

تمرین سوم علوم اعصاب محاسباتی

امیرکیوان محتشمی ۹۴۱۰۹۳۷۹

پیمان جبارزاده ۹۴۱۰۹۲۸۴

۲۰ اردیبهشت ۱۳۹۷

۱ آشنایی با مقاله

۱.۱ هدف پژوهش و تفاوت آن با کارهای قبلی

هدف پژوهش، بررسی کارایی یک سیستم BCI مبتنی بر P300، در هجی کردن کلمات است. سیگنال P300 زمانی در EEG ظاهر می‌شود، که یک اتفاق با احتمال کم (مثل ظاهر شدن کاراکتر مورد نظر روی صفحه) رخ دهد. این سیستم‌ها در مقایسه با سایر سیستم‌های BCI از سرعت و دقت بیشتری برخوردار هستند و در نتیجه گزینه‌ی مناسبی برای هجی کردن کلمات به نظر می‌رسند. پژوهش‌های مشابه قبلی عموماً با تعداد subject کم انجام شده‌اند. تمرکز این پژوهش‌ها بر بررسی اثر پارامترهای مختلف (مثل سایز ماتریس حروف) بر کارایی سیستم است. در این پژوهش، پارامترها ثابت هستند و هدف بررسی کارایی سیستم برای تعداد زیادی از افراد است. علاوه بر این در این پژوهش، زمان کوتاه‌تری به نسبت سایر پژوهش‌ها صرف آموزش می‌شود. به نظر می‌رسد یکی از اهداف این آزمایش، بررسی امکان استفاده از چنین سیستم‌هایی در عمل است.

۲.۱ توصیف آزمایش و تفاوت دو پروتکل انجام آن

این آزمایش در دو قالب RC و SC انجام شده است. هر فرد، به انتخاب خود در یکی یا هر دوی این قالب‌ها شرکت کرده است. در هر دو قالب، از افراد خواسته شده تا رو به روی یک لپ تاپ بنشینند و تعداد دفعاتی را بشمارند که حرف مورد نظرشان روشن می‌شود. هدف از شمارش، حفظ تمرکز افراد است. هر فرد ابتدا سعی کرده است تا حروف کلمه‌ی WATER را یکی یکی هجی کند و در این مدت فیدبکی از سیستم دریافت نمی‌کند. پس از پردازش داده‌های جمع‌آوری شده، هر فرد سعی کرده است تا حروف کلمه‌ی LUCAS را هجی کند. در این مرحله، سیستم حرفی را که تشخیص داده را به فرد نمایش می‌دهد.

در هر دو پروتکل، یک ماتریس شامل حروف بزرگ انگلیسی و اعداد صفر تا نه روی صفحه قرار دارد که عناصر آن خاکستری هستند و هرگاه یک عنصر روشن شود، به رنگ سفید در می‌آید. البته در شکل مقاله، برای پروتکل SC به جای عدد صفر حرف _ قرار دارد اما در متن مقاله گفته شده که ماتریس‌ها شامل اعداد صفر تا نه هستند.

در پروتکل RC، در هر مرحله یک سطر یا یک ستون از ماتریس روشن می‌شود اما در پروتکل SC در هر مرحله تنها یکی از حروف روشن می‌شود. این دو پروتکل در جزئیاتی مثل زمان روشن ماندن عناصر نیز با هم تفاوت‌هایی دارند. برای مثال در پروتکل RC هر سطر یا ستون به مدت ۱۰۰ میلی ثانیه روشن می‌ماند اما هر حرف در پروتکل SC تنها ۶۰ میلی ثانیه روشن می‌ماند. سیستم برای تشخیص یک حرف، به صورت تصادفی یک سطر یا ستون (در مدل RC) یا حرف (در مدل SC) را روشن می‌کند. در مدل سطری این کار تا زمانی ادامه پیدا می‌کند که هر سطر و هر ستون حداقل ۱۵ بار روشن شده باشند. در مدل SC این کار تا زمانی ادامه پیدا می‌کند که هر حرف حداقل ۱۵ بار روشن شده باشد.

۳.۱ شیوه جداسازی target و non-target

داده‌ی این آزمایش، سیگنال‌های EEG در ۸ نقطه هستند که با فرکانس ۲۵۶ هرتز جمع‌آوری می‌شوند. این داده‌ها ابتدا از طریق یک فیلتر میان‌گذر به بازه‌ی 0.5 تا 30 هرتز محدود می‌شوند و سپس فرکانس آن‌ها از طریق down-sampling به ۶۴ هرتز کاهش می‌یابد.

در این آزمایش یک LDA برای هر فرد به صورت جداگانه (و همچنین هر پروتکل) ایجاد می‌شود. به کمک این LDA در هر بار روشن شدن بخشی از ماتریس، مشخص می‌شود آیا حرف مورد نظر فرد روشن شده است (target) یا خیر (non-target). درواقع خروجی LDA برای کلاس target

یک عدد است. سیستم برای هر حرف در مدل SC یا هر سطر و هر ستون در مدل RC پانزده نمونه به دست آورده و خروجی LDA آن‌ها را با هم جمع می‌کند. بنابراین در مدل SC، به ازای هر کاراکتر و در مدل RC به ازای هر سطر و هر ستون یک عدد دارد. حال در مدل SC کاراکتر با بیشترین عدد و در مدل RC محل تلاقی سطر و ستون با بیشترین عدد به عنوان کاراکتر تشخیص داده می‌شود.

