

TRƯỜNG ĐẠI HỌC KINH TẾ - LUẬT
KHOA TÀI CHÍNH – NGÂN HÀNG



THẨM ĐỊNH RỦI RO TÍN DỤNG BẰNG MÔ HÌNH
HỒI QUY LOGISTIC VÀ DECISION TREE
Môn: Mô hình rủi ro tín dụng

Nhóm 5

Danh sách thành viên:

Vũ Ngọc Lâm

Nguyễn Trúc Linh

Nguyễn Thuy Hồng Thanh

Phạm Kim Hoàng

Đỗ Thị Lan Phương

Giảng viên:

Phạm Thị Thanh Xuân

Thành phố Hồ Chí Minh, 2022

1. Giới thiệu

Tín dụng là huyết mạch đối với sự phát triển kinh tế của một đất nước và là một trong những điều kiện quan trọng thúc đẩy phát triển kinh tế của một quốc gia nói chung và là nguồn tài chính quan trọng cho tiêu dùng, đầu tư sản xuất kinh doanh cho các cá nhân và hộ gia đình nói riêng. Khi dư âm của dịch COVID-19 vẫn diễn biến phức tạp, Việt Nam đã và đang phải thực hiện “mục tiêu kép” đó là vừa chống dịch, vừa phấn đấu hoàn thành các mục tiêu phát triển kinh tế - xã hội. Các ngân hàng đã nhanh chóng vào cuộc thực hiện hỗ trợ, tháo gỡ khó khăn cho các khách hàng bị ảnh hưởng bởi dịch Covid-19 bao gồm cơ cấu lại thời hạn trả nợ, miễn giảm lãi, phí, cho vay mới... theo Thông tư 01/2020/TT-NHNN và Thông tư 03/2021/TT-NHNN. Từ quý 2 năm 2021, các ngân hàng lần lượt đưa ra đa dạng các gói vay tiêu dùng với lãi suất thấp nhằm hỗ trợ phục hồi sản xuất kinh doanh cho các khách hàng cá nhân.

Có thể nói rằng, tín dụng nói chung và tín dụng cá nhân nói riêng là một trong các hoạt động trọng yếu nhất của ngân hàng thương mại vì hoạt động này mang lại tỷ trọng lớn về doanh thu, lợi nhuận. Vì vậy, ảnh hưởng của rủi ro này xảy ra sẽ mang lại hệ lụy dài lâu và rất nặng nề. Rủi ro tín dụng cá nhân nổi lên là một vấn đề đáng lo ngại mà nguyên nhân chính là từ những hạn chế trong khả năng nhận diện và đo lường khả năng trả nợ của khách hàng. Trên thực tế, hiện nay vẫn có rất nhiều khoản thẩm định vay được duyệt sơ sài và cảm tính dẫn đến những yếu tố phát sinh rủi ro như khách hàng không có khả năng trả được nợ, khách hàng có hành vi lừa đảo hay tài sản đảm bảo không có giá trị đáng kể. Vì vậy, nhằm nâng cao chất lượng tín dụng cá nhân, ta cần nghiên cứu nhiều về các rủi ro có mà tín dụng gây nên và một trong số đó là khả năng trả nợ đúng hạn của khách hàng.

Đã có nhiều các bài nghiên cứu về đo lường khả năng trả nợ vay của các khách hàng cá nhân tại ngân hàng nhưng chưa có nhiều bài kết hợp sử dụng mô hình “Cây quyết định” và phương pháp định lượng, định tính để phân tích. Nhờ tính ứng dụng của mô hình “Cây quyết định” vào lĩnh vực y học và tài chính cho độ chính xác cao, bài nghiên cứu này sẽ cũng sẽ sử dụng phương pháp ấy nhằm đánh giá khả năng trả nợ đúng hạn của các khách hàng cá nhân tại ngân hàng.

2. Khung lý thuyết

2.1 Khái niệm Khả năng trả nợ

Hiện tại, theo cơ sở pháp lý tại Việt Nam vẫn chưa có định nghĩa thống nhất về khái niệm “có khả năng trả nợ” mà vẫn chỉ là những dấu hiệu cho thấy khách hàng “không có khả năng trả nợ”, thông qua các phương pháp loại trừ, lược bỏ, nghiên cứu chúng ta có thể kết luận rằng ngoài khách hàng “không có khả năng trả nợ” là những khách hàng “Có khả năng trả nợ”

Theo Basel II, có hai trạng thái được dùng làm cơ sở để đánh giá khả năng trả nợ (Nguyễn Đăng Dòn, 2016):

- + *Khách hàng không có khả năng thanh toán đầy đủ khi đến hạn (chưa tính đến việc ngân hàng bán tài sản (nếu có) để hoàn trả)*
- + *Khách hàng có các khoản nợ xấu có thời gian quá hạn trên 90 ngày. Trong đó, những khoản thấu chi được xem là quá hạn khi khách hàng vượt hạn mức hoặc được thông báo một hạn mức nhỏ hơn dư nợ hiện tại.*

Bên cạnh đó, căn cứ theo định nghĩa của Quỹ tiền tệ quốc tế (IMF) thì: “*Nợ xấu là khoản nợ khi quá hạn trả lãi và/hoặc gốc trên 90 ngày; hoặc các khoản lãi chưa trả từ 90 ngày trở lên đã được nhập gốc, tái cấp vốn hoặc đồng ý chậm theo thỏa thuận, hoặc các khoản phải thanh toán đã quá hạn 90 ngày nhưng có lý do để chắc chắn để nghi ngờ về khả năng khoản vay sẽ không được thanh toán đầy đủ*” (Nguyễn Đăng Dòn, 2016).

Còn tại Việt Nam, nợ xấu (NPL) theo khoản 8 Điều 3 Chương I Thông tư số 11/2021/TT-NHNN là nợ nhóm 3, 4 và 5, trong đó có quy định tại Điều 11 Mục 1 Chương II. Có một điều rõ ràng là khả năng trả nợ của khách hàng là việc khách hàng có khả năng trả nợ đầy đủ và đúng hạn cho người cho vay hay không (Ngân hàng nhà nước, 2013) Trong các nợ nhóm được định nghĩa như sau:

- + ***Nhóm nợ 3 (nợ dưới tiêu chuẩn):*** Các khoản nợ được tổ chức tín dụng (TCTD), chi nhánh ngân hàng đánh giá là không có khả năng thu hồi nợ gốc và lãi khi đến hạn. Các khoản nợ này được TCTD, chi nhánh ngân hàng nước ngoài đánh giá là có khả năng tổn thất.
- + ***Nhóm nợ 4 (Nợ nghi ngờ):*** Các khoản nợ được TCTD, chi nhánh ngân hàng đánh giá là có khả năng tổn thất cao.
- + ***Nhóm nợ 5 (Nợ có khả năng mất vốn):*** Các khoản nợ được TCTD, chi nhánh ngân hàng đánh giá là không còn khả năng thu hồi, mất vốn.

Cũng theo Thông tư số 11/2021/TT-NHNN, nợ nhóm 2 (nợ cần chú ý) là các khoản nợ được TCTD, chi nhánh ngân hàng đánh giá là có khả năng thu hồi đầy đủ cả nợ gốc và lãi nhưng có dấu hiệu khách hàng suy giảm khả năng trả nợ. Như vậy nếu khách hàng phát sinh nợ nhóm 2 vẫn được hiểu là khách hàng vẫn còn khả năng trả nợ, dù khả năng trả nợ bị suy yếu trước mắt.

Tuy nhiên, Bài nghiên cứu chỉ thiên về hướng đánh giá các nhân tố có thể ảnh hưởng tới nợ tín dụng của khách hàng nên để bài nghiên cứu được tron tru hơn thì bài nghiên cứu chỉ dùng mức độ phân loại là có khả năng nợ xấu và không có khả năng nợ xấu

2.2 Lược khảo các mô hình thẩm định rủi ro tín dụng

Sự chính xác khi thẩm định khả năng trả nợ đúng hạn của khách hàng để ra quyết định cho vay sẽ góp phần nâng cao hiệu quả kinh doanh cũng như giảm thiểu rủi ro phi hệ thống cho các ngân hàng hay tổ chức tài chính. Để quản lý tốt hoạt động tín dụng, việc hiểu được những nhân tố nào có thể ảnh hưởng và ảnh hưởng như thế nào đến khả năng trả nợ của từng khách hàng đã trở thành thử thách đòi hỏi sự nghiên cứu và quan tâm sâu sắc của các ngân hàng và tổ chức thuộc lĩnh vực tài chính tín dụng.

Phan Thị Hằng Nga & Nguyễn Thị Ngọc Quỳnh (2020) khi thu thập dữ liệu và đánh giá khả năng trả nợ của các khách hàng cá nhân của Ngân hàng Thương mại cổ phần Sài Gòn Thương Tín (Sacombank) - Chi nhánh Tân Bình, nhận thấy có 5 nhân tố gồm: Thu nhập; Thời gian làm công việc hiện tại; Lãi suất; Quy mô khoản vay và Lịch sử tín dụng là có tác động đến tín nhiệm trả nợ của người vay. Kết quả của tác giả kết luận được sự tồn tại mối quan hệ đồng biến ở hai biến thời gian làm công việc hiện tại và thu nhập, số còn lại là nghịch biến. Trần Thanh Phong và cộng sự (2020) đã thực hiện nghiên cứu về khả năng trả nợ của khách hàng cá nhân đã phát sinh quan hệ tín dụng liên tục từ 3 năm trở lên ở Ngân hàng Nông nghiệp và phát triển (Agribank) Tân Hưng, tỉnh Long An. Kết quả cho thấy có 6 trên 8 biến sử dụng là có ý nghĩa thống kê đến khả năng trả nợ đúng hạn bao gồm: Nhà ở; Mục đích khoản vay; Số người phụ thuộc; Thu nhập; Tuổi và cuối cùng là Quy mô khoản vay. Trong đó, tác giả cho rằng việc sở hữu một tài sản đảm bảo như nhà ở được xem là yếu tố quan trọng nhất để nâng cao khả năng trả nợ của khách hàng. Đáng chú ý, kết quả về sự phát triển đồng biến của biến “Quy mô khoản vay” ở nghiên cứu này là khác biệt so với nghiên cứu đã lược khảo trước đó và trái ngược so với suy nghĩ thông thường rằng vay càng nhiều sẽ càng khó trả nợ. Bên cạnh các biến thông dụng, còn có nhiều yếu tố khác có khả năng ảnh hưởng đến việc trả nợ của các cá nhân đã được chỉ ra như: Độ tuổi; Giới tính; Tình trạng hôn nhân. Trình độ học vấn; Chức vụ công tác, Thu nhập hàng tháng của khách hàng; Số sản phẩm khách hàng nắm giữ; Thời hạn cho vay; Thời gian còn lại khoản vay; Lãi suất cho vay và dư nợ gốc còn lại của khoản vay(Nguyễn Hải Trường, 2017)

Từ những bài nghiên cứu đã lược khảo, tác giả lựa chọn các biến được cho là đáng chú ý và có khả năng ảnh hưởng đến khả năng trả nợ đúng hạn của khách hàng cá nhân bao gồm: Thu nhập; Số tiền vay; Quan hệ tín dụng; Số năm làm việc; Tài sản đảm bảo và Trình độ học vấn. Sáu thuộc tính này sẽ là cơ sở dữ liệu để tác giả đánh giá, thảo luận và xây dựng nên mô hình dự đoán khả năng trả nợ của khách hàng, cụ thể là thuật toán Decision Tree ở bài nghiên cứu này.

