

با اسمه تعالی



دانشگاه صنعتی شریف

دانشکده مهندسی برق

## درس علوم اعصاب محاسباتی

گزارش پروژه درس

سری دوم

یافتن معیاری برای عمق خواب مبتنی بر سیگنال‌های EEG

استاد درس:

دکتر کربلایی آقاجان

امیرحسین افشار راد

۹۵۱۰۱۰۷۷

طaha انتصاری

۹۵۱۰۱۱۱۷

۱۳۹۷ بهار

**قسمت اول) آشنایی با مقاله پژوهش اصلی****Terminology:****Slow-wave=:** موج آرام (فرکانس پایین)=SW**Slow Wave Power=:** انرژی موج آرام (فرکانس پایین)=SWP**Slow Wave microcontinuity=:** SW% = مقداری از انرژی موج آرام که در آینده نزدیک ۰/۰۲ ثانیه ادامه دارد**Non Rapid Eye Movement=:** NREM**Electroencephalogram=:** EEG**۱ - با مطالعه abstract و introduction، هدف کلی این مقاله را توضیح دهید.**

هدف این پژوهش یافتن رابطه بین گین مدار فیدبک SW و اثر سن فرد، زمان شب و داروی خواب می‌باشد. همچنین این پژوهش در صدد بررسی این است که آیا گین فیدبک میتواند نمایانگری از اثر جنسیت بر روی SWP باشد یا نه. پژوهش‌ها نمایانگر این می‌باشند که SW% بهتر میتواند عمق خواب را تخمین بزند.

در این پژوهش تاکید بر روی استفاده از SW% با دلیل عدم تاثیرپذیری این ویژگی سیگنال از ویژگی‌های آناتومی بدن و جنسیت فرد می‌باشد. همچنین روش SWP هم در حالت‌های ممکن به صورت "نسبی" پیاده سازی شده است. اما استفاده از روش نسبی SWP به خاطر نبود دلیل دقیق فیزیولوژی پشت بسیاری از این مقادیر، قابل استناد نمی‌باشند.

**۲ - آزمایش انجام شده برای این پژوهش را توضیح دهید.**

برای آزمایش ۲۲ نفر مرد و زن از بازه‌های سنی متفاوت انتخاب شده‌اند. داده از افراد در طی ۲ شب گرفته شده است. هر فرد در یکی از شب‌ها از داروی خواب temazepam استفاده کرده است.

در روز بعدی از افراد نیز در مورد کیفیت خوابشان سوال‌هایی پرسش شده است که این پرسش‌ها برای بررسی برداشت خود افراد از خواب با تحلیل‌های حاصل از پردازش سیگنال‌ها برای بررسی اثر داروی خواب می‌باشد. که البته جواب این سوال منفی می‌باشد و تا کنون همبستگی بین برداشت خود فرد از خواب و طول کلی خواب NREM مشاهده نشده است. البته تاثیر داروی خواب را میتوان در تفاوت‌های موجود در SW% دید.

**۳ - تعبیر مقاله از عمق خواب را توضیح دهید. همچنین تفاوتش را با تعاریف پیشین، و ایراد آن تعاریف را توضیح دهید.**

پژوهش با مدل کردن تولید موج آرام توسط شکل موجود در سوال بعد، عمق خواب را متناظر با ضرب فیدبک شبکه نورونی بیان می‌کند. در واقع عمق خواب متناظر است با افزایش فیدبک شبکه نورونی که باعث تولید SW می‌شود.

نتایج پژوهش بیانگر این است که SW%， با توجه به این که بر پایه مدلی فیزیولوژیک از خواب می‌باشد، عمق خواب را بهتر از SWP نشان میدهد.

از آنجایی که تغییر عمق خواب با تغییرات سنکرون در دامنه فرکانس‌های پایین سیگنال EEG همراه می‌باشد، بسیاری از پژوهش‌های پیشین از انرژی موج فرکانس کم(SWP) استفاده کرده‌اند. این در حالی می‌باشد که SWP از آناتومی بدن تاثیر می‌پذیرد. همچنین جنسیت افراد تاثیر بسزایی در مقدار این ویژگی دارد به طوریکه زن‌ها دارای مقدار ۲ برابر مقدار مرد‌ها می‌باشند در حالی که اثبات شده است که این نمایانگر عمق خواب افراد نیست (ارجاع اثبات به مقالات ارجاعی ۵ و ۶). همچنین این عدم قطعیت هنگامی که اثرات

"درون فردی" بر روی کل زمان خواب NREM مطالعه می‌شوند نیز موجود می‌باشد. محققان با ذکر اینکه امکان گریز از این مشکلات با بررسی نسبی مقادیر SWP وجود دارد این نکته را بیان می‌کنند که این مقادیر نسبی پایه فیزیولوژی ندارند.

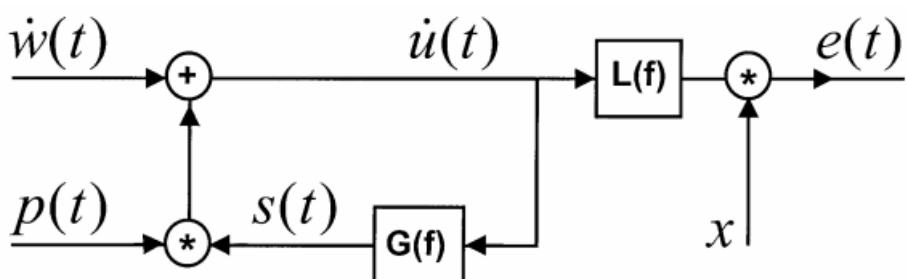
جدا از توضیح بالا مبنی بر عدم وجود علت برای برخی از این مقادیر نسبی، به دلیل اینکه این مقادیر نسبی در حوزه فرکانس حساب می‌شوند، عدم دقت زمانی کافی باعث کم شدن دقت و در نتیجه تخمین بد از جواب حقیقی می‌شود. البته این دقت زمانی کم برای روش گین فیدبک تاثیری ندارد چرا که این مدل تماماً در حوزه زمان پیاده سازی می‌شود.

۴- قسمت دوم مقاله (THE FEEDBACK MODEL) را به طور کلی مطالعه کنید. (ممکن است بعضی از مفاهیم به کل برایتان جدید باشد، برای همین نیازی نیست که دقیق شوید). کلیت مدل پیشنهادی را توضیح دهید. دلیل نویسنده‌گان مقاله برای انتخاب این مدل چه بوده است؟ به طور شهودی ارتباط این مدل با عمق خواب و سیگنال EEG را توضیح دهید با توجه به ویژگی‌های فیزیولوژیکی، مدل ما باستی ۳ ویژگی را داشته باشد:

۱. اینکه مقدار گین فیدبک با عمق خواب متناسب می‌باشد.

۲. اثرات تصادفی و غیر پیش بینی شده بیرونی که مدل را تحت تاثیر قرار میدهند را در نظر بگیرد

۳. گین فیدبک بزرگتر متناظر است با اینکه SWP بزرگتر باشد.



مدلی که پیاده سازی شده است به صورت بالا می‌باشد. در مدل بالا هر اثر نویز با ورودی  $W(t)$  که به صورت نویز سفید گوسی فرض شده است مدل می‌شود.

فیلتر  $G(f)$  که موج آرام را از خروجی استخراج می‌کند و پس از آن با ضرب شدن این مقدار در مقدار گین فیدبک  $p(t)$ ، فیدبک نهایی مدل آماده می‌باشد.  $p(t)$  نمایانگر عمق خواب می‌باشد که مشخص می‌کند که چه کسری از سیگنال کنونی موجود در مغز، در آینده نزدیک ادامه می‌یابد.

سیگنال موجود در مغز پس از عبور از یک فیلتر پایین گذر  $L(f)$  و یک ضریب نامرتبط به خواب،  $x$ ، (ویژگی‌های آناتومی سر) در الکتروود های EEG ضبط می‌شود.

در مدل بالا تابع  $G(f)$  یک فیلتر میان گذر خطی می‌باشد که مطابق آنچه که در صورت مقاله ذکر شده تعریف می‌شود. برای این پژوهش مورد نظر فرکانس های قطع پایین و بالا به ترتیب  $0/5$  و  $2$  هرتز می‌باشند در مورد این فیلتر ذکر شده است که تنها با شبی  $6db/oct$  افت می‌کند. همان‌گونه که مقاله توضیح میدهد، علت عدم افت شدید اندازه این فیلتر، ماهیت خود داده های موج آرام می‌باشد، چرا که در بر داشتن اطلاعاتی از محدوده‌های فرکانسی مجاور نیز می‌تواند حاوی اطلاعات مفیدی باشد.

6. M. S. Mourtazaev, B. Kemp, A. H. Zwinderman, and H. A. C. Kamphuisen, “Age and gender affect different characteristics of slow waves in the sleep EEG,”

10. Accurate measurement of flash-evoked alpha attenuation B.Kemp

به نظر میرسد از اولین استفاده های این مدل فیدبک در ارجاع فوق (شماره ۱۰) صورت گرفته باشد.

قسمت دوم) آشنایی با دیتاست

(۲-۲)

تابع Feature Extraction را نوشتیم که فایل آن در کنار فایل اصلی ضمیمه شده است. ورودی و خروجی‌های این تابع به شرح زیر است:

خروجی‌ها:

- t : زمان آغاز هر بازه ۱۰ ثانیه‌ای
- state : وضعیت خواب در طول هر کدام از بازه‌های ۱۰ ثانیه‌ای
- X : ماتریسی با ۱۰ سطر و N ستون که N تعداد بازه‌های ۱۰ ثانیه‌ای مربوط به سیگنال اصلی می‌باشد. سطرهای اول تا چهارم انرژی باندهای دلتا، تتا، آلفا، و بتای مربوط به الکترود Fpz و سطرهای پنجم تا هشتم مشابه چهار سطر اول، این بار مربوط به الکترود Oz هستند که هر درایه توان متوسط مربوط را در بازه ۱۰ ثانیه‌ای متناظر نشان می‌دهد. دو سطر آخر نیز توان متوسط مربوط به سیگنال‌های EOG و EMG می‌باشند.

ورودی:

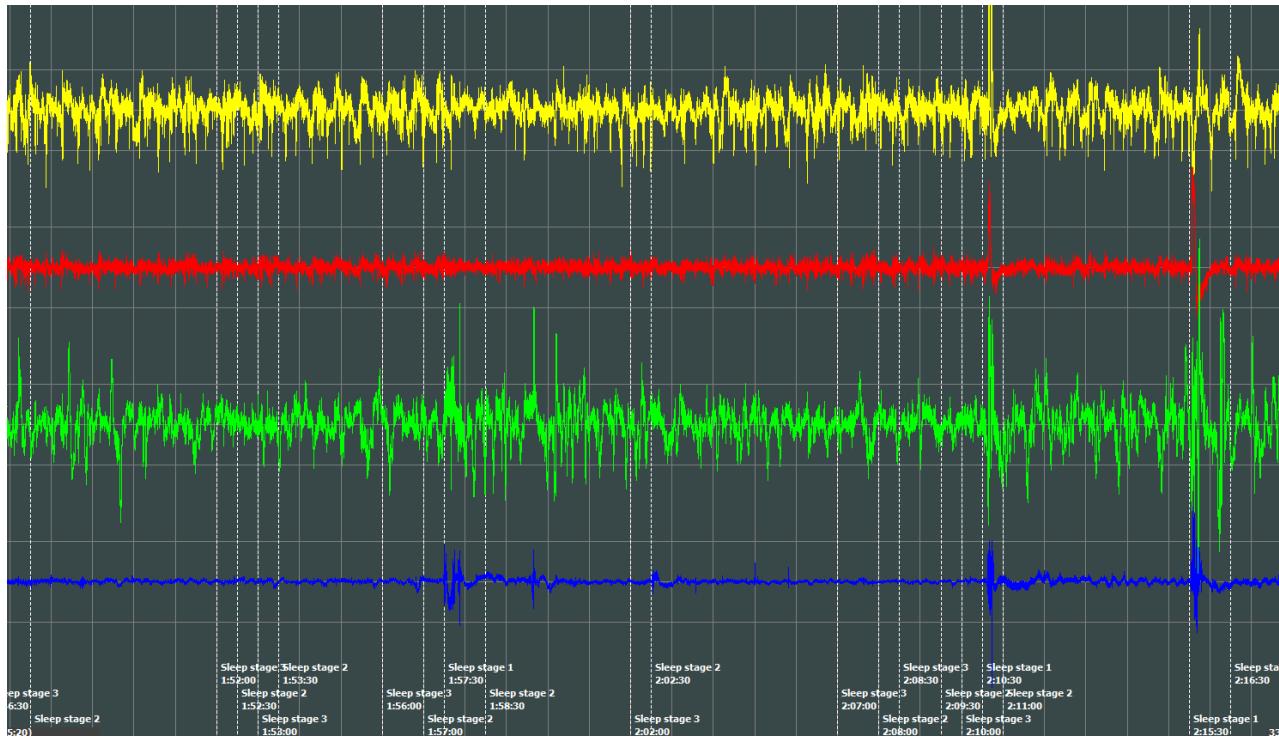
- edf\_path: آدرس فایل edf مربوطه می‌باشد.
- hyp\_path: آدرس فایل هایپنوگرام مربوطه می‌باشد.

نکته ۱: برای محاسبه توان متوسط سیگنال در یک بازه، از رابطه‌ی زیر استفاده شده است:

$$\bar{P} = \frac{1}{N} \sum_{n=1}^N x[n]^2$$

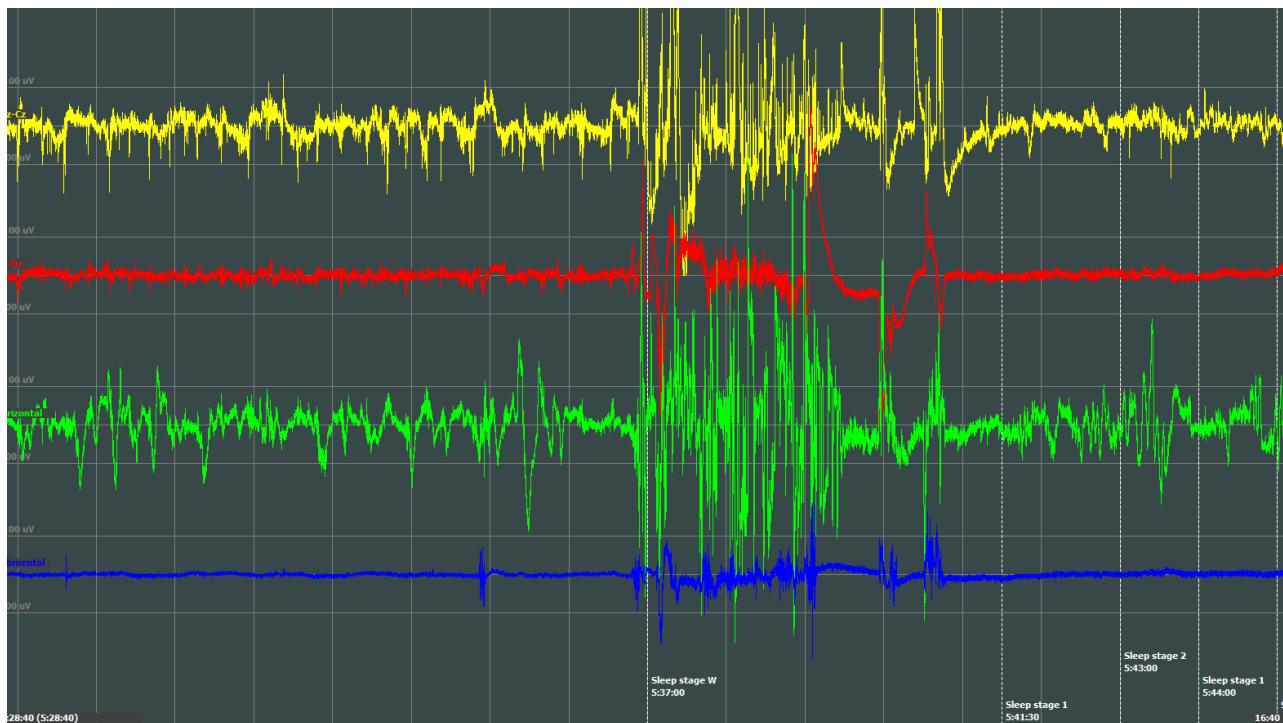
نکته ۲: برای انتخاب طول پنجره‌های زمانی، گزینه‌های متعددی وجود دارد. برای انجام این انتخاب، باید سعی کرد به گونه‌ای عمل کرد که هم نتایج خوبی بگیریم و هم تا حد امکان، دشواری بی‌مورد به کار خود اضافه نکنیم. با بررسی زمان‌های تغییر وضعیت خواب (خروجی تابع AnnotExtract)، می‌بینیم که اولاً همه‌ی زمان‌ها مضارب ۱۰ هستند، و ثانیاً عدد بزرگتری وجود ندارد که همه‌ی این زمان‌ها بر آن بخش‌پذیر باشند، لذا با انتخاب یک عدد دیگر، برخی از زمان‌های تغییر وضعیت خواب جزء مضارب صحیح آن عدد نخواهند بود و این عملکرد کار ما را دچار مشکل می‌کند. از طرفی، انتخاب شمارنده‌های ۱۰، مانند ۲ یا ۵ نیز اگرچه امکان‌پذیر است، اما ممکن است کیفیت کار ما را دچار مشکل کند، چرا که اگر پنجره‌های زمانی بیش از حد کوچک شوند، محتوای انرژی آن‌ها ممکن است فاقد اعتبار لازم باشد و به دلیل کوچک بودن بازه زمانی، مقادیر انرژی باندهای مختلف و مقایسه‌ی آن‌ها با یکدیگر، نتیجه‌ی دقیقی به ما ندهد. بنابراین عدد ۱۰ را به عنوان طول پنجره زمانی انتخاب می‌کنیم. (در سوال پنجم، این عدد را تغییر می‌دهیم تا بینیم آیا کوچک کردن این بازه، تأثیری - چه مثبت، چه منفی - روی عملکرد ما خواهد داشت یا نه).

به طور کلی، محیط برنامه به شکل زیر است:



که البته مستقیماً از روی شکل‌های زمانی سیگنال‌ها، نمی‌توان اطلاعات چندانی به دست آورد. در این جا، یک نکته‌ی استنباط شده از سیگنال‌های زمانی را ذکر می‌کنیم، و سپس به ارائه اطلاعاتی از مراجع دیگر می‌برداریم.

به استناد به سیگنال‌های مشاهده شده در نرم افزار، هرگاه تغییر حالت از یکی از سطوح خواب به بیداری صورت بگیرد، در تمامی سیگنال‌ها، نوسانات پردامنه، شدید و نامنظمی مشاهده می‌شود که می‌تواند ملاک بسیار دقیقی برای تشخیص زمان ورود به حالت بیداری از طریق سیگنال‌های زمانی باشد. نمونه‌ای از این سیگنال‌ها را در شکل زیر مشاهده می‌کنید:



گزارش کار پروژه درس علوم اعصاب محاسباتی سری دوم  
در مورد سیگنال های دریافتی، معمولاً از سیگنال EOG برای شناسایی حرکت پلکها (حالت های REM و SEM) استفاده می شود.

از سیگنال EMG معمولاً برای شناسایی حرکات عضلانی، و در نتیجه برای شناسایی حالت بیداری استفاده می شود.  
از سیگنال EEG برای استخراج اطلاعات ۴ لایه خواب استفاده می شود، به طوری که باند دلتا عموماً برای لایه های ۳ و ۴ (در ادامه این پژوهش، این مسئله تأیید می شود)، و به عنوان مثال از sleep spindle های موجود در این سیگنال ها برای شناسایی لایه ۲ و نیز از باند تتا برای لایه اول استفاده می شود. البته روش یکتایی برای ارتباط برقرار کردن بین سیگنال ها و لایه های خواب وجود ندارد و می توان ویژگی های دیگری نیز ارائه کرد. آن چه گفته شد در جدول زیر خلاصه شده است:

**Table I: RK Rules and Sleep Stages**

Pattern	Stage				
	Awake	Stage 1	Stage 2	Stage 3 and 4	REM
EEG	Slow delta	No	No	No	No
	Theta	Don't care	Yes	Don't Care	Yes
	Sleep spindles	No	No	Yes	No
EOG : REM Signal		Yes	No	No	Yes
EMG: Muscle Tone		Yes	Don't care	Don't care	Don't care
					No

مراجع:

## EOG and EMG: Two Important Switches in Automatic Sleep Stage Classification

E Estrada, H Nazeran, J Barragan , J R Burk\*, E A Lucas\*, and K Behbehani\*\*

Department of Electrical and Computer Engineering, University of Texas at El Paso, El Paso TX, USA

\*Sleep Consultants Inc, Ft Worth Texas, USA

\*\*

Department of Bioengineering, University of Texas at Arlington, Arlington, Texas, USA

E-mail: nazeran@ece.utep.edu

یک تقسیم بندی دیگر برای اطلاعات EEG و مراحل خواب در جدول زیر آمده است:

STAGE	FREQUENCY (HZ)	AMPLITUDE (MICRO VOLTS)	WAVEFORM TYPE
awake	15-50	<50	
pre-sleep	8-12	50	alpha rhythm
1	4-8	50-100	theta
2	4-15	50-150	spindle waves
3	2-4	100-150	spindle waves and slow waves
4	0.5-2	100-200	slow waves and delta waves
REM	15-30	<50	

مراجع: <https://www.tuck.com/stages/>

## قسمت سوم) یافتن معیاری برای توصیف خواب مبتنی بر PCA

(۳-۱)

PCA روشی برای کاهش ابعاد فضای از دست دادن کمترین اطلاعات ممکن است. در این روش، ابتدا ماتریس کوواریانس متغیرهای تصادفی موجود محسوبه می‌گردد. (تعداد ابعاد فضای  $N$ ، همان تعداد متغیرهای تصادفی است، که در هر مسئله با توجه به نمونه‌گیری انجام شده، تعدادی از تحقیقات این متغیر تصادفی – تعدادی نقطه در این فضای  $N$  بعدی – در اختیار است)

سپس با محسوبه مقادیر و بردارهای ویژه ماتریس کوواریانس، می‌توان راستاهایی در فضای را یافت که اطلاعات در آن راستاهای بیشترین پخش شدگی را داشته باشند. واریانس در هر راستا، همان مقادیر ویژه هستند و بردار ویژه متناظر با هر مقدار ویژه، نشان‌دهندهٔ راستای متناظر در فضای مذکور است. بنابراین کافی است مقادیر ویژه را به ترتیب نزولی مرتب کنیم و بردار ویژه‌های متناظر را به عنوان جهت‌های ماکزیمم تغییرات معرفی کنیم. نهایتاً می‌توان حد آستانه‌ای (مثلاً ۹۰ درصد) را در نظر گرفت و تنها راستاهایی از فضای را نگه داشت که مجموع مقادیر ویژه (واریانس‌ها) برای آن‌ها، بیش از حد آستانه‌ی معین شده باشد. عموماً مشاهده می‌شود که در نظر گرفتن ۳ راستا، کافی است. با این کار، ابعاد فضای را کاهش داده‌ایم و کمتر از ۱۰ درصد اطلاعات از دست رفته است.

نهایتاً باید نقاط موجود در فضای را روی این ۳ (یا هر تعداد انتخابی دیگری) تصویر کنیم و اطلاعات را در فضای جدید (با بعد کاهش یافته) منعکس کنیم.

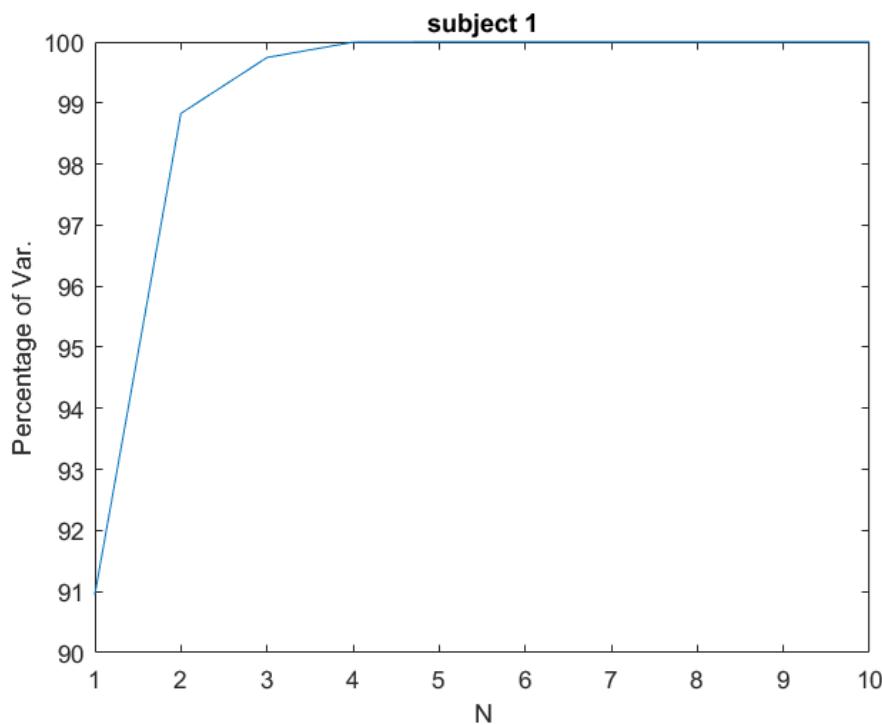
خروجی‌های تابع `pca` عبارتند از:

**coeff**: ماتریسی است که تمامی بردارهای ویژه راستاهای استخراج شده توسط PCA را به صورت ستونی و به ترتیب اهمیت (به ترتیب نزولی مقدار ویژه) در بر دارد.