۴.۱ نتایج پژوهش

این آزمایش نشان می‌دهد که امکان دستیابی به دقت بالا برای هجی کردن کلمات به کمک سیستم‌های BCI مبتنی بر P300 وجود دارد. 88.9% از ۸۱ نفری که پروتکل RC را آزمایش کردند و همچنین 77.9% از افرادی که پروتکل SC را آزمایش کردند، با دقت ۸۰ تا ۱۰۰ درصد (یعنی با حداکثر یک خطا) حرف مورد نظر خود را هجی کردند. با مقایسه‌ی این نتایج با نتایج پژوهش‌های گذشته، برتری روش‌های مبتنی بر P300 بر روش‌های دیگری مثل روش‌های مبتنی بر تصور حرکت مشخص می‌شود. همچنین به نظر می‌رسد دقت روش RC بهتر از SC است که البته ممکن است به دلیل طولانی تر بودن فرآیند انتخاب کاراکتر در روش SC باشد. همچنین در این آزمایش از طریق پرسش‌نامه تاثیرات احتمالی جنسیت، سطح تحصیلات، مدت زمان کار، کشیدن سیگار، نوشیدن قهوه و خواب شب گذشته بر دقت هجی کردن بررسی شده است که تنها مورد تاثیر گذار مدت خواب شب گذشته است. به نظر می‌رسد افرادی که شب گذشته کمتر از ۸ ساعت خوابیده‌اند دقت بهتری به دست آورده‌اند.

۲ آشنایی با دیتاست

۱.۲ توضیح فایل های دیتا

هر فایل دیتا که داده های یکی از سابجکتهای در آن قرار دارد، شامل دو بخش train و test است که بخش train برای آموزش classifier و بخش test برای تست classifier است. در هر یک از این بخش ها جدولی شامل ۱۱ سطر وجود دارد که ستون های این جدول، اطلاعات نمونه برداری های مختلف را ذخیره کرده است. اطلاعات سطرها نیز به صورت زیر است:

- سطر اول زمان نمونه برداری را نشان می دهد.
- سطر دوم تا نهم، فیچرهای مکان های مختلف نمونه برداری را نشان می دهد.
- سطر دهم محل فلش زدن را نشان می دهد. (عدد صفر به معنی خاموش بودن تمام جدول است)
- در صورتی که آزمایش RC باشد، اعداد آن بین ۱ تا ۱۲ است که اعداد ۱ تا ۶، ستون ها را از چپ به راست و اعداد ۷ تا ۱۲، سطرها را از بالا به پایین نشان می دهند.
- در صورتی که آزمایش SC باشد، اعداد آن بین ۱ تا ۳۶ خواهد بود که شماره خانه ی روشن در جدول را نشان می دهد.
- سطر یازدهم شامل اطلاعات خانه ی هدف است. در صورتی که خانه ی هدف روشن باشد، مقدار آن ۱ و در صورتی که خاموش باشد مقدار آن ۰ است.

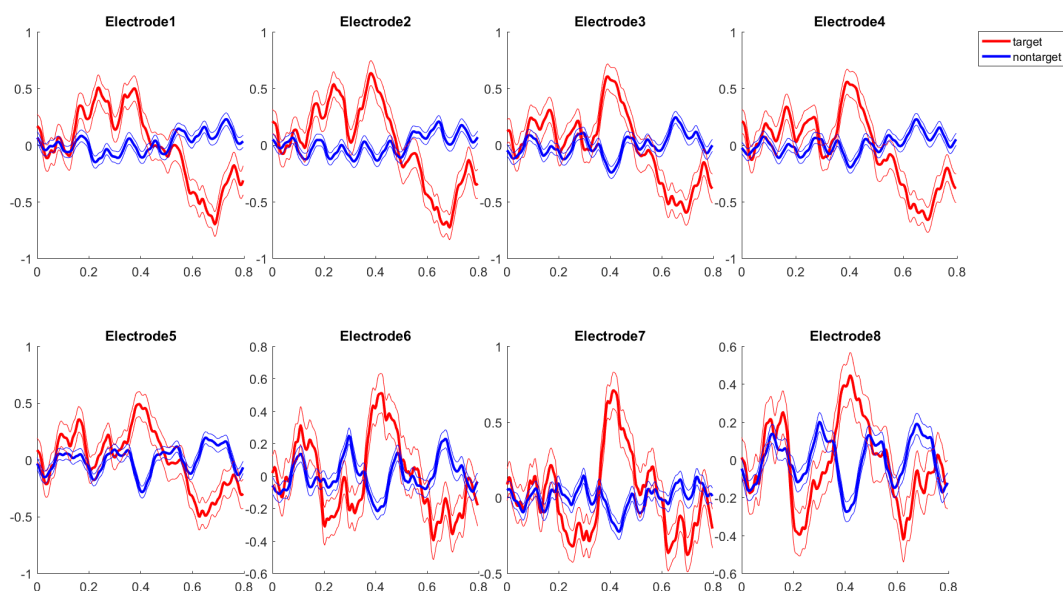
۲.۲ پروتکل سابجکتهای

سابجکتهایی که اعداد سطر دهم دیتای آنها بین ۰ تا ۱۲ است از پروتکل RC استفاده کرده اند و سایر سابجکتهای (که اعداد سطر دهم دیتای آنها بین ۰ تا ۳۶ است) از پروتکل SC استفاده کرده اند. با توجه به نکته گفته شده، نوع پروتکل ده سابجکت آزمایش به صورت زیر است (به کمک تابع DetermineProtocol اطلاعات زیر بدست آمده است):

