BƯỚC ĐẦU ĐÁNH GIÁ HIỆU QUẢ CỦA MÔ HÌNH HỌC MÁY ĐỂ DỰ ĐOÁN SỐC TIM Ở BỆNH NHÂN NHỒI MÁU CƠ TIM CẤP

Hoàng Văn Sỹ[1,2], Lý Quang Sang[1,2], Nguyễn Minh Kha[1,2], Nguyễn Nhật Tài[1,2], Trần Nguyễn Phương Hải[2]

TÓM TẮT

Mở đầu: Sốc tim là biến chứng nghiêm trọng nhất của nhồi máu cơ tim (NMCT) cấp, liên quan đến tỷ lệ tử vong nội viện cao. Việc dự đoán sớm nguy cơ sốc tim ngay khi nhập viện có ý nghĩa quan trọng trong định hướng điều trị. Mô hình học máy gần đây đã được ứng dụng rộng rãi trong dự đoán biến cố tim mạch.

Mục tiêu: Xây dựng và đánh giá hiệu suất của các mô hình học máy trong dự đoán sốc tim nội viện ở bệnh nhân NMCT cấp, từ đó lựa chọn mô hình có tiềm năng ứng dụng lâm sàng.

Đối tượng – Phương pháp nghiên cứu: Nghiên cứu cắt ngang mô tả, có theo dõi dọc trên 316 bệnh nhân NMCT cấp. Bệnh nhân được phân thành hai nhóm: có diễn tiến sốc tim và không sốc tim. Dữ liệu được thu thập từ thời điểm nhập viện bao gồm thông tin nhân trắc, tiền căn, lâm sàng, xét nghiệm, ECG và siêu âm tim. Sau khi xử lý, dữ liệu được chia thành tập Training (70%) và tập Testing (30%). Kỹ thuật SMOTE-Tomek được áp dụng trên tập Training để xử lý mất cân bằng mẫu. Các biến đầu vào được chọn lọc bằng phương pháp SHAP. Tám mô hình học máy được xây dựng và so sánh hiệu suất dự đoán trên tập Testing.

Kết quả: Trong số 316 bệnh nhân NMCT cấp, tỷ lệ sốc tim nội viện là 10,1%. Mô hình AdaBoost đạt hiệu suất dự đoán tốt nhất với AUC = 0,924, Recall = 0,60, và F1-score = 0,571; cho thấy khả năng cân bằng tốt giữa độ nhạy và độ chính xác. Logistic Regression và LASSO Logistic cũng thể hiện hiệu quả tốt với AUC cao lần lượt là 0,902 và 0,912; với cùng Recall = 0,60 và F1-score lần lượt là 0,600 và 0,571. Tất cả các mô hình học máy đều vượt trội so với thang điểm truyền thống như GRACE (AUC = 0,755) và ORBI đơn giản (AUC = 0,780).

Kết luận: Ba mô hình học máy AdaBoost, Logistic Regression và LASSO Logistic cho thấy hiệu suất cao trong dự đoán sớm sốc tim nội viện ở bệnh nhân NMCT cấp. Kết quả này gợi mở khả năng tích hợp các mô hình học máy vào hệ thống hỗ trợ ra quyết định lâm sàng nhằm tối ưu hóa chăm sóc cho nhóm bệnh nhân nguy cơ cao.

Từ khoá: Sốc tim, nhồi máu cơ tim cấp, học máy, dự đoán sớm, AdaBoost

SUMMARY

PRELIMINARY EVALUATION OF MACHINE LEARNING MODELS FOR PREDICTING CARDIOGENIC SHOCK IN PATIENTS WITH ACUTE MYOCARDIAL INFARCTION

Background: Cardiogenic shock (CS) is the most severe complication of acute myocardial infarction (AMI), associated with high in-hospital mortality. Early prediction of CS at the time of admission is critical for guiding timely treatment. In recent years, machine learning (ML) models have been increasingly applied to predict major cardiovascular events.

Objective: To develop and evaluate the performance of various machine learning models in predicting in-hospital CS among patients with AMI, and to identify the model with the highest clinical application potential.

