

ỨNG DỤNG PHÂN TÍCH DỮ LIỆU CỦA HỌC MÁY ĐỂ LỰA CHỌN, ĐÁNH GIÁ NHÀ CUNG CẤP TRONG HỆ THỐNG THÔNG TIN QUẢN LÝ CHUỖI CUNG ỨNG

PGS. TS. Nguyễn Thị Thu Thủy^{1*}

¹*Đại học Thương Mại*

*Tác giả liên hệ: Nguyễn Thị Thu Thủy, Email: nguyentthuthuy@gmail.com

THÔNG TIN CHUNG

Ngày nhận bài: 31/05/2023

Ngày nhận bài sửa: 16/06/2023

Ngày duyệt đăng: 24/06/2023

TỪ KHÓA

Chuỗi cung ứng;

Hỗ trợ ra quyết định;

Học máy;

Nhà cung cấp.

TÓM TẮT

Quá trình lựa chọn, đánh giá nhà cung cấp là một trong những công việc ra quyết định quan trọng của nhà quản lý. Quy trình lựa chọn nhà cung cấp thường được tích hợp trong hệ thống thông tin chuỗi cung ứng để có thể hỗ trợ doanh nghiệp có những lựa chọn cuối cùng cho danh sách các nhà cung cấp phù hợp nhất đối với doanh nghiệp. Trong điều kiện cạnh tranh trên thị trường ngày nay, các doanh nghiệp luôn phải tìm cách tối ưu hóa chi phí đầu vào để đạt được sản phẩm đầu ra cạnh tranh nhất. Việc xây dựng, lựa chọn và sử dụng hợp lý danh mục nhà cung cấp cũng là một trong những nhân tố cần thiết cho doanh nghiệp trong hoạt động kinh doanh của mình. Việc ứng dụng khoa học công nghệ, cụ thể là học máy trong hệ thống thông tin chuỗi cung ứng sẽ góp phần hỗ trợ nhà quản lý có được những lựa chọn đúng đắn dựa trên cơ sở khoa học dữ liệu thu thập trong hệ thống. Điều đó cũng làm minh bạch hóa quá trình lựa chọn nhà cung cấp trong chuỗi cung ứng của doanh nghiệp, tạo sức cạnh tranh về giá cả sản phẩm cho doanh nghiệp trên thị trường.

ABSTRACT

The supplier selection process is one of the important decision-making tasks of managers. The supplier selection process is often integrated in the supply chain management information system to help businesses make the final choices for a list of the most suitable suppliers for them. In today's competitive market conditions, businesses always have to find ways to optimize input costs to achieve the most competitive output products. The construction, selection, management and development of a supplier portfolio are also one of the factors that help businesses achieve their goals. The application of science and technology, especially machine learning in the supply chain information system, will assist managers to make the right choices which are based on the data science collected in the system. That also makes transparent the supplier selection process in the supply process of enterprises, creating competitiveness on product prices for businesses in the market.

1. GIỚI THIỆU

Trong chuỗi cung ứng, nhóm thành viên cung cấp đầu vào, các nguyên vật liệu, sản phẩm hoặc thông tin cho doanh nghiệp để doanh nghiệp có thể thực hiện một số quy trình tiếp theo như bán hàng hoặc cung cấp dịch vụ hoặc sản xuất sản phẩm, nhóm thành viên này được gọi là nhà cung cấp.

Quá trình lựa chọn nhà cung cấp có thể hiểu là quá trình sàng lọc để có thể đưa ra quyết định cuối cùng mang tính quyết định của nhà quản lý và thường bao gồm một số bước cơ bản như: Xác định tập nhà cung cấp tiềm năng; đề xuất các tiêu chí cho tập nhà cung cấp này; sau đó hệ thống cần đưa ra quy trình sàng lọc, lựa chọn và đánh giá sao cho phù hợp theo các tiêu chí đã được xây dựng trước đó. Dựa vào các kinh nghiệm quản lý của bản thân, đồng thời sử dụng các thông tin hỗ trợ từ hệ thống thông tin quản lý chuỗi cung ứng, nhà quản lý sẽ có lựa chọn cuối cùng cho danh sách các nhà cung cấp phù hợp nhất đối với doanh nghiệp. Trong điều kiện cạnh tranh ngày nay, việc tìm kiếm, thu thập, sàng lọc và lựa chọn được danh sách nhà cung cấp tốt cho doanh nghiệp là một trong yếu tố vô cùng quan trọng giúp cho doanh nghiệp có thể đạt được mục tiêu của mình. Bởi vì, nếu xem xét trên góc độ ảnh hưởng của nguyên vật liệu nhận được để doanh nghiệp có thể sản xuất ra sản phẩm thì các yếu tố như chất lượng nguyên vật liệu tốt, giá thành hợp lý hoặc rẻ, sự giao hàng đúng thời hạn, ... sẽ góp phần nâng giá trị của hàng hóa được sản xuất tại doanh nghiệp, và giảm giá bán trên thị trường do tiết kiệm chi phí đầu vào. Điều này cũng mang lại lợi thế cạnh tranh cho doanh nghiệp. Để hỗ trợ sự thành công của quan hệ đối tác, cần có sự hợp tác đầy đủ giữa các bên bao gồm: nhà sản xuất, nhà cung cấp và đôi khi nhà cung cấp của nhà cung cấp. Mỗi khi doanh nghiệp xác định được danh sách các nhà cung cấp với các thông tin về họ, doanh nghiệp sẽ tiến hành đánh giá các nhà cung cấp tiềm năng này theo các tiêu chí đã xác định.

