

Sparsification of the PLS Regression Algorithm using L2-Norm of Weighted Coefficients: Application in Emotion Recognition

Einizade, Aref¹ / Hajipour Sardouie, Sepide^{2*}

¹ - Ph.D. Student, Bioelectric Group, Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

² - Assistant Professor, Bioelectric Group, Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

DOI: 10.22041/IJBME.2020.124620.1589

Received: 13 April 2020

Revised: 3 August 2020

Accepted: 11 September 2020

KEYWORDS

EEG Signal
Emotion Recognition
Linear Regression
PLS Regression
Sparse Coefficients
Valence
Arousal

ABSTRACT

The brain electrical signal has been widely used in clinical and academic research, due to its ease of recording, non-invasiveness, and precision. One of the applications can be emotion recognition from the brain's electrical signal. Generally, two types of parameters (Valence and Arousal) are used to determine the type of emotion which in turn indicate "positive or negative" and "level of extroversion or excitement" for a specific emotion. The significance of emotion is determined by the effects of this phenomenon on daily tasks, especially in cases where the person is confronted with activities that require careful attention and concentration. In the emotion recognition problem, firstly, using proper emotion stimuli, different emotions are created for the subjects under study and the brain signals corresponding to each stimulus are recorded. The two main steps for solving the emotion recognition problem are extracting suitable features and using appropriate classification or regression methods. In previous studies, different visual and auditory have been used and various linear and nonlinear features and classifiers have been investigated. In this paper, the main goal was the improvement of linear regression algorithms to estimate the criteria for recognizing human emotions more efficiently. For this purpose we proposed a new algorithm that uses the sparseness of the mixing vector along with the linear regression cost function. The effectiveness of the proposed algorithm on simulated data has been investigated and its superiority to linear regression algorithms such as PLS, LASSO, SOPLS and Ridge was shown. Also, to apply the proposed algorithm on EEG data corresponding to emotion recognition, the DEAP dataset was used and the AR coefficients were extracted from the EEG signals. The results obtained from the proposed algorithm were compared with those of the other linear regression algorithms, which in total showed the relative superiority of the proposed method.

*Corresponding Author

Address	Bioelectric Group, Electrical Engineering Department, Sharif University of Technology, Tehran, Iran			
Postal Code	11155-4363	Tel	+98-21-66165928	
E-Mail	hajipour@sharif.edu		Fax	+98-21-66165928





تنک‌سازی رگرسیون PLS با استفاده از نرم-۲ ضرایب وزن‌دار: کاربرد در حل مسأله‌ی بازشناسی احساسات

عینی‌زاده، عارف^۱ / حاجی‌پور ساردوئی، سپیده^{۲*}

^۱ - دانشجوی دکتری مهندسی پزشکی، گروه بیوالکتریک، دانشکده‌ی مهندسی برق، دانشگاه صنعتی شریف، تهران، ایران
^۲ - استادیار، گروه بیوالکتریک، دانشکده‌ی مهندسی برق، دانشگاه صنعتی شریف، تهران، ایران

مشخصات مقاله

شناسه‌ی دیجیتال: 10.22041/IJBME.2020.124620.1589

پذیرش: ۲۱ شهریور ۱۳۹۹

بازنگری: ۱۳ مرداد ۱۳۹۹

ثبت در سامانه: ۲۵ فروردین ۱۳۹۹

واژه‌های کلیدی

سیگنال EEG
بازشناسی احساسات
رگرسیون خطی
رگرسیون PLS
ضرایب تنک
خوشایندی
انگیزختگی

چکیده

سیگنال الکتریکی مغز به دلیل سهولت در ثبت، غیرتهاجمی بودن و دقت بالا کاربرد بسیار گسترده‌ای در حوزه‌های کلینیکی و تحقیقات دانشگاهی پیدا کرده است که از آن جمله می‌توان به بازشناسی احساس از روی سیگنال الکتریکی مغز اشاره کرد. برای مشخص کردن نوع احساس برانگیخته عموماً از دو معیار خوشایندی و انگیزختگی استفاده شده که به ترتیب میزان مثبت یا منفی بودن و میزان برون‌فکنی یا هیجان ایجاد شده را برای یک احساس خاص مشخص می‌نمایند. اهمیت احساس با توجه به عکس‌العمل‌های ناشی از این پدیده در انجام وظایف روزانه به خصوص در فعالیت‌هایی که به دقت و تمرکز نیاز دارند مشخص می‌شود. در مسأله‌ی بازشناسی احساس ابتدا احساسات مختلف با استفاده از محرک‌های حسی مناسب برای سوژه‌های مورد بررسی ایجاد شده و سیگنال‌های مغزی متناظر با هر تحریک ثبت می‌شود. دو مرحله‌ی پردازش اصلی در حل مسأله‌ی بازشناسی احساسات، استخراج ویژگی‌های مناسب و استفاده از طبقه‌بندها یا رگرسیون‌های مناسب است. در پژوهش‌های پیشین محرک‌های مختلف بینایی و شنوایی مورد استفاده قرار گرفته و ویژگی‌ها و طبقه‌بندهای متنوع خطی و غیرخطی بررسی شده است. هدف این مقاله بهبود الگوریتم‌های رگرسیون خطی برای تخمین بهتر معیارهای بازشناسی احساسات انسانی است. بدین منظور یک الگوریتم جدید ارائه شده است که از تنک بودن بردار ترکیب در کنار تابع هزینه‌ی رگرسیون خطی استفاده می‌کند. کارایی الگوریتم ارائه شده روی داده‌های ساختگی مورد بررسی قرار گرفته و برتری آن نسبت به الگوریتم‌های رگرسیون خطی PLS، LASSO، SOPLS و Ridge نشان داده شده است. هم‌چنین برای اعمال الگوریتم ارائه شده روی داده‌های EEG مربوط به بازشناسی احساسات از مجموعه‌ی دادگان DEAP استفاده شده و ویژگی‌های ضرایب مدل AR از سیگنال‌های EEG استخراج شده است. نتایج به دست آمده از الگوریتم پیشنهادی با نتایج سایر الگوریتم‌ها مقایسه شده است که در مجموع برتری نسبی الگوریتم ارائه شده را نشان می‌دهد.

*نویسنده‌ی مسئول

نشانی: گروه بیوالکتریک، دانشکده‌ی مهندسی برق، دانشگاه صنعتی شریف، تهران، ایران

تلفن: ۹۸-۲۱-۶۶۱۶۵۹۲۸

۱۱۱۵۵-۴۳۶۳

دورنگار: ۹۸-۲۱-۶۶۱۶۵۹۲۸

hajipour@sharif.edu

پست الکترونیک



۱- مقدمه

در این پژوهش تغییرات در فعالیت مغز هنگام بروز احساسات مختلف مورد بررسی قرار گرفته است. در این زمینه تحقیقات مختلفی انجام شده که تقریباً تمام آن‌ها به موضوع طبقه‌بندی یا رگرسیون تبدیل می‌شود. در این راستا پردازش سیگنال‌های مغزی EEG با هدف تخمین پارامترهای مرتبط با تشخیص احساسات و بررسی الگوریتم‌های موثر در این مساله مورد بررسی و تحلیل قرار گرفته است. رویکردها و الگوریتم‌های مختلفی برای تخمین این پارامترها در مقالات مختلف معرفی شده که به صورت کلی به دو دسته‌ی خطی و غیرخطی قابل دسته‌بندی است. طراحی بهینه‌ی الگوریتم‌های خطی اغلب به دلیل پردازش سریع‌تر و پیاده‌سازی ساده‌تر می‌تواند به تنهایی و یا حتی به صورت بلوکی در ساختار الگوریتم‌های غیرخطی در افزایش کارایی سیستم تشخیص احساسات (افزایش دقت یا کاهش زمان پردازش) بسیار موثر باشد. در این تحقیق الگوریتم‌های خطی موجود و الگوریتم خطی پیشنهادی برای حل این مساله معرفی و بررسی شده و به مزایا و چالش‌های پیش روی این نوع الگوریتم‌ها پرداخته شده است.

نقش موثر احساسات در تصمیم‌گیری یکی از انگیزه‌های مهم دانشمندان برای تحقیق درباره‌ی مساله‌ی تشخیص احساسات است. بروز احساسات به صورت کلامی یا غیرکلامی مانند علائم فیزیولوژیکی می‌تواند نقطه‌ی شروع خوبی برای حل مساله از دریچه‌ی علوم مختلف باشد. مساله‌ی ارزیابی احساسات می‌تواند کاربردهای بسیار مهمی در بهبود کیفیت زندگی و افزایش توانایی انسان در برخورد با چالش‌های زندگی داشته باشد که از آن جمله می‌توان به کاربرد در رابط‌های مغز-رایانه^۱ (BCI)، دروغ‌سنجی، پیش‌گویی رفتار و نظارت بر حالت‌های حیاتی، سلامتی و بازتوانی و استفاده در ارزش‌یابی اشاره کرد [۱].

یکی از ساده‌ترین روش‌های دسترسی به خروجی‌های سیستم عصبی ثبت فعالیت‌های الکتریکی مغز با استفاده از الکترودهای روی پوست سر بوده که سیگنال‌های الکتروانسفالوگرام (EEG) نامیده می‌شوند. عمل کرد و حالت مغز را می‌توان با مطالعه روی منشا تولید این سیگنال و عوامل موثر بر تولید آن بررسی نمود.

