

Nghiên cứu năng lực sử dụng công cụ quản lý tài chính của sinh viên khối kinh tế Trường Đại học Nguyễn Tất Thành bằng học máy

Trịnh Khánh Linh

Khoa Tài chính - Kế toán, Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh, Việt Nam
linhtk@ntt.edu.vn

Tóm tắt

Sự phát triển của công nghệ tài chính tại Việt Nam làm gia tăng nhu cầu hiểu rõ hành vi tài chính cá nhân, đặc biệt ở sinh viên. Các nghiên cứu hiện có chủ yếu sử dụng mô hình hồi quy truyền thống, ít ứng dụng trí tuệ nhân tạo. Nghiên cứu ứng dụng mô hình học máy để mã hóa, huấn luyện và kiểm tra 805 mẫu khảo sát sinh viên khối ngành kinh tế, nhằm đánh giá hiệu suất của bảy thuật toán phân loại trong dự báo ảnh hưởng của các yếu tố nhân khẩu học, hiểu biết tài chính, kinh nghiệm làm thêm, ảnh hưởng gia đình và giáo dục tài chính nhà trường đến khả năng sử dụng công cụ quản lý tài chính. Kết quả cho thấy Logistic Regression, Support Vector Machine và AdaBoost đạt hiệu suất dự đoán cao, với độ chính xác 84,47 %. Phân tích cho thấy ảnh hưởng của gia đình, giáo dục tài chính nhà trường và hiểu biết tài chính là những yếu tố quan trọng nhất. Nghiên cứu cung cấp bằng chứng thực nghiệm về vai trò của trí tuệ nhân tạo, đồng thời đề xuất xây dựng công cụ dự báo và nhận diện hỗ trợ nhà quản lý giáo dục, cũng như cải tiến đào tạo nhằm nâng cao khả năng thích ứng với công nghệ tài chính của người học.

Nhận 17/07/2026
Được duyệt 03/08/2026
Công bố 28/08/2026

Từ khóa

Mô hình học máy; công nghệ tài chính; trí tuệ nhân tạo; kiến thức tài chính; quản lý giáo dục.

© 2026 Journal of Science and Technology - NTTU

1 Đặt vấn đề

Trong bối cảnh chuyển đổi số và sự phát triển nhanh chóng của công nghệ tài chính, các công cụ quản lý tài chính cá nhân như ứng dụng ngân hàng, ví điện tử, ứng dụng quản lý chi tiêu, nền tảng đầu tư trực tuyến và trợ lý tài chính thông minh đang trở thành một phần không thể thiếu trong hoạt động tài chính của người trẻ. Đối với sinh viên khối ngành kinh tế, khả năng sử dụng hiệu quả các công cụ này không chỉ phản ánh mức độ thích ứng với công nghệ tài chính mà còn thể hiện năng lực quản lý và ra quyết định tài chính trong tương lai. Vì vậy, việc nghiên cứu các yếu tố ảnh hưởng đến năng lực sử dụng công cụ quản lý tài chính của sinh viên có ý nghĩa quan trọng đối với các cơ sở giáo dục đại học

trong quá trình đào tạo nguồn nhân lực đáp ứng yêu cầu của nền kinh tế số [1].

Song song với sự phát triển của công nghệ tài chính, AI và học máy đang tạo ra những thay đổi đáng kể trong lĩnh vực phân tích dữ liệu giáo dục và tài chính. Không giống các phương pháp thống kê truyền thống chủ yếu tập trung vào kiểm định giả thuyết và phân tích mối quan hệ tuyến tính giữa các biến, các thuật toán học máy có khả năng học từ dữ liệu, phát hiện các mối quan hệ phi tuyến và thực hiện dự báo với độ chính xác cao hơn đối với các bài toán phân loại và nhận diện hành vi người dùng. Nhờ những ưu điểm này, học máy ngày càng được ứng dụng rộng rãi trong dự báo hành vi tài chính, đánh giá rủi ro tín dụng, phát

hiện gian lận tài chính và xây dựng các hệ thống hỗ trợ ra quyết định thông minh [2].

Các nghiên cứu cho thấy, phần lớn các công trình trong lĩnh vực giáo dục tài chính vẫn tập trung đo lường mức độ hiểu biết tài chính, hoặc phân tích các yếu tố ảnh hưởng đến hành vi quản lý tài chính bằng các mô hình hồi quy tuyến tính hoặc mô hình cấu trúc tuyến tính. Những phương pháp này phù hợp cho việc kiểm định lý thuyết nhưng còn hạn chế trong khai thác các tương tác phi tuyến giữa các biến cũng như chưa phát huy được khả năng dự báo của các thuật toán học máy. Bên cạnh đó, các nghiên cứu về năng lực sử dụng công cụ quản lý tài chính số của sinh viên còn tương đối, đặc biệt trong bối cảnh giáo dục đại học tại Việt Nam [3].

