ISSN 1859-3666
E-ISSN 2815-5726

MỤC LỤC

KINH TẾ VÀ QUẢN LÝ

QUẢN TRỊ KINH DOANH

ISSN 1859-3666
E-ISSN 2815-5726

MỨC ĐỘ HIỆU QUẢ CỦA CÁC MÔ HÌNH HỌC MÁY TREE-BASED TRONG PHÁT HIỆN GIAO DỊCH GIAN LẬN THẺ TÍN DỤNG

Nguyễn Minh Nhật*
Email: nhatnm@hub.edu.vn
Đào Lê Kiều Oanh*
Email: Oanhdlk@hub.edu.vn
*** Trường Đại học Ngân hàng TP. Hồ Chí Minh**

Ngày nhận: *03/09/2024* ***Ngày nhận lại:*** *25/10/2024* ***Ngày duyệt đăng:*** *28/10/2024*

Nghiên cứu này tập trung vào việc đánh giá và so sánh hiệu quả của các mô hình học máy dựa trên cây (Tree-based machine learning models) trong việc dự báo gian lận thẻ tín dụng. Các mô hình được xét gồm Decision Tree, Random Forest, Gradient Boosting Machines (GBM) và Extreme Gradient Boosting (XGBoost). Bộ dữ liệu sử dụng cho nghiên cứu này bao gồm 568,630 giao dịch thẻ tín dụng, với các thuộc tính từ V1 đến V28 được biến đổi thông qua phân tích thành phần chính (PCA) để bảo vệ thông tin cá nhân. Nghiên cứu này sử dụng ma trận nhầm lẫn (Confusion Matrix) và các chỉ số đánh giá như Độ chính xác, Độ nhạy (Recall), Precision và F1 Score để đánh giá hiệu quả của mỗi mô hình. Kết quả cho thấy rằng Random Forest và XGBoost đều có hiệu suất ấn tượng, đặc biệt Random Forest cho thấy sự vượt trội hơn trong việc giảm thiểu báo động giả và phát hiện chính xác các giao dịch gian lận. Mặc dù có một số hạn chế về khả năng giải thích các thuộc tính quan trọng do tính ẩn danh của dữ liệu, tuy nhiên nghiên cứu kỳ vọng cung cấp góc nhìn quan trọng về tiềm năng ứng dụng các mô hình học máy trong việc phát hiện gian lận thẻ tín dụng, từ đó có thể là kênh tham khảo hoặc hỗ trợ cho các tổ chức tín dụng trong hoạt động thực tiễn.

Từ khóa*: Học máy, Mô hình Tree-based, Random Forest, XGBoost, Gian lận thẻ tín dụng*

JEL Classifications*: C63, C45, G28.*

DOI*: 10.54404/JTS.2024.195V.04*

1. Đặt vấn đề

Gian lận thẻ tín dụng là một trong những thách thức nghiêm trọng nhất đối với ngành ngân hàng và tài chính, với ước tính chi phí toàn cầu của gian lận thanh toán trực tuyến sẽ đạt 260 tỷ đô la Mỹ vào năm 2025 (Juniper Research, 2022). Tại Việt Nam, nghiên cứu dữ liệu từ Visa cho thấy, trong Quý 3 năm 2023, tỷ lệ gian lận liên quan đến việc phát hành thẻ tại Việt Nam cao hơn so với mức trung bình của khu vực Châu Á - Thái Bình Dương. Hơn nữa, xu hướng này đang tăng lên một cách nhanh chóng (Hiệp hội Ngân hàng Việt Nam, 2024). Sự phát triển của công nghệ và sự phổ biến của giao dịch trực tuyến đã không chỉ đơn giản là mở rộng cơ hội cho tiêu dùng mà còn làm tăng khả năng xảy ra gian lận, đặt ra những thách thức lớn cho các tổ chức tài chính trong việc bảo vệ khách hàng và tài sản của họ.

Theo Dal Pozzolo và cộng sự (2018), các giao dịch gian lận thẻ tín dụng thường diễn ra

mà không có sự chấp thuận của chủ thẻ, với mục đích chiếm đoạt tài sản cá nhân một cách bất hợp pháp. Các giao dịch này được thực hiện thông qua việc sử dụng thông tin thẻ tín dụng bị đánh cắp hoặc làm giả, chủ yếu xảy ra trong môi trường mà không cần xác minh sự hiện diện của chủ thẻ. Hậu quả là những thiệt hại đáng kể không chỉ đối với cá nhân mà còn với các tổ chức tài chính phát hành thẻ. Trước thực trạng này, nhu cầu phát triển các công nghệ hiệu quả nhằm phát hiện và ngăn ngừa gian lận trở nên cấp thiết. Trong đó, công nghệ học máy nổi bật như một giải pháp tiềm năng, giúp tăng cường khả năng phát hiện các giao dịch đáng ngờ (Varmedja & cộng sự, 2019). Đặc biệt, các mô hình học máy Tree-based với kỹ thuật tiên tiến, mang đến khả năng tối ưu hóa trong việc phân loại và dự đoán gian lận, hỗ trợ đáng kể trong việc giảm thiểu rủi ro tài chính.

Tiếp cận theo nhóm mô hình học máy Tree-based bao gồm mô hình Decision Tree, Random Forest hay các mô hình Gradient Boosting, đã được chứng minh là rất hiệu quả trong việc phân tích và phân loại dữ liệu lớn trong việc phát hiện các giao dịch gian lận (Learning, 2023). Các mô hình này tận dụng lợi thế của việc kết hợp nhiều cây quyết định để tạo ra một mô hình tổng hợp mạnh mẽ hơn, giảm thiểu sai lệch và phương sai, đồng thời cải thiện độ chính xác của dự đoán. Đặc biệt, khả năng xử lý các tập dữ liệu lớn và không cân xứng - một đặc điểm thường thấy trong dữ liệu gian lận thẻ tín dụng - làm cho các phương pháp này trở nên vô cùng quý giá. Do đó, sự phức tạp và khả năng tự học của các mô hình này cung cấp một công cụ mạnh mẽ để giải mã các hành vi gian lận ngày càng tinh vi (Tanwar và cộng sự, 2023).

Tuy nhiên, hiện nay vẫn còn những tranh luận nhất định về tính hiệu quả của các mô hình học máy dựa trên cây trong việc phát hiện các giao dịch gian lận. Có nhiều nghiên cứu ủng hộ tính hiệu quả của mô hình rừng ngẫu nhiên như Udeze và cộng sự (2022), Tanwar và cộng sự (2023), nhưng cũng có những nghiên cứu lại ủng hộ mô hình Gradient Boosting như Faraji (2022), Learning (2023). Do đó, trong bài nghiên cứu này, tác giả sẽ tập trung phân tích tính hiệu quả và so sánh hiệu suất của bốn mô hình học máy dựa trên cây bao gồm cây quyết định, rừng ngẫu nhiên, GBM (Gradient Boosting Machines) và XGBoost (Extreme Gradient Boosting) trong việc dự báo gian lận thẻ tín dụng. Mức độ hiệu quả của các mô hình học máy dựa trên cây sẽ được đánh giá dựa trên 8 tiêu chí cụ thể được ước tính từ ma trận nhầm lẫn (Confusion matrix) của mô hình. Các mô hình này sẽ được kiểm định trên bộ dữ liệu lớn với 568,630 giao dịch trên thẻ tín dụng được thu thập đến thời điểm năm 2023 và công khai trên (Kaggle, 2023). Kết quả nghiên cứu kỳ vọng rằng sẽ giúp các tổ chức tài chính hiểu rõ hơn về hiệu suất của các phương pháp học máy dựa trên cây, để từ đó có những lựa chọn và cải tiến phù hợp trong hoạt động thực tiễn.

