pymongo.** 7](#_Toc124445470)

[**5)** **Các thử nghiệm và kết quả đạt được.** 16](#_Toc124445471)

[**6)** **So sánh ngoài lề mục thêm** 23](#_Toc124445472)

[**II)** **Mức độ đóng góp của từng thành viên?** 25](#_Toc124445473)

[**1)** **Đóng góp của từng thành viên:** 25](#_Toc124445474)

[**2)** **Nhận xét của từng thành viên đối với các thành viên còn lại.** 26](#_Toc124445475)

[**III)** **Những gì học được/ làm được/ đạt được của mỗi thành viên trong nhóm .** 27](#_Toc124445476)

[**1)** **Thành tựu chung của tất cả thành viên.** 27](#_Toc124445480)

[**2)** **Thành tựu cá nhân.** 27](#_Toc124445481)

[**IV)** **Trình bày những gì chưa học được/ chưa làm được/ chưa đạt được của mỗi thành viên trong nhóm.** 29](#_Toc124445482)

[**V)** **Dự định tương lai:** 30](#_Toc124445483)

[**VI)** **Áp dụng vào thực tế các kiến thức được học thông qua đồ án.** 31](#_Toc124445484)

1. **Phương pháp thực hiện đồ án.**
   1. **Bài toán đang giải quyết là gì?**

Nghiên cứu làm việc với cơ sở dữ liệu phi quan hệ bằng cách sử dụng PyMongo và các mô hình học máy để huấn luyện dựa trên iris\_train.csv và kiểm tra tính đúng đắn thông qua iris\_test\_label.csv sau đó sử dụng mô hình học máy để đưa ra kết quả dự đoán của file iris\_test\_input.csv và lưu vào cơ sở dữ liệu với collection name là iris\_predicted.

* 1. **Thông tin về dữ liệu?**

File iris\_train.csv gồm 6 cột vào 120 dòng.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Index | Sepal.length | Sepal.width | Petal.length | Petal.width | Variety |
| Cột index chỉ thứ tự của các quan sát. | Sepal.length chỉ chiều dài của đài hoa Iris. | Cột sepal.width chỉ chiều rộng của đài hoa Iris. | Cột petal.length chỉ chiều dài của cánh hoa Iris. | Cột petal.width chỉ chiều rộng của cánh hoa Iris. | Cột variety chỉ loại hoa Iris (0 - Iris Setosa, 1 - Iris Versicolour, 2- Iris Virginica). |

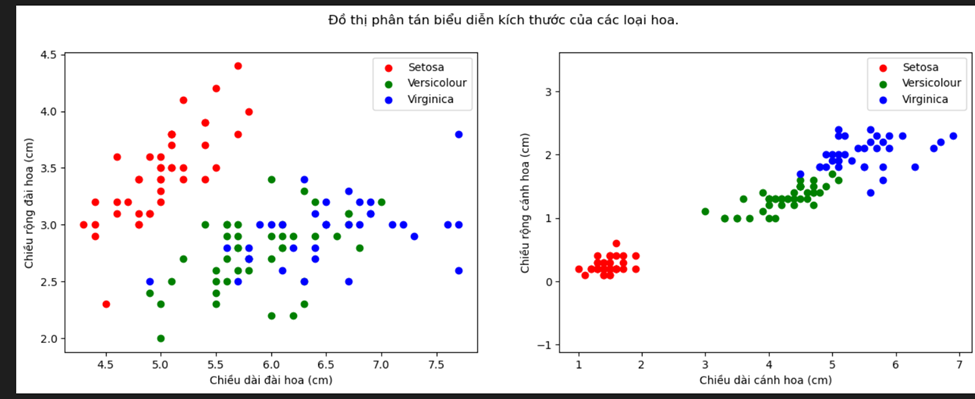
File iris\_test\_input.csv gồm 4 cột và 30 dòng.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Index | Sepal.length | Sepal.width | Petal.length | Petal.width |
| Cột index chỉ thứ tự của các quan sát. | Sepal.length chỉ chiều dài của đài hoa Iris. | Cột sepal.width chỉ chiều rộng của đài hoa Iris. | Cột petal.length chỉ chiều dài của cánh hoa Iris. | Cột petal.width chỉ chiều rộng của cánh hoa Iris. |

File iris\_test\_label.csv gồm 2 cột và 30 dòng.

|  |  |
| --- | --- |
| Index | Variety |
| Cột index chỉ thứ tự của các quan sát. | Cột variety chỉ loại hoa Iris (0 - Iris Setosa, 1 - Iris Versicolour, 2- Iris Virginica). |

Dữ liệu được cho file iris\_train dùng để huấn luyện gồm 4 thông tin có thể được thể hiện trong không gian 4 chiều để sử dụng các mô hình phân lớp.