Bài báo này đề xuất một cách tiếp cận dựa trên học máy nhằm dự báo năng lực sử dụng công cụ quản lý tài chính của sinh viên khối ngành kinh tế tại Trường Đại học Nguyễn Tất Thành (NTTU). Nghiên cứu tiến hành tiền xử lý dữ liệu, huấn luyện và đánh giá bảy thuật toán phân loại nhằm lựa chọn mô hình dự báo có hiệu suất tối ưu. Đồng thời, nghiên cứu phân tích mức độ quan trọng của các biến đầu vào để nhận diện các yếu tố ảnh hưởng mạnh nhất đến năng lực này. Đóng góp của nghiên cứu được thể hiện trên ba phương diện: (i) nghiên cứu mở rộng hướng tiếp cận từ các mô hình giải thích truyền thống sang mô hình dự báo dựa trên học máy trong lĩnh vực giáo dục tài chính; (ii) thực hiện so sánh hiệu suất của bảy thuật toán học máy trên cùng một bộ dữ liệu thực nghiệm nhằm lựa chọn mô hình phù hợp cho bài toán dự báo hành vi người học và (iii) cung cấp cơ sở thực nghiệm hỗ trợ các cơ sở giáo dục xây dựng hệ thống nhận diện sớm sinh viên còn hạn chế năng lực này, qua đó góp phần thiết kế các chương trình giáo dục tài chính số phù hợp trong bối cảnh chuyển đổi số giáo dục.

2 Lý thuyết nghiên cứu

2.1 Tổng quan tài liệu

Hành vi tài chính cá nhân được hiểu là toàn bộ các quyết định liên quan đến việc quản lý, phân bổ và sử dụng các nguồn lực tài chính nhằm đạt được các mục tiêu kinh tế trong ngắn hạn và dài hạn. Theo quan

điểm của kinh tế học hành vi, các quyết định tài chính của cá nhân không hoàn toàn dựa trên tính duy lý mà chịu ảnh hưởng bởi các yếu tố tâm lý, nhận thức và môi trường xã hội. Giả thuyết vòng đời hành vi của Shefrin và Thaler cho rằng, con người có xu hướng phân chia tài sản thành các tài khoản tinh thần khác nhau, từ đó hình thành các quyết định chi tiêu, tiết kiệm và đầu tư không hoàn toàn tối ưu về mặt kinh tế [4]. Đối với sinh viên đại học, đây là giai đoạn bắt đầu hình thành các thói quen quản lý tài chính độc lập, do đó hành vi tài chính ở giai đoạn này có ảnh hưởng lâu dài đến năng lực tài chính trong tương lai.

Hiểu biết tài chính được xem là nền tảng quan trọng của năng lực tài chính cá nhân không chỉ bao gồm kiến thức về các khái niệm tài chính như lãi suất, lạm phát, đa dạng hóa rủi ro hay tín dụng mà còn phản ánh khả năng vận dụng kiến thức đó vào quá trình ra quyết định tài chính trong thực tiễn. Dưới góc độ kinh tế học, Lusardi và Mitchell xem hiểu biết tài chính như một dạng vốn nhân lực giúp cá nhân nâng cao chất lượng các quyết định tài chính, tích lũy tài sản và cải thiện phúc lợi kinh tế trong suốt vòng đời [1]. Vì vậy, mức độ hiểu biết tài chính cao thường gắn liền với khả năng quản lý ngân sách hiệu quả, sử dụng hợp lý các sản phẩm tài chính và thích ứng tốt hơn với môi trường tài chính số.

Bên cạnh hiểu biết tài chính, quá trình xã hội hóa tài chính thông qua gia đình và nhà trường cũng có ảnh hưởng quan trọng đến việc hình thành năng lực tài chính của sinh viên. Gia đình là môi trường đầu tiên truyền đạt các giá trị, chuẩn mực và kinh nghiệm quản lý tài chính thông qua quan sát, hướng dẫn và thực hành trong cuộc sống hằng ngày. Theo Gudmunson và Danes [5], quá trình xã hội hóa tài chính trong gia đình góp phần hình thành thái độ tài chính, khả năng lập kế hoạch và các hành vi tài chính tích cực của người trẻ. Trong khi đó, giáo dục tài chính tại nhà trường đóng vai trò bổ sung thông qua việc cung cấp kiến thức, kỹ năng và cơ hội thực hành, giúp người học nâng cao năng lực ra quyết định tài chính và khả năng thích ứng với các sản phẩm tài chính hiện đại. Các nghiên cứu tổng quan cho thấy giáo dục tài chính có tác động tích cực đến cả hiểu biết tài chính và hành

vi tài chính, đặc biệt khi được triển khai gắn với các tình huống thực tiễn [6].

Trong bối cảnh chuyển đổi số, năng lực sử dụng công cụ quản lý tài chính còn chịu ảnh hưởng bởi mức độ chấp nhận công nghệ của người dùng. Theo mô hình chấp nhận công nghệ, nhận thức về tính hữu ích và nhận thức về tính dễ sử dụng là hai yếu tố quyết định ý định và hành vi sử dụng công nghệ [7]. Đối với các công cụ tài chính số như ngân hàng số, ví điện tử, ứng dụng quản lý chi tiêu hay nền tảng đầu tư trực tuyến, việc sử dụng hiệu quả không chỉ đòi hỏi khả năng tiếp cận công nghệ mà còn yêu cầu người dùng có hiểu biết tài chính, kỹ năng đánh giá rủi ro và khả năng bảo vệ thông tin cá nhân. Vì vậy, năng lực sử dụng công cụ quản lý tài chính có thể được xem là kết quả tổng hợp của hiểu biết tài chính, quá trình xã hội hóa tài chính và mức độ chấp nhận công nghệ [8].