3. Phương pháp, dữ liệu

3.1 Dữ liệu

Dữ liệu phục vụ mục đích học tập và nghiên cứu nhằm hiểu rõ bản chất và cách thức hoạt động của mô hình cây quyết định. Do đó, dữ liệu được khởi tạo có chủ đích đơn giản thông qua khảo sát. Nghiên cứu bao gồm 118 khách hàng cá nhân, cấu trúc mẫu đạt tỷ lệ cân bằng cao giữa hai nhóm bao gồm 1-Khách hàng trả được nợ và 0-Khách hàng không trả được nợ. Tỷ lệ giữa hai nhóm 1-0 là 65-53.

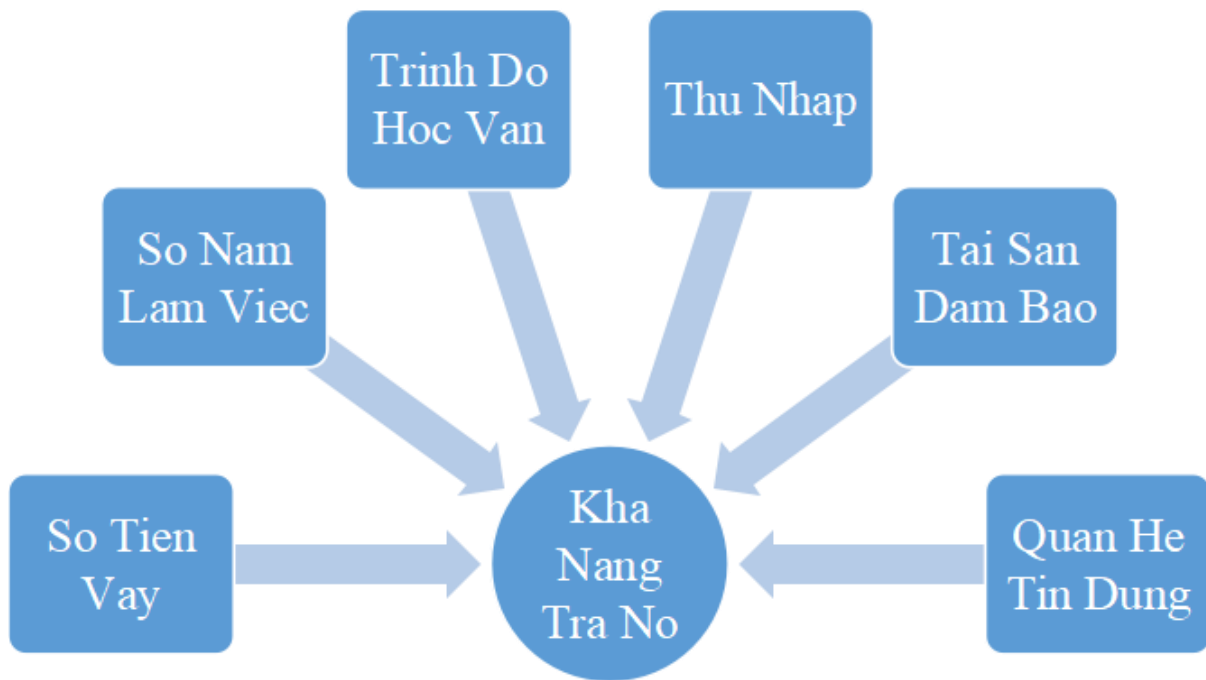
Bảng x. Mô tả biến sử dụng

Tên biến	Thang đo
Biến dự báo	
Khả năng trả nợ	1-Có khả năng trả nợ
	0-Không có khả năng trả nợ
Biến chứa thông tin dự báo	
Thu nhập	1-Thấp
	2-Trung bình
	3-Cao
	4-Rất cao
Số tiền vay	1-Nhỏ hơn khả năng trả nợ
	2-Lớn hơn khả năng trả nợ
Quan hệ tín dụng	1-Lần đầu
	2-Nhiều lần

Số năm làm việc	1-0-5 năm
	2-5-10 năm
	3-10-15 năm
	4-trên 15 năm
Tài sản đảm bảo	0-Không thuộc sở hữu người vay
	1-Thuộc sở hữu người vay
Trình độ học vấn	1-Phổ thông
	2-Đại học
	3-Sau đại học

3.2 Các phương pháp sàng lọc biến

Dựa vào khung lý thuyết về lý thuyết khả năng trả nợ cũng như các nghiên cứu thực nghiệm về các thuộc tính có tác động, ảnh hưởng trực tiếp đến khả năng trả nợ vay ở Phần 2, nhóm tác giả đề xuất mô hình nghiên cứu như hình sau:



Các biến ảnh hưởng trên đều được lựa chọn dựa trên mức độ ảnh hưởng, tầm quan trọng mà nhóm đo lường bằng các phương pháp máy học (ở phần sau sẽ đề cập) qua đó các biến được chọn đều là các biến thực sự đã có tác động đến khả năng trả nợ

3.3 Decision Tree (Cây quyết định)

3.3.1 Giới thiệu Decision Tree

Decision Tree là một thuật toán thuộc nhóm học có giám sát (supervised learning) giúp hỗ trợ việc đưa ra quyết định dựa trên mô hình cây quyết định và hệ quả. Khác với các thuật toán cùng nhóm, cây quyết định có thể được sử dụng để giải các bài toán hồi quy (regression) và phân loại (classification). Nhưng chủ yếu nó được áp dụng để giải quyết các bài toán phân loại. Dựa vào biến phụ thuộc, có thể chia mô hình Decision Tree thành hai loại:

- Cây phân loại: sử dụng các biến phụ thuộc là biến phân loại được phân tách thành một loạt các danh mục. Ví dụ, một cây phân loại có thể gán giá trị có là 1 và không có giá trị là 0. Sau đó, nó sẽ xác định chuỗi các nhánh và nút đại diện cho các lựa chọn và các yếu tố có thể theo sau từ quyết định có hoặc không
- Cây hồi quy: sử dụng các biến phụ thuộc là biến liên tục để dự đoán giá trị đầu ra dữ liệu. Ví dụ, có thể dự đoán giá của một chiếc ô tô dựa trên các yếu tố như giá hiện tại, hãng sản xuất, dịch vụ sửa chữa, mức độ tiêu hao nhiên liệu... Giá

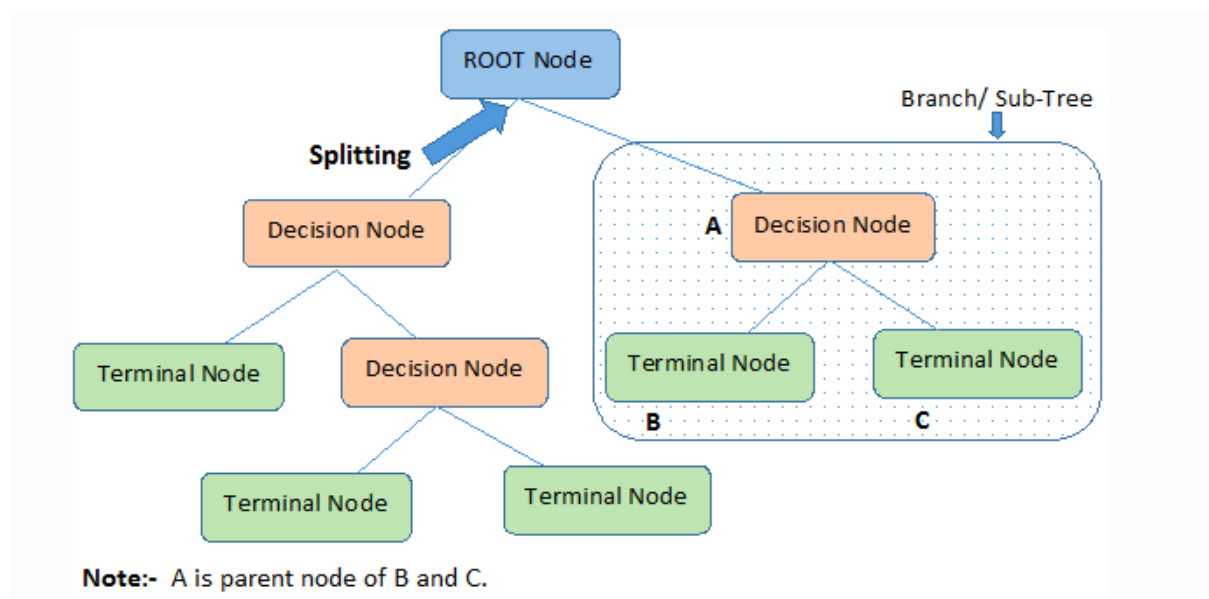
của ô tô là một biến số liên tục vì có thể liên tục tính toán lại nó dựa trên dữ liệu hoặc xu hướng mới có sẵn.

