





دانشگاه صنعتی شریف

دانشکده مهندسی برق

پایان نامه کارشناسی

گرایش الکترونیک

عنوان:

بررسی نتایج به دست آمده از طبقه‌بندی‌کننده‌های مختلف برای

یافتن روش مناسب در طبقه‌بندی سیگنال‌های مغزی در BCI

نگارش:

سپیده حاجی‌پور ساردوئی

استاد راهنما:

دکتر محمدباقر شمس‌الهی

مهر ۱۳۸۵

تقدیم به

پدر و مادر عزیزم

بدین وسیله از استاد گرامی جناب آقای دکتر محمدباقر شمس‌الهی که همواره از یاری‌ها و راهنمایی‌های ایشان بهره برده‌ام، سپاسگزاری می‌نمایم.

چکیده:

در سیستم‌های رابط مغز- رایانه یا *BCI*، سعی می‌شود به کمک اطلاعاتی که از سیگنال‌های مغزی یک فرد به دست می‌آید، خواسته‌های شخص به سیگنال‌های کنترلی برای دستگاه‌های خارجی مانند کامپیوتر تبدیل شود. این سیستم‌ها یک پل ارتباطی میان سیگنال‌های مغزی شخص و دنیای خارج برقرار می‌نمایند. برای انجام این فرآیند لازم است پردازش‌های مختلفی بر روی سیگنال‌های مغزی ثبت شده انجام گیرد و این سیگنال‌ها به روشی مناسب طبقه‌بندی گردند. روش‌های مختلفی برای طبقه‌بندی داده‌ها (به طور خاص سیگنال‌های مغزی) وجود دارد. تمایز این روش‌ها در نوع ویژگی‌ها و نوع طبقه‌بندی‌کننده‌های استفاده شده می‌باشد. ویژگی‌های استفاده شده به دسته ویژگی‌های اولیه‌ی استخراج شده از داده‌ها، روش‌های کاهش ویژگی، معیارهای انتخاب ویژگی و الگوریتم‌های جستجو برای یافتن دسته ویژگی بهینه بستگی دارند. بنابراین برای طبقه‌بندی یک دسته داده‌ی معین می‌توان الگوریتم‌های متفاوتی ارائه نمود که نتایج مختلفی ایجاد می‌نمایند. در این پایان‌نامه سعی شده است با مقایسه‌ی نتایج به دست آمده از طبقه‌بندی‌کننده‌ها و روش‌های مختلف انتخاب ویژگی‌های مناسب، الگوریتمی ارائه گردد که بتواند داده‌های ثبت شده را با دقت خوبی طبقه‌بندی نماید.

واژه‌های کلیدی:

رابط مغز- رایانه (*BCI*)، سیگنال‌های مغزی *EEG*، استخراج ویژگی، کاهش ویژگی، انتخاب ویژگی، طبقه‌بندی‌کننده

فهرست مطالب

فصل اول: مقدمه.....	۱
۱-۱ رابط مغز- رایانه.....	۱
۲-۱ سیگنال‌های مغزی <i>EEG</i>	۲
۳-۱ ساختار پایان‌نامه.....	۵
فصل دوم: ویژگی‌ها.....	۷
۱-۲ مقدمه.....	۷
۲-۲ ویژگی‌های آماری.....	۸
۳-۲ ویژگی‌های مبتنی بر آنتروپی.....	۱۰
۴-۲ ویژگی‌های وابسته به تبدیل‌های زمان- فرکانس.....	۱۳
۵-۲ ویژگی‌های وابسته به مدل‌های پارامتری.....	۱۴
۶-۲ ویژگی‌های وابسته به تبدیل‌های فرکانسی.....	۱۵
۷-۲ ویژگی‌های وابسته به انرژی باندهای فرکانسی.....	۱۵
۸-۲ جمع‌بندی.....	۱۶
فصل سوم: طبقه‌بندی‌کننده‌ها.....	۱۸
۱-۳ مقدمه.....	۱۸
۲-۳ طبقه‌بندی‌کننده <i>FLD</i>	۱۹
۳-۳ طبقه‌بندی‌کننده بیز.....	۲۰

۳-۴ طبقه‌بندی‌کننده‌ی ماشین بردار پشتیبان.....	۲۱
۳-۵ طبقه‌بندی‌کننده بر پایه‌ی تابع تفکیک‌کننده‌ی درجه دو.....	۲۳
۳-۶ طبقه‌بندی‌کننده بر پایه‌ی فاصله‌ی مایلانویس.....	۲۳
۳-۷ جمع‌بندی.....	۲۴
فصل چهارم: معیارها و روش‌های انتخاب ویژگی.....	۲۵
۴-۱ مقدمه.....	۲۵
۴-۲ کاهش تعداد ویژگی‌ها.....	۲۶
۴-۳ کاهش بعد به روش ترکیبی.....	۲۸
۴-۴ کاهش بعد به روش گزینش ویژگی‌ها.....	۲۹
۴-۴-۱ معیارهای انتخاب ویژگی.....	۲۹
۴-۴-۱-۱ معیارهای بدون سرپرست، غیروابسته به طبقه‌بندی‌کننده.....	۲۹
۴-۴-۱-۲ معیارهای با سرپرست، غیروابسته به طبقه‌بندی‌کننده.....	۳۰
۴-۴-۱-۳ معیارهای وابسته به طبقه‌بندی‌کننده.....	۳۲
۴-۴-۲ روش‌های جستجو برای گزینش ویژگی‌ها.....	۳۳
۴-۴-۲-۱ روش جستجوی فیلتری.....	۳۳
۴-۴-۲-۲ روش جستجوی پوششی.....	۳۴
۴-۴-۲-۲-۱ جستجوی پیشروی پی‌درپی.....	۳۶
۴-۴-۲-۲-۲ جستجوی شناور پی‌درپی.....	۳۶
۴-۴-۲-۲-۳ جستجوی پرتوی.....	۳۷
۴-۴-۲-۲-۴ جستجوی ژنتیک.....	۳۸

۴-۵	جمع‌بندی.....	۴۰
	فصل پنجم: آزمایش‌ها.....	۴۲
۱-۵	مقدمه.....	۴۲
۲-۵	داده‌های مورد آزمایش.....	۴۳
۳-۵	پیش پردازش داده‌ها.....	۴۴
۴-۵	استخراج ویژگی‌ها.....	۴۵
۵-۵	الگوریتم‌های کاهش ویژگی‌ها.....	۴۷
۱-۵-۵	روش گزینش ویژگی‌ها با تعداد ویژگی ثابت.....	۴۷
۱-۱-۵-۵	الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت.....	۴۸
۲-۱-۵-۵	بررسی نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت.....	۵۱
۲-۵-۵	روش گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده.....	۵۴
۱-۲-۵-۵	الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده.....	۵۴
۲-۲-۵-۵	بررسی نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده.....	۵۷
۳-۵-۵	روش ترکیب ویژگی‌ها با تعداد ویژگی ثابت.....	۵۹
۱-۳-۵-۵	الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت.....	۵۹
۲-۳-۵-۵	بررسی نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت.....	۶۲
۴-۵-۵	روش ترکیب ویژگی‌ها با تعداد ویژگی متغیر.....	۶۳

۵-۵-۴-۱ الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر.....	۶۳
۵-۵-۴-۲ بررسی نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر.....	۶۶
۵-۵-۵ مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی‌ها.....	۶۷
۵-۶ نتیجه‌گیری و مقایسه‌ی نتایج به دست آمده با نتایج دیگر گروه‌ها.....	۷۴
۵-۶-۱ ارائه‌ی الگوریتم کلی برای طبقه‌بندی داده‌ها.....	۷۴
۵-۶-۲ مقایسه‌ی نتیجه‌ی به دست آمده با نتایج دیگر گروه‌ها.....	۷۴
فصل ششم: جمع‌بندی، نتیجه‌گیری و پیشنهادها.....	۷۶
۶-۱ جمع‌بندی.....	۷۶
۶-۲ نتیجه‌گیری.....	۷۷
۶-۳ پیشنهادها.....	۷۸
ضمیمه.....	۸۰
مراجع.....	۸۵

فهرست شکل‌ها

- شکل (۱-۱) : بلوک دیگرام کلی یک سیستم *BCI*..... ۲
- شکل (۲-۱) : شمای کلی یک نورون [۵]..... ۳
- شکل (۳-۱) : سیستم استاندارد ۲۰-۱۰ برای قرارگیری الکترودها [۴]..... ۴
- شکل (۱-۲) : دیاگرام کاری یک ویولت دوطبقه [۹]..... ۱۴
- شکل (۱-۳) : طبقه‌بندی‌کننده‌ی ماشین بردار پشتیبان [۲]..... ۲۲
- شکل (۱-۴) : کاهش تعداد ویژگی‌ها به روش ترکیبی [۱۷]..... ۲۷
- شکل (۲-۴) : کاهش تعداد ویژگی‌ها به روش گزینشی [۱۷]..... ۲۷
- شکل (۳-۴) : تولید نسل جدید توسط عملگر همگذاری [۱۹]..... ۳۹
- شکل (۴-۴) : تولید نسل جدید توسط عملگر جهش [۱۹]..... ۳۹
- شکل (۱-۵) : نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت..... ۵۱
- شکل (۲-۵) : نمودار توزیع دسته ویژگی‌ها به ازای طبقه‌بندی‌کننده‌های مختلف با اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت..... ۵۲
- شکل (۳-۵) : توزیع کانال‌ها برای دسته ویژگی کلی با اعمال الگوریتم گزینش ویژگی با تعداد ویژگی ثابت (الف) *FLD*، (ب) *Bayes*، (ج) *SVM*، (د) *Quadratic*، (ه) *Mahalanobis*..... ۵۳
- شکل (۴-۵) : نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده..... ۵۷

شکل (۵-۵) : نمودار توزیع دسته ویژگی‌ها به ازای طبقه‌بندی‌کننده‌های مختلف با اعمال الگوریتم
گزینه‌ش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده ۵۸

شکل (۶-۵) : توزیع کانال‌ها برای دسته ویژگی کلی با اعمال الگوریتم گزینه‌ش ویژگی با تعداد ویژگی
وابسته به طبقه‌بندی‌کننده (الف) *FLD*، (ب) *Bayes*، (ج) *SVM*، (د) *Quadratic*، (ه) *Mahalanobis* ۵۹

شکل (۷-۵) : نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از
الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت ۶۲

شکل (۸-۵) : نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از
الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر ۶۶

شکل (۹-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی
طبقه‌بندی‌کننده‌ی *FLD* ۶۹

شکل (۱۰-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی
طبقه‌بندی‌کننده‌ی *Bayes* ۷۰

شکل (۱۱-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی
طبقه‌بندی‌کننده‌ی *SVM* ۷۰

شکل (۱۲-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی
طبقه‌بندی‌کننده‌ی *Quadratic* ۷۱

شکل (۱۳-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی
طبقه‌بندی‌کننده‌ی *Mahalanobis* ۷۱

شکل (۱۴-۵) : مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی حاصل از رأی‌گیری میان
طبقه‌بندی‌کننده‌ها ۷۲

فهرست جدول‌ها

- جدول (۱-۲): ویژگی‌های آماری..... ۸
- جدول (۲-۲): ویژگی‌های مبتنی بر آنتروپی..... ۱۰
- جدول (۳-۲): ویژگی‌های وابسته به مدل‌های پارامتری..... ۱۴
- جدول (۴-۲): ویژگی‌های وابسته به تبدیل‌های فرکانسی..... ۱۵
- جدول (۵-۲): طبقه‌بندی‌های فرکانسی برای سیگنال‌های مغزی..... ۱۶
- جدول (۱-۳): طبقه‌بندی‌کننده‌ها..... ۱۹
- جدول (۱-۴): معیارهای بدون سرپرست، غیروابسته به طبقه‌بندی‌کننده..... ۳۰
- جدول (۲-۴): معیارهای با سرپرست، غیروابسته به طبقه‌بندی‌کننده..... ۳۱
- جدول (۳-۴): معیارهای وابسته به طبقه‌بندی‌کننده..... ۳۲
- جدول (۴-۴): الگوریتم‌های جستجوی نمایی..... ۳۴
- جدول (۵-۴): الگوریتم‌های جستجوی پی‌درپی..... ۳۵
- جدول (۶-۴): الگوریتم‌های جستجوی تصادفی..... ۳۵
- جدول (۱-۵): شماره‌گذاری کانال‌های ثبت *EEG*..... ۴۴
- جدول (۲-۵): شماره‌گذاری کانال‌های تفاضلی تعریف شده..... ۴۷
- جدول (۳-۵): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت..... ۵۰
- جدول (۴-۵): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده..... ۵۶

جدول (۵-۵) : درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده

از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت..... ۶۱

جدول (۶-۵) : درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده با استفاده

از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر..... ۶۵

جدول (۷-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی

طبقه‌بندی‌کننده‌ی *FLD*..... ۶۷

جدول (۸-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی

طبقه‌بندی‌کننده‌ی *Bayes*..... ۶۸

جدول (۹-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی

طبقه‌بندی‌کننده‌ی *SVM*..... ۶۸

جدول (۱۰-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی

طبقه‌بندی‌کننده‌ی *Quadratic*..... ۶۸

جدول (۱۱-۵) : مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی

طبقه‌بندی‌کننده‌ی *Mahalanobis*..... ۶۸

جدول (۱۲-۵) : مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی حاصل از رأی‌گیری میان

طبقه‌بندی‌کننده‌ها..... ۶۹

فصل اول: مقدمه

۱-۱ رابط مغز- رایانه

بسیاری از افراد با ناتوانی‌های حرکتی و گفتاری نیاز دارند مقاصد خود را به گونه‌ای به دیگر افراد منتقل سازند و یا نیازهای حرکتی خویش را مرتفع سازند [۱]. رابط مغز- رایانه (*BCI*)^۱، راهی را برای ارتباط مغزی فرد با دنیای خارج ارائه می‌دهد. در اصل در *BCI* با داشتن سیگنال‌های مغزی یک فرد که می‌تواند به صورت‌های گوناگون همچون *EEG*^۲، *ECOG*^۳ یا *MEG*^۴ ثبت شده باشند، به تفسیر حالات یا مقاصد ذهنی فرد پرداخته می‌شود [۲].

برای ایجاد یک سیستم *BCI* موفق توجه به حداقل ۵ مورد لازم است :

۱. هدف از ایجاد سیستم

^۱ *Brain – Computer Interface*

^۲ *Electroencephalogram*

^۳ *Electrocorticogram*

^۴ *Magnetoencephalogram*

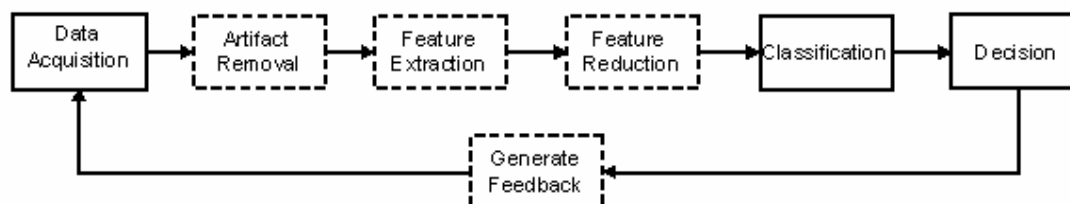
۲. شناخت سیگنال‌های فیزیولوژیکی وابسته

۳. ثبت داده از شخص

۴. استخراج ویژگی‌های مؤثر و مفید از سیگنال‌های خام

۵. طراحی رابط مناسب [۳]

استخراج ویژگی‌های مؤثر را می‌توان به سه بخش (۱) حذف نویزها و آرتیفکت‌ها، (۲) استخراج ویژگی‌ها و (۳) کاهش ویژگی‌ها تقسیم نمود. پس از استخراج ویژگی‌های مؤثر، لازم است برای اتخاذ تصمیم کنترلی مناسب، سیگنال‌ها بر اساس این ویژگی‌ها جداسازی و طبقه‌بندی شوند. علاوه بر اینها گاهی شخص مورد آزمایش تحت فیدبک قرار می‌گیرد و به شخص کمک می‌شود سیگنال‌های مغزی خود را کنترل نماید. به این نوع فیدبک، *Biofeedback* گفته می‌شود. با توجه به موارد مطرح شده، می‌توان یک سیستم رابط مغز- رایانه یا *BCI* را به صورت بلوک دیاگرام کلی شکل (۱-۱) مدل نمود.



شکل (۱-۱): بلوک دیاگرام کلی یک سیستم *BCI*

۲-۱ سیگنال‌های مغزی *EEG*

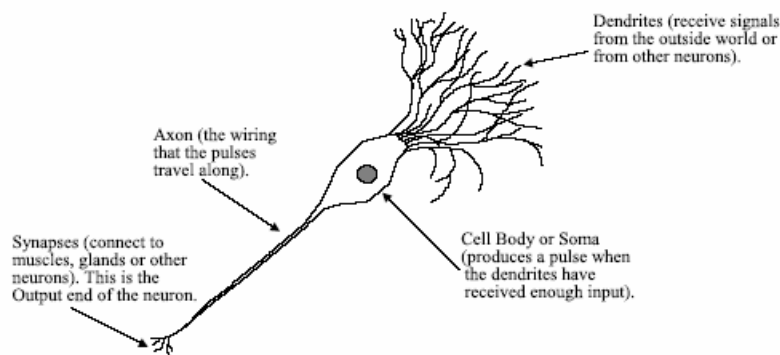
اطلاعات در مورد سیگنال‌های *EEG* می‌تواند از دو طریق به محققین *BCI* کمک کند:

(۱) انتخاب سیگنال مناسبی که بیشترین اطلاعات را به سیستم کنترلی منتقل نماید.

(۲) کمک به توسعه‌ی الگوریتم‌های پردازش سیگنال برای تشخیص سیگنال‌های وابسته.

پس تا زمانی که طبیعت اصلی سیگنال‌های EEG ناشناخته باشد، انتخاب سیگنال‌های مؤثر و روش‌های مناسب با مشکل همراه خواهد بود [۴].

فعالیت‌های الکتریکی مغز به وسیله‌ی بیلون‌ها سلول عصبی به نام نورون^۱ ایجاد می‌شود. هر یک از این نورون‌ها را می‌توان یک پردازشگر الکترونیکی دانست. شکل (۱-۲) شمای کلی یک نورون را نمایش می‌دهد.



شکل (۱-۲): شمای کلی یک نورون [۵]

دندریت‌ها^۲ سیگنال‌ها را از دیگر نورون‌ها یا محیط خارجی دریافت می‌کنند. سیگنال‌های رسیده با هم جمع می‌شوند. اگر این مجموع بیشتر از سطح پتانسیل معینی باشد (تحریک کافی به دندریت‌ها برسد)، بدنه‌ی سلول^۳ یک پالس الکتریکی تولید می‌کند که از طریق آکسون^۴ به دندریت‌های دیگر نورون‌ها و یا ماهیچه‌ها منتقل شده و آنها را تحریک می‌کند. زمانی که یک نورون تحریک بیشتری

^۱ Neuron

^۲ Dendrite

^۳ Cell Body

^۴ Axon

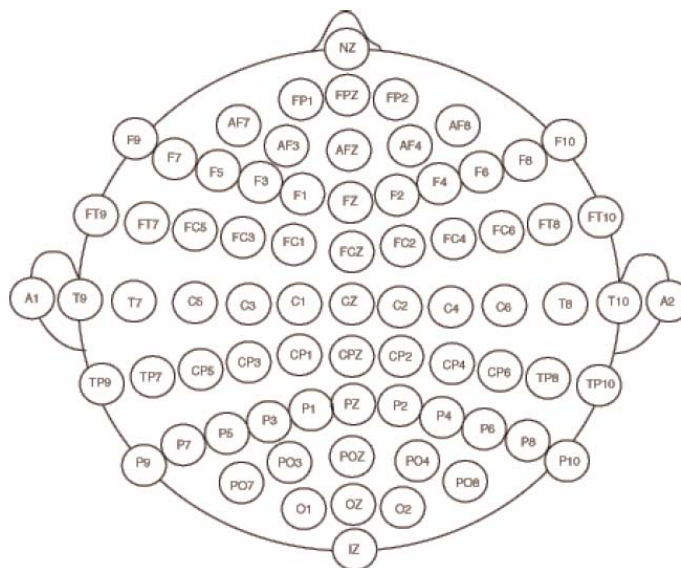
را از طریق دندریت‌ها دریافت نماید، پالس‌های بیشتری را در ثانیه ایجاد می‌کند. در نتیجه می‌توان گفت که اطلاعات در شبکه‌ی عصبی به طور فرکانسی مدوله شده‌اند [۳]، [۵].

EEG، نمایش پتانسیل‌های ولتاژ مغزی است که در طول زمان بر روی کاغذ نوشته می‌شود [۴].

ثبت سیگنال‌های *EEG* به صورت غیر تهاجمی بوده و از طریق الکترودهایی که بر روی پوست سر قرار می‌گیرند، انجام می‌شود. این الکترودها فعالیت‌های الکتریکی مغز را ثبت نموده و به صورت غیر فعال، به تقویت کننده‌ها انتقال می‌دهند. حساسیت این الکترودها بایستی در حد میکرو ولت باشد و نسبت سیگنال به نویز در حد امکان بزرگ باشد. تقویت کننده‌ها سیگنال‌ها را تقریباً ۱۰۰۰۰ برابر می‌کنند. سیگنال‌ها پس از گذشتن از مبدل آنالوگ به دیجیتال برای انجام پردازش‌های لازم و استخراج ویژگی‌ها به کامپیوتر داده می‌شوند [۴].

محل قرار گرفتن الکترودها معمولاً به صورت سیستم استاندارد بین‌المللی ۱۰-۲۰ می‌باشد که در

شکل (۳-۱) نشان داده شده است.



شکل (۳-۱) : سیستم استاندارد ۱۰-۲۰ برای قرارگیری الکترودها [۴]

۳-۱ ساختار پایان نامه

در فصل دوم این پایان نامه، ابتدا دسته ویژگی های مختلف که معمولاً در طبقه بندی سیگنال های مغزی مفید هستند، معرفی می شوند. سپس از هر دسته، ویژگی هایی که در این پروژه از آنها استفاده شده است، توضیح داده می شوند.

فصل سوم به معرفی طبقه بندی کننده های با سرپرست مختلف می پردازد. در این طبقه بندی ها، ابتدا توسط داده های آموزشی عمل یادگیری انجام شده و سپس داده های تست طبقه بندی می شوند. در این فصل روش عملکرد ۵ طبقه بندی کننده ی سریع (*Quadratic* ، *SVM* ، *Bayes* ، *FLD*) و (*Mahalanobis*) که در این پروژه استفاده شده اند، توضیح داده می شود.