**Nhận xét về file dữ liệu đã cho thì .**

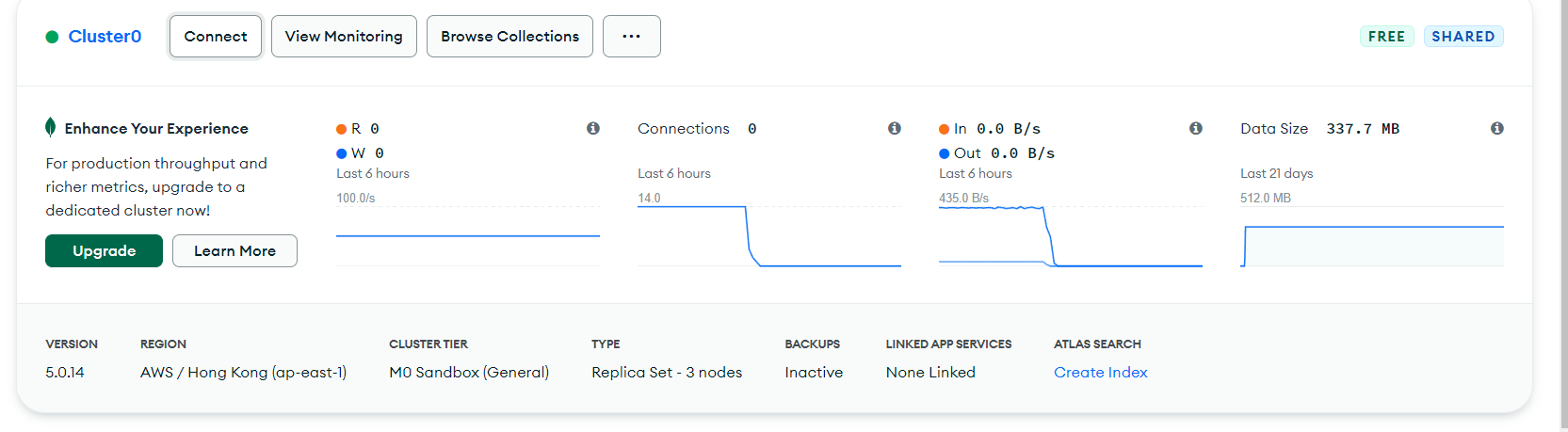
- Sự tập trung của hoa Setosa thì tập trung chủ yếu ở một phía ít bị lẫn lộn giữa các loại hoa còn lại nên khả năng dự đoán từ các thông tin từ file iris\_input\_test.csv thuộc nhóm hoa Setora thì có xác xuất đúng cực kì cao.

- Còn 2 loài hoa còn lại thì trong chiều dài của đài hoa thì các dữ liệu đã bị trộn lẫn vào nhau khó phân loại theo đồ thị bình thường nhưng còn trong, đồ thị chiều dài cánh hoa sự khác biệt này thu hẹp lại, khả năng dự đoán thông tin từ file iris\_input\_test.csv thuộc 2 nhóm hoa còn lại thì độ chính xác sẽ thấp hơn loài Setosa và chủ yếu phân loại phụ thuộc vào chiều dài cánh hoa, chiều rộng cánh hoa đã cho trước đó.

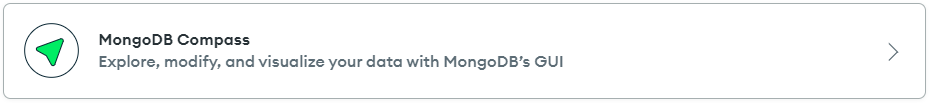
* 1. **Tạo cơ sở dữ liệu như thế nào.**

Sau khi cài đặt môi trường chúng em sử dụng Mongo atlas sau đó kết nối tới mongo DB.

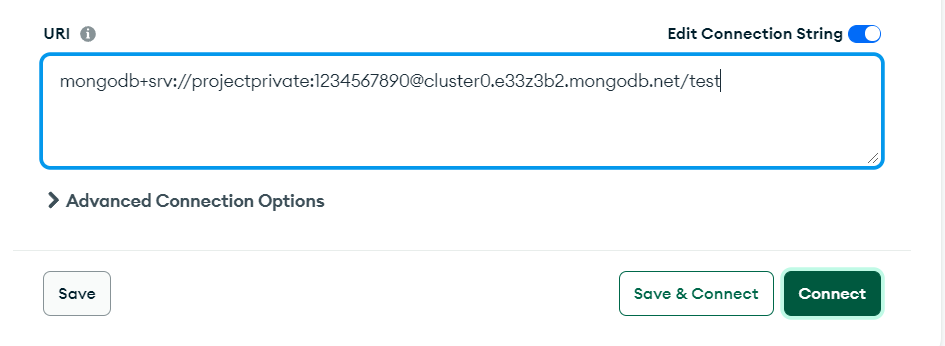
B1.Ấn connect



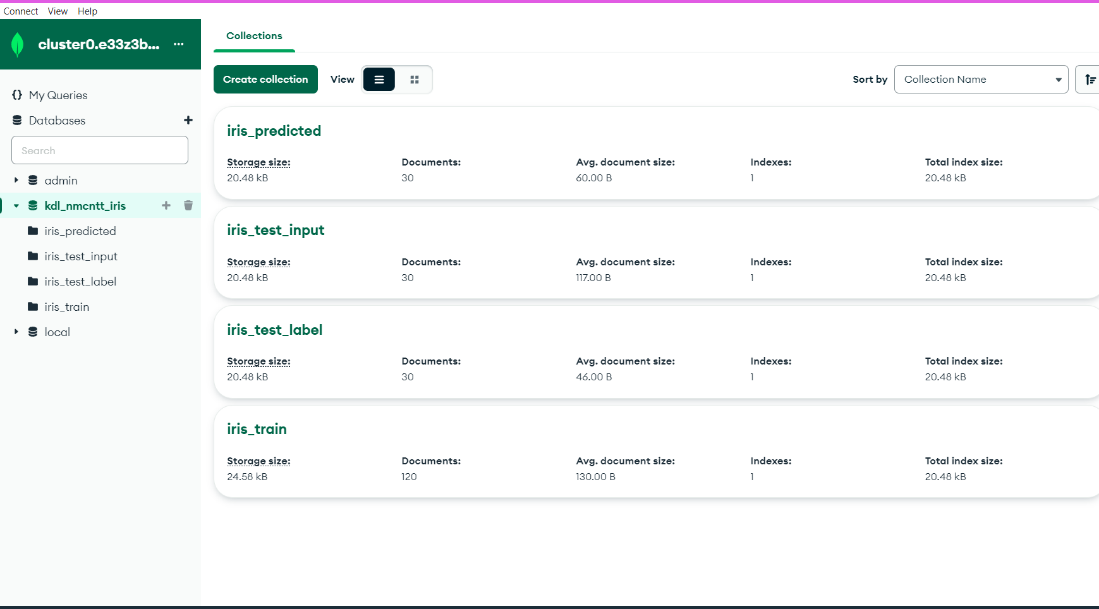
B2. Chọn Mongo DB compass.