Mục đích sử dụng: tạo ra một mô hình huấn luyện dữ liệu để dự đoán giá trị của một biến mục tiêu bằng cách học các quy tắc quyết định đơn giản được suy ra từ dữ liệu trước đó (training data).

Một số giả định khi xây dựng mô hình:

- Ban đầu, toàn bộ tập huấn luyện được xem là gốc (root)
- Giá trị đối tượng được ưu tiên phân loại.
- Các quan sát được phân phối một cách đệ quy trên cơ sở các giá trị thuộc tính.
- Thứ tự đặt các thuộc tính làm nút gốc hoặc nút bên trong của cây được thực hiện bằng cách sử dụng một số phương pháp thống kê.

3.3.2 Cấu trúc Decision Tree



Hình . Cấu trúc cơ bản của mô hình Decision Tree

Decision Tree là một cấu trúc phân cấp của các nút và các nhánh. Bao gồm các thành phần chính sau:

- Nút gốc (Root node): đại diện cho toàn bộ tập hợp hoặc mẫu và tiếp tục được chia thành hai hoặc nhiều tập hợp đồng nhất.
- Nút quyết định (Decision tree): khi một nút con tách thành các nút con nhỏ hơn, thì nó được gọi là nút quyết định.
- Nút lá/nút cuối (Leaf/Terminal node): các nút không phân chia hay chia tách thì được gọi là nút lá hoặc nút cuối.

Trong đó:

- Chia tách (Split): là một quá trình phân chia một nút thành hai hoặc nhiều nút con.

- Nhánh/Cây phụ (Branch/ Sub-tree): một phần phụ của toàn bộ cây được gọi là nhánh hoặc cây phụ.

3.3.3 Xây dựng Decision Tree

Trong Decision Tree, để dự đoán nhãn lớp cho một quan sát, chúng ta bắt đầu từ gốc của cây rồi so sánh các giá trị của thuộc tính gốc với thuộc tính của quan sát. Trên cơ sở so sánh, ta theo nhánh tương ứng với giá trị đó và nhảy đến nút tiếp theo. Đối với nút tiếp theo, thuật toán lại so sánh giá trị thuộc tính với các nút con khác và di chuyển xa hơn. Nó tiếp tục quá trình cho đến khi nó đạt đến nút lá của cây. Quy trình hoàn chỉnh có thể được hiểu rõ hơn bằng cách sử dụng thuật toán dưới đây:

Bước 1: Bắt đầu cây với nút gốc (gọi là S), nút này chứa tập dữ liệu hoàn chỉnh.

Bước 2: Tìm thuộc tính tốt nhất trong tập dữ liệu bằng cách sử dụng kỹ thuật Attribute selection measure (Phép đo lựa chọn thuộc tính).

Bước 3: Chia S thành các tập con chứa các giá trị có thể có cho các thuộc tính tốt nhất.

Bước 4: Tạo nút cây quyết định chứa thuộc tính tốt nhất.

Bước 5: Tạo một cách đệ quy cây quyết định mới bằng cách sử dụng các tập con của tập dữ liệu đã tạo ở bước 3. Tiếp tục quá trình này cho đến khi đạt đến một giai đoạn mà không thể phân loại thêm các nút và được gọi là nút cuối cùng (nút lá).

Điều kiện dừng của mô hình:

- Khi tất cả các quan sát đã thuộc về một lớp (nút lá).
- Khi không còn thuộc tính nào có thể dùng để chia tách quan sát nữa.
- Khi không còn lại quan sát nào tại nút.

3.3.4 Ưu điểm và hạn chế

Decision Tree được sử dụng khá phổ biến với một vài ưu điểm như sau:

- Đơn giản và dễ hiểu bởi trực quan hình ảnh rõ ràng, dễ hình dung và diễn giải. Đầu ra của cây có thể được giải thích dễ dàng.
- Có thể làm việc với biến liên tục và biến phân loại nên được sử dụng cho cả hai bài toán hồi quy và phân loại.
- Có thể tự động xử lý các giá trị bị thiếu (missing value), đặc biệt không cần phải chuẩn hóa dữ liệu đầu vào nên việc xử lý và làm sạch dữ liệu đầu vào đơn giản hơn các thuật toán khác.
- Thời gian huấn luyện mô hình ít hơn so với các mô hình khác như Random Forest,...

Tuy nhiên, Decision Tree vẫn có một vài hạn chế nhất định:

- Hạn chế lớn nhất là vấn đề overfitting. Để phù hợp với dữ liệu, mô hình học quá kỹ khi tiếp tục tạo ra các nút mới và cuối cùng cây trở nên quá phức tạp để diễn giải dẫn đến việc mất khả năng tổng quát hóa. Cho nên, mô hình không cho kết quả tốt ở tập kiểm nghiệm trong khi tập huấn luyện lại làm việc rất tốt.

- Tính ổn định mô hình không cao. Khi thêm một dữ liệu mới hay thay đổi nhỏ trong tập dữ liệu có thể làm cho cấu trúc của cây quyết định thay đổi hoàn toàn dẫn đến việc phải tạo lại cây và tất cả các nút.
- Không phù hợp với tập dữ liệu lớn. Nếu kích thước quá lớn thì cây đơn lẻ có thể được tạo nên phức tạp dẫn đến việc bị overfitting.
- Bị ảnh hưởng bởi các dữ liệu nhiễu (noise). Chỉ cần một vài dữ liệu nhiễu sẽ dẫn đến sai lệch trong dự đoán.

3.3.5 Ứng dụng mô hình Decision Tree

Áp dụng trong thị trường chứng khoán: Decision Tree được sử dụng để phân loại hiệu suất thể hiện của các cổ phiếu trên thị trường. Các cổ phiếu được phân loại đầu ra bằng ba mức độ “tốt”, “bình thường” và “kém”. Việc phân tích được hiệu suất thể hiện của cổ phiếu sẽ giúp cho nhà đầu tư có thêm nhiều lựa chọn cho quyết định của mình. Ngoài ra, với dữ liệu đầu vào là các chỉ số tài chính, nhà đầu tư có thể sử dụng cây quyết định để dự đoán xem cổ phiếu nào đang trong tình trạng khủng hoảng tài chính.

Nghiên cứu hoạt động trong việc lập kế hoạch hậu cần và quản trị chiến lược: Decision Tree có thể giúp xác định các chiến lược phù hợp sẽ giúp công ty đạt được các mục tiêu đã đề ra. Các lĩnh vực mà cây quyết định có thể được áp dụng bao gồm kỹ thuật, giáo dục, luật, kinh doanh, chăm sóc sức khỏe và tài chính.

3.4 Hồi quy logistic

3.4.1 Giới thiệu mô hình hồi quy Logistic

Mô hình hồi quy Logistic dự đoán một biến phụ thuộc bằng cách phân tích mối quan hệ giữa một hoặc nhiều biến độc lập hiện có và được sử dụng chủ yếu trong các bài toán phân loại. Hồi quy Logistic là một phương pháp phân tích thống kê để dự đoán xác suất xảy ra cho một quan sát bằng cách điều chỉnh dữ liệu bằng đường cong Logistic. Từ đó, mô hình cho ra kết quả nhị phân dựa trên xác suất của mỗi lớp. Với ngưỡng (threshold) là 0.5 thì các trường hợp có xác suất cao hơn ngưỡng này sẽ được phân loại là 1, trong khi đó các trường hợp thấp hơn sẽ được phân loại là 0. Ví dụ, một hồi quy Logistic có thể được sử dụng để dự đoán liệu một học sinh trung học có được nhận vào một trường đại học cụ thể hay không. Các kết quả nhị phân này cho phép các quyết định đơn giản giữa hai lựa chọn thay thế.

Mục đích sử dụng: được sử dụng để dự đoán xác suất một quan sát rơi vào một trong các loại của biến phụ thuộc dựa trên một hoặc nhiều biến độc lập có thể liên tục hoặc phân loại.

Một số giả định của mô hình:

- Giữa các biến độc lập không có hoặc có rất ít hiện tượng đa cộng tuyến
- Biến phụ thuộc là biến rời rạc
- Các biến độc lập phải có liên quan tuyến tính với tỷ lệ cược log
- Kích thước mẫu của mô hình được yêu cầu khá lớn

3.4.2 Phương trình của mô hình

Hồi quy Logistic sử dụng một phương trình biểu diễn rất giống phương trình hồi quy tuyến tính. Trong phương trình, các giá trị đầu vào được kết hợp bằng cách sử dụng trọng số hoặc giá trị hệ số để dự đoán giá trị đầu ra. Một điểm khác biệt chính so với hồi quy tuyến tính là giá trị đầu ra được mô hình hóa là giá trị nhị phân (0 hoặc 1) chứ không phải là giá trị số. Dưới đây là phương trình hồi quy Logistic:

$$\ln\left(\frac{p}{p-1}\right) = b_0 + b_1x$$

Từ phương trình này, xác suất của đầu ra được tính theo công thức:

$$p = \frac{1}{1 + e^{-y}}$$

Với $y = b_0 + b_1x$

Trong đó:

- p là xác suất xảy ra có giá trị trong khoảng từ 0 đến 1
- b_0 là hệ số chặn
- b_1 là hệ số góc của biến đầu vào x
- x là biến độc lập

3.4.3 Ưu điểm và hạn chế

Mô hình hồi quy Logistic được sử dụng phổ biến cho các bài toán phân loại bởi các ưu điểm vượt trội như sau:

- Dễ dàng xây dựng và huấn luyện dữ liệu hơn nhiều so với các mô hình học máy khác
- Là một trong những mô hình hoạt động tốt đối với trường hợp tập dữ liệu có thể phân tách tuyến tính
- Có thể giúp đo lường mức độ liên quan giữa các biến và tác động của chúng đối với kết quả đầu ra. Điều này có thể nhanh chóng xác định khi nào hai biến số có tương quan tích cực hoặc tiêu cực

Tuy nhiên, mô hình hồi quy Logistic cũng có một số hạn chế nhất định:

- Mô hình không thể dự đoán được kết quả đầu ra là biến liên tục
- Giả định tính tuyến tính giữa biến phụ thuộc và biến độc lập
- Nếu kích thước mẫu quá nhỏ, mô hình có thể không chính xác

3.4.4 Ứng dụng mô hình hồi quy Logistic

Chấm điểm tín dụng (credit scoring): Mục tiêu của các mô hình chấm điểm tín dụng là dự đoán mức độ tín nhiệm của khách hàng và xác định xem họ có thể đáp ứng một nghĩa vụ tài chính nhất định hay không. Mô hình hồi quy Logistic được sử dụng để dự đoán bằng cách sử dụng các yếu tố dự báo liên quan đến tình trạng cá nhân và lịch sử tài chính của họ. Từ đó, thiết lập mô hình ảnh hưởng của các biến khác nhau

đến mức độ tín nhiệm của người tiêu dùng. Đầu ra của mô hình là đưa ra quyết định có tiếp tục cấp tín dụng cho các khách hàng trong tương lai hay không.