در فصل چهارم، ابتدا دلایل نیاز به کاهش تعداد ویژگی ها بررسی شده و دو روش کلی کاهش ویژگی (ترکیبی و گزینشی) معرفی می گردند و با هم مقایسه می شوند. سپس معیارهای مناسب در گزینش ویژگی ها نام برده می شوند و معیارهای مختلفی که آزمایش ها بر اساس آنها انجام شده است، به طور خلاصه توضیح داده می شوند. در پایان فصل روش های مختلف جستجو در میان مجموعه ی تمام ویژگی ها برای یافتن دسته ویژگی بهینه، بررسی می گردند.

داده های مورد آزمایش، نحوه ی انجام آزمایش ها و نتایج به دست آمده از آنها در فصل پنجم ارائه می گردند. در این فصل، چهار الگوریتم مختلف برای کاهش ویژگی ها و طبقه بندی داده ها، معرفی گشته و نتایج به دست آمده از اعمال آنها بر داده ها، بررسی می شوند. در پایان فصل، با توجه به نتایج به دست آمده از آزمایش ها، الگوریتم کلی برای طبقه بندی داده ها ارائه می گردد و نتایج به دست آمده با نتایج دیگر گروه ها مقایسه می شود.

فصل ششم به جمع‌بندی پایان‌نامه و آزمایش‌های انجام شده می‌پردازد و پیشنهادهایی برای

ادامه‌ی کار ارائه می‌کند.

فصل دوم: ویژگی‌ها

۱-۲ مقدمه

در این فصل به معرفی تعدادی از ویژگی‌هایی که عموماً در طبقه‌بندی سیگنال‌های مغزی نقش مؤثری دارند، می‌پردازیم. برای بررسی آسان‌تر ۶ دسته‌ی مختلف از ویژگی‌ها را معرفی نموده و در هر دسته توضیح مختصری در مورد ویژگی‌هایی که در آزمایش‌ها از آنها استفاده شده است، خواهیم داد. دسته‌های مختلف ویژگی‌ها به صورت زیر می‌باشند:

۱. ویژگی‌های آماری

۲. ویژگی‌های مبتنی بر آنتروپی

۳. ویژگی‌های وابسته به تبدیل‌های زمان-فرکانس

۴. ویژگی‌های وابسته به مدل‌های پارامتری

^۱ Feature

۵. ویژگی‌های وابسته به تبدیل‌های فرکانسی

۶. ویژگی‌های وابسته به انرژی باندهای فرکانسی

که هر دسته از آنها در یک زیرفصل بررسی می‌گردند.

۲-۲ ویژگی‌های آماری^۱

این دسته ویژگی‌ها از روی خواص آماری سیگنال استخراج می‌شوند که با فرض ارگادیک^۲

بودن سیگنال، می‌توان آنها را از روی نمونه‌ای زمانی سیگنال تخمین زد [۲]. جدول (۱-۲) لیست

تعدادی از ویژگی‌های آماری که معمولاً در پردازش سیگنال‌های مغزی مؤثر هستند، نشان می‌دهد.

جدول (۱-۲): ویژگی‌های آماری

ویژگی‌های آماری	Features	استفاده شده در آزمایش‌ها
میانگین آماری	Statistical Mean	✓
واریانس آماری	Statistical Variance	✓
همبستگی میان سیگنال‌ها	Correlation	✓
مومنتوم	Moment	✓
کامیولنت	Comulant	
فرم فاکتور	Form Factor	✓

در ادامه توضیح مختصری در مورد ویژگی‌های آماری که در این پروژه استفاده شده‌اند، داده می‌شود. در

فرمول‌های ارائه شده، x_i ، x_j و ... نشانگر نمونه‌های سیگنال زمانی و N برابر با طول سیگنال است.

^۱ Statistical

^۲ ergodic

۱. میانگین آماری^۱: برابر است با مجموع همه نمونه‌های سیگنال تقسیم بر تعداد نمونه‌ها:

$$\bar{x}_i = \frac{1}{N} \sum_{n=1}^N x_i[n] \quad (1-2)$$

۲. واریانس آماری^۲: برابر است با میانگین آماری توان دوم اختلاف نمونه‌های سیگنال با میانگین آن:

$$\sigma_{x_i}^2 = \frac{1}{N} \sum_{n=1}^N (x_i[n] - \bar{x}_i[n])^2 \quad (2-2)$$

۳. همبستگی^۳ میان سیگنال‌ها: برابر است با میانگین آماری ضرب اختلاف دو سیگنال با میانگین آنها:

$$\sigma_{x_i, x_j} = \frac{1}{N} \sum_{n=1}^N (x_i[n] - \bar{x}_i)(x_j[n] - \bar{x}_j) \quad (3-2)$$

۴. مومنتوم^۴: با فرض ارگادیک بودن سیگنال، برابر است با میانگین ضرب سیگنال‌های تواندار، که x_i ها

می‌توانند نمونه‌های یک سیگنال که نسبت به هم تأخیر زمانی دارند بوده و یا اینکه نمونه‌های

سیگنال‌هایی باشند که از منابع مختلف دریافت شده‌اند (مانند سیگنال‌های مربوط به کانال‌های مختلف

ثبت) [۲]:

$$Mom[x_1^{k1}, x_2^{k2}, \dots, x_m^{km}] = \frac{1}{N} \sum_{n=1}^N (x_1[n])^{k1} \times (x_2[n])^{k2} \times \dots \times (x_m[n])^{km} \quad (4-2)$$

^۱ Statistical Mean

^۲ Statistical Variance

^۳ Correlation

^۴ Moment

۵. فرم فاکتور^۱: برابر است با نسبت پویایی مشتق یکم سیگنال به پویایی سیگنال که پویایی سیگنال به

صورت نسبت مجذور واریانس مشتق اول سیگنال به مجذور واریانس خود سیگنال تعریف می‌شود [۲]:

$$FF(x) = \frac{\sigma'_x / \sigma_x}{\sigma''_x / \sigma'_x} \quad (5-2)$$

۳-۲ ویژگی‌های مبتنی بر آنتروپی

آنتروپی معمولاً به عنوان معیاری برای شناخت پیچیدگی، بی‌نظمی و یا اطلاعات سیگنال معرفی

می‌شود. جدول (۲-۲) لیست تعدادی از ویژگی‌های مبتنی بر آنتروپی را نشان می‌دهد.

جدول (۲-۲): ویژگی‌های مبتنی بر آنتروپی

ویژگی‌های مبتنی بر آنتروپی	Features	استفاده شده در آزمایش‌ها
آنتروپی شنن	<i>Shannon Entropy</i>	✓
آنتروپی رنیه	<i>Renyi Entrop</i>	✓
آنتروپی تسالیس	<i>Tsallis Entropy</i>	✓
معیار پیچیدگی عصبی	<i>Neural Complexity Measure</i>	
آنتروپی تقریبی	<i>Approximation Entropy</i>	✓
معیار پیچیدگی L-Z	<i>Lemple-Ziv Complexity Measure</i>	✓

در ادامه توضیح کوتاهی در مورد ویژگی‌های مبتنی بر آنتروپی که در این پروژه از آنها استفاده شده

است، ارائه می‌شود.

در تمام روابط داده شده، $P(x)$ میزان احتمال رخداد x است.

۱. آنتروپی شنن^۲: آنتروپی شنن، به صورت زیر تعریف می‌شود [۶]:

^۱ Form Factor

^۲ Shannon Entropy

برای سیگنال پیوسته:

$$H(x) = - \int_{-\infty}^{+\infty} P(x) \log_2(P(x)) \quad (6-2)$$

برای سیگنال گسسته:

$$H(x) = - \sum_x P(x) \log_2(P(x)) \quad (7-2)$$

۲. آنتروپی رنیه^۱: آنتروپی رنیه بر حسب پارامتر $\alpha \neq 1$ ، به صورت زیر تعریف می شود:

$$H_\alpha(x) = \frac{1}{1-\alpha} \log_2 \left[\sum_x (P(x))^\alpha \right] \quad (8-2)$$

زمانی که α به سمت ۱ میل کند، آنتروپی رنیه با آنتروپی شنن یکی خواهد شد [۶].

۳. آنتروپی تساليس^۲: آنتروپی تساليس بر حسب پارامتر $q \neq 1$ ، به صورت زیر تعریف می شود:

$$H_q(x) = \frac{1}{1-q} \left[\sum_x (P(x))^q - 1 \right] \quad (9-2)$$

زمانی که q به سمت ۱ میل کند، آنتروپی تساليس با آنتروپی شنن یکی خواهد شد [۶].

۴. آنتروپی تقریبی^۳: m را عددی طبیعی و r را عددی حقیقی و مثبت در نظر می گیریم.

اگر $u(1), u(2), \dots, u(N)$ نمونه های زمانی سیگنال با فاصله های زمانی یکسان باشند. بردارهای $X(i)$ را

در R^m به صورت زیر تعریف می کنیم:

$$X(i) = [u(i), u(i+1), \dots, u(i+m-1)] \quad 1 \leq i \leq N-m+1 \quad (10-2)$$

^۱ Renyi Entropy

^۲ Tsallis Entropy

^۳ Approximation Entropy

همچنین تعاریف زیر را در نظر می گیریم:

$$d[X(i), X(j)] = \max(|u(i+k-1) - u(j+k-1)|) \quad k = 1, 2, \dots, m \quad (11-2)$$

$$C_i^m(r) = \frac{\text{number of } j \text{ such that } d[X(i), X(j)] \leq r}{(N-m+1)} \quad (12-2)$$

$$\phi^m(r) = \frac{\sum_{i=1}^{N-m+1} \log(C_i^m(r))}{(N-m+1)} \quad (13-2)$$

با توجه به روابط داده شده، آنتروپی تقریبی به صورت زیر تعریف می شود:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \quad (14-2)$$

معمولاً m را برابر ۲ و r را میان ۰,۱ و ۰,۲۵ از σ_x در نظر می گیرند [۷].

۵. معیار پیچیدگی $L-Z$ ^۱: برای به دست آوردن این معیار الگوریتم زیر را تا زمانی که به پایان رشته

برسیم، اجرا می نماییم:

(۱) به تعداد $\alpha - 1$ حد آستانه تعریف می کنیم تا به کمک آن سیگنال را به رشته ای از اعداد حسابی

تبدیل کنیم. برای نمونه اگر $\alpha = 2$ باشد، آن گاه در هر نقطه که سیگنال از آستانه بیشتر بود، به

آن مقدار ۱ و هر جا کمتر بود به آن مقدار صفر می دهیم.

(۲) c را برابر ۱ قرار می دهیم.

(۳) داده ی یکم از رشته را S نامیده و داده ی دوم را به عنوان Q در کنار آن قرار می دهیم.

(۴) رشته $SQ\pi$ را به صورت رشته ای که شامل S و Q بوده و داده ی آخر آن حذف شده تعریف

می کنیم.

^۱ Lemple-Ziv Complexity Measure

۵) اگر Q در $SQ\pi$ دیده شود، طول Q را با اضافه کردن یک داده‌ی بعدی از رشته‌ی (اصلی) بالا

می‌بریم و به مرحله‌ی ۳ بازمی‌گردیم. در غیر این صورت c را یک واحد افزایش داده و S را

برابر SQ قرار داده و داده‌ی بعدی رشته‌ی (اصلی) را Q می‌نامیم و به مرحله‌ی ۳ می‌رویم.

پس از اتمام الگوریتم فوق، معیار پیچیدگی $L-Z$ را برای سیگنال به صورت زیر تعریف می‌کنیم:

$$C = \frac{c}{N} \log_{\alpha}(N) \quad (15-2)$$

که در آن N برابر طول سیگنال است [۲].

۲-۴ ویژگی‌های وابسته به تبدیل‌های زمان-فرکانس

از جمله تبدیل‌های معروف زمان-فرکانس (زمان-مقیاس) می‌توان تبدیل ویولت^۱ را نام برد.

تبدیل ویولت، یک تبدیل انتگرالی است که از توابع پایه‌ای به نام ویولت استفاده می‌کند. مزیت این

تبدیل، بررسی کردن این سیگنال به صورت محلی می‌باشد که آن را مکان‌یابی زمان-فرکانس می‌نامند.

بنابراین سیگنال‌های غیرماندگار یا سیگنال‌هایی که مؤلفه‌های گذرای کوتاه مدتی دارند، توسط این روش

قابل تجزیه و تحلیل می‌باشند. در تجزیه و تحلیل ویولت، پارامتر مقیاس، جانشین پارامتر فرکانس در

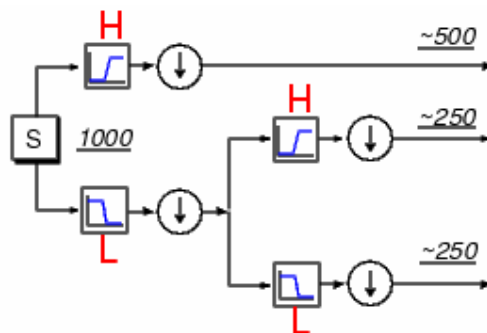
تجزیه و تحلیل زمان-فرکانس می‌شود [۸].

این تبدیل بر پایه‌ی گذراندن سیگنال از دو فیلتر بالاگذر و پایین‌گذر و سپس کاهش نمونه‌ها، به

دست می‌آید. این فرآیند می‌تواند چند بار (در چند طبقه) انجام بگیرد. با توجه به انواع فیلترهایی که

^۱ Wavelet Transform

برای این تبدیل می‌توان به کار برد، تبدیل ویولت انواع گوناگونی خواهد داشت [۲]. شکل (۱-۲) یک تبدیل ویولت دوطبقه را نشان می‌دهد.



شکل (۱-۲): دیاگرام کاری یک ویولت دوطبقه [۹]

۵-۲ ویژگی‌های وابسته به مدل‌های پارامتری

گاهی می‌توان یک سیگنال را با تعدادی پارامتر مدل نمود و این پارامترها را تحت شرایطی تخمین زد. لیست تعدادی از مدل‌های پارامتری در جدول (۳-۲) آمده است.

جدول (۳-۲): ویژگی‌های وابسته به مدل‌های پارامتری

ویژگی‌های پارامتری	Features	استفاده شده در آزمایش‌ها
مدل <i>AR</i>	<i>Autoregressive Model</i>	✓
مدل <i>MA</i>	<i>Moving Average Model</i>	
مدل <i>ARMA</i>	<i>Mixed Autoregressive Moving Average Model</i>	

در اینجا مدل *AR*^۱ که در آزمایش‌ها از آن استفاده شده است، معرفی می‌گردد:

در این مدل، دنباله‌ی زمانی توسط معادله دیفرانسیل خطی تخمین زده می‌شود:

$$x(n) = \sum_{i=1}^P \alpha(i)x(n-i) + e(n) \quad (۱۶-۲)$$

^۱ Autoregressive

که در این مدل، نمونه فعلی دنباله‌ی زمانی $x(n)$ ، یک تابع خطی از P نمونه‌ی قبلی زمانی و نویز سفید ورودی $e(n)$ می‌باشد [۱۰].

در مدل AR ، تعداد محدودی پارامتر برای تعریف ویژگی‌های طیف کافی است که می‌توان از

آنها به عنوان ویژگی استفاده نمود. همچنین در این مدل برخلاف مدل $ARMA$ ^۱ و دیگر مدل‌های غیرخطی که ممکن است مینیموم محلی اتفاق بیفتد، فقط یک جواب بهینه برای یک دسته داده وجود دارد [۱۱].

۶-۲ ویژگی‌های وابسته به تبدیل‌های فرکانسی

تبدیل‌های گوناگونی هستند که با تغییر حوزه‌ی سیگنال از زمان به فرکانس، اطلاعاتی از سیگنال

را در تعداد کمی المان قرار می‌دهند. لیست تعدادی از این تبدیل‌ها در جدول (۴-۲) آمده است.

جدول (۴-۲): ویژگی‌های وابسته به تبدیل‌های فرکانسی

ویژگی‌های وابسته به تبدیل‌های فرکانسی	Features	استفاده شده در آزمایش‌ها
تبدیل فوریه	<i>Discrete Fourier Transform</i>	
تبدیل کسینوسی	<i>Discrete Cosine Transform</i>	✓
تبدیل سینوسی	<i>Discrete Sine Transform</i>	✓

در این پروژه از تبدیل‌های کسینوسی (DCT)^۲ و سینوسی (DST)^۳ استفاده شده است.

۷-۲ ویژگی‌های وابسته به انرژی باندهای فرکانسی

^۱ *Mixed Autoregressive Moving Average*

^۲ *Discrete Cosine Transform*

^۳ *Discrete Sine Transform*

می‌توان انرژی سیگنال در باندهای مختلف فرکانسی را به عنوان ویژگی برای طبقه‌بندی به کار برد. با توجه به محدودیت فرکانسی سیگنال‌های مغزی و اینکه سیگنال مغزی در برخی از بازه‌ها فعالیت متفاوت و ویژه‌ای دارد، تقسیم‌بندی‌های مختلفی برای انتخاب بازه‌های خاص فرکانسی انجام شده است. چهار دسته از این تقسیم‌بندی‌ها در جدول (۵-۲) معرفی شده‌اند.

جدول (۵-۲): تقسیم‌بندی‌های فرکانسی برای سیگنال‌های مغزی

بازمهای فرکانسی (هرتز)							شماره دسته	
β		α		θ		δ	۱	
۲۰-۱۲/۵		۱۲/۵-۷/۵		۷/۵-۲/۵		۲/۵-۰		
β		α		θ		δ	۲	
۲۰-۱۴		۱۲-۷		۷-۴		۲-۰		
β_4	β_3	β_2	β_1	α	θ	δ	۳	
۲۰-۲۵/۵	۲۵-۱۸	۱۷/۵-۱۵	۱۵-۱۲	۱۲-۷/۵	۷-۲/۵	۲/۵-۰/۵		
β_4	β_3	β_2	β_1	α_2	α_1	θ	δ	۴
۴۰-۳۰	۲۰-۲۱	۲۱-۱۸/۵	۱۸/۵-۱۲/۵	۱۲-۱۰/۵	۱۰-۸/۵	۸-۶	۶-۱/۵	

۸-۲ جمع‌بندی

در این فصل به معرفی ویژگی‌هایی که معمولاً در طبقه‌بندی سیگنال‌های مغزی مؤثر هستند، پرداخته شد. این ویژگی‌ها در ۶ دسته‌ی ویژگی‌های آماری، مبتنی بر آنتروپی، وابسته به تبدیل‌های زمان-فرکانس، پارامتری، وابسته به تبدیل‌های فرکانسی و وابسته به انرژی باندهای فرکانسی بررسی

شدند. در فصل بعد طبقه‌بندی‌کننده‌هایی که معمولاً برای طبقه‌بندی سیگنال‌های مغزی کاربرد دارند، معرفی می‌شوند.

فصل سوم : طبقه‌بندی‌کننده‌ها

۱-۳ مقدمه

در این بخش به معرفی طبقه‌بندی‌کننده^۱های باسرپرست^۲ می‌پردازیم. در این طبقه‌بندی‌کننده‌ها، ابتدا به کمک داده‌های آموزشی^۳ و یا ویژگی‌های انتخاب شده از آنها، فرآیند یادگیری^۴ انجام شده و سپس داده‌های تست^۵ به وسیله‌ی آنها طبقه‌بندی می‌شوند. طبقه‌بندی‌کننده‌ها انواع مختلفی دارند و با توجه به روش آموزش آنها، سرعت آنها در کلاس‌بندی داده‌ها متفاوت می‌باشد. تعدادی از طبقه‌بندی‌کننده‌های معروف در جدول (۱-۳) معرفی شده‌اند.

^۱ Classifier

^۲ Supervised

^۳ Training Data

^۴ Learning

^۵ Test Data

جدول (۱-۳) : طبقه‌بندی‌کننده‌ها

طبقه‌بندی‌کننده	Classifier	استفاده شده در آزمایش‌ها
طبقه‌بندی‌کننده‌ی FLD	<i>Fisher Linear Discriminant</i>	✓
طبقه‌بندی‌کننده‌ی بیز	<i>Bayesian</i>	✓
طبقه‌بندی‌کننده بر پایه‌ی ماشین بردار پشتیبان	<i>Support Vector Machine</i>	✓
طبقه‌بندی‌کننده k تا نزدیک‌ترین همسایگی	<i>k-Nearest Neighbor</i>	
طبقه‌بندی‌کننده بر پایه‌ی شبکه‌های عصبی	<i>Neural Network (k-NN)</i>	
طبقه‌بندی‌کننده بر پایه‌ی شبکه‌های با تابع شعاعی	<i>Radial Basis Function (RBF)</i>	
طبقه‌بندی‌کننده بر پایه‌ی شبکه‌های تطابقی با استنتاج فازی	<i>Adaptive-Network-Based Fuzzy Interface System</i>	
طبقه‌بندی‌کننده بر پایه‌ی شبکه‌های فازی-عصبی	<i>Neuro-Fuzzy</i>	
طبقه‌بندی‌کننده بر پایه‌ی تابع تفکیک‌کننده‌ی درجه	<i>Quadratic</i>	✓
طبقه‌بندی‌کننده بر پایه‌ی فاصله‌ی ماحالانوبیس	<i>Mahalanobis</i>	✓

در ادامه به معرفی طبقه‌بندی‌کننده‌هایی که در این پروژه از آنها استفاده شده است، می‌پردازیم.

در تمام روابط داده شده، μ_i و Σ_i به ترتیب بردار میانگین و ماتریس کواریانس کلاس i ام می‌باشند.

