روشی جدید برای عضویت‌دهی به داده‌ها و شناسایی نوفه و داده‌های پرت با استفاده از ماشین بردار پشتیبان فازی

منا خداقلی[۱]، اردشیر دولتی[۲]، علی حسین زاده[۳]* و خشایارشمس‌الکتابی[۴]
[۱و۲و۳] گروه ریاضی، دانشکدۀ علوم پایه، دانشگاه شاهد، تهران، ایران
[۴] دانشجوی دکترای ریاضی، دانشگاه تربیت مدرس، تهران، ایران

چکیده

در این مقاله روشی جدید را با استفاده از ماشین بردار پشتیبان فازی به منظور عضویت‌دهی داده‌های آموزشی، براساس فاصله از ابرصفحه جداکننده معرفی می‌شود. در این روش، با استفاده از طبقه‌بندی ماشین بردار پشتیبان و نیز عضویت نخستین داده‌های آموزشی، یک تابع عضویت فازی به‌کمک اعداد فازی مثلثی متقارن برای تمام فضا معرفی می‌شود. مبتنی بر این روش، مقدار تابع عضویت فازی هر دادۀ جدیدی که می‌خواهد طبقه‌بندی شود، به‌گونه‌ای انتخاب می‌شود که کمترین میزان اختلاف را با عضویت اولیه داده‌های آموزشی و بیشترین میزان فازی‌سازی داشته باشد. نخست این مسأله به‌صورت یک مسألۀ بهینه‌سازی غیرخطی تعریف، سپس به‌کمک روش نقاط بحرانی، الگوریتمی کارا معرفی می‌شود و تابع عضویت نهایی داده‌های آموزشی به‌دست می‌آید؛ همچنین، در ادامه با مقایسۀ مقدار عضویت‌های اولیۀ داده‌های آموزشی با توزیع عضویت نهایی به‌دست‌آمده از روش پیشنهادی، میزان نوفه‌ای‌بودن دادۀ آموزشی بررسی می‌شود. در انتهای این مقاله نیز جهت فهم بهتر و نشان‌دادن کارایی الگوریتم، آزمایش‌هایی انجام و چگونگی رفتار الگوریتم پیشنهادی بر روی نمودار پیاده‌سازی و با یک روش پایه مقایسه می‌شود.

واژگان کلیدی: منطق فازی، یادگیری ماشین، داده‌کاوی، بردار پشتیبان فازی، تابع عضویت فازی

A New Method to Determine Data Membership and Find Noise and Outlier Data Using Fuzzy Support Vector Machine

Mona Khodagholi[1], Ardeshir Dolati[2], Ali Hoseinzadeh*[3] & Khashayar Shamsolketabi[4]
[1,2,3]Department of Mathematics, Shahed University, Tehran, Iran
[4]PHD student, Department of Mathematics, Tarbiat Modares University, Tehran, Iran

Abstract

Support Vector Machine (SVM) is one of the important classification techniques, has been recently attracted by many of the researchers. However, there are some limitations for this approach. Determining the hyperplane that distinguishes classes with the maximum margin and calculating the position of each point (train data) in SVM linear classifier can be interpreted as computing a data membership with certainty. A question may be raised here: how much the level of the certainty of this classification, based on hyperplane, can be trusted. In the standard SVM classification, the significance of error for different train data is considered equal and every datum is assumed to belong to just one class. However, in many cases some of train data, including outlier and vague data with no defined model, cannot be strictly considered as a member of a certain class. That means, a train datum may does not exactly belong to one class and its features may

* Corresponding author

*نویسندۀ عهده‌دار مکاتبات

show 90 percent membership of one class and 10 percent of another. In such cases, by using fuzzy SVM based on fuzzy logic, we can determine the significance of data in the train phase and finally determine relative class membership of data.

The method proposed by Lin and Wang is a basic method that introduces a membership function for fuzzy support vector machine. Their membership function is based on the distance between a point and the center of its corresponding class.

In this paper, we introduce a new method for giving membership to train data based on their distance from distinctive hyperplane. In this method, SVM classification together with primary train data membership are used to introduce a fuzzy membership function for the whole space using symmetrical triangular fuzzy numbers. Based on this method, fuzzy membership function value of new data is selected with minimum difference from primary membership of train data and with the maximum level of fuzzification. In the first step, we define the problem as a nonlinear optimization problem. Then we introduce an efficient algorithm using critical points and obtain final membership function of train data. According to the proposed algorithm, the more distant data from the hyperplane will have a higher membership degree. If a datum exists on the hyperplane, it belongs to both classes with the same membership degree. Moreover, by comparing the primary membership degree of train data and calculated final distribution, we compute the level of noise for train data. Finally, we give a numerical example for illustration the efficiency of the proposed method and comparing its results with the results of the Lin and Wang approach.

Keywords: Fuzzy Logic, Machine learning, Data mining, Fuzzy support vector machine, fuzzy membership function

۱- مقدمه

یادگیری، یک رفتار مهم و یکی از موضوعات بسیار نزدیک به هوش انسان است که به درجهٔ کامل‌بودن اطلاعات پیشین ما بستگی دارد. یادگیری عبارت است از به‌دست‌آوردن دانش یا فهم از طریق مطالعه، آموزش یا تجربه و درواقع توانایی بهبود عملکرد بر اساس تجربیات و مشاهدات گذشته است. بنابراین، توانایی یادگیری ابزار قدرتمندی است که به‌طور گسترده در هوش مصنوعی مطرح شده است [16].

هدف یادگیری ماشین، ساخت سامانه‌های هوشمند با قابلیت انعطاف و هوشمندی بالاست که قادر به استخراج دانش (یادگیری) از محیط بوده و با شبیه‌سازی رفتار انسان، از تجربیات گذشته استفاده و یک مسأله را حل می‌کند [6].

امروزه داده‌کاوی نیز به‌عنوان زیرمجموعه‌ای از یادگیری ماشین، نقشی حیاتی در بازیابی اطلاعات[1] برای طبقه‌بندی مجموعه‌های بزرگ مستندات متنی و غیرمتنی دارد. طبقه‌بندی فرایندی است که مجموعهٔ داده‌ها را به قسمت‌های مشخص تقسیم و سازمان‌ها را قادر می‌کند که در حل مسائل خاص در مجموعه‌های بزرگ و پیچیده به کشف الگوها دست یابند؛ بنابراین، این فرایند یکی از عملیات رایج و مورداستفاده در داده‌کاوی است [4,17]. بسیاری از دانش‌ها همچون پزشکی [20]، پردازش تصویر و تشخیص بیماری، روان‌شناسی، اقتصاد و جامعه‌شناسی، علم کنترل، آمار، بازشناسی الگو و ... را به‌عنوان زیرشاخه‌ای از علم طبقه‌بندی می‌توان در نظر گرفت [4].

هدف از طبقه‌بندی داده‌ها، سازماندهی و تخصیص داده‌های گسسته به طبقه‌های مجزاست. در این فرآیند بر اساس داده‌های توزیع‌شده، الگوی اولیه‌ای ایجاد می‌شود؛ سپس این الگو برای طبقه‌بندی داده‌های جدید مورداستفاده قرار می‌گیرد؛ به‌این‌ترتیب، با به‌کارگیری الگوی به‌دست‌آمده، تعلق داده‌های جدید به طبقهٔ معینی قابل پیش‌گویی است [17].

روش ماشین بردار پشتیبان (SVM)[2] یک فن طبقه‌بندی مهم است که در دههٔ اخیر توجه زیادی از دانشمندان را به خود جلب کرده‌است [1,2,3,7,8,18,19,23,24].

مبنای کاری دسته‌بندی‌کنندهٔ SVM دسته‌بندی خطی داده‌هااست. درواقع در این روش، خطّی که حاشیه اطمینان بیشتری داشته‌باشد، به‌عنوان ابرصفحهٔ جداکنندهٔ داده‌ها انتخاب می‌شود.

