**Logo

Description automatically generated**

**Báo cáo bài tập bonus**

SCIKIT-LEARN VÀ PCA, LDA

BÙI NGỌC CHÍNH-19127109

2022

**Mục lục**

[**Thư viện sử dụng: PyPI-PCA** 1](#_Toc101642067)

[**1.** **Nguồn** 1](#_Toc101642068)

[**2.** **Các chức năng chính được hỗ trợ:** 1](#_Toc101642069)

[**3.** **Cách cài đặt:** 1](#_Toc101642070)

[**4.** **Cách sử dụng:** 1](#_Toc101642071)

[**5.** **So sánh với model PCA của Scikit-Learn:** 2](#_Toc101642072)

[**6.** **Một số hình ảnh trực quan hóa từ PyPI PCA:** 2](#_Toc101642073)

# **Thư viện sử dụng: PyPI-PCA**

## **Nguồn**

<https://erdogant.github.io/pca/pages/html/index.html>

## **Các chức năng chính được hỗ trợ:**

* Biplot để vẽ biểu đồ tải
* Xác định phương sai được giải thích
* Trích xuất các tính năng hoạt động tốt nhất
* Biểu đồ phân tán với các tải
* Phát hiện ngoại lệ bằng cách sử dụng Hotelling T2 và / hoặc SPE / Dmodx

## **Cách cài đặt:**

A picture containing text

Description automatically generated

## **Cách sử dụng:**

Bước 1: import module pca vào chương trình:

Graphical user interface, text

Description automatically generated

Bước 2: Khởi tạo model PCA có sẵn:

Với hàm pca(n\_components=0.95): Khởi tạo để giảm dữ liệu lên đến số lượng các thành phần giải thích 95% phương sai

Với hàm pca(n\_components=3): Giảm dữ liệu xuống 3 thành phần chính

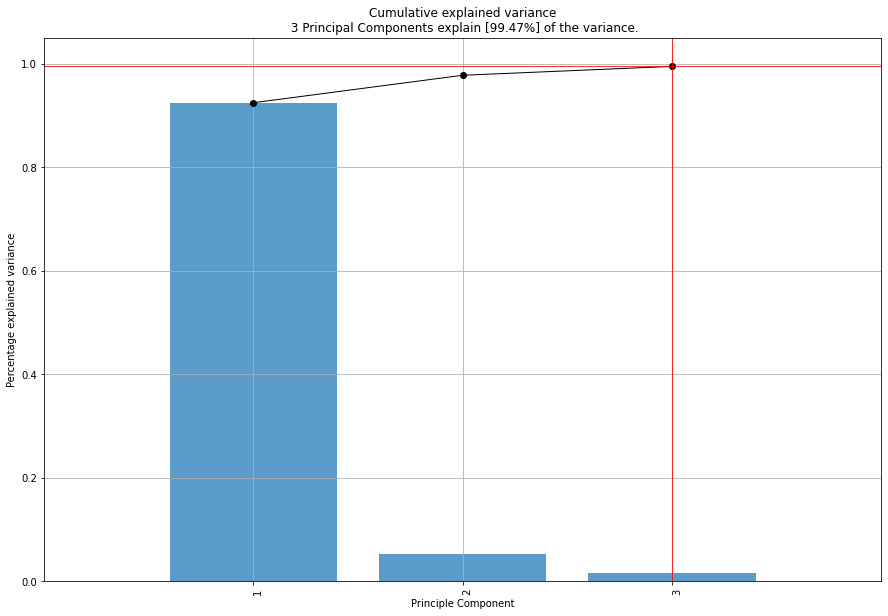
Bước 3: Thực hiện chuyển đổi dữ liệu vào model với hàm fit\_transform(<tập dữ liệu>)

