

۳- می‌خواهیم یک مسئله طبقه‌بندی دو کلاسه را با استفاده از ویژگی‌های CSP حل کنیم. طبقه‌بندی بر روی داده‌های EEG مربوط به ثبت‌های تصورات ذهنی (تصور حرکت پا و انجام عمل تفریق ذهنی) است. دو سری داده در فایل `CSPdata.mat` در اختیار شما قرار داده شده است. داده‌های آموزش در یک آرایه سه‌بعدی `TrainData` به ابعاد $(30 \times 256 \times 165)$ (تعداد کانال‌ها ۳۰، تعداد نمونه‌های زمانی ۲۵۶ و تعداد آزمایش‌ها ۱۶۵ است) قرار داده شده‌اند و برچسب‌های مربوط به داده‌های آموزشی در بردار `TrainLabel` قرار داده شده‌اند. داده‌های آزمون در آرایه سه‌بعدی `TestData` به ابعاد $(30 \times 256 \times 40)$ (تعداد آزمایش‌های تست ۴۵ است) قرار داده شده‌اند.

الف) ابتدا با استفاده از تمام داده‌های آموزشی، فیلترهای مکانی CSP را به دست آورده و بر روی داده‌های آموزشی اعمال کنید. سیگنال فیلتر شده متناظر با فیلتر اول و آخر را برای نمونه‌هایی از دو کلاس رسم کرده و با هم مقایسه کنید (مشابه صفحه ۴۳ اسلاید EEG Patterns).

ب) فیلترهای مکانی اول و آخر به دست آمده در قسمت الف) را با استفاده از تابع `plottopomap.m` (تمرین سری سوم) رسم کرده و با هم مقایسه کنید.

ج) با استفاده از روش 4-fold cross-validation داده‌های آموزش را به ۴ دسته تقسیم کنید. هر بار سه دسته را به عنوان داده‌های آموزشی در نظر گرفته و یک دسته را داده اعتبارسنجی قرار دهید. الگوریتم CSP را با استفاده از فقط بخش آموزشی اجرا کرده و ویژگی‌های CSP (واریانس کانال‌های جدید بعد از اعمال CSP) را از داده‌های آموزشی و داده‌های اعتبارسنجی استخراج کنید. از یک طبقه‌بند ساده (مانند kNN، SVM خطی یا LDA) استفاده کرده و آن را با استفاده از ویژگی‌های بخش آموزشی، آموزش داده و بر روی داده‌های اعتبارسنجی اعمال کنید و برچسب‌ها را تعیین کنید. این کار را چهار بار (هر بار متناظر با یک دسته از داده‌ها

به عنوان داده اعتبارسنجی) انجام داده و میانگین صحت را محاسبه نمایید. برای تعیین پارامتر مجهول تعداد فیلترهای CSP (برابر با تعداد ویژگی‌های مورد استفاده)، صحت 4-fold cross-validation را به ازای مقادیر مختلف تعداد فیلترها انجام داده و تعداد فیلتر بهینه را به دست آورید.

د) بهترین طبقه‌بند به دست آمده از قسمت (ج) را با استفاده از داده‌های آموزش، آموزش داده و بر داده‌های تست اعمال کرده و برچسب داده‌های تست را تعیین کنید و در یک بردار TestLabel ذخیره نمایید.

• توضیحاتی در مورد سیگنال‌های مورد بررسی:

- داده‌ها با استفاده از سیستم g.tec GAMMASys ثبت شده‌اند.

- از ۳۰ کانال اندازه‌گیری EEG که در موقعیت استاندارد جهانی سیستم ۲۰-۱۰ تعمیم‌یافته قرار گرفته‌اند، برای ثبت داده‌ها استفاده شده است.

- ۳۰ الکتروود ثبت:

AFz, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO3, PO4, O1, and O2.

- از فیلتر میان‌گذر بین ۰/۵ و ۱۰۰ هرتز استفاده شده و داده‌ها با فرکانس ۲۵۶ هرتز نمونه‌برداری شده‌اند. همچنین نویز برق شهر با فیلتر notch ۵۰ هرتز حذف شده است. آزمایش‌ها به بازه‌های ۱ ثانیه‌ای (۲۵۶ نمونه زمانی) تقسیم شده‌اند.

- هر آزمایش به صورت یک ماتریس (30×256) نشان داده می‌شود که ۳۰ تعداد کانال‌ها بوده و ۲۵۶ تعداد نمونه‌های زمانی است.

- تعداد کل داده‌ها: ۲۱۰ آزمایش

- داده‌های آموزشی: ۱۶۵ آزمایش

- کلاس مربوط به آنها مشخص شده است:

- ۱: تصور حرکت پا

- ۰: عمل تفریق ذهنی

- داده‌های آزمون: ۴۵ آزمایش

- هدف: تعیین برچسب داده‌های آزمون