

CẢI THIẾN MÔ HÌNH MỜ HƯỚNG DỮ LIỆU TSK VỚI TRI THỨC TIÊN NGHIỆM

Nguyễn Đức Hiền¹, Lê Mạnh Thanh²

Trường Đại học Khoa học, Đại học Huế

¹ Email: hiencit@gmail.com

² Email: lmthanh@hueuni.edu.vn

Ngày nhận bài: 26/12/2017; ngày hoàn thành phản biện: 9/01/2018; ngày duyệt đăng: 8/6/2018

TÓM TẮT

Bằng cách thỏa mãn các điều kiện đề đồng nhất các hàm đầu ra của mô hình mờ TSK (Takagi – Sugeno – Kang) và máy học véc-tơ hỗ trợ hồi quy, chúng ta có thể xây dựng một thuật toán cho phép trích xuất mô hình mờ TSK từ máy học véc-tơ hỗ trợ. Những nghiên cứu trước đây cho thấy mô hình mờ trích xuất được vẫn tồn tại những hạn chế chất định. Bài báo này đề xuất một mô hình mờ dựa trên sự tích hợp tri thức tiên nghiệm với mô hình mờ TSK trích xuất từ máy học véc-tơ hỗ trợ cho bài toán dự báo hồi quy. Mô hình này tiếp cận theo hướng trích xuất các tập luật mờ “có thể diễn dịch được” cho hệ dự báo dựa trên sự kết hợp các mô hình máy học thống kê. Bằng cách tích hợp tri thức tiên nghiệm với mô hình mờ dựa trên SVM (Support Vector Machine), hệ thống luật mờ trích xuất được sẽ giảm tính phức tạp. Hiệu quả của giải pháp đề xuất được đánh giá thông qua các kết quả thực nghiệm và có sự so sánh với một số mô hình khác.

Từ khóa: mô hình mờ; mô hình mờ TSK; luật mờ; máy học véc-tơ hỗ trợ; máy học véc-tơ hỗ trợ hồi quy; tri thức tiên nghiệm.

1 ĐẶT VẤN ĐỀ

Mô hình mờ được biết đến như là một mô hình khá hiệu quả trong việc xử lý những thông tin mơ hồ và không chắc đó chắn, đồng thời nó cũng thể hiện những lợi thế rõ ràng trong việc biểu diễn và xử lý tri thức. Hoạt động của mô hình mờ phụ thuộc chủ yếu vào hệ thống các luật mờ và quá trình suy diễn trên tập luật mờ đó. Đã có nhiều tác giả nghiên cứu và đề xuất các phương thức để xây dựng các mô hình mờ hướng dữ liệu. Nhìn chung các giải pháp này chủ yếu dựa trên kỹ thuật khai phá dữ liệu, các mô hình máy học thống kê như: Neural Network [6][7][8], Support Vector

Machine (SVM) [1][4][5][11][12][16], Self Organizing Map (SOM) [8], Cây quyết định [15], Đại số gia tử [17], và các thuật toán Phân cụm, Phân lớp, Hồi quy,... [3][6].

Việc trích xuất tập luật mờ tự động từ dữ liệu thông qua học máy sẽ có nhiều khiếm khuyết do dữ liệu ngẫu nhiên có thể bị lỗi (nhiều), thiếu tính đặc trưng, thiếu tính bao phủ. Vì vậy, việc hiểu được tập luật để hiệu chỉnh, bổ sung, tối ưu hóa là thật sự cần thiết. Các nghiên cứu nhằm tích hợp tri thức chuyên gia với mô hình mờ hướng dữ liệu có thể tìm thấy trong [2][9][12][13]. Tuy nhiên với một tập luật có số lượng lớn thì việc hiểu và diễn dịch được chúng, đồng thời có thể phân tích và tích hợp chúng với các tri thức có tính chất tinh túy của chuyên gia, quả thật rất khó khăn. Phân cụm tập luật đã đề xuất trong [16] là một trong những giải pháp cho vấn đề đơn giản hóa tập luật của các mô hình mờ hướng dữ liệu.

Nghiên cứu lý thuyết học dựa trên tri thức [14] cho thấy các kiểu khác nhau của tri thức tiên nghiệm (a priori knowledge) có thể sử dụng để cải thiện hiệu quả của mô hình máy học nói chung và mô hình mờ nói riêng. Bài báo này sẽ tập trung nghiên cứu về các kịch bản tích hợp tri thức tiên nghiệm trong việc học mô hình mờ để cải thiện mô hình mờ hướng dữ liệu trích xuất được.