Có nhiều tiêu chí được xây dựng dựa vào đặc thù của doanh nghiệp và mục đích kinh doanh của họ. Tuy nhiên, có thể thấy nổi bật ba tiêu chí chính khi xem xét đánh giá nhà cung cấp là: Chất lượng của hàng hóa, Chi phí/giá thành của hàng hóa, và tiêu chí Giao hàng.

Quá trình lựa chọn nhà cung cấp sẽ không kết thúc khi đã có đủ danh sách nhà cung cấp mà đó phải là một quá trình liên tục nhằm mục đích theo dõi, phát triển và nếu cần có thể thay thế các nhà cung cấp hiện tại bằng các nhà cung cấp mới mà họ có thể có lợi hơn về mặt đáp ứng các tiêu chí của doanh nghiệp. Ngày nay, với sự phát triển của nền công nghiệp 4.0, đã có rất nhiều thông tin về chuỗi cung ứng được tiếp cận một cách nhanh chóng. Công nghệ kỹ thuật số có thể cho phép linh hoạt ra quyết định bằng cách cung cấp dữ liệu thời gian thực cho tất cả các thành viên của chuỗi cung ứng. Ngoài ra, sự phát triển nhanh chóng của công nghệ thông tin cũng làm cho việc thu thập, truyền tải và lưu trữ thông tin trở nên dễ dàng hơn. Vì vậy, nó có thể coi là một phương pháp hiệu quả để đánh giá các nhà cung cấp trong xã hội ngày nay, nơi mà mọi thứ được định hình theo thông tin số. Cụ thể, để đạt được kết quả mong muốn thông qua việc sử dụng thông minh các dữ liệu thu thập được từ nhà cung cấp một cách hiệu quả, nghiên cứu sẽ áp dụng phương pháp học máy, là một kỹ thuật có thể phân tích các tập dữ liệu lớn, cho việc lựa chọn nhà cung cấp trong hệ thống thông tin chuỗi cung ứng.

Nghiên cứu được tổ chức như sau: Phần 2 trình bày các tài liệu liên quan hiện tại về các phương pháp lựa chọn nhà cung cấp và các nghiên cứu về việc lựa chọn nhà cung cấp ứng dụng học máy và các nội dung cơ bản bao trùm học máy, các kiểu học và nhiệm vụ học, Phần 3 trình bày một nghiên cứu trường hợp về việc sử dụng học máy trong lựa chọn và phân loại nhà cung cấp với các thuật toán học máy khác nhau. Các hàm ý quản lý và kết luận được trình bày ở phần thứ 4 của nghiên cứu.

2. TỔNG QUAN VỀ HỆ THỐNG THÔNG TIN QUẢN LÝ CHUỖI CUNG ỨNG VÀ ỨNG DỤNG PHÂN TÍCH DỮ LIỆU CỦA HỌC MÁY TRONG VIỆC LỰA CHỌN, ĐÁNH GIÁ NHÀ CUNG CẤP

2.1. Hệ thống thông tin quản lý chuỗi cung ứng

Hệ thống thông tin quản lý chuỗi cung ứng được hiểu là hệ thống thông tin liên quan đến quy trình sản xuất trong doanh nghiệp hoặc tạo ra các sản phẩm là hàng hóa hoặc dịch vụ. Nói cách khác, nó được hiểu là từ việc cung cấp các nguyên vật liệu thô cho quy trình sản xuất, và tiếp nối các hoạt động cho đến khi cung cấp sản phẩm cuối cùng đến tay người tiêu dùng (IBM, 2023). Doanh nghiệp sẽ tạo ra một hệ thống các các “mắt xích” trong chuỗi bắt đầu từ mạng lưới các nhà cung cấp, tổ chức sản xuất và logistic để giao dịch sản phẩm hoặc dịch vụ của doanh nghiệp với người dùng cuối.

Hệ thống thông tin chuỗi cung ứng có các thành phần cơ bản là:

- Xây dựng kế hoạch: xây dựng kế hoạch và quản lý các nguồn lực trong doanh nghiệp để đáp ứng nhu cầu của khách hàng. Khi doanh nghiệp thiết lập chuỗi cung ứng, họ sẽ xác định thông qua các số liệu để đo lường xem liệu chuỗi cung ứng có hiệu quả không hay có mang lại giá trị cho khách hàng không.

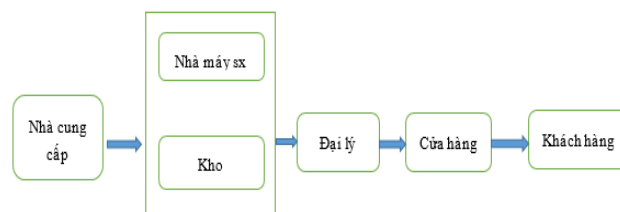
- Tìm nguồn cung: Bước này bao gồm việc lựa chọn nhà cung cấp. Tiếp theo đó, cần có các thiết lập cần thiết để theo dõi và quản lý các mối quan hệ với nhà cung cấp. Cụ thể bao gồm: đặt hàng hóa, nhận và quản lý hàng tồn kho, thanh toán cho nhà cung cấp. Trong thành phần này, việc lựa chọn được nhà cung cấp hiệu quả là một trong những yếu tố quan trọng để góp phần thúc đẩy các tiến trình khác của doanh nghiệp. Ở trong khuôn khổ nghiên cứu này, việc áp dụng công nghệ để đánh giá, tìm ra được tập các nhà cung cấp sẽ được tập trung nghiên cứu nhằm mục đích giúp cho doanh nghiệp có được tập các nhà cung

cấp tiềm năng đáng tin cậy phục vụ có hiệu quả nhất cho hoạt động quản lý chuỗi cung ứng của doanh nghiệp.