## **So sánh với model PCA của Scikit-Learn:**

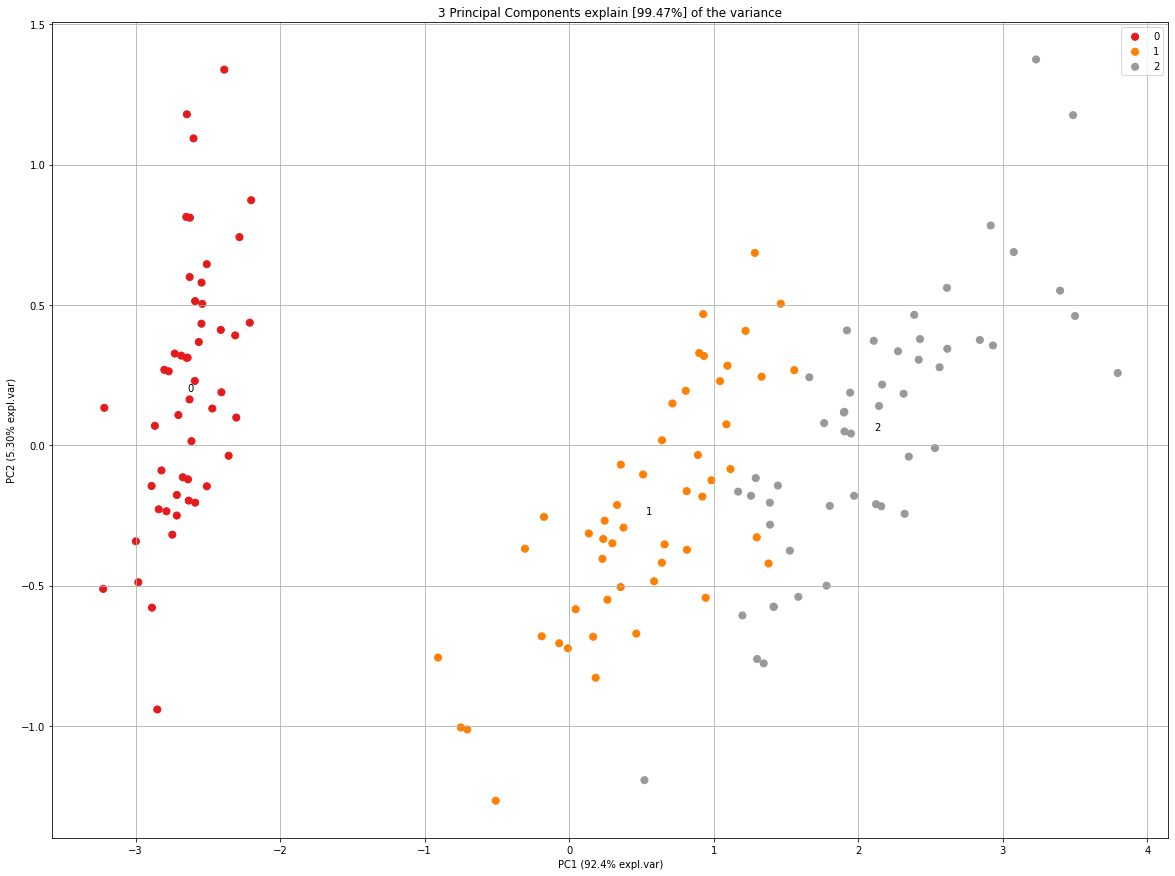
|  |  |  |
| --- | --- | --- |
| **Thuộc tính** | **Scikit-Learn PCA** | **PyPI PCA** |
| Độ phổ biến | Phổ biến hơn | Ít phổ biến hơn |
| Tốc độ thực thi | Chưa test | Chưa test |
| Có hỗ trợ trực quan hóa dữ liệu | Không | Có |
| Dễ sử dụng | Phải qua hàm scale rồi mới pca | Chỉ cần dùng hàm pca |

## 

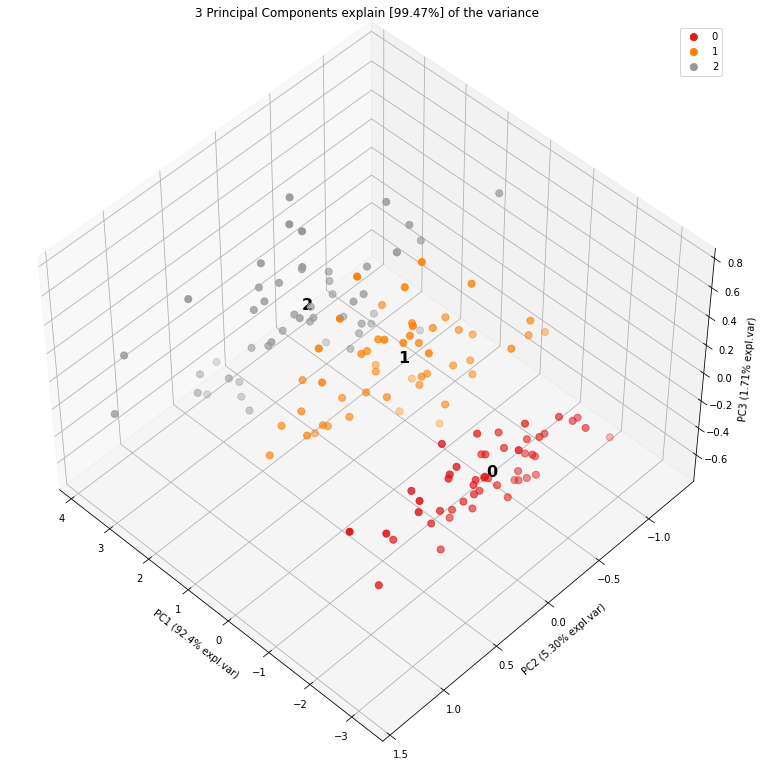
## **Một số hình ảnh trực quan hóa từ PyPI PCA:**