2.2 Khoảng trống nghiên cứu

Các nghiên cứu trước đây về giáo dục tài chính chủ yếu tập trung vào việc xác định các yếu tố ảnh hưởng đến hiểu biết tài chính, hành vi quản lý tài chính hoặc mức độ chấp nhận công nghệ tài chính thông qua các phương pháp phân tích truyền thống [9]. Bên cạnh đó, phần lớn các nghiên cứu hiện nay sử dụng hiểu biết tài chính, hành vi tài chính hoặc mức độ chấp nhận công nghệ làm biến phụ thuộc. Trong khi đó, năng lực sử dụng công cụ quản lý tài chính, bao gồm khả năng khai thác hiệu quả các ứng dụng quản lý chi tiêu, ngân hàng, ví điện tử và các nền tảng hỗ trợ ra quyết định tài chính vẫn chưa có nhiều nghiên cứu được xem xét như một cấu phần độc lập của năng lực tài chính số. Khoảng trống này đặc biệt rõ rệt trong bối cảnh giáo dục đại học tại Việt Nam, nơi quá trình chuyển đổi số đang diễn ra mạnh mẽ nhưng số lượng nghiên cứu ứng dụng AI trong giáo dục tài chính còn tương đối hạn chế [3].

Ở góc độ phương pháp, mặc dù nhiều nghiên cứu quốc tế đã chứng minh hiệu quả của các thuật toán học máy trong các lĩnh vực dự báo tín dụng, phát hiện gian lận, đánh giá rủi ro tài chính và phân tích hành vi người dùng, thì việc ứng dụng đồng thời nhiều thuật toán để dự báo năng lực sử dụng công cụ quản lý tài chính của sinh viên tại Việt Nam vẫn còn rất ít được

quan tâm. Đồng thời, các nghiên cứu hiện có cũng chưa làm rõ mức độ đóng góp của từng yếu tố đầu vào thông qua các kỹ thuật phân tích tầm quan trọng của biến, từ đó hạn chế khả năng hỗ trợ ra quyết định trong giáo dục [2, 10-14].

Xuất phát từ khoảng trống và yêu cầu nghiên cứu giải pháp nâng cao năng lực tài chính số cho nguồn nhân lực kinh tế trong thời đại chuyển đổi số, việc thực hiện nghiên cứu năng lực sử dụng công cụ quản lý tài chính của sinh viên khối ngành kinh tế tại NTTU bằng các phương pháp học máy có ý nghĩa cả về mặt lý luận và thực tiễn.

2.3 Các nghiên cứu liên quan đến học máy

Trong những năm gần đây, học máy đã trở thành một trong những hướng tiếp cận quan trọng trong lĩnh vực tài chính và giáo dục nhờ khả năng khai thác dữ liệu lớn, nhận diện các mẫu dữ liệu phức tạp và nâng cao độ chính xác dự báo so với các phương pháp thống kê truyền thống. Nhiều thuật toán phân loại như: Logistic Regression, Support Vector Machine (SVM), Random Forest, XGBoost và AdaBoost,... đã được ứng dụng thành công trong nhiều bài toán dự báo khác nhau, từ đánh giá rủi ro tín dụng đến phân tích hành vi người dùng và hỗ trợ ra quyết định tài chính [13].

Trong lĩnh vực tài chính, các nghiên cứu chủ yếu tập trung vào các bài toán đánh giá rủi ro và dự báo quyết định tài chính. Nuha và cộng sự [10] đã ứng dụng học máy để phát hiện gian lận trong hệ thống ngân hàng điện tử tại Bangladesh và chỉ ra rằng mức độ hiểu biết của người dùng có ảnh hưởng đáng kể đến nguy cơ trở thành nạn nhân của các hành vi lừa đảo trực tuyến. Tương tự, Haque và Hassan [12] đã chứng minh AdaBoost có khả năng nâng cao hiệu quả dự báo trong bài toán phê duyệt khoản vay ngân hàng, góp phần giảm thiểu rủi ro tín dụng và hỗ trợ tự động hóa quy trình ra quyết định.

Bên cạnh lĩnh vực tín dụng, nhiều nghiên cứu cũng cho thấy học máy có hiệu quả trong dự báo các hiện tượng kinh tế và hành vi người dùng. Sharma và cộng sự [11] so sánh nhiều thuật toán học máy trong dự báo giá nhà và kết luận rằng XGBoost đạt hiệu suất vượt trội khi xử lý các tập dữ liệu có đặc trưng phức tạp. Trong lĩnh vực viễn thông, Momin và cộng sự [14] áp

dụng học sâu để dự báo hành vi rời bỏ của khách hàng và ghi nhận khả năng nhận diện chính xác cao hơn so với các phương pháp phân loại truyền thống. Những kết quả này cho thấy học máy có khả năng xử lý hiệu quả các bài toán phân loại đa biến với dữ liệu có tính phi tuyến cao.

