



BÁO CÁO KIỂM TRA TRÙNG LẶP

Thông tin tài liệu

ĐỒ ÁN TỐT NGHIỆP Tên tài liệu:

Ngành CNTT Tác giả:

5 Điểm trùng lặp:

09:47 21/01/2024 Thời gian tải lên: 09:52 21/01/2024 Thời gian sinh báo cáo:

70/70 trang Các trang kiểm tra:



Kết quả kiểm tra trùng lặp



Có 5% nội dung trùng lặp



Có 95% nôi 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 luanvan.moet.gov.vn

Danh sách các câu trùng lặp

1. Trang 4: Nhiệm vụ Đồ án tốt nghiệp đã được Hội đồng thi tốt nghiệp của Khoa thông qua

Độ trùng lặp: 100%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Nhiệm vụ Đồ án tốt nghiệp đã được Hội đồng thi tốt nghiệp của Khoa thông qua

2. Trang 4: (Ký và ghi rõ Họ tên) Sinh viên đã hoàn thành và nộp bản Đồ án tốt nghiệp cho Hội đồng thi ngày

Độ trùng lặp: 100%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Họ tên) Sinh viên đã hoàn thành và nộp bản Đồ án tốt nghiệp cho Hội đồng thi</u> ngày tháng năm <u>Sinh viên</u> làm <u>Đồ án tốt nghiệp (Ký và ghi rõ Ho tên)</u>

3. Trang 5: Tuy nhiên, <u>sự phát triển của công nghệ, đặc biệt là trí tuệ nhân tạo (AI),</u> đã thay đổi <u>cách</u> các đội bóng quản lý và <u>phát triển</u> đội hình

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: sư phát triển của công nghê, đặc biệt là trí tuê nhân tạo (AI), cách

4. Trang 5: Điều này giúp tăng cường tính chính xác và khách quan trong quá trình đánh giá cầu

thủ

Độ trùng lặp: 58%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: tính chính xác và khách quan trong quá trình đánh giá

5. Trang 7: Tác giả xin cam đoan đây là Đồ án tốt nghiệp của bản thân Tác giả

Đô trùng lặp: 100%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Tác giả xin cam đoan đây là Đồ án tốt nghiệp của bản thân Tác giả</u>

6. Trang 7: Các kết quả trong Đồ án tốt nghiệp này là trung thực, và không sao chép từ bất kỳ một nguồn nào và dưới bất kỳ hình thức nào

Độ trùng lặp: 100%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Các kết quả trong Đồ án tốt nghiệp này là trung thực, và không sao chép từ bất kỳ một nguồn nào và dưới bất kỳ hình thức nào</u>

7. Trang 7: Việc tham khảo các nguồn tài liệu (nếu có) đã được thực hiện trích dẫn và ghi nguồn tài liệu tham khảo đúng quy đinh

Độ trùng lặp: 100%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Việc tham khảo các nguồn tài liêu (nếu có) đã được thực hiện trích dẫn và ghi

nguồn tài liệu tham khảo đúng quy định

8. Trang 8: Trong suốt quá trình học tập tại Khoa Công nghệ thông tin trường Đại học Thủy Lợi, em cảm thấy vô cùng may mắn khi <u>đã nhận được sự quan tâm, giúp đỡ, và sự</u> chỉ bảo nhiệt tình <u>của các thầy cô giáo, Trong</u> cả <u>học tâp và</u> đời sống

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Trong suốt quá trình học tập tại khoa Công nghệ thông tin trường Đại học Công nghệ Đại học</u> Quốc gia Hà Nội, <u>em đã nhận được sự quan tâm, giúp đỡ</u> tạo mọi điều kiện <u>của</u> nhà <u>trường của các thầy, giáo cô giáo</u> đến nay <u>em đã</u> hoàn thành chương <u>trình học tập và</u> luận văn tốt nghiệp <u>em</u> xin gửi lời <u>cảm</u>

9. Trang 8: <u>và</u> đặc biệt hơn, <u>trong</u> quãng <u>thời gian làm đồ án tốt nghiệp em đã nhận được sự hướng dẫn tận tình của</u> Thạc sĩ, <u>thầy giáo</u> Nguyễn Đắc Hiếu

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>em đã nhận được sự hướng dẫn tận tình của thầy giáo</u> Võ Văn Hường, đồng <u>thời em</u> cũng <u>nhận được</u> ý kiến đóng góp <u>của</u> các <u>thầy trong</u> Bộ mơn Otơ <u>Và</u> xe chun dụng Mặc dù <u>đã</u> cố gắng nhưng do kiến thức <u>Và thời gian làm đồ án tốt nghiệp.</u>