۱-۱- معیارهای بازنمایی احساسات

در زمینه‌ی بازشناسی احساسات تلاش‌های زیادی به منظور رسیدن به اجماع کلی روی تعیین یک مدل جامع برای بازنمایی معیارهای احساسی و بیان تمایز میان احساسات مختلف انجام

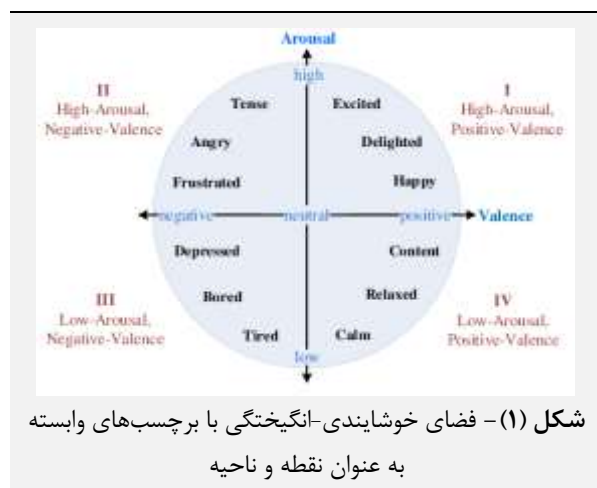
شده که در این راستا روی دو بعد مهم‌تر تمایز احساسات (خوشایندی^۲ و انگیزندگی^۳) توافق کلی صورت گرفته است [۱]. معیار خوشایندی معرف مثبت یا منفی بودن یک احساس در حالت کلی است. این معیار در واقع نشان دهنده‌ی میزان خوشایندی یک احساس بوده و قضاوت یک تحریک از ناخوشایندی تا رضایت بر اساس آن می‌باشد.

معیار انگیزندگی (برون‌فکنی، هیجان بیرونی) میزان هوشیاری و فعالیت یک احساس خاص را نشان می‌دهد. در واقع این معیار درجه‌ی تحریک شدن احساس به وسیله‌ی عوامل محیطی را در گستره‌ی آرامش تا برون‌فکنی (هیجان بیرونی) بیان می‌کند. علاوه بر موارد فوق معیارهای احساسی دیگری نیز تعریف شده که برای کامل شدن بحث در ادامه دو معیار تسلط^۴ و علاقه^۵ به اختصار معرفی شده است.

معیار تسلط نشان دهنده‌ی میزان تسلط و کنترل القا شده توسط یک حس خاص به سوژه است.

معیار علاقه نشان دهنده‌ی میزان پسندیده بودن یک احساس خاص پس از مشخص شدن مثبت یا منفی بودن آن است.

در تحقیقات مختلف معمولاً برای نمایش کلی فضای تمایز احساسات از فضای خوشایندی-انگیزندگی استفاده شده که این فضا با برچسب‌های وابسته به عنوان نقطه و ناحیه در شکل (۱) نشان داده شده است.



۱-۲- داده‌های مورد بررسی

در پژوهش‌های مختلف برای تحریک احساسات گوناگون در افراد مورد آزمایش از محرک‌های متنوع تصویری، صوتی و یا ویدئویی استفاده شده و سیگنال‌های مغزی متناظر با این تحریک‌ها ثبت شده است. در این پژوهش از بانک داده‌ی

^۱ Dominance

^۵ Liking

^۱ Brain Computer Interface (BCI)

^۲ Valence

^۳ Arousal



تخمینی بین X و Y حل مساله با فرض خطی بودن انجام می‌شود. در واقع در این نوع حل به جای یافتن نگاشت کلی f سعی می‌شود تا ماتریس B که رابطه‌ی خطی بین ورودی و خروجی ایجاد می‌کند به صورت زیر به دست آید.

$$X_{\text{trials} \times \text{features}} \times B_{\text{features} \times \text{measures}} \approx Y_{\text{trials} \times \text{measures}}$$

در مواردی که اطلاعات بیش‌تری در مورد رابطه‌ی بین داده‌ی ورودی X و خروجی Y وجود دارد (مانند تنک بودن^۲ ماتریس ترکیب B) می‌توان با بهره‌گیری از اثر اضافه کردن این اطلاعات تخمین بهتر و دقیق‌تری به دست آورد. هدف کلی این مقاله به کارگیری فرض تنک بودن ماتریس ترکیب برای بهبود روش‌های رگرسیون در حل مساله‌ی بازشناسی احساسات است. در ادامه در بخش ۲ مروری بر تحقیقات پیشین در زمینه‌ی بازشناسی احساسات انجام شده، در بخش ۳ ابتدا رگرسیون خطی ساده و PLS استاندارد و سپس الگوریتم پیشنهادی به تفصیل توضیح داده شده، در بخش ۴ نتایج اعمال الگوریتم پیشنهادی روی داده‌های ساختگی و واقعی ارائه و تحلیل شده و روند همگرایی الگوریتم پیشنهادی مورد بررسی قرار گرفته و در بخش ۵ مباحث بخش‌های پیشین مرور و جمع‌بندی شده و پیشنهادهایی نیز برای کارهای آینده بیان شده است.

۲- مروری بر پژوهش‌های پیشین

در پژوهش‌های حوزه‌ی بازشناسی احساس دو رویکرد کلی استفاده از یک بانک داده‌ی معتبر و یا استفاده از سیگنال‌های ثبت شده توسط خود محقق [۳] وجود دارد. بانک‌های داده‌ی معتبر و مرجعی در این زمینه فراهم شده است که از میان آن‌ها می‌توان به چهار مورد محبوب SEED، DREAMER [۶]، ASCRETAIN [۷] و DEAP اشاره کرد.

بانک داده‌ی SEED توسط مراجع مختلفی مورد استفاده قرار گرفته که در آن‌ها با استفاده از ویژگی‌های مختلفی مانند توان باندهای مختلف فرکانسی و استفاده از طبقه‌بندی‌های مختلف در حوزه‌ی یادگیری ماشین به پردازش سیگنال‌های EEG پرداخته شده است [۴، ۵]. بانک داده‌ی DREAMER علاوه بر سیگنال‌های EEG و سیگنال‌های ماهیچه‌ای و پوستی شامل سیگنال‌های قلبی (ECG) سوژه‌ها نیز می‌باشد. بانک داده‌ی ASCRETAIN از مجموعه‌های جدیدتر این حوزه بوده که در آن علاوه بر سیگنال‌های پرکاربرد این حوزه از تحلیل ویدئو نیز استفاده شده است. در ادامه به مقالاتی اشاره شده که در آن‌ها با استفاده از بانک داده‌ی DEAP (بانک داده‌ی مورد استفاده

DEAP [۲] استفاده شده که متشکل از سیگنال‌های EEG سطحی ۳۲ سوژه حین تماشای ۴۰ موزیک‌ویدئو است. این سیگنال‌ها از ۴۰ کانال (۳۲ کانال مربوط به الکترودهای روی پوست سر و ۸ کانال مربوط به الکترودهای چشمی و پوست صورت) با فرکانس نمونه‌برداری ۱۲۸ هرتز ثبت شده است. تمام این داده‌ها برای هر سوژه در یک تنسور به صورت $\text{data}_{40 \times 40 \times 8064}$ (۴۰ اول نشان دهنده‌ی تعداد ویدئوها و ۴۰ دوم نشان دهنده‌ی تعداد کانال‌ها) ذخیره شده است. هر موزیک‌ویدئو القاکننده‌ی یک حس خاص (غم، شادی و ...) بوده و مدت زمان هر کدام ۶۰ ثانیه است (عدد ۸۰۶۴ با توجه به زمان ۶۰ ثانیه به همراه ۳ ثانیه استراحت و فرکانس نمونه‌برداری ۱۲۸ هرتز به دست آمده است). هم‌چنین پس از تماشای هر ویدئو از سوژه‌ها خواسته شده تا به آن برای چهار معیار خوشایندی، انگیزتگی، تسلط و علاقه به صورت پیوسته از ۱ تا ۹ عددگذاری کنند که این مقادیر در ماتریس $\text{labels}_{40 \times 4}$ برای ۴۰ ویدئو و برای هر سوژه ذخیره شده است.

هدف هر الگوریتم طبقه‌بندی یا رگرسیون روی این بانک داده، تخمین چهار معیار ذکر شده (عددگذاری شده توسط خود سوژه‌ها) برای هر ویدئو با توجه به ویژگی‌های استخراج شده از سیگنال‌های EEG سطحی است. هم‌چنین معمولاً (از جمله در این پژوهش) فقط از کانال‌های ۱ تا ۳۲ (مربوط به الکترودهای روی پوست سر) برای پردازش استفاده شده است.

برای پیش‌پردازش سیگنال EEG نویز برق شهر و آرتیفکت‌های حرکتی با الگوریتم ICA حذف شده و فیلتر میان‌گذر باترورث ۴-۴۰ هرتز روی سیگنال‌های هر کانال اعمال شده است.