Các nội dung tiếp theo của bài nghiên cứu sẽ được trình bày với kết cấu như sau: (2) Khảo lược nghiên cứu; (3) Phương pháp nghiên cứu; (4) Kết quả nghiên cứu thực nghiệm; (5) Kết luận.

2. Khảo lược nghiên cứu

Các nghiên cứu về phát hiện gian lận thẻ tín dụng đã thu hút sự quan tâm rộng rãi từ cả cộng đồng nghiên cứu và ngành công nghiệp tài chính. Những nghiên cứu gần đây tập trung vào việc phát triển các phương pháp và công nghệ mới, bao gồm cả phương pháp học máy và trí tuệ nhân tạo, để nâng cao khả năng nhận diện và ngăn chặn gian lận hiệu quả. Bằng cách kết hợp các phương tiện phân tích dữ liệu, thuật toán học máy có thể học từ dữ liệu giao dịch và tự động phát hiện các biểu hiện của hoạt động gian lận, đem lại hiệu suất và độ chính xác cao trong việc bảo vệ tài chính cá nhân và tổ chức.

Awoyemi và cộng sự (2017) đã nghiên cứu hiệu suất của các kỹ thuật Naïve Bayes, k-nearest neighbor và hồi quy logistic trên bộ dữ liệu gian lận thẻ tín dụng có tính chất lệch cao, sử dụng một kỹ thuật kết hợp của việc lấy mẫu thiểu số và lấy mẫu dư thừa, và các kỹ thuật này được áp dụng trên dữ liệu thô và đã qua xử lý. Kết quả cho thấy hiệu suất tối ưu về độ chính xác cho các phân loại Naïve Bayes, k-nearest neighbor và hồi quy logistic lần lượt là 97.92%, 97.69% và 54.86%, với k-nearest neighbor thể hiện hiệu quả tốt hơn so với Naïve Bayes và hồi quy logistic. Jurgovsky và cộng sự (2018), trong bài nghiên cứu của mình đã định nghĩa vấn đề phát hiện gian lận như một nhiệm vụ phân loại chuỗi và sử dụng mạng LSTM (Long Short-Term Memory) để tích hợp các chuỗi giao dịch, đồng thời tích hợp các chiến lược tổng hợp đặc điểm tiên tiến nhất và báo cáo kết quả thông qua các chỉ số thu hồi truyền thống. So sánh với phân loại Random Forest cơ bản cho thấy LSTM cải thiện độ chính xác trong việc phát hiện gian lận trên các giao dịch ngoại tuyến khi chủ thẻ có mặt tại nhà cung cấp. Cả hai phương pháp học có trình tự và không có trình tự đều được hưởng lợi mạnh mẽ từ các chiến lược tổng hợp đặc điểm thủ công. Phân tích sau đó về các trường hợp tích cực cho thấy cả hai phương pháp có xu hướng phát hiện các hình thức gian lận khác nhau, điều này gợi ý một sự kết hợp của cả hai.

Dornadula và Geetha (2019) đã phát triển một phương pháp phát hiện gian lận mới cho dữ liệu giao dịch trực tuyến, bằng cách phân tích lịch sử giao dịch của khách hàng và rút ra các mẫu hành vi, sau đó phân loại chủ thẻ thành các nhóm dựa trên số tiền giao dịch và sử dụng các phân loại khác nhau để đào tạo cho từng nhóm một cách riêng biệt. Bài nghiên cứu đề xuất một cơ chế phản hồi để giải quyết vấn đề dựa trên bộ dữ liệu về gian lận thẻ tín dụng ở châu Âu. Thennakoon và cộng sự (2019) tiếp tục dành sự quan tâm khi tập trung vào bốn loại hình gian lận chính trong giao dịch thực tế, mỗi loại được giải quyết bằng cách sử dụng các mô hình học máy khác nhau và phương pháp tốt nhất được chọn thông qua đánh giá, cung cấp hướng dẫn toàn diện để chọn thuật toán tối ưu phù hợp với từng loại gian lận. Ngoài ra, các tác giả cũng đề cập đến phát hiện gian lận thẻ tín dụng thời gian thực, sử dụng phân tích dự báo từ các mô hình học máy được triển khai và một mô-đun API để xác định tính xác thực của một giao dịch cụ thể, đồng thời đánh giá một chiến lược mới hiệu quả cho việc giải quyết sự phân bố lệch của dữ liệu.

Maniraj và cộng sự (2019) hướng sự tập trung vào phân tích và tiền xử lý dữ liệu cũng như triển khai nhiều thuật toán phát hiện bất thường như Local Outlier Factor và Isolation Forest trên dữ liệu giao dịch thẻ tín dụng đã được biến đổi PCA, với mục tiêu phát hiện 100% các giao dịch gian lận và giảm thiểu phân loại gian lận không chính xác. Bagga và cộng sự (2020) đã cho thấy rằng việc phát hiện gian lận thẻ tín dụng đặc biệt khó khăn do hai vấn đề chính là sự thay đổi liên tục của hành vi gian lận và sự chênh lệch lớn trong dữ liệu được sử dụng. Các tác giả đã tiến hành so sánh hiệu suất của các phương pháp hồi quy logistic, K-nearest neighbors, Random Forest, Naive Bayes, perceptron đa tầng, AdaBoost, pipelining và học tập kết hợp trên dữ liệu gian lận thẻ tín dụng.

Bên cạnh đó, chủ đề phát hiện các giao dịch gian lận thẻ tín dụng cũng thu hút sự quan tâm của các nhà nghiên cứu trong nước, Nguyễn Thị Liên và cộng sự (2018) đã tiến hành nghiên cứu nghiên cứu trên bộ dữ liệu Châu Âu trên các mô hình phổ biến như mô hình Logistic, Mạng Bayesian, Decision Tree và phương pháp Stacking, từ đó đề xuất mô hình và phương pháp xử lý dữ liệu phù hợp cho các ngân hàng thương mại ở Việt Nam để phát hiện và kiểm soát gian lận thẻ tín dụng.

Trong dòng chảy của sự phát triển công nghệ và sự xuất hiện của các mô hình học máy hiện đại, có một xu hướng nổi bật được các nhà nghiên cứu trong thời gian gần đây rất quan tâm đó là áp dụng các mô hình học máy Tree-based để phát hiện các giao dịch gian lận thẻ tín dụng. Các mô hình này có những ưu điểm vượt trội như có khả năng xử lý dữ liệu phi tuyến tính, hiệu quả với dữ liệu có đặc tính phân tán, khả năng tự động xử lý dữ liệu phân loại và các biến dạng số, độ chính xác và độ tin cậy cao, giảm thiệu hiện tượng quá khớp (Overfitting) và có khả năng giải thích cao Faraji (2022).