۲-۳ طبقه‌بندی‌کننده‌ی FLD ^۱

FLD یک طبقه‌بندی‌کننده با تابع جداکننده‌ی خطی است که به صورت:

$$f(x) = w^T x + b \quad (۱-۳)$$

مشخص می‌شود. اگر داده‌های آموزشی به صورت مجموعه‌ی:

$$\Gamma = \{(x_1, y_1), \dots, (x_n, y_n)\} \quad (۲-۳)$$

تعریف شود به گونه‌ای که $y_i \in \{-1, 1\}$ و دو کلاس مختلف را مشخص نمایند، معیار تفکیک‌پذیری دو

کلاس به صورت زیر تعریف می‌گردد:

^۱ Fisher Linear Discriminant

$$F(w) = \frac{w^T S_B w}{w^T S_W w} \quad (3-3)$$

که در این رابطه، S_B ماتریس پخششی میان-کلاسی^۱ بوده و برابر است با:

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \quad (4-3)$$

و S_W ماتریس پخششی درون-کلاسی^۲ بوده و به صورت زیر تعریف می‌شود:

$$S_W = S_1 + S_2, \quad S_y = \sum_{i \in Y_y} (x_i - \mu_y)(x_i - \mu_y)^T, \quad y \in \{1, 2\} \quad (5-3)$$

در روش کلاسیک برای یافتن w مناسب برای ماکزیمم نمودن معیار تفکیک‌پذیری، به صورت زیر عمل می‌شود:

$$w = S_W^{-1}(\mu_1 - \mu_2)^T \quad (6-3)$$

برای کامل نمودن پارامترهای تابع جداکننده خطی، مقدار b از معادله‌ی زیر محاسبه می‌شود:

$$f(\mu_1) = -f(\mu_2) \quad (7-3)$$

پس از یافتن تابع جداکننده خطی، با قرار دادن بردارهای داده‌های تست در معادله‌ی به دست آمده، کلاس مربوط به هر داده تعیین می‌شود [۱۲]، [۱۳].

۳-۳ طبقه‌بندی‌کننده‌ی بیز^۳

این طبقه‌بندی‌کننده بر پایه‌ی احتمال شرطی رخداد در کلاس‌ها عمل می‌کند. در واقع کلاسی

برای رخداد x انتخاب می‌شود که احتمال تعلق به آن کلاس برای این رخداد، بیشتر باشد.

به صورت ریاضی می‌توان گفت:

^۱ Between-class scatter Matrix

^۲ Within-class scatter Matrix

^۳ Bayesian

$$\begin{cases} \text{if } P(x | w_1) > P(x | w_2) & \text{then } x \in w_1 \\ \text{if } P(x | w_1) < P(x | w_2) & \text{then } x \in w_2 \end{cases} \quad (۸-۳)$$

که در روابط بالا w_1 و w_2 کلاس‌های ممکن هستند.

برای به دست آوردن کلاس وابسته به داده‌ها، باید تخمینی از توزیع احتمالی کلاس‌ها داشته باشیم. چون عموماً از پارامترها و چگونگی توزیع احتمال کلاس‌ها آگاهی زیادی نداریم، از تخمین نرمال استفاده می‌کنیم.

با فرض توزیع نرمال می‌توان پارامترهای زیر را برای کلاس i تعریف نمود:

$$A_i = -0.5 \times \Sigma_i^{-1} \quad (۹-۳)$$

$$b_i = \Sigma_i^{-1} \times \mu_i \quad (۱۰-۳)$$

$$c_i = -0.5 \times \mu_i^T \times \Sigma_i^{-1} \times \mu_i - 0.5 \times \log|\Sigma_i| + \log(P_i) \quad (۱۱-۳)$$

$$d_i(x) = x^T \times A_i \times x + b_i^T \times x + c_i \quad (۱۲-۳)$$

با توجه به تعاریف بالا می‌توان رابطه‌ی (۸-۳) را به صورت زیر ساده نمود [۲]:

$$\begin{cases} \text{if } d_1(x) > d_2(x) & \text{then } x \in w_1 \\ \text{if } d_1(x) < d_2(x) & \text{then } x \in w_2 \end{cases} \quad (۱۳-۳)$$

۳-۴ طبقه‌بندی‌کننده‌ی ماشین بردار پشتیبان^۱

در طبقه‌بندی‌کننده‌ی ماشین بردار پشتیبان (SVM)، ابتدا داده‌های آموزشی را با توجه به کلاس مشخص آنها با یک ابرصفحه جدا می‌نماییم. چون از یک ابرصفحه استفاده می‌نماییم، روش طبقه‌بندی، یک روش خطی است. ابرصفحه‌ی انتخاب شده باید به گونه‌ای باشد که ماکزیمم فاصله با نزدیک‌ترین

^۱ Support Vector Machine (SVM)

داده‌ها از دو کلاس را داشته باشد. این خاصیت به منظور افزایش دقت طبقه‌بندی‌کننده در تفکیک داده‌های تست، استفاده می‌شود.

فرض می‌کنیم داده‌های آموزشی به صورت مجموعه‌ی:

$$\Gamma = \{(x_1, c_1), \dots, (x_n, c_n)\} \quad (14-3)$$

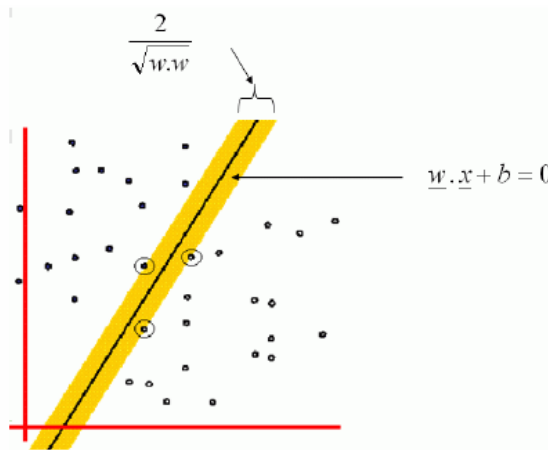
که $c_i \in \{-1, 1\}$ بوده و کلاس هر داده را مشخص می‌کند. با توجه به شکل (۱-۳) اگر ابرصفحه‌ی بهینه را به صورت:

$$w^T x + b = 0 \quad (15-3)$$

در نظر بگیریم، می‌توان نشان داد که پهنای مرز به صورت:

$$\frac{2}{|w|} \quad (16-3)$$

محاسبه می‌شود.



شکل (۱-۳): طبقه‌بندی‌کننده‌ی ماشین بردار پشتیبان [۲]

برای به دست آوردن ابرصفحه‌ای با ماکزیمم تفکیک‌پذیری داده‌ها، بایستی w کمینه شود. این

کار با حل معادلات درجه دوم و داشتن یک سری از شرط‌ها به صورت معادله و نامعادله انجام می‌شود

[۲]، [۱۴].

چون در بسیاری از موارد مرزی خطی که تمام داده‌ها را به درستی جدا کند، وجود ندارد، در حل مسئله‌ی ماشین بردار پشتیبان یک تابع خطا در طبقه‌بندی در نظر می‌گیرند که به نوعی نمایانگر مجموع فاصله‌های داده‌هایی است که به درستی طبقه‌بندی نشده‌اند. بنابراین مسئله به بیشینه کردن پهنای مرز به همراه کمینه کردن تابع خطا تبدیل خواهد شد [۲].

۳-۵ طبقه‌بندی‌کننده بر پایه‌ی تابع تفکیک‌کننده‌ی درجه دو^۱

در طبقه‌بندی‌کننده بر پایه‌ی تابع درجه دو، از یک مرز با معادله‌ی درجه دو، برای جداسازی داده‌های آموزشی استفاده می‌شود. بنابر این کلاس اختصاص یافته به هر داده بایستی تابعی از:

$$x^T \times A \times x + b^T \times x + c \quad (۳-۱۷)$$

باشد. در این طبقه‌بندی‌کننده فرض بر این است که توزیع احتمال داده‌ها به صورت نرمال است و داده‌ها در دو کلاس $y \in \{-1, 1\}$ جای گرفته‌اند. در این حالت نسبت احتمال حضور در کلاس‌ها با در نظر گرفتن یک مقدار آستانه‌ی t ، به صورت زیر به دست می‌آید [۱۵]:

$$Likelihood \text{ Ratio} = \frac{\sqrt{2\pi|\Sigma_1^{-1}|} \times \exp((x - \mu_1)^T \times \Sigma_1 \times (x - \mu_1))}{\sqrt{2\pi|\Sigma_2^{-1}|} \times \exp((x - \mu_2)^T \times \Sigma_2 \times (x - \mu_2))} \quad (۳-۱۸)$$

۳-۶ طبقه‌بندی‌کننده بر پایه‌ی فاصله‌ی ماهالانوبیس^۲

^۱ Quadratic

^۲ Mahalanobis

این طبقه‌بندی‌کننده بر اساس فاصله‌ی ماحالانویس تا مرکز کلاس‌ها عمل می‌نماید. ابتدا برای داده‌های آموزشی بردار میانگین هر یک از کلاس‌ها محاسبه می‌شود. سپس برای هر یک از داده‌های تست، فاصله‌ی ماحالانویس بین آن داده و میانگین کلاس‌ها تعیین می‌شود:

$$d_k(x, \bar{x}_k) = (x - \bar{x}_k)^T C_k^{-1} (x - \bar{x}_k) \quad (19-3)$$

که در رابطه‌ی داده شده، \bar{x}_k بردار میانگین و C_k ماتریس کواریانس کلاس k ام هستند. کلاسی که کمترین فاصله‌ی ماحالانویس تا داده‌ی مورد نظر را دارد، به عنوان کلاس مربوط به آن داده تعیین می‌گردد [۱۶].

۷-۳ جمع‌بندی

در این فصل طبقه‌بندی‌کننده‌های مختلفی که معمولاً در طبقه‌بندی سیگنال‌های مغزی کاربرد دارند، معرفی شدند. سپس توضیح مختصری در مورد ۵ طبقه‌بندی‌کننده که در این پروژه از آنها استفاده شده است، داده شد. مزیت این ۵ طبقه‌بندی‌کننده، سرعت بالای آنها در آموزش و طبقه‌بندی داده‌های تست می‌باشد. با توجه به مطالب فصل‌های دوم و سوم، در فصل بعد روش‌های انتخاب ویژگی‌های مؤثر در طبقه‌بندی داده‌ها معرفی می‌شوند.

فصل چهارم: معیارها و روش‌های انتخاب ویژگی

۴-۱ مقدمه

طبقه‌بندی^۱ به معنای مشخص نمودن کلاس برای هر داده می‌باشد. در مبحث طبقه‌بندی معمولاً تعدادی داده به عنوان داده‌های آموزشی^۲ معرفی می‌شوند که کلاس مربوط به هر کدام از آنها مشخص شده است. بقیه‌ی داده‌ها به عنوان داده‌های تست^۳ هستند که بایستی کلاس مربوط به آنها را تعیین نمود. برای این منظور از روی هر داده که می‌تواند یک سیگنال زمانی باشد، یک سری ویژگی^۴ تعیین می‌کنند و از این ویژگی‌ها برای طبقه‌بندی داده‌ها استفاده می‌کنند. تعدادی از ویژگی‌هایی که معمولاً در طبقه‌بندی سیگنال‌های مغزی مؤثرند، در فصل دوم معرفی شدند.

^۱ Classification

^۲ Training Data

^۳ Test Data

^۴ Feature

معمولاً تعداد ویژگی‌هایی که می‌توان برای هر داده مشخص نمود، بسیار زیاد است. بایستی تعداد ویژگی‌ها را به گونه‌ای کاهش داد که اولاً عملیات لازم برای طبقه‌بندی داده‌ها کاهش بیابد و ثانیاً ویژگی‌هایی که تأثیر بیشتری در طبقه‌بندی صحیح دارند انتخاب شوند. روش‌های کاهش بعد^۱ می‌توانند به صورت ترکیبی و یا گزینشی باشند. همچنین معیار^۲های مختلفی برای انتخاب ویژگی‌های مؤثر، وجود دارد.

در این فصل دو روش کاهش بعد (ترکیبی و گزینشی) معرفی شده و با هم مقایسه می‌شوند. در روش کاهش بعد گزینشی، معیارهایی که در انتخاب ویژگی‌های مؤثر کاربرد دارند، توضیح داده می‌شوند و در آخر، روش‌های جستجو^۳ برای یافتن دسته ویژگی‌های مؤثر معرفی می‌گردند.

۴-۲ کاهش تعداد ویژگی‌ها

همان گونه که گفته شد، لازم است برای افزایش سرعت و دقت طبقه‌بندی‌کننده‌ها، تعداد ویژگی‌ها را به گونه‌ای کاهش دهیم. دو روش برای کاهش ویژگی‌ها وجود دارد:

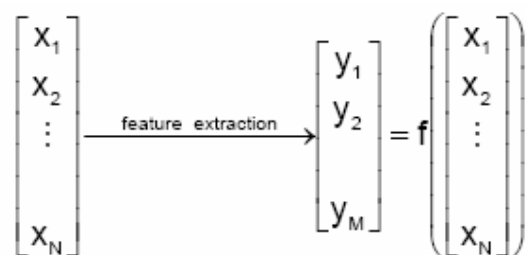
۱. کاهش بعد به روش ترکیبی:

در این روش مجموعه‌ای از ویژگی‌ها که از ترکیب ویژگی‌های اولیه تولید شده‌اند، انتخاب می‌شود. روش تعیین ویژگی‌های جدید از روی ویژگی‌های اولیه در شکل (۴-۱) نشان داده شده است.

^۱ Dimensionality reduction

^۲ Evaluation Measure

^۳ Search Algorithms



شکل (۴-۱) : کاهش تعداد ویژگی‌ها به روش ترکیبی [۱۷]

۲. کاهش بعد به روش گزینشی^۱:

در این روش زیرمجموعه‌ای از ویژگی‌های اولیه بدون اعمال نگاشت، انتخاب می‌شود. روش

تعیین ویژگی‌های جدید از روی ویژگی‌های اولیه در شکل (۴-۲) نشان داده شده است.



شکل (۴-۲) : کاهش تعداد ویژگی‌ها به روش گزینشی [۱۷]

می‌توان گفت کاهش بعد به روش گزینشی، حالت خاصی از کاهش بعد به روش ترکیبی است

که ماتریس نگاشت آن با تعدادی ۱ در قطر، تعریف می‌شود [۱۷].

در ادامه مقایسه‌ای بین دو روش کاهش ویژگی انجام می‌شود:

• ممکن است به دست آوردن ویژگی‌ها سخت و زمان‌بر باشد. در روش گزینش ویژگی‌ها، فقط

تعداد محدودی از ویژگی‌ها برای طبقه‌بندی استفاده می‌شوند و نیازی به تعیین تمام آنها برای

^۱ Feature Subset Selection

داده‌های تست نیست. اما در روش ترکیب ویژگی‌ها، بایستی تمام ویژگی‌ها برای داده‌های تست نیز معین گردند و سپس با ترکیب آنها ویژگی‌های مؤثر تولید شوند.

- با انتخاب ویژگی‌ها به روش گزینشی، می‌توان قوانین معناداری برای طبقه‌بندی‌کننده‌ها تعیین نمود. اما در روش ترکیب ویژگی‌ها، معنای ویژگی‌های اولیه در اثر اعمال نگاشت‌ها از بین می‌رود [۱۷].

۳-۴ کاهش بعد به روش ترکیبی

اگر فضای ویژگی‌ها را به صورت:

$$x_i \in R^N \quad (۱-۴)$$

تعیین کنیم، با اعمال روش ترکیب ویژگی‌ها، ویژگی‌های جدید را به صورت زیر ایجاد می‌کنیم:

$$y = f(x) : R^N \rightarrow R^M \quad M < N \quad (۲-۴)$$

به گونه‌ای که بردار ویژگی‌های نگاشت شده به فضای R^M اطلاعات و ساختمان ویژگی‌ها در فضای R^N را نگه دارد.

در حالت کلی، نگاشت بهینه $y = f(x)$ به صورت یک تابع غیرخطی خواهد بود. با توجه به

اینکه روش‌های سیستماتیک برای نگاشت‌های غیرخطی وجود ندارد و انتخاب زیرمجموعه‌ای از نگاشت‌ها به مسئله وابسته است، روش‌های کاهش بعد ترکیبی به نگاشت‌های خطی محدود می‌شود:

$$y = Wx \quad (۳-۴)$$

در این حالت Y یک ترکیب خطی از x خواهد بود. برای انتخاب نگاشت خطی مناسب، معیارهای متفاوتی وجود دارد. یکی از این معیارها قابلیت نمایش اطلاعات موجود در سیگنال اصلی در ابعاد پایین تر است. روش PCA ^۱ از این معیار برای کاهش ابعاد ویژگی ها استفاده می کند [۱۸].

۴-۴ کاهش بعد به روش گزینش ویژگی ها

در این روش می خواهیم از مجموعه ی ویژگی های:

$$X = \{x_i \mid i = 1, \dots, N\} \quad (۴-۴)$$

دسته ویژگی های:

$$Y_M = \{x_{i1}, x_{i2}, \dots, x_{iM}\} \quad M < N \quad (۵-۴)$$

را به گونه ای انتخاب کنیم که معیار معینی (در حالت ایده آل احتمال طبقه بندی صحیح)، برای آن بهینه گردد. در این روش باید به دو نکته توجه نمود:

۱. معیار مناسب برای انتخاب ویژگی ها

۲. روش جستجوی مناسب برای انتخاب ویژگی ها

در ادامه ابتدا معیارهای مختلف را بررسی نموده و سپس روش های جستجوی مناسب را معرفی می کنیم.

۴-۴-۱ معیارهای انتخاب ویژگی

۴-۴-۱-۱ معیارهای بدون سرپرست^۲، غیر وابسته به طبقه بندی کننده

^۱ Principal Component Analysis

^۲ Unsupervised

این معیارها به نوع طبقه‌بندی‌کننده و کلاس وابسته به هر داده، بستگی ندارند. در این معیارها به میزان ارتباط و شباهت میان دو ویژگی متفاوت توجه می‌شود، به گونه‌ای که ویژگی‌ها تا حد ممکن از یکدیگر دور باشند و وابستگی کمی به یکدیگر داشته باشند. چرا که در صورت شباهت و وابستگی زیاد، عملاً یکی از آنها قابل حذف خواهد بود [۲]. تعدادی از این معیارها در جدول (۴-۱) معرفی شده‌اند.

جدول (۴-۱) : معیارهای بدون سرپرست غیروابسته به طبقه‌بندی‌کننده

معیارهای بدون سرپرست، غیروابسته به طبقه‌بندی‌کننده
معیار بر پایه‌ی همبستگی
معیار بر پایه‌ی بیشینه‌ی نمای اطلاعات فشرده شده
معیار بر پایه‌ی خطای تخمین خطی
معیار بر پایه‌ی اطلاعات متقابل

در این پروژه از معیارهای بدون سرپرست و غیروابسته به طبقه‌بندی‌کننده، برای انتخاب ویژگی‌های مناسب استفاده نشده است.

۴-۱-۲-۴ معیارهای با سرپرست^۱، غیر وابسته به طبقه‌بندی‌کننده

این دسته از معیارها به کلاسی که برای هر داده در نظر گرفته می‌شود بدون توجه به نوع طبقه‌بندی‌کننده، بستگی دارند. در واقع در این معیارها می‌خواهیم ویژگی‌ها را به گونه‌ای انتخاب کنیم که داده‌های مربوط به یک ویژگی (و یا یک دسته ویژگی) خاص، برای کلاس‌های مختلف تا حد ممکن از یکدیگر دور باشند [۲]. لیست تعدادی از این معیارها که در انتخاب ویژگی‌ها مؤثرند، در جدول (۴-۲) مشخص شده است.

^۱ Supervised

جدول (۴-۲): معیارهای با سرپرست غیروابسته به طبقه‌بندی‌کننده

معیارهای با سرپرست، غیروابسته به طبقه‌بندی‌کننده
معیار بر پایه‌ی واگرایی
معیار بر پایه‌ی حد چرنوف و فاصله‌ی باتاچاریا
معیار بر پایه‌ی ماتریس‌های پخشی
معیار بر پایه‌ی همبستگی
معیار بر پایه‌ی پیشینه‌ی نمای اطلاعات فشرده شده
معیار بر پایه‌ی خطای تخمین خطی
معیار بر پایه‌ی اطلاعات متقابل

در این پروژه از معیار بر پایه‌ی ماتریس‌های پخشی برای کاهش اولیه‌ی تعداد ویژگی‌ها استفاده

شده است که در ادامه توضیح داده خواهد شد.

❖ معیار بر پایه‌ی ماتریس‌های پخشی

اگر μ_i را میانگین (و یا ماتریس میانگین) ویژگی (و یا دسته ویژگی) خاص که به کلاس وابسته‌اند و μ_0 را میانگین (یا ماتریس میانگین) همه‌ی ویژگی‌ها بدون در نظر گرفتن کلاس آنها در نظر بگیریم، ماتریس‌های پخشی درون کلاسی و میان کلاسی به صورت زیر تعریف می‌شوند:

۱. ماتریس پخشی درون کلاسی:

$$S_W = \sum_{i=1}^M P(w_i) S_i, \quad S_i = E\{(x - \mu_i)(x - \mu_i)^T\} \quad (۴-۶)$$

ماتریس پخشی درون کلاسی به نوعی میزان فشرده بودن کلاس‌های مختلف بر پایه‌ی دسته ویژگی خاص مورد بررسی را نشان می‌دهد.

۲. ماتریس پخشی میان کلاسی:

$$S_b = \sum_{i=1}^M P(w_i) (\mu_i - \mu_0)(\mu_i - \mu_0)^T, \quad \mu_0 = \sum_{i=1}^M P(w_i) \mu_i \quad (۴-۷)$$

ماتریس پخش درون کلاسی به نوعی میزان پراکنده بودن کلاس‌های مختلف بر پایه‌ی دسته ویژگی‌های خاص را نشان می‌دهد.

با توجه به روابط داده شده سه معیار زیر را تعریف می‌کنیم:

$$J_1 = \frac{\text{trace}\{S_b\}}{\text{trace}\{S_w\}} \quad , \quad J_2 = \frac{|S_b|}{|S_w|} \quad , \quad J_3 = \text{trace}\{S_w^{-1} S_b\} \quad (۸-۴)$$

برای هر سه معیار تعریف شده، هر چه میزان آنها بزرگتر باشد نمایانگر آن خواهد بود که به طور میانگین کلاس‌های مختلف مورد بررسی بر پایه‌ی دسته ویژگی خاص از هم دور بوده و تک تک این کلاس‌ها نیز فشرده می‌باشند [۲].

۴-۱-۳-۴ معیارهای وابسته به طبقه‌بندی‌کننده

در این معیارها به نوع طبقه‌بندی‌کننده و اثری که هر دسته ویژگی در خروجی طبقه‌بندی‌کننده خواهد داشت، کار داریم. می‌توان چهار دسته معیار وابسته به طبقه‌بندی‌کننده تعریف کرد که در جدول (۳-۴) مشخص شده‌اند.