10. Trang 8: em xin gửi lời biết ơn sâu sắc đến thầy, người <u>đã giúp đỡ</u> hỗ trợ <u>và cho em những lời khuyên hữu ích để em có thể hoàn thành đồ án tốt nghiệp của</u> mình hiệu quả <u>và</u> kịp tiến độ <u>của học</u> phần

Độ trùng lặp: 53%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>để Em có thể hoàn thành</u> đề tài nghiên cứu Đồng thời, <u>Em xin gửi lời</u> kính trọng <u>và</u> tự hào <u>của Em</u> khi được <u>học</u> tập tại Trường Đại <u>học</u> Nha Trang <u>Em xin gửi đến</u> sự <u>biết ơn sâu sắc</u> nhất <u>cho</u> cô Mai Thị Tuyết Nga <u>đã</u> tận tình hướng dẫn chỉ bảo <u>những</u> kinh nghiệm quý báu, <u>cho Em những lời khuyên hữu ích và</u> động viên <u>Em</u> trong suốt quá trình thực hiện <u>đồ án tốt nghiệp và Em</u> cũng <u>xin</u> được ghi nhớ tình cảm, sự <u>giúp đỡ.</u>

11. Trang 8: em rất mong nhân được sư góp ý của các thầy cô để đô án trở nên hoàn thiên hơn

Độ trùng lặp: 85%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: mong nhận được sự góp ý của các thầy, cơ để đồ án trở nên hoàn thiện hơn Một lần nữa <u>Em</u> xin gửi lời cảm ơn đến <u>thầy</u>. TS Nguyễn Hữu Luân và <u>các thầy</u>, <u>cô</u>

12. Trang 11: LightGBM 51 <u>Hình 4 7 Biểu đô</u> thể hiện <u>giữa giá trị</u> thực tế <u>và giá trị dự đoán của</u> <u>Mô Hình</u> Random Forest 51 v

Độ trùng lặp: 56%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: 951 Hình 4 8 đồ thị phân tán giữa các phần dư chuẩn hóa và giá trị dự đoán

<u>của mô</u>

13. Trang 15: CHƯƠNG 1 GIỚI THIỀU BÀI toán GIỚI THIỀU TỔNG QUAN 1 1 Lý do chon đề tài

Độ trùng lặp: 83%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: CHƯƠNG 1 GIỚI THIÊU TỔNG QUAN 1 1 Lý do chon đề tài Theo báo cáo

thanh <u>TOÁN,</u>

14. Trang 15: Với hơn 5 tỷ <u>người hâm mộ trên</u> toàn <u>thế giới bóng đá không chỉ là một môn thể thao</u> phổ biến nhất <u>trên thế giới mà còn là một hiện tương văn hóa manh mẽ</u>

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>là một hiện tượng văn hóa</u> đưa cầu thủ và <u>người hâm mộ</u> trở thành <u>một</u>) Stt <u>bóng</u> <u>đá</u> bằng tiếng Nhật Bản <u>bóng đá là môn thể thao</u> được yêu thích <u>trên</u> khắp <u>thế giới.</u> Nó <u>không chỉ là một</u> trò chơi đơn thuần <u>mà còn là một</u>

15. Trang 18: học máy là một lĩnh vực Trong trí tuệ nhân tạo (AI) mà Nó tập trung vào việc phát triển các phương pháp <u>và thuật toán để máy tính có thể học từ dữ liệu mà không cần phải được lập trình một cách</u> tường <u>minh</u>

Độ trùng lặp: 70%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>và Học máy Học máy</u> Machine learning <u>được</u> hiểu <u>là một lĩnh vực</u> con của <u>trí tuệ</u> <u>nhân</u> tạo(Artificial Intelligence) <u>nó</u> sử dụng những <u>thuật toán</u> cho phép <u>máy tính có thể Học từ dữ liệu để</u> thực hiện <u>các</u> công <u>việc</u> thay vì <u>được lập trình một cách</u> rõ ràng <u>Học máy là một</u> ứng dụng của <u>trí tuệ nhân tạo (AI)</u> cung cấp cho <u>các</u> khả năng tự động <u>Học</u> hỏi <u>và</u> cải thiện <u>từ</u> kinh nghiệm <u>mà không cần lập trình</u> rõ ràng <u>Học máy tập trung vào việc phát triển các</u>