Đặt phòng khách sạn (hotel booking): Mô hình hồi quy Logistic được sử dụng để dự đoán bằng cách đánh giá mối quan hệ giữa các biến độc lập như ngày đến, ngày đi, số lượng người lớn, số lượng trẻ em, phục vụ bữa ăn, loại phòng... và biến phụ thuộc là khách hàng có hủy đặt phòng hay không. Việc dự đoán xem khách đặt phòng khách sạn có hủy hay không là rất quan trọng trong việc thực hiện quản lý doanh thu và nguồn lực cho khách sạn.

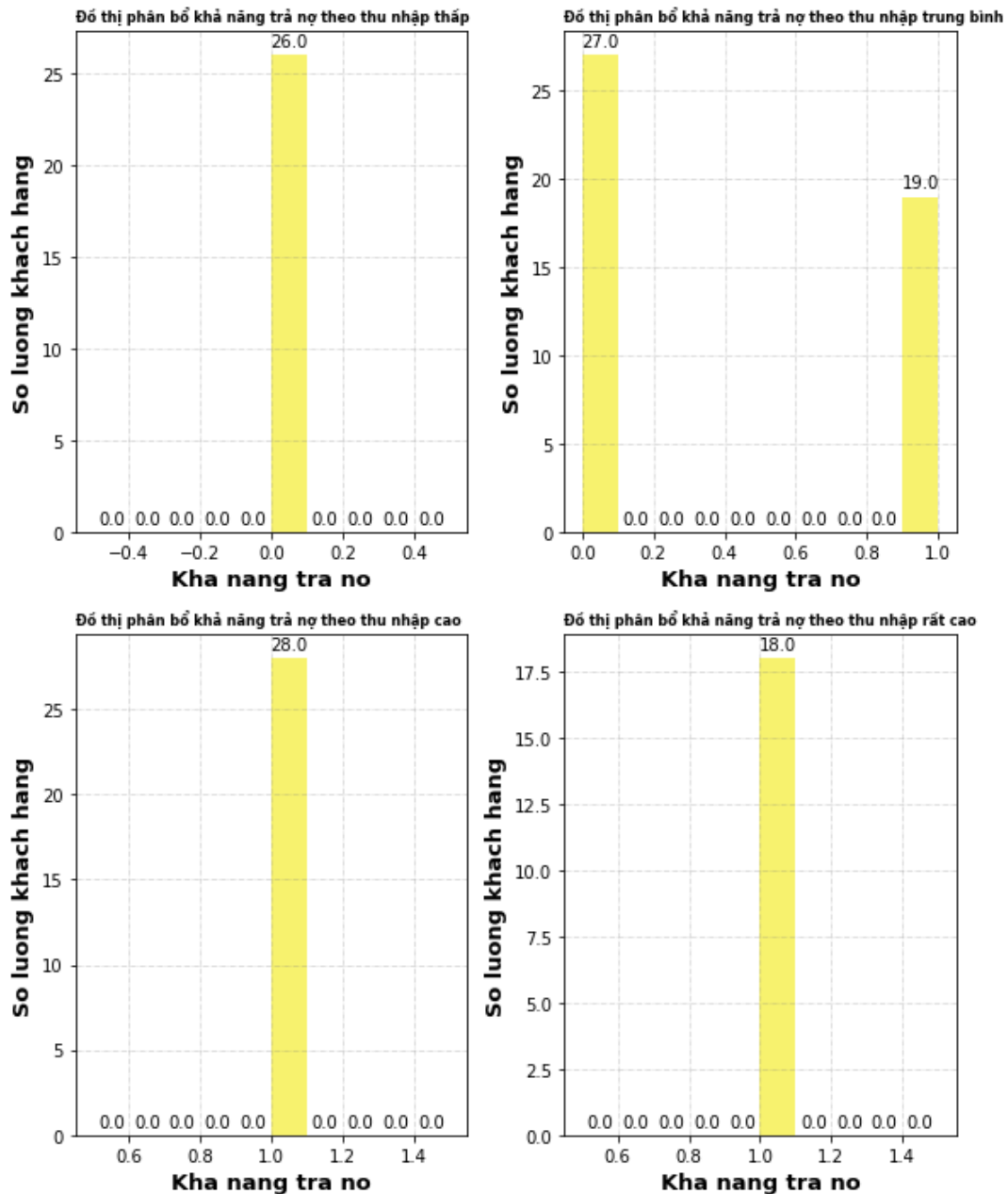
4. Kết quả

4.1 Thống kê mô tả

	Kha nang tra no	Thu nhap	So tien vay	Quan he tin dung	So nam lam viec	Tai san dam bao	Trinh do hoc van
count	118	118	118	118	118	118	118
unique	2	4	2	2	4	2	3
top	1	2	2	1	2	1	2
freq	65	46	62	63	48	60	60

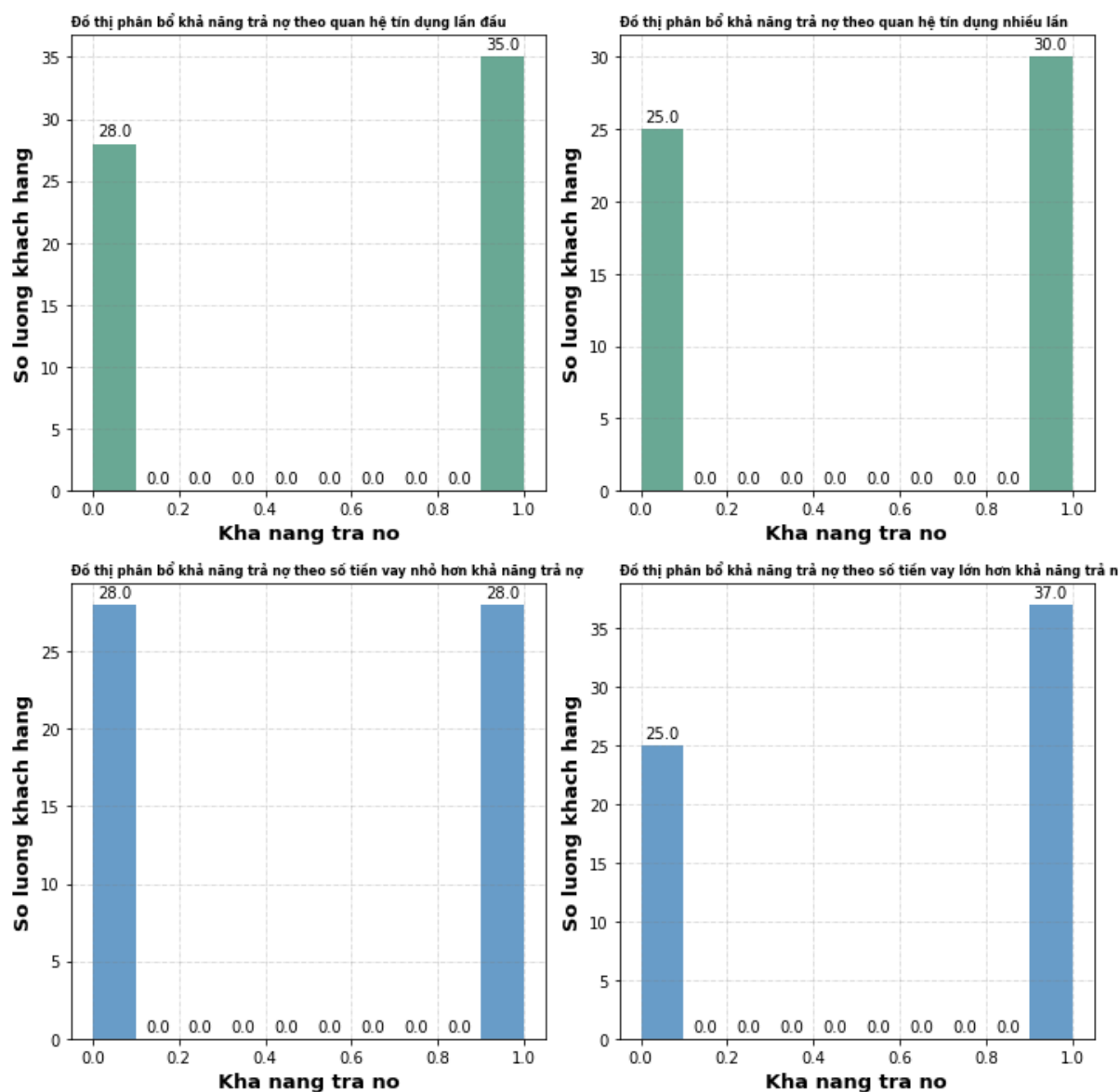
Hình 1: Thống kê mô tả cơ bản chỉ số các biến

Dựa vào hình trên, ta có thể kết luận rằng tất cả các biến có trong bài nghiên cứu này toàn bộ đều là biến category và không có biến numeric.