ماشین‌های بردار پشتیبان یک طبقه‌بند کاراست؛ به‌ویژه زمانی که تعداد داده‌ها کم و اشتراکی بین طبقه‌ها نباشد [3]. نخستین الگوریتم برای طبقه‌بندی الگوها را فیشر [9] ارائه کرد و معیار این الگوریتم برای بهینه‌بودن، کم‌کردن خطای طبقه‌بندی الگوهای آموزشی بوده‌است. در ادامه، پژوهش‌گرِ روسی به نام *وپنیک*[3] گام بسیار مهمی در طراحی دسته‌بندی‌کننده‌ها برداشت و نظریهٔ آماری یادگیری را به‌صورت محکم‌تری بنا نهاد و ماشین بردار پشتیبان (SVM) را بر این اساس ارائه داد [24]. روش SVM بر پایهٔ ایدهٔ ساختاری کمترین ریسک[4] است [7] و هر داده نیز در یک طبقه قرار می‌گیرد. اگر چه SVMها ابزار مناسبی برای طبقه‌بندی داده‌ها هستند، محدودیت‌هایی در این روش وجود دارد. دسته‌بندی‌کنندهٔ خطی SVM با مشخص‌کردن ابرصفحهٔ جداکنندهٔ طبقه‌ها با بیشینهٔ حاشیه و با محاسبهٔ

[1] Retrieval information

[2] Support vector machine

[3] Vapnic

[4] SRM

موقعیت یک نقطه (دادهٔ آموزشی) نسبت به ابرصفحه، عضویت طبقهٔ یک داده را به‌صورت قطعی تعیین می‌کند؛ درحالی‌که در بسیاری از کاربردها، تعدادی از نمونه‌های آموزشی، مانند داده‌های پرت، به‌طوردقیق نمی‌توانند در یک طبقه عضویت داشته‌باشند؛ که این از تأثیر داده‌های پرت است. همچنین، تعدادی از نقاط داده‌ای نیز به‌خاطر نوفه یا اغتشاشات داده‌ای معیوب شده‌اند که در این حالت SVM باید قادر باشد آن‌ها را حذف کند. *لی*[1] و همکاران انتخاب بردارهای نامزد را برای داده‌های پرت پیشنهاد دادند [14]. *ژانگ*[2] نیز یک روش SVM مرکزی معرفی کرد [27].

ماشین بردار پشتیبان فازی (FSVM)[3] ارتقایافتهٔ روش SVM و یکی دیگر از روش‌های حل مسائل با داده‌های پرت است [10-12,15].

درواقع اهمیت هر نمونه در تصمیم‌گیری متفاوت است. SVM فازی که مبتنی بر منطق فازی است، در مرحلهٔ آموزش، اهمیت هر نمونه در میان نمونه‌های طبقه را با استفاده از تابع عضویت فازی نسبت‌به داده‌های پرت، بررسی می‌کند و تصمیم‌گیری را انجام می‌دهد [15]. *ترانگ*[4] و همکاران، یک تابع عضویت فازی را برای طبقه‌بندی خطی و غیرخطی پیشنهاد دادند [22].

تأکید این روش‌ها بر رویکرد فازی برای کاهش حساسیت داده‌های کم‌اهمیت است [13]. این رویکرد یک مقدار عضویت فازی به‌عنوان وزن هر دادهٔ آموزشی تعیین و از این وزن برای کنترل اهمیت دادهٔ متناظر استفاده می‌کند. تانگ[5] [21]، نیز *لی*[6] و همکاران [13]، یک روش برای محاسبهٔ عضویت فازی برای مشخص‌کردن اهمیت نقاط داده‌ای معرفی کردند.

در بیشتر موارد، عضویت فازی براساس فاصله بین هر نقطه داده و مرکز طبقهٔ متناظرش تعریف می‌شود[10-12,15]. در این مقاله روش تازه‌ای مبتنی بر منطق فازی، به‌منظور عضویت‌دهی داده‌های آموزشی، بر اساس فاصله از ابرصفحهٔ جداکننده ارائه می‌شود. مبتنی بر این روش پیشنهادی، مقدار تابع عضویت فازیِ هر دادهٔ تازه‌ای برای طبقه‌بندی، به گونه‌ای انتخاب می‌شود که کمترین میزان اختلاف را با عضویت اولیهٔ داده‌های آموزشی و بیشترین میزان فازی‌سازی را داشته‌باشد. در این روش پیشنهادی پس از واردشدن دادهٔ آزمون به مجموعه، آن را با دقت بالایی عضویت‌دهی کرده و داده‌های پرت و نوفه را تشخیص می‌دهد و حذف می‌کند.

در ادامهٔ مقاله، در بخش دوم تعاریف پیش‌زمینه‌ای از نظریهٔ مجموعهٔ فازی بیان می‌شود. در بخش سوم نیز، به شرح اصول و پایه‌های یک طبقه‌بندی‌کننده از نوع ماشین بردار پشتیبان پرداخته می‌شود. در بخش چهارم شرح مختصری از بردار پشتیبان فازی ارائه و روش عضویت‌دهی *لین و وانگ*[7] به‌عنوان یکی از روش‌های عضویت‌دهی مبتنی بر فاصله بین هر نقطهٔ داده و مرکز طبقهٔ متناظرش بیان می‌شود. در بخش پنجم نیز به‌منظور عضویت‌دهی داده‌های آموزشی، روش و الگوی تازه‌ای را ارائه و در انتها به شرح الگوریتم و گام‌های روش پیشنهادی پرداخته، با ارائهٔ یک مثال عددی به مقایسهٔ روش پیشنهادی با روش لین و وانگ پرداخته و کارایی روش پیشنهادی بررسی می‌شود.

۲- پیش‌زمینه

در این قسمت مفاهیم پایه‌ای و مقدماتی از نظریهٔ مجموعهٔ فازی که در بخش‌های بعدی برای ما مفید هستند، با استفاده از [26] و [25] بیان می‌شوند.

زاده[8] در روش خود بیان کرده‌است: «به نظر می‌رسد ما به‌دنبال نوع تازه‌ای از ریاضیات هستیم تا بتواند در حل مسائلی استفاده شود که تئوری احتمالات برای آن‌ها کارساز نیست» [25]. پس مسأله‌ای که زاده بر روی آن تأکید دارد، جایگزین‌شدن برخی کاربردهای غیرقابل‌حل با تئوری احتمالات و منطق فازی است. زاده به‌جای تئوری احتمالات «یک تئوری امکان» را مطرح می‌کند.

منطق فازی در حالت کلی به‌صورت زیر بیان می‌شود: «بیان دانش و قواعد خاص انجام یک کار با عدم قطعیّت به‌صورت دقیق».

تعریف ۱-۲: اعداد فازی[9]**:** عدد فازی $\tilde{A}$ ، یک مجموعه فازی روی خط اعداد حقیقی است که در شرط نرمال‌بودن و تحدّب صدق کند.

اعداد فازی دارای انواع مختلفی هستند؛ در این مقاله تمرکز ما بر روی عدد فازی مثلثی و ویژگی‌های آن است.

تعریف ۲-۲: عدد فازی مثلثی: یک عدد فازی $\tilde{A}=(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ یک عدد فازی مثلثی است که به‌ترتیب $\boldsymbol{a}$ شاخه چپ و $\boldsymbol{b}$ مرکز و $\boldsymbol{c}$ شاخه راست است. اگر $\boldsymbol{b}=\dfrac{\boldsymbol{a}+\boldsymbol{c}}{2}$، آن‌گاه $\tilde{A}$ را عدد فازی مثلثی متقارن می‌گویند.

---
[1] Li
[2] Zhang
[3] Fuzzy support vector machine
[4] Trung
[5] Tang
[6] Lee
[7] Lin & Wang
[8] Zadeh
[9] Fuzzy Numbers

تعریف ۳-۲: تابع عضویت اعداد مثلثی: برای یک عدد فازی مثلثی مانند $\tilde{A}=(a,b,c)$ تابع عضویت آن به‌صورت زیر تعریف می‌شود:

$$\mu_{\tilde{A}}(x)=\begin{cases}\left(\dfrac{x-a}{b-a}\right), & for \quad a\le x\le b\\ \left(\dfrac{x-c}{b-c}\right), & for \ b\le x\le c\\ 0, & otherwise.\end{cases} \qquad (۱)$$

تعریف ۴-۲: بعد VC[1] یک مجموعه از توابع، بیشینهٔ تعداد نقاطی است که می‌توانند در روش‌های ممکن مجزا شوند. برای بعد R^n، بعد VC می تواند n+1 در نظر گرفته شود.

