HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

TRỊNH ĐÌNH MINH

HỆ THỐNG PHÁT HIỆN GIAN LẬN TÍN DỤNG THỜI GIAN THỰC VỚI MÔ HÌNH CẬP NHẬT LIÊN TỤC

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

HÀ NỘI – 2026

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

TRỊNH ĐÌNH MINH

HỆ THỐNG PHÁT HIỆN GIAN LẬN TÍN DỤNG THỜI GIAN THỰC VỚI MÔ HÌNH CẬP NHẬT LIÊN TỤC

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC:

TS. PHẠM HOÀNG DUY

HÀ NỘI – 2026

LỜI CAM ĐOAN

Tôi là Trịnh Đình Minh, học viên cao học ngành Hệ thống thông tin. Tôi xin cam đoan đề án "Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục" là do tôi nghiên cứu, tìm hiểu và phát triển dưới sự hướng dẫn của TS. Phạm Hoàng Duy, không phải sự sao chép từ các tài liệu, công trình nghiên cứu của người khác mà không ghi rõ trong tài liệu tham khảo. Tôi xin chịu trách nhiệm về lời cam đoan này.

Hà Nội, ngày 10 tháng 07 năm 2026

Học viên



Trịnh Đình Minh

MỤC LỤC

Lời cam đoan i
Danh mục các thuật ngữ, chữ viết tắt v
Danh mục các ký hiệu toán học vi
Danh sách bảng viii
Danh sách hình vẽ ix
MỞ ĐẦU 1
CHƯƠNG 1 - TỔNG QUAN VỀ PHÁT HIỆN BẤT THƯỜNG GIAO DỊCH TÍN DỤNG 3
1.1. Bài toán phát hiện giao dịch tín dụng bất thường 3
1.2. Các đặc trưng dữ liệu giao dịch tín dụng 4
1.2.1. Phân loại đặc trưng 4
1.2.2. Đặc điểm thống kê 5
1.3. Hệ thống phát hiện bất thường giao dịch tín dụng 6
1.3.1. Vai trò của hệ thống 6
1.3.2. Yêu cầu của hệ thống 6
1.3.3. Mô hình tổng quát 7
1.4. Các phương pháp phát hiện bất thường 7
1.4.1. Phương pháp dựa trên luật 8
1.4.2. Phương pháp học có giám sát 8
1.4.3. Phương pháp học không giám sát 10
1.4.4. Phát hiện chuỗi sự kiện bất thường 10
1.5. Các kỹ thuật xử lý luồng dữ liệu 11
1.6. Mô hình đề xuất 13
1.6.1. Lựa chọn mô hình học máy 13
1.6.2. Xử lý dữ liệu luồng thời gian thực 14
1.6.3. Xử lý sự kiện phức tạp (CEP) 14
1.6.4. Cơ chế cập nhật mô hình động 15
1.7. Kết luận chương 15

CHƯƠNG 2 - TRIỂN KHAI HỆ THỐNG VÀ CÁC KỸ THUẬT ĐƯỢC ÁP DỤNG 17
2.1. Kiến trúc tổng thể của hệ thống 17
2.1.1. Tầng thu thập dữ liệu 18
2.1.2. Tầng xử lý luồng và trích xuất đặc trưng 19
2.1.3. Tầng phân tích và chấm điểm rủi ro 19
2.1.4. Tầng ra quyết định và phân phối kết quả 19
2.1.5. Tầng cập nhật mô hình 19
2.2. Mô hình học máy có giám sát 20
2.3. Xử lý dữ liệu luồng 21
2.3.1. Apache Kafka 21
2.3.2. Apache Flink 22
2.4. Xử lý sự kiện phức tạp 23
2.4.1. Apache Flink CEP 23
2.4.2. Tích hợp CEP vào hệ thống xử lý dữ liệu 27
2.5. Cơ chế cập nhật mô hình động 28
2.6. Hậu xử lý của hệ thống 29
2.6.1. Tổng hợp điểm rủi ro và ra quyết định 29
2.6.2. Hành động ứng phó, cảnh báo và giám sát 30
2.7. Kết luận chương 32
CHƯƠNG 3 - THỬ NGHIỆM, ĐÁNH GIÁ VÀ KẾT LUẬN 33
3.1. Giới thiệu thử nghiệm 33
3.1.1. Tập dữ liệu 34
3.1.2. Môi trường thử nghiệm 35
3.1.3. Thiết lập chung và độ đo đánh giá 36
3.2. Kịch bản 1 – Năng lực phát hiện gian lận 37
3.2.1. Thiết lập thử nghiệm 37
3.2.2. Kết quả và nhận xét 42
3.3. Kịch bản 2 – Vận hành thời gian thực và cập nhật động 46

3.3.1. Thiết lập thử nghiệm 46
3.3.2. Kết quả và nhận xét 49
3.4. Kết luận chương 50
KẾT LUẬN VÀ KIẾN NGHỊ 52
DANH MỤC TÀI LIỆU THAM KHẢO 53

DANH MỤC CÁC THUẬT NGỮ, CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
ATM	Automated Teller Machine	Máy rút tiền tự động
AUC	Area Under the Curve	Diện tích dưới đường cong
CEP	Complex Event Processing	Xử lý sự kiện phức tạp
F1	F1 Score	Điểm F1
LSTM	Long Short-Term Memory	Bộ nhớ ngắn hạn – dài hạn
MCC	Matthews Correlation Coefficient	Hệ số tương quan Matthews
ML	Machine Learning	Học máy
POS	Point of Sale	Điểm bán hàng
PR-AUC	Area Under the Precision-Recall Curve	Diện tích dưới đường cong Precision–Recall
ROC	Receiver Operating Characteristic	Đường cong đặc trưng hoạt động của bộ phân loại
ROC-AUC	Area Under the ROC Curve	Diện tích dưới đường cong ROC

DANH MỤC CÁC KÝ HIỆU TOÁN HỌC

Ký hiệu	Ý nghĩa
$\text{Risk}(x)$	Điểm rủi ro tổng hợp của giao dịch x
$P(y = 1 \mid x)$	Xác suất giao dịch x là gian lận do mô hình học máy ước lượng
$S_{CEP}(x)$	Điểm rủi ro tổng hợp từ xử lý sự kiện phức tạp (CEP)
α, β	Trọng số của thành phần học máy và CEP trong công thức Risk
$\theta_{low}, \theta_{high}$	Ngưỡng rủi ro thấp và cao để quyết định hành động
$\text{Velocity}(x)$	Điểm mẫu vận tốc giao dịch (giao dịch dồn dập trong cửa sổ ngắn)
$\text{Frequency}(x)$	Điểm mẫu tần suất giao dịch trong cửa sổ trung hạn
$\text{Spike}(x)$	Điểm mẫu đột biến giá trị giao dịch
T_c	Tập mốc thời gian các giao dịch của thẻ c
t_{now}	Thời điểm hiện tại
t_i	Mốc thời gian của giao dịch thứ i
W_v, W_f	Độ rộng cửa sổ thời gian của mẫu vận tốc và mẫu tần suất
N_v, N_f	Ngưỡng số giao dịch của mẫu vận tốc và mẫu tần suất
a	Giá trị (số tiền) của giao dịch
μ_c, σ_c	Trung bình và độ lệch chuẩn giá trị giao dịch trong lịch sử thẻ c
z	Điểm chuẩn hóa (z-score) của giá trị giao dịch
$\sigma(z)$	Hàm sigmoid ánh xạ điểm z về khoảng [0, 1]
ε	Hằng số dương rất nhỏ để tránh chia cho 0
TP, TN, FP, FN	Số dương đúng, âm đúng, dương sai, âm sai

TPR, FPR	Tỉ lệ dương thật và tỉ lệ dương giả
M_{new}	Mô hình mới sau khi huấn luyện lại
$F_{1,min}$	Ngưỡng F1 tối thiểu để triển khai mô hình mới
L_i	Độ trễ đầu cuối của giao dịch thứ i
t_{prod}, t_{cons}	Thời điểm phát giao dịch và thời điểm nhận kết quả
$\overline{L}$, L_{p95}	Độ trễ trung bình và độ trễ phân vị 95
L_w	Độ rộng cửa sổ trượt khi tính thông lượng

DANH SÁCH BẢNG

Bảng 1.1 So sánh các nhóm phương pháp phát hiện gian lận 11
Bảng 1.2 So sánh các công cụ thu thập dữ liệu luồng 12
Bảng 1.3 So sánh các công cụ xử lý dữ liệu luồng 12
Bảng 2.1 Một số mẫu hành vi gian lận phổ biến được áp dụng 25
Bảng 3.1 Mô tả các trường dữ liệu của tập phát hiện gian lận thẻ tín dụng 2023 34
Bảng 3.2 Phân bố mẫu theo lớp của tập dữ liệu thử nghiệm 35
Bảng 3.3 Môi trường thử nghiệm hệ thống 35
Bảng 3.4 Các chỉ số đánh giá thử nghiệm 36
Bảng 3.5 Mô tả cách mô phỏng các yếu tố bất thường của giao dịch thẻ tín dụng 38
Bảng 3.6 Ma trận nhầm lẫn của mô hình trên tập kiểm thử 42
Bảng 3.7 Kết quả đánh giá mô hình CatBoost trên tập kiểm thử 44
Bảng 3.8 So sánh hiệu quả phát hiện gian lận khi kết hợp CEP 45
Bảng 3.9 Hiệu năng xử lý luồng của hệ thống 49
Bảng 3.10 Hiệu năng mô hình trước và sau khi cập nhật 50

DANH SÁCH HÌNH VẼ

Hình 1.1 Tổng quan về hệ thống phát hiện gian lận tín dụng 7
Hình 1.2 Kiến trúc tổng thể của hệ thống đề xuất 13
Hình 2.1 Kiến trúc tổng thể của hệ thống phát hiện gian lận tín dụng 18
Hình 2.2 Luồng xử lý Apache Kafka 22
Hình 2.3 Luồng xử lý Apache Flink 23
Hình 2.4 Quy trình cập nhật mô hình liên tục 29
Hình 3.1 Các chỉ số đánh giá của mô hình trên ứng dụng 43
Hình 3.2 Các chỉ số đánh giá giao dịch chế độ thời gian thực 44

MỞ ĐẦU

Trong nền kinh tế hiện đại, tín dụng giữ vai trò thiết yếu đối với hoạt động kinh doanh, đầu tư và tiêu dùng. Sự phát triển mạnh mẽ của thanh toán điện tử và thương mại trực tuyến khiến khối lượng giao dịch tín dụng tăng nhanh cả về số lượng lẫn tốc độ, đồng thời kéo theo sự gia tăng nghiêm trọng của các hành vi gian lận tín dụng cả về quy mô lẫn mức độ tinh vi. Gian lận không chỉ gây thiệt hại tài chính trực tiếp cho tổ chức tín dụng và chủ thẻ mà còn làm xói mòn niềm tin của khách hàng và ảnh hưởng đến sự ổn định của hệ thống thanh toán [8].

Các phương pháp phát hiện gian lận truyền thống dựa trên luật cố định hoặc mô hình thống kê đơn giản ngày càng tỏ ra kém hiệu quả trước khối lượng dữ liệu khổng lồ và các thủ đoạn gian lận thay đổi liên tục. Thực tiễn đó đặt ra hai thách thức cốt lõi đối với một hệ thống phát hiện gian lận hiện đại: một là phải xử lý và phân tích giao dịch theo thời gian thực với độ trễ thấp để kịp thời ngăn chặn; hai là phải liên tục cập nhật mô hình nhằm thích nghi với sự thay đổi không ngừng của hành vi gian lận.

Xuất phát từ những vấn đề lý luận và thực tiễn nêu trên, tác giả lựa chọn đề tài "Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục". Mục tiêu của đề án là nghiên cứu, xây dựng và đánh giá một hệ thống phát hiện gian lận hoạt động theo thời gian thực, đạt độ chính xác cao, có khả năng nhận diện các hành vi gian lận biểu hiện theo chuỗi và tự thích nghi với sự thay đổi của dữ liệu mà không làm gián đoạn dịch vụ.

Về phương pháp, đề án đề xuất một kiến trúc lai kết hợp bốn thành phần bổ trợ lẫn nhau: mô hình học máy có giám sát CatBoost làm lõi phân loại; hạ tầng xử lý dữ liệu luồng thời gian thực dựa trên Apache Kafka và Apache Flink; kỹ thuật xử lý sự kiện phức tạp (CEP) nhằm phát hiện các chuỗi hành vi bất thường mà mô hình xét từng giao dịch độc lập khó nhận diện; và cơ chế cập nhật mô hình động cho phép hệ thống tự huấn luyện lại và thay thế mô hình theo thời gian. Hệ thống được hiện thực hóa và thử nghiệm trên tập dữ liệu Phát hiện gian lận thẻ tín dụng năm 2023.

Kết quả thử nghiệm cho thấy hệ thống đạt được các mục tiêu đề ra: mô hình học máy cho chất lượng phân loại cao trên tập kiểm thử; trên luồng trực tuyến, việc kết hợp xử lý sự kiện phức tạp giúp giảm đáng kể số giao dịch gian lận bị bỏ sót so với khi chỉ sử dụng mô hình đơn lẻ; hệ thống xử lý ổn định theo thời gian thực với thông lượng cao và không thất thoát dữ liệu; đồng thời cơ chế cập nhật mô hình động vận hành liên tục mà không gây gián đoạn dịch vụ. Các số liệu định lượng chi tiết được trình bày và phân tích trong Chương 3.

Ngoài phần mở đầu và kết luận, nội dung của đề án được tổ chức thành ba chương. Chương 1 trình bày tổng quan về bài toán phát hiện gian lận giao dịch tín dụng và các phương pháp liên quan. Chương 2 trình bày kiến trúc tổng thể cùng các kỹ thuật được áp dụng để triển khai hệ thống. Chương 3 trình bày quá trình thử nghiệm, đánh giá kết quả và kết luận của đề án.

CHƯƠNG 1 - TỔNG QUAN VỀ PHÁT HIỆN BẤT THƯỜNG GIAO DỊCH TÍN DỤNG

1.1. Bài toán phát hiện giao dịch tín dụng bất thường

Tín dụng là một trong những trụ cột của hệ thống tài chính hiện đại, cho phép cá nhân và doanh nghiệp tiếp cận nguồn vốn để chi tiêu, đầu tư và kinh doanh. Cùng với sự bùng nổ của thanh toán điện tử và thương mại trực tuyến, khối lượng giao dịch tín dụng tăng nhanh cả về số lượng lẫn tốc độ, kéo theo sự gia tăng tương ứng của các hành vi gian lận. Gian lận tín dụng không chỉ gây thiệt hại tài chính trực tiếp cho ngân hàng và chủ thẻ mà còn làm xói mòn niềm tin của khách hàng và ảnh hưởng đến sự ổn định của cả hệ thống thanh toán [8].

Trong bài toán này, cần phân biệt hai khái niệm có liên quan nhưng không đồng nhất. Giao dịch bất thường (anomaly) là giao dịch lệch đáng kể so với hành vi tiêu dùng thông thường của chủ thẻ hoặc so với phần lớn các giao dịch hợp lệ khác. Giao dịch gian lận (fraud) là giao dịch được thực hiện một cách trái phép, không có sự đồng ý của chủ thẻ. Trên thực tế, không phải mọi giao dịch bất thường đều là gian lận, một khách hàng có thể chi tiêu đột biến khi đi du lịch và ngược lại, không phải mọi giao dịch gian lận đều có biểu hiện cực đoan rõ rệt, bởi kẻ gian thường cố tình mô phỏng hành vi bình thường để né tránh giám sát. Chính sự chồng lấn này khiến việc phát hiện gian lận trở thành một bài toán khó, đòi hỏi kết hợp nhiều góc nhìn phân tích khác nhau thay vì dựa vào một quy tắc đơn giản.

Mức độ nghiêm trọng của vấn đề thể hiện qua quy mô thiệt hại ngày càng lớn trên phạm vi toàn cầu, với tổn thất từ gian lận tín dụng được ghi nhận lên tới hàng chục tỷ đô-la mỗi năm và các thủ đoạn ngày càng tinh vi [8]. Điều này đặt ra nhu cầu cấp thiết về một hệ thống có khả năng phát hiện sớm và ngăn chặn giao dịch gian lận ngay tại thời điểm phát sinh, thay vì chỉ rà soát sau khi thiệt hại đã xảy ra. Một hệ thống như vậy phải đồng thời giải quyết hai mục tiêu mâu thuẫn: hạn chế tối đa giao dịch gian lận bị bỏ sót (âm tính giả) để giảm thiệt hại, đồng thời tránh cảnh báo nhầm các giao dịch hợp lệ (dương tính giả) gây phiền hà và làm giảm trải nghiệm của khách hàng.

Bài toán phát hiện bất thường giao dịch tín dụng về bản chất là một bài toán phân loại nhị phân trên dữ liệu lớn, nhưng mang ba đặc thù khiến nó khác biệt so với các bài toán phân loại thông thường: dữ liệu mất cân bằng nghiêm trọng khi tỷ lệ gian lận thường dưới 1%; yêu cầu xử lý theo thời gian thực với độ trễ thấp để kịp thời ngăn chặn; và hành vi gian lận thay đổi liên tục theo thời gian khiến mô hình nhanh chóng lỗi thời. Ba đặc thù này là sợi chỉ xuyên suốt, định hình toàn bộ các lựa chọn về dữ liệu, phương pháp và kiến trúc được trình bày trong các phần tiếp theo của chương.

1.2. Các đặc trưng dữ liệu giao dịch tín dụng

Dữ liệu là yếu tố cốt lõi quyết định hiệu quả của hệ thống phát hiện gian lận. Việc hiểu rõ cấu trúc, bản chất và các đặc trưng của dữ liệu giao dịch giúp lựa chọn mô hình phù hợp, thiết kế quy trình tiền xử lý hiệu quả và nâng cao khả năng phát hiện bất thường. Trong hệ thống phát hiện gian lận, mỗi bản ghi không chỉ đơn thuần mô tả một giao dịch riêng lẻ mà còn phản ánh hành vi tài chính của người dùng theo thời gian.

Một bản ghi giao dịch tín dụng thông thường bao gồm ba nhóm thông tin. Nhóm thông tin giao dịch gồm thời điểm thực hiện, số tiền, loại giao dịch (mua hàng, rút tiền, thanh toán trực tuyến), địa điểm hoặc quốc gia, và thiết bị hay kênh thanh toán được sử dụng. Nhóm thông tin khách hàng gồm mã khách hàng, lịch sử giao dịch, hạn mức tín dụng và mức độ rủi ro được ghi nhận trước đó. Cuối cùng là nhãn phân loại, với giá trị 0 ứng với giao dịch hợp lệ và 1 ứng với giao dịch gian lận, được sử dụng làm mục tiêu huấn luyện cho các mô hình học có giám sát.

1.2.1. Phân loại đặc trưng

Từ dữ liệu thô, các đặc trưng đưa vào mô hình thường được phân thành bốn nhóm. Đặc trưng định lượng phản ánh giá trị số trực tiếp của giao dịch như số tiền, thời gian kể từ giao dịch trước đó, số lượng giao dịch trong một khoảng thời gian hay giá trị trung bình trong 24 giờ gần nhất; các đặc trưng này thường được đưa vào mô hình sau khi chuẩn hóa thang đo. Đặc trưng thời gian nắm bắt tính chuỗi của hành vi gian lận, bao gồm khoảng cách thời gian giữa các giao dịch, tần suất giao

dịch trong một cửa sổ thời gian, hay việc giao dịch được thực hiện ngoài khung giờ hoạt động thông thường của khách hàng. Đặc trưng hành vi mô tả thói quen tiêu dùng đặc trưng của từng chủ thẻ như mức chi tiêu trung bình, độ lệch chuẩn giá trị giao dịch, vị trí địa lý và thiết bị thường dùng; bất kỳ sai lệch lớn nào so với hành vi nền này đều có thể là dấu hiệu gian lận. Cuối cùng, đặc trưng dẫn xuất được tạo ra trong quá trình tiền xử lý nhằm khuếch đại tín hiệu bất thường, ví dụ tỷ lệ thay đổi giá trị giao dịch so với trung bình, số giao dịch trong 5 phút gần nhất, số quốc gia khác nhau trong 24 giờ, hay điểm rủi ro tích lũy theo lịch sử.

