



BÁO CÁO KIỂM TRA TRÙNG LẬP

Thông tin tài liệu

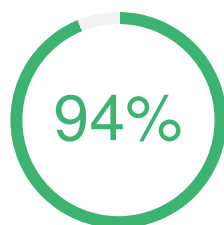
Tên tài liệu:	DATN_DoHoaiNam_2151160531 check
Tác giả:	Trần Mạnh Tuấn - HTTT
Điểm trùng lặp:	6
Thời gian tải lên:	11:11 06/07/2025
Thời gian sinh báo cáo:	11:12 06/07/2025
Các trang kiểm tra:	42/42 trang



Kết quả kiểm tra trùng lặp



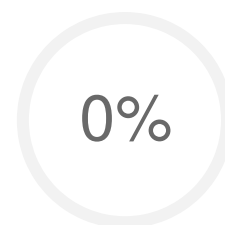
Có 6% nội dung trùng
lặp



Có 94% 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

luanvan.moet.gov.vn 123docz.net tailieu.vn

Danh sách các câu trùng lặp

Câu 1. Trang 1: phương pháp học máy truyền thống thường yêu cầu lượng lớn dữ liệu gán nhãn, trong khi việc gán nhãn thủ công lại tốn kém thời gian và công sức

Độ trùng lặp: 57%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: tốn kém thời gian và chi phí Ngoài ra, các Phương pháp học máy

Câu 2. Trang 3: các nguyên nhân gây sụt lở đất, bao gồm các yếu tố tự nhiên như địa chất, khí tượng, thủy văn, động đất, núi lửa, cũng như các hoạt động của con người

Độ trùng lặp: 68%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: yếu tố tự nhiên như địa chất, khí tượng, thủy văn, động đất núi lửa, và Các yếu tố hoạt động của con người Tại Việt Nam, hiện tượng, sụt lở đất

Câu 3. Trang 3: Việc phát hiện sớm các khu vực có nguy cơ sụt lở, là vô cùng cần thiết để kịp thời sơ tán, bảo vệ tính mạng và tài sản của người dân

Độ trùng lặp: 55%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: kịp thời phát hiện và xử lý những khu vực có nguy cơ sụt lở những đoạn đê xung yếu để bảo vệ tính mạng và tài sản của người dân /

Câu 4. Trang 5: Trên thế giới Các nghiên cứu chủ yếu tập trung vào Các nhóm phương pháp sau

Độ trùng lặp: 75%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Trên thế giới, các nghiên cứu chủ yếu tập trung vào các biến chứng sớm sau

Câu 5. Trang 5: phương pháp học máy, có giám sát (Supervised learning) Áp dụng các mô hình như SVM Random Forest, Decision Tree với đặc trưng được trích chọn thủ công (hand crafted features)

Độ trùng lặp: 50%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: với đặc trưng được trích xuất thủ công (hand crafted) và dựa trên học sâu (deep learning based) Đối với các Phương pháp học máy ý tưởng chủ đạo là sử dụng các đặc trưng được trích xuất thủ công đưa vào các mô hình dự đoán, phổ biến có thể kể đến SVM, [8, 42], rừng quyết định (Random Forest) [

Câu 6. Trang 5: Phương pháp bán giám sát và học chuyển giao kết hợp dữ liệu có nhãn và không nhãn giúp giảm chi phí gán nhãn và tăng tính tổng quát

Độ trùng lặp: 54%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: học có giám sát và học không giám sát trong đó một phần dữ liệu có nhãn và một phần không có nhãn Phương pháp này giúp giảm chi phí gán nhãn dữ liệu học tăng

Câu 7. Trang 6: Python là một ngôn ngữ lập trình mã nguồn mở, được sáng lập bởi Guido van Rossum vào năm 1991, nổi tiếng Với tính đơn giản, dễ đọc và hiệu quả

Độ trùng lặp: **61%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Python là một ngôn ngữ lập trình mã nguồn mở được phát triển bởi Guido van Rossum và ra mắt lần đầu vào những năm 1990 với cú pháp đơn giản, dễ hiểu và trực quan, Python nhanh chóng trở thành một trong những ngôn ngữ lập trình phổ biến nhất trên thế giới Python hỗ trợ nhiều mô hình lập trình bao gồm lập trình hướng đối tượng, lập trình hàm và lập trình cấu trúc, giúp người lập trình viết mã một cách linh hoạt và hiệu quả