16. Trang 18: Học máy có thể ứng dụng hầu hết các loại dữ liệu từ cấu trúc (bảng số liệu kinh tế, số liệu nông nghiệp,) và cả dữ liệu phi cấu trúc (hình ảnh, văn bản, tín hiệu âm thanh,)

Đô trùng lặp: 81%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>Học máy có thể ứng dụng hầu hết các loại dữ liệu từ cấu trúc (bảng số</u> liệu) <u>và cả dữ liệu phi cấu trúc (hình ảnh, văn bản,</u> tính <u>hiệu âm</u> thanh)

17. Trang 18: học máy được chia thành 3 loại chính học có giám sát, học không giám sát, và học tăng cường

Độ trùng lặp: 80%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Học máy được chia thành</u> ba <u>loại</u> là <u>Học có giám sát, Học không giám sát, và Học</u> tăng cường

18. Trang 18: Học có giám sát (supervised learning) là một dạng Học máy trong đó mô hình được huấn luyên trên dữ liệu có nhãn

Đô trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Học có giám sát (Supervised Learning) là dạng Học máy trong đó</u> tập <u>dữ liệu huấn luyên</u> đầu vào <u>được</u>

19. Trang 18: Sử dụng khi bạn muốn nhận dự đoán một kết quả đầu ra từ dữ liệu đầu vào và bạn có các cặp dữ liệu (đầu vào/đầu ra) tương ứng

Độ trùng lặp: 72%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Sử dụng khi muốn nhận dự đoán một kết quả đầu ra từ dữ liệu đầu vào và</u> c <u>các</u> c

p <u>dữ liệu (đầu vào/đầu ra)</u> t ơng <u>ứng</u>

20. Trang 19: Bagging <u>cho phép lựa chọn ngẫu nhiên một nhóm nhỏ các thuộc tính tại mỗi nút của cây để phân chia</u> dữ liệu, giúp <u>mô hình, có</u> khả năng <u>phân loại một cách</u> linh hoạt <u>và nhanh</u> chóng

Độ trùng lặp: 59%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>cách lựa chọn ngẫu nhiên một nhóm nhỏ các thuộc tính tại mỗi nút của cây để</u> <u>phân chia cho</u>

21. Trang 19: Đối với bài toán hồi quy, kết quả cuối cùng của mô hình RFR sẽ là trung bình của tất cả các kết quả dư báo của các cây (Hình 2 1)

Độ trùng lặp: 91%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Đối với bài toán hồi quy, kết quả cuối cùng của mô hình RFR sẽ là trung bình của tất cả các kết quả dư báo của các cây

22. Trang 19: Đối với bài toán hồi quy kết quả cuối cùng là giá tri trung bình của các dự đoán từ các cây như công thức sau = 1

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giá trị trung bình (average) cho bài toán hồi quy. Việc kết hợp các mô hình CART này để cho kết quả cuối cùng nên RF được gọi <u>là</u> phương pháp học tập thể <u>Đối với bài toán</u> phân loại, <u>cây</u> CART sử dụng <u>công thức</u> Gini <u>như là</u> một hàm điều kiện để tính <u>toán</u> điểm tách nút <u>của</u> <u>cây</u> Số lượng <u>cây là</u> không hạn chế, <u>các cây</u> trong RF được xây dựng

23. Trang 20: ban đầu nhắm đến giải quyết các bài toán phân loại và hồi quy XGBoost đã nhanh chóng trở thành một trong những thuật toán phổ biến nhất

Độ trùng lặp: 55%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: một trong những thuật toán học máy được sử dụng rất phổ biến bởi độ mạnh của nó trong việc giải quyết các bài toán phân loại và hồi quy.

24. Trang 21: <u>Tối ưu hóa phần cứng</u> XGBoost <u>được Tối ưu hóa</u> cho việc <u>sử dụng tài nguyên phần cứng</u> thông qua việc <u>nhận thức bộ nhớ cache</u> và <u>phân bổ bộ đệm nội bộ</u>

Độ trùng lặp: 60%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>Tối ưu hóa phần cứng</u> Thuật tốn này <u>được</u> thiết kế để <u>sử dụng</u> hiệu quả <u>tài</u> nguyên phần cứng Điều này <u>được</u> thực hiện nhờ <u>nhận thức bộ nhớ cache</u> bằng cách <u>phân bổ bộ</u> <u>đệm nội bộ</u>