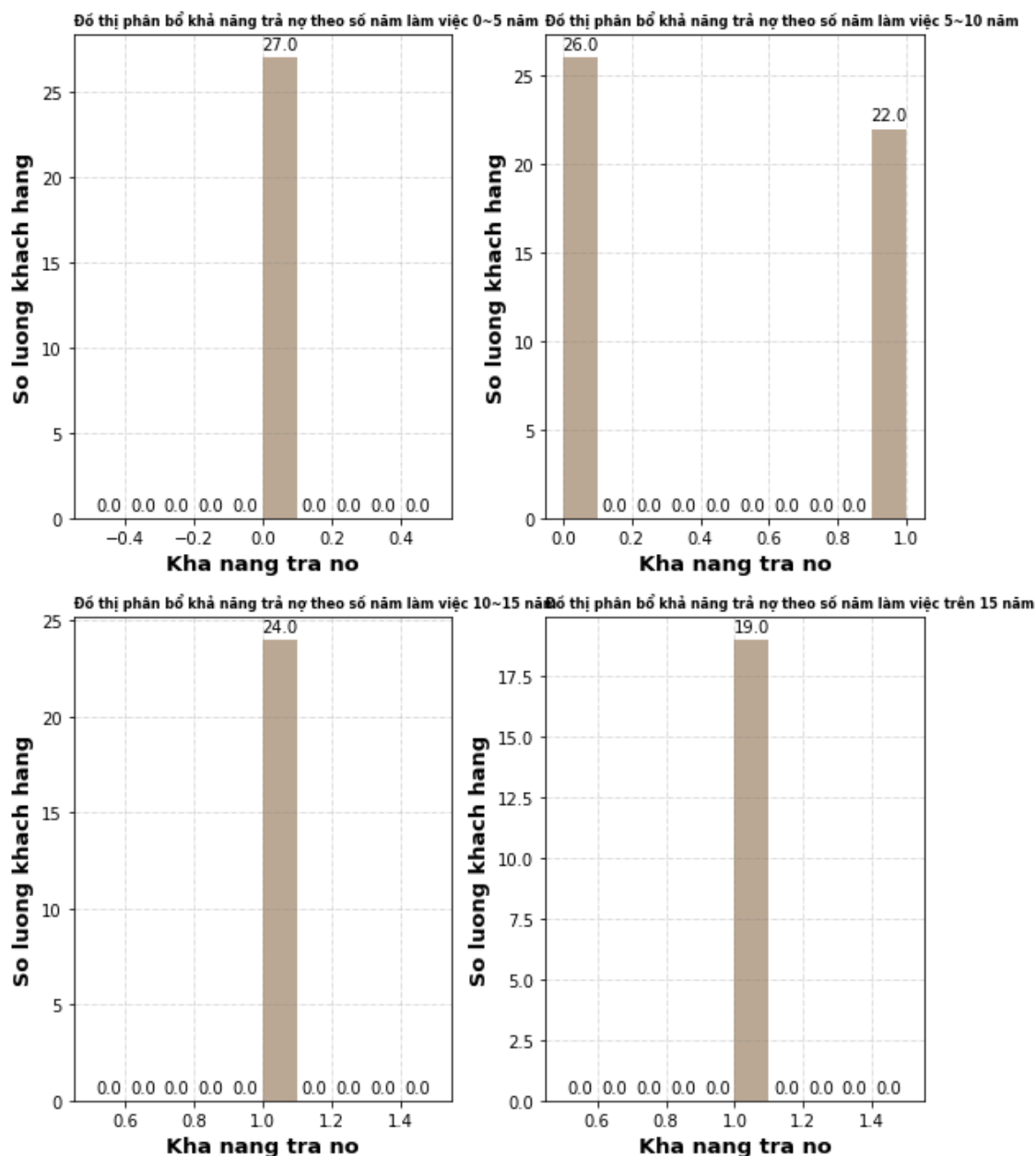
Hình 2: Biểu đồ phân tán trả nợ theo thu nhập

Nhìn vào 4 biểu đồ phân tán trên, ta thấy rõ được rằng số khách hàng thuộc thu nhập thấp hoàn toàn không có khả năng trả nợ và chiếm khoảng 22% (26/118) . Trái ngược vào điều này thì người thuộc nhập cao và rất cao thì hoàn toàn có khả năng trả nợ , trong khi đó tỉ lệ người trả nợ và không trả nợ ở mức thu nhập trung bình là khá cân bằng 27 so với 19 . Dựa vào đây ta cũng thu được kết quả quan trọng là số lượng không trả được nợ là 53/118 (44%) và người trả được nợ là 65/118 (56%)



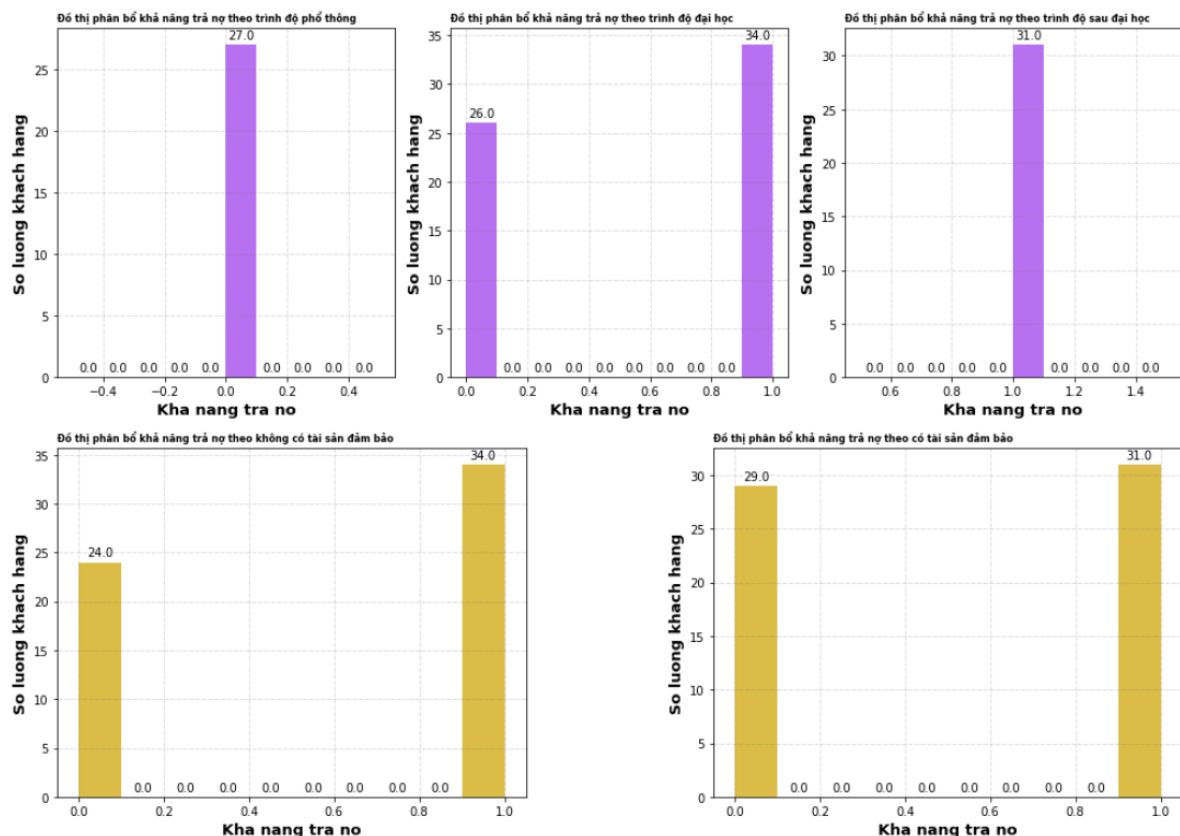
Hình 3: Biểu đồ phân tán khả năng trả nợ theo số tiền vay và quan hệ tín dụng

Nhìn vào 4 biểu đồ phân tán trên, ta thấy rõ được rằng số khách hàng thuộc quan hệ tín dụng lần hay khách hàng thuộc diện quan hệ tín dụng nhiều đều cho ra kết quả không trả nợ hoặc trả được nợ là như nhau (28 so với 35) và (25 so với 30) điều này cho thấy việc quan hệ tín dụng nhiều hay ít không ảnh hưởng nhiều tới khả năng trả nợ. Trái ngược với quan hệ tín dụng, thì nhìn vào biểu đồ ta thấy rõ sự chênh lệch giữa số tiền vay lớn hơn khả năng trả nợ và số tiền vay nhỏ hơn khả năng trả nợ. Ở số tiền vay nhỏ hơn khả năng trả nợ thì việc trả được nợ và việc không trả nợ có số lượng khách hàng cân bằng (= 28) còn ở ngược lại, số tiền vay lớn hơn khả năng trả nợ cho ra kết quả trái ngược khi người trả nợ lên cao hơn nhiều so với người không có khả năng trả nợ (37 so với 25). Vậy nhìn chung có thể nói số tiền vay có ảnh hưởng đến khả năng trả nợ.



Hình 4: Biểu đồ phân tán khả năng trả nợ theo số năm làm việc

Nhìn vào 4 biểu đồ phân tán trên, ta thấy rõ được rằng số khách hàng thuộc số năm làm việc dưới 5 năm đều không có khả năng trả nợ và chiếm khoảng 22% (27/118) . Trái ngược vào điều này thì người thuộc có số năm làm việc từ 10 năm trở lên thì hoàn toàn có khả năng trả nợ , trong khi đó tỉ lệ người trả nợ và không trả nợ ở người có số năm làm việc từ 5 đến 10 là khá cân bằng 26 so với 22 . Qua đó dẫn tới kết luận rằng người làm việc càng lâu thì càng có khả năng trả nợ cao