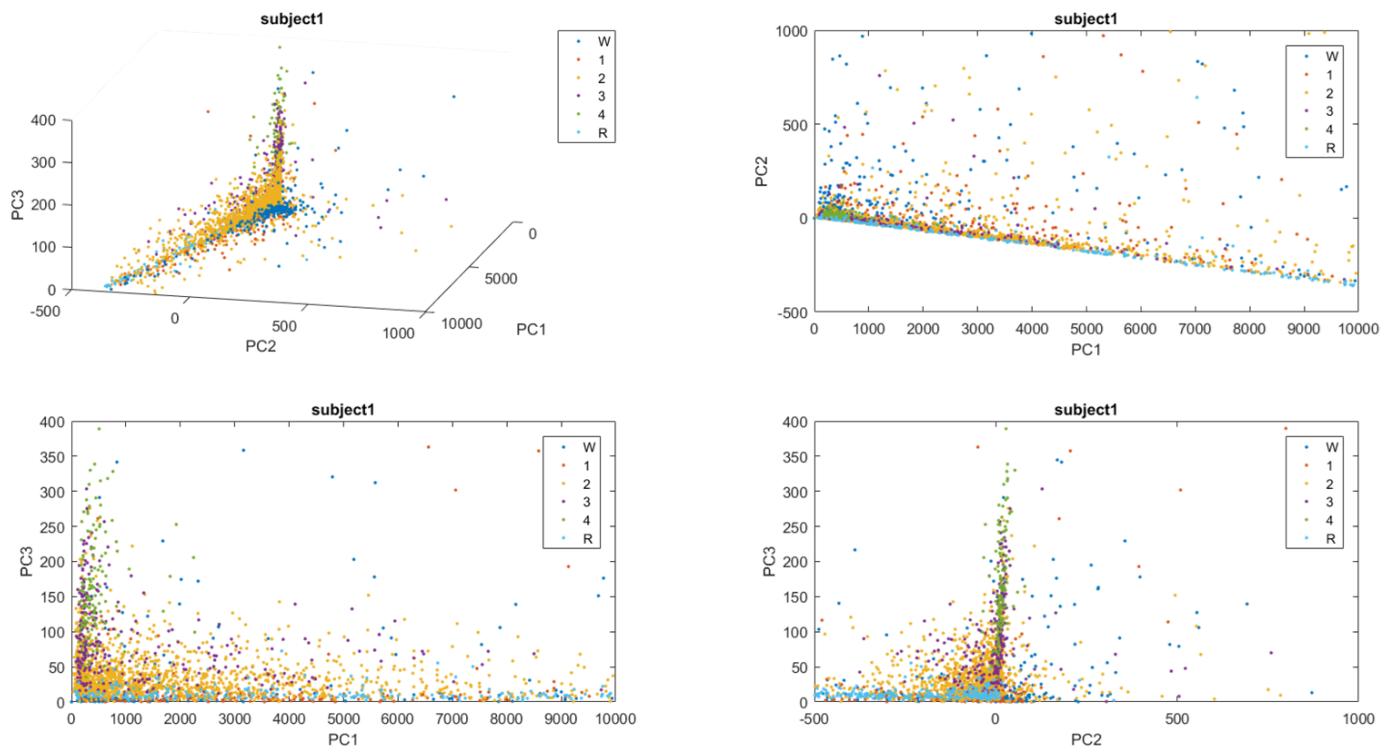
**score**: ماتریسی است که ابعاد آن، با ابعاد ماتریس ورودی تابع یکسان است و نقاط ورودی را در همان فضای اولیه (بدون کاهش بعد) اما با یک پایه‌ی متعامد جدید (که همان بردارهای ویژه هستند – می‌دانیم که به دلیل متقارن بودن ماتریس کوواریانس، این بردارها متعامد هستند و با در نظر گرفتن تمامی آن‌ها، پایه‌ای متعامد برای فضای اولیه خواهیم داشت –) توصیف می‌کند. بنابراین همان نقاط را در این ماتریس داریم، اما مختصات آن‌ها عوض شده است، چرا که پایه‌ی فضای را تغییر داده‌ایم.

**latent**: برداری به طول  $N$  است که مقدار ویژه‌ی متناظر با بردارهای ویژه موجود در ماتریس `coeff` را در بر دارد. مقادیر ویژه و بردارهای ویژه متناظر به ترتیب نزولی (به ترتیب کاهش اهمیت) مرتب شده‌اند.

نمودار تجمعی واریانس (انرژی) برای سابجکت اول به صورت زیر است:



مطابق نمودار فوق، حدود ۹۱ درصد انرژی در راستای اول، حدود ۹۸/۸ درصد انرژی در دو راستای اول، حدود ۹۹/۸ انرژی در سه راستای اول، و برای مابقی راستاهای، تقریباً تمامی انرژی در  $n$  راستای اول ذخیره شده است. این مشاهده با توضیحات بخش قبل نیز سازگار است و بیان می‌کند که انتخاب ۳ راستای اول می‌تواند برای کاهش بُعد فضای و از دست ندادن اطلاعات، انتخاب بسیار مناسبی باشد.



نمودارهای شکل فوق، شامل نمودار سه بعدی خواسته شده، و نیز تصویر نقاط در سه صفحه‌ی مختصات این نمودار هستند.

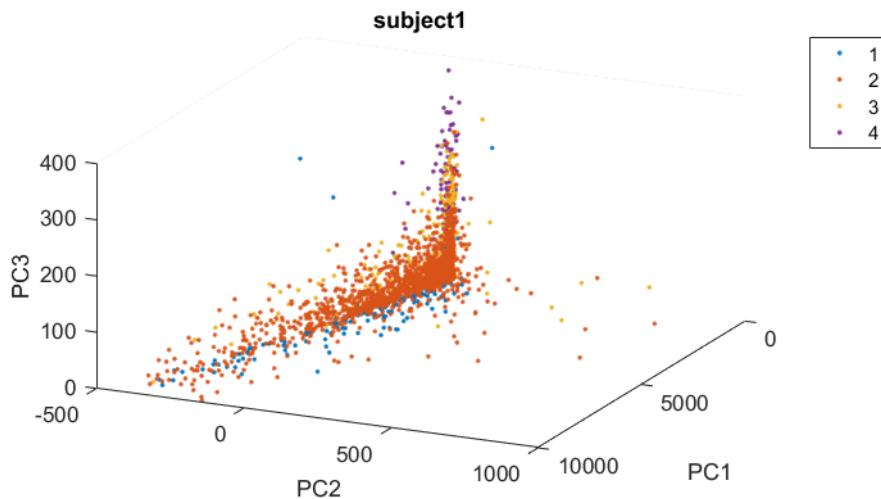
(تذکر: مقیاس‌های محورهای مختصات یکسان نیستند)

با توجه به نمودار فوق، مشهود است که نمی‌توان ادعا کرد هیچ کدام از این سه راستا یا ترکیبات آنها قادر به جداسازی همه‌ی ۶ حالت موجود هستند. با این وجود می‌توان مشاهده نمود که راستای PC1 تا حد قابل قبولی، همه‌ی دیتای مربوط به خواب REM را در بر دارد. (البته با توجه به نمودار رسم شده در صفحه‌ی PC1-PC2، به شکل دقیق‌تر می‌توان گفت که راستایی که اطلاعات REM را در بر دارد، خطی در این صفحه است که از مبدأ گذشته و زاویه‌ای در حدود 30° درجه دارد)

همچنین می‌توان ادعا نمود که وضعیتی نسبتا مشابه برای حالت بیداری نیز موجود است و نقاط مربوط به این وضعیت، اکثراً در صفحه‌ی PC1-PC2 پخش شده‌اند و راستای PC3، حجم کمی از این نقاط را در بر می‌گیرد.

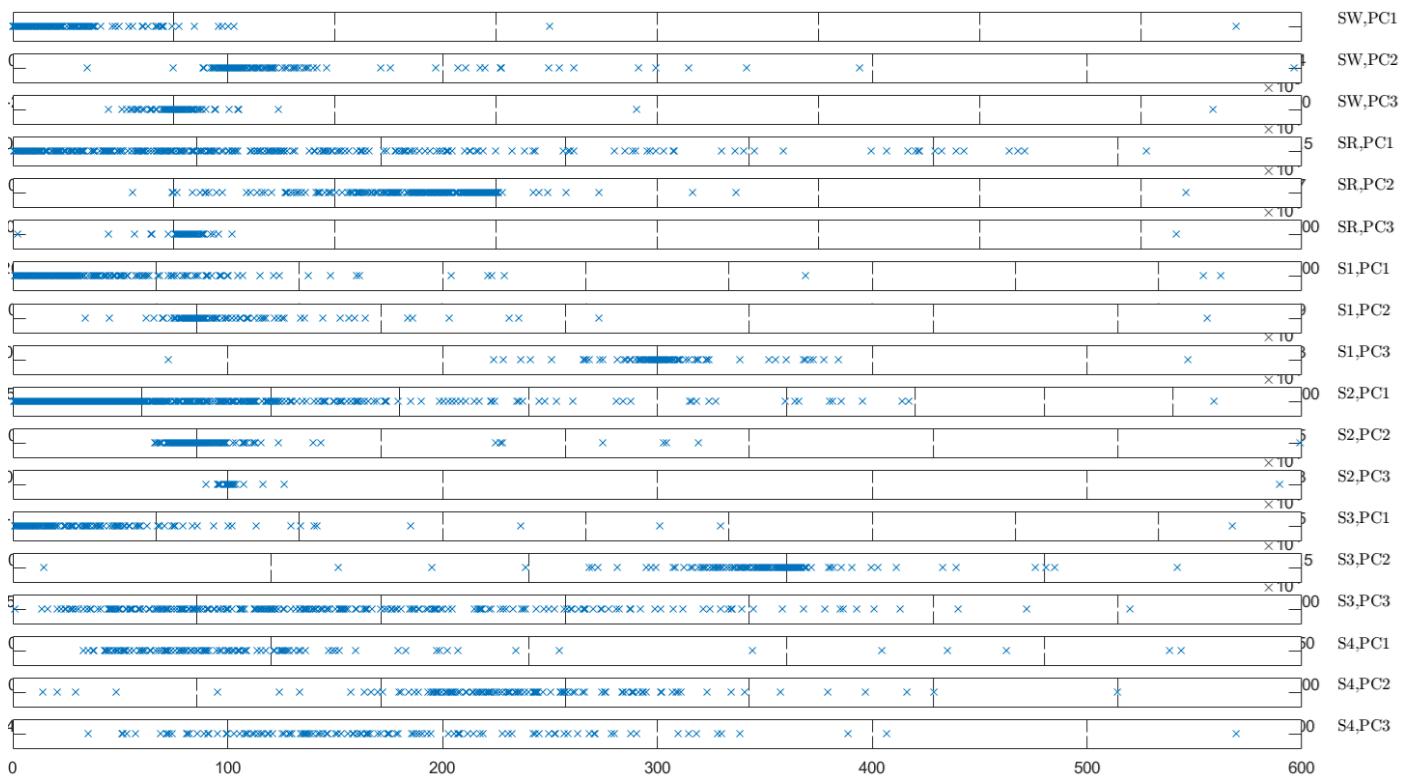
همچنین با دقت به راستای PC3 در نمودارهای مختلف، درمی‌یابیم که محتویات این راستا، عمده‌ای مربوط به خواب NREM (و به طور خاص، بیشتر شامل عمق ۳ و ۴) است و این راستا، عمق‌های مختلف NREM را می‌تواند تا حد قابل قبولی از هم جدا کند.

برای بررسی دقیق‌تر وضعیت‌های مختلف خواب NREM، می‌توانیم نمودار مربوط به این ۴ وضعیت را رسم کنیم.



آنچه مشخص است، راستای PC2 اطلاعات چندانی در بر ندارد، و دو راستای PC1 و PC3 در تمایز دادن داده‌های مربوط به لایه‌های مختلف NREM موفق‌تر هستند.

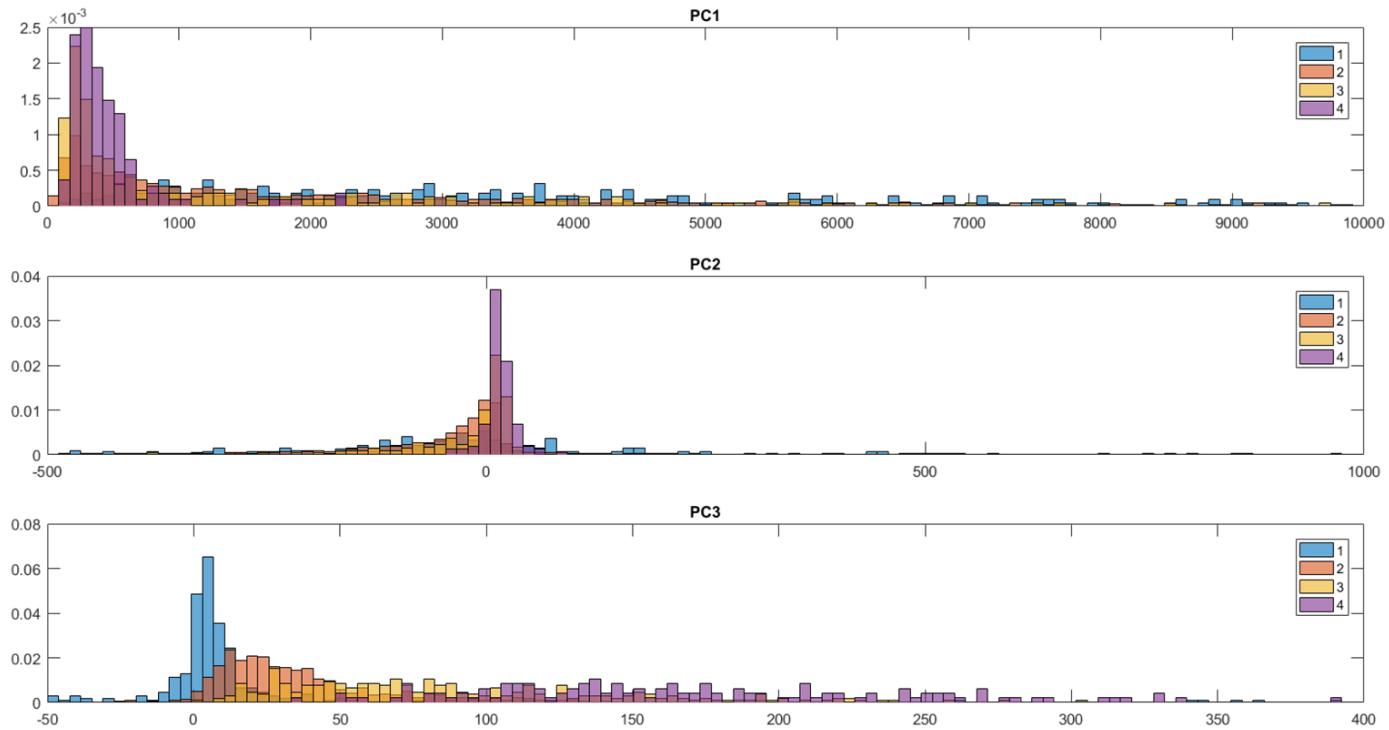
همچنین می‌توان در نمودارهایی یک‌بعدی، توزیع نقاط مربوط به هر یک از ۶ لایه را در راستای PC1، PC2، و PC3 (مجموعاً ۱۸ نمودار) مشاهده کرد. این نمودارها به شکل زیر هستند:



با توجه به نمودارهای فوق، می‌توان مشخص کرد که هر کدام از ۶ لایه‌ی خواب، در کدام یک از راستاهای کشیدگی بیشتری دارند. به عنوان مثال، وضعیت REM، لایه دوم، لایه سوم، و لایه چهارم به ترتیب در راستاهای PC1، PC2، و PC3 کشیدگی قابل توجهی دارند.

از طرف مقابل به عنوان مثال می‌توان مشاهده کرد که راستای PC3 برای لایه دوم، و نیز وضعیت REM تقریباً هیچ اطلاعاتی در بر ندارد.

اطلاعات موجود در نمودار قبل را می‌توان در قالب هیستوگرام نیز بیان کرد، که می‌تواند شهود بهتری از محل‌های تراکم نقاط در وضعیت‌های مختلف خواب در اختیار ما قرار دهد. هیستوگرام مذکور برای حالت‌های NREM به شکل زیر می‌باشد: (سه هیستوگرام برای سه راستای PC1، PC2، و PC3)



می‌بینیم که لایه‌های ۱، ۲، ۳، و ۴ به ترتیب بیشترین کشیدگی را در راستاهای PC3، PC1، PC2 دارند. در اینجا کشیدگی اطلاعات به یکنواخت بودن هیستوگرام و نداشتن قله تعییر می‌شود. نتایج مذکور با آنچه از تحلیل نمودارهای قبلی به دست آمد، منطبق است.

یک نکتهٔ مهم نیز آن است دو دسته از نقاط می‌توانند توسط یک راستا از هم جدا شوند، حتی اگر در آن راستا کشیدگی بسیار زیادی نداشته باشند. به عنوان مثال، با دقیق بیان می‌بینیم که تجمع ۴ حالت خواب در بخش‌های مختلفی اتفاق افتاد است، و این بدان معناست که راستای PC3 در جداسازی ۴ حالت NREM از یکدیگر نسبتاً موفق عمل کرده است.

در پایان، تکرار یکی از نکات مذکور در توضیحات این بخش خالی از لطف نیست، و آن نکته این است که ما در تحلیل‌های مطرح شده در این بخش، خود را محدود به سه راستای PC1، PC2، و PC3 کردیم، و طبعاً این محدودیت باعث می‌شود که نتوانیم داده‌های ۶ وضعیت مختلف را به خوبی جدا کنیم. در حالت کلی‌تر، ممکن است اطلاعات مربوط به یک وضعیت، در راستایی از این فضایی به صورت یک ترکیب خطی از سه راستای اصلی قابل بیان است، بیشترین کشیدگی و تمایز را داشته باشد، یا حتی ممکن است در ناحیه‌ای از این فضایی، تجمعی شاخص اما غیر خطی داشته باشد. با این حال، با توجه به این که در قسمت بعد، می‌خواهیم با بررسی منشأ سه راستای اصلی، ارتباطی بین سیگنال‌های مختلف EEG و لایه‌های خواب برقرار کنیم، در این بخش به همین تحلیل اکتفا می‌کنیم.

برای پاسخ به این سؤال، به ماتریس coeff مراجعه می‌کنیم (تنها ستون‌های اول تا سوم این ماتریس در این بخش برای ما اهمیت دارد):

	1	2	3
1	0.0062	0.0248	0.0274
2	2.8593e-04	8.0241e-04	4.8945e-06
3	2.1793e-04	0.0010	-1.9562e-04
4	7.3661e-04	0.0026	-4.7650e-04
5	0.0013	0.0412	0.9987
6	-4.3537e-05	1.2827e-04	0.0049
7	7.9358e-05	3.6283e-04	0.0013
8	4.0380e-04	0.0013	5.7752e-04
9	0.9993	-0.0369	6.0935e-05
10	0.0367	0.9982	-0.0419

مشاهده می‌شود که در هر سه راستای اول، تنها یکی از ابعاد فضای حالت اصلی نقش تأثیرگذار داشته و سایر ابعاد، تقریباً بی‌اثرند. به تعبیر دیگر، سه راستا از راستاهای فضای حالت اولیه هستند که واریانس بسیار زیادی دارند و می‌توان آن‌ها را به عنوان راستاهای مؤثر انتخاب کرد. با توجه به ضرایب، می‌بینیم که اثر راستاهای دیگر ناچیز است و می‌توان در عمل از آن‌ها صرف نظر کرد.

- راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
- راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
- راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)

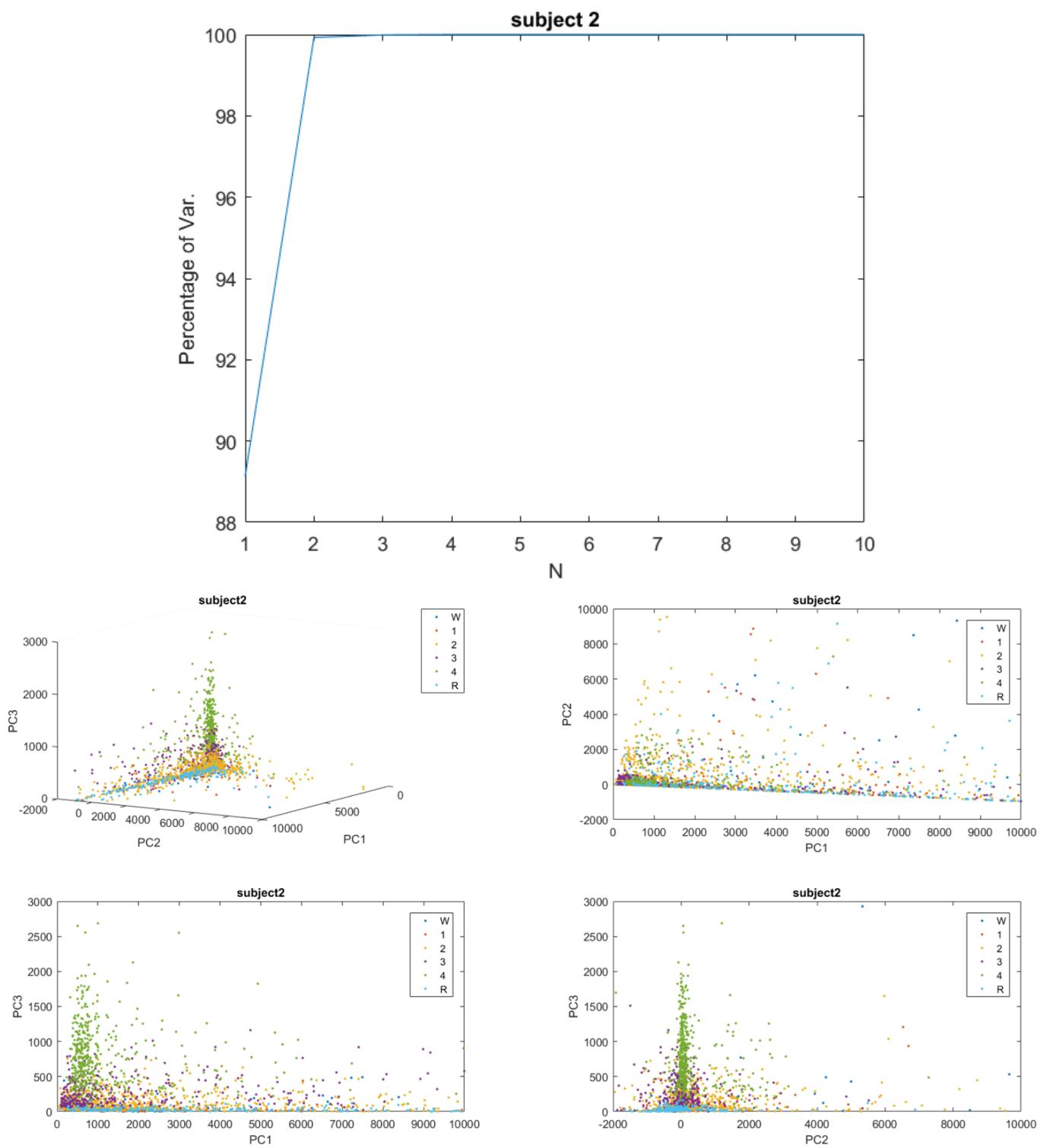
با توجه به این که ضریب متناظر برای سه پارامتر معرفی شده در سه راستا، نزدیک به ۱ است، در هر سه مورد، مؤلفه‌ی متناظر PCA با افزایش توان متوسط مربوطه افزایش می‌یابد و ارتباطی مستقیم و تقریباً از جنس رابطه‌ی همانی دارند. (اگر این ضریب نزدیک به ۱ - می‌بود، ارتباط آن‌ها بالا، اما به شکل معکوس بود)

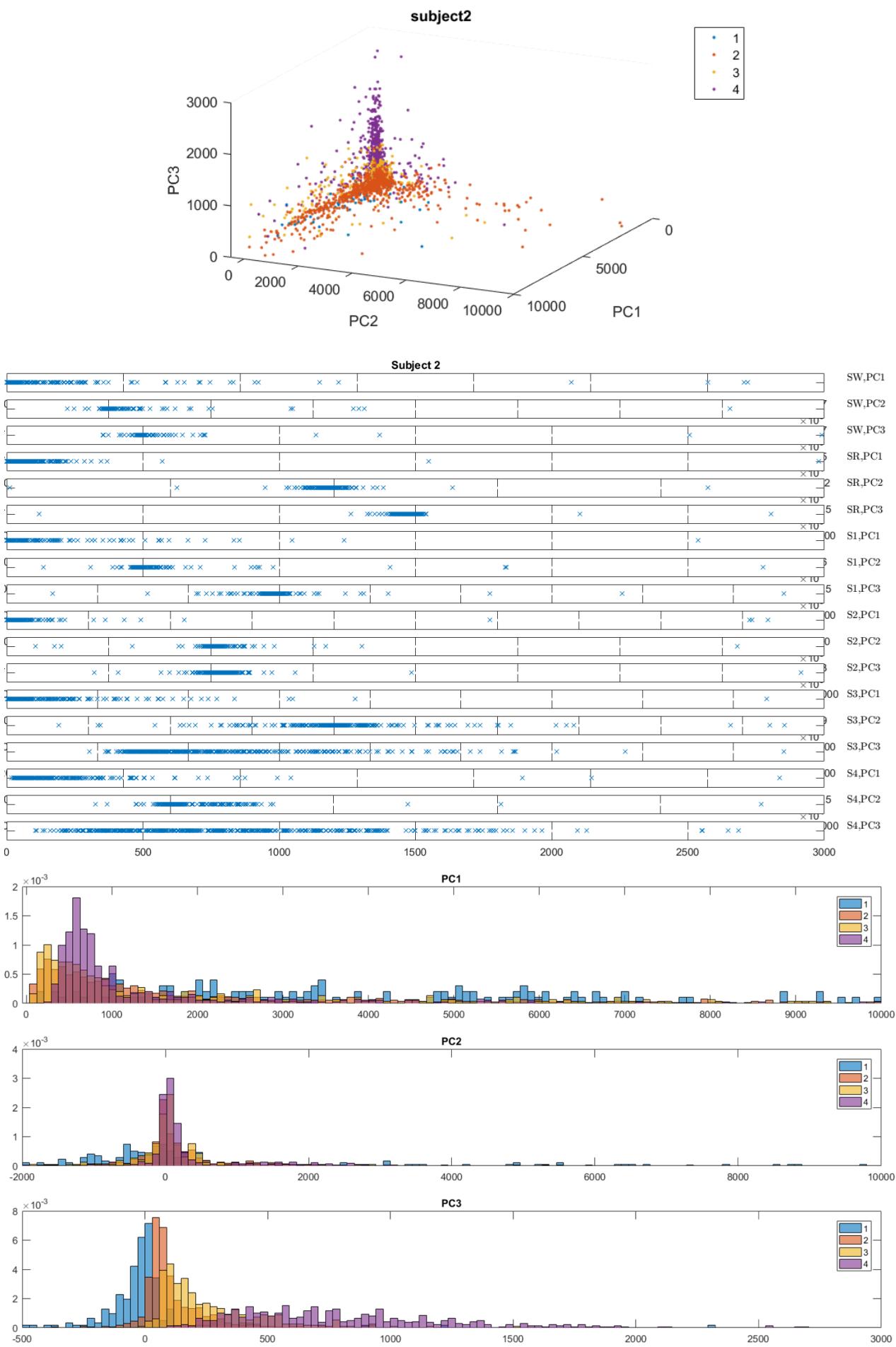
حال، می‌توان با توجه به تحلیل‌های انجام شده در این بخش و بخش قبل، نتیجه‌گیری‌هایی انجام داد:

- اعماق مختلف خواب NREM توسط متوسط باند فرکانسی دلتا (مربوط به الکترود Oz) می‌توانند تا حد قابل قبولی از یکدیگر جدا شوند.
- عمق اول و دوم خواب بیشتر در راستای توان متوسط سیگنال EOG دارای کشیدگی هستند و عمق سوم و چهارم بیشتر در راستای باند دلتای الکترود Oz، با این حال با توجه به هیستوگرام‌های رسم شده، باند فرکانسی دلتای الکترود Oz برای جداسازی مجموعه کلی آن‌ها انتخاب مناسب‌تری است.
- توان متوسط سیگنال EMG، اطلاعات چندانی به خصوص در مورد خواب NREM در اختیار ما قرار نمی‌دهد.
- خواب REM در راستایی که ترکیبی خطی از توان متوسط سیگنال EOG و EMG است، توزیع شده و کشیدگی بسیار زیادی دارد. در این راستا، سهم متوسط سیگنال EOG بیشتر است، اما نمی‌توان از اثر سیگنال EMG چشم‌پوشی کرد. بنابراین با در نظر گرفتن این راستا، می‌توان دیتای مربوط به خواب REM را تا حدی از سایرین جدا کرد. (البته با مشاهدات این ساچجکت، این کار کیفیت چندانی نخواهد داشت، چرا که اگرچه تمامی دیتای مربوطه به REM در این راستا متمرکز شده است، اما اطلاعات پرآکنده‌ای مربوط به سایر وضعیت‌ها نیز در آن موجود است).