Câu 8. Trang 6: Python được sử dụng rộng rãi trong nhiều lĩnh vực khác nhau nhờ vào tính linh hoạt và đa dạng của nó

Độ trùng lặp: **80%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: được sử dụng rộng rãi trong nhiều lĩnh vực khác nhau nhờ vào tính linh hoạt và độ chính xác của nó

Câu 9. Trang 7: Trong những năm gần đây, trí tuệ nhân tạo đã trở thành một xu hướng nổi bật

Độ trùng lặp: **87%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Trong những năm gần đây, ứng dụng trí tuệ nhân tạo (AI) đã trở thành một xu hướng nổi bật

Câu 10. Trang 7: Những công cụ này không chỉ giúp tối ưu hóa quy trình phát triển mà còn hỗ trợ tối đa trong việc kiểm thử và tinh chỉnh mô hình

Độ trùng lặp: **51%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: công cụ này không chỉ giúp tối ưu hóa quy trình làm việc mà còn hỗ trợ doanh nghiệp trong việc

Câu 11. Trang 7: Đặc biệt, cộng đồng Python luôn sẵn sàng chia sẻ kiến thức và kinh nghiệm, giúp người dùng giải quyết các vấn đề kỹ thuật một cách nhanh chóng và hiệu quả

Độ trùng lặp: **66%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: cộng đồng người dùng và nhà phát triển lớn mạnh, sẵn sàng hỗ trợ, chia sẻ kiến thức và kinh nghiệm, giúp giải quyết các vấn đề một cách nhanh chóng và hiệu quả Ngoài ra, các diễn đàn, tài liệu hướng dẫn và các tài nguyên học tập luôn sẵn

Câu 12. Trang 7: Tkinter là thư viện tiêu chuẩn của Python để xây dựng giao diện đồ họa người dùng (GUI)

Độ trùng lặp: **72%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: giao diện đồ họa người dùng (GUI Graphic User Interface) Tkinter là một phần của thư viện tiêu chuẩn của Python

Câu 13. Trang 9: AI không chỉ là một phần của khoa học máy tính mà còn là một cuộc cách mạng công nghệ với tiềm năng tuyệt vời Để thay đổi cách chúng ta làm việc, sống và tương tác với công nghệ

Độ trùng lặp: 67%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: công nghệ AI không chỉ là một khía cạnh mới của cuộc cách mạng công nghệ mà còn là một công cụ mạnh mẽ có thể thay đổi cách chúng ta làm việc, sống và tương tác với

Câu 14. Trang 9: Ví dụ, Trong lĩnh vực Y tế AI đã giúp cải thiện đáng kể quy trình chẩn đoán bệnh, lý và phát triển các phương pháp điều trị cá nhân hóa

Độ trùng lặp: 54%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: y tế, và phát triển các phương pháp điều trị cá nhân hóa

Câu 15. Trang 9: trong lĩnh vực giao thông công nghệ AI hỗ trợ trong việc phát triển xe tự lái an toàn và hiệu quả hơn, cùng với các Hệ thống quản lý giao thông thông minh để giảm thiểu ùn tắc và tai nạn giao thông

Độ trùng lặp: 55%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: lĩnh vực giao thông, và vận tải, AI là công nghệ cốt lõi Trong việc phát triển xe tự hành Nó giúp xe nhận diện và phản ứng với môi trường xung quanh một cách an toàn AI còn hỗ trợ tối ưu hóa các tuyến đường giao hàng và quản lý chuỗi cung ứng Từ đó giảm thiểu chi phí và thời gian vận chuyển hệ thống quản lý giao thông, thông minh sử dụng AI để điều chỉnh đèn giao thông, và dự báo các tình huống khẩn cấp giúp giảm thiểu tắc nghẽn và cải thiện an toàn giao thông, Tài chính và ngân hàng Ứng dụng Trong lĩnh vực

Câu 16. Trang 9: trong tương lai, sự phát triển của AI Dự kiến sẽ tiếp tục mở ra nhiều cơ hội mới và thúc đẩy sự tiến bộ của loài người thông qua việc áp dụng công nghệ vào các lĩnh vực khác nhau

