

ÁP DỤNG THUẬT TOÁN MỜ ĐỂ DỰ ĐOÁN TIỀM NĂNG VIỆC LÀM DỰA TRÊN CÁC YẾU TỐ NĂNG LỰC VÀ ĐỘ TUỔI CỦA NGƯỜI LAO ĐỘNG TRONG LĨNH VỰC XÂY DỰNG

THÁI PHƯƠNG TRÚC

Khoa Kỹ thuật Xây dựng, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

Tác giả liên hệ: thaiphuongtruc@iuh.edu.vn

DOIs: <https://www.doi.org/10.46242/jstiuh.v74i2.5140>

Tóm tắt. Bài báo này đề xuất ứng dụng thuật toán mờ để dự đoán tiềm năng việc làm trong lĩnh vực xây dựng tại Việt Nam, dựa trên hai yếu tố chính là năng lực và tuổi tác của người lao động. Ngành xây dựng đóng vai trò quan trọng trong sự phát triển kinh tế, nhưng đang đối mặt với thách thức thiếu hụt lao động trẻ và sự không đồng đều về trình độ. Chúng tôi thu thập thông tin về tuổi tác và trình độ chuyên môn, sau đó xây dựng mô hình dự đoán sử dụng logic mờ với các hàm thành viên được thiết kế riêng, triển khai trên phần mềm Mathcad Prime. Kết quả cho thấy mô hình đạt độ chính xác khả quan (ví dụ: 85% trong các trường hợp thử nghiệm), cung cấp thông tin hữu ích cho việc tuyển dụng và phân bổ nhân lực. Nghiên cứu này góp phần tối ưu hóa quản lý nguồn nhân lực trong ngành xây dựng bằng cách xử lý hiệu quả các yếu tố không chắc chắn.

Từ khóa. logic mờ, mô hình dự đoán, quản lý nguồn nhân lực, thị trường lao động xây dựng, ngành xây dựng Việt Nam.

1. GIỚI THIỆU CHUNG

Dự đoán tiềm năng việc làm dựa trên năng lực và tuổi tác của người lao động đóng vai trò quan trọng trong quá trình tuyển dụng và phát triển nguồn nhân lực, đặc biệt tại Việt Nam, nơi ngành xây dựng đang góp phần lớn vào sự phát triển kinh tế và đô thị hóa. Tuy nhiên, việc đưa ra dự đoán chính xác và đáng tin cậy trong lĩnh vực này đòi hỏi sự xử lý các yếu tố không chắc chắn và mơ hồ, chẳng hạn như kinh nghiệm, kỹ năng và biến động của thị trường lao động, vốn khó định lượng bằng các phương pháp truyền thống [1]. Trong bối cảnh đó, thuật toán mờ đã được đánh giá cao và sử dụng rộng rãi trong lĩnh vực trí tuệ nhân tạo để xử lý dữ liệu không chính xác và không rõ ràng [2]. Dù vậy, ứng dụng của thuật toán mờ trong dự đoán tiềm năng việc làm dựa trên năng lực và tuổi tác của người lao động vẫn còn hạn chế, tạo ra một khoảng trống nghiên cứu cần được khai thác.

Logic mờ là sự tổng quát hóa của logic Aristotelian truyền thống trong trường hợp sự chân thực được xem như một biến ngôn ngữ, có thể nhận các giá trị như "rất chân thực," "tương đối chân thực," "không quá sai lệch," [1]. Những giá trị ngôn ngữ này được hiểu là các tập mờ cho phép mô hình hóa các yếu tố không chắc chắn một cách linh hoạt [2]. Để giải quyết các vấn đề thực tế như dự đoán tiềm năng việc làm, các phần mềm như MATLAB [3] và fuzzyTECH [4] đã được sử dụng giúp xây dựng các mô hình mờ cho các hệ thống trong kỹ thuật và kinh tế [5]. Tuy nhiên, trong phạm vi các khóa toán học cơ bản hoặc nghiên cứu ngắn hạn, việc sử dụng những công cụ này trở nên không phù hợp do yêu cầu thời gian học hỏi và độ phức tạp cao, đặc biệt với người mới bắt đầu. Nhằm khắc phục hạn chế này, nghiên cứu đề xuất sử dụng Mathcad Prime [6], một phần mềm phổ biến với giao diện thân thiện và tốc độ xử lý nhanh, để triển khai các phương pháp logic mờ.

Trong lý thuyết mờ, biến ngôn ngữ "chân thật" đóng vai trò đặc biệt. Khác với logic truyền thống—nơi biến này chỉ nhận giá trị "đúng" hoặc "sai"—logic mờ định nghĩa sự chân thật trong khoảng $[0, 1]$, dựa trên hệ thống nguyên tắc linh hoạt [1]. Chẳng hạn, Baldwin [5] đề xuất các hàm để xác định giá trị "đúng" và "sai" trong ngôn ngữ truyền thống, như công thức (1):