۱-۳- تعریف مساله‌ی ریاضی برای رگرسیون

اگر X ماتریس ویژگی‌های استخراج شده از داده‌های خام و Y ماتریس برچسب‌ها (مقادیر پیوسته‌ای که هدف مساله‌ی ریاضی تخمین آن‌ها است) باشد، در حالت کلی هدف حل این مساله پیدا کردن تابعی مانند f است که توانایی پردازش و تبدیل داده‌ی ورودی X برای تخمین هر چه بهتر المان‌های Y را داشته باشد. این توضیحات را می‌توان به صورت رابطه‌ی زیر بیان کرد.

$$\text{ورودی: } X_{\text{trials} \times \text{features}} = [\underline{x}_1, \dots, \underline{x}_N]$$

$$\text{خروجی: } Y_{\text{trials} \times \text{measures}} = [\underline{y}_1, \dots, \underline{y}_M]$$

$$\text{هدف: } f(X) \approx Y$$

هم‌چنین صورت ساده شده‌ی مساله‌ی بالا که رگرسیون خطی^۱ نام دارد به این صورت تعبیر شده که برای تابع f و رابطه‌ی

^۲ Sparsity

^۱ Linear Regression

به بازه‌های زمانی کوتاه و تبدیل آن‌ها به بردار ویژگی متناظر با هر آزمایش، توان ۵ زیرباند مهم EEG برای تمام کانال‌ها و پنجره‌های زمانی ۱ ثانیه‌ای استخراج شده (۶۰ بازه‌ی زمانی برای هر ویدئو)، فاصله‌ی هر بردار ویژگی متناظر با هر بازه‌ی زمانی تا بردارهای ویژگی آموزشی مربوط به دو کلاس ۰ و ۱ برای هر ویدئو با استفاده از کرنل گوسی محاسبه شده، پنج فاصله‌ی کم‌تر این بردارها با بردارهای مربوط به سایر ویدئوها (برای هر یک از برچسب‌های ۰ و ۱) تعیین شده، به وسیله‌ی آن‌ها مفهومی شبیه به احتمال تعریف شده، بردار ویژگی برای هر ویدئو با در نظر گرفتن بردار احتمالات مربوط به تمام بازه‌های زمانی هر آزمایش (هر ویدئو) و با استفاده از روش‌های Voting، هیستوگرام و استخراج ضرایب دیریکله استخراج شده و با به کارگیری SVM-RBF دقت $0.76/0.9$ و $0.68/0.4$ به ترتیب برای HV-LV و HA-LA حاصل شده است. در مقاله‌ی [۱۵] تبدیل Scalogram برای تمام سیگنال‌های کانال‌های مختلف (برای یک ویدئو) محاسبه شده، تنسوری با ابعاد زمان، فرکانس و کانال به دست آمده، بعد زمان با پنجره‌هایی به طول ۱ ثانیه به ۶۰ بازه تقسیم شده (تجزیه‌ی تنسور اولیه به ۶۰ تنسور)، مقادیر بعد زمان در هر یک از این تنسورها با هم جمع شده (حذف بعد زمان)، در نتیجه ۶۰ ماتریس با ابعاد تعداد پنجره‌های زمانی و تعداد کانال‌ها پدید آمده، این ۶۰ ماتریس یک تنسور را برای یک ویدئو تشکیل داده، در نتیجه برای ۴۰ ویدئو یک تنسور ۴ بعدی به عنوان تنسور ویژگی در نظر گرفته شده و پس از طبقه‌بندی تنسورها به کمک یادگیری عمیق (CNN) دقت‌های $0.72/0.6$ و $0.74/0.12$ به ترتیب برای تشخیص خوشایندی و انگیزختگی به دست آمده است. در تحقیق [۲] نوع استخراج ویژگی مانند [۱۵] بوده، از طبقه‌بند Naive Bayes برای طبقه‌بندی استفاده شده و دقت‌های $0.57/0.6$ و 0.62 به ترتیب برای HV-LV و HA-LA گزارش شده است.

۳- مواد و روش‌ها

در این بخش ابتدا رگرسیون‌های خطی ساده و استاندارد PLS معرفی شده و سپس الگوریتم PLS پیشنهادی ارائه شده است.

۳-۱- رگرسیون خطی ساده

مسئله‌ی رگرسیون خطی را می‌توان به صورت پیدا کردن ماتریس ترکیب $B \in \mathbb{R}^{p \times m}$ به طوری که کم‌ترین خطا بین ماتریس برچسب $Y \in \mathbb{R}^{n \times m}$ و تخمین آن $\hat{Y} = XB$ ، $X \in \mathbb{R}^{n \times p}$

در این مقاله) به بررسی مسأله‌ی بازشناسی احساس پرداخته شده است. منظور از طبقه‌بندی دوکلاسه با این بانک داده این است که با توجه به High (≥ 5) یا Low (< 5) بودن معیار خوشایندی/انگیزختگی، برچسب‌های ۰ و ۱ برای این معیارها در نظر گرفته می‌شود (HV-LV یا HA-LA).

در پژوهش [۸] با استخراج ضرایب AR مرتبه‌ی ۱۰ از سیگنال EEG، کاهش ویژگی با روش SFS^۱ و طبقه‌بندی با روش KNN، دقت طبقه‌بندی دوکلاسه‌ی معیارهای خوشایندی و انگیزختگی (HV-LV و HA-LA) به ترتیب برابر با $0.72/0.33$ و $0.74/0.2$ به دست آمده است. در مقاله‌ی [۹] آنتروپی Renyi به عنوان ویژگی از سیگنال‌های EEG استخراج شده، از روش LDA برای کاهش بعد ویژگی استفاده شده، طبقه‌بندی با طبقه‌بند RF^۲ به همراه LightGBM و Xgboost انجام شده و دقت $0.79/0.6$ و $0.77/0.19$ به ترتیب برای HV-LV و HA-LA گزارش شده است. در تحقیق [۱۰] بازه‌ی زمانی از سیگنال اولیه برای تعیین موثرتر برچسب‌ها با روشی مبتنی بر Mutual-Information تشخیص داده شده و حجم داده‌ی موثر بسیار کاهش یافته، مشخصه‌های فرکانسی HOS^۳ و HOC^۴ به عنوان ویژگی استخراج شده، از طبقه‌بندهای SVM و Naive Bayes برای طبقه‌بندی استفاده شده و دقت $0.89/0.61$ و $0.89/0.4$ به ترتیب برای HV-LV و HA-LA حاصل شده است. در پژوهش [۱۱] ویژگی‌های مبتنی بر PSD استخراج شده، برای کاهش ویژگی از روش PCA استفاده شده، طبقه‌بندهای GELM و SVM برای طبقه‌بندی به کار گرفته شده و دقت طبقه‌بندی برای حالت ۴-کلاسه (LVHA، HVLA، LVLA و HVLA) برابر با $0.69/0.67$ گزارش شده است. در مقاله‌ی [۱۲] ویژگی‌هایی نظیر میانگین، میانه، ماکسیمم، مینیمم، انحراف معیار، واریانس، Skewness و Kurtosis از سیگنال هر کانال استخراج شده، از شبکه‌ی عمیق کانولوشنی (CNN) استفاده شده، آموزش شبکه به صورت Leave-One-Out (روی تمام سوزدها به جز یک سوزده) انجام شده و دقت طبقه‌بندی دوکلاسه برابر با $0.81/0.406$ و $0.73/0.36$ به ترتیب برای HV-LV و HA-LA و دقت طبقه‌بندی سه‌کلاسه برابر با $0.66/0.79$ و $0.57/0.58$ به ترتیب برای خوشایندی و انگیزختگی به دست آمده است. در تحقیق [۱۳] از روش HMM برای طبقه‌بندی طیف وسیعی از ویژگی‌های مختلف استفاده شده و دقت طبقه‌بندی دوکلاسه‌ی HV-LV و HA-LA به ترتیب برابر با 0.73 و $0.75/0.63$ گزارش شده است. در پژوهش [۱۴] با هدف تعریف ویژگی‌های مربوط

^۱ Higher Order Spectra

^۲ Higher Order Crossing

^۱ Sequential Forward Selection

^۲ Random Forest



بنابراین می‌توان تخمین برچسب‌ها برای داده‌ی جدید X_{new} را با استفاده از ماتریس ترکیب B حاصل از داده‌های آموزشی به صورت زیر به دست آورد.

$$\hat{Y}_{new} \approx X_{new}B \quad (5)$$

اثبات تمام روابط در مراجع [۱۶، ۱۷] ارائه شده است. در ادامه برای جلوگیری از طولانی شدن بحث، شبه کد این نوع رگرسیون در الگوریتم (۱) قرار داده شده است [۱۷].

الگوریتم (۱) - شبه کد رگرسیون PLS

ورودی: $X \in \mathbb{R}^{n \times p}$, $Y \in \mathbb{R}^{n \times m}$

R: تعداد مولفه‌های رگرسیون PLS

خروجی: $T \in \mathbb{R}^{n \times R}$, $U \in \mathbb{R}^{n \times R}$

۱- برای $r = 1:R$

۱-۱ مقداردهی اولیه‌ی u به صورت تصادفی

$$w = X^T u \quad 2-1$$

$$t = Xw, t \leftarrow t / \|t\| \quad 3-1$$

$$c = Y^T t \quad 4-1$$

$$u = Yc, u \leftarrow u / \|u\| \quad 5-1$$

۶-۱ تکرار گام‌های ۱-۲ تا ۵-۱ تا همگرا شدن

۲-۱ تکرار t -امین ستون T و u -امین ستون U

۳- بازسازی ماتریس‌های X و Y

$$X \leftarrow X - tt^T X$$

$$Y \leftarrow Y - uu^T Y$$

پایان

استفاده از رگرسیون PLS در مواردی مانند انعطاف‌پذیری بیش‌تر^{۱۱} در هنگام وجود مقادیر نویزی، تعداد نمونه‌ی آموزشی کم، تعداد ویژگی زیاد و نیز تفسیر بهتر مولفه‌ها می‌تواند مفید واقع شود [۱۶]. اما در کنار این مزیت‌ها بار محاسباتی و درجه‌ی آزادی الگوریتم از چالش‌های PLS به شمار می‌رود.