25. Trang 22: tạo cây quyết định (decision Trees) Mỗi cây quyết định được xây dựng bằng cách chia dữ liệu thành các nhóm dựa trên các đặc trưng của dữ liệu

Độ trùng lặp: 70%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>của cây quyết đinh (Decision</u> tree) ETR hoạt động <u>bằng cách Tạo</u> ra một tập hợp <u>các cây quyết định</u> ngẫu nhiên <u>Mỗi cây quyết định được xây dựng bằng cách chia dữ liệu thành các nút con dựa trên các</u>

26. Trang 22: <u>tối ưu</u> hóa <u>hàm mất mát</u> (Loss Function Optimization) Mục tiêu <u>là tố</u>i thiểu hóa <u>hàm mất mát là</u> sự <u>chênh lệch giữa giá tri dư đoán và giá tri thực tế</u>

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>hàm mất mát hàm mất mát hàm mất mát xác định độ chênh lệch giữa giá trị</u> đầu ra <u>dư đoán và giá tri</u> đầu ra <u>thực tế giá tri mất mát là</u>

27. Trang 22: <u>cross validation và</u> Early Stopping XGBoost thường <u>sử dụng</u> kỹ thuậ<u>t cross validation Để đánh giá hiệu suất của mô hình.</u>

Đô trùng lặp: 50%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: <u>hiệu suất của mô hình</u> chúng tôi <u>sử dụng</u> phương pháp five fold <u>Cross validation</u>

<u>và</u>

28. Trang 23: Leaf wise tree, growth lựa chọn nút để phát triển cây dựa trên tối ưu toàn bộ tree, trong khi level (depth) wise tree, growth tối ưu trên nhánh đang xét, do đó, với số node nhỏ, các cây xây dựng từ leaf wise tree, growth thường thực hiện tốt hơn level (depth) wise tree, growth (Hình 2 3)

Độ trùng lặp: 66%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>wise lựa chọn nút để phát triển cây dựa trên tối ưu toàn bộ tree trong khi</u> level <u>wise tối ưu trên nhánh đang xét, do đó với số node nhỏ, các tree xây dựng từ</u>

29. Trang 24: MSE (Mean <u>squared</u> Error) Đây <u>là trung bình giá trị bình phương của sai số giữa giá trị dư đoán và giá trị thực tế.</u>

Độ trùng lặp: 65%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>sai số giữa giá trị dự đoán và giá tri thực tế</u> SST <u>là</u> tổng <u>bình phương của</u> sự <u>sai</u> lệch <u>giữa giá tri thực tế và giá tri trung bình của</u> biến phụ thuộc Tại sao R <u>Squared</u>

30. Trang 24: MAE (Mean <u>Absolute</u> Error) <u>do độ lớn trung bình các lỗi trong một tập hợp các dự đoán mà không cần xem xét hướng của chúng</u>

Độ trùng lặp: 89%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Đo độ lớn trung bình của các lỗi trong một tập hợp các dự đoán mà không cần xem xét hướng của chúng

31. Trang 24: Đó là giá trị trung bình trên tập mẫu kiểm tra về sự khác biệt tuyệt đối giữa dự đoán và quan sát thực tế, trong Đó tất cả các khác biệt riêng lẻ có trong số bằng nhau

Độ trùng lặp: 87%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Đó là giá trị trung bình trên mẫu thử nghiệm về sự khác biệt tuyết đối giữa dự đoán và quan sát thực tế, trong Đó tất cả các khác biệt riêng lẻ có trong số bằng nhau

32. Trang 25: MAE <u>luôn không âm và giá trị 0 (hầu như không bao giờ đạt được trong thực tế) sẽ</u> chỉ ra sư phù hợp hoàn hảo với dữ liêu

Độ trùng lặp: 96%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>luôn không âm và giá trị 0 (hầu như không bao giờ đạt được trong thực tế) sẽ chỉ</u> ra sư phù hợp hoàn hảo với dữ liệu

33. Trang 25: R2 (Coefficient of Determination) Đây là độ đo đánh giá mức độ giải thích của mô hình hồi quy

Độ trùng lặp: 62%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: đo đánh giá mức đô giải thích của mô hình hồi quy

34. Trang 25: 2 <u>là tổng bình phương sai số</u> (Sum of Squared Errors), tức <u>là tổng của bình phương</u> hiệu giữa giá trị dư đoán và giá trị thực tế,

Độ trùng lặp: 60%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: bình phương sai số giữa giá tri dư đoán và giá tri thực tế SST <u>là tổng bình phương</u>

