

بهینه‌سازی وزن‌ها در کرنل مرکب برای طبقه‌بند مبتنی بر نمایش تنک کرنلی

طاهره زارع بیدکی^۱، دانشجوی دکتری؛ محمدتقی صادقی^۲، استادیار

۱- دانشکده مهندسی برق - دانشگاه یزد - یزد - ایران - t.zare@stu.yazd.ac.ir

۲- دانشکده مهندسی برق - دانشگاه یزد - یزد - ایران - m.sadeghi@yazd.ac.ir

چکیده: طبقه‌بند مبتنی بر نمایش تنک (SRC) یکی از الگوریتم‌های موفق در ترکیب مفاهیم مطرح در دو حوزه نمونه‌برداری فشرده و آموزش ماشین است. در SRC، هر نمونه بر اساس ترکیب خطی تنکی از نمونه‌های آموزشی نمایش داده می‌شود. با توجه به موفقیت‌های اولیه این الگوریتم، فرم کرنلیزه آن (KSRC) نیز ارائه شده که در آن داده‌ها با استفاده از تابع کرنل به طور غیر صریح به فضای ویژگی جدیدی با ابعاد بالاتر نگاشت یافته و سپس SRC اعمال می‌شود. در این چارچوب، انتخاب کرنل مناسب اهمیت زیادی دارد. استفاده از کرنل مرکب به جای استفاده از یک کرنل، بهترین راه حلی است که تاکنون برای این مسئله ارائه شده است. در این پژوهش، برای افزایش سرعت و دقت طبقه‌بند، از کرنل مرکب در ساختار KSRC استفاده شده است. کرنل مرکب، با جمع وزن‌دار کرنل‌های پایه ایجاد می‌شود که وزن کرنل‌ها با به‌کارگیری معیارهای متفاوتی تعیین می‌شود. در این مقاله، با هدف حداقل سازی خطای بازسازی در چارچوب KSRC، روش‌هایی برای بهینه سازی وزن کرنل‌ها ارائه می‌شود. بر این اساس، ابتدا روش‌های پایه برای ساخت کرنل مرکب که تا کنون برای کاربرد KSRC استفاده نشده است، مورد بررسی قرار خواهد گرفت. سپس الگوریتمی، نوین جهت بهینه‌سازی کرنل مرکب در چارچوب مسئله KSRC ارائه می‌شود. برای بررسی عملکرد الگوریتم‌های پیشنهادی از دادگان واقعی موجود در پایگاه داده UCI و مجموعه اعداد دست نویس MNIST استفاده شده است. نتایج حاصله موید عملکرد خوب الگوریتم پیشنهادی در شرایط مختلف و مقاوم‌تر بودن روش پیشنهادی در مواجهه با نویز است.

واژه‌های کلیدی: طبقه‌بند مبتنی بر نمایش تنک (SRC)، طبقه‌بند مبتنی بر نمایش تنک کرنلیزه شده (KSRC)، کرنل مرکب.

Optimization of the Weights in Multiple Kernel for Kernel Sparse Representation Based Classifier

T. Zare Bidaki¹, PhD Student; M. T. Sadeghi², Assistant Professor

1- Department of Electrical Engineering, University of Yazd, Yazd, Iran, Email: t.zare@stu.yazd.ac.ir

2- Department of Electrical Engineering, University of Yazd, Yazd, Iran, Email: m.sadeghi@yazd.ac.ir

Abstract: Sparse representation based classifier (SRC) is a well-known algorithm which combines the compressive sampling and machine learning concepts. In this classifier, each sample is represented by a linear and sparse combination of the associated training samples. Following the successful application of the SRC algorithm, the kernelized version of the classifier was also presented in which the data points are implicitly mapped into a high dimensional feature space. The SRC algorithm is then applied. Selection of a proper kernel is an important issue in such a kernel based algorithm. Using multiple kernel is a proper solution for this problem. In this study, in order to increase the accuracy and speed of the KSRC algorithm, we utilize a multiple kernel function within the framework of the KSRC. The multiple kernel is created by the weighted summation of a set of basis kernels where the kernel weights are determined using a set of different approaches. In this paper, we propose a novel method of determining the weights by using an optimization algorithm which is based on minimization of the reconstruction error of the KSRC. The proposed algorithm is evaluated considering real data sets from the UCI database and also hand written digits of the MNIST data sets. Our experimental results show the superiority of the proposed algorithm in different conditions. The proposed method is also more robust against additive noises.

Keywords: Multiple Kernel, Sparse Representation based Classifier (SRC), Kernel Sparse Representation based classifier (KSRC).

تاریخ ارسال مقاله: ۱۳۹۵/۰۳/۰۳

تاریخ اصلاح مقاله: ۱۳۹۵/۰۵/۲۳ و ۱۳۹۵/۰۷/۲۲

تاریخ پذیرش مقاله: ۱۳۹۵/۰۸/۲۷

نام نویسنده مسئول: محمدتقی صادقی

نشانی نویسنده مسئول: ایران - یزد - صفائیه - بلوار دانشگاه - دانشگاه یزد - دانشکده مهندسی برق.

۱- مقدمه

چنانچه اشاره شد، در الگوریتم‌های مبتنی بر کرنل از جمله KSRC انتخاب نوع تابع کرنل و پارامترهای تابع کرنل اهمیت زیادی دارد. در بسیاری از پژوهش‌های انجام شده در این حوزه، تابع کرنل با سعی و خطا و یا با اعمال فرآیند اعتبارسنجی متقابل^۲ از بین تعداد محدودی از توابع کرنل شناخته شده تعیین می‌شود. واضح است برای دستیابی به کرنل بهینه در چنین شرایطی بایستی مجموعه بزرگی از توابع کرنل متفاوت با پارامترهای متنوعی را در نظر گرفت که می‌تواند بسیار زمان‌بر باشد. علاوه بر این، گاهی اوقات توزیع داده‌های طبقات مختلف ممکن است به اندازه‌ای پیچیده باشد که یک تابع کرنل به‌تنهایی قادر به توصیف مناسب آن‌ها نباشد. از این رو، در سالهای اخیر، به کارگیری تابع کرنل مرکب و مسئله یادگیری کرنل مرکب مورد استقبال قرار گرفته است. در این رویکرد، تابع کرنل ترکیبی از توابع کرنل متفاوت در نظر گرفته شده و با استفاده از داده‌های آموزشی پارامترهای مؤثر بر ترکیب تعیین می‌شوند [۱۳ و ۱۴].

علی‌رغم عملکرد خوب کرنل‌های مرکب در کاربردهایی مانند SVM، کارهای اندکی برای استفاده از کرنل مرکب در چارچوب SRC صورت گرفته است [۱۵ و ۱۶]. در [۱۵]، وزن کرنل‌ها برای هر نمونه مورد آزمایش در جهت حداقل سازی خطای بازسازی بهینه می‌شود که یک فرآیند زمان‌بر در مرحله آزمایش است. همچنین در مقاله مذکور، ابتدا پارامترهای بهینه دو کرنل گوسی و چندجمله‌ای با استفاده از روش اعتبارسنجی متقابل، تعیین شده و سپس دو کرنل انتخاب شده در ساختار کرنل مرکب مورد استفاده قرار می‌گیرد که حالت بسیار ساده‌ای از کرنل‌های مرکب است. از دید کلی، استفاده از روش اعتبارسنجی متقابل برای تعیین پارامتر بهینه هر کرنل مغایر با مفاهیم اصلی در روشهای یادگیری کرنل مرکب است. در [۱۶]، از یک روش تکراری برای تعیین وزن کرنل‌ها در مرحله آموزش استفاده می‌شود. در روش پیشنهاد شده در [۱۶]، از میزان انطباق هر کرنل با کرنل ایده‌آل و اطلاعات تکمیلی به دست آمده از میزان دقت هر کرنل بر روی داده‌های آموزشی برای بهینه‌سازی وزن کرنل‌ها استفاده می‌شود. از آنجایی که در فرآیند بهینه‌سازی، دقت هر یک از نمونه‌های آموزشی در به روزرسانی وزن کرنل‌ها تأثیر دارد، الگوریتم پیشنهاد شده بسیار زمان‌بر است. همچنین از کرنل‌های مرکب در طبقه‌بند CRC نیز استفاده شده است [۱۷]. از آنجایی که مسئله CRC با نرم ۲، دارای جوابی تحلیلی است، در [۱۷] مسئله یافتن بهترین نمایش از نمونه و بهینه‌سازی وزن کرنل‌ها در دو گام صورت می‌گیرد. در گام نخست، بهترین نمایش از نمونه‌ها با مشتق‌گیری از مسئله بهینه‌سازی و در گام بعد، وزن کرنل‌ها با حل مسئله درجه دو^۴ محاسبه می‌شود.

در این پژوهش، به کارگیری کرنل مرکب در چارچوب KSRC، مورد بررسی قرار گرفته است. کرنل مرکب، با جمع وزن‌دار کرنل‌های پایه حاصل شده و وزن هر کرنل با استفاده از چند روش متفاوت تعیین می‌شود. در ساده‌ترین روش، از میانگین کرنل‌های پایه به عنوان

در سال‌های اخیر، طبقه‌بند مبتنی بر نمایش تنک^۱ (SRC) مورد توجه زیادی قرار گرفته است [۴-۱]. در این طبقه‌بند، بردارهای ویژگی مرتبط با داده‌های آموزشی طبقات مختلف دادگان در ماتریسی موسوم به ماتریس واژه‌نامه سازماندهی می‌شوند. در فرآیند طبقه‌بندی، نمایش تنک نمونه مورد آزمایش، با استفاده از ماتریس واژه‌نامه، تعیین می‌شود. سپس، داده مورد آزمایش با در نظر گرفتن ضرایب نمایش تنک مرتبط با هر طبقه، بازسازی می‌شود و داده متعلق به طبقه‌ای لحاظ می‌شود که کمترین خطای بازسازی را به همراه داشته باشد. واضح است که در چنین فرآیندی، انتخاب ماتریس واژه‌نامه اهمیت زیادی داشته و دقت و سرعت طبقه‌بند، تابع تعداد عناصر واژه‌نامه (اتم‌ها) و نوع ویژگی‌های استخراج شده از داده‌ها است. در این چارچوب، آموزش واژه‌نامه مورد توجه محققین قرار گرفته است. در فرآیند آموزش واژه‌نامه، عمدتاً به کاهش مؤثر تعداد اتم‌های ماتریس واژه‌نامه توجه شده است به‌گونه‌ای که بتوانیم با تعداد کمتری از عناصر پایه، نمایش تنک‌تر و در عین حال دقیق‌تری به دست آوریم.

بنابراین در مسئله SRC، هدف دستیابی به تنک‌ترین نمایش از نمونه است که برای این منظور از نرم ۱ در مسئله بهینه‌سازی استفاده شده است. برخی از محققین بر این باورند که نمایش تنک را نمی‌توان دلیل موفقیت الگوریتم SRC در طبقه‌بندی دانست و استفاده جمعی از نمونه‌های آموزشی در ساخت نمایش تنک، سبب دقت این الگوریتم در طبقه‌بندی داده‌ها است. از این رو، مسئله بهینه‌سازی فوق با نرم ۲ نیز با نام طبقه‌بند CRC^۲، مورد توجه قرار گرفت [۴]. در این پژوهش، تمرکز بر روی طبقه‌بند SRC است.