Hình 5: Biểu đồ phân tán khả năng trả nợ trình độ học vấn và tài sản đảm bảo

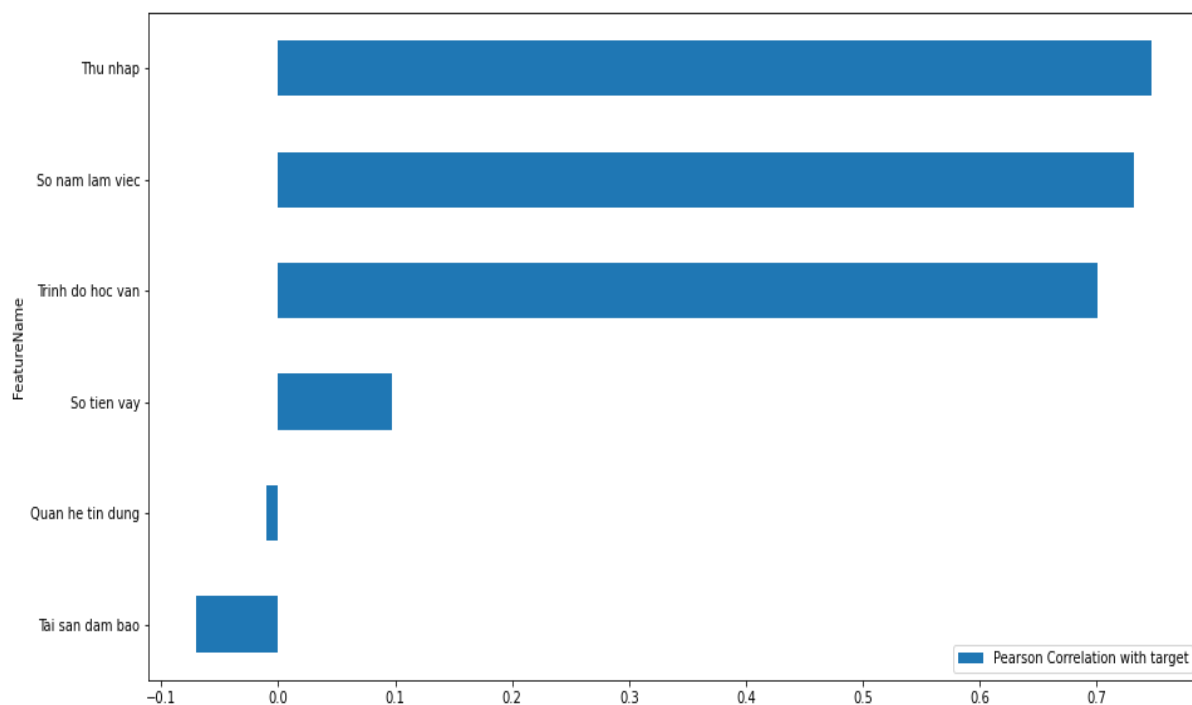
Nhìn vào 5 biểu đồ phân tán trên, ta thấy rõ được rằng số khách hàng có trình độ học vấn phổ thông đều không có khả năng trả nợ và chiếm khoảng 22% (27/118). Trái ngược vào điều này thì người có trình độ học vấn sau đại học thì hoàn toàn có khả năng trả nợ, trong khi đó tỉ lệ người trả nợ và không trả nợ ở người có trình độ học vấn đại học là cũng chênh lệch khá khá 36 so với 26. Qua đó dẫn tới kết luận rằng người có trình độ học vấn càng cao thì càng có khả năng trả nợ cao. Bên cạnh trình độ học vấn thì dựa vào biểu đồ phân bố của tài sản đảm bảo ta thấy việc có tài sản đảm bảo thì tỉ lệ không trả nợ và trả được nợ là 29 và 31 và trái lại không có tài sản đảm bảo là 24 và 34 qua đó việc có tài sản đảm bảo lại khiến con người ta có xu hướng quyết nợ nhiều hơn

4.2 Kết quả sàng lọc

4.2.1 Correlation

	Kha nang tra no	Thu nhap	So tien vay	Quan he tin dung	So nam lam viec	Tai san dam bao	Trinh do hoc van
Kha nang tra no	1.000000	0.747307	0.097153	-0.010130	0.732145	-0.069893	0.700619
Thu nhap	0.747307	1.000000	-0.016698	-0.029618	0.518561	-0.057353	0.538758
So tien vay	0.097153	-0.016698	1.000000	0.003460	0.078702	0.084005	0.070241
Quan he tin dung	-0.010130	-0.029618	0.003460	1.000000	-0.005358	-0.100792	-0.045230
So nam lam viec	0.732145	0.518561	0.078702	-0.005358	1.000000	-0.064739	0.509011
Tai san dam bao	-0.069893	-0.057353	0.084005	-0.100792	-0.064739	1.000000	-0.049235
Trinh do hoc van	0.700619	0.538758	0.070241	-0.045230	0.509011	-0.049235	1.000000

Hình 7: Mô tả sự tương quan giữa các biến



Hình 6: Mức độ tương quan của các biến đến với khả năng trả nợ

Nhìn vào 2 hình trên, nhóm thu được một vài kết quả như sau:

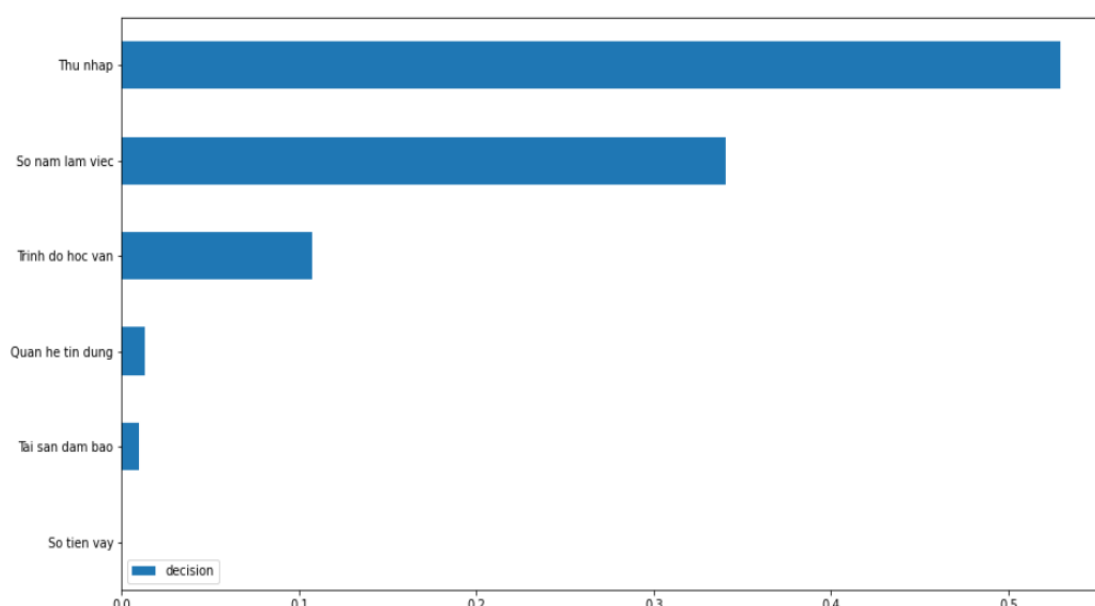
- + 3 biến ảnh hưởng mạnh nhất là Thu nhap, So nam lam viec, Trinh do hoc van
- + Có 4 biến tương quan dương là Thu nhap, So nam lam viec, Trinh do hoc van và So tien vay
- + Có 2 biến tương quan âm là Quan he tin dung và Tai san dam bao
- + Biến Quan he tin dung có độ tương quan thấp nhất nên có sự ảnh hưởng thấp nhất đến với khả năng trả nợ (<0.1). Qua đó có khả năng loại bỏ biến này là khỏi mô hình. Để xác định điều này thì nhóm sẽ xem xét độ quan trọng của biến có trong mô hình ở phần tiếp theo của bài
- + Biến Thu nhap có khả năng tương quan với biến So nam lam viec
- + Biến Thu nhap có khả năng tương quan với biến Trinh do hoc van
- + Biến Trinh do hoc van có khả năng tương quan với biến So nam lam viec

4.2.2 Feature importance trong mô hình nghiên cứu

	feature name	decision
0	Thu nhap	0.529587
1	So tien vay	0.000000
2	Quan he tin dung	0.013313
3	So nam lam viec	0.340404
4	Tai san dam bao	0.009509
5	Trinh do hoc van	0.107187

Hình 8: Mức độ quan trọng của các biến trong mô hình các nhân tố ảnh hưởng khả năng trả nợ

Bên cạnh việc chạy mô hình dự đoán, nhóm còn kiểm tra lại độ quan trọng của các biến có trong mô hình mà nhóm chọn và nhóm đã tìm ra được mức độ quan trọng của từng biến phụ thuộc đến khả năng trả nợ như Hình 9, theo đó các biến có mức độ quan trọng tới khả năng trả nợ lớn nhất chính là biến Thu Nhập và biến Số Tiền Vay ở trong mô hình nghiên cứu của bài viết này lại cho ra kết quả thấp nhất hoặc thậm chí có thể nói là việc Số tiền vay nhiều hay ít quan trọng đối với mô hình dự báo Khả Năng Trả nợ mà nhóm nghiên cứu



Hình 9: Sơ đồ mức độ quan trọng của các biến trong mô hình các nhân tố ảnh hưởng khả năng trả nợ

4.3 Kết quả Logistic

```
[[7 0]
 [0 5]]
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	7
1	1.00	1.00	1.00	5
accuracy			1.00	12
macro avg	1.00	1.00	1.00	12
weighted avg	1.00	1.00	1.00	12

Logistic Regression accuracy: 1.0

Hình 10: Kết quả Logistic theo tỉ lệ train:test là 9:1

Nhìn vào Confusion Matrix ta thấy:

+ True Positive = 7 và True Negative = 5, có nghĩa là mô hình đã dự đoán chính xác 5 khách hàng trả được nợ và 7 khách hàng không trả nợ được

+ False Positive = 0, nghĩa là 0 khách hàng thực sự không trả được nợ được mô hình phân loại không chính xác vào loại “trả nợ được”

+ False Negative = 0, có nghĩa là 0 khách hàng thực sự trả nợ được đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỉ lệ theo tỉ lệ Train:Test là 9:1 thì Kết quả mô hình Logistic Regression đạt được là quá tốt (đạt 100% ở cả 2 biến “Trả được nợ” và “Không trả được nợ”) điều này thì không đúng vì khó có mô hình phân loại nào có thể dự đoán tốt đến vậy do đó nhóm sẽ tiếp tục chạy mô hình với tỉ lệ Train:Test là 8:2

```
[[12  0]
 [ 0 12]]
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	12
1	1.00	1.00	1.00	12
accuracy			1.00	24
macro avg	1.00	1.00	1.00	24
weighted avg	1.00	1.00	1.00	24