- Tổ chức sản xuất hoặc triển khai dịch vụ: thành phần này được thực hiện tại doanh nghiệp nếu có.

- Phân phối sản phẩm/dịch vụ: Phối hợp với các hệ thống khác trong doanh nghiệp, như hệ thống Marketing, Sale,... để thực hiện các đơn đặt hàng từ phía khách hàng, lên lịch giao hàng hóa/dịch vụ, thiết lập hóa đơn và thực hiện quy trình thanh toán.

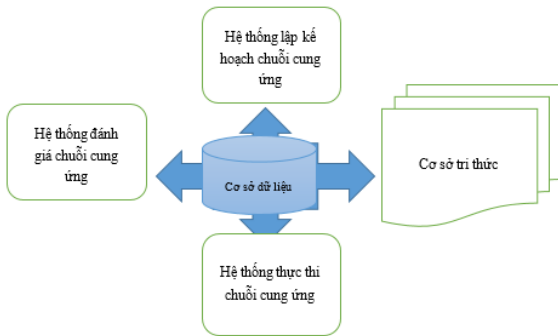
- Phản hồi/trả lại: Trong một số hệ thống chuỗi cung ứng, doanh nghiệp cần thiết lập một quy trình trả lại sản phẩm hoặc từ chối dịch vụ từ phía khách hàng.



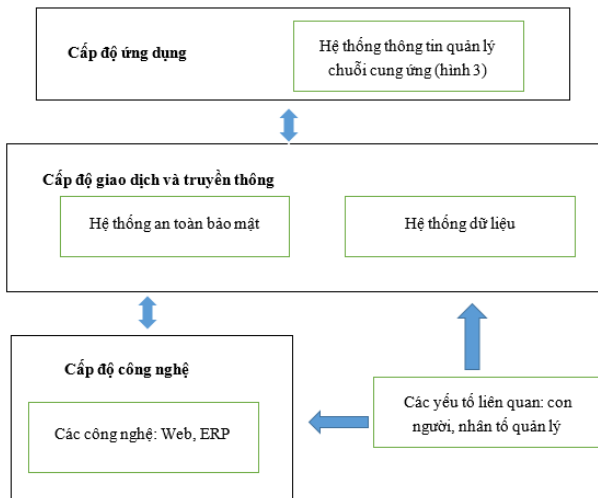
Hình 1. Quy trình chuỗi cung ứng cơ bản
(Andrii và cộng sự, 2018)

Quy trình chuỗi cung ứng được thể hiện ở Hình 1. Quy trình này có thể được xem như một luồng hệ thống thông tin các tác nghiệp mà nhằm quản lý hiệu quả các sản phẩm/nguyên liệu, hoặc dịch vụ từ nhà cung cấp tới các tổ chức/ doanh nghiệp trung gian để đến với khách hàng. Hay nói cách khác, đó là một hệ thống mạng lưới logistics nằm giữa nhà cung cấp đầu tiên đến khách hàng cuối cùng. Trong quy trình này, lựa chọn nhà cung cấp là bước đầu tiên cho các quy trình tác nghiệp khác trong doanh nghiệp. Nếu doanh nghiệp chọn được những nhà cung cấp thỏa mãn mục tiêu của mình, họ sẽ có sản phẩm tốt cung cấp đến khách hàng. Hệ thống thông tin chuỗi

cung ứng được thể hiện ở Hình 2, và Hình 3 dưới đây.



Hình 2. Mô hình hệ thống thông tin quản lý chuỗi cung ứng



Hình 3. Kiến trúc theo các cấp độ của hệ thống thông tin chuỗi cung ứng.

Triển khai mô hình hệ thống thông tin chuỗi cung ứng ở các cấp độ được thể hiện ở Hình 3. Ở cấp độ công nghệ, các nhân tố cấu thành bao gồm: các công nghệ hiện tại như công nghệ web, ERP, thực tế ảo (Reality)... Ngoài ra ở tầng cơ sở hạ tầng này còn bao gồm các yếu tố khác như: Các nhân tố con người, các quy trình quản lý,... Ở tầng truyền thông và giao dịch, các hệ thống đảm bảo cho việc truyền dữ liệu được tích hợp vào đây. Nó bao gồm: hệ thống bảo mật và hệ thống quản trị dữ liệu. Tầng ứng dụng là các giao

diện dành cho người dùng của hệ thống thông tin chuỗi cung ứng. Người dùng tương tác với hệ thống thông qua các giao diện và toàn bộ giao dịch này được đảm bảo an toàn thông tin trước khi thực hiện trên đường truyền như Internet hoặc các thiết bị truyền thông hay giao tiếp với hệ thống khác (như ERP).

2.2. Ứng dụng phân tích dữ liệu của học máy trong việc lựa chọn, đánh giá nhà cung cấp trong hệ thống thông tin chuỗi cung ứng

Như đã đề cập ở trên, việc đánh giá, lựa chọn nhà cung cấp là một trong những bước đầu tiên của quy trình chuỗi cung ứng, hay cụ thể là một trong các thành phần chính của hệ thống thông tin quản lý chuỗi cung ứng (Hình 2). Như vậy, việc chi tiết các bước cụ thể của thành phần này nhằm mục đích tối ưu hóa quy trình lựa chọn để đạt được mục tiêu cuối là tập các nhà cung cấp hiệu quả cho doanh nghiệp.