ابتدا نمودارهای بخش‌های قبل را برای هر کدام از سابجکت‌ها رسم می‌کنیم، سپس با بررسی نمودارها و نیز سه ستون اول ماتریس coeff، نتایج را برای آن‌ها بررسی می‌کنیم. در این بخش، به تفصیل دو بخش قبل برای هر سابجکت صحبت نخواهیم کرد و صرفاً به نتیجه‌گیری می‌پردازیم:

**سابجکت دوم:**





آن چه برای این سابجکت مشاهده می شود به طور خلاصه آن است که راستای PC3 جداکنندگی قابل توجهی برای اعماق مختلف خواب NREM دارد. این ویژگی در سابجکت قبلی نیز مشاهده شد.

همچنین خواب REM نیز کشیدگی نسبی در راستای PC1 دارد.

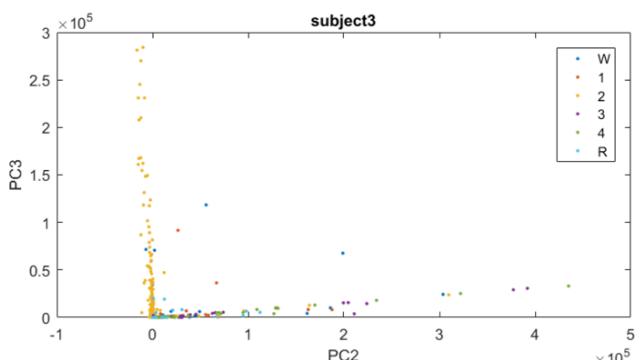
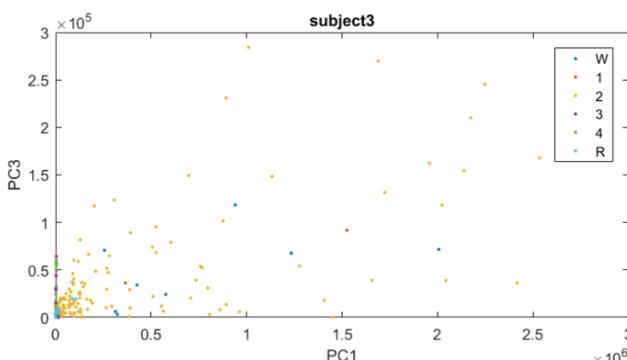
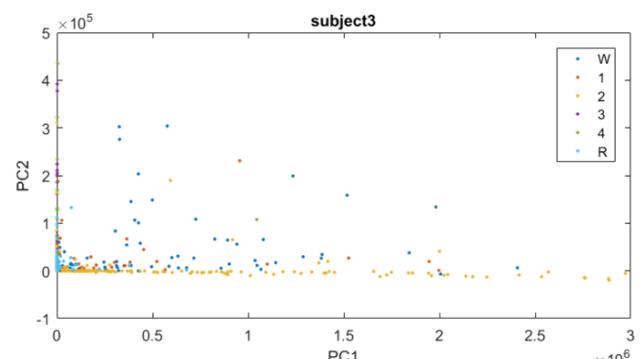
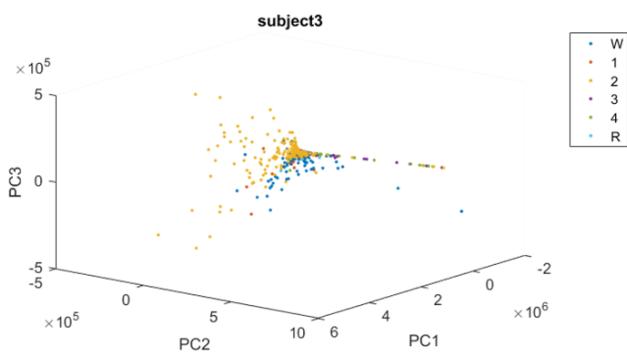
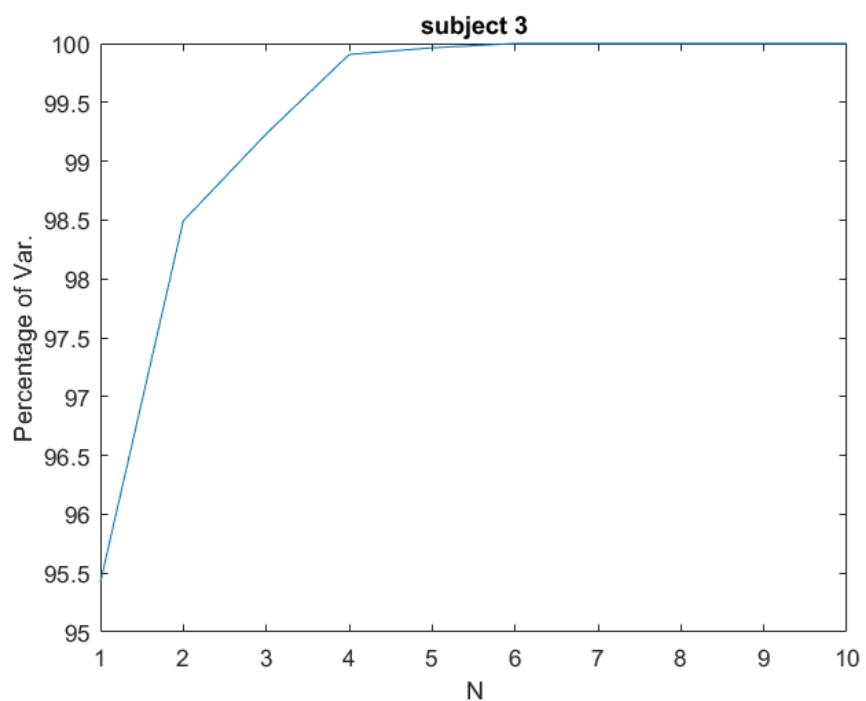
همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:

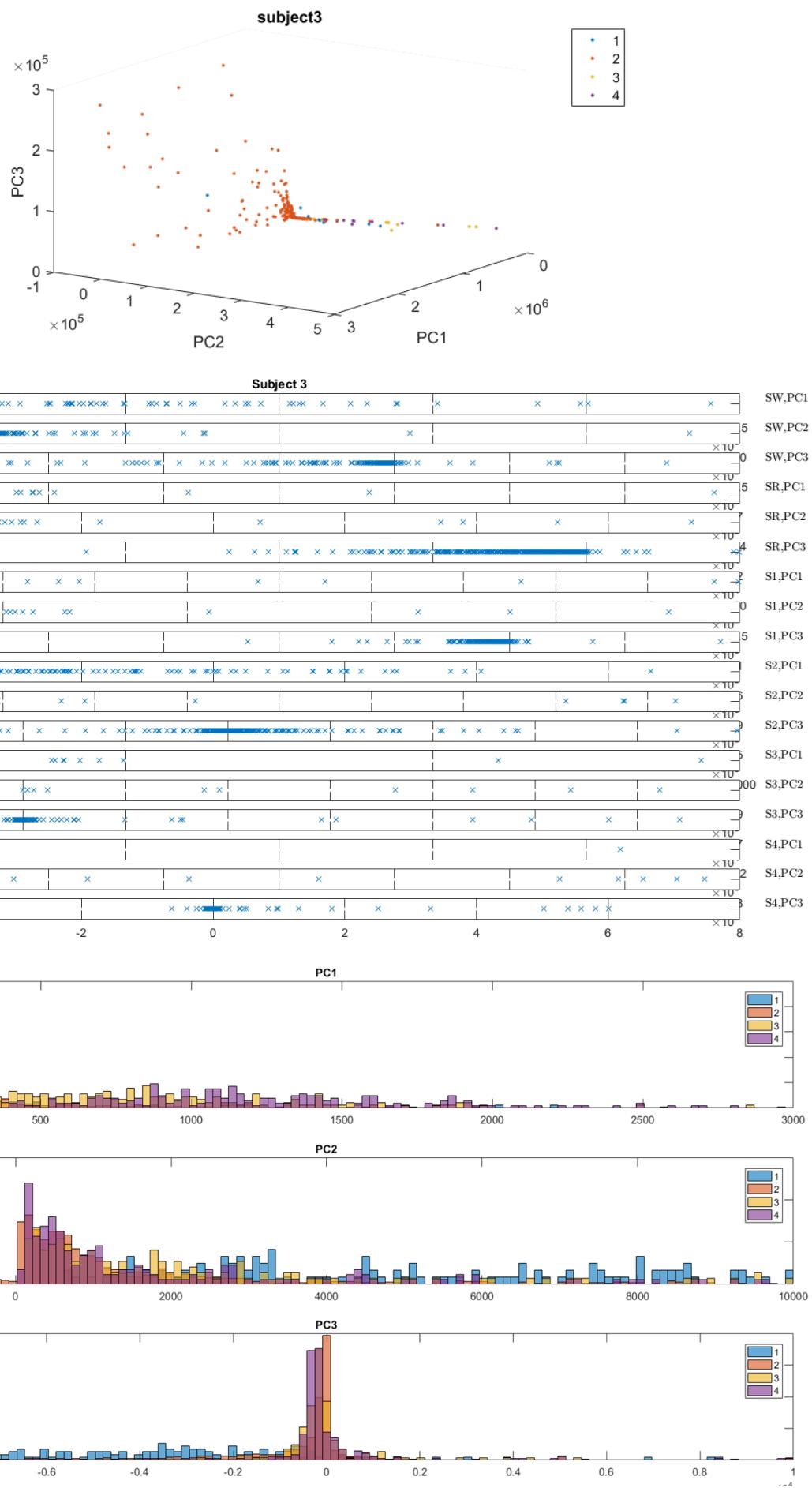
	1	2	3
1	-2.8850e-06	0.0020	0.1082
2	-1.1905e-05	3.6448e-05	0.0023
3	5.2723e-06	9.1051e-05	0.0023
4	9.5733e-05	2.0846e-04	0.0026
5	0.0058	0.0078	0.9941
6	4.3824e-06	8.5984e-06	0.0051
7	2.1221e-05	6.0468e-05	0.0022
8	5.3359e-05	1.0088e-04	0.0023
9	0.9952	-0.0972	-0.0050
10	0.0972	0.9952	-0.0085

و نتیجه به شرح زیر خواهد بود:

- راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
- راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
- راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)

حال، می توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (الکترود Oz) تأثیر بالایی در جداسازی لایه های خواب REM از یکدیگر دارد. همچنین خواب REM را نیز می توان تا حدی توان متوسط سیگنال EOG (به شکل دقیق تر، ترکیب خطی EOG و EMG) از بقیه جدا کرد.





این سابجکت کمی نسبت به دو مورد قبلی متفاوت است. مشاهده می شود که راستای PC1 (برخلاف دو سابجکت قبلی) در جداسازی خواب NREM موفق عمل کرده است. همچنین خواب REM بیشتر در راستای PC2 متتمرکز و قابل جداسازی است.

همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:

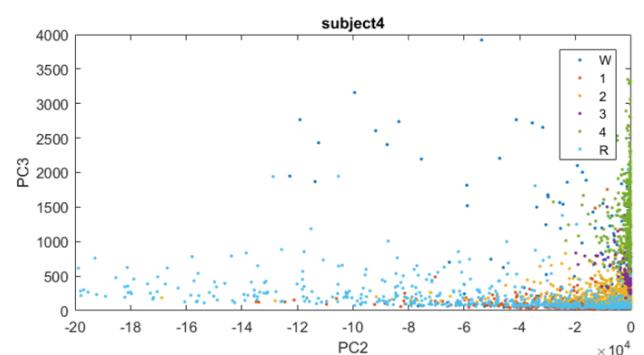
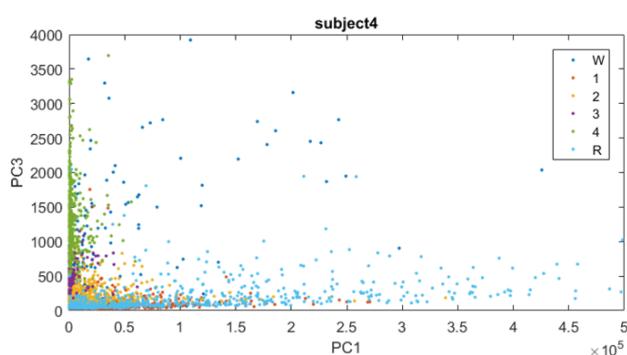
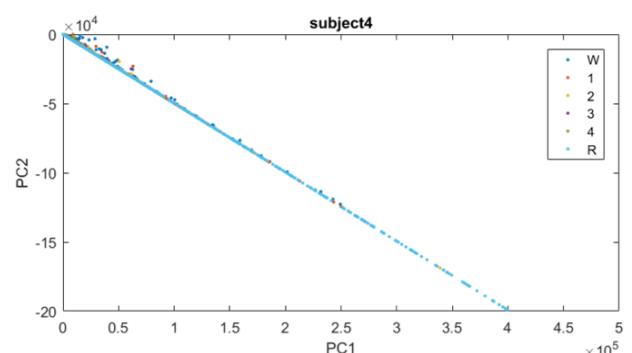
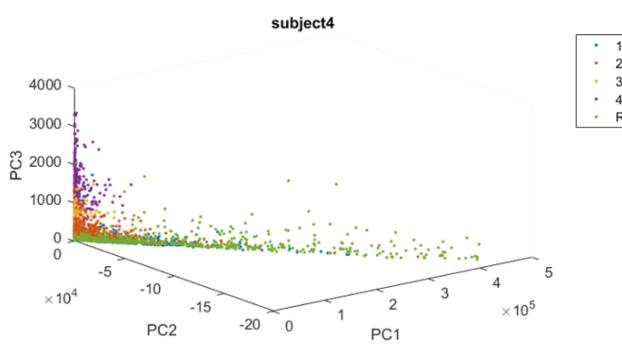
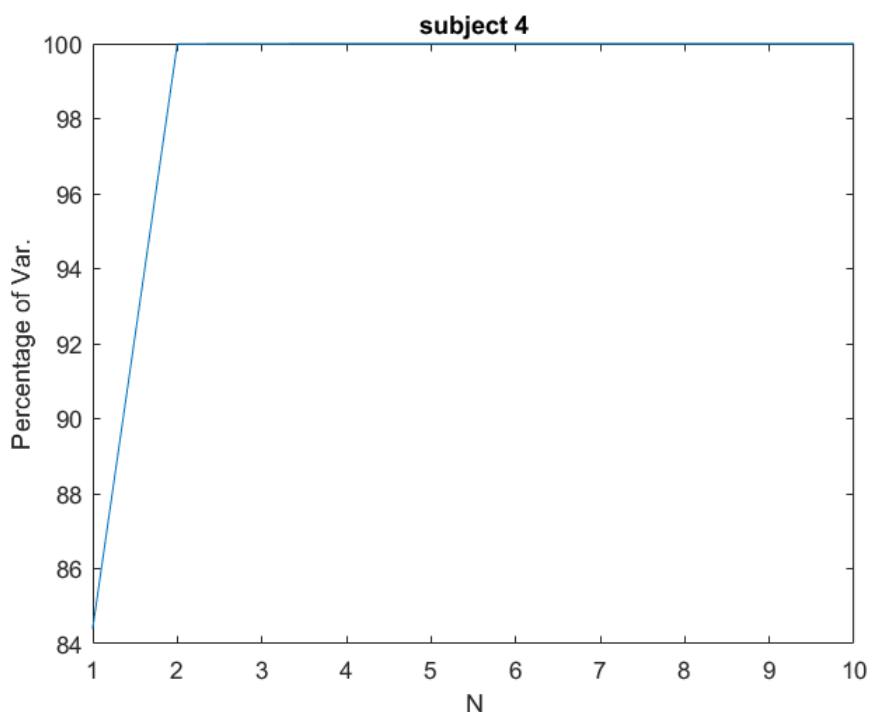
	1	2	3
1	3.4801e-04	0.0024	-0.0017
2	3.1203e-05	9.8502e-05	2.0001e-05
3	2.5711e-05	2.6941e-05	4.7290e-05
4	2.3105e-05	4.4121e-05	-9.8237e-05
5	0.9826	0.0010	-0.1778
6	0.1360	-0.0159	0.7111
7	0.1055	-0.0156	0.4888
8	0.0697	-0.0086	0.3696
9	-0.0011	0.1872	-0.2851
10	0.0037	0.9820	0.0771

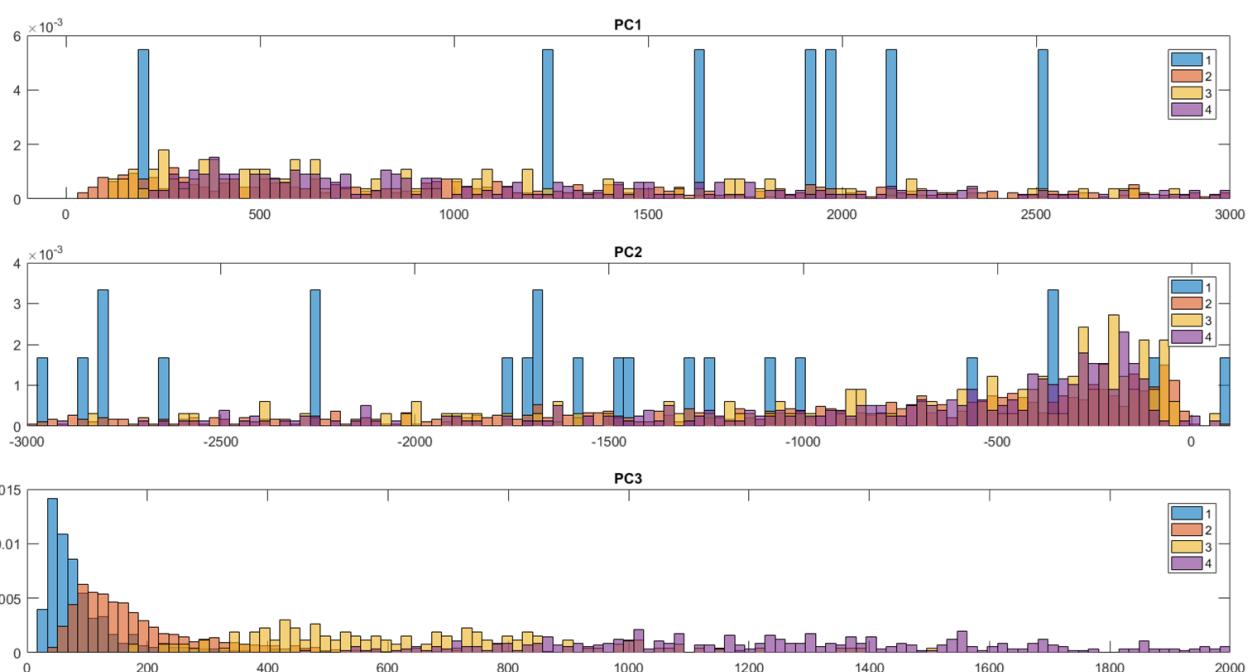
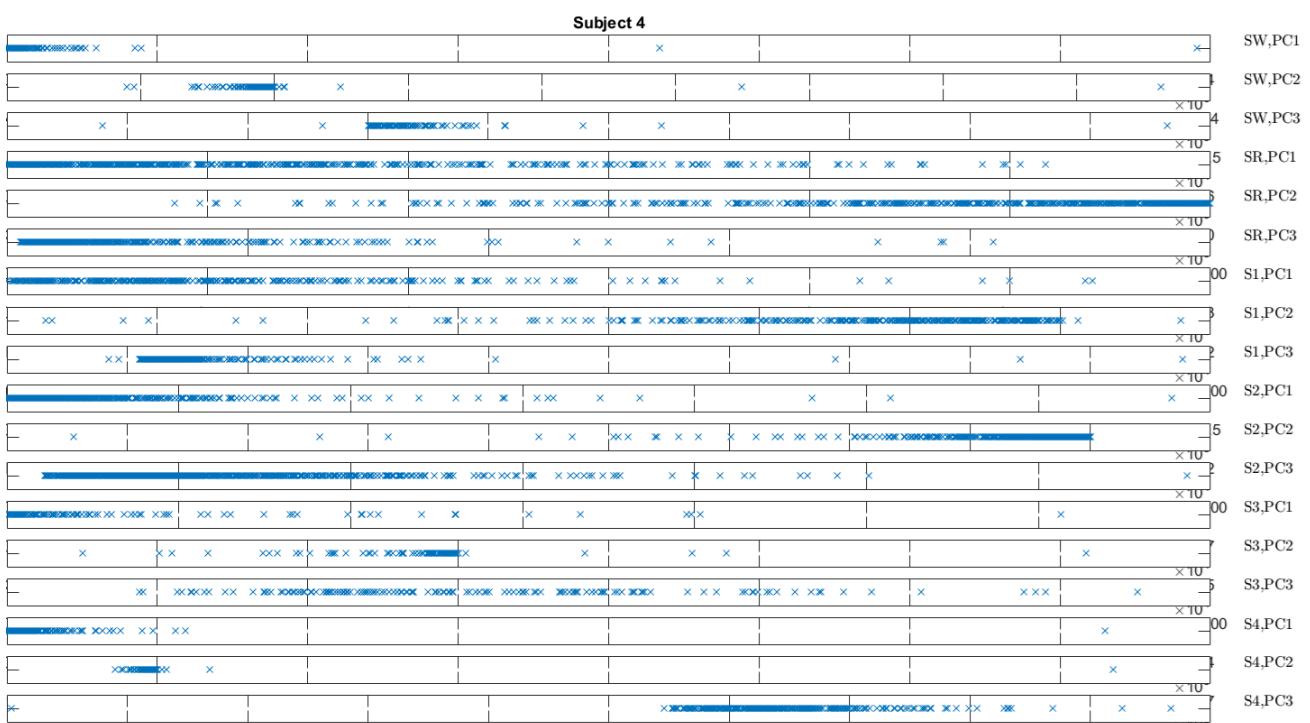
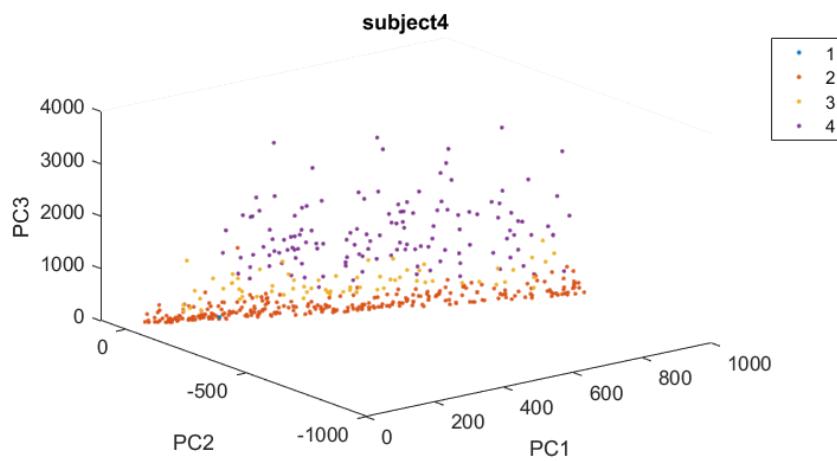
و نتیجه به شرح زیر خواهد بود:

- راستای اول (PC1) : توان متوسط باند فرکانسی دلتای الکترود Oz (سایر پارامترها، اثری ناچیز دارند)
- راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
- راستای سوم (PC3) : توان متوسط باند فرکانسی تتا الکترود Oz (با ضریب ۰/۷)، توان متوسط باند فرکانسی آلفای الکترود Oz (با ضریب ۰/۵)، توان متوسط باند فرکانسی بتای الکترود Oz (با ضریب ۰/۳۷)، و توان متوسط سیگنال EOG (با ضریب ۰/۲۹ –، یعنی اثر معکوس)

مشاهده می شود که توان متوسط باند فرکانسی دلتای الکترود Oz در جداسازی اعماق مختلف خواب NREM و توان متوسط سیگنال EMG در جداسازی خواب REM از سایرین می توانند قابل قبول باشند.