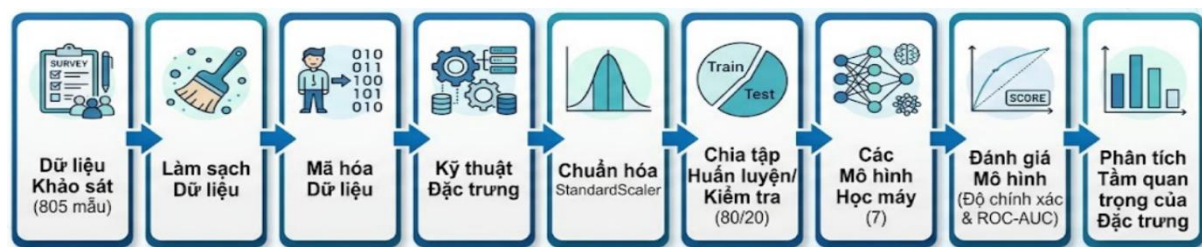
Đối với giáo dục tài chính, số lượng nghiên cứu ứng dụng học máy còn tương đối hạn chế. Levantesi và Zaccchia [2] sử dụng nhiều thuật toán học máy để xác định các yếu tố ảnh hưởng đến hiểu biết tài chính của người dân Ý và chỉ ra rằng độ tuổi, trình độ học vấn và hành vi tài chính là những biến dự báo quan trọng nhất. Kết quả nghiên cứu cũng khẳng định học máy có khả năng hỗ trợ các nhà hoạch định chính sách thiết kế các chương trình giáo dục tài chính hiệu quả hơn so với cách tiếp cận truyền thống. Tuy nhiên, nghiên cứu này chỉ tập trung vào việc dự báo hiểu biết tài chính mà chưa xem xét năng lực sử dụng các công cụ quản lý tài chính số của người học.

Tổng hợp các nghiên cứu trước cho thấy học máy đã được ứng dụng rộng rãi trong nhiều bài toán tài chính, song việc dự báo năng lực sử dụng công cụ quản lý tài chính của sinh viên trong môi trường giáo dục đại học vẫn còn là hướng nghiên cứu mới. Đồng thời, chưa có nhiều công trình thực hiện so sánh đồng thời nhiều thuật toán học máy trên cùng một bộ dữ liệu khảo sát nhằm lựa chọn mô hình tối ưu và phân tích mức độ quan trọng của các yếu tố ảnh hưởng. Vì vậy, nghiên cứu này kế thừa những thành tựu của các nghiên cứu trước đồng thời mở rộng ứng dụng học máy sang lĩnh vực giáo dục tài chính thông qua việc xây dựng mô hình dự báo và phân tích các yếu tố quyết định năng lực sử dụng công cụ quản lý tài chính của sinh viên.

3 Phương pháp nghiên cứu

3.1 Khung nghiên cứu

Nghiên cứu được triển khai theo quy trình của một bài toán phân loại nhị phân trong học máy, gồm các bước chính được mô tả trong Hình 1.



Hình 1 Quy trình nghiên cứu

3.2 Xây dựng bộ dữ liệu

Bộ dữ liệu được thu thập thông qua Google Forms từ sinh viên khối ngành kinh tế đang theo học tại NTTU. Sau quá trình kiểm tra và loại bỏ các bảng trả lời không hợp lệ, nghiên cứu thu được 805 quan sát hợp lệ phục vụ huấn luyện và đánh giá mô hình.

Bảng hỏi gồm hai nhóm thông tin. Nhóm thứ nhất bao gồm các đặc điểm cá nhân và học tập như giới tính, ngành học, khóa đào tạo, điểm trung bình tích lũy, tình trạng làm thêm và việc tham gia khóa học tài chính ngắn hạn. Nhóm thứ hai gồm các thang đo Likert năm mức về hiểu biết tài chính (FL), kinh nghiệm làm việc bán thời gian (PT), ảnh hưởng gia đình (FAM), giáo dục tài chính trường học (SCHOOL) và khả năng sử dụng công cụ quản lý tài chính (FT). Trong mô hình dự báo, FT được xác định là biến mục tiêu, trong khi các biến nhân khẩu học, học tập, FL,

PT, FAM và SCHOOL được sử dụng làm các biến dự báo.

3.3 Tiền xử lý và mã hóa dữ liệu

Trước khi huấn luyện mô hình, dữ liệu được kiểm tra nhằm phát hiện các giá trị thiếu, các phản hồi không đầy đủ và các quan sát không hợp lệ. Các biến định tính được chuyển đổi sang dạng số bằng phương pháp mã hóa nhãn. Chẳng hạn, biến giới tính được mã hóa nam = 0 và nữ = 1; các biến nhị phân khác được mã hóa theo nguyên tắc tương tự. Đối với các nhóm biến được đo lường bằng nhiều chỉ báo quan sát, điểm trung bình của các chỉ báo được sử dụng để hình thành đặc trưng đại diện cho từng khái niệm nghiên cứu. Biến mục tiêu FT được nhị phân hóa theo ngưỡng 3,5 dựa trên đặc điểm của thang đo Likert 5 điểm khi giá trị trung bình từ 3,5 trở lên phản ánh xu hướng đồng ý hoặc đánh giá tích cực của phản hồi,

phù hợp để phân biệt nhóm sinh viên có và chưa có năng lực này. Theo đó, $FT = 1$ biểu thị nhóm sinh viên có khả năng sử dụng công cụ quản lý tài chính,

trong khi $FT = 0$ biểu thị nhóm sinh viên chưa hoặc không có khả năng sử dụng.