۳- ماشین بردار پشتیبان

در این قسمت شرح مختصری از مسائل طبقه‌بندی SVM ارائه می‌شود [2,7,23,24].

فرض کنید تعدادی دادهٔ آموزشی به‌صورت $X=\{x_1,x_2,...,x_i,...,x_N\}$ داریم، که هر دادهٔ آموزشی یک بردار ویژگی d بُعدی بوده و دارای برچسب $y_i\in\{-1,1\}$ است. هدف، حل یک مسألهٔ دسته‌بندی دوطبقه‌ای به‌صورت بهینه است. اکنون فرض کنید بخواهیم این دو طبقه را با تابع تمایز $f\ (x)$ و به‌کمک ابر صفحهٔ H با معادلهٔ زیر از هم جدا کنیم:

$$H\ : w^T x + b = 0 \qquad (۲)$$
$$f\ (x) = sign(w\ x + b)$$

بردار وزن w، بردار نرمال ابرصفحه است و b مقدار بایاس وپنیک، ثابت کرد که بعد VC برای طبقه‌بندی‌کننده‌هایی از نوع ابرصفحات کانونی، دارای یک کران بالاست و این کران بالا با توان دوم نُرم بردار وزن، یعنی $\|w\|^2$، نسبت مستقیم دارد [24]. درواقع اگر $\|w\|^2$ محدود و کمینه شود، بعد VC طبقه‌بند را کمینه کرده‌ایم و تخمین ما از خطر واقعی به‌صورت احتمالی دقیق‌تر بوده و خاصیت تعمیم طبقه‌بند نیز بیشتر خواهدشد.

$$\|w\|=(\sum_{i=1}^{N}w_i^2)^{\frac{1}{2}} \qquad (۳)$$

رابطهٔ بین خاصیت تعمیم طبقه‌بند با نُرم بردار وزن $\|w\|$را می‌توان به طریق دیگر نیز توجیه کرد: فرض کنید داده‌های دو طبقه جدایی‌پذیر باشند و بردار ویژگی مرزی طبقهٔ نخست روی ابرصفحه بهینه H^+ و بردار ویژگی مرزی طبقهٔ دوم روی ابرصفحه H^- قرار گیرند. در این صورت، ابر صفحات H^+ و H^- به‌صورت زیر تعریف می‌شوند.

$$\begin{cases}H^+ : w\ x + b = +1,\\ H^- : w\ x + b = -1.\end{cases} \qquad (۴)$$

الگوهایی که بر روی ابرصفحات H^+ و H^- قرار می‌گیرند، بردار پشتیبان نامیده می‌شوند. ناحیهٔ بین این دو ابرصفحه را ناحیه مرزی یا حاشیه می‌نامند. همچنین، فاصلهٔ بین دو ابرصفحه که برابر $\frac{2}{\|w\|}$ است، زمانی بیشینه می‌شود که $\|w\|$ یا $\|w\|^2$ کمینه شود. هدف این است که الگوها درست طبقه‌بندی و بر روی ناحیهٔ مرزی یا خارج از آن واقع شوند، یعنی:

$$y_i(w\ x_i + b) \ge 0\ , \quad i=1,2,...,N \qquad (۵)$$

بنابراین، طراحی یک طبقه‌بندی‌کنندهٔ ابرصفحه‌ای با ناحیهٔ مرزی بهینه، به‌صورت مسألهٔ بهینه‌سازی زیر فرموله خواهدشد:

$$\begin{cases}minimize \quad \dfrac{\|w\|^2}{2}\\ s.t\\ y_i(w\ x_i + b) \ge 1\ , \quad i=1,2,...,N\end{cases} \qquad (۶)$$

الگوی بالا یک مسألهٔ بهینه‌سازی مقیّد از نوع محدب و درجهٔ دوم است. برای حلّ این مسأله تابع لاگرانژ را تشکیل داده و ضرایب را به‌دست می‌آوریم.

$$L(w,b,a)=\frac{1}{2}ww-\sum_{i=1}^{N}\alpha_i(y_i(wx_i+b)-1) \qquad (۷)$$

برای اینکه $(w\ ,b\ ,a)$ جواب مسأله باشد، باید در شرط KKT صدق کند. یعنی، به‌ازای جواب به‌دست‌آمده مشتق L نسبت به $w\ ,b\ ,a$ برابر صفر باشد. با مساوی صفر قراردادن مشتق به این معادلات خواهیم رسید:

$$\begin{cases}w=\sum_{i=1}^{N}\alpha_i y_i x_i\\ \sum_{i=1}^{N}\alpha_i y_i = 0\end{cases} \qquad (۸)$$

با قراردادن مقدار w از رابطهٔ بالا در مسألهٔ دوگان برای بهینه‌سازی مقید خواهیم داشت:

$$\begin{cases}\max imize \quad -\dfrac{1}{2}\sum_{i=1}^{N}\alpha_i\alpha_j y_i y_j x_i x_j+\sum_{i=1}^{N}\alpha_i\\ S.t \quad \alpha_i\ge 0 \qquad i=1,2,...,N,\\ \sum_{i=1}^{N}\alpha_i y_i = 0.\end{cases} \qquad (۹)$$

[1] Vapnic-Chervonen kis Dimension

پس از حل مسألۀ دوگان، ضرایب لاگرانژ $\alpha_i \geq 0$ به‌دست می‌آیند. در واقع هرکدام از ضرایب α_i متناظر با یکی از الگوهای x_i هستند. الگو های x_i که متناظر با ضرایب (مثبت) هستند، بردار پشتیبان sv_i نامیده می‌شوند.

مقدار b و w از رابطه (۱۰) به‌صورت زیر به‌دست می‌آیند:

$$\begin{cases} w = \sum_{i=1}^{N_{sv}} \alpha_i y_i sv_i \,, \\ b_j = y_j - \sum_{i=1}^{N_{sv}} \alpha_i y_i sv_i sv_j \,, \\ b = \frac{1}{N_{sv}} \sum_{i=1}^{N_{sv}} b_j \;. \end{cases} \tag{۱۰}$$

تابع تمایز برای طبقه‌بندی الگوی ورودی نیز به‌صورت زیر خواهد بود:

$$f(x) = sign(\sum_{i=1}^{N_{sv}} \alpha_i y_i sv_i sv_j) \tag{۱۱}$$

بنابراین اگر $f(x) > 0$ باشد، x متعلق به دسته نخست است و اگر $f(x) < 0$ باشد، x متعلق به دسته دوم است. یادآوری می‌شود که بنابر قیود مسأله، هیچ‌گاه $f(x) = 0$ نمی‌تواند اتفاق بیفتد.

۴- ماشین بردار پشتیبان فازی

اگرچه SVMها ابزار مناسبی برای طبقه‌بندی داده‌ها هستند، محدودیت‌هایی هم دارند. برای داده‌های ورودی $(x_1, y_1), \ldots, (x_l, y_l)$، طبق دسته‌بندی SVM استاندارد، این داده‌ها به‌طور دقیق به یکی از دو طبقه تعلق دارند. درحالی‌که در بسیاری از کاربردهای جهان واقعی، تأثیر نقاط آموزشی متفاوت است و ممکن است، چنین داده‌هایی با نسبت مشخصی به هر دو طبقه تعلق داشته‌باشند، یعنی، در مسائل طبقه‌بندی اغلب تعدادی از نمونه‌های آموزشی دارای اهمیت بیشتری از داده‌های دیگر هستند. ما نیازمند این هستیم که داده‌های آموزشی درست طبقه‌بندی و نباید برخی از این داده‌ها به‌عنوان نوفه مراقبت یا نادرست طبقه‌بندی شوند. در دسته‌بندی SVM استاندارد، اهمیت میزان خطا به‌ازای نمونه‌های آموزشی مختلف یکسان است، درحالی‌که به‌طورمنطقی نباید چنین باشد. با استفاده از منطق فازی می‌توان میزان اهمیت هر نمونه را در فاز آموزش وارد کرد [5]، یعنی هر نمونۀ آموزشی دقیقا متعلق به یک طبقه نیست و ممکن است با نسبت ۹۰٪ متعلق به یک طبقه و با نسبت ۱۰٪ به طبقۀ دیگر متعلّق باشد. به‌عبارت‌دیگر، یک تابع عضویت فازی $0 \leq \mu_i \leq 1$ وجود دارد که به هر دادۀ آموزشی x_i نسبت داده شده‌است. این عضویت فازی μ_i می‌تواند به‌عنوان تعلق یا کشش به سمت یک طبقه در مسائل طبقه‌بندی در نظر گرفته شود و $(1 - \mu_i)$ می‌تواند به‌عنوان بی‌معنی‌بودن در نظر گرفته شود. ماشین بردار پشتیبان فازی برای کاربردهایی که داده‌ها دارای ابهام بوده و الگوی مشخصی ندارند، مناسب است [15].

