

A retrospective study examining the use of machine learning models to diagnose chronic kidney disease in cats, utilizing hematological and blood biochemistry data

Dao H. Dang^{1*}, Chanh D. Tran¹, & Binh T. Tran²

¹Faculty of Veterinary Medicine, HUTECH University, Vietnam

²Department of Biochemistry, Faculty of Medicine, Ho Chi Minh City Medicine and Pharmacy University, Vietnam

ARTICLE INFO	ABSTRACT
Research Paper Received: August 12, 2025 Revised: September 24, 2025 Accepted: October 01, 2025 Keywords Blood biochemistry Cats Creatinine Kidney Machine learning *Corresponding author Dang Hoang Dao Email: dh.dao@hutech.edu.vn	The study was conducted from February 2025 to August 2025 at veterinary clinics in Ho Chi Minh City, Vietnam. This retrospective study analyzed hematological and blood biochemistry data from 60 cats (30 healthy and 30 with chronic kidney disease-CKD) to evaluate the efficacy of machine learning (ML) models in diagnosing CKD and identifying biomarkers. Parameters included red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), white blood cell count (WBC), neutrophil count (Neu#), eosinophil count (Eos#), lymphocyte count (Lym#), monocyte count (Mon#), platelet count (PLT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood urea nitrogen (BUN), creatinine (Crea), total protein (TP), albumin (ALB), and glucose (GLU). Data were preprocessed and split into 80% training and 20% testing sets. Seven ML models [Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Gradient Boosting, Support Vector Machine (SVM), and Naive Bayes] were assessed using accuracy, sensitivity, specificity, F1-score, and area under the curve (AUC). The SVM, Logistic Regression, and KNN achieved 92% accuracy, with SVM yielding the highest AUC (0.98). Decision Tree, Random Forest, and Naive Bayes performed less effectively (75 - 83%), often producing false positives. Creatinine was the most influential biomarker (importance score: 1,178), followed by Mon#, AST, Neu#, and WBC. The Crea, WBC, Neu#, ALT, and AST showed significant differences ($P < 0.01$) between healthy and CKD cats. In brief, ML, particularly SVM, holds promise for supporting the early diagnosis of CKD and enhancing clinical decision-making.

Cited as: Dang, D. H., Tran, C. D., & Tran, B. T. (2026). A retrospective study examining the use of machine learning models to diagnose chronic kidney disease in cats, utilizing hematological and blood biochemistry data. *The Journal of Agriculture and Development* 25(4), 59-73.

Nghiên cứu hồi cứu về ứng dụng mô hình học máy trong chẩn đoán bệnh thận mãn tính ở mèo dựa trên dữ liệu sinh lý và sinh hóa máu

Đặng Hoàng Đạo^{1*}, Trần Đình Chánh¹ & Trần Thanh Bình²

¹Khoa Thú Y, Trường Đại Học Công Nghệ TP.HCM, TP. Hồ Chí Minh

²Bộ Môn Hóa Sinh, Khoa Y, Đại Học Y Dược TP.HCM, TP. Hồ Chí Minh

THÔNG TIN BÀI BÁO	TÓM TẮT
Bài báo khoa học Ngày nhận: 12/08/2025 Ngày chỉnh sửa: 24/09/2025 Ngày chấp nhận: 01/10/2025 Từ khóa Creatinine Học máy Mèo Sinh hóa máu Thận *Tác giả liên hệ Đặng Hoàng Đạo Email: dh.dao@hutech.edu.vn	Nghiên cứu được thực hiện từ tháng 2/2025 đến 8/2025, tại các phòng khám thú y ở Thành phố Hồ Chí Minh, Việt Nam. Nghiên cứu hồi cứu phân tích dữ liệu sinh lý và sinh hóa máu của 60 mèo [30 khỏe mạnh và 30 mắc bệnh thận mãn tính (CKD)] để đánh giá hiệu quả học máy (ML) trong chẩn đoán CKD và xác định chỉ số sinh học. Các chỉ số gồm hồng cầu (RBC), Hemoglobin (HGB), hematocrit (HCT), bạch cầu (WBC), bạch cầu trung tính (Neu#), bạch cầu ưa axit (Eos#), tế bào lympho (Lym#), bạch cầu đơn nhân (Mon#), tiểu cầu (PLT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), ni-tơ trong urea huyết (BUN), Creatinine (Crea), protein tổng số (TP), albumin (ALB) và glucose (GLU). Dữ liệu được tiền xử lý, chia 80% huấn luyện và 20% kiểm tra. Bảy mô hình ML [Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Gradient Boosting, Support Vector Machine (SVM), Naive Bayes] được đánh giá qua độ chính xác, độ nhạy, độ đặc hiệu, F1-score và diện tích dưới đường cong (AUC). Mô hình SVM, Logistic Regression, KNN đạt độ chính xác 92%, với SVM có AUC cao nhất (0,98). Decision Tree, Random Forest, Naive Bayes kém hơn (75 - 83%), dễ tạo dương tính giả. Creatinine là chỉ số quan trọng nhất (hệ số ảnh hưởng: 1.178), tiếp theo là Mon#, AST, Neu#, WBC. Các chỉ tiêu Crea, WBC, Neu#, ALT và AST khác biệt đáng kể ($P < 0,01$) giữa mèo khỏe và mèo bệnh. Tóm lại, ML, đặc biệt SVM, hứa hẹn hỗ trợ chẩn đoán sớm CKD, nâng cao quyết định lâm sàng.

1. Đặt Vấn Đề

Bệnh thận mãn tính (CKD) là một bệnh lý phổ biến ở mèo, đặc biệt ở mèo cao tuổi, với tỷ lệ mắc lên đến 30 - 50% ở mèo trên 15 tuổi (Bartges, 2012). CKD ở mèo gây suy giảm tình trạng sức

khỏe tổng thể và làm rối loạn nhiều chỉ tiêu sinh lý và sinh hóa máu, bao gồm creatinine, urê, phosphate, canxi, cùng với các chỉ số huyết học liên quan đến tình trạng thiếu máu (Chakrabarti & ctv., 2013; Hokamp & Nabity, 2016). Tuy nhiên, những biến đổi này thường diễn ra âm thầm và

chỉ rõ rệt ở giai đoạn muộn, hạn chế khả năng chẩn đoán sớm bằng phương pháp truyền thống (Elliott & Barber, 1998; Sparkes & ctv., 2016).

Trong những năm gần đây, học máy (ML) đã trở thành công cụ tiềm năng trong y học thú y, cho phép phân tích lượng lớn dữ liệu lâm sàng để phát hiện các mẫu đặc trưng của bệnh (Banzato & ctv., 2018; Bradley & ctv., 2019). Các thuật toán như Random Forest, Support Vector Machine (SVM) và Gradient Boosting đã chứng minh hiệu quả trong phân loại và dự đoán tiến triển CKD ở mèo (Bradley & ctv., 2019; Biourge & ctv., 2020). ML không chỉ phân biệt mèo khỏe mạnh và mèo mắc CKD mà còn xác định các chỉ số sinh học quan trọng, hỗ trợ chẩn đoán sớm và tiên lượng bệnh (Vanden & ctv., 2025). Tuy vậy, hạn chế về bộ dữ liệu lớn và tính biến thiên sinh học giữa các cá thể vẫn là thách thức (Beam & Kohane, 2018; Esteva & ctv., 2019). Do đó, nghiên cứu này nhằm áp dụng ML để chẩn đoán CKD và xác định những chỉ số sinh lý, sinh hóa máu thay đổi rõ nhất, hướng tới phát triển hệ thống hỗ trợ ra quyết định lâm sàng chính xác và hiệu quả.