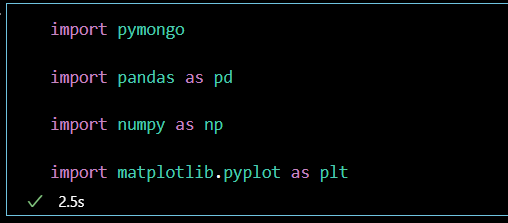
B3. Copy đoạn lệnh truy cập password và usename để thực hiện việc kết nối tới mongo compass.



Và ở đây chúng em đã tải file đồ án thực hành lên database



* 1. **Sử dụng Pymongo.**



**Sử dụng các thư viện Pymongo, Pandas,Numpy,Matplotlib trong Jupyter Notebook.**

Trong đó:

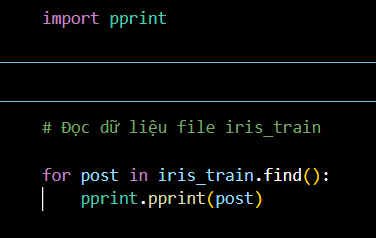
Pymongo là một thư viện Python cho phép bạn kết nối và sử dụng cơ sở dữ liệu MongoDB trong Python.

Pandas là một thư viện Python cho phép xử lý và phân tích dữ liệu. Nó cung cấp một loạt các công cụ và cú pháp dễ sử dụng để tải dữ liệu từ các nguồn khác nhau, truy vấn và xử lý dữ liệu, và thống kê và trực quan hóa dữ liệu.

**Truy xuất dữ liệu từ Database của MongoDB.**

NumPy là một thư viện Python cho phép thao tác với dữ liệu dạng mảng (array). Nó cung cấp một tập hợp các hàm và công cụ để xử lý các mảng và tính toán số học trên chúng. NumPy là một công cụ quan trọng trong các lĩnh vực như xử lý ảnh, xử lý dữ liệu và khoa học dữ liệu, và cũng được sử dụng rộng rãi trong các ứng dụng khác

Matplotlib là một thư viện Python cho phép bạn vẽ các đồ thị và biểu đồ.



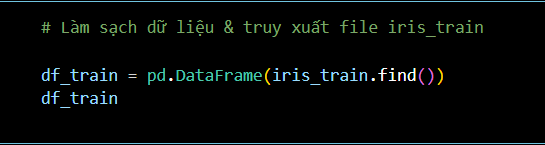
**Đọc dữ liệu từ file iris\_train.**

**Đọc dữ liệu từ file csv.**

**import pprint**

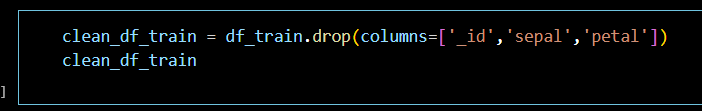
**for post in file.find():**

**pprint.pprint(post)**



**Làm sạch dữ liệu và truy xuất file csv bằng DataFrame của thư viên Pandas.**

Pandas là một thư viện Python cung cấp một số cấu trúc dữ liệu và các công cụ để xử lý dữ liệu. Trong đó, ***pd.DataFrame*** là một trong những cấu trúc dữ liệu chính của thư viện pandas. Nó cho phép bạn tạo ra một bảng 2 chiều, tương tự như một bảng trong Excel hoặc bảng SQL. Các cột của bảng được đặt tên và các hàng được đánh số.



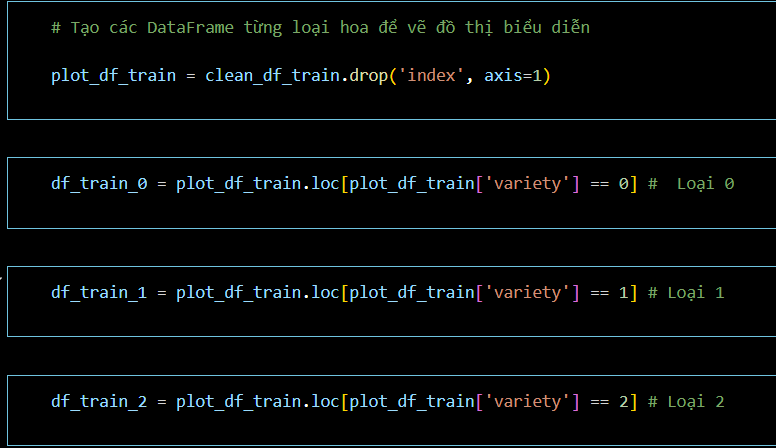
**Xóa các cột 'id', 'sepal', 'petal' bằng hàm drop trong pandas.**

Trong pandas, hàm drop() dùng để xóa các dòng hoặc cột trong một DataFrame. Cú pháp cơ bản của hàm như sau:

**DataFrame.drop(labels=None, axis=0, index=None, columns=None, level=None, inplace=False, errors='raise')**

**Trong đó:**

* labels: nhãn (hoặc một danh sách các nhãn) của dòng hoặc cột mà bạn muốn xóa.
* axis: trục mà bạn muốn xóa. Giá trị 0 tương đương với dòng và 1 tương đương với cột.
* index: chỉ số của dòng mà bạn muốn xóa.
* columns: tên của cột mà bạn muốn xóa.
* level: mức cấp độ của MultiIndex.
* inplace: Nếu được đặt thành True, sẽ thay đổi DataFrame gốc. Nếu False (mặc định) sẽ trả về một bản sao của DataFrame đã xóa.
* errors: cách thức xử lý lỗi khi xóa nhãn không tồn tại. Có thể là 'raise', 'ignore' hoặc 'warn'.