Bảng 1 Mô tả các biến trong mô hình

TT	Biến	Loại dữ liệu	Mô tả	Giá trị đại diện
1	Giới tính	object	Giới tính của sinh viên	Nam, nữ
2	Ngành học	object	Ngành đào tạo của sinh viên	Tài chính ngân hàng, Quản trị kinh doanh,...
3	Khóa học	64-bit integer	Khóa đào tạo	2017, 2019,...
4	Điểm trung bình (GPA)	object	Học lực	Dưới 6,0; từ 6,0 đến 6.9,...
5	Làm thêm	object	Trải nghiệm công việc bán thời gian	Có, không
6	Khóa học tài chính ngắn hạn	object	Đã học hay chưa học	Đã từng, chưa từng
7	Kiến thức tài chính (FL)	64-bit integer	Thang 5 điểm, từ mức độ hoàn toàn không đồng ý đến hoàn toàn đồng ý	1, 2, 3, 4, 5
8	Kinh nghiệm làm việc bán thời gian (PT)	64-bit integer		1, 2, 3, 4, 5
9	Ảnh hưởng gia đình (FAM)	64-bit integer		1, 2, 3, 4, 5
10	Giáo dục tài chính trường học (School)	64-bit integer		1, 2, 3, 4, 5
11	Khả năng sử dụng công cụ quản lý tài chính (FT)	64-bit integer		1, 2, 3, 4, 5

Sau khi mã hóa, các biến đầu vào được chuẩn hóa bằng StandardScaler, qua đó đưa dữ liệu về trung bình bằng 0 và độ lệch chuẩn bằng 1. Bước chuẩn hóa giúp hạn chế ảnh hưởng của sự khác biệt về thang đo và đặc biệt cần thiết đối với các thuật toán dựa trên khoảng cách hoặc biên phân tách như KNN và SVM. Cuối cùng, bộ dữ liệu được chia ngẫu nhiên thành tập huấn luyện chiếm 80 % và tập kiểm định chiếm 20 %. Tập huấn luyện được sử dụng để xây dựng mô hình, trong khi tập kiểm định được dùng để đánh giá khả năng dự báo trên các quan sát chưa xuất hiện trong quá trình huấn luyện

3.4 Mô hình học máy

Để đánh giá khả năng dự báo của các thuật toán học máy, nghiên cứu tiến hành xây dựng bảy mô hình phân loại bao gồm Logistic Regression, SVM, KNN, Decision Tree, Random Forest, AdaBoost và XGBoost. Các mô hình này được lựa chọn vì đại diện cho các nhóm thuật toán phổ biến trong học máy, bao gồm mô hình tuyến tính, mô hình dựa trên khoảng cách, mô hình cây quyết định và mô hình tăng cường. Việc lựa chọn đồng thời nhiều thuật toán giúp đánh giá khách quan hiệu quả dự báo của từng mô hình

cũng như xác định phương pháp phù hợp nhất đối với bộ dữ liệu khảo sát của nghiên cứu.

3.5 Tiêu chí đánh giá

Hiệu suất của các mô hình học máy được đánh giá thông qua hai chỉ số chính là Accuracy và diện tích dưới đường cong ROC (ROC-AUC). Đây là các chỉ số được sử dụng phổ biến trong các bài toán phân loại nhị phân nhằm đánh giá khả năng dự báo và năng lực phân biệt của mô hình [15]. Accuracy phản ánh tỷ lệ dự đoán chính xác trên tổng số quan sát, được xác định theo công thức (1):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Trong đó, TP (True Positive) là số trường hợp dự đoán đúng sinh viên có khả năng sử dụng công cụ quản lý tài chính, TN (True Negative) là số trường hợp dự đoán đúng sinh viên không có khả năng sử dụng, FP (False Positive) là số trường hợp dự đoán nhầm có khả năng sử dụng, còn FN (False Negative) là số trường hợp dự đoán nhầm không có khả năng sử dụng.

Bên cạnh đó, nghiên cứu sử dụng ROC-AUC để đánh giá khả năng phân biệt của mô hình giữa hai nhóm sinh viên. Giá trị AUC càng tiến gần 1 cho thấy mô

hình có khả năng phân loại tốt, trong khi giá trị xấp xỉ 0,5 phản ánh hiệu quả dự báo không khác biệt đáng kể so với phân loại ngẫu nhiên. Việc kết hợp Accuracy

và ROC-AUC giúp đánh giá toàn diện hơn hiệu suất của các thuật toán học máy và hỗ trợ lựa chọn mô hình phù hợp cho bài toán nghiên cứu.

Bảng 2 Mô hình bài toán phân loại

Kết quả dự báo	Diễn giải	Tỷ lệ phân bố (%)
FT = 1	Sinh viên có khả năng sử dụng công cụ quản lý tài chính.	68,57
FT = 0	Sinh viên chưa hoặc không có khả năng sử dụng công cụ quản lý tài chính.	31,43

Sau khi nhị phân hóa biến mục tiêu, tỷ lệ phân bố hai lớp ở Bảng 2 cho thấy dữ liệu có mức độ mất cân bằng lớp ở mức trung bình nhưng không nghiêm trọng. Do đó, bên cạnh Accuracy, nghiên cứu sử dụng thêm chỉ số ROC-AUC để đánh giá khả năng phân biệt của mô hình, qua đó giảm thiểu nguy cơ đánh giá quá cao hiệu suất dự báo do sự chênh lệch về phân bố lớp.