2. Vật Liệu và Phương Pháp Nghiên Cứu

2.1. Thời gian và địa điểm

Thời gian: Nghiên cứu được tiến hành từ tháng 2 đến 8 năm 2025.

Địa điểm nghiên cứu: Lấy mẫu và thu thập dữ liệu sinh lý, sinh hóa máu mèo tại các phòng khám Thú y khu vực Thành phố Hồ Chí Minh.

2.2. Thiết kế nghiên cứu

Nghiên cứu được thiết kế theo dạng mô tả - phân tích hồi cứu nhằm ứng dụng các mô hình học máy (Machine Learning) trong chẩn đoán

CKD và xác định các chỉ tiêu sinh lý, sinh hóa máu có ý nghĩa đặc trưng cho bệnh lý này.

Dữ liệu được thu thập từ hồ sơ xét nghiệm máu của mèo khỏe mạnh và mèo mắc bệnh thận đã được chẩn đoán xác định bằng phương pháp lâm sàng và cận lâm sàng tại các phòng khám Thú y khu vực Thành phố Hồ Chí Minh.

2.3. Đối tượng khảo sát

Đối tượng: Tất cả các thể mèo có độ tuổi dưới 6 tháng tuổi, từ 6 tháng tuổi đến 2 năm tuổi và trên 2 năm tuổi; về giới tính cả đực và cái; phương thức nuôi bao gồm nuôi nhốt và nuôi thả; giống mèo ghi nhận cả nội và ngoại đến khám và kiểm tra sức khỏe tại các phòng khám.

Tiêu chuẩn chọn mẫu gồm: Nhóm mèo khỏe - không có dấu hiệu lâm sàng, xét nghiệm sinh lý sinh và sinh hóa máu trong giới hạn tham chiếu và không có tiền sử bệnh thận; nhóm mèo bệnh CKD - được chẩn đoán xác định bởi bác sĩ thú y dựa trên triệu chứng lâm sàng như chán ăn, sụt cân, đa niệu, mất nước; về cận lâm sàng urê huyết cao, máu nhiễm toang, thiếu máu và đặc biệt protein niệu dai dẳng (0,2 - 0,4) làm gia tăng tốc độ tiến triển CKD do tăng áp lực cầu thận (Lees & ctv., 2005). Tiêu chuẩn loại trừ: mẫu thiếu > 30% dữ liệu hoặc mèo mắc bệnh hệ thống khác (viêm nhiễm nặng, bệnh gan mạn, bệnh chuyển hóa).

2.4. Biến số nghiên cứu

Nghiên cứu này tiến hành phân tích các chỉ tiêu huyết học bao gồm Red Blood Cell count (RBC), Hemoglobin (HGB), Hematocrit (HCT), White Blood Cell count (WBC), Neutrophil count (Neu#), Eosinophil count (Eos#), Lymphocyte count (Lym#), Monocyte count (Mon#) và Platelet count (PLT). Đồng thời, các

chỉ tiêu sinh hóa cũng được đánh giá bao gồm Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Blood Urea Nitrogen (BUN), Creatinine (Crea), Total Protein (TP), Albumin (ALB) và Glucose (GLU).

Biến mục tiêu (Label) được xác định dựa trên chẩn đoán lâm sàng: 0 = mèo khỏe, 1 = mèo bị bệnh thận mãn.

2.5. Tiền xử lý và làm sạch dữ liệu

Dữ liệu được xử lý bằng cách chuẩn hóa tên biến (loại bỏ khoảng trắng, ký tự đặc biệt, chú thích), kiểm tra và xử lý giá trị thiếu (tỷ lệ < 5% thay bằng trung vị, > 30% loại bỏ mẫu), mã hóa nhãn thành dạng nhị phân (0 = khỏe, 1 = bệnh), và chuẩn hóa các biến bằng StandardScaler để đưa về thang đo z-score, hỗ trợ tối ưu cho các mô hình nhảy với khoảng cách.

2.6. Chia tập dữ liệu

Tập dữ liệu gồm 60 mẫu (30 mèo khỏe và 30 mèo bệnh thận), được chia theo tỷ lệ 80% tập

huấn luyện (48 mẫu) và 20% tập kiểm tra (12 mẫu) bằng phương pháp train-test split để giữ nguyên tỉ lệ lớp (Han & ctv., 2022). Quá trình chia ngẫu nhiên được cố định (random_state = 48) nhằm đảm bảo tái lập. Trên tập huấn luyện, áp dụng k-fold cross-validation (k = 5) để kiểm soát overfitting và đánh giá độ ổn định của mô hình.

2.7. Xử lý dữ liệu

Dữ liệu được quản lý bằng Microsoft Excel và phân tích bằng phần mềm Python. Quá trình huấn luyện mô hình học máy được thực hiện với các thuật toán bao gồm: Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree và Random Forest, Gradient Boosting, Support Vector Machine (SVM), và Naive Bayes. Dùng t-test để so sánh các chỉ tiêu sinh lý sinh hóa của 2 nhóm biến độc lập (nhóm mèo khỏe và mèo bệnh).

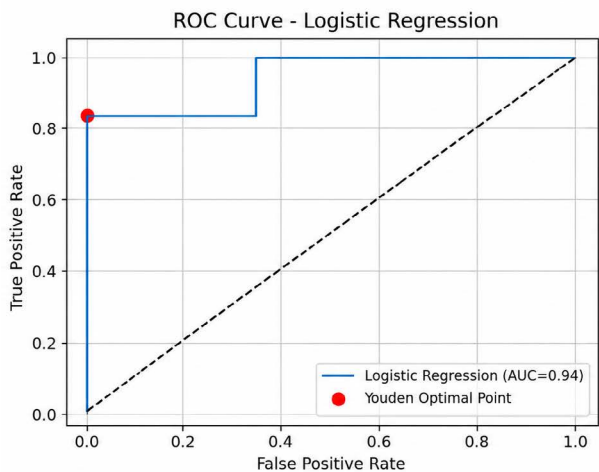
3. Kết Quả và Thảo Luận

Bảng 1. Kết quả đánh giá mô hình Logistic Regression trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	0,86	1,00	0,92	6,0
Mèo bệnh	1,00	0,83	0,91	6,0
Độ chính xác tổng thể			0,92	12,0
Trung bình Macro	0,93	0,92	0,92	12,0
Trung bình trọng số	0,93	0,92	0,92	12,0

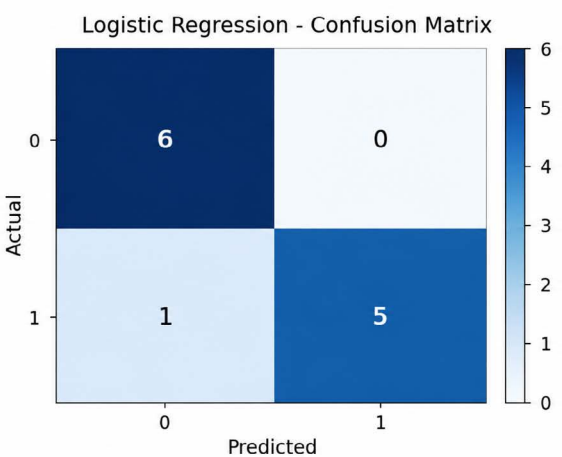
Qua Bảng 1, ta thấy mô hình Logistic Regression đạt độ chính xác tổng thể là 92% trên tập kiểm tra (11/12 mẫu được phân loại đúng). Precision cho nhóm mèo khỏe là 0,86 (6/7 mẫu được dự đoán là khỏe thực sự khỏe), trong khi Precision cho nhóm mèo bệnh là 1,00 (tất cả các mẫu được dự đoán là bệnh đều thực sự bệnh).