جدول (۳-۴): معیارهای وابسته به طبقه‌بندی‌کننده

معیارهای وابسته به طبقه‌بندی‌کننده
معیار بر پایه‌ی وابستگی میان ورودی و خروجی طبقه‌بندی‌کننده
معیار بر پایه‌ی درصد درستی طبقه‌بندی‌کننده
معیار بر پایه‌ی خطای طبقه‌بندی‌کننده
معیار بر پایه‌ی پارامترهای به دست آمده از آموزش طبقه‌بندی‌کننده

در این پروژه از معیار درصد درستی طبقه‌بندی‌کننده استفاده شده است که در ادامه توضیح داده خواهد شد.

❖ معیار بر پایه درصد درستی طبقه‌بندی‌کننده

به کمک این معیار می‌خواهیم به دست آوریم چند درصد از داده‌های ورودی با کمک یک طبقه‌بندی‌کننده‌ی خاص و با یک دسته ویژگی معین، به طور صحیح طبقه‌بندی می‌شوند. در واقع می‌خواهیم تأثیر یک دسته ویژگی را در درصد درستی طبقه‌بندی‌کننده تعیین نماییم [۲].

برای یافتن این دسته ویژگی که بهترین معیار بر پایه‌ی درصد درستی را داشته باشد، نیاز به یک جستجوی کامل میان تمام ویژگی‌ها می‌باشد که در بخش بعدی توضیح داده خواهد شد.

۴-۴-۲ روش‌های جستجو برای گزینش ویژگی‌ها

روش‌های جستجو برای یافتن دسته ویژگی‌های مناسب بر اساس معیارهای تعریف شده در بخش‌های قبلی را می‌توان به دو دسته‌ی کلی تقسیم نمود:

۱. روش جستجوی فیلتری^۱

۲. روش جستجوی پوششی^۲

که هر کدام از آنها در بخش‌های بعدی توضیح داده خواهند شد.

۴-۴-۱ روش جستجوی فیلتری

در این روش معیارها که معمولاً بر پایه‌ی محتوای اطلاعاتی داده‌ها هستند، بر روی تک تک ویژگی‌ها اعمال شده و ویژگی‌هایی که بهترین معیار را دارند، انتخاب می‌شوند. در واقع در این روش

^۱ Filter Method

^۲ Wrapper Method

ویژگی‌هایی انتخاب می‌شوند که اگر به تنهایی استفاده شوند، بهترین نتیجه را خواهند داد. معیارهای به کار رفته در این روش معمولاً معیارهای غیروابسته به طبقه‌بندی‌کننده می‌باشند [۲]، [۱۷].

۴-۲-۲ روش جستجوی پوششی

در این روش که معیارها معمولاً معیارهای وابسته به طبقه‌بندی‌کننده هستند، دسته ویژگی‌های مناسب با توجه به اثر آنها بر روی یکدیگر انتخاب می‌شوند. بنابراین در این روش لازم است که یک جستجوی کامل که حالت‌های کنار هم قرار گرفتن ویژگی‌ها را نیز پوشش دهد، برای یافتن دسته ویژگی‌ای که بهترین معیار را دارد، انجام شود [۲]، [۱۷].

روش‌های جستجو را می‌توان به سه دسته‌ی اصلی تقسیم نمود:

۱. الگوریتم‌های جستجوی نمایی^۱

در این روش‌ها معیار برای تعداد زیادی از زیرمجموعه‌ها که به صورت نمایی رشد می‌کنند، محاسبه می‌شود. این روش‌ها نتایج نسبتاً خوبی خواهند داشت، اما زمان‌بر هستند [۱۷]. جدول (۴-۴) تعدادی از الگوریتم‌های معروف جستجوی نمایی را نشان می‌دهد.

جدول (۴-۴): الگوریتم‌های جستجوی نمایی

الگوریتم‌های جستجوی نمایی	Exponential Algorithms	استفاده شده در آزمایش‌ها
جستجوی سرتاسری	<i>Exhaustive Search</i>	
روش شاخه و حد	<i>Branch and Bound</i>	
روش یکنواخت تقریبی با شاخه و حد	<i>Approximate Monotonicity with Branch and Bound</i>	
جستجوی پرتوی	<i>Beam Search</i>	✓

^۱ Exponential Algorithms

۲. الگوریتم‌های جستجوی پی‌درپی^۱

این روش‌ها ویژگی‌ها را به طور پی‌درپی اضافه کرده و یا خارج می‌کنند تا به یک دسته ویژگی با معیار مناسب برسند. این روش‌ها ممکن است به جای رسیدن به جواب بهینه، به یک مینیموم محلی ختم شوند [۱۷]. نمونه‌هایی از این روش‌ها در جدول (۵-۴) مشخص شده‌اند.

جدول (۵-۴): الگوریتم‌های جستجوی پی‌درپی

الگوریتم‌های جستجوی پی‌درپی	Sequential Algorithms	استفاده شده در آزمایش‌ها
جستجوی پیشروی پی‌درپی	<i>Sequential Forward Selection</i>	✓
جستجوی پسروی پی‌درپی	<i>Sequential Backward Selection</i>	
روش L -اضافه و R -کاهش	<i>Plus-L Minus-R Selection</i>	
جستجوی دوطرفه	<i>Bidirectional</i>	
جستجوی شناور پی‌درپی	<i>Sequential Floating Selection</i>	✓

۳. روش‌های جستجوی تصادفی^۲

این روش‌ها با انتخاب تصادفی دسته ویژگی‌ها و محاسبه‌ی معیار مورد نظر برای آنها، برای جلوگیری از خاتمه‌ی الگوریتم با یک مینیموم محلی، عمل می‌نمایند [۱۷]. سه نمونه از این روش‌ها در جدول (۶-۴) مشخص شده‌اند.

جدول (۶-۴): الگوریتم‌های جستجوی تصادفی

الگوریتم‌های جستجوی تصادفی	Randomized Algorithms	استفاده شده در آزمایش‌ها
تولید رندومی و انتخاب پی‌درپی	<i>Random Generation plus Sequential Selection</i>	
ذوب و انجماد شبیه‌سازی شده	<i>Simulated Annealing</i>	
الگوریتم ژنتیک	<i>Genetic Algorithm</i>	✓

^۱ Sequential Algorithms

^۲ Randomized Algorithms

در ادامه توضیح مختصری در مورد روش‌های جستجو که در این پروژه از آنها استفاده شده است، داده می‌شود.

۴-۲-۱-۲-۱ جستجوی پیشروی پی‌درپی^۱

این روش، که یکی از آسان‌ترین الگوریتم‌های جستجو می‌باشد، به صورت زیر عمل می‌نماید:

۱. از مجموعه‌ی تهی شروع می‌کنیم.

۲. ویژگی‌ای را به مجموعه ویژگی‌های فعلی اضافه می‌کنیم که به کمک آن بهترین معیار به دست آید.

۳. اگر به یک شرط نهایی (مثلاً تعداد ویژگی‌ها) برای اتمام الگوریتم نرسیده‌ایم، به مرحله ۲ بازگشته و در غیر این صورت الگوریتم پایان می‌یابد [۲]، [۱۷].

این روش زمانی خوب عمل می‌کند که تعداد اعضای زیرمجموعه‌ی نهایی نسبت به تعداد ویژگی‌های اولیه کم باشد. اشکال این روش این است که پس از اضافه شدن یک عضو به زیرمجموعه، امکان خروج آن وجود ندارد و همیشه جزء ویژگی‌های گزیده شده در نظر گرفته می‌شود [۲]، [۱۷].

۴-۲-۲-۲-۱ جستجوی شناور پی‌درپی^۲

نام‌گذاری این روش بر این اساس است که تعداد اعضای زیرمجموعه‌ی انتخاب شده در هر مرحله به صورت شناور می‌تواند کم یا زیاد شود.

^۱ Sequential Forward Selection

^۲ Sequential Floating Selection

الگوریتم جستجوی پیشروی شناور را می‌توان به صورت زیر خلاصه نمود:

۱. از مجموعه تهی شروع می‌کنیم.
۲. ویژگی‌ای را به مجموعه ویژگی‌های فعلی اضافه می‌کنیم که به کمک آن بهترین معیار به دست آید.
۳. ویژگی‌ای را انتخاب می‌کنیم که با حذف آن کمترین اثر بد بر روی معیار گذاشته شود. اگر این ویژگی، همان ویژگی اضافه شده در مرحله ۲ بود، تغییری در مجموعه نداده و به مرحله ۲ بازمی‌گردیم. در غیر این صورت ویژگی را از مجموعه حذف نموده و به مرحله ۴ می‌رویم.
۴. ویژگی‌ای را انتخاب می‌کنیم که با حذف آن کمترین اثر بد بر روی معیار گذاشته شود. اگر معیار به دست آمده برای دسته ویژگی‌ها با حذف این ویژگی از معیار به دست آمده قبل از حذف آن بهتر بود، ویژگی را حذف نموده و مرحله ۴ را تکرار می‌کنیم. در غیر این صورت به مرحله ۵ می‌رویم.
۵. اگر به یک شرط نهایی (مثلاً تعداد ویژگی‌ها) برای اتمام الگوریتم نرسیده‌ایم، به مرحله ۲ بازگشته و در غیر این صورت الگوریتم پایان می‌یابد [۲]، [۱۷].

۴-۲-۳ جستجوی پرتوی^۱

الگوریتم جستجوی پرتوی را می‌توان به صورت زیر خلاصه نمود:

۱. از مجموعه‌ی تهی شروع می‌کنیم.
۲. تمام زیر مجموعه‌های تک عضوی از مجموعه‌ی کل ویژگی‌ها را در نظر گرفته و n تا از آنها که بهترین معیار را می‌دهند، انتخاب می‌کنیم و در صف A قرار می‌دهیم. c را برابر ۱ قرار می‌دهیم.

^۱ Beam Search

۳. تمام زیرمجموعه‌های $c+1$ عضوی را که از اضافه نمودن یک ویژگی به یکی از زیرمجموعه‌های انتخاب شده در صف A به دست می‌آیند، در نظر گرفته و n تا از آنها که بهترین معیار را می‌دهند، انتخاب می‌کنیم و در صف A قرار می‌دهیم. به c ، ۱ واحد اضافه می‌کنیم. در این حالت A شامل n تا از زیرمجموعه‌های c عضوی که بهترین معیار را دارند، می‌باشد.

۴. اگر $c = m$ باشد (تعداد ویژگی‌های نهایی)، دسته ویژگی‌های گزینش شده برابر یکی از n زیرمجموعه‌ی موجود در A که بهترین معیار را دارد، خواهد بود. در غیر این صورت به مرحله‌ی ۳ بازمی‌گردیم.

در این الگوریتم اگر بر روی مقدار n محدودیتی قرار داده نشود، این روش همانند جستجوی سرتاسری خواهد شد [۲]، [۱۹].

۴-۲-۲-۴-۴ جستجوی ژنتیک^۱

در جستجوی ژنتیک برای معین نمودن هر زیرمجموعه از مجموعه‌ی کل ویژگی‌ها از یک رشته‌ی صفر و یک به طول تعداد کل ویژگی‌ها استفاده می‌شود که می‌توان آن را به صورت زیر نمایش داد:

$$X_i = [t_1, t_2, \dots, t_n] \quad t_j = 1 \text{ or } 0 \quad (۹-۴)$$

به این ترتیب که ۱ بودن t_j معادل وجود ویژگی j ام در زیرمجموعه است.

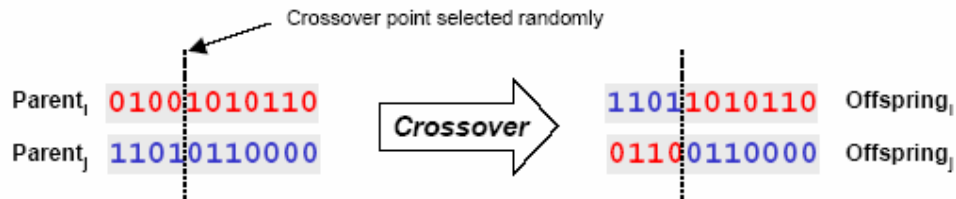
دو عملگر اصلی در جستجوی ژنتیک به صورت زیر تعریف می‌شوند:

❖ عملگر هم‌گذاری^۲:

^۱ Genetic Algorithm

^۲ Cross-Over

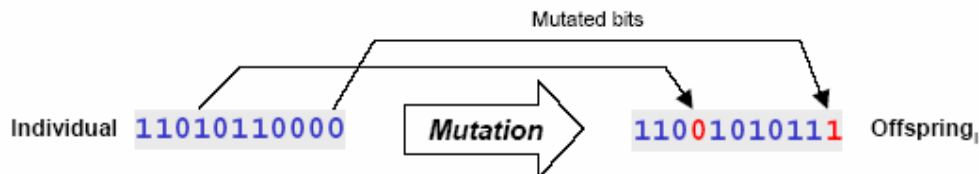
این عملگر با عمل بر روی رشته‌های دو زیرمجموعه از نسل مادر^۱ و جابه‌جا نمودن قسمتی از رشته‌ی صفر و یک مربوط به آنها، دو زیرمجموعه از نسل جدید^۲ را ایجاد می‌کند. روش تولید نسل جدید توسط این عملگر در شکل (۳-۴) نشان داده شده است.



شکل (۳-۴): تولید نسل جدید توسط عملگر همگذاری [۱۹]

❖ عملگر جهش^۳:

این عملگر با تغییر یک بیت از رشته‌ی یکی از زیرمجموعه‌های نسل مادر به صورت تصادفی، یک زیرمجموعه از نسل جدید تولید می‌نماید. روش تولید نسل جدید توسط این عملگر در شکل (۴-۴) نشان داده شده است.



شکل (۴-۴): تولید نسل جدید توسط عملگر جهش [۱۹]

حال با توجه به تعاریف داده شده، الگوریتم ژنتیک را می‌توان به طور مختصر به صورت زیر

بیان نمود:

۱. یک مجموعه از X_i ها به صورت تصادفی به عنوان نسل مادر انتخاب می‌کنیم.
۲. برای یافتن بهترین‌های نسل مادر، معیار مورد نظر را بر روی آنها اعمال می‌کنیم.

^۱ Parents

^۲ Offsprings

^۳ Mutation

۳. مراحل بعدی را تا زمانی که شرط نهایی برآورده شود تکرار می‌نماییم:

۳-۱. بهترین‌های نسل مادر را انتخاب می‌کنیم.

۳-۲. عملگر همگذاری را بر روی X_i های انتخاب شده برای تولید نسل نو اعمال می-

نماییم.

۳-۳. عملگر جهش را بر روی X_i های انتخاب شده برای تولید نسل جدید اعمال می-

نماییم.

۳-۴. مجموعه‌ی جدید را از اجتماع نسل مادر و نسل جدید، تولید می‌نماییم و به جای

نسل مادر قرار می‌دهیم.

۳-۵. معیار مورد نظر را بر روی مجموعه‌ی جدید بررسی می‌کنیم.

۴. بهترین عضو مجموعه، نمایانگر دسته ویژگی‌هایی خواهد بود که می‌خواهیم برگزینیم.

شرط نهایی برای خاتمه‌ی الگوریتم می‌تواند محدود کردن تعداد مراحل ایجاد نسل نو و یا

بررسی تکرار شدن فروان یکی از X_i ها در همه‌ی نسل‌ها باشد [۲]، [۱۹]، [۲۰].

۴-۵ جمع‌بندی

در این فصل دلایل نیاز به کاهش تعداد ویژگی‌ها بیان شدند و روش‌های مختلف کاهش بعد

(ترکیبی و گزینشی) معرفی گشتند. سپس معیارهای مختلف برای گزینش ویژگی‌ها توضیح داده شدند.

این معیارها به سه دسته‌ی اصلی (۱) معیارهای بدون سرپرست غیروابسته به طبقه‌بندی‌کننده، (۲) معیارهای

با سرپرست غیروابسته به طبقه‌بندی‌کننده و (۳) معیارهای وابسته به طبقه‌بندی‌کننده تقسیم شدند. در هر

دسته توضیح مختصری در مورد معیارهایی که در این پروژه از آنها استفاده شده است، داده شد. در پایان نیز روش‌های مختلف برای جستجو میان تمام ویژگی‌ها برای یافتن دسته ویژگی بهینه (با معیار بهینه) بررسی گشتند. این روش‌ها به سه دسته‌ی ۱) الگوریتم‌های جستجوی نمایی، ۲) الگوریتم‌های جستجوی پی‌درپی و ۳) الگوریتم‌های تصادفی تقسیم شده و الگوریتم‌های استفاده شده در پروژه توضیح داده شدند. در فصل بعد نحوه‌ی انجام آزمایش‌ها و نتایج به دست آمده از آنها بررسی می‌شوند و الگوریتم مناسبی برای به دست آوردن نتیجه‌ی مطلوب معرفی می‌گردد.

فصل پنجم : آزمایش‌ها

۵-۱ مقدمه

در این فصل ابتدا داده‌های مورد آزمایش را معرفی می‌کنیم. سپس ۴ الگوریتم مختلف برای انتخاب ویژگی‌ها معرفی می‌شوند. هر یک از این الگوریتم‌ها بر روی داده‌ها آزمایش شده و نتایج برای طبقه‌بندی‌کننده‌های مختلف تعیین می‌گردند. برای هر کدام از این ۴ روش، نتایج به دست آمده برای طبقه‌بندی‌کننده‌های مختلف، مقایسه می‌شوند. در پایان فصل نتایج به دست آمده توسط این ۴ الگوریتم مقایسه شده و مورد ارزیابی قرار می‌گیرند و الگوریتم کلی برای طبقه‌بندی داده‌ها ارائه می‌گردد.

۲-۵ داده‌های مورد آزمایش

داده‌های آزمایش شده، داده‌های دسته‌ی چهارم مسابقات *BCI2003* می‌باشند که توسط گروه برلین تهیه شده‌اند. این دسته داده‌ها از یک شخص عادی در حالت بدون فیدبک^۱ ثبت شده است. شخص مورد آزمایش بر روی یک صندلی عادی قرار گرفته است. دست‌ها بر روی میز تکیه داده شده و انگشت‌های او در موقعیت استاندارد برای تایپ بر روی کیبورد کامپیوتر قرار گرفته‌اند. وظیفه شخص فشردن ۴ کلید مشخص شده بر روی صفحه کلید با انگشتان کوچک یا اشاره دست راست یا چپ به خواست خود و با زمان بندی دلخواه است^۲ [۲۱]، [۲۲].

آزمایش از سه نوبت^۳ ۶ دقیقه‌ای تشکیل شده است. تمام نوبت‌ها در یک روز با چند دقیقه استراحت بین آنها انجام شده‌اند. عمل تایپ با سرعت متوسط یک کلید در هر ثانیه انجام می‌شود. داده‌ها به صورت ۴۱۶ دوره به طول ۵۰۰ میلی ثانیه هستند که ۱۳۰ میلی ثانیه قبل از فشردن کلید تمام می‌شوند. فرکانس نمونه برداری ۱۰۰۰ هرتز بوده و داده‌ها از فیلتر میان گذر ۰,۰۵ تا ۲۰۰ هرتز گذرانده شده‌اند [۲۱]، [۲۲].

داده‌ها با استفاده از تقویت کننده‌ی *NeuroScan* و الکترودهای *Ag/AgCl* ثبت شده‌اند. از ۲۸ کانال اندازه‌گیری *EEG* که در موقعیت استاندارد جهانی سیستم ۱۰-۲۰^۴ قرار گرفته‌اند، برای ثبت داده‌ها استفاده شده است [۲۱].

برای سادگی، کانال‌ها را به صورت نشان داده شده در جدول (۱-۵) شماره گذاری می‌کنیم.

^۱ *No-feedback*

^۲ *Self-paced key typing*

^۳ *Session*

^۴ *International 10 – 20 System*

جدول (۵-۱): شماره گذاری کانال‌های ثبت EEG

Number	1	2	3	4	5	6	7
Channel	F3	F1	Fz	F2	F4	FC5	FC3
Number	8	9	10	11	12	13	14
Channel	FC1	FCz	FC2	FC4	FC6	C5	C3
Number	15	16	17	18	19	20	21
Channel	C1	Cz	C2	C4	C6	CP5	CP3
Number	22	23	24	25	26	27	28
Channel	CP1	CPz	CP2	CP4	CP6	O1	O2

از این تعداد داده، ۳۱۶ دوره طبقه‌بندی شده و کلاس مربوط به آنها مشخص شده‌است (0: حرکت دست چپ، 1: حرکت دست راست). ۱۰۰ دوره‌ی دیگر به عنوان داده‌های تست داده شده‌اند.

۳-۵ پیش پردازش^۱ داده‌ها

با توجه به محدودیت باند فرکانسی سیگنال‌های مغزی و به منظور حذف نویزهای احتمالی (مانند برق شهر) ابتدا داده‌ها را از یک فیلتر پایین گذر ۴۵ هرتز می‌گذرانیم. برای پیاده‌سازی این فیلتر ابتدا داده‌ها را از حوزه‌ی زمان به حوزه‌ی فرکانس برده و پس از حذف فرکانس‌های بالا به حوزه‌ی زمان باز می‌گردانیم. سپس برای افزایش سرعت پردازش، نرخ نمونه برداری را کاهش داده و به ۱۰۰ هرتز می‌رسانیم.

^۱ Preprocessing

۵-۴ استخراج ویژگی‌ها^۱

پس از پیش پردازش داده‌ها، ویژگی‌ها را در ۷ دسته‌ی مختلف استخراج می‌کنیم:

دسته‌ی اول: ویژگی‌های آماری

- میانگین آماری برای داده‌های هر کانال
- واریانس آماری برای داده‌های هر کانال
- ضریب همبستگی بین داده‌های هر دو کانال
- مومنتوم بین داده‌های هر دو کانال
- فرم فاکتور برای داده‌های هر کانال

دسته‌ی دوم: ویژگی‌های مبتنی بر آنتروپی

- آنتروپی شنن (Shannon) برای داده‌های هر کانال
- آنتروپی رنیه (Renyi) برای داده‌های هر کانال ($\alpha = -5, -2, -1, 0.5, 1.5, 2, 3, 5$)
- آنتروپی تسالیس (Tsallis) برای داده‌های هر کانال ($q = -5, -2, -1, 0.5, 1.5, 2, 3, 5$)
- آنتروپی تقریبی برای داده‌های هر کانال
- معیار پیچیدگی Lample – Ziv برای داده‌های هر کانال

دسته‌ی سوم: ویژگی‌های وابسته به تبدیل Wavelet

^۱ Feature Extraction

- ضرایب تبدیل *Wavelet* (*haar, db2, db3, db4, db5, mexh*) برای داده‌های هر

کانال

دسته‌ی چهارم: ویژگی‌های وابسته به ضرایب پارامتری

- ضرایب مدل *AR* (*Auto Regressive*) مرتبه‌ی ۴، ۸، ۱۶ و ۳۲ برای داده‌های هر

کانال

دسته‌ی پنجم: ویژگی‌های وابسته به تبدیل‌های فرکانسی

- ضرایب تبدیل گسسته‌ی سینوسی (*DST*) برای داده‌های هر کانال

- ضرایب تبدیل گسسته‌ی کسینوسی (*DCT*) برای داده‌های هر کانال

دسته‌ی ششم: ویژگی‌های وابسته به انرژی سیگنال در باندهای ویژه‌ی فرکانسی

(برای استخراج ویژگی‌های این دسته، از داده‌های خام قبل از پیش پردازش استفاده می‌کنیم.)

