



BÁO CÁO KIỂM TRA TRÙNG LẶP

Thông tin tài liệu

Tên tài liệu: DATN_Pham_Van_Anh_2051063673_62HT(test)

Tác giả: Trần Mạnh Tuấn - HTTT

Điểm trùng lặp: 9

Thời gian tải lên: 07:27 07/01/2025

Thời gian sinh báo cáo: 07:37 07/01/2025

Các trang kiểm tra: 57/57 trang



Kết quả kiểm tra trùng lặp



Có 91% nôi

0%

0%

Có 9% nội dung trùng lặp Có 91% nội dung không trùng lặp Có 0% nội dung người dùng loại trừ

Có 0% nội dung hệ thống bỏ qua

Nguồn trùng lặp tiêu biểu

123docz.net tailieu.vn rdsic.edu.vn

Danh sách các câu trùng lặp

Câu 1. Trang 5: các lĩnh vực như Cơ sở dữ liệu học, máy, trí tuệ nhân tạo, lý thuyết thông, tin thống kê xác suất và tính toán hiệu năng cao, đóng vai trò cốt lõi trong việc phát triển các kỹ thuật Khai phá dữ liêu

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trong lĩnh vực khai phá dữ liệu, phần lớn được thừa kế từ <u>lĩnh vực cơ sở dữ liệu,</u> học máy, trí tuê nhân tạo, lý thuyết thông tin, xác suất thống kê và tính toán hiệu năng cao

Câu 2. Trang 6: học máy được chứng minh <u>là một công cụ quan trong trong việc phân tích các hệ thống phức tạp giúp nâng cao độ chính <u>Xác trong dư đoán</u></u>

Độ trùng lặp: 53%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>công cụ quan trọng trong việc phân tích</u> và <u>dự đoán</u> kết quả <u>trong các hệ thống phức tạp.</u> Thách Thức Khi Áp Dụng <u>xác</u> Suất Có Điều Kiện <u>xác</u> suất có điều kiện <u>là một công cụ</u> mạnh mẽ <u>trong</u> toán <u>Hoc</u>

Câu 3. Trang 6: trí tuệ nhân tạo (Artificial Intelligence) các kỹ thuật như Mạng Nơ ron nhân tạo

Độ trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Các kỹ thuật Trí tuê nhân tạo như mạng nơ ron nhân tạo (

Câu 4. Trang 6: cơ sở dữ liệu (Database) Hệ quản trị cơ sở dữ liệu (DBMS) và ngôn ngữ truy vấn, như SQL đóng vai trò quan trọng trong việc quản lý Lưu trữ và truy xuất dữ liệu một cách hiệu quả tạo nền tảng cho Các hoạt động khai phá dữ liệu

Đô trùng lặp: 64%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Hệ quản trị Cơ sở dữ liệu (DBMS)</u> đóng vai trò quan trọng trong việc quản lý, và vận hành <u>dữ liệu các</u> chức năng chính của DBMS bao gồm <u>lưu trữ và</u> tổ chức <u>dữ liệu</u> DBMS giúp <u>lưu trữ dữ liệu một cách</u> có cấu trúc, theo <u>các</u> mô hình cụ thể <u>như quan Hệ</u> phân cấp hoặc đồ thị <u>việc</u> tổ chức này giúp dễ dàng <u>truy xuất và quản lý, dữ liệu truy xuất dữ liệu</u> DBMS cung cấp <u>các</u> công cụ <u>và ngôn ngữ như SQL</u> để người dùng có thể <u>truy vấn</u> tìm kiếm <u>và</u> lấy <u>dữ liệu</u> từ <u>Cơ sở dữ liệu</u> môt cách hiệu quả.

Câu 5. Trang 6: Silberschatz <u>và</u> cộng sự (2010) <u>đã có những đóng góp quan trọng trong việc phát triển các phương pháp</u> xử lý <u>và</u> quản lý dữ liệu

Độ trùng lặp: 53%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đã có những đóng góp quan trọng trong việc phát triển các</u> công nghệ y tế mới <u>và</u> <u>các phương pháp</u>

Câu 6. Trang 7: dữ liệu này có thể gồm cả định dạng có cấu trúc như dữ liệu bảng và không có cấu trúc là văn bản, hình ảnh, hoặc video

Độ trùng lặp: 58%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>Dữ liệu này có thể là Dữ liệu có cấu trúc</u> (như cơ sở <u>Dữ liệu</u> SQL) <u>hoặc Dữ liệu</u> không có cấu trúc (như các têp văn bản, hình ảnh hoặc

Câu 7. Trang 8: Trích chọn dữ liệu là quá trình sàng lọc và lựa chọn những dữ liệu phù hợp, từ tập dữ liệu thô để chuẩn bị cho quá trình khai phá dữ liệu

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Trích chọn dữ liệu là</u> bước <u>chọn</u> ra <u>những tập dữ liệu phù hợp</u> cần được <u>khai phá</u> trong <u>tập dữ liệu</u> lớn Bước Tiền xử lý <u>dữ liệu là</u> bước làm sạch <u>dữ liệu</u> như xử lý <u>dữ liệu</u> thiếu, <u>dữ liệu</u> nhiễu, <u>dữ liệu</u> không nhất quán Bước Chuyển đổi <u>dữ liệu là</u> bước <u>chuẩn</u> hóa <u>dữ liệu để</u> đưa <u>dữ liêu</u> về dạng <u>phù hợp cho quá trình khai phá dữ liêu</u>

Câu 8. Trang 8: tiền xử lý dữ liệu là bước chuẩn bị và biến đổi <u>dữ liệu</u> ban đầu trước khi áp dụng các phương pháp phân tích hoặc xây dựng mô <u>hình</u>

Độ trùng lặp: 50%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>chuẩn bị và xử lý dữ liệu hình</u> ảnh <u>trước khi áp dụng các</u> thuâ t tốn <u>và phương</u> <u>pháp phân tích hình ảnh Mục đích của Tiền xử lý</u> ảnh <u>là</u>

Câu 9. Trang 9: Tri thức cuối cùng cần được trình bày một cách rõ ràng và dễ hiểu Để người dùng cuối có thể ứng dụng

Độ trùng lặp: 60%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>cần được trình bày một cách</u> trực quan <u>và dễ hiểu để người dùng cuối có thể</u>

Câu 10. Trang 9: phân tích hồi quy là một phương pháp thống kê hữu ích để khám phá mối liên hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập

Độ trùng lặp: 81%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: hồi quy là một phương pháp thống kê được sử dụng <u>để</u> xác định <u>mối</u> quan <u>hệ</u> giữa một biến phụ thuộc và một hoặc nhiều biến độc lập hồi quy</u> tuyến tính đơn <u>biến là</u> trường hợp đơn giản nhất trong <u>Phân tích hồi quy</u> trong đó chỉ có <u>một biến độc lập</u> được sử dụng <u>để</u> dự đoán <u>biến phụ thuộc phương</u> trình <u>hồi quy</u> tuyến tính được sử dụng <u>để</u> tìm ra <u>mối liên hệ</u>

Câu 11. Trang 9: Mục tiêu chính của phương pháp này <u>là</u> xây dựng <u>Một</u> mô hình <u>dự đoán giá tri</u> của biến phụ thuộc dựa trên các yếu tố tác động <u>từ các biến độc lập</u>

Độ trùng lặp: 61%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Mục tiêu chính của</u> phân tích hồi quy <u>là dự đoán giá trị của biến phụ thuộc dựa</u> trên thông tin có sẵn <u>từ các biến độc lập một</u> ví dụ cụ thể có thể <u>là</u> việc <u>dự đoán giá của một</u> sản phẩm <u>dựa trên các yếu tố</u>

Câu 12. Trang 9: Thu thập dữ liệu Tập hợp các thông tin liên quan đến biến phụ thuộc và các biến độc lập nhằm phục vụ cho quá trình phân tích và Xây dựng mô hình

Đô trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đến biến phu thuộc</u> 3 <u>Thu thập dữ liêu Thu thập thông tin</u> về <u>các biến phu thuộc</u> và các biến độc lập từ mẫu nghiên cứu hoặc nguồn dữ liệu 4 xây dựng mô hình hồi quy Sử dụng phương <u>trình</u>

Câu 13. Trang 10: tiền xử lý dữ liêu Tiến hành làm sach dữ liêu xử lý các giá tri bi thiếu, ngoại lê,

Độ trùng lặp: 80%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Tiền xử lý dữ liêu</u> Thực <u>hành làm sach dữ liêu xử lý các giá tri bi thiếu,</u> xác định

<u>giá tri ngoai</u>

Câu 14. Trang 10: hồi quy tuyến tính giả định rằng mối quan hệ giữa biến phu thuộc và các biến độc lập là tuyến tính tức là giá tri của biến phụ thuộc có thể được biểu diễn dưới dang một tổ hợp tuyến tính của các biến độc lập với các hệ số tương ứng

Đô trùng lặp: 76%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: trị của biến phu thuộc dựa trên các giá trị của các biến độc lập Mô hình Hồi quy tuyến tính giả đinh rằng có một mối liên hệ tuyến tính giữa biến phu thuộc và các biến độc lập Lý thuyết <u>Hồi quy tuyến tính</u> Lý thuyết <u>Hồi quy tuyến tính</u> dựa trên <u>một số giả đinh</u> cơ bản <u>một</u> trong những giả định quan trọng nhất <u>là mối quan hệ giữa biến phu thuộc và các biến độc lập là tuyến</u> tính tức là biến phu thuộc có thể được diễn giải bằng cách nhân các hệ số của các biến độc lập với giá tri tương ứng

Câu 15. Trang 11: là hê số góc (intercept), biểu thị giá tri của biến phu thuộc (y) khi tất cả các biến độc lập bằng 0

Đô trùng lặp: 69%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>là giá tri của biến phu thuộc</u> Y <u>khi tất cả các biến độc lập</u>

Câu 16. Trang 11: là Sai số ngẫu nhiên (error Term) phản ánh sự khác biệt giữa giá tri thực tế và giá tri dư đoán bao gồm các vếu tố không được mô hình hóa

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giữa giá tri thực tế và giá tri dư đoán, của biến phụ thuộc sai số thể hiện những yếu tố không được mô hình hóa

Câu 17. Trang 11: Phương pháp hồi quy, tuyến tính cho phép phân tích và đánh giá tác đông đồng thời của nhiều biến độc lập đến biến phu thuộc giúp hiểu rõ hơn về mối quan hệ giữa các biến và xây dưng mô hình dư đoán chính xác

Độ trùng lặp: 57%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>đánh giá tác động của các biến độc lập</u> lên <u>xác</u> suất xảy ra <u>của</u> sự kiện <u>quan</u> tâm Kết quả phân tích sẽ cung cấp các hệ số hồi quy cho phép nhà nghiên cứu hiểu rõ hơn về mối <u>quan hệ giữa các biến và</u> đưa ra <u>các dư đoán chính xác hơn các g</u>iả định <u>của hồi quy</u> nhị <u>phân hồi</u> quy logistic không yêu cầu mối quan hệ tuyến tính giữa biến phụ thuộc, và các biến độc lập các

<u>biến độc lập</u> trong <u>mô hình</u>

Câu 18. Trang 11: hồi quy phi tuyến là một phương pháp thống kê được sử dụng để mô hình hóa mối quan hệ không tuyến tính, giữa biến phụ thuộc và một hoặc nhiều biến độc lập

Độ trùng lặp: 89%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>là một phương pháp thống kê</u> dùng <u>để mô hình hóa mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập</u> Trong <u>Hồi quy tuyến tính</u> giả định rằng <u>mối quan hệ giữa biến phụ thuộc và</u> các <u>biến độc lập là tuyến tính</u> tức <u>là</u> có thể diễn giải bằng <u>một</u> đường thẳng Tuy nhiên, trong <u>một</u> số trường hợp, <u>mối quan hệ giữa</u> các <u>biến</u> này có thể <u>không tuyến tính</u> Trong trường hợp đó, <u>Hồi quy phi tuyến tính được sử dụng để</u> xây dựng <u>mô hình</u> mà có thể <u>mô</u> tả <u>được mối quan hệ không tuyến</u>

Câu 19. Trang 13: 1 3 3 3 phân tích thành phần CHÍNH (PCA Principal COMPONENT ANALYSIS)

Độ trùng lặp: 77%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Component Analysis) 48 3 1 Sơ lược về Phân tích thành phần chính PCA

Câu 20. Trang 13: Phân tích Thành phần Chính (PCA) là một kỹ thuật giảm chiều dữ liệu được sử dụng rộng rãi trong học máy và thống kê