1.2.2. Đặc điểm thống kê

Bên cạnh cách phân nhóm trên, dữ liệu giao dịch tín dụng còn mang một số đặc điểm thống kê đặc thù ảnh hưởng trực tiếp đến việc lựa chọn mô hình và chiến lược xử lý.

Thứ nhất là tình trạng mất cân bằng nghiêm trọng: tỷ lệ giao dịch gian lận thường rất thấp, phổ biến dưới 1% tổng số giao dịch. Hệ quả là một mô hình dự đoán toàn bộ giao dịch là hợp lệ vẫn đạt độ chính xác (accuracy) rất cao nhưng hoàn toàn vô nghĩa, bởi nó bỏ sót đúng nhóm cần phát hiện [1]. Vì vậy, việc đánh giá hệ thống không thể dựa vào độ chính xác tổng thể mà phải sử dụng các chỉ số phù hợp với dữ liệu mất cân bằng như độ chính xác dương (precision), độ nhạy (recall), điểm F1 và diện tích dưới đường cong ROC [9].

Thứ hai là tính phi tuyến: hành vi gian lận không tuân theo một phân phối chuẩn và các điểm dữ liệu gian lận thường phân tán, không tạo thành cụm rõ ràng trong không gian đặc trưng. Điều này khiến các phương pháp tuyến tính truyền thống khó đạt hiệu quả cao và đòi hỏi các mô hình có khả năng nắm bắt quan hệ phức tạp giữa các đặc trưng.

Thứ ba là sự thay đổi của hành vi gian lận theo thời gian (concept drift): quan hệ giữa đặc trưng và nhãn dịch chuyển do kẻ gian liên tục đổi chiến thuật, thay đổi thời điểm và hình thức hoạt động [10]. Đây là điểm khác biệt căn bản so với nhiều bài toán phân loại tĩnh, một mô hình đạt hiệu quả tốt ở hiện tại sẽ dần suy

giảm nếu không được cập nhật định kỳ, và đây chính là lý do hệ thống cần một cơ chế học liên tục thay vì huấn luyện một lần.

Thứ tư là tính phụ thuộc chuỗi: gian lận thường biểu hiện qua một chuỗi hành vi liên tiếp thay vì một giao dịch đơn lẻ, chẳng hạn nhiều giao dịch nhỏ dồn dập, thay đổi vị trí địa lý nhanh bất thường, hay tăng đột biến tần suất giao dịch. Đặc điểm này gợi ý rằng việc phân tích từng giao dịch độc lập là chưa đủ, và cần bổ sung cơ chế phát hiện theo chuỗi sự kiện.

Tổng hợp lại, dữ liệu giao dịch tín dụng có độ phức tạp cao với bốn đặc điểm nổi bật: mất cân bằng, phi tuyến, hành vi gian lận thay đổi theo thời gian và phụ thuộc chuỗi. Việc phân tích kỹ các đặc trưng này là cơ sở để lựa chọn phương pháp và thiết kế kiến trúc hệ thống ở các phần sau.

1.3. Hệ thống phát hiện bất thường giao dịch tín dụng

1.3.1. Vai trò của hệ thống

Hệ thống phát hiện gian lận tín dụng (Credit Fraud Detection System) là hệ thống có chức năng giám sát, phân tích và đánh giá các giao dịch nhằm xác định khả năng xảy ra hành vi gian lận. Hệ thống sử dụng các phương pháp phân tích dữ liệu, thống kê và học máy để phân loại giao dịch thành hai nhóm hợp lệ và gian lận; tùy theo mức độ rủi ro ước lượng được, hệ thống có thể cho giao dịch đi qua, đưa ra cảnh báo, yêu cầu xác thực bổ sung hoặc tự động từ chối [8]. Vai trò của hệ thống vì vậy không chỉ là phân loại chính xác mà còn là tuyến phòng thủ tự động giúp tổ chức tài chính giảm thiểu thiệt hại trong khi vẫn duy trì sự thông suốt của hoạt động thanh toán.

1.3.2. Yêu cầu của hệ thống

Việc thiết kế một hệ thống phát hiện gian lận tín dụng không chỉ dừng lại ở việc xây dựng một mô hình học máy có độ chính xác cao. Trong môi trường tài chính thực tế, hệ thống phải đáp ứng đồng thời nhiều yêu cầu về kỹ thuật, nghiệp vụ và vận hành, xuất phát từ đặc thù của dữ liệu giao dịch, mức độ rủi ro tài chính và yêu cầu trải nghiệm người dùng.

Trước hết, hệ thống cần đạt độ chính xác cao theo nghĩa cân bằng giữa độ nhạy và độ chính xác dương, bởi cả việc bỏ sót gian lận lẫn việc cảnh báo nhầm đều gây tổn thất. Hệ thống phải có khả năng xử lý thời gian thực với độ trễ thấp để kịp ngăn chặn giao dịch trước khi hoàn tất, và khả năng mở rộng để đáp ứng khối lượng giao dịch lớn, biến động theo thời điểm. Do hành vi gian lận thay đổi liên tục, hệ thống cần khả năng thích nghi thông qua cập nhật mô hình theo thời gian, đồng thời nên tích hợp đa phương pháp để bù đắp hạn chế của từng kỹ thuật đơn lẻ. Bên cạnh đó, khả năng giải thích quyết định giúp bộ phận nghiệp vụ tin tưởng và rà soát kết quả; độ tin cậy và an toàn dữ liệu bảo đảm hệ thống vận hành ổn định trên dữ liệu nhạy cảm; và cuối cùng là khả năng giám sát và cập nhật mô hình để theo dõi hiệu năng và can thiệp kịp thời khi chất lượng suy giảm.

1.3.3. Mô hình tổng quát

Về mặt cấu trúc, một hệ thống phát hiện gian lận tín dụng điển hình bao gồm bốn thành phần kế tiếp nhau. Thành phần thu thập và tiếp nhận dữ liệu đảm nhận việc tiếp nhận luồng giao dịch từ nhiều nguồn khác nhau. Thành phần xử lý và tiền xử lý dữ liệu thực hiện làm sạch, chuẩn hóa và trích xuất đặc trưng. Thành phần mô hình đánh giá rủi ro sử dụng các kỹ thuật phân tích và học máy để chấm điểm khả năng gian lận của từng giao dịch. Cuối cùng, thành phần cảnh báo và giám sát đưa ra quyết định và theo dõi hoạt động của toàn hệ thống. Hình 1.1 minh họa mô hình tổng quát này.



Hình 1.1 Tổng quan về hệ thống phát hiện gian lận tín dụng

1.4. Các phương pháp phát hiện bất thường

Phát hiện gian lận tín dụng về bản chất là bài toán phát hiện bất thường trong tập dữ liệu giao dịch tài chính. Như đã phân tích, không phải mọi giao dịch bất thường đều là gian lận và không phải gian lận nào cũng có biểu hiện cực đoan, do đó việc lựa chọn phương pháp phát hiện phù hợp đóng vai trò quyết định. Phần này trình bày các nhóm phương pháp tiêu biểu cùng với cách chúng đã được vận dụng trong lĩnh vực tài chính.

1.4.1. Phương pháp dựa trên luật

Phương pháp dựa trên luật sử dụng các quy tắc nghiệp vụ dạng nếu – thì được chuyên gia tài chính thiết lập trước, ví dụ cảnh báo khi số tiền vượt một ngưỡng cho trước, khi giao dịch phát sinh tại quốc gia chưa từng xuất hiện, hoặc khi số giao dịch trong một phút vượt quá một giới hạn nhất định. Các quy tắc này được đúc kết từ kinh nghiệm và thống kê lịch sử. Đây cũng là hướng tiếp cận xuất hiện sớm nhất trong thực tiễn ngân hàng: ngay từ thập niên 1990, các hệ chuyên gia dựa trên luật đã được xây dựng để cảnh báo gian lận tín dụng tiêu dùng ngay trong quá trình cấp phép giao dịch [16].

Ưu điểm của phương pháp là dễ triển khai, dễ hiểu và dễ giải thích, đồng thời phản ứng nhanh với các mẫu gian lận đã biết. Tuy nhiên, nó không phát hiện được các mẫu gian lận mới, dễ bị vô hiệu hóa khi kẻ gian thay đổi chiến thuật, và chi phí bảo trì tăng nhanh khi số lượng quy tắc lớn dần. Vì vậy, phương pháp dựa trên luật ngày nay thường được giữ lại như một lớp phòng thủ cơ bản, bổ trợ cho các mô hình học máy chứ không còn đóng vai trò trung tâm.

1.4.2. Phương pháp học có giám sát

Các phương pháp học có giám sát học quy luật phân loại trực tiếp từ dữ liệu đã được gắn nhãn (giao dịch hợp lệ hoặc gian lận). Đây là nhóm được sử dụng phổ biến và đạt hiệu quả cao nhất trong thực tế khi có sẵn nhãn, bao gồm ba hướng chính: mô hình thống kê, học máy truyền thống và học sâu.

a) Mô hình thống kê

Nhóm phương pháp thống kê giả định rằng các giao dịch hợp lệ tuân theo một phân phối nhất định, và những giao dịch có giá trị vượt quá ngưỡng xác định sẽ

bị xem là bất thường. Trong số này, hồi quy Logistic là kỹ thuật phân loại nhị phân phổ biến, ước lượng xác suất một giao dịch thuộc lớp gian lận dựa trên tổ hợp tuyến tính của các đặc trưng và đã được áp dụng rộng rãi cho dữ liệu thẻ tín dụng nhờ tốc độ huấn luyện nhanh và khả năng diễn giải tốt [3]. Hạn chế cốt lõi của nhóm này nằm ở giả định quan hệ tuyến tính giữa đặc trưng và nhãn, khiến chúng khó nắm bắt các tương tác phức tạp vốn phổ biến trong hành vi gian lận.

b) Học máy truyền thống

Khác với hướng tiếp cận dựa trên luật cố định, các thuật toán học máy cho phép mô hình tự học các quy luật phân loại từ dữ liệu. Cây quyết định phân chia dữ liệu theo các điều kiện tối ưu trên từng đặc trưng, có khả năng xử lý quan hệ phi tuyến và dễ diễn giải, và đã được sử dụng trong phát hiện gian lận thẻ tín dụng từ khá sớm [11]. Rừng ngẫu nhiên tập hợp nhiều cây quyết định nhằm giảm phương sai và tăng độ ổn định, cho hiệu quả tốt hơn cây đơn lẻ trên dữ liệu phi tuyến, đổi lại khó diễn giải và tốn chi phí tính toán hơn [23].

Nổi bật nhất trong nhóm này là tăng cường theo độ dốc (gradient boosting), kỹ thuật xây dựng các cây quyết định một cách tuần tự, trong đó mỗi mô hình sau học từ sai số của mô hình trước. Các biến thể hiện đại như XGBoost, LightGBM và CatBoost được đánh giá là đạt hiệu suất cao nhất trên dữ liệu dạng bảng trong nhiều nghiên cứu phát hiện gian lận, nhờ khả năng xử lý tốt dữ liệu mất cân bằng và quan hệ phi tuyến mà không đòi hỏi tiền xử lý phức tạp [21], [23]. Đây là cơ sở để đề án lựa chọn nhóm phương pháp này làm mô hình lõi, sẽ được luận giải ở mục 1.6.

c) Học sâu

Các mô hình học sâu nắm bắt những quan hệ phức tạp giữa các đặc trưng mà các mô hình tuyến tính khó biểu diễn [15]. Đáng chú ý trong bài toán này là các mạng hồi tiếp như RNN và biến thể LSTM, vốn được thiết kế để xử lý dữ liệu chuỗi và do đó phù hợp với bản chất phụ thuộc thời gian của giao dịch gian lận. Thay vì xét từng giao dịch độc lập, các mô hình này mô hình hóa toàn bộ chuỗi lịch sử của một thẻ, qua đó phát hiện những thay đổi hành vi mà mô hình tĩnh dễ bỏ sót [12]. Các cải tiến gần đây kết hợp thêm cơ chế chú ý (attention) với LSTM cho thấy hiệu

quả cao hơn nữa trên dữ liệu thẻ tín dụng [2]. Tuy vậy, học sâu đòi hỏi chi phí tính toán lớn, khó triển khai trong môi trường yêu cầu độ trễ thấp và khó giải thích quyết định, những rào cản đáng kể đối với hệ thống vận hành thời gian thực.

1.4.3. Phương pháp học không giám sát

Khi dữ liệu không có nhãn hoặc nhãn gian lận quá hiếm, các phương pháp học không giám sát trở nên hữu ích vì chúng phát hiện những điểm dữ liệu khác biệt so với phần lớn dữ liệu còn lại mà không cần biết trước đâu là gian lận. Các kỹ thuật tiêu biểu gồm phân cụm K-trung bình, máy véc-tơ hỗ trợ một lớp, rừng cô lập (isolation forest) và bộ tự mã hóa (autoencoder). Rừng cô lập [17] phát hiện bất thường bằng cách cô lập nhanh các điểm hiếm trong không gian đặc trưng thay vì mô hình hóa vùng dữ liệu bình thường; trong các nghiên cứu so sánh trên dữ liệu giao dịch thẻ tín dụng, rừng cô lập thường cho hiệu quả tốt nhất trong nhóm các phương pháp không giám sát [19]. Bộ tự mã hóa, được huấn luyện chỉ trên giao dịch hợp lệ để học lại mẫu hành vi bình thường, sẽ tạo sai số tái tạo lớn ở các giao dịch gian lận và nhờ đó nhận diện được chúng, một hướng tiếp cận rất phù hợp với dữ liệu mất cân bằng cực độ. Hạn chế chung của nhóm này là không phân biệt rõ giữa bất thường hợp lệ và gian lận thực sự, nên độ chính xác thường thấp hơn so với các phương pháp có giám sát.

1.4.4. Phát hiện chuỗi sự kiện bất thường

Như đã nêu, gian lận thường xảy ra theo chuỗi sự kiện thay vì một giao dịch đơn lẻ, chẳng hạn năm giao dịch nhỏ liên tiếp trong một phút, hai giao dịch tại hai quốc gia khác nhau trong thời gian ngắn, hay sự tăng đột biến tần suất giao dịch. Các hệ thống xử lý sự kiện phức tạp (Complex Event Processing - CEP) cho phép định nghĩa trước những mẫu chuỗi như vậy và phát hiện chúng ngay khi xảy ra trên luồng dữ liệu thời gian thực [7]. Khác với các mô hình học máy vốn đánh giá từng giao dịch độc lập, CEP nắm bắt mối liên hệ giữa nhiều giao dịch theo thời gian, nhờ đó bổ sung hiệu quả cho mô hình học máy và giảm tỷ lệ bỏ sót.

Bảng 1.1 tổng hợp và so sánh các nhóm phương pháp nêu trên theo các tiêu chí chính.

Bảng 1.1 So sánh các nhóm phương pháp phát hiện gian lận

Phương pháp	Khả năng phát hiện mẫu mới	Thời gian thực	Độ phức tạp	Khả năng giải thích
Dựa trên luật	Thấp	Cao	Thấp	Cao
Thống kê	Thấp	Cao	Thấp	Cao
Học máy truyền thống	Trung bình	Cao	Trung bình	Trung bình
Học sâu	Cao	Trung bình	Cao	Thấp
Học không giám sát	Trung bình	Trung bình	Trung bình	Thấp
Xử lý sự kiện phức tạp	Trung bình	Cao	Trung bình	Cao

1.5. Các kỹ thuật xử lý luồng dữ liệu

Các phương pháp phát hiện ở mục 1.4 mới chỉ trả lời câu hỏi phân tích bằng gì, trong khi yêu cầu xử lý thời gian thực đặt ra một câu hỏi không kém phần quan trọng là xử lý dữ liệu như thế nào. Trong mô hình xử lý theo lô (batch) truyền thống, dữ liệu được gom lại và phân tích định kỳ; cách làm này đơn giản nhưng không phù hợp với phát hiện gian lận, bởi một giao dịch chỉ có ý nghĩa ngăn chặn nếu được đánh giá ngay tại thời điểm phát sinh. Do đó, hệ thống cần chuyển sang mô hình xử lý luồng (stream processing), trong đó mỗi giao dịch được tiếp nhận và xử lý liên tục ngay khi xuất hiện, với độ trễ ở mức mili-giây đến vài giây.

Một kiến trúc xử lý luồng điển hình gồm hai tầng. Tầng truyền tải và quản lý thông điệp tiếp nhận giao dịch từ nhiều nguồn, lưu đệm và phân phối tới các thành phần phía sau, bảo đảm dữ liệu không bị mất khi tải tăng cao. Apache Kafka là nền tảng phổ biến nhất cho tầng này nhờ thông lượng cao, khả năng lưu trữ bền vững và mở rộng theo chiều ngang [14]. Các giải pháp tương tự gồm RabbitMQ, một message broker truyền thống mạnh về định tuyến linh hoạt nhưng thông lượng thấp hơn, và Apache Pulsar, nền tảng thế hệ mới tách biệt tầng tính toán và lưu trữ, hỗ trợ đa người thuê. Trong các phép đo thực nghiệm phổ biến, Kafka thường đạt thông lượng cao nhất và độ trễ thấp, đồng thời có hệ sinh thái tích hợp trưởng thành,

nên được lựa chọn rộng rãi cho các hệ thống giao dịch tài chính quy mô lớn. Bảng 1.2 So sánh ba nền tảng truyền tải tiêu biểu.

Bảng 1.2 So sánh các công cụ thu thập dữ liệu luồng

Tiêu chí	RabbitMQ	Apache Pulsar	Apache Kafka
Thông lượng lớn	Trung bình	Cao	Rất cao
Lưu trữ dữ liệu dài hạn	Hạn chế	Có	Có
Khả năng phát lại dữ liệu	Hạn chế	Có	Có
Khả năng mở rộng	Trung bình	Cao	Cao
Hệ sinh thái phổ biến	Cao	Trung bình	Rất cao

Tầng xử lý luồng thực hiện tính toán trên dòng dữ liệu liên tục: trích xuất đặc trưng theo cửa sổ thời gian, tính các đặc trưng hành vi động và áp dụng mô hình để chấm điểm rủi ro. Apache Flink là một trong những công cụ hàng đầu cho tầng này nhờ mô hình xử lý theo thời gian sự kiện (event time), cho phép tính toán chính xác trên dữ liệu đến không đúng thứ tự, cùng cơ chế quản lý trạng thái và bảo đảm xử lý chính xác một lần [5]. So với các giải pháp khác như Apache Spark Streaming (xử lý theo vi lô nên độ trễ cao hơn) hay Apache Storm (độ trễ thấp nhưng thiếu quản lý trạng thái mạnh), Flink cân bằng tốt giữa độ trễ thấp và khả năng xử lý trạng thái phức tạp, đặc biệt phù hợp với việc tích hợp cả mô hình học máy lẫn cơ chế CEP trên cùng một nền tảng. Đây là cơ sở để đề án lựa chọn cặp Kafka – Flink làm xương sống xử lý dữ liệu, với chi tiết triển khai được trình bày ở Chương 2. Bảng 1.3 so sánh ba công cụ xử lý luồng tiêu biểu.