یکی از کاربردهای این الگو، در طبقه‌بندی هوشمند متن است. برای مثال، اگر جمله «رئیس جمهور روسیه، آقای پوتین در مسابقات جودو و شنا حضور یافت و موفق به کسب مقام در این مسابقات شد.» را در نظر بگیریم، نمی‌توان به‌صورت قطعی طبقۀ این متن را سیاسی و یا ورزشی بیان کرد؛ بلکه باید این‌طور مشخص شود که این متن با درجۀ عضویت 0.8 به طبقۀ متون ورزشی و با درجۀ عضویت 0.2 به طبقۀ متون سیاسی تعلق دارد.

بنابراین، در ادامه ضمن معرفی میزان اهمیت نمونه‌ها، روش لین و وانگ به‌عنوان یکی از روش‌های پایه در عضویت‌دهی فازی معرفی، سپس به منظور بهبود عضویت‌دهی فازی یک روش پیشنهادی برای عضویت‌دهی به داده‌ها، مبتنی بر فاصله از ابرصفحۀ جداکننده ارائه می‌شود.

۴-۱- افزودن میزان اهمیت نمونه‌ها به الگو

SVM استاندارد، نمونه‌های آموزشی را به‌صورت زوج‌های (x_i, y_i) در نظر می‌گیرد به‌طوری‌که $y_i \in \{-1,1\}$. حال برای بررسی میزان اهمیت هر نمونه، نمونه‌ها به‌صورت سه‌تایی (x_i, y_i, μ_i) در نظر گرفته می‌شوند:

μ_i درواقع میزان درجۀ عضویت نمونه x_i به طبقۀ خودش است. برای بررسی اهمیت هر نمونه، فرمول SVM را به‌صورت زیر تغییر می‌دهیم [15]:

$$\begin{cases} Minimize \; \frac{1}{2} w \,.\; w^T + C \sum_{i=1}^{l} \mu_i \varepsilon_i \\ s.t \\ y_i (w . x_i + b) \geq 1 - \varepsilon_i \,,\; i = 1, \ldots N \\ \qquad \varepsilon_i \geq 0 \qquad i = 1, \ldots N \end{cases} \tag{۱۲}$$

با این تغییر، دوگانِ مسأله بهینه‌سازی در SVM به‌صورت زیر خواهد شد:

$$\begin{cases} \max imize\ -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j x_i x_j + \sum_{i=1}^{N}\alpha_i \\ S.t \quad 0\le\alpha_i\le\mu_i C\ , \\ \sum_{i=1}^{N}\alpha_i y_i = 0 \qquad i=1,2,...,N. \end{cases} \qquad (۱۳)$$

در واقع تفاوت SVM فازی با SVM استاندارد در این است که حد بالای ضرایب لاگرانژ α_i برابر $\mu_i C$ است. درحالی‌که در SVM استاندارد این حد برابر C است. بردارهای پشتیبان الگوهایی خواهند بود که ضرایب لاگرانژ متناظر آنها در رابطه $0\le\alpha_i\le\mu_i C$ صدق کند. تعدادی از بردارهای پشتیبان که ضرایب لاگرانژ متناظر آنها در رابطه $0\le\alpha_i\le\mu_i C$ صدق کند، برای محاسبه b استفاده می‌شود.

۲-۴- تابع عضویت لین و وانگ

یکی از روش‌های پایه که به معرفی تابع عضویت برای FSVM می‌پردازد، روش لین و وانگ است [15] . این روش براساس فاصلهٔ بین یک نقطه و مرکز طبقه، تابع عضویت فازی را تعریف می‌کند.

فرض کنید مجموعه‌ای از نمونه‌های آموزشی را به‌صورت $S=\{(x_1.y_1.s_1).\dots.(x_l.y_l.s_l)\}$ در اختیار داشته باشیم که در آن مرکز طبقهٔ 1+ را با x_+ و مرکز طبقهٔ 1- را با x_- نمایش می‌دهیم.

بنابراین شعاع طبقهٔ 1+ برابر

$$r_+ = \max\{|x_+ - x_i| \quad |y_i = +1\}$$

و نیز شعاع طبقهٔ 1- برابر

$$r_- = \max\{|x_- - x_i| \quad |y_i = -1\}$$

است.

تابع عضویت فازی معرفی‌شده توسط لین و وانگ برپایه مرکز و شعاع هر طبقه است که برابر با:

$$\mu_i = \begin{cases} 1-\frac{|x_+ - x_i|}{(r_+ + \delta)} & .y_i = +1 \\ 1-\frac{|x_- - x_i|}{(r_- + \delta)} & .y_i = -1 \end{cases} \qquad (۱۴)$$

است. به‌طوری‌که برای اینکه $\mu_i = 0$ نشود، فرض می‌کنیم $\delta > 0$ است.

۵- روش پیشنهادی

در روش ماشین بردار پشتیبان استاندارد، با مشخص‌کردن ابرصفحهٔ جداکنندهٔ طبقه‌ها و با محاسبهٔ موقعیت یک نقطه نسبت‌به ابرصفحه، تعیین می‌شود که این نقطه (داده) به کدام طبقه تعلق دارد.

اما سؤالی که پیش می‌آید این است که تا چه حد می‌توان به این تصمیم‌گیری مطلق بر اساس ابرصفحه اطمینان کرد. به‌طور مثال، یک نقطه روی خود ابرصفحه به کدام طبقه تعلق دارد؟

منطقی است که بگوییم عضویت یک نقطه روی ابرصفحه به اندازهٔ 0.5 به هر دو طبقه تعلق دارد؛ اما برای یک نقطه که بسیار نزدیک به ابرصفحه (یعنی مرز میان دو طبقه) است، چه تصمیمی باید گرفت؟ روش ماشین بردار پشتیبان کلاسیک پیشنهاد می‌کند که این نقطه به‌طوردقیق به یکی از طبقه‌ها می‌تواند تعلق داشته‌باشد؛ اما منطقی است که ما انتظار داشته‌باشیم در نزدیکی مرز میان دو طبقه، نقاط با توزیع مناسبی به هر دو طبقه تعلق داشته باشند و نه به‌طور کامل به یک طبقه.

بنابراین هدف این مقاله توزیع عضویت میان دو طبقه است؛ به‌طوری‌که عضویت تازه کمترین اختلاف را با عضویت اولیه داده‌های آموزشی داشته‌باشند و در نتیجه، به‌دست‌آوردن تابع عضویتی است که اثرات نوفه و داده‌های پرت را کاهش می‌دهد.

