

Trực quan có trọng số mờ SVM đôi bình phương nhỏ nhất

Ông Tanveer ^{1b}, Thành viên cao cấp, IEEE, MA Ganaie ^{1b}, A. Bhattacharjee và CT Lin ^{1b}, Nghiên cứu viên, IEEE

Tóm tắt—Thành viên mờ là một phương pháp hiệu quả được sử dụng trong các máy vector hỗ trợ song sinh (SVM) để giảm tác động của nhiễu và các giá trị ngoại lai trong các vấn đề phân loại. SVM song sinh mờ (TWSVM) gán trọng số thành viên để giảm tác động của các giá trị ngoại lai, tuy nhiên, nó bỏ qua vị trí của các mẫu dữ liệu đầu vào và do đó không phân biệt được giữa các vector hỗ trợ và nhiễu. Để khắc phục vấn đề này, TWSVM mờ trực quan đã kết hợp khái niệm số mờ trực quan với TWSVM để giảm tác động của các giá trị ngoại lai và phân biệt các vector hỗ trợ với nhiễu. Mặc dù có những lợi ích này, TWSVM và TWSVM mờ trực quan vẫn gặp một số nhược điểm như: 1) thông tin lân cận cục bộ bị bỏ qua giữa các điểm dữ liệu và 2) chúng giải quyết các vấn đề lập trình bậc hai (QPP), vốn không hiệu quả về mặt tính toán. Để khắc phục những vấn đề này, chúng tôi đề xuất một TWSVM bình phương tối thiểu có trọng số mờ trực quan mới cho các vấn đề phân loại. Phương pháp được đề xuất sử dụng thông tin lân cận cục bộ giữa các điểm dữ liệu và cũng sử dụng cả trọng số thành viên và không thành viên để giảm tác động của nhiễu và các giá trị ngoại lai. Phương pháp đề xuất giải quyết một hệ phương trình tuyến tính thay vì giải quyết QPP, điều này làm cho mô hình hiệu quả hơn. Chúng tôi đã đánh giá các TWSVMs bình phương tối thiểu có trọng số mờ trực quan được đề xuất trên một số tập dữ liệu chuẩn để chứng minh hiệu quả của mô hình được đề xuất. Phân tích thống kê được thực hiện để định lượng kết quả về mặt thống kê. Là một ứng dụng, chúng tôi đã sử dụng mô hình được đề xuất để chẩn đoán bệnh tâm thần phân liệt.

Thuật ngữ chỉ mục—Thành viên mờ, số mờ trực giác (IFN), SVM song sinh bình phương nhỏ nhất, máy vector hỗ trợ (SVM).

Bản thảo được nhận ngày 23 tháng 4 năm 2021; sửa đổi ngày 15 tháng 9 năm 2021 và ngày 13 tháng 2 năm 2022; chấp nhận ngày 28 tháng 3 năm 2022. Ngày xuất bản ngày 10 tháng 6 năm 2022; ngày của phiên bản hiện tại ngày 16 tháng 6 năm 2023. Công trình này được hỗ trợ một phần bởi Bộ Khoa học và Công nghệ, Ban Nghiên cứu Khoa học và Kỹ thuật (SERB) tài trợ cho các Dự án Nghiên cứu, Chính phủ Ấn Độ, thông qua Chương trình Giải thưởng Nghiên cứu Sự nghiệp Sớm theo Khoản tài trợ ECR/2017/000053, thông qua Chương trình Học bổng Ramanujan, theo Khoản tài trợ SB/S2/RJN-001/2016 và thông qua Chương trình Hệ thống Vật lý Mạng Liên ngành (ICPS) theo Khoản tài trợ DST/ICPS/CPS-Individual/2018/276; và một phần bởi Hội đồng Nghiên cứu Khoa học và Công nghệ (CSIR), New Delhi, Ấn Độ, thông qua Chương trình Nghiên cứu Bức tranh tương mở rộng (EMR) theo Khoản tài trợ 22(0751)/17/EMR-II và SERB thông qua Chương trình Hỗ trợ tập trung vào tác động nghiên cứu toán học (MATRICS) theo Khoản tài trợ MTR/2021/000787. Bài viết này được biên tập viên liên kết Y. Tan giới thiệu. (Tác giả liên hệ: M. Tanveer.)

M. Tanveer, MA Ganaie và A. Bhattacharjee thuộc Khoa Toán, Viện Công nghệ Ấn Độ Indore, Indore 453552, Ấn Độ (email: mtanveer@iiti.ac.in ; phd1901141006@iiti.ac.in ; avijit.bhattacharjee1996@gmail.com).

CT Lin làm việc tại Trung tâm Trí tuệ nhân tạo, Khoa Kỹ thuật và Công nghệ thông tin, Đại học Công nghệ Sydney, Ultimo, NSW 2007, Úc (email: chintenglin@gmail.com ; chin-teng.lin@uts.edu.au).

Bài viết này có bổ sung vật liệu cung cấp qua các tác giả và phiên bản màu của một hoặc nhiều hình ảnh có sẵn tại <https://doi.org/10.1109/TCYB.2022.3165879>.

Mã số đối tượng kỹ thuật số 10.1109/TCYB.2022.3165879

2168-2267/2022 IEEE. Được phép sử dụng cá nhân, nhưng việc xuất bản lại/phân phối lại phải có sự cho phép của IEEE.

Xem <https://www.ieee.org/publications/rights/index.html> để biết thêm thông tin.

Được cấp phép sử dụng giới hạn cho: Đại học Kỹ thuật Ostrava. Tải xuống vào ngày 08 tháng 10 năm 2024 lúc 14:36:37 UTC từ IEEE Xplore. Có áp dụng các hạn chế.

Tôi. TÔI GIỚI THIỆU

S Máy vector hỗ trợ (SVM) [1] đã tìm thấy ứng dụng trên nhiều lĩnh vực khác nhau, chẳng hạn như hồi quy [2], [3], dự đoán lỗi [4], chẩn đoán y khoa trước [5], phát hiện dị thường [6], và học đa góc nhìn [7]. SVM sử dụng khái niệm biên độ tối đa để giảm thiểu lỗi tổng quát. Các cách tiếp cận khác nhau, chẳng hạn như SVM tổng quát [8], học tập gia tăng SVM [9], học bán giám sát SVM [10], và SVM đa lớp [11] đã được đề xuất để cải thiện hiệu suất của SVM. SVM triển khai nguyên lý giảm thiểu rủi ro cấu trúc để đạt được hiệu suất tổng quát hóa tốt hơn. SVM giải quyết vấn đề lập trình bậc hai (QPP) để có được siêu phẳng tối ưu, tuy nhiên, việc giải quyết QPP hạn chế khả năng áp dụng của SVM do thời gian đào tạo cao hơn. Để khắc phục các vấn đề về tính toán, một số phương pháp đã được đưa ra, chẳng hạn như SVM cận trị riêng tổng quát (GEPSVM) [12] và SVM đôi (TWSVM) [13]. Trong GEPSVM và TWSVM, hai siêu phẳng không song song được tạo ra sao cho mỗi siêu phẳng gần hơn với các mẫu của một lớp và xa nhất có thể so với các mẫu của lớp kia. Không giống như giải một QPP lớn duy nhất trong SVM, TWSVM giải hai QPP có kích thước nhỏ hơn dẫn đến hiệu quả tốt hơn của TWSVM. TWSVM nhanh hơn SVM khoảng bốn lần [13]. Không giống như SVM thực hiện nguyên tắc giảm thiểu rủi ro cấu trúc, TWSVM thực hiện nguyên tắc giảm thiểu rủi ro thực nghiệm [14], [15]. Do đó, các vấn đề về khả năng đảo ngược cần được xử lý rõ ràng khi áp dụng TWSVM vào các vấn đề phân loại.

SVM chuẩn không khả thi khi các vector hỗ trợ được trộn lẫn với nhiễu hoặc tập dữ liệu bị nhiễm các giá trị ngoại lai. Để giảm tác động của nhiễu, SVM mất bản lề bị cắt cụt [16] đã được đề xuất. SVM với hàm mất mát pinball [17] tối đa hóa khoảng cách phân vị để giảm tác động của nhiễu và các giá trị ngoại lai. Để duy trì tính thừa thớt của các mô hình, *l*-mất mát pinball không nhạy cảm SVM [17] bỏ qua các mẫu trong vùng không nhạy cảm. SVM mất mát pinball tổng quát [18] đã định nghĩa lại vùng không nhạy cảm để khắc phục các vấn đề về nhiễu và các giá trị ngoại lệ. Trong các biến thể TWSVM, nhiều cách tiếp cận, chẳng hạn như mất mát pinball chung TWSVM [19], *L*₁ chuẩn bình phương nhỏ nhất TWSVM [20], biên độ tham số đôi SVM [21], và biên độ tham số đôi SVM với tổn thất pinball [22] đã được đề xuất để xử lý các vấn đề về nhiễu và các giá trị ngoại lệ. Ngoài các hàm mất mát khác nhau, thành viên mờ

trọng số được sử dụng trong SVM mờ (FSVM) [23]–[26] và mờ nhạt TWSVM [27] để giảm tác động của tiếng ồn và các giá trị ngoại lệ. Một tỷ lệ mất cân bằng kết hợp thành viên mờ mới lạ cho

gán trọng số mờ trong một mô hình được gọi là TWSVM bình phương mờ tối thiểu mạnh mẽ cho việc học mất cân bằng lớp (RFLSTSVM-CIL) [28]. Khoảng cách tối đa của các quả cầu đôi SVM [29] đã được sử dụng để xử lý các vấn đề mất cân bằng. Theo một nghiên cứu gần đây [30], TWSVM bình phương nhỏ dựa trên năng lượng mạnh mẽ (RELSTSVM) [31] nổi lên như là bộ phân loại tốt nhất trong số các mô hình dựa trên TWSVM. Để biết chi tiết về các mô hình TWSVM, chúng tôi giới thiệu độc giả đến [32].