Điều này cho thấy mô hình có khả năng phân loại mèo bệnh tốt hơn mèo khỏe. Recall cho nhóm mèo khỏe là 1,00 (tất cả mèo khỏe đều được phân loại đúng), trong khi Recall cho nhóm mèo bệnh là 0,83 (5/6 mèo bệnh được phân loại đúng). F1-score cho cả hai nhóm đều cao (0,92 và 0,91), cho thấy sự cân bằng giữa Precision và Recall.



Hình 1. Đường cong ROC (Receiver Operating Characteristic) của mô hình Logistic Regression (Ngưỡng tối ưu (Youden) = 0,8; J = 0,83; Se = 0,83; Spe = 1,0).

Đường cong ROC (Receiver Operating Characteristic) của mô hình Logistic Regression cho thấy khả năng phân biệt tốt giữa mèo khỏe và mèo bệnh, với diện tích dưới đường cong (AUC) là 0,94 (Hình 1). Ngưỡng Youden tối ưu là 0,8; tương ứng với độ nhạy là 0,83 và độ đặc hiệu là 1,0. Điều này có nghĩa là, ở ngưỡng này, mô hình có thể phát hiện 83% số mèo bệnh và



Hình 2. Ma trận nhầm lẫn của mô hình Logistic Regression.

loại bỏ hoàn toàn các trường hợp dương tính giả (mèo khỏe bị chẩn đoán nhầm là bệnh).

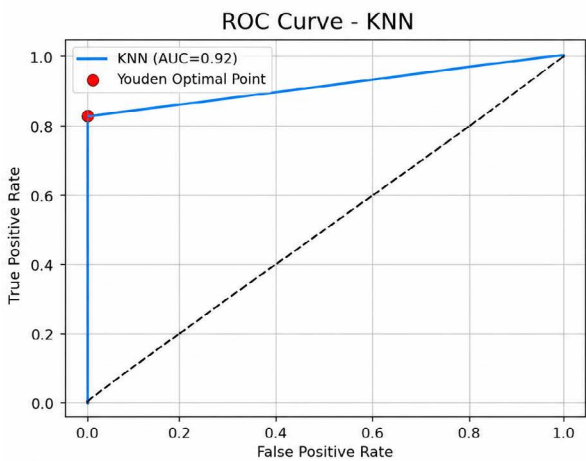
Qua Hình 2, cho thấy ma trận nhầm lẫn cho thấy mô hình Logistic Regression phân loại đúng 6/6 mèo khỏe và 5/6 mèo bệnh. Mèo bệnh bị phân loại nhầm là mèo khỏe. Điều này có thể là do sự tương đồng về các chỉ số sinh lý và sinh hóa máu giữa mèo bệnh ở giai đoạn đầu và mèo khỏe.

Bảng 2. Kết quả đánh giá mô hình K-Nearest Neighbors (KNN) trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	0,86	1,00	0,92	6,0
Mèo bệnh	1,00	0,83	0,91	6,0
Độ chính xác tổng thể			0,92	12,0
Trung bình Macro	0,93	0,92	0,92	12,0
Trung bình trọng số	0,93	0,92	0,92	12,0

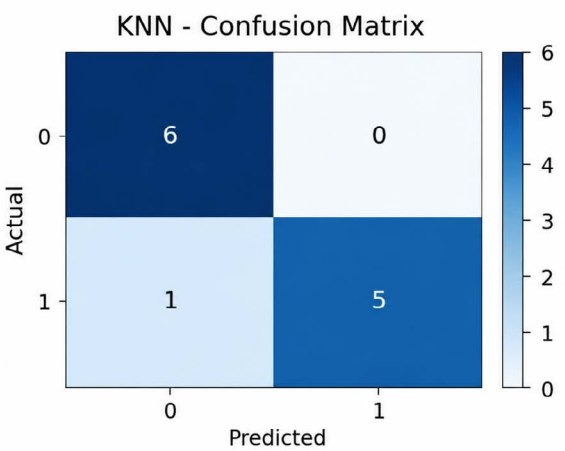
Tương tự như Logistic Regression, mô hình KNN cũng đạt độ chính xác tổng thể cao (92%) trên tập kiểm tra thể hiện qua (Bảng 2). Precision và Recall đều tốt cho cả hai nhóm,

với F1-score tương đương (khoảng 0,91 - 0,92). Điều này cho thấy KNN là một lựa chọn thay thế khả thi cho Logistic Regression trong việc chẩn đoán CKD ở mèo.



Hình 3. Đường cong ROC (Receiver Operating Characteristic) của mô hình K-Nearest Neighbors (KNN) (Ngưỡng tối ưu (Youden) = 0,6; J = 0,83; Se = 0,83; Spe = 1,0).

Hình 3 cho thấy đường cong ROC của mô hình KNN có AUC là 0,92, thấp hơn một chút so với Logistic Regression (0,94). Ngưỡng Youden tối ưu là 0,6 với độ nhạy 0,83 và độ đặc hiệu 1,0. Mặc dù, AUC thấp hơn, nhưng KNN vẫn có khả năng phân biệt tốt giữa mèo khỏe và mèo bệnh.



Hình 4. Ma trận nhầm lẫn của mô hình K-Nearest Neighbors (KNN).

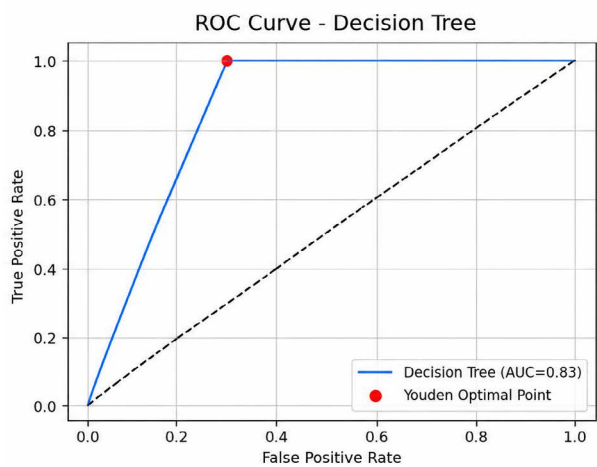
Hình 4 cho thấy ma trận nhầm lẫn của KNN tương tự như Logistic Regression, với một mèo bệnh bị phân loại nhầm là mèo khỏe. Điều này cho thấy cả hai mô hình đều có xu hướng mắc lỗi tương tự.

