

بهینه سازی طبقه بند ماشین بردار پشتیبان با استفاده از

الگوریتم بهینه سازی صفحات شیبدار

ایمان بهروان<sup>1</sup>، سید حمید ظهیری<sup>2</sup>

<sup>1</sup>دانشگاه بیرجند، دانشکده مهندسی برق و کامپیوتر، بیرجند، [i.behravan@gmail.com](mailto:i.behravan@gmail.com)، شماره ی فیش(هزینه

داوری): 426811224589

<sup>2</sup>دانشگاه بیرجند، دانشکده مهندسی برق و کامپیوتر، بیرجند، [shzahiri@birjand.ac.ir](mailto:shzahiri@birjand.ac.ir)

چکیده- ماشین بردار پشتیبان یکی از روش های یادگیری با نظارت است که از آن برای طبقه بندی و رگرسیون استفاده می شود. عملکرد این طبقه بند به پارامترهای متعددی وابسته است که نسبت دادن مقادیر مختلفی به این پارامترها روی خروجی آن تاثیر گذار است. از جمله این پارامترها، پارامتر  $C$  و  $\text{Sigma}$  (در تابع کرنل) می باشند. علاوه بر این نوع تابع کرنل به کار گرفته شده نیز تاثیر مستقیم روی نرخ بازشناسی دارد. استفاده از الگوریتم های هوشمند جهت یافتن مقادیر بهینه پارامترهای ذکر شده و انتخاب تابع کرنل مناسب می تواند برای رسیدن به نرخ بازشناسی مناسب راه حل مناسبی باشد. علاوه بر این انتخاب ویژگی هم که یکی از مهمترین مراحل پیش پردازش در علم بازشناسی الگو می باشد، می تواند در رسیدن به این هدف یعنی نرخ بازشناسی بالا کمک قابل توجهی نماید. در این مقاله از الگوریتم بهینه سازی صفحات شیبدار برای یافتن مقادیر مناسب طبقه بند، انتخاب تابع کرنل مناسب و انتخاب ویژگی های موثر به طور همزمان، برای دستیابی به نرخ بازشناسی بالا استفاده شده است.

کلید واژه ها- ماشین بردار پشتیبان، الگوریتم بهینه سازی صفحات شیبدار، انتخاب ویژگی

## 1-مقدمه

داده ها به فضای با ابعاد بی نهایت نیز یکی دیگر از عواملی است که قبل از اجرای الگوریتم طبقه بند باید توسط کاربر تنظیم شود. مسلماً جواب های مختلفی برای این مساله ( یافتن مقادیر مختلف برای پارامترها و انتخاب تابع کرنل مناسب ) وجود دارد اما دستیابی به بهترین جواب یعنی جوابی که بهترین نرخ بازشناسی را روی یک مجموعه داده ارائه دهد هدف ما است. که این مساله می تواند با به کارگیری الگوریتم های هوشمند حل شود. عامل بسیار مهم دیگری که می تواند در رسیدن به این هدف یعنی نرخ بازشناسی بیشتر تاثیر گذار باشد، انتخاب ویژگی<sup>3</sup> است. با انتخاب ویژگی مناسب از میان ویژگی های استخراج شده می توان علاوه بر دست یافتن به دقت طبقه بندی بالا، هزینه های محاسباتی و پیچیدگی سیستم را کاهش داد و از استخراج و اندازه گیری ویژگی های غیر ضروری در مرحله ی آزمایش خودداری نمود. روش های متعددی برای جست و جو در فضای ویژگی جهت یافتن ویژگی های مناسب ارائه شده است. یکی از این روش ها روش جست و جوی مستقیم متوالی<sup>5</sup> (SFS) نام دارد [2و3]. در این