<u>cúa</u>

35. Trang 25: 2 <u>là tổng bình phương</u> độ <u>lệch</u> (Sum of <u>squared</u> Total), tức <u>là tổng của bình phương</u> hiệu g<u>iữa giá trị thực tế và giá trị trung bình của biến phụ thuộc</u>

Độ trùng lặp: 61%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>là tổng bình phương của</u> sự sai <u>lệch giữa giá trị thực tế và giá trị trung bình của</u> biến phu thuộc Tai sao R <u>Squared</u>

36. Trang 25: RMSE (Root Mean Squared Error) Đây Là căn bậc hai của trung bình giá trị bình phương của sai số giữa giá trị dư đoán và giá trị thực tế

Độ trùng lặp: 74%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>sai số giữa giá trị dự đoán và giá tri thực tế</u> trong bài toán <u>dự đoán RMSE</u> tính <u>căn</u> <u>bâc hai trung bình bình phương của</u> hiệu <u>giữa giá tri dư đoán và giá tri thực tế</u>

37. Trang 25: RMSE có thể được sử dụng để so sánh các mô hình khác nhau

Độ trùng lặp: 92%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: có thể được sử dung để so sánh các mô hình khác nhau

38. Trang 26: Pandas là một thư viện Python cung cấp các cấu trúc dữ liệu và công cụ phân tích dữ liêu manh mẽ đặc biệt là DataFrame

Độ trùng lặp: 64%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>phân tích dữ liệu</u> Nó <u>cung cấp một</u> loạt <u>các thư viện và</u> framework <u>mạnh mẽ</u>, để thực hiện <u>các</u> tác vụ xử lý <u>dữ liệu các thư viện</u> phổ biến bao gồm NumPy <u>thư viện</u> xử lý mảng <u>và</u> ma trận số học trong <u>Python Pandas thư viện cung cấp cấu trúc dữ liệu và công cụ phân tích mạnh mẽ</u>.

39. Trang 27: Gaussian Process <u>là một trong những mô hình phổ biến được sử dụng</u> cho mục đích này

Độ trùng lặp: 64%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: những mô hình phổ biến được sử dụng trong

40. Trang 28: Quá trình này lặp đi lặp lại cho đến khi đạt được sự hội tụ hoặc đã <u>đạt đến</u> số <u>lượng</u> lần <u>lặp</u> quy <u>đinh</u>

Độ trùng lặp: 65%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: Quá trình này lặp đi lặp lại cho đến khi đạt được sự hội tụ Tổng năng lượng xác

<u>định</u>

41. Trang 30: Bước 2 <u>chia dữ liệu Ở</u> bước <u>chia dữ liệu</u> này, <u>ta sẽ sử dụng hàm</u> train_test split <u>của scikit</u> learns <u>để chia dữ liệu thành 80% dữ liệu huấn luyện và 20% dữ liệu kiểm tra</u>

Đô trùng lặp: 60%

Nguồn: Dữ liệu nôi sinh

Nội dung nguồn: <u>ta Chia dữ liệu thành</u> tập <u>huấn luyện và</u> thử nghiệm <u>để</u> chúng <u>ta</u> có thể <u>điều</u> chỉnh <u>và</u> đánh giá mô hình Chúng <u>ta sẽ sử dụng hàm</u> train_test_split () từ <u>sciki</u>t learning <u>và sử dụng 80% dữ liệu</u> đào tạo (train) <u>và 20% để kiểm tra</u> (

42. Trang 55: Z score là một phép đo thống kê đo khoảng cách của một giá trị từ trung bình của tập dữ liêu, tính theo đơn vị độ lệch chuẩn

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Z score Z score là một</u> biện pháp <u>thống kê</u> định lượng <u>khoảng cách</u> (đo bằng <u>độ</u> <u>lệch</u> chuẩn) <u>từ một</u> điểm <u>dữ liệu.</u> bất kỳ đến <u>giá trị trung bình của một</u>

43. Trang 56: Về <u>mã hóa dữ liệu</u> tôi <u>sử dụng</u> LabelEncoder, một thư viện <u>của</u> Python <u>được</u> dùng <u>để chuyển đổi các giá tri của biến phân loại thành các biến số</u> nguyên

Độ trùng lặp: 50%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: thường <u>được sử dụng để mã hóa các biến phân loại</u> và sẽ <u>chuyển đổi các giá tri</u> <u>của biến thành các giá tri</u> dưới dạng <u>số mã hóa số</u> học <u>sử dụng các</u> thuật toán phức tạp <u>để mã hóa dữ liêu</u>