نکته) یک دلیل تفاوت این سابجکت با سایرین، می تواند سن وی باشد که ۱۸ سال است، در حالی که سایر سابجکت ها در حدود ۳۰ سال دارند.





با توجه به شکل‌ها، به خصوص هیستوگرام نهایی، مشاهده می‌شود که راستای PC3 به خوبی حالت‌های NREM را از یکدیگر جدا کرده است.

در خصوص جداسازی حالت REM نیز دو راستای PC1 و PC2 مناسب هستند (چراکه REM در این دو راستا کشیده شده است، در حالی که سایر حالات تقریباً در مبدأ متمرکز هستند)

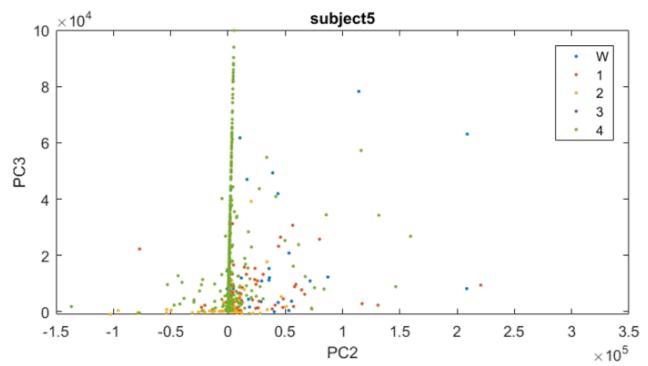
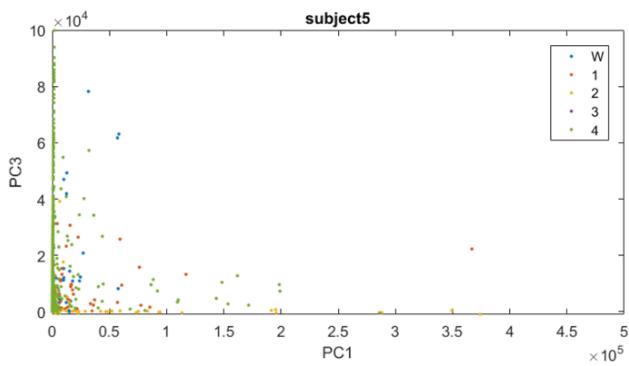
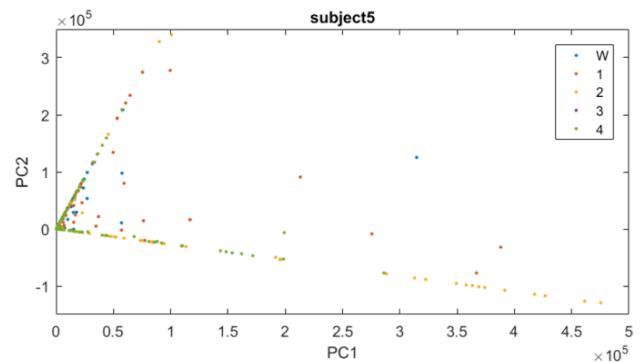
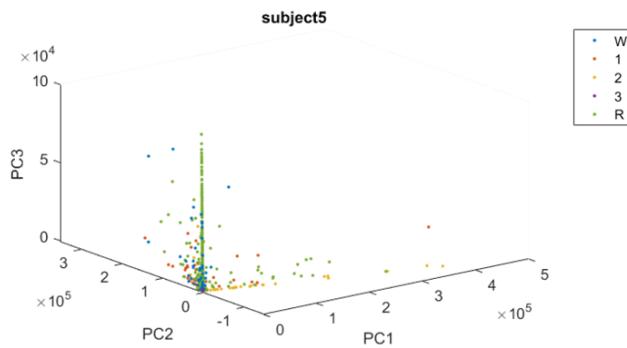
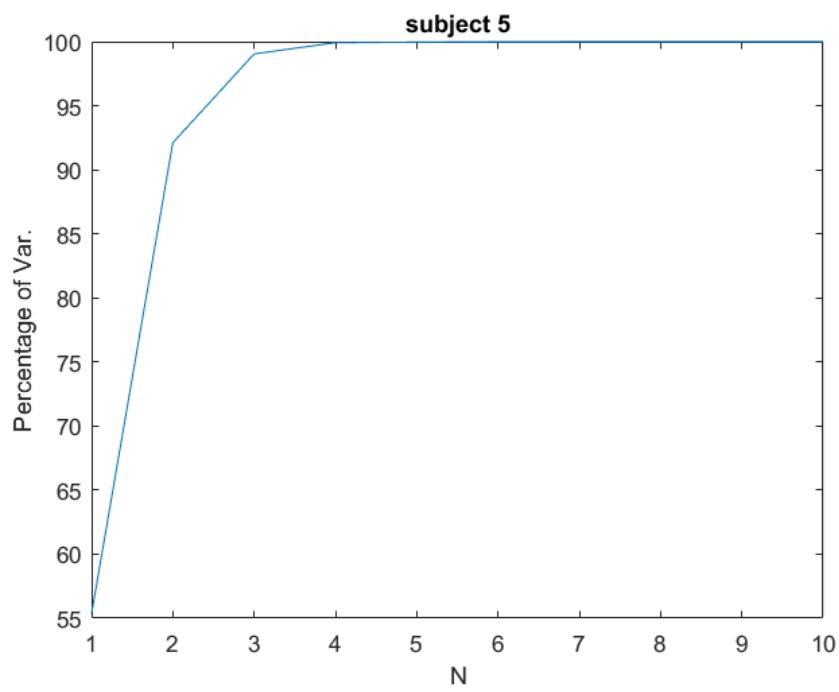
همچنین سه بردار ویژه اصلی برای این سابجکت به صورت زیر است:

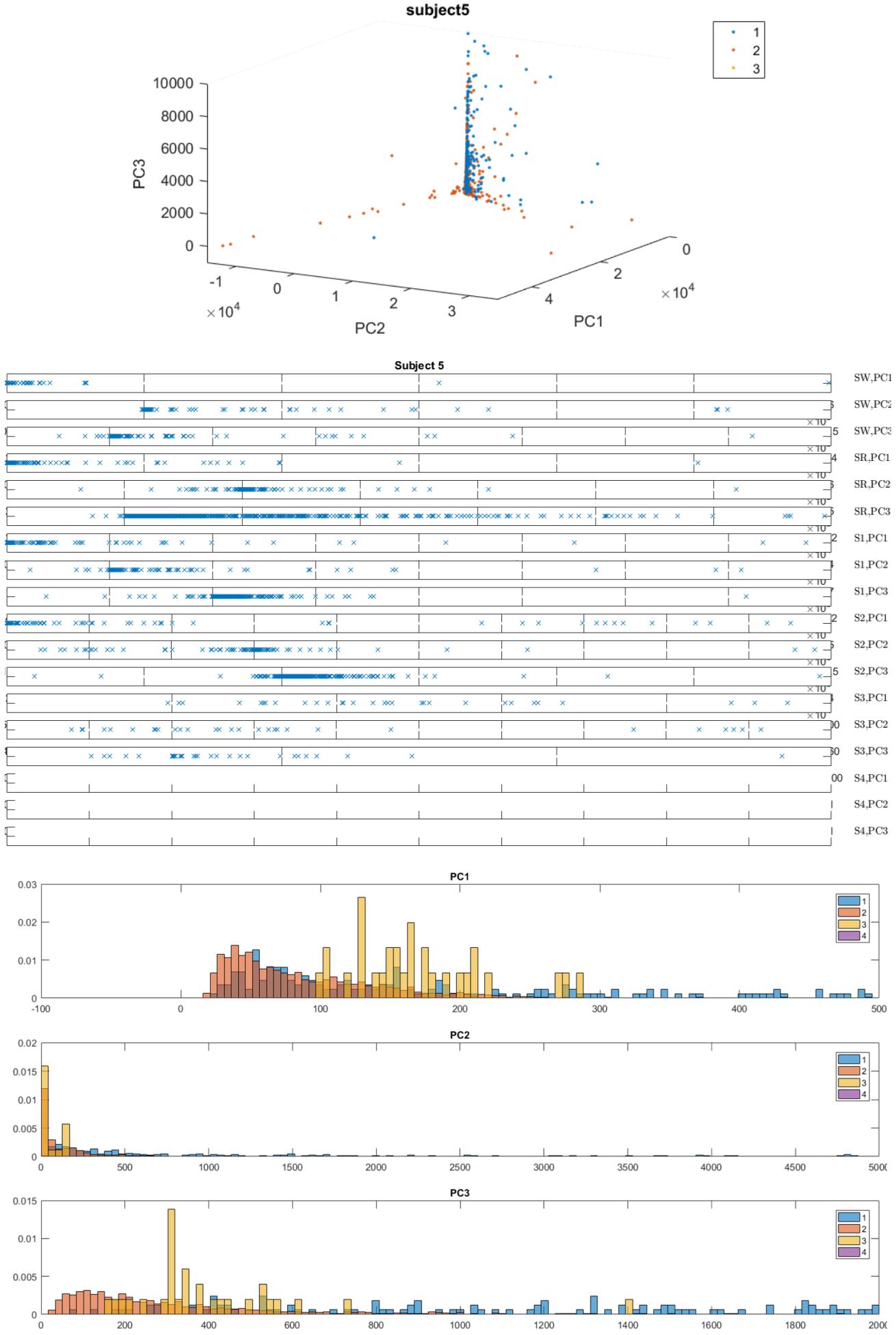
	1	2	3
1	0.0027	0.0056	0.8426
2	1.8280e-04	3.4280e-04	0.0365
3	1.6006e-04	3.1977e-04	0.0456
4	1.3364e-04	1.6063e-04	0.0336
5	0.0027	0.0059	0.5286
6	1.7490e-04	3.6341e-04	0.0337
7	1.4344e-04	2.4934e-04	0.0362
8	1.9859e-04	2.2170e-04	0.0591
9	0.8951	-0.4458	1.7487e-04
10	0.4458	0.8951	-0.0087

و نتیجه به شرح زیر خواهد بود:

- راستای اول (PC1) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)
- راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
- راستای سوم (PC3) : توان متوسط باند فرکانسی دلتای الکترودهای Fpz (با ضریب ۰/۸۵) و Oz (با ضریب ۰/۵۳)

حال، می‌توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (هم در الکترود Oz و هم در الکترود Fpz) تأثیر بالایی در جداسازی لایه‌های خواب REM از یکدیگر دارد. همچنین خواب REM را نیز می‌توان تا حدی توسط توان متوسط سیگنال‌های EOG و EMG از بقیه جدا کرد.





مشاهده می شود که (با توجه به هیستوگرام فوق) لایه های خواب NREM در راستای PC1 بهترین تغکیک پذیری را دارند، اما در مورد خواب REM، در این سایجکت، چیزی که مشاهده می شود آن است که خواب REM تقریباً به تمامی در راستای PC3 توزیع شده است و می توان به کمک این راستا، آن را از خواب NREM جدا کرد.

همچنین سه بردار ویژه اصلی برای این سایجکت به صورت زیر است:

	1	2	3
1	0.0011	0.0017	0.0075
2	1.5225e-04	3.1763e-04	8.3047e-04
3	1.5206e-04	2.5944e-04	1.0369e-04
4	4.3017e-04	0.0012	7.8301e-04
5	0.9609	-0.2641	-0.0055
6	0.0673	-0.0164	0.0094
7	0.0388	-0.0090	0.0161
8	0.0240	-0.0049	0.0137
9	0.0188	0.0540	0.9981
10	0.2642	0.9628	-0.0571

و نتیجه به شرح زیر خواهد بود:

- راستای اول (PC1) : توان متوسط باند فرکانسی دلتای الکترود Oz (با ضریب ۰/۹۶) و توان متوسط سیگنال EMG (با ضریب ۰/۲۶)
- راستای دوم (PC2) : توان متوسط سیگنال EMG (سایر پارامترها، اثری ناچیز دارند)
- راستای سوم (PC3) : توان متوسط سیگنال EOG (سایر پارامترها، اثری ناچیز دارند)

حال، می توان نتیجه گرفت که توان متوسط باند فرکانسی دلتا (الکترود Oz) تأثیر بالایی در جداسازی لایه های خواب NREM از یکدیگر دارد. همچنین خواب REM را نیز می توان تا حدی توان متوسط سیگنال EOG از بقیه جدا کرد.

در بخش سوم، یافتن معیاری برای توصیف خواب مبتنی بر PCA، ۱۰ ویژگی (توان متوسط ۴ باند فرکانسی الکترود Fpz و Oz و توان متوسط سیگنال‌های EOG و EMG) برای خواب استخراج کردیم و روی آنها PCA زدیم. در نتیجه، ۱۰ راستا در فضای حالت اولیه به دست آورده‌یم که با توجه به این که PCA ابعاد را بر حسب واریانس داده‌ها مرتب می‌کند، صرفاً سه بعد اول را مدنظر قرار دادیم و به بررسی این مسئله پرداختیم که هر کدامیک از این راستاهای می‌توانند حالت‌های مختلف خواب را از یکدیگر جدا کنند. آنچه مشاهده شد، تقریباً بین تمامی سابجکت‌های مورد بررسی یکسان بود (نتایج سابجکت‌های مختلف با هم سازگار بودند) و نتیجه این بود که توان متوسط باند دلتا (در الکترود Oz) معیار مناسبی برای جداسازی لایه‌های خواب NREM است و توان متوسط سیگنال EOG (و نیز EMG) معیار مناسبی برای جداسازی خواب REM از NREM است.

البته یک نکته قابل ذکر این است که الزاماً برای تمامی سابجکت‌ها، ملاک‌های ذکر شده در جهت‌های یکسانی در فضای ۳ بعدی ناشی از PCA نبودند، به عنوان مثال توان متوسط باند دلتا در ۳ مورد در راستای PC3 و در ۲ مورد در راستای PC1 قرار داشت، لذا نمی‌توان در مورد ارتباط راستاهای لایه‌های خواب ادعای دقیقی کرد، اما در تمامی این موارد، ارتباط ویژگی مربوطه در فضای حالت اولیه با لایه‌های خواب مشترک و سازگار و مطابق توضیحات فوق بودند.

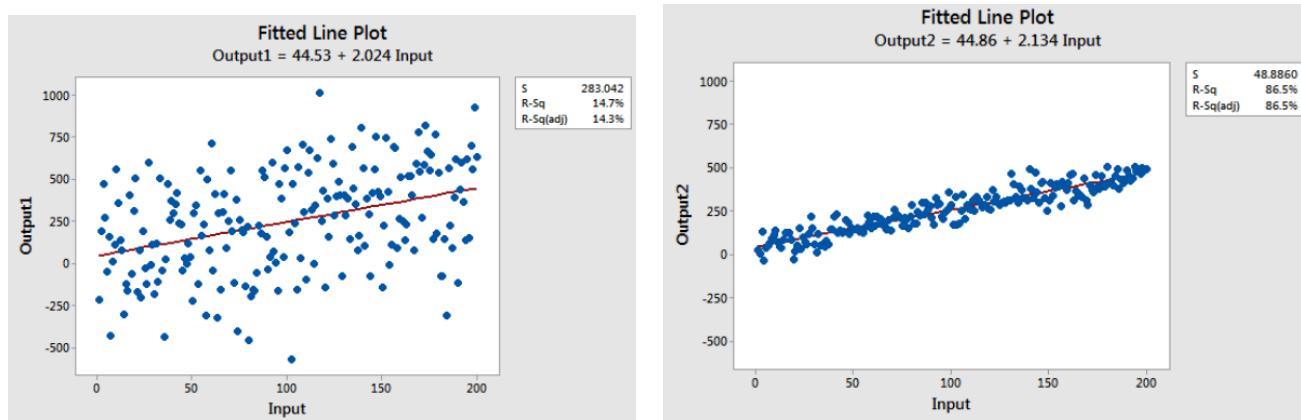
## قسمت چهارم) یافتن معیاری برای عمق خواب مبتنی بر Linear Regression

۱ - (۴)

در این بخش می‌خواهیم با فیت کردن یک مدل خطی، اثر متغیرها را بر عمق خواب بستجیم. ابتدا به توضیح دو مشخصه در هر مدل خطی می‌پردازیم:

- **p-value:** در هر مدل خطی، برای هر یک از متغیرهای ورودی، یک محاسبه می‌شود که یک مؤلفه‌ی آماری است که از آزمون فرضیه‌ای با فرض اولیه‌ی «بی‌اثر بودن این متغیر بر روی پاسخ» یا به عبارت دیگر، «صفر بودن ضریب این متغیر» ناشی می‌شود. بنابراین، هر چقدر p-value کمتر باشد، احتمال بی‌اثر بودن متغیر متناظر کمتر است و این یعنی این متغیر و ضریب به دست آمده برای آن، قابل اعتماد هستند. در حالت کلی، می‌توانیم از بازه اطمینان ۹۵٪ استفاده کنیم و متغیرهایی با p-値‌های کمتر از ۰.۰۵ را بپذیریم.

- **R-squared:** برای هر مدل خطی فیت شده، R-squared یک عدد است که مشخص می‌کند این مدل، تا چه میزان دیتای اصلی را خوب توصیف می‌کند و با کیفیت بالاتری بر روی آن فیت شده است. به عنوان نمونه دو نمودار زیر را در نظر بگیرید:



در هر دو دیتاست، یک خط فیت شده است، اما مشخصاً در یکی از آن‌ها، این خط توصیف‌کننده‌ی بسیار بهتری از دیتای اصلی است، در مقایسه با دیگری. مشاهده می‌شود که در نمونه‌ی سمت راست، مقدار R-squared برابر ۰/۸۶ و برای نمونه‌ی سمت چپ، برابر ۰/۱۴ است. اگر R-squared در یک مدل خطی برابر ۱ باشد، یعنی مدل خطی کاملاً ارتباط بین متغیرها و پاسخ را توصیف می‌کند.

حال می‌خواهیم به بررسی این موضوع بپردازیم که در یک مدل خطی، کدامیک از متغیرها اثر بیشتری (چه اثر مثبت، چه اثر منفی) روی پاسخ دارد و کدام متغیرها تقریباً بی‌اثرند. ابتدا بیشترین چیزی که به ذهن می‌رسد، آن است که ضریب هر متغیر را بررسی کنیم و هر کدام که اندازه‌ی بزرگ‌تری داشت، اثر بیشتری دارد. ابتدا بینیم که هر ضریب به چه معنی است: ضریب حاصل از یک مدل خطی برای هر متغیر به معنی «تغییرات پاسخ، به ازای یک واحد از تغییر هر متغیر مشخص، در حالی که سایر متغیرها ثابت هستند» می‌باشد. این ادعایکه با مقایسه ضریب متغیرها می‌توانیم اهمیت آن‌ها را دریابیم، در حالت کلی غلط است. به عنوان مثال، فرض کنید در یک مدل خطی، یکی از پارامترهای ورودی وزن افراد است. چنان‌چه دیتای مربوط به وزن بر حسب کیلوگرم یا بر حسب گرم بیان شود، بدیهی است اثر وزن در پاسخ تغییر نخواهد کرد، اما ضریب مربوط به وزن، در این دو حالت، از مرتبه ۱۰۰۰ با هم تفاوت خواهد داشت.

برای حل این مشکل، یک روش رایج آن است که دیتای هر کدام از متغیرهای ورودی به صورت استاندارد شده (z-score) ارائه شود.

البته به این روش نیز ایراداتی وارد است، و گفته می‌شود که نمی‌توان بدون در نظر گرفتن ماهیت مسئله و پارامترهای آن، به یک پاسخ مشخص رسید، چرا که ممکن است یکی از متغیرها، جدا از ضریب و اطلاعات آماری، قابلیت تغییر چندانی نداشته باشد (مثلاً غلظت یک ماد شیمیایی در خون) و لذا برای کنترل یک پدیده، نمی‌توان از این متغیر، هرچند اثرگذاری آن از نظر فرمولی بالا باشد، استفاده

کرد. البته این ایراد در این جا به کار ما وارد نیست، چراکه هدف کار ما تغییر دادن عمق خواب با استفاده از تغییر دادن توان متوسط سیگنال‌های ورودی نیست، و ما صرفاً می‌خواهیم مشاهدات خود را تحلیل کنیم و از روی آن استنتاجی به دست آوریم.

در این جا یک روش دیگر را نیز برای مشکلی که در رابطه با تفاوت بُعد دیتای ورودی مطرح شد، بیان می‌کنیم. روش دوم، بررسی تغییرات R-squared در اثر اضافه شدن یک متغیر، در حالتی است که سایر متغیرها ثابت هستند. این روش نیز روش رایجی است که در آن، مهم‌ترین متغیر را متغیری در نظر می‌گیرند که با اضافه کردن آن، بیشترین تغییرات مثبت را در **R-squared** مشاهده کنیم.

در این جا به این بحث پایان می‌دهیم و در انجام تحلیل خود، از روش اول (محاسبه‌ی z-score و استفاده از دیتای استاندارد شده) استفاده می‌کنیم. البته داده‌های ما مشکل تفاوت ابعادی ندارند (همگی توان متوسط هستند) اما چون مرتبه بزرگی برخی سیگنال‌ها با سایرین متفاوت است، این کار می‌تواند مفید باشد.

مراجع این بخش:

1. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients>
2. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-a-regression-model-with-low-r-squared-and-low-p-values>
3. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-most-important-predictor-variables-in-regression-models>
4. <http://blog.minitab.com/blog/adventures-in-statistics-2/how-high-should-r-squared-be-in-regression-analysis>

	1 Estimate	2 SE	3 tStat	4 pValue
1 (Intercept)	2.1018	0.0118	177.7937	0
2 Oz_delta	0.1158	0.0169	6.8531	9.2587e-12
3 Oz_theta	0.0141	0.0228	0.6167	0.5375
4 Oz_alpha	0.0685	0.0230	2.9761	0.0030
5 Oz_beta	-0.0842	0.0162	-5.1910	2.2764e-07
6 Fpz_delta	-0.2865	0.0260	-11.0064	1.7200e-27
7 Fpz_theta	0.4749	0.0274	17.3414	2.1758e-63
8 Fpz_alpha	-0.1513	0.0249	-6.0862	1.3519e-09
9 Fpz_beta	-0.0169	0.0189	-0.8945	0.3712
10 EOG	-0.0663	0.0127	-5.2214	1.9359e-07
11 EMG	-0.0134	0.0124	-1.0794	0.2805

• موثر ترین متغیرها به ترتیب:

- باند تتابی الکترود Fpz (اثر مثبت)
- باند دلتای الکترود Fpz (اثر منفی)
- باند آلفای الکترود Fpz (اثر منفی)
- باند دلتای الکترود Oz (اثر مثبت)
- مابقی متغیرها تقریباً بی‌اثرند

همچنین چهار متغیری که نام برده شدند، p-value‌های بسیار ناچیزی دارند که همان‌طور که توضیح داده شد، اثر آن‌ها می‌تواند از نظر آماری قابل اعتنا باشد.

#### (۴-۲)

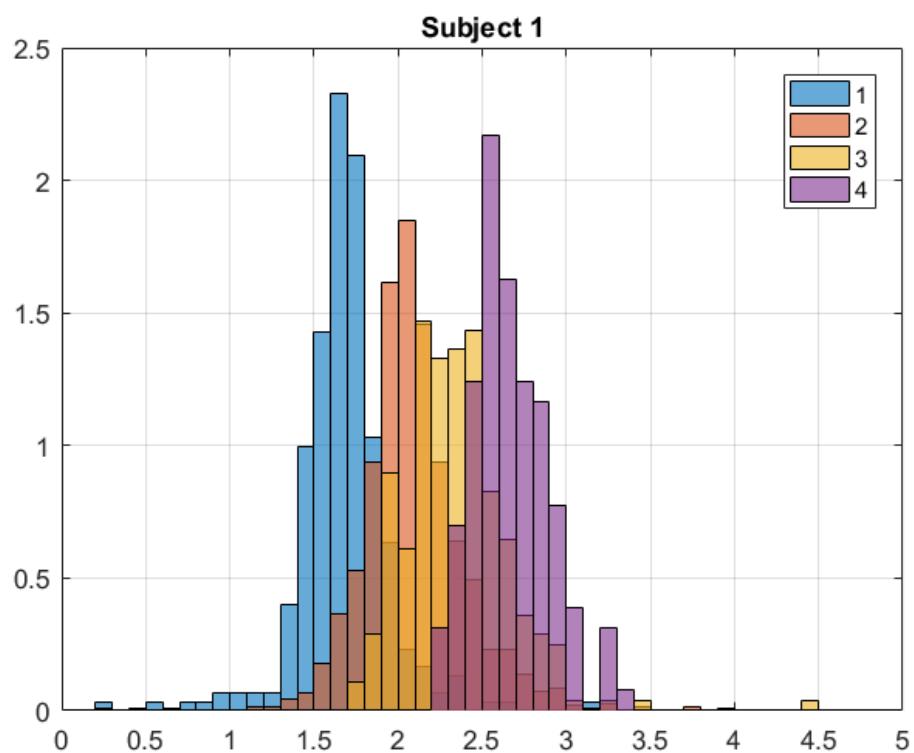
توضیحات کامل در مورد مفهوم R-squared در بخش قبلی بیان شد.