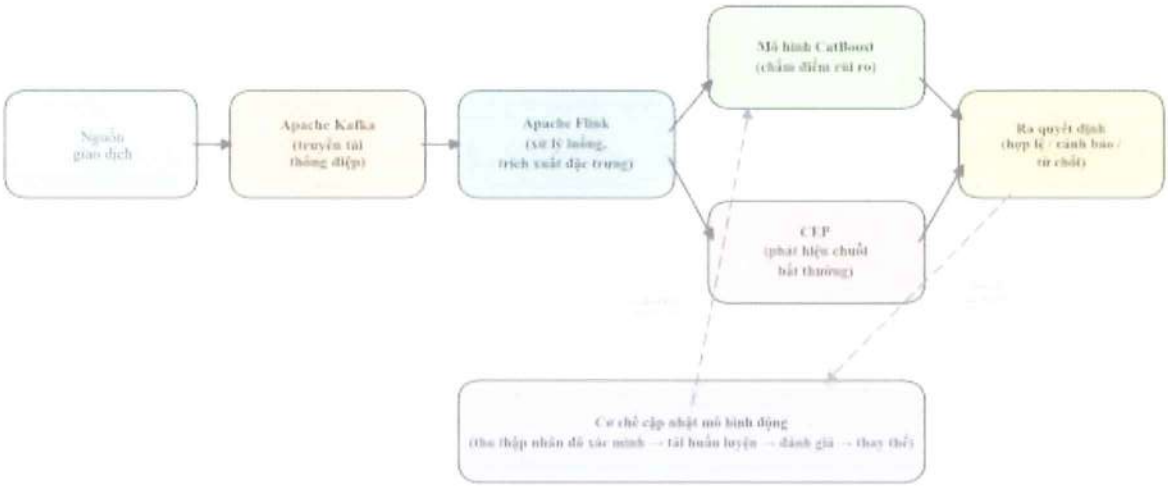
Bảng 1.3 So sánh các công cụ xử lý dữ liệu luồng

Tiêu chí	Storm	Spark Streaming	Flink
Xử lý thời gian thực thực sự	Có	Lô dữ liệu nhỏ	Có
Độ trễ	Thấp	Trung bình	Rất thấp
Quản lý trạng thái	Hạn chế	Tốt	Rất tốt
Chính xác một lần	Khó	Có	Hỗ trợ đầy đủ
Cửa sổ và sự kiện thời gian	Cơ bản	Tốt	Rất mạnh

Khả năng xử lý phức tạp	Trung bình	Cao	Rất cao

1.6. Mô hình đề xuất

Từ các yêu cầu hệ thống ở mục 1.3 và các nhóm phương pháp ở mục 1.4, có thể thấy không một phương pháp đơn lẻ nào đáp ứng đầy đủ các tiêu chí về độ chính xác, khả năng xử lý thời gian thực, khả năng mở rộng và tính thích nghi với sự thay đổi của hành vi gian lận theo thời gian. Phương pháp dựa trên luật minh bạch nhưng cứng nhắc; mô hình học máy mạnh về độ chính xác nhưng xét từng giao dịch độc lập; học sâu nắm bắt chuỗi tốt nhưng nặng và khó triển khai thời gian thực; còn cơ chế cập nhật lại nằm ngoài bản thân các thuật toán phân loại. Do đó, đề án đề xuất một kiến trúc lai kết hợp bốn thành phần bổ trợ lẫn nhau: một mô hình học máy có giám sát làm lõi phân loại, một nền tảng xử lý dữ liệu luồng thời gian thực, một cơ chế xử lý sự kiện phức tạp để phát hiện chuỗi bất thường, và một cơ chế cập nhật mô hình động để thích nghi với sự thay đổi của dữ liệu. Hình 1.2 minh họa kiến trúc tổng thể được đề xuất.



Hình 1.2 Kiến trúc tổng thể của hệ thống đề xuất

1.6.1. Lựa chọn mô hình học máy

Để chọn mô hình lõi, đề án cân nhắc giữa ba hướng. Hồi quy Logistic tuy nhanh và dễ diễn giải nhưng giả định quan hệ tuyến tính và khó xử lý tương tác phức tạp giữa các đặc trưng, nên không phù hợp với dữ liệu gian lận có tính phi

tuyến cao. Các mô hình học sâu như LSTM nắm bắt mẫu phức tạp nhưng đòi hỏi chi phí tính toán lớn, khó đạt độ trễ thấp và khó giải thích quyết định, không lý tưởng cho một hệ thống vận hành gần thời gian thực. Trong khi đó, các thuật toán tăng cường theo độ dốc trên cây quyết định cho hiệu suất cao và ổn định trên dữ liệu dạng bảng, xử lý tốt dữ liệu mất cân bằng và phi tuyến mà không cần tiền xử lý phức tạp.

Trên cơ sở đó, đề án lựa chọn CatBoost làm mô hình phân loại chính. CatBoost là một thuật toán tăng cường theo độ dốc trên cây quyết định, nổi bật ở cơ chế tăng cường theo thứ tự (ordered boosting) giúp giảm hiện tượng quá khớp, cùng khả năng xử lý đặc trưng phân loại hiệu quả và độ ổn định cao mà không đòi hỏi tinh chỉnh siêu tham số quá phức tạp [21]. So với rừng ngẫu nhiên và XGBoost, CatBoost cho sự cân bằng tốt giữa độ chính xác, độ ổn định và chi phí vận hành, phù hợp với yêu cầu triển khai thực tế của đề án.

1.6.2. Xử lý dữ liệu luồng thời gian thực

Mô hình học máy chỉ phát huy giá trị nếu được đặt trong một hạ tầng xử lý đủ nhanh. Như đã phân tích ở mục 1.5, đề án sử dụng Apache Kafka làm tầng truyền tải để tiếp nhận giao dịch từ nhiều nguồn, bảo đảm tính bền vững của dữ liệu và phân phối tới các thành phần xử lý, đồng thời cho phép mở rộng theo chiều ngang khi khối lượng giao dịch tăng [14]. Apache Flink đảm nhận tầng xử lý luồng, thực hiện việc tiếp nhận từng giao dịch ngay khi phát sinh, xây dựng đặc trưng theo cửa sổ thời gian, tính toán các đặc trưng hành vi động và gọi mô hình học máy để chấm điểm rủi ro [5]. Sự kết hợp này cho phép hệ thống xử lý liên tục theo thời gian sự kiện, đáp ứng yêu cầu độ trễ thấp đặc thù của bài toán phát hiện gian lận [22].

1.6.3. Xử lý sự kiện phức tạp (CEP)

Một hạn chế cố hữu của mô hình học máy là nó đánh giá mỗi giao dịch một cách độc lập, trong khi nhiều hành vi gian lận chỉ bộc lộ khi xét nhiều giao dịch trong mối liên hệ thời gian. Để khắc phục, đề án tích hợp cơ chế xử lý sự kiện phức tạp nhằm phát hiện các chuỗi bất thường ngay trên luồng dữ liệu. Cụ thể, CEP định nghĩa trước các mẫu chuỗi đáng ngờ, như nhiều giao dịch liên tiếp trong thời gian

ngắn, thay đổi vị trí địa lý đột ngột, hay tăng đột biến giá trị giao dịch, và liên tục đối chiếu dòng giao dịch với các mẫu này theo cửa sổ thời gian. Khi một chuỗi khớp với mẫu định nghĩa, hệ thống phát tín hiệu cảnh báo bổ sung bên cạnh điểm rủi ro của mô hình học máy. Nhờ đó, CEP giúp giảm tỷ lệ âm tính giả và phát hiện được những mẫu hành vi mà mô hình đơn lẻ có thể bỏ sót; sự kết hợp giữa học máy và CEP tạo nên một hệ thống mạnh hơn so với chỉ sử dụng một trong hai.

1.6.4. Cơ chế cập nhật mô hình động

Thành phần cuối cùng giải quyết trực tiếp đặc điểm hành vi gian lận thay đổi theo thời gian. Do hành vi gian lận thay đổi liên tục, một mô hình cố định sẽ suy giảm hiệu quả theo thời gian; vì vậy hệ thống cần khả năng tự cập nhật để duy trì độ chính xác [10]. Cơ chế cập nhật động được đề xuất hoạt động theo một vòng khép kín: hệ thống thu thập các giao dịch đã được xác minh nhãn, tích lũy chúng vào tập dữ liệu huấn luyện bổ sung, rồi tái huấn luyện mô hình theo định kỳ. Mô hình mới sau khi huấn luyện được đánh giá và so sánh với mô hình đang vận hành; chỉ khi đạt hiệu quả cao hơn, nó mới được đưa vào thay thế. Quy trình này cho phép hệ thống liên tục thích nghi với các mẫu gian lận mới mà không cần can thiệp thủ công, biến khả năng học liên tục thành một đặc tính vận hành thường trực thay vì một thao tác rời rạc.

Tóm lại, kiến trúc lai được đề xuất gồm: CatBoost làm mô hình phân loại chính, Apache Kafka làm tầng truyền tải và quản lý sự kiện, Apache Flink làm tầng xử lý luồng, cơ chế CEP để phát hiện chuỗi bất thường, và cơ chế cập nhật mô hình động để thích nghi với sự thay đổi của hành vi gian lận theo thời gian. Chi tiết triển khai từng thành phần được trình bày trong Chương 2.

1.7. Kết luận chương

Chương 1 đã trình bày tổng quan về bài toán phát hiện gian lận trong giao dịch tín dụng, bắt đầu từ việc đặt vấn đề và làm rõ nhu cầu phát hiện, ngăn chặn giao dịch bất thường theo thời gian thực. Trên cơ sở phân tích các đặc trưng và đặc điểm thống kê của dữ liệu giao dịch, mất cân bằng, phi tuyến, hành vi gian lận thay đổi theo thời gian và phụ thuộc chuỗi, chương đã chỉ ra những thách thức cốt lõi và

các yêu cầu đặt ra đối với một hệ thống phát hiện hiệu quả. Tiếp đó, chương khảo sát các nhóm phương pháp phát hiện bất thường cùng cách chúng được vận dụng trong lĩnh vực tài chính, và giới thiệu các kỹ thuật xử lý luồng dữ liệu làm nền tảng cho yêu cầu thời gian thực. Từ những phân tích này, chương đã đề xuất một kiến trúc lai kết hợp mô hình học máy CatBoost, hạ tầng xử lý luồng Apache Kafka và Apache Flink, cơ chế xử lý sự kiện phức tạp và cơ chế cập nhật mô hình động. Kiến trúc đề xuất sẽ được triển khai chi tiết trong Chương 2 và đánh giá thực nghiệm trong Chương 3.

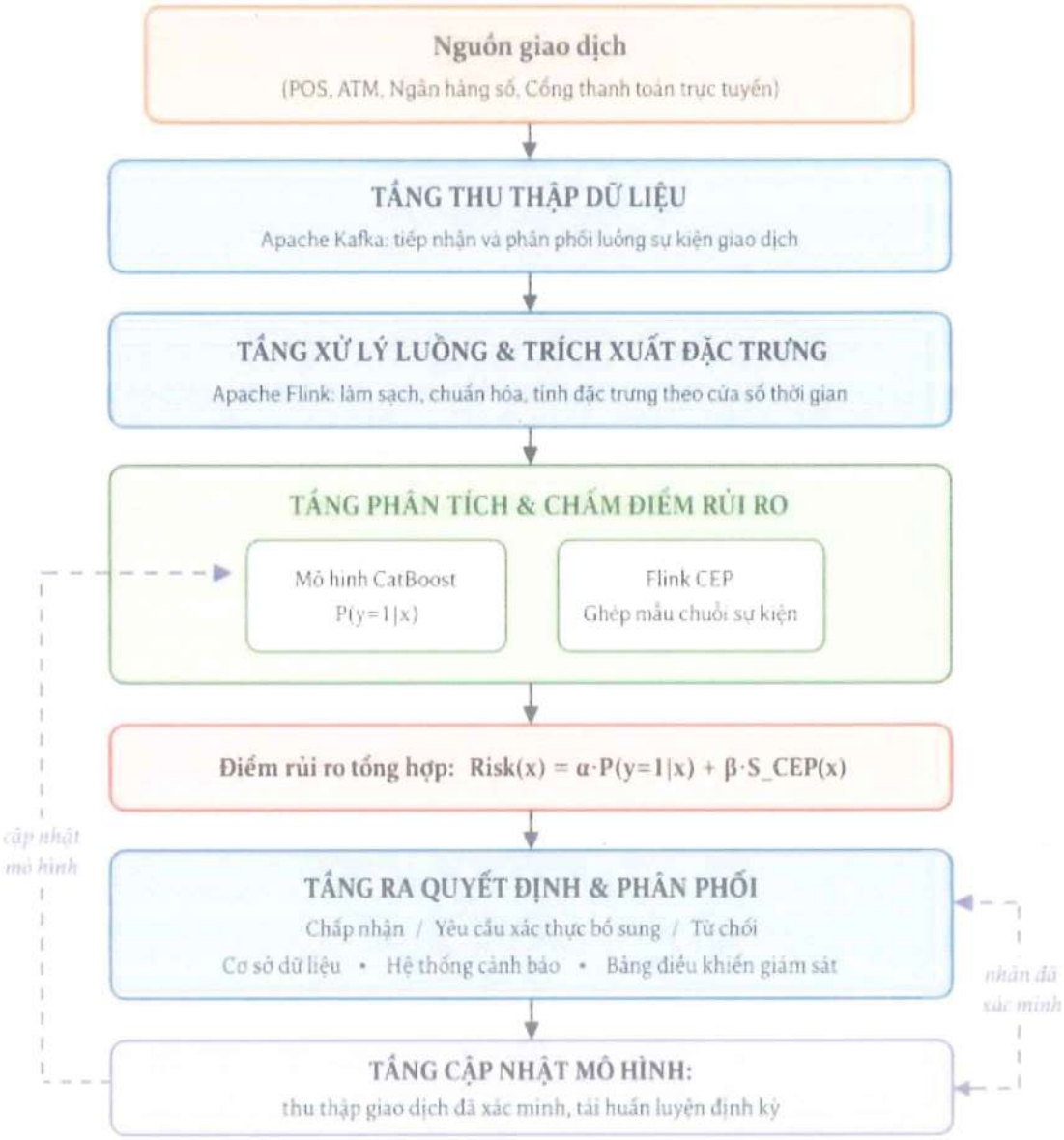
CHƯƠNG 2 - TRIỂN KHAI HỆ THỐNG VÀ CÁC KỸ THUẬT ĐƯỢC ÁP DỤNG

Chương 1 đã đề xuất một kiến trúc lai kết hợp bốn thành phần: mô hình học máy có giám sát, xử lý dữ liệu luồng thời gian thực, xử lý sự kiện phức tạp và cơ chế cập nhật mô hình động. Chương 2 hiện thực hóa kiến trúc đó. Trước hết, mục 2.1 trình bày kiến trúc tổng thể và cách bốn thành phần phối hợp để giải quyết bài toán; các mục tiếp theo lần lượt đi sâu vào từng thành phần theo đúng thứ tự đã nêu ở Chương 1: mô hình học máy có giám sát (2.2), xử lý dữ liệu luồng thời gian thực (2.3), xử lý sự kiện phức tạp (2.4) và cơ chế cập nhật mô hình động (2.5); cuối cùng, mục 2.6 trình bày khâu hậu xử lý gồm tổng hợp tín hiệu, ra quyết định và giám sát.

2.1. Kiến trúc tổng thể của hệ thống

Trước khi đi sâu vào từng thành phần, mục này trình bày kiến trúc tổng thể của hệ thống nhằm cung cấp một cái nhìn bao quát về cách các thành phần phối hợp với nhau. Hệ thống được tổ chức để giải quyết đồng thời bốn vấn đề đã nêu ở Chương 1: phân loại giao dịch bằng mô hình học máy có giám sát, xử lý dữ liệu luồng theo thời gian thực, phát hiện chuỗi hành vi bất thường bằng xử lý sự kiện phức tạp, và thích nghi với sự thay đổi của hành vi gian lận theo thời gian thông qua cơ chế cập nhật mô hình động. Các mục tiếp theo của chương sẽ lần lượt trình bày chi tiết cách mỗi vấn đề được giải quyết.

Hệ thống được thiết kế theo kiến trúc phân tầng và hướng sự kiện, trong đó mỗi tầng đảm nhiệm một nhóm chức năng riêng biệt và giao tiếp với các tầng còn lại thông qua luồng dữ liệu. Cách tổ chức này giúp tách biệt rõ ràng trách nhiệm giữa các thành phần, đồng thời cho phép mở rộng hoặc thay thế từng tầng mà không ảnh hưởng đến toàn bộ hệ thống. Kiến trúc tổng thể của hệ thống được minh họa trong Hình 2.1.



Hình 2.1 Kiến trúc tổng thể của hệ thống phát hiện gian lận tín dụng

Kiến trúc của hệ thống bao gồm năm tầng chức năng chính, được mô tả chi tiết như sau:

2.1.1. Tầng thu thập dữ liệu

Tầng thu thập đảm nhiệm việc tiếp nhận dữ liệu giao dịch phát sinh liên tục từ nhiều nguồn khác nhau như cổng thanh toán trực tuyến, ứng dụng ngân hàng số, máy ATM, thiết bị POS và các dịch vụ thanh toán trung gian. Mỗi giao dịch được chuyển thành một sự kiện và đưa vào Apache Kafka thông qua các chủ đề dữ liệu tương ứng. Kafka đóng vai trò tầng đệm trung gian, giúp tách biệt nguồn phát sinh

dữ liệu với hệ thống xử lý phía sau, đảm bảo tính bền vững của dữ liệu và khả năng mở rộng khi khối lượng giao dịch tăng cao.

2.1.2. Tầng xử lý luồng và trích xuất đặc trưng

Dữ liệu từ Kafka được Apache Flink tiếp nhận và xử lý theo thời gian thực. Tại tầng này, hệ thống thực hiện làm sạch dữ liệu, chuẩn hóa định dạng và loại bỏ các bản ghi lỗi. Dựa trên cơ chế quản lý trạng thái và xử lý theo cửa sổ thời gian, Flink tính toán các đặc trưng hành vi của từng giao dịch và từng khách hàng, chẳng hạn số lượng giao dịch trong một khoảng thời gian, tổng giá trị giao dịch, tần suất giao dịch thất bại liên tiếp hoặc sự thay đổi vị trí địa lý. Các đặc trưng này là đầu vào quan trọng cho tầng phân tích và chấm điểm rủi ro.

2.1.3. Tầng phân tích và chấm điểm rủi ro

Tầng phân tích là nơi đưa ra đánh giá về mức độ rủi ro của giao dịch dựa trên hai nguồn tín hiệu bổ trợ cho nhau. Thứ nhất, mô hình học máy CatBoost nhận véc-tơ đặc trưng từ tầng xử lý luồng và tính toán xác suất gian lận cho từng giao dịch. Thứ hai, thư viện Flink CEP phân tích chuỗi sự kiện nhằm phát hiện các mẫu hành vi bất thường mà mô hình đơn lẻ khó nhận diện. Kết quả từ hai nguồn này được tổng hợp thành điểm rủi ro chung, làm cơ sở cho quyết định phân loại cuối cùng.

2.1.4. Tầng ra quyết định và phân phối kết quả

Dựa trên điểm rủi ro tổng hợp, hệ thống đưa ra quyết định đối với từng giao dịch như chấp nhận, từ chối, yêu cầu xác thực bổ sung hoặc gắn cờ nghi ngờ để giám sát. Kết quả xử lý tiếp tục được phân phối qua Kafka đến các thành phần liên quan, bao gồm cơ sở dữ liệu lưu trữ lịch sử giao dịch, hệ thống cảnh báo thời gian thực và bảng điều khiển giám sát dành cho quản trị viên.

2.1.5. Tầng cập nhật mô hình

Để thích nghi với hiện tượng hành vi gian lận thay đổi theo thời gian, hệ thống duy trì một vòng cập nhật mô hình hoạt động song song với luồng xử lý chính. Các giao dịch đã được xác minh được thu thập và bổ sung vào tập dữ liệu huấn luyện, sau đó mô hình CatBoost được tái huấn luyện theo chu kỳ và đánh giá hiệu năng trước khi thay thế mô hình hiện tại. Nhờ tách biệt khỏi luồng suy luận

thời gian thực, quá trình cập nhật không làm gián đoạn hoạt động phát hiện gian lận của hệ thống.