SVM có độ phức tạp tính toán lớn, do đó, SVM hạ xuống tọa độ (CDSVM) [33] sử dụng phương pháp giảm tọa độ để tăng tốc quá trình tối ưu hóa. CDSVM cho rằng tất cả các mẫu đều quan trọng như nhau đối với quá trình tối ưu hóa, điều này có thể không khả thi khi có nhiều và các giá trị ngoại lai. Để đưa ra trọng số thích hợp, phương pháp giảm tọa độ TWSVM mờ (CDFTWSVM) [34] đã cân nhắc các mẫu một cách thích hợp thông qua tư cách thành viên mờ để giảm tác động của nhiễu và các giá trị ngoại lai. Các hàm thành viên mờ tập trung vào khoảng cách của mỗi mẫu từ trung tâm lớp và có thể bị nhầm lẫn khi phân biệt giữa các vectơ hỗ trợ và các giá trị ngoại lai. Do đó, FSVM dựa trên tư cách thành viên kép [35] được xây dựng. Mặc dù FSVM thành viên kép cho thấy hiệu suất tốt hơn so với FSVM, tuy nhiên, các mẫu gần trung tâm lớp được chỉ định thành viên thấp hơn so với các mẫu xa trung tâm lớp [35]. Sử dụng số mờ trực giác (IFN) [36] và hàm hạt nhân [37], mỗi mẫu đào tạo được gán một trọng số để có hiệu suất tổng quát hóa tốt hơn. TWSVM mờ trực quan (IFTWSVM) [38] gán một cặp hàm thành viên và không thành viên để giảm tác động của nhiễu và xác định chính xác nhiễu từ các vectơ hỗ trợ. Tuy nhiên, IFTWSVM có một số nhược điểm: 1) bỏ qua cấu trúc lân cận cục bộ của các mẫu dữ liệu đào tạo; 2) có độ phức tạp tính toán lớn vì phải giải quyết hai QPP; và 3) cần có hộp công cụ bên ngoài để giải quyết các QPP. Để khắc phục những vấn đề này, chúng tôi đề xuất TWSVM bình phương tối thiểu có trọng số mờ trực quan (IFW-LSTSVM). IFW-LSTSVM sử dụng trọng số cục bộ trong lớp thông qua *tô-lý* thuật trọng số dựa trên láng giềng gần nhất (KNN) để kết hợp cấu trúc lân cận cục bộ và giải hệ phương trình tuyến tính thay vì giải QPP, dẫn đến giải pháp nhanh hơn. IFW-LSTSVM được đề xuất sử dụng hàm thành viên để đo khoảng cách của mỗi mẫu từ tâm lớp và hàm không thành viên để đo mối quan hệ giữa số mẫu trong lân cận và số mẫu không điều hòa. Các ưu điểm của mô hình IFW-LSTSVM được đề xuất được liệt kê như sau.

- 1) Không giống như các mô hình TWSVM và IFTWSVM liên quan đến hộp công cụ bên ngoài để giải quyết QPP, mô hình IFW-LSTSVM được đề xuất giải quyết hệ phương trình tuyến tính không yêu cầu bất kỳ hộp công cụ bên ngoài nào.
- 2) Các mô hình TWSVM, IFTWSVM, RFLSTSVM-CIL và RELSTSVM bỏ qua thông tin lân cận cục bộ, tuy nhiên, mô hình IFW-LSTSVM được đề xuất khai thác thông tin lân cận cục bộ để cải thiện hiệu suất tổng quát hóa.
- 3) Mô hình RFLSTSVM-CIL sử dụng tỷ lệ mất cân bằng và khoảng cách của các mẫu để tính toán trọng số thành viên mờ, tuy nhiên, mô hình được đề xuất

Mô hình IFW-LSTSVM sử dụng khoảng cách giữa các mẫu cũng như tính không đồng nhất của các mẫu để tính toán giá trị điểm mờ.

- 4) Các mô hình TWSVM và RELSTSVM cho rằng tất cả các mẫu đều quan trọng như nhau để tạo ra các siêu phẳng tối ưu, tuy nhiên, giả định như vậy có thể không đúng trong các tình huống thực tế. Do đó, mô hình IFW-LSTSVM được đề xuất sử dụng các giá trị điểm mờ để đưa ra trọng số thích hợp cho các mẫu.
- 5) Các mô hình TWSVM và RFLSTSVM-CIL giảm thiểu rủi ro thực nghiệm có thể dẫn đến các vấn đề về quá khớp. Do đó, IFW-LSTSVM được đề xuất giảm thiểu rủi ro về mặt cấu trúc để tránh các vấn đề về quá khớp và cải thiện hiệu suất tổng quát hóa.

Phần còn lại của bài viết này được tổ chức như sau: Phần II thảo luận về việc xây dựng IFTWSVM, Mục III thảo luận về khái niệm tập mờ trực quan, phương pháp tạo trọng số và xây dựng mô hình IFW-LSTSVM được đề xuất cho các trường hợp tuyến tính và phi tuyến tính. Mục IV thảo luận về độ phức tạp tính toán của các mô hình IFW-LSTSVM được đề xuất và Mục V đánh giá hiệu suất của các mô hình IFW-LSTSVM được đề xuất và các mô hình cơ sở. Cuối cùng, chúng tôi kết thúc bài viết này trong Phần VI.

II. RVUI MỪNG TORK

Chúng tôi thảo luận về việc xây dựng TWSVM mờ trực quan [38] trong phần này.

A. Máy hỗ trợ vectơ song sinh mờ trực quan

Ý tưởng chính của IFTWSVM [38] được kích hoạt bởi IFN [37] và TWSVM mờ [27] các khái niệm. Không giống như SVM dựa trên mờ, IFTWSVM tạo ra các siêu phẳng phân tách tối ưu thông qua việc giải quyết hai QPP có kích thước nhỏ hơn thay vì một QPP lớn duy nhất. Để trợ cấp cho hiệu ứng của các giá trị ngoại lai, các mẫu được cân nhắc dựa trên khoảng cách của chúng từ tâm cũng như tính không đồng nhất của các mẫu. IFTWSVM với hạt nhân phi tuyến tính tìm thấy các siêu phẳng được đưa ra bởi

$$K(x, A_t)v_1 + b_1 = 0, \text{ và } K(x, A_t)v_2 + b_2 = 0 \quad (1)$$

Ở đây x là một mẫu dữ liệu, $MỘT$ là tập huấn luyện, chúng tôi T là T và b lần lượt là trọng số và độ lệch cho $T=1, 2$.

Công thức nguyên thủy của IFTWSVM phi tuyến tính [38] được định nghĩa là

$$\text{phút } \frac{1}{2} \|K(X_1, MỘT_T)v_1 + v_1 b_1\|^2 + \frac{C_1 \|chúng tôi\|_2 + C_2}{2S} \xi \\ \text{thứ } -K(X_2, MỘT_T)v_1 + v_1 b_1 + \xi \geq v_2, \xi \geq 0 \quad (2)$$

Và

$$\text{phút } \frac{1}{2} \|K(X_2, MỘT_T)v_2 + v_2 b_2\|^2 + \frac{C}{2} \|chúng tôi\|_2 + C_4 S \eta \\ \text{thứ } K(X_1, MỘT_T)v_2 + v_2 b_2 + \eta \geq v_1, \eta \geq 0. \quad (3)$$

Đây, $C_1, C_2, C_3, C_4 > 0$ là siêu tham số; v_1 và v_2 là các vectơ có số chiều thích hợp; X_1 và X_2 là các mẫu tương ứng với lớp dương và lớp âm; $MỘT = [X_1; X_2]$; ξ và η là các biến lỏng lẻo; và S là giá trị điểm số của IFN, đối với $T=1, 2$.

Thuật ngữ đầu tiên trong mỗi hàm mục tiêu làm cho các siêu phẳng gần với các mẫu của lớp tương ứng, thuật ngữ thứ hai giảm thiểu rủi ro về mặt cấu trúc và thuật ngữ thứ ba đảm bảo rằng các mẫu có trọng số ít nhất ở khoảng cách 1 từ các mẫu của lớp khác.

Đối ngẫu Wolfe tương ứng của (2) được đưa ra như

$$\begin{aligned} & \text{tối đa} \quad \alpha - \frac{1}{2} \left(\frac{1}{C_1} P + \frac{1}{C_2} S \right) \\ & \text{số } 0 \leq \alpha \leq C_2 S_2 \end{aligned} \quad (4)$$

Ở đây α là vector của hệ số nhân Lagrange, $R = [K(X_1, MỘT_t), \text{và}_1]$ và $P = [K(X_2, MỘT_t), \text{và}_2]$.

Tương tự như vậy, người ta có thể thu được Wolfe kép của bài toán (3) BẢNG

$$\begin{aligned} & \text{tối đa} \quad \beta - \frac{1}{2} \left(\frac{1}{C_1} P + \frac{1}{C_2} S \right) \\ & \text{số } 0 \leq \beta \leq C_4 S_1 \end{aligned} \quad (5)$$

Ở đây β là vector của hệ số nhân Lagrange.