۳-۳- الگوریتم پیشنهادی PLSS

۳-۳-۱- تاثیر نرم‌نرم-ضرایب روی جواب نهایی

هدف این بخش دستیابی به یک الگوریتم دقیق‌تر و دارای قابلیت تعمیم^{۱۲} بیش‌تر با استفاده از اطلاعاتی اضافه مانند تنک بودن ماتریس B است. ایده‌ی اولیه برای دستیابی به یک راه حل ساده‌تر و کم‌هزینه‌تر برای حل این‌گونه مسائل و به کار

وجود داشته باشد بیان کرد که m و p به ترتیب بیان‌گر تعداد آزمایش‌ها، تعداد ویژگی‌ها و تعداد معیار‌هایی است که باید تخمین زده شود. بیان ریاضی این توضیحات به صورت زیر بوده که در آن X^+ شبه معکوس یا معکوس از چپ^۳ ماتریس X است.

$$\min_B \|Y - XB\|_F^2 \rightarrow B = X^+ Y \quad (1)$$

$$X^+ = (X^T X)^{-1} X^T Y$$

۳-۲- رگرسیون PLS استاندارد

رگرسیون PLS ابتدا به منظور در نظر گرفتن همبستگی خطی^۴ بین score-های خروجی و ویژگی‌های ورودی توسعه داده شده و اهمیت و سودمندی آن توسط افرادی چون ولد^۵ و کوالسکی^۶ در تجزیه‌های شیمی نشان داده شده است. در سال‌های اخیر از رگرسیون PLS در کاربردهای مختلفی مانند تجزیه‌های شیمیایی^۷، ارزیابی حسگرها^۸، کنترل فرایندهای صنعتی و بررسی و نمایش عمل‌کرد مغز استفاده شده است [۱۶]. در واقع نکته‌ی اساسی در این نوع رگرسیون جست‌وجوی بردارهای مخفی^۹ است به طوری که پس از تجزیه‌ی ماتریس ورودی $X \in \mathbb{R}^{n \times p}$ و ماتریس خروجی $Y \in \mathbb{R}^{n \times m}$ بیش‌ترین کوواریانس بین مولفه‌های X و Y برقرار باشد. بیان ریاضی این توضیحات به صورت زیر است.

$$X = TP^T + E = \sum_{r=1}^R t_r p_r^T + E \quad (2)$$

$$T = [t_1, t_2, \dots, t_R] \in \mathbb{R}^{n \times R}, T^T T = I$$

$$Y = UQ^T + F = \sum_{r=1}^R u_r q_r^T + F \quad (3)$$

$$U = [u_1, u_2, \dots, u_R] \in \mathbb{R}^{n \times R}, U^T U = I$$

در این روابط ماتریس‌های T و U به ترتیب ماتریس ویژگی‌های مخفی متعامد یک‌ای^{۱۰} X و Y هستند. در واقع ستون‌های ماتریس‌های T و U در این نوع رگرسیون به صورتی محاسبه شده که ماکسیمم کوواریانس را با یکدیگر داشته باشند. خروجی‌های الگوریتم PLS همین ماتریس‌های T و U می‌باشد. در نهایت تخمین ماتریس ترکیب (B) از روی داده‌ها (X) و برچسب‌های آموزشی (Y) و ماتریس‌های T و U (خروجی‌های الگوریتم بالا) به صورت زیر به دست می‌آید.

$$B = X^T U (T^T X X^T U)^{-1} T^T Y \quad (4)$$

^۷ Chemometrics

^۸ Sensory Evaluation

^۹ Latent Vector

^{۱۰} Orthonormal Latent Variables

^{۱۱} Robustness

^{۱۲} Generalization

^۱ Trials

^۲ Measures

^۳ Left inverse/Moore-Penrose pseudoinverse

^۴ Collinearity

^۵ Wold

^۶ Kowalski

واقع میزان اهمیت ویژگی‌های استخراج شده‌ی مختلف روی رسیدن به برچسب داده‌ی دیده نشده را نشان می‌دهد. به عبارت دیگر کم بودن نرم-۰ این بردار نشان دهنده‌ی تعداد درایه‌های صفر زیاد در آن و اهمیت کم ویژگی‌های متناظر با درایه‌های صفر در رسیدن به یک جواب دقیق‌تر است. بنابراین با اضافه کردن یک عبارت در حل مساله‌ی بهینه‌سازی اولیه که منجر به کم کردن نرم-۰ بردار ترکیب در رگرسیون شود، علاوه بر پایین آوردن بار محاسباتی، قابلیت تعمیم برای پیش‌بینی برچسب داده‌ی دیده نشده نیز افزایش می‌یابد [۱۸].

نکته‌ی مهم این است که حل مساله‌ی رگرسیون با نرم-۰ بردار ترکیب (به صورت مرسوم) راه حل ساده‌ای ندارد. در نتیجه برای راحت‌تر شدن حل مساله باید به جای ترم مرسوم $\|b\|_0$ یک ترم معادل قرار داده شود. بدین منظور در این مقاله یک ترم نرم-۲ (رابطه‌ی ۸) قرار داده شده و مساله به صورت تکرارشونده حل شده که مقادیر λ_i در هر تکرار تغییر داده شده و در واقع در هر تکرار در مورد مقدار λ_i برای $i=1:p$ تصمیم‌گیری شده است. در نتیجه مساله‌ی بهینه‌سازی اولیه به صورت رابطه‌ی (۹) تبدیل شده است.

$$\sum_{i=1}^p \lambda_i |b_i|^2 \quad (۸)$$

$$b = \underset{b}{\operatorname{argmin}} \|y - Xb\|_F^2 + \sum_{i=1}^p \lambda_i \|b_i\|_2^2 \quad (۹)$$

با توجه به رابطه‌ی (۹) می‌توان بیان کرد که افزایش ضریب λ_i منجر به کاهش اندازه‌ی b_i و در نتیجه کاهش نرم-۰ بردار ترکیب b می‌شود که می‌توان از این نکته جهت کاهش نرم-۰ بردار ترکیب استفاده کرد. اما با توجه به مشخص نبودن درایه‌های صفر b باید از یک الگوریتم مناسب جهت تعیین درایه‌هایی از b که باید صفر شوند استفاده شود. الگوریتم پیشنهاد شده در این مقاله دو شرط را در یک ساختار تکرارشونده بررسی می‌کند. در واقع در هر تکرار از الگوریتم اگر مقادیر فعلی ضرایب متناظر با هر b_f به صورت λ_f باشد، برای هر $f \in \{1, \dots, p\}$ دو شرط زیر مورد بررسی قرار می‌گیرد.

۱- آیا افزایش λ_f (صفر در نظر گرفتن سایر ضرایب $\lambda_j=0; j \neq f$) باعث کاهش خطای میانگین مربعات^۱ در تخمین داده‌های اعتبارسنجی می‌شود؟

۲- با افزایش λ_f (با صفر در نظر گرفتن سایر ضرایب $\lambda_j=0; j \neq f$) و حل مساله‌ی PLS آیا درایه‌ی متناظر در بردار ترکیب (b_f) از یک آستانه‌ی مشخص (Thr) کوچک‌تر است؟

نبردن روش‌های پیچیده‌ی بهینه‌سازی، از مرجع [۱۸] الهام گرفته شده است. در مرجع [۱۸] برای رسیدن به یک پاسخ هموار (تقریب ماتریس Y به صورت یک ماتریس هموار) و نزدیک شدن تخمین ماتریس Y به تقریب هموار آن، به جای حل مستقیم مساله‌ی بهینه‌سازی به صورت زیر عمل شده است.

$$B = \underset{B}{\operatorname{argmin}} \|Y - BX\|_F^2 + \lambda \|BX - B\tilde{X}\|_F^2 \quad (۶)$$

$\tilde{X} = X$ تقریب هموار ماتریس

در این رابطه عبارت $\|Y - BX\|_F^2$ مربوط به حل مساله‌ی رگرسیون اصلی بوده، عبارت $\|BX - B\tilde{X}\|_F^2$ برای اضافه کردن شرط همواری به پاسخ Y قرار داده شده و از ضریب λ برای تعیین میزان اهمیت این دو عبارت استفاده شده است. در مرجع [۱۸] به جای حل مساله‌ی بهینه‌سازی، با توجه به این که هر دو ترم مورد استفاده در تابع هزینه به صورت نرم-۲ است، مساله به صورت زیر به یک مساله‌ی رگرسیون خطی ساده‌ی اصلاح شده تبدیل شده است.

$$B = \underset{B}{\operatorname{argmin}} \|\tilde{Y} - B\tilde{X}\|_F^2 \quad (۷)$$

$$\tilde{X} = \begin{bmatrix} X \\ \sqrt{\lambda}(X - \tilde{X}) \end{bmatrix}; \tilde{Y} = \begin{bmatrix} Y \\ 0 \end{bmatrix}$$

در این رابطه تعداد ستون ماتریس‌های \tilde{X} و X یک‌سان و برابر با تعداد ویژگی‌های مورد استفاده بوده و به همین ترتیب تعداد ستون دو ماتریس \tilde{Y} و Y نیز یک‌سان است. در نتیجه با این تغییر متغیر به جای حل مساله‌ی رگرسیون با نرم پنالتی (رابطه‌ی ۶)، مساله به صورت تابع هزینه‌ی نرم-۲ (رابطه‌ی ۷) حل می‌شود.

