HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



TRẦN THỊ HƯƠNG GIANG

ỨNG DỤNG GRAPH NEURAL NETWORK VÀ KỸ THUẬT XỬ LÝ DỮ LIỆU MẤT CÂN BẰNG TRONG PHÁT HIỆN GIAO DỊCH GIAN LẬN TÀI CHÍNH

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 8.48.01.01

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: PGS. TS. Đỗ Xuân Chợ

Phản biện 1: TS. Nguyễn Thanh Tuyên

Phản biện 2: PGS. TS. Nguyễn Linh Giang

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

Vào lúc: 10 giờ 00 ngày 22 tháng 01 năm 2026

Có thể tìm hiểu đề án tốt nghiệp tại:
– Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

NỘI DUNG TÓM TẮT

1. CƠ SỞ LÝ THUYẾT

1.1. Tổng quan về gian lận giao dịch tài chính

Trong các hệ thống tài chính – ngân hàng hiện đại, gian lận giao dịch tài chính là một trong những rủi ro nghiêm trọng nhất, gây thiệt hại lớn về kinh tế và làm suy giảm niềm tin của người dùng đối với các dịch vụ tài chính số. Gian lận giao dịch tài chính có thể được hiểu là các hành vi cố ý thực hiện giao dịch nhằm chiếm đoạt tài sản, làm sai lệch trạng thái tài chính hoặc khai thác các điểm yếu trong quy trình xử lý giao dịch của hệ thống.

Các hành vi gian lận có thể xuất hiện dưới nhiều hình thức khác nhau, bao gồm chuyển tiền trái phép, sử dụng tài khoản trung gian để che giấu dòng tiền, gian lận thẻ, lợi dụng hệ thống thanh toán điện tử hoặc thao túng chuỗi giao dịch để tránh bị phát hiện. Trong môi trường tài chính số, nơi giao dịch được xử lý tự động với tốc độ cao, các hành vi gian lận ngày càng trở nên tinh vi và khó nhận diện hơn.

Một đặc điểm quan trọng của gian lận giao dịch là tính không hiển nhiên ở mức giao dịch đơn lẻ. Nhiều giao dịch gian lận riêng lẻ có thể không khác biệt rõ ràng so với các giao dịch hợp lệ nếu chỉ xét các thuộc tính cục bộ như giá trị giao dịch hoặc thời điểm thực hiện. Các dấu hiệu gian lận thường chỉ bộc lộ khi giao dịch được xem xét trong mối quan hệ với các giao dịch khác, với lịch sử hoạt động của tài khoản hoặc trong một khoảng thời gian đủ dài.

Do đó, bài toán phát hiện gian lận giao dịch không chỉ là bài toán phân loại các giao dịch độc lập, mà còn là bài toán phân tích hành vi, trong đó cần xem xét bối cảnh và mối quan hệ giữa các thực thể tham gia giao dịch.

1.2. Cấu trúc và đặc điểm của dữ liệu giao dịch tài chính

- **Quy mô lớn và phát sinh liên tục**

Dữ liệu giao dịch tài chính thường có quy mô rất lớn, được sinh ra liên tục theo thời gian và yêu cầu xử lý gần thời gian thực. Trong các hệ thống thanh toán điện tử hoặc ngân hàng số, mỗi ngày có thể phát sinh hàng triệu đến hàng chục triệu giao dịch. Điều này đặt ra yêu cầu cao về hiệu năng xử lý và khả năng mở rộng đối với các hệ thống giám sát gian lận.

Tính chất phát sinh liên tục của dữ liệu cũng khiến việc lưu trữ, cập nhật mô hình và đánh

giá hiệu quả trở nên phức tạp hơn so với các bài toán học máy truyền thống trên dữ liệu tĩnh.

- **Mất cân bằng dữ liệu nghiêm trọng**

Một trong những thách thức lớn nhất của dữ liệu gian lận là sự mất cân bằng cực lớn giữa các lớp. Trong hầu hết các tập dữ liệu giao dịch thực tế, số lượng giao dịch gian lận chỉ chiếm một tỷ lệ rất nhỏ so với tổng số giao dịch hợp lệ, thường dưới 1% và trong nhiều trường hợp dưới 0,1%.

Sự mất cân bằng này khiến nhiều mô hình học máy có xu hướng thiên lệch về lớp đa số, dẫn đến việc bỏ sót các trường hợp gian lận hiếm gặp. Đồng thời, các thước đo đánh giá tổng quát như accuracy hoặc ROC-AUC có thể không phản ánh đúng hiệu quả phát hiện gian lận trong thực tế vận hành.

- **Phụ thuộc quan hệ và theo chuỗi**

Giao dịch tài chính không tồn tại độc lập mà thường có mối liên hệ với các giao dịch khác thông qua tài khoản gửi, tài khoản nhận hoặc thông qua chuỗi hành vi theo thời gian. Các mẫu gian lận phức tạp thường xuất hiện dưới dạng chuỗi giao dịch liên tiếp, sự phối hợp giữa nhiều tài khoản hoặc các mạng lưới trung gian.

Đặc điểm này cho thấy dữ liệu giao dịch mang bản chất quan hệ (relational) và theo chuỗi (temporal), khiến các mô hình giả định tính độc lập giữa các mẫu dữ liệu trở nên kém hiệu quả trong nhiều trường hợp.