- سابجکتهای ۱ و ۲ از نوع SC است.
- سابجکتهای ۳ تا ۱۰ از نوع RC است.

۳ بررسی ERP به روش کلاسیک

در این بخش، سابیجت سوم به عنوان نمونه انتخاب شده است.



شکل ۱: سابیجت سوم

۱.۳ بررسی نوسان‌ها

با بررسی ۸ تصویر حاصل از نمونه انتخاب شده، متوجه می‌شویم که non-target نوسان کمی دارد و از الگو ثابتی پیروی می‌کند. در مقابل آن target دارای نوسان‌های زیادی است.

۲.۳ بهترین الکتروود و زمان

تقریباً همه الکتروودها در جداسازی نمودار target و non-target شبیه هم عمل می‌کنند. اما الکتروود ۷ اندکی بهتر از بقیه عمل کرده است و در زمان‌های ۰/۴۵ و ۰/۷ ثانیه، تفاوت زیادی را ایجاد کرده است.

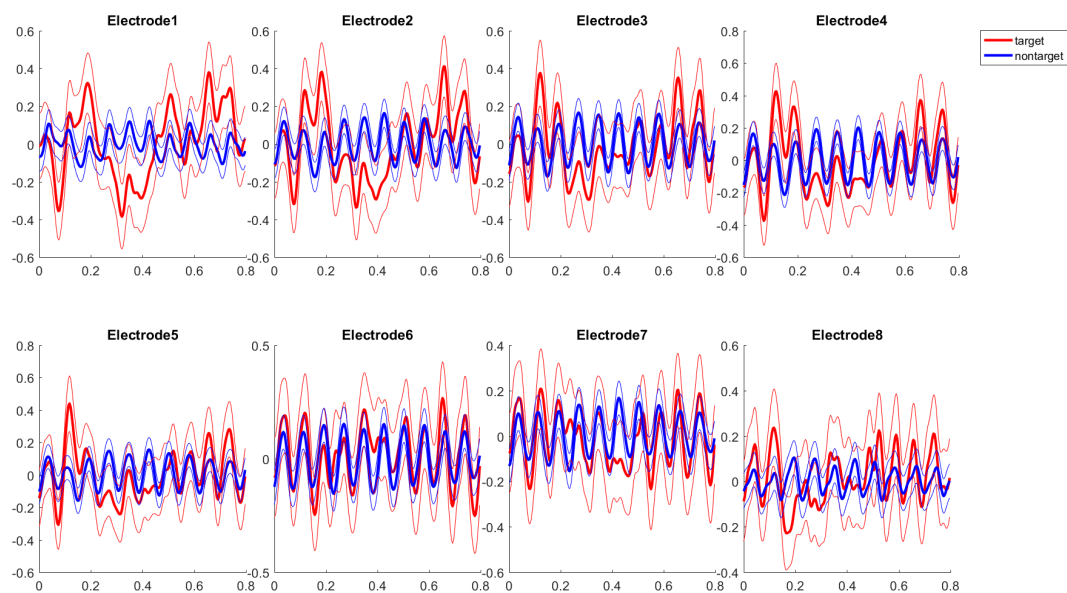
۳.۳ بررسی تمام سابیجت‌ها

۱.۳.۳ نوسان‌ها

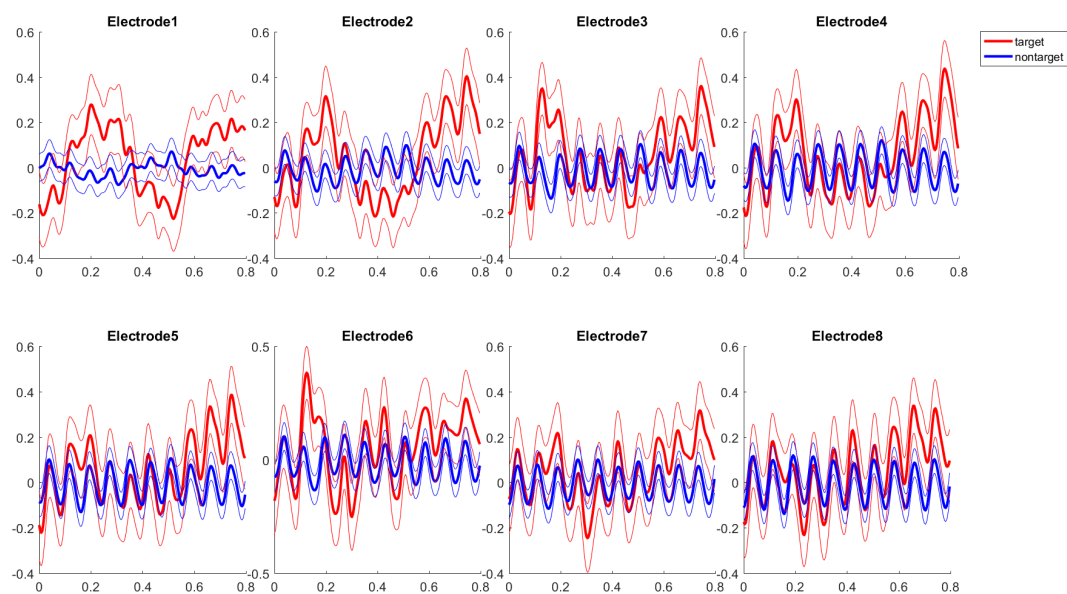
در تمام سابیجت‌ها نمودار target دارای نوسان‌های زیادی است. اما نمودار non-target در سابیجت‌های مختلف کمی متفاوت است. در بعضی سابیجت‌ها مثل سابیجت سوم نوسانات کمی دارد، و در بعضی سابیجت‌ها مثل سابیجت هفتم نوسانات زیادی دارد. اما در تمام سابیجت‌ها، نوسانات non-target از الگو ثابتی پیروی می‌کند.

۲.۳.۳ بهترین الکتروود و زمان

بهترین الکتروود و زمان در سابیجت‌های مختلف متفاوت است. برای مثال در سابیجت ششم بهترین الکتروود، الکتروود اول که در ثانیه‌های ۰/۲ و ۰/۵ بهترین تفاوت را ایجاد کرده است. در مقابل آن سابیجت نهم وجود دارد که بهترین الکتروود آن، الکتروود هشتم است و در ثانیه‌های ۰/۱ و ۰/۲ بهترین تفاوت را ایجاد می‌کند. هر دو این سابیجت‌ها نتایج متفاوتی با سابیجت سوم دارند که در قسمت‌های قبل‌تر بررسی شده است. ولی به‌طور کلی الکتروودهای ۲ و ۷ در همه سابیجت‌ها تفاوت خوبی را ایجاد کرده‌اند.