در این مقاله نرم-۰ ستون‌های ماتریس ترکیب به عنوان یک ترم پنالتی به مساله‌ی رگرسیون خطی اضافه شده است. همچنین با الهام از ایده‌ی مرجع [۱۸]، به جای حل پیچیده‌ی مساله‌ی کمینه کردن نرم-۰، مساله به حل یک سری مسائل نرم-۲ کمینه‌ی تکرارشونده تبدیل شده است.

در این بخش برای بیان ساده‌تر الگوریتم پیشنهادی ماتریس برچسب‌ها فقط با یک ستون در نظر گرفته شده (بردار برچسب $y \in \mathbb{R}^n$ به جای ماتریس برچسب $Y \in \mathbb{R}^{n \times m}$) و در نتیجه هدف این بخش تخمین بردار ترکیب $b \in \mathbb{R}^p$ است.

نرم-۰ یک بردار در واقع تعداد درایه‌های غیرصفر آن را نشان می‌دهد. به همین صورت با در نظر گرفتن بردار ترکیب b در مساله‌ی رگرسیون خطی ساده می‌توان بیان کرد که بردار تنک به دست آمده از حل مساله‌ی رگرسیون با نرم پنالتی نرم-۰ در

^۱ Mean Square Error (MSE)

اما مهم‌ترین چالش روش حل به‌صرفه‌تر مسأله‌ی بهینه‌سازی مطرح شده در رابطه‌ی (۹) می‌باشد. برای حل این چالش می‌توان مسأله را به صورت ساده‌شده‌ی زیر نوشت.

$$\mathbf{b} = \underset{\mathbf{b}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{X}\mathbf{b}\|_F^2 + \|\mathbf{A}\mathbf{b}\|_F^2 \quad (10)$$

$A_{p \times p}$: diagonal & $a_{ii} = \sqrt{\lambda_i}$

می‌توان مسأله‌ی بالا را به یک مسأله‌ی رگرسیون خطی ساده به صورت زیر تبدیل کرد.

$$\mathbf{b} = \underset{\mathbf{b}}{\operatorname{argmin}} \|\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\mathbf{b}\|_F^2; \tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}; \tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} \quad (11)$$

لازم به ذکر است که ماتریس 0 در رابطه‌ی بالا دارای p سطر و ستون‌هایی برابر با تعداد ستون‌های y بوده (در این حالت 1 ستون) و ماتریس A در رابطه‌ی (10) تعریف شده است. حال برای حل مسأله‌ی رگرسیون (رابطه‌ی 11) با استفاده از رگرسیون PLS می‌توان مشابه الگوریتم (1) عمل کرد. در واقع با در نظر گرفتن $\tilde{\mathbf{X}}$ و $\tilde{\mathbf{y}}$ در رابطه‌ی (11) و اعمال الگوریتم (1) برای این دو ماتریس می‌توان به جای رگرسیون خطی ساده از PLS استفاده کرده و از مزایای آن بهره برد. شبه کد الگوریتم پیشنهادی PLSS در الگوریتم (2) نشان داده شده است.

الگوریتم (2) - شبه کد الگوریتم PLSS

ورودی:

$$\lambda_{\text{increase}}, \text{MaxIterations}, \text{Thr}_{\text{cte}}, \mathbf{y} \in \mathbb{R}^n, \mathbf{X} \in \mathbb{R}^{n \times p}$$

خروجی: $\mathbf{b}_{\text{Result}} \in \mathbb{R}^p$

۱- مقداردهی اولیه:

$$\lambda_i = 0; \forall i = 1:p, \text{Thr} = \text{Thr}_{\text{cte}}, \text{MSE}_{\text{min}} = \text{Inf}$$

۲- برای $\text{iteration} = 1:\text{MaxIterations}$

۱-۲- برای $f = 1:p$

$$\lambda_f^{\text{temporary}} \leftarrow \lambda_f + \lambda_{\text{increase}}; \quad 1-1-2$$

۲-۱-۲- محاسبه‌ی $\text{MSE}_{\text{validation}}$ و بردار ضرایب b با

رگرسیون PLS:

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}; [A]_{i,j} = \begin{cases} 0; & i \neq f \text{ \& } j \neq f \\ \lambda_f^{\text{temporary}}; & i = j = f \end{cases}$$

۲-۱-۳- اگر $\text{MSE}_{\text{validation}} < \text{MSE}_{\text{min}}$ & $|\mathbf{b}_f| < \text{Thr}$

$$\text{MSE}_{\text{min}} \leftarrow \text{MSE}_{\text{validation}} \quad \lambda_f \leftarrow \lambda_f^{\text{temporary}}$$

پایان شرط

پایان برای f

$$\text{Thr} \leftarrow \text{Thr} / \text{Thr}_{\text{cte}} \quad 2-2$$

پایان برای Iteration

۳- محاسبه‌ی $\mathbf{b}_{\text{Result}}$ با رگرسیون PLS

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}; [A]_{i,j} = \begin{cases} 0; & i \neq j \\ \lambda_i; & i = j \end{cases}$$

در صورتی که دو شرط بالا برقرار باشد λ_f افزایش داده می‌شود. در مورد به کارگیری این دو شرط در حل مسأله‌ی رگرسیون خطی باید 6 نکته‌ی زیر در نظر گرفته شود.

۱- اگر مقدار نویز موجود در داده‌ی اولیه قابل توجه باشد شرط تصمیم‌گیری طبق MSE داده‌ی اعتبارسنجی به تنهایی قابل اطمینان نیست. به همین دلیل وجود یک شرط تصمیم‌گیری معتبر دیگر (شرط دوم) در کنار شرط MSE داده‌ی اعتبارسنجی اجتناب‌ناپذیر به نظر می‌رسد.

۲- درایه‌هایی از بردار b که در واقع صفر هستند در حل مسأله‌ی رگرسیون نرم-۲ احتمالاً مقادیر کوچکی دارند. بنابراین در صورتی که قدرمطلق مقدار فعلی درایه‌ی b_f از یک آستانه‌ی مشخص کم‌تر باشد می‌توان آن را نامزد صفر بودن در نظر گرفت (در صورتی که شرط اول نیز برقرار باشد) و ضریب متناظر با آن (λ_f) را افزایش داد. این افزایش ضریب می‌تواند موجب کوچک‌تر شدن قدرمطلق درایه‌ی b_f در تکرارهای بعدی شود.

۳- در بررسی شرط دوم از یک آستانه‌ی متغیر که مقدار آن در زمان کاهش می‌یابد استفاده می‌شود. با توجه به این واقعیت که سرعت نسبی کاهش قدرمطلق b_f -هایی که واقعا در تخمین دقیق‌تر کم‌اهمیت‌تر هستند (درایه‌های صفر در بردار b) نسبت به سایر b_f -ها بیشتر است، مقادیر قدرمطلق b_f -هایی که در تکرارهای قبل ضریب λ_f آن‌ها به اشتباه بزرگ شده است دیگر از آستانه‌ی جدید کوچک‌تر نبوده و از کاهش اشتباه قدرمطلق آن‌ها جلوگیری می‌شود. در واقع این آستانه (و مقادیر کاهش یافته‌ی آن در تکرارهای بعد) به عنوان یک صافی در کنار شرط MSE در صورت ناکارآمدی شرط اول در صورت وجود نویز زیاد، عمل کرده و خطاهای شرط اول را کاهش می‌دهد. مقدار آستانه و مقدار کاهش آن در زمان (در تکرارهای الگوریتم) به داده وابسته بوده و با ارزیابی متقاطع تعیین می‌شود.

۴- در حل مسأله‌ی رگرسیون ابتدا داده‌های آموزشی اولیه به دو بخش آموزشی و اعتبارسنجی تقسیم شده و سپس در هر تکرار الگوریتم از بخش آموزشی برای حل مسأله‌ی بهینه‌سازی و از بخش اعتبارسنجی برای تصمیم‌گیری در مورد تغییر ضرایب λ_f استفاده می‌شود.

۵- تعداد تکرارهای اجرای الگوریتم وابسته به داده است. در هر تکرار با بررسی دو شرط ارائه شده در مورد افزایش یا عدم تغییر هر یک از ضرایب λ_f به صورت جداگانه تصمیم گرفته می‌شود.

۶- پارامترهای الگوریتم نظیر تعداد تکرار (MaxIterations)، روند کاهشی آستانه (Thr_{cte}) و روند افزایشی λ_f -ها ($\lambda_{\text{increase}}$) را می‌توان به صورت بهینه با cross-validation تعیین کرد.