Tuy nhiên hiện nay vẫn còn những tranh luận nhất định về tính hiệu quả của các mô hình học máy Tree-based trong việc xử lý bài toán phát hiện các giao dịch gian lận. Chẳng hạn như nghiên cứu của Jain và cộng sự (2020) đã thực hiện so sánh hiệu quả giữa các thuật toán học máy Tree-based, kết quả cho thấy rằng thuật toán Random Forest cho độ chính xác cao nhất so với Decision Tree và XGBoost. Udeze & cộng sự (2022) cũng có kết quả nghiên cứu tương tự khi áp dụng các thuật toán học máy dựa trên cây trong phát hiện giao dịch gian lận thẻ tín dụng trong điều kiện bộ dữ liệu có sự mất cân bằng lớn. Ngược lại với kết quả của các nghiên cứu trên, Faraji (2022) đưa ra bằng chứng rằng XGBoost cho kết quả tốt nhất so với Random Forest trên các tiêu chí đánh giá bao gồm Độ chính xác (Accuracy), Độ Nhạy (Recall), Precision và F1 score. Learning (2023) cũng có kết quả tương tự khi chứng minh rằng XGBoost cho kết quả tốt hơn các mô hình học máy dựa trên cây khác trên các tiêu chí như AUC (diện tích dưới đường cong), độ chính xác, giá trị dự đoán dương, độ nhớ và F1 score.

Dựa trên kết quả khảo lược trên, chúng ta có thể nhận thấy rằng hiện nay vẫn tồn tại những tranh luận nhất định về tính hiệu quả của mô hình học máy Tree-based và việc lựa chọn mô hình nào là mô hình tối ưu trong việc dự báo các giao dịch gian lận thẻ tín dụng. Do đó, trong nghiên cứu này, nhóm tác giả sẽ tập trung trả lời hai câu hỏi quan trọng sau:

Câu hỏi 1: Mức độ dự báo chính xác các giao dịch gian lận thẻ tín dụng của các mô hình học máy Tree-based như thế nào?

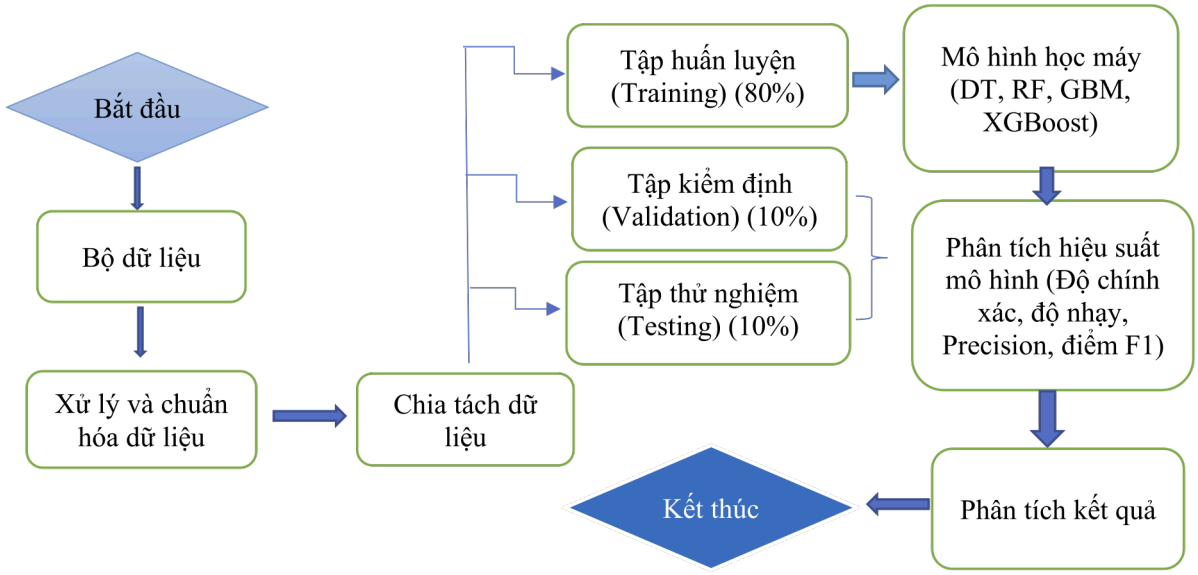
Câu hỏi 2: Trong số các mô hình học máy Tree-based được lựa chọn để nghiên cứu, đâu là mô hình tốt nhất để dự báo giao dịch gian lận thẻ tín dụng?

3. Phương pháp nghiên cứu

3.1 Quy trình nghiên cứu

Hình 1 mô tả cơ bản về các bước trong quy trình nghiên cứu. Đầu tiên, quy trình nghiên cứu được thực hiện với việc thu thập và xử lý bộ dữ liệu liên quan đến các giao dịch thẻ tín dụng. Dữ liệu sau đó được làm sạch và chuẩn hóa để loại bỏ các thông tin trùng lắp hoặc các yếu tố gây nhiễu nhằm chuẩn bị cho các bước tiếp theo. Sau khi chuẩn hóa, dữ liệu được chia tách thành ba phần cụ thể, trong đó tập huấn luyện (Training data) chiếm 80%, tập kiểm định (Validation data) và tập thử nghiệm (testing data) mỗi tập chiếm 10%. Tỷ lệ chia tách này được tác giả tham khảo trong các nghiên cứu Nuthalapati (2023), Khalid và cộng sự (2024) trong việc dự báo gian lận thẻ tín dụng với ưu điểm của tỷ lệ này là sự cân bằng giữa việc huấn luyện và đánh giá hiệu suất của các mô hình học máy. Tập huấn luyện sẽ được sử dụng để huấn luyện các mô hình học máy được sử dụng trong nghiên cứu này như Decision Tree, Random Forest, GBM và XGBoost học cách phân biệt được những giao dịch gian lận hay hợp pháp. Bên cạnh đó, quy trình chia tách dữ liệu cũng giúp tối ưu hóa hiệu suất và giảm thiểu nguy cơ quá khớp (overfitting) hoặc dưới khớp (underfitting) của mô hình dự báo.

Sau khi hoàn thành giai đoạn huấn luyện, các mô hình học máy sẽ được đánh giá hiệu suất thông qua các chỉ số quan trọng như độ chính xác, độ nhạy, Precision và điểm F1. Các



(Nguồn: Tác giả)

Hình 1: *Mô tả về quy trình nghiên cứu*

chỉ số này giúp nhóm nghiên cứu có thể kiểm tra được tính chính xác của mô hình trong việc phân loại các giao dịch là gian lận hay hợp pháp. Đồng thời, tập kiểm định được sử dụng để điều chỉnh các siêu tham số trong mô hình một cách hợp lý và cải thiện hiệu quả phân loại trước khi đưa đến bước đánh giá cuối cùng. Cuối cùng, tập thử nghiệm sẽ giúp đánh giá một cách khách quan về hiệu suất tổng thể của các mô hình dự báo khi áp dụng vào tình huống thực tế. Kết quả của quá trình phân tích này sẽ cung cấp thông tin chi tiết về khả năng phát hiện các giao dịch gian lận thẻ tín dụng của các mô hình và đóng góp vào việc lựa chọn mô hình phù hợp nhất để tiến hành triển khai trong môi trường thực tiễn.

3.2. Nhóm mô hình học máy Tree-Based

Nhóm mô hình học máy Tree-Based bao gồm các thuật toán dựa trên cấu trúc cây để thực hiện các nhiệm vụ học có giám sát như phân loại và hồi quy. Mô hình này dựa trên việc chia tập dữ liệu thành các nhóm nhỏ hơn bằng cách sử dụng các quyết định dựa trên giá trị của các thuộc tính (features). Các thuật toán này có ưu điểm là dễ hiểu, dễ giải thích và hiệu quả trong việc xử lý các loại dữ liệu khác nhau. Mô hình Decision Tree, Random Forest và GBM (Gradient Boosting Machines) và XGBoost (Extreme Gradient Boosting) là những mô hình nổi bật của hướng nghiên cứu này (Learning, 2023).