4 Kết quả và thảo luận

Bảng 3 Kết quả độ chính xác của từng mô hình và giá trị ROC-AUC

TT	Mô hình	Độ chính xác (%)	Giá trị ROC-AUC
1	Logistic Regression	84,47	0,89
2	Support Vector Machine(SVM)	84,47	0,86
3	AdaBoost	84,47	0,92
4	K-Nearest Neighbors (KNN)	83,85	0,85
5	Random Forest	81,37	0,91
6	XGBoost	78,88	0,88
7	Decision Tree	78,26	0,77

Mặc dù dữ liệu có sự chênh lệch nhất định giữa hai lớp nhưng kết quả cho thấy, các mô hình có mức độ phức tạp vừa phải như Logistic Regression và SVM hoạt động hiệu quả hơn các mô hình dựa trên cây quyết định. Điều này cho thấy bộ dữ liệu nghiên cứu có khả năng phân tách tương đối rõ sau khi được chuẩn hóa, giúp các mô hình tuyến tính đạt hiệu suất dự báo cao mà không cần sử dụng các thuật toán phức tạp hơn. Đồng thời, Logistic Regression có ưu thế về khả năng diễn giải, phù hợp với các bài toán hỗ trợ ra quyết định trong giáo dục. Ngoài ra, chỉ số giá trị

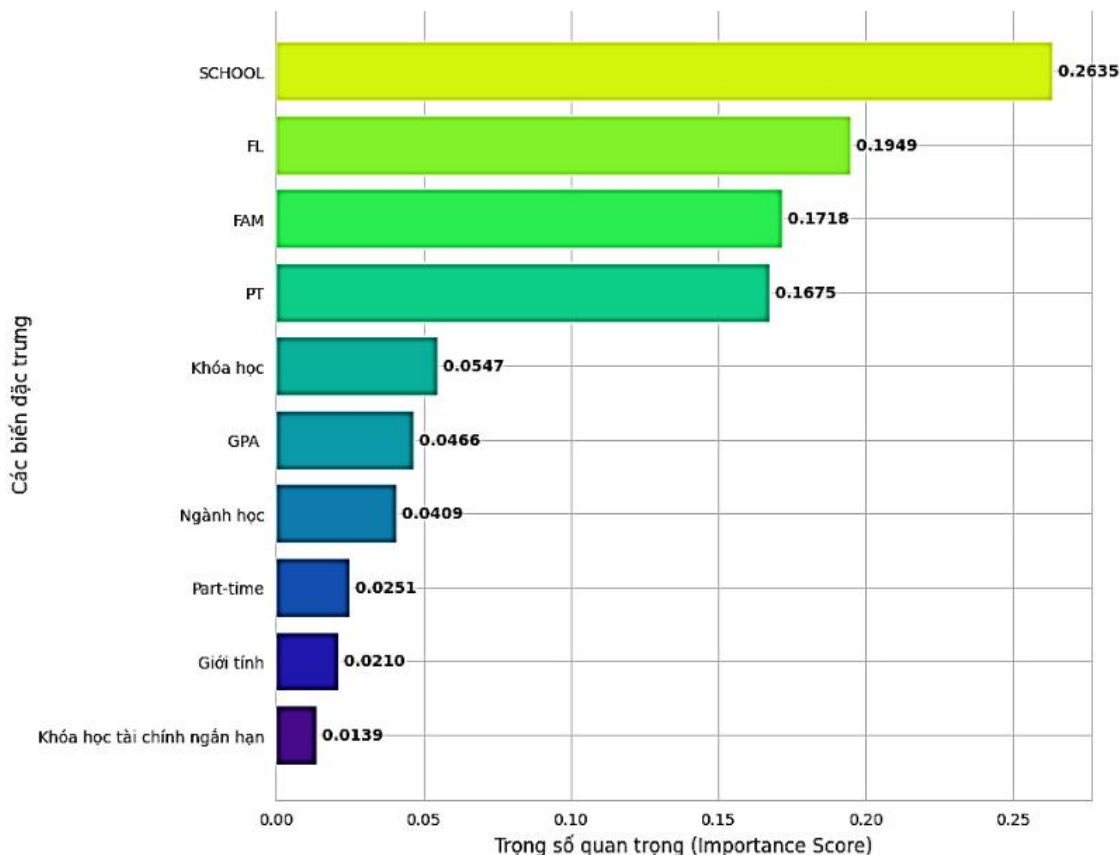
4.1 So sánh hiệu suất các mô hình học máy

Bảy thuật toán học máy được xây dựng và đánh giá trên cùng một bộ dữ liệu nhằm xác định mô hình phù hợp cho bài toán dự báo năng lực sử dụng công cụ quản lý tài chính của sinh viên. Kết quả trong Bảng 3 cho thấy, Logistic Regression, SVM và AdaBoost đều đạt độ chính xác cao nhất (84,47 %).