1.3. Cấu trúc của một giao dịch tài chính

Một giao dịch điển hình

Một giao dịch tài chính điển hình trong hệ thống thanh toán điện tử thường bao gồm các thành phần cơ bản sau:

- Tài khoản gửi và tài khoản nhận
- Giá trị giao dịch
- Thời điểm phát sinh giao dịch
- Loại giao dịch (chuyển tiền, thanh toán, rút tiền, v.v.)
- Trạng thái giao dịch và các thông tin liên quan đến số dư trước và sau giao dịch

Các thành phần này tạo thành cơ sở dữ liệu ban đầu để trích xuất đặc trưng phục vụ cho các mô hình phát hiện gian lận.

Ngữ cảnh giao dịch theo tài khoản và theo thời gian

Ngoài các thuộc tính cục bộ, giao dịch còn mang thông tin ngữ cảnh quan trọng khi được xem xét trong lịch sử hoạt động của tài khoản. Các đặc trưng như tần suất giao dịch, tổng giá trị giao dịch trong một khoảng thời gian, sự thay đổi hành vi theo thời gian hoặc mối quan hệ giữa các tài khoản đóng vai trò quan trọng trong việc nhận diện các hành vi bất thường.

Do đó, việc khai thác thông tin ngữ cảnh và quan hệ là yếu tố then chốt trong các hệ thống phát hiện gian lận hiện đại.

1.4. Hoạt động bất thường trong giao dịch tài chính

Khái niệm hoạt động bất thường

Hoạt động bất thường được hiểu là các giao dịch hoặc chuỗi giao dịch có đặc điểm lệch đáng kể so với mô hình hành vi thông thường của hệ thống. Cần lưu ý rằng không phải mọi hoạt động bất thường đều là gian lận; tuy nhiên, các giao dịch gian lận thường nằm trong tập con của các hoạt động bất thường có tính chủ ý.

Hoạt động bất thường trong giao dịch tài chính có thể được phân loại thành nhiều dạng:

- **Bất thường điểm**: giao dịch riêng lẻ có giá trị hoặc thời điểm bất thường
- **Bất thường theo ngữ cảnh**: giao dịch chỉ trở nên bất thường trong bối cảnh cụ thể
- **Bất thường tập thể**: một nhóm giao dịch phối hợp tạo thành hành vi bất thường
- **Thay đổi chế độ**: sự thay đổi kéo dài trong hành vi giao dịch của tài khoản

Các hình thức gian lận giao dịch phổ biến

Trong thực tế, gian lận giao dịch tài chính có thể xuất hiện dưới nhiều hình thức như:

- Gian lận chuyển tiền thông qua các tài khoản trung gian
- Chia nhỏ giao dịch nhằm tránh các ngưỡng giám sát
- Lợi dụng các khung thời gian ít được giám sát
- Thao túng hành vi giao dịch để đánh lạc hướng hệ thống phát hiện

Các hình thức này thường để lại dấu vết dưới dạng chuỗi giao dịch liên quan, mức độ giao dịch tăng đột biến hoặc các mối quan hệ bất thường giữa các tài khoản.

1.5. Các phương pháp phát hiện gian lận giao dịch

Cách tiếp cận dựa trên học máy truyền thống

Các phương pháp học máy truyền thống là một trong những hướng tiếp cận sớm và phổ biến nhất trong phát hiện gian lận giao dịch. Các mô hình như Logistic Regression, Decision Tree, Random Forest hoặc Gradient Boosting thường được huấn luyện trên dữ liệu dạng bảng với các đặc trưng thủ công.

Ưu điểm của các phương pháp này là dễ triển khai, khả năng giải thích tương đối tốt và hiệu quả ổn định trong nhiều bối cảnh. Tuy nhiên, hạn chế lớn là khả năng khai thác các mối quan hệ phức tạp giữa các giao dịch và tài khoản còn hạn chế.

Cách tiếp cận dựa trên học sâu và học biểu diễn

Học sâu cho phép mô hình tự động học các biểu diễn phi tuyến từ dữ liệu giao dịch, giảm phụ thuộc vào thiết kế đặc trưng thủ công. Các kiến trúc như MLP, RNN hoặc Transformer đã được áp dụng để mô hình hóa hành vi giao dịch và các mẫu theo thời gian.

Tuy nhiên, các mô hình học sâu thuần túy vẫn gặp khó khăn khi dữ liệu mất cân bằng nghiêm trọng hoặc khi các mối quan hệ giữa giao dịch không được biểu diễn tường minh.

Cách tiếp cận phát hiện bất thường

Một hướng tiếp cận khác là xem gian lận như một dạng bất thường hiếm gặp. Các phương pháp như Isolation Forest, One-Class SVM hoặc Autoencoder thường được sử dụng trong bối cảnh nhãn gian lận khan hiếm.

Ưu điểm của các phương pháp này là không yêu cầu nhiều dữ liệu có nhãn, nhưng chúng thường gặp thách thức về tỷ lệ cảnh báo sai cao và khó phân biệt giữa bất thường ngẫu nhiên và gian lận có chủ ý.