Bảng 3. Kết quả đánh giá mô hình Decision Tree trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	1,00	0,67	0,80	6,0
Mèo bệnh	0,75	1,00	0,86	6,0
Độ chính xác tổng thể			0,83	12,0
Trung bình Macro	0,88	0,83	0,83	12,0
Trung bình trọng số	0,88	0,83	0,83	12,0

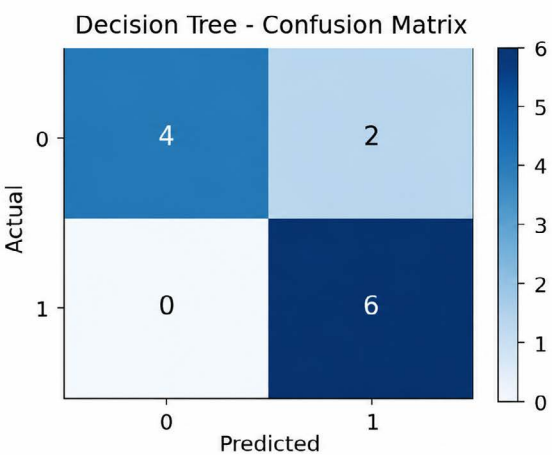
Mô hình Decision Tree có độ chính xác tổng thể thấp hơn (83%) so với Logistic Regression và KNN được trình bày qua (Bảng 3). Mặc dù, Recall cho nhóm mèo bệnh là 1,00 (tất cả mèo bệnh đều được phân loại đúng), nhưng Precision cho

nhóm này chỉ là 0,75 (một số mèo khỏe bị phân loại nhầm là bệnh). Điều này, cho thấy mô hình Decision Tree có xu hướng overfitting và tạo ra nhiều dương tính giả.



Hình 5. Đường cong ROC (Receiver Operating Characteristic) của mô hình Decision Tree (Ngưỡng tối ưu (Youden = 1,0; J = 0,67; Se = 1,0; Spe = 0,67).

Hình 5 cho thấy rằng đường cong ROC của mô hình Decision Tree có AUC là 0,83, thấp hơn đáng kể so với Logistic Regression và KNN. Ngưỡng Youden tối ưu là 1,0, với độ nhạy 1,0 và độ đặc hiệu 0,67. Điều này, khẳng định rằng mô hình Decision Tree có thể phát hiện tất cả mèo bệnh, nhưng đồng thời cũng phân loại nhầm nhiều mèo khỏe.



Hình 6. Ma trận nhầm lẫn của mô hình Decision Tree.

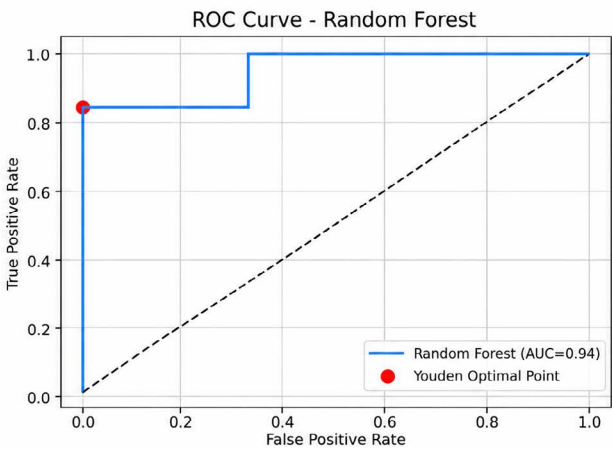
Hình 6 cho thấy ma trận nhầm lẫn của mô hình Decision Tree đã phân loại đúng tất cả mèo bệnh, nhưng lại phân loại nhầm 2 mèo khỏe là bệnh. Điều này khẳng định rằng mô hình Decision Tree có xu hướng tạo ra nhiều dương tính giả.

Bảng 4. Kết quả đánh giá mô hình Random Forest trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	0,80	0,67	0,73	6,0
Mèo bệnh	0,71	0,83	0,77	6,0
Độ chính xác tổng thể			0,75	12,0
Trung Bình Macro	0,76	0,75	0,75	12,0
Trung bình trọng số	0,76	0,75	0,75	12,0

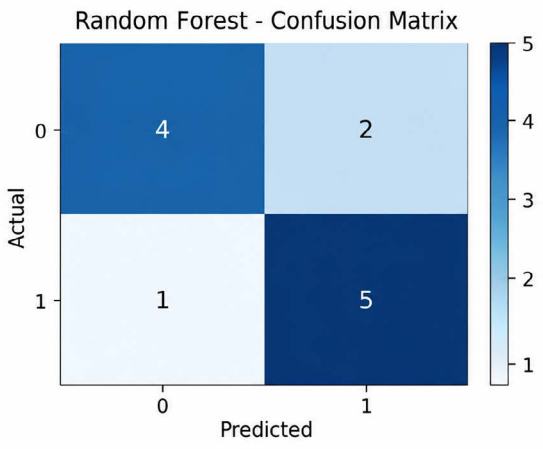
Tương tự như mô hình Decision Tree, mô hình Random Forest cũng có độ chính xác tổng thể tương đối thấp (75%) được trình bày qua

(Bảng 4). Precision và Recall đều không cao, cho thấy mô hình không hoạt động tốt trên tập dữ liệu này.



Hình 7. Đường cong ROC (Receiver Operating Characteristic) của mô hình Random Forest (Ngưỡng tối ưu (Youden) = 0,88; J = 0,83; Se = 0,83; Spe = 1,0).

Hình 7 cho thấy đường cong ROC của Random Forest có AUC là 0,94; cao hơn Decision Tree nhưng vẫn thấp hơn Logistic Regression và KNN. Điều này, cũng cho thấy rằng mô hình Random Forest có khả năng phân biệt giữa mèo khỏe và



Hình 8. Ma trận nhầm lẫn của mô hình Random Forest.

mèo bệnh tốt hơn mô hình Decision Tree, nhưng vẫn không hiệu quả bằng các mô hình tuyến tính.

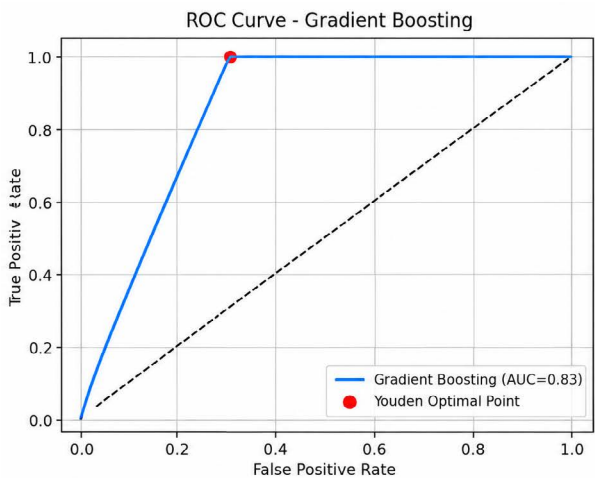
Hình 8 cho thấy ma trận nhầm lẫn của Random Forest mô hình phân loại nhầm cả mèo khỏe và mèo bệnh, cho thấy hiệu suất không ổn định.