44. Trang 56: Mục tiêu của mô hình là dự đoán giá trị của biến Mục tiêu fee dựa trên các biến độc lập trong bô dữ liêu

Độ trùng lặp: 54%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Mục tiêu của mô hình</u> hồi quy <u>là</u> ước lượng (hay <u>dự</u> đoán) <u>giá trị của biến</u> phụ thuôc trên cơ sở các giá tri đã biết <u>của biến</u> độc lập

45. Trang 56: Sau đó, tập dữ liệu được chia thành tập Huấn luyện (X_train, y_train) <u>và tập Kiểm tra</u> (X_test, y_test) <u>Sử dung</u> hàm train test split

Độ trùng lặp: 60%

Nguồn: *Dữ liệu nội sinh* Nội dung nguồn: <u>sử dung</u>

46. Trang 57: Các <u>phương</u> pháp <u>đánh giá</u> như <u>Mean Squared Error</u>, (MSE), R <u>Squared Mean Absolute Error</u>, (MAE), <u>và</u> Root <u>Mean Squared Error</u>, (RMSE) <u>được sử dụng để đánh giá</u> hiệu suất <u>của mô hình</u> trên tập kiểm tra

Độ trùng lặp: 51%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Mean Squared Error</u> RMSE), <u>và</u> Sai số tuyệt đối trung bình (Mean <u>Absolute Error</u> MAE), <u>được sử dung để đánh giá</u>

47. Trang 58: Tiếp theo, <u>dữ liệu được</u> chia thành <u>tập Huấn luyện</u> (X_train, y_train) <u>và tập Kiểm tra</u> (X_test, y_test) bằng cách <u>Sử dụng</u> hàm train_test_split

Đô trùng lặp: 54%

Nguồn: *Dữ liệu nội sinh* Nội dung nguồn: <u>sử dung</u>

48. Trang 59: <u>dữ liệu được chia thành hai phần chính tập dữ liệu huấn luyện</u> (X_train, y_train) <u>và tập dữ liệu kiểm tra</u> (X_test, y_test)

Độ trùng lặp: 67%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: <u>Dữ liệu được chia thành hai phần chính tập Dữ liệu huấn luyên</u> chứa hồ sơ (khoảng %) <u>và tập Dữ liệu kiểm tra</u>

49. Trang 59: Giảm giá trị <u>này có thể làm tăng độ chính xác của</u> lá <u>nhưng cũng có thể</u> dẫn đến overfitting

Độ trùng lặp: 57%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: này có thể làm tăng đô chính xác của hệ thống nhưng cũng có thể

50. Trang 65: <u>hình</u> 4 31 Biểu đồ thể hiện <u>giữa giá trị thực tế và giá trị dự đoán của mô hình</u> LightGBM

Độ trùng lặp: 61%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: giữa giá trị thực tế và giá trị dư đoán của mô Hình

51. Trang 66: Tổng quan phân phối giữa giá trị dự đoán và giá trị thực tế của cả ba mô Hình có Sư tương đồng

Độ trùng lặp: 56%

Nguồn: *Dữ liệu nội sinh*

Nội dung nguồn: <u>sự tương quan, giữa giá tri dự đoán và giá tri thực tế của</u> các biến phụ thuộc Y, Y <u>và</u> Y <u>hình</u> % INH phóng thích theo thời gian <u>của</u> các cơng thức , , , <u>hình hình</u> chụp TEM <u>của</u> tiểu <u>phân</u> nano PLGA cơng thức <u>hình hình</u> chụp SEM <u>của</u> CT <u>hình hình</u> chụp SEM cơng thức phun sấy <u>hình Tổng</u>

52. Trang 69: Do thời gian có hạn, và trình độ hiểu biết của còn hạn, chế, nên đô án này không thể tránh khỏi những thiếu sót

Độ trùng lặp: 76%

Nguồn: Dữ liệu nội sinh

Nội dung nguồn: còn han chế nên đồ án này không thể tránh khỏi những thiếu sót

53. Trang 69: Rất mong sẽ nhận được những lời khuyên, <u>lời</u> đánh giá <u>từ quý</u> thầy, cô <u>để đô án</u> này được hoàn thiên hơn

Độ trùng lặp: 57%

Nguồn: Dữ liêu nôi sinh

Nội dung nguồn: mong nhận được những lời góp ý từ quý Thầy/Cô để đồ án được hoàn thiện hơn

--- Hết ---