Cách tiếp cận khai thác cấu trúc quan hệ và học trên đồ thị

Gần đây, các phương pháp học trên đồ thị, đặc biệt là Graph Neural Network, đã thu hút nhiều sự quan tâm trong phát hiện gian lận giao dịch. Bằng cách biểu diễn dữ liệu giao dịch dưới dạng đồ thị, các mô hình này cho phép khai thác trực tiếp các mối quan hệ giữa các thực thể, làm nổi bật các mẫu gian lận mang tính liên kết.

Hướng tiếp cận này được xem là phù hợp với bản chất quan hệ của dữ liệu giao dịch tài

chính, tuy nhiên việc xây dựng đồ thị phù hợp và xử lý các vấn đề như dữ liệu thưa và mất cân bằng vẫn là những thách thức cần được giải quyết.

1.6. Tổng quan các công trình liên quan

Các công trình nghiên cứu về phát hiện gian lận giao dịch tài chính có thể được phân nhóm thành ba hướng chính: (i) học máy truyền thống trên dữ liệu bảng, (ii) các phương pháp dựa trên học biểu diễn, và (iii) các mô hình khai thác cấu trúc quan hệ và đồ thị.

Nhiều nghiên cứu đã chỉ ra rằng các mô hình ensemble dựa trên cây quyết định thường đạt hiệu quả ổn định trong bối cảnh dữ liệu mất cân bằng. Các công trình khác cho thấy việc sử dụng mạng nơ-ron giúp cải thiện khả năng học các quan hệ phi tuyến, nhưng vẫn gặp hạn chế khi không khai thác được cấu trúc quan hệ giữa các giao dịch.

Song song đó, các nghiên cứu gần đây nhấn mạnh vai trò của embedding và graph-based learning trong việc phát hiện các hành vi gian lận mang tính liên kết. Tuy nhiên, phần lớn các công trình tập trung vào từng thành phần riêng lẻ, và chưa có nhiều nghiên cứu kết hợp một cách hệ thống các hướng tiếp cận này trong một pipeline thống nhất, đặc biệt trong bối cảnh dữ liệu có tỷ lệ gian lận cực kỳ thấp.

Kết luận phần cơ sở lý thuyết

Phần cơ sở lý thuyết đã trình bày bối cảnh và đặc điểm của bài toán phát hiện gian lận giao dịch tài chính, làm rõ cấu trúc và tính chất của dữ liệu giao dịch, cũng như các hướng tiếp cận phổ biến trong nghiên cứu hiện nay. Những phân tích này cho thấy rằng bài toán phát hiện gian lận là một bài toán phức tạp, đòi hỏi sự kết hợp của nhiều phương pháp nhằm khai thác đồng thời đặc trưng cục bộ, thông tin ngữ cảnh và cấu trúc quan hệ.

Trên nền tảng lý thuyết này, các chương tiếp theo sẽ tập trung vào việc thiết kế pipeline phát hiện gian lận cụ thể, triển khai thực nghiệm và đánh giá hiệu quả của các phương pháp được đề xuất.

2. PHƯƠNG PHÁP NGHIÊN CỨU VÀ THIẾT KẾ THỰC NGHIỆM

2.1. Giới thiệu chung

Trên cơ sở các khái niệm và hướng tiếp cận đã được trình bày ở Chương 1, chương này tập trung mô tả phương pháp nghiên cứu và thiết kế pipeline phát hiện giao dịch gian lận được sử dụng trong đề án. Trọng tâm của chương là chuyển hóa các nền tảng lý thuyết thành một quy

trình xử lý và mô hình hóa cụ thể, có thể triển khai, đánh giá và so sánh trên dữ liệu giao dịch tài chính quy mô lớn.

Trong bối cảnh dữ liệu giao dịch tài chính có cấu trúc phức tạp, tồn tại mối quan hệ giữa nhiều thực thể và đồng thời chịu sự mất cân bằng nghiêm trọng của nhãn gian lận, việc xây dựng một pipeline phát hiện gian lận không thể chỉ dừng lại ở lựa chọn một mô hình đơn lẻ. Thay vào đó, đề án tiếp cận bài toán theo hướng pipeline nhiều tầng, trong đó mỗi thành phần đảm nhiệm một vai trò riêng nhằm từng bước làm giàu thông tin, khai thác cấu trúc quan hệ và cải thiện khả năng phân biệt các giao dịch gian lận hiếm gặp.

Cụ thể, pipeline phát hiện gian lận trong đề án được thiết kế theo chuỗi xử lý logic bắt đầu từ dữ liệu giao dịch thô và kết thúc ở quyết định phân loại giao dịch. Pipeline bao gồm các giai đoạn chính: tiền xử lý và xây dựng đặc trưng từ dữ liệu giao dịch, tổ chức lại không gian dữ liệu thông qua gom cụm tài khoản, biểu diễn giao dịch dưới dạng đồ thị, học biểu diễn quan hệ bằng mô hình học sâu theo cấu trúc đồ thị, xử lý mất cân bằng dữ liệu trong không gian embedding và cuối cùng là phân loại giao dịch gian lận bằng mô hình học máy truyền thống. Cách tổ chức này cho phép kết hợp ưu điểm của nhiều hướng tiếp cận khác nhau trong một khung thống nhất.