3.2.1. Decision Tree

Decision Tree là một mô hình phân loại trong học máy, nơi các quyết định được thực hiện dựa trên thuộc tính của dữ liệu. Trong bối cảnh phát hiện gian lận thẻ tín dụng, Decision Tree thường sử dụng giá trị Entropy hoặc Gini để tối ưu hóa quá trình phân loại, nhằm phân biệt giao dịch gian lận với giao dịch hợp pháp. Trong bài nghiên cứu này, tác giả sẽ sử dụng chỉ số Gini để tối ưu quá trình phân loại, do Gini cung cấp cách tính đơn giản và thường làm cho mô hình Decision Tree hiệu quả hơn về mặt tính toán.

Chỉ số Gini, hay Gini Impurity, là một công thức đo lường xác suất một mẫu ngẫu nhiên được phân loại sai nếu nó được gắn nhãn một cách ngẫu nhiên dựa trên phân phối nhãn trong tập con đó. Chỉ số Gini được định nghĩa cho tập dữ liệu S như sau:

$$G(S) = 1 - \sum_{i=1}^{n} p_i^2 \quad (1)$$

Trong đó, p_i là tỷ lệ mẫu thuộc lớp i trong tập S (lớp i trong trường hợp này là giao dịch "Gian lận" hoặc giao dịch "Hợp pháp"). Gini có giá trị càng thấp thì tập dữ liệu càng trở nên đồng nhất hơn.

Trong mô hình Decision Tree, việc chọn thuộc tính để phân chia dữ liệu ở mỗi nút dựa trên giá trị Gini Impurity thấp nhất sau phân chia. Gini Gain, đo lường sự giảm Gini Impurity, là một chỉ số được sử dụng trong việc xây dựng mô hình để đánh giá mức độ cải thiện (hoặc giảm thiểu) của độ không thuần khiết (impurity) sau khi dữ liệu được phân chia dựa trên một thuộc tính cụ thể. Gini Gain được tính bằng cách lấy Gini Impurity ban đầu của tập dữ liệu trừ đi trọng số trung bình của Gini Impurity của các tập con sau phân chia:

$$\text{Gini Gain}(S,A) = G(S) - \sum_{v \in Value(A)} \frac{|S_v|}{|S|} G(S_v) \quad (2)$$

Trong đó: S_v là tập con của S khi thuộc tính A có giá trị v.

Gini Gain cho biết mức độ mà mỗi thuộc tính góp phần làm giảm độ không đồng nhất trong tập dữ liệu khi nó được sử dụng để phân chia dữ liệu tại một nút. Thuộc tính với Gini Gain cao nhất là thuộc tính tối ưu nhất để tạo nút phân chia tiếp theo, vì nó tạo ra các tập con có độ đồng nhất cao nhất.

Quá trình xây dựng cây sẽ được lặp đi lặp lại qua các thuộc tính của dữ liệu cho đến khi: (i) Mỗi nút lá đạt đến mức độ đồng nhất nhất định hay tất cả các giao dịch tại một nút thuộc cùng một lớp "Gian lận" hay "Hợp pháp" (Gini = 0); (ii) Cây đạt đến độ sâu đã được xác định trước; (iii) Số lượng giao dịch tại một nút dưới một ngưỡng nhất định.

3.2.2. Random Forest

Mô hình Random Forest là một kỹ thuật kết hợp (Ensemble) dựa trên việc kết hợp nhiều cây quyết định để giảm nguy cơ quá khớp (overfitting), cải thiện độ chính xác và khả năng tổng quát hóa của mô hình. Random Forest được tạo thành từ nhiều cây quyết định $T_b(X)$, mỗi cây b là một hàm của tập dữ liệu huấn luyện X, được xây dựng từ mẫu tái chọn (bootstrap sample) của tập X. Số lượng cây B và cách thức hoạt động của từng cây trong rừng được định nghĩa như sau:

$$RF = \{T_1(X), T_2(X), \ldots, T_B(X)\} \quad (3)$$

Trong đó, mỗi cây $T_b(X)$ được xây dựng theo quy tắc như sau: (i) Mỗi cây T_b được xây dựng từ một mẫu tái chọn (bootstrap sample) của tập dữ liệu gốc X, được ký hiệu là X_b (X_b = BootstrapSample(X)); (ii) Khi xây dựng mỗi nút của cây, một tập hợp con m của thuộc tính được chọn ngẫu nhiên từ tổng số p thuộc tính của tập dữ liệu ($m \leq p$). Các thuộc tính được đánh giá dựa trên chỉ số Gini Impurity để chọn điểm phân chia tối ưu.

Trong bài toán phát hiện gian lận thẻ tín dụng, dự đoán của mô hình rừng ngẫu nhiên cho một giao dịch mới x được thực hiện bằng cách lấy bình chọn từ đa số các cây:

$$\hat{y} = \text{Model}\{T_1(x), T_2(x), \ldots, T_B(x)\} \quad (4)$$

3.2.3. Gradient Boosting Machines (GBM)

GBM hoạt động trên nguyên tắc của tăng cường gradient, nơi từng cây quyết định liên tiếp được huấn luyện để giảm thiểu lỗi của mô hình hiện tại. Việc ứng dụng mô hình GBM trong việc phát hiện gian lận thẻ tín dụng có thể được mô tả như sau:

(i) Xác định hàm mất mát

Hàm mất mát L(y, f(x)) đo lường sự sai lệch giữa giá trị thực tế y và giá trị dự đoán f(x). Trong trường hợp phát hiện gian lận thẻ tín dụng, một lựa chọn phổ biến là hàm mất mát logistic, được định nghĩa là:

$$L(y, f(x)) = \log(1+\exp(-yf(x))) \quad (5)$$

Trong đó: y là nhãn lớp, được mã hóa thành giá trị 1 cho giao dịch gian lận và -1 cho

giao dịch hợp pháp, và f(x) là dự đoán mô hình tại điểm dữ liệu x.

(ii) Huấn luyện cây quyết định tuần tự

GBM sẽ bắt đầu bằng một mô hình ban đầu rất đơn giản, thường là một dự đoán hằng số (x), và lặp lại các bước sau:

- Tính toán sai số (Residuals) cho mỗi điểm dữ liệu, dựa trên gradient của hàm mất mát:

$$r_i = - [\frac{dL(y_i,f(x_i))}{df(x_i))}]_{f(x)= f_{t-1}(x)} \quad (6)$$

Với i từ 1 đến n và n là số lượng mẫu dữ liệu

- Huấn luyện một cây quyết định mới (x) để dự đoán Residuals

$$\eta: f_t(x) = f_{t-1}(x) + \eta . h_t(x) \quad (7)$$

Cập nhật mô hình dự đoán bằng cách thêm cây mới với một tốc độ học Quá trình huấn luyện này tiếp tục cho đến khi số lượng cây đạt giới hạn xác định trước hoặc khi cải thiện trong hàm mất mát dưới một ngưỡng nhất định theo mục tiêu. Mô hình GBM có thể cung cấp một hiệu suất phân loại tốt vì nó tinh chỉnh mô hình dựa trên sai sót từ dữ liệu trước đó và cố gắng giảm thiểu chúng trong các lần lặp tiếp theo.