Methods: This was a cross-sectional descriptive study with longitudinal follow-up on 316 patients diagnosed with AMI. Patients were classified into two groups: with and without CS during hospitalization. Data were collected at admission, including medical history, clinical features, laboratory results, ECG, and echocardiography. After preprocessing, the dataset was randomly split into training (70%) and testing (30%) sets. The SMOTE-Tomek technique was applied to the training set to address class imbalance. Feature selection was performed using SHAP. Eight ML models were developed and their predictive performance was compared on the testing set.

Results: Among the 316 patients with AMI, the incidence of in-hospital CS was 10.1%. The AdaBoost model demonstrated the best predictive performance with an AUC of 0.924, recall of 0.60, and F1-score of 0.571, indicating a good balance between sensitivity

[1]*Đại học Y dược TP. HCM*
[2]*Bệnh viện Chợ Rẫy*
Tác giả chịu trách nhiệm chính: Hoàng Văn Sỹ
Email: hoangvansy@gmail.com
Mã bài: N925
Ngày nhận bài: 27/6/2026
Ngày phản biện khoa học: 24/6/2026
Ngày duyệt bài: 6/8/2026

and overall accuracy. Logistic Regression and LASSO Logistic also showed strong performance, with high AUCs of 0.902 and 0.912, respectively; both models achieved a recall of 0.60 and F1-scores of 0.600 and 0.571. All machine learning models outperformed traditional risk scores such as GRACE (AUC = 0.755) and the simplified ORBI score (AUC = 0.780).

Conclusion: The three machine learning models—AdaBoost, Logistic Regression, and LASSO Logistic—demonstrated high performance in the early prediction of in-hospital CS among patients with AMI. These findings highlight the potential for integrating machine learning models into clinical decision support systems to optimize care for high-risk patients.

Keywords: Cardiogenic shock, acute myocardial infarction, machine learning, early prediction, AdaBoost

I. ĐẶT VẤN ĐỀ

Tuổi, tiền căn đột quỵ/TIA, ngưng tim lúc khởi phát, NMCT thành trước, phân độ killip, tần số tim, huyết áp tâm thu, đường huyết. Sốc tim là biến chứng nặng nề và là nguyên nhân hàng đầu gây tử vong ở bệnh nhân nhồi máu cơ tim (NMCT) cấp. Tỷ lệ sốc tim dao động từ 5–10% trong nhóm bệnh nhân NMCT cấp, với tỷ lệ tử vong nội viện lên đến 40–50%, bất chấp những tiến bộ trong can thiệp mạch vành và điều trị hỗ trợ tuần hoàn.[1] Việc nhận diện sớm nguy cơ diễn tiến đến sốc tim có vai trò then chốt trong việc triển khai các biện pháp điều trị kịp thời nhằm cải thiện tiên lượng.

Hiện nay, một số thang điểm lâm sàng như GRACE hay ORBI được sử dụng để tiên đoán biến cố nặng trong hội chứng vành cấp, bao gồm cả sốc tim.[2] Tuy nhiên, các mô hình này chủ yếu được xây dựng bằng phương pháp hồi quy truyền thống, vốn chưa đủ linh hoạt để mô hình hóa các mối liên hệ phi tuyến giữa nhiều biến số phức tạp trong bối cảnh lâm sàng thực tế. Bên cạnh đó, nhiều yếu tố nguy cơ chỉ được ghi nhận muộn sau khi bệnh nhân đã có biến chứng, làm giảm hiệu quả cảnh báo sớm.

Trí tuệ nhân tạo (Artificial Intelligence - AI), đặc biệt là các thuật toán học máy (machine learning – ML), đang nổi lên như một công cụ tiềm năng trong việc xây dựng các mô hình tiên đoán y khoa có độ chính xác cao.[3] ML cho phép xử lý và học từ một lượng lớn dữ liệu đa biến, từ đó tạo ra các mô hình tiên đoán phù hợp với các đặc điểm lâm sàng đa dạng và phi tuyến của bệnh nhân. Một số nghiên cứu gần đây đã ứng dụng ML để dự đoán sốc tim nội viện sau NMCT và cho kết quả khả quan.[4]

Tại Việt Nam, chưa có nghiên cứu nào ứng dụng các thuật toán học máy để xây dựng mô hình dự đoán sốc tim ở bệnh nhân NMCT cấp trên dữ liệu thực tế. Do đó, chúng tôi thực hiện nghiên cứu này nhằm xây dựng và đánh giá hiệu suất các mô hình học máy trong tiên đoán sốc tim nội viện ở bệnh nhân NMCT cấp.