از سوی دیگر، با توجه به موفقیت روش‌های مبتنی بر کرنل در تعمیم الگوریتم‌های خطی نوع کرنلیزه طبقه‌بند مبتنی بر نمایش تنک (KSRC) نیز مورد توجه قرار گرفته است [۱۱-۱۶]. در روش‌های مبتنی بر کرنل، در ظاهر، داده‌ها تحت اثر نگاشتی عموماً غیرخطی به فضایی با ابعاد بیشتر انتقال داده شده و الگوریتم خطی مربوطه در فضای جدید اجرا می‌شود. اما در عمل، با استفاده از تابع کرنل، این نگاشت به صورت غیرصریح انجام می‌شود و بدون افزایش پیچیدگی محاسباتی، نسخه غیرخطی الگوریتم حاصل می‌شود. در روش‌های مبتنی بر کرنل، ماتریس گرام که با اعمال تابع کرنل به زوج داده‌های آموزشی تولید می‌شود و ابعاد آن برابر با تعداد داده‌های آموزشی و نه تعداد ابعاد بردار ویژگی‌هاست، نقشی کلیدی دارد [۱۲]. واضح است که در چنین الگوریتم‌هایی، انتخاب تابع کرنل مناسب اهمیت ویژه‌ای دارد. در KSRC چنانچه متعاقباً تشریح خواهد شد، عملاً ماتریس گرام جایگزین ماتریس واژه‌نامه نمایش تنک می‌شود که با توجه به توضیحات پیشین در مورد نقش واژه‌نامه در این طبقه‌بند، تأثیر تابع کرنل بر عملکرد سامانه طبقه‌بند مشخص می‌شود.

سال ۲۰۰۹، طبقه‌بند مبتنی بر نمایش تنک (SRC) معرفی شد که در آن، از تکنیک نمایش تنک سیگنال برای طبقه‌بندی تصاویر چهره استفاده شد [۴-۱]. ایده اصلی در این طبقه‌بند، یافتن طبقه‌ای است که نمونه ورودی را با دقت خوبی به صورت یک ترکیب خطی تنک از نمونه‌های آموزشی مربوطه توصیف نماید. در ادامه، این طبقه‌بند با جزئیات بیشتر معرفی می‌شود.

فرض کنید مجموعه داده‌های آموزشی به صورت $\{(\mathbf{x}_i, y_i) \mid (\mathbf{x}_i \in \mathbb{R}^m, y_i \in \{1, \dots, c\}, i = 1, \dots, n)\}$ باشد که c تعداد کلاس‌ها، m بعد فضای ویژگی داده‌های \mathcal{X} و y_i برچسب متناظر با بردار ویژگی \mathbf{x}_i می‌باشد. در فرآیند طبقه‌بندی، هدف تعیین برچسب کلاس، y ، برای نمونه آزمون ورودی، \mathbf{x} ، از میان c کلاس موجود است.

در طبقه‌بند مبتنی بر نمایش تنک، نمونه‌های متعلق به کلاس z ام در ستون‌های ماتریس \mathbf{D}_j نهاده می‌شود؛ یعنی $\mathbf{D}_j = \{\mathbf{x}_{j,1}, \mathbf{x}_{j,2}, \dots, \mathbf{x}_{j,n_j}\} \in \mathbb{R}^{m \times n_j}$ که n_j تعداد نمونه‌های آموزشی متعلق به کلاس z ام و $\mathbf{x}_{j,i}$ نمونه i ام از این کلاس است. ماتریس $\mathbf{D} = \{\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_c\} \in \mathbb{R}^{m \times n}$ تعریف می‌شود که $n = \sum_{j=1}^c n_j$ است.

در مرحله طبقه‌بندی، نمونه آزمون ورودی به صورت ترکیب خطی تنکی از تمام نمونه‌های آموزشی نمایش داده می‌شود:

$$\mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \quad (1)$$

که $\boldsymbol{\alpha} \in \mathbb{R}^n$ بردار ضرایب نمایش تنک است. اگر نمونه آزمون \mathbf{x} متعلق به کلاس z ام باشد، انتظار می‌رود که در $\boldsymbol{\alpha}$ تنها مولفه‌های متناظر با کلاس z ام مقدار غیرصفر داشته باشند. به عبارتی

$$\boldsymbol{\alpha} = [0, \dots, 0, \alpha_{j,1}, \dots, \alpha_{j,n_j}, 0, \dots, 0]^T \quad (2)$$

که $\alpha_{j,i} \in \mathbb{R}$ ضریب متناظر با نمونه آموزشی $\mathbf{x}_{j,i}$ است. به عبارتی می‌توان \mathbf{x} را با ترکیب خطی نمونه‌های آموزشی کلاس z ام توصیف نمود. بنابراین انتظار می‌رود که بردار ضرایب $\boldsymbol{\alpha}$ برداری تنک باشد. مسئله یافتن بردار ضرایب $\boldsymbol{\alpha}$ در الگوریتم SRC یک مسئله محدب به صورت زیر است:

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_1 \\ \text{subject to} \quad & \mathbf{x} = \mathbf{D}\boldsymbol{\alpha} \end{aligned} \quad (3)$$

که $\|\cdot\|_1$ نرم ۱ را نشان می‌دهد و به عنوان تقریبی از نرم صفر در نظر گرفته می‌شود. برای داده نویزی، نمایش تنک به صورت زیر در نظر گرفته می‌شود:

$$\mathbf{x} = \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad (4)$$

که $\boldsymbol{\varepsilon} \in \mathbb{R}^m$ بردار نویز با انرژی محدود $\|\boldsymbol{\varepsilon}\|_2 < \sigma$ است که $\|\cdot\|_2$ نرم ۲ را نشان می‌دهد. به عبارتی $\boldsymbol{\varepsilon}$ همان نویز و با خطای موجود در فرآیند بازسازی را نشان می‌دهد. بنابراین مسئله (۳) را می‌توان به صورت زیر تغییر داد:

کرنل مرکب استفاده می‌شود که به عبارتی وزن تمامی کرنل‌ها مشابه فرض می‌شود. وزن کرنل‌ها را همچنین می‌توان در چارچوب روش‌هایی مانند استفاده از دقت هر کرنل بر روی داده‌های آموزشی و یا استفاده از میزان انطباق هر کرنل با کرنل ایده‌آل تعیین نمود. اما در هیچ یک از این روش‌ها، وزن کرنل‌ها با توجه به هدف طبقه‌بند اصلی طراحی نمی‌شود. در این مقاله، الگوریتم‌های نوینی ارائه می‌شود که در آن‌ها رابطه بهینه سازی وزن کرنل‌ها در تابع هدف طبقه‌بند مبتنی بر نمایش تنک لحاظ شده و وزن کرنل‌ها با هدف حداقل سازی خطای بازسازی تعیین می‌شود. در یکی از روش‌های پیشنهادی، بدون در نظر گرفتن برچسب داده‌های آموزشی، وزن کرنل‌ها با هدف حداقل سازی خطای بازسازی نمونه‌های ارزیابی بهینه می‌شود. اما برچسب کلاس نمونه‌های آموزشی حاوی اطلاعاتی هستند که می‌توانند در بهینه‌سازی وزن کرنل‌ها برای افزایش دقت طبقه‌بند نهایی کمک کنند، از این رو در الگوریتم پیشنهادی دوم از اطلاعات برچسب کلاس‌ها در بهینه سازی وزن کرنل‌ها استفاده می‌شود.

به طور کلی نوآوری‌های این مقاله عبارتند از:

- استفاده از کرنل مرکب در چارچوب طبقه‌بند مبتنی بر نمایش تنک کرنلی با در نظر گرفتن کرنل میانگین، کرنل میانگین وزن‌دار مبتنی بر دقت عملکرد هر کرنل و وزن دهی بر اساس انطباق کرنلی.
- آموزش کرنل مرکب در راستای حداقل سازی خطای بازسازی نمایش تنک.
- استفاده از فیلتر انطباق کرنلی برای حذف کرنل‌های نامناسب پیش از مرحله آموزش کرنل مرکب.

سازماندهی این مقاله به قرار زیر است: در بخش ۲ به توضیح اجمالی مفاهیم اصلی طبقه‌بند مبتنی بر نمایش تنک (SRC) می‌پردازیم. پس از آن، در بخش ۳، نگاهی کرنلی و مفاهیم مهم مطرح در یادگیری کرنل مرکب تشریح می‌شود. در بخش ۴، پس از تشریح فرآیند کار در طبقه‌بند مبتنی بر نمایش تنک (SRC)، نوع کرنلی این طبقه‌بند (KSRC) معرفی می‌شود. در بخش ۵، روش‌های پایه برای تولید کرنل مرکب در ساختار طبقه‌بند مبتنی بر نمایش تنک کرنلی معرفی می‌شود و در بخش ۶ به معرفی روش‌های پیشنهادی برای تولید کرنل مرکب در ساختار طبقه‌بند مبتنی بر نمایش تنک کرنلی، خواهیم پرداخت. در بخش ۷، شرح آزمایش‌های انجام شده بر روی داده‌های مصنوعی و واقعی، جهت ارزیابی روش‌های پیشنهادی، ارائه شده و مسئله همگرایی الگوریتم‌های پیشنهادی و مقاومت آن‌ها در مواجهه با نویز بررسی می‌شود. در نهایت در بخش ۸، نتایج کلی و پیشنهادات ارائه می‌شود.

۲- طبقه‌بند مبتنی بر نمایش تنک (SRC)

همانطور که قبلاً اشاره شد، نمونه برداری فشرده یکی از حوزه‌های فعال تحقیقاتی است که در سال‌های اخیر در بسیاری از کاربردها از جمله کاربردهای آموزش ماشین مورد استفاده قرار گرفته است. در

خواهد داشت. برای تعیین کرنل مناسب و یا پارامترهای آن یکی از روش‌هایی که ممکن است پیشنهاد شود، ارزیابی سیستم با استفاده از روش اعتبار سنجی متقابل است. اما استفاده از این روش نه تنها زمان‌بر است بلکه تضمینی برای یافتن کرنل بهینه وجود ندارد. از طرفی در برخی از موارد استفاده از یک کرنل، پاسخگوی پیچیدگی داده‌ها یا مسئله مورد بررسی نبوده و به همین دلیل در سال‌های اخیر استفاده از ترکیب مناسب تعدادی از توابع کرنل به جای استفاده از یک تابع کرنل یا به عبارت دیگر، به کارگیری کرنل مرکب پیشنهاد شده است [۱۳]. همانطور که اشاره شد، کرنل‌های مرکب جایگزینی برای انتخاب یک کرنل بهینه در فرآیندی زمان‌بر می‌باشند. کرنل‌های مرکب را می‌توان به طرق مختلف و از دیدگاه‌های مختلف مورد بررسی قرار داد. در حالت کلی کرنل مرکب به صورت زیر تعریف می‌شود [۱۴]:

$$k_{\beta}(\mathbf{x}_i, \mathbf{x}_j) = f_{\beta} \left(\left\{ k_m(\mathbf{x}_i^m, \mathbf{x}_j^m) \right\}_{m=1}^M \right) \quad (6)$$

که در اینجا M تابع کرنل $\{k_m(\dots)\}_{m=1}^M$ به عنوان کرنل‌های پایه (توابع کرنل اولیه) در نظر گرفته شده و $f_{\beta}(\cdot)$ تابع ترکیب این کرنل‌ها است که می‌تواند خطی و یا غیر خطی برحسب کرنل‌های پایه باشد. با توجه به تعریف ارائه شده برای کرنل مرکب در رابطه (۶)، کرنل مرکب را می‌توان به دو طریق و یا با دو هدف ساخت [۱۴].