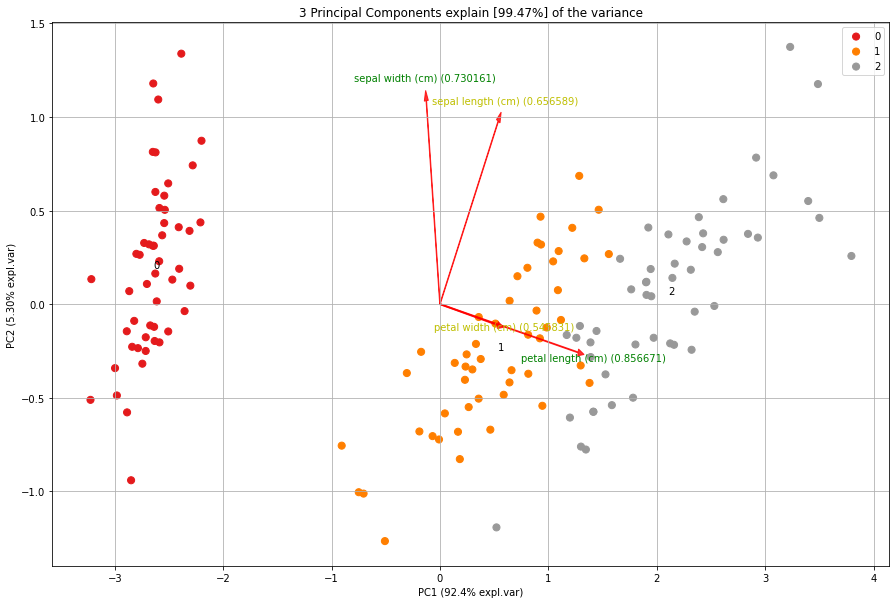
Hình 1. Biểu đồ thể hiện với 3 thành phần chính đã giải thích được hơn 99% phương sai



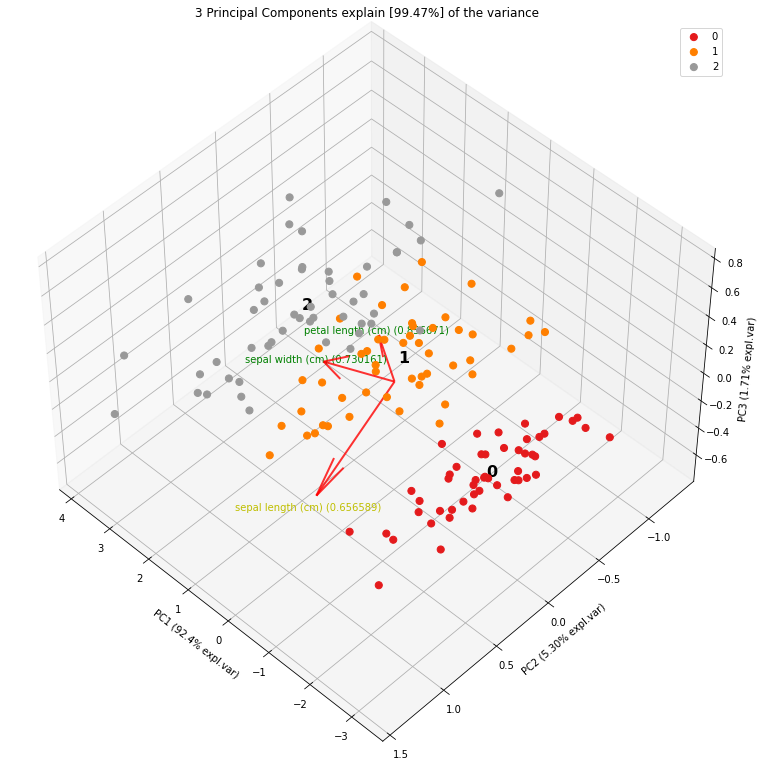
Hình 2. Biểu đồ Scatter 2D thể hiện 3 thành phần chính



Hình 3. Biểu đồ Scatter 3D thể hiện 3 thành phần chính



Hình 4. Biplot 2D



Hình 5. Biplot 3D