Độ trùng lặp: 68%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>được sử dụng trong</u> các lĩnh vực này 1 <u>Phân tích thành phần chính (PCA) Phân tích thành phần chính (PCA) là một kỹ thuật giảm chiều dữ liệu</u>

Câu 21. Trang 13: Phương pháp này <u>chuyển đổi các biến</u> gốc có <u>liên quan thành một tập hợp các thành phần chính không tương quan</u> giúp <u>Giữ lại phần</u> lớn <u>thông tin quan trọng của dữ liệu</u>

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: thành một tập hợp các biến mới không tương quan được gọi là các thành phần chính giữ lại thông tin quan trọng

Câu 22. Trang 14: PCA <u>là một công cụ mạnh mẽ đặc biệt hữu ích trong việc</u> giảm chiều <u>dữ liệu</u>, phân tích, mô hình <u>hóa</u> và <u>trưc quan hóa</u>

Độ trùng lặp: 53%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>là một công cu mạnh mẽ, trong việc trực quan hóa dữ liêu, đặc biệt hữu ích trong</u>

<u>viêc</u>

Câu 23. Trang 14: Python là một ngôn ngữ lập trình thông dịch, hướng đối tượng, và bậc cao, được phát triển bởi Guido van Rossum và ra mắt vào năm 1991

Độ trùng lặp: 75%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Python là một ngôn ngữ lập trình bậc cao, mã nguồn mở, hướng đối tượng được

phát triển bởi Guido van Rossum và lần đầu tiên ra mắt vào năm 1991

Câu 24. Trang 14: Python hiện <u>là một trong những ngôn ngữ lập trình phổ biến nhất trên thế giới,</u> đặc biệt <u>được</u> ưa chuộng <u>trong</u> các <u>lĩnh vực như</u>

Độ trùng lặp: 64%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Python là một trong những ngôn ngữ lập trình phổ biến nhất trên thế giới, được</u> sử

dụng rộng rãi <u>trong</u> nhiều <u>lĩnh vực như</u>

Câu 25. Trang 14: Python được sử dụng rộng rãi trong các lĩnh vực học máy. (Machine Learning) và trí tuê nhân tạo, (Al) nhờ vào hệ sinh thái phong phú các thư viên và công cụ mạnh mẽ như

Độ trùng lặp: 58%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Python được sử dụng rộng rãi trong các lĩnh vực như trí tuệ nhân tạo học máy</u> phân tích dữ liệu, <u>và phát triển web nhờ vào hệ sinh thái thư viên phong phú</u>

Câu 26. Trang 15: Đa năng o phù hợp với nhiều ứng dụng khác nhau, từ phát triển web, phân tích dữ liêu, đến các <u>lĩnh vực như học máy trí tuệ nhân tạo, và tư động hóa</u>

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>học máy, và tự động hóa</u> Tại sao nên <u>học</u> Python Dễ <u>học</u> Cú pháp Python rõ ràng <u>và</u> dễ hiểu, rất <u>Phù hợp</u> cho người mới bắt đầu <u>học</u> lập trình <u>Đa dụng</u> Python có thể được sử dụng trong nhiều lĩnh vực như phân tích dữ liệu, trí tuệ nhân tạo phát triển web, và nhiều ứng dụng khác

Câu 27. Trang 15: o <u>sự hỗ trợ mạnh mẽ từ các thư viện như NumPy, Pandas, Matplotlib</u> TensorFlow <u>và</u>

Độ trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nôi dung nguồn: Sư hỗ trơ mạnh mẽ từ các thư viên như Pandas, NumPy, Matplotlib, và

Câu 28. Trang 15: PyTorch, giúp người dùng thực hiện các tác vụ học máy, học sâu và phân tích dữ liêu một cách hiệu quả

Đô trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giúp người dùng thực hiện các tác vụ tự động và phân tích dữ liệu một cách hiệu

<u>quả</u>

Câu 29. Trang 15: khả năng Tích hợp cao o <u>Python dễ dàng Tích hợp với các ngôn ngữ và công cụ khác</u> như C++, <u>Java, và</u> R

Độ trùng lặp: 60%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: với các ngôn ngữ và công

Câu 30. Trang 16: o Python sử dụng cơ chế quản lý bộ nhớ tự động điều này có thể ảnh hưởng

đến hiệu suất

Độ trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Python sử dung cơ chế quản lý bô nhớ tư đông, (garbage collection), điều này có

thể gây ra vấn đề về <u>hiệu suất</u>

Câu 31. Trang 17: xử lý dữ liệu thô o Kiểm tra và Loại bỏ giá trị bị thiếu hoặc không hợp lệ trong

dữ liêu (ví dụ các thời điểm

Độ trùng lặp: 56%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>loại bỏ giá trị thiếu Kiểm tra và loại bỏ các giá trị thiếu trong dữ liệu</u> có thể thay thế bằng <u>giá trị</u> trung bình <u>hoặc giá trị</u> gần nhất <u>loại bỏ</u> ngoại <u>lệ</u> Xác định <u>và loại bỏ các giá trị</u> ngoại <u>lệ không hợp lệ hoặc không</u>

Câu 32. Trang 17: o chuẩn hóa dữ liệu để đảm bảo các biến đầu vào có cùng quy mô và giảm độ chênh lệch giữa các đặc tính (normalization hoặc standardization)

Độ trùng lặp: 53%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: để đảm bảo các biến đầu vào có cùng

Câu 33. Trang 18: dữ liệu huấn luyện và kiểm tra o chia dữ liệu thành tập huấn luyện (training set) và kiểm tra (test set) Ví du 80% dữ liêu để huấn luyên và 20% để kiểm tra

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>huấn luyện (training set) và</u> số mẫu được chọn cho <u>tập kiểm tra (test set) ví dụ</u> Percentage split = % nghĩa là % mẫu dùng <u>tập huấn luyện và</u> % mẫu dùng <u>kiểm tra</u> Thực nghiệm <u>Dữ liệu</u> với J trên Weka <u>Dữ liệu</u> thực nghiệm <u>Chia thành</u> phần, <u>huấn luyện và kiểm</u> thử Tiến hành dùng <u>Dữ liêu huấn luyên để</u>

Câu 34. Trang 18: <u>chỉ số đánh giá mô hình</u> o MAE, <u>MSE</u> RMSE <u>đo lường sai số giữa giá tri dự</u> đoán và thực tế.