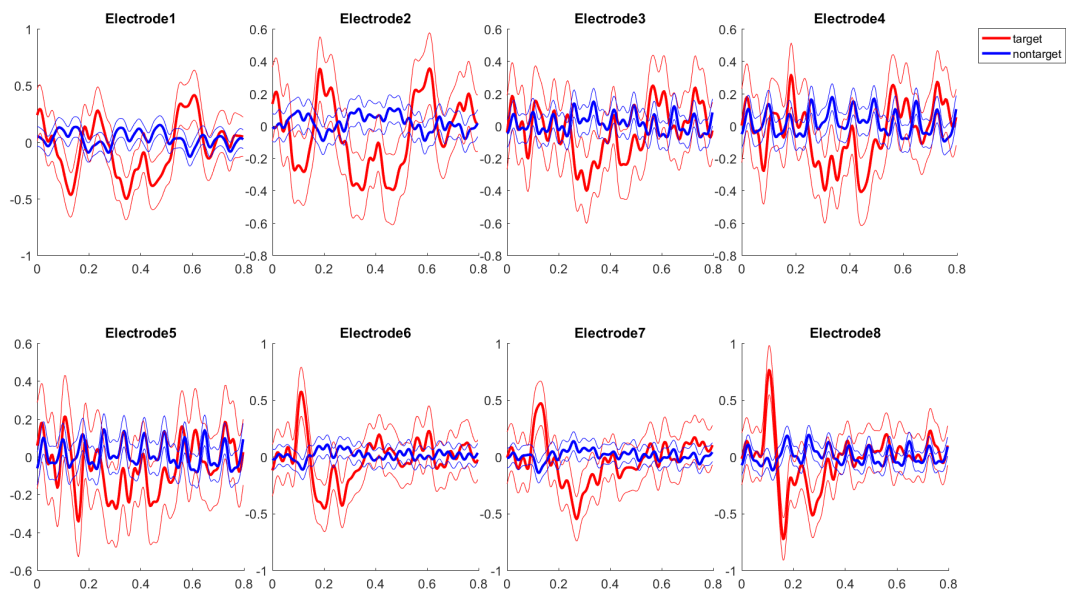
شکل ۲: سابیجت هفتم



شکل ۳: سابیجت ششم

۳.۳.۳ تاثیر پارادیم‌های مختلف

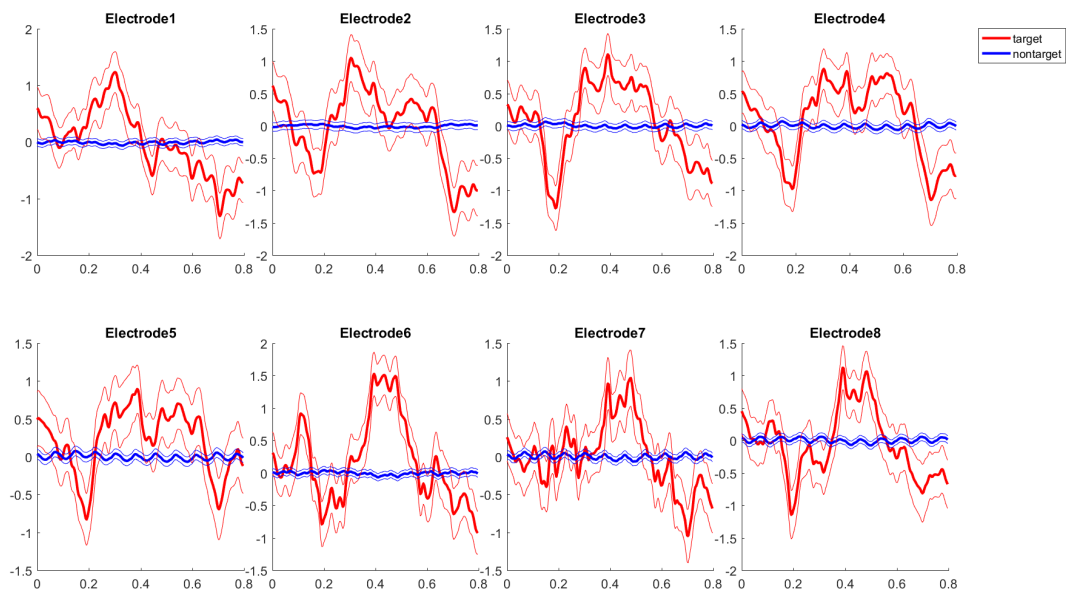
در بین سابیجت‌ها، سابیجت ۱ و ۲ که از پارادیم SC استفاده می‌کنند، تفاوت بیش‌تری بین نمودار target و non-target آن‌ها وجود دارد و برخلاف مقاله، بهتر از RC عمل می‌کنند.



شکل ۴: سابیجت نهم

۴.۳.۳ بهترین سابیجت

با توجه به شکل سابیجت‌ها، سابیجت اول بهترین عملکرد را انتظار داریم داشته باشد.



شکل ۵: سابیجت اول

۴.۳ خلاصه

با مشاهده تمام ده سابجکت، متوجه می‌شویم که نمودار target نوسانی تر از سیگنال non-target بوده و این موضوع در روش‌های SC بهتر دیده می‌شود.

در اکثر سابجکت‌ها، الکترودهای ۲، ۷ و ۸ بهتر عمل می‌کنند و تفاوت بیش‌تری را نشان می‌دهند. به‌طور خاص، الکتروود دوم در ثانیه ۰/۲ در اکثر سابجکت‌ها تفاوت خوبی را نشان می‌دهد.