۱- تولید کرنل مرکب با استفاده از ترکیب توابع کرنل مختلف: کرنل‌های مختلف را می‌توان معادل با معیارهای متفاوتی برای سنجش شباهت داده‌ها دانست. در این صورت به جای جستجو برای پیدا کردن تابع کرنلی که بهترین نمایش از شباهت میان نمونه‌ها را ارائه می‌دهد، می‌توان از یک الگوریتم یادگیری برای انتخاب و یا ترکیب توابع کرنل مختلف (نمایش‌های مختلف از شباهت) استفاده نمود. در این حالت کرنل‌های پایه ممکن است توابع مختلفی مانند کرنل گوسی، خطی و چند جمله‌ای باشند و یا توابع کرنل با استفاده از یک نوع تابع کرنل ولی به ازای پارامترهای مختلف تولید شوند مانند کرنل‌های گوسی با مقادیر واریانس مختلف. در رابطه (۶)، زیر نویس m برای توابع کرنل پایه، $k_m(\dots)$ به همین دلیل انتخاب شده است.

۲- تولید کرنل مرکب با استفاده از منابع ورودی متفاوت: توابع کرنل مختلف ممکن است به واسطه نمایش‌های مختلف ناشی از داده‌های منابع ورودی و یا مودالیتی‌های متفاوت ایجاد شوند. در این حالت ترکیب کرنل‌ها راهی برای ترکیب منابع اطلاعاتی متفاوت است. استفاده از بالانویس m برای نمونه‌های \mathbf{x}_i^m بیانگر فضاهای ورودی مختلف می‌باشد. در توصیف منابع ورودی متفاوت، می‌توان از کرنل‌های متفاوت و در نتیجه معیارهای شباهت سنجی متفاوت نیز استفاده نمود.

تاکنون روش‌های زیادی برای ترکیب توابع کرنل پایه ارائه شده است که در آنها از توابع مختلف $f_{\beta}(\cdot)$ برای ساخت کرنل‌های مرکب و سپس یادگیری این کرنل‌ها استفاده شده است. در این مقاله، ترکیب

$$\begin{aligned} \min_{\mathbf{a}} \quad & \|\mathbf{a}\|_1 \\ \text{subject to} \quad & \|\mathbf{x} - \mathbf{D}\mathbf{a}\|_2 \leq \sigma \end{aligned} \quad (5)$$

مسئله فوق، فرم استاندارد از مسئله بازسازی تنک در حوزه نمونه‌برداری فشرده است و به نام مسئله حداقل سازی نرم ۱ با قیدهای درجه دو شناخته می‌شود [۲].

با وجود کارایی بالای الگوریتم SRC در طبقه‌بندی برخی از داده‌ها، این طبقه‌بند کارایی خود را هنگام مواجهه با داده‌هایی با پراکندگی غیرخطی و همچنین داده‌هایی که در آن، نمونه‌های کلاس‌های مختلف در راستاهای مشابهی توزیع شده‌اند، از دست می‌دهد. از این رو استفاده از کرنل‌ها و نگاشت داده‌ها به فضایی با ابعاد بالاتر می‌تواند گامی مؤثر در جهت افزایش قدرت تفکیک پذیری این الگوریتم بشمار آید. بنابراین در بخش بعد به معرفی کرنل‌ها و مفاهیم اصلی مرتبط با آنها خواهیم پرداخت و سپس در بخش ۴ به معرفی طبقه‌بند مبتنی بر نمایش تنک کرنلیز KSRC خواهیم پرداخت که در آن برای رفع مشکل مورد اشاره، از نگاشت کرنلی استفاده می‌شود.

۳- نگاشت کرنلی و کرنل مرکب

کلمه کرنل به تنهایی معانی متعددی در ریاضیات و علم بازشناسی آماری الگو دارد اما در کاربرد یادگیری ماشین، منظور از کرنل، عموماً تابع (نیمه) معین مثبتی است که می‌توان آن را به عنوان یک معیار سنجش شباهت بین دو نمونه در نظر گرفت. در واقع، شباهت میان زوج نمونه‌ها توسط یک مقدار حقیقی به دست آمده از تابع کرنل مشخص می‌شود. تابع کرنل به صورت $k(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$ تعریف می‌شود که در آن $\phi: \mathbb{R}^m \rightarrow H$ نگاشت غیر خطی از فضای اولیه به فضای ویژگی ناشی از کرنل (فضای هیلبرت H) است و $\langle u, v \rangle$ نشان‌دهنده ضرب نقطه‌ای بین دو بردار داده دلخواه u و v می‌باشد. بنابراین محاسبه تابع کرنل را می‌توان به مثابه انجام عملیات ضرب نقطه‌ای دو داده در فضای هیلبرت (فضای ویژگی) متناظر با آن کرنل در نظر گرفت و این یکی از مهمترین خصوصیات توابع کرنل است که سبب معرفی ترفند کرنل شده است. فرض کنید تابع کرنل k داده شده باشد، به ماتریس \mathbf{K} که اندازه آن $n \times n$ بوده و هر المان آن با رابطه $\mathbf{K}_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$ به دست می‌آید، ماتریس گرام حاصل از تابع کرنل k گفته می‌شود [۱۲]. توابع کرنل چندجمله‌ای، گوسی و سیگموئید نمونه‌هایی از پرکاربردترین توابع کرنل هستند. بنابراین با توجه به ترفند کرنل نیازی به دانستن تابع نگاشت ϕ نبوده و انتخاب تابع کرنل برای نگاشت داده‌ها کافی است.

علی‌رغم عملکرد خوب توابع کرنل در کاربردهای مختلف، انتخاب کرنل مناسب و پارامترهای آن (به عنوان مثال درجه چندجمله‌ای در کرنل چندجمله‌ای و یا پارامتر واریانس در کرنل گوسی) یکی از مهم‌ترین مسائل در روش‌های کرنلی است چرا که تابع کرنل باید متناسب با توزیع داده‌های ورودی انتخاب شود و نوع کرنل انتخاب شده تأثیر زیادی در کیفیت عملکرد الگوریتم یادگیری مورد بررسی

که \mathbf{K} ماتریس کرنل نمونه‌های آموزشی و یا ماتریس گرام است. $\mathbf{k}(\mathbf{x}, \cdot)$ نیز بردار حاصل از اعمال تابع کرنل به نمونه آزمون و نمونه‌های آموزشی است. با انتخاب تابع کرنل، مقادیر ماتریس \mathbf{K} و بردار \mathbf{k} معلوم بوده و بنابراین مسئله فوق مانند مسئله SRC است با این تفاوت که در KSRC، مقادیر متناظر با میزان شباهت بردارهای ویژگی داده‌ها که با تابع کرنل اندازه‌گیری می‌شوند، جایگزین مقادیر بردارهای ویژگی نمونه‌ها می‌شود.

در مرحله طبقه‌بندی، ابتدا ضرایب نمایش تنک با استفاده از کرنل مرکب به دست آمده توسط یکی از روش‌های شرح داده شده در بخش قبل، محاسبه می‌شود. متغیر δ_i برای هر کلاس به صورت زیر تعریف می‌شود:

$$\delta_i(\mathbf{a}) = [0, \dots, 0, \alpha_{i,1}, \dots, \alpha_{i,n_i}, 0, \dots, 0]^T \in \mathbb{R}^{n \times 1} \quad (13)$$

به عبارتی تنها ضرایب مربوط به کلاس i ام حفظ شده و بقیه ضرایب با صفر جایگزین می‌شوند. برچسب کلاس برای نمونه تست \mathbf{x}_{test} با حداقل سازی خطای بازسازی به صورت زیر تخمین زده می‌شود:

$$\hat{y} = \arg \min_{i=1, \dots, C} r_i(\mathbf{x}_{test}) = \|\mathbf{k}(\mathbf{x}_{test}, \cdot) - \mathbf{K}\delta_i\|_2 \quad (14)$$

همانطور که قبلاً نیز اشاره شد، علی‌رغم موفقیت روش‌های کرنلی، انتخاب کرنل بهینه مسئله اصلی در چنین روش‌هایی خواهد بود. بنابراین در طبقه‌بند KSRC نیز از کرنل مرکب استفاده شده است. در ادامه به بررسی برخی از روش‌های پایه استفاده از کرنل مرکب خواهیم پرداخت و سپس ایده اصلی در این پژوهش را مورد بررسی قرار خواهیم داد.

۵- روش‌های پایه در ساخت کرنل مرکب

ترکیب‌های مختلف خطی و غیر خطی را می‌توان برای ساخت کرنل مرکب در نظر گرفت. در این پژوهش، ترکیب خطی کرنل‌های پایه یعنی جمع وزن‌دار کرنل‌ها در نظر گرفته شده و هدف تعیین وزن کرنل‌ها در این ترکیب است. در ادامه، ساختارهای به کار گرفته شده جهت تعیین وزن هر کرنل معرفی می‌شود.

۵-۱- میانگین‌گیری کرنل‌ها

یکی از ساختارها، میانگین‌گیری ساده کرنل‌های پایه است به عبارتی:

$$\mathbf{K}_\beta = \frac{1}{M} \sum_{m=1}^M \mathbf{K}_m \quad (15)$$

علی‌رغم سادگی کرنل میانگین، نتایج خوبی از به کارگیری این کرنل گزارش شده است [۱۸]. همانطور که از رابطه فوق پیداست، در این تابع کرنل، به هر یک از کرنل‌های پایه سهم یکسانی اختصاص داده شده است. بنابراین اگر تمامی کرنل‌های پایه، دقت و یا کارایی خوبی داشته باشند، این کرنل مرکب نیز دقت خوبی خواهد داشت اما در

خطی کرنل‌های پایه یعنی جمع وزن‌دار کرنل‌ها در نظر گرفته شده است.

۴- طبقه‌بند مبتنی بر نمایش تنک کرنلی (KSRC)

با توجه به خواص نگاشت غیر صریح داده‌ها با استفاده از توابع کرنل، نسخه کرنلیزه الگوریتم‌هایی که در آنها فرآیند آموزش با محاسبه ضرب نقطه‌ای بین نمونه‌های آموزشی انجام می‌پذیرد، قابل ارائه است. به عبارت دیگر الگوریتم‌های کرنلی، الگوریتم‌هایی هستند که در آن‌ها تنها به ماتریس کرنل به منظور آموزش نیاز است. بر همین اساس فرم کرنلی الگوریتم SRC نیز ارائه شده است که در ادامه به بررسی آن خواهیم پرداخت.