**Tạo các DataFrame từng loại hoa để vẽ đồ thị biểu diễn**.

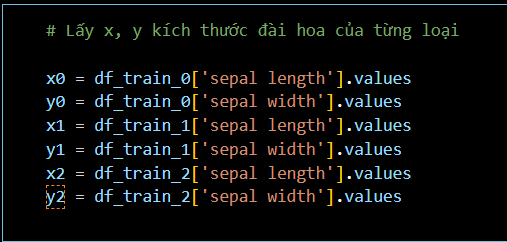
Trong pandas, hàm loc dùng để lấy ra một hoặc nhiều hàng (row) hoặc cột (column) trong DataFrame theo một label hoặc nhiều label.

Cú pháp cơ bản của hàm loc là **DataFrame.loc[row\_indexer, column\_indexer]**, trong đó:

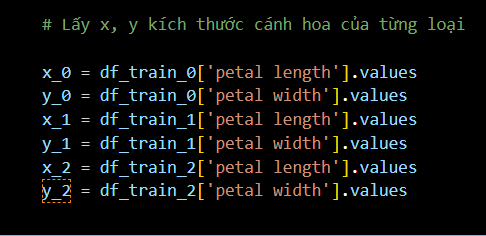
row\_indexer là một label hoặc một list các label của hàng cần lấy

column\_indexer là một label hoặc một list các label của cột cần lấy.

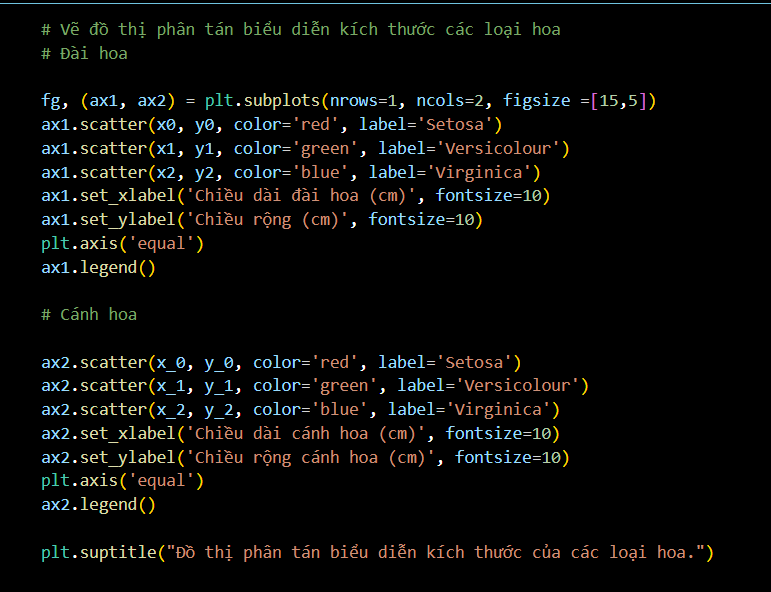
Trong phạm vi đồ án, ta sẽ tạo ra các DataFrame của từng loại hoa bằng cách dựa vào cột variety để phân loại. Các loại hoa có variety là '0' sẽ là loại 0, variety là '1' sẽ là loại 1, variety là '2' sẽ là loại 2.



**Lấy x,y kích thước đài hoa.**

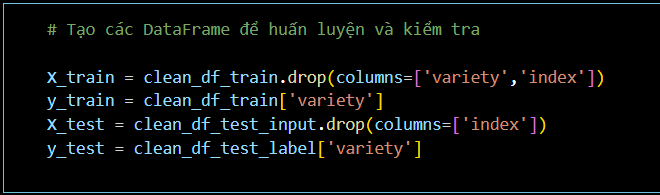


**Lấy x,y kích thước cánh hoa.**



**Vẽ đồ thị phân tán biểu diễn kích thước của các loại hoa.**

Hàm **plt.subplots()** trong thư viện matplotlib là một hàm để tạo ra một figure và các subplots trong đó. Hàm này trả về hai giá trị: một figure object và một axes object (hoặc một mảng các axes objects nếu có nhiều hơn một subplot).Chỉ định số lượng subplots trên một hàng và số lượng hàng bằng cách sử dụng các tham số nrows và ncols.Figure có một thuộc tính rất quan trọng đó là figsize. Thuộc tính này quy định kích thước cửa sổ figure.



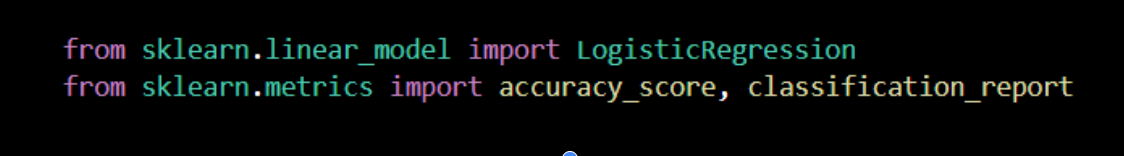
**Tạo các DataFrame để huấn luyện và kiểm tra.**

Trong đồ án, **X\_train** là một DataFrame được tạo từ DataFrame **Clean\_df\_train** bằng cách bỏ đi 2 cột "**variety**" và "**index**".