Người ta có thể có được $b_{\alpha 1} = [\text{chúng tôi } b_1]$ và $b_{\alpha 2} = [\text{chúng tôi } b_2]$ BẢNG

$$b_{\alpha 1} = \begin{pmatrix} -R_t R + C_1 I - P_t \\ a, b_{\alpha 2} = \begin{pmatrix} P_t P + C_2 I - R_t \\ \beta \end{pmatrix}$$

Các mẫu dữ liệu thử nghiệm được gán nhãn lớp dựa trên mức độ gần của chúng với các siêu phẳng tối ưu của mỗi lớp. Do đó, điểm dữ liệu thử nghiệm mới $x \in R^n$ được giao/đã gán nhãn lớp như sau: $\arg \min_{i=1,2} (|\text{chúng tôi } \text{Tôi}(x, A_i) + b_{\text{Tôi}}| / \text{chúng tôi } \text{Tôi}(K(A_i, A_i) \text{Tôi}))$.

III. P. ĐƯỢC ĐÓNG CỐ TORK

Được thúc đẩy bởi IFTWSVM [38] và bình phương nhỏ nhất TWSVM [39], chúng tôi đề xuất một mô hình IFW-LSTSVM mới cho các vấn đề phân loại.

1) Tập hợp mờ trực quan [38]: Xét một tập hợp không rỗng X , một tập hợp mờ Z trong vũ trụ được định nghĩa là

$$Z = \{(z, \mu_Z(\text{chữ cái})) \mid z \in X\} \quad (6)$$

Ở đây $\mu_Z: X \rightarrow [0,1]$ và $\mu_Z(\text{tiếng Anh})$ là mức độ thành viên của $z \in X$. Không giống như các tập mờ xem xét tư cách thành viên cho mỗi mẫu, các tập mờ trực giác xem xét cả tư cách thành viên và không thành viên cho mỗi mẫu. Về mặt toán học, một tập mờ trực giác được đưa ra như

$$Z = \{(z, \mu_Z(z), \nu_Z(\text{tiếng Anh})) \mid z \in X\} \quad (7)$$

Ở đây $\mu_Z(\text{tiếng Anh})$ và $\nu_Z(\text{tiếng Anh})$ là các mức độ thành viên và các giá trị không phải thành viên của $z \in X$, tương ứng, $\mu_Z: X \rightarrow [0,1]$, $\nu_Z: X \rightarrow [0,1]$, và $0 \leq \mu_Z(\text{tiếng Anh}) + \nu_Z(\text{tiếng Anh}) \leq 1$. Mức độ do dự của $z \in X$ được đưa ra như

$$\pi_Z(\text{tiếng Anh}) = 1 - \mu_Z(\text{tiếng Anh}) - \nu_Z(\text{chữ cái}). \quad (8)$$

Một IFN được định nghĩa là $s = (\mu_s, \nu_s)$, Ở đây $\mu_s \in [0,1]$, $\nu_s \in [0,1]$, và $0 \leq \mu_s + \nu_s \leq 1$. $s = (0,1)$ và $s = (1,0)$ lần lượt là IFN nhỏ nhất và lớn nhất. Đối với một $s = (\mu_s, \nu_s)$, giá trị điểm cho IFN được đưa ra là

$$S(\theta) = \mu_s - \nu_s.$$

Tuy nhiên, đối với một số IFN, hàm số điểm không thể xác định được. Do đó, để tránh vấn đề này, hàm sau được sử dụng:

$$h(\theta) = \mu_s + \nu_s. \quad (10)$$

Từ (8) và (10), chúng ta thu được

$$h(\theta) + \pi(\theta) = 1. \quad (11)$$

Nếu như $S(\theta_1) = S(\theta_2)$ và $h(\theta_1) < h(\theta_2)$ sau đó $s_1 < \theta_2$.

Sử dụng (9), hàm điểm được đưa ra như sau

$$H(\theta) = \frac{1 - \nu(\theta)}{2 - \mu(\theta) - \nu(\theta)}. \quad (12)$$

Mối quan hệ giữa tư cách thành viên và không phải thành viên được định nghĩa như sau:

$$S(\theta_1) < S(\theta_2) \Rightarrow H(\theta_1) < H(\theta_2) \quad (13)$$

$$S(\theta_1) = S(\theta_2), h(\theta_1) < h(\theta_2) \Rightarrow H(\theta_1) < H(\theta_2). \quad (14)$$

2) Trọng số trong lớp: Chúng tôi kết hợp cấu trúc cục bộ nội lớp bằng cách sử dụng kỹ thuật trọng số dựa trên KNN [40]. Cho bất kỳ cặp điểm nào (x_i, y_i) trong cùng một lớp, trọng lượng ma trận có thể được định nghĩa là

$$T_{ij} = \begin{cases} \frac{\|x_i - x_j\|_2}{\sigma}, & \text{nếu như } x_i \text{ là tổng người hàng xóm gần nhất của } x_j \\ 0, & \text{nếu không} \end{cases}$$

Ở đây σ là tham số hạt nhân.

Sử dụng (15) trọng số nội lớp có thể được tìm thấy như sau:

$\rho(c) = \frac{1}{T_{ij}}$ Ở đây $\rho(c)$ đại diện cho trọng lượng của T mẫu thứ của lớp thứ.

A. Tuyến tính IFW-LSTSVM

Linear IFW-LSTSVM tạo ra hai siêu phẳng không song song cho các vấn đề lớp nhị phân. Về mặt toán học, các siêu phẳng được đưa ra dưới dạng

$$x + b_1 = 0 \text{ và } w^T x + b_2 = 0. \quad (16)$$

Để giải quyết các vấn đề về lớp nhị phân, chúng tôi tạo ra mặt phẳng gần cho mỗi lớp sao cho các mẫu của một lớp nhất định gần với các mẫu của mặt phẳng tương ứng. Về mặt toán học, các vấn đề tối ưu hóa của IFW-LSTSVM tuyến tính được đề xuất được đưa ra như sau

$$\begin{aligned} & \text{phút} \quad \frac{1}{2} \|\rho_1(x_1 + \text{và } b_1)\|_2^2 + \frac{C_1}{2} \|S_2\|_2^2 + \frac{C_2}{2} \|\text{chúng tôi } b_1\|_2^2 \\ & \text{thứ} \quad (X_2 \text{ chúng tôi } + \text{và } b_1) - \xi = \text{và} \end{aligned} \quad (17)$$

Và

$$\begin{aligned} & \text{phút} \quad \frac{1}{2} \|\rho_2(x_2 \text{ chúng tôi } + \text{và } b_2)\|_2^2 + \frac{C_3}{2} \|S_1\|_2^2 + \frac{C_4}{2} \|\text{chúng tôi } b_2\|_2^2 \\ & \text{thứ} \quad (X_1 \text{ chúng tôi } + \text{và } b_2) - \eta = \text{và} \end{aligned} \quad (18)$$

Đây, ρ_1 và ρ_2 là ma trận trọng số nội lớp cho từng lớp tương ứng.

Thuật ngữ đầu tiên trong mỗi bài toán tối ưu hóa làm cho mặt phẳng gần với các mẫu có trọng số của lớp tương ứng, thuật ngữ thứ hai đảm bảo rằng các mẫu có trọng số cách ít nhất 1 so với các mẫu của lớp khác và

(9) thuật ngữ cuối cùng giảm thiểu rủi ro về cấu trúc.

Thay thế các ràng buộc về mặt bằng nhau trong hàm mục tiêu tương ứng của (17), chúng ta có

$$\text{phút} \quad \frac{1}{2} \| \rho_1 (X_1 v_1 + e) \|_2 + \frac{C_1}{2} \| S_2 (X_2 v_2 + e) \|_2 \quad (19)$$

Lấy độ dốc của (19) liên quan đến (wrt) $chúng\ tô\ i$, b_1 và sau đó bằng không, chúng ta thu được

$$\begin{aligned} & (\rho_1 X_1)_t \rho_1 (X_1 chúng\ tô\ i + v_1 b_1) + C_1 (S_2 X_2)_t \\ & \times (S_2 (X_2 chúng\ tô\ i + v_2 b_1 + v_2)) + C_2 chúng\ tô\ i = 0 \\ & (\rho_1 v_1)_t \rho_1 (X_1 chúng\ tô\ i + v_1 b_1) + C_1 (S_2 v_2)_t \\ & \times (S_2 (X_2 chúng\ tô\ i + v_2 b_1 + v_2)) + C_2 b_1 = 0. \end{aligned}$$

Viết lại hai phương trình trên dưới dạng ma trận, ta thu được

$$\begin{bmatrix} chúng\ tô\ i \\ b_1 \end{bmatrix} = -C_1 \begin{bmatrix} T_1 R_1 + C_1 T_1 \\ S_2 v_2 \end{bmatrix} \quad (20)$$

Ở đây $R = \rho_1 X_1$, $T = \rho_1 v_1$, $S_2 X_2 S_2 v_2$, và T_1 là ma trận đơn vị có chiều thích hợp.

Theo cách tương tự, chúng ta thu được siêu phẳng thứ hai như

$$\begin{bmatrix} chúng\ tô\ i \\ b_2 \end{bmatrix} = -C_3 \begin{bmatrix} R_2 R_2 + T_2 T_2 + C_4 T_2 I \\ R_2 S_1 v_1 \end{bmatrix} - 1 R_2 S_1 v_1$$

Ở đây $R = S_1 X_1 S_1 v_1$ và $T = \rho_2 X_2 \rho_2 v_2$.

Từ (20) và (21), rõ ràng là IFW-LSTSVM được đề xuất giải quyết hệ phương trình để có được các siêu phẳng phân tách tối ưu mà không cần bất kỳ hộp công cụ bên ngoài nào.

Một điểm dữ liệu thử nghiệm mới $x \in R^M$ nhãn lớp từ hàm quyết định được đưa ra bởi: $\arg \min_{T \in \{1, 2\}} (\| chúng\ tô\ i - T_1 x + b_1 \|_2, \| chúng\ tô\ i - T_2 x + b_2 \|_2)$.