ساخته شده و با ضرب آن در ماتریس ورودی به صورت رابطه‌ی (۱۲) برچسب‌های خروجی بدون نویز به دست می‌آید.

$$y_{\text{Original}}^{\text{trials} \times 1} = X_{\text{trials} \times \text{features}} \times b_{\text{features} \times 1} \quad (12)$$

حال برای مدل‌سازی و شبیه‌سازی واقعی‌تر، مقادیر نویز سفید گاوسی به برچسب‌های خروجی اضافه می‌شود (رابطه‌ی ۱۳).

$$y_{\text{Noisy}} = y_{\text{Original}} + \alpha \times \text{noise} \quad (13)$$

با تغییر مناسب α می‌توان SNR-های مختلفی از مساله به دست آورد. در ادامه نتایج حاصل برای SNR-های ۵-، ۵، ۱۰ و ۱۵ دسی‌بل ارائه شده است. با توجه به امکان تغییر تعداد ویژگی‌ها در ماتریس ورودی، دو حالت مختلف متناظر با دو میزان مختلف تنک بودن در بردار ضرایب داده‌های ورودی بررسی شده است. برای ارزیابی و مقایسه‌ی الگوریتم‌ها از خطای میانگین مربعات (MSE) تخمین ضرایب رگرسیون و برای ارائه‌ی نتایج از پنج روش PLS، PLSS (الگوریتم پیشنهادی)، LASSO، Sparse Orthonormalized PLS (SOPLS) [۲۰] و Ridge Regression (L2-Norm) استفاده شده است.

الگوریتم LASSO از معروف‌ترین الگوریتم‌های حل مساله‌ی نرم-۱ بوده و تابع هزینه‌ی آن به صورت زیر است که در آن t مقدار مطلوبی برای پذیرش نرم-۱ جواب (β) می‌باشد. جزئیات بیشتر در مورد این الگوریتم در مقاله‌ی [۱۹] ارائه شده است.

$$\min_{\beta \in R^p} \left\{ \frac{1}{N} \|y - X\beta\|^2 \right\} \text{ subject to } \|\beta\|_1 \leq t \quad (14)$$

الگوریتم SOPLS [۲۰] یکی از روش‌های معروف برای استفاده از اطلاعات نرم-۱ یا نرم-۲ ضرایب در حل مساله‌ی PLS است. میانگین MSE تخمین ضرایب رگرسیون (b) در مقادیر ۸/۱۵ و ۵/۱۵ برای Sparsity برای داده‌ی $X_{300 \times 15}$ به همراه پارامتر بهینه (Thr_{cte}) در چهار SNR مختلف در جداول (۱) و (۲) ارائه شده که با توجه به آن می‌توان نکات زیر را نتیجه‌گیری نمود.

۱- در حضور نویز خیلی زیاد (SNR حدود ۵- دسی‌بل) الگوریتم پیشنهادی PLSS در بهبود تخمین ضرایب بهتر عمل کرده است.

۲- در حضور نویزهای کمتر (SNR در حدود ۵ تا ۱۵ دسی‌بل) الگوریتم پیشنهادی PLSS عمل کرد قابل قبولی نسبت به سایر الگوریتم‌ها داشته است.

۳- با افزایش میزان تنک بودن بردار ترکیب دقت تخمین افزایش یافته است.

همان‌گونه که پیش‌تر بیان شد این الگوریتم دارای سه پارامتر Thr_{cte} ، MaxIterations و $\lambda_{\text{increase}}$ است. پارامتر MaxIterations بیان‌گر تعداد تکرار حلقه‌ی اصلی الگوریتم بوده و از دو پارامتر دیگر برای ایجاد تغییر در مقادیر λ_f و بررسی اجزای این تغییر در هر گام از الگوریتم استفاده می‌شود. پارامتر $\lambda_{\text{increase}}$ مقدار ثابتی داشته که در هر گام از الگوریتم به مقدار λ_f ($f=1, \dots, p$) افزوده می‌شود. افزایش مقادیر λ_f تابع هزینه را در جهت کوچک‌تر شدن مقدار ضریب متناظر در بردار ضرایب b (b_f) سوق می‌دهد. البته در هر گام برای هر f باید کم‌تر بودن خطای اعتبارسنجی حاصل از این تغییر از یک آستانه‌ی ثابت MSE_{min} مورد بررسی قرار گیرد. قدرمطلق مقدار ضریب b_f نیز با یک آستانه‌ی متغیر Thr مقایسه می‌شود. اعمال تغییر در λ_f منوط به برقراری هر دو شرط است. برای سخت‌گیرانه‌تر شدن شرط مربوط به ضرایب b_f در هر گام از پارامتر Thr_{cte} برای کوچک کردن مقدار آستانه‌ی Thr استفاده می‌شود. به طوری که در هر تکرار آستانه‌ی Thr بر مقدار ثابت Thr_{cte} تقسیم می‌شود.

۴- یافته‌ها و بحث

در این بخش به بررسی و تحلیل نتایج الگوریتم پیشنهادی روی داده‌های ساختگی و واقعی و بیان تفاوت‌های الگوریتم‌های مختلف در کاربردهای مختلف پرداخته شده است.

۴-۱- نتایج الگوریتم پیشنهادی روی داده‌های

ساختگی

در این بخش نتایج برای نسبت سیگنال به نویزهای (SNR^1) مختلف روی داده‌ی مورد نظر و با پنج الگوریتم مختلف شامل الگوریتم پیشنهادی (PLSS) بیان شده است. در ساخت داده‌ها، نحوه‌ی تقسیم آن‌ها به داده‌های آموزشی، اعتبارسنجی و آزمون و نحوه‌ی به کارگیری الگوریتم‌ها شرایط زیر اعمال شده است.

۱- دوسوم داده‌ها به عنوان داده‌ی آموزشی، یک‌سوم آن‌ها به عنوان داده‌ی آزمون و یک‌پنجم داده‌های آموزشی نیز به عنوان داده‌ی اعتبارسنجی (در صورت نیاز برای تعیین پارامتری در یک الگوریتم خاص) در نظر گرفته شده است.

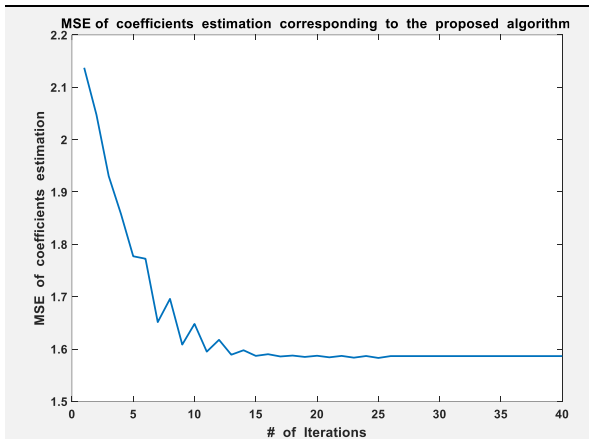
۲- میزان MSE گزارش شده، میانگین MSE برای ۵ مرتبه تولید داده‌ی تصادفی مختلف است (برای اطمینان از نتایج).

۳- منظور از $\text{Sparsity} = m/n$ این است که تعداد m درایه از n درایه‌ی بردار ترکیب (b) برابر با صفر می‌باشد.

ابتدا ماتریس ورودی $X_{\text{trials} \times \text{features}}$ با استفاده از توزیع گاوسی ساخته شده، سپس بردار رگرسیون $b_{\text{features} \times 1}$ از توزیع گاوسی

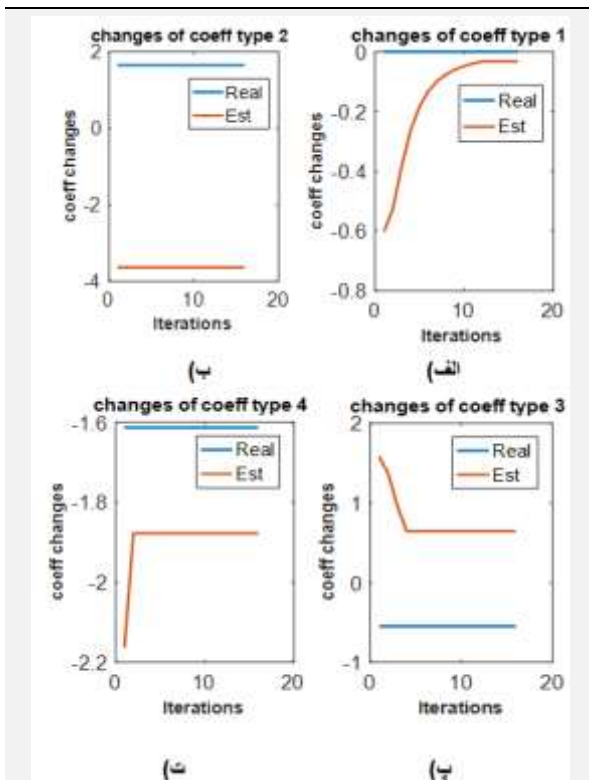
¹ Signal to Noise Ratio

تخمین را بهبود داده است. همچنین از تکرار حدود ۱۵ به بعد همگرایی قابل قبولی اتفاق افتاده است.



شکل (۲) - بررسی روند همگرایی الگوریتم پیشنهادی با ملاحظه‌ی روند تغییر MSE تخمین ضرایب رگرسیون

روند تغییر و اصلاح ضرایب رگرسیون در تکرارهای مختلف نیز مورد بررسی قرار گرفته است. با توجه به بررسی حالات مختلفی از شبیه‌سازی‌ها می‌توان چهار الگوی مختلف برای این تغییر و اصلاح برای ضرایب الگوریتم پیشنهادی در نظر گرفت (شکل ۳).