ماشین بردار پشتیبان<sup>1</sup> اولین بار توسط vapnik در سال 1995 ارائه شد و تاکنون در مسائل مختلفی همچون بازشناسی الگو<sup>2</sup> و طبقه بندی متن<sup>3</sup> به کار برده شده است [1]. شهرت SVM به خاطر موفقیت آن در تشخیص حروف دستنویس است که با شبکه های عصبی با دقت بالا برابری می کند. مبنای عملکرد این طبقه بند دسته بندی خطی داده هاست. و در مسائلی که داده ها به صورت خطی جداپذیر نیستند، از توابعی به نام کرنل برای منتقل کردن داده ها به فضای با ابعاد بی نهایت استفاده می شود. این طبقه بند به دنبال ابر صفحه ایست که نه تنها دو کلاس را از هم جدا کند بلکه بیشترین حاشیه ی اطمینان را نیز داشته باشد. پارامترهای مختلفی روی عملکرد این طبقه بند تاثیر گذار هستند که برای دسترسی به خروجی مطلوب کاربر باید از قبل آن ها را تنظیم کند. از جمله این پارامترها، پارامتر  $C$  (penalty factor) و پارامتر  $\text{sigma}$  می باشند که پارامتر  $\text{sigma}$  در تابع کرنل قرار دارد. علاوه بر این، انتخاب تابع کرنل مناسب برای انتقال

<sup>3</sup> Text categorization

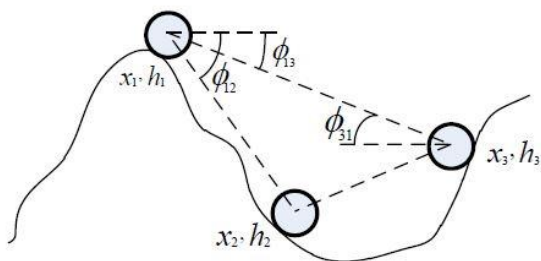
<sup>4</sup> Feature selection

<sup>5</sup> Sequential forward search

<sup>1</sup> Support vector machines (SVM)

<sup>2</sup> Pattern recognition

بدین صورت است که ابتدا توپ هایی به صورت تصادفی در فضای پاسخ پخش می شوند. توپی که ارتفاع کمتری دارد، برانزده تر است. یعنی به نقطه ی بهینه که همانا پایین ترین نقطه ی سطح شیبدار است، نزدیک تر می باشد. هر توپ به توپ دیگر نیرو وارد می کند که باعث می شود موقعیت توپ تغییر کند. نیروی وارد شده به یک توپ برابر مجموع نیروهای وارد شده به آن توپ از طرف توپ های با ارتفاع کمتر ( برانزده تر) می باشد. یعنی نیروی وارد شده از طرف توپی که در ارتفاع بالاتری قرار دارد در نظر گرفته نمی شود. هر توپ در فضای جست و جو دارای سه مختصات می باشد: موقعیت، ارتفاع و زاویه هایی که با دیگر توپ ها می سازد. موقعیت هر توپ، یک جواب در فضای جست و جو و ارتفاع آن به وسیله ی تابع برانزندی بدست می آید. شکل 1 یک فضای جست و جو با سه توپ را نشان می دهد.



شکل 1- یک مثال از فضای جست و جو با سه توپ و تخمینی از صفحه ی شیبدار

همان طور که در شکل 1 نشان داده شده است، برای داشتن تخمینی از صفحات شیبدار که توپ ها بر روی آن قرار دارند، از خطوط مستقیمی استفاده شده که مراکز توپ ها را به هم وصل می کند. این خطوط، با خط راست افقی عبوری از مرکز توپ، زاویه ای را تشکیل می دهد که به وسیله ی آن جهت و میزان شتاب اختصاص یافته به هر توپ تعیین می شود.

اگر سیستمی با N توپ را در نظر بگیریم، موقعیت توپ  $i$ ام که جوابی از مساله می باشد به صورت زیر تعیین می شود:

$$x_i = (x_i^1, x_i^2, \dots, x_i^d, \dots, x_i^n) \quad i=1,2,3,\dots,N \quad (1)$$

و

روش جست و جوی فضای ویژگی، با یک مجموعه تھی از ویژگی ها شروع شده و ویژگی های مناسب به طور متوالی به آن اضافه می شود.