Quy trình xử lý dữ liệu xuyên suốt hệ thống được thực hiện tuần tự qua các bước sau:

1. Giao dịch phát sinh từ các nguồn được đẩy vào Apache Kafka dưới dạng sự kiện;
2. Apache Flink tiếp nhận luồng sự kiện, thực hiện tiền xử lý và trích xuất đặc trưng;
3. Đặc trưng được đưa vào mô hình CatBoost để tính xác suất gian lận, đồng thời luồng sự kiện được chuyển tới Flink CEP để kiểm tra các mẫu hành vi bất thường;
4. Điểm dự đoán của mô hình và kết quả ghép mẫu của CEP được tổng hợp thành điểm rủi ro chung;
5. Hệ thống ra quyết định phân loại và phân phối kết quả tới cơ sở dữ liệu, hệ thống cảnh báo và bảng điều khiển giám sát;
6. Song song, các giao dịch đã được xác minh được đưa vào vòng cập nhật mô hình để tái huấn luyện định kỳ.

2.2. Mô hình học máy có giám sát

Đề án lựa chọn CatBoost, một thuật toán tăng cường độ dốc trên cây quyết định, làm mô hình phân loại chính trong hệ thống phát hiện gian lận [21].

Thuật toán được lựa chọn do có khả năng xử lý tốt dữ liệu dạng bảng, hoạt động hiệu quả với dữ liệu mất cân bằng và không yêu cầu quá nhiều bước tiền xử lý phức tạp. Bên cạnh đó, mô hình cho hiệu suất cao và ổn định trong nhiều bài toán phát hiện gian lận.

Tuy nhiên, đặc điểm của dữ liệu giao dịch tài chính là liên tục thay đổi theo thời gian, dẫn đến sự thay đổi mối quan hệ giữa dữ liệu đầu vào và đầu ra. Do đó, mô hình cần được cập nhật định kỳ nhằm duy trì khả năng phát hiện các hình thức gian lận mới. Vì vậy, đề án áp dụng cơ chế cập nhật mô hình động để tự động tái huấn luyện và thay thế mô hình khi cần thiết. [10]

Trong hệ thống, mô hình CatBoost được huấn luyện trên 29 đặc trưng đầu vào và xuất ra xác suất gian lận P(y=1|x); xác suất này là một trong hai nguồn tín hiệu được tổng hợp trong công thức điểm rủi ro (công thức 2.7), bên cạnh điểm xử lý sự kiện phức tạp.

2.3. Xử lý dữ liệu luồng

Trong bài toán phát hiện gian lận tín dụng, dữ liệu giao dịch thường phát sinh liên tục với tốc độ cao từ nhiều nguồn khác nhau như hệ thống thanh toán, ứng dụng ngân hàng số, cổng giao dịch trực tuyến, máy POS, ATM hoặc các dịch vụ trung gian thanh toán. Đặc điểm của loại dữ liệu này là có tính thời gian thực, khối lượng lớn, cấu trúc đa dạng và yêu cầu xử lý trong thời gian ngắn. Do đó, việc áp dụng mô hình xử lý dữ liệu luồng là cần thiết nhằm phát hiện kịp thời các hành vi bất thường, giảm thiểu rủi ro tài chính và hỗ trợ hệ thống cảnh báo sớm.

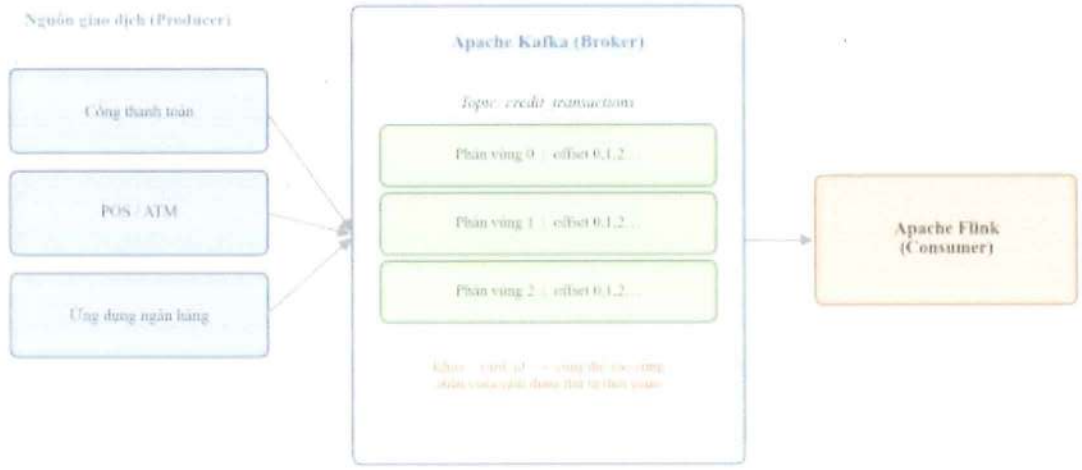
Xử lý dữ liệu luồng là phương pháp xử lý dữ liệu ngay khi dữ liệu được sinh ra, thay vì lưu trữ toàn bộ dữ liệu rồi mới xử lý theo từng lô. Trong hệ thống phát hiện gian lận tín dụng, cách tiếp cận này cho phép phân tích từng giao dịch gần như theo thời gian thực, từ đó đưa ra quyết định như chấp nhận giao dịch, từ chối giao dịch, yêu cầu xác thực bổ sung hoặc gửi cảnh báo đến bộ phận giám sát. [22]

2.3.1. Apache Kafka

Trong hệ thống, Apache Kafka đóng vai trò nền tảng truyền tải trung gian theo mô hình xuất bản, đăng ký: các nguồn giao dịch ghi dữ liệu vào các chủ đề (topic), còn tầng xử lý Apache Flink đọc và xử lý dữ liệu từ đó. Nhờ tách rời nguồn phát sinh dữ liệu khỏi tầng xử lý, hai phía có thể vận hành và mở rộng độc lập; đồng thời dữ liệu được lưu lại theo chính sách giữ lại trong một khoảng thời gian định trước, cho phép đọc lại lịch sử giao dịch phục vụ huấn luyện lại mô hình mà không cần một kho lưu trữ riêng. [14]

Mỗi chủ đề được chia thành nhiều phân vùng (partition) là đơn vị song song của Kafka, giúp tăng thông lượng theo chiều ngang. Đề án chọn khóa phân vùng là mã thẻ (card_id), nhờ đó toàn bộ giao dịch của một thẻ luôn nằm trên cùng một

phân vùng và được xử lý đúng thứ tự thời gian; đây là tiền đề để tầng xử lý sự kiện phức tạp tính chính xác các đặc trưng phụ thuộc trình tự như vận tốc, tần suất và đột biến giá trị giao dịch. Hệ thống áp dụng mức bảo đảm phân phối tối thiểu một lần ở khâu ghi kết quả để không bỏ sót bản ghi, và sử dụng chế độ KRaft để quản lý siêu dữ liệu cụm thay cho ZooKeeper nhằm đơn giản hóa triển khai. Kiến trúc Kafka trong đề án được minh họa ở Hình 2.2.



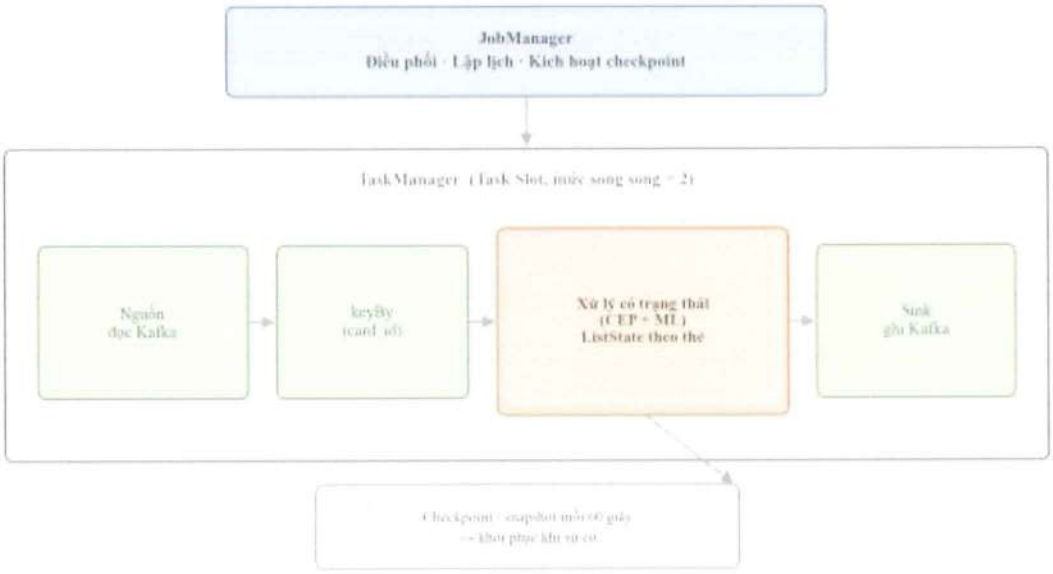
Hình 2.2 Luồng xử lý Apache Kafka

2.3.2. Apache Flink

Apache Flink là công cụ xử lý luồng phân tán có trạng thái, xử lý từng sự kiện ngay khi đến nên đạt độ trễ thấp mà vẫn quản lý được trạng thái và bảo đảm tính nhất quán của dữ liệu, phù hợp với bài toán phát hiện gian lận vốn dựa trên chuỗi hành vi giao dịch. Trong đề án, chương trình Flink được tổ chức thành một đồ thị xử lý gồm: nguồn đọc dữ liệu từ Kafka, phép phân vùng theo khóa keyBy(card_id), một hàm xử lý có trạng thái thực hiện đồng thời xử lý sự kiện phức tạp và chấm điểm bằng mô hình học máy, và cuối cùng là đích ghi kết quả trở lại Kafka; mức song song được đặt bằng 2. [5]

Trên luồng đã phân khóa, Flink lưu trạng thái theo từng thẻ bằng ListState để giữ danh sách mốc thời gian và giá trị của các giao dịch gần nhất, làm dữ liệu nền cho các mẫu xử lý sự kiện phức tạp. Flink phân biệt thời gian sự kiện với thời gian

xử lý và dùng cơ chế dấu mốc thời gian để xử lý đúng ngay cả khi giao dịch đến trễ hoặc không đúng thứ tự, đồng thời hỗ trợ các cửa sổ thời gian phục vụ tính đặc trưng theo khoảng thời gian. Khả năng chịu lỗi được bảo đảm bằng cơ chế điểm kiểm tra định kỳ 60 giây: khi xảy ra sự cố, hệ thống khôi phục cả trạng thái lẫn vị trí đọc, bảo đảm không bỏ sót và không tính trùng giao dịch. Kiến trúc xử lý của Flink trong đề án được minh họa ở Hình 2.3.



Hình 2.3 Luồng xử lý Apache Flink

2.4. Xử lý sự kiện phức tạp

Mô hình học máy đánh giá từng giao dịch riêng lẻ thông qua véc-tơ đặc trưng tĩnh, do đó khó nhận diện các hành vi gian lận biểu hiện dưới dạng chuỗi sự kiện theo thời gian. Trong thực tế, nhiều mẫu gian lận chỉ trở nên đáng ngờ khi xem xét tổng thể lịch sử giao dịch của một thẻ, ví dụ một loạt giao dịch dồn dập trong khoảng thời gian ngắn hoặc giao dịch có giá trị đột biến so với hành vi chi tiêu thông thường. Để phát hiện các mẫu hành vi này, đề án áp dụng kỹ thuật xử lý sự kiện phức tạp (Complex Event Processing, CEP) trên nền Apache Flink.

2.4.1. Apache Flink CEP

Bên cạnh việc sử dụng mô hình học máy để đánh giá từng giao dịch riêng lẻ, nhiều hành vi gian lận trong thực tế lại xuất hiện dưới dạng chuỗi sự kiện liên tiếp

thay vì chỉ một giao dịch đơn lẻ. Vì vậy, việc phân tích mối quan hệ giữa các sự kiện theo thời gian là cần thiết nhằm nâng cao khả năng phát hiện gian lận.

Xử lý sự kiện phức tạp (CEP) là kỹ thuật cho phép phát hiện các mẫu sự kiện phức tạp thông qua việc phân tích luồng dữ liệu liên tục theo thời gian thực. CEP không chỉ xem xét từng sự kiện riêng lẻ mà còn đánh giá mối liên hệ giữa nhiều sự kiện nhằm phát hiện các hành vi bất thường. [7]

Trong đề án này, thư viện CEP của Apache Flink được sử dụng để phát hiện các mẫu giao dịch nghi ngờ. Flink CEP cung cấp cơ chế xây dựng các mẫu sự kiện (Pattern) dựa trên: [7]

1. Thứ tự xuất hiện của các sự kiện;
2. Điều kiện ràng buộc giữa các sự kiện;
3. Khoảng thời gian xảy ra giữa các sự kiện;
4. Các thuộc tính của từng giao dịch.

Thông qua cơ chế ghép mẫu, Flink CEP có khả năng phát hiện các chuỗi hành vi bất thường ngay khi chúng xuất hiện trong luồng dữ liệu.

Apache Flink hỗ trợ CEP qua hai phương án triển khai:

Thư viện FlinkCEP với ngôn ngữ Java/Scala cung cấp một bộ DSL khai báo khuôn mẫu, tự động khớp mẫu trên dòng dữ liệu đã được phân vùng theo khoá và sinh kết quả qua hàm chọn mẫu.

PyFlink DataStream API cho phép thực hiện chính cùng mô hình CEP nhưng dưới dạng hàm xử lý có quản lý trạng thái: pattern được hiện thực hoá bằng logic Python trên trạng thái có quản lý khuôn mẫu, ràng buộc cửa sổ thời gian và hàm chọn được cài đặt tường minh.

Do đề án xây dựng pipeline phát hiện gian lận hoàn toàn bằng Python để đồng nhất với tầng mô hình học máy tránh phải duy trì thêm một môi trường Java tương đương, giải pháp được lựa chọn là CEP trên PyFlink. Cách tiếp cận này vẫn đảm bảo các tính chất quan trọng của một hệ CEP đúng nghĩa khi triển khai trên Apache Flink.

a) Định nghĩa khuôn mẫu

Bảng 2.1 Một số mẫu hành vi gian lận phổ biến được áp dụng

Mẫu hành vi	Mô tả
Giao dịch liên tiếp trong thời gian ngắn	Khách hàng thực hiện nhiều giao dịch với giá trị lớn trong khoảng thời gian vài phút, bất thường so với thói quen sử dụng thông thường.
Thay đổi vị trí địa lý bất thường	Hai giao dịch liên tiếp xuất hiện tại các khu vực địa lý cách xa nhau trong khoảng thời gian không đủ để di chuyển, gợi ý khả năng thông tin thẻ bị đánh cắp.
Gia tăng đột biến giá trị giao dịch	Giá trị giao dịch tăng đột biến so với lịch sử chi tiêu thông thường của khách hàng, vượt ngưỡng thống kê cho phép.
Xác thực thất bại nhiều lần liên tiếp	Nhiều lần xác thực hoặc thanh toán thất bại liên tiếp trong thời gian ngắn, có thể là dấu hiệu của hành vi dò quét thông tin thẻ.

Dựa trên đặc điểm dữ liệu Phát hiện gian lận thẻ tín dụng năm 2023 (28 đặc trưng đã ẩn danh hoá V1–V28 và trường tổng tiền giao dịch) và các mẫu hành vi gian lận phổ biến (Bảng 2.1), đề án triển khai ba khuôn mẫu CEP. Mỗi pattern là một hàm thuần, lấy đầu vào là trạng thái của một thẻ và giao dịch đang xử lý, trả về điểm số trong khoảng [0, 1]. Ba khuôn mẫu này do tác giả đề xuất, kế thừa chiến lược tổng hợp giao dịch [11] trong khung xử lý sự kiện phức tạp [7].

Mẫu 1 – Số lượng giao dịch trong thời gian ngắn. Phát hiện trường hợp một thẻ phát sinh nhiều giao dịch trong thời gian rất ngắn. Cửa sổ trượt 60 giây, ngưỡng cảnh báo 5 giao dịch:

$$\text{Velocity}(x) = \min\left(\frac{|\{t_i \in T_c : t_{now} - t_i < W_v\}|}{N_v}, 1\right) \quad (2.1)$$

trong đó T_c là tập mốc thời gian giao dịch của thẻ c, $W_v = 60$ giây, $N_v = 5$.

Mẫu 2 – Giá trị giao dịch tăng đột biến. Phát hiện giao dịch có giá trị lệch khỏi mức chi tiêu trung bình của thẻ. Lịch sử 20 giao dịch gần nhất của thẻ được dùng để tính trung bình và độ lệch chuẩn; z-score được ép về khoảng [0, 1] qua hàm sigmoid.

$$\text{Spike}(x) = \sigma(z) \quad (2.2)$$

$$z = \frac{a - \mu_c}{\sigma_c + \varepsilon} \quad (2.3)$$

[4]

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.4)$$

[4]

với a là giá trị giao dịch hiện tại, μ_c, σ_c là trung bình và độ lệch chuẩn lịch sử của thẻ c, $\varepsilon = 10^{-9}$ là số nhỏ để tránh chia cho không.

Mẫu 3 – Tần suất trung hạn. Phát hiện thẻ có tần suất giao dịch cao bất thường trong khoảng thời gian dài hơn pattern Velocity. Cửa sổ 5 phút, ngưỡng 15 giao dịch.

$$\text{Frequency}(x) = \min\left(\frac{|\{t_i \in T_c : t_{now} - t_i < W_f\}|}{N_f}, 1\right) \quad (2.5)$$

với $W_f = 300$ giây, $N_f = 15$. So sánh với mẫu 2, mẫu này có cửa sổ dài và ngưỡng cao hơn, nhằm phát hiện hành vi tần suất cao ổn định thay vì burst tức thời.

Tổng hợp ba mẫu. Điểm CEP cuối cùng được tính bằng phép hợp lớn nhất, tức một giao dịch chỉ cần khớp với một pattern là đã được CEP đánh giá đáng ngờ:

$$S_{CEP}(x) = \max\{\text{Velocity}(x),\ \text{Spike}(x),\ \text{Frequency}(x)\} \quad (2.6)$$

Cờ mẫu phát hiện gian lận được bật khi có ít nhất một pattern vượt ngưỡng cảnh báo riêng, lần lượt là $\theta_v = 0{,}6$, $\theta_s = 0{,}8$, $\theta_f = 0{,}6$ tương ứng cho Velocity, Spike, Frequency.

b) Luồng sự kiện

Luồng sự kiện là tập dữ liệu đầu vào mà Flink CEP tiếp nhận và áp dụng khuôn mẫu đã định nghĩa nhằm tìm kiếm các chuỗi sự kiện thỏa mãn điều kiện. Toàn bộ logic CEP được đóng gói trong một hàm xử lý quy trình của PyFlink. Luồng giao dịch từ Kafka được phân vùng theo mã số thẻ trước khi đi vào hàm xử lý, đảm bảo các giao dịch của cùng một thẻ được điều phối tới cùng một luồng và do đó truy cập đúng phần trạng thái của thẻ đó.

Việc dùng trạng thái này thay cho cấu trúc dữ liệu phía ứng dụng cho phép Flink quản lý trạng thái như một phần của bản chụp điểm kiểm tra. Khi quản lý tác vụ khởi động lại hoặc thay đổi mức song song, trạng thái CEP của mỗi thẻ được khôi phục tự động từ điểm kiểm tra gần nhất, đảm bảo tính nhất quán của các mẫu tích luỹ theo thời gian.