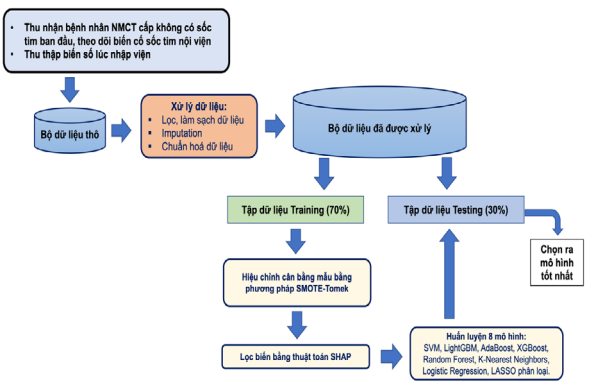
II. ĐỐI TƯỢNG-PHƯƠNG PHÁP

2.1. Thiết kế nghiên cứu

Đây là nghiên cứu mô tả cắt ngang có theo dõi dọc, được tiến hành tại Khoa Tim mạch Can thiệp – Bệnh viện Chợ Rẫy từ tháng 9 năm 2023 đến tháng 4 năm 2025. Tiêu chuẩn đưa vào bao gồm: bệnh nhân ≥18 tuổi nhập viện với chẩn đoán nhồi máu cơ tim cấp (STEMI hoặc NSTEMI), không có sốc tim tại thời điểm nhập viện ban đầu, và đồng ý tham gia nghiên cứu. Nghiên cứu loại trừ những bệnh nhân có sốc tim ngay khi nhập viện.

Nghiên cứu được thông qua bởi Hội đồng đạo đức trong nghiên cứu Y sinh học tại Đại học Y Dược TP. Hồ Chí Minh (Quyết định số 2247/ĐHYD-HĐĐĐ)

2.2. Quy trình nghiên cứu



Hình 1. Quy trình nghiên cứu

Tất cả bệnh nhân thỏa tiêu chuẩn nhận vào và không có tiêu chuẩn loại trừ được theo dõi diễn tiến trong thời gian nằm viện và được phân thành hai nhóm: có diễn tiến đến sốc tim và không sốc tim.

Tiêu chuẩn chẩn đoán sốc tim: huyết áp tâm thu ≤ 90 mmHg kéo dài ≥ 30 phút hoặc cần sử

dụng thuốc vận mạch/tăng co bóp để duy trì huyết áp tâm thu > 90 mmHg, kèm theo ít nhất một dấu hiệu giảm tưới máu mô như thay đổi tri giác, chi lạnh ẩm, thiểu niệu (<30 ml/giờ) hoặc lactate ≥ 2 mmol/L.

Dữ liệu thu thập bao gồm: thông tin nhân trắc học, tiền căn bệnh lý; đặc điểm lâm sàng, kết quả xét nghiệm huyết học – sinh hóa, điện tâm đồ, và siêu âm tim lúc nhập viện.

Dữ liệu thu thập được xử lý gồm: làm sạch, loại bỏ giá trị ngoại lai, điền dữ liệu thiếu bằng phương pháp imputation và chuẩn hóa dữ liệu theo phân phối chuẩn.

Tập dữ liệu sau đó được chia ngẫu nhiên thành tập dữ liệu Training (70%) và tập dữ liệu Testing (30%), đảm bảo phân bố tỷ lệ sốc tim – không sốc tim tương đương.

Tỷ lệ sốc tim thấp dẫn đến vấn đề mất cân bằng mẫu, dễ gây hiện tượng quá khớp (overfitting) khi xây dựng mô hình học máy. Để xử lý vấn đề này, kỹ thuật SMOTE-Tomek được áp dụng riêng trên tập dữ liệu Training.[5]

Biến đầu vào được chọn lọc bằng phương pháp SHAP (SHapley Additive exPlanations), chọn ra 20 biến quan trọng nhất được giữ lại để xây dựng mô hình.

Các mô hình học máy được sử dụng và huấn luyện trên tập dữ liệu Training: Support Vector Machine (SVM), Light Gradient Boosting Machine (LightGBM), AdaBoost, XGBoost, Random Forest, K-Nearest Neighbors (KNN) và Logistic Regression, LASSO phân loại.