روش دیگر جست و جوی برگشتی<sup>1</sup> (SBS) نام دارد [4و5و6و7]. در این روش خلاف روش قبل، فرایند انتخاب ویژگی با انتخاب یک مجموعه کامل از ویژگی ها شروع شده و در هر مرحله بعضی از ویژگی ها از آن حذف می شوند. عیب این روش ها این است که در آن ها کل فضای ویژگی با هم در نظر گرفته نمی شود. به عبارت دیگر ویژگی که در یک مرحله حذف می شود، ممکن است در کنار ویژگی های دیگر موجب بهبود کارایی سیستم شود. به همین دلیل الگوریتم های جست و جوی سراسری بیشتر مورد علاقه هستند. در روش های جست و جوی سراسری کل فضای پاسخ (فضای ویژگی) با هم در نظر گرفته می شود و یک الگوریتم ابتکاری (مانند الگوریتم ژنتیک<sup>2</sup> یا هوش جمعی) در هر تکرار تعدادی زیرمجموعه ویژگی تولید می کند و در اختیار تابع ارزیابی قرار می دهد. در واقع الگوریتم های جست و جوی سراسری با بررسی موثر فضای ویژگی، بدون جست و جوی کل آن نسبت به یافتن زیر مجموعه هایی کارآمد و مطلوب از بردارهای ویژگی اقدام می کنند. الگوریتم های ابتکاری متعددی به منظور انتخاب ویژگی به کار گرفته شده اند از جمله: الگوریتم ژنتیک [8و9و10]، الگوریتم کلونی مورچگان<sup>3</sup> [11]، الگوریتم بهینه سازی گروه ذرات<sup>4</sup> [12و13] و الگوریتم باینری جست و جوی هارمونی<sup>5</sup> [14]. در این مقاله برای دسترسی به نرخ بازشناسی بیشتر، از الگوریتم بهینه سازی صفحات شیبدار<sup>6</sup> [15] برای تنظیم پارامترهای طبقه بند، انتخاب تابع کرنل مناسب و همچنین انتخاب ویژگی به طور همزمان، استفاده شده است. علاوه بر این برای ارزیابی زیرمجموعه ویژگی های انتخاب شده روش «یک در مقابل همه»<sup>7</sup> [16و17] به کار گرفته شده است.

## 2- الگوریتم بهینه سازی صفحات شیبدار (IPO)

الگوریتم بهینه سازی صفحات شیبدار الهام گرفته از حرکت اجسام کروی روی سطوح شیبدار بدون اصطکاک است، که در این حالت همه ی اجسام تمایل دارند به پایین ترین نقطه ی سطح برسند. در این الگوریتم عوامل جست و جو توپ نام دارند، یعنی تعدادی توپ که در فضای پاسخ قرار گرفته اند، به دنبال نقطه ی بهینه می باشند. فضای پاسخ در این الگوریتم به صورت شیبدار در نظر گرفته شده است. مکانیزم عملکرد این الگوریتم

<sup>5</sup> Binary harmony search

<sup>6</sup> Inclined planes optimization

<sup>7</sup> One-vs-all

<sup>1</sup> Sequential backward search

<sup>2</sup> Genetic algorithm

<sup>3</sup> Ant colony optimization

<sup>4</sup> Particle swarm optimization

همچنین در رابطه ی 6، ثابت گرانش برای کاهش پیچیدگی حذف شده است. الگوریتم IPO از معادلات حرکت با شتاب ثابت برای بروزرسانی موقعیت توپ ها استفاده می کند:

$$x_i^d(t+1) = k1.rand1 \cdot a_i^d(t) \cdot \Delta t^2 + k2.rand2 \cdot v_i^d(t) \cdot \Delta t \quad (8)$$

به طوری که rand1 و rand2 دو ثابت تصادفی هستند که به طور یکنواخت در بازه ی [0 1] توزیع شده اند تا به الگوریتم IPO خاصیت تصادفی داده شود. در رابطه ی 8،  $v_i^d(t)$  سرعت توپ  $i$ ام در بعد  $d$ ام و در تکرار  $t$ ام است که از رابطه ی زیر محاسبه می شود:

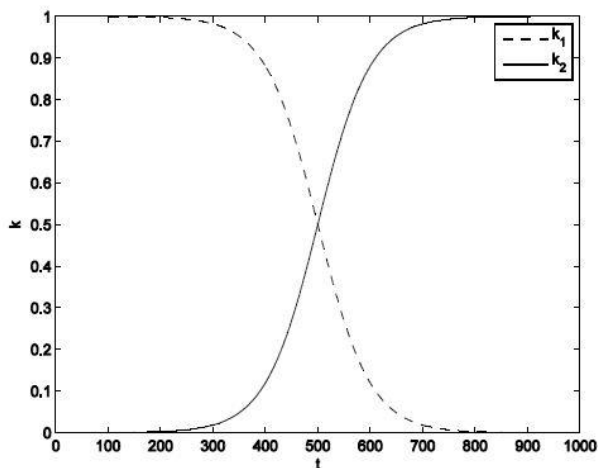
$$v_i^d(t) = \frac{x_{best}^d(t) - x_i^d(t)}{\Delta t} \quad (9)$$

در این رابطه  $x_{best}$ ، توپ با کمترین ارتفاع (برازندگی) در کل تکرارها تا تکرار فعلی است. در رابطه ی 8 برای کنترل فرایند جست و جوی الگوریتم، از دو ثابت مهم  $k1$  و  $k2$  استفاده شده است. این دو ثابت تابعی از زمان (تکرار الگوریتم) هستند به طوری که با گذشت زمان  $k1$  باید از مقدار اولیه ی خود کاهش یابد و  $k2$  افزایش یابد. این دو پارامتر معمولاً با استفاده از روابط زیر مقدار دهی می شوند:

$$k_1 = \frac{c_1}{1 + \exp((r - shift1) \times scale1)} \quad (10)$$

$$k_2 = \frac{c_2}{1 + \exp((r - shift2) \times scale2)} \quad (11)$$

شکل 3 نمودار تغییرات این دو ثابت را نشان می دهد:



شکل 3- نمودار تغییرات دو ثابت  $k1$  و  $k2$

مقادیر  $c1$ ،  $c2$ ،  $shift1$ ،  $shift2$ ،  $scale1$  و  $scale2$  به صورت تجربی بدست می آیند. معادلات مربوط به جابجایی و سرعت از فرمول های دینامیک حرکت با شتاب ثابت در فیزیک کلاسیک برگرفته شده اند:

$$x_j^{min} \leq x_j \leq x_j^{max} \quad (2)$$

به طوری که  $x_i^d$  بعد  $d$ ام توپ  $i$ ام می باشد. در یک زمان مشخص زاویه ی بین توپ  $i$ ام و توپ  $j$ ام در بعد  $d$  یعنی  $\varphi_{ij}^d$  به صورت زیر محاسبه می گردد:

$$\varphi_{ij}^d(t) = \tan^{-1} \left( \frac{f_j(t) - f_i(t)}{x_i^d(t) - x_j^d(t)} \right) \quad (3)$$

$$d = 1, 2, \dots, n$$

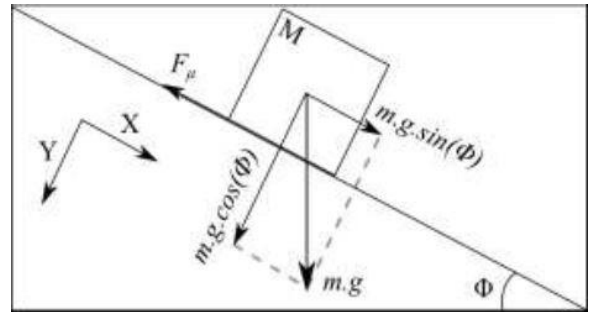
$$i, j = 1, 2, \dots, N, i \neq j$$

در رابطه ی 3،  $f_i(t)$  برازندگی توپ  $i$ ام در زمان  $t$  است. برای هر توپ شتاب های ناشی از صفحات شیبدار مختلف با یکدیگر جمع شده تا شتاب کل اعمالی به هر توپ محاسبه شود. بر اساس قانون دوم نیوتون داریم:

$$\sum f = m \cdot a \quad (4)$$

و با توجه به شکل 2 شتاب اعمالی از سطح شیبدار به هر توپ از رابطه ی زیر محاسبه می گردد:

$$a = g \cdot \sin(\varphi) \quad (5)$$



شکل 2- نیروی اعمالی به جسمی با جرم  $M$  روی سطح شیبداری با زاویه ی  $\Phi$

در این الگوریتم، شتاب وارد بر توپ  $i$ ام در بعد  $d$ ام با استفاده از رابطه ی زیر محاسبه می گردد:

$$a_i^d(t) = \sum U(f_j(t) - f_i(t)) \cdot \sin(\varphi_{ij}^d(t)) \quad (6)$$

در رابطه ی 6،  $U$  تابع پله ی واحد می باشد:

$$U(w) = \begin{cases} 1 & w > 0 \\ 0 & w \leq 0 \end{cases} \quad (7)$$

می باشد. که به ترتیب متناظر با توابع خطی، POLYNOMIAL، RBF و QUADRATIC می باشند. درایه های دوم و سوم مقادیر پارامترهای C و SIGMA را جهت آموزش طبقه بند مشخص می کنند. سه درایه  $F1$ ،  $F2$  و  $F3$  هم که برای انتخاب ویژگی استفاده می شوند، صفر و یا یک می باشند که صفر معادل عدم حضور و یک معادل حضور این ویژگی از مجموعه داده ی آموزشی، برای آموزش دسته بند دودویی می باشد. جدول 2 مثالی را برای یک توپ نشان می دهد:

جدول 2- مثالی برای یک توپ

3	5.5	7	1	0	1
---	-----	---	---	---	---

در این مثال طبقه بند ماشین بردار پشتیبان، با تابع کرنل RBF، مقدار  $C=5.5$ ، مقدار  $\sigma=7$  و مجموعه داده ای که از ویژگی های 1 و 3 داده های آموزشی تشکیل شده آموزش می بیند. حال نوبت مشخص کردن مقدار برانزنگی برای هر توپ می باشد. همان طور که گفته شد، هر توپ معادل یک طبقه بند دودویی است که با تابع کرنل مشخص شده، مقادیر پارامترهای آن و مجموعه داده ای که با ویژگی های انتخاب شده از داده های آموزشی شکل گرفته، آموزش می بیند و یک کلاس را از سایر کلاس ها جدا میکند. پس از آموزش این طبقه بند، مقدار برانزنگی توپ مربوط به آن، برابر میزان نرخ بازشناسی طبقه بند روی داده های آموزش می باشد. پس از هزار بار اجرای حلقه ی اصلی الگوریتم، بهترین توپ، یعنی توپی که بیشترین نرخ بازشناسی را روی داده های آموزش دارد، بدست آورده و نرخ بازشناسی آن را روی داده های آزمایش بدست می آوریم. که این عدد معادل نرخ بازشناسی روی یک کلاس خاص می باشد. بر اساس روش طبقه بندی ((یک در مقابل همه))، به اندازه ی تعداد دسته های مجموعه داده، الگوریتم IPO اجرا شده و بهترین توپ برای هر کلاس بدست می آید. به عنوان مثال برای داده های glass که دارای 6 کلاس می باشد، 6 توپ بدست می آوریم که هر کدام متناظر با یک طبقه بند دودویی است که یک کلاس را از سایر کلاس ها جدا می کند. در نهایت با ترکیب نتایج این طبقه بندها روی داده های آزمایش، به داده های آزمایش برچسب زده شده و نرخ بازشناسی روی مجموعه داده ی مورد نظر، ارائه می شود. به عبارت دیگر هر کدام از دسته بندهای بدست آمده نظر خود را راجع به داده های آزمایش وارد شده اعلام می کنند یعنی بیان می کنند که داده های مذکور به کلاسی که توسط آن ها از سایر کلاس ها جدا می شود تعلق دارند یا نه و در نهایت مشخص می شود که هر کدام از این داده ها متعلق به کدام کلاس هستند.

