HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Đỗ Việt Anh

NGHIÊN CỨU VÀ ỨNG DỤNG CÁC PHƯƠNG PHÁP HỌC MÁY ĐỂ PHÁT HIỆN GIAN LẬN GIAO DỊCH TRONG HỆ THỐNG THANH TOÁN NGÂN HÀNG

Chuyên ngành: Hệ thống thông tin
Mã số: 8.48.01.04 (Hệ thống thông tin)

TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

HÀ NỘI - NĂM 2025

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: ……

(Ghi rõ học hàm, học vị)

Phản biện 1: ……

Phản biện 2: ……

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: …… giờ …… ngày …… tháng ….. năm ……

Có thể tìm hiểu đề án tốt nghiệp tại:

– Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

Mở đầu

1. Lý do chọn đề tài

Gian lận tài chính gia tăng mạnh cùng sự phát triển của ngân hàng số. Các phương pháp dựa trên luật truyền thống không còn đáp ứng được trước hành vi gian lận biến đổi nhanh, nên cần nghiên cứu các mô hình học máy và học sâu để nâng cao năng lực phát hiện sớm.

2. Tổng quan vấn đề nghiên cứu

Dữ liệu gian lận có tỷ lệ mất cân bằng lớn và phụ thuộc thời gian. Mô hình truyền thống khó mô hình hóa hành vi, trong khi LSTM cho thấy hiệu quả cao trong phân tích chuỗi giao dịch nhưng chưa được đánh giá đầy đủ trong bối cảnh Việt Nam.

3. Đối tượng và phạm vi nghiên cứu

Nghiên cứu tập trung vào dữ liệu giao dịch điện tử và các mô hình phân loại gian lận; phạm vi giới hạn trong mô phỏng dữ liệu, tiền xử lý, huấn luyện và đánh giá mô hình. Không xem xét khía cạnh nghiệp vụ hay pháp lý.

4. Phương pháp nghiên cứu

Kết hợp nghiên cứu lý thuyết và thực nghiệm trên bộ dữ liệu mô phỏng gồm 200.000 giao dịch; so sánh

Logistic Regression, Random Forest, XGBoost và LSTM theo các chỉ số đánh giá chuẩn.

5. Cấu trúc đề án

Gồm bốn chương: cơ sở lý luận; đề xuất mô hình; thực nghiệm và đánh giá; cài đặt, thử nghiệm và bàn luận.

CHƯƠNG 1: CƠ SỞ LÝ LUẬN

1.1. Giới thiệu chương

Chương này trình bày nền tảng lý thuyết phục vụ bài toán phát hiện gian lận giao dịch trong ngân hàng, bao gồm tổng quan về gian lận tài chính, các phương pháp học máy ứng dụng trong phát hiện gian lận và các nghiên cứu liên quan.

1.2. Gian lận tài chính trong hệ thống ngân hàng

Gian lận tài chính gia tăng cùng sự phát triển của ngân hàng số. Hành vi gian lận đa dạng, biến đổi liên tục, đến từ cả bên trong và bên ngoài, và thường thể hiện qua chuỗi giao dịch thay vì từng giao dịch đơn lẻ. Dữ liệu gian lận có đặc điểm mất cân bằng mạnh, hành vi thay đổi theo thời gian và khác biệt giữa các kênh giao dịch. Các hệ thống dựa trên luật tĩnh không còn đáp ứng được do khó thích ứng với sự thay đổi hành vi gian lận. Do đó, cần các phương pháp phân tích dữ liệu hiện đại có khả năng học và thích nghi.

1.3. Học máy trong phát hiện gian lận giao dịch

Học máy cho phép mô hình tự học từ dữ liệu lịch sử để phân loại giao dịch gian lận và hợp lệ. Các mô hình như

Logistic Regression, Random Forest và XGBoost được sử dụng rộng rãi trong lĩnh vực tài chính nhờ khả năng phân loại phi tuyến và xử lý dữ liệu lớn. Tuy nhiên, các mô hình truyền thống thường xem giao dịch độc lập, trong khi gian lận thực tế mang tính chuỗi. Mô hình học sâu, đặc biệt là LSTM, cho phép phân tích quan hệ theo thời gian và nhận diện các hành vi bất thường xuất hiện thông qua chuỗi giao dịch liên tục.