Sau khi phát hiện thành công một mẫu sự kiện, chức năng chọn mẫu sẽ thực hiện các hành động tiếp theo như:

1. Gắn cờ giao dịch nghi ngờ;
2. Gửi dữ liệu tới hệ thống xử lý tiếp theo.

2.4.2. Tích hợp CEP vào hệ thống xử lý dữ liệu

Trong hệ thống đề xuất, Flink CEP được triển khai như một thành phần bổ sung cho mô hình học máy, hoạt động song song nhằm tăng cường khả năng phát hiện gian lận theo thời gian thực. Sự kết hợp này cho phép hệ thống vừa khai thác năng lực phân tích đặc trưng của học máy, vừa tận dụng khả năng nhận diện chuỗi hành vi của CEP.

Kết quả ghép mẫu của Flink CEP được biểu diễn dưới dạng một điểm rủi ro, sau đó được tổng hợp cùng xác suất dự đoán của mô hình CatBoost để tạo thành điểm rủi ro chung của giao dịch. Công thức tổng hợp hai nguồn tín hiệu và cơ chế ra quyết định dựa trên điểm rủi ro được trình bày chi tiết ở mục 2.6.

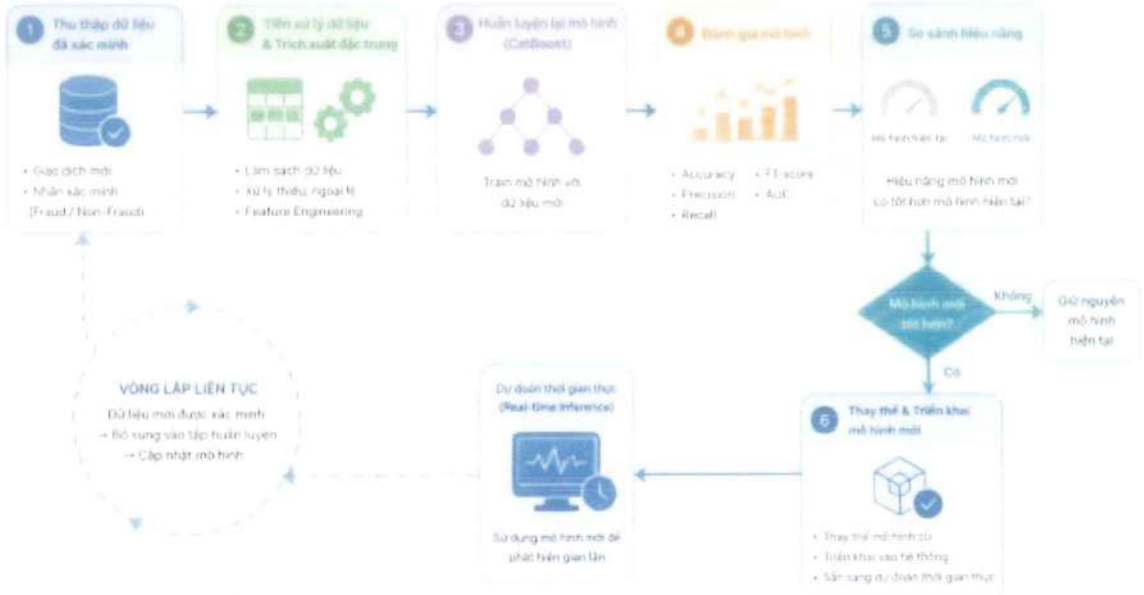
2.5. Cơ chế cập nhật mô hình động

Có nhiều phương pháp cập nhật mô hình khác nhau như cập nhật thủ công, học trực tuyến và huấn luyện lại định kỳ. Trong đề án này, cơ chế huấn luyện lại được lựa chọn do đảm bảo sự cân bằng giữa hiệu năng, độ ổn định và tính khả thi khi triển khai.

Phương pháp này cho phép hệ thống thu thập dữ liệu giao dịch mới đã được xác minh, thực hiện huấn luyện lại theo chu kỳ và đánh giá hiệu năng của mô hình mới trước khi thay thế mô hình hiện tại. Nhờ đó, hệ thống có khả năng thích nghi với sự thay đổi của dữ liệu mà không ảnh hưởng đến quá trình suy luận thời gian thực.

Quy trình cập nhật mô hình liên tục, được minh họa ở Hình 2.4, bao gồm các bước sau:

1. Thu thập các giao dịch đã được xác minh và bổ sung vào tập dữ liệu huấn luyện.
2. Thực hiện tiền xử lý dữ liệu và trích xuất đặc trưng.
3. Kích hoạt quá trình huấn luyện lại mô hình theo chu kỳ.
4. Đánh giá mô hình mới thông qua các chỉ số độ chính xác
5. Nếu mô hình mới đạt hiệu năng cao hơn mô hình hiện tại, hệ thống sẽ thực hiện thay thế mô hình và triển khai vào môi trường xử lý thời gian thực.



Hình 2.4 Quy trình cập nhật mô hình liên tục

Cơ chế cập nhật mô hình giúp hệ thống duy trì hiệu quả phát hiện gian lận trong thời gian dài, đồng thời nâng cao khả năng thích ứng trước các hình thức gian lận mới xuất hiện.

2.6. Hậu xử lý của hệ thống

2.6.1. Tổng hợp điểm rủi ro và ra quyết định

Sau khi mô hình học máy và Flink CEP hoàn tất việc đánh giá một giao dịch, hệ thống chuyển sang giai đoạn hậu xử lý nhằm tổng hợp các tín hiệu rủi ro và đưa ra quyết định cuối cùng. Thay vì chỉ dựa trên một nguồn tín hiệu duy nhất, hệ thống kết hợp hai thành phần gồm xác suất gian lận từ mô hình CatBoost và kết quả ghép mẫu từ Flink CEP để tính điểm rủi ro tổng hợp cho giao dịch.

Điểm rủi ro tổng hợp Risk(x) được tính bằng cách kết hợp tuyến tính có trọng số điểm số từ hai nguồn phân loại, dựa theo quy tắc tổng (sum rule) trong kết hợp bộ phân loại [13] - một trong những cách kết hợp được chứng minh là đơn giản và ổn định nhất. Theo đó, xác suất gian lận từ mô hình học máy CatBoost [21] được kết hợp với điểm rủi ro thu được từ xử lý sự kiện phức tạp (CEP) thông qua hai trọng số không âm α và β nhằm cân bằng mức đóng góp của hai nguồn tín hiệu, theo công thức (2.7):

$$\text{Risk}(x) = \alpha \cdot P(y = 1|x) + \beta \cdot S_{CEP}(x) \quad (2.7)$$

[13]

Trong đó P là xác suất gian lận do mô hình CatBoost dự đoán và S_CEP là điểm rủi ro thu được khi giao dịch khớp với một mẫu hành vi bất thường trong Flink CEP; α và β là hai trọng số không âm thể hiện mức độ đóng góp của hai nguồn tín hiệu, thỏa mãn ràng buộc α + β = 1. Một giao dịch được đánh dấu là gian lận nếu thỏa mãn một trong các điều kiện sau (cơ chế OR): xác suất gian lận từ mô hình vượt ngưỡng cao, có ít nhất một mẫu CEP khớp, hoặc điểm rủi ro tổng hợp đạt ngưỡng thấp. Cơ chế này cho phép hệ thống đạt độ chính xác cao hơn so với từng phương pháp đơn lẻ, đồng thời giảm thiểu tỷ lệ bỏ sót giao dịch gian lận.

Các trọng số α, β có thể được điều chỉnh theo đặc thù dữ liệu và yêu cầu nghiệp vụ của tổ chức tín dụng. Quyết định phân loại cuối cùng được đưa ra bằng cách so sánh điểm rủi ro tổng hợp với một ngưỡng θ cho trước:

$$\hat{y} = 1 \, nếu \, Risk(x) \geq \theta; \; \hat{y} = 0 \, nếu \, Risk(x) < \theta \quad (2.8)$$

Trong đó $\hat{y} = 1$ tương ứng với giao dịch bị đánh dấu là nghi ngờ gian lận và $\hat{y} = 0$ tương ứng với giao dịch hợp lệ. Ngưỡng θ được lựa chọn dựa trên sự cân bằng giữa tỷ lệ phát hiện gian lận và tỷ lệ cảnh báo sai, phù hợp với mức độ chấp nhận rủi ro của tổ chức.

2.6.2. Hành động ứng phó, cảnh báo và giám sát

Các thành phần hậu xử lý trình bày dưới đây, gồm hành động ứng phó theo mức rủi ro và cơ chế cảnh báo, giám sát, được mô tả ở mức thiết kế và là định hướng để hoàn thiện, mở rộng hệ thống trong tương lai, việc triển khai và đánh giá đầy đủ các thành phần này nằm ngoài phạm vi thử nghiệm của đề án. Tùy theo giá trị điểm rủi ro, hệ thống có thể áp dụng các hành động ứng phó khác nhau:

- Điểm rủi ro thấp: giao dịch được chấp nhận và xử lý bình thường;
- Điểm rủi ro trung bình: giao dịch được yêu cầu xác thực bổ sung, ví dụ xác thực một lần (OTP) hoặc xác thực sinh trắc học;

- Điểm rủi ro cao: giao dịch bị tạm giữ hoặc từ chối, đồng thời sinh cảnh báo gửi tới bộ phận giám sát.

Kết quả của quá trình hậu xử lý, cùng với phản hồi xác minh từ bộ phận nghiệp vụ, được lưu lại và đưa trở lại vòng cập nhật mô hình. Cơ chế phản hồi này giúp hệ thống liên tục học từ các giao dịch mới và duy trì hiệu quả phát hiện trong dài hạn.

Bên cạnh khả năng phát hiện gian lận, hệ thống cần được trang bị cơ chế cảnh báo và giám sát nhằm bảo đảm hoạt động ổn định, phát hiện kịp thời các giao dịch rủi ro cao và theo dõi sức khỏe của toàn bộ hệ thống theo thời gian thực.

Cơ chế cảnh báo được thiết kế theo nhiều mức độ ưu tiên tương ứng với điểm rủi ro của giao dịch. Khi một giao dịch vượt ngưỡng rủi ro hoặc khớp với một mẫu hành vi gian lận nghiêm trọng, hệ thống sinh cảnh báo và phân phối qua Kafka tới các kênh tiếp nhận như bảng điều khiển giám sát, thư điện tử hoặc dịch vụ nhắn tin nội bộ. Mỗi cảnh báo bao gồm thông tin về giao dịch, điểm rủi ro, lý do gắn cờ và mức độ ưu tiên, giúp bộ phận giám sát nhanh chóng xác minh và xử lý.

Cơ chế giám sát theo dõi đồng thời hai nhóm chỉ số. Nhóm thứ nhất là các chỉ số vận hành của hệ thống, bao gồm thông lượng xử lý, độ trễ đầu cuối, tỷ lệ lỗi và tình trạng hoạt động của Kafka và Flink. Nhóm thứ hai là các chỉ số chất lượng của mô hình, bao gồm tỷ lệ giao dịch bị gắn cờ, độ chính xác trên dữ liệu đã xác minh và mức độ thay đổi của hành vi gian lận theo thời gian. Khi một chỉ số vượt ngưỡng cho phép, hệ thống phát cảnh báo vận hành để quản trị viên can thiệp hoặc kích hoạt quá trình tái huấn luyện mô hình.

Các chỉ số giám sát được trực quan hóa trên bảng điều khiển thời gian thực, cho phép quản trị viên quan sát toàn cảnh hoạt động của hệ thống, theo dõi xu hướng gian lận và đánh giá hiệu quả của mô hình. Nhờ đó, hệ thống không chỉ phát hiện gian lận mà còn bảo đảm khả năng vận hành liên tục, ổn định và thích nghi trong môi trường tài chính có yêu cầu cao về độ tin cậy.

2.7. Kết luận chương

Chương 2 đã trình bày chi tiết quá trình hiện thực hóa kiến trúc phát hiện gian lận tín dụng thời gian thực được đề xuất ở Chương 1. Trước hết, chương giới thiệu kiến trúc tổng thể phân tầng của hệ thống cùng cách bốn thành phần phối hợp để giải quyết bài toán. Tiếp theo, chương lần lượt trình bày từng thành phần: mô hình học máy có giám sát CatBoost làm lõi phân loại; tầng xử lý dữ liệu luồng thời gian thực dựa trên Apache Kafka và Apache Flink; kỹ thuật xử lý sự kiện phức tạp với Flink CEP để phát hiện các chuỗi hành vi bất thường; và cơ chế cập nhật mô hình động giúp hệ thống thích nghi với sự thay đổi của hành vi gian lận theo thời gian. Cuối cùng, chương mô tả khâu hậu xử lý dựa trên điểm rủi ro tổng hợp cùng hệ thống cảnh báo và giám sát. Các nội dung trình bày trong Chương 2 là cơ sở cho việc thử nghiệm và đánh giá hệ thống ở Chương 3.

CHƯƠNG 3 - THỬ NGHIỆM, ĐÁNH GIÁ VÀ KẾT LUẬN

3.1. Giới thiệu thử nghiệm

Chương 2 đã trình bày cách hiện thực hóa hệ thống phát hiện gian lận tín dụng thời gian thực, giải quyết đồng thời bốn bài toán cốt lõi. Nhiệm vụ của chương này là kiểm chứng bằng thực nghiệm rằng hệ thống giải quyết được cả bốn bài toán đó. Cụ thể, bốn bài toán cùng thành phần đảm nhiệm và tiêu chí đánh giá tương ứng như sau:

(1) Phân loại chính xác từng giao dịch là hợp lệ hay gian lận. Đây là nhiệm vụ của mô hình học máy có giám sát (CatBoost), được đánh giá bằng các độ đo Precision, Recall, điểm F1 và AUC.

(2) Phát hiện các chuỗi hành vi gian lận theo thời gian mà mô hình xét từng giao dịch độc lập dễ bỏ sót. Đây là nhiệm vụ của kỹ thuật xử lý sự kiện phức tạp (CEP), được đánh giá qua mức cải thiện độ bao phủ khi bổ sung CEP vào mô hình học máy.

(3) Tiếp nhận và xử lý giao dịch liên tục với độ trễ thấp, không thất thoát dữ liệu dưới tải cao. Đây là nhiệm vụ của tầng xử lý dữ liệu luồng thời gian thực dựa trên Apache Kafka và Apache Flink, được đánh giá bằng thông lượng, độ trễ và tỉ lệ thất thoát giao dịch.

(4) Tự thích nghi với sự thay đổi của hành vi gian lận theo thời gian mà không gián đoạn dịch vụ. Đây là nhiệm vụ của cơ chế cập nhật mô hình động, được đánh giá qua chất lượng mô hình trước và sau các chu kỳ huấn luyện lại tự động.

Bốn bài toán trên được kiểm chứng qua hai kịch bản thử nghiệm bổ trợ nhau, mỗi kịch bản trả lời một câu hỏi nghiên cứu trung tâm. Kịch bản 1 tập trung vào năng lực phát hiện, qua đó kiểm chứng hai bài toán (1) và (2); Kịch bản 2 tập trung vào năng lực vận hành, qua đó kiểm chứng hai bài toán (3) và (4).

Kịch bản 1 trả lời câu hỏi *"Hệ thống phát hiện gian lận tốt đến mức nào?"*, tập trung đánh giá chất lượng phát hiện ở hai cấp độ: từ một mô hình học máy đơn lẻ đến kiến trúc kết hợp đa nguồn tín hiệu gồm xử lý sự kiện phức hợp (CEP) và học máy (ML).

Kịch bản 2 trả lời câu hỏi *"Hệ thống vận hành ổn định ra sao trong điều kiện thời gian thực và dữ liệu biến đổi liên tục?"*, tập trung đánh giá hiệu năng xử lý luồng và khả năng tự thích nghi của hệ thống thông qua cơ chế cập nhật mô hình động.

Hệ thống được thử nghiệm và đánh giá trên bộ dữ liệu, thiết bị và các độ đo phát hiện gian lận được trình bày dưới đây.

3.1.1. Tập dữ liệu

Đề án sử dụng tập dữ liệu "Phát hiện gian lận thẻ tín dụng năm 2023" (*Credit Card Fraud Detection Dataset 2023*) được công bố trên nền tảng Kaggle. Đây là tập dữ liệu chứa các giao dịch thẻ tín dụng của các chủ thẻ tại châu Âu trong năm 2023, được sử dụng rộng rãi trong các nghiên cứu về phát hiện gian lận. Tập dữ liệu gồm khoảng 568.630 bản ghi giao dịch, mỗi bản ghi tương ứng với một giao dịch và được gán nhãn là hợp lệ hoặc gian lận. [18]

Nhằm bảo đảm tính riêng tư của khách hàng, các đặc trưng gốc của giao dịch đã được ẩn danh và biến đổi thành các đặc trưng số V1 đến V28. Bên cạnh đó, tập dữ liệu còn chứa trường giá trị giao dịch và nhãn phân loại. Cấu trúc các trường dữ liệu được mô tả trong Bảng 3.1.

Bảng 3.1 Mô tả các trường dữ liệu của tập phát hiện gian lận thẻ tín dụng 2023

Trường dữ liệu	**Kiểu**	**Mô tả**
id	Số nguyên	Mã định danh duy nhất của mỗi giao dịch
V1 – V28	Số thực	28 đặc trưng đã được ẩn danh và biến đổi từ thông tin giao dịch gốc nhằm bảo vệ dữ liệu cá nhân
Amount	Số thực	Giá trị của giao dịch
Class	Nhị phân	Nhãn phân loại: 0 là giao dịch hợp lệ, 1 là giao dịch gian lận

Một đặc điểm quan trọng của tập dữ liệu là tính cân bằng giữa hai lớp giao dịch hợp lệ và gian lận, khác với nhiều tập dữ liệu gian lận thực tế thường có tỷ lệ mất cân bằng rất cao. Đặc điểm này thuận lợi cho quá trình huấn luyện và đánh giá mô hình. Tuy nhiên, để phản ánh sát hơn điều kiện thực tế, đề án vẫn chú trọng các độ đo phù hợp với bài toán phát hiện gian lận như độ bao phủ, điểm F1 và AUC bên cạnh độ chính xác tổng thể. Số lượng mẫu cụ thể của từng lớp trên toàn bộ tập dữ liệu, tập huấn luyện và tập kiểm thử được trình bày trong Bảng 3.2.

Bảng 3.2 Phân bố mẫu theo lớp của tập dữ liệu thử nghiệm

Tập dữ liệu	Giao dịch hợp lệ	Giao dịch gian lận	Tổng cộng
Toàn bộ	284.315	284.315	568.630
Huấn luyện (80%)	227.452	227.452	454.904
Kiểm thử (20%)	56.863	56.863	113.726

3.1.2. Môi trường thử nghiệm

Hệ thống được triển khai và thử nghiệm trong môi trường có cấu hình phần cứng và phần mềm được mô tả trong Bảng 3.3. Toàn bộ các thành phần (Apache Kafka, Apache Flink, dịch vụ huấn luyện lại và giao diện giám sát) được đóng gói và vận hành dưới dạng các container độc lập, cho phép tái lập môi trường một cách nhất quán.

Bảng 3.3 Môi trường thử nghiệm hệ thống

Thành phần	**Cấu hình / Phiên bản**
Bộ xử lý (CPU)	Intel Core i5-10400F @ 2.90GHz (6 nhân, 12 luồng)
Bộ nhớ (RAM)	16 GB
Hệ điều hành	Windows 11 Pro (build 10.0.26200)
Ngôn ngữ lập trình	Python 3.10.20
Thư viện mô hình	CatBoost 1.2.7 (scikit-learn 1.4.2, pandas 2.2.1)
Nền tảng truyền tải	Apache Kafka 7.6.0 (6 phân vùng)
Nền tảng xử lý luồng	Apache Flink 1.18.1 (mức song song 2,

	chu kỳ điểm kiểm tra 60 s)
Công cụ triển khai	Docker 29.3.1 và Docker Compose

3.1.3. Thiết lập chung và độ đo đánh giá

Tập dữ liệu sử dụng là *Phát hiện gian lận thẻ tín dụng năm 2023* phân bố cân bằng giữa hai lớp (gian lận và hợp lệ). Dữ liệu được chia thành hai phần theo tỉ lệ 80/20 bằng phương pháp lấy mẫu phân tầng nhằm bảo toàn tỉ lệ lớp ở cả hai tập, với hạt giống ngẫu nhiên cố định để bảo đảm khả năng tái lập kết quả. Tập huấn luyện được dùng để xây dựng mô hình phân loại ban đầu; tập kiểm thử đảm nhận đồng thời hai vai trò: vừa phục vụ đánh giá ngoại tuyến, vừa được phát lại vào hệ thống hàng đợi thông điệp để mô phỏng luồng giao dịch trực tuyến theo thời gian thực.