Quá trình lựa chọn, đánh giá nhà cung cấp là một loạt các bước mua để xác định các yêu cầu về sản phẩm hoặc dịch vụ và kết hợp chúng với khả năng và giá cả đưa ra của nhà cung cấp. Nó bao gồm việc xác định các nhà cung cấp tiềm năng, nhận báo giá hoặc đề xuất với giá thầu cạnh tranh, đánh giá các nhà cung cấp bằng cách liên hệ với các tài liệu tham khảo và áp dụng danh sách kiểm tra tiêu chí lựa chọn nhà cung cấp của công ty và ký kết hợp đồng.

5 bước cơ bản trong quá trình lựa chọn nhà cung cấp bao gồm:

Bước 1: Xác định nhu cầu kinh doanh

Bước 2: Kiểm tra các nhà cung cấp tiềm năng

Bước 3: Đặt tiêu chí đánh giá

Bước 4: Kết nối và đánh giá nhà cung cấp

Bước 5: Hoàn thiện lựa chọn nhà cung cấp

Việc thực hiện quy trình lựa chọn nhà cung cấp được thể hiện thông qua cách cập nhật chính sách của công ty, sự chuẩn bị danh sách kiểm tra

tiêu chí lựa chọn nhà cung cấp, cách thức chọn và sử dụng hệ thống mua bán hiệu quả. Để quản lý quy trình làm việc, hệ thống sẽ sử dụng danh sách nhà cung cấp được phê duyệt mà đã được xác định trong các RFI (yêu cầu thông tin), RFQ (yêu cầu báo giá) hoặc RFP (yêu cầu đề xuất). Hệ thống còn đưa ra các nhà cung cấp tiềm năng để người dùng có thể kiểm tra và đánh giá, đồng thời chọn nhà cung cấp đủ điều kiện nhất trên cơ sở nhiều tiêu chí đã đặt ra của doanh nghiệp.

Từ 5 bước cơ bản nói trên, 11 bước chi tiết trong quy trình lựa chọn nhà cung cấp được thể hiện như sau (Tipalti, 2023).

Bước 1: Khởi động việc sử dụng hệ thống thông tin quản lý chuỗi cung ứng

Bước 2: Xác định các yêu cầu về sản phẩm và nhà cung cấp, đồng thời chuẩn bị RFI, RFQ hoặc RFP để thu hút giá thầu

Bước 3: Tìm kiếm danh sách các nhà cung cấp và đánh giá nhà cung cấp được phê duyệt của doanh nghiệp.

Bước 4: Xác định bất kỳ nhà cung cấp tiềm năng nào khác có thể thay thế

Bước 5: Tham khảo danh sách tiêu chí lựa chọn nhà cung cấp và các chính sách chung của doanh nghiệp.

Bước 6: Xây dựng danh sách thu gọn các nhà cung cấp là đủ điều kiện nhất

Bước 7: Xem bản demo trình diễn (nếu có) từ nhà cung cấp, tiến hành gặp gỡ các nhà cung cấp tiềm năng trong trường hợp mua với số lượng lớn

Bước 8: Tham khảo các tài liệu liên quan đến nhà cung cấp thông qua ý kiến tham khảo của khách hàng. Có thể xem tham khảo bất kỳ nghiên cứu trường hợp nào có liên quan được lấy trên trang web của nhà cung cấp.

Bước 9: Kiểm tra sự ổn định tài chính của các nhà cung cấp và nhà cung cấp tiềm năng

Bước 10: Đàm phán các điều khoản, giá cả, giảm giá và ngày giao hàng

Bước 11: Hợp đồng với nhà cung cấp được chọn

Việc ứng dụng học máy trong việc lựa chọn nhà cung cấp đã được đề cập ở nhiều nghiên cứu khác nhau. Phương pháp MCDM (multiple choice decision methods) (Cengiza và cộng sự, 2017) đã được sử dụng rộng rãi trong các nghiên cứu lựa chọn nhà cung cấp dựa trên bảng hỏi các nhà cung cấp mua hàng từ phía doanh nghiệp. Các nhà quản lý mua hàng đánh giá các nhà cung cấp bằng cách tiêu chí đánh giá đưa ra nhận định chủ quan thông qua các bảng hỏi (Akarte và cộng sự, 2001). Trong các nghiên cứu, phương pháp này được thực hiện thông qua hai giai đoạn. Giai đoạn 1, các tiêu chí đánh giá được thiết lập và xây dựng thành bảng hỏi, và tiến hành khảo sát. Giai đoạn 2 sẽ sử dụng một số kỹ thuật đánh giá như AHP (mô hình trọng số) và DEA (mô hình phân tích bao số liệu- Data Envelopment Analysis) hay kỹ thuật học máy (Machine Learning). Như vậy việc áp dụng học máy có thể trợ giúp cho các bước 3,4,5,6 để có được tập danh sách các nhà cung cấp rút gọn nhất thỏa mãn các tiêu chí đưa ra của doanh nghiệp.