1.4. Các nghiên cứu liên quan

1.4.1. Các nghiên cứu khai phá dữ liệu và mô hình học máy truyền thống

Nhiều nghiên cứu tập trung vào các mô hình phân loại có giám sát như Logistic Regression hoặc cây quyết định. Các mô hình này dễ triển khai và minh bạch nhưng hạn chế với dữ liệu mất cân bằng và khó mô tả hành vi thay đổi theo thời gian.

1.4.2. Xử lý dữ liệu mất cân bằng và học máy lai

Các kỹ thuật như SMOTE giúp cải thiện khả năng học lớp gian lận hiếm. Một số nghiên cứu kết hợp phương pháp không giám sát và có giám sát nhằm tăng khả năng tổng quát hóa. Tuy nhiên, hầu hết vẫn chỉ xử lý đặc trưng

tĩnh của giao dịch.

1.4.3. Các tiếp cận học sâu và mô hình chuỗi

Các mô hình RNN, đặc biệt LSTM, được chứng minh hiệu quả trong việc học hành vi theo chuỗi nhờ khả năng ghi nhớ phụ thuộc dài hạn. Một số nghiên cứu còn kết hợp mô hình với hệ thống tương tác người – máy để hỗ trợ ra quyết định, tuy nhiên chi phí tính toán và yêu cầu dữ liệu lớn là thách thức.

1.4.4. Nhận xét và khoảng trống nghiên cứu

Các nghiên cứu quốc tế cho thấy ứng dụng mô hình chuỗi thời gian mang lại hiệu quả cao, nhưng tại Việt Nam, việc đánh giá các mô hình sâu trên dữ liệu mô phỏng theo hành vi giao dịch còn hạn chế. Cần một nghiên cứu hệ thống so sánh mô hình truyền thống và mô hình chuỗi trên cùng bộ dữ liệu để chỉ ra lợi thế của từng phương pháp.

1.5. Kết luận chương

Chương 1 cung cấp nền tảng lý luận cho bài toán phát hiện gian lận: gian lận tài chính phức tạp, dữ liệu có tính chuỗi, mô hình truyền thống bị hạn chế và mô hình học sâu, đặc biệt LSTM, là hướng tiếp cận đầy tiềm năng. Đây là cơ sở cho việc đề xuất bộ mô hình trong chương tiếp

theo.

CHƯƠNG 2: ĐỀ XUẤT MÔ HÌNH

2.1. Giới thiệu chương

Chương này trình bày bộ mô hình được sử dụng trong nghiên cứu, bao gồm các mô hình học máy truyền thống và mô hình học sâu theo chuỗi thời gian. Mục tiêu là làm rõ cơ sở lựa chọn mô hình và đặc điểm chính của từng phương pháp phục vụ cho bài toán phát hiện gian lận giao dịch.

2.2. Cơ sở lựa chọn mô hình

Việc lựa chọn mô hình dựa trên đặc thù dữ liệu giao dịch: tỷ lệ gian lận rất thấp, quan hệ đặc trưng mang tính phi tuyến và hành vi giao dịch phụ thuộc theo thời gian. Do đó, nghiên cứu kết hợp mô hình truyền thống (để đánh giá hiệu quả với đặc trưng tĩnh) và mô hình LSTM (để khai thác chuỗi giao dịch). Bộ mô hình được chọn nhằm bảo đảm khả năng so sánh giữa hai hướng tiếp cận.

2.3. Nhóm mô hình học máy truyền thống

Nhóm này gồm Logistic Regression, Random Forest và XGBoost, đại diện cho các phương pháp phân loại phổ biến trong lĩnh vực tài chính và phù hợp làm chuẩn so sánh với mô hình học sâu.

2.3.1. Logistic Regression

Là mô hình tuyến tính đơn giản, dễ diễn giải và phù hợp làm baseline. Tuy nhiên, mô hình khó mô tả các quan hệ phi tuyến và phụ thuộc vào chất lượng đặc trưng đầu vào.

2.3.2. Random Forest

Là phương pháp ensemble dựa trên nhiều cây quyết định, cho phép mô hình hóa quan hệ phi tuyến và đánh giá độ quan trọng đặc trưng. Tuy nhiên, mô hình dễ thiên lệch về lớp đa số khi dữ liệu mất cân bằng nặng.