#### 4- آزمایشات صورت گرفته و نتایج

$$x = \frac{1}{2}.a.t^2 + v_0.t + x_0 \quad (12)$$

$$v = \frac{\Delta x}{\Delta t} \quad (13)$$

در ادامه شبه کد الگوریتم IPO آورده شده است:

- 1- تولید جمعیت اولیه به صورت تصادفی و مقداردهی اولیه به پارامترهای  $k1$  و  $k2$  و سایر پارامترها
- 2- ارزیابی برانزنگی (ارتفاع) هر توپ
- 3- محاسبه ی زاویه، شتاب و سرعت هر توپ
- 4- بروزرسانی موقعیت هر توپ
- 5- محاسبه و تعیین مقدار  $k1$  و  $k2$
- 6- بازگرداندن توپ هایی که از بازه ی فضای جست و جو خارج شده اند به داخل فضای جست و جو
- 7- پایان اگر شرط توقف مشاهده شود در غیر این صورت، برو به مرحله ی 2

### 3- نحوه ی استفاده از الگوریتم بهینه سازی صفحات شیبدار برای یافتن مقادیر بهینه طبقه بند ماشین بردار پشتیبان و انتخاب ویژگی

ابتدا باید پارامترهای الگوریتم IPO مقداردهی اولیه شوند. برای تمام داده ها جمعیت اولیه 50 در نظر گرفته شده و مقادیر سایر پارامترها به شرح زیر می باشد:

$$C1 = 0.3893, C2 = 2.9512, \text{Shift}1 = 534.5566$$

$$\text{Shift}2 = 367.0491, \text{Scale}1 = 0.1452, \text{Scale}2 = 0.9001$$

که این داده ها به صورت تجربی بدست آمده اند. شرط توقف الگوریتم، تکرار حلقه ی اصلی آن به تعداد 1000 بار، در نظر گرفته شده است. پس از این باید ساختار هر توپ و تابع برانزنگی آن مشخص شود. هر توپ یک بردار است که تعداد درایه های آن برابر تعداد ویژگی های مجموعه داده ی مورد نظر به اضافه ی 3 می باشد. چرا که علاوه بر انتخاب ویژگی ما به دنبال یافتن مقادیر مناسبی برای پارامترهای C و sigma و همچنین انتخاب تابع کرنل مناسب می باشیم. جدول 1 ساختار کلی هر توپ را برای یک مجموعه داده با سه ویژگی نشان می دهد:

#### جدول 1- ساختار یک توپ برای یک مجموعه داده ی با 3 ویژگی

Kernel function	c	sigma	F1	F2	F3
-----------------	---	-------	----	----	----

درایه ی اول (KERNEL FUNCTION)، 1، 2، 3 و یا 4 است که هر کدام متناظر با استفاده از یک تابع کرنل خاص جهت آموزش طبقه بند

ویژگی های انتخاب شده	2,3,4,6,8,9,10,12,13	1,3,4,5,6,7,8,9,10,11,12,13	1,3,7,11,12,13
تابع کرنل	RBF	RBF	polynomial
مقدار پارامتر C	175.755	5.1337	152.0165
مقدار پارامتر sigma	14.5782	4.7528	1.0554
نرخ بازشناسی روی داده های آموزش	100%	100%	100%
نرخ بازشناسی روی داده های آزمایش	100%	100%	100%

در این بخش ابتدا در جدول 3 ویژگی های سه مجموعه داده ی Iris، Wine و Glass آورده شده و سپس نتایج حاصل از اجرای الگوریتم IPO برای یافتن طبقه بندهای دودویی مربوط به هر کلاس از مجموعه داده، در جدول های 4 و 5 و 6 ارائه شده است. در واقع این جدول ها بهترین توپ هایی را که هر کلاس از مجموعه داده ی مورد نظر را از سایر داده های آن جدا می کند، نشان می دهد.