Việc sử dụng học máy trong mô hình lựa chọn nhà cung cấp là để đo lường mức độ hiệu quả của nhà cung cấp. Các kỹ thuật học máy cũng đã được sử dụng kết hợp với AHP, trong đó AHP được sử dụng để tính trọng số tiêu chí và kỹ thuật học máy dùng để lựa chọn nhà cung cấp phù hợp (Tang và cộng sự, 2013; Lakshmanpriya và cộng sự, 2013). Tương tự, Kar (2015) kết hợp AHP với học máy và sử dụng Tập mờ Lý thuyết (FST), trong khi Bottani và Rizzi (2008) kết hợp FST với AHP và phân tích cụm để giảm số lượng lựa chọn thay thế và chọn cụm tốt nhất có thể cho quá trình lựa chọn. Ngoài ra, một giải pháp đề xuất khác cho việc lựa chọn nhà cung cấp là sự kết hợp của DEA, cây quyết định (DT) và kỹ thuật hồi quy của mạng nơ ron (Wu, 2009). DEA được sử dụng để chia các nhà cung cấp thành các nhóm

tùy thuộc vào hiệu quả điểm số; sau đó, dữ liệu được sử dụng để đào tạo cây quyết định và kỹ thuật hồi quy mạng nơ ron.

Đặc điểm chung của các phương pháp kể trên là việc đề xuất các phương pháp tích hợp giữa các kỹ thuật MCDM như AHP và DEA, sau đó sử dụng kết quả của các phương pháp này để các kỹ thuật học máy học các dữ liệu để đưa ra mức độ dự báo/phòng đoán (prediction) cuối cùng cho việc lựa chọn nhà cung cấp.

3. NGHIÊN CỨU THỰC NGHIỆM ỨNG DỤNG LỰA CHỌN, ĐÁNH GIÁ NHÀ CUNG CẤP

3.1. Phân tích dữ liệu

Dữ liệu thực hiện trong nghiên cứu này là các dữ liệu thứ cấp được thu thập được từ Nong &Ho (2019). Thông tin có thể nhóm thành ba nhóm tiêu chí:

- Yếu tố tổ chức
- Thước đo hiệu suất
- Quốc gia cung ứng.

cung ứng của các nhà cung cấp. Nhóm cuối cùng, quốc gia tìm nguồn cung ứng, được sử dụng để đánh giá môi trường vĩ mô của các quốc gia của nhà cung cấp, có thể ảnh hưởng đến thương mại giữa người mua và nhà cung cấp.

Các nhóm tiêu chí này đã đã được áp dụng với phương pháp AHP để tạo thành 42 nhóm con và 8 nhóm chính liên quan đến Chất lượng (Quality); Chi phí (Cost); Vận chuyển (Delivery), Dịch vụ (Service); Khả năng đáp ứng (Capablity); Mối quan hệ (Relationship); Trách nhiệm xã hội (corporate social responsibility -CSR); Xuất xứ (Source Country).

Dữ liệu thu thập được là 362 dòng và 9 thuộc tính đã được mã hóa thành các giá trị số. Biến quyết định được thể hiện thông qua biến y để xác định mức quyết định tương ứng với sự lựa chọn của học máy là có/không. Nhân của việc phân lớp được thực hiện dựa trên đánh giá của các chuyên gia với thang đo likert từ 1 đến 5 tương ứng với mức từ rất không đồng ý đến rất đồng ý (mức 5).

Bảng 1. Thống kê mô tả của các biến

Index	QA	CO	DE	SE	CA	PR	CS	SC	SD	y
Count	362.0	362.0	362.0	362.0	362.0	362.0	362.0	362.0	362.0	362.0
Mean	3.1	3.7	3.7	2.9	3.7	4.2	3.8	4.2	3.8	0.824
Std	0.6	0.6	0.6	0.8	0.7	0.8	0.7	0.7	0.5	0.4
min	1.0	2.0	1.0	1.0	2.0	1.0	1.0	2.0	1.0	0.0
max	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	1.0

Trong các nhóm trên, các yếu tố tổ chức chủ yếu tập trung vào khả năng, trách nhiệm xã hội và đặc điểm của nhà cung cấp. Nhóm này bao gồm khả năng của nhà cung cấp, mối quan hệ với người mua và trách nhiệm xã hội của công ty. Nhóm chỉ số hiệu suất bao gồm chất lượng, chi phí, phân phối và dịch vụ dùng để đo sự hiệu quả

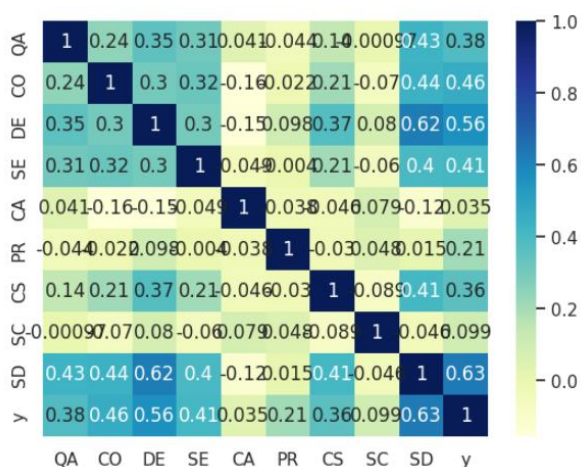
Dữ liệu để gán nhãn cho biến y căn cứ vào giá trị của cột SD cụ thể như sau”

Nếu các chuyên gia cho điểm nhà cung cấp từ 1 đến 3 điểm, độ đánh giá được gán nhãn cho nhà cung cấp đó là “0” (tương ứng với mức độ chấp nhận rủi ro là không). Ngược lại, nhãn được gán là “1” tương ứng với điểm đánh giá lớn hơn 3

(được 4 hoặc 5 điểm) (tương ứng với việc chấp nhận rủi ro là yes, hay nói cách khác đây là nhóm nhà cung cấp có độ rủi ro thấp). Việc gán nhãn này cho người được hỗ trợ trong hệ thống biết được mức độ xem xét các nhà cung cấp này với các mức độ chấp nhận độ rủi ro khác nhau để có các quyết định quản lý tương ứng.