Bảng 5. Kết quả đánh giá mô hình Gradient Boosting trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	1,00	0,67	0,80	6,0
Mèo bệnh	0,75	1,00	0,86	6,0
Độ chính xác tổng thể			0,83	12,0
Trung bình Macro	0,88	0,83	0,83	12,0
Trung bình trọng số	0,88	0,83	0,83	12,0

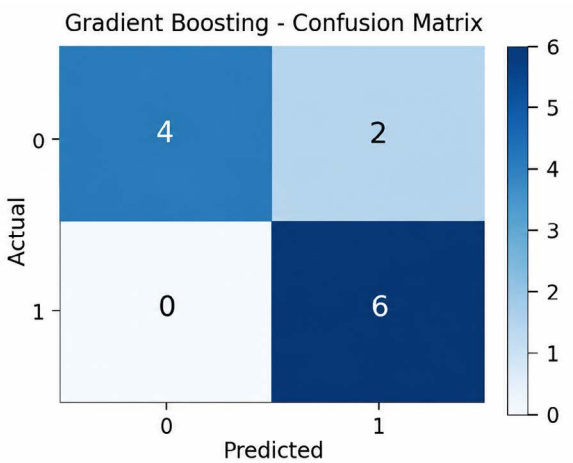
Bảng 5 cho thấy mô hình Gradient Boosting có kết quả tương tự Decision tree, độ chính xác

83%, Recall mèo bệnh 100% nhưng độ chính xác (Precision) của nhóm này chỉ 0,75; cho thấy có tình trạng dương tính giả.



Hình 9. Đường cong ROC (Receiver Operating Characteristic) của mô hình Gradient Boosting (Ngưỡng tối ưu (Youden) = 1,0; J = 0,67; Se = 1,0; Spe = 0,67).

Hình 9 cho thấy đường cong ROC của Gradient Boosting có AUC là 0,83, tương tự mô hình Decision Tree. Ngưỡng Youden tối ưu là 1,0, với độ nhạy 1,0 và độ đặc hiệu 0,67. Điều này cho thấy Gradient Boosting tương tự mô hình Decision Tree có thể phát hiện tất cả mèo bệnh, nhưng đồng thời cũng phân loại nhầm nhiều mèo khỏe.



Hình 10. Ma trận nhầm lẫn của mô hình Gradient Boosting.

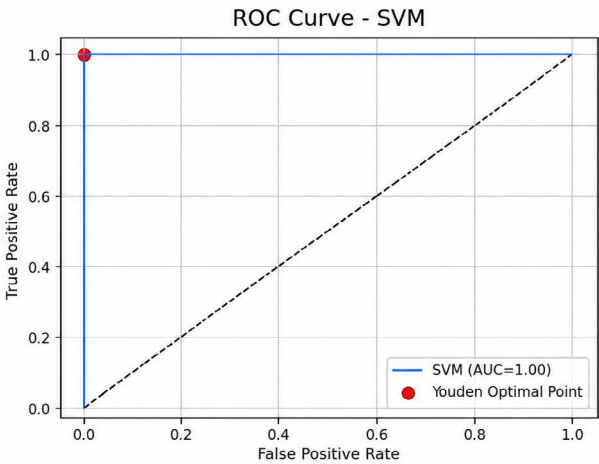
Ma trận nhầm lẫn của Gradient Boosting (Hình 10) cho thấy mô hình phân loại đúng tất cả mèo bệnh, nhưng lại phân loại nhầm 2 mèo khỏe là bệnh. Điều này khẳng định rằng Gradient Boosting có xu hướng tạo ra nhiều dương tính giả.

Bảng 6. Kết quả đánh giá mô hình Support Vector Machine (SVM) trên tập kiểm tra

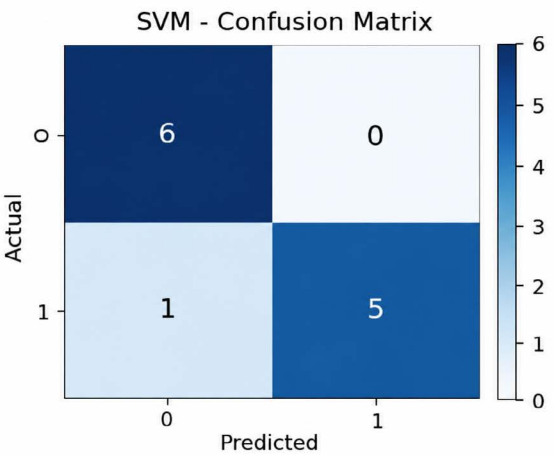
Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	0,86	1,00	0,92	6,0
Mèo bệnh	1,00	0,83	0,91	6,0
Độ chính xác tổng thể			0,92	12,0
Trung bình macro	0,93	0,92	0,92	12,0
Trung bình trọng số	0,93	0,92	0,92	12,0

Qua kết quả Bảng 6 cho thấy mô hình Support Vector Machine (SVM) đạt độ chính xác tổng thể cao (92%) trên tập kiểm tra, tương tự như Logistic Regression và KNN. Precision

và Recall đều tốt cho cả hai nhóm, với F1-score tương đương (khoảng 0,91 - 0,92). Điều này cho thấy SVM là một lựa chọn tốt cho việc chẩn đoán CKD ở mèo.



Hình 11. Đường cong ROC (Receiver Operating Characteristic) của mô hình Support Vector Machine (SVM) (Ngưỡng tối ưu (Youden) = 0,4; J = 1,0; Se = 1,0; Spe = 1,0).



Hình 12. Ma trận nhầm lẫn của mô hình Support Vector Machine (SVM).

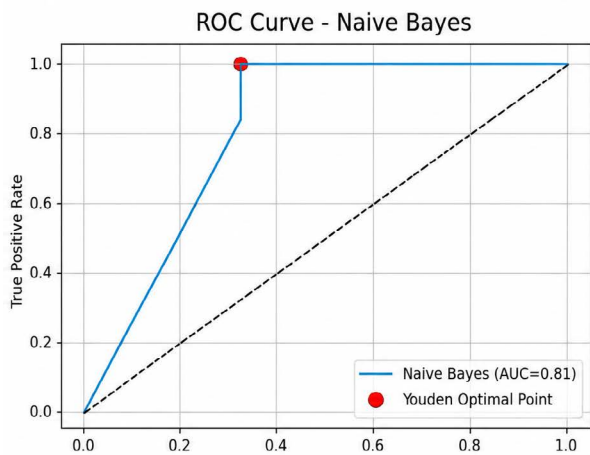
Đường cong ROC của mô hình SVM (Hình 11) có AUC là 1,00, cao nhất trong số các mô hình. Điều này cho thấy SVM có khả năng phân biệt tốt nhất giữa mèo khỏe và mèo bệnh.