Các phần tiếp theo của bài báo bao gồm: phần 2 tìm hiểu về các nguồn tri thức tiên nghiệm trong việc học mô hình mờ. Trong phần 3, chúng tôi trình bày những kịch bản khác nhau của việc tích hợp tri thức tiên nghiệm trong việc học mô hình mờ. Phần 4 trình bày một số kết quả thực nghiệm với một số ví dụ tích hợp tri thức tiên nghiệm trong việc huấn luyện mô hình mờ. Cuối cùng, trong phần 5, chúng tôi nêu lên một số kết luận và định hướng nghiên cứu tiếp theo.

2 CÁC NGUỒN TRI THỨC TIÊN NGHIỆM TRONG VIỆC HỌC MÔ HÌNH MỜ

Tri thức tiên nghiệm về hệ thống được nghiên cứu có thể ở dưới nhiều dạng khác nhau. Một khác biệt đầu tiên về tri thức của một mô hình là tri thức mô tả cơ chế hoạt động của mô hình và tri thức tinh túy có từ kinh nghiệm của chuyên gia. Cả hai kiểu khác nhau của tri thức này đều có thể kết hợp với nhau trong một mô hình mờ [9].

Tri thức về qui trình sẵn có có thể được sử dụng để mô tả hệ thống phi tuyến phức tạp như là một bộ thu thập giản đơn, ví dụ như các hệ thống tuyến tính chỉ có giá trị trong chế độ hoạt động nhất định nào đó. Những thông tin này có thể mã hóa dưới dạng các qui tắc mờ. Các biến đặc trưng cho sự thay đổi các chế độ hoạt động trở thành một phần của các đối tượng trong hệ thống các qui tắc mờ, và hàm thành viên được định nghĩa để xác định cho mỗi mô hình thành phần của một miền nhất định.

Đối với vấn đề mô hình hóa hệ thống, các kiểu khác nhau của tri thức tiên nghiệm gồm có:

- Tầm quan trọng của dữ liệu: trong nhiều ứng dụng thực tế, những mẫu dữ liệu nhất định có thể là ngoại lai và một số có thể bị nhiễu. Do vậy, mô hình xây dựng được từ dữ liệu có thể bị nhiễu hay mất ổn định.

- Hành vi của các máy học: trong một quá trình học tập, không gian giả thuyết của máy học cần được hạn chế trước. Ví dụ, đối với mô hình mạng nơ-ron hồi quy, người ta phải xác định các nguyên mẫu của một vấn đề hồi quy và thiết kế trước các cấu trúc liên kết mạng của một mạng nơ-ron.

- Mục tiêu của các máy học: tiêu chí như sự ổn định, độ bền vững, thời gian khởi động, thời gian thiết lập là các kiến thức phải có trước cho một nhà thiết kế hệ thống.

3 CÁC KỊCH BẢN HỌC CÓ TRI THỨC TIỀN NGHIỆM

Trong phần này chúng ta sẽ chứng tỏ vai trò của tri thức tiên nghiệm với việc học một mô hình mờ. Ở đây chúng ta cần làm rõ mối quan hệ logic giữa giả thuyết (Hypothesis), những mô tả mẫu (Descriptions) (dưới dạng các thuộc tính), và kết quả dự đoán (Predictions). Cho Descriptions là hội của tất cả các mô tả mẫu trong tập huấn luyện, và cho Predictions là hội của tất cả các tiên đoán. Khi đó, Hypothesis “giải thích các quan sát” phải thỏa mãn điều kiện sau (ký hiệu \models có nghĩa là kế thừa logic) [14]:

$$\text{Hypothesis} \wedge \text{Descriptions} \models \text{Predictions}$$

Xét trong trường hợp học mô hình mờ, chúng ta có thể định nghĩa khái niệm Hypothesis như sau:

Định nghĩa 1 (Hypothesis). Cho $D=\{X,Y\}$ là một tập dữ liệu huấn luyện (các quan sát), một mô hình mờ M là được gọi là Hypothesis nếu điều kiện sau thỏa mãn:

$$(\forall X \in D) (M(X)=Y)$$

Ở các nội dung tiếp theo chúng ta sẽ giải thích và định nghĩa 3 phương thức học mô hình mờ theo [14]: Học dựa trên giải thích (EBL – Explanation-based learning), Học dựa trên sự thích hợp (RBL – Relevance-based learning) và Học quy nạp dựa trên tri thức (KBIL – Knowledge-based inductive learning).