شکل (۳) - چهار الگوی تغییر یا اصلاح ضرایب مشاهده شده در تکرارهای مختلف توسط الگوریتم پیشنهادی، الف) ضرایب واقعا تنک، تشخیص و اصلاح به میزان قابل قبول، ب) ضرایب غیرتنک، میزان اصلاح بسیار کم، پ) ضرایب غیرتنک، میزان اصلاح نسبتا کم، ت) ضرایب غیرتنک، میزان اصلاح قابل قبول

جدول (۱) - MSE پنج الگوریتم برای تخمین بردار ترکیب (b)

داده‌ی ساختگی $X_{300 \times 15}$ با $(\text{Thr}_{\text{cte}}=2)$ و $\text{Sparsity}=5/15$

روش	dB -5	dB 5	dB 10	dB 15
PLS	۱/۸۷	۰/۵۹	۰/۳۳	۰/۱۸
PLSS	۱/۷۲	۰/۵۵	۰/۳	۰/۱۸
LASSO	۱/۸۷	۰/۶۲	۰/۳۴	۰/۱۹
SOPLS	۱/۸۲	۰/۵۶	۰/۳۲	۰/۱۸
L2-Norm	۱/۷۸	۰/۵۷	۰/۳۲	۰/۱۸

جدول (۲) - MSE پنج الگوریتم برای تخمین بردار ترکیب (b)

داده‌ی ساختگی $X_{300 \times 15}$ با $(\text{Thr}_{\text{cte}}=4)$ و $\text{Sparsity}=15/8$

روش	dB -5	dB 5	dB 10	dB 15
PLS	۱/۶۴	۰/۵۲	۰/۲۹	۰/۱۶
PLSS	۱/۶۱	۰/۴۹	۰/۲۶	۰/۱۴
LASSO	۱/۸۴	۰/۵۱	۰/۳	۰/۱۶
SOPLS	۱/۶۴	۰/۵۱	۰/۲۹	۰/۱۵
L2-Norm	۱/۶۳	۰/۵۲	۰/۲۹	۰/۱۶

۲-۴ - بررسی همگرایی الگوریتم پیشنهادی

همان‌گونه که در بخش‌های پیش توضیح داده شد تابع هدف الگوریتم به صورت رابطه‌ی (۹) است و می‌توان آن را به صورت زیر بازنویسی کرد.

$$L(\mathbf{b}) = \|\mathbf{y} - \mathbf{X}\mathbf{b}\|_F^2 + \sum_{i=1}^p \lambda_i \|\mathbf{b}_i\|_2^2 \quad (15)$$

$$= \|\mathbf{y} - \mathbf{X}\mathbf{b}\|_F^2 + \|\sqrt{\Lambda}\mathbf{b}\|_2^2$$

$$\Lambda_{ii} = \lambda_i \geq 0$$

با فرض ثابت بودن مقادیر λ_i مشتق دوم L نسبت به \mathbf{b} برابر با $(\mathbf{X}^T\mathbf{X} + \sqrt{\Lambda})$ بوده که با توجه به نامنفی بودن مقادیر روی قطر اصلی ماتریس $\sqrt{\Lambda}$ ، جمع دو ماتریس مثبت معین برابر با یک ماتریس مثبت معین جدید شده و بیانگر تحدب تابع هدف است. از آنجا که الگوریتم پیشنهادی یک الگوریتم ابتکاری با گام‌های تکراری بوده و مقادیر λ_i در تکرارهای مختلف تغییر می‌کنند نمی‌توان در مورد همگرایی آن حکم کلی داد اما می‌توان رفتار آن را در گام‌های مختلف الگوریتم بررسی کرد.

برای تحلیل همگرایی ابتدا تعداد تکرارهای الگوریتم پیشنهادی برابر با ۴۰ در نظر گرفته شده ($\text{MaxIterations}=40$)، سپس الگوریتم پیشنهادی ۵۰ مرتبه روی داده‌های تصادفی جدید اجرا شده و بردار MSE بین تخمین ضرایب و ضرایب اصلی ذخیره شده است. بردار میانگین این بردارهای MSE در هر تکرار محاسبه و در شکل (۲) ارائه شده است. مشاهده می‌شود که الگوریتم نسبت به مقدار اولیه‌ی بردار ضرایب رگرسیون نتیجه‌ی

هم‌چنین پارامتر Thr_{cte} برای پیش‌بینی Arousal، Valence، Dominance و Liking به ترتیب برابر با ۷، ۶، ۵ و ۵ قرار داده شده است. تمام این پارامترها بر اساس Cross Validation روی داده‌های آموزشی به دست آمده است.

جدول (۳) - میانگین دقت (%) تخمین چهار معیار تمایز احساسات روی ۳۲ سوژه برای پنج الگوریتم به همراه پارامتر بهینه برای الگوریتم PLSS

روش	Valence	Arousal	Dominance	Liking
PLS	۶۱/۵۶	۶۱/۷۲	۶۲/۱۹	۶۴/۱۴
PLSS	۶۱/۹۵	۶۱/۸	۶۳/۵۹	۶۴/۶۱
LASSO	۵۹/۰۶	۶۰/۳۹	۶۲/۳۴	۶۴/۴۵
SOPLS	۶۱/۸	۶۱/۴۱	۶۳/۷۵	۶۴/۳۸
L2-Norm	۵۸/۹۸	۵۷/۳۴	۶۱/۰۹	۶۰/۲۳

با توجه به بیش‌تر بودن تعداد ویژگی‌ها نسبت به نمونه‌ها و مدیریت خوب PLS و مشتقات آن در این‌گونه موارد، طبق انتظار و بر اساس نتایج به دست آمده الگوریتم‌های مبتنی بر PLS به طور کلی عمل کرد بهتری نسبت به الگوریتم LASSO داشته است. هم‌چنین الگوریتم پیشنهادی PLSS در مجموع بهترین عمل کرد را نسبت به سایر الگوریتم‌ها از خود نشان داده است. تمام مقادیر بهینه‌ی پارامترهای آزاد برای تمام الگوریتم‌ها در این مقاله با Cross Validation روی داده‌های آموزشی به دست آمده است. به عبارت دقیق‌تر به منظور تعیین تعداد مولفه‌های PLS به این صورت عمل شده که پس از اجرای Cross Validation روی داده‌ی آموزشی میانگین عمل‌کرد الگوریتم PLS پایه (و نه پیشنهادی) برای هر کدام از مولفه‌ها به دست آمده و مولفه‌ی با مقدار کم‌تر که تغییر عمل‌کرد PLS از آن مولفه به بعد قابل ملاحظه نباشد به عنوان تعداد مولفه‌ی بهینه انتخاب شده است. لازم به ذکر است که تعداد مولفه‌ها با استفاده از PLS پایه به دست آمده، برای الگوریتم پیشنهادی ثابت در نظر گرفته شده و در طول تکرارها تغییری نداشته است. در انتها برای بررسی قابلیت اعتماد به نتایج به دست آمده از تکرارهای مختلف روی داده‌های ساختگی و واقعی، p-value حاصل از تست آماری One-Way ANOVA به ترتیب در جدول‌های (۴) و (۵) ارائه شده است. این تست آماری روی MSE حاصل از تکرارهای مختلف الگوریتم پیشنهادی و الگوریتم پایه‌ی PLS اجرا شده است. مشاهده می‌شود که مقادیر p-value در تمام موارد مربوط به داده‌های ساختگی و واقعی از $\alpha = 0/05$ کوچک‌تر است.

الگوی الف: این رفتار برای ضرایبی که در نسخه‌ی اصلی ضرایب واقعا تنک بوده‌اند مشاهده شده است. در واقع الگوریتم پیشنهادی با دقت قابل قبولی این نوع ضرایب را تشخیص داده و سعی در تنک‌سازی آن‌ها داشته است. الگوی ب: برخی از ضرایبی که در واقع تنک نبوده‌اند از شروط الگوریتم عبور نکرده و بنابراین دچار تغییرات و اصلاحات بسیار کمی در طول الگوریتم شده یا بدون تغییر باقی مانده‌اند. الگوی پ: برای تعداد بسیار کمی از ضرایبی که تنک نبوده‌اند الگوریتم سعی در اصلاح مقدار واقعی آن‌ها داشته که بعضا دچار خطا نیز شده و منجر به تنک شدن این ضرایب شده است. الگوی ت: در انتها الگوریتم برای برخی از ضرایبی که تنک نبوده‌اند سعی در اصلاح مقدار واقعی آن‌ها داشته که در الگوی (ت) تا حد قابل قبول و خوبی موفق بوده است. در انتها باید ذکر شود که در تمام الگوهای ارائه شده نیز روند همگرایی قابل ملاحظه است.