Ma trận nhầm lẫn của SVM (Hình 12) tương tự như Logistic Regression và KNN, với 1 mèo bệnh bị phân loại nhầm là mèo khỏe.

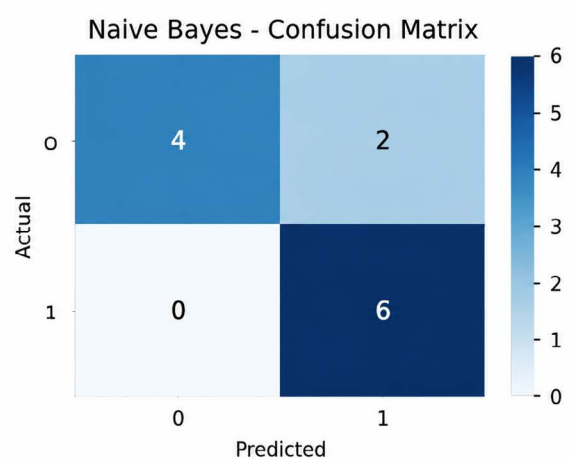
Mô hình Naïve Bayes có độ chính xác tổng thể thấp hơn (83%) so với Logistic Regression, KNN và SVM (Bảng 7). Mặc dù, Recall cho nhóm mèo bệnh là 1,00, nhưng Precision cho nhóm này chỉ là 0,75. Điều này cho thấy Naïve Bayes có xu hướng tạo ra nhiều dương tính giả.

Bảng 7. Kết quả đánh giá mô hình Naïve Bayes trên tập kiểm tra

Đặc tính	Độ chính xác	Độ nhớ	Điểm F1	Tập kiểm tra
Mèo khỏe	1,00	0,67	0,80	6,0
Mèo bệnh	0,75	1,00	0,86	6,0
Độ chính xác tổng thể			0,83	12,0
Trung bình macro	0,88	0,83	0,83	12,0
Trung bình trọng số	0,88	0,83	0,83	12,0



Hình 13. Đường cong ROC (Receiver Operating Characteristic) của mô hình Naïve Bayes (Ngưỡng tối ưu (Youden) = 0,51; J = 0,67; Se = 1,0; Spe = 0,67).

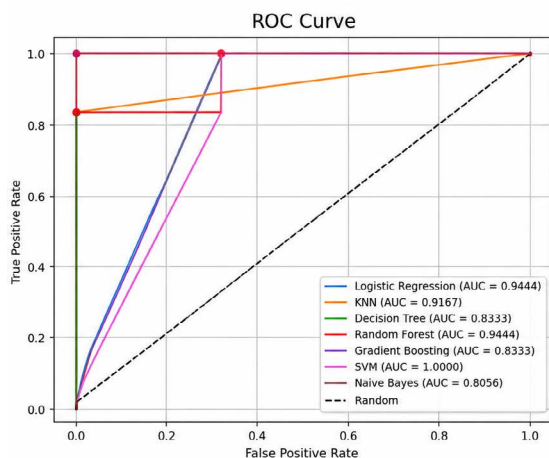


Hình 14. Ma trận nhầm lẫn của mô hình Naive Bayes.

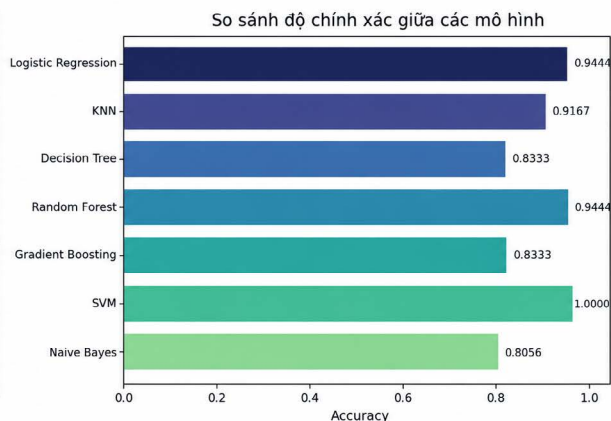
Hình 13 cho thấy đường cong ROC của Naïve Bayes có AUC là 0,81, thấp nhất trong số các mô hình. Điều này, cho thấy mô hình Naïve Bayes có khả năng phân biệt kém nhất giữa mèo khỏe và mèo bệnh. Hình 14 cho thấy ma trận nhầm lẫn của Naïve Bayes đã phân loại nhầm nhiều mèo khỏe là bệnh.

Hình 15 cho thấy Đường cong ROC của SVM đạt AUC = 1,00, vượt trội so với Logistic

Regression (0,94), KNN (0,92), Random Forest (0,94), Decision Tree (0,83), Gradient Boosting (0,83), và Naive Bayes (0,81) trong chẩn đoán CKD ở mèo. SVM tối ưu hóa phân biệt với TPR = 1,0 và FPR = 0,0 (ngưỡng Youden = 1,0), trong khi các mô hình khác bị hạn chế bởi overfitting hoặc giả định không phù hợp (John và ctv., 2010).



Hình 15. So sánh đường cong ROC (Receiver Operating Characteristic) của các mô hình.



Hình 16. So sánh độ chính xác của các mô hình.

Hình 16 cho thấy SVM cao nhất về mức độ chính xác $\sim 0,92$, phù hợp AUC = 1,00 (Hình 15), nhờ tối ưu hóa siêu phẳng (Vapnik, 2013). Logistic Regression và KNN đạt $\sim 0,92$, Random Forest và Gradient Boosting $\sim 0,83 - 0,85$,

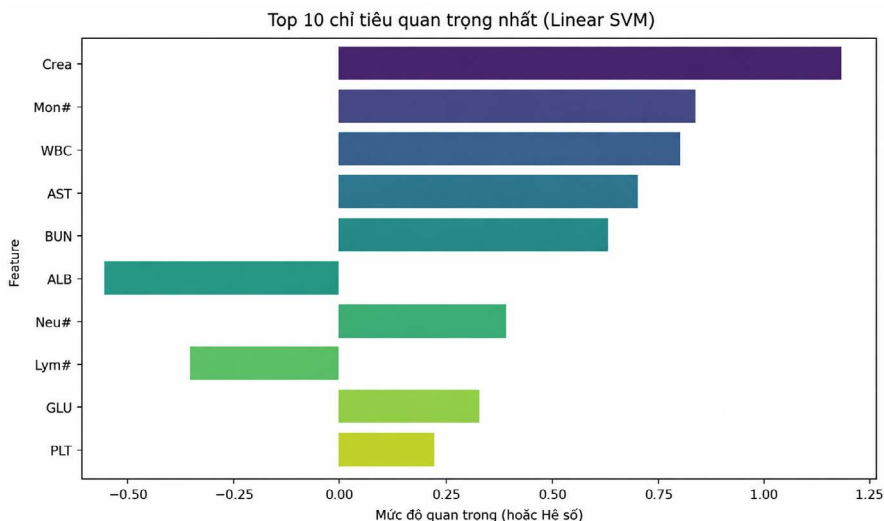
Decision Tree và Naive Bayes $\sim 0,83$. SVM tối ưu cho CKD ở mèo (độ chính xác $\sim 0,92$), với Crea, WBC, Neu#, ALT, AST ($P < 0.01$) quan trọng (Chakrabarti & ctv., 2013).