- انرژی سیگنال در ۴ باند فرکانسی:

$$[\delta(0-3.5\text{Hz}), \theta(3.5-7.5\text{Hz}), \alpha(7.5-13.5\text{Hz}), \beta(13.5-20\text{Hz})]$$

- انرژی سیگنال در ۷ باند فرکانسی:

$$[\delta(0.5-3.5\text{Hz}), \theta(3.5-7\text{Hz}), \alpha(7.5-13\text{Hz}), \beta_1(13-15\text{Hz}), \\ \beta_2(15-17\text{Hz}), \beta_3(18-25\text{Hz}), \beta_4(25.5-30\text{Hz})]$$

- انرژی سیگنال در ۸ باند فرکانسی:

$$[\delta(1.5-6\text{Hz}), \theta(6-8\text{Hz}), \alpha_1(8.5-10\text{Hz}), \alpha_2(10.5-12\text{Hz}), \\ \beta_1(12.5-18\text{Hz}), \beta_2(18.5-21\text{Hz}), \beta_3(21-30\text{Hz}), \beta_4(30-40\text{Hz})]$$

- انرژی سیگنال در ۱۰ باند فرکانسی:

$[(0-10\text{Hz}), (10-20\text{Hz}), (20-30\text{Hz}), (30-40\text{Hz}), (40-50\text{Hz}),$
 $(50-60\text{Hz}), (60-70\text{Hz}), (70-80\text{Hz}), (80-90\text{Hz}), (90-100\text{Hz})]$

دسته‌ی هفتم: ویژگی‌های آماری مربوط به تفاضل داده‌های کانال‌های متقارن

(برای استخراج ویژگی‌های این دسته، ابتدا ۱۱ سری داده را به صورت تفاضل داده‌های کانال‌های متقارن

ایجاد می‌کنیم. برای سادگی کانال‌های تفاضلی جدید را به صورت نمایش داده شده در جدول (۲-۵)

شماره گذاری می‌کنیم:

جدول (۲-۵): شماره‌گذاری کانال‌های تفاضلی تعریف شده

Number	1	2	3	4	5	6
Channel	F1-F2	F3-F4	FC1-FC2	FC3-FC4	FC5-FC6	C1-C2
Number	7	8	9	10	11	
Channel	C3-C4	C5-C6	CP1-CP2	CP3-CP4	CP5-CP6	

سپس برای داده‌های کانال‌های تفاضلی ویژگی‌های آماری را محاسبه می‌کنیم. (

- میانگین آماری برای داده‌های کانال‌های تفاضلی
- واریانس آماری برای داده‌های کانال‌های تفاضلی
- ضریب همبستگی بین داده‌های هردو کانال تفاضلی
- فرم فاکتور برای داده‌های هر یک از کانال‌های تفاضلی

پس از استخراج ویژگی‌ها، برای افزایش دقت طبقه‌بندی‌کننده‌ها ویژگی‌ها را نرمال می‌کنیم.

۵-۵ الگوریتم‌های کاهش ویژگی‌ها

۵-۵-۱ روش گزینش ویژگی‌ها با تعداد ویژگی ثابت

۵-۱-۱ الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت

در این روش برای انتخاب ویژگی‌های مؤثر برای هر طبقه‌بندی‌کننده، به صورت زیر عمل

می‌کنیم:

۱. با توجه به حجم زیاد ویژگی‌ها در هر دسته با استفاده از یکی از معیارهای با سرپرست

غیروابسته به طبقه‌بندی‌کننده، تعداد ویژگی‌ها را کاهش می‌دهیم.

❖ در آزمایش‌های انجام شده از معیار وابسته به ماتریس پراکندگی (*Scatter Matrix*)

استفاده شده است.

❖ از هر دسته ۵۰ ویژگی که بهترین معیار را دارند، انتخاب شده است.

۲. داده‌های آموزش (*training data*) را به دو قسمت تقسیم می‌کنیم. دو سوم از داده‌ها را به

عنوان داده‌های آموزشی - آموزشی (*train-train*) و بقیه داده‌ها را به عنوان داده‌های

آموزشی - تست (*train-test*) در نظر می‌گیریم. برای هر طبقه‌بندی‌کننده، با استفاده از

روش‌های جستجو، در هر دسته از ویژگی‌ها مجموعه‌ای از ویژگی‌ها را که خطای

طبقه‌بندی‌کننده را کمینه می‌کند، تعیین می‌کنیم. برای به دست آوردن خطای طبقه‌بندی‌کننده،

به ازای هر مجموعه از ویژگی‌ها، طبقه‌بندی‌کننده را به کمک داده‌های آموزشی - آموزشی،

آموزش داده و با داده‌های آموزشی - تست، امتحان می‌کنیم. تعداد ویژگی‌های مطلوب را در

این روش ثابت در نظر می‌گیریم.

❖ در آزمایش‌های انجام شده، ۲۱۰ سری داده‌ی اول را به عنوان داده‌های آموزشی-آموزشی

در نظر گرفته و ۱۰۶ سری داده‌ی دیگر را، داده‌های آموزشی- تست قرار می‌دهیم.

❖ تعداد ویژگی‌های مطلوب را برای هر یک از ۷ دسته‌ی تعریف شده، ۲۵ ویژگی در نظر

می‌گیریم.

❖ برای یافتن مجموعه ویژگی‌های مطلوب، از روش جستجوی شناور پی‌درپی

(*Sequential Floating Search*) یا روش جستجوی پرتوی (*Beam Search*)

استفاده می‌کنیم.

۳. پس از تعیین مجموعه ویژگی‌های تعیین شده برای تمام دسته‌ها، مجموعه‌ی به دست آمده را

به طبقه‌بندی‌کننده داده و به روش مشابه قسمت ۲ (تعیین داده‌های آموزشی- آموزشی و

آموزشی- تست و محاسبه خطای طبقه‌بندی‌کننده)، با کمینه کردن خطا، مجموعه ویژگی‌های

مطلوب را تعیین می‌کنیم. تعداد ویژگی‌های مطلوب در این مرحله نیز ثابت است.

❖ تعداد ویژگی‌های مطلوب ۲۵ در نظر گرفته شده است.

۴. ویژگی‌های مطلوب تعیین شده را از داده‌های آموزشی و داده‌های تست استخراج کرده و پس

از آموزش طبقه‌بندی‌کننده با همه داده‌های آموزشی (آموزشی- آموزشی و آموزشی- تست)،

داده‌های تست را طبقه‌بندی می‌کنیم.

نتایج به دست آمده برای ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic* و *Mahalanobis* در جدول (۳-۵)، نشان داده شده‌اند.

جدول (۳-۵): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

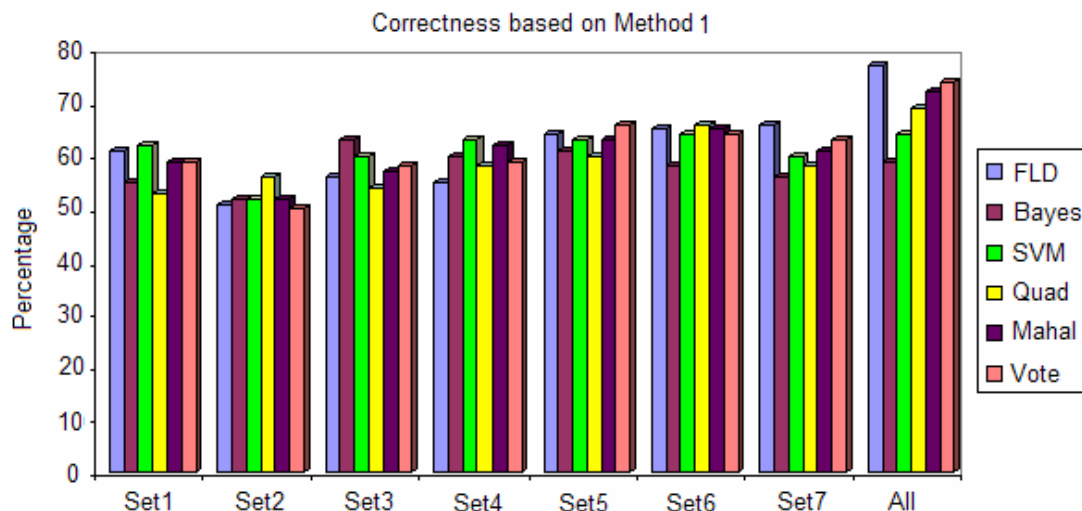
با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All(25)	Best
FLD	61	51	56	55	64	65	66	77	77
Bayes	55	52	63	60	61	58	56	59	63
SVM	62	52	60	63	63	64	60	64	64
Quad	53	56	54	58	60	66	58	69	69
Mahal	59	52	57	62	63	65	61	72	72
Best	62	56	63	62	64	66	66	77	
Vote	59	50	58	59	66	64	63	74	

■ Best result of all classifiers
 ■ Vote of all classifiers
 ■ Best result of each row

ابتدا برای هر کدام از طبقه‌بندی‌کننده‌ها الگوریتم بالا را اجرا نموده و درصد درستی را محاسبه می‌کنیم. سپس به ازای هر کدام از ۱۰۰ داده‌ی تست، کلاس مورد نظر را با رأی‌گیری میان ۵ طبقه‌بندی‌کننده معین می‌کنیم، به این صورت که کلاسی را که حداقل ۳ طبقه‌بندی‌کننده تعیین کرده‌اند، به آن نسبت می‌دهیم. درصد درستی نتایج به دست آمده با این روش در سطر *Vote* در جدول (۳-۵) نشان داده شده‌اند. در سطر *Best* در جدول (۳-۵)، بهترین نتیجه از بین نتایج به دست آمده از ۵ طبقه‌بندی‌کننده، نشان داده شده است.

نمودار درصد درستی به ازای هر کدام از طبقه‌بندی‌کننده‌ها و درصد درستی حاصل از رأی‌گیری میان آنها در شکل (۱-۵) نمایش داده شده است.



شکل (۵-۱): نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت

۵-۱-۵-۲ بررسی نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی با تعداد ویژگی ثابت

نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی با تعداد ویژگی ثابت، را می‌توان به

صورت زیر خلاصه نمود:

- به جز طبقه‌بندی‌کننده *Bayes*، برای بقیه‌ی طبقه‌بندی‌کننده‌ها بهترین درصد درستی به ازای دسته ویژگی‌های کلی (انتخاب شده از هر ۷ دسته ویژگی تعریف شده)، به دست آمده است. برای طبقه‌بندی‌کننده *Bayes*، بهترین نتیجه به ازای دسته ویژگی سوم (ویژگی‌های وابسته به ضرایب *Wavelet*) به دست آمده است.

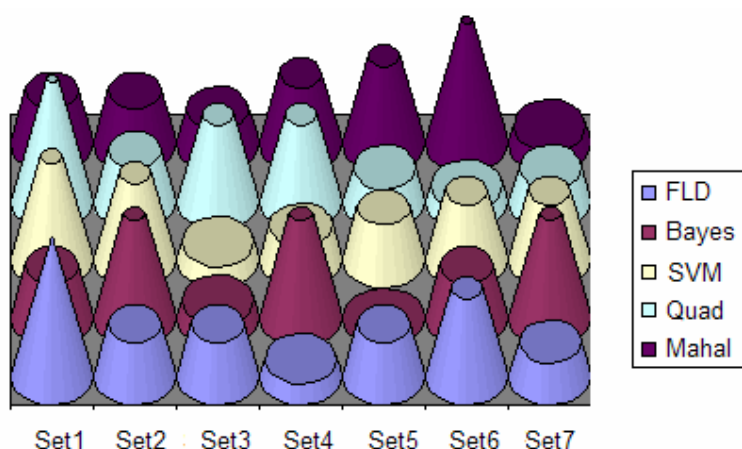
- برای ۷ دسته ویژگی بهترین نتیجه در بین ۵ طبقه‌بندی‌کننده متغیر بوده و نتیجه‌ی خاصی برای تمایز یکی از طبقه‌بندی‌کننده‌ها نسبت به بقیه، به دست نمی‌آید.

- برای دسته ویژگی کلی، بهترین نتیجه به ازای طبقه‌بندی‌کننده *FLD*، به دست آمده است.

- برای هر یک از ۷ دسته ویژگی به جز دسته پنجم، نتیجه‌ی به دست آمده از رأی‌گیری از بهترین نتیجه به دست آمده توسط ۵ طبقه‌بندی‌کننده کمتر است و در اکثر دسته‌ها در حدود میانگین قرار می‌گیرد.

- برای دسته ویژگی کلی، نتیجه‌ی به دست آمده از رأی‌گیری از بهترین نتیجه به دست آمده از ۵ طبقه‌بندی‌کننده (طبقه‌بندی‌کننده‌ی FLD ، 77%) کمتر است. اما این نتیجه از نتایج به دست آمده از بقیه طبقه‌بندی‌کننده‌ها بهتر است.

- برای تعیین توزیع هر یک از دسته ویژگی‌ها برای دسته ویژگی کلی، نسبت ویژگی‌های انتخاب شده از هر دسته را به تعداد کل ویژگی‌های انتخاب شده (۲۵ ویژگی) محاسبه می‌کنیم. نمودار توزیع هر یک از ۷ دسته ویژگی، به ازای طبقه‌بندی‌کننده‌های مختلف در شکل (۵-۲) نشان داده شده است.



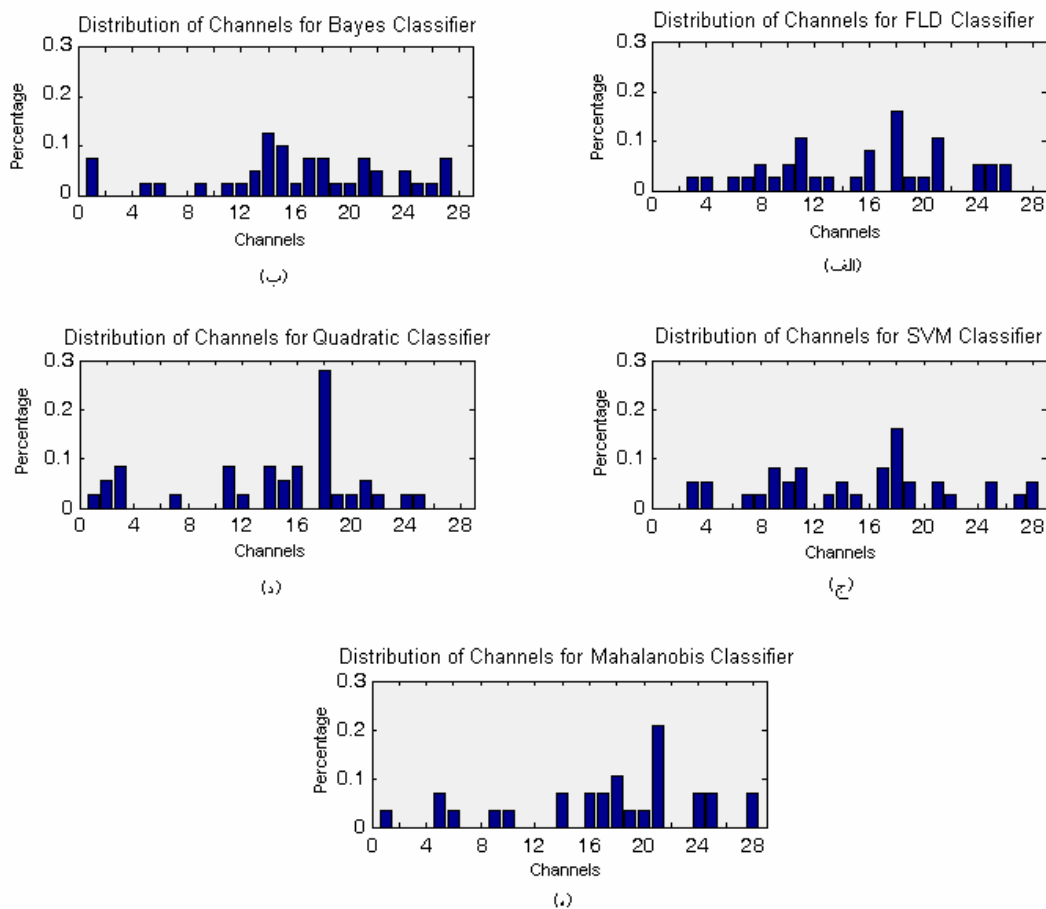
شکل (۵-۲): نمودار توزیع دسته ویژگی‌ها به ازای طبقه‌بندی‌کننده‌های مختلف

با اعمال الگوریتم گزینش ویژگی با تعداد ویژگی ثابت

برای طبقه‌بندی‌کننده‌های FLD ، SVM و $Quadratic$ دسته ویژگی اول (ویژگی‌های آماری) بیشترین توزیع را داشته است. برای طبقه‌بندی‌کننده‌ی $Bayes$ ، دسته ویژگی‌های دوم (ویژگی‌های مبتنی بر آنتروپی)، چهارم (ویژگی‌های وابسته به ضرایب پارامتری) و هفتم (ویژگی‌های آماری مربوط به

تفاضل داده‌های کانال‌های متقارن) توزیع مشابهی داشته‌اند. برای طبقه‌بندی‌کننده‌ی *Mahalanobis*، دسته ویژگی ششم (ویژگی‌های وابسته به انرژی سیگنال در باندهای ویژه فرکانسی) بیشترین توزیع را داشته است.

- نمودار توزیع ۲۸ کانال مختلف برای دسته ویژگی‌های کلی به ازای طبقه‌بندی‌کننده‌های مختلف، در شکل (۳-۵) نشان داده شده است.



شکل (۳-۵): توزیع کانال‌ها برای دسته ویژگی کلی با اعمال الگوریتم گزینش ویژگی با تعداد ویژگی ثابت :

(الف) *FLD*، (ب) *Bayes*، (ج) *SVM*، (د) *Quadratic*، (ه) *Mahalanobis*

برای سه طبقه‌بندی‌کننده‌ی *FLD*، *SVM* و *Quadratic*، کانال شماره ۱۸ (کانال C4) بیشترین توزیع را داشته است.

۵-۵-۲ روش گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده

۵-۵-۱ الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده

در این روش برای انتخاب ویژگی‌های مؤثر برای هر طبقه‌بندی‌کننده، به صورت زیر عمل

می‌کنیم:

۱. با توجه به حجم زیاد ویژگی‌ها در هر دسته با استفاده از یکی از معیارهای با سرپرست غیر

وابسته به طبقه‌بندی‌کننده، تعداد ویژگی‌ها را کاهش می‌دهیم.

❖ در آزمایش‌های انجام شده از معیار وابسته به ماتریس پراکندگی (*Scatter Matrix*)

استفاده شده است.

❖ از هر دسته ۵۰ ویژگی که بهترین معیار را دارند انتخاب شده است.

۲. داده‌های آموزش (*trainning data*) را به دو قسمت تقسیم می‌کنیم. دو سوم از داده‌ها را به

عنوان داده‌های آموزشی - آموزشی (*train-train*) و بقیه‌ی داده‌ها را به عنوان داده‌های

آموزشی - تست (*train-test*) در نظر می‌گیریم. برای هر طبقه‌بندی‌کننده، با استفاده از

روش‌های جستجو، در هر دسته از ویژگی‌ها مجموعه‌ای از ویژگی‌ها را که خطای

طبقه‌بندی‌کننده را کمینه می‌کند، تعیین می‌کنیم. برای به دست آوردن خطای طبقه‌بندی‌کننده،

به ازای هر مجموعه از ویژگی‌ها، طبقه‌بندی‌کننده را به کمک داده‌های آموزشی - آموزشی،

آموزش داده و با داده‌های آموزشی - تست، امتحان می‌کنیم. تعداد ویژگی‌های مطلوب را در

این روش متغیر در نظر می‌گیریم، به گونه ای که خطای طبقه‌بندی‌کننده به ازای این مجموعه ویژگی‌ها از حد آستانه‌ی معینی کمتر گردد.

❖ در آزمایش‌های انجام شده، ۲۱۰ سری داده‌ی اول را به عنوان داده‌های آموزشی-آموزشی در نظر گرفته و ۱۰۶ سری داده‌ی دیگر را، داده‌های آموزشی- تست قرار می‌دهیم.

❖ تعداد ویژگی‌های مطلوب را برای هر یک از ۷ دسته‌ی تعریف شده، حداکثر ۳۵ ویژگی در نظر می‌گیریم.

❖ برای یافتن مجموعه ویژگی‌های مطلوب، از روش جستجوی ژنتیک (*Genetic Search*) یا روش جستجوی پرتوی (*Beam Search*) استفاده می‌کنیم.

۳. پس از تعیین مجموعه ویژگی‌های تعیین شده برای تمام دسته‌ها، مجموعه‌ی به دست آمده را به طبقه‌بندی‌کننده داده و به روش مشابه قسمت ۲ (تعیین داده‌های آموزشی- آموزشی و آموزشی- تست و محاسبه‌ی خطای طبقه‌بندی‌کننده) ، با کمینه کردن خطا، مجموعه ویژگی‌های مطلوب را تعیین می‌کنیم. تعداد ویژگی‌های مطلوب در این مرحله نیز متغیر است و با شرط حد پایین خطای طبقه‌بندی‌کننده تعیین می‌شود.

۴. ویژگی‌های مطلوب تعیین شده را از داده‌های آموزشی و داده‌های تست استخراج کرده و پس از آموزش طبقه‌بندی‌کننده با همه داده‌های آموزشی (آموزشی- آموزشی و آموزشی- تست)، داده‌های تست را طبقه‌بندی می‌کنیم.

نتایج به دست آمده برای ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic* و *Mahalanobis* در جدول (۵-۴)، نشان داده شده‌اند.