Khi được phát lại, mỗi giao dịch được bổ sung ba thuộc tính phục vụ phân tích: một mã giao dịch duy nhất, một mã thẻ dùng để phân vùng dữ liệu, và một nhãn nguồn phân biệt giao dịch thuộc luồng nền bình thường hay thuộc luồng tấn công mô phỏng. Ba thuộc tính này là điều kiện cần thiết để nền tảng xử lý luồng phân vùng dữ liệu theo từng thẻ, nhờ đó duy trì đúng trật tự thời gian của các giao dịch trên cùng một thẻ phục vụ CEP và để tách bạch các chỉ số đánh giá theo từng nguồn phát sinh. Toàn bộ các độ đo đánh giá sử dụng trong thử nghiệm được tổng hợp trong Bảng 3.4.

Bảng 3.4 Các chỉ số đánh giá thử nghiệm

Chỉ số	**Cách xác định**
Thông lượng, độ trễ trung bình, phân vị 95	Từ tầng giám sát và nhật ký xử lý luồng
Tỉ lệ thất thoát thông điệp	So sánh số giao dịch phát ra với số bản ghi kết quả nhận được
F_1, Precision, Recall, ROC-AUC trước cập nhật	Tính trên cửa sổ trượt ngay trước thời điểm phát sự kiện cập nhật
Các chỉ số tương ứng sau cập nhật	Lấy từ thông tin đính kèm sự kiện cập nhật mô hình
Độ trễ phát hiện thay đổi	Khoảng cách thời gian giữa thời điểm

	dịch vụ huấn luyện lại phát sự kiện và thời điểm tầng xử lý luồng nạp xong mô hình mới
Số giao dịch bị gián đoạn trong cửa sổ hoán đổi	Đếm số giao dịch xử lý từ lúc phát sự kiện đến khi mô hình mới có hiệu lực, kỳ vọng bằng 0 nhờ thao tác thay thế nguyên tử

Toàn bộ quy trình thử nghiệm được tiến hành tuần tự như sau. Trước hết, mô hình học máy gốc được huấn luyện và đánh giá ngoại tuyến để thiết lập đường nền. Tiếp theo, ở Kịch bản 1, dữ liệu được phát đồng thời theo hai luồng, luồng nền bình thường và luồng tấn công mô phỏng thẻ bị chiếm đoạt, và hệ thống được vận hành theo cấu hình kết hợp học máy với CEP để đo hiệu quả phát hiện tổng thể. Sau cùng, ở Kịch bản 2, hệ thống được nén tải để đo hiệu năng xử lý luồng và được kích hoạt dịch vụ huấn luyện lại tự động để đánh giá khả năng cập nhật mô hình động.

3.2. Kịch bản 1 – Năng lực phát hiện gian lận

Kịch bản này nhằm định lượng chất lượng phát hiện gian lận của kiến trúc kết hợp mô hình học máy CatBoost với xử lý sự kiện phức tạp theo công thức điểm rủi ro (2.7), ở hai cấp độ: chất lượng của riêng mô hình học máy và chất lượng của hệ thống kết hợp đa nguồn tín hiệu.

3.2.1. Thiết lập thử nghiệm

Dữ liệu dùng cho kịch bản này là tập kiểm thử chiếm 20% bộ dữ liệu Phát hiện gian lận thẻ tín dụng năm 2023 (mục 3.1.1), được phát lại theo thời gian thực qua Apache Kafka để mô phỏng luồng giao dịch trực tuyến. Lập luận của kịch bản như sau: nếu hệ thống vận hành trên dòng dữ liệu có nhãn thật này mà đạt độ bao phủ cao trong khi giữ tỷ lệ cảnh báo sai thấp, đồng thời độ bao phủ tăng lên khi bổ sung xử lý sự kiện phức tạp, thì có thể khẳng định hệ thống đáp ứng mục tiêu phân loại chính xác từng giao dịch (bài toán 1) và phát hiện chuỗi hành vi gian lận theo thời gian (bài toán 2).

a) Mô phỏng hai luồng dữ liệu thử nghiệm

Để mô phỏng sát sao nhất luồng dữ liệu giao dịch tín dụng cùng với tỉ lệ gian lận kèm các hành vi bất thường, dữ liệu được phát đồng thời theo hai luồng độc lập vào cùng một hệ thống xử lý luồng (với mức song song và chu kỳ điểm kiểm tra cố định), nhằm tái hiện môi trường vận hành thực tế trong đó giao dịch hợp lệ và hành vi tấn công đan xen.

Luồng nền (giao dịch bình thường). Luồng này mô phỏng tải giao dịch thông thường: các giao dịch được phát đều đặn, mỗi giao dịch được gán ngẫu nhiên cho một thẻ trong một tập thẻ lớn; cứ sau mỗi 20–30 giao dịch hợp lệ lại xen một giao dịch gian lận thật lấy ngẫu nhiên từ dữ liệu. Do mỗi thẻ chỉ xuất hiện rải rác, luồng nền hầu như không tự hình thành mẫu CEP nào; vai trò của nó là cung cấp một dòng dữ liệu có nhãn thực tế (gồm cả giao dịch hợp lệ lẫn gian lận) để đo các chỉ số phát hiện trên điều kiện gần với thực tế.

Luồng tấn công (mô phỏng kích hoạt các giao dịch bất thường đợt cho là gian). Luồng này chỉ phát các giao dịch gian lận thật và được thiết kế để chủ động kích hoạt đầy đủ ba mẫu CEP. Tại mỗi đợt tấn công, hệ thống tập trung tấn công vào một thẻ duy nhất (thuộc một vùng mã thẻ riêng, tách biệt với luồng nền) bằng một chuỗi giao dịch liên tiếp với nhịp phát chậm vừa phải, rồi luân phiên sang thẻ kế tiếp. Trong mỗi đợt tấn công, phần lớn giao dịch có giá trị thấp và ổn định để tạo nền lịch sử cho thẻ, xen vào đó là một vài giao dịch có giá trị rất cao nhằm tạo đột biến. Cách bố trí này khiến ba mẫu CEP lần lượt được kích hoạt theo tiến trình của đợt tấn công, như tóm tắt trong Bảng 3.5.

Bảng 3.5 Mô tả cách mô phỏng các yếu tố bất thường của giao dịch thẻ tín dụng

Mẫu CEP	**Cơ chế kích hoạt trong đợt tấn công**	**Điều kiện**
Vận tốc	Nhiều giao dịch dồn vào một thẻ trong cửa sổ ngắn	Kích hoạt từ những giao dịch đầu của đợt tấn công (cửa sổ 60 s)

Tần suất	Số giao dịch tích lũy trên thẻ vượt ngưỡng trong cửa sổ trung hạn	Kích hoạt khi đợt tấn công kéo dài (cửa sổ 300 s)
Đột biến giá trị	Chèn giao dịch giá trị rất cao trên nền lịch sử giá trị thấp	Kích hoạt tại các giao dịch đột biến

Vì toàn bộ giao dịch trong luồng tấn công đều là gian lận thật và có xác suất mô hình rất cao, chúng vừa được mô hình học máy dự đoán chính xác, vừa minh họa rõ ràng đóng góp của tín hiệu CEP, qua đó cho phép đánh giá riêng rẽ giá trị mà tầng CEP bổ sung cho hệ thống.

b) Tham số xử lý sự kiện phức tạp

Ba mẫu CEP được hiện thực hóa dưới dạng các hàm cho điểm trong khoảng [0,1], tính trên lịch sử giao dịch của từng thẻ. Ba mẫu điểm CEP dưới đây do tác giả đề xuất, kế thừa chiến lược tổng hợp giao dịch [11] trong khung xử lý sự kiện phức tạp [7].

Mẫu vận tốc giao dịch (Velocity) đo mức độ dồn dập của giao dịch trên một thẻ trong cửa sổ thời gian ngắn W_v. Gọi T_c là tập thời điểm các giao dịch gần đây của thẻ c và t_{now} là thời điểm hiện tại, điểm vận tốc được chuẩn hóa theo ngưỡng số giao dịch N_v:

$$\text{Velocity}(x) = min(\frac{|\{t_i \in T_c : t_{now} - t_i < W_v\}|}{N_v}, 1), W_v = 60 \text{ s}, N_v = 5 \qquad (3.1)$$

Mẫu tần suất (Frequency) có cấu trúc tương tự nhưng quan sát trên cửa sổ thời gian dài hơn W_f với ngưỡng N_f, nhằm phát hiện hành vi gian lận kéo dài:

$$\text{Frequency}(x) = min(\frac{|\{t_i \in T_c : t_{now} - t_i < W_f\}|}{N_f}, 1), W_f = 300 \text{ s}, N_f = 15 \qquad (3.2)$$

Mẫu đột biến giá trị (Spike) đánh giá mức bất thường của giá trị giao dịch a so với lịch sử chi tiêu của thẻ. Trước hết, giá trị được chuẩn hóa thành điểm z dựa

trên trung bình μ_c và độ lệch chuẩn σ_c của thẻ tính trên tối đa $H = 20$ giao dịch gần nhất (ε là hằng số nhỏ tránh chia cho 0); sau đó điểm z được ánh xạ về khoảng [0,1] qua hàm sigmoid:

$$\text{Spike}(x) = \sigma(z) = \frac{1}{1+e^{-z}}, z = \frac{a-\mu_c}{\sigma_c+\varepsilon} \quad (3.3)$$

[4]

Khi lịch sử của thẻ chưa đủ tối thiểu ba giao dịch, độ lệch chuẩn chưa đáng tin cậy nên điểm đột biến được gán giá trị nền trung tính 0,5 thay vì tính theo công thức trên.

Điểm CEP tổng hợp được lấy bằng giá trị lớn nhất trong ba mẫu, theo nguyên tắc một dấu hiệu bất thường đủ mạnh là đủ để nâng mức cảnh báo:

$$S_{CEP}(x) = max\{\text{Velocity}(x), \text{Spike}(x), \text{Frequency}(x)\} \quad (3.4)$$

c) Tham số mô hình học máy và tổng hợp rủi ro

Về thành phần đánh giá từ mô hình học máy, mô hình học máy được huấn luyện trên 29 đặc trưng đầu vào (28 đặc trưng ẩn danh đã qua biến đổi thành phần chính và một đặc trưng giá trị giao dịch). Hai chỉ số phản ánh riêng hiệu quả của mô hình học máy là ROC-AUC và PR-AUC. Diện tích dưới đường cong ROC (ROC-AUC) được dùng để đánh giá khả năng phân biệt hai lớp ở mọi ngưỡng quyết định. Đường cong ROC được dựng từ quan hệ giữa tỉ lệ dương thật (TPR) và tỉ lệ dương giả (FPR):

$$\text{TPR} = \frac{TP}{TP+FN} \quad (3.5)$$

[9]

$$\text{FPR} = \frac{FP}{FP+TN} \quad (3.6)$$

[9]

Ngoài ra, diện tích dưới đường cong Precision–Recall (PR-AUC) cũng được tính nhằm phản ánh sát hơn năng lực phát hiện trên lớp thiểu số. Kết quả của pha này được trực quan hóa qua ma trận nhầm lẫn, đường cong ROC và đường cong Precision–Recall, đồng thời mô hình huấn luyện được lưu lại để phục vụ pha kiểm thử trực tuyến tiếp theo.

Hệ thống được vận hành với bộ trọng số $\alpha = 0{,}625$ cho mô hình học máy và $\beta = 0{,}375$ cho xử lý sự kiện phức tạp (CEP), thỏa mãn ràng buộc $\alpha + \beta = 1$.

Ở cả hai cấu hình, một giao dịch được xác định là gian lận khi và chỉ khi điểm rủi ro tổng hợp đạt ngưỡng thấp θ_{low}. Cách định nghĩa này bảo đảm tính nhất quán giữa dự đoán và hành động: mọi giao dịch bị gắn cờ gian lận đều dẫn tới một hành động can thiệp, và ngược lại:

$$\text{is_fraud}(x) = \mathbb{1}[\text{Risk}(x) \geq \theta_{low}] \tag{3.7}$$

Trên cơ sở điểm rủi ro tổng hợp, hành động cuối cùng đối với mỗi giao dịch được quyết định theo ba mức ngưỡng: từ chối khi rủi ro cao, yêu cầu xác minh bổ sung khi rủi ro trung bình, và chấp nhận khi rủi ro thấp:

$$\text{action}(x) = \begin{cases} \text{từ chối} & \text{Risk}(x) \geq \theta_{high} \\ \text{yêu cầu xác minh} & \theta_{low} \leq \text{Risk}(x) < \theta_{high} \\ \text{chấp nhận} & \text{Risk}(x) < \theta_{low} \end{cases} \tag{3.8}$$

với hai ngưỡng hành động $\theta_{low} = 0{,}40$ và $\theta_{high} = 0{,}70$.

Dữ liệu sau phân tích, mỗi giao dịch sau khi xử lý sinh ra một bản ghi kết quả chứa đầy đủ các thành phần: nhãn dự đoán, xác suất gian lận, điểm CEP tổng hợp, điểm rủi ro cuối cùng và hành động được khuyến nghị. Trên cơ sở các bản ghi này, hệ thống được đánh giá theo hiệu quả phát hiện tổng thể của cấu hình kết hợp thông qua bộ độ đo đã trình bày (Precision, Recall, F_1, MCC, ROC-AUC, PR-AUC), tính trên cửa sổ trượt gồm các mẫu có nhãn thực tế gần nhất. Trong đó, các chỉ số dựa trên ngưỡng (Precision, Recall, F_1, MCC) phản ánh quyết định hợp nhất theo $\text{Risk}(x)$ - tức năng lực của toàn hệ thống kết hợp học máy và CEP, còn ROC-AUC và PR-AUC đo năng lực phân biệt của thành phần học máy. Kết quả được tổng hợp ở phần kết quả và nhận xét của kịch bản, kèm biểu đồ phân bố điểm rủi ro và phân bố ba loại hành động (chấp nhận, yêu cầu xác minh, từ chối) phản ánh năng lực phân loại chung của hệ thống.

3.2.2. Kết quả và nhận xét

Hiệu quả phát hiện gian lận của hệ thống được đánh giá dựa trên ma trận nhầm lẫn, trong đó các giao dịch được phân thành bốn nhóm: dương đúng (TP) là giao dịch gian lận được dự đoán đúng, dương sai (FP) là giao dịch hợp lệ bị dự đoán nhầm là gian lận, âm đúng (TN) là giao dịch hợp lệ được dự đoán đúng, và âm sai (FN) là giao dịch gian lận nhưng không được phát hiện. Ma trận nhầm lẫn của mô hình trên tập kiểm thử (113.726 giao dịch) được trình bày trong Bảng 3.6.

Bảng 3.6 Ma trận nhầm lẫn của mô hình trên tập kiểm thử

	Dự đoán: Hợp lệ	**Dự đoán: Gian lận**
Thực tế: Hợp lệ	TN = 56.799	FP = 64
Thực tế: Gian lận	FN = 0	TP = 56.863

Từ ma trận nhầm lẫn, các độ đo đánh giá được xác định và diễn giải như sau. Độ chính xác dương tính (Precision) phản ánh tỷ lệ giao dịch thực sự gian lận trong số các giao dịch bị hệ thống gắn cờ, chỉ số này càng cao thì tỷ lệ cảnh báo sai càng thấp. Độ bao phủ (Recall) phản ánh tỷ lệ giao dịch gian lận được phát hiện trên tổng số giao dịch gian lận thực tế, chỉ số này càng cao thì số gian lận bị bỏ sót càng ít. Điểm F1 là trung bình điều hòa của Precision và Recall, thể hiện sự cân bằng giữa hai chỉ số trên [20]. Hệ số tương quan Matthews (MCC) được dùng như một độ đo tổng hợp cân bằng cho dữ liệu mất cân bằng [6]. Độ chính xác tổng thể (Accuracy) phản ánh tỷ lệ giao dịch được phân loại đúng trên toàn bộ tập dữ liệu. Ngoài ra, đề án sử dụng chỉ số AUC, diện tích dưới đường cong ROC, để đánh giá khả năng phân biệt giữa hai lớp giao dịch ở mọi ngưỡng phân loại; giá trị AUC càng gần 1 thể hiện khả năng phân biệt càng tốt. [9]

Do dữ liệu gian lận có đặc thù riêng và hậu quả của việc bỏ sót giao dịch gian lận thường nghiêm trọng, đề án đặc biệt chú trọng các độ đo Recall, F1 và AUC nhằm hạn chế bỏ sót gian lận, đồng thời kiểm soát tỷ lệ cảnh báo sai thông qua Precision.

Trước khi đi vào phân tích định lượng, phần này giới thiệu khái quát giao diện giám sát thời gian thực của hệ thống, công cụ được sử dụng để theo dõi luồng

giao dịch, quan sát kết quả phát hiện và thu thập các số liệu đánh giá được trình bày ở những phần tiếp theo.

Hình 3.1 và Hình 3.2 minh họa giao diện giám sát thời gian thực của hệ thống, hiển thị các chỉ số phát hiện (Precision, Recall, F1, ROC-AUC, PR-AUC, MCC) được tính trực tiếp trên luồng giao dịch. Hai hình này mang tính minh họa, cho thấy hệ thống đã được hiện thực hóa và vận hành trên luồng giao dịch; chúng đóng vai trò là công cụ trực quan để thu thập số liệu trực tuyến, còn kết quả đánh giá định lượng được trình bày và phân tích qua các bảng số liệu ở những phần tiếp theo, không rút ra trực tiếp từ hai hình này.



Hình 3.1 Các chỉ số đánh giá của mô hình trên ứng dụng



Hình 3.2 Các chỉ số đánh giá giao dịch chế độ thời gian thực

Trên cơ sở dữ liệu thu thập từ giao diện giám sát này, các phần tiếp theo phân tích chi tiết hiệu quả phát hiện gian lận, hiệu năng xử lý thời gian thực và khả năng cập nhật mô hình động của hệ thống

a) Hiệu quả phát hiện của mô hình học máy

Kết quả đánh giá mô hình CatBoost trên tập dữ liệu kiểm thử (113.726 giao dịch, phân bố cân bằng giữa hai lớp) theo các độ đo, được tổng hợp trong Bảng 3.7.