Thống kê mô tả của các nhóm biến đầu vào và ra của mô hình có thể xem ở Bảng 1 và mối tương quan giữa các biến có thể xem ở Hình 4 dưới đây.

Dữ liệu ở Bảng 1 cho thấy các biến không có sự biến động bất thường và giá trị lớn nhất là 5, giá trị nhỏ nhất là 1 (tương ứng với 5 mức đo likert ở bảng hỏi). Theo Hình 4, sự đậm nhạt của màu sắc thể hiện các biến có mối tương quan chặt chẽ hay không. Màu càng đậm (gần 1) nghĩa là các biến có mối tương quan chặt với nhau. Rõ ràng hầu hết các biến là độc lập, ngoại trừ biến SD là biến quyết định liên quan đến biến y, nên thể hiện mối liên quan chặt chẽ. Ở Hình 4 cũng chỉ ra mối tương quan tốt giữa biến DE (vận chuyển) với biến quyết định SD và y. Điều đó cũng cho thấy việc vận chuyển (logistics) đúng thời gian của nhà cung cấp ảnh hưởng quan trọng đến việc lựa chọn nhà cung cấp.



Hình 4. Mối tương quan giữa các biến trong mô hình

3.2. Mô hình thuật toán

Kỹ thuật hồi quy tuyến tính bội (LR) là một công cụ chủ yếu khi muốn xem xét mối tương quan giữa các biến và được sử dụng rộng rãi đối với các dữ liệu mang tính thống kê (như phiếu điều tra) (Adam, 2018). Cây quyết định là đại diện cho việc xây dựng các tập luật dựa trên các giá trị có sẵn trong tập dữ liệu ban đầu, và phù hợp với các bài toán hỗ trợ ra quyết định. Hơn nữa, nghiên cứu kế thừa các phương pháp đã được công bố trước đó về việc vận dụng học máy trong việc hỗ trợ ra quyết định lựa chọn nhà cung cấp đã được đề cập ở mục 2.2. Vì vậy, trong nghiên cứu này, các kỹ thuật học máy bao gồm hồi quy tuyến tính bội (Linear Regression-LR) và cây quyết định (Random Forest Tree, RFT) được sử dụng với bộ dữ liệu đã được lựa chọn thông qua phương pháp AHP. Cụ thể thuật toán (pseudo code) của LR và RFT được thể hiện như sau:

Thuật toán hồi quy

Thuật toán hồi quy là thuật toán khá cơ bản trong machine learning. Mục tiêu của bài toán là được sử dụng cho vấn đề liên quan đến dự đoán, dự báo.

Phương pháp hồi quy có thể hiểu là một phương pháp thống kê trong đó phương pháp sẽ xây dựng mối quan hệ giữa các biến độc lập (thuộc tính trong tập dữ liệu) và biến phụ thuộc (thường là một biến) và biến phụ thuộc này chính là biến cần dự báo.

Trong machine learning, các bài toán chủ yếu được xử lý là bài toán liên quan đến học có giám sát, không giám sát và bán giám sát. Thuật toán hồi quy là thuộc về bài toán học có giám sát. Nghĩa là, bài toán đưa vào học máy được xây dựng dựa trên tập dữ liệu mà ở đó tập dữ liệu này đã được gán nhãn (output, đích). Việc còn lại là huấn luyện máy học để tạo ra mối quan hệ giữa các biến input với biến output cần tìm. Hay nói cách khác là dựa trên tập dữ liệu huấn luyện để xây dựng hàm phụ thuộc giữa biến phụ thuộc và các biến độc lập.

Chi tiết các bước được thể hiện như sau:

- Bước 1: Xác định tập các biến phụ thuộc $x_1, x_2, x_3, \dots, x_n$ trong đó x_i là nhóm yếu tố đầu vào thu thập được từ tập dữ liệu.
- Bước 2: Xác định biến phụ thuộc y là biến quyết định và là đầu ra của mô hình
- Bước 3: Xây dựng hàm số tuyến tính bội:

$$y = x_1 * \beta_1 + x_2 * \beta_2 + \dots + x_n * \beta_n + \epsilon$$
(1).

Trong đó các β_i là trọng số ảnh hưởng tác động vào các nhân tố đầu vào.

- Bước 4: Tính các giá trị $y_i = \sum_{j=1}^n x_{ij} * \beta_j + \epsilon_i$
- Bước 5: Tính sai số (độ lỗi) RMSE (Root Mean Square Error) của mô hình theo công thức:

$$RMSE = \sqrt{\frac{\sum (y_{thực\ t\hat{e}} - y_{d\hat{u}\hat{a}\hat{o}})^2}{N}}$$
(2),
trong đó N là tổng số quan sát. Hoặc tính các độ đo phân lớp khác dựa vào ma trận confusion matrix.

Thuật toán RFT

Random Forest Tree là một thuật toán học máy học có giám sát rất phổ biến trong lĩnh vực học máy. RFT dành cho bài toán có biến mục tiêu đã được gán nhãn. Thuật toán này được sử dụng để giải các bài toán hồi quy (với biến mục tiêu số) và phân loại (biến mục tiêu phân loại). Đây là một phương pháp tập hợp. Nghĩa là thuật toán kết hợp các dự đoán từ các mô hình khác mà chủ yếu mỗi mô hình nhỏ hơn này chính là một nhân tố trong quần thể rừng (forest) và chính là một cây quyết định.