جدول (۵-۴): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

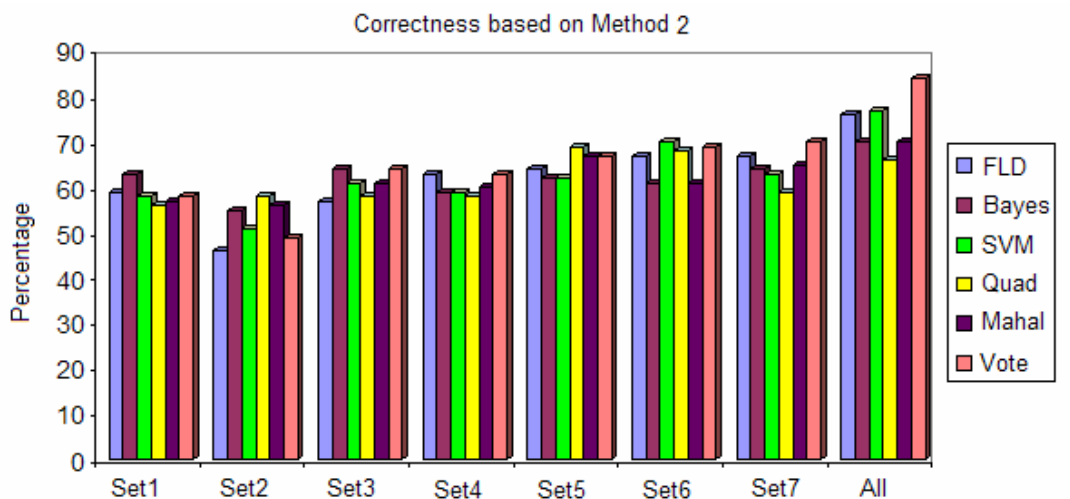
با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All	Best
FLD	59	46	57	63	64	67	67	76	76
Bayes	63	55	64	59	62	61	64	70	70
SVM	58	51	61	59	62	70	63	77	77
Quad	56	58	58	58	69	68	59	66	69
Mahal	57	56	61	60	67	61	65	70	70
Best	63	58	64	63	69	70	67	77	
Vote	58	49	64	63	67	69	70	84	

■ Best result of all classifiers
 ■ Vote of all classifiers
 ■ Best result of each row

پس از اعمال الگوریتم توضیح داده شده برای هر کدام از طبقه‌بندی‌کننده‌ها، به ازای هر کدام از ۱۰۰ داده‌ی تست، کلاس مورد نظر را با رأی‌گیری میان ۵ طبقه‌بندی‌کننده معین می‌کنیم، به این صورت که کلاسی را که حداقل ۳ طبقه‌بندی‌کننده تعیین کرده‌اند، به آن نسبت می‌دهیم. درصد درستی نتایج به دست آمده با این روش در سطر *Vote* در جدول (۵-۴) نشان داده شده‌اند. در سطر *Best* در جدول (۵-۴)، بهترین نتیجه از بین نتایج به دست آمده از ۵ طبقه‌بندی‌کننده، نشان داده شده است.

نمودار درصد درستی به ازای هر کدام از طبقه‌بندی‌کننده‌ها و درصد درستی حاصل از رأی‌گیری میان آنها در شکل (۵-۴) نمایش داده شده است.



شکل (۵-۴): نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده

۵-۲-۲ بررسی نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی با تعداد ویژگی وابسته به

طبقه‌بندی‌کننده

نتایج به دست آمده از اعمال الگوریتم گزینش ویژگی با تعداد ویژگی وابسته به طبقه‌بندی‌کننده،

را می‌توان به صورت زیر خلاصه نمود:

• به جز طبقه‌بندی‌کننده *Quadratic*، برای بقیه‌ی طبقه‌بندی‌کننده‌ها بهترین درصد درستی به ازای دسته

ویژگی‌های کلی (انتخاب شده از هر ۷ دسته ویژگی تعریف شده)، به دست آمده است. برای

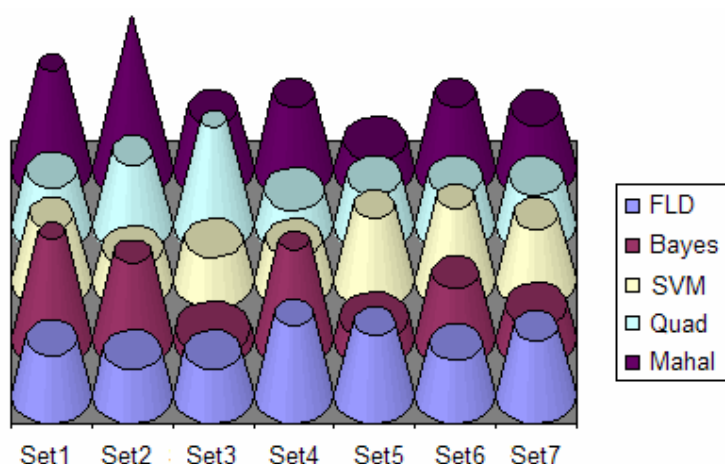
طبقه‌بندی‌کننده *Quadratic*، بهترین نتیجه به ازای دسته ویژگی پنجم (ویژگی‌های وابسته به

تبدیل‌های فرکانسی) به دست آمده است.

• برای ۷ دسته ویژگی بهترین نتیجه در بین ۵ طبقه‌بندی‌کننده متغیر بوده و نتیجه خاصی برای تمایز یکی

از طبقه‌بندی‌کننده‌ها نسبت به بقیه، به دست نمی‌آید.

- برای دسته ویژگی کلی، بهترین نتیجه به ازای طبقه‌بندی‌کننده‌ی SVM (۷۷٪)، به دست آمده است. نتیجه به دست آمده از طبقه‌بندی‌کننده‌ی FLD (۷۶٪) نیز، دومین نتیجه‌ی مطلوب است.
- برای هر یک از ۷ دسته ویژگی به جز دسته‌ی هفتم، درصد درستی به دست آمده از رأی‌گیری از بهترین نتیجه به دست آمده توسط ۵ طبقه‌بندی‌کننده کوچک‌تر یا مساوی است.
- برای دسته ویژگی کلی، نتیجه‌ی به دست آمده از رأی‌گیری (۸۴٪) از بهترین نتیجه‌ی به دست آمده از ۵ طبقه‌بندی‌کننده (۷۷٪) بیشتر است.
- برای تعیین توزیع هر یک از دسته ویژگی‌ها برای دسته ویژگی کلی، نسبت ویژگی‌های انتخاب شده از هر دسته را به تعداد کل ویژگی‌های انتخاب شده محاسبه می‌کنیم. نمودار توزیع هر یک از ۷ دسته ویژگی، به ازای طبقه‌بندی‌کننده‌های مختلف در شکل (۵-۵) نشان داده شده است.



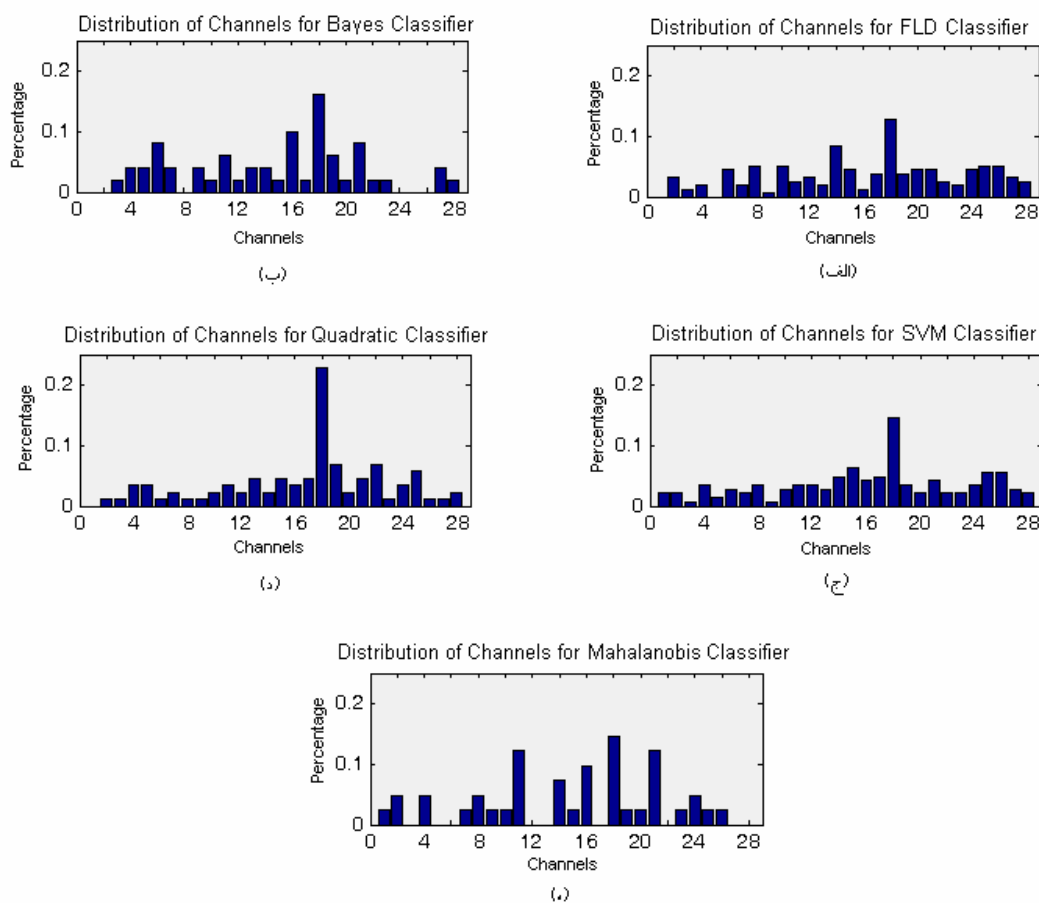
شکل (۵-۵): نمودار توزیع دسته ویژگی‌ها به ازای طبقه‌بندی‌کننده‌های مختلف

با اعمال الگوریتم گزینش ویژگی با تعداد ویژگی وابسته به طبقه‌بندی‌کننده

پراکندگی دسته ویژگی‌های انتخاب شده برای طبقه‌بندی‌کننده‌های مختلف، متفاوت است و نمی‌توان نتیجه‌ی معینی به دست آورد.

• نمودار توزیع ۲۸ کانال مختلف برای دسته ویژگی‌های کلی به ازای طبقه‌بندی‌کننده‌های مختلف، در

شکل (۵-۶) نشان داده شده است.



شکل (۵-۶): توزیع کانال‌ها برای دسته ویژگی کلی با اعمال الگوریتم گزینش ویژگی با تعداد ویژگی وابسته به طبقه‌بندی‌کننده:

(الف) FLD، (ب) Bayes، (ج) SVM، (د) Quadratic، (ه) Mahalanobis

برای هر ۵ طبقه‌بندی‌کننده کانال شماره ۱۸ (کانال C4) بیشترین توزیع را داشته است.

۵-۵-۳ روش ترکیب ویژگی‌ها با تعداد ویژگی ثابت

۵-۵-۳-۱ الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت

در این روش برای تولید ویژگی‌های ترکیبی برای هر طبقه‌بندی‌کننده، به صورت زیر عمل

می‌کنیم:

۱. با توجه به حجم زیاد ویژگی‌ها در هر دسته با استفاده از یکی از معیارهای با سرپرست غیر

وابسته به طبقه‌بندی‌کننده، تعداد ویژگی‌ها را کاهش می‌دهیم.

❖ در آزمایش‌های انجام شده از معیار وابسته به ماتریس پراکندگی (*Scatter Matrix*)

استفاده شده است.

❖ از هر دسته ۵۰ ویژگی که بهترین معیار را دارند انتخاب شده است.

۲. با اعمال روش *PCA*، ویژگی‌های دسته داده‌های آموزش را که در قسمت قبل انتخاب شده‌اند،

ترکیب کرده و تعداد ثابتی ویژگی ترکیبی تولید می‌کنیم.

❖ تعداد ویژگی‌های ترکیبی مطلوب را برای هر یک از ۷ دسته تعریف شده، ۲۵ ویژگی در

نظر می‌گیریم.

❖ ماتریس تبدیل به دست آمده در روش *PCA* برای داده‌های آموزش را بر داده‌های تست

اعمال کرده و ویژگی‌های ترکیبی مربوط به آنها را تعیین می‌کنیم.

۳. پس از تعیین مجموعه ویژگی‌های تعیین شده برای تمام دسته‌ها، روش *PCA* را بر روی

مجموعه به دست آمده اعمال کرده و ویژگی‌های ترکیبی نهایی (تعداد ثابت) را تولید می‌کنیم.

❖ تعداد ویژگی‌های ترکیبی مطلوب را در این مرحله نیز، ۲۵ ویژگی در نظر می‌گیریم.

❖ ماتریس تبدیل به دست آمده در روش *PCA* برای داده‌های آموزش را بر مجموعه

ویژگی به دست آمده از داده‌های تست در قسمت قبل اعمال کرده و ویژگی‌های ترکیبی

مربوط به آنها را تعیین می‌کنیم.

۴. طبقه‌بندی کننده را با ویژگی‌های ترکیبی به دست آمده از داده‌های آموزشی، آموزش داده و

داده‌های تست را به کمک آن طبقه‌بندی می‌کنیم.

نتایج به دست آمده برای ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic* و

Mahalanobis در جدول (۵-۵)، نشان داده شده‌اند.

جدول (۵-۵): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All(25)	Best
FLD	61	58	56	59	56	69	63	62	69
Bayes	56	52	58	58	59	56	62	61	62
SVM	60	56	63	61	59	71	66	64	71
Quad	56	52	58	58	59	56	62	61	62
Mahal	57	54	57	58	61	58	65	59	65
Best	61	58	63	61	61	71	66	64	
Vote	58	57	59	58	60	61	66	61	

■ Best result of all classifiers

■ Vote of all classifiers

■ Best result of each row

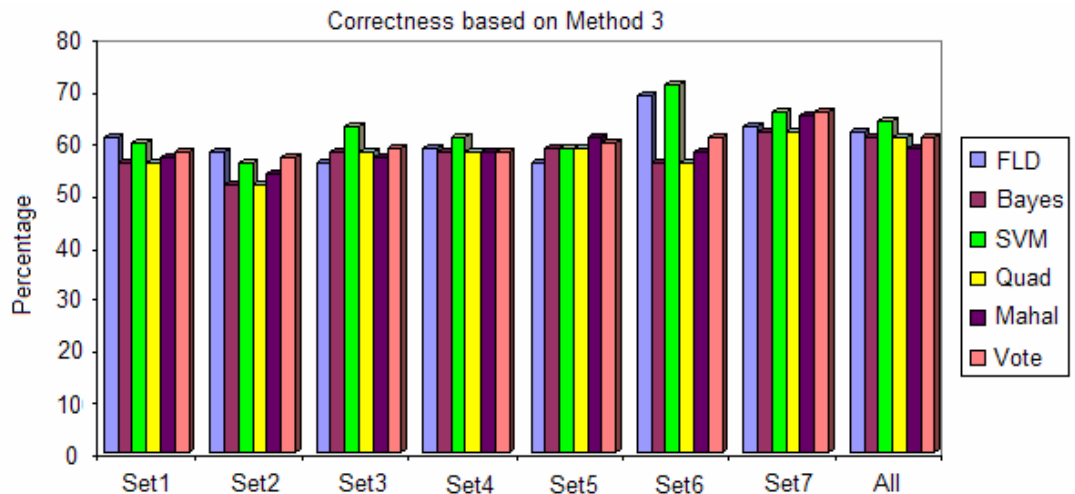
پس از اعمال الگوریتم توضیح داده شده برای هر کدام از طبقه‌بندی‌کننده‌ها، به ازای هر کدام از ۱۰۰

داده‌ی تست، کلاس مورد نظر را با رأی‌گیری میان ۵ طبقه‌بندی‌کننده معین می‌کنیم، به این صورت که

کلاسی را که حداقل ۳ طبقه‌بندی‌کننده تعیین کرده‌اند، به آن نسبت می‌دهیم. درصد درستی نتایج به دست

آمده با این روش در سطر *Vote* در جدول (۵-۵) نشان داده شده‌اند. در سطر *Best* در جدول (۵-۵)، بهترین نتیجه از بین نتایج به دست آمده از ۵ طبقه‌بندی‌کننده، نشان داده شده است.

نمودار درصد درستی به ازای هر کدام از طبقه‌بندی‌کننده‌ها و درصد درستی حاصل از رأی‌گیری میان آنها در شکل (۷-۵) نمایش داده شده است.



شکل (۷-۵): نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت

۵-۳-۲ بررسی نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت

نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی با تعداد ویژگی ثابت، را می‌توان به

صورت زیر خلاصه نمود:

- برای هیچ کدام از طبقه‌بندی‌کننده‌ها نتیجه‌ی به دست آمده برای دسته ویژگی کلی، بهترین نتیجه نیست.
- برای دسته ویژگی‌های سوم، چهارم، ششم و هفتم، نتیجه‌ی به دست آمده توسط طبقه‌بندی‌کننده *SVM*، بهترین نتیجه است. برای دسته ویژگی‌های اول و دوم، نتیجه به دست آمده توسط

طبقه‌بندی کننده‌ی *FLD* بهترین نتیجه بوده و برای دسته ویژگی پنجم بهترین نتیجه توسط طبقه‌بندی کننده‌ی *Mahalanobis*، به دست آمده است.

- برای دسته ویژگی کلی، بهترین نتیجه به ازای طبقه‌بندی کننده‌ی *SVM* (۶۴٪)، به دست آمده است.
- برای هر یک از ۷ دسته ویژگی، درصد درستی به دست آمده از رأی‌گیری از بهترین نتیجه‌ی به دست آمده توسط ۵ طبقه‌بندی کننده کوچک‌تر یا مساوی است.
- برای دسته ویژگی کلی، نتیجه به دست آمده از رأی‌گیری (۶۱٪) از بهترین نتیجه‌ی به دست آمده از ۵ طبقه‌بندی کننده (۶۴٪) کمتر است.

۵-۵-۴ روش ترکیب ویژگی‌ها با تعداد ویژگی متغیر

۵-۵-۴-۱ الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر

در این روش برای تولید ویژگی‌های ترکیبی برای هر طبقه‌بندی کننده، به صورت زیر عمل

می‌کنیم:

۱. با توجه به حجم زیاد ویژگی‌ها در هر دسته با استفاده از یکی از معیارهای با سرپرست غیر

وابسته به طبقه‌بندی کننده، تعداد ویژگی‌ها را کاهش می‌دهیم.

❖ در آزمایش‌های انجام شده از معیار وابسته به ماتریس پراکندگی (*Scatter Matrix*)

استفاده شده است.

❖ از هر دسته ۵۰ ویژگی که بهترین معیار را دارند انتخاب شده است.

۲. با اعمال روش PCA ، ویژگی‌های دسته داده‌های آموزش را که در قسمت قبل انتخاب شده‌اند، ترکیب می‌کنیم به گونه‌ای که واریانس بین ویژگی‌های ترکیبی تولید شده از حد معینی کوچکتر نباشد.

❖ تعداد ویژگی‌های ترکیبی مطلوب در این حالت متغیر بوده و توسط مینیوم واریانس مشخص می‌شود.

❖ مینیوم واریانس را در آزمایش‌ها ۰,۰۱ در نظر می‌گیریم.

❖ ماتریس تبدیل به دست آمده در روش PCA برای داده‌های آموزش را بر داده‌های تست اعمال کرده و ویژگی‌های ترکیبی مربوط به آنها را تعیین می‌کنیم.

۳. پس از تعیین مجموعه ویژگی‌های تعیین شده برای تمام دسته‌ها، روش PCA را بر روی مجموعه‌ی به دست آمده اعمال کرده و ویژگی‌های ترکیبی نهایی (تعدادمتغیر) را براساس مینیوم واریانس تولید می‌کنیم.

❖ مینیوم واریانس را در این مرحله نیز، ۰,۰۱ در نظر می‌گیریم.

❖ ماتریس تبدیل به دست آمده در روش PCA برای داده‌های آموزش را بر مجموعه ویژگی‌های به دست آمده از داده‌های تست در قسمت قبل اعمال کرده و ویژگی‌های ترکیبی مربوط به آنها را تعیین می‌کنیم.

۴. طبقه‌بندی کننده را با ویژگی‌های ترکیبی به دست آمده از داده‌های آموزشی، آموزش داده و داده‌های تست را به کمک آن طبقه‌بندی می‌کنیم.

نتایج به دست آمده برای ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic* و

Mahalanobis در جدول (۶-۵)، نشان داده شده‌اند.

جدول (۶-۵): درصد درستی هر یک از طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All	Best
FLD	64	51	62	55	53	67	63	74	74
Bayes	55	47	62	58	54	58	66	57	66
SVM	63	55	61	57	56	63	60	74	74
Quad	55	48	62	58	53	58	66	58	66
Mahal	54	52	64	57	52	59	62	59	64
Best	64	55	64	58	56	67	66	74	
Vote	59	58	62	58	52	62	64	58	

■ Best result of all classifiers
 ■ Vote of all classifiers
 ■ Best result of each row

پس از اعمال الگوریتم توضیح داده شده برای هر کدام از طبقه‌بندی‌کننده‌ها، به ازای هر کدام از ۱۰۰

داده‌ی تست، کلاس مورد نظر را با رأی‌گیری میان ۵ طبقه‌بندی‌کننده معین می‌کنیم، به این صورت که

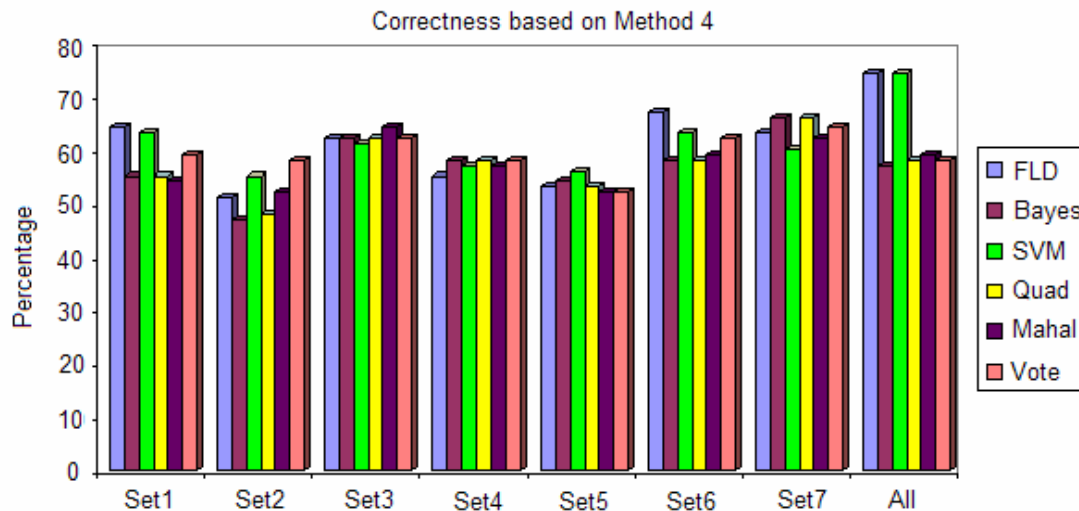
کلاسی را که حداقل ۳ طبقه‌بندی‌کننده تعیین کرده‌اند، به آن نسبت می‌دهیم. درصد درستی نتایج به دست

آمده با این روش در سطر *Vote* در جدول (۶-۵) نشان داده شده‌اند. در سطر *Best* در جدول (۶-۵)،

بهترین نتیجه از بین نتایج به دست آمده از ۵ طبقه‌بندی‌کننده، نشان داده شده است.