3.2.4. XGBoost (Extreme Gradient Boosting)

XGBoost là một biến thể cải tiến của GBM, được thiết kế để tối ưu hóa cả về hiệu năng lẫn tốc độ, đồng thời có thể xử lý quy mô dữ liệu lớn một cách hiệu quả. XGBoost bao gồm nhiều cải tiến kỹ thuật nhằm tăng cường hiệu suất và tính khả thi trong thực tiễn.

XGBoost mở rộng khả năng của GBM bằng cách giới thiệu một hàm mất mát chính xác hơn và kỹ thuật tối ưu hóa hiệu quả hơn. XGBoost sử dụng hàm mất mát thường là hàm mất mát log-likelihood, tương tự như GBM, nhưng với việc bổ sung thêm các thành phần regularization:

$$L(y,f(x)) = \sum_{i=1}^{n} l(y_i, f(x_i)) + \sum_{k=1}^{K} \Omega(f_k)) \quad (8)$$

Trong đó:

$l(y_i, f(x_i))$ là hàm mất mát Logistic

$\Omega(f_k)$ là hàm regularization cho mỗi cây f_k trong mô hình, thường bao gồm cả L1(Lasso) và L2 (ridge) regularization:

$$\Omega(f) = \gamma T + ½ \lambda \sum_{j=1}^{T} w_j^2$$

Với γ và λ lần lượt là các tham số regularization và T là số lượng nút lá trong cây, w_i là giá trị của nút lá.

XGBoost cung cấp khả năng điều chỉnh qua các tham số regularization, làm cho mô hình của nó có khả năng chống lại hiện tượng overfitting tốt hơn và thường cung cấp hiệu suất tối ưu hơn trên các tập dữ liệu lớn và phức tạp.

3.3. Các tiêu chí đánh giá hiệu quả của mô hình

3.3.1. Ma trận nhầm lẫn (Confusion matrix)

Confusion Matrix cung cấp cái nhìn toàn diện về hiệu suất mô hình phân loại, không chỉ tập trung vào tổng số lượng các dự đoán đúng mà còn làm nổi bật các loại lỗi phân loại cụ thể. Ma trận nhầm lẫn thường được biểu diễn dưới dạng 2x2, bao gồm các thành phần sau (Bảng 1):

Thông tin trong ma trận nhầm lẫn được mô tả cụ thể như sau:

True Positive (TP): Số lượng giao dịch mà mô hình dự đoán chính xác là gian lận. Nghĩa là, các giao dịch này thực sự là gian lận và mô hình đã thành công trong việc phát hiện chúng.

False Positive (FP): Số lượng giao dịch mà mô hình sai lầm phân loại là gian lận, trong khi thực tế chúng là hợp pháp. Đây là lỗi "Báo động giả", có thể gây ra phiền toái và chi phí không cần thiết cho khách hàng và ngân hàng.

True Negative (TN): Số lượng giao dịch mà mô hình dự đoán chính xác là hợp

Bảng 1: *Ma trận nhầm lẫn trong trường hợp gian lận thẻ tín dụng*

		Mô hình dự báo	
		Hợp pháp	Gian lận
Dữ liệu thực tế	Hợp pháp	TN	FP
	Gian lận	FN	TP

(Nguồn: (Learning, 2023))

pháp. Đây là trường hợp lý tưởng nơi mô hình xác định đúng các giao dịch không phải là gian lận.

False Negative (FN): Số lượng giao dịch mà mô hình sai lầm phân loại là hợp pháp, trong khi thực tế chúng là gian lận. Đây là lỗi "Bỏ sót" và nó là loại lỗi nguy hiểm nhất trong tình huống này vì nó cho phép các hoạt động gian lận tiếp tục không bị phát hiện.

3.3.2. Các chỉ số đánh giá từ ma trận nhầm lẫn

4. Kết quả nghiên cứu thực nghiệm

4.1. Dữ liệu nghiên cứu

Bộ dữ liệu nghiên cứu là một tập dữ liệu lớn bao gồm 568,630 giao dịch thẻ tín dụng, các giao dịch được thu thập đến thời điểm 2023 và được công khai trên Kaggle (2023). Mỗi giao dịch được mô tả bởi 31 thuộc tính, trong đó 28 thuộc tính (từ V1 đến V28) là kết quả của quá trình phân tích thành phần chính (PCA), một phương pháp giảm chiều dữ liệu để ẩn danh thông tin nhạy cảm. Hai thuộc tính còn lại là "Amount", thể hiện số tiền giao dịch, và "Class", chỉ ra liệu giao dịch có phải là gian lận (giá trị 1) hay không (giá trị 0). Thuộc tính "id" được sử dụng để định danh duy nhất cho mỗi giao dịch.

Phân tích thống kê mô tả của bộ dữ liệu cho thấy, giá trị trung bình của cột "Amount" là 12,041.96, với giá trị dao động từ 50.01 đến 24,039.93. Các thuộc tính từ V1 đến V28 có trung bình xấp xỉ 0 và độ lệch chuẩn là 1, cho thấy dữ liệu đã được chuẩn hóa trước khi phân tích. Cột "Class" có giá trị trung bình là 0.5, điều này cho thấy tập dữ liệu có thể đã được cân bằng giữa các trường hợp gian lận và hợp pháp. Điều này rất hữu ích trong việc huấn luyện các mô hình học máy, đặc biệt là các mô hình học máy thuộc nhóm mô hình cây vì nó ảnh hưởng đến các mô hình học và dự đoán. (Bảng 3)

Ngôn ngữ được sử dụng để phân tích dữ liệu là ngôn ngữ lập trình Python, các mô hình học máy dựa trên cây được tác giả tham khảo trong thư viện Scikit-learn, một trong những thư viện máy học phổ biến với nhiều công cụ hữu ích và dễ sử dụng. Các hàm lệnh liên quan đến các mô hình được sử dụng trong thư viện này bao gồm: DecisionTreeClassifier(); RandomForestClassifier(); GradientBoostingClassifier(); XGBClassifier().

4.2. Phân tích kết quả nghiên cứu

Trong quá trình xây dựng mô hình dự báo gian lận thẻ tín dụng, tác giả sử dụng 80% bộ dữ liệu tương đương với 454.904 giao dịch bao gồm cả giao dịch gian lận và hợp pháp để huấn luyện mô hình, tỷ lệ dữ liệu còn lại

Bảng 2: *Mô tả các chỉ số đánh giá từ ma trận nhầm lẫn*

Số thứ tự	Chỉ số đánh giá	Công thức tính toán	Ý nghĩa
01	Độ chính xác (Accuracy)	$\frac{TP+TN}{TP+TN+FP+FN}$	Độ chính xác cho biết tỷ lệ phần trăm các giao dịch được mô hình phân loại chính xác bao gồm các giao dịch gian lận và giao dịch hợp pháp.
02	Độ Nhạy (Recall)	$\frac{TP}{TP+FN}$	Độ nhạy cho thấy khả năng của mô hình trong việc phát hiện đúng các giao dịch gian lận.
03	Precision	$\frac{TP}{TP+FP}$	Precision cho thấy tỷ lệ các giao dịch được dự đoán là gian lận thực sự là gian lận. Tỷ lệ Precision cao giúp giảm thiểu sự bất tiện và chi phí phát sinh do việc xử lý những báo động giả, từ đó nâng cao trải nghiệm của khách hàng.
04	F1 Score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	F1 Score cân bằng giữa Recall và Precision. Trong bối cảnh gian lận thẻ tín dụng, chỉ số này đặc biệt hữu ích khi bạn cần một mô hình cân bằng giữa việc không bỏ sót các giao dịch gian lận và không gây ra quá nhiều báo động giả.