Bảng 3.7 Kết quả đánh giá mô hình CatBoost trên tập kiểm thử

Độ đo	**Giá trị**
Độ chính xác tổng thể (Accuracy)	0,9994
Precision	0,9989
Recall	1,0000
Điểm F1	0,9994
AUC	1,0000

Nhận xét: Mô hình CatBoost đạt chất lượng phân loại rất cao trên tập kiểm thử, với độ bao phủ tuyệt đối (Recall = 1,0000, không bỏ sót giao dịch gian lận nào

trong tổng số 56.863 giao dịch gian lận) và độ chính xác gần như tuyệt đối (Precision = 0,9989, chỉ 64 trên 56.927 cảnh báo là sai). Giá trị AUC bằng 1,0000 cho thấy mô hình phân tách hai lớp gần như hoàn hảo ở mọi ngưỡng. Kết quả này phản ánh đặc điểm của tập dữ liệu *Phát hiện gian lận thẻ tín dụng năm 2023* (cân bằng lớp, đặc trưng đã qua biến đổi thành phần chính có tính phân tách cao), đồng thời khẳng định CatBoost là lựa chọn phù hợp làm thành phần học máy nền tảng của hệ thống.

b) Hiệu quả của hệ thống kết hợp trên luồng trực tuyến

Để đánh giá hiệu quả trong điều kiện vận hành thực tế, hệ thống được chạy ở chế độ trực tuyến với hai luồng dữ liệu đan xen (luồng giao dịch nền và luồng tấn công mô phỏng thẻ bị chiếm đoạt) như mô tả ở phần thiết lập của kịch bản này. Các chỉ số được tính trên hơn 30.000 giao dịch có nhãn thực tế thu thập từ luồng kết quả. Do mỗi bản ghi kết quả lưu đồng thời xác suất của mô hình học máy và điểm rủi ro hợp nhất, có thể tách bạch hiệu quả của riêng thành phần học máy và của toàn hệ thống kết hợp ngay trên cùng một phiên chạy, không cần vận hành lại. Kết quả được trình bày trong Bảng 3.8.

Bảng 3.8 So sánh hiệu quả phát hiện gian lận khi kết hợp CEP

Phương pháp	Precision	Recall	F1	AUC
Chỉ sử dụng CatBoost	0,9938	0,9767	0,9852	0,9992
CatBoost kết hợp CEP	0,9541	0,9878	0,9707	0,9992

Nhận xét: Trên luồng trực tuyến, việc bổ sung tầng xử lý sự kiện phức tạp (CEP) làm tăng rõ rệt độ bao phủ: Recall tăng từ 0,9767 lên 0,9878, tương ứng giảm số giao dịch gian lận bị bỏ sót gần một nửa (từ 165 xuống còn 86 trường hợp). Đây chính là giá trị cốt lõi của CEP, phát hiện được các chuỗi hành vi gian lận theo thời gian (giao dịch dồn dập, tần suất cao, đột biến giá trị) mà mô hình học máy xét từng giao dịch độc lập khó nhận diện. Đổi lại, độ chính xác giảm từ 0,9938 xuống 0,9541 do tầng CEP tạo thêm một số cảnh báo sai, khiến điểm F1 tổng thể giảm nhẹ. Chỉ số

AUC không thay đổi (0,9992) ở cả hai trường hợp, phù hợp với bản chất của chỉ số này: AUC đo năng lực phân biệt của xác suất do mô hình học máy sinh ra, trong khi CEP chỉ tham gia ở khâu ra quyết định theo ngưỡng. Như vậy, việc kết hợp CEP thể hiện sự đánh đổi hợp lý giữa độ bao phủ và độ chính xác, ưu tiên giảm thiểu bỏ sót gian lận, phù hợp với định hướng nghiệp vụ của hệ thống.

Cũng cần lưu ý sự chênh lệch giữa kết quả ngoại tuyến (Bảng 3.7) và trực tuyến (Bảng 3.8). Riêng thành phần CatBoost, trên tập kiểm thử tĩnh đạt Recall = 1,0000 và Precision = 0,9989, nhưng trên luồng trực tuyến chỉ đạt Recall = 0,9767 và Precision = 0,9938, thấp hơn một chút. Khác biệt này đến từ điều kiện đánh giá: Bảng 3.7 được đo trên tập kiểm thử tĩnh, cân bằng lớp và cùng phân phối với dữ liệu huấn luyện, là điều kiện gần như lý tưởng; trong khi Bảng 3.8 được đo trên luồng phát lại theo thời gian thực, nơi luồng giao dịch nền và luồng tấn công đan xen với thành phần, thứ tự và tỷ lệ lớp khác với tập kiểm thử, đồng thời các chỉ số được tính trên cửa sổ trượt gồm các mẫu có nhãn gần nhất thay vì toàn bộ tập dữ liệu. Mức chênh lệch nhỏ cho thấy mô hình vẫn duy trì chất lượng cao khi chuyển từ điều kiện đánh giá lý tưởng sang môi trường vận hành thực tế.

3.3. Kịch bản 2 – Vận hành thời gian thực và cập nhật động

Kịch bản này đánh giá hai khía cạnh vận hành: thứ nhất, hệ thống xử lý được bao nhiêu giao dịch mỗi giây và mất bao lâu để trả về kết quả; thứ hai, hệ thống có thể tự cập nhật mô hình khi dữ liệu thay đổi mà không phải dừng dịch vụ hay không.

3.3.1. Thiết lập thử nghiệm

Dữ liệu dùng cho kịch bản này cũng là tập kiểm thử của bộ dữ liệu năm 2023, được phát lại qua Apache Kafka với nhịp phát tăng dần nhằm tạo tải cao. Lập luận của kịch bản: nếu hệ thống xử lý được lượng lớn giao dịch mỗi giây với độ trễ trong ngưỡng cho phép và không thất thoát dữ liệu, đồng thời dịch vụ huấn luyện lại tự động cập nhật mô hình mà không làm gián đoạn luồng, thì có thể khẳng định hệ thống đáp ứng mục tiêu vận hành thời gian thực (bài toán 3) và tự thích nghi với thay đổi của hành vi gian lận theo thời gian (bài toán 4).

a) Hiệu năng xử lý luồng

Ban đầu dữ liệu được phát ở nhịp ổn định để đo độ trễ trong điều kiện bình thường. Sau đó, nhịp phát được tăng dần (rút ngắn khoảng cách giữa các giao dịch) cho đến khi hệ thống đạt giới hạn, nhằm xác định ngưỡng tải tối đa mà hệ thống vẫn xử lý ổn định.

Hai đại lượng được đo là độ trễ và thông lượng.

Độ trễ là thời gian từ lúc một giao dịch được phát đi đến lúc nhận được kết quả xử lý. Mỗi giao dịch được đóng dấu hai mốc thời gian: thời điểm phát t_{prod} và thời điểm nhận kết quả t_{cons}. Độ trễ của giao dịch thứ i là hiệu của hai mốc này:

$$L_i = t_{cons,i} - t_{prod,i} \tag{3.9}$$

Trên toàn phiên, độ trễ được tổng hợp bằng hai con số: giá trị trung bình $\overline{L}$ và phân vị thứ 95 L_{p95}. Phân vị 95 là mức độ trễ mà 95% giao dịch nằm dưới nó (chỉ 5% giao dịch chậm hơn); con số này cho biết hệ thống phản hồi tệ đến đâu trong những tình huống xấu, điều mà giá trị trung bình thường che giấu:

$$\overline{L} = \frac{1}{n}\sum_{i=1}^{n} L_i \tag{3.10}$$

$$L_{p95} = \text{percentile}_{95}\{L_1, L_2, \ldots, L_n\} \tag{3.11}$$

Thông lượng là số giao dịch hệ thống xử lý xong trong một giây. Để theo dõi theo thời gian thực, thông lượng được tính trên một cửa sổ trượt rộng $L_w = 10$ giây: đếm số giao dịch hoàn tất trong cửa sổ rồi chia cho độ rộng cửa sổ:

$$\text{Throughput}(t) = \frac{|\{i : t - L_w \le t_{cons,i} \le t\}|}{L_w}, L_w = 10\ \text{s} \tag{3.12}$$

Cuối cùng, để kiểm tra hệ thống có làm mất giao dịch dưới tải cao hay không, ta so sánh tổng số giao dịch đã phát với số kết quả thực sự nhận được; tỉ lệ chênh lệch chính là tỉ lệ thất thoát.

b) Cập nhật mô hình động và hoán đổi nóng

Pha này kiểm tra khả năng tự học lại của hệ thống. Trong thực tế, hành vi gian lận thay đổi theo thời gian nên một mô hình cố định sẽ dần kém chính xác. Vì vậy hệ thống có thêm một dịch vụ huấn luyện lại chạy song song, hoạt động theo ba bước:

1. Thu thập dữ liệu mới. Dịch vụ liên tục gom các giao dịch đã có nhãn từ luồng và tích lũy vào bộ đệm.

2. Huấn luyện lại định kỳ. Sau mỗi chu kỳ thời gian định trước (được rút ngắn để phục vụ thử nghiệm) và khi đã đủ số mẫu tối thiểu, một mô hình mới được huấn luyện trên dữ liệu vừa gom. Mô hình chỉ học trên 29 đặc trưng gốc; điểm CEP là nguồn tín hiệu song song, không dùng làm đặc trưng đầu vào.

3. Quyết định triển khai. Mô hình mới chỉ được đưa vào thay thế nếu đạt chất lượng tối thiểu, đo bằng F_1:

$$\text{Deploy}(M_{new}) \Leftrightarrow F_1(M_{new}) \geq F_{1,min}, F_{1,min} = 0{,}90 \quad (3.13)$$

Nếu mô hình mới không đạt ngưỡng, hệ thống giữ nguyên mô hình cũ và ghi nhận một sự kiện "bỏ qua cập nhật".

Điểm cốt lõi của pha này là cơ chế hoán đổi nóng (hot-swap): khi mô hình mới được chấp nhận, nó được ghi ra một tệp tạm rồi thay thế nguyên tử lên mô hình đang chạy. "Nguyên tử" nghĩa là thao tác thay thế diễn ra tức thời và trọn vẹn, không có khoảnh khắc nào hệ thống đọc phải một mô hình dở dang. Tầng xử lý luồng nhận ra có mô hình mới ở giao dịch kế tiếp và tự nạp lại mà không cần dừng dịch vụ, nhờ đó luồng giao dịch không bị gián đoạn trong suốt quá trình cập nhật.

Kết quả được tổng hợp cùng các chỉ số F_1 và AUC trước và sau cập nhật, kèm trục thời gian các sự kiện huấn luyện lại nhằm minh họa tính liên tục của dịch vụ.

3.3.2. Kết quả và nhận xét

a) Hiệu năng xử lý thời gian thực

Về khả năng xử lý thời gian thực, hệ thống được đánh giá thông qua thông lượng và độ trễ khi mô phỏng luồng giao dịch liên tục qua Kafka và Flink. Độ trễ đầu cuối được đo bằng cách bơm các giao dịch thăm dò có đóng dấu thời gian và ghi nhận thời điểm kết quả tương ứng quay về; thông lượng tối đa được đo bằng cách phát một loạt giao dịch dồn dập và quan sát tốc độ hệ thống trả kết quả. Kết quả được trình bày trong Bảng 3.9.

Bảng 3.9 Hiệu năng xử lý luồng của hệ thống

Chỉ số hiệu năng	**Giá trị**
Thông lượng tối đa (giao dịch/giây)	≈ 623
Tỷ lệ thất thoát thông điệp	0% (1.500/1.500 giao dịch)
Độ trễ trung bình (ms)	≈ 2.030
Độ trễ phân vị 95 (ms)	≈ 3.070

Nhận xét: Hệ thống xử lý ổn định loạt 1.500 giao dịch dồn dập mà không thất thoát bản ghi nào (tỷ lệ mất = 0%), với thông lượng tối đa đạt khoảng 623 giao dịch/giây, cao hơn nhiều lần so với nhịp phát của luồng thử nghiệm, cho thấy hệ thống còn nhiều dư địa tải. Độ trễ đầu cuối trung bình khoảng 2 giây và phân vị 95 khoảng 3 giây; mức độ trễ này chủ yếu đến từ các chặng truyền qua Kafka và bước suy luận mô hình thực hiện bằng Python trên từng giao dịch trong Flink. Với bài toán sàng lọc gian lận gần thời gian thực (ngưỡng đáp ứng dưới vài giây), mức độ trễ này là chấp nhận được; nếu cần giảm thêm, có thể tăng mức song song của Flink hoặc tối ưu bước suy luận theo lô.

b) Hiệu quả của cơ chế cập nhật mô hình động

Cuối cùng, hiệu quả của cơ chế cập nhật mô hình động được đánh giá qua quá trình hệ thống tự huấn luyện lại liên tục trên dữ liệu luồng tích lũy. Trong suốt phiên vận hành kéo dài khoảng 15 giờ, dịch vụ huấn luyện lại đã thực hiện 12 chu kỳ cập nhật (mỗi chu kỳ khoảng một giờ), cả 12 mô hình mới đều vượt ngưỡng chất lượng tối thiểu và được triển khai thành công bằng cơ chế hoán đổi nóng, không có

chu kỳ nào bị bỏ qua và không làm gián đoạn dịch vụ. Bảng 3.10 so sánh hiệu năng của mô hình ở lần cập nhật đầu tiên và lần cập nhật mới nhất.

Bảng 3.10 Hiệu năng mô hình trước và sau khi cập nhật

Mô hình	Số mẫu tích lũy	F1	AUC
Lần cập nhật đầu (v1)	3.164	0,9835	0,9982
Lần cập nhật mới nhất (v12)	26.857	0,9937	1,0000

Nhận xét: Khi lượng dữ liệu luồng tích lũy tăng dần (từ khoảng 3.000 lên gần 27.000 mẫu), chất lượng mô hình được cải thiện ổn định: F1 tăng từ 0,9835 lên 0,9937 và AUC tăng từ 0,9982 lên 1,0000. Điều này cho thấy cơ chế cập nhật động giúp mô hình liên tục thích nghi với dữ liệu mới mà vẫn duy trì chất lượng cao. Cơ chế kiểm soát chất lượng (chỉ triển khai mô hình mới khi đạt ngưỡng F1 tối thiểu) bảo đảm hệ thống không bao giờ bị thay thế bằng một mô hình kém hơn, trong khi cơ chế hoán đổi nóng bảo đảm việc cập nhật diễn ra liền mạch. Cần lưu ý rằng trong thử nghiệm này, dữ liệu phát lại vừa được dùng để đánh giá vừa được tích lũy để huấn luyện lại, nên các chỉ số sau cập nhật chủ yếu phản ánh khả năng thích nghi của mô hình với phân phối dữ liệu của phiên chạy; việc kiểm chứng trên một tập dữ liệu độc lập là hướng cần bổ sung trong các nghiên cứu tiếp theo.

3.4. Kết luận chương

Chương 3 đã trình bày quá trình thử nghiệm và đánh giá hệ thống phát hiện gian lận tín dụng theo thời gian thực được đề xuất. Trên cơ sở tập dữ liệu *Phát hiện gian lận thẻ tín dụng năm 2023*, chương đã thiết kế hai kịch bản thử nghiệm nhằm lần lượt trả lời hai câu hỏi nghiên cứu trung tâm: hệ thống phát hiện gian lận tốt đến mức nào, và hệ thống vận hành ổn định ra sao trong điều kiện thời gian thực và dữ liệu biến đổi liên tục.

Kết quả thử nghiệm cho thấy hệ thống đạt được các mục tiêu đặt ra. Về năng lực phát hiện, mô hình học máy CatBoost đạt chất lượng gần như tuyệt đối trên tập kiểm thử (F1 và AUC xấp xỉ 1,0); trên luồng trực tuyến, kiến trúc kết hợp học máy

với xử lý sự kiện phức tạp (CEP) duy trì độ bao phủ cao (Recall ≈ 0,99) và đặc biệt giúp giảm gần một nửa số giao dịch gian lận bị bỏ sót so với khi chỉ dùng mô hình đơn lẻ, khẳng định giá trị của việc bổ sung tín hiệu CEP trong phát hiện các hành vi gian lận theo chuỗi. Về hiệu năng, hệ thống đạt thông lượng khoảng 623 giao dịch/giây, không thất thoát dữ liệu và độ trễ ở mức vài giây, đáp ứng yêu cầu xử lý gần thời gian thực. Về khả năng thích nghi, cơ chế cập nhật mô hình động vận hành ổn định qua nhiều chu kỳ huấn luyện lại liên tục mà không gây gián đoạn dịch vụ, giúp mô hình liên tục cải thiện theo dữ liệu mới.

Bên cạnh các kết quả đạt được, hệ thống vẫn còn nhiều hướng có thể tiếp tục hoàn thiện trong tương lai, như tinh chỉnh siêu tham số mô hình, tăng mức song song của Apache Kafka và Apache Flink, bổ sung các mẫu CEP mới và làm giàu đặc trưng từ nhiều nguồn dữ liệu hơn. Những kết quả và định hướng này là cơ sở để đưa ra các kết luận và kiến nghị chung của đề án ở chương tiếp theo.

KẾT LUẬN VÀ KIẾN NGHỊ

Đề án đã nghiên cứu bài toán phát hiện gian lận giao dịch tín dụng trong môi trường dữ liệu lớn và thời gian thực. Trên cơ sở khảo sát các đặc điểm của dữ liệu giao dịch tín dụng và phân tích các phương pháp phát hiện bất thường truyền thống lẫn hiện đại, đề án đã đề xuất và hiện thực hóa một kiến trúc lai cho hệ thống phát hiện gian lận hoạt động theo thời gian thực.

Hệ thống được xây dựng kết hợp bốn thành phần bổ trợ lẫn nhau: mô hình học máy có giám sát CatBoost làm lõi phân loại, hạ tầng xử lý dữ liệu luồng dựa trên Apache Kafka và Apache Flink, kỹ thuật xử lý sự kiện phức tạp (CEP) để phát hiện các chuỗi hành vi bất thường, và cơ chế cập nhật mô hình động giúp hệ thống tự thích nghi với sự thay đổi của hành vi gian lận theo thời gian. Kết quả thử nghiệm cho thấy hệ thống đạt độ chính xác và độ bao phủ cao trong phát hiện gian lận, trong đó việc bổ sung CEP giúp giảm đáng kể số giao dịch gian lận bị bỏ sót so với khi chỉ sử dụng mô hình đơn lẻ; đồng thời hệ thống xử lý ổn định theo thời gian thực với thông lượng cao, không thất thoát dữ liệu, và cơ chế cập nhật mô hình động vận hành liên tục qua nhiều chu kỳ mà không làm gián đoạn dịch vụ.

Kết quả nghiên cứu của đề án có thể được ứng dụng trong việc xây dựng và triển khai các hệ thống giám sát giao dịch tín dụng thời gian thực, góp phần nâng cao hiệu quả phát hiện gian lận, giảm thiểu rủi ro tài chính và hỗ trợ quá trình ra quyết định trong lĩnh vực ngân hàng và tài chính.

Trong thời gian tới, đề án có thể được phát triển theo một số hướng. Thứ nhất, kiểm chứng hệ thống trên các tập dữ liệu giao dịch thực tế có mức mất cân bằng cao và khó phân tách hơn nhằm đánh giá khả năng tổng quát hóa sát với điều kiện vận hành thật. Thứ hai, mở rộng tầng phân tích bằng các mô hình học sâu xử lý chuỗi và làm giàu đặc trưng từ nhiều nguồn dữ liệu như thông tin thiết bị, vị trí địa lý và hành vi người dùng. Thứ ba, bổ sung và tinh chỉnh các mẫu CEP cùng cơ chế cập nhật mô hình nhằm nâng cao độ chính xác và giảm tỷ lệ cảnh báo sai. Thứ tư, mở rộng quy mô hệ thống theo chiều ngang để đáp ứng khối lượng giao dịch lớn hơn trong môi trường sản xuất.

DANH MỤC TÀI LIỆU THAM KHẢO

[1] M. AlEmad (2022), "Credit Card Fraud Detection Using Machine Learning," luận văn thạc sĩ, *Rochester Institute of Technology*.

[2] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari (2021), "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," *Journal of Big Data*, 8, art. 151.

[3] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland (2011), "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, 50(3), pp. 602-613.

[4] C. M. Bishop (2006), *Pattern Recognition and Machine Learning*, Springer, ch. 4.

[5] P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, and K. Tzoumas (2015), "Apache Flink: Stream and batch processing in a single engine," *IEEE Data Engineering Bulletin*, 38(4), pp. 28-38.

[6] D. Chicco and G. Jurman (2020), "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, 21, art. 6.

[7] G. Cugola and A. Margara (2012), "Processing flows of information: From data stream to complex event processing," *ACM Computing Surveys*, 44(3), art. 15, pp. 1-62.

