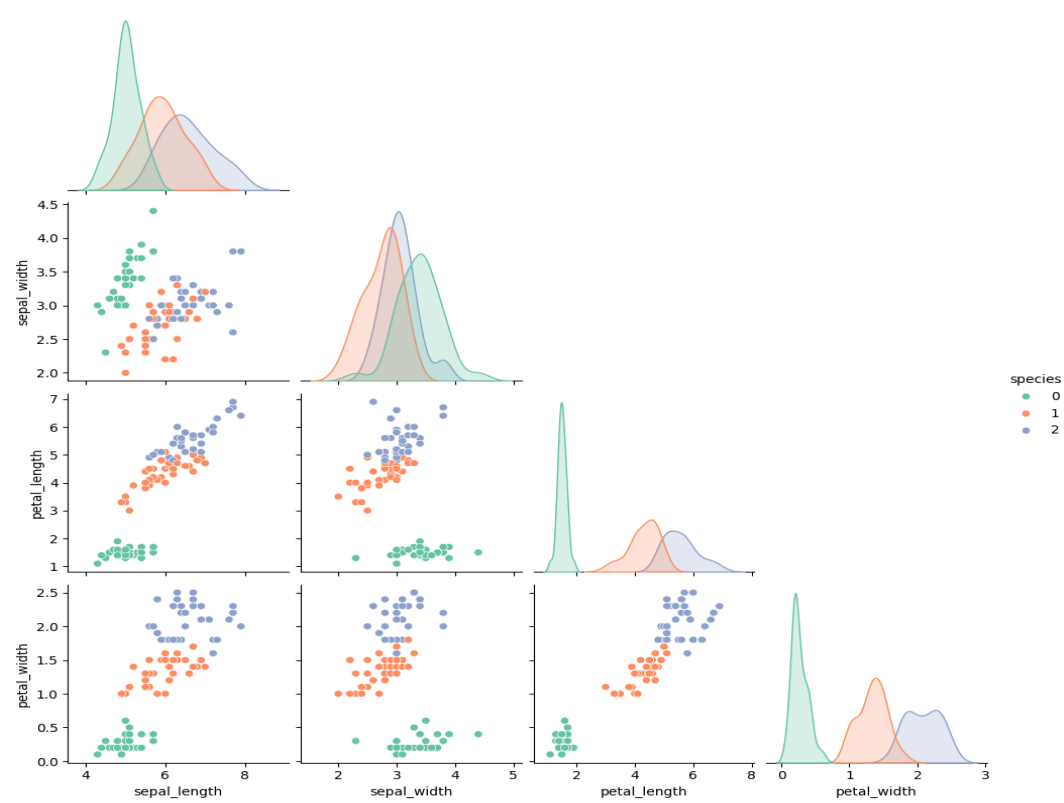


# PHÂN TÍCH & DỰ ĐOÁN DỮ LIỆU IRIS

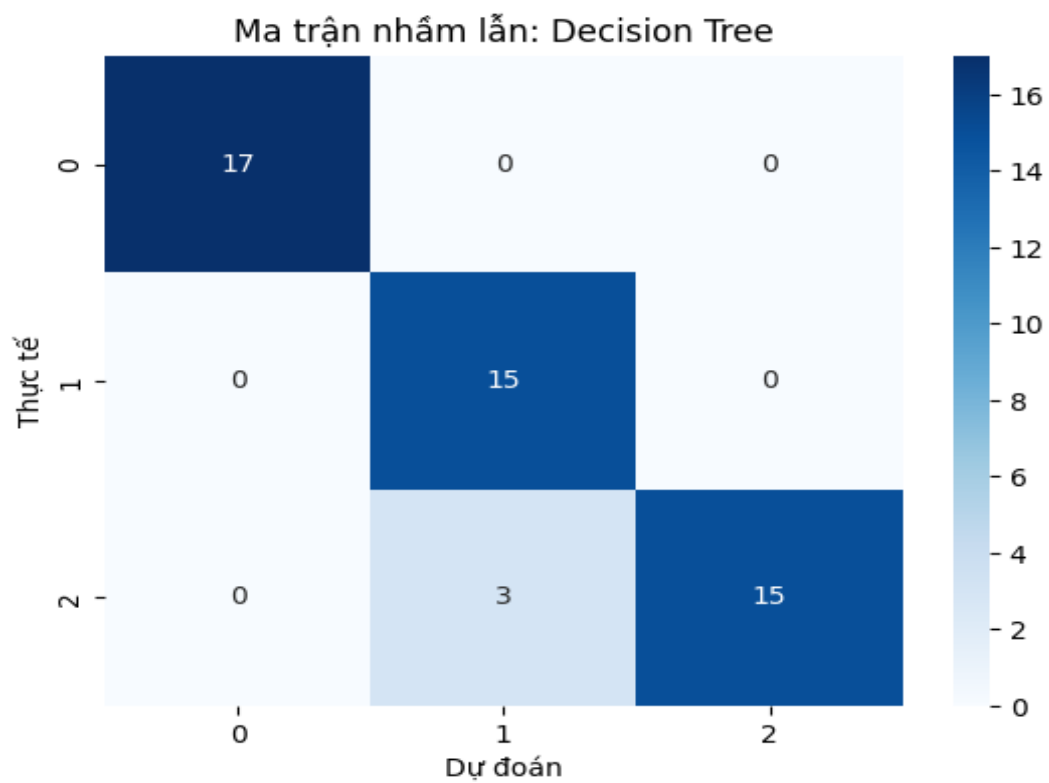
Để có cái nhìn trực quan về tập dữ liệu, chúng ta cần hiển thị dữ liệu. Có rất nhiều phương pháp trực quan hóa dữ liệu, tuy nhiên, đối với tập Iris, tôi dùng sơ đồ phân tán để hiển thị dữ liệu

Biểu đồ phân tán (Scatter Plot) thể hiện mối quan hệ giữa các cặp đặc trưng trong tập dữ liệu Iris. Mỗi điểm dữ liệu đại diện cho một bông hoa, được tô màu theo loài (species).



Báo cáo này trình bày kết quả huấn luyện và so sánh 5 mô hình Machine Learning trên tập dữ liệu IRIS.

Mô hình	■■ chính xác
Logistic Regression	0.92
Decision Tree	0.94
KNN	0.92
NB	0.92
SVM	0.92



Mô hình có hiệu suất cao nhất là: Decision Tree.