Khác với Chương 1 – nơi các phương pháp được trình bày ở mức khái niệm và tổng quan nhằm làm rõ bối cảnh nghiên cứu, Chương 2 đi sâu vào cách các phương pháp đó được tích hợp và triển khai cụ thể trong pipeline đề xuất. Các quyết định thiết kế, từ việc lựa chọn chiến lược tiền xử lý, cách xây dựng đồ thị giao dịch cho đến kiến trúc mô hình và chiến lược xử lý mất cân bằng dữ liệu, đều xuất phát từ đặc thù của dữ liệu giao dịch tài chính và mục tiêu phát hiện gian lận trong điều kiện vận hành thực tế.



Mục tiêu của chương này là:

- Trình bày rõ cấu trúc và luồng xử lý của pipeline phát hiện gian lận được đề xuất,
- Mô tả chi tiết các bước xử lý dữ liệu và mô hình hóa tương ứng với từng giai đoạn trong pipeline,
- Thiết lập một khung thực nghiệm nhất quán, làm cơ sở cho việc triển khai, đánh giá và so sánh hiệu quả của các pipeline ở Chương 3.

Trên cơ sở phần giới thiệu này, các mục tiếp theo của chương sẽ lần lượt trình bày chi tiết từng thành phần trong pipeline, bắt đầu từ khảo sát dữ liệu và tiền xử lý, sau đó là biểu diễn đồ thị, học biểu diễn và thiết kế thực nghiệm.

2.2. Dữ liệu nghiên cứu

Tập dữ liệu PaySim

Đề án sử dụng tập dữ liệu PaySim, là dữ liệu mô phỏng giao dịch tài chính dựa trên nhật ký của một hệ thống thanh toán di động. PaySim được xây dựng nhằm tái hiện các đặc trưng phổ biến của dữ liệu giao dịch trong môi trường tài chính thực tế, đồng thời cung cấp nhãn gian lận phục vụ nghiên cứu.

Tập dữ liệu bao gồm hơn **6,3 triệu giao dịch**, trong đó mỗi bản ghi tương ứng với một giao dịch và được gán nhãn gian lận hoặc không gian lận. Các thuộc tính chính của dữ liệu bao gồm tài khoản gửi, tài khoản nhận, giá trị giao dịch, loại giao dịch, thời điểm phát sinh và thông tin số dư trước – sau giao dịch.

Một đặc điểm quan trọng của PaySim là **tỷ lệ giao dịch gian lận nhỏ hơn 0,1%**, phản ánh mức độ mất cân bằng rất cao thường gặp trong các hệ thống tài chính thực tế. Chính đặc điểm này khiến PaySim trở thành bộ dữ liệu phù hợp để đánh giá hiệu quả của các phương pháp phát hiện gian lận trong bối cảnh khó khăn.

Đặc điểm dữ liệu và thách thức

Việc sử dụng PaySim đặt ra một số thách thức chính:

- Dữ liệu có **quy mô lớn**, yêu cầu xử lý hiệu quả về bộ nhớ và thời gian,
- Nhãn gian lận **cực kỳ hiếm**, khiến nhiều mô hình dễ bị thiên lệch,
- Mức độ liên kết giữa các tài khoản trong dữ liệu mô phỏng **thấp hơn dữ liệu thực**, đòi hỏi các bước xử lý bổ sung để khai thác cấu trúc quan hệ.

Những đặc điểm này trực tiếp ảnh hưởng đến cách thiết kế pipeline tiền xử lý và biểu

diễn dữ liệu trong đề án.

2.3. Tiền xử lý dữ liệu

Mục tiêu tiền xử lý

Mục tiêu của bước tiền xử lý là chuyển dữ liệu giao dịch thô thành dạng đầu vào phù hợp cho pipeline học biểu diễn và phân loại, đồng thời:

- Giảm nhiễu kỹ thuật và ảnh hưởng của ngoại lai,
- Hạn chế rò rỉ thông tin từ tương lai,
- Chuẩn hóa dữ liệu để các mô hình học máy và học sâu có thể hội tụ ổn định.

Làm sạch dữ liệu và xử lý giá trị thiếu

Các bản ghi thiếu thông tin cốt lõi phục vụ phân tích (ví dụ tài khoản gửi, tài khoản nhận hoặc giá trị giao dịch) được loại bỏ khỏi tập dữ liệu. Đối với các thuộc tính số còn lại, giá trị thiếu được thay thế bằng **giá trị trung vị (median)** của thuộc tính tương ứng nhằm giảm ảnh hưởng của các phân bố lệch và ngoại lai, vốn rất phổ biến trong dữ liệu tài chính.

Chuẩn hóa thang đo và xử lý phân bố lệch

Dữ liệu giao dịch tài chính thường có phân bố không đối xứng và chứa các giá trị ngoại lai lớn. Để giảm ảnh hưởng của hiện tượng này:

- Các thuộc tính có phân bố lệch mạnh (ví dụ giá trị giao dịch) được áp dụng **biến đổi log**
- Các thuộc tính số sau đó được chuẩn hóa về cùng thang đo để hỗ trợ quá trình huấn luyện mô hình.

Các bước này giúp tránh việc một số đặc trưng chi phối hoàn toàn quá trình học của mô hình.

Xây dựng đặc trưng thời gian

Thời gian phát sinh giao dịch là nguồn thông tin quan trọng trong phát hiện gian lận. Đề án xây dựng các đặc trưng thời gian nhằm mô hình hóa nhịp độ và tính chu kỳ của giao dịch, bao gồm các biểu diễn chuẩn hóa theo trục thời gian và các đặc trưng mã hóa chu kỳ. Các đặc trưng này giúp mô hình học được các mẫu hành vi gắn với thời điểm phát sinh giao dịch.