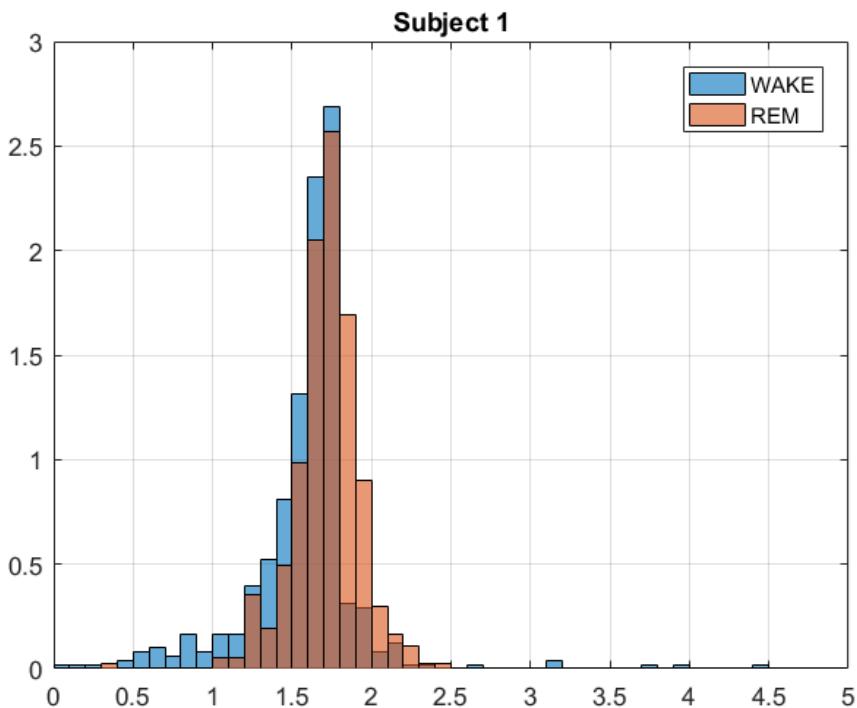
برای این سابجکت، مقدار R-squared برابر با  $0.32$  است. با توجه به توضیحاتی که در بخش قبل داده شد، می‌توان گفت که مدل توصیف‌کننده‌ی چندان جالبی نیست، چراکه مقادیر R-squared بین صفر و یک تغییر می‌کند و مقدار  $0.32$  نسبتاً پایین است. البته این سؤال که «R-squared خوب چقدر است؟» سؤال درستی نیست، چراکه بسته به کاربرد ما از مدل، ممکن است مقادیر مختلفی از این پاسخ برای ما کافی باشد، اما در حالت کلی می‌توان گفت که مدل ما،  $32\%$  از تغییرات پاسخ، توسط متغیرها توصیف می‌شود.

#### (۴-۳)

آن چه تا به اینجا به دست آورده‌ایم، بیشتر یک مدل توصیف‌کننده است تا پیش‌بینی‌کننده. علت این امر آن است که برای ما، آن‌چه مهم‌تر بوده است، آن است که کدام متغیرها دیتای موجود را بهتر توصیف می‌کنند و اثر مشخص‌تری دارند (سؤال اول) و نیز آن که این مدل، در کل، چه میزان از تغییرات پاسخ (عمق خواب) را به درستی توسط متغیرهای موجود توصیف می‌کند. (سؤال دوم)

در واقع، ما تا به اینجا سعی کردیم پارامترهای مؤثر را شناسایی کنیم تا بتوانیم ملاکی برای عمق خواب بیابیم، که این کار بیشتر در محدوده‌ی توصیف قرار می‌گیرد و نه پیش‌بینی. (البته این کار، می‌تواند مقدمه‌ای باشد برای انجام یک فعالیت پیش‌بینی‌کننده)



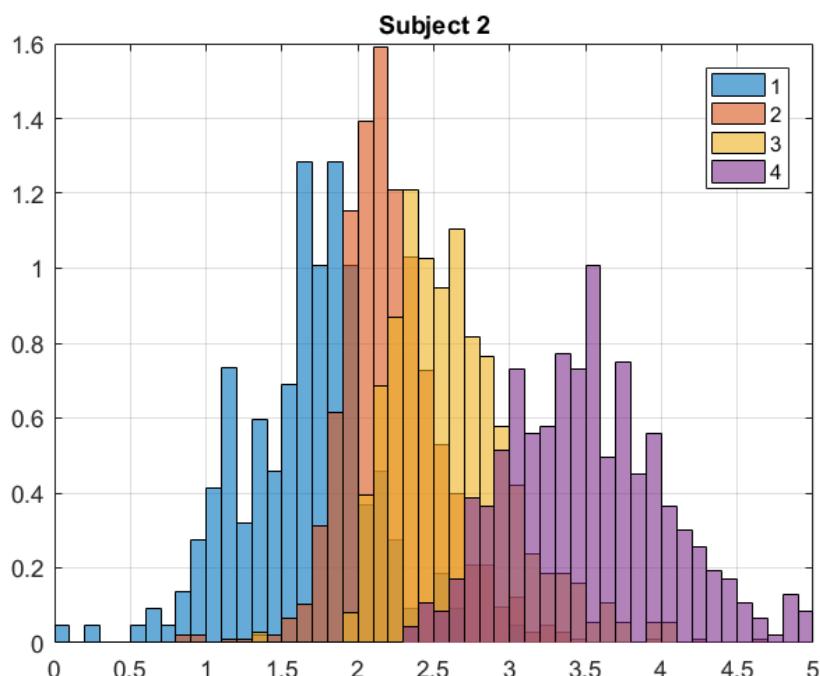


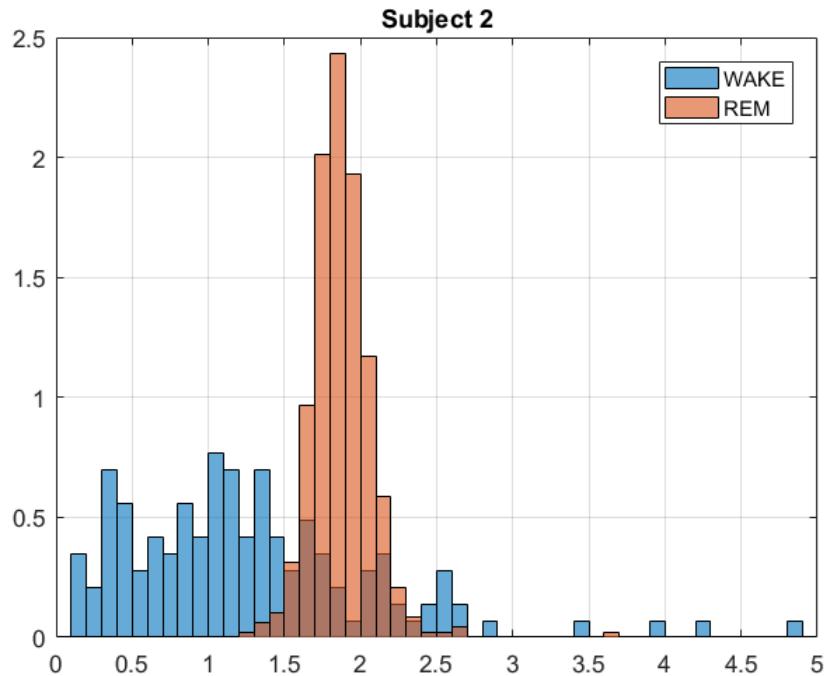
مشاهده می شود که هر دو وضعیت خواب REM و بیداری، عمق پیش‌بینی شده‌ای دارند که اکثریت تمرکز آن در محدوده ۱ تا ۲ است. این پیش‌بینی به شکل حدودی منطقی است، چراکه باید از عمق‌های بالای خواب فاصله‌ی زیادی داشته باشیم که در اینجا نیز چنین مشاهده می شود. به شکل دقیق‌تر، انتظار داریم که عمق پیش‌بینی شده برای بیداری، کمترین مقدار ممکن را به خود گرفته، و الزاماً کمتر از ۱ باشد، که این انتظار در اینجا برآورده نشده، البته می‌بینیم که توزیع بیداری، در مقایسه با توزیع REM بیشتر به سمت عمق‌های کم‌گراییش دارد که این نتیجه‌ی درستی است.

البته طبعاً انتظار نداریم که این مدل خطی همه چیز را به درستی توصیف کند، و ضمناً با بررسی سابجکت‌های دیگر در بخش‌های بعدی، به جمع‌بندی بهتری خواهیم رسید.

	1 Estimate	2 SE	3 tStat	4 pValue
1 (Intercept)	2.5162	0.0125	200.6715	0
2 Oz_delta	0.3485	0.0196	17.7966	3.7903e-66
3 Oz_theta	-0.0117	0.0219	-0.5332	0.5939
4 Oz_alpha	0.0861	0.0433	1.9873	0.0470
5 Oz_beta	-0.0731	0.0345	-2.1182	0.0343
6 Fpz_delta	0.2339	0.0205	11.3980	2.9798e-29
7 Fpz_theta	0.2551	0.0177	14.3892	7.2810e-45
8 Fpz_alpha	-0.0162	0.0277	-0.5833	0.5598
9 Fpz_beta	-0.3313	0.0201	-16.5138	9.9808e-58
10 EOG	0.0013	0.0172	0.0778	0.9380
11 EMG	-0.0241	0.0132	-1.8335	0.0669

- موثر ترین متغیرها به ترتیب:
  - باند دلتای الکترود Oz (اثر مثبت)
  - باند بتای الکترود Fpz (اثر منفی)
  - باند تتای الکترود Fpz (اثر مثبت)
  - باند دلتای الکترود Fpz (اثر مثبت)
  - مابقی متغیرها تقریباً بی اثرند
- مقدار R-squared برابر با ۰/۶۳ است، و مدل برای این سابجکت توصیف نسبتاً خوبی دارد، به طوری که ۶۳ درصد تغییرات حالت خواب توسط متغیرها توصیف می شود.

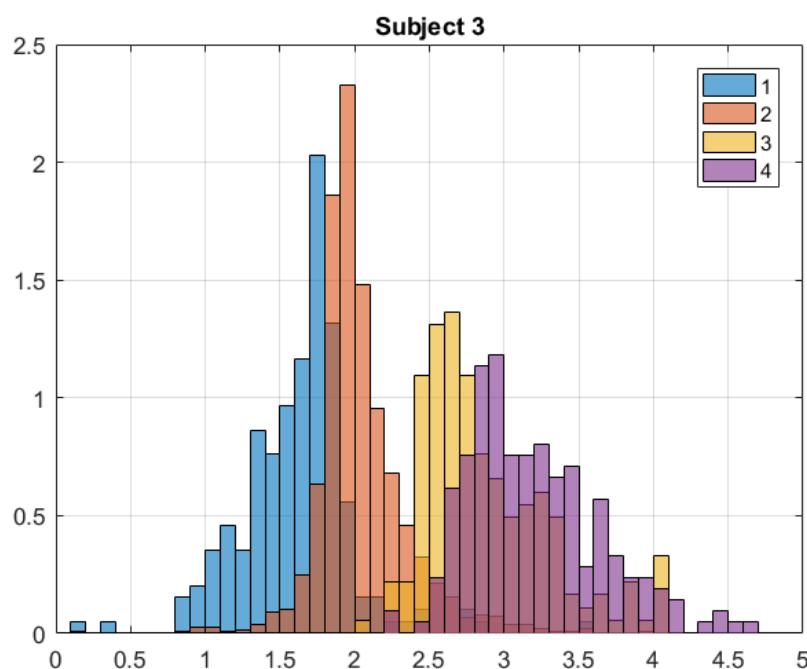


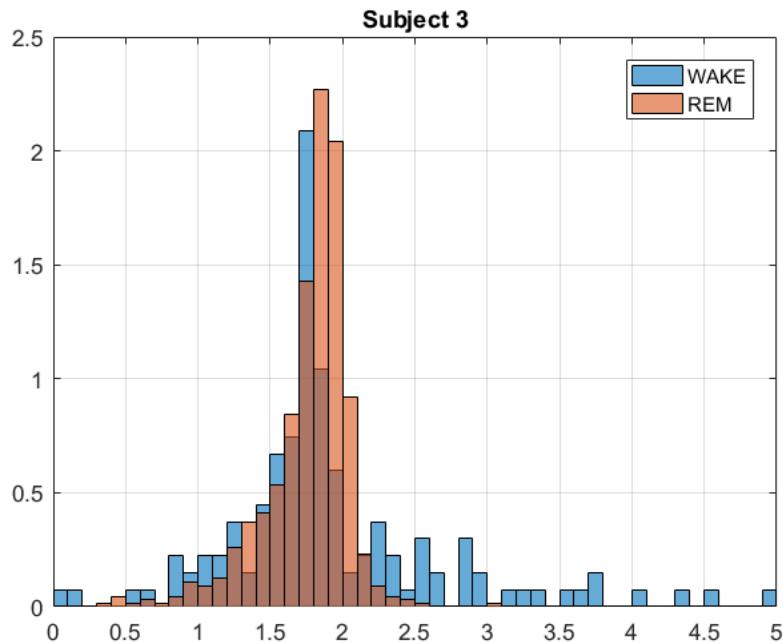


میزان جداشدگی حالت بیداری و REM و عمق پیش‌بینی شده برای آن‌ها در این ساچگشت، خوب و منطقی است، به طوری که حالت بیداری کمترین عمق، و به خصوص، عمقی کمتر از ۱ را به خود اختصاص می‌دهد، و دیتای REM نیز عمقی در حد فاصل ۱ تا ۲ دارد که با انتظارات ما، تناسب نسبی دارد، به خصوص که عمق هیچ کدام از این دو وضعیت، به حالت‌های عمیق خواب نزدیک نمی‌شود.

	1 Estimate	2 SE	3 tStat	4 pValue
1 (Intercept)	2.1866	0.0099	220.1455	0
2 Oz_delta	0.3834	0.0141	27.1023	9.4781e-140
3 Oz_theta	0.1740	0.0172	10.1269	1.3756e-23
4 Oz_alpha	-0.0686	0.0136	-5.0345	5.1828e-07
5 Oz_beta	-0.1393	0.0122	-11.4515	1.5697e-29
6 Fpz_delta	-0.0629	0.0257	-2.4523	0.0143
7 Fpz_theta	-0.1718	0.0473	-3.6337	2.8584e-04
8 Fpz_alpha	0.1185	0.0492	2.4109	0.0160
9 Fpz_beta	0.0511	0.0389	1.3133	0.1892
10 EOG	-0.1458	0.0105	-13.9521	1.8928e-42
11 EMG	0.0330	0.0101	3.2762	0.0011

- موثر ترین متغیرها به ترتیب:
  - باند دلتای الکترود Oz (اثر مثبت)
  - باند تتای الکترود Oz (اثر مثبت)
  - باند تتای الکترود Fpz (اثر منفی)
  - توان متوسط سیگنال EOG (اثر منفی)
  - باند بتای الکترود Oz (اثر منفی)
  - باند آلفای الکترود Fpz (اثر مثبت)
  - مابقی متغیرها تقریباً بی اثرند
- مقدار R-squared برابر با ۰/۵۹ است، و مدل برای این سابجکت توصیف نسبتاً خوبی دارد، به طوری که ۵۹ درصد تغییرات حالت خواب توسط متغیرها توصیف می شود.

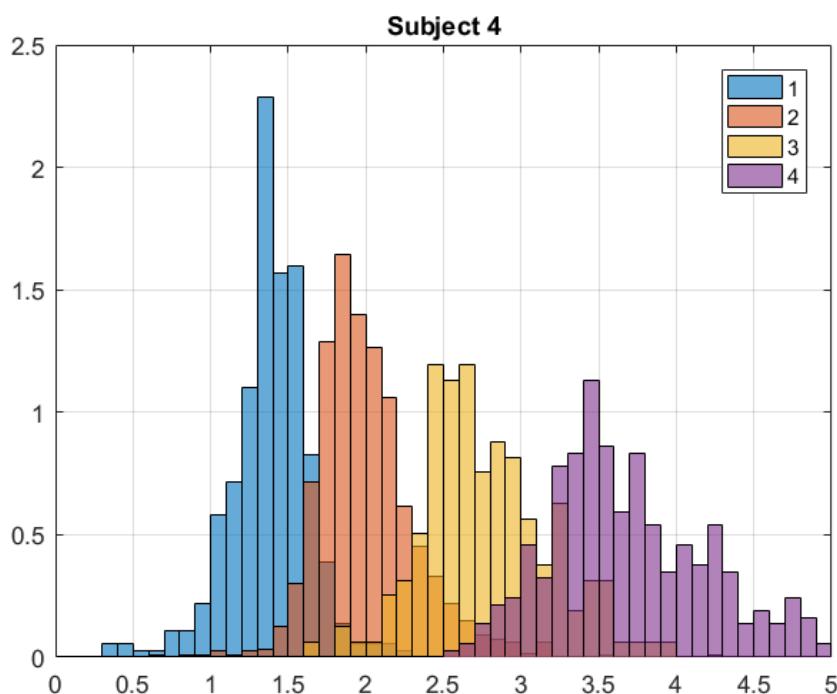


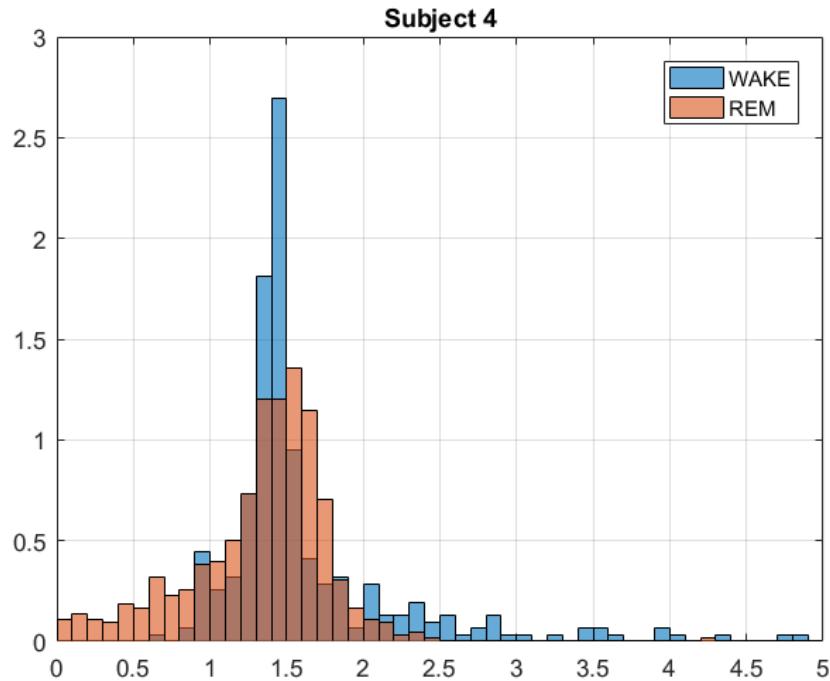


میزان جداشدگی حالت بیداری و REM و عمق پیش‌بینی شده برای آنها در این ساچگشت، نسبت به موارد قبلی خیلی جالب نیست، به خصوص که بیداری و REM تقریباً بر هم منطبق هستند، و نیز دیتای بیداری حضوری (اگرچه کم) در اعمق بالاتر نیز دارد. با این وجود، تمرکز هر دو دسته داده در حد فاصل اعمق ۱ تا ۲ است، که این با موارد قبلی (به خصوص در مورد خواب REM) سازگار است.

	1 Estimate	2 SE	3 tStat	4 pValue
1 (Intercept)	2.2669	0.0094	240.8793	0
2 Oz_delta	0.2561	0.0194	13.1994	2.9207e-38
3 Oz_theta	0.0710	0.0171	4.1639	3.2556e-05
4 Oz_alpha	0.0139	0.0248	0.5588	0.5763
5 Oz_beta	-0.0722	0.0127	-5.6902	1.4486e-08
6 Fpz_delta	0.3191	0.0204	15.6777	1.9130e-52
7 Fpz_theta	0.5684	0.0306	18.5998	1.4880e-71
8 Fpz_alpha	-0.0998	0.1173	-0.8507	0.3951
9 Fpz_beta	-0.3972	0.1171	-3.3919	7.0707e-04
10 EOG	-0.1492	0.0102	-14.6625	2.1331e-46
11 EMG	0.0321	0.0095	3.3735	7.5581e-04

- موثر ترین متغیرها به ترتیب:
  - باند بتای الکترود Fpz (اثر مثبت)
  - باند بتای الکترود Fpz (اثر منفی)
  - باند دلتای الکترود Fpz (اثر مثبت)
  - باند دلتای الکترود Oz (اثر مثبت)
  - توان متوسط سیگنال EOG (اثر منفی)
  - باند آلفای الکترود Fpz (اثر منفی)
  - مابقی متغیرها تقریباً بی اثرند
- مقدار R-squared برابر با ۰/۸ است، و مدل برای این سابجکت توصیف خوبی دارد، به طوری که ۸۰ درصد تغییرات حالت خواب توسط متغیرها توصیف می شود.

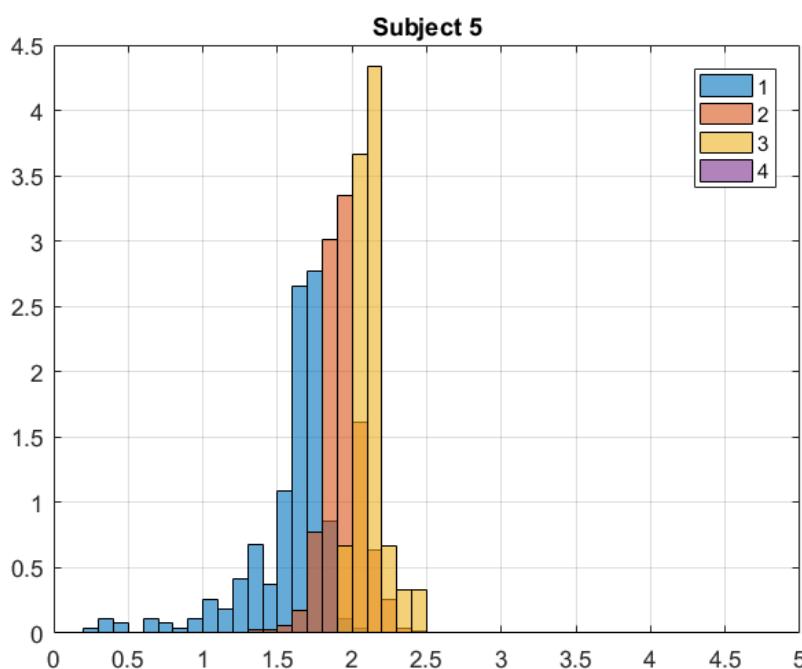


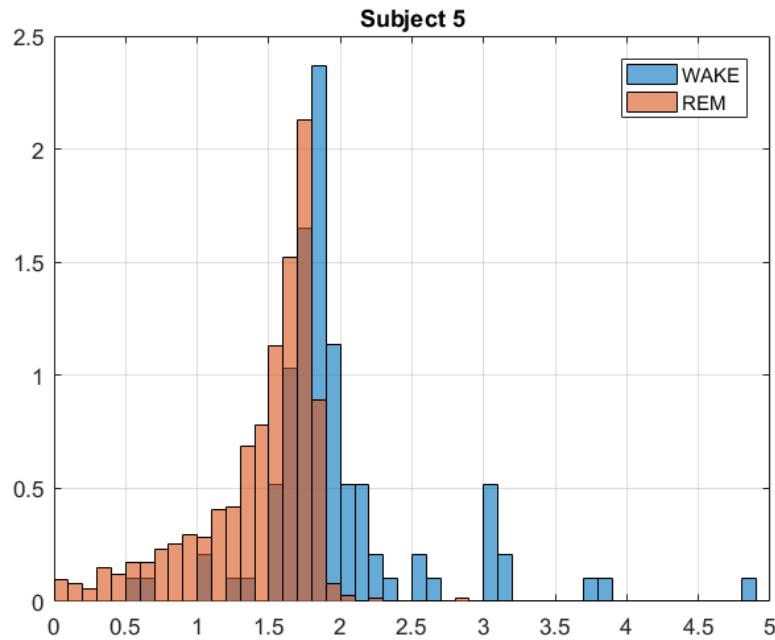


مجدداً در این حالت نیز تمرکز عمق پیش‌بینی شده، هم برای REM، و هم برای بیداری، در محدوده ۱ تا ۲ است، با این تفاوت که این بار، کشیدگی REM به سمت اعمق کمتر و کشیدگی بیداری به سمت اعمق بیشتر است، که این نتیجه نادرست است و نشان می‌دهد که پیش‌بینی انجام شده، ایراد دارد.

	1 Estimate	2 SE	3 tStat	4 pValue
1 (Intercept)	1.8852	0.0064	292.2990	0
2 Oz_delta	0.0850	0.0082	10.3588	1.5180e-24
3 Oz_theta	0.0627	0.0104	6.0071	2.2288e-09
4 Oz_alpha	0.0727	0.0115	6.3392	2.8309e-10
5 Oz_beta	-0.1634	0.0123	-13.2475	1.6903e-38
6 Fpz_delta	-0.0202	0.0116	-1.7380	0.0824
7 Fpz_theta	-0.1892	0.0248	-7.6425	3.2477e-14
8 Fpz_alpha	0.3697	0.0411	8.9990	5.0580e-19
9 Fpz_beta	-0.1395	0.0422	-3.3059	9.6313e-04
10 EOG	-0.1605	0.0093	-17.2808	1.4127e-62
11 EMG	-0.0271	0.0069	-3.9375	8.5072e-05

- موثر ترین متغیرها به ترتیب:
  - باند آلفای الکترود Fpz (اثر مثبت)
  - باند بتای الکترود Fpz (اثر منفی)
  - باند بتای الکترود Oz (اثر منفی)
  - توان متوسط سیگнал EOG (اثر منفی)
  - باند بتای الکترود Fpz (اثر منفی)
  - مابقی متغیرها تقریباً بی اثرند
- مقدار R-squared برابر با ۰/۳۵ است، و مدل برای این سابجکت توصیف چندان جالبی ندارد، به طوری که ۳۵ درصد تغییرات حالت خواب توسط متغیرها توصیف می شود.





مجدداً در این حالت نیز تمرکز عمق پیش‌بینی شده، هم برای REM، و هم برای بیداری، در محدوده ۱ تا ۲ است، با این تفاوت که این بار، کشیدگی REM به سمت اعمق کمتر و کشیدگی بیداری به سمت اعمق بیشتر است، که این نتیجه نادرست است و نشان می‌دهد که پیش‌بینی انجام شده، ایراد دارد.