Hiệu suất dự đoán của các mô hình được đánh giá trên tập dữ liệu Testing thông qua các chỉ số: độ chính xác (accuracy), diện tích dưới đường cong ROC (AUC), recall, precision, và F1 score. Mô hình có hiệu suất cao nhất sẽ được lựa chọn để đề xuất ứng dụng trong lâm sàng.

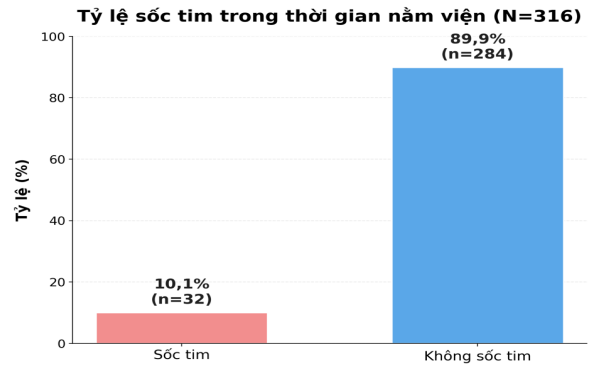
2.3. Phương pháp thống kê

Dữ liệu đầu vào cho các mô hình học máy được xử lý tiền xử lý bao gồm: điền dữ liệu thiếu bằng phương pháp imputation (trung bình hoặc giá trị phổ biến nhất), chuẩn hóa theo phân phối chuẩn (chuẩn hóa Z-score: $\mu = 0$, $\sigma^2 = 1$), và khắc phục mất cân bằng lớp bằng kỹ thuật SMOTE-Tomek.

Hiệu suất dự đoán của mô hình được đánh giá trên tập kiểm tra (test set) thông qua các chỉ số: độ chính xác (accuracy), diện tích dưới đường cong ROC (AUC), độ nhạy (recall), độ chính xác dương tính (precision) và F1-score. F1-score là trung bình điều hòa giữa recall và precision, đặc biệt hữu ích trong các bài toán phân loại với dữ liệu mất cân bằng.

III. KẾT QUẢ

Tổng cộng 316 bệnh nhân NMCT cấp nhập viện tại Khoa Tim mạch Can thiệp – Bệnh viện Chợ Rẫy từ tháng 09/2023 đến tháng 04/2025 được đưa vào phân tích, trong đó 32 trường hợp (10,1%) diễn tiến thành sốc tim nội viện.



Biểu đồ 1. Tỷ lệ sốc tim xảy ra trong thời gian nằm viện

Nhóm sốc tim có điểm GRACE và ORBI đơn giản cao hơn có ý nghĩa thống kê so với nhóm không sốc tim (đều p<0,001) (Bảng 1).

Bảng 1. So sánh điểm GRACE và điểm ORBI đơn giản (không tính đến biến số liên quan thủ thuật) giữa nhóm có và không có sốc tim

Đặc điểm	Chung (N=316)	Sốc tim (n=32)	Không sốc tim (n=284)	P
Điểm GRACE	155,5 ± 31,5	184,1 ± 34,2	152,3 ± 29,5	<0,001
Điểm ORBI đơn giản (không tính biến số thủ thuật)	2 (1-5)	6 (4-10)	2 (1-4)	<0,001

Sau khi chia theo tỷ lệ 70:30, tập Training gồm 221 ca (22 sốc tim), còn tập Testing có 95 ca (10 sốc tim). Tỷ lệ sốc tim ở cả hai tập tương đối cân đối, đảm bảo tính đại diện (Bảng 2).

Để cải thiện khả năng học của mô hình với tập Training mất cân đối, kỹ thuật SMOTE-Tomek được áp dụng, tạo ra một tập dữ liệu cân bằng với 199 ca sốc tim và 199 ca không sốc tim, giúp tối ưu hiệu suất dự đoán trên các mô hình học máy.