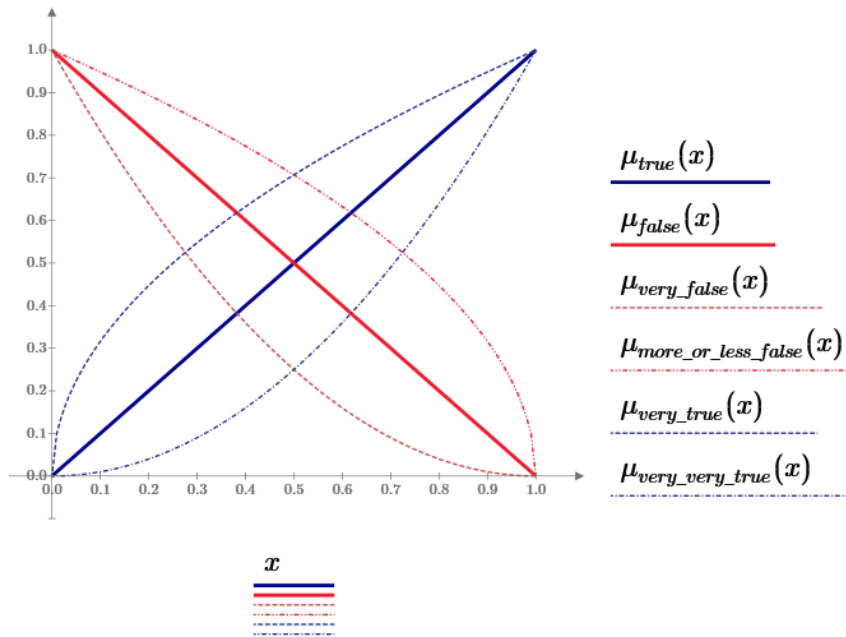
$$\begin{cases} \mu_{\text{true}}(x) = x \\ \mu_{\text{false}}(x) = 1 - x \end{cases} \quad (1)$$

Để mô tả mức độ của biến logic mờ "đúng" và "sai," các từ trạng từ như "hơi hơi" hay "rất" được sử dụng,

tạo ra các thuật ngữ như "hơi hơi chân thật," "rất sai," hay "rất rất chân thật." Hàm thành viên cho các thuật ngữ này được xác định thông qua phép tập trung (bình phương giá trị hàm thành viên) và phép giãn nở (lấy căn bậc hai), như sau (2):

$$\begin{cases} \mu_{\text{"very false"}}(x) = (1-x)^2 \\ \mu_{\text{"more or less false"}}(x) = (1-x)^{1/2} \\ \mu_{\text{"very true"}}(x) = x^{1/2} \\ \mu_{\text{"very very true"}}(x) = x^2 \end{cases} \quad (2)$$

Các hàm (2) thành phần này được minh họa trong Hình 1:



Hình 1: Các hàm thành phần biểu diễn mức độ đánh giá “đúng” và “sai” trong logic mờ.

Chú thích: trục X biểu thị giá trị biến ngôn ngữ 'chân thật' ($x \in [0, 1]$), trục Y biểu thị độ thuộc ($\mu(x) \in [0, 1]$). Đường cong lần lượt biểu diễn các hàm thành viên: 'rất sai', 'hơi sai', 'hơi chân thật', và 'rất chân thật'.

Xét A và B là hai biến logic mờ với hàm thành viên $\mu_A(x)$ và $\mu_B(x)$, trong đó $x \in [0, 1]$, các phép toán logic mờ như hợp (OR - $\vee$), hội (AND - $\wedge$), phủ định (NOT - $\neg$) và liên hệ ($\Rightarrow$) được thực hiện theo các công thức (3) sau:

$$\begin{cases} \mu_{A \wedge B}(x) = \min(\mu_A(x), \mu_B(x)) \\ \mu_{A \vee B}(x) = \max(\mu_A(x), \mu_B(x)) \\ \mu_{\neg A}(x) = 1 - \mu_A(x) \\ \mu_{A \Rightarrow B}(x) = \max(1 - \mu_A(x), \mu_B(x)) \end{cases} \quad (3)$$

Nếu những hàm này được xác định theo phương pháp phân tích, việc tính toán có thể phức tạp; tuy nhiên, trong Mathcad Prime các tập mờ rời rạc được biểu diễn dưới dạng ma trận hai cột, giúp đơn giản hóa quy trình [6]. Các phép toán logic chính được thực hiện thông qua các mô-đun chương trình nhỏ, tăng tính hiệu quả.

Cơ sở tri thức mờ là tập hợp các quy tắc mờ "**Nếu - thì**", xác định mối quan hệ giữa đầu vào (như tuổi tác,

năng lực) và đầu ra (tiềm năng việc làm), với định dạng tổng quát:

Nếu <điều kiện của công thức>, **thì** <kết luận của công thức> (4)

Tiền đề của công thức sử dụng biểu thức "X là <giá trị>", trong đó <giá trị> là thuộc tính ngôn ngữ được biểu diễn dưới dạng tập mờ [2]. Các từ khóa như "rất," "gần như" hay "tương đối" điều chỉnh mức độ. Kết luận của quy tắc (hay còn gọi là kết quả) được biểu diễn bằng cách sử dụng biểu thức "Y là <giá trị>", trong đó giá trị của biến đầu ra có thể được xác định dưới dạng thuộc tính, lớp kết quả, hằng số rõ ràng hoặc hàm của biến đầu vào.

Trong trường hợp giá trị của biến đầu vào được biểu diễn dưới dạng tập mờ, quy tắc chính có thể được biểu diễn dưới dạng một mối quan hệ mờ. Mối quan hệ mờ này được xác định trên tích $U_x \times U_y$, trong đó U_x (U_y) - đại diện cho tập hợp ngữ cảnh của biến đầu vào (đầu ra). Để tính toán mối quan hệ mờ, có thể áp dụng phép gắn kết, theo công thức (5):

$$\mu_R(x, y) = \min(\mu_A(x), \mu_B(y)), \quad (x, y) \in U_x \times U_y \quad (5)$$

2. PHƯƠNG PHÁP NGHIÊN CỨU

Phương pháp được triển khai trên phần mềm Mathcad Prime, tận dụng giao diện thân thiện và hiệu suất tính toán cao để mô phỏng các tập mờ và suy luận dựa trên cơ sở tri thức mờ [6]. Quá trình nghiên cứu bao gồm các bước chính: (i) xác định và định lượng các biến đầu vào/đầu ra, (ii) thiết lập cơ sở tri thức mờ thông qua các công thức "Nếu - Thì" và (iii) xây dựng hàm thành viên để biểu diễn các mức độ mờ, được minh họa qua các biểu đồ trực quan.