Phân chia dữ liệu theo thứ tự thời gian

Sau khi hoàn tất tiền xử lý, dữ liệu được phân chia theo thứ tự thời gian thành ba tập:

- **Tập huấn luyện (70%)**: các giao dịch sớm nhất,
- **Tập kiểm định (15%)**: các giao dịch tiếp theo,
- **Tập kiểm tra (15%)**: các giao dịch muộn nhất.

Cách phân chia này phản ánh đúng kịch bản triển khai thực tế, trong đó mô hình chỉ có thể học từ dữ liệu quá khứ để dự đoán các giao dịch phát sinh trong tương lai.

2.4. Biểu diễn dữ liệu và xây dựng đồ thị

Động cơ sử dụng gom cụm tài khoản

Một thách thức lớn khi áp dụng các mô hình học trên đồ thị trực tiếp lên dữ liệu PaySim là **đồ thị tài khoản nguyên bản rất rời rạc**, do số lượng tài khoản lớn và nhiều tài khoản chỉ xuất hiện với tần suất thấp. Đồ thị thưa làm suy giảm hiệu quả lan truyền thông tin của các mô hình GNN.

Để khắc phục hạn chế này, đề án áp dụng bước **gom cụm tài khoản (account clustering)** dựa trên các đặc trưng hành vi tổng hợp. Cách tiếp cận này giúp giảm số lượng node trong đồ thị, tăng mật độ kết nối và làm nổi bật các mẫu hành vi ở mức nhóm thay vì mức cá thể.

Xây dựng đồ thị giao dịch theo cụm

Sau bước gom cụm, dữ liệu giao dịch được biểu diễn dưới dạng **đồ thị có hướng**, trong đó:

- Mỗi node đại diện cho một cụm tài khoản,
- Mỗi cạnh đại diện cho một giao dịch giữa cụm nguồn và cụm đích.

Node mang các đặc trưng tổng hợp mô tả hành vi giao dịch của cụm, trong khi cạnh mang thông tin liên quan đến giao dịch. Biểu diễn này cho phép mô hình hóa trực tiếp các mối quan hệ giao dịch giữa các nhóm tài khoản.

2.5. Học biểu diễn và mô hình phân loại

Học embedding bằng mạng nơ-ron đa tầng

Trước khi áp dụng mô hình đồ thị, các đặc trưng node được mã hóa thông qua một mạng nơ-ron đa tầng nhằm ánh xạ dữ liệu bảng sang không gian biểu diễn liên tục. Bước này đóng vai trò tiền xử lý học được, giúp mô hình tự động trích xuất các tổ hợp phi tuyến giữa các đặc trưng hành vi.

Học theo cấu trúc đồ thị

Trên đồ thị đã xây dựng, mô hình học theo cấu trúc đồ thị được sử dụng để lan truyền và tổng hợp thông tin giữa các node, từ đó học được biểu diễn quan hệ giữa các cụm tài khoản. Việc học biểu diễn đồ thị cho phép mô hình khai thác các mối liên hệ mà dữ liệu bảng truyền thống khó thể hiện.

Xử lý mất cân bằng và phân loại

Do tỷ lệ gian lận cực thấp, đề án triển khai hai hướng thực nghiệm song song:

- Pipeline không áp dụng kỹ thuật cân bằng dữ liệu,
- Pipeline có áp dụng kỹ thuật cân bằng dữ liệu trên không gian embedding.

Bộ phân loại được huấn luyện trên embedding giao dịch nhằm phát hiện các giao dịch gian lận hiếm gặp.

2.6. Thiết kế thực nghiệm và đánh giá

Các pipeline được huấn luyện và đánh giá theo cùng một chiến lược nhằm đảm bảo tính công bằng khi so sánh. Các thước đo đánh giá tập trung vào khả năng phát hiện lớp gian lận hiếm, bao gồm PR-AUC, F1-score, Precision và Recall, bên cạnh các thước đo tổng quát như ROC-AUC.

Kết luận chương

Chương này đã trình bày phương pháp nghiên cứu và thiết kế thực nghiệm của đề án, từ lựa chọn dữ liệu, tiền xử lý, biểu diễn đồ thị đến kiến trúc mô hình và chiến lược đánh giá. Những nội dung này tạo nền tảng trực tiếp cho Chương 3, nơi các kết quả thực nghiệm sẽ được trình bày và phân tích chi tiết.

3. TRIỂN KHAI THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

3.1. Giới thiệu chung

Chương này trình bày quá trình triển khai thực nghiệm nhằm đánh giá hiệu quả của pipeline phát hiện giao dịch gian lận tài chính được đề xuất ở Chương 2. Trọng tâm của chương là kiểm chứng vai trò của **biểu diễn đồ thị** và **kỹ thuật xử lý mất cân bằng dữ liệu** trong bối cảnh dữ liệu giao dịch có tỷ lệ gian lận cực thấp.