**Y\_train** là một DataFrame được tạo từ DataFrame **Clean\_df\_train** với một cột "**variety**".

**X\_test** là một DataFrame được tạo từ DataFrame **Clean\_df\_test\_input** bằng cách bỏ đi một cột "**index**".

**y\_test** là một DataFrame được tạo từ DataFrame **Clean\_df\_test\_label** với một cột "**variety**".



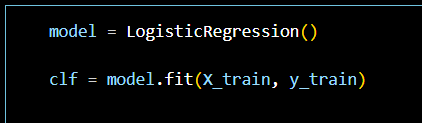
**Sử dụng mô hình LogisticRegression.**

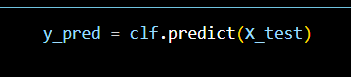
"**from sklearn.linear\_model import LogisticRegression**" dùng để nhập thư viện Logistic Regression từ thư viện scikit-learn (sklearn).

"**from sklearn.metrics import accuracy\_score, classification\_report**" nhập hai hàm từ thư viện scikit-learn (sklearn) về để sử dụng.Trong đó:

**accuracy\_score** là hàm dùng để tính toán độ chính xác của một mô hình dự đoán. Nó cần nhận hai tham số: dự đoán (predictions) và thật sự (ground truth).

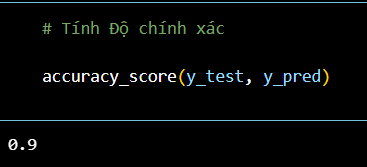
**classification\_report** là hàm dùng để tạo báo cáo phân loại cho một tập dự đoán. Nó cần nhận ba tham số: dự đoán, thật sự và tên của các lớp. Nó sẽ trả về các chỉ số như precision, recall, f1-score, support.

**Tạo mô hình học máy từ 2 DataFrame là X\_train và y\_train.**

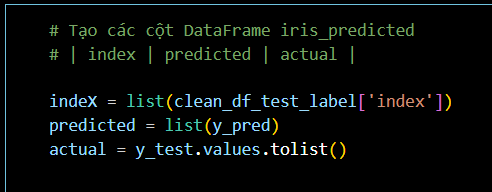


**Hàm predict đưa ra dự đoán từ DataFrame X\_test.**

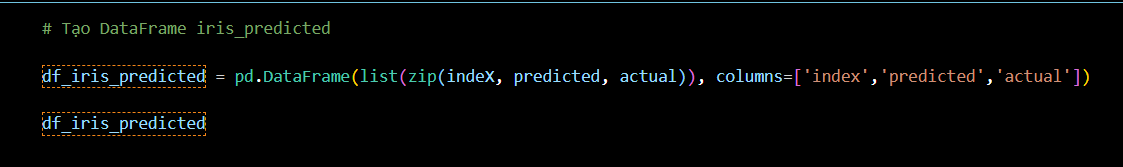
Đưa ra dự đoán cho X\_test dự vào mô hình clf được tạo ra từ trước đó, với X\_test là từ file iris\_input\_test.csv

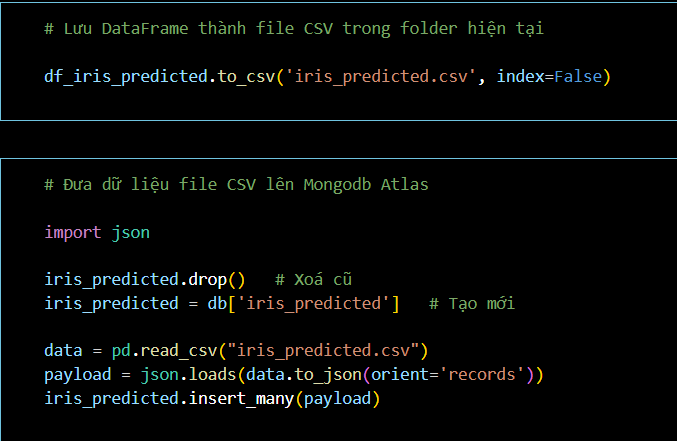


Độ chính xác của mô hình này sau khi đưa file iris\_test\_lable vào kiểm tra là 90%. Trong đó y\_pred là DataFrame được dự đoán bởi mô hình LogisticRegression dựa vào DataFrame y\_test cho trước để kiểm tra.