[8] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi (2014), "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert Systems with Applications*, 41(10), pp. 4915-4928.

[9] T. Fawcett (2006), "An introduction to ROC analysis," *Pattern Recognition Letters*, 27(8), pp. 861-874.

[10] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia (2014), "A survey on concept drift adaptation," *ACM Computing Surveys*, 46(4), art. 44, pp. 1-37.

[11] S. Jha, M. Guillen, and J. C. Westland (2012), "Employing transaction aggregation strategy to detect credit card fraud," *Expert Systems with Applications*, 39(16), pp. 12650-12657.

[12] J. Jurgovsky, M. Granitzer, K. Ziegler, S. Calabretto, P.-E. Portier, L. He-Guelton, and O. Caelen (2018), "Sequence classification for credit-card fraud detection," *Expert Systems with Applications*, 100, pp. 234-245.

[13] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas (1998), "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3), pp. 226-239.

[14] J. Kreps, N. Narkhede, and J. Rao (2011), "Kafka: A distributed messaging system for log processing," in *Proceedings of the NetDB Workshop*, pp. 1-7.

[15] Y. LeCun, Y. Bengio, and G. Hinton (2015), "Deep learning," *Nature*, 521(7553), pp. 436-444.

[16] K. J. Leonard (1995), "The development of a rule based expert system model for fraud alert in consumer credit," *European Journal of Operational Research*, 80(2), pp. 350-356.

[17] F. T. Liu, K. M. Ting, and Z.-H. Zhou (2008), "Isolation Forest," in *Proc. 8th IEEE International Conference on Data Mining (ICDM)*, pp. 413-422.

[18] Nelgiriyewithana (2023), *Credit Card Fraud Detection Dataset 2023*, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/nelgiriyewithana/credit-card-fraud-detection-dataset-2023, truy nhập ngày 15/08/2024.

[19] S. Ounacer, H. A. El Bour, Y. Oubrahim, M. Y. Ghoumari, and M. Azzouazi (2018), "Using isolation forest in anomaly detection: the case of credit card transactions," *Periodicals of Engineering and Natural Sciences*, 6(2), pp. 394-400.

[20] D. M. W. Powers (2011), "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, 2(1), pp. 37-63.

[21] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin (2018), "CatBoost: unbiased boosting with categorical features," in *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 6639-6649.

[22] Rajeshwari U and Sathish Babu B (2016), "Real-time credit card fraud detection using Streaming Analytics," in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, pp. 439-444.

[23] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi (2018), "Credit Card Fraud Detection Using AdaBoost and Majority Voting," *IEEE Access*, 6, pp. 14277-14284.

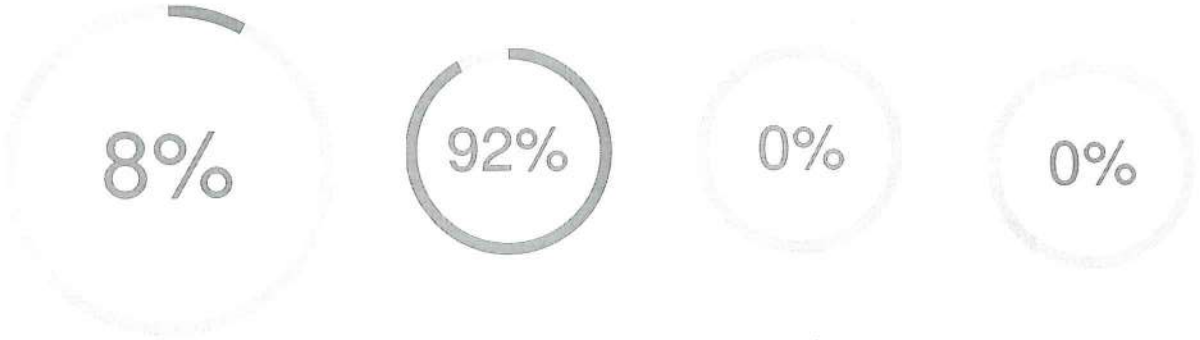
KiểmTraTàiLiệu

BÁO CÁO KIỂM TRA TRÙNG LẶP

Thông tin tài liệu

Tên tài liệu:	TRỊNH ĐÌNH MINH - Luận văn (sau bv)
Tác giả:	Minh Trịnh
Điểm trùng lặp:	8
Thời gian tải lên:	22:12 09/07/2026
Thời gian sinh báo cáo:	22:20 09/07/2026
Các trang kiểm tra:	66/66 trang

Kết quả kiểm tra trùng lặp



Có 8% nội dung trùng lặp

Có 92% 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

tailieu.vn arxiv.org 123docz.net

Phạm Hoàng Duy

Trịnh Đình Minh

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BÁO CÁO GIẢI TRÌNH
SỬA CHỮA, HOÀN THIỆN ĐỀ ÁN TỐT NGHIỆP

Họ và tên học viên: Trịnh Đình Minh
Chuyên ngành: Hệ thống thông tin
Khóa: 2023 đợt 1

Tên đề tài: Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục

Người hướng dẫn khoa học: TS. Phạm Hoàng Duy

Ngày bảo vệ: 08/07/2026

Các nội dung học viên đã sửa chữa, bổ sung trong đề án tốt nghiệp theo ý kiến đóng góp của Hội đồng chấm đề án tốt nghiệp:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Bổ sung trích dẫn các công thức	Học viên đã rà soát, bổ sung thêm vào đầy đủ trích dẫn công thức

Hà Nội, ngày 10 tháng 07 năm 2026

Ký xác nhận của

CHỦ TỊCH HỘI ĐỒNG CHẤM ĐỀ ÁN	THƯ KÝ HỘI ĐỒNG	NGƯỜI HƯỚNG DẪN KHOA HỌC	HỌC VIÊN
PGS. TSKH. Hoàng Đăng Hải	TS. Đỗ Tiến Dũng	TS. Phạm Hoàng Duy	Trịnh Đình Minh

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BIÊN BẢN
HỌP HỘI ĐỒNG CHẤM ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

Căn cứ quyết định số Quyết định số 1895/QĐ-HV ngày 29 tháng 06 năm 2026 của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ. Hội đồng đã họp vào hồi..9.....giờ..45.....phút, ngày 08 tháng 07 năm 2026 tại Học viện Công nghệ Bưu chính Viễn thông để chấm đề án tốt nghiệp thạc sĩ cho:

Học viên: **Trịnh Đình Minh**

Tên đề án tốt nghiệp: Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục

Chuyên ngành: Hệ thống thông tin Mã số: **8.48.01.04**

Các thành viên của Hội đồng chấm đề án tốt nghiệp có mặt: ..5..../ 05

TT	HỌ VÀ TÊN	TRÁCH NHIỆM TRONG HĐ	GHI CHÚ
1	**PGS. TSKH. Hoàng Đăng Hải**	Chủ tịch	
2	**TS. Đỗ Tiến Dũng**	Thư ký	
3	**TS. Nguyễn Xuân Thắng**	Phản biện 1	
4	**TS. Đặng Hoàng Long**	Phản biện 2	
5	**TS. Bùi Hải Phong**	Uỷ viên	

Các nội dung thực hiện:

1. Chủ tịch Hội đồng điều khiển buổi họp. Công bố quyết định của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ.
2. Người hướng dẫn khoa học hoặc thư ký đọc lý lịch khoa học và các điều kiện bảo vệ đề án tốt nghiệp của học viên. (có bản lý lịch khoa học và kết quả các môn học cao học của học viên kèm theo).
3. Học viên trình bày tóm tắt đề án tốt nghiệp.
4. Phản biện 1 đọc nhận xét (có văn bản kèm theo)
5. Phản biện 2 đọc nhận xét (có văn bản kèm theo)
6. Các câu hỏi của thành viên Hội đồng:

Vì sao Chương chỉ dùng Accuracy?
Nếu dữ liệu Phương sai mô thức, vị trí địa lý thì xử lý như thế nào?
Giải Cơ trong số A, B, D?

7. Trả lời của học viên:

Tự giới thiệu đề tài, các phần chính đồ án, các câu hỏi Recall, mục đích đồ án, nội dung chính, bổ sung 2 câu hỏi do kinh nghiệm bản thân

8. Thư ký đọc nhận xét về quá trình thực hiện đề án tốt nghiệp của học viên *(có văn bản kèm theo).*

9. Hội đồng họp riêng:

- Bầu Ban kiểm phiếu:
  1. Trưởng Ban kiểm phiếu: PGS.TS. Hoàng Đăng Hải
  2. Uỷ viên Ban kiểm phiếu: TS. ~~Nguyễn~~ Đặng Hoàng Ing
  3. Uỷ viên Ban kiểm phiếu: TS. Đỗ Tiến Dũng
- Hội đồng chấm đề án tốt nghiệp bằng bỏ phiếu kín.
- Ban kiểm phiếu làm việc:
- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu *(có Biên bản họp Ban kiểm phiếu kèm theo)*
- **Điểm trung bình của đề án tốt nghiệp:** 7,8

Kết luận:

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ đề án tốt nghiệp:

Trích dẫn công thức

2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên: Đồng ý

3. Đề án tốt nghiệp có thể phát triển thành đề tài nghiên cứu cho NCS.....

Buổi làm việc kết thúc vào...10:15... cùng ngày.

Chủ tịch	**Thư ký**
PGS.TS.KH. Hoàng Đăng Hải	Đỗ Tiến Dũng

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

(Dùng cho người phản biện)

Tên đề tài đề án tốt nghiệp: Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục

Chuyên ngành: Hệ thống thông tin

Mã chuyên ngành: 8.48.01.04

Họ và tên học viên: Trịnh Đình Minh

Họ và tên người nhận xét: Nguyễn Xuân Thắng

Học hàm, học vị: Tiến sĩ

Chuyên ngành: Công nghệ thông tin

Cơ quan công tác: Trường Đại học Hà Nội

Số điện thoại: 0912543560 E-mail: nxthang@hanu.edu.vn

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Đề tài có cơ sở khoa học và thực tiễn rõ ràng trong bối cảnh các giao dịch tín dụng, thanh toán thẻ và ngân hàng số phát sinh với khối lượng lớn, tốc độ cao và mức độ rủi ro ngày càng phức tạp. Các phương pháp phát hiện gian lận dựa trên luật cố định hoặc mô hình thống kê đơn giản khó thích ứng với các thủ đoạn gian lận thay đổi liên tục, trong khi yêu cầu nghiệp vụ đòi hỏi hệ thống phải phát hiện sớm, giảm bỏ sót gian lận nhưng vẫn hạn chế cảnh báo sai.

Về mặt khoa học, tiếp cận giải quyết vấn đề dựa trên các đặc trưng quan trọng của dữ liệu giao dịch tín dụng: mất cân bằng lớp, quan hệ phi tuyến, sự thay đổi của hành vi gian lận theo thời gian và tính phụ thuộc chuỗi của hành vi gian lận. Đây là những thách thức phù hợp với hướng nghiên cứu và triển khai hệ thống thông tin hiện đại, đòi hỏi kết hợp giữa học máy, xử lý dữ liệu dòng thời gian và cơ chế cập nhật mô hình.

Hướng tiếp cận của đề án không chỉ tập trung vào độ chính xác của mô hình phân loại, mà còn quan tâm đến yêu cầu vận hành gần thời gian thực, khả năng phát hiện chuỗi hành vi bất thường và khả năng thích nghi khi phân phối dữ liệu thay đổi. Vì vậy đề tài có tính cấp thiết và ý nghĩa ứng dụng.

II/ Nội dung của đề án tốt nghiệp, các kết quả đã đạt được:

Nội dung đề án được tổ chức tương đối logic, gồm phần tổng quan về bài toán phát hiện bất thường trong giao dịch tín dụng, phần triển khai hệ thống và các kỹ thuật áp dụng, sau đó là phần thử nghiệm, đánh giá và kết luận. Học viên đã trình bày được đặc điểm dữ liệu giao dịch, các yêu cầu của hệ thống phát hiện gian lận, các nhóm phương pháp phát hiện bất thường và giải thích được lý do lựa chọn kiến trúc lai cho bài toán.

Về phương pháp và thiết kế hệ thống, đề án đề xuất kiến trúc kết hợp bốn thành phần chính: mô hình học máy có giám sát CatBoost làm lõi phân loại, Apache Kafka làm tầng truyền tải dữ liệu luồng, Apache Flink làm tầng xử lý luồng có trạng thái, cơ chế xử lý sự kiện phức tạp (CEP) để phát hiện các chuỗi hành vi bất thường và vòng cập nhật mô hình động để thích nghi với dữ liệu mới. Cách tổ chức thành các tầng thu thập, xử lý luồng, phân tích rủi ro, ra quyết định và cập nhật mô hình là rõ ràng, phù hợp với một hệ thống vận hành theo hướng sự kiện.

Về thực nghiệm, học viên đã sử dụng tập dữ liệu phát hiện gian lận thẻ tín dụng năm 2023, xây dựng các kịch bản đánh giá năng lực phát hiện gian lận, vận hành thời gian thực và cập nhật mô hình động. Các kết quả trên tập kiểm thử cho thấy hiệu quả và tính khả thi của hệ thống.

Đề án cũng đánh giá được các chỉ số vận hành của hệ thống: thông lượng tối đa khoảng 623 giao dịch/giây, tỷ lệ thất thoát thông điệp 0%, độ trễ trung bình khoảng 2,03 giây. Cơ chế cập nhật mô hình động được kiểm chứng qua nhiều chu kỳ huấn luyện lại, trong đó mô hình mới được triển khai bằng hoán đổi nóng mà không làm gián đoạn luồng xử lý; F1 tăng từ 0,9835 ở lần cập nhật đầu lên 0,9937 ở lần cập nhật mới nhất.

Bên cạnh các kết quả đạt được, đề án đã nhận diện một số hướng cần tiếp tục hoàn thiện như kiểm chứng trên dữ liệu thực tế có mức mất cân bằng cao và khó phân tách hơn, làm giàu đặc trưng từ nhiều nguồn dữ liệu, bổ sung mẫu CEP ... Đây là các nhận xét phù hợp, thể hiện học viên nhìn nhận được giới hạn của thử nghiệm hiện tại.

III/ Những vấn đề cần giải thích thêm:

1. Vì sao trong bài toán phát hiện gian lận tín dụng không nên chỉ sử dụng chỉ số Accuracy để đánh giá mô hình, và các chỉ số Precision, Recall, F1 có ý nghĩa gì trong bối cảnh dữ liệu mất cân bằng?

2. Trong hệ thống đề xuất, việc bổ sung CEP giúp tăng Recall nhưng làm giảm Precision so với chỉ dùng CatBoost. Khi triển khai trên thực tế, học viên sẽ lựa chọn hoặc điều chỉnh ngưỡng tổng hợp điểm rủi ro như thế nào để cân bằng giữa giảm bỏ sót gian lận và hạn chế cảnh báo sai?

IV/ Kết luận:

Đồng ý cho học viên bảo vệ đề án tốt nghiệp

Ngày 04 tháng 06 năm 2026

NGƯỜI NHẬN XÉT

Nguyễn Xuân Thắng

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

(Dùng cho người phản biện)

Tên đề tài đề án tốt nghiệp: Hệ thống phát hiện gian lận tín dụng thời gian thực với mô hình cập nhật liên tục

Chuyên ngành: Hệ thống thông tin

Mã chuyên ngành: 8.48.01.04

Họ và tên học viên: Trịnh Đình Minh

Họ và tên người nhận xét: Đặng Hoàng Long

Học hàm, học vị: Tiến sĩ

Chuyên ngành: Khoa học máy tính

Cơ quan công tác: Học viện Công nghệ Bưu chính Viễn thông

Số điện thoại: 0335229386 E-mail: longdh@ptit.edu.vn

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Đề tài có cơ sở khoa học và thực tiễn. Gian lận trong giao dịch thẻ tín dụng gây thiệt hại tài chính lớn và ngày càng tinh vi, trong khi các phương pháp phát hiện truyền thống dựa trên luật cố định hoặc thống kê đơn giản tỏ ra kém hiệu quả trước khối lượng dữ liệu lớn và hành vi gian lận thay đổi liên tục. Bài toán vì vậy đặt ra hai yêu cầu cốt lõi: xử lý giao dịch theo thời gian thực với độ trễ thấp và liên tục cập nhật mô hình để thích nghi với dữ liệu mới. Hướng tiếp cận của đề án, kết hợp mô hình học máy, hạ tầng xử lý luồng (Kafka/Flink), xử lý sự kiện phức tạp (CEP) và cơ chế cập nhật mô hình động, là hướng có ý nghĩa và thực tế.

II/ Nội dung của đề án tốt nghiệp, các kết quả đã đạt được:

- Khảo sát tổng quan bài toán phát hiện gian lận giao dịch tín dụng, phân tích đặc trưng và đặc điểm thống kê của dữ liệu, các nhóm phương pháp phát hiện bất thường và các công cụ xử lý luồng dữ liệu.
- Đề xuất một kiến trúc lai phân tầng kết hợp bốn thành phần bổ trợ nhau: CatBoost làm lõi phân loại, Apache Kafka và Apache Flink làm hạ tầng xử lý luồng, Flink CEP để phát hiện chuỗi hành vi bất thường, và cơ chế cập nhật mô hình động.
- Hiện thực hóa hệ thống bằng Python/PyFlink, đóng gói bằng Docker; xây dựng ba mẫu CEP (vận tốc, đột biến giá trị, tần suất) và công thức tổng hợp điểm rủi ro từ hai nguồn tín hiệu học máy và CEP.
- Thiết kế hai kịch bản thử nghiệm trên bộ dữ liệu Credit Card Fraud 2023: kịch bản 1 đánh giá năng lực phát hiện (việc bổ sung CEP giúp giảm gần một nửa số giao dịch gian lận bị bỏ sót so với chỉ dùng mô hình đơn lẻ); kịch bản 2 đánh giá hiệu năng vận hành (thông lượng khoảng 623 giao dịch/giây, không thất thoát dữ liệu, độ trễ ở mức vài giây) và cơ chế cập nhật mô hình động qua nhiều chu kỳ hoán đổi nóng không gián đoạn dịch vụ.

III/ Những vấn đề cần giải thích thêm:

- Bộ dữ liệu Credit Card Fraud 2023 chỉ gồm các đặc trưng ẩn danh V1–V28 và Amount, không có mã thẻ, mốc thời gian hay vị trí địa lý. Vậy mã thẻ và mốc thời gian phục vụ CEP được sinh/mô phỏng như thế nào, và điều này ảnh hưởng ra sao đến tính đại diện của kết quả đánh giá tầng CEP?
- Bộ dữ liệu sử dụng có phân bố cân bằng 50/50, trong khi phần tổng quan nhấn mạnh đặc thù mất cân bằng nghiêm trọng (tỷ lệ gian lận dưới 1%) của bài toán. Các kết quả gần như tuyệt đối (Recall = 1,0; AUC = 1,0) phản ánh độ khó thực tế đến đâu, và hệ thống dự kiến ứng xử ra sao trên dữ liệu mất cân bằng cao?
- Các trọng số α, β trong công thức điểm rủi ro (2.7) và các ngưỡng quyết định θ được lựa chọn, hiệu chỉnh dựa trên cơ sở nào? Đã có khảo sát độ nhạy của kết quả theo các giá trị này chưa?
- Bảng 2.1 nêu bốn mẫu hành vi (bao gồm thay đổi vị trí địa lý và xác thực thất bại liên tiếp) nhưng phần hiện thực chỉ triển khai ba mẫu. Lý do và hướng bổ sung mẫu còn lại là gì?

IV/ Kết luận:

Đề án đáp ứng yêu cầu của một đề án tốt nghiệp trình độ thạc sĩ. Đồng ý cho phép học viên bảo vệ đề án tốt nghiệp.

Ngày .8.. tháng .7... năm 2026

NGƯỜI NHẬN XÉT

Đặng Hoàng Long