Các thí nghiệm được thiết kế xoay quanh **hai pipeline có cùng kiến trúc học biểu diễn**, nhưng khác nhau ở bước xử lý mất cân bằng dữ liệu, bao gồm:

- **Pipeline baseline**: NodeMLP → GAT → Random Forest, huấn luyện trực tiếp trên dữ liệu gốc không cân bằng.
- **Pipeline chính**: NodeMLP → GAT → Random Forest kết hợp với SMOTE trên không gian embedding của giao dịch.

Thông qua việc so sánh hai pipeline này, chương nhằm làm rõ:

- Sự khác biệt giữa các thước đo tổng quát (ROC-AUC) và các thước đo tập trung vào lớp gian lận hiếm (PR-AUC, F1-score).
- Tác động thực tế của kỹ thuật cân bằng dữ liệu khi kết hợp với biểu diễn đồ thị trong bài toán phát hiện gian lận.

3.2. Môi trường và công cụ thực nghiệm

Toàn bộ quá trình tiền xử lý dữ liệu, xây dựng đồ thị, huấn luyện và đánh giá mô hình trong đề án được triển khai trên nền tảng **Google Colaboratory (Google Colab)**. Việc sử dụng Colab giúp chuẩn hóa môi trường thực nghiệm, giảm phụ thuộc vào cấu hình phần cứng cục bộ và hỗ trợ khả năng tái lập kết quả.

Các thí nghiệm được thực hiện trong môi trường Python, tận dụng hệ sinh thái thư viện mã nguồn mở cho học máy và học sâu. Đối với các bước huấn luyện mô hình học sâu như NodeMLP và GAT, đề án sử dụng GPU do Colab cung cấp nhằm tăng tốc tính toán và rút ngắn thời gian huấn luyện.

Các thư viện chính được sử dụng bao gồm:

- **NumPy, Pandas** cho xử lý và phân tích dữ liệu,
- **Scikit-learn** cho các bước tiền xử lý và mô hình học máy truyền thống,
- **PyTorch và PyTorch Geometric** cho xây dựng và huấn luyện mô hình đồ thị,
- **Imbalanced-learn** cho kỹ thuật cân bằng dữ liệu SMOTE,
- **Matplotlib** cho trực quan hóa kết quả thực nghiệm.

Việc triển khai trên Google Colab phản ánh đúng bối cảnh nghiên cứu học thuật, trong đó trọng tâm đặt vào đánh giá phương pháp và phân tích mô hình, thay vì tối ưu hạ tầng triển khai.

3.3. Thiết lập thực nghiệm

Dữ liệu và phân chia tập

Thực nghiệm được tiến hành trên tập dữ liệu PaySim, bao gồm hơn 6,3 triệu giao dịch, trong đó tỷ lệ giao dịch gian lận nhỏ hơn 0,1%. Sau bước tiền xử lý và xây dựng đặc trưng (trình bày tại Chương 2), dữ liệu được phân chia theo thứ tự thời gian thành ba tập:

- Tập huấn luyện (Train): 70% dữ liệu sớm nhất,
- Tập kiểm định (Validation): 15% dữ liệu tiếp theo,
- Tập kiểm tra (Test): 15% dữ liệu muộn nhất.

Cách phân chia này bảo đảm rằng mô hình chỉ được huấn luyện trên dữ liệu quá khứ và được đánh giá trên các giao dịch phát sinh sau đó, tránh rò rỉ thông tin và phản ánh đúng kịch bản triển khai thực tế.

Chiến lược huấn luyện và chọn ngưỡng

Hai pipeline được huấn luyện theo cùng một chiến lược để công bằng khi so sánh:

- Kiến trúc học biểu diễn (NodeMLP và GAT) được giữ cố định.
- Các siêu tham số của Random Forest được giữ cố định giữa hai pipeline.
- Ngưỡng phân loại được xác định dựa trên **F1-score tối đa trên tập kiểm định**.

Ngưỡng này sau đó được áp dụng trực tiếp lên tập kiểm tra, vốn chưa từng được mô hình quan sát trong quá trình huấn luyện hoặc điều chỉnh.

3.4. Kết quả thực nghiệm và phân tích

Pipeline không áp dụng SMOTE (Baseline)

Pipeline baseline được thiết kế để phản ánh hiệu quả của mô hình khi **không sử dụng kỹ thuật xử lý mất cân bằng dữ liệu**. Kết quả trên tập kiểm tra cho thấy pipeline baseline đạt **ROC-AUC tương đối cao**, phản ánh khả năng phân biệt tổng quát giữa hai lớp giao dịch gian lận và không gian lận.

Tuy nhiên, các thước đo tập trung vào lớp thiểu số cho thấy hiệu năng rất thấp do số lượng giao dịch gian lận chiếm tỷ lệ cực hiếm trong tập dữ liệu. Mô hình có xu hướng ưu tiên dự đoán lớp không gian lận, dẫn đến khả năng phát hiện gian lận hiếm gặp gần như không đáng kể.

Kết quả định lượng:

- ROC-AUC ≈ 0.82–0.86
- PR-AUC gần bằng 0
- F1-score ≈ 0.00

- Recall gần như bằng 0 tại ngưỡng phân loại 0.5

Kết quả này minh họa rõ hạn chế của việc chỉ sử dụng các thước đo tổng quát trong bối cảnh dữ liệu cực kỳ mất cân bằng.