B. IFW-LSTSVM phi tuyến tính

Các mẫu dữ liệu có thể không thể tách biệt tuyến tính, do đó, chúng tôi sử dụng thủ thuật kernel để chiếu dữ liệu vào không gian chiều cao hơn. Chúng tôi sử dụng kernel Gaussian để chiếu dữ liệu vào không gian chiều cao hơn. Về mặt toán học, IFW-LSTSVM với kernel phi tuyến tìm thấy các siêu phẳng được đưa ra bởi

$$K(x, A_t) v_1 + b_1 = 0 \text{ và } K(x, A_t) chúng\ tô\ i + b_2 = 0$$

đây $K(\cdot, \cdot)$ biểu thị hàm hạt nhân.

Các vấn đề tối ưu hóa của IFW-LSTSVM phi tuyến tính được đề xuất được đưa ra như sau

$$\text{phút} \quad \frac{1}{2} \| \rho_1 (K(X_1, A_t) v_1 + v_1 b_1) \|_2 + \frac{C_1}{2} \| S_2 \xi \|_2 + \frac{C_2}{2} \| chúng\ tô\ i \|_2 + b_1 \quad (22)$$

Và

$$\text{phút} \quad \frac{1}{2} \| \rho_2 K(X_2, MỘT\ t) v_2 + e b_2 \|_2 + \frac{C_3}{2} \| S_1 \|_2 + \frac{C_4}{2} \| chúng\ tô\ i \|_2 + b_2 + \eta = v_1.$$

Thuật toán 1 IFW-LSTSVM

1: Đầu vào: Các mẫu đào tạo X_1, X_2 .

2: Đầu ra: Các vectơ trọng số và độ lệch cho mỗi lớp tức là, $chúng\ tô\ i, b_1, b_2$ với $T \in \{1, 2\}$.

3: Tính toán ρ_1 cho các mẫu của mỗi lớp, cho $T \in \{1, 2\}$.

4: Tính toán giá trị điểm S_2 cho các mẫu của mỗi lớp, cho $T \in \{1, 2\}$, tương ứng.

5: Giải các phương trình (20) và (21) hoặc phương trình (24) và (25) để thu được các siêu phẳng tối ưu cho từng lớp trường hợp tuyến tính và phi tuyến tính.

6: Trả về các siêu phẳng tối ưu $chúng\ tô\ i, b_1, b_2$ cho mỗi lớp, cho $T \in \{1, 2\}$.

Ở đây, mỗi hạng tử của hàm mục tiêu có động cơ tương tự như trong trường hợp tuyến tính. Hơn nữa, mỗi ký hiệu có cùng ý nghĩa như trong trường hợp tuyến tính.

Tương tự như trường hợp tuyến tính, chúng ta có

$$\begin{bmatrix} chúng\ tô\ i \\ b_1 \end{bmatrix} = -C_1 \begin{bmatrix} T_1 T_1 + R_1 R_1 + C_1 T_1 I \\ S_2 v_2 \end{bmatrix} \quad (24)$$

Ở đây $R = \rho_1 K(X_1, MỘT\ t)$, $\rho_1 v_1$ và $T = [S_2 K(X_2, MỘT\ t), S_2 v_2]$.

Tương tự như vậy, siêu phẳng tương ứng với siêu phẳng kia

(21) lớp được đưa ra như

$$\begin{bmatrix} chúng\ tô\ i \\ b_2 \end{bmatrix} = -C_3 \begin{bmatrix} R_2 R_2 + T_2 T_2 + C_4 T_2 I \\ R_2 S_1 v_1 \end{bmatrix} - 1 R_2 S_1 v_1 \quad (25)$$

Ở đây $R = [S_1 K(X_1, MỘT\ t), S_1 v_1]$ và $T = [\rho_2 K(X_2, MỘT\ t), \rho_2 v_2]$.

Một điểm dữ liệu thử nghiệm mới $x \in R^N$ có được nó là nhãn lớp từ $\arg \min_{T \in \{1, 2\}} (\| chúng\ tô\ i - T_1 x + b_1 \|_2, \| chúng\ tô\ i - T_2 x + b_2 \|_2)$.

Thuật toán cho IFW-LSTSVM được trình bày tóm tắt trong Thuật toán 1.

IV. CTÍNH TOÁNCSự phức tạp của

CÁCĐƯỢC ĐÓNG CỐIFW-LSTSVM

Trong phần này, chúng tôi thảo luận về độ phức tạp tính toán của mô hình IFW-LSTSVM được đề xuất.

A. Độ phức tạp tính toán

Chúng tôi sử dụng lớn-Ồ ký hiệu để phân tích độ phức tạp thời gian của mô hình của chúng tôi. Hãy mà tổng số mẫu đào tạo và $t \in (N/2)$ là số lượng mẫu có trong mỗi lớp. Để tính toán mức độ thành viên, mô hình IFW-LSTSVM được đề xuất bao gồm tính toán tâm lớp, tính toán bán kính lớp, khoảng cách của mỗi mẫu từ tâm lớp và phép đo mức độ thành viên của mỗi mẫu yêu cầu $O(1) + O(1) + O(m) + O(m)$ hoạt động. Để đo lường mức độ không phải là thành viên, $O(m) + O(m)$

(22) các hoạt động được yêu cầu. Do đó, mô hình IFW-LSTSVM được đề xuất sử dụng $O(m)$ các hoạt động để gán các giá trị điểm. Tính toán KNN, chúng ta có thể tính toán các trọng số nội lớp trong $O(2 \text{ tô\ i\ nh\ t\ ký\ (nam)})$ [40]. Tương tự như LSTSVM, mô hình IFW-LSTSVM được đề xuất liên quan đến việc tính toán ma trận nghịch đảo của thứ tự t . So với LSTSVM, không cần phải tốn nhiều chi phí tính toán trong khi giải quyết tối ưu hóa (23) các vấn đề của mô hình IFW-LSTSVM đề xuất.

V.E.THỬ NGHIỆM KẾT QUẢ

Trong phần này, chúng tôi phân tích các kết quả thử nghiệm của mô hình IFW-LSTSVM được đề xuất và các mô hình cơ sở đã cho (ở đây là TWSVM, RFLSTSVM-CIL, IFTWSVM và RELSTSVM).

A. Thiết lập thử nghiệm

Để đánh giá hiệu suất của IFW-LSTSVM được đề xuất và các mô hình cơ sở, chúng tôi sử dụng các tập dữ liệu chuẩn từ tập dữ liệu nghề cá [41], UCI [42], và KEEL [43] repository. Các thí nghiệm được thực hiện trên máy có MATLAB R2017a trên hệ thống có 2 Bộ xử lý Intel Xeon, 128 GB RAM và 4 TB dung lượng lưu trữ thứ cấp. Bộ dữ liệu được chia thành tỷ lệ 70 : 30 để đào tạo và thử nghiệm các mô hình tương ứng. Chúng tôi đã sử dụng phương pháp tìm kiếm lưới để điều chỉnh các siêu tham số thông qua phương pháp xác thực chéo năm lần. Đối với IFTWSVM và mô hình IFW-LSTSVM được đề xuất, chúng tôi đặt các tham số là $C_1=C_3$ và $C_2=C_4$ và đối với TWSVM chúng tôi đã sử dụng $C_1=C_2$. Để phân loại dữ liệu phi tuyến tính, chúng tôi sử dụng hạt nhân Gaussian ($K(x_1, x_2) = \exp(-||x_1 - x_2||^2 / \sigma^2)$), ở đây σ là tham số hạt nhân. Các siêu tham số tương ứng với các mô hình khác nhau được chọn từ phạm vi sau: $\sigma = [2^{-5}, 2^{-4}, \dots, 2^4, 2^5]$, $C = [10^{-5}, 10^{-4}, 10^{-3}, \dots, 10^3, 10^4, 10^5]$, vì $\gamma = 1, 2, 3, 4$. Ngoài ra, chúng tôi thiết lập phạm vi $C_0 = [0.5, 1, 1.5, 2, 2.5]$ cho RFLSTSVM-CIL và $E_1 = E_2 = 0.6$ cho mô hình RELSTSVM. Chúng tôi đã sử dụng chuẩn hóa điểm Z để chuẩn hóa các tập dữ liệu.

Chúng tôi đã đánh giá hiệu suất của IFW-LSTSVM được đề xuất và các mô hình cơ sở đã cho bằng cách sử dụng các số liệu khác nhau, chẳng hạn như diện tích dưới đường cong đặc tính hoạt động của máy thu (ROC) (AUC) hoặc độ chính xác, độ nhạy, độ đặc hiệu và các biện pháp khác được xác định là

$$\begin{aligned} \text{Độ chính xác, AUC} &= \frac{TP + TN}{TP + FP + TN + FN} \\ \text{Độ nhạy} &= \frac{TP}{TP + FN}, \text{Độ đặc hiệu} = \frac{TN}{TN + FP} \\ \text{Đo F} &= \frac{2 \times \text{Độ chính xác} \times \text{Nhớ lại}}{\sqrt{\text{Độ chính xác} + \text{Thu hồi}}} \\ \text{G-trung bình} &= \frac{\text{Độ chính xác} \times \text{Nhớ lại}}{\text{Độ chính xác} + \text{Nhớ lại}} \end{aligned} \quad (26) \quad (27) \quad (28) \quad (29)$$

Ở đây TP , TN , FP , và FN lần lượt là kết quả dương tính thật, kết quả âm tính thật, kết quả dương tính giả và kết quả âm tính giả.