Bảng 8. Top 10 chỉ số ảnh hưởng nhất đến dự đoán bệnh thận theo mô hình Support Vector Machine

Xếp hạng	Chỉ số	Hệ số ảnh hưởng
1	Creatinine	1,178
2	Bạch cầu đơn nhân (Mon#)	0,824
3	Bạch cầu (WBC)	0,806
4	Aspartate aminotransferase (AST)	0,716
5	Ni-tơ trong ure huyết (BUN)	0,646
6	Albumin (ALB)	-0,527
7	Bạch cầu trung tính (Neu#)	0,415
8	Tế bào lympho (Lym#)	-0,355
9	Glucose (GLU)	0,336
10	Tiểu cầu (PLT)	0,222

Bảng 8 cho thấy mô hình SVM chỉ ra rằng Crea là hệ số ảnh hưởng lớn nhất đến khả năng dự đoán bệnh thận ở mèo (Hệ số ảnh hưởng = 1,178). Kết quả này hoàn toàn phù hợp với vai trò đã được công nhận rộng rãi của Crea trong đánh giá chức năng thận và chẩn đoán bệnh thận

mạn tính (CKD) (Hall & ctv., 2014). Crea là một sản phẩm thải của quá trình chuyển hóa cơ bắp, được đào thải chủ yếu qua thận. Khi chức năng thận suy giảm, nồng độ Crea trong máu tăng lên, phản ánh khả năng lọc của cầu thận bị suy yếu.



Hình 17. Mức độ quan trọng của các chỉ dấu sinh học trong dự đoán bệnh thận của mô hình Support Vector Machine (SVM).

Qua Hình 17, cho thấy Crea cao hơn hẳn so với các chỉ số khác về mức độ ảnh hưởng việc dự đoán CKD. Điều này khẳng định vai trò then chốt của Crea trong chẩn đoán CKD và cho thấy mô hình SVM đã học được mối quan hệ này từ dữ liệu.

Qua kết quả Bảng 9, cho thấy Crea là chất chỉ dấu sinh học nổi bật giúp phân biệt mèo bị CKD và mèo khỏe mạnh, với mức tăng đáng kể ($P < 0,001$) từ $2,42 \pm 4,52 \mu\text{mol/L}$ lên $17,43 \pm 18,34 \mu\text{mol/L}$. Về mặt sinh lý thì Crea là một sản phẩm phụ của quá trình chuyển hóa creatine phosphate trong cơ bắp, nơi cung cấp năng lượng cho các hoạt động cơ cơ và được thải ra ngoài qua quá trình lọc cầu thận (glomerular filtration)

mà không bị tái hấp thu đáng kể ở ống thận. Sự gia tăng này phản ánh tình trạng azotemia và suy giảm glomerular filtration rate (GFR), củng cố vai trò của Crea như chỉ dấu tiêu chuẩn cho CKD ở mèo, dù có hạn chế ở giai đoạn tiền lâm sàng do biến thiên cá thể (Lawson & ctv., 2015). Ở mèo, đặc biệt những con trên 15 tuổi (30 - 50% mắc CKD) (Bartges, 2012), nồng độ creatinine tăng cao ($P < 0,001$) phản ánh tình trạng suy giảm GFR, thường đi kèm với tổn thương mạn tính ở cầu thận hoặc ống thận. Trong bối cảnh ứng dụng machine learning (ML), Crea có thể là chỉ dấu sinh học quan trọng và khi kết hợp với các chỉ dấu khác như symmetric dimethylarginine (SDMA) sẽ nâng cao khả năng phát hiện sớm (Bradley & ctv., 2019).

Bảng 9. So sánh giá trị trung bình của 10 chỉ tiêu sinh lý và sinh hóa máu quang trọng giữa mèo khỏe mạnh và mèo mắc bệnh thận

Chỉ tiêu ¹	Mèo khỏe ($\bar{X} \pm \text{SD}$)	Mèo bệnh ($\bar{X} \pm \text{SD}$)	Chênh lệch tuyệt đối	P^2
PLT (109/L)	209,73 $\pm$ 98,63	283,83 $\pm$ 100,26	74,10	0,005
ALT (U/L)	65,50 $\pm$ 28,60	117,37 $\pm$ 97,59	51,87	0,009
AST (U/L)	23,40 $\pm$ 11,32	70,50 $\pm$ 78,54	47,10	0,003
GLU (mmol/L)	103,62 $\pm$ 30,47	139,42 $\pm$ 99,92	35,80	0,069
Crea ($\mu\text{mol/L}$)	2,42 $\pm$ 4,52	17,43 $\pm$ 18,34	15,02	< 0,001
WBC (109/L)	11,14 $\pm$ 3,50	21,49 $\pm$ 10,83	10,35	< 0,001
Neu# (109/L)	7,38 $\pm$ 3,33	17,48 $\pm$ 11,20	10,11	< 0,001
HGB (g/L)	133,00 $\pm$ 30,03	123,17 $\pm$ 43,38	9,83	0,312
TP (g/L)	11,18 $\pm$ 16,94	7,13 $\pm$ 1,29	4,05	0,202
BUN (mmol/L)	21,36 $\pm$ 8,50	23,72 $\pm$ 8,43	2,36	0,284

¹PLT: platelet, ALT: alanine aminotransferase, AST: aspartate aminotransferase, GLU: glucose, Crea: creatinine, WBC: white blood cell, Neu#: neutrophil count, HGB: hemoglobin, TP: total protein, BUN: blood urea nitrogen.

²Independent t-test: của 2 nhóm biến độc lập phân phối chuẩn.

Ngoài ra, số lượng bạch cầu (WBC) và bạch cầu trung tính (Neu#) cũng tăng rõ rệt ($P < 0,001$), WBC đóng vai trò quan trọng trong hệ miễn dịch, giúp bảo vệ cơ thể khỏi nhiễm trùng và viêm, WBC và Neu# chiếm ưu thế thì gợi ý tình trạng viêm hệ thống hoặc phản ứng miễn dịch mạn tính liên quan đến tổn thương thận, có thể do tác động của uremic toxins (Hokamp & Nabity, 2016). Kết quả này tương đồng với các báo cáo về vai trò của viêm mạn trong tiến triển CKD ở mèo, nơi neutrophil-to-lymphocyte ratio (NLR) tăng cao ở giai đoạn azotemic (Chakrabarti & ctv., 2013). Ở mèo bị CKD, WBC tăng cao không chỉ phản ánh viêm mà còn liên quan đến stress oxy hóa và tổn thương mô thận. Tuy nhiên, mức tăng có thể biến thiên do nhiễm trùng đồng mắc, đòi hỏi phân tích kỹ lưỡng để tránh chẩn đoán sai và việc tích hợp thông tin viêm vào ML có thể cải thiện khả năng dự đoán và phân biệt CKD với các bệnh lý viêm khác (Sparkes & ctv., 2016).