3.1 Học dựa trên sự giải thích (EBL)

Phương pháp học EBL là một phương thức trích xuất những luật mờ chung từ các quan sát riêng lẻ. Ý tưởng cơ bản của EBL là sử dụng tri thức tiên nghiệm để xây dựng cấu trúc ban đầu của Hypothesis, rồi sau đó xác lập Hypothesis chính thức dựa vào các quan sát thực nghiệm. Cụ thể theo định nghĩa của [14] như sau:

$$\begin{aligned} & \text{Background} \models \text{Hypothesis} \\ & \text{Hypothesis} \wedge \text{Descriptions} \models \text{Predictions} \end{aligned}$$

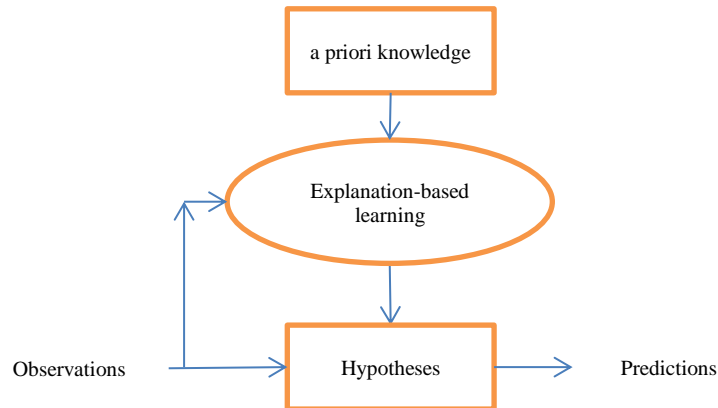
Cải thiện mô hình mờ hướng dữ liệu TSK với tri thức tiên nghiệm

Hình 1 mô tả phương pháp học EBL. Xét trong trường hợp học mô hình mờ, EBL được định nghĩa như sau:

Định nghĩa 2 (EBL). Cho B là một tri thức nền về một mô hình mờ M. Khi đó, một thuật toán A được gọi là một thuật toán học EBL nếu thỏa mãn điều kiện sau:

$$B \models M,$$

$$(\forall X \in D)(M(X)=Y).$$



Hình 1. Mô hình học EBL

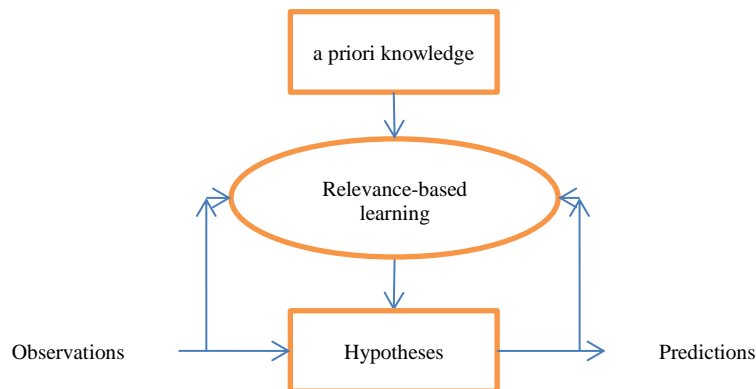
Theo phương pháp học này, đối với việc học mô hình mờ thì việc lựa chọn trước các nguyên mẫu về phân lớp hay hồi quy cho mô hình, rồi sau đó tiến hành huấn luyện mô hình, được xem là hình thức học EBL.