Bảng 2. Phân bố số lượng bệnh nhân sốc tim và không sốc tim theo từng tập dữ liệu

	Dữ liệu chung (n=316)	**Tập dữ liệu Training (n=221)**	**Tập dữ liệu Testing (n=95)**
Không sốc tim	284	199	85
Sốc tim	32	22	10

Trong tổng số 71 biến đầu vào, phương pháp SHAP đã chọn ra 20 biến quan trọng nhất có ảnh hưởng lớn đến khả năng dự đoán sốc tim. Các biến được chọn bao gồm: các biến lâm sàng (tuổi, STEMI thành trước, STEMI thất phải, huyết áp tâm thu, Killip II/III), xét nghiệm máu (hemoglobin, lympho, tiểu cầu, BUN, eGFR, kali, cholesterol toàn phần, triglyceride và troponin I), ECG (block nhánh phải), siêu âm tim (LVEF, RLVĐ thành trước, RLVĐ thất phải, tràn dịch màng ngoài tim), và điểm ORBI đơn giản.

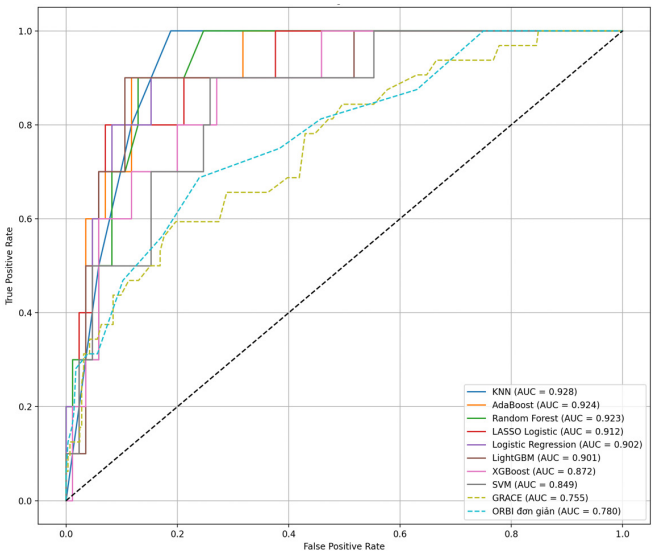
Trên tập Training (sau khi cân bằng bằng SMOTE-Tomek), hầu hết mô hình đạt AUC và F1-score gần 1, cho thấy khả năng học mạnh nhưng tiềm ẩn nguy cơ quá khớp. Trên tập kiểm tra, AdaBoost cho hiệu suất tổng thể tốt nhất với AUC = 0,924, Recall = 0,60, và F1 = 0,571, cho thấy mô hình cân bằng tốt giữa độ nhạy và độ chính xác. LASSO Logistic và Logistic Regression cũng là lựa chọn hiệu quả với AUC cao, Recall = 0,60, F1 lần lượt là 0,571 và 0,600, đồng thời có ưu điểm là mô hình đơn giản, dễ triển khai. (Bảng 3).

Mô hình KNN đạt AUC cao nhất (0,928) nhưng chỉ đạt Recall = 0,50 và F1 = 0,500, cho thấy hiệu quả phát hiện ca sốc tim còn hạn chế. LightGBM có AUC tốt (0,901) nhưng Recall và F1 thấp hơn nên không nổi bật. Các mô hình khác như Random Forest, XGBoost và SVM tuy có AUC cao, nhưng thể hiện tình trạng overfitting rõ rệt khi chỉ đạt Recall = 0,30 và F1-score thấp (≤ 0,353), không phù hợp để ứng dụng trong bối cảnh cần ưu tiên phát hiện sớm sốc tim.

Bảng 3. Hiệu suất chẩn đoán của các mô hình trên tập dữ liệu Training và Testing

Mô hình	**Trên tập dữ liệu Training**					**Trên tập dữ liệu Testing**				
	AUC	Accuracy	Recall	Precision	F1	AUC	Accuracy	Recall	Precision	F1
KNN	1,000	0,967	1,000	0,939	0,968	0,928	0,895	0,5	0,500	0,500
AdaBoost	1,000	0,995	0,990	1,000	0,995	0,924	0,905	0,6	0,545	0,571
Random Forest	1,000	1,000	1,000	1,000	1,000	0,923	0,884	0,3	0,429	0,353
LASSO Logistic	0,998	0,982	0,990	0,975	0,983	0,912	0,905	0,6	0,545	0,571
Logistic Regression	0,996	0,970	0,965	0,975	0,970	0,902	0,916	0,6	0,600	0,600
LightGBM	1,000	1,000	1,000	1,000	1,000	0,901	0,916	0,5	0,625	0,556
XGBoost	1,000	1,000	1,000	1,000	1,000	0,872	0,874	0,3	0,375	0,333
SVM	1,000	0,998	0,995	0,985	0,990	0,849	0,884	0,3	0,429	0,353