**Tạo các cột DataFrame iris\_predicted.**

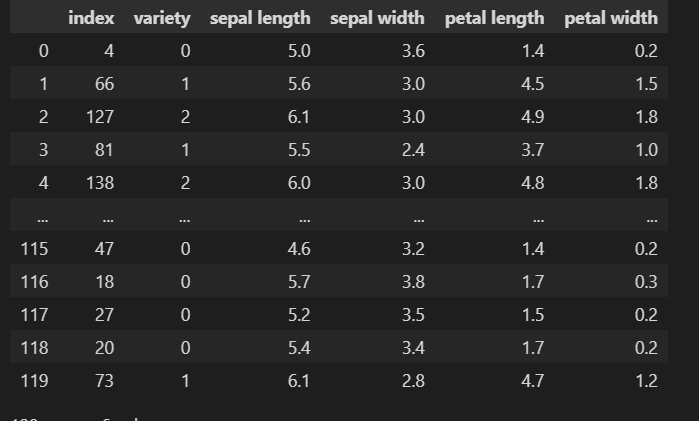


 **Lưu dữ liệu vào file iris\_predicted.csv , tạo Database mới và đưa file csv vào Database này.**

* 1. **Các thử nghiệm và kết quả đạt được.**

**Dữ liệu sau khi được lọc.**

**-iris\_train.csv**

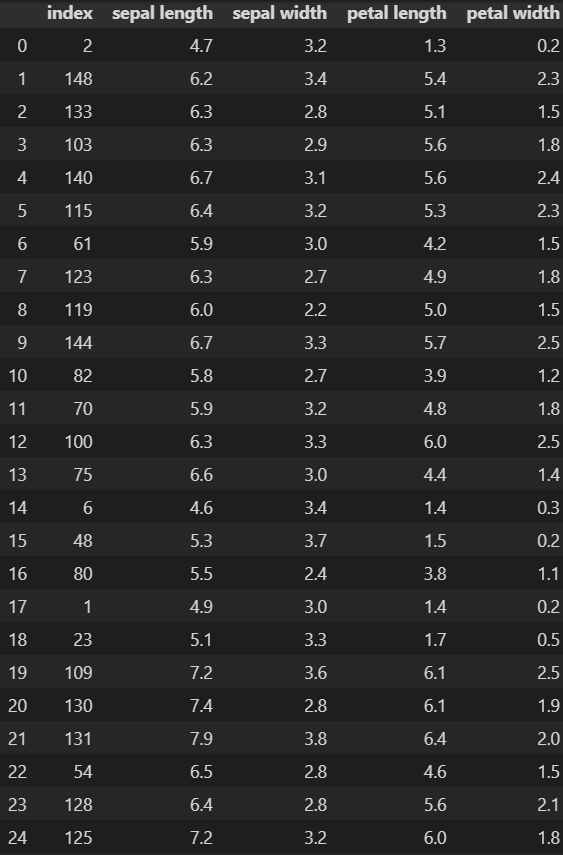


**Mô tả dữ liệu iris train.csv.**

**-Dữ liệu iris\_input\_test.csv sau khi được lọc.**

**-iris\_input\_test.csv**

**Sau khi được lọc.**





**-File iris\_test\_label.csv sau khi được lọc.**



**Đầu ra của sau khi áp dụng mô hình Support vector machine vào file dự doán.**





**Đầu ra của sau khi áp dụng mô hình logistic regression vào file dự đoán.**





**Sự khác nhau giữa mô hình Support vector maching và Logistic regression trong việc phân loại.**

|  |  |
| --- | --- |
| Logistic Regression | Support Vector Machine |
| Thuật toán: sử dụng hàm sigmod :  Sử dụng mô hình Logistic regression trong học máy cho độ chính xác 90%        Mô hình phân loại logistic regression phân loại các loài hoa có độ chính xác cao nhưng thấp hơn SMV. | Thuật toán tìm “đường bay” hyper-plane để phân chia các lớp: Sử dụng mô hình support vector machine trong học máy cho độ chính xác 96,666666666%      SMV phân loại các loài hoa cho độ chính xác cao hơn. |

* 1. **So sánh ngoài lề mục thêm**

**Sự khác nhau giữa fit với partial fit .**

Trong mô hình học máy, fit là một phương thức được sử dụng để huấn luyện mô hình trên tập dữ liệu huấn luyện. Nó sẽ sử dụng dữ liệu huấn luyện để tìm tham số mô hình cho phù hợp và tính toán các giá trị dự báo cho dữ liệu huấn luyện.

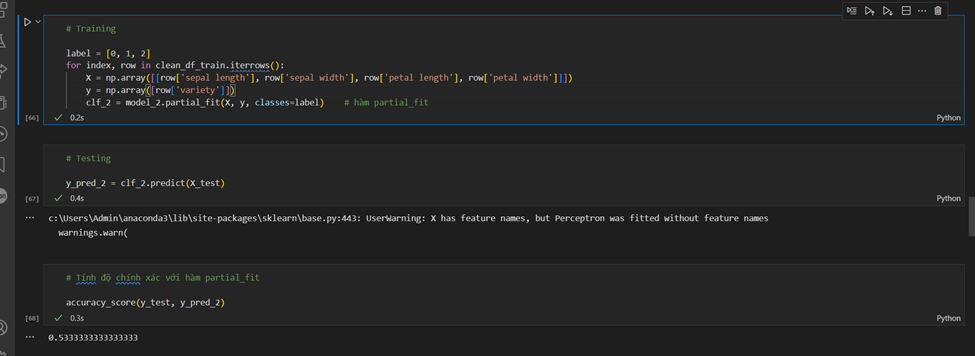
Trong khi đó, partial\_fit là một phương thức tương tự như fit, nhưng nó cho phép bạn huấn luyện mô hình trên một tập dữ liệu nhỏ hơn mà không cần phải truyền toàn bộ dữ liệu huấn luyện vào một lần. Điều này có thể hữu ích khi bạn cần huấn luyện mô hình trên một số lượng dữ liệu lớn mà không muốn sử dụng tất cả dữ liệu đó một lần.

**Hàm Partial\_fit**

Thuật toán perceptron: Cho hai class được gán nhãn, hãy tìm một đường phẳng sao cho toàn bộ các điểm thuộc class 1 nằm về 1 phía, toàn bộ các điểm thuộc class 2 nằm về phía còn lại của đường phẳng đó. Với giả định rằng tồn tại một đường phẳng như thế.

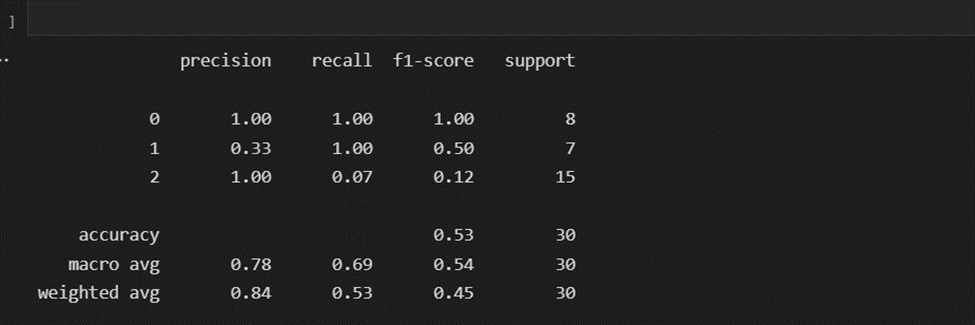
Phương thức trên khi sử dụng trong support vector maching hoặc logistic regression thì bị lỗi.

