**Mô tả đồ án môn học**

1. **Mô tả đề tài:**

Công cụ phân tích dữ liệu mà nhóm em đang phát triển là một giải pháp toàn diện, linh hoạt, dễ sử dụng, được thiết kế để hoạt động trên ngôn ngữ Python, nền tảng Notebook và tích hợp vào các phần mềm. Với việc kết hợp các chức năng chính như chuẩn bị dữ liệu, khám phá dữ liệu, tiền xử lý, phân tích dữ liệu, đánh giá và tối ưu hóa mô hình, trực quan hóa kết quả và ứng dụng machine learning, công cụ này đáp ứng được nhu cầu của người dùng trong việc xử lý và phân tích dữ liệu một cách hiệu quả và chính xác.

Bằng cách tận dụng sức mạnh của Pandas Dataframe, các kỹ thuật trực quan hóa dữ liệu phổ biến và các mô hình machine learning, công cụ này giúp người dùng dễ dàng chuẩn bị và phân tích dữ liệu từ nhiều nguồn khác nhau, đồng thời đưa ra dự đoán và quyết định kinh doanh chính xác. Đặc biệt, công cụ này phù hợp với cả những người mới bắt đầu và những chuyên gia dữ liệu, giúp họ tiết kiệm thời gian và nỗ lực trong quá trình phân tích dữ liệu.

Nói tóm lại, công cụ phân tích dữ liệu này mang đến một giải pháp toàn diện và dễ sử dụng cho người dùng để giải quyết các vấn đề liên quan đến dữ liệu trong công việc kinh doanh, nghiên cứu và phát triển. Với sự tích hợp mạnh mẽ của Python và Notebook, công cụ này hứa hẹn sẽ trở thành một công cụ đắc lực cho các nhà phân tích dữ liệu, chuyên gia và nhà quản lý.

1. **Chức năng chính:**
2. Chuẩn bị dữ liệu: Công cụ sẽ xác định và thu thập dữ liệu từ các nguồn khác nhau, sau đó lưu trữ chúng vào các đối tượng Pandas Dataframe trong Python, giúp dễ dàng quản lý và xử lý dữ liệu.
3. Khám phá dữ liệu: Công cụ giúp trực quan hóa dữ liệu thông qua các biểu đồ, biểu đồ hộp, biểu đồ phân phối và thống kê mô tả cơ bản, giúp người dùng hiểu rõ hơn về đặc điểm và xu hướng của dữ liệu.
4. Tiền xử lý dữ liệu: Công cụ hỗ trợ xử lý dữ liệu để chuẩn bị cho việc phân tích, bao gồm việc loại bỏ giá trị null, xử lý ngoại lệ và chuẩn hóa dữ liệu.
5. Phân tích dữ liệu: Công cụ cho phép người dùng tạo ra các mô hình dự đoán giá trị kinh doanh thông qua các phương pháp như hồi quy tuyến tính, phân loại, cây quyết định và mạng neuron.
6. Đánh giá và tối ưu hóa mô hình: Công cụ đánh giá hiệu suất của mô hình thông qua các chỉ số đo lường và tối ưu hóa mô hình để cải thiện dự đoán, giúp người dùng đạt được kết quả tốt nhất.
7. Trực quan hóa kết quả: Công cụ trực quan hóa kết quả của mô hình, giúp người dùng dễ dàng nhận thức được hiệu quả của các dự đoán và đưa ra quyết định kinh doanh phù hợp.
8. Xử dụng machine learning: Công cụ sử dụng các kỹ thuật machine learning để dự đoán, giúp người dùng đạt được kết quả chính xác và hiệu quả hơn.