چنانچه یادآور شدیم در SRC، نمونه آزمون به صورت ترکیب خطی از نمونه‌های آموزشی بیان می‌شود. پس در فرم کرنلیزه این الگوریتم نیز بایستی نگاشت یافته نمونه آزمون در فضای ویژگی جدید به صورت ترکیبی خطی از نگاشت یافته نمونه‌های آموزشی توصیف شود [۹ و ۱۰]، یعنی:

$$\Phi(\mathbf{x}) = \sum_{i=1}^n \alpha_i \Phi(\mathbf{x}_i) = \Phi \mathbf{a} \quad (7)$$

که $\mathbf{a} = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ بردار ضرایب تنک و α_i مؤلفه متناظر با نمونه $\Phi(\mathbf{x}_i)$ است. ماتریس نمونه‌ها در فضای هیلبرت H با بعد d نیز به صورت زیر نمایش داده می‌شود:

$$\Phi = [\Phi(\mathbf{x}_1), \Phi(\mathbf{x}_2), \dots, \Phi(\mathbf{x}_n)] \in \mathbb{R}^{d \times n} \quad (8)$$

بنابراین مسئله یافتن نمایش تنک در فضای ویژگی به صورت زیر تعریف می‌شود:

$$\min_{\mathbf{a}} \|\mathbf{a}\|_1 \quad \text{subject to } \Phi(\mathbf{x}) = \Phi \mathbf{a} \quad (9)$$

و برای حالتی که نویز (و یا خطا) در فرآیند بازسازی در نظر گرفته شود، مسئله فوق به صورت زیر خواهد بود:

$$\min_{\mathbf{a}} \|\mathbf{a}\|_1 \quad \text{subject to } \|\Phi(\mathbf{x}) - \Phi \mathbf{a}\|_2 \leq \sigma \quad (10)$$

می‌توان نشان داد که برای هر $\sigma \geq 0$ باید $\zeta \geq 0$ وجود داشته باشد، به طوری که وقتی رابطه $\|\Phi(\mathbf{x}) - \Phi \mathbf{a}\|_2 \leq \sigma$ برقرار است، رابطه $\|\Phi^T \Phi(\mathbf{x}) - \Phi^T \Phi \mathbf{a}\|_2 \leq \zeta$ نیز برقرار باشد [۹].

بنابراین مسئله نمایش تنک به صورت زیر خواهد بود:

$$\min_{\mathbf{a}} \|\mathbf{a}\|_1 \quad \text{subject to } \|\Phi^T \Phi(\mathbf{x}) - \Phi^T \Phi \mathbf{a}\|_2 \leq \zeta \quad (11)$$

رابطه فوق را می‌توان به صورت زیر نوشت:

$$\min_{\mathbf{a}} \|\mathbf{a}\|_1 \quad \text{subject to } \|\mathbf{k}(\mathbf{x}, \cdot) - \mathbf{K} \mathbf{a}\|_2 \leq \zeta \quad (12)$$

واضح است که در این روش، وزن کرنل‌ها مستقل از الگوریتم نمایش ضرایب تنک محاسبه می‌شود. در نهایت، برای طبقه‌بندی داده‌ها از الگوریتم KSRC استفاده می‌شود با این تفاوت که به جای استفاده از یک کرنل تنها، از کرنل مرکب با وزن‌های به دست آمده از رابطه (۱۹) استفاده می‌شود.

۳-۵- وزن‌دهی بر مبنای کیفیت عملکرد هر کرنل

برای محاسبه وزن کرنل‌ها در ترکیب خطی، می‌توان از کیفیت عملکرد هر کرنل به تنهایی نیز استفاده کرد که در این صورت، وزن کرنل‌ها به صورت زیر محاسبه می‌شود [۲۰]:

$$\beta_m = \frac{\pi_m - \nu}{\sum_{m=1}^M (\pi_m - \nu)} \quad \forall m=1, 2, \dots, M \quad (20)$$

که π_m بیانگر میزان دقت طبقه‌بند وقتی از تابع کرنل k_m استفاده شود، است و ν حد آستانه متناظر با حداقل دقت قابل قبول است. کرنل‌های با دقتی کمتر از مقدار آستانه در ترکیب کرنلی لحاظ نمی‌شوند. در الگوریتم ۱، فرآیند تعیین وزن کرنل‌ها به‌طور خلاصه نشان داده شده است. در این حالت نیز، پس از محاسبه وزن کرنل‌ها، از این کرنل مرکب، در چارچوب مسئله KSRC استفاده می‌شود. در بخش بعد، روش‌های پیشنهادی در این مقاله برای بهینه‌سازی وزن کرنل‌ها معرفی می‌شود.

الگوریتم ۱. وزن‌دهی مبتنی بر دقت عملکرد، در کرنل مرکب ورودی: مجموعه ماتریس‌های گرام متناظر با توابع کرنل پایه $\{\mathbf{K}_m\}_{m=1}^M$.

خروجی: بردار وزن کرنل‌ها β .

گام اول: محاسبه ضرایب نمایش تنک برای داده‌های آموزشی. برای هر کرنل، نمایش تنک هر یک از نمونه‌های آموزشی با استفاده از رابطه (۱۲) محاسبه می‌شود. برای این منظور، سطر و ستون مربوط به آن نمونه از ماتریس واژنامه حذف می‌شود.

گام دوم: محاسبه وزن کرنل‌ها برای هر کرنل، برچسب کلاس هر یک از نمونه‌ها با حداقل سازی خطای بازسازی $\arg \min_{i=1, \dots, c} r_i^m(\mathbf{x}) = \|\mathbf{k}_m(\mathbf{x}, \cdot) - \mathbf{K}_m \delta_i\|_2$ تخمین زده می‌شود. سپس با توجه به برچسب‌های تخمین زده شده برای نمونه‌های آموزشی، دقت طبقه‌بندی هر کرنل محاسبه می‌شود و در نهایت وزن کرنل‌ها از رابطه (۲۰) محاسبه می‌شود.

شکل ۱: وزن‌دهی مبتنی بر دقت عملکرد در کرنل مرکب

۶- روش‌های پیشنهادی در ساخت کرنل مرکب

در بخش قبل روش‌های پایه برای ساخت کرنل مرکب ارائه شد. اگر چه این روش‌ها به عنوان روش‌های پایه ارائه شد اما تا کنون از این روش‌ها به طور مستقیم در ساخت کرنل مرکب برای طبقه‌بند KSRC استفاده نشده است. پس از معرفی این روش‌های پایه، در این بخش روش‌های

صورتی که کرنل‌های نویزی و یا کرنل‌های با دقت پایین نیز در مجموعه کرنل‌های پایه وجود داشته باشد، کرنل مرکب فوق، از آنجایی که به همه کرنل‌ها وزن یکسانی اختصاص می‌دهد، سبب کاهش دقت خواهد شد. بنابراین در حالت کلی، جمع وزن‌دار کرنل‌ها اولویت دارد. جمع وزن‌دار کرنل‌ها به صورت زیر تعریف می‌شود:

$$\mathbf{K}_\beta = \sum_{m=1}^M \beta_m \mathbf{K}_m \quad (16)$$

که β_m وزن متناظر با کرنل \mathbf{K}_m است. روش‌های مختلفی برای تعیین وزن کرنل‌ها وجود دارد که در ادامه به معرفی برخی از آن‌ها خواهیم پرداخت.

۵-۲- وزن‌دهی با استفاده از انطباق کرنلی

در [۱۹]، انطباق میان دو ماتریس کرنل که با نام انطباق کرنلی شناخته می‌شود، معرفی شد که معیاری بر گرفته از معیار فاصله کسینوسی^۵ بوده و برای محاسبه میزان شباهت دو ماتریس استفاده می‌شود. رابطه انطباق کرنلی میان دو ماتریس کرنل \mathbf{K}_1 و \mathbf{K}_2 به صورت زیر است:

$$A(\mathbf{K}_1, \mathbf{K}_2) = \frac{\langle \mathbf{K}_1, \mathbf{K}_2 \rangle_F}{\sqrt{\langle \mathbf{K}_1, \mathbf{K}_1 \rangle_F \langle \mathbf{K}_2, \mathbf{K}_2 \rangle_F}} \quad (17)$$

که در رابطه فوق، $\langle \mathbf{K}_1, \mathbf{K}_2 \rangle_F$ نمایشی از ضرب فروبنیوس^۶ بین دو ماتریس \mathbf{K}_1 و \mathbf{K}_2 به صورت

$$\langle \mathbf{K}_1, \mathbf{K}_2 \rangle_F = \sum_{i=1}^n \sum_{j=1}^n k_1(\mathbf{x}_i, \mathbf{x}_j) k_2(\mathbf{x}_i, \mathbf{x}_j)$$

انطباق هر کرنل با کرنل ایده‌آل سنجیده می‌شود. ماتریس کرنل ایده‌آل به صورت $\mathbf{K}_y = \mathbf{y}\mathbf{y}^T$ تعریف می‌شود که در آن بردار برچسب طبقه متناظر با نمونه‌های آموزشی است و $\mathbf{K}_y(i, j)$ زمانی برابر با یک است که برچسب کلاس دو نمونه یکسان باشد و در غیر اینصورت $\mathbf{K}_y(i, j)$ برابر با صفر خواهد بود. فرض کنید داده‌های آموزشی از ۳ طبقه تشکیل شده و در هر طبقه ۲ نمونه وجود داشته باشد، ماتریس کرنل ایده‌آل برای این داده‌ها به صورت زیر خواهد بود:

$$\mathbf{K}_y = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (18)$$

پس از معرفی کرنل ایده‌آل و رابطه انطباق کرنلی، وزن کرنل‌ها را می‌توان بر اساس میزان انطباق آنها با کرنل ایده‌آل و به صورت زیر محاسبه کرد:

$$\beta_m = \frac{A(\mathbf{K}_m, \mathbf{K}_y)}{\sum_{h=1}^M A(\mathbf{K}_h, \mathbf{K}_y)} \quad \forall m=1, 2, \dots, M \quad (19)$$

نمونه آموزشی، برابر با صفر خواهد شد و مقدار ضریب مربوط به همان نمونه آموزشی برابر با یک خواهد بود. همچنین در رابطه فوق از قید نرم ۲ برای بردار وزن کرنل‌ها به صورت $\sum_{m=1}^M \beta_m^2 = 1$ استفاده شده است. این انتخاب علاوه بر اینکه مثبت بودن وزن کرنل‌ها و در نتیجه معتبر بودن کرنل مرکب را تضمین می‌کند، باعث ساده شدن روابط بعدی و در نتیجه وجود حل تحلیلی برای وزن کرنل‌ها می‌شود. تابع لاگرانژ مسئله فوق به صورت زیر خواهد بود:

$$J = \|\Phi(\mathbf{X}) - \Phi\mathbf{U}\|_2 + \lambda \|\mathbf{U}\|_1 + \gamma(\beta^T \beta - 1) \quad (22)$$

λ مقدار ثابتی است که بین میزان تنگی و مقدار خطای بازسازی مصالحه ایجاد می‌کند. در رابطه فوق β بردار وزن کرنل‌ها به صورت $\beta = [\beta_1, \beta_2, \dots, \beta_M]$ و همچنین γ ضریب لاگرانژ است. مسئله فوق معادل با n مسئله لاگرانژ مستقل است بنابراین خواهیم داشت:

$$\begin{aligned} J &= \|\Phi(\mathbf{X}) - \Phi\mathbf{U}\|_2 + \lambda \|\mathbf{U}\|_1 + \gamma(\beta^T \beta - 1) \\ &= \sum_{i=1}^n \|\Phi(\mathbf{x}_i) - \Phi\mathbf{a}_i\|_2 + \sum_{i=1}^n \lambda_i \|\mathbf{a}_i\|_1 + \gamma(\beta^T \beta - 1) \\ &= \sum_{i=1}^n [\mathbf{K}(\mathbf{x}_i, \mathbf{x}_i) - 2\mathbf{a}_i^T \mathbf{K}(\mathbf{x}_i, \cdot) + \mathbf{a}_i^T \mathbf{K} \mathbf{a}_i] \\ &\quad + \sum_{i=1}^n \lambda_i \|\mathbf{a}_i\|_1 + \gamma(\beta^T \beta - 1) \end{aligned} \quad (23)$$