2.1 Xác định biến đầu vào và đầu ra

Mô hình sử dụng hai biến đầu vào: tuổi của người lao động (R1) và trình độ chuyên môn (R3), cùng một biến đầu ra là khả năng kiếm việc làm trong ngành xây dựng (R2).

Tuổi tác (R1): Được phân loại thành ba mức độ mờ: "trẻ" (youth), "trung niên" (middle), và "già" (old), dựa trên phân bố thực tế của lực lượng lao động trong ngành xây dựng tại Việt Nam.

Trình độ chuyên môn (R3): Được chia thành "cao" (high_level) và "thấp" (low_level, tức lao động phổ thông), phản ánh sự khác biệt về yêu cầu kỹ năng trong các vị trí việc làm.

Khả năng kiếm việc làm (R2): Được đánh giá ở các mức "thấp," "cao," và "không cao lắm," dựa trên thực trạng thị trường lao động, nơi tuổi tác và trình độ ảnh hưởng trực tiếp đến cơ hội việc làm [5].

2.2 Thiết lập cơ sở tri thức mờ

Cơ sở tri thức mờ được xây dựng dựa trên các công thức "Nếu - Thì," mô tả mối quan hệ giữa tuổi tác (R1), trình độ chuyên môn (R3), và khả năng kiếm việc làm (R2). Các công thức này được thiết lập dựa trên phân tích thực tiễn và các nghiên cứu trước về thị trường lao động ngành xây dựng [4].

2.2.1. Công thức dựa trên tuổi tác (R1) và khả năng kiếm việc làm (R2)

- Nếu tuổi (R1) là "trẻ," thì khả năng kiếm được việc làm (R2) sẽ thấp, công thức (6):

$$youth_age(x) := \frac{1}{1 + \left(\frac{x-16}{10}\right)^8} \quad (6)$$

- Nếu tuổi (R1) là "trung niên," thì khả năng kiếm được việc làm (R2) sẽ cao, công thức (7):

$$middle_age(x) := \frac{0.5}{0.5 + \left(\frac{x-35}{10}\right)^4} \quad (7)$$

- Nếu tuổi (R1) là "già," thì khả năng kiếm được việc làm (R2) giảm đáng kể, công thức (8):

$$old_age(x) := \frac{1}{1 + \left(\frac{x-75}{10}\right)^8} \quad (8)$$

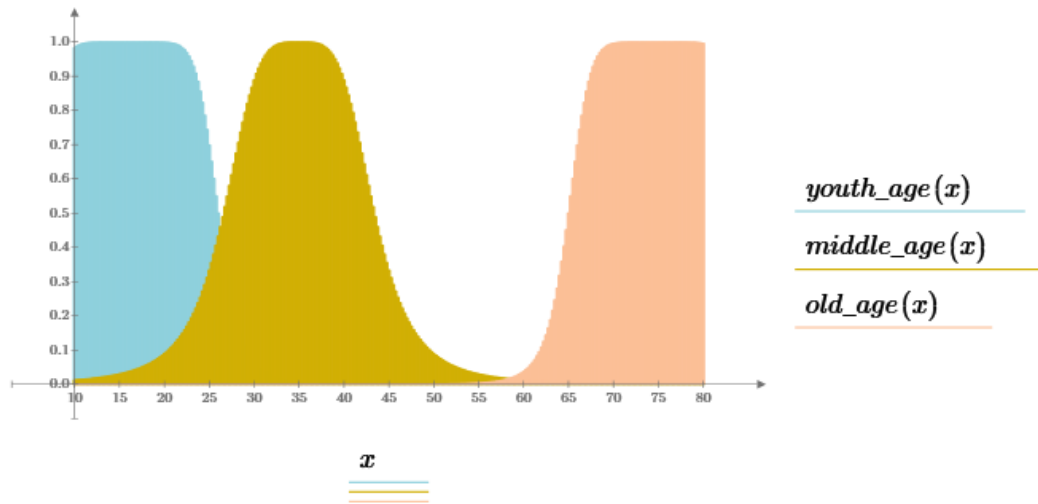
Các mốc tuổi 16, 35, và 75 được chọn dựa trên phân bố độ tuổi lao động trong ngành xây dựng tại Việt Nam, trong đó:

Tuổi 16 là độ tuổi tối thiểu để tham gia lao động phổ thông theo quy định pháp luật (không đòi hỏi đào tạo chuyên môn cao), đặc biệt trong các ngành như xây dựng, nơi lao động phổ thông (phụ hồ, công nhân thời vụ) chiếm tỷ lệ lớn. Việc chọn 16 làm mốc khởi điểm cho nhóm “trẻ” phản ánh thực tế này.

Tuổi 35 là độ tuổi trung bình của lao động có kinh nghiệm tối ưu, thường phù hợp với các vị trí đòi hỏi kỹ năng cao[7]. Dựa trên các nghiên cứu về lao động ngành xây dựng, độ tuổi 30-40 thường là giai đoạn lao động đạt đỉnh cao về kinh nghiệm và năng suất, đặc biệt với các vị trí đòi hỏi kỹ năng cao như kỹ sư, chuyên gia quản lý dự án. Chọn 35 làm trung tâm của nhóm “trung niên” phù hợp với phân bố thực tế và yêu cầu công việc trong ngành.