سابجکت پنجم	سابجکت چهارم	سابجکت سوم	سابجکت دوم	سابجکت اول	
$Fpz.\alpha +$	$Fpz.\theta +$	$Oz.\delta +$	$Oz.\delta +$	$Fpz.\theta +$	متغیر اول
$Fpz.\theta -$	$Fpz.\beta -$	$Oz.\theta +$	$Fpz.\beta -$	$Fpz.\delta -$	متغیر دوم
$Oz.\beta -$	$Fpz.\delta +$	$Fpz.\theta -$	$Fpz.\theta +$	$Fpz.\alpha -$	متغیر سوم
$EOG -$	$Oz.\delta +$	$EOG -$	$Fpz.\delta +$	$Oz.\delta +$	متغیر چهارم
$Fpz.\beta -$	$EOG -$	$Oz.\beta -$			متغیر پنجم
	$Fpz.\alpha -$	$Fpz.\alpha -$			متغیر ششم

جدول فوق، نتایج بخش اول تحلیل هر سابجکت (کدام متغیرها اثر بیشتری بر عمق خواب دارند) را به شکل خلاصه شده نمایش می‌دهد.

باند تای الکترود  $Fpz$  در تمامی ستون‌های این جدول حضور دارد، و این یعنی به نظر، یکی از مهم‌ترین مؤلفه‌های تشخیص عمق خواب است. البته در دو مورد اثر منفی و در سه مورد اثر مثبت دارد که این کمی نتیجه‌گیری کلی ما را تحت تأثیر قرار می‌دهد، چراکه نمی‌توان به تنهایی، رابطه‌ی مشخصی بین انرژی باند تای الکترود  $Fpz$  و عمق خواب تعیین کرد.

باند دلتای الکترود  $Oz$  و باند آلفای الکترود  $Fpz$  در چهار ستون از پنج ستون جدول فوق حضور دارند. باند آلفای  $Fpz$  در سه مورد اثر منفی و در یک مورد اثر مثبت داشته است، و باند دلتای الکترود  $Oz$  در تمامی موارد اثر مثبت داشته است. همچنین در یک موردی هم که این متغیر به عنوان متغیرهای تأثیرگذار شناسایی نشد، باز هم اثر آن مثبت (هر چند ناچیز) بوده است.

در ادامه، باند دلتای الکترود  $Fpz$ ، باند بتای الکترود  $Fpz$  و سیگنال  $EOG$  نیز سه حضور در جدول فوق داشته اند.

در جدول فوق، ۲۵ متغیر در ۵ ستون نام برده شده اند که سه متغیر فوق (باند تای الکترود  $Oz$  و باند آلفای الکترود  $Fpz$ )، در ۱۳ مورد از این ۲۵ مورد مشاهده می‌شوند، یعنی بیش از ۵۰ درصد موارد را این سه به خود اختصاص داده‌اند و می‌توان این سه را به عنوان پارامترهای برتر در نظر گرفت.

بنابراین به عنوان جمع‌بندی می‌توان گفت که این مشاهدات برای سابجکت‌های مختلف یکسان نیست، اما می‌توان ویژگی‌هایی را معرفی کرد که تقریباً در تمامی سابجکت‌ها اثرگذار بوده‌اند و می‌توان ادعا کرد که این ویژگی‌ها، "ویژگی‌های مؤثر" هستند.

در مورد بخش آخر، یعنی پیش‌بینی عمق خواب REM و حالت بیداری نیز، با توجه به مشاهدات انجام شده، باید گفت که این کار چندان دقیق نیست، اما نتایج در حالت کلی عمدتاً در محدوده عمق ۱ تا ۲ توزیع شده‌اند، که از نظر فاصله داشتن با اعماق بالا قابل قبول است، اما به خصوص برای بیداری انتظار داریم که عمق بسیار پایین و به ویژه کمتر از ۱ داشته باشد، که به جز یک سابجکت، سایرین چنین نتیجه‌هایی ارائه ندادند و این یعنی نمی‌توان از این برونویابی، انتظار چندانی داشت. همچنین یک نکته دیگر آن است که در اکثر مشاهدات، تمایز چندانی بین حالت بیداری و REM مشاهده نمی‌شد.

بخشی از استدلال‌های این بخش، در صفحه‌ی قبل بیان شد. با استناد به توضیحات صفحه‌ی قبل و جدول رسم شده، به این نتیجه می‌رسیم که پارامترهای تأثیرگذار برای سابجکت‌های مختلف به طور دقیق یکسان نیستند، به خصوص که ترتیب اثرگذاری آن‌ها دقیقاً بر هم منطبق نیست، اما با بررسی پارامترهای مؤثر، می‌بینیم که سه پارامتر وجود دارند که تقریباً در تمامی سابجکت‌ها اثرگذار هستند. این سه عبارتند از باند دلتای الکترود Oz، باند دلتای الکترود Fpz، و باند بتای الکترود Fpz.

از میان این سه، باند دلتای الکترود Oz تنها موردی است که در تمامی موارد، اثر آن از نظر مستقیم یا معکوس بودن یکسان است، و برای تمامی پنج سابجکت مورد بررسی، اثر مشتبی بر عمق خواب داشته است، یعنی با افزایش مقدار آن، عمق خواب نیز زیاد شده است.

این نتیجه، دقیقاً همان نتیجه‌ای است که در قسمت تحلیل PCA نیز به دست آمده بود. (و نیز مقالات علمی منتشرشده نیز تا حد قابل قبولی آن را تأیید می‌کنند)

در حالت کلی نیز باید توجه داشت که این دو تحلیل (Linear Regression و PCA) ماهیت‌هایی متفاوت و فرض‌هایی متفاوت دارند و نباید انتظار داشت که نتایج آن‌ها کاملاً منطبق باشند.

ضمیناً در تحلیل PCA، توان متوسط سیگنال EOG به عنوان پارامتر جداکننده خواب REM شناسایی شد که در این بخش، ما به طور کلی بخش مربوط به خواب REM را از دیتاست حذف کردیم و انتظار نداریم نتیجه‌ی متناظری به دست بیاید.

در مورد R-squared، به طور متوسط مقادیری بیش از ۵۰٪ مشاهده می‌شد که می‌توانست قابل قبول باشد.

نهایتاً در مورد پیش‌بینی عمق برای بیداری و REM، دیدیم که نتایج چندان جالب توجه نیستند و انتظاراتی از این قبیل که REM و بیداری جدا شوند، یا این که عمق بیداری تفاوت معنی‌داری با صفر نداشته باشد را برآورده نمی‌کنند، و عمق پیش‌بینی شده برای هر دو، در فاصله ۱ تا ۲ مترمکز است؛ اما از این نظر که با اعمق بالا تفاوت معنی‌داری مشاهده می‌شود، نتایج درست و معقول هستند.

**قسمت پنجم) یک سؤال دلخواه!**

در این بخش به بررسی چند سؤال مختلف پرداخته ایم.

سوال اول) می خواهیم با استفاده از طبقه بندی داده های موجود با توجه به ویژگی های استخراج شده (ماتریس X)، و با اعتبار سنجی Cross Validation ببینیم که این ویژگی چه میزان موفق به جداسازی حالت های مختلف خواب می شود.

برای پاسخ گویی به این قسمت، از روش طبقه بندی خطی SVM استفاده می کنیم. برای این کار، یک راه استفاده از زوج های توابع متلب مثل fitcsvm و svmtrain یا svmlclassify است، که البته این توابع فرایند طبقه بندی را برای دو کلاس انجام می دهند، لذا باید برای n کلاس، ابتدا n طبقه بندی کننده دو کلاسه آموزش دهیم، سپس دیتای تست را به تمامی آنها بدھیم، و هر کدام که دیتا را جزء گروه خود اعلام کرد، جواب مسأله را به ما می دهد. (وظیفه هر طبقه بندی کننده، جداسازی دیتای یک گروه از سایر گروه ها است) بعد از این کار، باید با روش اعتبار سنجی Cross Validation، ببینیم که ویژگی های استخراج شده، تا چه حد می توانند دیتای مربوط به دسته های مختلف را از هم جدا کنند. روش k-fold Cross Validation داده های آموزش را به k دسته تقسیم می کند، سپس فرایند طبقه بندی را k بار انجام می دهد که هر بار، یک دسته به عنوان دیتای تست، و k-1 دسته به عنوان دیتای آموزش به کار می روند. نهایتاً درصد پاسخ گویی صحیح در این k آزمایش را میانگین می گیرد و به عنوان درصد پاسخ گویی نهایی اعلام می کند. در این مسأله ما داده های جدیدی برای طبقه بندی نداریم، لذا صرفاً نتیجه های فرایند Cross Validation است که حائز اهمیت است.

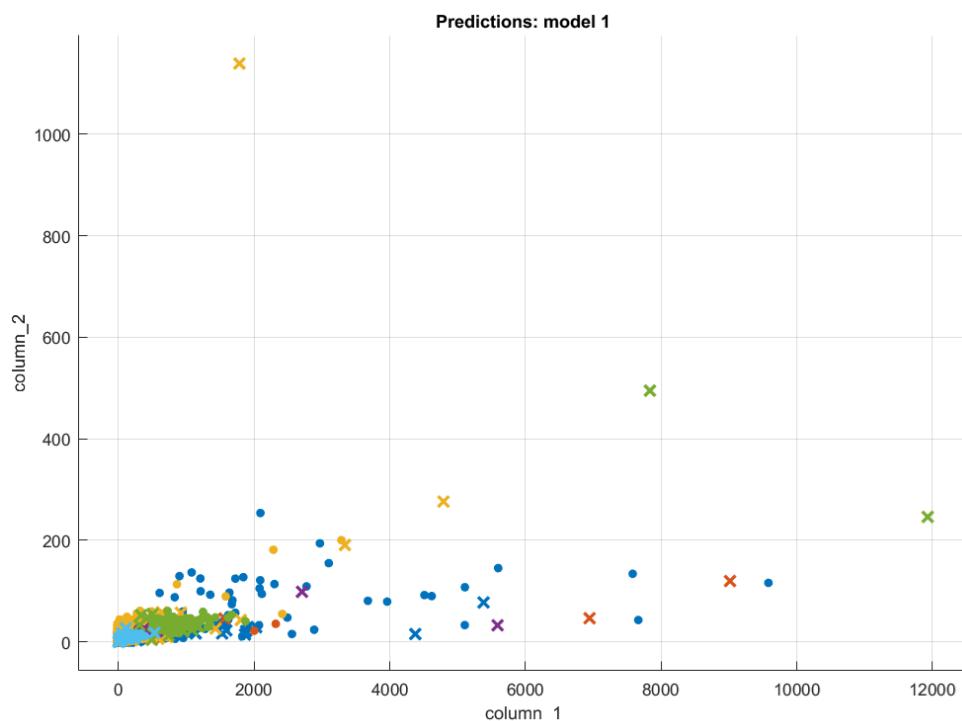
یک نکته در مورد Cross Validation و این که درصد پاسخ آن چقدر قابل اعتماد است وجود دارد، و آن نکته این است که گاهی ویژگی های استخراج شده، اثری (مانند یک آماره) از کل دیتای آموزش را در خود دارند، لذا اگرچه در فرایند اعتبار سنجی، داده ها به دو دسته تقسیم شده اند و ما به ظاهر، از داده های تست کاملاً بی اطلاعیم، اما ویژگی های استخراج شده دارای یک بایاس مرتبط با مجموعه کلی داده ها هستند، که می تواند باعث شود نتیجه Cross Validation بهتر از نتیجه های شود که در صورت آزمایش روی داده های جدید به دست آید. اگر این مشکل وجود نداشته باشد، می توان ادعا کرد که این اعتبار سنجی کاملاً مشابه مسأله ای است که در آن، دیتای تست جدید و ناشناخته ای در اختیار ما قرار داده باشند و بخواهیم آنها را طبقه بندی کنیم.

همان طور که اشاره شد، در اینجا ما صرفاً می خواهیم ببینیم که این ویژگی ها چه میزان قادر به جداسازی هستند و صرفاً نتیجه فرایند Cross Validation برای ما اهمیت دارد. برای پیاده سازی این فرایند، از ابزار Classification Learner مطلب استفاده می کنیم و این کار را برای هر ۵ سابجکت انجام می دهیم و برای هر یک، نمودار دیتای درست و اشتباه تشخیص داده شده در دو راستا، نتیجه Cross Validation و درصد صحیت، و confusion matrix (نشان دهنده این است که تشخیص های اشتباه بین کدام دو گروه اتفاق افتاده اند) را گزارش می کنیم، و در نهایت به جمع بندی می پردازیم.

ضمیناً در تمامی نمودارهای پراکنده رسم شده، راهنمای نمودار به صورت زیر است که شماره کلاس ها، متناظر با بردار state است که در طول پروره با آن کار کرده ایم.

Classes	
<input checked="" type="checkbox"/>	0
<input checked="" type="checkbox"/>	1
<input checked="" type="checkbox"/>	2
<input checked="" type="checkbox"/>	3
<input checked="" type="checkbox"/>	4
<input checked="" type="checkbox"/>	6

الف) نمودار پراکندگی در دو راستا (ضربدرها نقاطی هستند که اشتباه تشخیص داده شده اند)



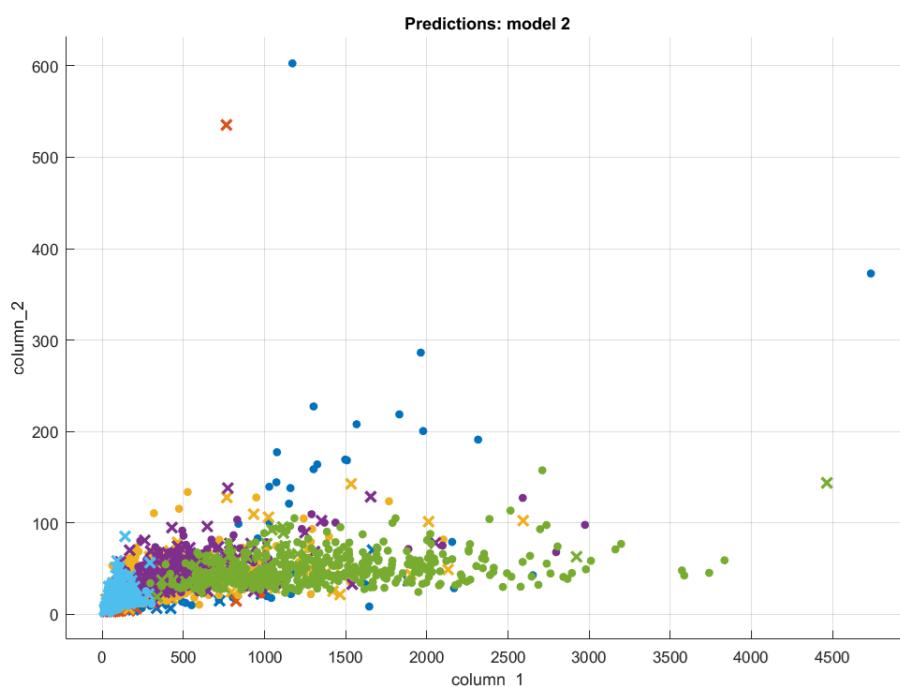
ب) حاصل Cross Validation



ج) Confusion Matrix

Model 1						
True class	0	1	2	3	4	5
0	423	9	48		2	5
1	157	10	127		1	8
2	81	7	1477	2	11	9
3			238	6	35	
4			48	7	74	
5	24	1	63			278

الف) نمودار پراکندگی در دو راستا (ضربدرها نقاطی هستند که اشتباه تشخیص داده شده اند)



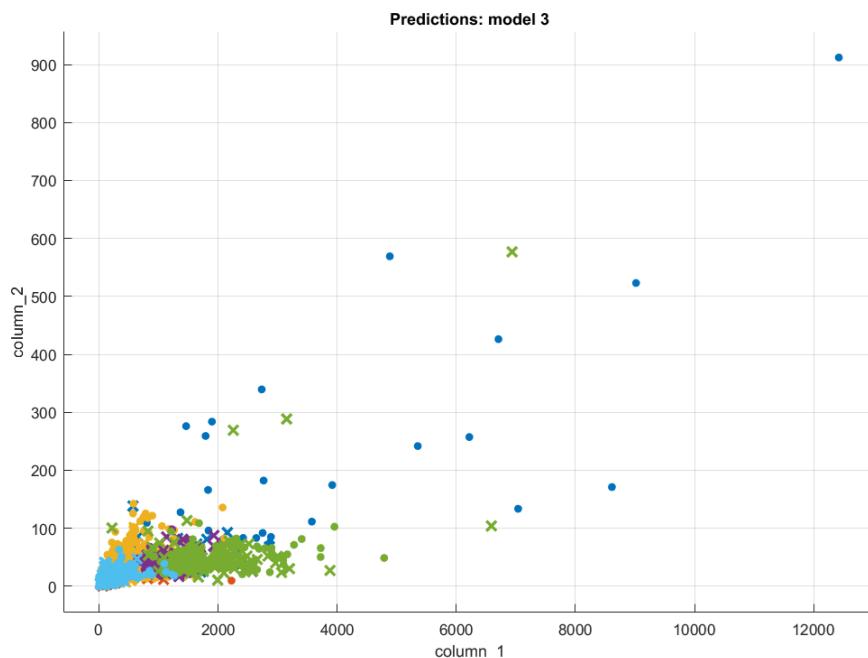
ب) حاصل



Confusion Matrix (ج)

Model 2						
True class	0	1	2	3	4	5
0	77	49	21	1	1	14
1	15	72	33	1		101
2	10	8	875	55	5	106
3	1		178	173	28	1
4			7	42	422	
5	1	14	37			425

الف) نمودار پراکندگی در دو راستا (ضربدرها نقاطی هستند که اشتباه تشخیص داده شده اند)



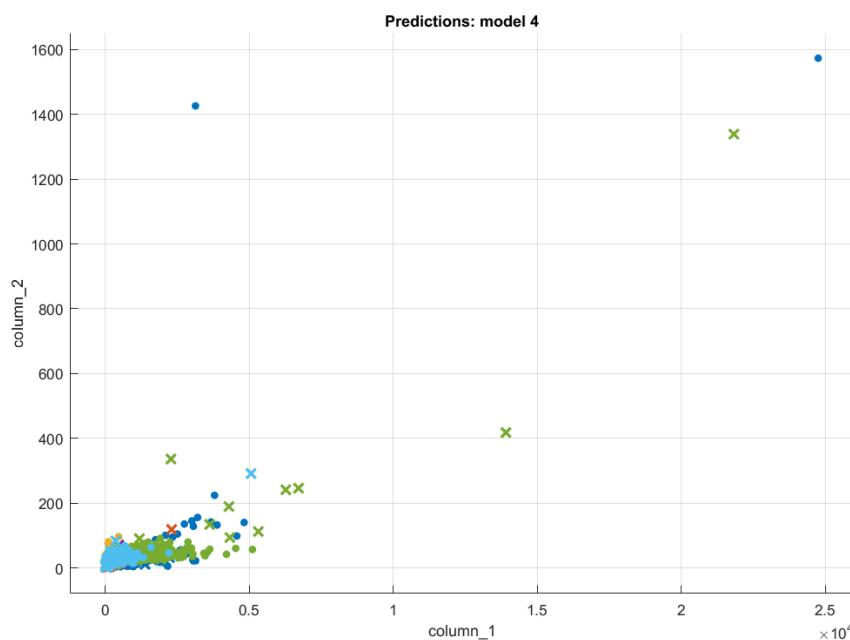
ب) حاصل Cross Validation



ج) Confusion Matrix

Model 3						
True class	0	1	2	3	4	6
0	50	25	16	1	8	51
1	13	29	28		2	126
2	4	9	1413	28	17	137
3	1		65	56	61	
4	2		10	53	148	
6	2	8	45	1		598

الف) نمودار پراکندگی در دو راستا (ضربدرها نقاطی هستند که اشتباه تشخیص داده شده اند)



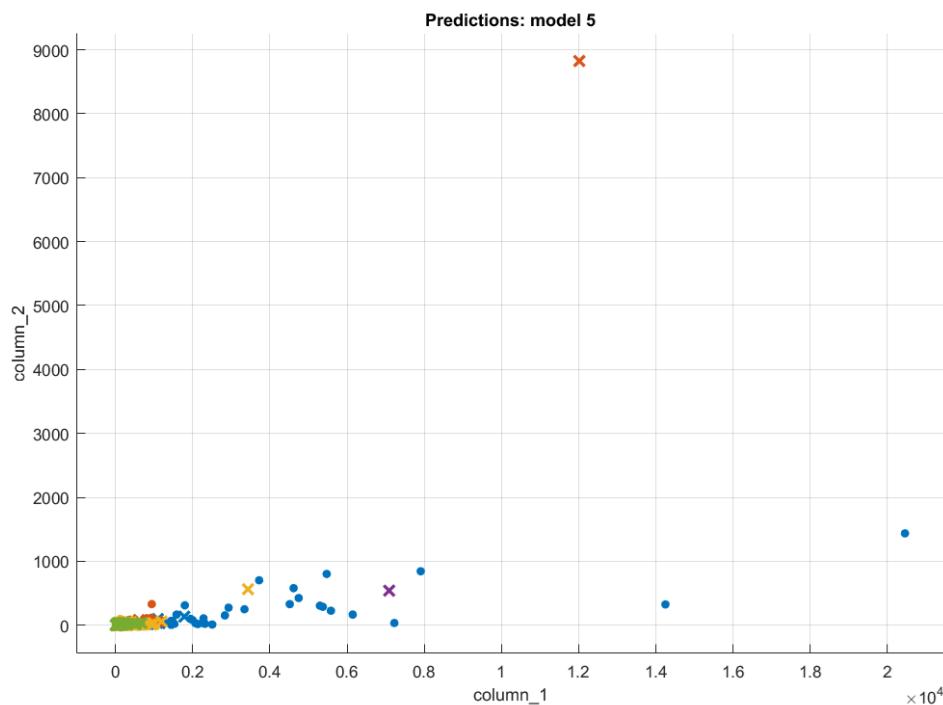
ب) حاصل Cross Validation



ج) Confusion Matrix

Model 4						
True class	0	1	2	3	4	6
0	246	61	1	1	11	5
1	40	241	27			55
2	6	20	1070	17	3	72
3	4		39	81	35	
4	4			32	345	
6	10	34	47		1	586

الف) نمودار پراکندگی در دو راستا (ضربدر ها نقاطی هستند که اشتباه تشخیص داده شده اند)



ب) حاصل Cross Validation

<b>5 ★ SVM</b>	<b>Accuracy: 81.4%</b>
Last change: Linear SVM	10/10 features

Confusion Matrix (ج)

		Model 5				
		0	1	2	3	6
True class	0	34	13	60	1	7
	1	11	31	80		145
2	1	6	1689		71	
3			30			
6	3	12	116		682	

نتیجه‌گیری:

- ۱ - متوسط درصد پاسخ‌گویی (که توسط 5-fold Cross Validation محاسبه شده است) برابر ۷۷٪، و محدوده تغییرات درصد پاسخ‌گویی بین ۷۲ تا ۸۳ درصد است، که این یعنی درصد پاسخ‌گویی بالا و قابل قبول است، و می‌توان ادعا کرد که ۱۰ ویژگی استخراج شده، توانایی جداسازی نسبتاً خوبی برای ۶ حالت خواب (و بیداری) دارند.
- ۲ - با توجه به Confusion Matrix‌ها، می‌بینیم که داده‌های عمق سوم و چهارم خواب، به شکل معنی‌داری از داده‌های عمق اول و دوم، و نیز حالت REM و بیداری جدا می‌شوند، به این معنی که میزان اشتباه میان این دو دسته (یعنی وضعیتی که داده‌ای متعلق به یکی از این دو دسته باشد و به اشتباه، جزء دسته دیگر شناخته شود) بسیار ناچیز است، اما میزان اشتباه در داخل هر یک از این دو دسته، به نسبت زیاد است.
- ۳ - یک نکته‌ی دیگر که جالب توجه است، آن است که این روش دسته‌بندی، عملکرد قابل توجهی برای شناسایی عمق دوم خواب دارد، به طوری که در هر پنج سابجکت، درصد پاسخ صحیح برای این عمقدار، با اختلاف از سایرین بیشتر است.