Sử dụng partial\_fit trong thuật toán perceptron.



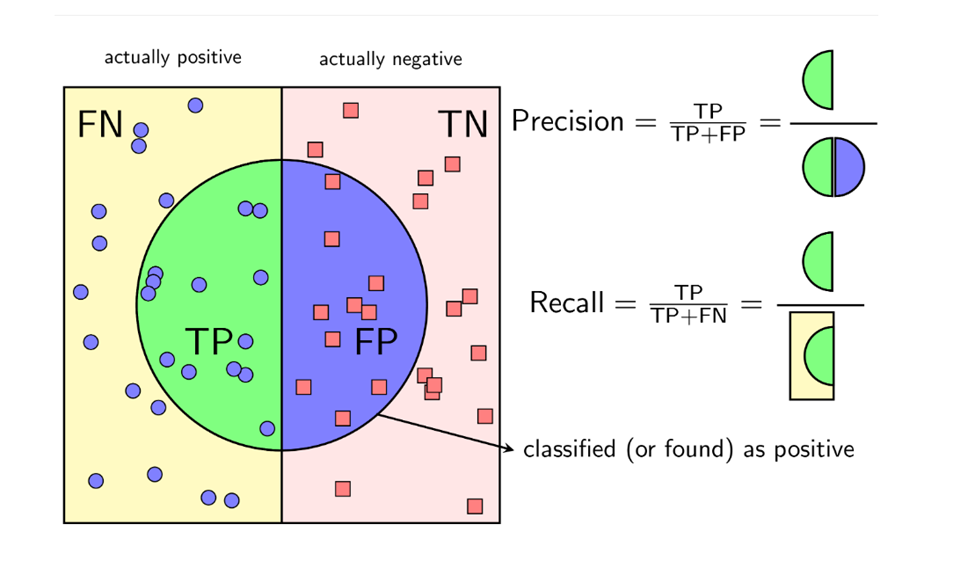
*Kết quả độ chính xác 53,3333333%.*

*Báo cáo cụ thể:*



Với bài toán phân loại mà tập dữ liệu của các lớp là chênh lệch nhau rất nhiều, có một phép đó hiệu quả thường được sử dụng là Precision-Recall.

Ghi chú thêm về partial\_fit().



F1 càng cao độ phân lớp càng tốt

Support là số lượng đối tượng được báo cáo.

1. **Mức độ đóng góp của từng thành viên?**
   1. **Đóng góp của từng thành viên:**

Trảo An Huy:

Tham gia xây dựng kế hoạch qui trình scrum, tham gia việc giám sát tiến độ của các thành viên, viết báo cáo hàng tuần, xây dựng cuộc họp, hổ trợ hướng dẫn các thành viên thực hiện cài đặt môi trường, tìm hiểu về các thuật toán trên web, hổ trợ làm web, hổ trợ sửa lỗi code, viết báo cáo nội dung, viết kết quả đồ án.

Trần Gia Huy:

Thiết kế website,tham gia đưa dữ liệu lên database,tìm hiểu cấu trúc hàm,thuật toán,đồ thị có trong code để phân tích chuyên sâu trong website.

Đinh Xuân Khang: viết code của bài.

Nguyễn Lê Đăng Khoa:

Nghiên cứu thuật ngữ,Test code và Viết báo cáo.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Người nhận xét  Người được  nhận xét | Trảo An Huy | Trần Gia Huy | Đinh Xuân Khang | Nguyễn Lê Đăng Khoa |
| Trảo An Huy |  | Tốt | Tốt | Có trách nhiệm, nghiêm túc trong công việc. |
| Trần Gia Huy | Tốt |  | Tốt | Đẹp trai, có trách nhiệm,có gu thẩm mĩ. |
| Đinh Xuân Khang | Tốt | Tốt |  | Đẹp trai,thân thiện,tinh thần trách nhiệm cao, |
| Nguyễn Lê Đăng Khoa | Tốt | Tốt | Tốt |  |

* 1. **Nhận xét của từng thành viên đối với các thành viên còn lại.**

1. **Những gì học được/ làm được/ đạt được của mỗi thành viên trong nhóm .**

* 1. **Thành tựu chung của tất cả thành viên.**
* Hiểu biết sơ bộ về ngôn ngữ lập trình Python, thư viện , scikit-learn.
* Cách thức hoạt động và sử dụng của Anacoda.
* Thảo luận nhóm, đóng góp ý kiến cá nhân.
* Phân tích, tổng hợp và viết tài liệu kỹ thuật theo mẫu cho trước theo cá nhân hoặc cộng tác nhóm.
* Hiểu được quy trình làm việc của một quy trình cụ thể.
  1. **Thành tựu cá nhân.**

Trảo An Huy:

* Phương pháp quản lí nhóm theo quy trình scrum.
* Hiểu được thuật toán cơ bản trong học máy,
* Học được rất nhiều về các lệnh python.
* Học được cách tạo kế hoạch tiếp theo cho cả nhóm, phân công chức năng cho các thành viên.

Trần Gia Huy:

* Các thiết kế và tìm hiểu nội dung cho website.
* Hiểu về các thuật toán máy học, mongodb để truy xuất và xử lý số liệu.
* Học được thêm về các hàm và thuật toán trong code.
* Các thao tác để có thể thuận lợi trong làm việc nhóm.
* Cách thiết kế và tìm hiểu nội dung cho website, sử dụng hình ảnh minh họa và nội dụng chi tiết dễ hiểu cho 1 website.

Đinh Xuân Khang:

* Lập trình python.
* Biết được quy trình của 1 Data Scientist.
* Xử lý, làm sạch dữ liệu.
* Áp dụng mô hình học máy để đưa ra dự đoán.