ROC-AUC duy trì trong khoảng 0,77-0,92 cho thấy khả năng phân biệt tốt giữa hai nhóm sinh viên.

4.2 Phân tích mức độ quan trọng của các biến

Để làm rõ vai trò của các yếu tố ảnh hưởng, nghiên cứu sử dụng Random Forest để phân tích mức độ quan trọng của các biến. Kết quả tại Hình 2 cho thấy, giáo dục tài chính nhà trường (SCHOOL) là yếu tố quan trọng nhất (26,35 %), tiếp theo là hiểu biết tài chính (FL) (19,49 %), ảnh hưởng gia đình (FAM) (17,18 %) và kinh nghiệm làm thêm (PT) (16,75 %). Trong khi đó, các đặc điểm nhân khẩu học như giới tính, GPA và khóa học có mức độ đóng góp thấp hơn.



Hình 2 Mức độ quan trọng của các biến quan sát

Như vậy, năng lực sử dụng công cụ quản lý tài chính chịu ảnh hưởng chủ yếu bởi các yếu tố có thể can thiệp thông qua giáo dục và trải nghiệm thực tiễn. Điều này phù hợp với các nghiên cứu trước về vai trò của giáo dục tài chính [6], hiểu biết tài chính [2] và xã hội hóa tài chính trong gia đình [5]. Đồng thời, kết quả cũng khẳng định rằng việc kết hợp kiến thức tài chính với trải nghiệm thực tế giúp sinh viên nâng cao khả năng khai thác hiệu quả các công cụ tài chính số.

4.3 Thảo luận và đề xuất

Nghiên cứu này so sánh đồng thời bảy thuật toán học máy trên cùng một bộ dữ liệu giúp xác định Logistic Regression, SVM và AdaBoost là các mô hình có hiệu suất dự báo tốt nhất, trong khi phân tích mức độ quan trọng của các biến giúp xác định các yếu tố ảnh hưởng đối với hành vi sử dụng công cụ quản lý tài chính. Đặc biệt, Logistic Regression và SVM đạt kết quả cao hơn Random Forest và XGBoost, cho thấy bộ dữ liệu sau khi chuẩn hóa và xây dựng các biến tổng hợp có xu hướng phân tách tương đối rõ, phù hợp với các mô hình tuyến tính. Bên cạnh đó, việc các mô

hình cây được huấn luyện với tham số mặc định cũng có thể làm hạn chế khả năng khai thác tối đa hiệu quả của thuật toán. Kết quả này cho thấy hiệu suất của mô hình không chỉ phụ thuộc vào mức độ phức tạp của thuật toán mà còn chịu ảnh hưởng bởi đặc điểm dữ liệu và quy trình xây dựng mô hình.

Trên cơ sở đó, nghiên cứu khẳng định tiềm năng ứng dụng học máy trong dự báo hành vi sử dụng công cụ quản lý tài chính của sinh viên, góp phần hỗ trợ các cơ sở giáo dục trong việc nhận diện nhóm sinh viên cần được tăng cường giáo dục tài chính. Đồng thời, nhà trường cần đẩy mạnh tích hợp nội dung quản lý tài chính cá nhân vào chương trình học theo hướng thực hành, kết hợp các hoạt động trải nghiệm và ứng dụng công nghệ tài chính nhằm nâng cao năng lực này và khả năng thích ứng với môi trường tài chính số của sinh viên.

5 Kết luận

Nghiên cứu này ứng dụng các thuật toán học máy nhằm dự báo năng lực sử dụng công cụ quản lý tài

chính của sinh viên khối ngành kinh tế NTTU. Bày mô hình phân loại đã được xây dựng và so sánh để xác định thuật toán phù hợp nhất cho bài toán dự báo. Kết quả thực nghiệm cho thấy Logistic Regression, SVM và AdaBoost đạt hiệu suất cao nhất với độ chính xác 84,47 %, đồng thời duy trì khả năng phân biệt tốt giữa hai nhóm thông qua giá trị ROC-AUC trong khoảng 0,77-0,92. Kết quả này cho thấy các phương pháp học máy có tiềm năng hỗ trợ hiệu quả cho các bài toán dự báo trong lĩnh vực giáo dục tài chính.

Phân tích mức độ quan trọng của các biến cho thấy giáo dục tài chính nhà trường, hiểu biết tài chính, ảnh hưởng của gia đình và kinh nghiệm làm thêm là những yếu tố dự báo ảnh hưởng nhất đối với năng lực sử dụng công cụ quản lý tài chính của sinh viên. Kết quả này không chỉ khẳng định vai trò của giáo dục tài chính trong bối cảnh chuyển đổi số mà còn cung cấp cơ sở thực nghiệm để nhà trường xây dựng các

chương trình đào tạo và hỗ trợ phù hợp nhằm nâng cao năng lực của người học theo hướng cá nhân hóa. Bên cạnh những kết quả đạt được, nghiên cứu vẫn còn hạn chế do dữ liệu thu thập tại một cơ sở đào tạo và được đánh giá trong cùng điều kiện thực nghiệm với các tham số mặc định của mô hình. Do đó, khả năng khái quát hóa và tính ổn định của kết quả vẫn cần được kiểm chứng trên các bộ dữ liệu đa dạng hơn. Trong nghiên cứu tiếp theo, cần mở rộng phạm vi khảo sát, áp dụng các kỹ thuật tối ưu hóa tham số, Stratified K-fold Cross-Validation và các phương pháp học sâu nhằm nâng cao độ chính xác, tính ổn định và khả năng khái quát của mô hình.