Độ trùng lặp: 65%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giữa giá tri dự đoán và giá tri thực tế sau khi bình phương <u>sai số và</u> tính toán căn bậc hai của giá trị này <u>MSE</u>, <u>Chỉ</u> một trong những <u>Chỉ số</u> được sử dụng phổ biến để <u>đánh giá</u> mức độ chính sát của <u>mô hình dự đoán MSE</u>,

Câu 35. Trang 18: o sử dụng cross validation (kiểm định chéo) để kiểm tra khả năng tổng quát hóa của mô hình

Độ trùng lặp: 68%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Sử dụng cross validation để</u> xác <u>định</u> kích thước tập <u>kiểm tra</u> tối ưu, nhằm nâng cao <u>khả năng tổng quát hóa của mô</u> hình

Câu 36. Trang 19: Dự báo lưu lượng nước xả của hồ Tua Srah có thể được xây dưng bằng cách

sử dung các phương pháp học sâu

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: có thể được xây dựng bằng cách sử dụng các phương pháp như hồi

Câu 37. Trang 19: Trong bước này, cần xử lý các vấn đề như o dữ liệu thiếu Có thể xử lý dữ liệu thiếu bằng cách điền vào các giá trị thay thế, ví dụ như giá trị trung bình, giá trị trung vị hoặc sử dung các mô hình dư đoán để ước tính giá trị thiếu

Độ trùng lặp: 58%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>Trong bước này, cần</u> xác định <u>các giá trị thiếu</u> và quyết định <u>cách xử lý</u> chúng <u>Có</u> <u>thể sử dụng các</u> phương pháp <u>như điền giá trị trung bình, giá trị trung vi, hoặc sử dụng mô hình dự đoán để điền các giá trị thiếu 3 <u>xử lý Dữ liêu</u></u>

Câu 38. Trang 20: huấn luyện và đánh giá mô hình o huấn luyện dữ liệu huấn luyện được chia thành các tập chuỗi thời gian, từ đó điều chỉnh các tham số của mô hình sao cho tối ưu

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Huấn luyện và điều chỉnh các</u> siêu <u>tham số</u> Cuối cùng, <u>tập</u> kiểm tra (10%) <u>được</u> sử dụng để <u>đánh giá</u> hiệu suất cuối cùng <u>của mô hình</u> trên <u>Dữ liệu</u> chưa từng thấy <u>hình</u> 2 Giai đoạn xử lý <u>tập Dữ liệu hình</u> ảnh nấm bằng Roboflow Annotate <u>tập Huấn luyện được</u> sử dụng để phù hợp với <u>mô hình tập</u> xác thực <u>được</u> sử dụng để <u>điều chỉnh các tham số của mô hình và đánh giá</u>

Câu 39. Trang 20: o <u>Đánh giá hiệu suất của mô hình được Đánh giá</u> trên <u>dữ liệu kiểm tra</u> thông qua <u>các chỉ số như</u> MAE, RMSE hoặc R², nhằm <u>đảm bảo tính</u> khả thi <u>và đô tin cây của mô hình</u>

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>liệu kiểm tra và dữ liệu</u> xác thực để <u>đảm bảo tính</u> tổng quát <u>và độ tin cậy của mô</u> <u>hình</u> 4 <u>Đánh giá mô hình</u> Để <u>Đánh giá Hiệu suất của mô hình các chỉ số như độ</u> chính xác, precision, recall <u>và</u> F1 score <u>được</u>

Câu 40. Trang 21: trong mạng nơ ron (Neural Network), một loạt các thuật toán được sử dụng để xác định và nhận diện các mối quan hệ trong các tập dữ liệu

Độ trùng lặp: 82%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: một loạt các thuật toán được sử dụng để xác định và nhận ra các mối quan hệ Trong các tập dữ liệu

Câu 41. Trang 21: Mạng Nơ ron <u>có khả năng thích ứng linh hoạt với mọi thay đổi</u> ngay <u>từ đầu vào,</u> vì vậy <u>không cần phải thiết kế lại</u> các <u>tiêu chí đầu ra mà</u> vẫn <u>có</u> thể đưa <u>ra kết quả tối ưu</u> 2 3 2 Kiến trúc <u>Mang Noron</u>

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nôi dung nguồn: có khả năng thích ứng linh hoat với moi thay đổi từ đầu vào, giúp dư đoán kết quả

tối ưu mà không cần phải thiết kế lai tiêu chí đầu ra

Câu 42. Trang 22: Một mạng nơ ron cơ bản bao gồm ba lớp nơ ron nhân tạo, được liên <u>kết với</u> nhau <u>ron</u> qua <u>lớp đầu vào.</u>

Độ trùng lặp: 63%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: nơ ron nhân tạo được kết nối <u>với</u> các hệ số (trọng số) tạo thành kết cấu nơ ron

Một mạng nơ ron cơ bản bao gồm Một lớp đầu vào

Câu 43. Trang 22: Các nút đầu vào sẽ xử lý dữ liệu, phân tích hoặc phân loại và sau đó chuyển dữ liêu, sang lớp tiếp theo

Độ trùng lặp: 91%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Các nút đầu vào xử lý dữ liệu, phân tích hoặc phân loại và sau đó chuyển dữ liêu,

sang lớp tiếp theo

Câu 44. Trang 22: Mỗi lớp ẩn sẽ phân tích dữ liệu đầu ra từ lớp trước, xử lý thêm và truyền dữ liệu

sang lớp tiếp theo

Độ trùng lặp: 75%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Mỗi lớp ẩn phân tích dữ liệu đầu ra từ lớp trước, xử lý dữ liệu đó sâu hơn và

chuyển <u>dữ liệu sang lớp tiếp theo</u>

Câu 45. Trang 22: Lớp đầu ra (Output Layer) Lớp này cung cấp kết quả cuối cùng sau khi tất cả

dữ liêu đã được xử lý bởi mang nơ ron

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>ra kết quả cuối cùng của tất cả dữ liêu được xử lý bởi mang nơ ron</u>

Câu 46. Trang 22: Ví dụ trong bài toán phân loại nhị phân (có/không), lớp đầu ra sẽ chỉ có một nút

với <u>kết quả</u> là <u>1 hoặc 0</u>

Độ trùng lặp: 60%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: phân loại nhi phân (có/không), lớp đầu ra sẽ có một nút, đầu ra nút,