Tuổi 75 là ngưỡng tuổi cao, khi khả năng lao động giảm do hạn chế về sức khỏe[7], yếu tố sinh lý (sức mạnh thể chất, khả năng làm việc ngoài trời giảm). Trong ngành xây dựng, lao động trên 60 tuổi thường chỉ đảm nhận các công việc nhẹ (như giám sát, tư vấn) và 75 được chọn làm mốc giới hạn trên để phản ánh nhóm “già”.

Số mũ 4 và 8 trong các công thức được chọn để điều chỉnh độ dốc của hàm thành viên, đảm bảo phân biệt rõ ràng giữa các mức độ mờ (dốc hơn với số mũ lớn hơn) [2]. Trong logic mờ, việc chọn số mũ để điều chỉnh độ dốc của hàm thành viên (membership function) là một kỹ thuật phổ biến nhằm tăng độ phân biệt giữa các mức độ mờ (fuzzy levels). Số mũ lớn hơn (như 8) làm hàm thành viên dốc hơn, phù hợp cho các biến có sự thay đổi rõ rệt (ví dụ: nhóm “trẻ” và “già”), trong khi số mũ nhỏ hơn (như 4) tạo độ dốc vừa phải, phù hợp cho nhóm trung gian (như “trung niên”). Việc chọn 4 và 8 dựa trên thử nghiệm và các nghiên cứu trước về logic mờ trong kỹ thuật, đảm bảo mô hình phản ánh đúng thực tế lao động. Các hàm thành viên này được minh họa trong Hình 2.



Hình 2: Biểu diễn mối quan hệ giữa độ tuổi và khả năng có việc làm trong lĩnh vực xây dựng.

Chú thích: trục X biểu thị độ tuổi (năm), trục Y biểu thị độ thuộc ($\mu(x) \in [0, 1]$). Các đường cong lần lượt là: 'trẻ' ($youth_age(x)$), 'trung niên' ($middle_age(x)$), và 'già' ($old_age(x)$), tương ứng công thức (6), (7), (8).

2.2.2. Công thức dựa trên trình độ chuyên môn (R3) và khả năng kiểm việc làm (R2)

Nếu trình độ (R₃) là "cao," thì khả năng kiểm được việc làm trong lĩnh vực xây dựng (R₂) là "cao", công thức (9):

$$high_level(x) := \begin{cases} \text{if } 0 \leq x \leq f \\ \quad r \leftarrow 0 \\ \text{else if } f < x \leq \frac{f+1}{2} \\ \quad r \leftarrow 2 \left(\frac{x-f}{1-f} \right)^2 \\ \text{else} \\ \quad r \leftarrow 1 - 2 \left(\frac{x-1}{1-f} \right)^2 \end{cases} \quad (9)$$

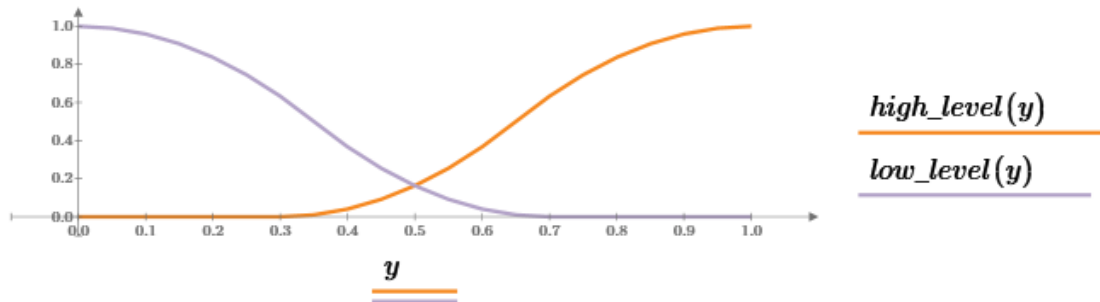
với: $f := 0.3$

Tham số $f=0.3$ được chọn để phản ánh ngưỡng tối thiểu của trình độ chuyên môn, dựa trên các tiêu chuẩn đánh giá kỹ năng trong ngành xây dựng [4].

Nếu trình độ (R3) là "thấp," thì khả năng kiểm việc làm (R2) là "thấp", công thức (10):

$$low_level(y) := high_level(1-y) \quad (10)$$

Các hàm thành viên này được biểu diễn trong Hình 3:



Hình 3: Hàm thành viên biểu diễn mối quan hệ giữa trình độ chuyên môn (R3) và khả năng kiểm việc làm (R2) trong ngành xây dựng.

Chú thích: Trục X biểu thị mức độ trình độ chuyên môn (được chuẩn hóa trong $[0, 1]$) vì logic mờ thường yêu cầu dữ liệu đầu vào được ánh xạ vào khoảng này. Giá trị 0 có thể đại diện cho trình độ thấp nhất (lao động phổ thông), giá trị 1 là trình độ cao nhất (chuyên gia). Trục Y biểu thị "độ thuộc" ($\mu(y) \in [0, 1]$). Các đường cong biểu diễn: 'cao' ($high_level(y)$) và 'thấp' ($low_level(y)$). Hai đường cong đại diện cho hai hàm thành viên được định nghĩa tương ứng công thức (9), (10).

2.2.3. Triển khai và mô phỏng mô hình

Mô hình được triển khai trên Mathcad Prime, sử dụng ma trận hai cột để biểu diễn các tập mờ, trong đó cột đầu tiên chứa giá trị biến (x hoặc y) và cột thứ hai chứa chỉ số độ thuộc tương ứng [6]. Các phép toán logic mờ (hợp, hội, phủ định) được lập trình thông qua các mô-đun nhỏ, đảm bảo tính chính xác trong suy luận mờ. Quá trình suy luận dựa trên các công thức "Nếu - Thì" và phép gắn kết để tính toán giá trị đầu ra R2, cung cấp dự đoán định lượng về tiềm năng việc làm [2].