**سؤال دوم) خوشبندی**

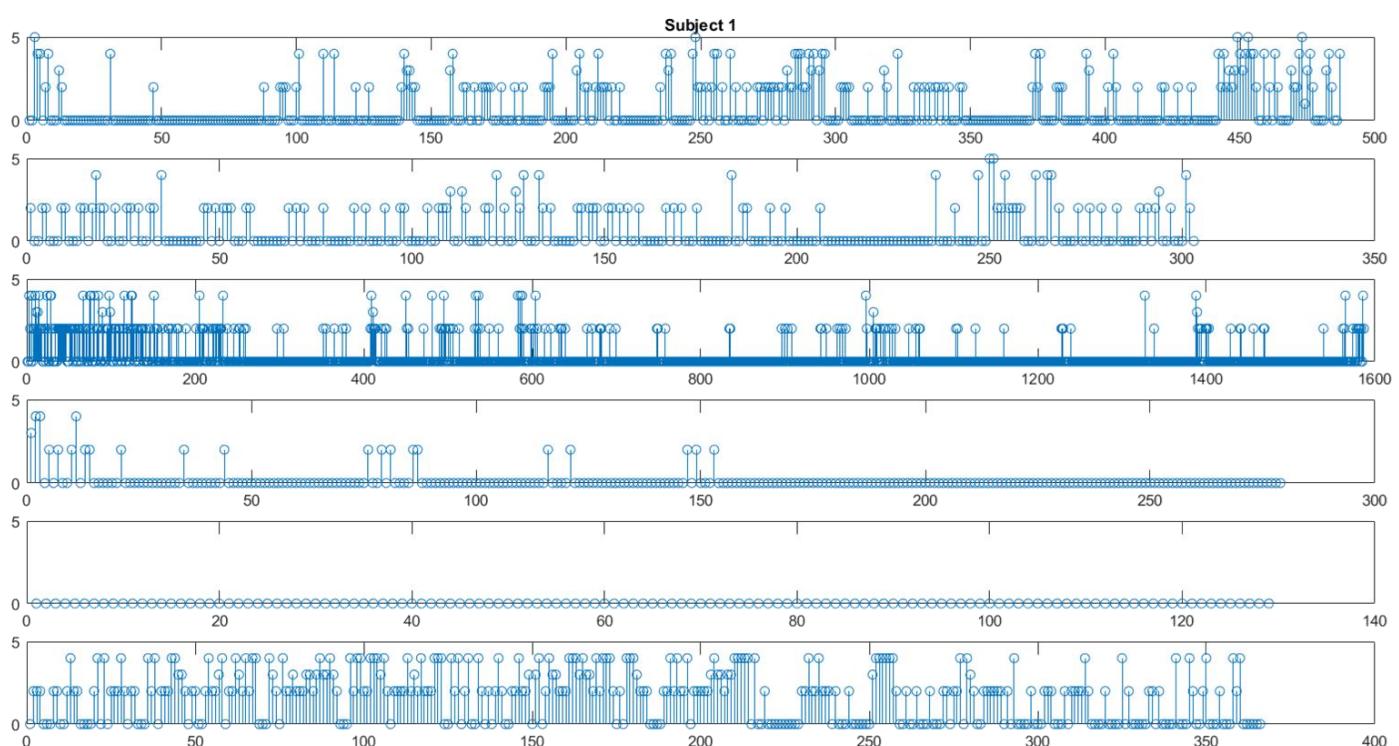
در این بخش می خواهیم دیتای به دست آمده از استخراج ویژگی های مربوطه را خوشبندی کنیم. این کار که نوعی unsupervised learning است، از جهاتی مشابه بخش قبل است، اما در اینجا ما صرفاً دیتا را در اختیار الگوریتم قرار می دهیم تا آن را به چند دسته تقسیم کند.

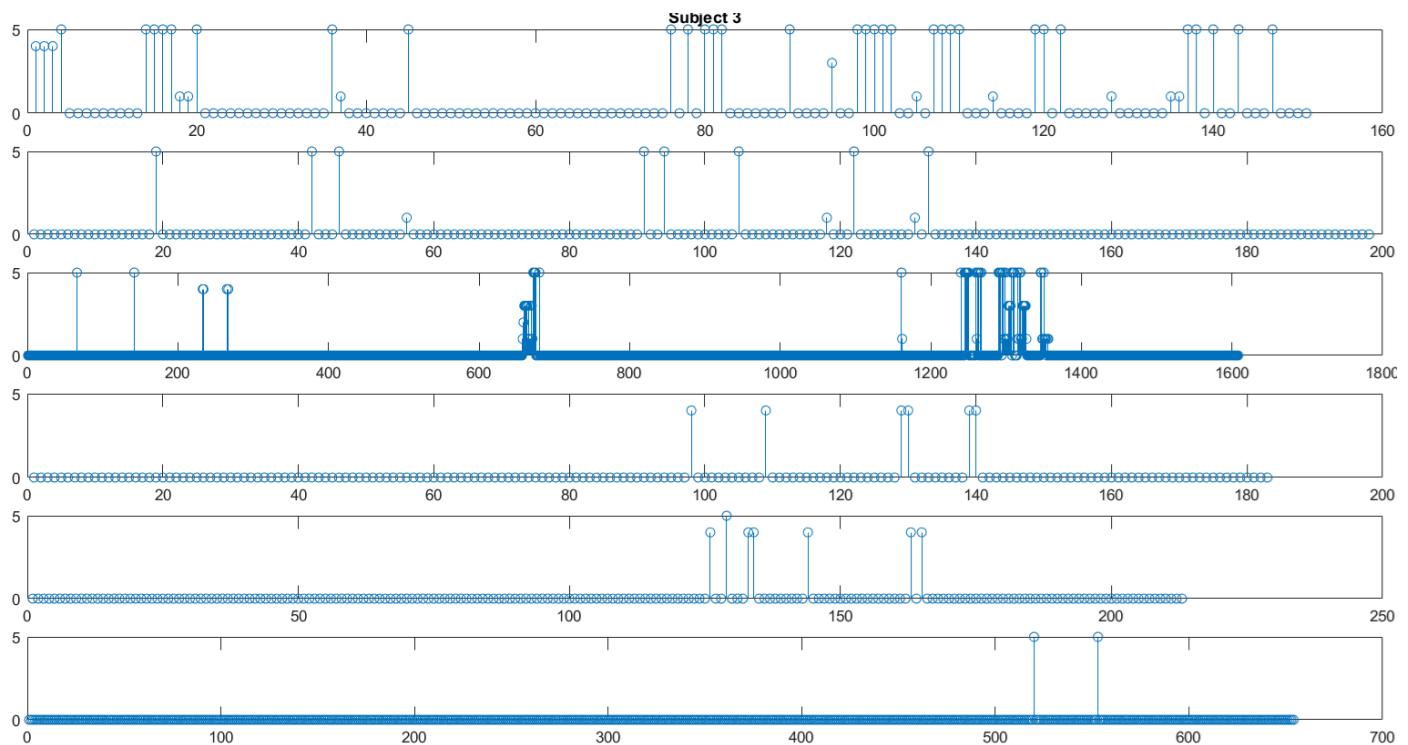
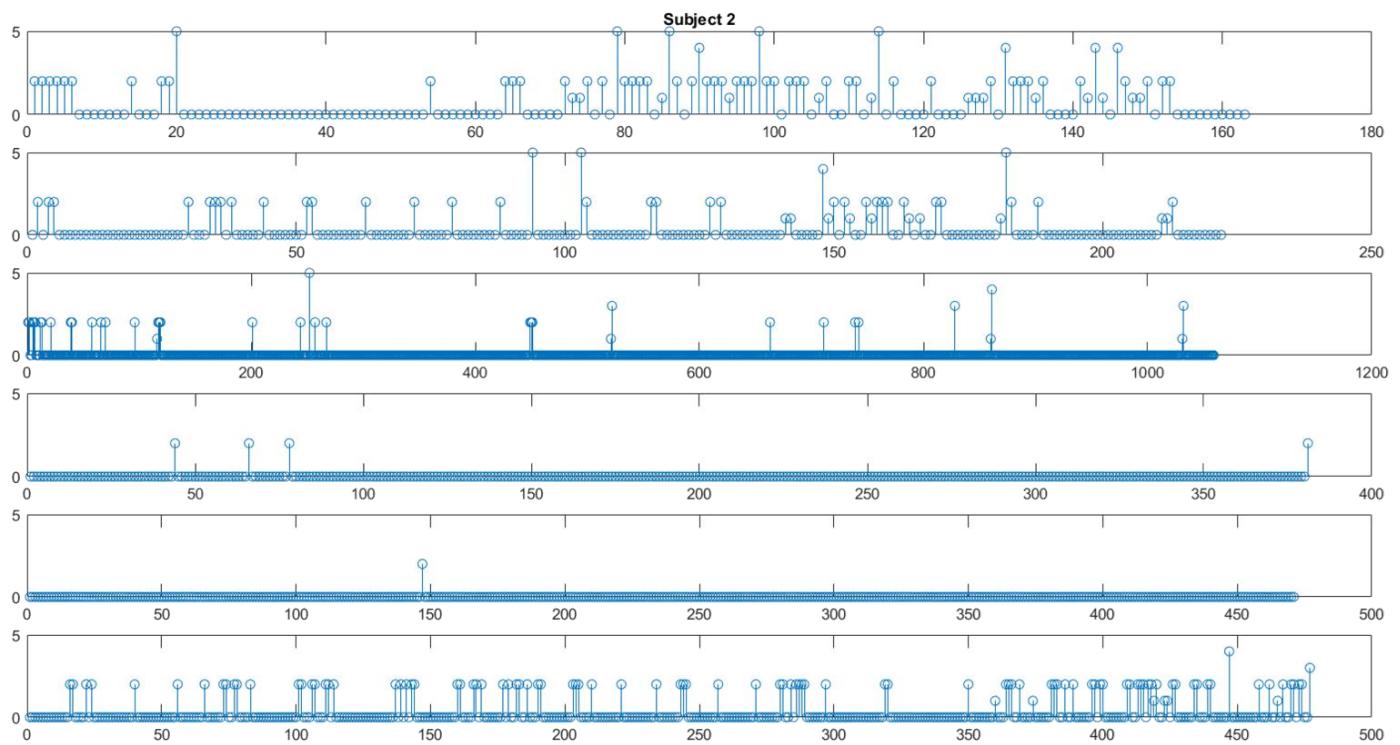
از الگوریتم k-means استفاده می کنیم و ماتریس ویژگی ها ( $X$ ) را به عنوان ورودی در نظر می گیریم. همچنین با توجه با دانش اولیه خود،  $k = 6$  را انتخاب می کنیم.

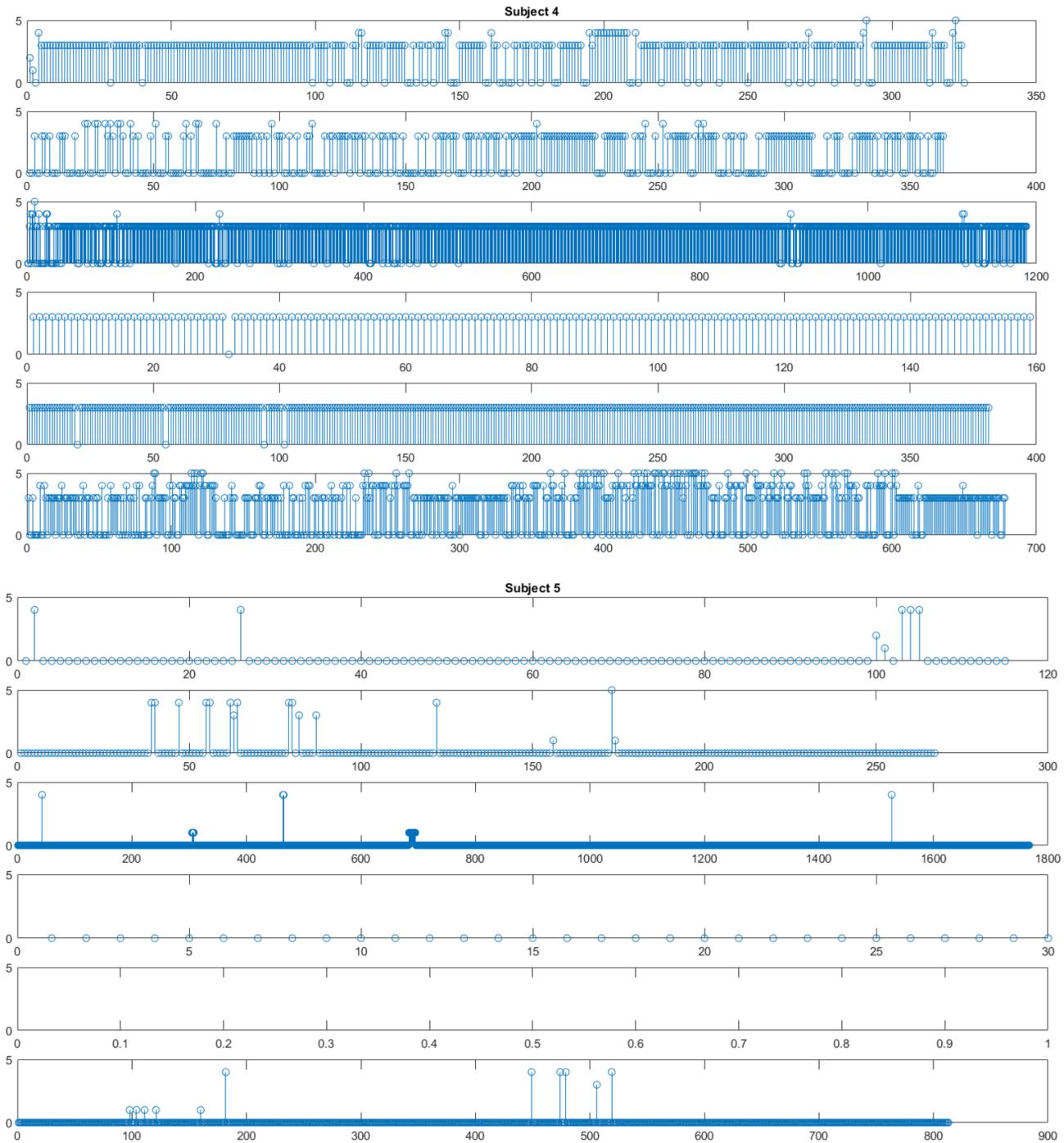
پس از انجام خوشبندی، 6 نمودار رسم می کنیم که در هر کدام، شماره گروه جدید (حاصل از خوشبندی) برای اعضای یک گروه (که از قبل می دانیم) رسم شده است.

اگر پاسخ کاملاً ایدهآل باشد، باید نمودارهایی داشته باشیم که هر کدام، به شکل خط افقی  $y = k$  هستند که مقدار  $k$  در هر نمودار متفاوت است. (این یعنی گروه ها دقیقاً طبق انتظارات اولیه ما جدا شده اند)

در ادامه نمودارهای متناظر با خوشبندی را برای هر پنج سابجکت مشاهده می کنید:







آنچه در نمودارهای فوق مشاهده می‌شود، اطلاعات جالبی نیست، چراکه تقریباً در هر شکل، همه نمودارها نشان‌دهنده‌ی یک مقدار غالب هستند، و سایر مقادیر به شکل پراکنده مشاهده می‌شوند.

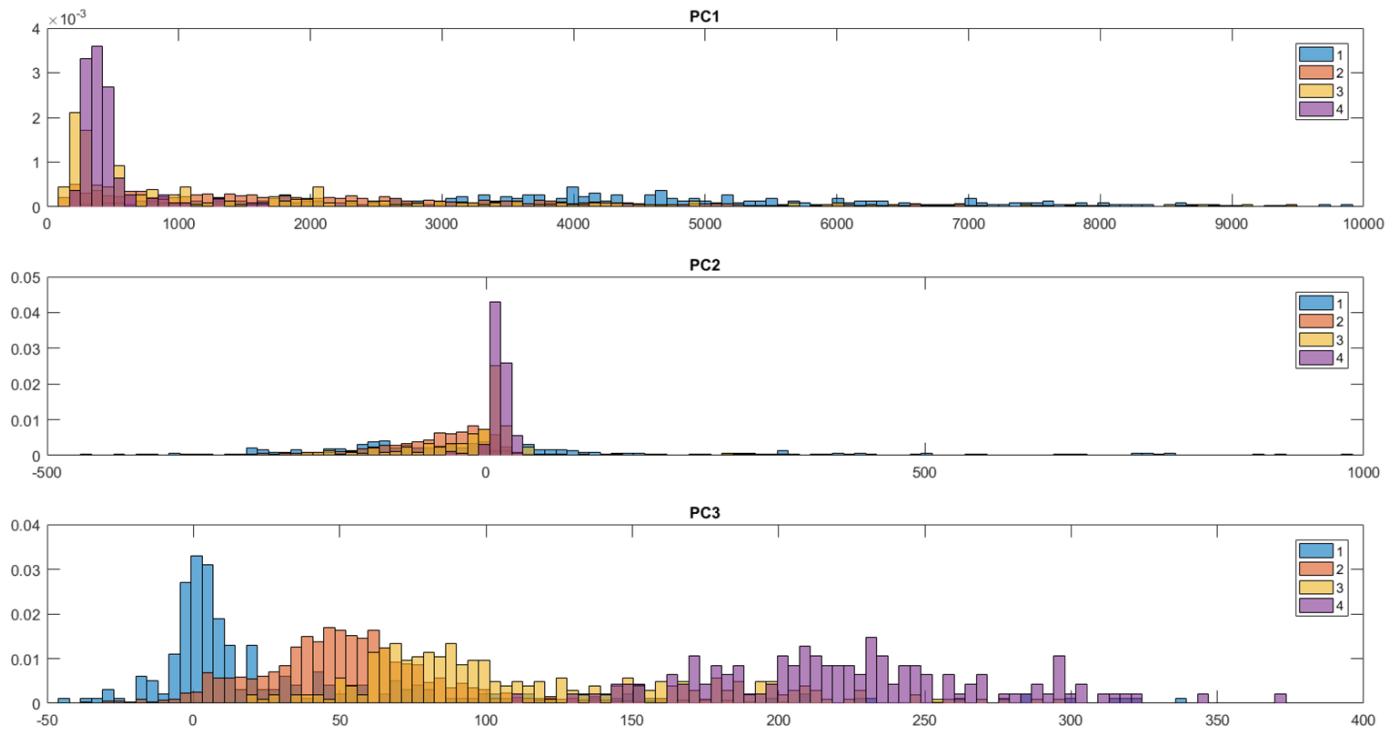
به طور خلاصه باید گفت که پاسخ به این سؤال که «آیا با ویژگی استخراج شده، و با استفاده از روش خوشه‌بندی، می‌توان عمق‌های مختلف خواب و نیز حالت REM و بیداری را از هم جدا کرد؟» با روشی که ما عمل کردیم، منفی است.

همچنین این ایده را روی ماتریس X که فقط شامل ۴ عمق خواب NREM باشد نیز تست کردیم، و نیز سعی کردیم در حالت ساده‌تر، صرفاً دیتای عمق ۱ و ۲ را از ۳ و ۴ جدا کنیم (که قبلاً مشاهده کرده بودیم به آسانی انجام می‌شود) که موفقیت‌آمیز نبود.

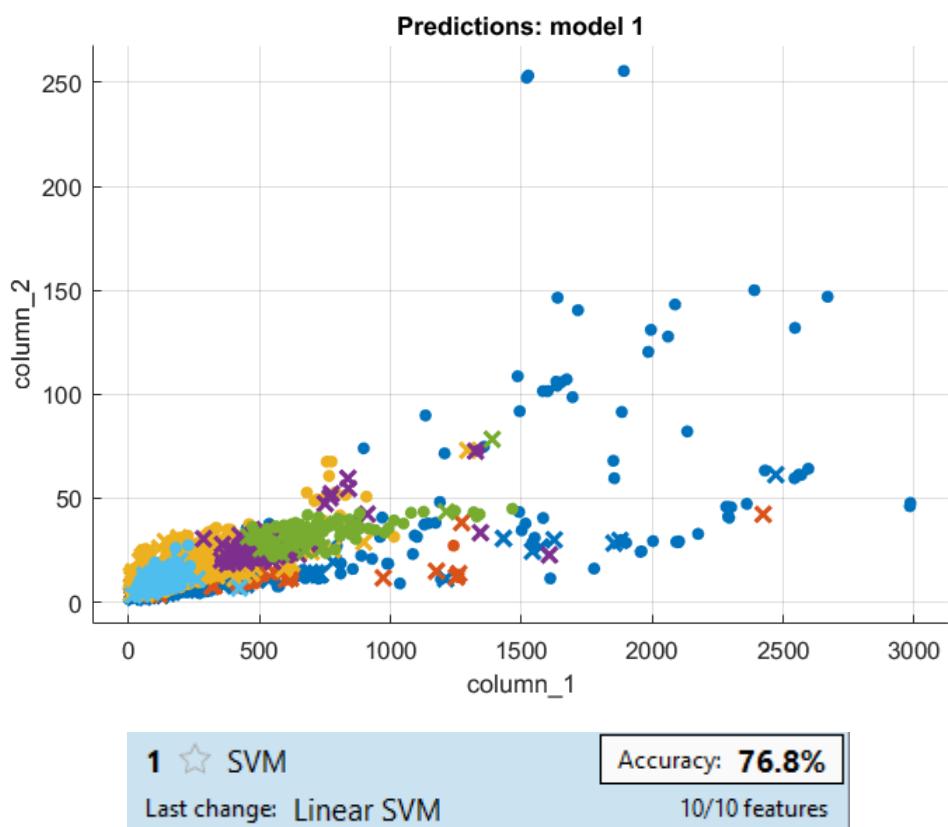
سؤال سوم) آیا حافظه‌دار در نظر گرفتن وضعیت‌های خواب اثر مثبتی بر تشخیص عمق خواب دارد؟ به عبارت دیگر، اگر عمق خواب، نه تنها تابعی از ویژگی‌های استخراج شده در همان بازه، بلکه تابعی از زمان‌های قبل نیز باشد، آیا می‌توان به نحو بهتری وضعیت‌های خواب را از هم جدا کرد؟

برای پاسخ به این سؤال، ماتریس ویژگی جدیدی را تشکیل می‌دهیم که در آن، برای هر ستون، میانگینی از  $k$  بازه ۱۰ ثانیه‌ای قبل گرفته شده و در جای آن ستون قرار داده شده است. در اینجا،  $k$  را برابر با ۵ در نظر می‌گیریم.

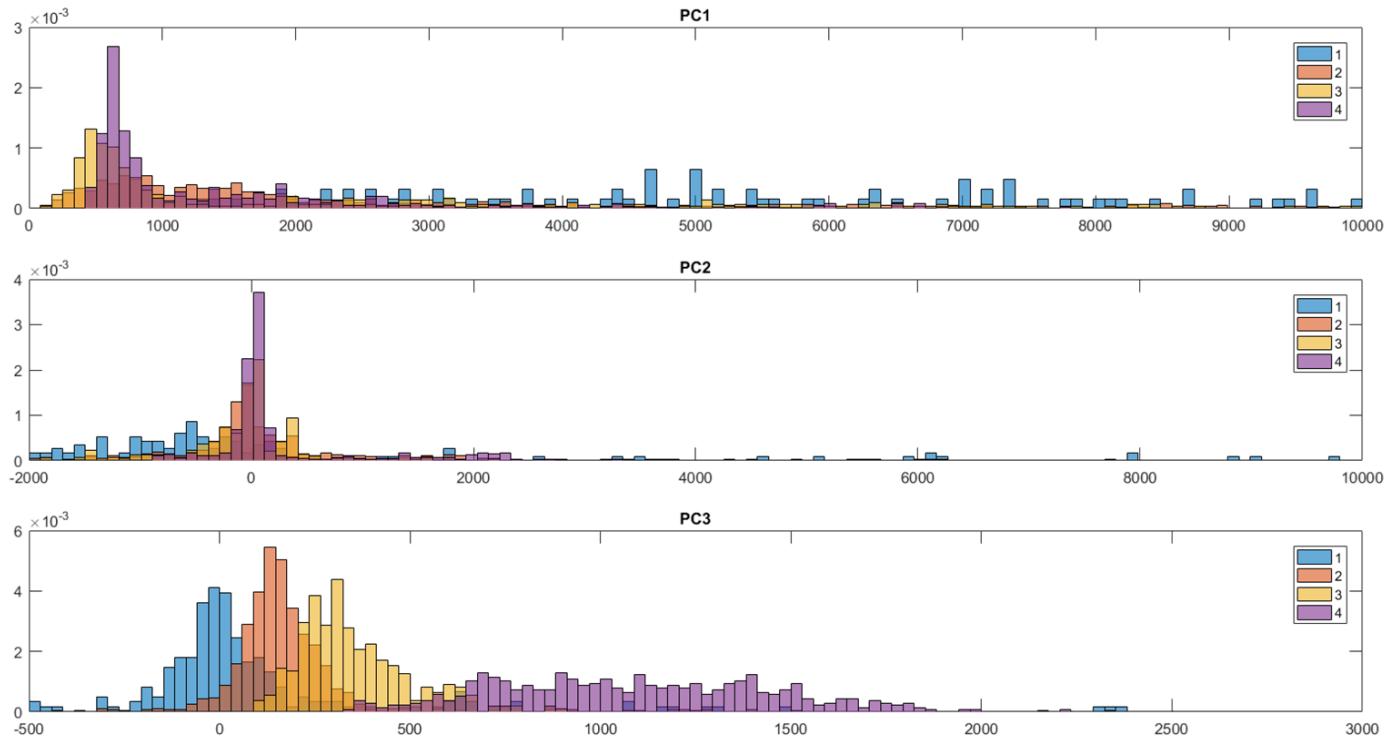
با انجام این تغییر در ماتریس  $X$ ، می‌توانیم تمامی آنچه که از ابتدای این گزارش نوشته‌ایم را مجدداً انجام دهیم و نتایج را مقایسه کنیم. اما برای جلوگیری از طولانی شدن (و البته به دلیل کمبود وقت) صرفاً دو بخش رسم هیستوگرام‌های بخش PCA که از روی آن، میزان جداکنندگی را برای عمق‌های چهارگانه‌ی خواب شناسایی می‌کردیم، و بخش SVM (که در سؤال اول قسمت پنجم مطرح کردیم) را مجدداً انجام می‌دهیم.