3.2 Học dựa trên sự thích hợp (RBL)

Theo phương pháp học này, tri thức tiên nghiệm sẽ kết hợp với những quan sát thực nghiệm để cho phép máy học có thể rút ra những qui tắc mới giải thích cho các thực nghiệm trên [14]:

$$\text{Background} \wedge \text{Descriptions} \wedge \text{Predictions} \models \text{Hypothesis}$$

$$\text{Hypothesis} \wedge \text{Descriptions} \models \text{Predictions}$$



Hình 2. Mô hình học RBL

Hình 2 mô tả phương thức học RBL. Xét trong trường hợp học mô hình mờ thì RBL được định nghĩa như sau:

Định nghĩa 3 (RBL). Cho B là kiến thức nền về một mô hình mờ M. Một thuật toán A được gọi là một thuật toán học RBL nếu điều kiện sau được thỏa mãn:

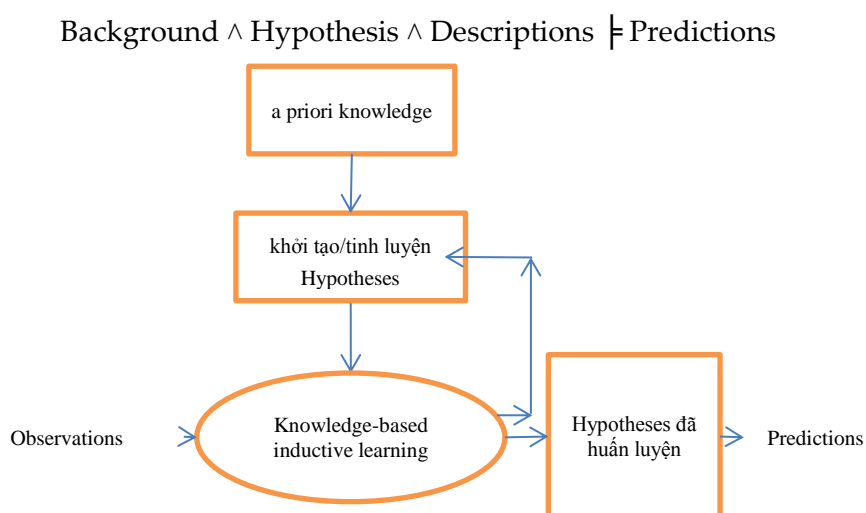
$$(\forall X \in D)(M(X)=Y),$$

$$B \wedge Y \models M.$$

Theo phương pháp học này, đối với việc học mô hình mờ thì việc sử dụng những tri thức liên quan về việc xác định cấu trúc mô hình, các thuộc tính về chức năng của mô hình để cải thiện độ vững chắc của mô hình.

3.3 Học quy nạp dựa trên tri thức (KBIL)

KBIL là một phương thức học theo kiểu gia tăng, trong đó tri thức tiên nghiệm và giả thiết mới học được sẽ kết hợp với nhau để giải thích cho các thực nghiệm. Cụ thể theo định nghĩa của [14] như sau:



Hình 3. Mô hình học KBIL

Hình 3 mô tả phương pháp học KBIL. Theo phương pháp này, ban đầu tri thức nền và những quan sát thực nghiệm được dùng để đưa ra Hypotheses ban đầu. Sau đó, kết hợp các quan sát thực nghiệm và tri thức nền để củng cố Hypotheses để có thể gia tăng tính chính xác của dự đoán. Xét trong trường hợp học mô hình mờ thì KBIL được định nghĩa như sau:

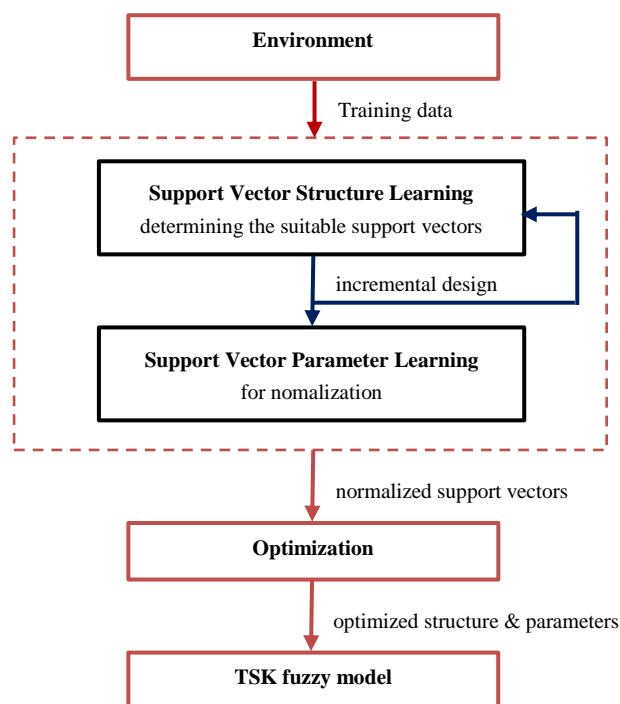
Định nghĩa 4 (KBIL). Cho B là kiến thức nền về một mô hình mờ M. Một thuật toán A được gọi là một thuật toán học KBIL nếu điều kiện sau được thỏa mãn:

$$(\forall X \in D)(B \wedge M(X)=Y)$$

Chú ý rằng, bất kỳ Hypotheses nào sinh ra dựa vào KBIL cũng phải phù hợp với tri thức tiên nghiệm cũng như với những quan sát thực nghiệm mới, phạm vi của Hypotheses sẽ được thu gọn để chỉ chứa quy tắc thật sự đã biết. Do vậy, trong trường hợp học mô hình mờ, việc áp dụng các thuật toán tối ưu hóa mô hình như gradient descent chính là một hình thức học dựa vào KBIL, bởi vì Hypotheses (mô hình) sẽ tăng dần sự thích nghi theo quá trình học và phạm vi của Hypotheses sẽ thu gọn để phù hợp với những phản hồi từ quan sát thực nghiệm.

4 Một số kết quả thực nghiệm và bàn luận

Ở nghiên cứu này chúng tôi thực nghiệm xây dựng mô hình mờ TSK trên cơ sở trích xuất tập luật mờ từ máy học véc-tơ hỗ trợ hồi quy (kịch bản EBL); trong đó có tối ưu hóa các tham số hàm thành viên Gauss của mô hình bằng phương pháp Gradient descent (kịch bản KBIL), và tối ưu hóa cấu trúc mô hình bằng cách thay đổi tham số ϵ và rút gọn các luật mờ tương tự (kịch bản RBL). Sơ đồ khối của mô hình thực nghiệm được biểu diễn ở hình 4.



Hình 4. Mô hình thực nghiệm trích xuất mô hình mờ TSK từ SVM

4.1 Ví dụ một mô hình hồi qui phi tuyến

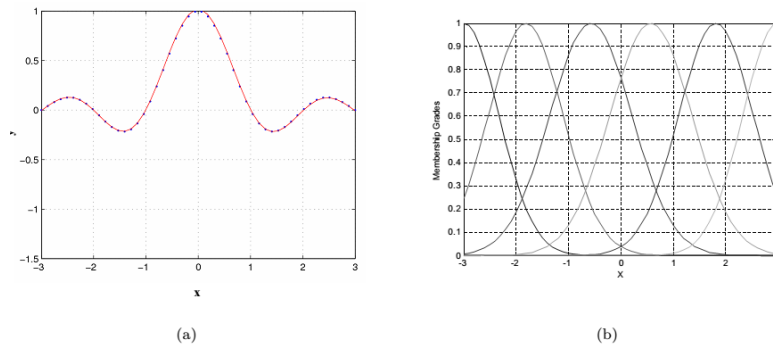
Để chứng tỏ hiệu quả của tri thức tiên nghiệm trong việc huấn luyện mô hình mờ, chúng tôi đã tích hợp tri thức tiên nghiệm về cấu trúc mô hình vào mô hình mờ TSK trích xuất từ máy học véc-tơ hỗ trợ (SVMs) [16], cụ thể là thông qua việc điều chỉnh tham số ϵ chúng ta có thể giới hạn số lượng SVs; và đây cũng chính là số lượng

luật mờ trích xuất được. Chúng tôi tiến hành thực nghiệm giải quyết bài toán hồi quy phi tuyến sau:

$$\text{Sinc}(x) = \begin{cases} \frac{\sin(x)}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \quad (1)$$

Tập dữ liệu huấn luyện được xác định trong phạm vi từ -3π đến $+3\pi$.

Trong quá trình xác định cấu trúc SVM, chúng tôi sử dụng tham số ε để điều chỉnh số lượng SVs. Chúng tôi cố định tham số $C = 10$. Khi giá trị $\varepsilon = 0.0$, sẽ có 50 SVs nhận được từ mô hình, đồng nghĩa với việc chúng ta nhận được 50 luật. Sau đó chúng tôi tăng giá trị tham số ε lên, khi $\varepsilon = 0.1$, thì có 6 SVs nhận được tương ứng với 6 luật mờ trích xuất. Hình 4 thể hiện kết quả của mô hình thực nghiệm. Như vậy chúng ta có thể điều chỉnh giá trị tham số ε , tức điều chỉnh số lượng SVs để tối ưu hóa vị trí của SVs, đồng nghĩa với việc tối ưu hóa phân bố và số lượng luật mờ.



