|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TRƯỜNG ĐẠI HỌC ĐIỆN LỰC**  **KHOA CÔNG NGHỆ THÔNG TIN**    **BÁO CÁO THỰC TẬP TỐT NGHIỆP**  **HỆ THỐNG THÔNG TIN QUẢN LÝ**  **ĐỀ TÀI:**   |  |  | | --- | --- | | **Sinh viên thực hiện** | **: TRẦN TUẤN ANH** | | **Giảng viên hướng dẫn** | **: LÊ HOÀN** | | **Ngành** | **: CÔNG NGHỆ THÔNG TIN** | | **Chuyên ngành** | **: CÔNG NGHỆ PHẦN MỀM** | | **Lớp** | **: D14CNPM1** | | **Khóa** | **: 2019-2024** |   ***Hà Nội,5 tháng 5 năm 2023*** |

**PHIẾU CHẤM ĐIỂM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STT** | **Họ và tên sinh viên** | **Nội dung thực hiện** | **Điểm** | **Chữ ký** |
| 1 | TRẦN TUẤN ANH  (19810310026) |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |

|  |  |  |
| --- | --- | --- |
| **Họ và tên giảng viên** | **Chữ ký** | **Ghi chú** |
| Giảng viên chấm 1: |  |  |
| Giảng viên chấm 2: |  |  |

**MỤC LỤC**

**DANH SÁCH CÁC HÌNH ẢNH**

[**Hình 3.1: Tương quan của khai phá dữ liệu 11**](#_Toc106108541)

[**Hình 3.2: Dữ liệu thử nghiệm sử dụng để ước lượng độ chính xác của quy tắc phân loại 12**](#_Toc106108542)

[**Hình 3.3: Dữ liệu thử nghiệm được sử dụng để ước tính độ chính xác của quy tắc phân lớp 13**](#_Toc106108543)

[**Hình 3.4: Mô tả thuật toán Gaussian Naive Bayes 15**](#_Toc106108544)

[**Hình 3.5: Mô tả thuật toán Bernoulli Naive Bayes 15**](#_Toc106108545)

[**Hình 3.6: Mô tả thuật toán Multinomial Naive Bayes 16**](#_Toc106108546)

[**Hình 4.1: Cấu trúc thư mục của dự án 17**](#_Toc106108547)

[**Hình 4.2: Kết quả đọc file CSV 17**](#_Toc106108548)

[**Hình 4.3: Kết quả sau khi chuyển đổi dữ liệu sang dạng số 18**](#_Toc106108549)

[**Hình 4.4: Kết quả trích chọn cột dùng để làm dữ liệu huấn luyện và nhãn 18**](#_Toc106108550)

[**Hình 4.5: Kết quả sau khi chia tỉ lệ giữa dữ liệu để huấn luyện và dữ liệu để học 19**](#_Toc106108551)

[**Hình 4.6: Số các nhãn được đếm trong tập huấn luyện 19**](#_Toc106108552)

[**Hình 4.7: Các nhãn dự đoán được sau khi dùng mô hình GaussianNB 20**](#_Toc106108553)

[**Hình 4.8: Tỉ lệ chính xác khi đoán nhãn khi dùng mô hình GaussianNB + Báo cáo của việc huấn luyện: 20**](#_Toc106108554)

[**Hình 4.9: Báo cáo sau khi dùng mô hình GaussianNB để huấn luyện 20**](#_Toc106108555)

[**Hình 4.10: Ma trận nhầm lẫn sau khi dùng mô hình GaussianNB để huấn luyện 21**](#_Toc106108556)

[**Hình 4.11: Các nhãn dự đoán được sau khi dùng mô hình BernoulliNB 22**](#_Toc106108557)

[**Hình 4.12: Tỉ lệ chính xác khi đoán nhãn khi dùng mô hình BernoulliNB 22**](#_Toc106108558)

[**Hình 4.13: Báo cáo sau khi dùng mô hình BernoulliNB để huấn luyện 22**](#_Toc106108559)

[**Hình 4.14: Ma trận nhầm lẫn sau khi dùng mô hình BernoulliNB để huấn luyện 23**](#_Toc106108560)

[**Hình 4.15: Các nhãn dự đoán được sau khi dùng mô hình MultinomialNB 24**](#_Toc106108561)

[**Hình 4.16: Tỉ lệ chính xác khi đoán nhãn khi dùng mô hình MultinomialNB 24**](#_Toc106108562)

[**Hình 4.17: Báo cáo sau khi dùng mô hình MultinomialNB để huấn luyện 24**](#_Toc106108563)

[**Hình 4.18: Ma trận nhầm lẫn sau khi dùng mô hình MultinomialNB để huấn luyện 25**](#_Toc106108564)

**DANH SÁCH CÁC BẢNG**

[Bảng 4.1: Bảng so sánh Accuracy 25](#_Toc106108592)

[Bảng 4.2: So sánh Precision của các mô hình 26](#_Toc106108593)

[Bảng 4.3: So sánh Recall của các mô hình 27](#_Toc106108594)

[Bảng 4.4: So sánh f1-score của các mô hình 27](#_Toc106108595)

# LỜI CẢM ƠN

Chúng em xin chân thành cảm ơn!