Lời cảm ơn

Chúng tôi xin cảm ơn Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh đã hỗ trợ cho nghiên cứu này.

Tài liệu tham khảo

1. Lusardi, A., & Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52(1), 5-44. <https://doi.org/10.1257/jel.52.1.5>.
2. Levantesi, S., & Zaccchia, G. (2021). Machine Learning and Financial Literacy: An Exploration of Factors Influencing Financial Knowledge in Italy. *Journal of Risk and Financial Management*. <https://doi.org/10.3390/jrfm14030120>.
3. Khúc Thế Anh., & Trần Tuấn Vinh. (2022). Impact of Financial Literacy on Vietnamese Students' Spending Management. *VNU Journal of Economics and Business*, 2(4). <https://js.vnu.edu.vn/EAB/article/view/4800>.
4. Shefrin, H. M., & Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic Inquiry*. <https://doi.org/10.1111/j.1465-7295.1988.tb01520.x>.
5. Gudmunson, C. G., & Danes, S. M. (2011). Family Financial Socialization: Theory and Critical Review. *Journal of Family and Economic Issues*, 32(4), 644-667. <https://doi.org/10.1007/s10834-011-9275-y>.
6. Kaiser, T., & Menkhoff, L. (2017). Does Financial Education Impact Financial Literacy and Financial Behavior, and If So, When? *The World Bank Economic Review*. <https://doi.org/10.1093/wber/lhx018>.
7. Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*. <https://doi.org/10.2307/249008>.
8. Đình Thế Hùng., Trần Nguyễn Nhật Nam., & Trần Vũ Hoàng Nam. (2025). The Impact of Digital Financial Literacy on Fintech Adoption among Students in Hanoi. *Asian Business Research Journal*, 10(6), 21-27. <https://doi.org/10.55220/25766759.471>.

9. Kurniawan, M. R. F., & Retnasih, N. R. (2024). The Impact of Financial Literacy and Fintech Adoption on Personal Financial Management Among University Students. *Proceeding ICAMEKA: International Conference Accounting, Management & Economics Uniska*, 1, 639-652. <https://doi.org/10.32503/icameka.v1i.53>.
10. Nuha, M., Mahmud, S., & Sattar, A. (2021). A Case Study and Fraud Rate Prediction in e-Banking Systems Using Machine Learning and Data Mining. In S. Borah, R. Pradhan, N. Dey, & P. Gupta (Eds.), *Soft Computing Techniques and Applications* (Vol. 1248, pp. 71-83). Springer Singapore. https://doi.org/10.1007/978-981-15-7394-1_6.
11. Sharma, H., Harsora, H., & Ogunleye, B. (2024). An Optimal House Price Prediction Algorithm: XGBoost. *Analytics*, 3(1), 30-45. <https://doi.org/10.3390/analytics3010003>.
12. Haque, F. M. A., & Hassan, M. M. (2024). *Bank Loan Prediction Using Machine Learning Techniques* (arXiv:2410.08886). arXiv. <https://doi.org/10.48550/arXiv.2410.08886>.
13. Bazarbash, M. (2019). FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk. IMF Working Papers, *IMF Working Papers*, Article 2019/109. <https://ideas.repec.org/p/imf/imfwpa/2019-109.html>.
14. Momin, S., Bohra, T., & Raut, P. (2020). Prediction of Customer Churn Using Machine Learning. In A. Haldorai, A. Ramu, S. Mohanram, & C. C. Onn (Eds.), *EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing* (pp. 203-212). Springer International Publishing. https://doi.org/10.1007/978-3-030-19562-5_20.
15. Powers, D. M. W. (2020). *Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation* (arXiv:2010.16061). arXiv. <https://doi.org/10.48550/arXiv.2010.16061>.

Research on the Ability to Use Financial Management Tools of Economics Students at Nguyen Tat Thanh University Using Machine Learning Approaches

Trinh Khanh Linh

Faculty of Finance and Accounting, Nguyen Tat Thanh University, Ho Chi Minh City, Viet Nam
linhtk@ntt.edu.vn

Abstract The development of financial technology in Viet Nam has increased the need to understand personal financial behavior, particularly among university students. Existing studies have mainly relied on traditional regression models, with limited application of artificial intelligence. This study applied machine learning models to encode, train, and test 805 survey responses from economics students to evaluate the performance of seven classification algorithms in predicting the effects of demographic factors, financial literacy, part-time work experience, family influence, and financial education at school on students' ability to use financial management tools. The results showed that Logistic Regression, Support Vector Machine, and AdaBoost achieved high predictive performance, with an accuracy of 84.47%. The analysis indicated that family influence, financial education at school, and financial literacy were the most important factors. The study provided empirical evidence of the role of artificial intelligence and proposed developing predictive and identification tools to support educational managers, while also providing a basis for improving training programs to enhance students' adaptability to financial technology.

Keywords Machine learning; fintech; artificial intelligence; financial literacy; educational management.

