**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BỘ MÔN CƠ SỞ TRÍ TUỆ NHÂN TẠO**

---o0o---



**Lab Bonus: Decision Tree**

**GIÁO VIÊN HƯỚNG DẪN THỰC HÀNH :**

**cô Nguyễn Thị Thu Hằng và thầy Nguyễn Ngọc Đức**

# **MỤC LỤC**

[**MỤC LỤC** 2](#_Toc185520684)

[**THÔNG TIN SINH VIÊN** 3](#_Toc185520685)

[**MỨC ĐỘ HOÀN THÀNH** 3](#_Toc185520686)

[**Kết quả thống kê và nhận xét** 5](#_Toc185520687)

[**1/ Breast Cancer Datasets** 5](#_Toc185520688)

[**2/ Wine Quality Dataset** 6](#_Toc185520689)

[**3/ Iris Dataset** 7](#_Toc185520690)

[**Tài liệu tham khảo** 8](#_Toc185520691)

# **THÔNG TIN SINH VIÊN**

|  |  |  |
| --- | --- | --- |
| STT | MSSV | Họ và tên |
| 1 | [22120376](mailto:%32%32%31%32%30%33%376@%73%74%75d%65n%74%2ehc%6d%75%73%2e%65%64u%2e%76%6e) | Nguyễn Đức Toàn |
| 2 | 22120430 | Lê Hoàng Việt |
| 3 | [22120435](mailto:2%32%31%320%34%335@%73%74%75%64%65%6et%2e%68%63m%75%73.%65%64u%2e%76n) | Nguyên Quốc Vinh |

# **MỨC ĐỘ HOÀN THÀNH**

|  |  |  |
| --- | --- | --- |
| **Task** | **Mức độ hoàn thành** | **Thành viên** |
| Chuẩn bị dữ liệu (Data Preparation) | 100% | Lê Hoàng Việt |
| Tổng hợp báo cáo | 100% | Lê Hoàng Việt |
| Xây dựng Decision Tree | 100% | Nguyễn Quốc Vinh |
| Ảnh hưởng của độ sâu | 100% | Nguyễn Quốc Vinh |
| Đánh giá mô hình | 100% | Nguyễn Đức Toàn |
| Phân tích trên các dataset khác | 100% | Nguyễn Đức Toàn |
| Mức độ hoàn thành các Task/ Đồ án: 100% | | |

# **Kết quả thống kê và nhận xét**

## **1/ Breast Cancer Datasets**

**Độ chính xác:** 0.90

**Cây quyết định:**

A diagram of a company

Description automatically generated with medium confidence

**Confusion Matrix:**

A diagram of confusion matrix

Description automatically generated

**Nhận xét:** Mô hình Decision Tree cho bộ dữ liệu Breast Cancer đạt độ chính xác 0.90. Mô hình này phân loại chính xác hầu hết các mẫu, nhưng vẫn có một số mẫu bị phân loại sai, đặc biệt là các mẫu thuộc lớp Malignant.

## **2/ Wine Quality Dataset**

**Độ chính xác:** 0.83

**Cây quyết định:**

A screenshot of a computer

Description automatically generated

**Confusion Matrix:**

A chart with different colored squares

Description automatically generated with medium confidence

**Nhận xét:** Mô hình Decision Tree cho bộ dữ liệu Wine Quality đạt độ chính xác 0.83. Mô hình này gặp khó khăn trong việc phân loại các mẫu thuộc lớp Low và High, dẫn đến nhiều mẫu bị phân loại sai.

## **3/ Iris Dataset**

**Độ chính xác:** 0.97

**Cây quyết định:**

A diagram of a number of numbers

Description automatically generated with medium confidence

**Confusion Matrix:**

**A diagram of a confusion matrix

Description automatically generated**

**Nhận xét:** Mô hình Decision Tree cho bộ dữ liệu Iris đạt độ chính xác cao nhất là 0.97. Mô hình này phân loại chính xác hầu hết các mẫu, chỉ có một mẫu thuộc lớp Iris-versicolor bị phân loại sai thành Iris-virginica.

# **Tài liệu tham khảo**

**UCI Machine Learning Repository:** Breast Cancer Wisconsin (Diagnostic) Data Set

**UCI Machine Learning Repository:** Wine Quality Data Set

**UCI Machine Learning Repository:** Iris Data Set

**ChatGPT OpenAi**