با در نظر گرفتن کرنل مرکب به صورت رابطه (۱۴) و جایگذاری آن در رابطه (۲۱) خواهیم داشت:

$$\begin{aligned} J &= \sum_{i=1}^n \sum_{m=1}^M \beta_m \mathbf{K}_m(\mathbf{x}_i, \mathbf{x}_i) - 2 \sum_{i=1}^n \mathbf{a}_i^T \sum_{m=1}^M \beta_m \mathbf{K}_m(\mathbf{x}_i, \cdot) \\ &\quad + \sum_{i=1}^n \mathbf{a}_i^T \sum_{m=1}^M \beta_m \mathbf{K}_m \mathbf{a}_i + \sum_{i=1}^n \lambda_i \|\mathbf{a}_i\|_1 + \gamma(\beta^T \beta - 1) \end{aligned} \quad (24)$$

با مشتق‌گیری از رابطه فوق نسبت به وزن کرنل m ام یعنی β_m خواهیم داشت:

$$\begin{aligned} \frac{\partial J}{\partial \beta_m} &= \sum_{i=1}^n \mathbf{K}_m(\mathbf{x}_i, \mathbf{x}_i) - 2 \sum_{i=1}^n \mathbf{a}_i^T \mathbf{K}_m(\mathbf{x}_i, \cdot) \\ &\quad + \sum_{i=1}^n \mathbf{a}_i^T \mathbf{K}_m \mathbf{a}_i + 2\gamma\beta_m = 0 \end{aligned} \quad (25)$$

بنابراین وزن کرنل β_m به صورت زیر محاسبه می‌شود:

$$\beta_m = \frac{-1}{2\gamma} \eta_m \quad (26)$$

که

$$\eta_m = \sum_{i=1}^n \mathbf{K}_m(\mathbf{x}_i, \mathbf{x}_i) - 2 \sum_{i=1}^n \mathbf{a}_i^T \mathbf{K}_m(\mathbf{x}_i, \cdot) + \sum_{i=1}^n \mathbf{a}_i^T \mathbf{K}_m \mathbf{a}_i \quad (27)$$

همچنین با مشتق‌گیری از تابع لاگرانژ در رابطه (۲۲) نسبت به ضریب لاگرانژ γ و استفاده از رابطه (۲۴) خواهیم داشت:

پیشنهادی برای ساخت کرنل را معرفی خواهیم نمود که در آنها، وزن کرنل‌ها در چهارچوب مسئله KSRC تعیین می‌شود.

۶-۱- بهینه‌سازی وزن کرنل‌ها در چارچوب مسئله نمایش تنک بدون استفاده از برچسب کلاس داده‌های آموزشی

در بخش قبل، روش‌های پایه میانگین‌گیری، وزن‌دهی با استفاده از انطباق کرنلی و همچنین وزن‌دهی بر مبنای کیفیت عملکرد هر کرنل مورد بررسی قرار گرفت. در [۱۶] از انطباق کرنلی و دقت هر کرنل در یک فرآیند تکراری برای محاسبه وزن کرنل‌ها استفاده شده است. برای این منظور، ابتدا کرنل‌ها با توجه به میزان انطباق کرنلی، مرتب شده و سپس برچسب کلاس نمونه‌های آموزشی با استفاده از هر یک از کرنل‌ها و همچنین با استفاده از کرنل مرکب به دست آمده از مرحله قبل، تخمین زده می‌شود. دقت هر کرنل با توجه به نمونه‌هایی که توسط کرنل مرکب به اشتباه طبقه‌بندی شده‌اند، مشخص می‌شود. سپس از بین کرنل‌ها، کرنلی که بیشترین دقت (منظور از دقت، دقت به دست آمده بر روی داده‌های آموزشی است که کرنل مرکب به اشتباه طبقه‌بندی نموده است) و همچنین بالاترین درجه انطباق با کرنل ایده‌آل را داشته باشد، انتخاب می‌شود. سپس وزن کرنل انتخاب شده و کرنل مرکب مرحله قبل به روز رسانی می‌شود و این فرآیند تا برقراری شرط توقف تکرار می‌شود. به عبارتی کرنل‌هایی که دارای اطلاعات تکمیلی در طبقه‌بندی نمونه‌های آموزشی هستند، وزن بیشتری می‌گیرند.

در این پژوهش، بهینه‌سازی وزن کرنل‌ها در فرآیند طبقه‌بندی مبتنی بر نمایش تنک داده‌ها ارائه شده است. براین اساس، هدف بهینه‌سازی وزن کرنل‌ها در جهت حداقل سازی خطای بازسازی نمایش تنک در نظر گرفته شده است. برای این منظور فرض کنید، نمونه‌های آموزشی در فضای ویژگی با $\Phi(\mathbf{x}_i)$ نمایش داده شود. حال با در نظر گرفتن کرنل مرکب مطابق رابطه (۱۶)، می‌توان مسئله بهینه‌سازی زیر را در نظر گرفت:

$$\begin{aligned} \min_{\mathbf{a}, \beta} & \|\Phi(\mathbf{X}) - \Phi\mathbf{U}\|_2 \\ \text{subject to} & \|\mathbf{a}_i\|_1 \leq \zeta \quad i = 1, \dots, n \\ & \sum_{m=1}^M \beta_m^2 = 1 \end{aligned} \quad (21)$$

که در رابطه فوق $\Phi(\mathbf{X}) = [\Phi(\mathbf{x}_1), \Phi(\mathbf{x}_2), \dots, \Phi(\mathbf{x}_n)]$ ماتریسی است که از کنار هم قرار دادن نمونه‌های آموزشی به دست می‌آید. همچنین $\mathbf{U} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] \in \mathbb{R}^{n \times n}$ ، ماتریسی است که از قرار دادن بردار ضرایب نمایش تنک نمونه‌های آموزشی در ستون‌های آن به دست می‌آید. قابل ذکر است که برای محاسبه ضرایب نمایش تنک مربوط به هریک از نمونه‌های آموزشی، ستون مربوط به آن نمونه آموزشی در واژه‌نامه با یک بردار صفر جایگزین می‌شود زیرا در غیراینصورت، نمایش تنک به دست آمده در تمام مولفه‌ها غیر از ضریب مربوط به همان

۷- شرح آزمایش‌ها

$$\gamma = \frac{\sqrt{\sum_{m=1}^M (\eta_m)^2}}{2} \quad (28)$$

به منظور بررسی کیفیت عملکرد طبقه‌بند مبتنی بر نمایش تنک کرنلی شده با استفاده از کرنل‌های مرکب مختلف، آزمایش‌ها بر روی دادگان واقعی و مصنوعی انجام شد. در این بخش، ابتدا دادگان مورد استفاده در این پژوهش را معرفی کرده و سپس نحوه انجام آزمایش‌ها و پارامترهای مورد استفاده در انجام آزمایش‌ها را تشریح می‌کنیم. در نهایت نتایج به دست آمده از هر روش ارائه و تحلیل می‌شود.

۷-۱- معرفی داده‌های مورد استفاده در آزمایش‌ها

در انجام آزمایش‌ها از دو مجموعه داده‌های واقعی استفاده شد. مجموعه اول دادگان، داده‌های جمع آوری شده در دانشگاه کالیفرنیا در ایروین، UCI^{۱۰} است. پایگاه داده UCI شامل مجموعه بزرگی از داده‌ها است که از آن‌ها برای کاربردهای مختلف طبقه‌بندی، خوشه‌بندی و رگرسیون استفاده می‌شود. ما از مجموعه‌های Iris و Sonar، Wine، Ionosphere، Heart، Soybean و Iris برای انجام آزمایش‌ها استفاده کرده‌ایم. همچنین با در نظر گرفتن کاربرد بازشناسی اعداد دست‌نویس، چگونگی عملکرد الگوریتم‌های پیشنهادی بر روی پایگاه داده MNIST نیز مورد بررسی قرار گرفت. مشخصات این داده‌ها در جدول ۱ خلاصه شده است. در این جدول، n تعداد کل نمونه‌ها، m بعد داده‌ها در فضای اولیه و c تعداد طبقات داده‌ها را نشان می‌دهد.

جدول ۱: مشخصات داده‌های مورد استفاده در آزمایش‌ها

Data Set	n	m	c
Soybean	47	35	4
Heart	270	13	2
Ionosphere	351	34	2
Wine	178	13	3
Sonar	208	60	2
Iris	150	4	3
MNIST	70000	784	10

۷-۲- چگونگی انجام آزمایش‌ها و نتایج حاصل

در انجام آزمایش‌ها، ابتدا هر مجموعه از داده نرمالیزه می‌شوند به گونه‌ای که بردارهای ویژگی جدید دارای متوسط صفر و انحراف معیار یک باشند. سپس داده‌ها به صورت تصادفی به ۱۰ بسته تقسیم شده و هر بار ۹ بسته به عنوان داده‌های آموزشی و ۱ بسته به عنوان داده آزمایشی در نظر گرفته می‌شود. نتایج گزارش شده، میانگین نتایج به دست آمده از ۱۰ بار تکرار الگوریتم‌ها است. پارامتر حد آستانه، ν ، در روش تعیین وزن کرنل‌ها بر اساس دقت برابر با $\nu = \min(\pi_m) - 0.1$ در نظر گرفته شد. واضح است که با انتخاب ν به اینصورت، وزن هیچ یک از کرنل‌ها صفر نشده و هیچ کرنلی حذف نمی‌شوند. برای بررسی

در رابطه (۲۶)، فرض شده است که بردار ضرایب نمایش تنک نمونه‌های آموزشی معلوم است، بنابراین بهینه سازی وزن کرنل‌ها با لحاظ نمودن یک فرآیند تکراری، در دو مرحله صورت می‌گیرد:

✓ محاسبه ضرایب تنک با فرض معلوم بودن وزن کرنل‌ها

✓ محاسبه وزن کرنل‌ها با ضرایب تنک به دست آمده از مرحله قبل

این مراحل تا برقراری شرط توقف تکرار می‌شوند. در این مقاله، شرط همگرایی به صورت $\|\beta^{j+1} - \beta^j\| \leq \tau$ در نظر گرفته شده است. با فرض معلوم بودن وزن کرنل‌ها، ضرایب تنک از رابطه (۱۲) وقتی از کرنل مرکب استفاده شود، محاسبه می‌شود.

۶-۲- بهینه سازی وزن کرنل‌ها در چارچوب مسئله نمایش تنک با استفاده از برچسب کلاس داده‌های آموزشی

همانطور که مشخص است در رابطه وزن فوق، از برچسب نمونه‌های آموزشی استفاده نشده است، برای استفاده از برچسب کلاس نمونه‌های آموزشی در رابطه وزن، می‌توان به جای \mathbf{a}_i از $\mathbf{a}'_i = \mathbf{a}_i \cdot \mathbf{\theta}_i$ استفاده نمود که در آن $\mathbf{\theta}_i = [0, \dots, 0, 1, \dots, 1, 0, \dots, 0]$ می‌باشد که محل‌های غیر صفر $\mathbf{\theta}_i$ متناظر با نمونه‌های آموزشی مربوط به کلاس نمونه i ام است. با استفاده از \mathbf{a}'_i ، از اطلاعات موجود در برچسب کلاس نمونه‌های آموزشی نیز در بهینه سازی وزن کرنل‌ها استفاده می‌شود. خلاصه فرآیند آموزش وزن کرنل‌ها و ضرایب تنک با استفاده از برچسب کلاس نمونه‌های آموزش و بدون استفاده از آن در الگوریتم ۲ ارائه شده است.

