

GỢI Ý MỘT SỐ CHỦ ĐỀ PHÂN TÍCH DỮ LIỆU

STT	Chủ đề	Website dữ liệu	Ghi chú
1	Phân tích dữ liệu hình ảnh y khoa: sử dụng các kỹ thuật học máy để phân tích hình ảnh y khoa: X-quang, MRI, và CT-scan. Phân tích dữ liệu hình ảnh y khoa: chẩn đoán bệnh sớm, phân tích hiệu quả điều trị, và nghiên cứu các bệnh lý mới.	https://radiopaedia.org/	Các nhóm có thể chọn theo dữ liệu hình ảnh khác nhau - X-quang, MRI,... - Bệnh viêm phổi,...
2	Phân tích dữ liệu hành vi sức khỏe: sử dụng để dự đoán nguy cơ mắc bệnh: bệnh tim mạch, bệnh tiểu đường, và ung thư. Các mô hình dự đoán này có thể được sử dụng để xác định những người có nguy cơ mắc bệnh cao, từ đó có thể đưa ra các biện pháp can thiệp sớm để ngăn ngừa bệnh tật.	https://www.kaggle.com/	
3	Phân tích dữ liệu để hiểu nhu cầu và hành vi của khách hàng. Phân tích dữ liệu có thể giúp các doanh nghiệp hiểu rõ hơn về khách hàng và từ đó có thể đưa ra các quyết định kinh doanh sáng suốt hơn.	https://www.kaggle.com/	
4	Mua sắm trực tuyến: Phân tích các mẫu mua hàng trực tuyến để tìm hiểu các luật kết hợp giữa các mặt hàng. - Dữ liệu về lịch sử mua hàng: sử dụng để xác định các mặt hàng thường được mua cùng nhau, các mặt hàng thường được mua trước hoặc sau một mặt hàng khác, và các mặt hàng thường được mua bởi các khách hàng có chung sở thích. - Dữ liệu về dữ liệu vị trí: sử dụng để xác định các mặt hàng thường được mua ở cùng một địa điểm, các mặt hàng thường được mua bởi các khách hàng ở cùng một khu vực, và các mặt hàng thường được mua vào cùng một thời điểm. - Dữ liệu về dữ liệu thanh toán: sử dụng để xác định các mặt hàng thường được mua bởi cùng một khách hàng, các mặt hàng thường được mua cùng với một phương thức thanh toán cụ thể, và các mặt hàng thường được mua vào cùng một thời điểm trong năm.	https://www.kaggle.com/	

STT	Chủ đề	Website dữ liệu	Ghi chú
5	<p>Phân tích dữ liệu doanh nghiệp để cải thiện hiệu quả của các chiến dịch tiếp thị và tăng doanh số bán hàng: Nghiên cứu các luật kết hợp trong các chiến dịch tiếp thị để hiểu cách các yếu tố tiếp thị ảnh hưởng đến việc mua hàng của khách hàng.</p> <ul style="list-style-type: none"> - Phân tích dữ liệu về hiệu quả của chiến dịch tiếp thị để xác định giá cả nào có hiệu quả nhất trong việc thúc đẩy doanh số bán hàng. - Phân tích dữ liệu từ dữ liệu khách hàng để xác định quảng cáo nào có hiệu quả nhất trong việc thu hút khách hàng tiềm năng. - Phân tích dữ liệu từ dữ liệu thị trường để xác định khuyến mãi nào có hiệu quả nhất trong việc thúc đẩy doanh số bán hàng trong các dịp lễ. 	https://www.kaggle.com/	
6	<p>Phân tích dữ liệu sản xuất và chuỗi cung ứng để tìm hiểu mối quan hệ giữa các yếu tố như nguyên liệu, quy trình sản xuất và chất lượng sản phẩm là một chủ đề nghiên cứu quan trọng trong lĩnh vực sản xuất và chuỗi cung ứng - giúp các nhà sản xuất và nhà quản lý chuỗi cung ứng hiểu rõ hơn về cách các yếu tố khác nhau ảnh hưởng đến chất lượng sản phẩm, có thể cải thiện hiệu quả sản xuất và giảm thiểu rủi ro.</p>	https://www.kaggle.com/	
7	<p>Nghiên cứu các luật kết hợp trong dữ liệu hành vi người dùng trên mạng xã hội để hiểu mối quan hệ giữa các hoạt động như like, comment và chia sẻ là một chủ đề nghiên cứu quan trọng trong lĩnh vực khoa học dữ liệu - giúp các nhà nghiên cứu hiểu rõ hơn về cách người dùng tương tác với nhau và với nội dung trên mạng xã hội, từ đó có thể phát triển các chiến lược tiếp thị và truyền thông hiệu quả hơn.</p> <ul style="list-style-type: none"> - Tìm hiểu mối quan hệ giữa việc like và comment - phát hiện ra rằng những người like một bài đăng có nhiều khả năng comment hơn những người không like. - Tìm hiểu mối quan hệ giữa việc comment và chia sẻ - phát hiện ra rằng những người comment một bài đăng có nhiều khả năng chia sẻ hơn những người không comment. 	https://www.kaggle.com/	