جدول 3 - مشخصات مجموعه داده های IRIS، WINE و GLASS

تعداد کلاس ها	تعداد ویژگی ها	تعداد نمونه ها	مجموعه داده ها
3	4	150	Iris
3	13	178	Wine
6	9	214	Glass

جدول 6- بهترین توپ های بدست آمده برای داده های GLASS

کلاس	1	2	3	5	6	7
ویژگی ها	1,3,4,6,8	2,3,4,8,9	2,3,4,5,6,8,9	2,3,5,7	1,2,3,5,6,7,8,9	تمام ویژگی ها
تابع کرنل	polynomial	RBF	RBF	RBF	quadratic	quadratic
C	163.4515	106.6469	71.1181	196.0424	61.4832	0.1672
sigma	6.0216	0.1867	0.1226	3.6808	-	0.0270
نرخ بازشناسی روی داده های آموزشی	100%	100%	100%	100%	100%	100%
نرخ بازشناسی روی داده های آزمایشی	81.31%	74.77%	92.52%	99.07%	100%	96.26%

جدول 4- بهترین توپ های بدست آمده برای مجموعه داده های IRIS

کلاس	1	2	3
ویژگی های انتخاب شده	2و3	تمام ویژگی ها	تمام ویژگی ها
تابع کرنل	polynomial	RBF	خطی
مقدار پارامتر C	21.88723	61.6025	84.7313
مقدار پارامتر sigma	4.8125	0.5872	9.5197
نرخ بازشناسی روی داده های آموزش	100%	100%	100%
نرخ بازشناسی روی داده های آزمایشی	100%	97.33%	97.33%

با توجه به ستون دوم جدول 4، طبقه بندی که نمونه های کلاس اول را از نمونه های سایر کلاس ها جدا می کند، SVM ایست که تابع کرنل آن polynomial می باشد و پارامترهای C و sigma در آن به ترتیب 21.8872 و 4.8125 می باشند. علاوه بر این، این طبقه بند با مجموعه داده ای که از ویژگی های 2 و 3 داده های آموزشی تشکیل شده آموزش می بیند. نرخ بازشناسی این طبقه بند روی داده های آموزش و آزمایش 100% می باشد. در تمام آزمایشات نیمی از داده ها برای آموزش و نیمی دیگر برای آزمایش به کار گرفته شده اند.

در ادامه جدول های 5 و 6 این اطلاعات را برای داده های Wine و Glass ارائه می دهند.

جدول 5- بهترین توپ ها بدست آمده برای داده های WINE

کلاس	1	2	3
------	---	---	---

حال نوبت به برچسب زدن به داده های آزمایش با استفاده از ترکیب نتایج طبقه بندهای دودویی می باشد. در واقع هر داده ی آزمایش توسط هر کدام از طبقه بندهای دودویی مورد بررسی قرار می گیرد و هر کدام از این طبقه بندها به تعبیری، نظر خود را راجع به نمونه ی مذکور اعلام می دارند. به عنوان مثال دسته بند  $f_i$ ، اعلام می کند که نمونه ی مذکور به کلاس  $i$  تعلق دارد یا نه. با ترکیب نتایج طبقه بندها، شماره ی کلاس هر داده مشخص شده و سپس نرخ بازشناسی برای کل مجموعه داده مشخص می شود. جدول 7 نرخ بازشناسی روی داده های Iris، Wine و Glass را نشان می دهند.

جدول 7- نرخ بازشناسی روی کل مجموعه داده های  
GLASS,WINE,IRIS

مجموعه داده	نرخ بازشناسی روی کل مجموعه داده
Iris	98%
Wine	100%
Glass	76.17%

6- نتیجه گیری

در این مقاله از یک الگوریتم بهینه سازی هوش جمعی برای بهبود عملکرد طبقه بند ماشین بردار پشتیبان استفاده کردیم. در واقع نکته ی اصلی استفاده از الگوریتم بهینه سازی صفحات شیبدار برای این منظور بود که اتفاقی تازه می باشد. علاوه بر این استفاده از این الگوریتم برای فرایند انتخاب ویژگی همزمان با تنظیم پارامترهای طبقه بند مذکور سبب بهبود نرخ بازشناسی شده است. مقایسه ی نرخ بازشناسی در کارهای انجام شده در [18و19] با نتایج موجود در جدول 7 موبد این مطلب می باشد.