Đáng chú ý, men gan ALT và AST cũng tăng có ý nghĩa thống kê ($P \leq 0,009$), ALT và AST là các enzyme chuyển aminase, chủ yếu được tìm thấy trong gan, nhưng cũng hiện diện ở mức thấp trong thận, cơ tim, và cơ xương, vấn đề này có thể chỉ ra rằng CKD có thể đồng mắc tổn thương gan hoặc rối loạn chuyển hóa, phù hợp với nhận định của Twedt (2006). Trong ML, các chỉ số này có thể đóng vai trò đặc trưng phụ, giúp giảm dương tính giả khi phân biệt CKD đa cơ quan tới bệnh ở gan.

Ngược lại, GLU, HGB, TP và BUN không khác biệt đáng kể ($P > 0,05$), GLU là nguồn năng lượng chính, được điều hòa bởi insulin và bài tiết qua thận khi vượt ngưỡng tái hấp thu (khoảng 180 mg/dL ở mèo) và có thể biến thiên do chế độ ăn hoặc stress cũng như biến động bởi nhiều yếu tố ngoài thận (Bartges, 2012). Phân tử HGB

vận chuyển oxy trong máu, thường giảm trong CKD do thiếu máu (anemia) liên quan đến thiếu erythropoietin từ thận tổn thương, Giảm HGB phản ánh giai đoạn muộn của CKD, nhưng không phải chỉ số chẩn đoán sớm (Chakrabarti và ctv., 2013). BUN là sản phẩm chuyển hóa protein, được thận bài tiết, là chất chỉ dấu sinh học phụ cho CKD, nhạy hơn creatinine ở giai đoạn sớm, nhưng dễ bị ảnh hưởng bởi chế độ ăn (Lawson và ctv., 2015). Trong ML, giảm trọng số hoặc loại bỏ các đặc trưng này giúp tránh overfitting, tập trung vào chỉ dấu đặc hiệu hơn (Biourge & ctv., 2020).

4. Kết Luận

Nghiên cứu cho thấy các mô hình của ML đặc biệt là Support Vector Machine, Logistic Regression và K-Nearest Neighbors, đạt độ chính xác cao (92%) trong chẩn đoán bệnh thận mạn tính ở mèo, với SVM có khả năng phân biệt vượt trội ($AUC = 1,00$). Crea là chỉ số sinh học quan trọng nhất, cùng với WBC, Neu#, ALT và AST, cho thấy sự khác biệt đáng kể ($P < 0,01$) giữa mèo khỏe và mèo bệnh. Các mô hình Decision Tree, Random Forest và Naive Bayes kém hiệu quả hơn (75 - 83%), dễ tạo dương tính giả.

Lời Cam Đoan

Nhóm nghiên cứu không có sự tranh chấp và cam đoan chưa từng công bố bài báo này trước đó.

Tài Liệu Tham Khảo (References)

- Banzato, T., Bedin, M., Burti, S., Zotti, A., & Bonsembiante, F. (2018). Machine learning classification of feline polycystic kidney disease using ultrasonographic texture analysis. *Journal of Veterinary Internal Medicine* 32(6), 2043-2049. <https://doi.org/10.1111/jvim.15323>.

- Bartges, J. W. (2012). Chronic kidney disease in dogs and cats. *Veterinary Clinics of North America: Small Animal Practice* 42(4), 669-692. <https://doi.org/10.1016/j.cvsm.2012.04.008>.
- Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Journal of The American Medical Association* 319(13), 1317-1318. <https://doi.org/10.1001/jama.2017.18391>.
- Biourge, V., Delmotte, S., Feugier, A., Bradley, R., McAllister, M., & Elliott, J. (2020). An artificial neural network-based model to predict chronic kidney disease in aged cats. *Journal of Veterinary Internal Medicine* 34(5), 1920-1931.
- Bradley, R., Tagkopoulos, I., Kim, M., Kokkinos, Y., Plevritis, S., Rubin, D. L., & Yu, K. (2019). Predicting early risk of chronic kidney disease in cats using routine clinical laboratory tests and machine learning. *Journal of Veterinary Internal Medicine* 33(6), 2644-2656. <https://doi.org/10.1111/jvim.15623>.
- Chakrabarti, S., Syme, H. M., & Elliott, J. (2013). Clinicopathological variables predicting progression of azotemia in cats with chronic kidney disease. *Journal of Veterinary Internal Medicine* 26(2), 275-281. <https://doi.org/10.1111/j.1939-1676.2011.00874.x>.
- Elliott, J., & Barber, P. J. (1998). Feline chronic renal failure: Clinical findings in 80 cases diagnosed between 1992 and 1995. *Journal of Small Animal Practice* 39(2), 78-85. <https://doi.org/10.1111/j.1748-5827.1998.tb03598.x>.
- Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G. S., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine* 25(1), 24-29. <https://doi.org/10.1038/s41591-018-0316-z>.
- Han, J., Pei, J., & Tong, H. (2022). *Data mining: Concepts and techniques*. Waltham, Massachusetts, United States of America: Morgan Kaufmann.
- Hokamp, J. A., & Nabity, M. B. (2016). Renal biomarkers in domestic species. *Veterinary Clinical Pathology* 45(1), 28-56. <https://doi.org/10.1111/vcp.12333>.
- Lawson, J., Elliott, J., Wheeler-Jones, C., Syme, H., & Jepson, R. (2015). Renal fibrosis in feline chronic kidney disease: Known mediators and mechanisms of injury. *The Veterinary Journal* 203(1), 18-26. <https://doi.org/10.1016/j.tvjl.2014.10.009>.
- Lees, G. E., Brown, S. A., Elliott, J., Grauer, G. F., & Vaden, S. L. (2005). Assessment and management of proteinuria in dogs and cats: 2004 ACVIM Forum consensus statement. *Journal of Veterinary Internal Medicine* 19(3), 377-385. [https://doi.org/10.1892/0891-6640\(2005\)19\[377:AAMOPI\]2.0.CO;2](https://doi.org/10.1892/0891-6640(2005)19[377:AAMOPI]2.0.CO;2).
- John, L. Z. Q. (2010). The elements of statistical learning: data mining, inference, and prediction. *Journal of The Royal Statistical Society Series A: Statistics in Society* 173(3), 693-694. https://doi.org/10.1111/j.1467-985X.2010.00646_6.x.
- Sparkes, A. H., Caney, S., Chalhoub, S., Elliott, J., Finch, N., Gajanayake, I., Langston, C., Lefebvre, H. P., White, J., & Quimby, J. (2016). ISFM consensus guidelines on the diagnosis and management of feline chronic kidney disease. *Journal of Feline Medicine and Surgery* 18(3), 219-239. <https://doi.org/10.1177/1098612X16631234>.
- Twedt, D. C. (2006). Update on feline hepatobiliary disease. *Journal of Feline Medicine and Surgery* 8(5), xi-xiv. <https://doi.org/10.1016/S1098-612X-06-00125-2>.
- Vanden, B. E., Van Mulders, L., De Paepe, E., Paepe, D., Daminet, S., & Vanhaecke, L. (2025). Early detection of feline chronic kidney disease via 3-hydroxykynurenine and machine learning. *Scientific Reports* 15, 6875. <https://doi.org/10.1038/s41598-025-90019-x>.
- Vapnik, V. (2013). *The nature of statistical learning theory*. New York, USA: Springer Science & Business Media.