Câu 47. Trang 22: Nhận dạng chữ viết tay Mạng nơ ron nhân tạo (Neural Network) được sử dụng để chuyển đổi các ký tự viết tay thành các ký tự kỹ thuật số giúp máy tính dễ dàng Nhận dạng và

xử lý <u>các ký tự</u> này

Độ trùng lặp: 68%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Nhận dạng chữ viết tay Mạng nơ ron nhân tạo được sử dụng để chuyển đổi các

ký tự viết tay thành các ký tự kỹ thuật số. mà máy

Câu 48. Trang 23: nén hình ảnh Neural Network có thể sử dụng để nén hình ảnh bằng cách lưu trữ, mã hoá và tái tạo các hình ảnh một cách hiệu quả

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>hình ảnh Neural Network</u> được <u>sử dụng</u> nhiều trong việc <u>lưu trữ, mã</u> hóa <u>và tạo</u>

Nén hình ảnh Con người <u>có thể tái tạo và</u> tối ưu kích thước dữ liệu <u>bằng cách</u>

Câu 49. Trang 23: Ngoài các ứng dụng trên Neural network còn được sử dụng rộng rãi trong nhiều lĩnh vực và công nghệ khác như thị giác máy tính, trò chơi điện tử, dịch tự động lọc Mạng xã hội, nhân dạng giong nói, và chẩn đoán y tế

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: được sử dụng trên, nhiều công nghệ và ứng dụng khác nhau như trò chơi điện tử,

thị giác máy tính, nhận dạng giọng nói lọc mạng xã hội, dịch tư động, và chẩn đoán y tế

Câu 50. Trang 23: mang LSTM (Long short term memory) được thiết kế đặc biệt để xử lý và học từ

Độ trùng lặp: 77%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: LSTM (Long Short Term Memory) là một loại Mạng nơ ron hồi quy sâu, được thiết

kế đặc biệt để xử lý và

Câu 51. Trang 25: <u>kết quả cuối cùng của Random Forest được đưa ra bằng cách cây quyết đinh</u> học <u>Một</u> phần <u>dữ liệu</u> khác <u>nhau và kết quả cuối cùng được</u> tổng <u>hợp để đưa ra dự đoán</u> chính xác hơn

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đưa ra dự đoán cuối cùng một</u> ví dụ về Bagging trong <u>cây quyết đinh</u> là <u>Random Forest Random Forest</u> là <u>một</u> mô hình <u>dự đoán được</u> xây dựng <u>bằng cách</u> sử dụng nhiều <u>cây quyết</u> <u>định</u> độc lập với <u>nhau. và Kết hợp dự đoán của</u> chúng <u>để đưa ra Kết quả cuối cùng</u> Các <u>cây quyết</u> <u>đinh</u> trong <u>Random Forest được</u> huấn luyện với các tập <u>dữ liệu</u>

Câu 52. Trang 25: Random Forest <u>là một</u> mô hình <u>học máy mạnh mẽ, sử dụng</u> sự <u>kết hợp của nhiều cây quyết đinh</u> (Decision Trees) <u>để cải thiên đô chính xác của các dư đoán</u>

Độ trùng lặp: 58%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: học máy kết hợp nhiều cây quyết đinh để cải thiên đô chính xác

Câu 53. Trang 26: tạo các tập dữ liệu con (Bootstrap Sampling) o từ tập dữ liệu huấn luyện ban đầu Random Forest áp dụng phương pháp Bootstrap

Độ trùng lặp: 50%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>Từ tập dữ liệu huấn luyện ban đầu,</u> để <u>Tạo</u> ra <u>các tập dữ liệu con</u> (

Câu 54. Trang 26: Sampling để tạo ra các tập dữ liệu con bằng cách <u>lấy mẫu ngẫu nhiên với sự</u> thay thế từ dữ liêu gốc

thay the ta da lieu god

Độ trùng lặp: 59%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: các tập dữ liêu con, được lấy mẫu ngẫu nhiên với sư thay thế từ

Câu 55. Trang 26: (Bootstrap Aggregating), trong đó nhiều cây quyết định được xây dựng trên các

mẫu <u>dữ liệu con</u> khác nhau

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trong đó các cây quyết định được xây dựng trên các tập dữ liêu con

Câu 56. Trang 26: xây dựng cây quyết định o Mỗi cây được xây dựng trên tập dữ liệu con với một tâp hợp các đặc trưng, chọn lọc

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Xây dựng cây quyết định trên tập dữ liệu được gọi là học bằng cây quyết định hay đơn giản chỉ là cây quyết định Mỗi nút bên trong của cây tương ứng <u>với một</u> 19 <u>đặc trưng các</u> nút lá đại diện cho <u>các</u> phân loại và <u>các</u> cành đại diện cho <u>các</u> kết <u>hợp</u> của <u>các đặc trưng</u>

Câu 57. Trang 27: các tham số của mô hình LSTM được khởi tạo, bao gồm được sử dụng làm đặc trưng đại diện cho dữ liệu <u>đầu vào</u>

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đại diện cho mô hình</u> ResNet 18 <u>bao gồm Các tham số của mô hình</u> mạng nơ ron tích chập, <u>được sử dụng</u> để biến đối ảnh <u>đầu vào của</u> đối tượng 0 () thành vector <u>đặc trưng</u>

Câu 58. Trang 30: dự đoán cuối cùng Random Forest sử dụng các đặc trưng, được trích xuất từ LSTM để đưa ra dư đoán cuối cùng

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>trích xuất đặc trưng</u> sau đó <u>các đặc trưng</u> này <u>được</u> tổng hợp qua <u>các</u> lớp dày <u>đặc</u> để đưa ra Dư đoán cuối cùng

Câu 59. Trang 30: mô hình này kết hợp các dự đoán từ nhiều cây quyết định, giúp nâng cao độ chính xác và giảm thiểu hiện tương overfitting

Độ trùng lặp: 55%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: <u>này giảm thiểu hiện tượng</u> quá khớp bằng cách <u>kết hợp</u> trung bình <u>các dự đoán từ</u> <u>nhiều cây quyết đinh, giúp Mô hình</u>

Câu 60. Trang 30: <u>xử lý</u> hiệu quả <u>dữ liệu chuỗi thời gian LSTM là một loại mạng nơ ron hồ</u>i tiếp (RNN) <u>được thiết kế đặc biệt để xử lý và</u> học từ <u>dữ liệu chuỗi thời gian</u> nhờ <u>Khả năng ghi nhớ thông</u> tin dài han và ngắn han