مراجع

- [13] Y.Liu, G.Wang, "an Improved Particle Swarm Optimization for Feature Selection", *Journal of Bionic Engineering*, 8 (2001) 191-200.
- [14] علی سجاد زاده، سید حمید ظهیری، سید محمود رضوی، "انتخاب ویژگی با استفاده از الگوریتم باینری جست و جوی هارمونی مبتنی بر روش طبقه بندی OVA"، *مجله علمی-پژوهشی رایانش نرم و فناوری اطلاعات* بابل، دانشگاه صنعتی نوشیروانی بابل، جلد 1، شماره 1، 1392
- [15] محمد حامد مظفری، حامد عبدی، سید حمید ظهیری، "الگوریتم جدید بهینه سازی سیستم صفحات شیبدار"، *مجله علمی-پژوهشی رایانش نرم و فناوری اطلاعات*، دانشگاه صنعتی نوشیروانی بابل، جلد 1، شماره 1، 1391
- [16] L. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 12, no. 10, pp. 993-1001, Oct. 1990.
- [17] A. Krogh and J. Vedelsby, "Neural network ensemble, cross validation, and active learning," in *Proc. Advances in Neural Information Processing Systems*, vol. 7, pp. 231-238, 1995.
- [18] S.H.Zahiri and S.A.seyedin, "Intelligent Particle Swarm Classifier", *Iranian Journal of Electrical and Computer engineering*, Vol.4, No.1, Winter-Spring 2005
- [19] S.H.Zahiri, "learning automata based qualifier", *Pattern Recognition Letters* 29 (2008) 40-48
- [1] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, New York, 2000.
- [2] S.Guan, J.Liu, Y. Qi, "An incremental approach to contribution-based feature selection", *Journal of Intelligence System* 13 (1) (2004).
- [3] J.Liu, C.Wang, "Using support vector machines with a novel hybrid feature selection method for diagnosis of erythematous-squamous diseases", *Expert Systems with Applications*, 38 (2011) 5809-5815.
- [4] E. Gasca, J.S. Sanchez, R. Alonso, "Eliminating redundancy and irrelevance using a new MLP-based feature selection method", *Pattern Recognition* 39 (2006) 313-315.
- [5] S.Abe, "Modified backward feature selection by cross validation", *proceedings of the European Symposium on Artificial Neural Networks* (2005) 163-168.
- [6] C. Hsu, H. Huang, D. Schuster, "The ANNIGMA-wrapper approach to fast feature selection for neural nets", *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics* 32 (2) (2002) 207-212.
- [7] E. Romero, J.M.Sopena, "Performing feature selection with multilayer perceptron's", *IEEE Transactions on Neural Networks* 19 (3) (2008) 431-441.
- [8] I. Oh, J. Lee, B. Moon, "Hybrid genetic algorithms for feature selection", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (11) (2004) 1424-1437.
- [9] Y.L. Wu, C.Y.Tang, M.K. Hor, P.F.Wu, "Feature selection using genetic algorithm and cluster validation", *Expert Systems with Applications* 38 (2011) 2727-2732.
- [10] C.L.Huang, C.J.Wang, "A GA-based feature selection and parameters optimization for support vector machines", *Expert Systems with Applications* 31 (2006) 231-240
- [11] L. Ke, Z. Feng, Z.Ren, "An efficient ant colony optimization approach to attribute reduction in rough set theory", *pattern Recognition Letters* 29 (2008) 1351-1357.
- [12] Li-Yeh Chuang, Cheng-Hong Yang, Junge-Chike Li, "Chaotic maps based on binary particle swarm optimization for feature selection", *Applied Soft Computing* 11 (2011) 239-248.