این روش کاربردهای بسیار زیادی در علوم پزشکی، مهندسی و ... دارد. به‌عنوان نمونه، روش پیشنهادی به‌عنوان یک روش پایه در مواردی مانند پزشکی، برای تشخیص بیماری با استفاده از داده‌های آزمایشگاهی بیمار و نیز با استفاده از مرز سلامتی و بیماری به عضویت‌دهی داده‌ها می‌پردازد. پس از عضویت‌دهی اهمیت داده‌های آزمایشگاهی متناسب با افزایش فاصله از مرز سلامتی و بیماری مشخص می‌شود و تشخیص بیماری و یا سلامتی بسیار دقیق‌تر انجام می‌شود. در ادامه الگوی ریاضی این روش بیان می‌شود:

۱-۵- بیان ریاضی الگو

مقادیر A و B به‌صورت زیر در نظر گرفته می‌شود:

$$B=\{q^j\}_{j=1}^{l} \text{ و } A=\{p_i\}_{i=1}^{k}$$

که در آن $p_i.q^j \in R^N$. برای این نقاط آموزشی عضویت‌های پیش‌فرض μ_i و μ^j را به‌ترتیب اختصاص می‌دهیم. حال به‌کمک روش‌های کلاسیک ماشین بردار پشتیبان، ابر صفحه $H=a_0+a_1x_1+\dots+a_Nx_N$ را مشخص می‌کنیم. هدف، به‌دست آوردن توزیع فازی است؛ به‌طوری‌که عضویت تازه‌ای را برای تصمیم‌گیری در مورد داده‌ها به ما پیشنهاد می‌دهد.

این فازی‌سازی به‌کمک اعداد مثلثی متقارن $(-r,0,r)$ خواهدبود که در آن هدف مسأله، پیداکردن r است.

ابتدا توضیح می‌دهیم که اگر چنین rای را به‌دست آوریم عضویت داده‌ها چگونه متناسب با ابر صفحهٔ جداکننده H و عدد فازی بالا صورت خواهد گرفت.

در این صورت تابع عضویت فازی x به‌صورت زیر تعریف می‌شود:

$$\mu_r(x)=\begin{cases}\frac{1}{2}+\frac{|H(x)|}{2r} & ,|H(x)|<r\\ 1 & ,|H(x)|\geq r\end{cases} \tag{۱۵}$$

روش بالا، برخلاف روش ۰ و ۱ کلاسیک ماشین بردار پشتیبان یک توزیع عضویت پیوسته و معقول را برای تصمیم‌گیری به ما می‌دهد. بنابراین، تابع هدف به‌گونه‌ای تعریف می‌شود که توزیع عضویت نهایی کمترین اختلاف را با توزیع اولیه عضویت داده‌های آموزشی داشته‌باشد. همچنین، برای افزایش خصوصیات فازی در میان انتخاب‌هایی با کمترین اختلاف در عضویت، بزرگ‌ترین r انتخاب می‌شود.

بنابراین مسألهٔ بهینه‌سازی زیر برای محاسبه r پیشنهاد می‌شود.

$$Max\ \min_r\{\sum_{i=1}^{k}\|\mu_r(p_i)-\mu_i\|^2 + \sum_{j=1}^{l}\|\mu_r(q^j)-\mu^j\|^2\} \tag{۱۶}$$

۲-۵- روش حل

برای حل مسألهٔ بهینه‌سازی که پیشتر به آن اشاره شد، ازآنجاکه هیچ قیدی در این مسأله وجود ندارد، کافی‌ست نقاط بحرانی یک تابع یک‌متغیره بررسی و در میان این نقاط، نقطهٔ بهینه محاسبه شود.

خوشبختانه تابع هدف دارای تعدادی متناهی نقطهٔ بحرانی است که الگوریتم کارا و مناسبی برای محاسبهٔ آن پیشنهاد می‌کند.

ابتدا تأکید می‌شود که $r=0$ نقطه ابتدایی بازه است، زیرا $r\geq 0$. با بازنویسی تابع هدف سعی در تعیین بقیهٔ نقاط بحرانی می‌شود:

$$\begin{aligned}
\varphi(r)&=\{\sum_{i=1}^{k}\|\mu_r(p_i)-\mu_i\|^2+\sum_{j=1}^{l}\|\mu_r(q^j)-\mu^j\|^2\}\\
&=\{\sum_{|H(p_i)|<r}\|\mu_r(p_i)-\mu_i\|^2\\
&+\sum_{|H(q^j)|<r}\|\mu_r(q^j)-\mu^j\|^2\\
&+\sum_{|H(p_i)|\geq r}\|\mu_r(p_i)-\mu_i\|^2\\
&+\sum_{|H(q^j)|\geq r}\|\mu_r(q^j)-\mu^j\|^2\}\\
&=\{\sum_{|H(p_i)|<r}\left\|\frac{1}{2}+\frac{|H(p_i)|}{2r}-\mu_i\right\|^2\\
&+\sum_{|H(q^j)|<r}\left\|\frac{1}{2}+\frac{|H(q^j)|}{2r}-\mu^j\right\|^2\\
&+\sum_{|H(p_i)|\geq r}\|1-\mu_i\|^2+\sum_{|H(q^j)|\geq r}\|1-\mu^j\|^2\}\\
&=\{\sum_{|H(p_i)|<r}(\frac{1}{4}+\frac{|H(p_i)|^2}{4r^2}+\mu_i^2\\
&+\frac{|H(p_i)|}{2r}-\mu_i\frac{|H(p_i)|}{2r}-\mu_i)\\
&+\sum_{|H(q^j)|<r}(\frac{1}{4}+\frac{|H(q^j)|^2}{4r^2}+(\mu^j)^2\\
&+\frac{|H(q^j)|}{2r}-\mu^j\frac{|H(q^j)|}{2r}-\mu^j)\\
&+\sum_{|H(p_i)|\geq r}\|1-\mu_i\|^2+\sum_{|H(q^j)|\geq r}\|1-\mu^j\|^2\}
\end{aligned} \tag{۱۷}$$

حال باید نقاطی که تابع هدف در آن تغییر ضابطه می‌دهد و امکان آن وجود دارد که هموار نباشد، به‌عنوان نقطه‌هایی که ممکن است، بحرانی باشند، فهرست شوند. با توجه به ضابطهٔ تابع هدف مشخص می‌شود که به‌ازای rهایی که با مقدار تابع مشخص ابرصفحه در یکی از داده‌های نمونه برابری می‌کند، احتمال شکستگی (از بین رفتن همواری) وجود دارد؛ پس $\{r\,|\,\exists i, r=|H(p_i)|\}$ همچنین $\{r\,|\,\exists j, r=|H(q^j)|\}$ انتخابی دیگر از نقاط بحرانی به ما می‌دهند.

با محاسبهٔ مشتق تابع هدف، بقیهٔ نقاط بحرانی شناسایی می‌شود.

$$\begin{aligned}
\dot{\varphi}(r)&=\sum_{|H(p_i)|<r}(-\frac{|H(p_i)|^2}{2r^3}-\frac{|H(p_i)|}{2r^2}+\mu_i\frac{|H(p_i)|}{r^2})\\
&+\sum_{|H(q^j)|<r}(-\frac{|H(q^j)|^2}{4r^2}-\frac{|H(q^j)|}{2r^2}+\mu^j\frac{|H(q^j)|}{r^2})\\
&=(\sum_{|H(p_i)|<r}(\mu_i|H(p_i)|-\frac{|H(p_i)|}{2})\\
&+\sum_{|H(q^j)|<r}(\mu^j|H(q^j)|-\frac{|H(q^j)|}{2}))\frac{1}{r^2}\\
&-(\sum_{|H(p_i)|<r}\frac{|H(p_i)|^2}{2}+\sum_{|H(q^j)|<r}\frac{|H(q^j)|^2}{2})\frac{1}{r^3}
\end{aligned} \tag{۱۸}$$

برای به‌دست آوردن نقاطی که دارای مشتق احتمالی صفر هستند (با شرط $r>0$)، در عبارت بالا، طرفین در r^3 ضرب و مقدار r به‌دست می‌آید.

$$r=\frac{((\sum_{|H(p_i)|<r}\mu_i|H(p_i)|-\frac{|H(p_i)|}{2})+(\sum_{|H(q^j)|<r}\mu^j|H(q^j)|-\frac{|H(q^j)|}{2}))}{(\sum_{|H(p_i)|<r}\frac{|H(p_i)|^2}{2}+\sum_{|H(q^j)|<r}\frac{|H(q^j)|^2}{2})} \tag{۱۹}$$

$$\{|H(p_1)|,\ldots,|H(p_k)|,|H(q^1)|,,\ldots,|H(q^k)|\} = \{h_1 < \ldots < h_n\}.$$

حال با تعریف

$$I_t = \{i \mid |H(p_1)| \le h_t\},\ t = 1,\ldots,n$$
$$J_t = \{j \mid |H(q^j)| \le h_t\},\ t = 1,\ldots,n \qquad (۲۰)$$

می‌توان مقادیر بحرانی را این‌گونه محاسبه کرد:

$$r_t = \frac{\left(\sum_{i\in I_t} \mu_i |H(p_i)| - \frac{|H(p_i)|}{2}\right) + \left(\sum_{j\in J_t} \mu^j |H(q^j)| - \frac{|H(q^j)|}{2}\right)}{\left(\sum_{i\in I_t} \frac{|H(p_i)|^2}{2} + \sum_{j\in J_t} \frac{|H(q^j)|^2}{2}\right)} \qquad (۲۱)$$

که درآن $t = 1,\ldots,n$ است.