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>là một loại mạng nơ ron hồi</u> quy <u>và được thiết kế đặc biệt để Xử lý dữ liệu chuỗi</u>

thời gian khả năng ghi nhớ thông tin dài han của LSTM làm

Câu 61. Trang 30: trích xuất đặc trưng mạnh mẽ LSTM tự động học và trích xuất các đặc trưng quan trong từ dữ liêu, chuỗi thời gian thông qua các hidden states

Độ trùng lặp: 52%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: và Trích xuất các đặc trưng quan trong từ dữ liệu

Câu 62. Trang 33: bước 2 kiểm tra và xử lý dữ liệu thiếu, bị thiếu, này được kiểm tra và điền vào bằng giá tri trung bình, của các mẫu trong ngày tương ứng

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>xử lý dữ liệu thiếu</u> Trước khi làm sạch <u>dữ liệu</u> ta cần <u>Kiểm tra và xử lý các giá tri</u> thiếu Có thể sử dụng <u>các</u> phương pháp như <u>điền giá tri trung bình giá tri trung</u>

Câu 63. Trang 35: 3 3 1 Sai số trung bình tuyệt đối (Mean Absolute Error MAE) giữa giá tri dự đoán và giá tri thực tế

Độ trùng lặp: 62%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trung bình tuyệt đối (Mean Absolute Percentage Error) 3 Căn bậc hai Sai số

trung bình

Câu 64. Trang 35: Định nghĩa MAE (Mean Absolute Error) là giá trị, trung bình của các sai số tuyết đối

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: số trung bình của các sai số

Câu 65. Trang 35: Ý nghĩa <u>MAE cho</u> biết <u>sai số trung bình tuyệt đối</u> giữa <u>giá trị dự đoán và giá trị thực tế của</u> lưu lượng xả, giúp <u>đánh giá độ chính xác của mô hình dư</u> báo

Độ trùng lặp: 58%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đánh giá độ chính xác</u> khác <u>của mô hình</u> gồm <u>sai số</u> căn quân phương RMSE <u>của</u> toàn tập dữ liệu kiểm chứng <u>cho giá trị dự đoán và giá trị thực tế</u> thấp hơn <u>giá trị</u> RMSE <u>của</u> tập huấn luyện <u>và</u> đạt 75 3522 W <u>sai số trung bình tuyệt đối MAE</u>

Câu 66. Trang 35: Định nghĩa MSE (Mean Squared Error) là giá trị trung bình của bình phương sai số giữa giá trị dự đoán và giá trị thực tế

Độ trùng lặp: 77%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trung bình của bình phương sai số giữa giá tri dư đoán và giá tri thực tế

Câu 67. Trang 36: 3 3 Căn bâc hai của sai số bình phương trung bình (Root Mean Squared Error

Đô trùng lặp: 87%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Căn bậc hai của sai số bình phương trung bình (Root Mean Squared</u> Error) 1 <u>sai số bình phương trung bình (Mean Squared</u> Error) <u>3</u>

Câu 68. Trang 38: Visual Studio Code (VS Code) là Một trình soạn thảo mã nguồn mở mạnh mẽ, và phổ biến hỗ trợ nhiều ngôn ngữ lập trình bao gồm cả Python

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Visual Studio Code (VS Code) một trình soạn thảo mã nguồn mở</u> do Microsoft phát triển, <u>hỗ trợ nhiều ngôn ngữ lập trình và</u> có <u>nhiều</u> tiện ích <u>mở</u> rộng Sublime Text <u>một trình soạn thảo mã nguồn</u>

Câu 69. Trang 38: Python nổi bật với cú pháp đơn giản, dễ học và đặc biệt <u>là hệ sinh thái thư viện phong phú</u> dành <u>cho</u> Machine

Độ trùng lặp: 55%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>cú pháp đơn giản, dễ</u> hiểu <u>và</u> một <u>hệ sinh thái thư viện phong phú</u> Dưới đây <u>là</u> những điểm <u>nổi bật</u> về cách <u>Python</u> được sử dụng trong <u>học</u>, máy <u>thư viện học</u>, máy mạnh mẽ <u>Python</u> cung cấp nhiều <u>thư viên</u> chuyên dụng <u>cho</u>

Câu 70. Trang 38: Pandas là một thư viện mạnh mẽ để xử lý và phân tích dữ liệu dạng bảng, cung cấp cấu trúc dữ liệu DataFrame, linh hoạt cho phép thao tác biến đổi và phân tích dữ liệu một cách hiệu quả

Độ trùng lặp: 62%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Pandas là một thư viện mạnh mẽ và linh hoạt, cho xử lý và phân tích dữ liệu</u> trong Python Với các <u>cấu trúc dữ liệu</u> như Series <u>và DataFrame</u> cùng nhiều chức năng hữu ích, <u>Pandas</u> giúp <u>cho</u> việc <u>thao tác, và phân tích dữ liệu</u>

Câu 71. Trang 38: Numpy cung cấp các đối tượng mảng đa chiều (ndarray) hiệu suất cao và các hàm toán học để thao tác với các mảng này

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>cung cấp các mảng đa chiều hiệu suất cao và các hàm toán học</u> cho phép bạn thực hiện <u>các</u> phép tính khoa <u>học với</u> tốc độ nhanh hơn nhiều so <u>với các mảng</u> Python thông thường Nguồn <u>NumPy</u>

Câu 72. Trang 39: Pandas và NumPy được sử dụng trong các bước tiền xử lý, xử lý, và phân tích dữ liêu,

Đô trùng lặp: 64%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>được sử dụng trong phân tích dữ liệu trong</u> đó <u>dữ liệu được</u> chuyển qua <u>các bước</u> <u>tiền xử lý.</u> khám phá, biến đổi <u>và</u>

Câu 73. Trang 39: Quá trình thực nghiệm bao gồm các bước tiền xử lý dữ liệu chuẩn bị dữ liệu cho LSTM, xây dưng và huấn luyên mô hình LSTM, Trích xuất đặc trưng, huấn luyên mô hình Random