۴ الگوریتم P300-Speller

۱.۴ مقایسه‌ی دقت‌ها

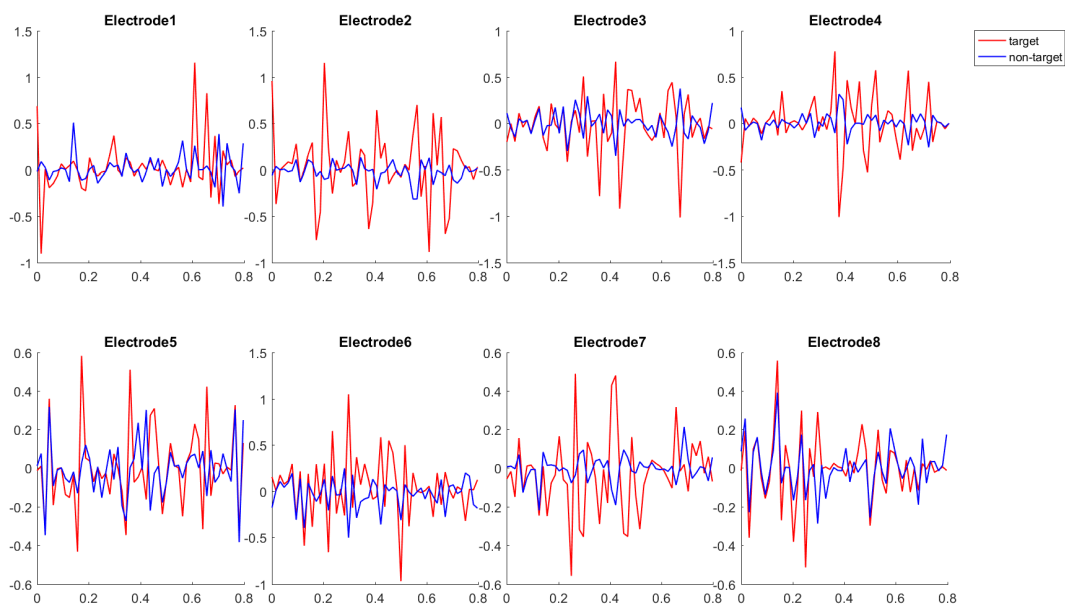
دقت روی Test، Train و Cross Validation به ترتیب تقریباً برابر ۹۷٪، ۷۵٪ و ۷۳٪ است. تفاوت دقت تمرین با تست، نشان‌دهنده‌ی over-fitting است. اگر عدم توانایی generalization به خاطر fit شدن روی نویز داده‌ی تمرین باشد، یکی از روش‌ها برای رفع این مشکل می‌تواند استفاده از regularization باشد که در قسمت آخر تمرین به بررسی آن می‌پردازیم. لازم به ذکر است، درصد بالای دقت ناشی از بالانس نبودن دیتاست است. در دیتاست تعداد داده‌های با لیبل صفر، به مراتب بیشتر از داده‌های با لیبل یک است. برای مثال در داده‌های پروتکل RC، از میان ۹۰۰ آزمایش، ۷۵۰ آزمایش (۸۳٪) لیبل صفر دارند.