الگوریتم ۲. روش پیشنهادی ساخت کرنل مرکب در چهارچوب

طبقه‌بند KSRC

ورودی: مجموعه کرنل‌های پایه $\{\mathbf{K}_m\}_{m=1}^M$.

خروجی: ماتریس ضرایب تنک \mathbf{U} ، بردار وزن کرنل‌ها β .

مقادیر اولیه: $\beta^{(0)} = \frac{1}{M}$ و متغیر تکرار $J = 1$

گام اول: محاسبه ضرایب نمایش تنک

محاسبه ضرایب تنک $\mathbf{U}^{(J)}$ در رابطه (۱۲) وقتی کرنل مرکب با بردار وزن $\beta^{(J-1)}$ ساخته شده باشد.

گام دوم: به روزرسانی وزن کرنل‌ها

محاسبه وزن کرنل‌ها $\beta^{(J)}$ در رابطه (۲۶) با استفاده از ضرایب نمایش تنک به دست آمده از گام قبل $\mathbf{U}^{(J)}$ و $\mathbf{U}^{(J)} = \mathbf{U}^{(J-1)} \mathbf{\theta}$.

تنظیم $J = J + 1$ و تکرار الگوریتم تا رسیدن به شرط توقف $\|\beta^{j+1} - \beta^j\| \leq \tau$.

شکل ۲: روش پیشنهادی ساخت کرنل مرکب در چارچوب طبقه‌بند

KSRC

نتایج به دست آمده در فاز اول آزمایش‌ها در جدول شماره ۲ خلاصه شده است. همانطور که از نتایج به دست آمده مشخص است در درجه اول، استفاده از KSRC منجر به بهبود قابل توجه نتایج در مقایسه با SRC می‌شود که تأکیدی بر عملکرد خوب روش مبتنی بر کرنل است. همچنین می‌توان دید که با استفاده از کرنل مرکب نتایج بهتری در مقایسه با بهترین کرنل به تنهایی، به دست می‌آید. توجه شود که برای هر مجموعه داده، بهترین کرنل، کرنل متفاوتی ممکن است باشد و در عمل با استفاده از یک کرنل به تنهایی نتایج ضعیف‌تری به دست می‌آید. واضح است که با در نظر گرفتن بخشی از داده‌های آموزشی به‌عنوان داده‌های ارزیابی، می‌توان بهترین کرنل را تعیین نمود اما این فرآیند، زمان‌بر بوده و لزوماً همیشه بهترین کرنل در مرحله ارزیابی، منجر به بهترین نتایج در مرحله آزمون نمی‌شود. منظور اینکه، نتایج مرتبط با KSRC(Best)، در واقع حد بالای کیفیت مورد انتظار با در نظر گرفتن یک کرنل از میان کرنل‌های موجود را ارائه می‌دهد. حال آنکه الگوریتم‌های مبتنی بر کرنل مرکب، در شرایط مختلف منجر به نتایج بهتری در مقایسه با الگوریتم SRC و KSRC(Best) شده‌اند. با توجه به نتایج جدول ۲، در میان الگوریتم‌های KSRC مبتنی بر کرنل مرکب پیاده‌سازی شده نیز، به طور متوسط الگوریتم MKSRC(KA) دارای بهترین نتایج بوده است. همچنین با توجه به نتایج به دست آمده برای الگوریتم مبتنی بر دقت یعنی MKSRC(ACC)، می‌توان چنین استنباط نمود که دقت کرنل‌های استفاده شده تقریباً مشابه بوده است. از این رو در فاز بعدی، مجموعه بزرگتری از کرنل‌ها با پارامترهای مختلف در نظر گرفته شد، که در ادامه به بررسی این فاز از آزمایش‌ها خواهیم پرداخت.

۷-۴- فاز دوم آزمایش‌ها

در این فاز از آزمایش‌ها، از مجموعه بزرگتری از کرنل‌های پایه برای ساخت کرنل مرکب استفاده شده که این کرنل‌ها عبارتند از:

- کرنل گوسی با پارامترهای $\sigma = \{0.1, 1, 10, 100\}$.
- کرنل چندجمله‌ای با درجه‌های $d = \{2, 3, \dots, 10\}$ و $a = 1$.
- کرنل چندجمله‌ای با $a = \{0.1, 5, 10, 20, 100\}$ و $d = 2$.
- کرنل لاپلاسی به صورت $k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\delta}\right)$ با پارامتر $\delta = \{0.1, 1, 10, 100\}$.
- کرنل حلقوی به صورت $k(\mathbf{x}_i, \mathbf{x}_j) = \tanh(v\langle \mathbf{x}_i, \mathbf{x}_j \rangle + \tau)$ با $v = 1$ و $\tau = \{0.1, 0.2, 0.3, 0.4, 0.5\}$.
- کرنل حلقوی با $v = \{0.2, 0.4, 0.6, 0.8\}$ و $\tau = 0.5$.

نتایج به دست آمده در این فاز از آزمایش‌ها در جدول ۳ خلاصه شده است. کاهش دقت و افت نتایج در الگوریتم‌های مبتنی بر کرنل مرکب، در این فاز از آزمایش‌ها در مقایسه با نتایج جدول ۲، بیانگر وجود کرنل‌های نامناسب در میان مجموعه کرنل‌های پایه انتخابی

عملکرد طبقه‌بند مبتنی بر نمایش تنک در استفاده از کرنل‌های مرکب، آزمایش‌ها در فازهای مختلفی صورت گرفت که در ادامه به بررسی هر فاز و تحلیل نتایج آن خواهیم پرداخت. الگوریتم‌های مقایسه شده در این پژوهش عبارتند از:

- (۱) الگوریتم SRC.
- (۲) الگوریتم KSRC با استفاده از بهترین کرنل از میان کرنل‌های انتخاب شده با نام KSRC(Best).
- (۳) الگوریتم KSRC با استفاده از کرنل مرکب حاصل از میانگین‌گیری ساده کرنل‌ها با نام KSRC(Mean).
- (۴) الگوریتم KSRC با استفاده از کرنل مرکب حاصل از ترکیب خطی وزن‌دار کرنل‌ها که وزن‌ها با استفاده از روش انطباق کرنلی محاسبه می‌شود، با نام MKSRC(KA).
- (۵) الگوریتم KSRC با استفاده از کرنل مرکب حاصل از ترکیب خطی وزن‌دار کرنل‌ها که وزن‌ها با استفاده از دقت هر کرنل محاسبه می‌شود، با نام MKSRC(ACC).
- (۶) الگوریتم KSRC با استفاده از کرنل مرکب حاصل از ترکیب خطی وزن‌دار کرنل‌ها که وزن‌ها با استفاده از رابطه (۲۶) محاسبه می‌شود، با نام MKSRC(EV).
- (۷) الگوریتم KSRC با استفاده از کرنل مرکب حاصل از ترکیب خطی وزن‌دار کرنل‌ها که وزن‌ها با به کارگیری برچسب کلاس نمونه‌های آموزشی در فرآیند آموزش وزن‌ها، محاسبه می‌شود، با نام MKSRC(EV-Label).

۷-۳- فاز اول آزمایش‌ها

در فاز اول آزمایش‌ها، فرض شد که مجموعه کرنل‌های پایه از کرنل گوسی با پارامترهای $\sigma = \{0.1, 1, 10, 100\}$ و کرنل چندجمله‌ای با درجه‌های $d = \{2, 3, \dots, 10\}$ تشکیل شده است. روابط کرنل گوسی و کرنل چندجمله‌ای به صورت زیر است:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (29)$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + a)^d$$

که در کرنل چندجمله‌ای، پارامتر a ثابت و برابر ۱ در نظر گرفته شد. در روش تعیین وزن کرنل‌ها با استفاده از دقت کرنل، به علت حجم بالای محاسبات، از ۳۰ درصد از نمونه‌های آموزشی به عنوان داده‌های ارزیابی و برای تعیین وزن کرنل‌ها استفاده شده است. در این روش، با استفاده از هر کرنل به صورت جداگانه، داده‌های ارزیابی طبقه‌بندی شده و با استفاده از آن دقت هر کرنل و وزن متناظر با آن محاسبه می‌شود.

۷-۵- فاز سوم آزمایش‌ها

همانطور که اشاره شد، از نتایج به دست آمده در فاز قبلی آزمایش‌ها می‌توان چنین استنباط نمود که مجموعه‌ای از کرنل‌های نویزی و یا کرنل‌های نامناسب نیز در میان مجموعه کرنل‌های فاز قبل وجود دارد. بنابراین در این فاز از آزمایش‌ها، کرنل‌ها را بر اساس میزان انطباق آن‌ها با کرنل ایده‌آل با استفاده از رابطه (۱۷)، فیلتر نموده و کرنل‌هایی که میزان انطباق آن‌ها از حد آستانه‌ای بیشتر باشد برای ساخت کرنل مرکب مورد استفاده قرار می‌گیرند. در اینجا حد آستانه برابر با متوسط انطباق به دست آمده از مقایسه هر کرنل با کرنل ایده‌آل در نظر گرفته شده است.

است. در این فاز از آزمایش‌ها می‌توان به تأثیر مجموعه کرنل‌های انتخاب شده بر روی دقت الگوریتم‌های مبتنی بر کرنل پی برد. همانطور که در بخش ۶ نیز شرح داده شد، در الگوریتم‌های پیشنهادی از قید محدب برای وزن کرنل‌ها استفاده شده است که سبب می‌شود هیچ کرنلی در فرآیند آموزش حذف نشود حتی اگر در فرآیند آموزش وزنی بسیار کوچک به آن اختصاص داده شود. از طرفی با مقایسه نتایج مربوط به بهترین کرنل (KSRC(Best)) در دو جدول ۲ و ۳ می‌توان دید که کرنل‌های با دقت بالاتر نیز در میان مجموعه کرنل‌های استفاده شده در فاز دوم در مقایسه با فاز اول وجود دارد. بنابراین باید کرنل‌های نامناسب موجود در این مجموعه را با استفاده از فیلتری حذف نمود. از این رو در فاز بعدی آزمایش‌ها به طراحی فیلتری برای حذف کرنل‌های نامناسب خواهیم پرداخت.