Forest, và đánh giá hiệu suất của mô hình

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đặc trưng, huấn luyện mô hình</u> học máy, kiểm tra <u>mô hình</u> với <u>các</u> mẫu <u>dữ liệu,</u> mới, <u>và đánh giá</u> kết quả Thu thập <u>dữ liệu, thực</u> hiện <u>các bước tiền xử lý dữ liệu, gán nhãn cho các</u> mẫu <u>dữ liệu, trích</u> chọn <u>đặc trưng, trích xuất đặc trưng, và</u> lựa chọn <u>đặc trưng, phù hợp cho</u> bài toán Sử dụng <u>các</u> thuật toán học máy để <u>xây dưng</u>

Câu 74. Trang 39: MinMaxScaler giúp tăng tốc đô hôi tu và cải thiên hiệu suất của mô hình

Độ trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: tăng tốc đô hôi tu và cải thiên hiệu suất mô hình

Câu 75. Trang 40: <u>sử dụng mô hình LSTM đã huấn luyện</u> Dùng <u>mô hình LSTM đã huấn luyện để dư đoán trên tâp huấn luyên và tâp kiểm tra</u>

Độ trùng lặp: 58%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Sử dụng mô hình đã huấn luyện để dự đoán</u> mua hàng <u>trên tập</u> dữ liệu <u>kiểm tra</u>

<u>và</u>

Câu 76. Trang 41: Dự đoán Sử dụng mô hình Random Forest <u>đã huấn luyện để Dự đoán</u> lưu_lượng_xả trên tập <u>kiểm tra</u>

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Dư đoán Sử dung mô hình đã huấn luyên để Dư đoán</u> nhãn cho bộ <u>kiểm tra</u>

Câu 77. Trang 41: R2 (R squared) Hệ số xác định Đánh giá mức độ phù hợp của mô hình với dữ

<u>liêu</u>

Độ trùng lặp: 81%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: mức đô phù hợp của mô hình với dữ liêu

Câu 78. Trang 41: MSE (Mean Squared Error) sai số bình phương trung bình đo lường trung bình bình phương của sai số giữa giá trị dự đoán và giá trị thực tế

Đô trùng lặp: 73%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trung bình bình phương của Sai số giữa giá tri dư đoán và giá tri thực tế Mean

Absolute Error (MAE) MAE cũng đo lường

Câu 79. Trang 41: RMSE (root mean squared error) Căn bậc hai của sai số bình phương trung bình có cùng đơn vi đo với biến mục tiêu

Đô trùng lặp: 71%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Căn bậc hai của sai số bình phương trung bình (Root Mean Squared Error)</u> nó <u>cùng đơn vi đo với</u> các quan sát, nên <u>có</u>

Câu 80. Trang 41: MAE (Mean Absolute Error) sai số tuyệt đối trung bình đo lường trung bình của giá tri tuyệt đối của sai số giữa giá tri dư đoán và giá tri thực tế

Độ trùng lặp: 84%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: đo lường trung bình, của giá trị tuyệt đối của Sai số giữa giá trị dự đoán và giá trị

<u>thực tế</u>

Câu 81. Trang 41: Vẽ biểu đô so sánh giá trị lưu_lượng_xả thực tế và giá trị dự đoán từ mô Hình

Random

Độ trùng lặp: 63%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Vẽ biểu đồ so sánh giá trị dự đoán và giá trị thực tế</u> (sử dụng 200 <u>giá trị</u> đầu) với đường cam là đường <u>biểu</u> diễn <u>giá trị thực tế và</u> đường xanh là đường <u>biểu</u> diễn <u>giá trị dự đoán</u> được <u>hình</u> 3 28 <u>biểu đồ so sánh qiá trị dự đoán và thực tế</u> khi áp dụng <u>mô hình</u>

Câu 82. Trang 46: <u>ứng dụng được</u> xây dựng bằng <u>Python một ngôn ngữ Lập trình mạnh mẽ, phổ biến và dễ sử dụng trong các lĩnh vực khoa học, dữ liêu và học, máy</u>

Độ trùng lặp: 63%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>dữ liệu và biến trong Python</u>, Bài 5 Toán tử <u>và</u> điều kiện <u>trong Python</u>, Bài 6 Hàm <u>trong Python</u>, Bài 7 Xử lý chuỗi <u>trong Python</u>, Bài 8 <u>lập trình</u> hướng đối tượng <u>trong Python</u>, Bài 9 Modules <u>và</u> Packages <u>trong Python</u>, Bài 10 Xử lý lỗi <u>và</u> ngoại lệ <u>trong Python</u>, Bài tập <u>Python</u>, cơ bản <u>Python</u>, cơ bản <u>Python</u>, là <u>một ngôn ngữ lập trình mạnh mẽ</u>, phổ biến và dễ học được sử dụng rộng rãi <u>trong</u> nhiều <u>lĩnh vưc</u> như phát triển web, <u>khoa học dữ liêu học máy</u>

Câu 83. Trang 46: Python cung cấp nhiều thư viện hỗ trợ mạnh mẽ cho việc Xử lý dữ liệu và phát triển các mô hình học máy giúp tối ưu hóa, quá trình triển khai và phân tích

Đô trùng lặp: 62%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>và thư viện hỗ trợ mạnh mẽ cho việc phát triển các mô hình học máy.</u> Nó bao gồm <u>các</u> API <u>cho việc</u> xây dựng <u>và</u> huấn luyện <u>các</u> mạng neural, <u>các</u> thuật toán <u>tối ưu hóa</u> quản <u>lý</u> <u>dữ liệu và</u> công cụ đánh giá hiệu suất <u>xử lý dữ liệu</u> lớn TensorFlow <u>hỗ trợ xử lý dữ liệu</u>

Câu 84. Trang 46: Python cung cấp nhiều thư viện hỗ trợ cho việc phát triển các ứng dụng khoa học và kỹ thuật bao gồm

Độ trùng lặp: 66%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>cung cấp nhiều thư viện và</u> công cụ để <u>phát triển các ứng dụng</u> máy tính, <u>bao</u> gồm các ứng dụng desktop <u>và các ứng dụng</u> server side <u>phát triển các</u> hệ thống nhúng Java được sử <u>dụng</u> để <u>phát triển các</u> hệ thống nhúng, <u>bao gồm các</u> thiết bị điện tử, <u>các</u> bộ điều khiển <u>và các</u> thiết bị khác <u>phát triển các</u> hệ thống <u>khoa học và kỹ thuật</u>, Java được sử <u>dụng</u> để <u>phát triển các</u> <u>ứng dụng và</u> công cụ trong cáclĩnh vực như <u>khoa học và kỹ thuật</u>, <u>bao gồm</u>