۲.۴ تشخیص سیستم

کلمه‌ی تشخیص داده شده LUKAP است.

۳.۴ مهم‌ترین الکتروود و بازه‌ی زمانی

سیگنال‌های ERP مربوط به target و non-target را در ضرایب LDA ضرب می‌کنیم و نمودار آن را رسم می‌کنیم.



شکل ۶: بررسی موثرترین الکتروود در LDA در سابجکت سوم

براساس این نمودار بهترین الکتروودها که تفاوت دو نوع سیگنال در آن مشخص است، الکتروودهای ۲ و ۷ هستند. این تفاوت در بازه‌ی زمانی نزدیک به ۲۰۰ میلی ثانیه بعد از نمایش مشهودتر است.

۴.۴ نتایج برای سایر سابجکت‌ها

نتیجه جدول زیر است:

۱.۴.۴ هیستوگرام درصد صحت

نمودارهای مشابه برای تمامی سابجکت‌ها در پوشه‌ی output/phase4 قرار دارد.

train_accr	test_accr	cross_accr	train_word	test_word
0.985555556	0.963333333	0.952592593	lukas	f0kr0
0.99037037	0.972222222	0.967777778	lukas	lukas
0.968888889	0.753333333	0.732222222	lukas	lukap
0.981111111	0.794444444	0.764444444	lukas	luias
0.98	0.762222222	0.797777778	lukas	water
0.951111111	0.743333333	0.71	lukas	wcw2f
0.963333333	0.77	0.762222222	lukas	eater
0.984444444	0.825555556	0.792222222	lukas	water
0.978888889	0.831111111	0.81	lukas	water
0.981111111	0.825555556	0.757777778	lukas	wgter

شکل ۷: جدول نتیجه برای تمامی سابجکت‌ها

۲.۴.۴ تعداد تشخیص‌های درست

سیستم برای سابجکت‌های ۲، ۵، ۸، و ۹ به درستی کلمه‌ها را تشخیص می‌دهد. کلمه‌ی سابجکت ۲، lukas و کلمه‌ی سایر سابجکت‌ها water است.

۳.۴.۴ یکسان بودن بهترین الکتروود

در تمامی سابجکت‌ها، در همان دو الکتروود و در همان بازه‌ی زمانی نزدیک به ۲۰۰ میلی ثانیه بعد از روشن شدن، تفاوت نسبتاً مشهودی بین دو نوع سیگنال وجود دارد. با این وجود برای مثال در سابجکت چهار، تفاوت مشهودتری در الکتروود سوم دیده می‌شود اما از آنجا که این تفاوت در سایر سابجکت‌ها قابل مشاهده نیست، به نظر همان فعالیت الکتروودهای ۲ و ۷ در زمان ۲۰۰ میلی ثانیه، معیار بهتری هستند. تفاوت بین این دو سیگنال در این دو الکتروود و در این دو زمان در نتایج قسمت سوم نیز قابل مشاهده است. لازم به ذکر است منظور از زمان ۲۰۰ میلی ثانیه، بازه‌ی حدوداً ۱۰۰ میلی‌ثانیه‌ای حول زمان ۲۰۰ میلی ثانیه است.