جدول ۲: نتایج به دست آمده در فاز اول آزمایش‌ها

Method	Soybean	Heart	Ionosphere	Wine	Sonar	Iris
SRC	98.0000	77.7778	92.6256	89.8366	83.1667	93.7355
KSRC(Best)	100.0000	82.5926	88.7963	96.6667	89.4236	98.0000
KSRC(mean)	100.0000	81.4815	93.7037	98.3333	88.0451	98.0000
KSRC(KA)	100.0000	81.8519	94.4444	98.8889	89.3734	98.0000
KSRC(ACC)	100.0000	81.4815	92.9630	98.3333	85.0877	97.3333
KSRC(EV)	100.0000	79.6296	90.7407	96.6667	81.7043	98.0000
KSRC(EV-Label)	100.0000	83.7037	90.7407	97.7778	85.6140	96.6667

جدول ۳: نتایج به دست آمده در فاز دوم آزمایش‌ها

Method	Soybean	Heart	Ionosphere	Wine	Sonar	Iris
SRC	98.0000	77.7778	92.6256	89.8366	83.1667	93.7355
KSRC(Best)	100.0000	83.3333	89.8148	99.4444	89.3734	97.3333
KSRC(mean)	100.0000	72.2222	87.3148	95.4861	83.5088	97.3333
KSRC(KA)	100.0000	75.1852	88.7963	96.0417	79.7995	97.3333
KSRC(ACC)	94.0000	71.8519	83.3333	94.9306	75.4637	98.0000
KSRC(EV)	92.0000	70.3704	83.7963	88.8194	67.2180	97.3333
KSRC(EV-Label)	100.0000	76.6667	93.3333	95.0000	85.6140	96.6667

الگوریتم‌های مبتنی بر کرنل مرکب و به خصوص الگوریتم (KSRC(EV-Label)) MKSRC شده است چرا که در الگوریتم‌های پیشنهادی، وزن کرنل‌ها در جهت حداقل سازی خطای بازسازی نمایش تنک بهینه می‌شوند.

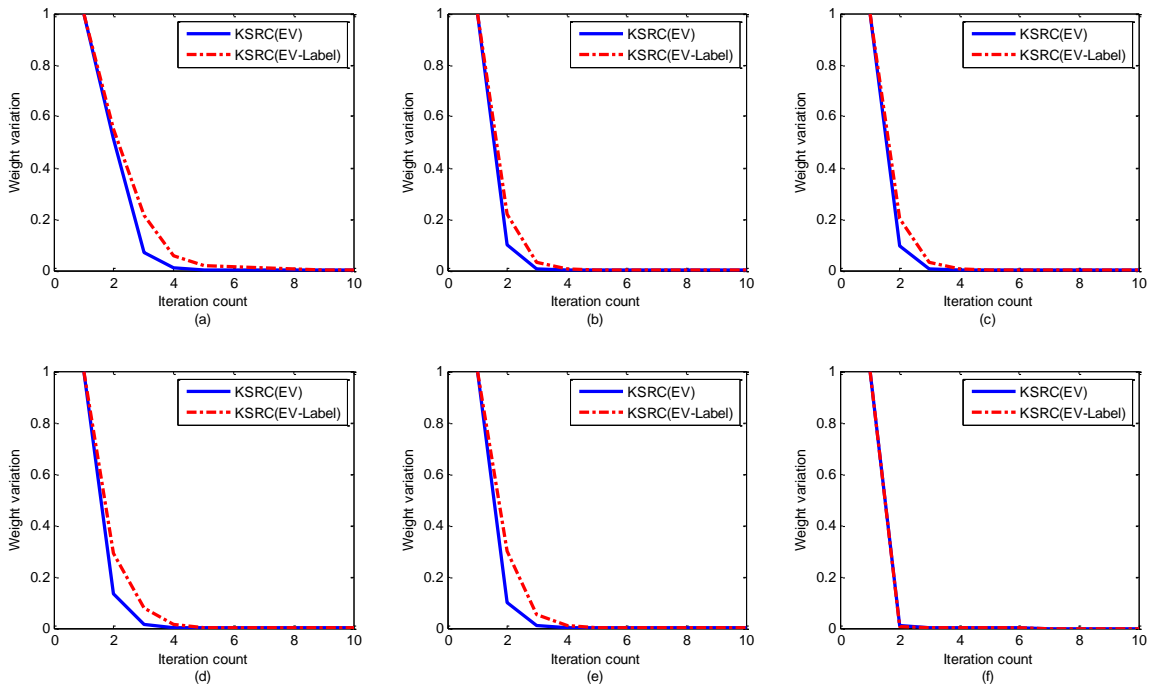
نتایج به دست آمده در این فاز از آزمایش‌ها در جدول ۴ خلاصه شده است. همانطور که انتظار می‌رفت، حذف کرنل‌های نامناسب با استفاده از فرآیند فیلتر کردن شرح داده شده، سبب بهبود قابل توجهی در نتایج

تصادفی با انحراف معیارهای مختلف باشند، نشان داده شده است. در این آزمایش‌ها انحراف معیار نویز گوسی از صفر (بدون نویز) تا دو با اندازه گام 0.2 تغییر داده شده است. چنانچه قبلاً ذکر شد، داده‌های ورودی ابتدا نرمالیزه شده‌اند. همانطور که نتایج ارائه شده نشان می‌دهد، روش‌های مبتنی بر کرنل، در شرایط نویزی عملکرد بهتری در مقایسه با الگوریتم پایه SRC دارند. همچنین در بین روش‌های مبتنی بر کرنل، روش پیشنهادی (KSRC(EV-Label)) به طور متوسط عملکرد بهتری در مقایسه با سایر روش‌ها دارد.

همگرایی بردار وزن کرنل‌ها برای الگوریتم‌های پیشنهادی در فاز سوم (به عنوان فاز نهایی آزمایش‌ها بر روی مجموعه داده‌های UCI) مورد بررسی قرار گرفت. برای این منظور، نمودار $\|\beta^J - \beta^{J-1}\|_2$ برای تمامی مجموعه داده‌های UCI محاسبه و در شکل ۳ نشان داده شده است. همانطور که در شکل ۳ مشاهده می‌شود، بردار وزن کرنل‌ها در هر دو روش پیشنهادی به سرعت همگرا می‌شوند. همچنین عملکرد الگوریتم‌های پیشنهادی در شرایط نویزی نیز مورد آزمایش قرار گرفت. در شکل ۴ نتایج به دست آمده از اعمال الگوریتم‌های مختلف، وقتی نمونه‌های آزمون آغشته به نویز گوسی

جدول ۴: نتایج به دست آمده در فاز سوم آزمایش‌ها

Method	Soybean	Heart	Ionosphere	Wine	Sonar	Iris
SRC	98.0000	77.7778	92.6256	89.8366	83.1667	93.7355
KSRC(Best)	100.0000	83.3333	89.8148	99.4444	89.3734	97.3333
KSRC(mean)	100.0000	78.5185	94.8148	96.6667	86.0401	97.3333
KSRC(KA)	100.0000	78.8889	93.9815	96.6667	87.4687	98.0000
KSRC(ACC)	94.0000	81.1111	93.7037	98.33333	86.9925	98.0000
KSRC(EV)	100.0000	80.7407	94.9523	97.7778	86.5639	98.0000
KSRC(EV-Label)	100.0000	82.5926	93.7037	98.6667	88.2341	98.0000



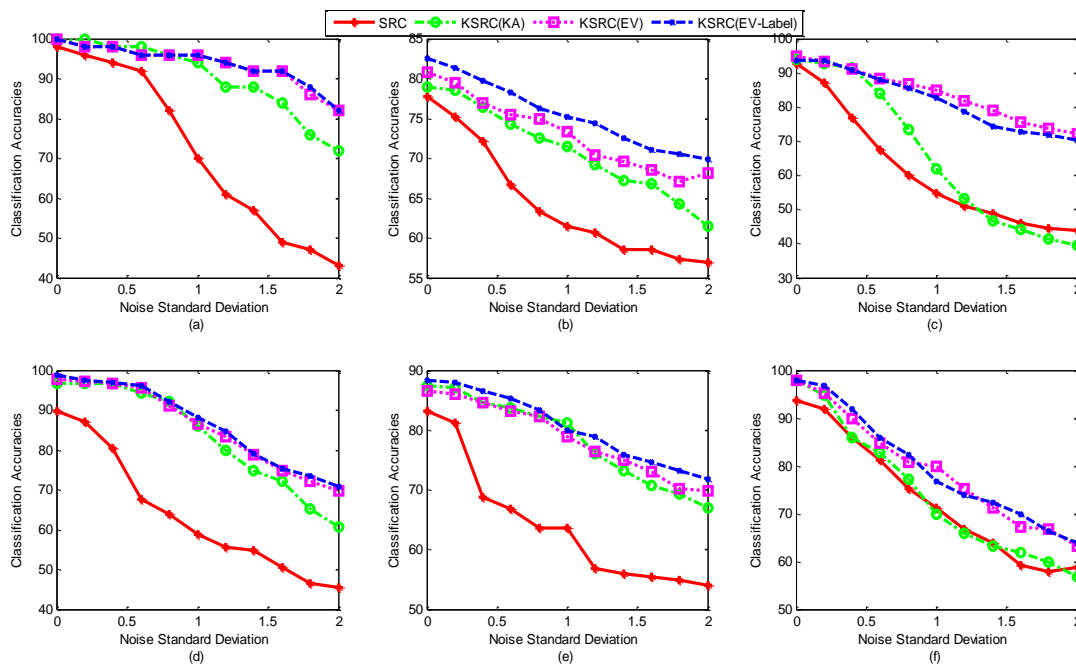
شکل ۳: تغییرات وزن کرنل‌ها در تکرارهای مختلف $(\|\beta^J - \beta^{J-1}\|_2)$ برای پایگاه داده UCI: (a) داده Soybean، (b) داده Heart، (c) داده Iris

(d) داده Wine، (e) داده Sonar، (f) داده Iris

۷-۶- فاز چهارم آزمایش‌ها

داده‌های آموزشی در نظر گرفته شد. در مواردی که به داده ارزیابی نیز احتیاج است، از هر کلاس به طور تصادفی ۱۰ نمونه به عنوان داده‌های ارزیابی در نظر گرفته شده است. همچنین از هر کلاس ۱۰۰ نمونه به طور تصادفی به عنوان مجموعه داده‌های آزمایشی در نظر گرفته شده است. آزمایش‌ها ۱۰ بار تکرار شده و متوسط نتایج ارائه می‌شود. همچنین در این فاز از آزمایش‌ها نیز، از کرنل‌ها و فیلتر معرفی شده در فاز سوم آزمایش‌ها استفاده شده است. نتایج به دست آمده بر روی این پایگاه داده نیز در جدول ۵ خلاصه شده است.