Phân tích F-measure và G-mean của các mô hình phân loại đã cho và mô hình IFW-LSTSVM được đề xuất trên các tập dữ liệu KEEL và UCI với hạt nhân Gaussian được đưa ra trong Hình.-s1 và Hình.-s2, tương ứng (Ở đây, s1 và s2 biểu thị chỉ số 1 và 2 của tệp bổ sung). Người ta có thể thấy rằng mô hình được đề xuất có tính cạnh tranh với các mô hình phân loại cơ sở đã cho.

Như một ứng dụng, chúng tôi đã sử dụng mô hình đề xuất để chẩn đoán bệnh nhân tâm thần phân liệt. Theo F-phân phối cho sự xuất sắc trong nghiên cứu y sinh học (COBRE) (http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html). Dữ liệu liên quan đến 72 đối tượng mắc bệnh Tâm thần phân liệt (38.1±13.9 tuổi, độ tuổi từ 18-65) và 74 đối tượng kiểm soát khỏe mạnh (35.8±11.5 tuổi, độ tuổi từ 18-65

BẢNG I

DCHẨN ĐOÁN BỆNH SẴN CHIZOPENIA

Model	(AUC, Time(s))	(Sens., Spec.)	(F-measure, G-mean)
TWSVM	(69.96, 1.9)	(81.58, 58.33)	(73.81, 74.15)
RFLSTSVM-CIL	(71.27, 0.26)	(84.21, 58.33)	(75.29, 75.72)
IFTWSVM	(68.79, 1.94)	(73.68, 63.89)	(70.89, 70.94)
RELSTSVM	(71.27, 0.19)	(84.21, 58.33)	(75.29, 75.72)
IFW-LSTSVM	(72.81, 0.41)	(78.95, 66.67)	(75, 75.09)

Sens. denotes Sensitivity, Spec. denotes Specificity and bold face denotes best performance.

năm). Xử lý hình ảnh được thực hiện thông qua gói CAT12 hợp công cụ lập bản đồ tham số kỹ thuật (SPM) phiên bản 12 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). 3-D TQuét MRI có trọng số 1 được phân chia thành chất xám (GM), chất trắng (WM), và dịch não tủy, hộp sọ, da đầu và khoang khí. Với đăng ký giải phẫu diffeomorphic chiều cao thông qua thuật toán đại số Lie mũ (DARTEL), hình ảnh GM đã được chuẩn hóa thành không gian của Viện Thần kinh Montreal (MNI). Với hạt nhân Gaussian toàn chiều rộng-nửa cực đại 8 mm, hình ảnh GM được làm mịn đã được tạo ra. Vì kích thước tập dữ liệu tương đối Nhỏ, do đó chúng tôi chia dữ liệu thành tỷ lệ đào tạo và kiểm tra là 50:50. Chúng tôi đã sử dụng xác thực chéo 10 lần để có được các tham số tối ưu.

B. Chẩn đoán bệnh tâm thần phân liệt

Để thể hiện ứng dụng của mô hình đề xuất vào các ứng dụng thực tế, chúng tôi đã đánh giá mô hình đề xuất để chẩn đoán bệnh tâm thần phân liệt. Bảng TÔI cung cấp hiệu suất của các mô hình cơ sở và mô hình IFW-LSTSVM được đề xuất. Người ta có thể thấy rằng IFTWSVM đạt được AUC với 68.79%, tiếp theo là TWSVM với 69.96%. Cả hai mô hình RFLSTSVM-CIL và RELSTSVM đều đạt được 71.27%

(26) AUC. Tuy nhiên, trong số các mô hình được đưa ra, mô hình IFW-LSTSVM được đề xuất đạt được 72 cao nhất.81% AUC. Điều này chứng minh rằng mô hình IFW-LSTSVM được đề xuất là tốt hơn so với các mô hình cơ sở.

C. Kết quả thực nghiệm và thảo luận

Trong phần này, chúng tôi phân tích kết quả thử nghiệm của các mô hình cơ sở đã cho và mô hình IFW-LSTSVM được đề xuất với hạt nhân tuyến tính và Gaussian trên các tập dữ liệu chuẩn UCI và KEEL.

Chúng tôi đánh giá kết quả thực nghiệm theo thống kê thông qua thử nghiệm Friedman và thử nghiệm hậu học. Friedman $\sum_{j=1}^n R_j^2 = \frac{12}{n(n+1)} \frac{\sum_{j=1}^n R_j^2 - \frac{(\sum_{j=1}^n R_j)^2}{n}}{n-1}$ được đưa ra bởi $\chi^2 = \frac{12}{n(n+1)} \frac{\sum_{j=1}^n R_j^2 - \frac{(\sum_{j=1}^n R_j)^2}{n}}{n-1}$, ở đây t là số lượng các thuật toán và M là số lượng tập dữ liệu.

Thống kê tốt hơn F được đưa ra như sau: $F = \frac{1}{(N-1)} \frac{\sum_{j=1}^n R_j^2 - \frac{(\sum_{j=1}^n R_j)^2}{n}}{n-1}$, ở đây F được phân phối với $(N-1, (N-1)(t-1))$ dữ liệu độ được sử dụng trong của tự do.

Theo giả thuyết Null của phép thử Friedman, tất cả các mô hình đều hoạt động như nhau và thứ hạng trung bình của chúng bằng nhau. Giả thuyết Null bị bác bỏ nếu giá trị tính toán của F lớn hơn giá trị quan trọng của F bằng phân phối. Nếu giả thuyết không bị bác bỏ, kiểm định Nemenyi post hoc được sử dụng

để đưa ra so sánh từng cặp giữa các mô hình khác nhau. Chúng tôi đưa ra phân tích này cho từng hạt nhân tuyến tính và phi tuyến tính trên các tập dữ liệu UCI và KEEL như sau.

1) *Bộ dữ liệu KEEL với Kernel tuyến tính*: Hiệu suất của các mô hình phân loại trên các tập dữ liệu chuẩn KEEL với hạt nhân tuyến tính được đưa ra trong Bảng-sI. (Ở đây, sI biểu thị TỔI chỉ mục thứ của tập bổ sung). Bảng cung cấp AUC, thời gian tính bằng giây, độ nhạy và độ đặc hiệu của các mô hình. Người ta có thể thấy từ Bảng-sI trong tài liệu bổ sung, hiệu suất của IFW-LSTSVM được đề xuất tốt hơn so với các mô hình cơ sở. Ngoài ra, thứ hạng trung bình của IFW-LSTSVM được đề xuất tốt hơn so với các mô hình TWSVM và RFLSTSVM-CIL. Để phân tích hiệu suất của TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất về mặt thống kê, chúng tôi sử dụng kiểm định Friedman. Thứ hạng trung bình của các mô hình phân loại TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất là 3.13, 3.33, 2.78, 2.8, và 2.98, tôn trọng

một cách tích cực. Với các phép tính đơn giản, chúng ta thu được $\chi^2 = 2.6768$ và $F = 0.6577$. F được phân phối với $(tôi-1)N$ và $(tôi-1)(N-1)$ mức độ tự do. Từ F bảng phân phối, với $tôi=5$ và $N=20$, giá trị quan trọng $F(4,76) = 2.495$. Vì đã tính toán F giá trị nhỏ hơn giá trị quan trọng của F bảng phân phối, tức là $0.6577 < 2.495$, do đó, kiểm tra Friedman không cho thấy sự khác biệt đáng kể giữa IFW-LSTSVM được đề xuất và các phương pháp hiện có. Tuy nhiên, người ta có thể thấy rằng IFW-LSTSVM được đề xuất đạt được độ chính xác trung bình cao nhất so với các phương pháp cơ sở.

2) *Bộ dữ liệu KEEL với hạt nhân Gaussian*: Hiệu suất của các mô hình phân loại trên các tập dữ liệu chuẩn KEEL với hạt nhân Gaussian được đưa ra trong Bảng-II. Bảng cung cấp AUC, thời gian tính bằng giây, độ nhạy và độ đặc hiệu của các mô hình. Từ Bảng-II, người ta có thể thấy rằng hiệu suất của IFW-LSTSVM được đề xuất tốt hơn so với các mô hình cơ sở. So với các mô hình cơ sở, thứ hạng trung bình của IFW-LSTSVM được đề xuất tốt hơn. Kiểm định Friedman được sử dụng để phân tích hiệu suất của TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất về mặt thống kê. Thứ hạng trung bình của các mô hình phân loại TWSVM, IFTWSVM và IFW-LSTSVM là 3.25, 3.55, 3.25, 2.73, và 2.23, tương ứng. Với đơn giản tính toán, chúng tôi thu được $\chi^2 = 9.2264$ và $F = 2.4769$. F được phân phối với $(tôi-1)N$ và $(tôi-1)(N-1)$ mức độ tự do. Với $tôi=5$, $N=20$, và $F(4,76) = 2.495$. Từ $2.4769 < 2.495$, do đó, bài kiểm tra Friedman không phát hiện ra sự khác biệt đáng kể giữa các mô hình. Tuy nhiên, IFW-LSTSVM được đề xuất đạt được độ chính xác trung bình tốt hơn và thứ hạng trung bình thấp hơn so với các mô hình cơ sở.

3) *Bộ dữ liệu UCI với Kernel tuyến tính*: Bảng-sII trong tài liệu bổ sung đưa ra kết quả thử nghiệm của các mô hình cơ sở trên các tập dữ liệu chuẩn UCI với hạt nhân tuyến tính. Bảng-sII trong tài liệu bổ sung đưa ra AUC, thời gian tính bằng giây, độ nhạy và độ đặc hiệu của các mô hình. Từ Bảng-sII trong tài liệu bổ sung, hiệu suất của IFW-LSTSVM được đề xuất tốt hơn so với các mô hình cơ sở. Ngoài ra, thứ hạng trung bình của IFW-LSTSVM được đề xuất là thấp nhất so với các mô hình cơ sở. Phân tích thống kê của TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và đề xuất

IFW-LSTSVM được thực hiện thông qua thử nghiệm Friedman. Xếp hạng trung bình của các mô hình phân loại TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất là 4.04, 2.86, 2.71, 2.82, và 2.57, tương ứng. Sau khi sim-tính toán nhiều lần, chúng ta thu được $\chi^2 = 7.8546$ và $F = 2.1209$.