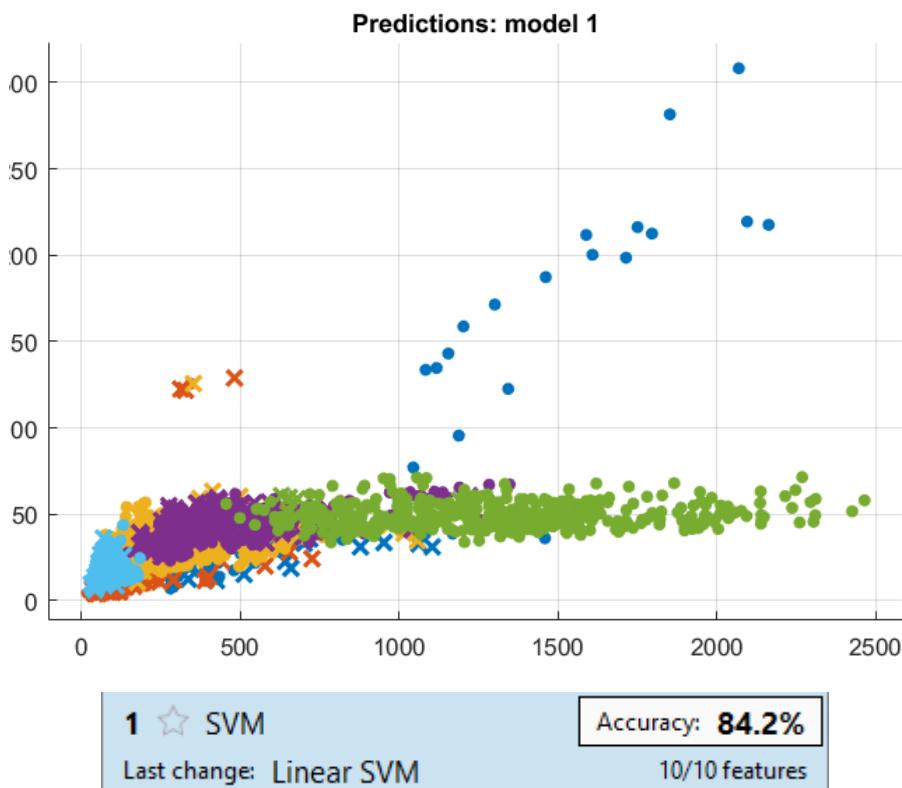
در این نمودار، مشاهده می شود که راستای جداکننده (PC3) به شکل قابل توجهی، بهتر عمل می کند و لایه های خواب را بهتر جدا می کند. (برای مقایسه می توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



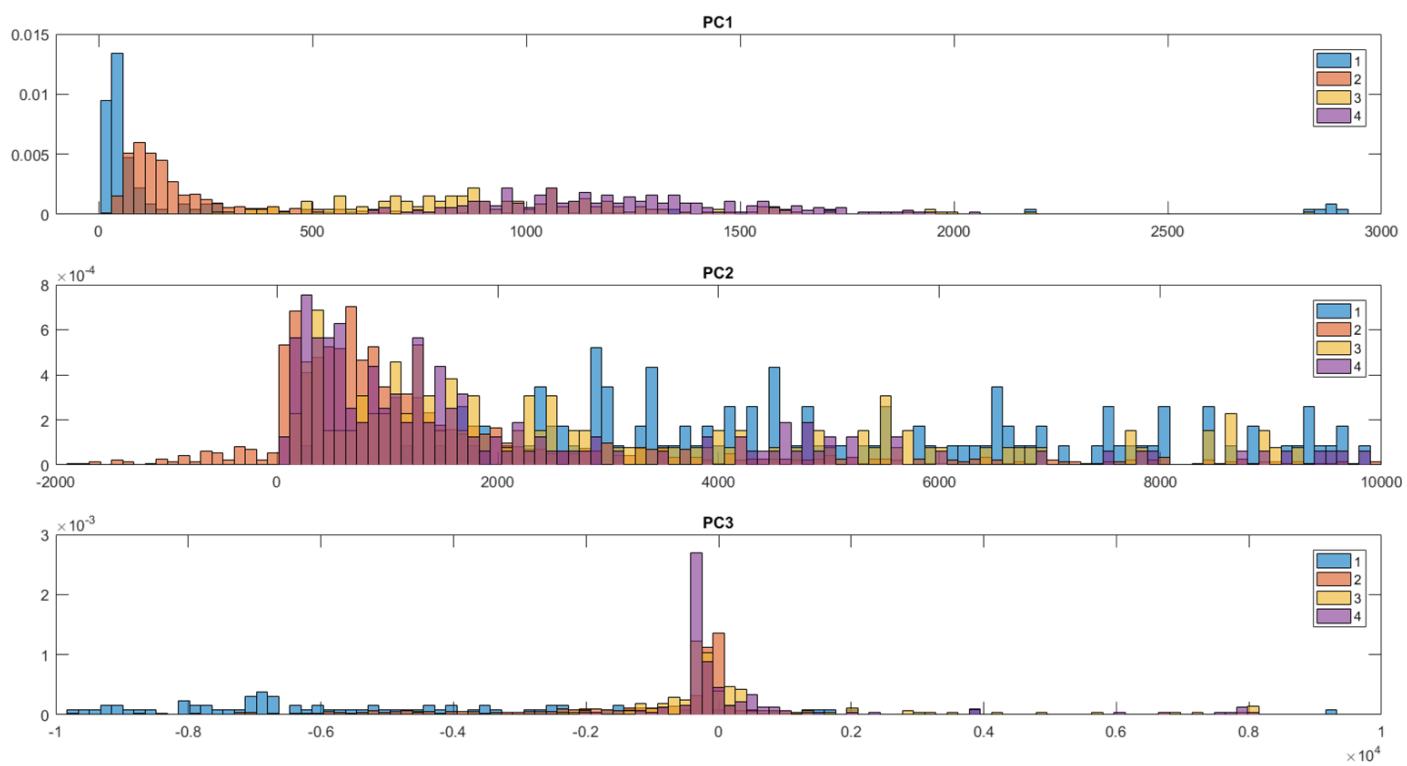
مشاهده می شود که درصد جواب صحیح Cross Validation به اندازه ۴/۸ % درصد افزایش یافته است.



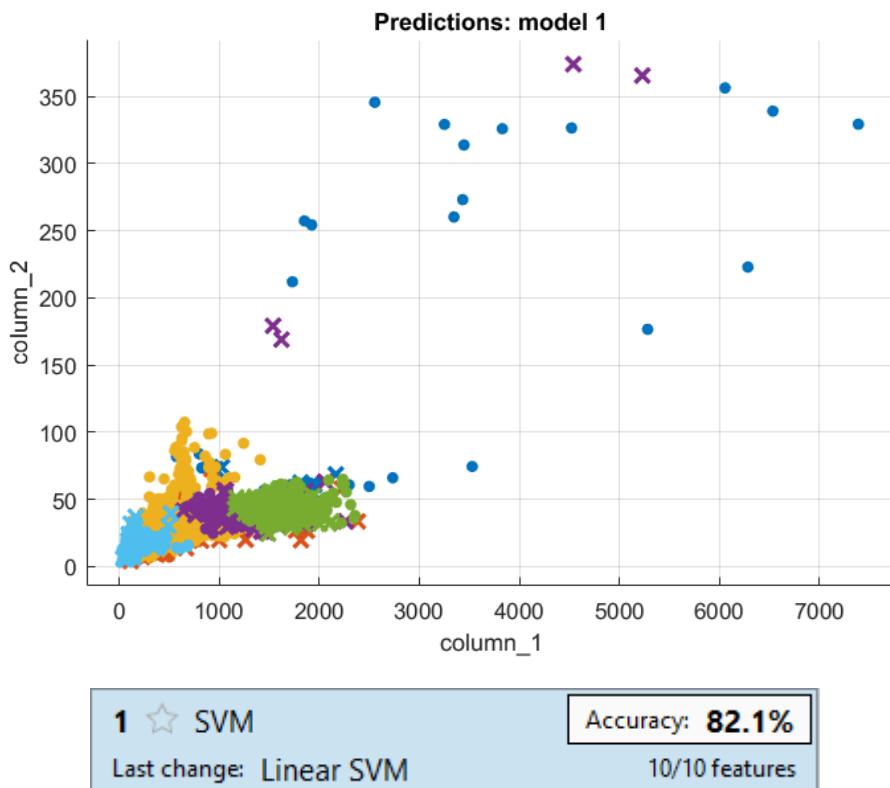
در این نمودار، مشاهده می شود که راستای جداسازننده (PC3) به شکل قابل توجهی، بهتر عمل می کند و لایه های خواب را بهتر جدا می کند. (برای مقایسه می توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



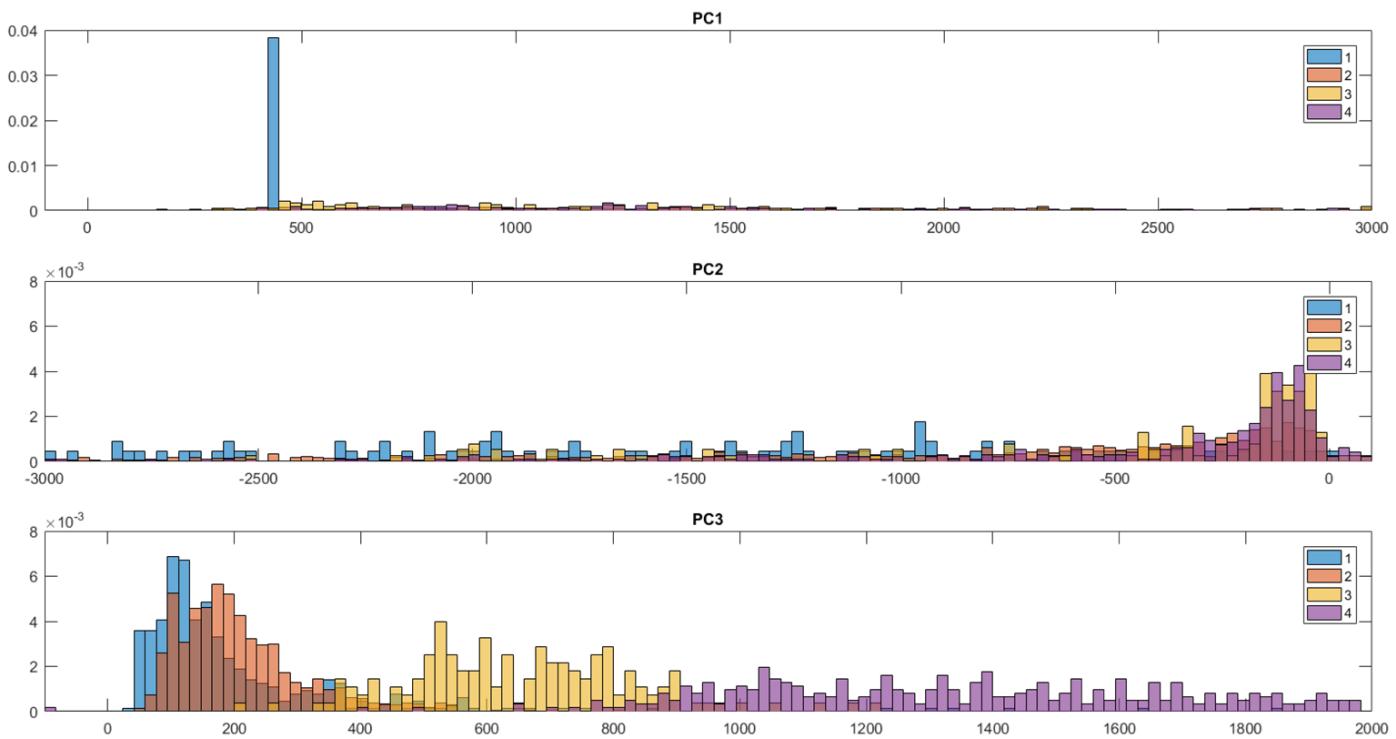
مشاهده می شود که درصد جواب صحیح Cross Validation به اندازه ۱۰/۵ % درصد افزایش یافته است.



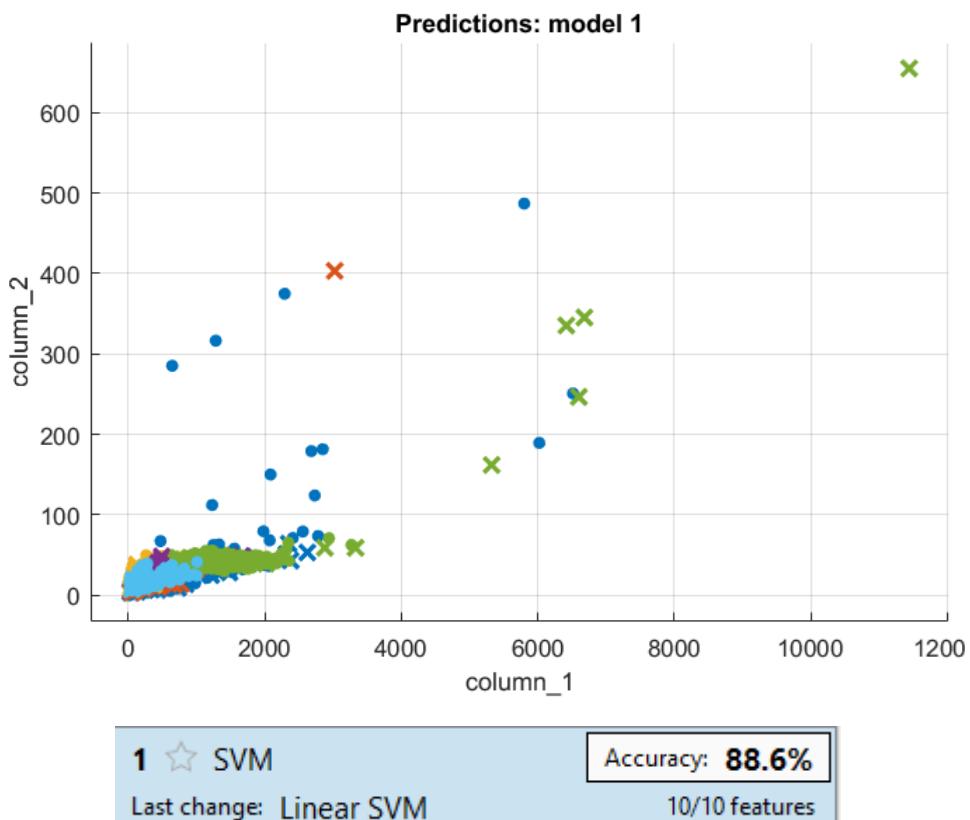
در این نمودار، مشاهده می شود که راستای جداکننده (PC1) به شکل قابل توجهی، بهتر عمل می کند و لایه های خواب را بهتر جدا می کند. (برای مقایسه می توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



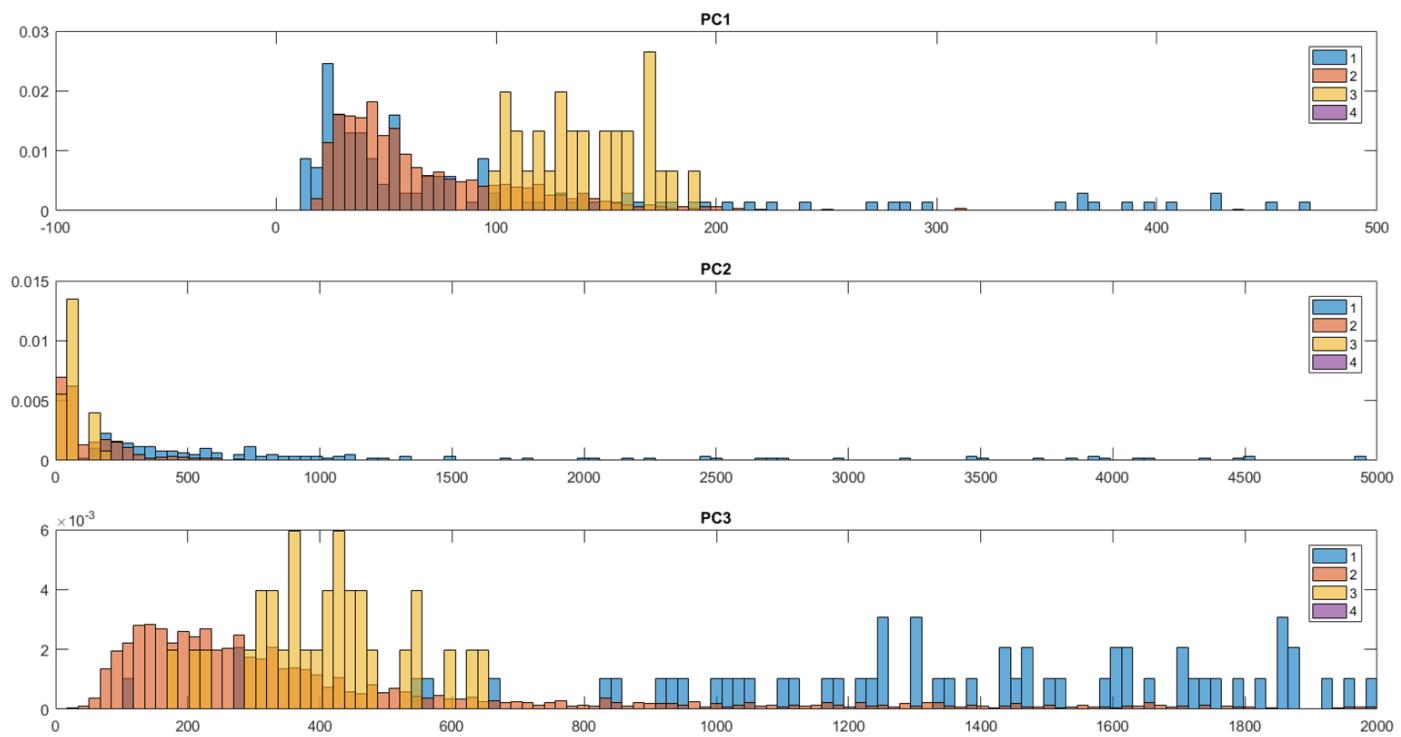
مشاهده می شود که درصد جواب صحیح Cross Validation به اندازه ۵/۸ % درصد افزایش یافته است.



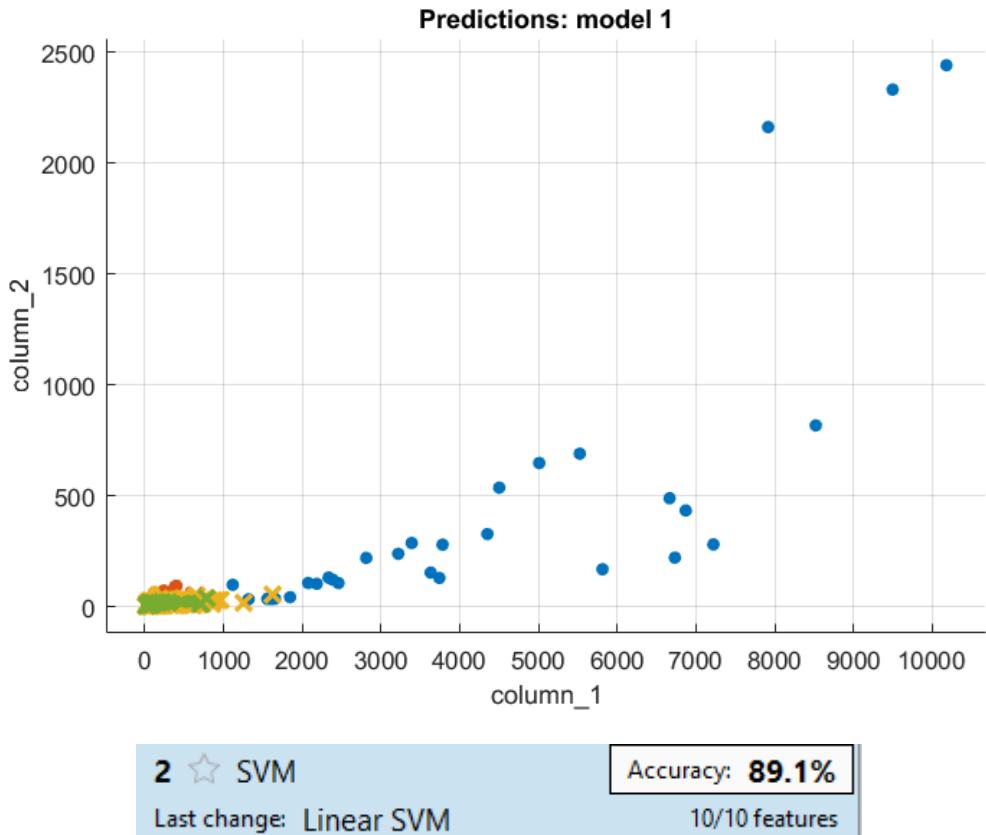
در این نمودار، مشاهده می شود که راستای جدائنتد (PC3) به شکل قابل توجهی، بهتر عمل می کند و لایه های خواب را بهتر جدا می کند. (برای مقایسه می توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



مشاهده می شود که درصد جواب صحیح Cross Validation به اندازه  $5/6$  % درصد افزایش یافته است.



در این نمودار، مشاهده می شود که راستای جدائنتد (PC1) به شکل قابل توجهی، بهتر عمل می کند و لایه های خواب را بهتر جدا می کند. (برای مقایسه می توانید به نمودار متناظر در بخش مربوط به PCA مراجعه کنید)



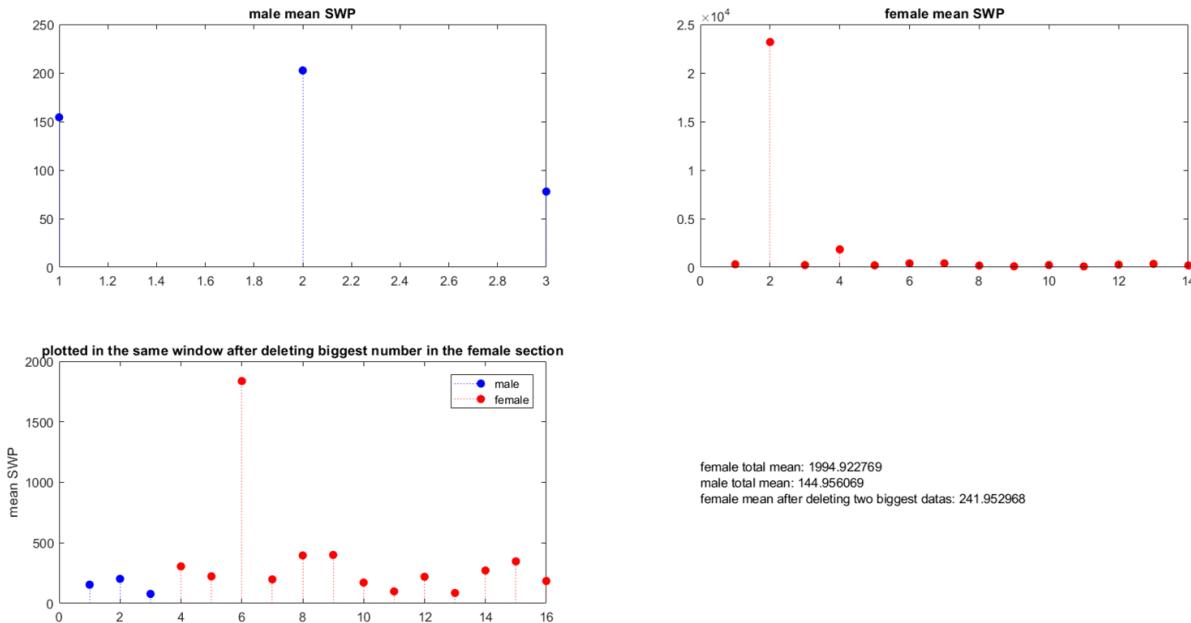
مشاهده می شود که درصد جواب صحیح Cross Validation به اندازه ۷/۷ درصد افزایش یافته است.

همان‌طور که ذکر شد، به دلیل کمبود وقت، صرفاً دو نمونه از فعالیت‌های انجام شده در طول پژوهه را با فرض حافظه‌دار بودن وضعیت خواب تکرار کردیم، که در تمامی موارد شاهد افزایش چشمگیری در پاسخگویی (هم در جداسازی لایه‌های خواب توسط PCA، و هم در درصد پاسخگویی Cross Validation) شاهد بودیم که حتی یک نمونه هم پسرفت مشاهده نشد.

بنابراین تقریباً با قطعیت نتیجه می‌گیریم که وضعیت خواب در هر بازه (۱۰ ثانیه‌ای)، تنها به پارامترهای سیگнал EEG در آن ۱۰ ثانیه وابسته نیست، بلکه عملکردی حافظه‌دار دارد و باید چند بازه زمانی قبل‌تر را نیز در نظر گرفت، تا بتوان توصیف دقیق‌تر و بهتری از عمق خواب ارائه داد.

## سوال چهارم) مقایسه تاثیر جنسیت در اندازه SWP

برای بررسی تاثیر جنسیت بر روی SWP ابتدا نیاز به جداسازی این داده ها از الکترود های موجود بود. SWP بنا بر تعریف شامل باند فرکانسی ۰/۵ تا ۲ هرتز می باشد. با اعمال این فیلتر بر روی ۲ الکترود موجود برای هر فرد و میانگین گیری، در نهایت عددی را برای مقایسه بدست آوردیم. از آن جا که در داده های اولیه که در گزارش کار ذکر شده بودند تنها ۵ فرد موجود بودند و در بین این ۵ فرد تنها یک مرد وجود داشت، داده های اضافی را از سایت دانلود کرده ایم. لیست تمامی این افراد در آخر آمده است. البته مشکل تنوع بعد از این افزایش داده نیز بهبود نیافت و در بین ۱۷ نفر، ۳ مرد و بقیه زن بودند.



اگر بخواهیم بر همین تعداد داده‌ی کم تکیه کنیم و نتیجه گیری بکنیم، میتوان گفت که مقدار SWP برای زن‌ها بیشتر از مرد‌ها می‌باشد، چرا که با وجود کم کردن داده‌های بزرگ از سری داده‌های زن‌ها، باز هم دارای میانگینی بیشتر از مرد‌ها می‌باشند.

۲ نکته‌ی دیگر از نمودار قابل استخراج است، یکی اینکه ادعا‌ی مقاله بر اینکه SWP تغییرات زیادی دارد و از عوامل مختلفی همچون جنسیت و آناتومی سر تاثیر می‌پذیرد، صحت دارد. اما نکته‌ی دیگر اینکه انتظار اینکه برای تمامی افراد، SWP زن‌ها بیشتر از مرد‌ها باشد، غلط می‌باشد و به نظر می‌رسد همواره برقرار نیست. البته در بررسی بالا اثر سن افراد دخیل داده نشده است و برای بررسی دقیق‌تر، علاوه بر تنوع داده‌ی بیشتر، نیاز داریم سن افراد را نیز وارد معادلات بکنیم.

لازم به ذکر است که تمامی این داده‌ها از افراد در هنگام خواب معمولی و بدون داروی خواب گرفته شده است و نگرانی از اثر داروی خواب بر SWP وارد نیست.

لیست تمامی افراد استفاده شده در این سری داده‌ها: (پنج نفر اول همان پنج نفر ذکر شده در دستور کار می‌باشند.)

ST7011J0

ST7022J0

ST7041J0

ST7052J0

ST7061J0

ST7082J0

ST7101J0

ST7112J0

ST7151J0

ST7162J0

ST7171J0

ST7182J0

ST7192J0

ST7201J0

ST7212J0

ST7221J0

ST7241J0