Hình 5. Kết quả mô hình đã tối ưu hóa (RMSE = 0.0183)

4.2 Ví dụ mô hình dự đoán chuỗi hồi qui hỗn loạn Mackey-Glass

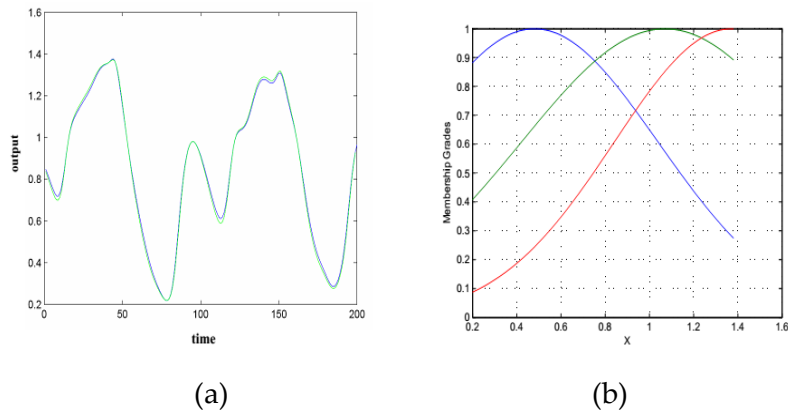
Với thực nghiệm này chúng tôi muốn chứng tỏ rằng, mô hình mờ dựa trên SVMs với sự tích hợp tri thức tiên nghiệm thật sự mang lại hiệu quả. Dữ liệu được lựa chọn để thử nghiệm là dữ liệu chuỗi thời gian Mackey-Glass (Mackey-Glass time series). Dữ liệu chuỗi thời gian Mackey-Glass được sinh theo công thức sau [10]:

$$\dot{x} = \frac{ax(t-\tau)}{1+x^b(t-\tau)} - cx(t) \quad (2)$$

Trong đó ta chọn $\tau = 30$, $a = 0.2$, $b = 10$, và $c = 0.1$. Với 1000 mẫu dữ liệu sinh ra, 800 mẫu dữ liệu được sử dụng để huấn luyện cho máy học Véc-tơ hỗ trợ và sinh ra các luật mờ, 200 mẫu dữ liệu còn lại được sử dụng để thử nghiệm suy luận dựa trên tập luật mờ trích xuất được. Thuộc tính đầu vào được lựa chọn là các giá trị $x(t-1)$, $x(t-2)$, thuộc tính đầu ra cần dự đoán là giá trị $x(t)$. Như vậy mô hình có 02 đầu vào và 01 đầu ra.

Trong thực nghiệm này chúng tôi cũng giữ cố định tham số $C = 10$ và điều chỉnh giá trị tham số $\varepsilon = 0.1$, số lượng SVs tương ứng của mô hình là 3. Hình 5b thể hiện phân bố của các hàm thành viên tương ứng với biến đầu vào $x(t-2)$.

Cải thiện mô hình mờ hướng dữ liệu TSK với tri thức tiên nghiệm



Hình 6. Kết quả dự đoán trên 200 mẫu dữ liệu thử nghiệm (RMSE = 0.0092)

Bên cạnh việc thực nghiệm dự đoán giá trị $x(t)$ bộ dữ liệu thử nghiệm (500 mẫu dữ liệu) sử dụng tập luật mờ sản xuất được từ mô hình đề xuất (mô hình f-SVM*), chúng tôi cũng thử nghiệm dự đoán trên cùng bộ dữ liệu đó với các mô hình được đề xuất bởi các tác giả khác, bao gồm mô hình ANFIS, RBF và SVM. Hiệu quả của các mô hình được so sánh và đánh giá dựa trên thông số RMSE, thông số này đo lường độ lệch giữa giá trị $x(t)$ thực tế và giá trị dự đoán được. So sánh các giá trị của RMSE trong Bảng 1 ta có thể nhận thấy rằng mô hình f-SVM* cho kết quả dự đoán tương đương với mô hình ANFIS và tốt hơn so với các mô hình RBF và SVM.