Độ trùng lặp: 57%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Trong lĩnh vực công nghệ AI tại Việt Nam dự báo xu hướng công nghệ AI tại Việt Nam Xu hướng phát triển công nghệ AI tại Việt Nam dự kiến sẽ tiếp tục tăng trưởng với việc đẩy mạnh nghiên cứu và phát triển công nghệ AI Trong nhiều lĩnh vực khác nhau sự phát triển của AI Trong kết nối IoT, blockchain và big data dự kiến sẽ đem lại nhiều tiềm năng và triển vọng cho Việt Nam việc thúc đẩy sự hợp tác giữa các công ty, tổ chức, trường đại học và chính phủ sẽ đóng vai trò quan trọng Trong việc phát triển và ứng dụng công nghệ AI tại Việt Nam công nghệ AI tại Việt Nam sẽ phổ biến rộng rãi Trong nhiều năm tới Kết luận công nghệ AI tại Việt Nam đang có sự tiến bộ đáng kể và triển vọng rất lớn việc phát triển và áp dụng công

Câu 17. Trang 9: Xử lý ngôn ngữ tự nhiên (Natural Language Processing NLP) nghiên cứu và ứng dụng để máy tính có thể hiểu, đánh giá, và tạo ra ngôn ngữ tự nhiên như con người bao gồm Xử lý văn bản, giọng nói, và ngữ pháp.

Độ trùng lặp: 60%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: tự nhiên (Natural Language Processing NLP) là một lĩnh vực Nghiên cứu và ứng dụng của trí tuệ nhân tạo liên quan đến tương tác giữa con người, và máy tính thông qua ngôn ngữ

tự nhiên. Mục tiêu chính của NLP là giúp máy tính hiểu, diễn giải và tạo ra ngôn ngữ tự nhiên một cách tự động. NLP bao gồm nhiều phần tử và công nghệ khác nhau, bao gồm Xử lý ngôn ngữ tự nhiên. Đây là quá trình Xử lý và phân tích ngôn ngữ tự nhiên. Điều này bao gồm việc tách từ, phân tích cú pháp, phân loại ngữ nghĩa, trích xuất thông tin và truy vấn ngôn ngữ tự nhiên. hiểu, ngôn ngữ tự nhiên. Mục tiêu là giúp máy tính hiểu, ý nghĩa của ngôn ngữ tự nhiên. Điều này bao gồm hiểu, ý nghĩa của câu, xác định ngữ cảnh, nhận dạng tác giả hoặc người, nói, và

Câu 18. Trang 9: thị giác máy (Computer Vision) lĩnh vực nghiên cứu và phát triển các thuật toán cho máy tính, nhận diện và hiểu hình ảnh và video giúp máy tính, nhận biết đối tượng, hành động, và môi trường xung quanh

Độ trùng lặp: 52%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: máy tính nghiên cứu và phát triển các phương pháp và công nghệ để máy tính có thể hiểu và xử lý thông tin từ hình ảnh và video, như con người các phương pháp và công nghệ trong Thị giác máy tính nhằm mục đích phân tích và trích xuất thông tin từ hình ảnh và video, nhận diện và phân loại đối tượng, và hiểu và tương tác với môi trường xung quanh

Câu 19. Trang 10: học máy (Machine Learning) phân nhánh của trí tuệ nhân tạo, tập trung vào việc phát triển các thuật toán và kỹ thuật để máy tính có thể học hỏi từ dữ liệu và tự động cải thiện hiệu suất của chúng

Độ trùng lặp: 69%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: trí tuệ nhân tạo tập trung vào việc phát triển các thuật toán và mô hình có khả năng Học hỏi từ dữ liệu. Thay vì lập trình trực tiếp các quy tắc cụ thể để giải quyết một vấn đề, machine learning cho phép máy tính tự Học thông qua việc Phân tích và rút ra những quy luật từ dữ liệu đầu vào. Công nghệ này cũng cho phép máy tính có khả năng cải thiện chính bản thân chúng dựa trên dữ liệu mẫu (training data) và những kinh nghiệm đã Học được. Máy Học là gì là thách thức của nhiều người dùng kỹ thuật số. Mục tiêu của máy Học là tạo ra các hệ thống có khả năng tự cải thiện hiệu suất theo thời gian và có khả năng tự động hóa quá trình Học hỏi. Bằng cách sử dụng các phương pháp thống kê, toán Học và khai phá dữ liệu machine learning giúp máy tính phát hiện và Học các mẫu, quy luật ẩn và tạo ra các mô hình dự đoán hoặc Phân loại dữ liệu mới. Phân loại máy Học có thể