(Nguồn: (Learning, 2023))

tương đương với 113.726 giao dịch (trong đó tỷ lệ giao dịch gian lận và hợp pháp sẽ là 50:50) sẽ được sử dụng để kiểm định và thử nghiệm mô hình. Kết quả kiểm định của các mô hình học máy dựa trên cây sẽ được trình bày cụ thể trong Bảng 4.

Kết quả nghiên cứu cho thấy rằng, các mô hình học máy đã cho thấy hiệu quả đáng kể trong việc phát hiện gian lận thẻ tín dụng, với mô hình Random Forest dẫn đầu về độ chính xác và khả năng giảm báo động giả. XGBoost cũng thể hiện hiệu suất ấn tượng, nhấn mạnh

Bảng 3: *Bảng mô tả thống kê dữ liệu*

Thông số	Giá trị	V1 đến V28	Amount	Class
Số lượng	568630	568630	568630	568630
Trung bình	-	0 (khoảng)	12041.96	0.5
Độ lệch chuẩn	-	1	6919.64	0.5
Giá trị nhỏ nhất	-	Khoảng -10 đến -21	50.01	0
25%	-	-0.29 đến -0.55	6054.89	0
Trung vị	-	Khoảng 0.08 đến -0.01	12030.15	0.5
75%	-	0.44 đến 0.56	18036.33	1
Giá trị lớn nhất	-	Khoảng 4 đến 43	24039.93	1

(Nguồn: Tính toán của tác giả)

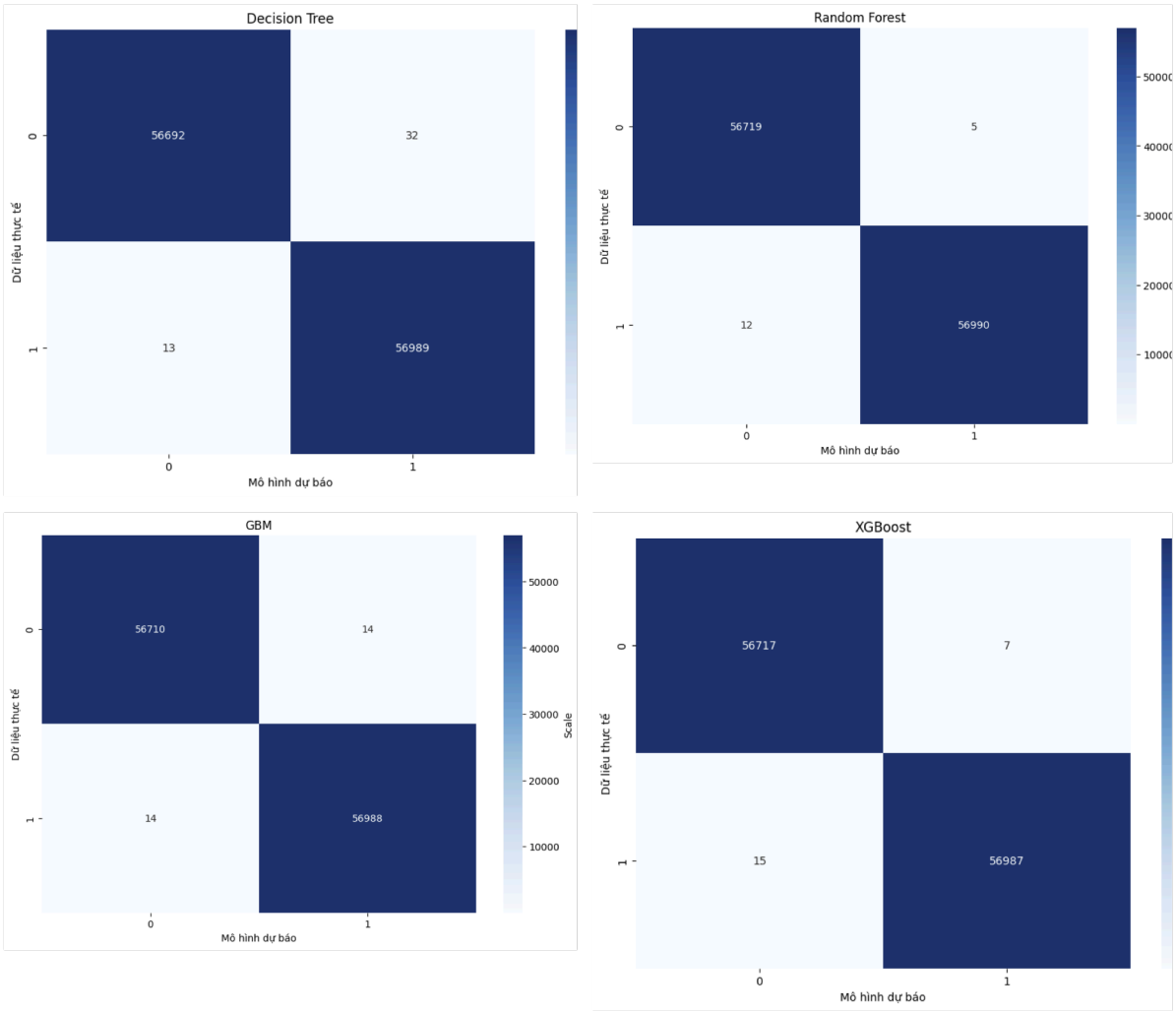
Bảng 4: *Kết quả của các mô hình trên tập dữ liệu kiểm định*

Mô hình	TN	FN	FP	TP	Accuracy	Recall	Precision	F1 Score
Decision Tree	56.692	13	32	56.989	99,92%	99,98%	99,94%	99,96%
Random Forest	56.719	12	5	56.990	99,985%	99,98%	99,99%	99,99%
GBM	56.710	14	14	56.988	99,97%	99,975%	99,98%	99,975%
XGBoost	56.717	15	7	56.987	99,98%	99,974%	99,988%	99,98%

(Nguồn: Tính toán của tác giả)

vào khả năng xử lý dữ liệu lớn và độ tin cậy cao. GBM cung cấp một sự cân bằng tốt giữa hiệu quả và chi phí, trong khi mô hình Decision Tree mang lại một giải pháp đơn giản và nhanh chóng cho các yêu cầu phù hợp với hệ thống và nguồn lực sẵn có.

Mô hình Random Forest đã thể hiện hiệu suất xuất sắc nhất trong số các mô hình được xét với độ chính xác tổng thể đạt tới 99.985%, thể hiện khả năng phân biệt hiệu quả giữa các giao dịch gian lận và hợp pháp. Mô hình này đã chính xác xác định 56.990 trường hợp gian lận và chỉ bỏ qua 12 trường hợp, phản ánh một độ nhạy (Recall) ấn tượng là 99,98%. Điều đáng chú ý, chỉ có 5 giao dịch hợp pháp được phân loại nhầm là gian lận (FP), cho thấy một tỷ lệ Precision rất cao là 99,99%. Điều này chỉ ra rằng mô hình có khả năng xác định đáng tin cậy các giao dịch gian lận mà không gây ra nhiều báo động giả. F1 Score, đo lường sự cân bằng giữa Precision và Recall, đạt mức 99,99%, củng cố hiệu quả



(Nguồn: Từ kết quả kiểm định của các mô hình)

Hình 2: *Mô tả ma trận nhầm lẫn của các mô hình học máy dựa trên cây*

của mô hình trong việc xử lý dữ liệu gian lận một cách chính xác và hiệu quả. Phân tích này không chỉ khẳng định tính ưu việt của mô hình Random Forest trong bối cảnh phát hiện gian lận thẻ tín dụng mà còn làm nổi bật tiềm năng của nó trong ứng dụng thực tế, đặc biệt trong lĩnh vực tài chính và ngân hàng, nơi đòi hỏi độ chính xác và độ tin cậy cao.