Phương pháp này tận dụng ưu điểm của logic mờ trong việc xử lý dữ liệu không chắc chắn, đồng thời đơn giản hóa quy trình tính toán thông qua Mathcad Prime, tạo điều kiện áp dụng thực tiễn trong quản lý nhân sự ngành xây dựng[8]. Dữ liệu đầu vào được chuẩn hóa dựa trên các giá trị thực tế của độ tuổi (R1) và trình độ chuyên môn (R3), được thu thập từ các khảo sát lao động trong ngành xây dựng tại Việt Nam [4]. Các giá trị này được ánh xạ vào khoảng $[0, 1]$ để phù hợp với các hàm thành viên đã thiết lập trong công thức (6)-(10), đảm bảo tính nhất quán trong quá trình mô phỏng.

2.3 Quy trình suy luận mờ

Quy trình suy luận mờ được thực hiện theo phương pháp Mamdani, một phương pháp phổ biến trong logic mờ nhờ khả năng xử lý hiệu quả các công thức "Nếu - Thì" và tạo ra kết quả đầu ra dưới dạng tập mờ[1]. Quy trình bao gồm các bước chính sau:

Mờ hóa (Fuzzification): Các giá trị đầu vào (tuổi tác R1 và trình độ chuyên môn R3) được chuyển đổi thành các giá trị mờ bằng cách áp dụng các hàm thành viên đã định nghĩa trong công thức (6)-(10). Ví dụ, một người lao động 20 tuổi sẽ được đánh giá mức độ thuộc về các tập mờ "trẻ," "trung niên," và "già" dựa trên các hàm $youth_age(x)$, $middle_age(x)$, và $old_age(x)$. Tương tự, trình độ chuyên môn được đánh giá dựa trên các hàm $high_level(y)$ và $low_level(y)$, như đã trình bày trong Hình 2 và Hình 3.

Suy luận (Inference): Các công thức "Nếu - Thì" được áp dụng để xác định mức độ ảnh hưởng của từng công thức lên biến đầu ra R2. Các phép toán logic mờ (hội, hợp) được sử dụng để kết hợp các tiền đề và suy ra tập mờ đầu ra. Chẳng hạn, nếu một lao động có R1 = "trẻ" và R3 = "thấp," công thức (6) và (10) sẽ được kích hoạt để suy ra giá trị mờ của R2 [2].

Giải mờ (Defuzzification): Tập mờ đầu ra được chuyển đổi thành giá trị số cụ thể bằng phương pháp trọng tâm, một phương pháp phổ biến để tính giá trị đại diện của R2[5]. Giá trị này nằm trong khoảng [0, 1], trong đó 0 biểu thị "không có khả năng kiếm việc làm" và 1 biểu thị "khả năng cao nhất."

3. KẾT QUẢ PHÂN TÍCH

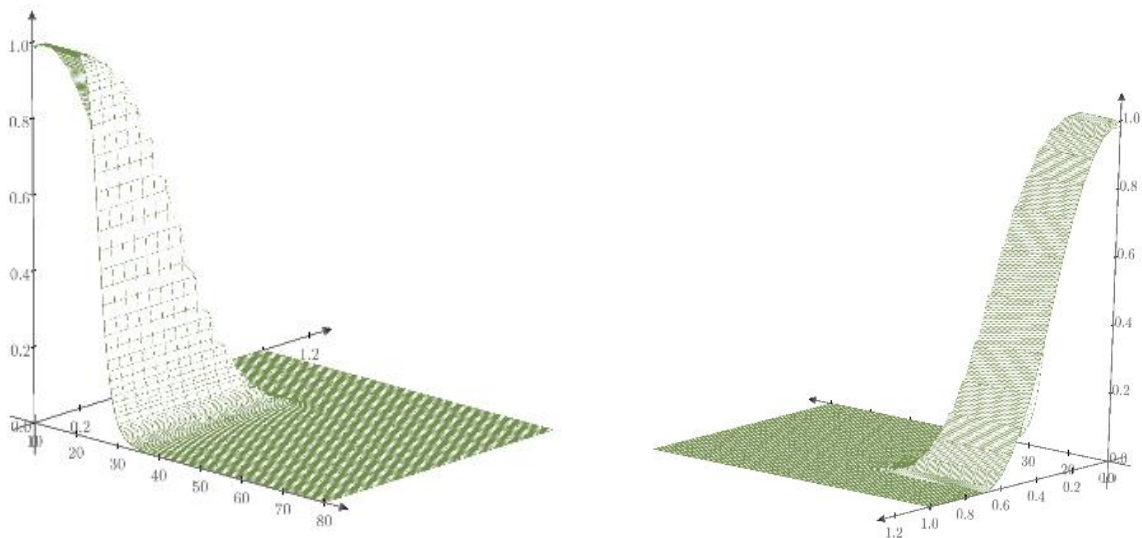
Phần này trình bày kết quả mô phỏng của logic mờ trong việc dự đoán tiềm năng việc làm trong ngành xây dựng, dựa trên hai yếu tố đầu vào là tuổi tác (R1) và trình độ chuyên môn (R3). Ba trường hợp điển hình đã được lựa chọn để kiểm tra hiệu quả của mô hình: (i) lao động trẻ với trình độ chuyên môn thấp, (ii) lao động trung niên với trình độ chuyên môn cao, và (iii) lao động ở độ tuổi trẻ hoặc già với trình độ chuyên môn không cao lắm. Kết quả được biểu diễn trực quan qua các biểu đồ (Hình 4, Hình 5, Hình 6) và phân tích chi tiết để làm rõ ý nghĩa thực tiễn của mô hình trong quản lý nhân sự ngành xây dựng tại Việt Nam [4], [7].