نمودار درصد درستی به ازای هر کدام از طبقه‌بندی‌کننده‌ها و درصد درستی حاصل از رأی‌گیری میان

آنها در شکل (۸-۵) نمایش داده شده است.



شکل (۵-۸): نمودار درصد درستی طبقه‌بندی‌کننده‌ها برای دسته ویژگی‌های تعریف شده

با استفاده از الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر

۵-۴-۵-۵ بررسی نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر

نتایج به دست آمده از اعمال الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی متغیر را می‌توان به

صورت زیر خلاصه نمود:

- برای طبقه‌بندی‌کننده‌های *FLD* و *SVM*، نتیجه‌ی به دست آمده برای دسته ویژگی کلی، بهترین نتیجه است. بهترین درصد درستی برای طبقه‌بندی‌کننده‌های *Bayes* و *Quadratic*، نتیجه‌ی به دست آمده از دسته ویژگی هفتم (ویژگی‌های آماری مربوط به تفاضل داده‌های کانال‌های متقارن) و برای طبقه‌بندی‌کننده‌ی *Mahalanobis*، نتیجه‌ی به دست آمده از دسته ویژگی سوم (ویژگی‌های وابسته به ضرایب *Wavelet*) می‌باشد.

- برای ۷ دسته ویژگی بهترین نتیجه در بین ۵ طبقه‌بندی‌کننده متغیر بوده و نتیجه خاصی برای تمایز یکی از طبقه‌بندی‌کننده‌ها نسبت به بقیه، به دست نمی‌آید.

● برای دسته ویژگی کلی، بهترین نتیجه به ازای طبقه‌بندی‌کننده‌های *FLD* و *SVM* (۷۴٪)، به دست آمده است.

● برای همه دسته ویژگی‌ها به جز دسته ویژگی دوم، درصد درستی به دست آمده از رأی‌گیری از بهترین نتیجه‌ی به دست آمده توسط ۵ طبقه‌بندی‌کننده کوچک‌تر بوده یا با آن مساوی است.

● برای دسته ویژگی کلی، نتیجه‌ی به دست آمده از رأی‌گیری (۵۸٪) از بهترین نتیجه به دست آمده از ۵ طبقه‌بندی‌کننده (۷۴٪) کمتر است.

۵-۵-۵ مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی‌ها

در این بخش نتایج به دست آمده در بخش‌های ۵-۵-۱ تا ۵-۵-۴، برای هر طبقه‌بندی‌کننده، با یکدیگر مقایسه می‌شوند. مقایسه‌ی نتایج به دست آمده برای طبقه‌بندی‌کننده‌های *FLD*، *Bayes*، *SVM*، *Quadratic*، *Mahalanobis* و حاصل رأی‌گیری میان آنها، به ترتیب در جدول‌های (۵-۷) تا (۵-۱۲) نشان داده شده است.

جدول (۵-۷): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *FLD*

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	61	51	56	55	64	65	66	77
Method2	59	46	57	63	64	67	67	76
Method3	61	58	56	59	56	69	63	62
Method4	64	51	62	55	53	67	63	74
Best	64	58	57	63	64	69	67	77

■ Best result of four methods

جدول (۸-۵): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Bayes*

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	55	52	63	60	61	58	56	59
Method2	63	55	64	59	62	61	64	70
Method3	56	52	58	58	59	56	62	61
Method4	55	47	62	58	54	58	66	57
Best	63	55	64	60	62	61	66	70

■ Best result of four methods

جدول (۹-۵): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *SVM*

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	62	52	60	63	63	64	60	64
Method2	58	51	61	59	62	70	63	77
Method3	60	56	63	61	59	71	66	64
Method4	63	55	61	57	56	63	60	74
Best	63	56	63	63	63	71	66	77

■ Best result of four methods

جدول (۱۰-۵): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Quadratic*

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	53	56	54	58	60	66	58	69
Method2	56	58	58	58	69	68	59	66
Method3	56	52	58	58	59	56	62	61
Method4	55	48	62	58	53	58	66	58
Best	56	58	62	58	69	68	66	69

■ Best result of four methods

جدول (۱۱-۵): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Mahalanobis*

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	59	52	57	62	63	65	61	72
Method2	57	56	61	60	67	61	65	70
Method3	57	54	57	58	61	58	65	59
Method4	54	52	64	57	52	59	62	59
Best	56	56	64	62	67	65	65	72

■ Best result of four methods

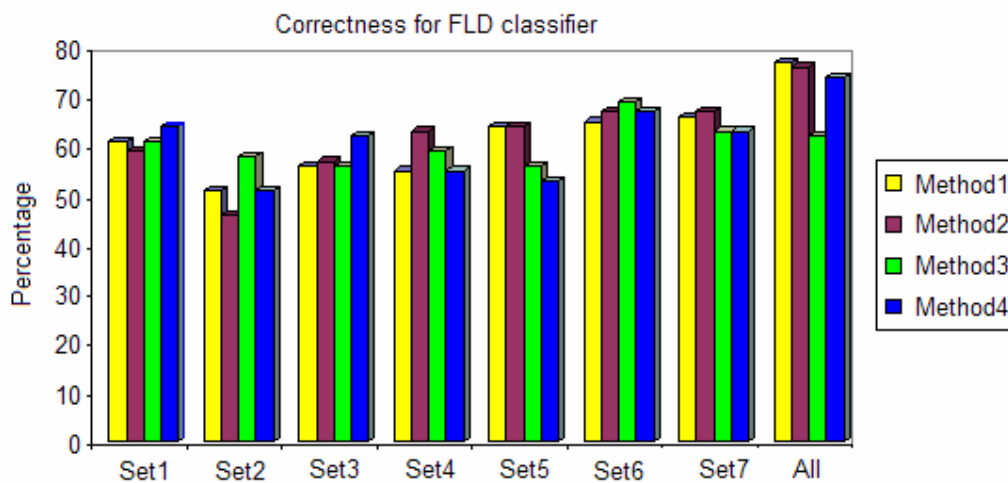
جدول (۱۲-۵): مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی میان طبقه‌بندی‌کننده‌ها

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	All
Method1	59	50	58	59	66	64	63	74
Method2	58	49	64	63	67	69	70	84
Method3	58	57	59	58	60	61	66	61
Method4	59	58	62	58	52	62	64	58
Best	59	58	64	63	67	69	70	84

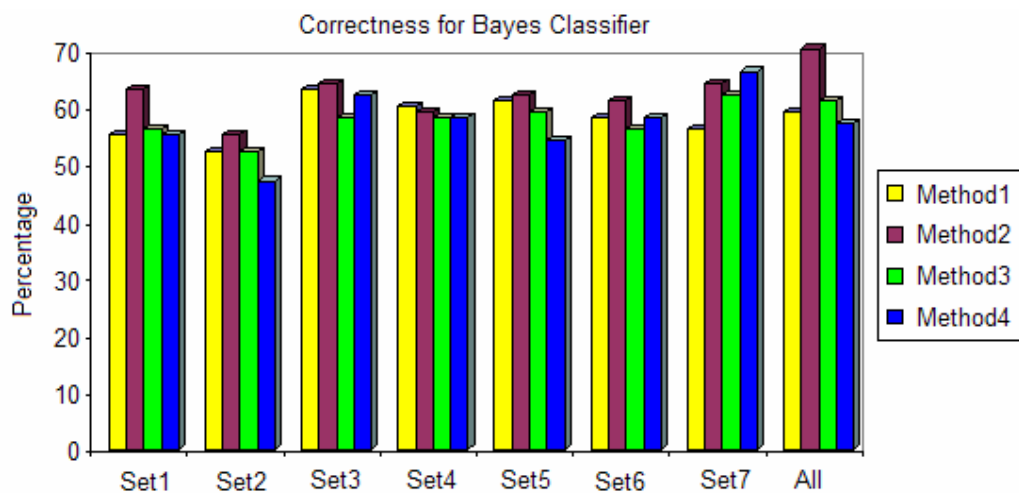
■ Best result of four methods

نمودار مقایسه‌ی نتایج به دست آمده برای ۵ طبقه‌بندی‌کننده و حاصل رأی‌گیری میان آنها، به ترتیب در

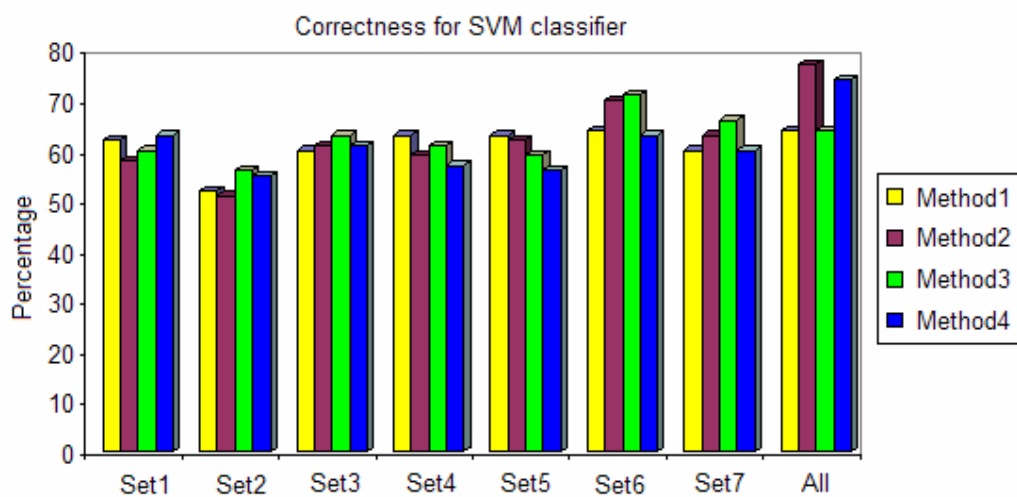
شکل‌های (۹-۵) تا (۱۴-۵) نمایش داده شده است.



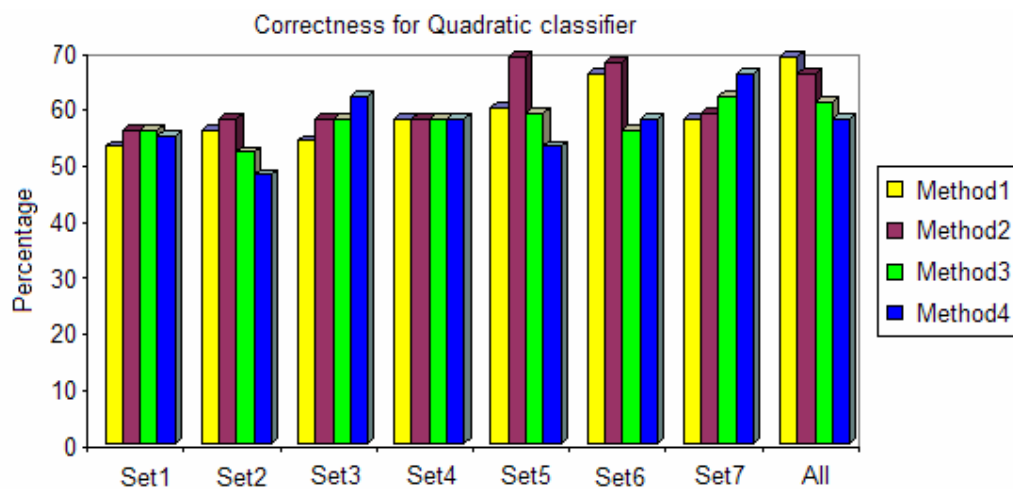
شکل (۹-۵): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی FLD



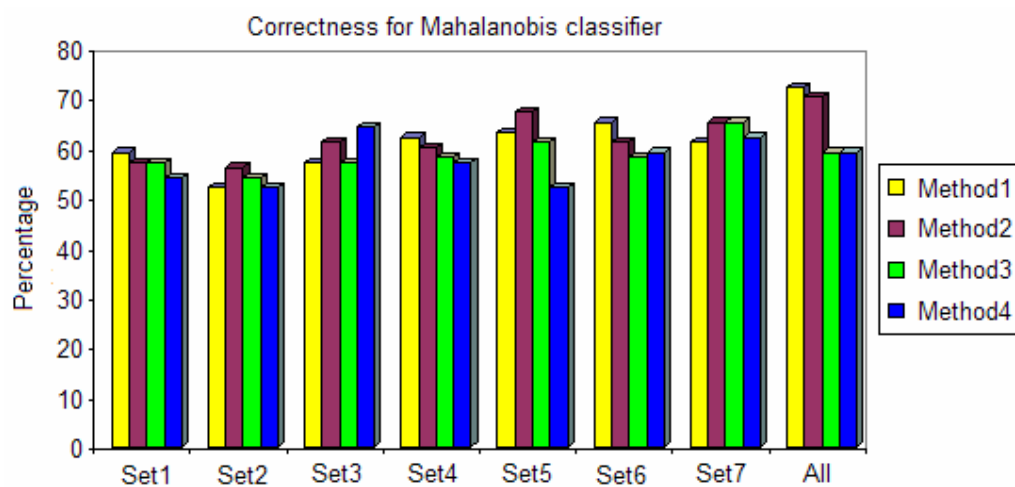
شکل (۵-۱۰): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Bayes*



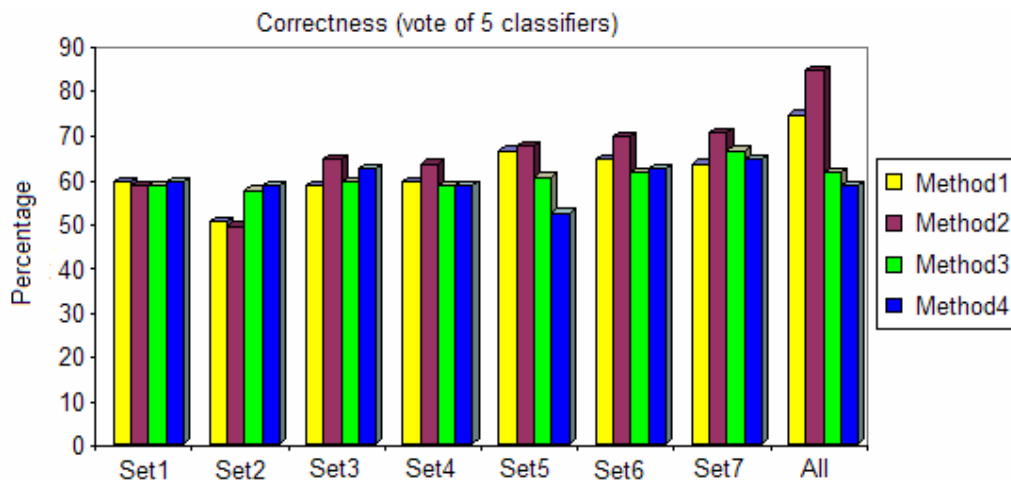
شکل (۵-۱۱): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *SVM*



شکل (۵-۱۲): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Quadratic*



شکل (۵-۱۳): مقایسه‌ی نتایج به دست آمده از اعمال چهار روش کاهش ویژگی بر روی طبقه‌بندی‌کننده‌ی *Mahalanobis*



شکل (۵-۱۴): مقایسه‌ی نتایج به دست آمده از چهار روش کاهش ویژگی حاصل از رأی‌گیری میان طبقه‌بندی‌کننده‌ها

با توجه به جدول‌ها و نمودارهای ارائه شده، می‌توان نتایج به دست آمده را به صورت زیر خلاصه نمود:

- برای طبقه‌بندی‌کننده‌ی *FLD*، مؤثرترین روش برای ۷ دسته ویژگی، الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده است. بهترین نتیجه برای دسته ویژگی کلی از اعمال روش گزینش ویژگی با تعداد ویژگی ثابت، به دست آمده است و الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده با ۱٪ اختلاف دوم است.

- برای طبقه‌بندی‌کننده‌ی *Bayes*، مؤثرترین روش برای ۷ دسته ویژگی، الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده است. بهترین نتیجه برای دسته ویژگی کلی از اعمال همین روش، به دست می‌آید.

- برای طبقه‌بندی‌کننده‌ی *SVM*، مؤثرترین روش برای ۷ دسته ویژگی، الگوریتم ترکیب ویژگی‌ها با تعداد ویژگی ثابت است. بهترین نتیجه برای دسته ویژگی کلی از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده، به دست می‌آید.

- برای طبقه‌بندی‌کننده‌ی *Quadratic*، مؤثرترین روش برای ۷ دسته ویژگی، الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده است. بهترین نتیجه برای دسته ویژگی کلی از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت، به دست می‌آید و نتیجه به دست آمده از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده، ۳٪ پایین‌تر است.

- برای طبقه‌بندی‌کننده‌ی *Mahalanobis*، الگوریتم‌های گزینش ویژگی‌ها با تعداد ویژگی ثابت و تعداد ویژگی وابسته به طبقه‌بندی‌کننده، به نسبت یکسان روش‌های مؤثر برای ۷ دسته ویژگی بوده‌اند. بهترین نتیجه برای دسته ویژگی کلی از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی ثابت، به دست آمده و نتیجه‌ی به دست آمده از اعمال الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده، ۲٪ پایین‌تر است.

- با توجه به نتایج به دست آمده از رأی‌گیری میان طبقه‌بندی‌کننده‌ها، الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده، مؤثرترین روش برای ۷ دسته ویژگی بوده و برای دسته ویژگی کلی نیز بهترین نتیجه را تولید می‌کند. این نتیجه (۸۴٪) بهترین نتیجه‌ی ایجاد شده میان همه الگوریتم‌ها و طبقه‌بندی‌کننده‌ها می‌باشد.

۵-۶ نتیجه‌گیری و مقایسه‌ی نتایج به دست آمده با نتایج دیگر گروه‌ها

۵-۶-۱ ارائه‌ی الگوریتم کلی برای طبقه‌بندی داده‌ها

با بررسی نتایج به دست آمده می‌توان الگوریتم کلی برای کاهش ویژگی‌ها و تولید نتیجه‌ی مناسب را به صورت زیر ارائه نمود:

۱) الگوریتم گزینش ویژگی‌ها با تعداد ویژگی وابسته به طبقه‌بندی‌کننده (بخش (۵-۵-۲-۱)) را بر

روی ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic*، *Mahalanobis* اعمال

نموده و داده‌های تست را توسط هر کدام از طبقه‌بندی‌کننده‌ها به طور جداگانه طبقه‌بندی

می‌کنیم.

۲) به ازای هر کدام از داده‌های تست، کلاس مورد نظر را با رأی‌گیری میان ۵ طبقه‌بندی‌کننده معین

می‌کنیم، به این صورت که کلاسی را که حداقل ۳ طبقه‌بندی‌کننده تعیین کرده‌اند، به آن نسبت

می‌دهیم.

۵-۶-۲ مقایسه‌ی نتیجه به دست آمده با نتایج دیگر گروه‌ها

نتیجه‌ی به دست آمده برای داده‌های دسته چهارم مسابقات *BCI2003*، توسط الگوریتم ارائه شده، ۸۴٪ است.

بهترین نتیجه به دست آمده در مسابقه توسط *Zhiguang Zhang* و همکارانش از دانشگاه

Tsinghua به دست آمده است که برابر ۸۴٪ می‌باشد. ویژگی‌های استفاده شده توسط این گروه از

ترکیب *Common Subspace Decomposition (CSSD)* و *Fisher discriminant* تولید شده‌اند

و از طبقه‌بندی‌کننده‌ی *Perceptron* استفاده شده است. بیشترین درصد درستی بعدی به دست آمده برابر ۸۱٪ است [۲۳].

با توجه به بررسی انجام شده توسط *Blankertz* و تیم او بر روی نتایج به دست آمده از گروه‌های مختلف که خطای طبقه‌بندی داده‌های تست آنها کمتر از ۳۰٪ بوده است، مشخص شده است که ۱۰٪ از داده‌ها توسط حداقل ۶ روش از ۸ روش به طور اشتباه طبقه‌بندی شده‌اند. این نتیجه مشخص می‌کند که (۱) آرتیفکت‌ها می‌توانند سیگنال‌های حرکتی *EEG* را غیر قابل بازیافت سازند یا اینکه (۲) حرکت‌ها به طور کامل در پارامترهای انتخاب شده سیگنال‌های *EEG* منعکس نمی‌شوند [۲۲].

فصل ششم: جمع‌بندی، نتیجه‌گیری و پیشنهادها

۶-۱ جمع‌بندی

فصل اول به عنوان مقدمه‌ای بر موضوع ارتباط مغز با کامپیوتر، ارائه شد. در این فصل سیگنال‌های مغزی معرفی گشتند و نحوه‌ی انجام پروژه و ساختار پایان‌نامه توضیح داده شد.

در فصل دوم، دسته ویژگی‌های مختلف که معمولاً در طبقه‌بندی سیگنال‌های مغزی مفید هستند، معرفی شدند و از هر دسته، ویژگی‌هایی که در این پروژه از آنها استفاده شده است، توضیح داده شدند.

در فصل سوم به معرفی طبقه‌بندی‌کننده‌های با سرپرست مختلف پرداخته شد و روش عملکرد ۵ طبقه‌بندی‌کننده‌ی سریع (*FLD*، *Bayes*، *SVM*، *Quadratic* و *Mahalanobis*) استفاده شده در آزمایش‌ها، توضیح داده شد.

در فصل چهارم، ابتدا دلایل نیاز به کاهش تعداد ویژگی‌ها بررسی شدند. سپس دو روش کلی کاهش ویژگی (ترکیبی و گزینشی) معرفی گشتند و با هم مقایسه شدند. پس از آن معیارهای مناسب در

گزینش ویژگی‌ها نام برده شده و معیارهای مختلفی که آزمایش‌ها بر اساس آنها انجام شده‌اند، به طور خلاصه توضیح داده شدند. در پایان فصل نیز روش‌های مختلف جستجو در میان مجموعه‌ی تمام ویژگی‌ها برای یافتن دسته ویژگی بهینه، مورد بررسی قرار گرفتند.