Bảng 1. So sánh kết quả các mô hình qua thông số RMSE

Số luật mờ	Mô hình áp dụng			
	ANFIS	RBF	SVM	f-SVM*
170	$<10^{-10}$	0.0042	0.0540	$<10^{-10}$
6	0.0034	0.0082	0.0509	0.0076
5	0.0041	0.0086	0.0635	0.0090
4	0.0050	0.0091	0.0748	0.0091
3	0.0074	0.0141	0.1466	0.0092
2	0.0087	0.0191	0.1955	0.0088

5 KẾT LUẬN

Trong bài báo này chúng tôi nghiên cứu một hướng tiếp cận tích hợp tri thức tiên nghiệm với việc học mô hình mờ để có thể trích xuất được hệ thống mờ tốt hơn từ dữ liệu huấn luyện. Với các kịch bản khác nhau của việc học mô hình có sự kết hợp của

tri thức tiên nghiệm, chúng ta có thể xây dựng các mô hình mờ hướng dữ liệu trong đó có sự tích hợp các kiểu tri thức tiên nghiệm để cải tiến hiệu quả của mô hình.

Với một số kết quả thực nghiệm trên mô hình mờ hướng dữ liệu trích xuất từ máy học véc-tơ hỗ trợ; đồng thời kết hợp với một số tri thức tiên nghiệm thích hợp được lựa chọn đã chứng tỏ được tính khả thi và hiệu quả của việc tích hợp tri thức tiên nghiệm cải tiến hiệu quả mô hình mờ hướng dữ liệu.

Trong định hướng nghiên cứu tiếp theo, chúng tôi sẽ xây dựng các mô hình mờ tích hợp để giải quyết một số bài toán dự báo trong thực tế. Các mô hình mờ tích hợp này có thể được xây dựng dựa trên cơ sở lựa chọn các tri thức tiên nghiệm thích hợp với mô hình để có thể trích xuất được tập luật mờ “tốt”; đồng thời từ đó có thể tối ưu hóa tập luật và lựa chọn bổ sung một số luật thu thập được từ các chuyên gia.

TÀI LIỆU THAM KHẢO

- [1] David Martens, Johan Huysmans, Rudy Setiono, Jan Vanthienen, Bart Baesens,: Rule Extraction from Support Vector Machines - An Overview of issues and application in credit scoring, *Studies in Computational Intelligence (SCI)* 80, 33–63, 2008.
- [2] Dou, W., Ruan, S., Chen, Y., Bloyet, D., and Constans, J.-M.: A framework of fuzzy information fusion for the segmentation of brain tumor tissues on RM images, *Image and Vision Computing*, vol. 25, no. 2, 164-171, 2007.
- [3] George Bojadjev, Maria Bojadjev,: *Fuzzy logic for Business, Finance, and Management*, World Scientific Publishing Co. Pte. Ltd, 2007.
- [4] J.-H Chiang and P.-Y Hao,: Support vector learning mechanism for fuzzy rule-based modeling: a new approach. *IEEE Trans. On Fuzzy Systems*, vol. 12, pp. 1-12, 2004.
- [5] J.L. Castro, L.D. Flores-Hidalgo, C.J. Mantas and J.M. Puche,: Extraction of fuzzy rules from support vector machines. Elsevier. *Fuzzy Sets and Systems*, 158, 2057 – 2077, 2007.
- [6] John Yen, Reza Langari,: *Fuzzy logic: Intelligence, Control, and Information*, Prentice hall, Uper dadle river, New Jersey, 1999.
- [7] J.-S. R. Jang and C.-T. Sun,: Functional equivalence between radial basis function networks and fuzzy inference systems, *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 156-159, 1993.
- [8] Juan C. Figueroa-García, Cynthia M. Ochoa-Rey, José A. Avellaneda-González,: Rule generation of fuzzy logic systems using a self-organized fuzzy neural network, *Neurocomputing– ELSEVIER*, 151, 955–962, 2015.
- [9] M Delgado, F Herrera, E Herrera-Viedma, L Martin,: *Combining Numerical and Linguistic Information in Group Decision making*, *Journal of Information Sciences*, 1996.
- [10] M.C. MacKey and L. Glass,: Oscillation and chaos in physiological control systems, *Science*, vol. 197, pp. 287–289, 1997.