Nguyễn Lê Đăng Khoa:

* Biết thêm một số kiến thức về Machine learning,ngôn ngữ lập trình Python,thư viện Scikit-learn,Matplotlib.
* Sử dụng được cơ bản các công cụ:Google colab, Jupyter Notebook.
* Biết cách làm việc nhóm, phân công công việc như thế nào, giải quyết các vấn đề ra sao.

1. **Trình bày những gì chưa học được/ chưa làm được/ chưa đạt được của mỗi thành viên trong nhóm.**

Trảo An Huy:

* Chưa học được việc phân công hợp lí giữa các thành viên, khó khăn trong việc giải quyết các vấn đề nội bộ.
* Tìm hiểu không được trọn vẹn thêm các mô hình.
* Thôi thúc các thành viên một cách thô bạo.
* Làm nhiều điều hơn với mô hình , thử nghiệm nhiều hơn về các mô hình khác nhau.

Trần Gia Huy:

* Chưa hiểu rõ về các thuật toán máy học
* Chưa biết cách tạo được 1 website bằng html và css.
* Chưa có được sự nhanh nhạy trong cách dùng ngôn từ để mô tả thuật toàn và các hàm.
* Áp dụng rập khuôn các thao tác trong hàm cho mọi bài toán

Đinh Xuân Khang:

* Chưa hiểu sâu, rõ về các mô hình máy học.

Nguyễn Lê Đăng Khoa:

* Chưa hoàn thành tốt nhiệm vụ được giao (Không hoàn thành đúng hạn).
* Chưa học được cách xây dựng, bắt đầu giải quyết một bài toán lập trình từ đâu.
* Vẫn còn mắc nhiều lỗi trong quá trình tìm hiểu về đồ án cũng như các mô hình học máy áp dụng cho bài toán cần giải quyết của đồ án.

1. **Dự định tương lai:**

Trảo An Huy:

Tìm hiểu nhiều hơn về các thuật toán, các hàm, nghiên cứu kĩ hơn về dữ liệu quan hệ (SQL) dữ liệu phi quan hệ (NoSQL) ,

Tìm hiểu thêm các cách thức quản lí nhóm, quản lí cá nhân, xây dựng lộ trình học tập về những điều em yêu thích về dữ liêu, nghiên cứu sâu vào cách các lệnh python hoạt động và tối ưu chúng.

Tìm hiểu cách xây dựng hệ thống dữ liệu .

Cập nhật thêm các mô hình có sẳn mới nhất.

Áp dụng các kiến học máy vào các dự án về thị giác máy tính phân loại thực tế như phân loại đồ vật, hoặc áp dụng vào các bài toán kinh doanh, trong thực tiễn, nghiên cứu kĩ về database, clouding computing, lưu dữ liệu lên đám mây.

Trần Gia Huy:

Cải thêm về kiến thức Python,MongoDB và Jupyter Notebooks

Áp dụng thuật toán và các kiến thức đã học được trong đồ án để thực hành với những lượng dữ liệu khác trong tương lai

Học cách tạo ra 1 website bằng ngôn ngữ HTML và CSS.

Tìm hiểu thêm về SQL và các ngôn ngữ quản lý dữ liệu khác để tìm ra điều phù hợp với mình nhất

Tìm hiểu thêm các mô hình máy học thuật toán khác trong việc phân tích dữ liệu

- Áp dụng kỹ năng mềm và kiến thức trong quá trình làm việc nhóm vào các công việc chung sau này.

- áp dụng kiến thức đồ án như thuật toán máy học và

Đinh Xuân Khang: Cải thiện ngoại ngữ, viết code. Và sẽ học thêm nhiều hơn nữa về mảng này

Nguyễn Lê Đăng Khoa:

Cải thiện năng lực ngoại ngữ

-Học thêm một số ngôn ngữ lập trình mới: Python,Java,C#...

-Học thêm một số thứ về việc viết chương trình:OOP,Cấu trúc dữ liệu và giải thuật,Kiến trúc máy tính,Cơ sở dữ liệu…

-Có lẽ kì tới em sẽ đi làm thêm để tìm kiếm thêm các mối quan hệ, tích luỹ kiến thức,kinh nghiệm, trải nghiệm và quan trọng nhất là có tiền để làm điều mình thích.

1. **Áp dụng vào thực tế các kiến thức được học thông qua đồ án.**

Trảo An Huy:

* Áp dụng các kiến học máy vào các dự án về thị giác máy tính phân loại thực tế như phân loại đồ vật, hoặc áp dụng vào các bài toán kinh doanh, trong thực tiễn, nghiên cứu kĩ về database, clouding computing, lưu dữ liệu lên đám mây.

Trần Gia Huy:

* Áp dụng kỹ năng mềm và kiến thức trong quá trình làm việc nhóm vào các công việc chung sau này.
* Áp dụng kiến thức đồ án như thuật toán máy học và mongodb vào các bài toán dữ liệu.

Đinh Xuân Khang: Áp dụng được kỹ năng làm việc nhóm với các thành viên

Đinh Xuân Khang: Áp dụng được kỹ năng làm việc nhóm với các thành viên, thực hiện code .

Nguyễn Lê Đăng Khoa:

* Cải thiện năng lực ngoại ngữ
* Học thêm một số ngôn ngữ lập trình mới: Python,Java,C#...
* Học thêm một số thứ về việc viết chương trình:OOP,Cấu trúc dữ liệu và giải thuật,Kiến trúc máy tính,Cơ sở dữ liệu…
* Có lẽ kì tới em sẽ đi làm thêm để tìm kiếm thêm các mối quan hệ, tích luỹ kiến thức,kinh nghiệm, trải nghiệm và quan trọng nhất là có tiền để làm điều mình thích.