Phân tích đường cong ROC cho thấy hầu hết các mô hình học máy đều có khả năng phân biệt sốc tim tốt hơn so với thang điểm GRACE (AUC = 0,755) và ORBI đơn giản (AUC = 0,780). Điều này cho thấy so với các thang điểm truyền thống, các mô hình học máy có tiềm năng ứng dụng tốt hơn với độ chính xác cao trong việc dự đoán biến cố sốc tim nội viện ở bệnh nhân nhồi máu cơ tim cấp.



Biểu đồ 2. Đường cong ROC của các mô hình và của điểm GRACE và điểm ORBI đơn giản trên trên tập dữ liệu Testing

IV. BÀN LUẬN

Trong nghiên cứu này, chúng tôi chọn được ba mô hình tốt nhất gồm AdaBoost, Logistic Regression, và LASSO Logistic – đại diện cho các hướng tiếp cận mô hình hóa khác nhau nhưng đều thể hiện hiệu quả cao trong dự đoán sốc tim nội viện.

Mô hình AdaBoost trong nghiên cứu của chúng tôi đạt hiệu suất phân loại cao và giữ được độ ổn định giữa tập Training và Testing (AUC = 0,924; Recall = 0,60; F1 = 0,571), cho thấy khả năng ứng dụng tốt trong dự đoán sốc tim nội viện. Ưu điểm của AdaBoost nằm ở cơ chế điều chỉnh trọng số theo từng vòng lặp, giúp tối ưu hóa lỗi và tăng cường độ nhạy trong phát hiện các trường hợp biến cố ít gặp như sốc tim nội viện.

Logistic Regression được đưa vào so sánh như một mô hình tuyến tính cổ điển, có ưu điểm rõ ràng về tính minh bạch, dễ triển khai và giải thích. Mặc dù đơn giản, mô hình vẫn đạt AUC = 0,902 và Recall = 0,60 – tương đương với các mô hình phức tạp hơn, đồng thời F1 = 0,6 là cao nhất trong tất cả mô hình khảo sát. Điều này cho thấy hồi quy logistic vẫn là lựa chọn mạnh mẽ trong môi trường lâm sàng yêu cầu rõ ràng về giải thích mô hình.

LASSO Logistic Regression là phiên bản được chuẩn hóa với phạt L1, cho phép lựa chọn biến tự động. Mô hình này không những duy trì hiệu năng cao (AUC = 0,912; Recall = 0,60; F1 = 0,571), mà còn mang lại mô hình gọn nhẹ hơn, phù hợp với các ứng dụng thực tế cần tiết kiệm thời gian và tài nguyên xử lý.

Sự khác biệt về mô hình tối ưu giữa các nghiên cứu là điều dễ lý giải. Nghiên cứu của Bai và cộng sự xác định LASSO là mô hình tốt nhất nhờ đơn giản, ít overfitting;[6] trong khi nghiên cứu của Böhm và cộng sự (STOP SHOCK) lựa chọn hồi quy logistic với chuẩn hóa elastic net là mô hình tối ưu nhờ tính đơn giản, dễ hiệu chỉnh và khả năng hiệu năng cao khi được huấn luyện trên dữ liệu lớn từ bệnh án điện tử.[7] Mỗi mô hình phù hợp với đặc điểm dữ liệu khác nhau: dân số, cỡ mẫu, cách chọn biến, phân chia dữ liệu và đặc điểm kỹ thuật đều ảnh hưởng đến kết quả. Trong nghiên cứu này, dữ liệu được thu thập chủ động từ đầu, hạn chế biến mất và cho phép xây dựng mô hình dựa trên các thông số sẵn có tại nhập viện – một lợi thế đáng kể so với nhiều nghiên cứu quốc tế phụ

thuộc vào dữ liệu bệnh án điện tử, vốn dễ thiếu hoặc không đồng nhất.