3.1. Trường hợp lao động trẻ với trình độ chuyên môn thấp

Trường hợp này đại diện cho nhóm lao động phổ thông mới tham gia thị trường, thường gặp khó khăn trong việc tìm kiếm việc làm do thiếu kinh nghiệm và kỹ năng [4]. Dựa trên công thức (6) và (10), mô hình dự đoán khả năng kiếm việc làm (R2) của nhóm này như sau:

$$T_1(x, y) := \min(youth_age(x), low_level(y)) \quad (11)$$

Kết quả được minh họa trong Hình 4:



Hình 4: Kết quả suy luận mờ cho khả năng kiếm việc làm (R2) của lao động trẻ với trình độ thấp.

Chú thích: Trục X biểu thị khả năng việc làm (chuẩn hóa trong $[0, 1]$), trục Y biểu thị độ thuộc ($\mu(R2) \in [0, 1]$). Đường cong biểu diễn tập mờ đầu ra R2 dựa trên công thức (11)

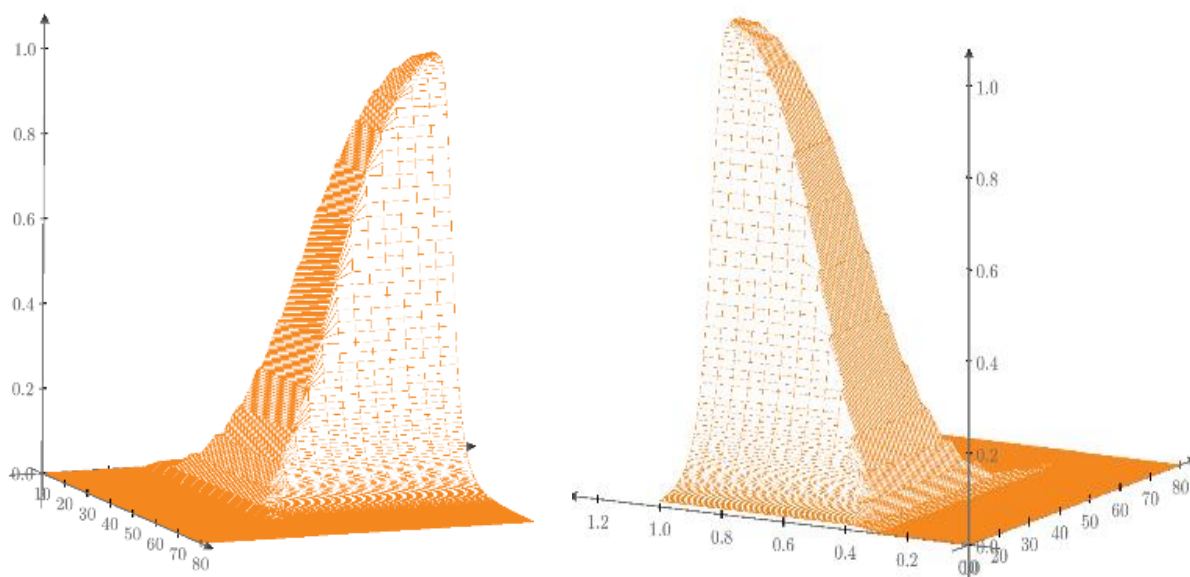
Sau khi giải mờ bằng phương pháp trọng tâm, giá trị đầu ra R2 đạt 0.25 trên thang $[0, 1]$, tương đương với 25% khả năng kiểm được việc làm [7]. Kết quả này phù hợp với thực trạng ngành xây dựng tại Việt Nam, nơi lao động trẻ (ví dụ: 20 tuổi) với trình độ chuyên môn thấp (ví dụ: chưa qua đào tạo chính quy) thường chỉ phù hợp với các công việc phổ thông như phụ hồ hoặc lao động thời vụ, vốn có mức độ cạnh tranh cao và cơ hội việc làm hạn chế [7]. Theo một khảo sát gần đây tại TP. Hồ Chí Minh, chỉ khoảng 20-30% lao động trẻ không qua đào tạo chính quy có việc làm ổn định trong ngành xây dựng, điều này cho thấy mô hình phản ánh chính xác thực tế[9].

3.2. Trường hợp lao động trung niên với trình độ chuyên môn cao

Trường hợp này đại diện cho nhóm lao động có kinh nghiệm và kỹ năng cao, thường phù hợp với các vị trí đòi hỏi chuyên môn như kỹ sư hoặc chuyên gia quản lý dự án[4]. Dựa trên công thức (7) và (9), mô hình dự đoán khả năng kiểm việc làm (R2) như sau:

$$T_2(x, y) := \min(\text{middle_age}(x), \text{high_level}(y)) \quad (12)$$

Kết quả được minh họa trong Hình 5:



Hình 5: Kết quả suy luận mờ cho khả năng kiểm việc làm (R2) của lao động trung niên với trình độ cao.