۴.۴.۴ تفاوت بین پروتکل‌ها

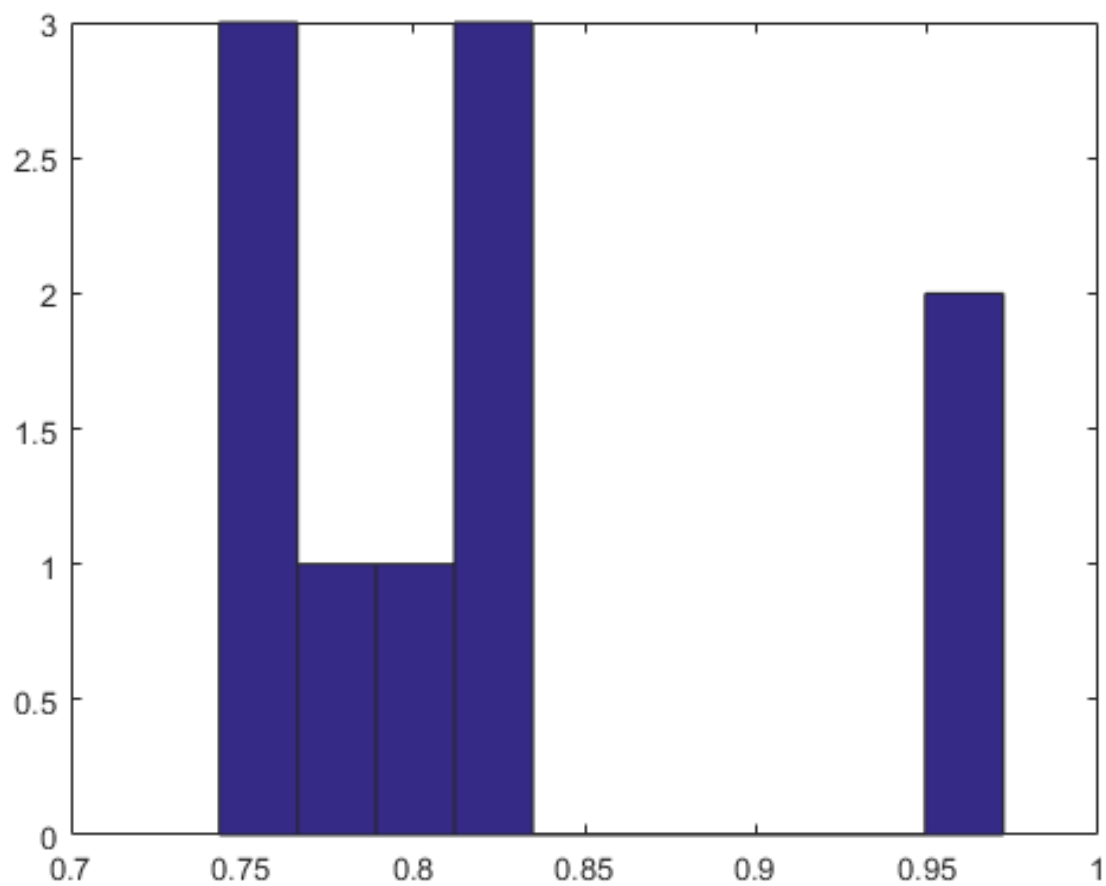
یکی از دو نفری که در تست SC شرکت کرده است، کلمه را به درستی هجی کرده است اما نفر دیگر حروف کاملاً نامرتبیطی را هجی کرده است. سه نفر از هشت نفری که تست RC را انجام داده‌اند، کلمه را کاملاً درست هجی کرده‌اند و چهار نفر دیگر با یک غلط کلمه را هجی کرده‌اند. با توجه به این موارد و با این‌که درصد بیشتری از افراد آزمایش اول، کلمه را به طور کامل درست هجی کرده‌اند (۵۰٪ در مقابل ۳۷٪)، آزمایش RC پروتکل مناسب‌تری باشد چرا که ۷ نفر از ۸ نفر (۸۷٪) کلمه را با حداکثر یک غلط هجی کرده‌اند.

۵.۴.۴ مقایسه با جواب بخش قبل

خیر. در بخش قبل به نظر می‌رسید که باید امکان جداسازی بهتری برای پروتکل SC وجود داشته باشد.

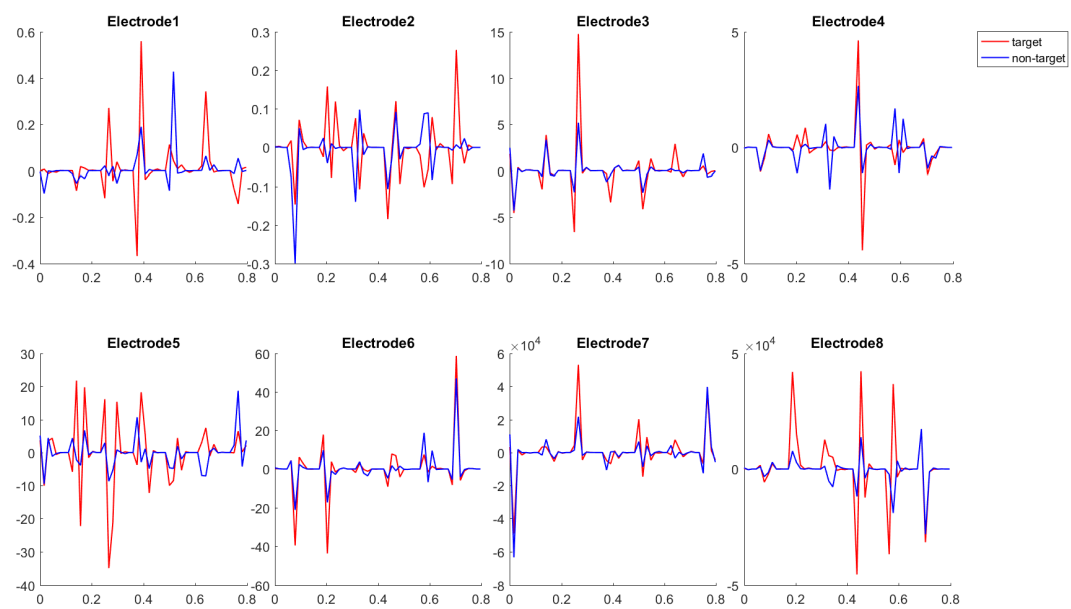
۵.۴ خلاصه

با وجود داده‌ی کم و مشکلات داده، همچنان سیستم برای ۸۰٪ از سابجکت‌ها توانست کلمه را با حداکثر یک غلط تشخیص دهد. تفاوت بین دقت تمرین و تست نشان‌دهنده‌ی overfitting است. در قسمت بعد به regularization به عنوان یکی از راه‌حل‌های این مشکل پرداخته شده است. در