Điểm chung quan trọng là cả nghiên cứu của chúng tôi và của tác giả Bai đều cho thấy các mô hình học máy vượt trội hơn GRACE và ORBI – những thang điểm hồi quy truyền thống, vốn khó phản ánh được tính chất phi tuyến và tương tác biến phức tạp của bệnh lý sốc tim.[6] Các mô hình học máy cho phép tích hợp nhiều loại dữ liệu đầu vào, tăng độ chính xác mà không phụ thuộc vào biến can thiệp hoặc theo dõi sau nhập viện.

Tuy nhiên, để mô hình được triển khai thực tế, cần tiến hành kiểm định ngoài trên các quần thể độc lập với quy mô lớn hơn và thiết kế giao diện ứng dụng lâm sàng thân thiện. Chỉ khi đảm bảo tính tổng quát, tính khả thi và khả năng tích hợp vào hệ thống lâm sàng, mô hình mới thực sự trở thành công cụ hỗ trợ ra quyết định hiệu quả trong điều trị bệnh nhân NMCT cấp có nguy cơ sốc tim.

V. HẠN CHẾ

Nghiên cứu đơn trung tâm với cỡ mẫu còn hạn chế, chưa thực hiện kiểm định ngoài, do đó khả năng tổng quát của mô hình còn cần được xác minh. Ngoài ra, mô hình chưa được tích hợp vào hệ thống lâm sàng để đánh giá tính khả thi trong thực tế.

VI. KẾT LUẬN

Nghiên cứu cho thấy các mô hình học máy, đặc biệt là AdaBoost, Logistic Regression và LASSO Logistic, có hiệu suất cao trong dự đoán sớm sốc tim nội viện ở bệnh nhân nhồi máu cơ tim cấp. Việc khai thác các thông số lâm sàng, cận lâm sàng và hình ảnh học ngay tại thời điểm nhập viện giúp mô hình dễ tích hợp vào quy trình phân tầng nguy cơ và hỗ trợ quyết định điều trị sớm. Trong đó, AdaBoost nổi bật với độ chính xác cao và khả năng cân bằng tốt giữa độ nhạy và độ đặc hiệu, trong khi Logistic Regression và LASSO Logistic có ưu điểm là mô hình đơn giản, dễ triển khai trong thực hành lâm sàng. Tuy nhiên, để áp dụng rộng rãi, cần thêm các nghiên cứu kiểm định ngoài và đánh giá tính khả thi khi tích hợp vào hệ thống chăm sóc lâm sàng thực tế.

NGUỒN TÀI TRỢ: Nghiên cứu này nhận được sự hỗ trợ từ Sở Khoa học và Công nghệ Thành phố Hồ Chí Minh (theo hợp đồng số 02/2025/HĐ-QKHCN ngày 07/02/2025).

TÀI LIỆU THAM KHẢO

1. **Vallabhajosyula S, Ahmad K.** The Price We Pay: Cardiogenic Shock After Acute Myocardial Infarction. American College of Cardiology Foundation Washington DC; 2024. p. 101048.
2. **Auffret V, Cottin Y, Leurent G, et al.** Predicting the development of in-hospital cardiogenic shock in patients with ST-segment elevation myocardial infarction treated by primary percutaneous coronary intervention: the ORBI risk score. *European heart journal*. 2018;39(22):2090-2102.
3. **Bommu R.** Machine Learning Applications in Cardiology: A Viable Practical Solution for Developing Countries. *The Metascience*. 2024;2(1):10-20.
4. **Aleman R, Patel S, Sleiman J, et al.** Cardiogenic shock and machine learning: a systematic review on prediction through clinical decision support softwares. *Journal of Cardiac Surgery*. 2021;36(11):4153-4159.
5. **Or B.** Improving Requirements Classification with SMOTE-Tomek Preprocessing. *arXiv preprint arXiv:250106491*. 2025;
6. **Bai Z, Hu S, Wang Y, et al.** Development of a machine learning model to predict the risk of late cardiogenic shock in patients with ST-segment elevation myocardial infarction. *Annals of Translational Medicine*. 2021;9(14):1162.
7. **Böhm A, Segev A, Jajcay N, et al.** Machine learning-based scoring system to predict cardiogenic shock in acute coronary syndrome. *European Heart Journal-Digital Health*. 2025:ztaf002.