Logistic Regression accuracy: 1.0

Hình 11: Kết quả Logistic theo tỉ lệ train:test là 8:2

Nhìn vào Confusion Matrix ta thấy:

+ True Positive = 12 và True Negative = 12, có nghĩa là mô hình đã dự đoán chính xác 12 khách hàng trả được nợ và 12 khách hàng không trả nợ được

+ False Positive = 0, nghĩa là 0 khách hàng thực không trả được nợ được mô hình phân loại không chính xác vào loại “ trả nợ được”

+ False Negative = 0, có nghĩa là 0 khách hàng thực sự trả nợ được đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỉ lệ theo tỉ lệ Train:Test là 8:2 thì Kết quả mô hình Logistic Regression đạt được là quá tốt (đạt 100% ở cả 2 biến “Trả được nợ” và “Không trả được nợ”) điều này thì không đúng vì khó có mô hình phân loại nào có thể dự đoán tốt đến vậy do đó nhóm sẽ tiếp tục chạy mô hình với tỉ lệ Train:Test là 7:3

```

[[19  0]
 [ 1 16]]

```

	precision	recall	f1-score	support
0	0.95	1.00	0.97	19
1	1.00	0.94	0.97	17
accuracy			0.97	36
macro avg	0.97	0.97	0.97	36
weighted avg	0.97	0.97	0.97	36

Logistic Regression accuracy: 0.9722222222222222

Hình 12: Kết quả Logistic theo tỉ lệ train:test là 7:3

Nhìn vào Confusion Matrix ta thấy:

- + True Positive = 19 và True Negative = 16, có nghĩa là mô hình đã dự đoán chính xác 16 khách hàng trả được nợ và 19 khách hàng không trả được nợ
- + False Positive = 0, nghĩa là 0 khách hàng thực sự không trả được nợ được mô hình phân loại không chính xác vào loại “trả nợ được”
- + False Negative = 1, có nghĩa là 1 khách hàng thực sự trả nợ được đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỉ lệ theo tỉ lệ Train:Test 7:3 thì Kết quả mô hình Logistic Regression đạt được là quá tốt (đạt 97,22%) Trong đó :

- Đối với lớp 0 -không trả được nợ :

- + Precision = 95% nghĩa là mô hình dự đoán đúng 95% trường hợp khách hàng không trả được nợ trong tổng số trường hợp được xếp vào loại 0.
- + Recall = 100% có nghĩa là trong tất cả các trường hợp khách hàng không trả được nợ, mô hình dự đoán hết 100% trường hợp, 0% khách hàng không trả được nợ nhưng bị mô hình bỏ sót.

- Đối với lớp 1 -trả được nợ:

- + Precision = 100% nghĩa là mô hình dự đoán đúng 100% trường hợp khách hàng trả được nợ trong tổng số trường hợp được xếp vào loại 1.
- + Recall = 94% có nghĩa là trong tất cả các trường hợp khách hàng trả được nợ, mô hình dự đoán hết 94% trường hợp, 6% khách hàng trả được nợ nhưng bị mô hình bỏ sót.

Nhìn chung kết quả thu được của mô hình Logistic Regression từ 3 cách chia tập Train:Test đều cho ra tỉ lệ dự đoán khá cao (đều trên 97%) qua đó có thể mô hình đang bị Overfitting (trường hợp đúng với thực tế quá mức). Do đó nhóm đã tiếp tục chạy mô hình Cây quyết định và thu được vài kết quả.

4.4 Kết quả Decision Tree

```
[[7 0]
 [0 5]]
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	7
1	1.00	1.00	1.00	5
accuracy			1.00	12
macro avg	1.00	1.00	1.00	12
weighted avg	1.00	1.00	1.00	12

Decision Tree accuracy: 1.0

Hình 13: Kết quả Decision Tree theo tỷ lệ train:test là 9:1

Nhìn vào Confusion Matrix ta thấy:

+ True Positive = 7 và True Negative = 5, có nghĩa là mô hình đã dự đoán chính xác 5 khách hàng trả được nợ và 7 khách hàng không trả nợ được

+ False Positive = 0, nghĩa là 0 khách hàng thực sự không trả được nợ được mô hình phân loại không chính xác vào loại “trả nợ được”

+ False Negative = 0, có nghĩa là 0 khách hàng thực sự trả nợ được đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỷ lệ theo tỷ lệ Train:Test là 9:1 thì Kết quả mô hình Decision Tree đạt được là quá tốt (đạt 100% ở cả 2 biến “Trả được nợ” và “Không trả được nợ”) điều này thì không đúng vì khó có mô hình phân loại nào có thể dự đoán tốt đến vậy do đó nhóm sẽ tiếp tục chạy mô hình với tỷ lệ Train:Test là 8:2

```
[[12 0]
 [0 12]]
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	12
1	1.00	1.00	1.00	12
accuracy			1.00	24
macro avg	1.00	1.00	1.00	24
weighted avg	1.00	1.00	1.00	24

Decision Tree accuracy: 1.0

Hình 14: Kết quả Decision Tree theo tỷ lệ train:test là 8:2

Nhìn vào Confusion Matrix ta thấy:

+ True Positive = 12 và True Negative = 12, có nghĩa là mô hình đã dự đoán chính xác 12 khách hàng trả được nợ và 12 khách hàng không trả nợ được

+ False Positive = 0, nghĩa là 0 khách hàng thực sự không trả được nợ được mô hình phân loại không chính xác vào loại “trả nợ được”

+ False Negative = 0, có nghĩa là 0 khách hàng thực sự trả nợ được đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỉ lệ theo tỉ lệ Train:Test là 8:2 thì Kết quả mô hình Decision Tree đạt được là quá tốt (đạt 100% ở cả 2 biến “Trả được nợ” và “Không trả được nợ”) điều này thì không đúng vì khó có mô hình phân loại nào có thể dự đoán tốt đến vậy do đó nhóm sẽ tiếp tục chạy mô hình với tỉ lệ Train:Test là 7:3

```
[[19  0]
 [ 1 16]]
```

	precision	recall	f1-score	support
0	0.95	1.00	0.97	19
1	1.00	0.94	0.97	17
accuracy			0.97	36
macro avg	0.97	0.97	0.97	36
weighted avg	0.97	0.97	0.97	36

Decision Tree accuracy: 0.9722222222222222

Hình 15: Kết quả Decision Tree theo tỷ lệ train:test là 7:3

Nhìn vào Confusion Matrix ta thấy:

+ True Positive = 19 và True Negative = 16, có nghĩa là mô hình đã dự đoán chính xác 16 khách hàng trả được nợ và 19 khách hàng không trả được nợ

+ False Positive = 0, nghĩa là 0 khách hàng thực sự không trả được nợ được mô hình phân loại không chính xác vào loại “trả được nợ”

+ False Negative = 1, có nghĩa là 1 khách hàng thực sự trả được nợ đã được mô hình phân loại không chính xác vào loại “không trả được nợ”

Nhìn chung khi chia tỉ lệ theo tỉ lệ Train:Test 7:3 thì Kết quả mô hình Decision Tree đạt được là quá tốt (đạt 97,22%) Trong đó :

- Đối với lớp 0 -không trả được nợ:

+ Precision = 95% nghĩa là mô hình dự đoán đúng 95% trường hợp khách hàng không trả được nợ trong tổng số trường hợp được xếp vào loại 0.

+ Recall = 100% có nghĩa là trong tất cả các trường hợp khách hàng không trả được nợ, mô hình dự đoán hết 100% trường hợp, 0% khách hàng không trả được nợ nhưng bị mô hình bỏ sót.

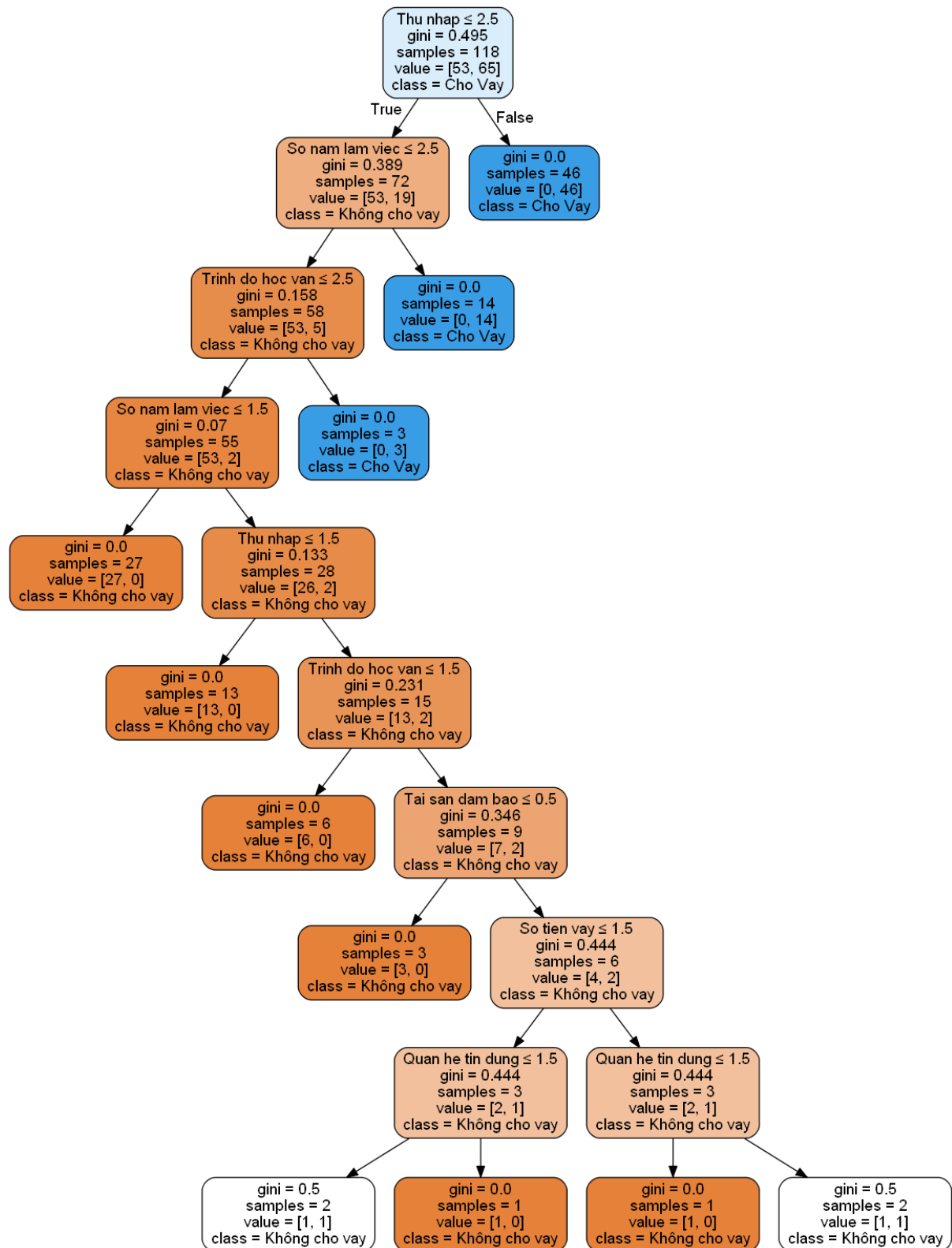
- Đối với lớp 1 -trả được nợ:

+ Precision = 100% nghĩa là mô hình dự đoán đúng 100% trường hợp khách hàng trả được nợ trong tổng số trường hợp được xếp vào loại 1.