2.3.3. XGBoost

Là thuật toán boosting mạnh, có khả năng xử lý dữ liệu mất cân bằng thông qua điều chỉnh trọng số và thường đạt hiệu quả cao trong bài toán gian lận. Tuy nhiên, mô hình vẫn dựa trên đặc trưng tĩnh của từng giao dịch.

2.4. Nhóm mô hình học sâu theo chuỗi thời gian

2.4.1. LSTM (Long Short-Term Memory)

LSTM được thiết kế để xử lý dữ liệu chuỗi và có khả năng ghi nhớ phụ thuộc dài hạn giữa các giao dịch. Mô hình phù hợp với bài toán phát hiện gian lận vì nhiều hành vi bất thường chỉ lộ rõ khi phân tích chuỗi giao dịch thay vì

từng giao dịch độc lập.

2.5. Lý do lựa chọn bộ mô hình trong nghiên cứu

Bộ mô hình được lựa chọn bảo đảm tính đại diện: Logistic Regression thể hiện baseline; Random Forest và XGBoost đại diện cho nhóm mô hình cây quyết định phi tuyến; LSTM đại diện cho các mô hình chuỗi có khả năng học hành vi theo thời gian. Sự kết hợp này cho phép đánh giá toàn diện ưu – nhược điểm của từng hướng tiếp cận.

2.6. Kết luận chương

Chương 2 đã mô tả các mô hình được sử dụng và cơ sở lựa chọn chúng. Các mô hình truyền thống phục vụ việc đánh giá với đặc trưng tĩnh, trong khi LSTM khai thác đặc tính chuỗi của dữ liệu. Bộ mô hình này là nền tảng cho quá trình thực nghiệm và so sánh hiệu quả ở Chương 3.

CHƯƠNG 3: THỰC NGHIỆM VÀ ĐÁNH GIÁ (BẢN TÓM TẮT)

3.1. Giới thiệu chương

Chương này trình bày bộ dữ liệu sử dụng, quy trình xử lý dữ liệu, cấu hình huấn luyện, các chỉ số đánh giá và kết quả thực nghiệm của bốn mô hình: Logistic Regression, Random Forest, XGBoost và LSTM.

3.2. Bộ dữ liệu sử dụng trong nghiên cứu

Dữ liệu được mô phỏng gồm 200.000 giao dịch của khoảng 10.000 người dùng, tỷ lệ gian lận chiếm khoảng 1,2%. Mỗi giao dịch chứa nhóm đặc trưng giao dịch, hành vi và ngữ cảnh; dữ liệu được tổ chức theo chuỗi để phục vụ mô hình LSTM.

3.3. Tiền xử lý dữ liệu

Các bản ghi không hợp lệ hoặc thiếu thông tin được loại bỏ; dữ liệu số được chuẩn hóa; các trường phân loại được mã hóa phù hợp cho từng mô hình. Chuỗi giao dịch của mỗi người dùng được sắp xếp theo thời gian. Vấn đề mất cân bằng được xử lý thông qua trọng số và tham số mô hình thay vì can thiệp trực tiếp vào dữ liệu.

3.4. Phân chia dữ liệu huấn luyện và kiểm thử

Dữ liệu được chia theo tỷ lệ 70% – 15% – 15% cho huấn luyện, kiểm định và kiểm thử. Với mô hình LSTM, toàn bộ chuỗi của một người dùng chỉ nằm trong một tập dữ liệu nhằm tránh rò rỉ thông tin theo thời gian.

3.5. Cấu hình huấn luyện và tham số mô hình

Các mô hình được huấn luyện trên cùng môi trường tính toán. Logistic Regression sử dụng regularization L2; Random Forest tối ưu số cây và độ sâu; XGBoost tinh chỉnh learning rate, số cây và trọng số lớp; LSTM gồm 1–2 lớp LSTM với dropout và được tối ưu bằng Adam. Bảng tham số tối ưu được tổng hợp tại Bảng 3.1.

3.6. Các chỉ số đánh giá mô hình

Các mô hình được đánh giá bằng Precision, Recall, F1-score, ROC-AUC, kết hợp quan sát ma trận nhầm lẫn và đường cong Precision–Recall, nhằm phản ánh hiệu quả trong bối cảnh dữ liệu mất cân bằng mạnh.