Câu 85. Trang 46: o NumPy thư viện cơ bản cho tính toán khoa học trong Python cung cấp các công cụ mạnh mẽ để làm việc với mảng đa chiều, (ndarray) và các hàm toán học hiệu suất cao

Độ trùng lặp: 72%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: thư viện cơ bản cho tính toán khoa học và xử lý mảng (arrays) trong Python. Nó cung cấp các công cu mạnh mẽ để làm việc với mảng đa chiều và

Câu 86. Trang 46: Nó hỗ trợ các phép toán số học đại số tuyến tính, thống kê và xử lý ma trận giúp tối <u>ưu</u> hóa <u>hiệu suất và</u> khả năng <u>tính</u>, toán <u>trong các ứng dung khoa học dữ liêu và học máy</u>

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>đại số tuyến tính, thống kê, và phép</u> biến đổi Fourier NumPy được sử <u>dụng</u> rộng rãi <u>trong các ứng dụng khoa học, dữ liệu và tính,</u> tốn <u>số</u> như phân tích <u>dữ liệu xử lý</u> ảnh <u>và</u> âm thanh, <u>và học, máy các ưu</u> điểm của NumPy bao gồm <u>hiệu suất tính,</u> tốn cao, <u>hỗ trợ</u> nhiều <u>phép</u> <u>tính,</u> tốn trên <u>ma trân,</u>

Câu 87. Trang 46: o <u>Pandas thư viện mạnh mẽ cho việc xử lý và phân tích dữ liệu</u> dạng bảng, <u>cho phép thao tác biến đổi và</u> trích <u>xuất</u> thông tin <u>từ dữ liệu dễ dàng</u>

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Pandas cho phép xử lý dữ liệu dễ dàng và hiệu quả hơn, từ việc nhập và xuất dữ liệu xử lý dữ liệu đến việc thực hiện các thao tác, phân tích và trực quan hóa dữ liệu Các Ứng Dụng Chính Của Pandas xử lý dữ liệu Pandas giúp làm sạch, biến đổi và thao tác, với dữ liệu một cách dễ dàng và hiệu quả phân tích dữ liệu Pandas cung cấp các công cụ mạnh mẽ để phân tích dữ liêu từ

Câu 88. Trang 46: o <u>Scikit learn thư viện mạnh mẽ</u> trong Python dành <u>cho học</u> máy, <u>cung cấp các thuật toán phân loại, hồi quy, phân cụm, và</u> nhiều phương pháp <u>tiền xử lý dữ liệu.</u>

Đô trùng lặp: 58%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: tiền xử lý dữ liệu trích xuất đặc trưng, và các thuật toán phân loại, hồi quy, phân

<u>cụm, và học</u>

Câu 89. Trang 49: <u>phương pháp có thể sử dụng kỹ thuậ</u>t crawling (thu thập dữ liệu tự động) <u>đ</u>ể mô phỏng <u>dữ liêu</u> từ <u>các trang web trưc tuyến</u>

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nôi dung nguồn: thập dữ liệu trưc tuyến Người quản tri trạng web có thể sử dung

Câu 90. Trang 50: o Quá trình Huấn luyện bao gồm các bước Tiền xử lý dữ liệu Chuẩn bị dữ liệu cho LSTM, Huấn luyện mô hình LSTM, trích xuất đặc trưng, và Huấn luyện mô hình Random Forest

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>tiền xử lý dữ liệu, chuẩn bi dữ liêu,</u> để loại bỏ nhiễu <u>và</u> gắn nhãn cần thiết <u>huấn</u>

<u>luyện mô hình</u>

Câu 91. Trang 50: Tiền xử lý dữ liệu dữ liệu đầu vào, được Tiền xử lý (ví dụ chuẩn hóa) trước khi

<u>đưa vào, mô hình</u> Độ trùng lặp: 74%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Tiền xử lý dữ liệu dữ liêu đầu vào thường là hình ảnh, được chuẩn hóa hoặc

chuẩn bị trước khi đưa vào mô hình

Câu 92. Trang 52: cải thiên đô chính xác o Tối ưu hóa tham số Tiếp tục nghiên cứu và Tối ưu

hóa các tham số của mô hình LSTM

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Cải thiện độ chính xác của ước lượng hướng Tối ưu hóa tham số của thuật toán

Sensor Fusion Một <u>số</u> thuật toán Sensor Fusion cho phép <u>Tối ưu hóa các tham s</u>ố

Câu 93. Trang 52: o Áp dụng kỹ thuật song song khai thác các kỹ thuật tính toán song song để rút

ngắn thời gian huấn luyên mô hình

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: tính toán song song giúp rút ngắn thời gian huấn luyên các mô hình

Câu 94. Trang 54: Áp dung các kiến thức lý thuyết đã học để giải quyết vấn đề thực tiễn trong

ngành thủy điện

Độ trùng lặp: 63%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>dung các kiến thức lý thuyết đã học để giải quyết các vấn đề thực</u>

Câu 95. Trang 55: phát triển ứng dung o Hoàn thiên ứng dung cải thiên giao diên người dùng và

bổ sung thêm các tính năng hữu

Độ trùng lặp: 53%

Nguồn: Dữ liệu nội sinh

Nôi dung nguồn: Phát triển các tính năng mới nhằm nâng cao trải nghiêm người dùng. Đề xuất các

cách để cải thiên giao diên người dùng và

Câu 96. Trang 55: o Nghiên cứu và triển khai mô hình trên nền tảng web hoặc di động, giúp người

dùng có thể dễ dàng truy cập và sử dụng ứng dụng một cách thuận tiên và hiệu quả

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nôi dung nguồn: có thể dễ dàng truy cập và sử dung ứng dung một cách thuận tiên TIẾP TUC MỞ

RỘNG tiên ÍCH SỐ Với mini app trên nền tảng Zalo, yêu cầu người dùng

Câu 97. Trang 56: "Áp dụng mô hình MIKE SHE kết hợp sử dụng sản phẩm mưa dự báoIFS dự báo lưu lượng đến hồ lưu vực sông Trà Khúc sông Vệ " Retrieved from https://www.researchgate

net/publication/330382123

Độ trùng lặp: 73%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: SHE kết hợp sử dụng sản phẩm mưa dự báo IFS dự báo lưu lượng đến hồ lưu

<u>vực</u>

--- Hết ---