F được phân phối với $(tôi-1)N$ và $(tôi-1)(N-1)$ mức độ tự do. Với $tôi=5$, $N=14$, và $F(4,52) = 2.545$. Từ $2.1209 < 2.545$, do đó, kiểm tra Friedman không cho thấy sự khác biệt đáng kể giữa IFW-LSTSVM được đề xuất và các phương pháp hiện có. Tuy nhiên, người ta có thể thấy rằng IFW-LSTSVM được đề xuất đạt được độ chính xác trung bình cao nhất và thứ hạng trung bình thấp hơn so với các phương pháp cơ sở.

4) *Bộ dữ liệu UCI với hạt nhân Gaussian*: Bàn-III cung cấp hiệu suất của các mô hình phân loại trên các tập dữ liệu chuẩn UCI với hạt nhân Gaussian. Bảng cung cấp AUC, thời gian tính bằng giây, độ nhạy và độ đặc hiệu của các mô hình. Ngoại trừ mô hình RELSTSVM, thứ hạng trung bình của IFW-LSTSVM được đề xuất là tốt hơn. Chúng tôi sử dụng kiểm định Friedman để đánh giá hiệu suất của TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất về mặt thống kê. Thứ hạng trung bình của các mô hình phân loại TWSVM, RFLSTSVM-CIL, IFTWSVM, RELSTSVM và IFW-LSTSVM được đề xuất là 3.46, 3.71, 3.14, 2.29, và 2.39,

tương ứng. Với các phép tính đơn giản, chúng ta thu được $\chi^2 = 8.6884$ và $F = 2.3873$. F được phân phối với $(tôi-1)N$ và $(tôi-1)(N-1)$ mức độ tự do. Với $tôi=5$, $N=14$, và $F(4,52) = 2.545$. Từ $2.3873 < 2.545$, do đó, kiểm định Friedman không cho thấy sự khác biệt đáng kể giữa IFW-LSTSVM được đề xuất và các phương pháp hiện có. Từ bảng kết quả, người ta có thể quan sát thấy rằng IFW-LSTSVM được đề xuất đạt được độ chính xác trung bình cao nhất so với các mô hình cơ sở. Ngoài ra, thứ hạng trung bình của các mô hình IFW-LSTSVM được đề xuất tốt hơn các mô hình cơ sở ngoại trừ mô hình RELSTSVM.

D. Độ nhạy của siêu tham số

Trong phần này, chúng tôi phân tích tác động của siêu tham số C_1 và C_2 về hiệu suất tổng quát của IFW-LSTSVM được đề xuất. Chúng tôi đã đánh giá độ nhạy trên cả tập dữ liệu UCI và dữ liệu KEEL trong Hình.1 và 2, tương ứng. Trong Hình.1(a), người ta có thể thấy rằng hiệu suất tốt hơn ở các giá trị thấp hơn của C_1 . Tuy nhiên, trong Hình.1(b) hiệu suất tốt hơn ở các giá trị cao hơn của C_1 và tốt hơn ở các giá trị thấp hơn C_2 . Tương tự như vậy, trong Hình.1(c) hiệu suất của IFW-LSTSVM được đề xuất tốt hơn ở phạm vi trung bình của C_1 giá trị. Tương tự như trong Hình.2, người ta có thể thấy rằng việc thay đổi các siêu tham số C_1 và C_2 dẫn đến hiệu suất khác nhau của mô hình. Do đó, các siêu tham số của mô hình IFW-LSTSVM được đề xuất cần được lựa chọn cẩn thận để có hiệu suất tổng quát hóa tối ưu.

E. Ảnh hưởng của số lượng láng giềng gần nhất đến hiệu suất của mô hình IFW-LSTSVM được đề xuất

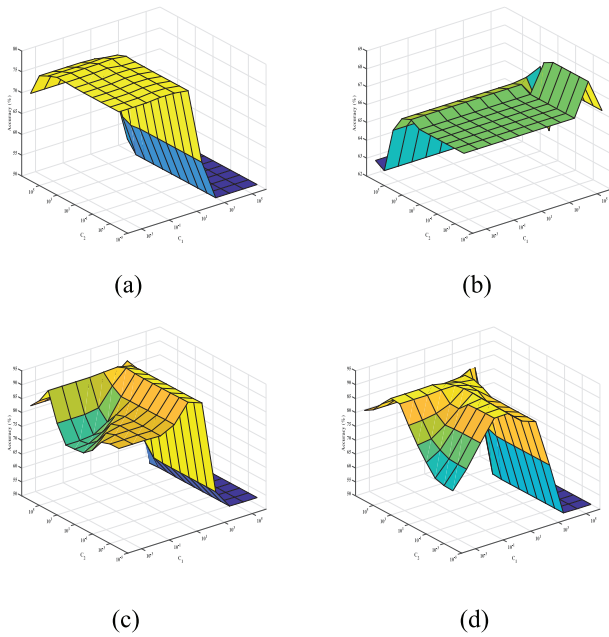
Trong phần này, chúng tôi phân tích tác động của một số hàng xóm gần nhất (k) về hiệu suất của mô hình IFW-LSTSVM được đề xuất. Trong Hình.3(a), người ta có thể thấy rằng ở các giá trị cao hơn của $tôi$ hiệu suất tốt hơn so với các giá trị thấp hơn của $tôi$. Tuy nhiên, trường hợp ngược lại xuất hiện trong Hình.3(b)