3.7. Kết quả thực nghiệm

Kết quả trên tập kiểm thử được trình bày tại Bảng 3.2. LSTM đạt hiệu năng vượt trội với Precision, Recall, F1-score và ROC-AUC đều cao nhất (AUC ≈ 0,97). Các

mô hình truyền thống cho kết quả hạn chế, đặc biệt Recall thấp. Hình ROC-AUC và PRC cho thấy sự phân tách lớp tốt của LSTM.

3.8. Phân tích và đánh giá kết quả

Logistic Regression cho kết quả trung bình và tạo nhiều cảnh báo sai. Random Forest bỏ sót hầu hết giao dịch gian lận do lệch lớp. XGBoost cải thiện đáng kể nhưng vẫn kém trong xử lý quan hệ theo thời gian. LSTM khai thác hiệu quả chuỗi hành vi, phát hiện gian lận chính xác và cân bằng tốt giữa hai chỉ số Precision – Recall. Đây là mô hình phù hợp nhất cho bài toán.

3.9. Kết luận chương

Chương 3 khẳng định vai trò quan trọng của việc mô hình hóa chuỗi thời gian. LSTM cho hiệu năng vượt trội so với các mô hình truyền thống và được lựa chọn làm mô hình triển khai thử nghiệm ở chương tiếp theo.

CHƯƠNG 4: CÀI ĐẶT, THỬ NGHIỆM VÀ BÀN LUẬN

4.1. Giới thiệu chương

Chương này trình bày quá trình triển khai mô hình LSTM tối ưu trong môi trường giả lập gần thời gian thực, bao gồm kiến trúc triển khai, mô phỏng xử lý giao dịch, đánh giá khả năng đáp ứng và phân tích ý nghĩa thực tiễn trong bối cảnh hệ thống ngân hàng.

4.2. Kiến trúc triển khai

Mô hình LSTM được đóng gói dưới dạng dịch vụ suy luận độc lập, giao tiếp với hệ thống mô phỏng qua API REST. Kiến trúc gồm: tầng tiếp nhận yêu cầu, tầng tiền xử lý và xây dựng chuỗi giao dịch, tầng suy luận và tầng phản hồi. Lịch sử giao dịch được lưu trữ theo người dùng để tái tạo chuỗi thời gian phục vụ mô hình.

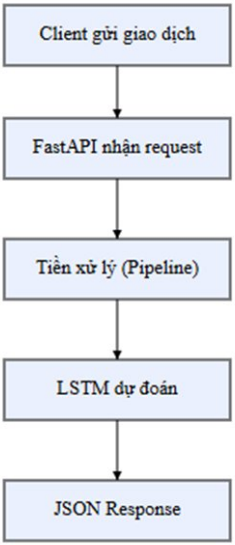


Hình 4.1. Kiến trúc tổng thể hệ thống phát hiện gian lận

4.3. Mô phỏng thử nghiệm trong môi trường giả lập

Trình mô phỏng sinh giao dịch với tốc độ từ 10 đến 200 giao dịch/giây và gửi tới API mô hình. Mỗi yêu cầu được xử lý qua các bước: xác thực dữ liệu đầu vào, truy vấn lịch sử giao dịch gần nhất, chuẩn hóa đặc trưng, suy luận bằng mô hình và trả về điểm rủi ro. Thời gian phản hồi trung bình ổn định ở mức vài chục mili-giây cho mỗi giao dịch.

Hệ thống được kiểm thử liên tục nhiều giờ và không phát sinh lỗi bộ nhớ hoặc tắc nghẽn kết nối, cho thấy dịch vụ LSTM hoạt động ổn định. Các chỉ số Precision và Recall ghi nhận trong môi trường giả lập gần tương đương kết quả huấn luyện, chứng tỏ mô hình giữ được hiệu năng khi triển khai dưới dạng API.