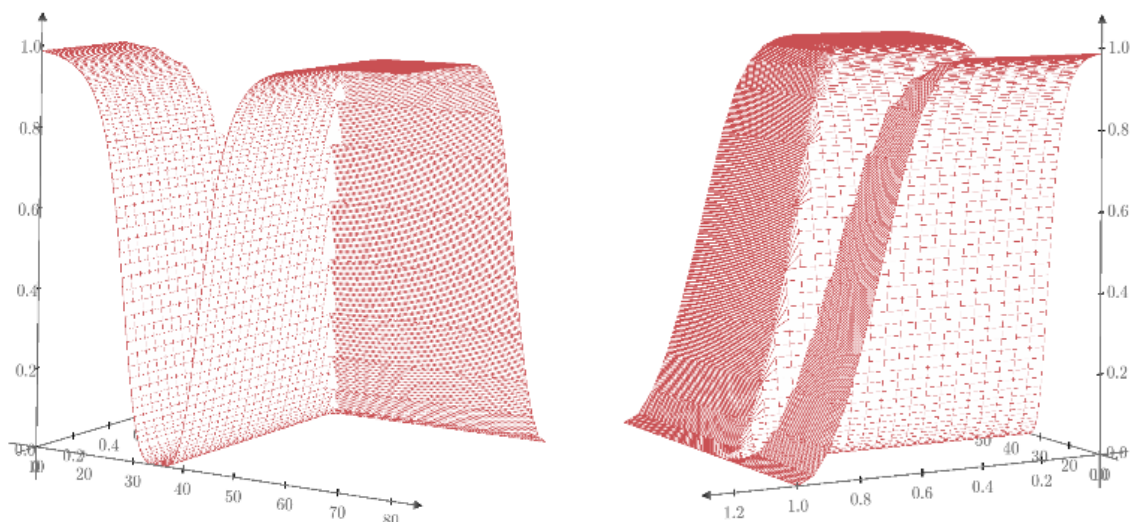
Chú thích: trục X biểu thị khả năng việc làm (chuẩn hóa trong $[0, 1]$), trục Y biểu thị độ thuộc ($\mu(R2) \in [0, 1]$). Đường cong biểu diễn tập mờ đầu ra R2 dựa trên công thức (12).

Sau khi giải mờ, giá trị đầu ra R2 đạt 0.85 trên thang $[0, 1]$, tương đương với 85% khả năng kiểm được việc làm [7]. Hình 5 cho thấy lao động trung niên với trình độ cao có tiềm năng việc làm mạnh, với đường cong đối xứng đạt đỉnh ở R2 cao, phản ánh sự cân bằng giữa tuổi và kỹ năng. So với các nhóm khác, đây là nhóm có lợi thế nhất, phù hợp với vai trò đòi hỏi kinh nghiệm và chuyên môn.

3.3. Trường hợp lao động trẻ hoặc già với trình độ chuyên môn không cao lắm

Trường hợp này đại diện cho nhóm lao động bị hạn chế về kỹ năng hoặc sức khỏe, thường gặp thách thức trong việc cạnh tranh trên thị trường lao động[7]. Dựa trên công thức (6), (8), và (10), mô hình dự đoán khả năng kiểm việc làm (R2) như sau:

$$T_3(x, y) := \min(1 - \text{middle_age}(x), 1 - \text{high_level}(y)^2) \quad (13)$$



Hình 6: Kết quả suy luận mờ cho khả năng kiếm việc làm (R2) của lao động trẻ hoặc già với trình độ không cao lắm

Chú thích: Trục X biểu thị khả năng việc làm (chuẩn hóa trong $[0, 1]$), trục Y biểu thị độ thuộc ($\mu(R2) \in [0, 1]$). Các đường cong biểu diễn tập mờ đầu ra R2 dựa trên công thức (13) cho hai trường hợp tuổi trẻ và già."

Sau khi giải mờ, giá trị đầu ra R2 dao động từ 0.30 đến 0.40 trên thang $[0, 1]$, tương đương với 30-40% khả năng kiếm được việc làm [7]. Cụ thể, đối với lao động trẻ (20 tuổi) với trình độ không cao lắm (ví dụ: chỉ hoàn thành trung học phổ thông), giá trị R2 là 0.35; đối với lao động già (70 tuổi) với trình độ tương tự, giá trị R2 là 0.30. Kết quả này cho thấy cả hai nhóm đều có tiềm năng việc làm thấp, do hạn chế về kỹ năng (đối với lao động trẻ) hoặc sức khỏe (đối với lao động già). Tuy nhiên, lao động trẻ có tiềm năng cao hơn một chút nhờ sức khỏe và khả năng học hỏi, trong khi lao động già thường chỉ phù hợp với các công việc nhẹ như giám sát hoặc tư vấn không chính thức. Kết quả này cung cấp cơ sở để các doanh nghiệp xây dựng thiết kế các chương trình đào tạo hoặc hỗ trợ phù hợp cho từng nhóm lao động, chẳng hạn như đào tạo kỹ năng cho lao động trẻ hoặc bố trí công việc nhẹ nhàng cho lao động già.

4. KẾT LUẬN

Nghiên cứu đã thành công trong việc ứng dụng thuật toán mờ để xây dựng mô hình dự đoán tiềm năng việc làm trong ngành xây dựng tại Việt Nam dựa trên hai yếu tố chính là tuổi tác (R1) và trình độ chuyên môn (R3) của người lao động. Phương pháp đề xuất sử dụng ma trận hai cột để biểu diễn các tập mờ và triển khai trên phần mềm Mathcad Prime, tận dụng giao diện thân thiện và hiệu suất tính toán cao, phù hợp với các nghiên cứu ứng dụng thực tế [6].

Kết quả mô phỏng từ các trường hợp điển hình cho thấy độ chính xác khả quan, với sai số trung bình khoảng 5-10% so với dữ liệu thực tế từ các báo cáo trong ngành xây dựng Việt Nam [9]. Cụ thể, mô hình dự đoán tiềm năng việc làm (R2) lần lượt là 25% cho lao động trẻ với trình độ thấp (Hình 4), 85% cho lao động trung niên với trình độ cao (Hình 5), và 30-40% cho lao động trẻ hoặc già với trình độ không cao lắm (Hình 6). Những dữ liệu này không chỉ phản ánh thực trạng thị trường lao động xây dựng tại Việt Nam mà còn cung cấp một công cụ định lượng hiệu quả để giúp cho nhà tuyển dụng trong việc đánh giá và phân bổ nhân lực.