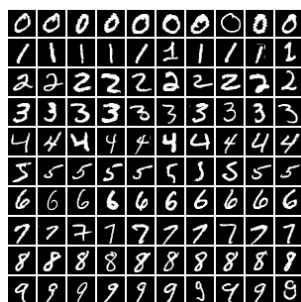
برای بررسی عملکرد روش‌های پیشنهادی بر روی مجموعه دادگان تصویری، آزمایش‌ها بر روی پایگاه داده اعداد دست نویس MNIST نیز تکرار شده است. این پایگاه داده شامل ۶۰۰۰۰ نمونه آموزشی و ۱۰۰۰۰ نمونه آزمایشی است. تصاویر در این پایگاه داده، از نوع تصاویر خاکستری نرمالیزه شده با ابعاد ۲۸×۲۸ است. در شکل ۵ مجموعه‌ای از اعداد دست نویس پایگاه داده MNIST نشان داده شده است. در انجام آزمایش‌ها، از هر کلاس ۵۰ نمونه و در مجموع ۵۰۰ نمونه از مجموعه داده‌های آموزشی به طور تصادفی انتخاب و به عنوان مجموعه



شکل ۴: مقایسه مقاومت روش‌های مختلف در مواجهه با نویز گوسی برای پایگاه داده UCI: (a) داده Soybean، (b) داده Heart، (c) داده Iris، (d) داده Wine، (e) داده Sonar، (f) داده Ionosphere

جدول ۵: نتایج به دست آمده بر روی پایگاه داده اعداد دست‌نویس MNIST

SRC	KSRC(Best)	KSRC(mean)	KSRC(KA)	KSRC(ACC)	KSRC(EV)	KSRC(EV-Label)
88.1000	90.0000	90.1000	90.2000	89.2000	90.0000	91.1000



شکل ۵: نمونه‌هایی از تصاویر پایگاه داده اعداد دست‌نویس MNIST

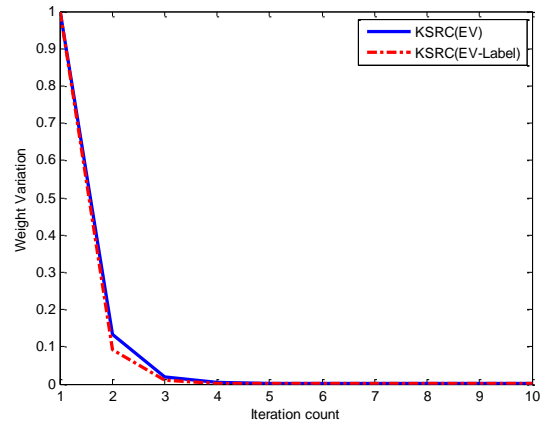
همانطور که از نتایج پیداست، استفاده از کرنل در الگوریتم‌های KSRC، سبب بهبود نتایج در مقایسه با الگوریتم SRC شده است. علاوه بر بهبود نتایج، استفاده از کرنل برای داده‌های تصویری همچون MNIST که بعد نمونه‌ها زیاد است، سبب کاهش بعد ویژه‌نامه به تعداد نمونه‌ها و در نتیجه تسریع عملکرد الگوریتم‌های مبتنی بر نمایش تنک می‌شود. همانطور که از نتایج مشخص است، الگوریتم پیشنهادی KSRC(EV-Label) نتایج بهتری از KSRC با بهترین کرنل و همچنین سایر الگوریتم‌های KSRC دارد.

حاصل از اعمال تابع کرنل به داده‌ها، به جای بردار ویژگی نمونه‌ها در ساخت واژه‌نامه، کاهش بعد واژه‌نامه به تعداد نمونه‌ها است. اما انتخاب نوع کرنل می‌تواند تأثیر بسزایی در نتایج نهایی طبقه‌بند KSRC داشته باشد. با توجه به مشکل انتخاب کرنل بهینه برای الگوریتم‌های مبتنی بر کرنل، در این مقاله استفاده از کرنل مرکب در چارچوب طبقه‌بند مبتنی بر نمایش تنک پیشنهاد گردید. برای این منظور از کرنل‌های مرکب ساده مانند میانگین‌گیری ساده تا جمع وزن‌دار کرنل‌ها مورد بررسی قرار گرفت. برای جمع وزن‌دار کرنل‌ها، فرآیند بهینه‌سازی وزن کرنل‌ها در چارچوب الگوریتم KSRC ارائه شد و روابط مربوط به بهینه‌سازی وزن‌ها استخراج گردید. آزمایش‌ها بر روی مجموعه‌های متنوعی از داده‌ها، شامل داده‌های پایگاه داده UCI و همچنین مجموعه اعداد دست‌نویس MNIST مورد بررسی قرار گرفت. وجود کرنل‌های نامناسب در میان مجموعه کرنل‌های پایه، وقتی از قید محدب برای آموزش وزن کرنل‌ها استفاده شود، می‌تواند سبب کاهش دقت طبقه‌بندی شود زیرا هیچ کرنلی حتی با وجود وزن کوچک حذف نشده و در کرنل مرکب تأثیر خواهد داشت. از این رو فیلتری برای حذف اولیه کرنل‌های نامناسب ارائه شد. کرنل‌های معتبر از دید این فیلتر، وارد ساختار کرنل مرکب و فرآیند آموزش و طبقه‌بندی می‌شوند. از نتایج به دست آمده بعد از اعمال این فیلتر، می‌توان عملکرد خوب طبقه‌بندی مبتنی بر نمایش تنک در مقایسه با SRC را مشاهده نمود. در میان مجموعه طبقه‌بند مبتنی بر نمایش تنک کرنلی با کرنل‌های متفاوت، الگوریتم‌های پیشنهادی و به خصوص الگوریتم KSRC(EV-Label) نتایج بهتری در مقایسه با سایر الگوریتم‌های KSRC داشته است چرا که در این الگوریتم علاوه بر اینکه وزن‌ها در چهارچوب مسئله نمایش تنک و در جهت کاهش خطای بازسازی بهینه می‌شوند، از برجسب کلاس نمونه‌ها نیز در فرآیند آموزش استفاده می‌شود. همچنین مقاومت الگوریتم‌های مختلف در مواجهه با نویز گوسی جمع‌شونده نیز مورد بررسی قرار گرفت. نتایج نشان داد که الگوریتم پیشنهادی KSRC(EV-Label) در مجموع دارای عملکرد بهتری در مقایسه با سایر روش‌ها در شرایط نویزی دارد. در آینده قصد داریم، علاوه بر آموزش وزن کرنل‌ها، آموزش واژه‌نامه مبتنی بر کرنل را نیز در نظر بگیریم که تعمیمی از الگوریتم پیشنهادی خواهد بود.

مراجع

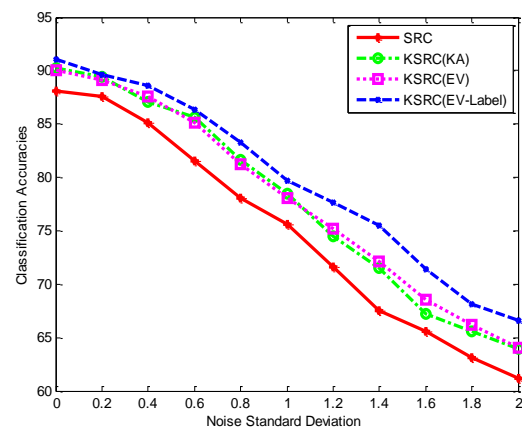
- [1] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [2] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer-Verlag, 2010.
- [3] V. M. Patel and R. Chellappa, "Sparse representations, compressive sensing and dictionaries for pattern recognition," *In Proceedings of Asian Conference on Pattern Recognition*, pp. 325-329, 2011.

برای بررسی همگرایی وزن کرنل‌ها در الگوریتم‌های پیشنهادی، نمودار $\|\beta^J - \beta^{J-1}\|_2$ برای پایگاه داده اعداد دست‌نویس MNIST در شکل ۶ رسم شده است. همانطور که در شکل ۶ دیده می‌شود، وزن کرنل‌ها بعد از تکرار چهارم همگرا شده است.



شکل ۶: تغییرات وزن کرنل‌ها در تکرارهای مختلف الگوریتم $\|\beta^J - \beta^{J-1}\|_2$ برای پایگاه داده اعداد دست‌نویس MNIST

مشابه با آزمایش‌های گزارش شده پیشین، میزان مقاومت الگوریتم‌های مختلف در شرایط نویزی نیز مورد بررسی قرار گرفت. در اینجا نیز، انحراف معیار نویز گوسی در محدوده صفر تا دو در نظر گرفته شد. شکل ۷ در برگیرنده نتایج به دست آمده است. همانطور که مشخص است، مجدداً روش پیشنهادی KSRC(EV-Label) عملکرد بهتری در مقایسه با سایر روش‌ها در شرایط نویزی دارد.



شکل ۷: مقایسه مقاومت روش‌های مختلف در مواجهه با نویز گوسی برای پایگاه داده اعداد دست‌نویس MNIST

۸- نتیجه‌گیری و پیشنهادات

در این مقاله ساختار طبقه‌بند مبتنی بر نمایش تنک و نوع کرنلی آن مورد بررسی قرار گرفت. یکی از مزایای مهم استفاده از بردار تشابه

- [13] G. R. G. Lanckriet, N. Cristianini, P. L. Bartlett, L. E. Ghaoui, and M. I. Jordan, "Learning the kernel matrix with semidefinite programming," *Journal of Machine Learning Research*, vol. 5, no. 1, pp. 27–72, 2004.
- [14] M. Gönen and E. Alpaydm, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, no. 1, pp. 2211–2268, 2011.
- [15] H. Zheng, F. Liu, and ZH. Jin, "Multiple Kernel Sparse Representation Based Classification". *Pattern Recognition*, vol. 321, no. 1, pp. 48-55, 2012.
- [16] A. Shrivastava, V.M. Patel, and R. Chellappa, "Multiple kernel learning for sparse representation-based classification", *IEEE Transactions on Image Processing* vol. 23, no. 7, pp. 3013-3024, 2014.
- [17] W. Liu, Z. Yu, Y. Wen, M. Yang, and Y. Zou, "Multi-kernel collaborative representation for image classification", *In Proceedings of International Conference on Image Processing*, 2015.
- [18] P. Pavlidis, J. Weston, J. Cai and W. N. Grundy, "Gene Functional Classification from Heterogeneous Data," *In Proceedings of the 5th Annual International Conference on Computational Molecular Biology*, 2001.
- [19] N. Cristianini, J. Shawe-Taylor, A. Elisseeff, and J. S. Kandola, "On kernel-target alignment," *In Advances in Neural Information Processing Systems*, 2001.
- [20] H. Tanabe, T. B. Ho, C. H. Nguyen, and S. Kawasaki, "Simple but Effective Methods for Combining Kernels in Computational Biology," *In Proceedings of IEEE International Conference on Research, Innovation and Vision for the Future*, 2008.
- [4] L. Zhang and M. Yang et al., "Sparse representation or collaborative representation: Which helps face recognition?," *In Proceedings of International Conference on Computer Vision*, 2011.
- [5] V. M. Patel, R. Chellappa, and M. Tistarelli, "Sparse representations and random projections for robust and cancelable biometrics," *In Proceedings of 11th International Conference of Control*, 2010.
- [6] S. Gao, I. W. Tsang, and L.-T. Chia, "Kernel sparse representation for image classification and face recognition," *In Proceedings of 11th European Conference on Computer Vision*, vol. 6314, no. 1, pp. 1-14, 2010.
- [7] H. Li, Y. Gao, and J. Sun, "Fast kernel sparse representation," *In Proceedings of International Conference of Digital Image Computing: Technology Application*, pp. 72–77, 2011.
- [8] L. Zhang et al., "Kernel sparse representation-based classifier," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1684–1695, 2012.
- [9] H. V. Nguyen, V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Kernel dictionary learning," *In Proceedings of IEEE ICASSP*, pp. 2021–2024, 2012.
- [10] J. Yin, X. Liu, Z. Jin, and W. Yang, "Kernel sparse representation based classification," *Neurocomputing*, vol. 77, no. 1, pp. 120-128, 2012.
- [11] S. Gao, I. W. Tsang, and L.-T. Chia, "Sparse representation with kernels," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 423–434, 2013.
- [12] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, Cambridge, UK, 2004.

زیر نویس‌ها

¹ Sparse Representation Classifier

² Collaborative Representation based Classifier

³ Cross validation

⁴ Quadratic Program

⁵ Kernel Alignment

⁶ Cosine distance measure

⁷ Frobenius product

⁸ University of California at Irvine