+ Recall = 94% có nghĩa là trong tất cả các trường hợp khách hàng trả được nợ, mô hình dự đoán hết 94% trường hợp, 6% khách hàng trả được nợ nhưng bị mô hình bỏ sót.

Sau khi có kết quả từ mô hình Decision Tree, nhóm nhận thấy kết quả thu được là giống nhau nên từ đó có thể kết luận có thể tập dữ liệu còn quá nhỏ để chạy mô hình có lẽ đây là hạn chế của bài nghiên cứu này, nhóm cần thu thập thêm dữ liệu cho bài sau

4.5 Quy tắc cho vay



Theo kết quả cây quyết định, thuộc tính quan trọng nhất để quyết định khả năng trả nợ là “Thu nhập”. Mô hình cho thấy có đến 46/118 thu thập cao và rất cao đều có thể cho vay. Điều này gợi ý ngân hàng có thể ra quyết định nhanh chóng, hạn chế các thủ tục rườm rà, dễ dàng tiếp nhận hồ sơ vay vốn đối với các khách hàng có thu nhập cao và rất cao. Đối với nhóm khách hàng có thu nhập thuộc hai nhóm còn lại (thấp và trung bình), yếu tố tiếp theo mà mô hình gợi ý là “Số năm làm việc” để ra quyết định cho vay hay không. Điều này tương ứng đối với những khách hàng có thu nhập thấp và trung bình, nhưng nếu số năm làm việc trên 10 năm, số tiền tiết kiệm của họ có thể đủ để ngân hàng xem xét và ra quyết định cho vay. Trường hợp này mô hình cho ra 14/72 trường hợp, ở trường hợp ngược lại, thì tiếp tục xét đến “Trình độ học vấn”, đối với “Trình độ học vấn sau đại học, tài chính có thể sẽ tự chủ hơn đối với sinh viên và học sinh, ngân hàng có thể xem xét và cho vay, tuy nhiên chỉ có 3/58 trường hợp là thuộc điều kiện trên. Đối với sinh viên và học sinh, quy trình sẽ phức tạp hơn đôi chút, số năm làm việc chưa đến 5 năm (có thể là thực tập hoặc đi làm thêm) thì mô hình khuyến nghị ngân hàng sẽ quyết định không cho vay đối với trường hợp này, và mô hình cho ra 27/55 trường hợp. Có 28/55 ở tình huống ngược lại, mô hình một lần nữa xem xét đến thu nhập, nếu số năm làm việc trên 5 năm nhưng thu nhập lại rất thấp, có 13/28 trường hợp này và ngân hàng không nên cho vay. Và nếu thu nhập trung bình nhưng chỉ đang là học sinh cũng sẽ được khuyến nghị là không cho vay. Và ở tình huống ngược lại, lần này, tiêu chí tài sản đảm bảo được sử dụng, nếu như khách hàng có tài sản đảm bảo nhưng không có quyền sở hữu cũng sẽ không được vay. Tài sản thuộc quyền sở hữu nhưng số tiền vay lớn hơn khả năng trả nợ cũng không được xét đến. Tiêu chí cuối cùng chính là quan hệ tín dụng, tuy nhiên thì đến bước này tất cả các trường hợp còn lại đều không được xem xét cho vay.

5. Kết luận

Tài chính - Ngân hàng là lĩnh vực chính yếu trong việc vận hành nền kinh tế của một quốc gia, do đó, tập trung củng cố các chức năng của tổ chức tài chính hay bộ máy ngân hàng như nghiên cứu hiện tại của tác giả là vô cùng cần thiết. Gia tăng sự chính xác khi thẩm định tín nhiệm của khách hàng cá nhân sẽ góp phần làm cho hoạt động tín dụng trở nên thuận lợi và hiệu quả hơn. Trong bài nghiên cứu này, tác giả cho thấy mức độ ảnh hưởng đến khả năng trả nợ từ cao đến thấp của các biến được sử dụng lần lượt là: Thu nhập; Số năm làm việc; Trình độ học vấn; Quan hệ tín dụng; Tài sản đảm bảo; và Số tiền vay.

Đánh giá được khả năng trả nợ của một khách hàng giúp các ngân hàng nhận diện được tệp khách hàng tiềm năng hoặc không có khả năng trả nợ, giảm thiểu rủi ro và xử lý kịp thời các khoản vay có vấn đề. Thông qua kết quả có được từ mô hình cây quyết định và Logistic, tác giả có đề xuất một khuyến nghị như sau:

- + Về “Thu nhập”: Chú trọng xem xét yếu tố này đầu tiên khi quyết định cho vay. Xây dựng chính sách ưu đãi cho các khách hàng có thu nhập cao như hạ lãi suất vay để khai thác khách hàng tiềm năng.
- + Về “Số năm làm việc”: Thông thường thâm niên làm việc tỷ lệ thuận với mức độ ổn định tài chính của một cá nhân, vì vậy ngân hàng nên thiết kế mức ưu đãi cũng như mức độ được xem xét cho vay dựa theo từng mức thời gian làm việc như < 10 năm, 10-20 năm, >20 năm.
- + Về “Trình độ học vấn”: Tương tự như số năm làm việc, trình độ học vấn càng cao thì càng nên được hưởng những quyền lợi nhất định cao hơn so với các mức thấp hơn.
- + Về “Quan hệ tín dụng”: Khi tra CIC hoặc xếp hạng tín dụng nội bộ, nếu phát hiện lịch sử quan hệ tín dụng từng có nợ trễ hạn, ngân hàng nên xem xét không cấp tín dụng để tránh rủi ro về sau, xem xét cả những trường hợp quá hạn hoặc chậm thanh toán thẻ tín dụng. Đối với những khách hàng này nếu cấp tín dụng cần hạn chế quy mô khoản vay, yêu cầu tỷ lệ tài sản đảm bảo cao hơn hoặc có mức lãi suất cao hơn để bù đắp rủi ro nếu không may có thể xảy ra nợ quá hạn hoặc xấu nhất là mất khả năng trả nợ.
- + Về “Tài sản đảm bảo”: Những khách hàng có các chỉ số khác thấp và tiềm ẩn rủi ro thì cần phải chú ý đến tỷ lệ bảo đảm của tài sản thế chấp của họ. Sau khi nhận thấy tỷ lệ này có khả năng bù đắp được khoản vay đạt mức yêu cầu thì tiến hành xem xét các yếu tố tiếp.
- + Về “Số tiền vay”: Kết quả cho thấy biến này tác động ít và không có ý nghĩa mạnh mẽ trong việc ra quyết định cho vay đối với khách hàng cá nhân. Tuy nhiên vẫn là một yếu tố nên cân nhắc để so sánh cùng với “Tài sản đảm bảo” hoặc “Thu nhập”.

Tham khảo:

1. Ngân hàng Nhà nước Việt Nam (2021). Thông tư số 11/2021/TT-NHNN: Quy định về phân loại tài sản nợ, phương pháp trích lập dự phòng rủi ro và việc sử dụng dự phòng để xử lý rủi ro trong hoạt động của tổ chức tín dụng, chi nhánh ngân hàng nước ngoài.
2. Nguyễn Đăng Dờn (2016). Giáo trình Quản trị kinh doanh ngân hàng II. Thành phố Hồ Chí Minh: Nhà xuất bản Kinh tế thành phố Hồ Chí Minh.
3. Tạp chí Thị trường Tài chính Tiền tệ số 10/2021 (2021), Giải pháp giảm thiểu rủi ro tín dụng của các NHTM Việt Nam trong bối cảnh COVID-19.
4. Trần Thanh Phong và cộng sự (2020). Đánh giá khả năng trả nợ của khách hàng cá nhân–Nghiên cứu trường hợp Ngân hàng Nông nghiệp và Phát triển Nông thôn Việt Nam–Chi nhánh huyện Tân Hưng, tỉnh Long An. Tạp chí Nghiên cứu Tài chính-Marketing, (57).
5. Nguyễn Hải Trường (2017). Đánh giá khả năng trả nợ và phân loại khách hàng vay tín chấp tại Ngân hàng Thương mại Cổ phần Việt Nam Thịnh vượng.

6. Phan Thị Hằng Nga & Nguyễn Thị Ngọc Quỳnh (2020). Phân tích các yếu tố ảnh hưởng đến khả năng trả nợ của khách hàng cá nhân tại Ngân hàng Thương mại cổ phần Sài Gòn Thương Tín -Chi nhánh Tân Bình. Available at: <https://www.tapchicongthuong.vn/bai-viet/phan-tich-cac-yeu-to-anh-huong-den-kha-nang-tra-no-cua-khach-hang-ca-nhan-tai-ngan-hang-thuong-mai-co-phan-sai-gon-thuong-tin-chi-nhanh-tan-binh-72927.html>.