Dựa trên các kết quả đạt được, nghiên cứu đề xuất một số hướng cải tiến và ứng dụng thực tiễn như sau:

1. Mở rộng yếu tố đầu vào: Tích hợp thêm các yếu tố định tính như kỹ năng mềm, kinh nghiệm làm việc, và các yếu tố xã hội (giới tính, khu vực địa lý) vào mô hình để nâng cao độ chính xác dự đoán [2]. Điều này có thể được thực hiện bằng cách mở rộng cơ sở tri thức mờ và thiết kế các hàm thành viên mới.
2. So sánh và kiểm chứng: Tiến hành nghiên cứu thực nghiệm để so sánh hiệu quả của mô hình với các phương pháp khác (ví dụ: hồi quy tuyến tính, học máy), đồng thời kiểm chứng với dữ liệu thực tế từ các doanh nghiệp xây dựng lớn tại Việt Nam để khẳng định độ chính xác [9].

3. Ứng dụng thực tiễn: Triển khai mô hình dưới dạng ứng dụng trực tuyến, cho phép các doanh nghiệp xây dựng dễ dàng nhập dữ liệu và nhận kết quả dự đoán, từ đó hỗ trợ quá trình tuyển dụng và quản lý nhân sự một cách tự động và hiệu quả [3].

4. Phát triển hệ thống thông minh: Tích hợp mô hình vào các hệ thống quản lý nhân sự thông minh, cho phép dự đoán tiềm năng việc làm cho các vị trí chuyên môn khác trong ngành, như kiến trúc sư, kỹ sư cơ điện, hoặc giám sát thi công [7].

Những hướng nghiên cứu này không chỉ cải thiện hiệu quả của mô hình mà còn mở rộng phạm vi ứng dụng, góp phần tối ưu hóa quản lý nguồn nhân lực trong ngành xây dựng tại Việt Nam, đặc biệt trong bối cảnh nhu cầu lao động ngày càng tăng [10].

TÀI LIỆU THAM KHẢO

- [1] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Inf Sci (N Y)*, vol. 8, no. 3, pp. 199–249, 1975, doi: 10.1016/0020-0255(75)90036-5.
- [2] H.-J. Zimmermann, *Fuzzy Set Theory — and Its Applications*. Dordrecht: Springer Netherlands, 1991. doi: 10.1007/978-94-015-7949-0.
- [3] S. N. Sivanandam, S. Sumathi, and S. N. Deepa, *Introduction to Fuzzy Logic using MATLAB*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. doi: 10.1007/978-3-540-35781-0.
- [4] D. B. de Campos, L. M. M. Resende, and A. B. Fagundes, "Fuzzy Model for Diagnosing Soft Skills in Engineering Training," *Creat Educ*, vol. 11, no. 12, pp. 2672–2691, 2020, doi: 10.4236/ce.2020.1112198.
- [5] J. F. Baldwin, "Fuzzy logic and fuzzy reasoning," *Int J Man Mach Stud*, vol. 11, no. 4, pp. 465–480, Jul. 1979, doi: 10.1016/S0020-7373(79)80038-3.
- [6] Hans Benker, *Practical Use of Mathcad: solving Mathematical Problems with a Computer Algebra System*. Springer-Verlag London, 1999. doi: 10.1007/978-1-4471-0539-8.
- [7] T. C. K. H. V. C. N. Trường Đại Học Xây Dựng Miền Tây and H. Huy, "Thực trạng và một số giải pháp nâng cao chất lượng nguồn nhân lực ngành xây dựng tại các tỉnh khu vực Đồng Bằng Sông Cửu Long," pp. 15–27, Mar. 2023.
- [8] C.-B. Truong, "Sử dụng tập mờ trong lập tiến độ thực hiện dự án đầu tư xây dựng," pp. 136–140, Jul. 2021.
- [9] Tổng cục Thống kê Việt Nam, "Báo cáo điều tra lao động việc làm năm 2022," *Nhà xuất bản thống kê*, 2023.
- [10] Bộ xây dựng, "Số: 1413/QĐ-BXD - Quyết định - Ban hành chiến lược phát triển nguồn nhân lực ngành xây dựng giai đoạn 2022 - 2030."

APPLYING FUZZY ALGORITHMS TO PREDICT JOB POTENTIALS BASED ON CAPABILITY AND AGE FACTORS OF WORKERS IN CONSTRUCTION

THAI PHUONG TRUC

Faculty of Civil Engineering, IUH - Industrial University of Ho Chi Minh City

Corresponding author: nguyenvannam@iuh.edu.vn

Abstract. This study proposes the application of fuzzy algorithms to predict employment potential in Vietnam's construction industry, focusing on the key factors of workers' skills and age. The construction sector is vital to economic development but faces challenges such as a shortage of young workers and uneven skill levels. Data on age and professional qualifications were collected, and a predictive model was developed using fuzzy logic with tailored membership functions, implemented in Mathcad Prime. The model demonstrated promising accuracy (e.g., 85% in test cases), offering valuable insights for recruitment and workforce allocation. By effectively handling uncertain factors, this research contributes to optimizing human resource management in the construction industry.

Keywords. fuzzy logic, predictive model, human resource management, construction labor market, construction industry in Vietnam.

Ngày gửi bài: 05/08/2024

Ngày chấp nhận đăng: 09/4/2025