Pipeline áp dụng SMOTE (Nhánh chính)

Khi áp dụng kỹ thuật SMOTE trên không gian embedding của giao dịch, số lượng các mẫu gian lận trong tập huấn luyện được mở rộng đáng kể, giúp bộ phân loại Random Forest học được ranh giới phân lớp hiệu quả hơn.

Kết quả trên tập kiểm tra cho thấy sự tăng trưởng rõ rệt ở các thước đo đánh giá tập trung vào lớp gian lận hiếm gặp:

- ROC-AUC = 0.8117
- PR-AUC = 0.1670
- F1-score = 0.2054
- Precision = 0.2428
- Recall = 0.1780

Các kết quả này cho thấy pipeline có áp dụng SMOTE không chỉ cải thiện khả năng phát hiện gian lận mà còn giữ cho tỷ lệ cảnh báo sai ở mức chấp nhận được. Mô hình có khả năng nhận diện đáng kể số lượng giao dịch gian lận hiếm gặp, trong khi số lượng false positive vẫn được kiểm soát (237 trên hơn 100 nghìn giao dịch).

So với pipeline baseline, **PR-AUC tăng từ mức gần như không đáng kể lên 0.1670**, chứng minh vai trò rõ rệt của kỹ thuật xử lý mất cân bằng dữ liệu trong bối cảnh này.

So sánh hai pipeline

Bảng 1: Bảng so sánh hai pipeline

Chỉ số	Không SMOTE	Có SMOTE
ROC-AUC	0.82–0.86	0.8117
PR-AUC	< 0.03	0.1670
F1	0.00	0.2054

So sánh trực tiếp hai pipeline cho thấy việc sử dụng kỹ thuật cân bằng dữ liệu đã cải thiện đáng kể các thước đo quan trọng trong bài toán gian lận có tỷ lệ dương tính rất thấp. Mặc dù hai pipeline có ROC-AUC gần tương đương, sự khác biệt thể hiện rõ ở PR-AUC và F1-score – những thước đo phản ánh trực tiếp hiệu quả phát hiện gian lận trong thực tế.

Kết quả này khẳng định rằng ROC-AUC không phải là thước đo duy nhất phản ánh chất lượng mô hình trong bối cảnh dữ liệu mất cân bằng nghiêm trọng, và các thước đo tập trung vào lớp thiểu số cần được xem xét đồng thời trong quá trình đánh giá.

3.5. So sánh với các nghiên cứu liên quan

Dựa trên kết quả thực nghiệm, pipeline đề xuất trong đề án cho thấy hiệu năng cạnh tranh so với nhiều nghiên cứu gần đây trong lĩnh vực phát hiện gian lận và bất thường.

So với các phương pháp học máy truyền thống trên dữ liệu bảng, pipeline dựa trên biểu diễn đồ thị cho thấy ưu thế rõ rệt trong việc khai thác quan hệ giữa các thực thể giao dịch. So với các hướng tiếp cận dựa trên embedding hoặc mô hình đồ thị kết hợp mô hình ngôn ngữ lớn, pipeline của đề án đạt hiệu quả cao trong bối cảnh dữ liệu cực kỳ mất cân bằng mà không cần đến các kiến trúc phức tạp hơn.

Những so sánh này cho thấy hướng tiếp cận kết hợp biểu diễn đồ thị và xử lý mất cân bằng dữ liệu là phù hợp và có giá trị thực tiễn đối với bài toán phát hiện gian lận giao dịch tài chính.

3.6. Kết luận chương

Chương này đã trình bày quá trình triển khai thực nghiệm và đánh giá hiệu quả của pipeline phát hiện giao dịch gian lận tài chính dựa trên học theo cấu trúc đồ thị. Thông qua việc so sánh hai pipeline có và không áp dụng kỹ thuật cân bằng dữ liệu, chương đã làm rõ vai trò của xử lý mất cân bằng trong việc nâng cao hiệu quả phát hiện gian lận hiếm gặp.

Kết quả thực nghiệm cho thấy việc kết hợp học biểu diễn bằng GAT với kỹ thuật SMOTE giúp cải thiện đáng kể các thước đo quan trọng như PR-AUC và F1-score, trong khi các thước đo tổng quát như ROC-AUC chưa phản ánh đầy đủ hiệu quả phát hiện gian lận trong bối cảnh dữ liệu cực lệch.

Những phân tích trong chương này tạo nền tảng cho chương kết luận, nơi các kết quả đạt được sẽ được tổng hợp, đánh giá hạn chế và đề xuất các hướng mở rộng cho nghiên cứu trong tương lai.

4. KẾT LUẬN VÀ KIẾN NGHỊ

4.1. Kết luận

Đề án đã nghiên cứu và triển khai một pipeline phát hiện giao dịch gian lận tài chính dựa

trên học theo cấu trúc đồ thị, trong bối cảnh dữ liệu giao dịch có quy mô lớn và tỷ lệ gian lận cực kỳ thấp. Trên cơ sở tổng hợp các hướng tiếp cận hiện đại trong phát hiện gian lận, đề án tập trung khai thác mối quan hệ giữa các thực thể giao dịch thông qua biểu diễn đồ thị, đồng thời kết hợp các kỹ thuật học biểu diễn và xử lý mất cân bằng dữ liệu nhằm nâng cao hiệu quả phát hiện các trường hợp gian lận hiếm gặp.