BẢNG II
PHIẾU SUẤT CỦA CLASSIFICATION TỚI MODELS TRÊN KEEL DỮ LIỆU TÀI NGUYÊN ỨNG DỤNG

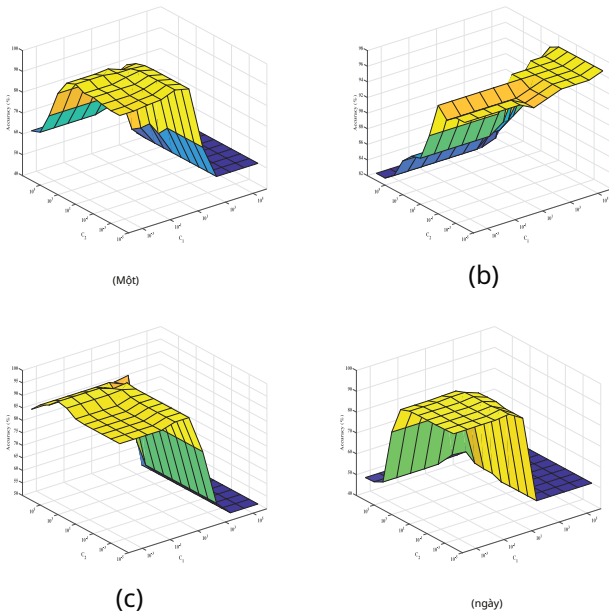
Datasets	TWSVM [13] (AUC, Time (s), Sens., Spec.) (C_1, μ)	RFLSTVSVM-CIL [28] (AUC, Time (s), Sens., Spec.) (C_0, C_1, μ)	IFTWSVM [38] (AUC, Time (s), Sens., Spec.) (C_1, C_2, μ)	RELSTVSVM [31] (AUC, Time (s), Sens., Spec.) (C_1, C_2, μ)	Proposed IFW-LSTVSVM (AUC, Time (s), Sens., Spec.) (C_1, C_2, k, μ)
aus (690 × 14)	(84.2, 0.097, 84.04, 84.35) (0.01, 32)	(83.95, 0.039, 86.17, 81.74) (1, 1, 32)	(84.82, 0.085, 86.17, 83.48) (1000, 10, 4)	(87, 0.048, 86.17, 87.83) (0.1, 1, 8)	(85.5, 0.075, 84.04, 86.96) (0.00001, 10, 4, 8)
brwiscosin (683 × 9)	(96.05, 0.12, 93.33, 98.76) (0.01, 8)	(98.76, 0.039, 100, 97.52) (1.5, 0.1, 16)	(98.76, 0.098, 100, 97.52) (1000, 100000, 4)	(98.76, 0.05, 100, 97.52) (0.00001, 100000, 4)	(99.07, 0.067, 100, 98.14) (0.01, 0.00001, 1, 16)
checkerboard_Data (690 × 14)	(84.2, 0.095, 84.04, 84.35) (0.01, 32)	(83.95, 0.04, 86.17, 81.74) (1, 1, 32)	(84.82, 0.081, 86.17, 83.48) (1000, 10, 4)	(87, 0.048, 86.17, 87.83) (0.1, 1, 8)	(85.5, 0.073, 84.04, 86.96) (0.00001, 10, 4, 8)
ecoli-0-1-4-7_vs_5-6 (332 × 6)	(87.5, 0.043, 75, 100) (0.1, 32)	(84.81, 0.01, 75, 94.62) (1, 0.1, 32)	(86.42, 0.028, 75, 97.85) (10, 0.001, 0.5)	(86.96, 0.008, 75, 98.92) (0.1, 10, 2)	(87.5, 0.018, 75, 100) (0.00001, 0.00001, 1, 0.5)
glass2 (214 × 9)	(90.32, 0.03, 100, 80.65) (10, 32)	(88.71, 0.005, 100, 77.42) (0.5, 0.1, 32)	(80.65, 0.023, 100, 61.29) (0.00001, 0.001, 32)	(80.65, 0.005, 100, 61.29) (0.00001, 0.01, 16)	(95.97, 0.01, 100, 91.94) (0.1, 0.00001, 4, 4)
haberman (306 × 3)	(57.38, 0.023, 28, 86.76) (0.1, 16)	(64.24, 0.008, 52, 76.47) (1.5, 0.01, 4)	(55.88, 0.031, 100, 11.76) (1000, 0.1, 0.125)	(61.38, 0.01, 36, 86.76) (0.1, 0.001, 32)	(62.65, 0.012, 40, 85.29) (0.00001, 0.001, 3, 4)
monk2 (601 × 7)	(90.16, 0.086, 92.19, 88.14) (0.01, 4)	(86.48, 0.026, 89.06, 83.9) (0.5, 0.001, 4)	(87.06, 0.058, 95.31, 78.81) (100, 0.00001, 8)	(85.28, 0.037, 87.5, 83.05) (0.001, 0.00001, 8)	(98.37, 0.047, 98.44, 98.31) (0.1, 0.00001, 5, 8)
new-thyroid1 (215 × 5)	(94.44, 0.023, 88.89, 100) (0.001, 1)	(93.57, 0.006, 88.89, 98.25) (0.5, 0.01, 1)	(94.44, 0.026, 88.89, 100) (0.0001, 0.0001, 4)	(99.12, 0.005, 100, 98.25) (0.01, 0.00001, 16)	(94.44, 0.01, 88.89, 100) (0.00001, 0.00001, 1, 4)
pima (768 × 8)	(77.07, 0.128, 63.29, 90.85) (0.00001, 2)	(62.77, 0.046, 98.73, 26.8) (0.5, 0.1, 32)	(75.82, 0.091, 79.75, 71.9) (0.001, 1, 2)	(77.84, 0.06, 75.95, 79.74) (0.01, 1, 32)	(76.78, 0.08, 81.01, 72.55) (0.001, 0.1, 1, 2)
ripley (1250 × 2)	(90.74, 0.344, 87.43, 94.05) (1, 2)	(91.73, 0.118, 93.19, 90.27) (1.5, 10, 1)	(92.03, 0.232, 91.62, 92.43) (10, 100, 0.125)	(90.73, 0.151, 88.48, 92.97) (10, 100, 0.25)	(92.57, 0.184, 91.62, 93.51) (10000, 0.001, 3, 0.25)
shuttle-c0-vs-c4 (1829 × 9)	(98.65, 0.793, 97.3, 100) (0.00001, 2)	(98.65, 0.276, 97.3, 100) (0.5, 0.0001, 16)	(100, 0.602, 100, 100) (0.0001, 0.00001, 16)	(98.65, 0.229, 97.3, 100) (0.00001, 0.00001, 4)	(99.81, 0.403, 100, 99.61) (0.00001, 0.00001, 1, 0.125)
sonar (208 × 60)	(80.2, 0.021, 68.97, 91.43) (0.001, 16)	(80.2, 0.006, 68.97, 91.43) (0.5, 0.001, 16)	(78.77, 0.022, 68.97, 88.57) (0.001, 0.01, 32)	(83.35, 0.004, 72.41, 94.29) (0.001, 0.00001, 16)	(81.63, 0.011, 68.97, 94.29) (0.001, 0.0001, 1, 16)
vehicle2 (846 × 18)	(100, 0.153, 100, 100) (10, 16)	(100, 0.063, 100, 100) (0.5, 1, 8)	(98.19, 0.139, 100, 96.37) (0.001, 0.01, 32)	(100, 0.048, 100, 100) (1, 0.00001, 16)	(99.48, 0.112, 100, 98.96) (1, 0.00001, 2, 16)
votes (435 × 16)	(96.87, 0.036, 96.3, 97.44) (0.01, 16)	(96.23, 0.019, 96.3, 96.15) (0.5, 10, 32)	(95.58, 0.037, 96.3, 94.87) (1, 1, 32)	(96.23, 0.011, 96.3, 96.15) (0.01, 0.00001, 32)	(96.23, 0.021, 96.3, 96.15) (0.1, 0.01, 1, 16)
vowel (988 × 10)	(79.63, 0.222, 59.26, 100) (0.0001, 1)	(81.48, 0.089, 62.96, 100) (0.5, 0.0001, 2)	(94.82, 0.202, 92.59, 97.05) (0.00001, 0.001, 1)	(81.3, 0.081, 62.96, 99.63) (0.00001, 0.00001, 2)	(97.23, 0.145, 96.3, 98.15) (0.00001, 0.00001, 1, 1)
yeast-0-5-6-7-9_vs_4 (528 × 8)	(76.96, 0.087, 58.82, 95.1) (1, 16)	(75.15, 0.027, 70.59, 79.72) (1, 0.01, 32)	(71.37, 0.066, 58.82, 83.92) (100, 10, 16)	(73.26, 0.024, 64.71, 81.82) (0.00001, 1000, 4)	(79, 0.043, 70.59, 87.41) (0.00001, 0.01, 5, 16)
yeast1 (2968 × 8)	(99.12, 1.991, 98.56, 99.67) (0.01, 0.5)	(100, 0.676, 100, 100) (1, 0.0001, 0.5)	(100, 1.341, 100, 100) (0.001, 0.00001, 0.5)	(99.92, 0.778, 100, 99.84) (0.00001, 0.001, 0.25)	(94.13, 1.026, 93.14, 95.12) (0.01, 0.0001, 1, 1)
yeast2vs8 (483 × 8)	(70.42, 0.096, 50, 90.85) (10, 8)	(72.89, 0.017, 50, 95.77) (1, 0.1, 32)	(86.8, 0.054, 75, 98.59) (10, 1, 16)	(74.3, 0.017, 50, 98.59) (0.00001, 10, 4)	(75, 0.028, 50, 100) (0.00001, 10, 3, 16)
yeast3 (1484 × 8)	(89.67, 0.559, 86.15, 93.19) (1, 16)	(74.9, 0.194, 96.92, 52.88) (1.5, 1, 32)	(81.75, 0.456, 67.69, 95.81) (10000, 10000, 32)	(93.63, 0.168, 95.38, 91.88) (0.0001, 0.01, 16)	(91.36, 0.285, 87.68, 95.03) (0.001, 0.1, 6, 16)
yeast5 (1484 × 8)	(90.25, 1.001, 83.33, 97.16) (1, 8)	(92.21, 0.192, 87.5, 96.93) (2.5, 0.001, 8)	(94.53, 0.478, 91.67, 97.4) (10000, 10000, 16)	(92.45, 0.167, 87.5, 97.4) (1, 10000, 16)	(84.83, 0.295, 70.83, 98.82) (10000, 100000, 2, 32)
Average Accuracy	86.69	85.53	87.13	87.39	88.85
Average Rank	3.25	3.55	3.25	2.73	2.23