شکل ۸: هیستوگرام درصد صحت سابی‌ها

تمامی سابی‌ها، در همان دو الکتروود و در همان بازه‌ی زمانی نزدیک به ۲۰۰ میلی ثانیه بعد از روشن شدن، تفاوت نسبتاً مشهودی بین دو نوع سیگنال وجود دارد.

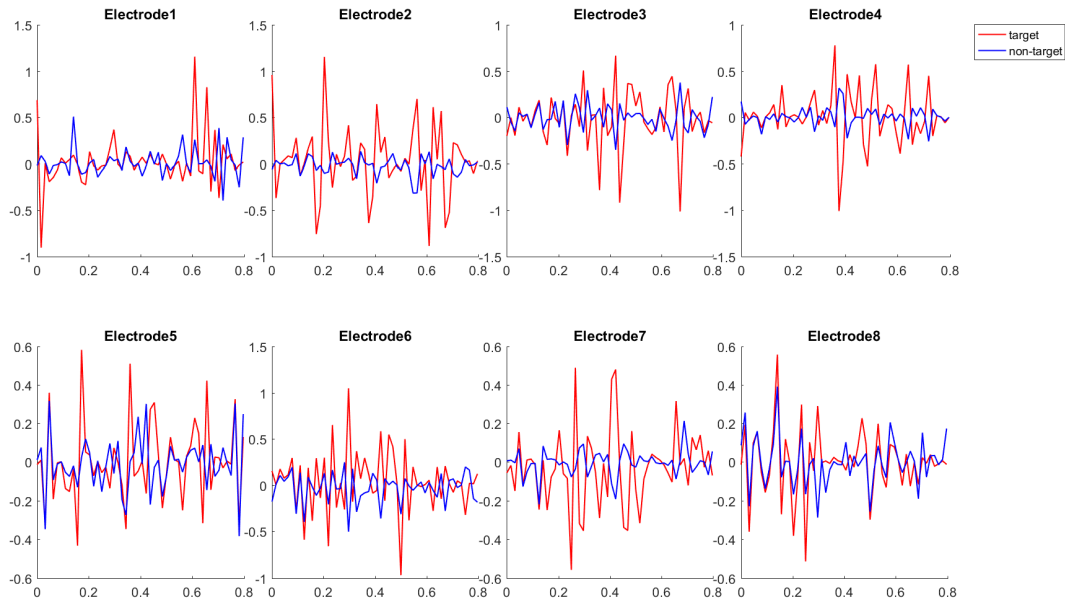


شکل ۹: بررسی موثرترین الکترود در LDA در سابجکت چهارم

۵ استفاده از Regularization

همانطور که در بخش قبل دیدیم، دقت train خیلی بهتر از test بود که علت آن overfitting می‌تواند باشد. برای مقابله با آن از Regularization استفاده می‌کنیم.

اگر به نمودارهای سابیجت سوم توجه کنیم، متوجه می‌شویم تفاوت بین target و non-target در هنگام استفاده از Regularization بهتر قابل مشاهده است.



شکل ۱۰: بررسی تفاوت target و nontarget در LDA در سابیجت سوم با استفاده از Regularization

علاوه بر این، اگر به خروجی تمام سابیجت‌ها توجه کنید، مشاهده خواهید کرد که به کمک Regularization نتایج بهتری به دست آورده ایم. در بخش قبل ۷۴٪ حروف درست تشخیص داده شده بودند و فقط ۴ نفر توانسته بودند کلمه را کامل تشخیص دهند، اما در روش جدید ۹۰٪ حروف درست تشخیص داده شده اند و ۷ نفر توانسته بودند کلمه را کامل تشخیص دهند. البته در قسمت قبل، داده‌های آموزش کامل تشخیص داده شده بود، اما در این بخش یکی از داده‌های آموزش lxkas تشخیص داده شده است، که یکی از نتایج جلوگیری از overfitting می‌تواند باشد.

train_accr	test_accr	cross_accr	train_word	test_word
0.978889	0.967037	0.96963	lukas	l0k9s
0.983333	0.972222	0.972963	lukas	lukas
0.871111	0.855556	0.821111	lxkas	lukas
0.938889	0.87	0.881111	lukas	lukas
0.896667	0.858889	0.854444	lukas	water
0.87	0.808889	0.815556	lukas	waxex
0.906667	0.815556	0.837778	lukas	2ater
0.918889	0.856667	0.854444	lukas	water
0.912222	0.881111	0.878889	lukas	water
0.923333	0.876667	0.877778	lukas	water

شکل ۱۱: جدول نتیجه برای تمامی سابجکت‌ها