Hình 4.2. Luồng xử lý yêu cầu từ client đến mô hình LSTM

4.4. Đánh giá khả năng triển khai thực tế vào hệ thống ngân hàng

Kết quả cho thấy mô hình đáp ứng yêu cầu về tốc độ, khả năng mở rộng và tính ổn định của hệ thống phát hiện gian lận. Mô hình có thể tích hợp như một dịch vụ độc lập, không ảnh hưởng đến hệ thống lõi. Risk score do mô hình trả về hỗ trợ phân loại mức độ rủi ro và linh hoạt cấu hình ngưỡng cảnh báo. Kiến trúc triển khai cho phép mở rộng theo chiều ngang khi lưu lượng tăng cao.

4.5. Hạn chế và hướng cải tiến

Một số hạn chế gồm: dữ liệu mô phỏng chưa phản ánh đầy đủ hành vi gian lận thực tế; chi phí huấn luyện LSTM cao; khả năng giải thích mô hình hạn chế; yêu cầu tối ưu hạ tầng lưu trữ để giảm độ trễ truy vấn chuỗi. Các hướng cải tiến gồm: sử dụng dữ liệu thực đã ẩn danh, thử nghiệm mô hình Transformer hoặc hybrid, bổ sung kỹ thuật giải thích như SHAP/Attention, và xây dựng pipeline retraining – monitoring.

4.6. Kết quả và bàn luận

4.6.1. So sánh tổng quan các mô hình

Kết quả từ Chương 3 cho thấy LSTM vượt trội hơn mô hình truyền thống cả về Precision, Recall và ROC-AUC, nhờ khả năng khai thác quan hệ theo thời gian.

4.6.2. Bàn luận chi tiết theo từng mô hình

Logistic Regression tạo nhiều cảnh báo sai; Random Forest bỏ sót phần lớn gian lận; XGBoost cải thiện nhưng vẫn xem dữ liệu dưới dạng tĩnh; LSTM khai thác hành vi chuỗi và nhận diện tốt mẫu gian lận phức tạp.

4.6.3. Ý nghĩa đối với nghiệp vụ ngân hàng

Mô hình giúp giảm thiểu tổn thất, phát hiện gian lận

sớm, giảm số lượng cảnh báo sai và có thể tích hợp linh hoạt vào quy trình kiểm soát rủi ro nhờ cơ chế risk score.

4.6.4. Hạn chế nghiên cứu

Dữ liệu mô phỏng, chi phí tính toán cao và hạn chế giải thích mô hình ảnh hưởng một phần tới khả năng triển khai thực tế.

4.6.5. Hướng nghiên cứu tiếp theo

Mở rộng sang dữ liệu thực, thử nghiệm kiến trúc hiện đại hơn và hoàn thiện cơ chế giám sát – cập nhật mô hình.

4.7. Kết luận chương

Chương 4 khẳng định mô hình LSTM có khả năng triển khai trong môi trường gần thời gian thực, đáp ứng yêu cầu nghiệp vụ và mang lại hiệu quả vượt trội trong phát hiện gian lận. Các đánh giá kỹ thuật và nghiệp vụ là cơ sở quan trọng để đề xuất mô hình cho hệ thống ngân hàng.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Đề tài đã nghiên cứu và ứng dụng các phương pháp học máy và học sâu trong phát hiện gian lận giao dịch của hệ thống ngân hàng, với trọng tâm là mô hình học sâu LSTM. Trên cơ sở phân tích lý thuyết, tổng quan nghiên cứu và thực nghiệm trên bộ dữ liệu mô phỏng, luận văn đã xây dựng và đánh giá hiệu quả của bốn mô hình: Logistic Regression, Random Forest, XGBoost và LSTM. Kết quả cho thấy mô hình LSTM vượt trội so với các mô hình truyền thống ở tất cả các chỉ số đánh giá quan trọng, nhờ khả năng khai thác quan hệ phụ thuộc theo thời gian trong chuỗi giao dịch – yếu tố trọng tâm mà các mô hình học máy truyền thống không thể mô hình hóa.

Bên cạnh hiệu quả dự đoán, mô hình LSTM còn chứng minh khả năng triển khai thực tế thông qua quá trình mô phỏng trong môi trường gần thời gian thực. Các thử nghiệm cho thấy mô hình đáp ứng tốt các yêu cầu về tốc độ, tính ổn định và độ chính xác, qua đó có thể trở thành một thành phần quan trọng trong hệ thống phòng chống gian lận của ngân hàng. Kiến trúc triển khai được đề xuất

trong nghiên cứu cũng phù hợp với mô hình hệ thống ngân hàng hiện đại, cho phép tích hợp linh hoạt và mở rộng khi cần thiết.