- [11] Nahla Barakat, Andrew P. Bradley,: Rule extraction from support vector machines: A review, *Neurocomputing – ELSEVIER*, 74, 178–190, 2010.
- [12] S. Chen, J. Wang and D. Wang,: Extraction of fuzzy rules by using support vector machines. *IEEE, Computer society*, pp. 438-441, 2008.
- [13] Serge Guillaume, Luis Magdalena,: Expert guided integration of induced knowledge into a fuzzy knowledge base, *Soft Comput, Springer-Verlag*, 10: 773–784, 2006
- [14] Stuart Rusell, Perter Norvig,: *Artificial Intelligence: A Modern Approach*, Second Editor, Prentice Hall - Series in Artificial Intelligence, 2003
- [15] Xianchang Wang, Xiaodong Liu, Witold Pedrycz, Lishi Zhang,: Fuzzy rule based decision trees, *Pattern Recognition– ELSEVIER*, 48, 50–59, 2015.
- [16] Duc-Hien Nguyen, Manh-Thanh Le,: A two-stage architecture for stock price forecasting by combining SOM and fuzzy-SVM, *International Journal of Computer Science and Information Security (IJCSIS)*, USA, ISSN: 1947-5500. No: Vol. 12 No. 8, 20-25, 2014.
- [17] Dương Thăng Long, *Luận án tiến sĩ Toán học: Phương pháp xây dựng hệ mờ dạng luật với ngữ nghĩa dựa trên đại số gia tử và ứng dụng trong bài toán phân lớp*, Viện Công nghệ Thông tin - Viện Khoa học và Công nghệ Việt Nam, 2010.

IMPROVING DATA DRIVEN FUZZY MODELS WITH A PRIORI KNOWLEDGE

Nguyen Duc Hien¹, Le Manh Thanh²

University of Sciences, Hue University

¹ Email: hiencit@gmail.com

² Email: lmthanh@hueuni.edu.vn

ABSTRACT

By uniformly satisfying these conditions between TSK fuzzy models and Support vector machines for regression, we can construct an algorithm to extract TSK fuzzy model from Support vector machines. The previous study showed that the extracted fuzzy model still exists certain limitations. This paper proposed a fuzzy model based on the intergartion of a priori knowledge and Support vector machines-based TSK fuzzy model for regression prediction. The extraction of intepretable fuzzy rules from data based on the combination of statistical machine learning models is the foundation of this proposed approach. By integrating a priori knowlegde with a SVM-based fuzzy model, the extracted fuzzy set is optimized and reduces the complication. The effectiveness of the proposed solutions is evaluated through experimental results and a comparison with the results of some other models.

Keywords: fuzzy model; TSK fuzzy model; fuzzy rules; support vector machine; support vector machine for regression; a priori knowledge.



Nguyễn Đức Hiến sinh ngày 13/12/1973 tại Đà Nẵng. Năm 1996, ông tốt nghiệp cử nhân ngành Vật lý – Tin học tại Trường Đại học Sư phạm Huế. Năm 2003, ông tốt nghiệp thạc sĩ chuyên ngành Công nghệ thông tin tại Đại học Đà Nẵng. Từ năm 2007 đến nay, ông giảng dạy tại Trường CĐ Công nghệ thông tin, Đại học Đà Nẵng. Từ năm 2012 đến nay, ông là nghiên cứu sinh tiến sĩ chuyên ngành Khoa học máy tính tại Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: máy học, đặc biệt là mô hình mờ hướng dữ liệu.



Lê Mạnh Thạnh sinh ngày 01/06/1953 tại Quảng Trị. Năm 1977, ông tốt nghiệp kỹ sư ngành Toán công trình tại Trường Đại học Bách khoa Hà Nội. Năm 1994, ông tốt nghiệp tiến sĩ chuyên ngành Khoa học máy tính tại trường Đại học tổng hợp Budapest, Hungary (ELTE). Năm 2002, ông được nhà nước công nhận Phó giáo sư chuyên ngành Khoa học máy tính. Từ năm 1978 đến nay, ông là giảng viên tại Trường Đại học Tổng hợp Huế, nay gọi là Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: cơ sở dữ liệu, hệ thống thông tin, chương trình logic và các lĩnh vực liên quan.