Chi tiết các bước thực hiện như sau:

- Bước 1: Chọn một bộ dữ liệu bất kỳ từ tập dữ liệu huấn luyện.
- Bước 2: Xây dựng cây quyết định dựa trên mối liên kết với bộ dữ liệu đã chọn.
- Bước 3: Chọn số cây trong rừng cây
- Bước 4: Lặp lại các bước 1 và 2.
- Bước 5: Khi có bộ dữ liệu mới, chỉ ra dự đoán của các cây nói trên và gán giá trị tìm được (đầu ra) cho điểm dữ liệu này thông qua việc đạt được nhiều số phiếu bầu chọn của các cây (Vote).

Đối với thuật toán RFT, các cây quyết định con được xây dựng để ngăn chặn việc chia tách các quyết định ở một mức quá sâu khi chỉ sử dụng cây quyết định (decision tree-DT). Đây cũng thể hiện sự hiệu quả của mô hình cây quyết định so với giả thuyết sử dụng tuyến tính bội.

3.3. Kết quả

Nghiên cứu thực nghiệm bộ dữ liệu với hai thuật toán LR và RFT đã phân tích ở mục 3.2. Chương trình được cài đặt với ngôn ngữ Python cho các thuật toán nói trên với 362 dòng dữ liệu, 8 biến đầu vào (input) và 01 biến đầu ra (output). Kết quả tính toán hiệu suất của mô hình được thực hiện thông qua việc đánh giá precision; recall; f1-score; và độ chính xác (accuracy).

	precision	recall	f1-score
support			
0	0.89	1.00	0.94
16			
1	1.00	0.98	0.99
93			
accuracy			0.98
109			
macro avg	0.94	0.99	0.97
109			
weighted avg	0.98	0.98	0.98
109			

Hình 5. Kết quả thực nghiệm dữ liệu với LR.

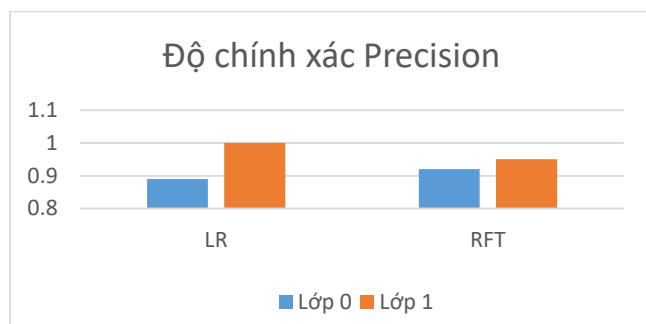
	precision	recall	f1-score	support
0	0.92	0.69	0.79	16
1	0.95	0.99	0.97	93
accuracy			0.94	109
macro avg	0.93	0.84	0.88	109
weighted avg	0.94	0.94	0.94	109

Hình 6. Kết quả thực nghiệm dữ liệu với RFT

Hình 5 và 6 chỉ ra độ chính xác của mô hình tương ứng với 02 thuật toán LR và RFT là 0.98 và 0.94. Điều này có nghĩa tập dữ liệu thực nghiệm này khá tương thích với thuật toán hồi quy bội và khả năng dự báo chính xác của mô hình là 98% trong khi rừng cây quyết định (RFT) có khả năng dự báo với độ chính xác là 94%.

Mặc dù độ chính xác của RFT là thấp hơn so với hồi quy bội LR, tuy nhiên nếu nhìn vào tham số Precision (độ chính xác của việc dự báo đối với từng lớp) thì rõ ràng Precision của RFT

đối với các lớp quyết định (no: 0; yes: 1) là cao hơn so với Precision của hồi quy bội (xem Hình 7). Điều này chỉ ra rằng việc dùng hồi quy bội sẽ có thể dẫn đến việc dự báo thiên lệch (lệch về phía số mẫu nhiều hơn trong tập dữ liệu), và khuynh hướng thiên lệch này sẽ dẫn đến dự báo sai nếu ở tập dữ liệu mà sự mất cân bằng là lớn. Trong nghiên cứu này không đề cập đến việc mất cân bằng dữ liệu trong học máy.



Hình 7. So sánh độ chính xác của từng lớp đối với 2 thuật toán.

4. KẾT LUẬN

Mục đích của quy trình lựa chọn nhà cung cấp là để có thể lựa chọn được nhà cung cấp phù hợp với yêu cầu doanh nghiệp, đảm bảo chất lượng chuỗi cung ứng. Việc lựa chọn nhà cung cấp là một trong những bước quan trọng trong hệ thống thông tin chuỗi cung ứng. Đầu vào của quy trình lựa chọn là tập các nhà cung cấp đã có và các nhà cung cấp tiềm năng. Mục đích của việc lựa chọn này là để đảm bảo sau quá trình lựa chọn, doanh nghiệp sẽ có một tập các nhà cung cấp có thể đáp ứng các yêu cầu do doanh nghiệp đề ra trên các góc độ như kỹ thuật, dịch vụ, tài chính và thương mại,...