پس کافی‌ست مسأله بهینه‌سازی برای مقادیر متناهی زیر بررسی شود:

$$\{0,|H(p_1)|,\ldots,|H(p_k)|, |H(q^1)|,\ldots,|H(q^l)|,r_1,\ldots,r_t\} \qquad (۲۲)$$

با محاسبهٔ تابع هدف یعنی φ روی مقادیر بالا، یک مجموعه از مقادیر کمینه برای r به‌دست می‌آید، یعنی مقادیری که φدر آنها کمینه است. اگر این مجموعه مقادیری که φرا کمینه می‌کند $S = \{s_1,\ldots,s_2\}$ باشد، جواب نهایی مسأله برای r مقدار $\max\{s_1,\ldots,s_2\}$ است.

۳-۵- الگوریتم و گام‌های حل

در این قسمت مراحل اجرای روش پیشنهادی، به‌صورت گام‌به‌گام توضیح داده خواهدشد:

گام ۱. بعد از گرفتن داده‌های طبقهٔ A و B، مسأله با روش ماشین بردار پشتیبان حل و ابرصفحه یافته می‌شود.

گام ۲. مقادیر $|H(q^\xi)|,|H(p_\nu)|$ با استفاده از ابرصفحه به‌دست می‌آید.

گام ۱-۲. مقادیر به‌صورت صعودی مرتب می‌شود.

گام ۳. مجموعه‌های اندیس I_t و J_t مشخص می‌شود.

گام ۱-۳. با استفاده از مقادیر عضویتμ_i وμ^j و مجموعه‌های اندیس I_t و J_t نقاط بحرانی r_t به‌دست می‌آید.

گام ۴. نقاط بحرانی به‌صورت زیر فهرست می‌شود:

$$\{0,|H(p_1)|,\ldots,|H(p_\xi)|, |H(q^1)|,\ldots,|H(q^\nu)|,r_1,\ldots,r_t\}$$

گام ۵. مقدار عضویت$\mu_r(p_i)$ و $\mu_r(q^j)$ به‌ازای نقاط بحرانی محاسبه می‌شود.

گام ۶. مقدار

$$\min\{\sum_{i=1}^{\xi}\|\mu_r(p_i)-\mu_i\|^2+\sum_{j=1}^{\nu}\|\mu_r(q^j)-\mu^j\|^2\}$$

به‌دست می‌آید.

گام ۷. از مقادیر به‌دست‌آمده در گام پیش max گرفته و مقدار r مشخص می‌شود.

برای حالتی که داده‌ها همگی غیر فازی باشند، می‌توان مقدار r را به‌سادگی از قضیهٔ زیر محاسبه کرد.

قضیه ۱: مقادیر دو طبقهٔ A و B به‌صورت زیر در نظر گرفته می‌شود:

$$B = \{q^j\}_{j=1}^{m} \text{ و } A = \{p_i\}_{i=1}^{n}$$

که با ابرصفحهٔ H از هم جدا شده‌باشند و $\{\mu_i\}_{i=1}^{n}$ و $\{\mu^j\}_{j=1}^{m}$ مقادیر عضویت متناظر داده‌های این طبقه‌ها باشند. اگر $\mu_i = \mu^j = 1$ برای $1 \le i \le n$ و $1 \le j \le m$، آنگاه r بهینه از رابطه زیر به‌دست می‌آید:

$$r = \min\{|H(p_1)|,\ldots,|H(p_k)|,|H(q^1)|,,\ldots,|H(q^k)|\} = \min\{h_1 < \ldots < h_n\}.$$

اثبات:

فرض می‌کنیم r مانند بالا تعریف شده‌باشد. در این‌صورت برای هر s که $0 \le s \le r$ داریم:

$$\mu_S(p_i) = \mu_i = 1,\ \ 1 \le i \le n$$
$$\mu_S(q^j) = \mu_j = 1,\ \ 1 \le j \le n$$

بنابراین، برای تابع هدف داریم $\varphi(s) = 0$؛ بنابراین، به‌ازای $0 \le s \le r$ کمینه است.

مقدار s را به این صورت در نظر می‌گیریم که $s > r$ باشد. دراین‌صورت یک داده مانند x وجود دارد به‌طوری‌که $x \in \{p_1,p_2,\ldots,p_n,q^1,q^2,\ldots,q^m\}$و $|H(x)| < s$

بنابراین داریم:

$$\mu_s(x) = 0.5 + \frac{|H(x)|}{2s} < 1$$

که این نشان می‌دهد $\varphi(s) > 0$؛ بنابراین، برای $s > r$، تابع φ کمینه نخواهدبود. در نتیجه، این نشان می‌دهدr جواب بهینه است.

۴-۵- شناسایی نوفه و داده‌های پرت

از آنجا که توزیع عضویت برای داده‌های فضا به‌گونه‌ای صورت گرفته که کمترین میزان اختلاف با عضویت نخستین داده‌های آموزشی را داشته باشد، بنابراین منطقی است، داده‌هایی که

دارای اختلاف عضویتی نامعقول هستند، به‌عنوان نوفه یا دادهٔ پرت در نظر گرفته شوند.

متناسب با کاربردهای گوناگون، می‌توان میزانی از اختلاف را به‌عنوان ε در نظر گرفت و ε-نوفه را معرفی کرد.

تعریف ۱-۴-۵: در نظر می‌گیریم که x یک داده تمرینی با عضویت نخستین μ باشد. اگر توزیع عضویت بهینه به‌کمک r به‌دست آمده باشد، آنگاه x یک $\varepsilon-$ نوفه برای $\varepsilon > 0$ نامیده می‌شود، هرگاه داشته‌باشیم:

$$\left|\mu - \mu_r(x)\right| \geq \varepsilon$$

بنابراین، می‌توان با استفاده از روش پیشنهادی و با مقایسهٔ عضویت نخستین داده‌های آموزشی و درجهٔ عضویت نهایی آن از رابطهٔ (۱۵)، یک داده پرت یا نوفه را تشخیص داد.

۶- مثال عددی

در این بخش از مقاله با ارائهٔ مثالی، به مقایسهٔ روش پیشنهادی و روش لین و وانگ پرداخته می‌شود.

داده‌های مورداستفاده در این بخش در جداول (۱و۲) بیان شده‌اند.

(جدول-۱): داده‌های طبقهٔ A
(Table-1): Class A data

داده‌های طبقهٔ A		
X	Y	درجه عضویت
9.5	9.2	0.94
4	3	0.94
10.5	13.2	0.99
8	10	0.85
0	20	0.65
6	20	0.9
10	0	0.79
4	25	0.89
5	12	0.75

(جدول-۲): داده‌های طبقهٔ B
(Table-2): class B data

داده‌های طبقهٔ B		
X	Y	**درجه عضویت**
1	2	0.57
-2	5	0.66
-2	-5	0.8
-6	15	0.69
-5	7	0.76
0	-5	0.72
1	-3	0.64
-7	3	0.9

گام ۱ : با استفاده از برنامه‌ریزی درجهٔ دوم مسأله را حل و ابرصفحهٔ بهینه به‌دست می‌آید.

$$H(x,y) = 0.6\,x + 0.2\,y - 2$$

گام ۲: مقادیر $|H(q^{\xi})|,|H(p_v)|$ با استفاده از ابرصفحه، در جدول (۳) فهرست می‌شود.