در فصل پنجم داده‌های مورد آزمایش و نحوه‌ی انجام آزمایش‌ها توضیح داده شدند. در این فصل، چهار الگوریتم مختلف برای کاهش ویژگی‌ها و طبقه‌بندی داده‌ها، معرفی گشته و نتایج به دست آمده از اعمال آنها بر داده‌ها، بررسی شدند. در پایان فصل، با توجه به نتایج به دست آمده از آزمایش‌ها، الگوریتم کلی برای طبقه‌بندی داده‌ها ارائه گشت و نتایج به دست آمده با نتایج دیگر گروه‌ها مقایسه شد.

۶-۲ نتیجه‌گیری

با توجه به توضیحاتی که در چهار فصل نخست این پایان‌نامه داده شدند، روش‌های مختلفی برای طبقه‌بندی داده‌ها (به طور خاص سیگنال‌های مغزی) وجود دارد. تمایز این روش‌ها در نوع ویژگی‌ها و نوع طبقه‌بندی‌کننده(ها)ی استفاده شده می‌باشد. ویژگی‌های استفاده شده به دسته ویژگی‌های اولیه‌ی استخراج شده از داده‌ها، روش‌های کاهش ویژگی، معیارهای انتخاب ویژگی و الگوریتم‌های جستجو برای یافتن دسته ویژگی بهینه بستگی دارند. بنابراین برای طبقه‌بندی یک دسته داده‌ی معین می‌توان الگوریتم‌های متفاوتی ارائه نمود که نتایج مختلفی ایجاد می‌نمایند.

هدف اصلی در این پایان‌نامه ارائه‌ی روشی است که با استفاده از ابزارهای تعریف شده، علاوه بر دقت در طبقه‌بندی داده‌ها و تولید نتایج مطلوب، تا حد امکان از سرعت بالایی بهره‌مند باشد. ایده‌ی اصلی این پایان‌نامه، از پایان‌نامه‌ی کارشناسی ارشد آقای احسان اربابی [۲] به دست آمده است. در بخش آخر پایان‌نامه‌ی آقای اربابی پیشنهاد می‌شود، الگوریتم ارائه شده بر روی دیگر داده‌های *BCI* نیز

بررسی گردد و نتایج به دست آمده با نتایج ایشان مقایسه گردد. در ابتدای انجام این پروژه، الگوریتم مذکور (الگوریتم اول در فصل پنجم این پایان نامه الگوریتم مطرح شده توسط آقای اربابی می‌باشد) بر روی داده‌های دسته‌ی چهارم مسابقه‌ی *BCI2003* بررسی گردید که نتایج به دست آمده در مقایسه با نتایجی که ایشان از دیگر داده‌ها به دست آورده بودند، مناسب نبودند. سپس الگوریتم‌های دیگری که با اعمال تغییرات جزئی یا کلی از الگوریتم اول به دست آمده بودند، بر روی این داده‌ها مورد بررسی قرار گرفتند. با توجه به نتایج به دست آمده از چهار الگوریتم مختلف برای کاهش ویژگی‌ها و طبقه‌بندی داده‌ها که در فصل پنجم ارائه شدند، الگوریتمی کلی برای طبقه‌بندی داده‌ها به دست آمد. در این الگوریتم از کاهش ویژگی‌ها به روش گزینشی با تعداد ویژگی وابسته به طبقه‌بندی‌کننده و ۵ طبقه‌بندی‌کننده‌ی سریع *FLD*، *Bayes*، *SVM*، *Quadratic* و *Mahalanobis*، استفاده شد. برای به دست آوردن نتیجه‌ی نهایی میان نتایج به دست آمده از این پنج طبقه‌بندی‌کننده، رأی‌گیری انجام شد.

همان‌گونه که در فصل قبل نیز گفته شد، نتیجه‌ی نهایی به دست آمده از این الگوریتم (درصد درستی ۸۴٪) برابر با بهترین نتیجه‌ی به دست آمده توسط دیگر گروه‌ها می‌باشد. نتیجه‌ی به دست آمده با توجه به دلایل ارائه شده مبنی بر مناسب نبودن داده‌ها به حد کافی [۲۲]، نتیجه‌ی مطلوبی است.

۳-۶-۳ پیشنهادها

با توجه به محدودیت‌های زمانی و عملی، در این پروژه فقط به بررسی یکی از دسته‌های ارائه شده در زمینه‌ی *BCI* پرداخته شد. تنوع داده‌های *BCI*، امکان انجام تحقیقات و پژوهش‌های متنوعی را پدید می‌آورد. این داده‌ها در زمینه‌های مختلفی همچون تصور حرکت، تصور اعمال ذهنی و تشخیص

حروف توسط گروه‌های مختلف ثبت شده‌اند. انجام آزمایش‌های گوناگون بر روی این داده‌ها می‌تواند نتایج مفیدی در زمینه‌ی ارتباط مغز با کامپیوتر در پی داشته باشد.

همچنین با توجه به توضیحات ارائه شده در بخش قبلی در مورد تنوع روش‌های طبقه‌بندی، پیشنهاد می‌شود داده‌های استفاده شده در این پایان‌نامه، توسط روش‌های دیگر نیز بررسی گردند و ویژگی‌ها و طبقه‌بندی‌کننده‌های دیگر نیز برای یافتن نتایج بهتر مورد استفاده قرار گیرند.

همچنین پیشنهاد می‌شود الگوریتم کلی ارائه شده در این پایان‌نامه، برای تعیین قابلیت‌های آن، بر روی دیگر داده‌ها نیز پیاده‌سازی گردد و از نتایج به دست آمده برای بهبود عملکرد این الگوریتم، استفاده شود.

ضمیمه

با مطالعه‌ی دسته داده‌های مختلف که در مسابقات *BCI2003* و *BCI2005* داده شده‌اند، می‌توان انواع داده‌ها را به صورت زیر دسته‌بندی نمود:

(در هر یک از این دسته‌ها یک سری داده به عنوان داده‌های آموزشی برای طراحی طبقه‌بندی‌کننده و سری دیگر به عنوان داده‌های تست، داده شده‌اند که بایستی کلاس صحیح برای آنها مشخص گردد).

۱. تنظیم شخصی پتانسیل کرتکس کند^۱

SCP^۲ ایجاد یک شیفت مثبت یا منفی در سیگنال *EEG* است که ممکن است از ۳۰۰ میلی ثانیه تا چند ثانیه طول بکشد. در هنگام استراحت سیگنال *EEG* به سمت مقادیر مثبت و در هنگام شروع به فعالیت‌های خاص به سمت مقادیر منفی جابه جا می شود. دامنه این تغییرات می‌تواند در اختیار فرد باشد و فرد با تمرین مقدار آن را تا حدی کم و زیاد کند. در سری داده‌های (*Ia*) و (*Ib*) مسابقه‌ی *BCI2003*، از شخص خواسته می شود با کنترل *SCP* خود یک نشانه موس را روی صفحه‌ی مانیتور بالا و پایین ببرد. پتانسیل مثبت کرتکس باعث حرکت به سمت پایین و پتانسیل منفی باعث حرکت به سمت بالا می‌شود. هدف طبقه‌بندی داده‌های تست به دو دسته‌ی بالا و پایین بر اساس برچسب^۳‌های مشخص شده برای داده‌های آموزشی است [۲۴]، [۲۵].

^۱ Self-regulation of SCPs

^۲ Slow Cortical Potential

^۳ Label

۲. تنظیم شخصی فرکانس‌های باند μ یا β ^۱

در سری داده‌ی (IIa) از مسابقه‌ی BCI2003، از شخص مورد آزمایش خواسته می‌شود حرکت عمودی یک نشانگر را با کمک تغییر دامنه‌ی فرکانس‌های μ/β خود برای رسیدن به یک هدف عمودی معین در سمت راست صفحه کنترل کند. در ثبت داده‌ها از فیدبک بصری^۲ برای کنترل تغییر دامنه‌ی فرکانس‌های μ/β توسط شخص استفاده شده است. مسئله تعیین ۴ کلاس مختلف برای مکان‌های عمودی هدف از پیش معین شده بر روی صفحه، برای داده‌های تست می‌باشد [۲۶].

۳. نوشتن متن به کمک پتانسیل برانگیخته‌ی P300^۳

پتانسیل برانگیخته‌ی P300، ایجاد یک الگوی خاص بر روی سیگنال EEG تقریباً ۳۰۰ میلی ثانیه پس از دریافت تحریک می‌باشد. در ثبت این داده‌ها، شخص با یک ماتریس ۶×۶ از حروف و اعداد روبه‌روست که سطرها و ستون‌های آن به طور تصادفی برجسته^۴ می‌شوند. شخص بایستی در هر مرحله توجه خود را بر روی کاراکتر خواسته شده معطوف نماید. زمانی که سطر یا ستون برجسته شده شامل کاراکتر مورد نظر باشد، در سیگنال EEG شخص، الگویی شبیه پتانسیل برانگیخته‌ی P300، مشاهده می‌گردد.

دسته داده‌های (IIb) در مسابقه‌ی BCI2003 و (II) در مسابقه‌ی BCI2005، شامل سیگنال‌های ثبت شده برای نوشتن متن به کمک پتانسیل برانگیخته‌ی P300، می‌باشند. در این داده‌ها، هدف تعیین کلاس مربوط به هر کاراکتر از بین ۳۶ کلاس ممکن است [۲۷]، [۲۸].

^۱ self-regulation of mu -and/or beta- rhythm

^۲ Visual feedback

^۳ P300 speller paradigm

^۴ evoked potential

^۵ intensify

۴. تصور حرکت^۱

در ثبت این داده‌ها، از شخص مورد آزمایش خواسته می‌شود تصور کند یکی از اعضای بدن (معمولاً دست راست، دست چپ، پا یا زبان) خود را حرکت می‌دهد. دسته داده‌های مختلفی برای سیگنال‌های مربوط به تصور حرکت ثبت شده اند که در ادامه دسته‌های مشخص شده در مسابقات *BCI2003* و *BCI2005* معرفی می‌گردند:

• تصور حرکت، حالت غیر ایستا^۲: دسته داده‌های (*III*) در مسابقه‌ی *BCI2003* و (*IIIb*) در مسابقه‌ی *BCI2005*

در ثبت این دسته داده، از شخص خواسته می‌شود با کمک تصور حرکت دست راست یا دست چپ، نشانگر را به سمت مشخص شده بر روی صفحه‌ی مانیتور، حرکت دهد. در طول انجام آزمایش فیدبک بصری موجود است. هدف طراحی یک طبقه‌بندی‌کننده با الگوریتمی علی^۳ برای تعیین دو کلاس تصور حرکت دست راست و چپ، برای داده‌های تست می‌باشد [۲۹]، [۳۰].

• تصور حرکت، حالت چند کلاسه^۴: دسته داده‌های (*IIIa*) در مسابقه‌ی *BCI2005*

در ثبت این داده‌ها از شخص خواسته می‌شود به یکی از ۴ حرکت دست راست، دست چپ، پا و یا زبان فکر کند. در طول انجام آزمایش فیدبکی به شخص منتقل نمی‌گردد. طبقه‌بندی طراحی شده باید بتواند داده‌های تست را در یکی از ۴ گروه ذکر شده قرار دهد [۳۱].

• تصور حرکت، انتقال جلسه به جلسه^۵: دسته داده‌ی (*I*) در مسابقه‌ی *BCI2005*

^۱ Motor imagery

^۲ non-stationary

^۳ causal algorithm

^۴ multi class

^۵ session to session transfer

هدف از این دسته داده‌ها، دسته‌بندی تصور حرکت انگشت کوچک دست چپ و حرکت زبان (دو دسته) در شرایط بدون فیدبک برای یک شخص واحد در دو روز مختلف است به گونه‌ای که داده‌های آموزشی در روز (جلسه) اول و داده‌های تست در روز (جلسه) دوم ثبت شده‌اند [۳۲].

• تصور حرکت، تعداد داده‌ی آموزشی کم^۱: دسته داده‌ی (*IVa*) در مسابقه‌ی *BCI2005*

در این دسته داده‌ها هدف دسته‌بندی تصور حرکت برای سه کلاس دست راست، دست چپ و پا در شرایط بدون فیدبک و داده‌های آموزشی کم می‌باشد. روش پیشنهاد شده برای جبران داده‌های آموزشی کم، استفاده از داده‌های ثبت شده برای افراد دیگر، برای آموزش طبقه‌بندی‌کننده برای شخص مورد آزمایش است [۳۳].

• تصور حرکت، طبقه‌بندی بدون مشخص نمودن زمان تغییر وضعیت^۲: دسته داده‌ی (*IVb*) در مسابقه‌ی

BCI2005

در این دسته داده‌ها هدف دسته‌بندی تصور حرکت برای دو کلاس دست چپ و پا می‌باشد. داده‌های تست پیوسته بوده (مشخص نیست در چه زمانی تصور حرکت تغییر کرده است) و خروجی بایستی در هر لحظه از زمان به صورت (-1) (دست چپ)، (1) (پا) و (0) (عدم وابستگی به هیچ کدام از دو کلاس) باشد [۳۴].

۵. تصور اعمال ذهنی^۳

در این دسته داده‌ها هدف طبقه‌بندی داده‌های ثبت شده‌ی مختلفی است که در آنها شخص مورد آزمایش به تصور اعمال ذهنی مختلف (نامه‌نگاری، شمارش اعداد، ضرب اعداد و...) می‌پردازد. در مسابقات *BCI* برگزار شده، دسته داده‌ی (*V*) در مسابقه‌ی *BCI2005* سیگنال‌های ثبت شده مربوط به

^۱ *small training sets*

^۲ *uncued classifier application*

^۳ *Mental imagery*

تصور اعمال ذهنی (همراه با تصور حرکت) می‌باشد. کلاس‌های مورد نظر در این دسته داده، تصور حرکت مکرر دست راست، تصور حرکت مکرر دست چپ و ساختن کلماتی که با یک حرف رندوم یکسان آغاز می‌شوند، در شرایط بدون فیدبک، می‌باشند [۳۵].

۶. تایپ حروف به طور دلخواه^۱

در دسته داده‌های (IV) مسابقه‌ی BCI2003 که از یک شخص عادی در حالت بدون فیدبک ثبت شده است، شخص مورد آزمایش در موقعیت استاندارد برای تایپ بر روی کیبورد کامپیوتر قرار گرفته است. وظیفه شخص فشردن ۴ کلید مشخص شده بر روی صفحه کلید با انگشتان کوچک یا اشاره دست راست یا چپ به خواست خود و با زمان بندی دلخواه است. عمل تایپ با سرعت متوسط یک کلید در هر ثانیه انجام می‌شود. داده‌ها به صورت دوره‌هایی به طول ۵۰۰ میلی ثانیه هستند که ۱۳۰ میلی ثانیه قبل از فشردن کلید تمام می‌شوند. هدف، تعیین کلاس مربوط به داده‌های تست (0: حرکت دست چپ، 1: حرکت دست راست) می‌باشد [۲۱].

^۱ Self-paced IS

مراجع

[1] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer and J. R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System", *IEEE Trans. On Biomedical Eng.*, Vol. 51, No. 6, pp. 1034–1043, June 2004.

[۲] احسان اربابی، "بررسی وابستگی میان ویژگی‌های مؤثر به کار رفته در روش‌های مختلف طبقه‌بندی

سیگنال‌های مغزی در BCI"، پایان‌نامه کارشناسی ارشد، دانشکده مهندسی برق، دانشگاه صنعتی شریف،

مهر ۱۳۸۵

[3] J. Lehtonen, "EEG-based Brain Computer Interfaces", Department of Electrical and Communications Engineering, Helsinki University of Technology, May 2002.

[4] J. D. Baliss, "A Flexible Brain-Computer Interface", Department of Computer Science, University of Rochester, 2001.

[5] C. McLeod and G. Maxwell, "Practical Neural Networks (1)", *Elektor Electronics*, 2003.

[6] S. Tong, A. Bezerianos, A. Malhotra, Y. Zhu and N. Thakor, "Parameterized entropy analysis of EEG following hypoxic–ischemic brain injury", *Physics Letters A*, vol. 314, pp. 354–361, 2003.

[7] S. M. Pincus, "Approximate Entropy as A Measure of System Complexity", *Proc. Nati. Acad. Sci. USA*, Vol. 88, pp. 2297–2301, March 1991.

[۸] محمد صالح نامبخش، "شبکه‌های عصبی مصنوعی Wavelet و کاربردهای آن"، انتشارات ناقوس.

[9] MATLAB Help, MATLAB Version 7.0, The MathWorks, Inc.

[10] G. Mohammadi, P. Shoushtari, B. Molaee Ardekani and M. B. Shamsollahi, "Person Identification by Using AR Model for EEG Signals", *Enformatika, Trans. On Engineering, Computing and Technology*, Vol. 11, pp. 281–285, Feb. 2006.

[11] A. Schlögl, G. Müller, C. Neuper, G. Krausz, B. Graimann and G. Pfurtscheller, "Adaptive autoregressive parameters used in BCI research", Institute for Biomedical Engineering, University of Technology Graz, Austria.

- [12] H. Mirghasemi, R. Fazel-Rezai and M. B. Shamsollahi, "Analysis of P300 Classifiers in Brain Computer Interface Speller".
- [13] H. Mirghasemi, M. B. Shamsollahi and R. Fazel-Rezai, "Assessment of Preprocessing on Classifiers Used in the P300 Speller Paradigm".
- [14] Support vector machine, Wikipedia, the free encyclopedia, [online]. Available: http://en.wikipedia.org/wiki/Support_vector_machine
- [15] Quadratic Classifier, Wikipedia, the free encyclopedia, [online]. Available: http://en.wikipedia.org/wiki/Quadratic_classifier
- [16] D. Mery, "Classification of Potential Defects in The Automatic Inspection of Aluminium Castings Using Statistical Pattern", *NDT.net*, Vol. 7 , No.12 , Dec. 2002 , [online]. Available: <http://www.ndt.net/article/ecndt02/232/232.htm>
- [17] R. Gutierrez-Osuna, "Introduction to Pattern Analysis, Lecture11: Sequential Feature Selection", Texas A&M University, [online]. Available: http://courses.cs.tamu.edu/rgutier/cpsc689_f06/111.pdf
- [18] R. Gutierrez-Osuna, "Introduction to Pattern Analysis, Lecture 9 : Principal Components Analysis", Texas A&M University, [online]. Available: http://courses.cs.tamu.edu/rgutier/cpsc689_f06/19.pdf
- [19] R. Gutierrez-Osuna, "Introduction to Pattern Recognition, Lecture 17 : Feature Subset Selection II", Write State University, [online]. Available: http://courses.cs.tamu.edu/rgutier/cs790_w02/117.pdf
- [20] J. Yang and V. Honavar, "Feature Subset Selection Using A Genetic Algorithm", Artificial Intelligence Research Group, Department of Computer Science, Iowa State University, [online]. Available: <http://www.cs.iastate.edu/~honavar/Papers/bookfinal.ps>
- [21] Data set <self-paced1s>, Fraunhofer-FIRST, Intelligent Data Analysis Group (Klaus-Robert Müller), and Freie Universität Berlin, Department of Neurology, Neurophysics Group (Gabriel Curio), online. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/berlin_desc.html
- [22] B. Blankertz, K. R. Muller, G. Gurio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schogl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroder and N. Birbaumer, "The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG

Signal Trials”, IEEE Trans. On Biomedical Eng., Vol.51, No.6 , pp. 1044 – 1051, June 2004 .

[23] BCI Competition 2003 , Final Results , [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/results/

[24] Data set Ia: <self-regulation of SCPs>, University of Tuebingen, [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/tuebingen_desc_i.html

[25] Data set Ib: <self-regulation of SCPs>, University of Tuebingen, [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/tuebingen_desc_ii.html

[26] Wadsworth BCI Dataset, NIPS* 2001 BCI Classification Contest, Documentation, Wadsworth Center, NYS Department of Health, (Jonathan R. Wolpaw), [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/albany_desc/albany_desc_i.html

[27] Documentatio, 2nd Wadsworth BCI Dataset (P300 Evoked Potentials), Data Acquired Using BCI 2000 P3 Speller Paradigm, BCI Classification Contest November 2002 , Wadsworth Center, NYS Department of Health, (Jonathan R. Wolpaw), [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/albany_desc/albany_desc_ii.html

[28] Documentation Wadsworth BCI Dataset (P300 Evoked Potentials) Data Acquired Using BCI 2000's P3 Speller Paradigm, BCI Competition III Challenge 2004 , Wadsworth Center, NYS Department of Health, (Jonathan R. Wolpaw, Gerwin Schalk, Dean Krusienski) [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_II.pdf

[29] Data set: BCI-experiment , Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz. (Gert Pfurtscheller), [online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_ii/Graz_description.doc

[30] Dataset IIIb: Non-stationary 2 -class BCI data, Laboratory of Brain-Computer Interfaces (BCI-Lab), Graz University of Technology, (Gert Pfurtscheller, Alois Schlögl), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_IIIb.pdf

[31] Dataset IIIa: 4 -class EEG data, Laboratory of Brain-Computer Interfaces (BCI-Lab), Graz University of Technology, (Gert Pfurtscheller, Alois Schlögl), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_IIIa.pdf

[32] Data set I <motor imagery in ECoG recordings, session-to-session transfer>, University of Tübingen, Germany, Dept. of Computer Engineering (Prof. Rosenstiel) and

Institute of Medical Psychology and Behavioral Neurobiology (Niels Birbaumer), and Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany (Bernhard Schölkopf), and Universität Bonn, Germany, Dept. of Epileptology (Prof. Elger), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_I.html

[33] Data set IVa <motor imagery, small training sets> , Fraunhofer FIRST, Intelligent Data Analysis Group (Klaus-Robert Müller, Benjamin Blankertz), and Campus Benjamin Franklin of the Charité - University Medicine Berlin, Department of Neurology, Neurophysics Group (Gabriel Curio), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_IVa.html

[34] Data set IVb <motor imagery, uncued classifier application>, Fraunhofer FIRST, Intelligent Data Analysis Group (Klaus-Robert Müller, Benjamin Blankertz), and Campus Benjamin Franklin of the Charité - University Medicine Berlin, Department of Neurology, Neurophysics Group (Gabriel Curio), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_IVb.html

[35] Data set V <mental imagery, multi-class>, IDIAP Research Institute (Silvia Chiappa, José del R. Millán), [online]. Available: http://ida.first.fhg.de/projects/bci/competition_iii/desc_V.html