Câu 20. Trang 10: robot học (Robotics) Áp dụng trí tuệ nhân tạo và các kỹ thuật lập trình để thiết kế, xây dựng, và điều khiển robot có khả năng thực hiện các tác vụ phức tạp một cách tự động

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: thiết kế, xây dựng vận hành và sử dụng Robot để thực hiện các nhiệm vụ một cách tự động hoặc bán tự động. Sau đây là tổng quan về Robot. 2.1 Định nghĩa Robot kết hợp các yếu tố của kỹ thuật cơ khí, kỹ thuật điện, khoa học máy tính và trí tuệ nhân tạo để tạo ra những cỗ máy có khả năng tương tác với thế giới vật lý. Robot là những cỗ máy có thể lập trình có thể thực hiện các hành động phức tạp một cách chính xác. 2.2 các thành phần của Robot. Phần cứng các thành phần vật lý của Robot bao gồm bộ truyền động cảm biến, bộ điều khiển

Câu 21. Trang 10: Học sâu (Deep Learning) Phân nhánh của Học máy, tập trung vào việc xây dựng và huấn luyện các mạng nơ ron sâu để giải quyết các vấn đề phức tạp như nhận dạng hình ảnh, nhận dạng giọng nói, và xử lý ngôn ngữ tự nhiên

Độ trùng lặp: 65%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Học sâu (Deep Learning) Một nhánh của Học máy sử dụng các mạng nơ ron nhân tạo để mô phỏng cách thức hoạt động của não người, hỗ trợ giải quyết các vấn đề phức tạp, như nhận dạng hình ảnh, giọng nói, và xử lý ngôn ngữ tự nhiên

Câu 22. Trang 10: tự động hóa (Automation) là sự kết hợp mạnh mẽ giữa trí tuệ nhân tạo (AI) và tự động hóa quy trình robot (RPA) mang đến giải pháp tự động hóa quy trình và công việc, giúp tăng năng suất và giảm thiểu sự can thiệp của con người

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giữa Tự động hóa quy trình robot (RPA), với các công nghệ hiện đại như trí tuệ nhân tạo (AI), học máy (ML) và xử lý ngôn ngữ Tự nhiên (NLP) Khác với các giải pháp Tự động hóa

Câu 23. Trang 14: Phân cụm (Clustering) là phương pháp nhóm các điểm dữ liệu vào các cụm sao cho các điểm trong cùng một cụm có độ tương đồng lớn hơn so với các điểm ở cụm khác

Độ trùng lặp: 60%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: là nhóm các điểm dữ liệu thành K cụm khác nhau, sao cho các điểm trong cùng một cụm có độ tương đồng cao hơn với nhau so với các điểm

Câu 24. Trang 16: CS3FC (Confidence Semi supervised FCM) đã được phát triển mạnh mẽ trong những năm gần đây

Độ trùng lặp: 66%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: phát triển mạnh mẽ trong những năm gần đây

Câu 25. Trang 17: tiền xử lý ảnh Đây là bước quan trọng nhằm loại bỏ các yếu tố gây nhiễu, và chuẩn hóa, dữ liệu đầu vào

Độ trùng lặp: 59%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Tiền xử lý dữ liệu là đảm bảo tính chính xác và đáng tin cậy của dữ liệu đầu vào. Khi dữ liệu không được Tiền xử lý đúng cách, có thể dẫn đến các vấn đề như dữ liệu nhiễu dữ liệu thiếu, dữ liệu không đồng nhất, và có thể ảnh hưởng đến kết quả xử lý dữ liệu sau này. các phương pháp Tiền xử lý dữ liệu bao gồm 1 Làm sạch dữ liệu Đây là quá trình loại bỏ các giá trị nhiễu dữ liệu trùng lặp, dữ liệu không chính xác và dữ liệu không đầy đủ trong tập dữ liệu. Việc làm sạch dữ liệu giúp đảm bảo tính toàn vẹn và chính xác của dữ liệu. 2 chuẩn hóa dữ liệu chuẩn hóa dữ liệu là

Câu 26. Trang 18: đánh giá kết quả Cuối cùng kết quả sẽ được so sánh với dữ liệu tham chiếu (nếu có), tính các chỉ số đánh giá như Accuracy, IoU, F1 score