BẢNG III
PHIẾU SUẤT CỦA CLASSIFICATION TÔI MODELS TRÊN UC1-DTÁP ĐOÀN TÔI GNGƯỜI ỦCKERNEL

Datasets	TWSVM [13] (AUC, Time (s), Sens., Spec.) (C_1, μ)	RFLSTSV-CIL [28] (AUC, Time (s), Sens., Spec.) (C_0, C_1, μ)	IFTWSVM [38] (AUC, Time (s), Sens., Spec.) (C_1, C_2, μ)	RELSTSV [31] (AUC, Time (s), Sens., Spec.) (C_1, C_2, μ)	Proposed IFW-LSTSV (AUC, Time (s), Sens., Spec.) (C_1, C_2, k, μ)
bank (4521 × 15)	(72.85, 4.099, 62.03, 83.67) (0.00001, 32)	(75.94, 1.712, 74.05, 77.83) (2, 0.01, 16)	(76.57, 3.012, 67.72, 85.42) (0.001, 10, 16)	(78.37, 2.134, 76.58, 80.17) (0.001, 10, 8)	(76.51, 2.471, 71.52, 81.5) (0.00001, 1, 3, 32)
blood (748 × 3)	(63.79, 0.109, 58.06, 69.51) (0.1, 16)	(64.28, 0.039, 37.1, 91.46) (1, 1, 8)	(68.18, 0.083, 51.61, 84.76) (0.00001, 0.1, 8)	(71.28, 0.051, 64.52, 78.05) (0.01, 1, 16)	(66.82, 0.074, 62.9, 70.73) (1, 0.0001, 5, 2)
chess-krvkp (3196 × 35)	(98.53, 1.649, 98.08, 98.98) (0.1, 16)	(97.48, 0.648, 96.8, 98.17) (2, 1, 16)	(79.3, 1.251, 98.93, 59.67) (0.1, 1, 8)	(99.16, 0.95, 98.93, 99.39) (0.01, 0.0001, 32)	(98.95, 1.061, 98.72, 99.19) (1, 0.00001, 2, 32)
ilpd-indian-liver (583 × 8)	(67.75, 0.06, 85.11, 50.39) (1, 8)	(66.39, 0.021, 80.85, 51.94) (2, 0.1, 4)	(63.4, 0.048, 70.21, 56.59) (0.0001, 10, 4)	(66.11, 0.021, 70.21, 62.02) (0.1, 0.01, 32)	(68.42, 0.034, 87.23, 49.61) (1000, 100000, 1, 32)
mushroom (8124 × 20)	(100, 14.821, 100, 100) (0.00001, 2)	(99.96, 7.121, 100, 99.92) (0.5, 0.00001, 2)	(100, 10.153, 100, 100) (0.00001, 0.00001, 16)	(100, 9.531, 100, 100) (0.00001, 0.00001, 2)	(99.84, 7.366, 99.92, 99.76) (0.00001, 0.00001, 1, 2)
oocytes_merlucius_nucleus_4d (1022 × 40)	(80.86, 0.184, 75.25, 86.47) (100, 32)	(81.03, 0.067, 89.11, 72.95) (1, 1, 32)	(67.08, 0.167, 94.06, 40.1) (0.00001, 0.01, 32)	(83.87, 0.096, 84.16, 83.57) (0.1, 0.0001, 32)	(83.34, 0.168, 80.2, 86.47) (1, 0.01, 3, 16)
ringnorm (7400 × 19)	(98.03, 11.068, 98.62, 97.44) (0.00001, 2)	(98.2, 5.555, 98.34, 98.06) (0.5, 10000, 2)	(98.5, 6.906, 97.97, 99.03) (0.1, 10, 4)	(98.73, 7.24, 98.43, 99.03) (0.01, 10, 8)	(97.87, 5.945, 99.08, 96.65) (0.001, 0.01, 3, 2)
spambase (4601 × 56)	(93.41, 3.888, 93.08, 93.74) (0.01, 32)	(82.84, 1.641, 96.54, 69.14) (2, 0.01, 8)	(91.37, 4.826, 93.65, 89.1) (0.0001, 10, 16)	(94.12, 2.301, 92.31, 95.94) (0.01, 0.0001, 32)	(94.01, 4.511, 92.31, 95.71) (0.1, 0.001, 4, 32)
statlog-australian-credit (690 × 13)	(52.38, 0.092, 38.57, 66.19) (0.1, 16)	(51.2, 0.032, 67.14, 35.25) (1.5, 0.1, 32)	(50.31, 0.056, 14.29, 86.33) (100, 1, 2)	(52.65, 0.044, 64.29, 41.01) (10, 1000, 16)	(53.43, 0.061, 45.71, 61.15) (0.1, 10000, 2, 16)
statlog-german-credit (1000 × 23)	(67.8, 0.17, 61.04, 74.55) (0.1, 32)	(66.98, 0.068, 70.13, 63.84) (1.5, 0.01, 32)	(67.96, 0.128, 64.94, 70.98) (0.1, 100, 32)	(67.9, 0.081, 68.83, 66.96) (0.00001, 10, 32)	(69.7, 0.114, 67.53, 71.88) (0.1, 1, 6, 16)
statlog-heart (270 × 12)	(81.4, 0.03, 88.89, 73.91) (0.0001, 4)	(82.79, 0.006, 91.67, 73.91) (1.5, 0.0001, 4)	(81.7, 0.016, 91.67, 71.74) (1, 0.001, 16)	(80.5, 0.006, 80.56, 80.43) (1, 0.1, 32)	(82.79, 0.009, 91.67, 73.91) (10, 0.0001, 1, 4)
titanic (2201 × 2)	(70.94, 0.586, 51.89, 90) (0.1, 32)	(72.67, 0.26, 65.57, 79.78) (1.5, 100000, 0.5)	(69.76, 0.47, 49.53, 90) (0.01, 10000, 8)	(70.94, 0.406, 51.89, 90) (1, 0.00001, 0.03125)	(72.67, 0.468, 65.57, 79.78) (0.00001, 1, 6, 0.03125)
twonorm (7400 × 19)	(97.26, 11.314, 96.68, 97.83) (0.00001, 4)	(97.21, 5.522, 96.68, 97.74) (0.5, 0.0001, 4)	(97.71, 6.776, 97.31, 98.1) (10, 100, 16)	(97.62, 7.212, 97.13, 98.1) (0.00001, 10, 32)	(97.62, 6.424, 96.86, 98.37) (1000, 100000, 2, 16)
vertebral-column-2classes (310 × 5)	(82.25, 0.029, 85.71, 78.79) (0.01, 32)	(77.65, 0.008, 75, 80.3) (1, 10, 16)	(83.28, 0.018, 89.29, 77.27) (0.00001, 0.001, 32)	(79.92, 0.009, 75, 84.85) (1, 10, 8)	(80.95, 0.012, 78.57, 83.33) (1, 0.0001, 1, 4)
Average Accuracy	80.52	79.62	78.22	81.51	81.64
Average Rank	3.46	3.71	3.14	2.29	2.39

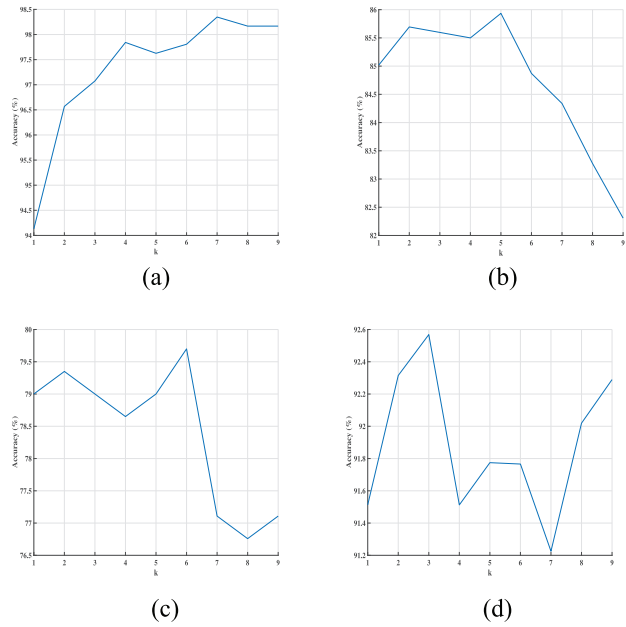


Hình 1. Hiệu ứng của các tham số C_1 và C_2 về hiệu suất của đề xuất Mô hình IFW-LSTSVM trên các tập dữ liệu UCI. (a) Bank. (b) Blood. (c) Chesskrvkp. (d) Spambase.



Hình 2. Hiệu ứng của các tham số C_1 và C_2 về hiệu suất của mô hình IFW-LSTSVM được đề xuất trên các tập dữ liệu KEEL. (a) Vehicle2. (b) Votes. (c) Vowel. (d) Yeast5.

trong đó hiệu suất tốt hơn đạt được ở các giá trị thấp hơn của $t_{\text{tôi}}$ và giảm khi tăng $t_{\text{tôi}}$. Từ Hình.3(c), người ta có thể thấy rằng hiệu suất đang thay đổi ở các mức độ khác nhau $t_{\text{tôi}}$. Trong Hình.3(d), người ta có thể quan sát thấy rằng hiệu suất ban đầu của mô hình IFW-LSTSVM được đề xuất tăng lên và sau đó giảm xuống khi tăng $t_{\text{tôi}}$. Tuy nhiên, hiệu suất sau đó lại tăng lên ở các giá trị cao hơn của $t_{\text{tôi}}$. Do đó, siêu tham số $t_{\text{tôi}}$ cần được lựa chọn tối ưu để mô hình IFW-LSTSVM được đề xuất có hiệu suất tốt hơn.



Hình 3. Hiệu ứng của tham số t về hiệu suất của IFW được đề xuất-Mô hình LSTSVM. (a) Nam men1. (b) Aus. (c) Nam men-0-5-6-7-9_vs_4. (d) Ripley.

VI. CBAO GỒM

Trong bài viết này, chúng tôi đề xuất các TWSVMs bình phương tối thiểu có trọng số mờ trực quan (IFW-LSTSVM) cho các vấn đề phân loại. Các mô hình TWSVM, RFLSTSVM-CIL và RELSTSVM coi rằng tất cả các mẫu dữ liệu đều quan trọng như nhau và bỏ qua vùng lân cận của các mẫu. Tuy nhiên, mô hình IFW-LSTSVM được đề xuất kết hợp thông tin trước về vùng lân cận của một mẫu và đưa ra trọng số thích hợp cho từng mẫu, giúp cải thiện hiệu suất khi có nhiễu và giá trị ngoại lai. Cả hai mô hình TWSVM và IFTWSVM đều yêu cầu hộp công cụ bên ngoài để có được các siêu phẳng phân tách tối ưu trong khi mô hình IFW-LSTSVM được đề xuất giải hệ phương trình tuyến tính không yêu cầu bất kỳ hộp công cụ bên ngoài nào. Mặc dù mô hình IFTWSVM sử dụng trọng số mờ để gán trọng số thích hợp cho các mẫu dựa trên khoảng cách mẫu của chúng từ tâm, tuy nhiên, nó lại bỏ qua tính không đồng nhất của các mẫu. Mô hình IFW-LSTSVM được đề xuất gán điểm mờ dựa trên khoảng cách mẫu từ tâm cũng như tính không đồng nhất của các mẫu. Vì cả tư cách thành viên và không phải thành viên đều góp phần vào điểm số mờ nên điều này dẫn đến hiệu suất được cải thiện. Đánh giá mô hình IFW-LSTSVM được đề xuất trên một số tập dữ liệu chuẩn cho thấy hiệu quả của mô hình. Người ta có thể thấy rằng mô hình IFW-LSTSVM được đề xuất cho thấy khả năng khái quát hóa tốt hơn trên các tập dữ liệu chuẩn UCI và KEEL. Ngoài ra, hiệu suất của mô hình IFW-LSTSVM được đề xuất tốt hơn trong chẩn đoán bệnh tâm thần phân liệt. Trong tương lai, người ta có thể khám phá việc sử dụng các tập mờ có giá trị khoảng, các phần mở rộng của IFN, chẳng hạn như IFN tam giác, IFN hình thang và IFN hình thang có giá trị khoảng. Ngoài ra, người ta có thể mở rộng thuật toán này cho các vấn đề phân loại đa lớp.

MỘT LỜI CẢM ƠN

Các tác giả xin chân thành cảm ơn Viện Công nghệ Ấn Độ Indore đã cung cấp cơ sở vật chất và hỗ trợ.