۳-۴- نتایج الگوریتم پیشنهادی روی داده‌های واقعی

در این بخش نتایج الگوریتم پیشنهادی روی داده‌های واقعی بررسی شده و بدین منظور از بانک داده‌ی DEAP [۲] استفاده شده است. برای سیگنال‌های EEG ثبت شده از ۳۲ کانال برای هر ویدئو ضرایب AR مرتبه‌ی ۱۰ (۱۰ عدد) استخراج شده و بنابراین برای ۴۰ ویدئو (برای هر سوژه) یک ماتریس $X_{40 \times 320}$ به عنوان ماتریس ویژگی‌ها در نظر گرفته شده است. از ماتریس برچسب Y هر بار برای تخمین یک معیار به صورت تک‌بعدی (بردار y) استفاده شده است. در واقع با در نظر گرفتن چهار معیار Arousal، Valence، Dominance و Liking، مساله چهار بار و هر بار برای یکی از معیارها حل شده است. پس از تخمین پیوسته‌ی هر یک از معیارها، مقادیر کوچک‌تر از ۵ به کلاس صفر و مقادیر بزرگ‌تر از ۵ به کلاس ۱ نسبت داده شده و این روند برای برچسب‌های اصلی نیز اجرا شده است. صحت تخمین هر معیار با توجه به درصد صحت مقادیر گسسته‌ی (۰، ۱) بردار تخمینی به دست آمده و بردار واقعی برای داده‌های آزمون و در حالت 5-Fold Cross Validation گزارش شده است. دقت تخمین برچسب‌های خروجی (۴ معیار اصلی ذکر شده برای تمایز احساسات) به همراه پارامتر بهینه‌ی Thr_{cte} الگوریتم PLSS در جدول (۳) ارائه شده است. مقادیر سایر پارامترهای الگوریتم پیشنهادی برای گزارش نتایج ساختگی و واقعی به صورت زیر می‌باشد.

$$\lambda_{increase} = 5, MaxIterations = 15$$

یکی از ایده‌ها برای تحقیقات آینده می‌تواند آزمایش و بهبود الگوریتم پیشنهادی روی این بانک داده باشد. هدف مراجع مختلف از کار روی این بانک داده تخمین مسیر سه‌بعدی حرکت دست میمون با استفاده از سیگنال مغزی ثبت شده از روی کورتکس (ECoG) است. هم‌چنین برای استفاده از الگوریتم پیشنهادی در موارد غیرخطی می‌توان از نسخه‌های کرنلیزه^۱ شده‌ی الگوریتم فعلی استفاده کرد تا انعطاف مورد نظر برای استفاده در این موارد را داشته باشد. بنابراین یکی دیگر از حوزه‌های پژوهش بعدی تلاش برای ارائه‌ی نسخه‌های غیرخطی از این الگوریتم می‌باشد.

۶- مراجع

[۱] عبدالصالحی، مهدی. "بررسی تغییرات کمیت‌های غیر خطی سیگنال EEG در احساسات شادی، غم و خنثی" رساله دکتری مهندسی پزشکی بیوالکترونیک، دانشگاه آزاد اسلامی واحد علوم تحقیقات، سال تحصیلی ۹۳-۱۳۹۲.

- [2] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., & Patras, I. (2012). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18-31.
- [3] Mehmood, R. M., Du, R., & Lee, H. J. (2017). Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors. *Cities*, 4, 5.
- [4] Zheng, W. (2017). Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis. *IEEE Transactions on Cognitive and Developmental Systems*, 9(3), 281-290.
- [5] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*.
- [6] Katsigiannis, S., & Ramzan, N. (2018). DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE journal of biomedical and health informatics*, 22(1), 98-107.
- [7] Subramanian, R., Wache, J., Abadi, M. K., Vieriu, R. L., Winkler, S., & Sebe, N. (2018). ASCERTAIN: Emotion and personality recognition using commercial sensors. *IEEE Transactions on Affective Computing*, (2), 147-160.
- [8] Hatamikia, S., Maghooli, K., & Nasrabadi, A. M. (2014). The emotion recognition system based on autoregressive model and sequential forward feature selection of

جدول (۴) - مقدار p-value محاسبه شده با استفاده از تست آماری One-Way ANOVA برای الگوریتم پیشنهادی روی داده‌های ساختگی به ازای SNR-های مختلف

Sparsity	-۵ dB	۵ dB	۱۰ dB	۱۵ dB
۵/۱۵	۰/۰۲۱	۰/۰۰۵۶۵	۰/۰۰۳۶	۰/۰۰۹۹۹
۸/۱۵	۰/۰۱۶۷	۰/۰۰۹	۰/۰۰۶۹۶	۰/۰۰۷۷۴

جدول (۵) - مقدار p-value محاسبه شده با استفاده از تست آماری One-Way ANOVA برای الگوریتم پیشنهادی روی داده‌های واقعی

Liking	Dominance	Arousal	Valence
$۳/۸۹ \times ۱۰^{-۵}$	۰/۰۰۴۱	$۱/۲۶ \times ۱۰^{-۵}$	۰/۰۰۴۸

۵- نتیجه‌گیری

در این مقاله یک الگوریتم رگرسیون خطی برای تخمین بهتر معیارهای بازشناسی احساسات انسانی ارائه شده است که از تنک بودن بردار ترکیب در کنار تابع هزینه‌ی رگرسیون خطی استفاده می‌کند. کارایی این الگوریتم روی داده‌های ساختگی بررسی شده و برتری آن نسبت به الگوریتم‌های رگرسیون خطی Sparse Orthonormalized PLS (SOPLS), LASSO, PLS و Ridge Regression نشان داده شده است. هم‌چنین برای اعمال الگوریتم ارائه شده روی داده‌های EEG مربوط به بازشناسی احساسات از مجموعه‌ی دادگان DEAP استفاده شده و ویژگی ضرایب مدل AR از سیگنال‌های EEG استخراج شده است. نتایج به دست آمده از الگوریتم پیشنهادی با نتایج به دست آمده از سایر الگوریتم‌های رگرسیون خطی مقایسه شده که در مجموع نشان دهنده‌ی برتری نسبی الگوریتم ارائه شده است. لازم به ذکر است که هدف اصلی این تحقیق رسیدن به دقت بسیار بالا روی این بانک داده نبوده و بیش‌تر بررسی عمل‌کرد الگوریتم خطی پیشنهادی در کنار سایر الگوریتم‌های خطی معروف روی یک بانک داده‌ی واقعی و معتبر مورد نظر بوده است. البته از آن‌جا که در مسائل بازشناسی احساسات و کار با این بانک داده عملاً باید برچسب‌های علامت‌گذاری شده توسط سوژه‌ها تخمین زده شود، رسیدن به دقت بالا حتی در موارد طبقه‌بندی دوکلاسه کار دشواری است. در مرجع [۱۵] با استفاده از الگوریتم غیرخطی قدرتمند CNN در تخمین دوکلاسه‌ی معیار Valence دقت ۷۲٪ گزارش شده است که به نوعی سختی کار با این بانک داده را نشان می‌دهد. با توجه به وجود مقالات مختلف مبنی بر استفاده از روش‌های مبتنی بر PLS روی بانک داده‌ی Food Tracking Task [۲۱]

^۱ Kernelized



- [15] Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., & Hu, B. (2016, December). Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network. In *Bioinformatics and Biomedicine (BIBM)*, 2016 IEEE International Conference on (pp. 352-359). IEEE.
- [16] Zhao, Q., Caiafa, C. F., Mandic, D. P., Chao, Z. C., Nagasaka, Y., Fujii, N., ... & Cichocki, A. (2013). Higher order partial least squares (HOPLS): a generalized multilinear regression method. *IEEE transactions on pattern analysis and machine intelligence*, 35(7), 1660-1673.
- [17] Rosipal, R., & Trejo, L. J. (2001). Kernel partial least squares regression in reproducing kernel hilbert space. *Journal of machine learning research*, 2(Dec), 97-123.
- [18] Elisseyev, A., & Aksenova, T. (2016). Penalized multi-way partial least squares for smooth trajectory decoding from electrocorticographic (ECoG) recording. *PloS one*, 11(5), e0154878.
- [19] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- [20] Van Gerven, M. A., Chao, Z. C., & Heskes, T. (2012). On the decoding of intracranial data using sparse orthonormalized partial least squares. *Journal of neural engineering*, 9(2), 026017.
- [21] Shimoda, K., Nagasaka, Y., Chao, Z. C., & Fujii, N. (2012). Decoding continuous three-dimensional hand trajectories from epidural electrocorticographic signals in Japanese macaques. *Journal of neural engineering*, 9(3), 036015.
- electroencephalogram signals. *Journal of medical signals and sensors*, 4(3), 194.
- [9] Xie, O., Liu, Z. T., & Ding, X. W. (2018, July). Electroencephalogram Emotion Recognition Based on A Stacking Classification Model. In *2018 37th Chinese Control Conference (CCC)* (pp. 5544-5548). IEEE.
- [10] Piho, L., & Tjahjadi, T. (2018). A mutual information based adaptive windowing of informative EEG for emotion recognition. *IEEE Transactions on Affective Computing*.
- [11] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*.
- [12] Tripathi, S., Acharya, S., Sharma, R. D., Mittal, S., & Bhattacharya, S. (2017, February). Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset. In *AAAI* (pp. 4746-4752).
- [13] Chen, J., Hu, B., Xu, L., Moore, P., & Su, Y. (2015, November). Feature-level fusion of multimodal physiological signals for emotion recognition. In *Bioinformatics and Biomedicine (BIBM)*, 2015 IEEE International Conference on (pp. 395-399). IEEE.
- [14] Rozgić, V., Vitaladevuni, S. N., & Prasad, R. (2013, May). Robust EEG emotion classification using segment level decision fusion. In *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on (pp. 1286-1290). IEEE.