STT	Chủ đề	Website dữ liệu	Ghi chú
	- Tìm hiểu mối quan hệ giữa việc like, comment và chia sẻ - phát hiện ra rằng những người like, comment và chia sẻ một bài đăng có nhiều khả năng tương tác với bài đăng đó hơn những người chỉ like hoặc comment.		
8	<p>Phân tích dữ liệu thời tiết để tìm hiểu mối quan hệ giữa các yếu tố như nhiệt độ, độ ẩm và mưa là một chủ đề nghiên cứu quan trọng trong lĩnh vực khoa học dữ liệu - giúp các nhà nghiên cứu hiểu rõ hơn về cách các yếu tố thời tiết khác nhau ảnh hưởng lẫn nhau, từ đó có thể dự báo thời tiết chính xác hơn và phát triển các giải pháp ứng phó với biến đổi khí hậu.</p> <p>- Tìm hiểu mối quan hệ giữa nhiệt độ và độ ẩm - phát hiện ra rằng nhiệt độ và độ ẩm có mối quan hệ nghịch biến, nghĩa là khi nhiệt độ tăng, độ ẩm thường giảm.</p> <p>- Tìm hiểu mối quan hệ giữa độ ẩm và mưa - phát hiện ra rằng độ ẩm cao có liên quan đến lượng mưa nhiều hơn.</p> <p>- Tìm hiểu mối quan hệ giữa nhiệt độ, độ ẩm và mưa - phát hiện ra rằng nhiệt độ và độ ẩm cao có liên quan đến lượng mưa nhiều hơn.</p>		
9	<p>Nghiên cứu về việc sử dụng luật kết hợp để dự đoán giá cổ phiếu và xác định các mô hình dự báo giá cổ phiếu:</p> <p>- Phát triển các luật kết hợp mới: Các nhà nghiên cứu có thể phát triển các luật kết hợp mới để cải thiện độ chính xác của các mô hình dự đoán giá cổ phiếu.</p> <p>- Tối ưu hóa các luật kết hợp hiện có: Các nhà nghiên cứu có thể tối ưu hóa các luật kết hợp hiện có để cải thiện hiệu suất của chúng.</p> <p>- Xác định các mô hình dự báo giá cổ phiếu hiệu quả: Các nhà nghiên cứu có thể xác định các mô hình dự báo giá cổ phiếu hiệu quả bằng cách sử dụng phân tích luật kết hợp.</p>		

Ngoài ra, các nhóm có thể chọn chủ đề theo chương 2 đến chương 11 (Learning Data Mining with Python của tác giả Robert Layton).

Quy cách đồ án:

- Bài báo cáo chi tiết (file mềm .DOCx) theo cấu trúc tham khảo link: <https://www.studocu.com/vn/document/truong-dai-hoc-kinh-te-thanh-pho-ho-chi-minh/data-science/do-an-khoa-hoc-du-lieu-thay-nguyen-manh-tuan/42760794>
- Sildes trình chiếu (.PPTx);
- Chương trình demo quá trình phân tích dữ liệu (gồm cả tiền xử lý và biến đổi dữ liệu);
- Mỗi nhóm trình bày 15 phút;

Lưu ý: Trong báo cáo chi tiết cần thể hiện bảng phân công công việc trong nhóm.