Trong hệ thống chuỗi cung ứng, việc lựa chọn nhà cung cấp được thực hiện thông qua việc thỏa mãn một tập tiêu chí do doanh nghiệp đề ra trong một khoảng thời gian nào đó. Có nhiều phương pháp để có thể lựa chọn nhà cung cấp, ví dụ như kỹ thuật AHP hay DEA đã đề cập ở trên. Việc sử dụng các phương pháp thường được kết hợp với học máy để đạt được hiệu quả cao hơn trong việc

đưa ra quyết định lựa chọn nhà cung cấp cuối cùng từ nhà quản lý. Dựa trên các kinh nghiệm nghiên cứu đã được xuất bản được đề cập ở phần tổng quan, chúng tôi đề xuất việc ứng dụng học máy trong hệ thống thông tin chuỗi cung ứng để giảm cả số lượng nhà cung cấp và số lượng tiêu chí. Mặc dù nhược điểm của phương pháp phân lớp của học máy có hạn chế là kết quả nhận xác định là không đổi với tập dữ liệu đầu vào. Tuy nhiên xét ở góc độ hỗ trợ trong bài toán quản lý của việc đánh giá mức độ rủi ro như trong bài báo này, nó có thể hỗ trợ cho nhà quản lý dự báo mức độ chấp nhận rủi ro (no; yes) trong các trường hợp tương tự (với tập dữ liệu đầu vào tương tự).

Trong thực tế, số lượng nhà cung cấp ở chuỗi cung ứng của một doanh nghiệp không quá nhiều và các doanh nghiệp sử dụng phương pháp AHP hoặc thống kê để lựa chọn nhà cung cấp. Tuy nhiên, việc xây dựng tập dữ liệu huấn luyện cho hệ thống thông tin chuỗi cung ứng thường được lấy từ lịch sử giao dịch trước đó. Mỗi lần giao dịch được hiểu trên hệ thống là một quan sát cho tập huấn luyện để máy học. Chính vì vậy, hệ thống có thể đưa ra thông tin hỗ trợ đánh giá việc chấp nhận nhà cung cấp hay không khi có tập dữ liệu mới (như trong bài báo này là no, hay yes).

Trong các phương pháp được sử dụng với tập dữ liệu minh họa ở trên, kết quả dự báo đều đạt ở mức cao. Điều đó chứng tỏ sự hiệu quả của ứng dụng các kỹ thuật hiện đại trong bài toán thực tế của hoạt động doanh nghiệp, mà cụ thể là trong hệ thống thông tin chuỗi cung ứng. Trong hệ thống này, việc lựa chọn ra được tập các nhà cung ứng tốt sẽ giúp cho các bước tiếp theo hoạt động hiệu quả và đưa ra sản phẩm hoặc dịch vụ có chất lượng với giá cả hợp lý. Đây cũng là yếu tố để nâng sức cạnh tranh của doanh nghiệp trên thị trường. Hướng nghiên cứu tiếp theo sẽ tập trung vào khai thác các luật rút ra được trong quá trình xây dựng cây quyết định ở phương pháp RFT để chỉ ra các mối liên quan ở góc độ quản lý giữa các thuộc tính (biến) trong mô hình. Việc tổ chức

các kinh nghiệm của các chuyên gia thành một cơ sở tri thức trong hệ thống cũng hỗ trợ đắc lực cho việc ra quyết định của nhà quản lý chính xác và nhanh chóng hơn trong hệ thống. Việc tổ chức lại và bổ sung các thành phần của hệ thống thông tin để nó trở thành hệ thống thông tin hỗ trợ ra quyết định sẽ dành cho các nghiên cứu tiếp theo.

TÀI LIỆU THAM KHẢO

A. K. Kar, "A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network," *Journal of Computational Science*, vol. 6, pp. 23-33, 2015.

A.E. Cengiza, O. Aytekinb, I. Ozdemirb, H. Kusanb, A. Cabuk, A Multi-Criteria Decision Model for Construction Material Supplier Selection. doi: 10.1016/j.proeng.2017.07.202

A. Adam. "Exploring Machine Learning for Supplier Selection - A case study at Bufab Sweden". 2018. Tại website: <https://www.diva-portal.org/smash/get/diva2:1217866/FULLTEXT01.pdf>

B. Andrii., Vira S., Olha B. (2018). Information systems for supply chain management: uncertainties, risks and cyber security. ICTE in Transportation and Logistics 2018 (ICTE 2018). ScienceDirect. Available online at www.sciencedirect.com Procedia Computer Science 149 (2019) 65–70.

C. Lakshmanpriya, N. Sangeetha, and C. Lavanpriya, "Vendor selection in manufacturing industry using AHP and ANN," *The SIJ Transactions on Industrial, Financial & Business Management*, vol. 1, no. 1, pp. 29-34, 2013.

D. Wu, "Supplier selection: A hybrid model using DEA, decision tree and neural network," *Expert Systems with Applications*, vol. 36, no. 5, pp. 9105-9112, 2009

E. Bottani and A. Rizzi, "An adapted multi-criteria approach to suppliers and products selection—An application oriented to lead-time

reduction," *International Journal of Production Economics*, vol. 111, no. 2, pp. 763-781, 2008

IBM. (2023).

<https://www.ibm.com/topics/supply-chain-management>

M. Akarte, N. Surendra, B. Ravi, and N. Rangaraj, "Web based casting supplier evaluation using analytical hierarchy process," *Journal of the Operational Research Society*, vol. 52, no. 5, pp. 511-522, 2001.

T. N. M., Nong, & T. P. Ho, "Supplier selection criteria in Vietnam: A case study in textile and apparel industry". *Journal of Asian Business and Economic Studies*, 26(S02), Pp: 71–100. 2019.

S. Tang, N. Hakim, W. Khaksar, M. Ariffin, S. Sulaiman, and P. Pah, "A hybrid method using analytical hierarchical process and artificial neural network for supplier selection," *International Journal of Innovation, Management and Technology*, vol. 4, no. 1, pp. 109-111, 2013.

Tipalti, (2023). Vendor selection Process. Tại website: <https://tipalti.com/vendor-selection-process/>