گام ۱-۲: اعداد به‌دست آمده در گام ۲ به‌صورت صعودی در جدول (۴) مرتب می‌شود.

گام۳ : در این گام مقادیر r_t به‌صورت زیر به‌دست می‌آید.

r_t = {0.0358, 0.382, 0.449, 0.0542, 0.0675, 0.0843, 0.1037, 0.1216, 0.1486}

گام ۴: با استفاده از این گام نقاط بحرانی فهرست می‌شوند.

Critical Points =
{1.0000, 2.0000, 2.2000, 2.6000, 3.0000, 3.4000, 3.6000, 4.0000, 4.2000, 4.8000, 5.4000, 5.5400, 5.6000, 6.9400, 0.0358, 0.0382, 0.0449, 0.0542, 0.0675, 0.0843, 0.1037, 0.1216, 0.1486, 0}

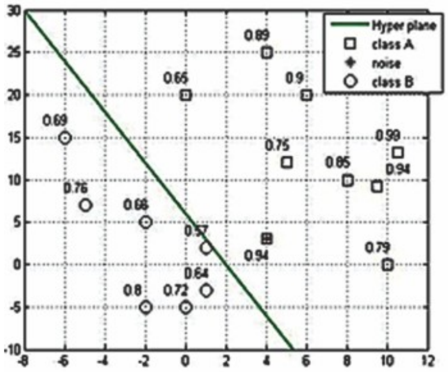
(جدول-۳): مقادیر $|H(q^{\xi})|,|H(p_v)|$
(Table-3): value of $|H(q^{\xi})|,|H(p_v)|$

$\lvert H(q^{\xi})\rvert,\lvert H(p_v)\rvert$
5. 5400
1.0000
6.9400
4.8000
2.0000
5.6000
4.0000
5.4000
3.4000
1.0000
2.2000
4.2000
2.6000
3.6000
3.0000
2.0000

(جدول-۴): مقادیر $|H(q^{\xi})|,|H(p_v)|$ **, به‌صورت صعودی**
(Table-4): Ascending order value of $|H(q^{\xi})|,|H(p_v)|$

$\lvert H(q^{\xi})\rvert,\lvert H(p_v)\rvert$
1.0000
2.0000
2.2000
2.6000
3.0000
3.4000
3.6000
4.0000
4.2000
4.8000
5.4000
5.5400
5.6000
6.9400

گام ۵: با استفاده از نقاط بحرانی به‌دست‌آمده درگام پیش، مقادیر تابع هدف را را به‌دست‌آورده و در جدول (۵) فهرست می‌شود.



(شکل-۱): نمایش داده‌ها همراه با مقدار عضویت- داده نوفه.
(Figure-1): data and noise data membership degree

(جدول-۵): مقادیر تابع هدف با استفاده از نقاط بحرانی
(Table-5): Target function values using critical points

ردیف	نقاط بحرانی	مقدار تابع هدف
1	1.0000	0.9832
2	2.0000	0.8632
3	2.2000	0.8043
4	3.0000	0.5879
5	3.4000	0.5014
6	3.6000	0.4592
7	4.0000	0.3789
8	4.2000	0.3416
9	4.8000	0.2492
10	5.4000	0.1910
11	5.5400	0.1802
12	5.6000	0.1756
13	6.9400	0.1373
14	0.0358	0.9832
15	0.0382	0.9832
16	0.0449	0.9832
17	0.0542	0.9832
18	0.0675	0.9832
19	0.0843	0.9832
20	0.1037	0.9832
21	0.1216	0.9832
22	0.1486	0.9832
23	0	0.9832

گام ۶: با استفاده از MaxMin مقدار $r = 6.9400$ می‌شود.

در انتها دوباره نقاط آموزشی به‌عنوان دادهٔ آزمون وارد شده و برای آن درجهٔ عضویت جدید (درجه عضویت نهایی) با استفاده از تابع عضویت (۱۵) به‌دست می‌آید (جدول۶).

(جدول-۶): محاسبه درجهٔ عضویت نهایی از روش پیشنهادی
(Table-6): calculating final membership degree using the proposed method

x	y	درجه عضویت اولیه	درجه عضویت نهایی از روش پیشنهادی
9.5	9.2	0.9400	0.8991
4	3	0.9400	0.5720
10.5	13.2	0.9900	1.0000
8	10	0.8500	0.8458
0	20	0.6500	0.6441
6	20	0.6500	0.6441
10	0	0.7900	0.7882
4	25	0.8900	0.8890
5	12	0.7500	0.7450
1	2	0.5700	0.5720
-2	5	0.6600	0.6585
-2	-5	0.8000	0.8026
-6	15	0.6900	0.6873
-5	7	0.7600	0.7594
0	-5	0.7200	0.7161
1	-3	0.6400	0.6441
-7	3	0.9000	0.9276

داده‌ها همراه با مقدار عضویت اولیهٔ آن‌ها و همچنین ابرصفحهٔ جداکنندهٔ این داده‌ها با استفاده از نرم‌افزار متلب در شکل (۱) رسم شده‌است.

$\varepsilon = 0.05$ را در نظر می‌گیریم. اکنون با توجه به قسمت ۴-۵ و نیز مقادیر به‌دست‌آمده درگام ۶ مشاهده می‌شود که داده (4,3) باتوجه‌به‌اینکه در مجاورت ابرصفحه قرار دارد، باید مقدارعضویتی نزدیک به 0.6 داشته‌باشد، درحالی‌که پس از آموزش به‌عنوان نوفه تشخیص داده شده و حذف می‌شود. همچنین :

$$\left|\mu - \mu_r(x)\right| = \left|0.9400 - 0.5720\right| = 0.36 \geq \varepsilon$$

حال مقدار عضویت نهایی داده‌ها که از رابطهٔ (۱۵) به‌دست آمده‌است، با مقدار عضویت روش ارائه‌شده توسط لین و وانگ در جدول (۷) مقایسه می‌شود:

(جدول-۷): مقایسهٔ درجهٔ عضویت نهایی از روش پیشنهادی با روش لی و وانگ
(Table-7): comparing final membership degree using the proposed and Lin and Wang methods

X	Y	درجه عضویت نهایی از روش پیشنهادی	درجه عضویت نهایی از روش لین و وانگ
9.5	9.2	0.8991	0.7562
4	3	0.5720	0.8549
10.5	13.2	1.0000	0.7149
8	10	0.8458	0.8211
0	20	0.6441	0.6433
6	20	0.6441	0.6321
10	0	0.7882	0.6852
4	25	0.8890	0.5029
5	12	0.7450	0.8534
1	2	0.5720	0.8326
-2	5	0.6585	0.8572
-2	-5	0.8026	0.6165
-6	15	0.6873	0.6866
-5	7	0.7594	0.7935
0	-5	0.7161	0.6303
1	-3	0.6441	0.6910
-7	3	0.9276	0.7045

با مقایسهٔ روش پیشنهادی و روش لین و وانگ، مشاهده می‌شود که روش ما بهینه‌تر و کاراتر است.

اختلاف درجهٔ عضویت نهایی دو روش لین و وانگ و روش پیشنهادی با درجهٔ عضویّت اولیه در جدول (۸) فهرست شده‌است.

(جدول-۸): اختلاف درجه عضویت نهایی دو روش لین و وانگ و روش پیشنهادی با درجه عضویت اولیه

(Table-8): difference in final membership degrees from Lin and Wang and the proposed methods from primary membership degrees

x	Y	اختلاف درجه عضویت نهایی از روش پیشنهادی و درجه عضویت اولیه	اختلاف درجه عضویت نهایی از روش لین و وانگ و درجه عضویت اولیه
9.5	9.2	0.0409	0.1538
4	3	0.3680	0.0851
10.5	13.2	0.0100	0.1351
8	10	0.0042	0.0289
0	20	0.0059	0.0067
6	20	0.0059	0.0179
10	0	0.0018	0.1048
4	25	0.0010	0.3871
5	12	0.0050	0.1034
1	2	0.0020	0.2626
-2	5	0.0015	0.1972
-2	-5	0.0026	0.1835
-6	15	0.0027	0.0032
-5	7	0.0006	0.0335
0	-5	0.0039	0.0897
1	-3	0.0041	0.0510
-7	3	0.0276	0.1955

۶-۱- مزایا و ویژگی‌های روش ما نسبت به روش کلاسیک:

الف) در میان روش‌های موجود، روش این مقاله نخستین روشی است که عضوت اولیه داده‌های تمرینی را در تابع عضویت اثر داده و نزدیک ترین تابع را برای تقریب این عضویت‌ها پیشنهاد می‌دهد.