Về mặt phương pháp, đề án đã xây dựng một pipeline nhiều tầng bao gồm: tiền xử lý dữ liệu giao dịch, gom cụm tài khoản để tăng tính kết nối của đồ thị, học embedding phi tuyến bằng mạng nơ-ron đa tầng, mô hình hóa quan hệ giữa các cụm tài khoản bằng Graph Attention Network, xử lý mất cân bằng dữ liệu trong không gian embedding bằng SMOTE và phân loại giao dịch bằng Random Forest. Cách tiếp cận này cho phép khai thác đồng thời thông tin cục bộ của giao dịch và thông tin quan hệ giữa các thực thể, vượt qua hạn chế của các phương pháp chỉ dựa trên dữ liệu bảng truyền thống.

Kết quả thực nghiệm trên tập dữ liệu PaySim cho thấy sự khác biệt rõ rệt giữa hai pipeline có và không áp dụng kỹ thuật cân bằng dữ liệu. Pipeline baseline, mặc dù đạt ROC-AUC tương đối cao, gần như không phát hiện được các giao dịch gian lận hiếm gặp khi xét theo các thước đo tập trung vào lớp thiểu số như PR-AUC và F1-score. Ngược lại, pipeline có áp dụng SMOTE trong không gian embedding đã cải thiện đáng kể khả năng phát hiện gian lận, thể hiện qua PR-AUC và F1-score tăng rõ rệt, trong khi vẫn giữ được số lượng cảnh báo sai ở mức chấp nhận được.

Các kết quả này khẳng định rằng trong bối cảnh dữ liệu gian lận có tỷ lệ dương tính rất thấp, việc chỉ dựa vào các thước đo tổng quát như ROC-AUC là không đủ để đánh giá hiệu quả của mô hình. Việc kết hợp biểu diễn đồ thị với kỹ thuật xử lý mất cân bằng dữ liệu đóng vai trò quan trọng trong việc làm rõ ranh giới phân lớp và nâng cao khả năng phát hiện các hành vi gian lận hiếm gặp trong thực tế vận hành.

Tổng thể, đề án đã chứng minh tính khả thi và hiệu quả của hướng tiếp cận dựa trên học theo cấu trúc đồ thị kết hợp với xử lý mất cân bằng dữ liệu trong bài toán phát hiện giao dịch gian lận tài chính. Pipeline đề xuất không chỉ có giá trị về mặt học thuật, mà còn mang tiềm năng ứng dụng thực tiễn trong các hệ thống giám sát giao dịch tài chính, nơi yêu cầu phát hiện sớm gian lận đi kèm với kiểm soát chặt chẽ tỷ lệ cảnh báo sai.

4.2. Kiến nghị và hướng phát triển

Bên cạnh các kết quả đạt được, đề án vẫn tồn tại một số hạn chế và mở ra các hướng phát triển tiềm năng trong tương lai.

Thứ nhất, dữ liệu sử dụng trong đề án là dữ liệu mô phỏng (PaySim), do đó mức độ liên kết và đa dạng hành vi chưa phản ánh đầy đủ độ phức tạp của dữ liệu giao dịch trong các hệ thống tài chính thực tế. Trong các nghiên cứu tiếp theo, việc áp dụng pipeline đề xuất trên dữ liệu thực hoặc dữ liệu có cấu trúc quan hệ phong phú hơn sẽ giúp đánh giá toàn diện hơn hiệu quả của phương pháp.

Thứ hai, bước gom cụm tài khoản trong đề án được thực hiện với số cụm cố định và dựa trên các đặc trưng hành vi tổng hợp. Các nghiên cứu tiếp theo có thể xem xét các chiến lược gom cụm thích nghi hơn, hoặc các phương pháp xây dựng đồ thị khác nhằm phản ánh tốt hơn cấu trúc quan hệ động trong dữ liệu giao dịch.

Thứ ba, pipeline hiện tại sử dụng SMOTE để xử lý mất cân bằng dữ liệu trong không gian embedding. Mặc dù mang lại hiệu quả rõ rệt, việc lựa chọn tỷ lệ oversampling và chiến lược sinh mẫu tổng hợp vẫn có thể ảnh hưởng đến chất lượng mô hình. Các hướng mở rộng có thể bao gồm việc nghiên cứu các kỹ thuật cân bằng dữ liệu khác, hoặc các phương pháp học chi phí (cost-sensitive learning) nhằm giảm phụ thuộc vào oversampling.

Thứ tư, trong bối cảnh triển khai thực tế, yêu cầu về minh bạch và khả năng giải thích mô hình ngày càng trở nên quan trọng. Do đó, việc tích hợp các kỹ thuật giải thích mô hình (Explainable AI) để phân tích vai trò của từng đặc trưng, từng cụm tài khoản hoặc từng mối quan hệ trong quá trình dự đoán là một hướng phát triển cần được quan tâm.

Cuối cùng, pipeline đề xuất có thể được mở rộng sang các bài toán liên quan như giám sát giao dịch trong chống rửa tiền (AML), phát hiện gian lận trong thanh toán điện tử hoặc các hệ thống tài chính phi tập trung. Với cấu trúc mô-đun và khả năng khai thác quan hệ, hướng tiếp cận dựa trên học theo cấu trúc đồ thị hứa hẹn mang lại nhiều giá trị trong các hệ thống giám sát tài chính hiện đại.