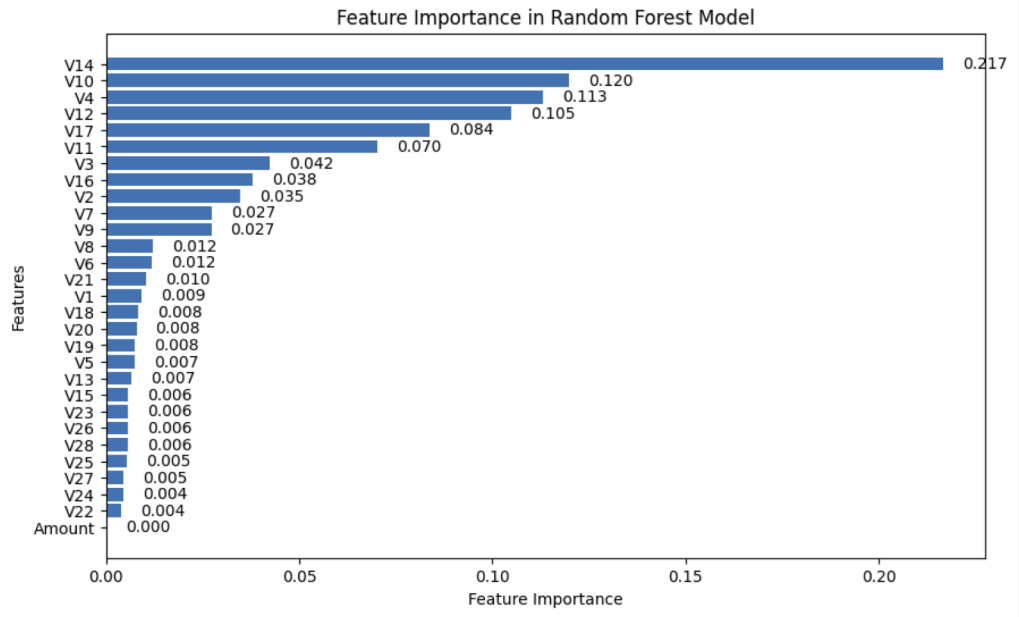
XGBoost cũng cho thấy hiệu suất rất cao với độ chính xác 99.98%, gần ngang bằng Random Forest nhưng với sô lượng FP thâp hơn là 7. Mô hình này có 56,987 TP và Precision đạt 99.988%, minh họa khả năng phát hiện chính xác các trường hợp gian lận mà ít ảnh hưởng đến các giao dịch hợp pháp. Độ nhạy 99.974% cùng F1 Score 99.98% cho thấy mô hình này cũng rất cân bằng, phù hợp cho các ứng dụng cần độ tin cậy cao.

GBM thể hiện độ chính xác 99.97% với TP là 56,988 và FP là 14. Mặc dù không đạt

được Precision cao như hai mô hình trên, 99.98% của GBM vẫn là rất ấn tượng, cho thấy hiệu quả trong việc giảm báo động giả. Độ nhạy 99.975% cùng F1 Score 99.975% cũng phản ánh khả năng cân bằng tốt giữa việc phát hiện và không bỏ sót các giao dịch gian lận. Mô hình này là một lựa chọn tốt khi cần một giải pháp vừa hiệu quả vừa kinh tế.

Mô hình Decision Tree có độ chính xác thấp nhất trong số các mô hình được phân tích là 99.92%, với TP là 56,989 và FP là 32. Tuy Precision 99.94% không cao bằng các mô hình kia, nhưng vẫn cho thấy khả năng phân biệt khá tốt giữa gian lận và hợp pháp. Độ nhạy 99.98% và F1 Score 99.96% chứng minh mô hình này vẫn hiệu quả trong việc phát hiện các trường hợp gian lận, phù hợp cho các ứng dụng cần giải pháp nhanh chóng và không quá phức tạp.

Kết quả nghiên cứu một lần nữa ủng hộ quan điểm của Jain và cộng sự (2020), Udeze và cộng sự (2022) và Tanwar và cộng sự (2023) khi cho rằng mô hình Random Forest cho kết quả tốt hơn so với các mô hình Gradient Boosting như GBM hay XGBoost trong quá trình phát hiện các giao dịch gian lận thẻ tín dụng. Nguyên nhân đến từ việc mô hình Random Forest ít nhạy cảm với việc chọn tham số, tuy nhiên các mô hình Gradient Boosting lại yêu cầu việc điều chỉnh kỹ lưỡng hơn các tham số như tỷ lệ học tập và số lượng cây để tránh hiện tượng quá khớp và đảm bảo mô hình hoạt động tốt. Bên cạnh đó, mô hình Random Forest cũng thường xử lý tốt với các loại dữ liệu và phân phối khác nhau, nhờ vào cơ chế bỏ phiếu của nhiều cây quyết định. Điều này giúp mô hình tổng quát hóa tốt hơn trên dữ liệu mới, trong khi đó các mô hình Gradient Boosting có thể có hiệu suất cao trên



(Nguồn:Tác giả tổng hợp)

Hình 3: *Xếp hạng tầm quan trọng của các thuộc tính đặc trưng trong mô hình Random Forest*

tập huấn luyện nhưng đôi khi không tổng quát hóa tốt trên dữ liệu mới do mô hình có thể quá phù hợp với dữ liệu huấn luyện.

Phân tích sâu hơn về các thuộc tính dự báo quan trọng trong mô hình Random Forest (Hình 3), chúng ta có thể thấy được rằng thuộc tính V14 có tầm quan trọng lớn nhất trong việc phát hiện các giao dịch gian lận trong thẻ tín dụng, với giá trị đóng góp tương ứng là 21.7%. Các thuộc tính V10, V4 và V12 cũng có mức độ quan trọng đáng kể khi lần lượt đóng góp với giá trị tương ứng là 12%, 11.3% và 10.5%, cho thấy đây là những yếu tố có tính dự báo và mức độ ảnh hưởng lớn trong mô hình. Các thuộc tính tiếp theo, bao gồm V17 và V11, có giá trị lần lượt là 8.4% và 7%, cũng đóng vai trò quan trọng nhưng không mạnh mẽ bằng các thuộc tính kể trên. Các thuộc tính có tầm quan trọng thấp hơn như V24, V22 và Amount có giá trị gần như bằng 0.000, cho thấy chúng ít có sự ảnh hưởng đến khả năng phát hiện gian lận thẻ tín dụng của mô hình. Thông qua biểu đồ này, chúng ta có thể thấy rõ những thuộc tính quan trọng nhất cần tập trung khi ứng dụng mô hình trong việc phát hiện các giao dịch gian lận thẻ tín dụng.