ب) هنگامی که دادهٔ جدیدی به مسأله داده می‌شود، می‌توان به دادهٔ جدید عضویتی نظیر کرد و تشخیص داد که با چه درجه عضویتی متعلق به طبقه است.

ج) اگر عضویت اولیه دادهٔ تمرینی، اختلاف زیادی با عضویت محاسبه‌شده توسط تابع عضویت داشته‌باشد، می‌توان به آن نوفه اطلاق و آن را حذف کرد.

۷- نتیجه‌گیری

در این مقاله روش جدیدی برای طبقه‌بندی داده‌ها با استفاده از ماشین بردار پشتیبان فازی ارائه شد که در آن می‌توان برای داده‌های عضو طبقهٔ درجهٔ عضویتی بر اساس فاصله از ابرصفحهٔ جداکننده، نظیر کرد و هر دادهٔ ورودی را با توجه به عضویت و میزان تعلّقش به طبقه به‌شکلی درست طبقه‌بندی کرد. این روش در مقایسه با سایر روش‌های طبقه‌بندی ماشین بردار پشتیبان بسیار کاراتر و بهینه‌تر است. همچنین در این روش می‌توان داده‌های نوفه‌ای را به‌راحتی تشخیص داد و حذف کرد.

8- References ۸- مراجع

[۱] محمدامین، فولادی کاظم. بازشناسی برخط حروف مجزای دست‌نویس فارسی بر اساس تشخیص گروه بدنه اصلی با استفاده از ماشین بردار پشتیبان .پردازش علائم و داده‌ها. ۱۳۹۱; ۹ (۱) :۵۹-۶۸

[1] Mehralian M A, kazem fouladi K. the recognition of online handwritten persian characters based on their main bodies using svm. JSDP. 2012; 9 (1) :59-68

[۲] منتظر غلامعلی، شایسته‌فر محمد. شناسایی پلاک خودروهای ایرانی با الگوریتم ماشین بردار پشتیبانی فازی .پردازش علائم و داده‌ها. ۱۳۹۴; ۱۲ (۱) :۴۷-۵۶.

[2] Montazer G A, shayestehfar M. iranian license plate identification with fuzzy support vector machine . JSDP. 2015; 12 (1) :47-56

[3] Abe S. Pattern Classification: Neuro-Fuzzy Methods And Their Comparison. springer-verlag, London, UK, 2001.

[4] Alpayden E. Introduction To Machine Learning. The MIT Press, 2010.

[5] Bezdek J. C. Fuzzy Mathematics In Pattern Classification. Phd Dissertation, Cornell University, Ithaca, NY, 1973.

[6] Bishop CM. Pattern Recognition And Machine Learning. springer; 2006.

[7] Burges J. C. a tutorial on support vector machinees for pattern recognition. Data Mining and Knowledge Discovery 1998; 2(2): 121-167.

[8] Cortes C., Vapnik V. support-vector networks. Machine Learning 1995; 20(3): 273-297.

[9] Fisher RA. the use of multiple measurements in taxonomic problems. Annals of eugenics. 1936 Sep;7(2):179-88.

[10] Huang H. P., Liu Y. H. fuzzy support vector machine for pattern recognition and data mining. Int J Fuzzy Syst 2002; 4(3): 826–835.

[11] Inoue T., Abe S. fuzzy support vector machines for pattern classification. In Proceeding of IJCNN 2001; 2: 1449–1454.

[12] Jiang X. F., Yi, Z., Lv J. C. fuzzy svm with a new fuzzy membership function. Neural Compute 2006; 15(3-4): 268–276.

[13] Lee, G.H., Taur, J. S., Tao, C.W. a robust fuzzy support vector machine for two-class pattern classification. International Journal of Fuzzy Systems 2006; 8(2): 76-87.

[14] Li M. Q., Chen F. Z., Kou J. S. candidate vectors selection for training support vector machines. IEEE Computer Society, Third International

Conference on Natural Computation (ICNC) 2007; 1: 538-542.

[15] Lin C. F., Wang S.D. fuzzy support vector machines. IEEE Trans. on Neural Networks 2002; 13(2): 464-471.

[16] Mitchell T. Machine Learning. McGraw Hill, ISBN 0-07-042807-7, 1997.

[17] Nilsson N. J. Introduction To Machine Learning. Robotics Laboratory Department of Computer Science Stanford University Stanford, CA 94305, 2005.

[18] Pontil M., Verri A. properties of support vector machines. Neural Computation 1998; 10(4): 955-974.

[19] Schölkopf B., Burges J. C., Smola A. Advances In Kernel Methods: Support Vector Learning. Cambridge. MA: MIT Press, 1999.

[20] Shiry S., Sadatpoor S. S. Using Machine Learning Techniques In Homeopathy. Amir Kabir University Press, 2010.

[21] Tang M. E. fuzzy svm with a new fuzzy membership function to solve the two-class problems. Neural Processing Letters 2011; 34(3): 209-219.

[22] Trung L. e., Tran D., Wanli M. A., Sharma D. a new fuzzy membership computation method for fuzzy support vector machines, Communications and Electronics (ICCE), Third International Conference on 2010; 153 – 157.

[23] Vapnik V. N. Statistical Learning Theory. New York: Wiley, 1998.

[24] Vapnik V. N. The Nature Of Statistical Learning Theory. New York: Springer-Verlag, 1995.

[25] Zadeh L. A. Fuzzy Set A Basis For A Thory Of Possibility. In Fuzzy Set and System 1987; 1: 3-28.

[26] Zadeh L. A. fuzzy sets. information and control, 1965; 8(3): 338-353.

[27] Zhang X.G. using class-center vectors to build support vector machines. In: Proceeding of the IEEE signal processing society workshop 1999; 3–11.

مونا خداقلی کارشناسی خود را در رشته ریاضیات و کاربردها در دانشگاه قم به پایان رسانده و کارشناسی ارشد خود را نیز در رشتۀ ریاضی کاربردی گرایش بهینه‌سازی ترکیبیاتی در سال ۱۳۹۶ از دانشگاه شاهد اخذ کرده‌است.

نشانی رایانامۀ ایشان عبارت است از:

monakhodagholy@gmail.com

اردشیردولتی کارشناسی خود را از دانشگاه صنعتی اصفهان و کارشناسی ارشد و دکترای خود را از دانشگاه صنعتی امیرکبیر اخذ کرده‌است. ایشان هم‌اکنون عضو هیأت علمی گروه علوم کامپیوتر دانشگاه شاهد هستند. زمینه‌های پژوهشی ایشان بهینه‌سازی ترکیبیاتی، شبکه و نظریۀ گراف است.

نشانی رایانامۀ ایشان عبارت است از:

dolati@shahed.ac.ir

علی حسین‌زاده کارشناسی خود را در دانشگاه نیشابور به پایان رسانده و کارشناسی ارشد خود را نیز از دانشگاه جامع امام حسین(ع) اخذ کرده‌است. ایشان هم اکنون دانشجوی دکترای ریاضی گرایش بهینه‌سازی ترکیبیاتی در دانشگاه شاهد است.

نشانی رایانامۀ ایشان عبارت است از:

hoseinzadeh1393@gmail.com

خشایار شمس‌الکتابی کارشناسی و کارشناسی ارشد خود را در رشتۀ ریاضی در دانشگاه آزاد واحد جنوب به پایان رسانده و اکنون دانشجوی دکتری ریاضی دانشگاه تربیت مدرس است.

نشانی رایانامۀ ایشان عبارت است از:

khashuyur@gmail.com