Nhìn chung, luận văn đã hoàn thành đầy đủ các mục tiêu đề ra: khảo sát cơ sở lý luận, đề xuất mô hình, thực nghiệm, triển khai thử nghiệm và đánh giá mức độ phù hợp của mô hình học sâu trong bài toán phát hiện gian lận giao dịch ngân hàng.

2. Kiến nghị

Dựa trên các kết quả thực nghiệm và triển khai thử nghiệm, luận văn đưa ra một số kiến nghị nhằm nâng cao khả năng ứng dụng mô hình vào hệ thống ngân hàng trong thực tế:

Thứ nhất, cần mở rộng thử nghiệm trên dữ liệu giao dịch thực tế đã được ẩn danh hóa. Dữ liệu mô phỏng dù phản ánh nhiều đặc điểm quan trọng nhưng chưa thể hiện đầy đủ độ phức tạp và mức độ tinh vi của gian lận thật. Việc thử nghiệm trên dữ liệu thực sẽ giúp mô hình học được hành vi trong môi trường ngân hàng thực tiễn và tăng độ tin cậy khi triển khai.

Thứ hai, đối với quá trình vận hành thực tế, ngân

hàng cần xây dựng hệ thống giám sát mô hình (model monitoring) để quan sát độ chính xác theo thời gian và phát hiện sự suy giảm hiệu năng. Việc retrain mô hình định kỳ hoặc sử dụng các kỹ thuật học liên tục sẽ giúp mô hình thích nghi tốt hơn với các kiểu gian lận mới xuất hiện.

Thứ ba, cần xem xét áp dụng các kỹ thuật giải thích mô hình (Explainable AI) như Attention, LIME hoặc SHAP. Đây là yêu cầu quan trọng trong môi trường ngân hàng, nơi bộ phận kiểm toán và kiểm soát rủi ro cần hiểu rõ vì sao mô hình gán nhãn một giao dịch là gian lận để phục vụ công tác điều tra hoặc xử lý khiếu nại.

Thứ tư, để triển khai ở quy mô lớn, ngân hàng nên nghiên cứu tích hợp mô hình LSTM trong kiến trúc xử lý dữ liệu thời gian thực thông qua các nền tảng xử lý thời gian thực như Apache Kafka, Spark Streaming hoặc Flink, kết hợp với hệ thống điều phối luồng xử lý (workflow orchestration) như Apache Airflow. Điều này giúp giảm độ trễ, hạn chế truy vấn lặp và bảo đảm khả năng xử lý lượng giao dịch lớn trong các khung giờ cao điểm.

Cuối cùng, ngân hàng nên xem xét mở rộng mô hình sang các kênh giao dịch khác như thẻ, ATM, Mobile

Banking, Internet Banking và điểm chấp nhận POS. Việc mở rộng mô hình sang nhiều kênh giao dịch khác nhau sẽ giúp xây dựng một hệ thống phát hiện gian lận hợp nhất (Unified Fraud Detection System), từ đó nâng cao hiệu quả giám sát rủi ro trên toàn bộ hệ sinh thái ngân hàng và giảm thiểu các lỗ hổng giữa các kênh giao dịch.

3. Lời kết

Đề tài đã khẳng định tiềm năng và hiệu quả của các mô hình học sâu, đặc biệt là LSTM, trong việc phát hiện gian lận giao dịch ngân hàng. Kết quả nghiên cứu mở ra hướng ứng dụng thực tiễn với giá trị cao đối với lĩnh vực tài chính – ngân hàng, góp phần hoàn thiện hệ thống phòng chống gian lận và nâng cao chất lượng dịch vụ. Những nghiên cứu tiếp theo sẽ tiếp tục cải thiện và mở rộng mô hình theo hướng thông minh, linh hoạt và phù hợp hơn với biến động của môi trường số. Bằng cách kết hợp giữa phân tích lý thuyết, thực nghiệm mô hình và thử nghiệm triển khai, đề tài đã đóng góp một hướng tiếp cận mang tính khả thi và có thể áp dụng trong thực tiễn cho các hệ thống phát hiện gian lận trong ngân hàng.