5. Kết luận

Trong bối cảnh phát triển nhanh chóng của công nghệ và sự phổ biến của các giao dịch trực tuyến, việc phát hiện và ngăn chặn gian lận thẻ tín dụng đã trở thành một thách thức đáng kể đối với ngành tài chính và ngân hàng. Nhóm mô hình học máy Tree-based, bao gồm Decision Tree, Random Forest, GBM và XGBoost, đã chứng minh hiệu quả ấn tượng trong việc giải quyết vấn đề này. Sự ưu việt của các mô hình này không chỉ đến từ khả năng phân loại dữ liệu lớn và phức tạp mà còn bởi khả năng tự động hóa và xử lý các giao dịch không cân xứng, một đặc điểm phổ biến trong dữ liệu gian lận.

Các kết quả thực nghiệm từ bộ dữ liệu lớn đã cung cấp bằng chứng rõ ràng về khả năng của các mô hình này trong việc cải thiện độ chính xác và giảm báo động giả. Random Forest đặc biệt nổi bật với hiệu suất cao, cho thấy sự cân bằng tốt giữa độ nhạy và độ chính xác, làm nổi bật khả năng của nó trong việc phát hiện các giao dịch gian lận mà không gây ra nhiều báo động giả. Điều này không chỉ củng cố vị trí của học máy trong chiến lược chống gian lận của các ngân hàng mà còn mở ra hướng đi mới cho việc ứng dụng công nghệ vào các giải pháp an ninh tài chính.

Một hạn chế đáng chú ý trong nghiên cứu này đó là không xác định được chính xác ý nghĩa của các thuộc tính từ V1 đến V28 trong bộ dữ liệu nghiên cứu. Nguyên nhân là các thuộc tính này đã được bên cung cấp dữ liệu mã hóa nhằm bảo vệ thông tin cá nhân của các khách hàng. Điều này dẫn đến việc tác giả không thể phân tích sâu hơn về tầm quan trọng của các thuộc tính trong mô hình dự báo. Trong tương lai, việc sử dụng bộ dữ liệu với các thuộc tính được diễn giải chi tiết hơn hứa hẹn sẽ cung cấp cái nhìn sâu sắc và định hướng rõ ràng cho các nghiên cứu trong việc phát hiện giao dịch gian lận.◆

***Tài liệu tham khảo*:**

Awoyemi, J. O., Adetunmbi, A. O., & Oluwadare, S. A. (2017). Credit card fraud detection using machine learning techniques:

A comparative analysis. *2017 International Conference on Computing Networking and Informatics (ICCNI)*, 1-9. https://doi.org/10.1109/ICCNI.2017.8123782.

Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). Credit Card Fraud Detection using Pipeling and Ensemble Learning. *Procedia Computer Science*, *173*, 104-112. https://doi.org/10.1016/j.procs.2020.06.014.

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(8), 3784-3797. https://doi.org/10.1109/TNNLS.2017.2736643.

Dornadula, V. N., & Geetha, S. (2019). Credit Card Fraud Detection using Machine Learning Algorithms. *Procedia Computer Science*, *165*, 631-641. https://doi.org/10.1016/j.procs.2020.01.057

Faraji, Z. (2022). A Review of Machine Learning Applications for Credit Card Fraud Detection with A Case study. *SEISENSE Journal of Management*, *5*(1), 49-59. https://doi.org/10.33215/sjom.v5i1.770.

Hiệp hội Ngân hàng Việt Nam. (2024). *Thị trường và xu hướng rủi ro, gian lận thanh toán thẻ*. https://vnba.org.vn/vi/thi-truong-va-xu-huong-rui-ro—gian-lan-thanh-toan-the-13799.htm.

Jain, V., Agrawal, M., & Kumar, A. (2020). Performance analysis of machine learning algorithms in credit cards fraud detection. *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, 86-88. https://doi.org/10.1109/ICRITO48877.2020.9197762.

Juniper Research. (2022). *Contactless Payments Transaction Values to Surpass $10 Trillion Globally by 2027*. https://www.juniper-research.com/press/contactless-payments-transaction-values-to-surpass/.

Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He, L., & Caelen, O. (2018). Sequence Classification for Credit-Card Fraud Detection. *Expert Systems with Applications*, *100*. https://doi.org/10.1016/j.eswa.2018.01.037.

Kaggle. (2023). *Credit Card Fraud Detection* [Dataset]. https://www.kaggle.com/datasets/nelgiriye-withana/credit-card-fraud-detection-dataset-2023/data.

Khalid, A., Owoh, N., Uthmani, O., Ashawa, M., Osamor, J., & John, A. (2024). Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach. *Big Data and Cognitive Computing*, *8*, 6. https://doi.org/10.3390/bdcc8010006.

Learning, I. T.-B. M. (2023). Credit Card Detection by Applying Interpretable Tree-Based Machine Learning Models. *2023 4th International Conference on E-Commerce and Internet Technology (ECIT 2023)*, *18*, 266.

Maniraj, S. P., Saini, A., Ahmed, S., & Sarkar, S. (2019). Credit card fraud detection using machine learning and data science. *International Journal of Engineering Research*, *8*(9), 110-115.

Nguyễn Thị Liên, Nguyễn Thị Thu Trang, & Nguyễn Chiến Thắng. (2018). Phương pháp học máy trong phát hiện gian lận thẻ tín

dụng - Một nghiên cứu thực nghiệm. *Tạp Chí Kinh Tế & Phát Triển*, *256*, 118-126.

Nuthalapati, A. (2023). Smart Fraud Detection Leveraging Machine Learning For Credit Card Security. *Educational Administration: Theory and Practice*, *29*, 433-443. https://doi.org/10.53555/kuey.v29i2.6907.

Tanwar, J., Singh, S., Kumar, A., Mittal, M., Singh, L., & Tripathi, S. (2023). Analysis of tree-based machine learning techniques for credit card fraud detection. In *Advancements in Cybercrime Investigation and Digital Forensics* (pp. 247-263). Apple Academic Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003369479-12/analysis-tree-based-machine-learning-techniques-credit-card-fraud-detection-jitender-tanwar-shubham-singh-akash-kumar-mandeep-mittal-leena-singh-sudhanshu-tripathi?context=ubx&refId=7329f564-74ee-4f30-9610-7c04af5942fa.

Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., & Kuruwitaarachchi, N. (2019). Real-time credit card fraud detection using machine learning. *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 488–493. https://doi.org/10.1109/CONFLUENCE.2019.8776942.

Udeze, C., Eteng, I., & Ibor, A. (2022). Application of Machine Learning and Resampling Techniques to Credit Card Fraud Detection. *Journal of the Nigerian Society of Physical Sciences*, 769. https://doi.org/10.46481/jnsps.2022.769.

Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., & Anderla, A. (2019). *Credit Card Fraud Detection - Machine Learning methods* (p. 5). https://doi.org/10.1109/INFOTEH.2019.8717766.

Summary

This study focuses on evaluating and comparing the effectiveness of tree-based machine learning models in predicting credit card fraud. The models considered include Decision Trees, Random Forests, Gradient Boosting Machines (GBM), and Extreme Gradient Boosting (XGBoost). The dataset used for this research includes 568,630 credit card transactions, with attributes from V1 to V28 transformed through Principal Component Analysis (PCA) to protect personal information. This study utilizes a confusion matrix and performance metrics such as Accuracy, Recall, Precision, and F1 Score to assess the effectiveness of each model. The results indicate that both Random Forest and XGBoost perform impressively, with Random Forest demonstrating superior capabilities in minimizing false alarms and accurately detecting fraudulent transactions. Despite some limitations in interpreting important attributes due to the anonymity of the data, this research provides significant insights into the potential application of machine learning models in detecting credit card fraud, paving the way for future studies using datasets with more clearly explained attributes.