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Cuối cùng, kết quả phân loại sẽ được so sánh với các vùng mẫu và dữ liệu tham chiếu sử dụng ma trận sai số các chỉ số sử dụng cho Đánh giá

Câu 27. Trang 23: Độ chính xác (Accuracy) tỷ lệ tổng số mẫu được phân loại đúng trên tổng số

mẫu

Độ trùng lặp: **87%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: số mẫu được phân loại đúng trên tổng số mẫu

Câu 28. Trang 23: Độ chính xác dương tính (Precision) tỷ lệ các mẫu được gán nhãn dương tính đúng trên tổng số mẫu được gán nhãn dương tính

Độ trùng lặp: **62%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: các mẫu được gán nhãn đúng chia cho tổng số dữ liệu kiểm thử (gồm cả mẫu mã độc và lành tính) (2 2) + Tỷ lệ dương tính đúng (

Câu 29. Trang 23: F1 Score trung bình điều hòa giữa Precision và Recall, đánh giá cân bằng giữa hai chỉ số này

Độ trùng lặp: **83%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Trung bình điều hòa giữa Precision và Recall, nhằm cân bằng giữa hai chỉ số này Weighted F1 Score Trung bình

Câu 30. Trang 25: thu thập và tiền xử lý dữ liệu dữ liệu đầu vào là các ảnh vệ tinh, cần được tiền xử lý để giảm nhiễu và chuẩn hóa thông tin

Độ trùng lặp: **52%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: xử lý dữ liệu dữ liệu Thu thập được thường cần được tiền xử lý để loại bỏ nhiễu và chuẩn hóa

Câu 31. Trang 25: thuật toán phân cụm bán giám sát mờ phần này sử dụng lý thuyết mờ để phân cụm các điểm dữ liệu trong không gian đặc trưng

Độ trùng lặp: **53%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: để phân cụm trong phân đoạn ảnh cuối cùng Tuy nhiên Thuật toán này có một số nhược điểm chưa sử dụng các đặc trưng của ảnh nha khoa và khi phân cụm chưa sử dụng thành Phần không gian trong hàm mục tiêu để tăng độ chính xác 2 2 Thuật toán phân cụm bán giám sát mờ

Câu 32. Trang 25: Trích xuất và phân loại đối tượng Sau khi phân cụm các đối tượng trong ảnh, sẽ được Trích xuất và phân loại thành các lớp đối tượng khác nhau như đất đai, nước

Độ trùng lặp: **56%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: đối tượng trong ảnh Chẳng hạn như màu sắc, hình ảnh hoặc các đặc tính khác để có thể phân biệt các đối tượng khác nhau Bước 5 phân loại đối tượng

Câu 33. Trang 32: Các Thư viện sử dụng Ngôn ngữ C gồm Các Thư viện chuẩn stdio h, stdlib h, math h, string h, time h, conio h

Độ trùng lặp: **55%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Các thư viện C chuẩn (C Standard Library) <assert h> <ctype h> <errno h> <float h> <limits h> <locale h> <math h> <setjmp h> <signal h> <stdarg h> <stddef h> <stdio h> <stdlib h> <string h> <time h>

Câu 34. Trang 38: Ứng dụng hỗ trợ người dùng thực hiện phân cụm dữ liệu đánh giá và trực quan hóa kết quả qua giao diện thân thiện

Độ trùng lặp: **56%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: và giao diện thân thiện Phần mềm này hỗ trợ người dùng trong việc thực hiện các phân tích thống kê phức tạp và trực quan hóa kết quả

Câu 35. Trang 41: Mặc dù đã cố gắng trong quá trình thực hiện để hoàn thành đồ án tốt nhất, do năng lực và trình độ vẫn còn hạn, chế nên không tránh khỏi những thiếu sót

Độ trùng lặp: **64%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: Mặc dù đã cố gắng trong quá trình thực hiện để khóa luận có tính khoa học và thực tiễn cao nhất, song do thời gian có hạn trình độ chuyên môn và vốn kiến thức còn hạn chế nên không tránh khỏi những hạn chế thiếu sót

Câu 36. Trang 41: Kính mong thầy cô phản hồi, chỉ bảo thêm để đồ án của em được hoàn thiện hơn

Độ trùng lặp: **72%**

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: chỉ bảo thêm để đồ án của em được hoàn

--- Hết ---