

به نام خدا



دانشگاه تهران

پردیس دانشکده‌های فنی

دانشکده برق و کامپیوتر



درس مبانی علوم شناختی

تمرین شماره ۵

نام و نام خانوادگی : مهیار ملکی

شماره دانشجویی : ۸۱۰۱۰۰۴۷۶

۱۴۰۱ ماه تیر

فهرست

سوال اول – پیش‌پردازش اسپایک‌ها: مرتب‌سازی.	۴
۱ - شروع.	۱/۱
۲ - فیلتر کردن داده‌ها.	۲/۱
۳ - شناسایی اسپایک‌ها	۳/۱
۴ - استخراج ویژگی‌ها	۴/۱
۵ - خوشبینی اسپایک‌ها	۵/۱
پرسش n.	۸
پرسش o.	۸
پرسش p.	۹
سوال دوم – کاوش داده‌ها: سیگنال‌های زمانی گستته	۱۰
۱۰ Raster plot	۱۰
۱۱ Peri-Stimulus Time Histogram (PSTH)	۱۱
۱۲ Firing Rate	۱۲
سوال سوم – تک نورون‌ها: اندازه‌گیری اطلاعات.	۱۳
۱۳ ۱	۱۳
۱۳ ۲	۱۳
۱۳ ۳	۱۳
۱۳ ۴	۱۳
۱۳ ۵	۱۳
سوال چهارم – جمعیت نورون‌ها: اندازه‌گیری اطلاعات.	۱۴

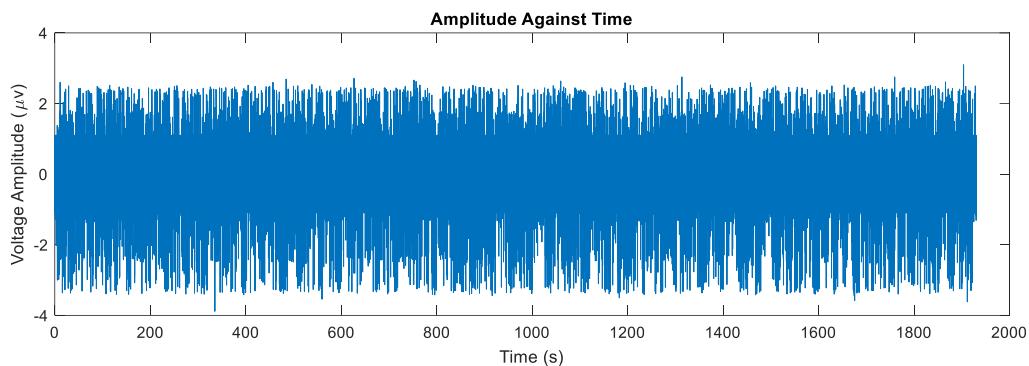
فهرست اشکال

..... ۴ شکل ۱- نمودار دامنه ولتاژ سیگنال زمانی بر اساس زمان
..... ۴ شکل ۲- هیستوگرام ولتاژ سیگنال ضبط شده
..... ۵ شکل ۳- مقایسه سیگنال نویزی و فیلترشده (پنجره ۲ ثانیه ای)
..... ۶ شکل ۴- شکل موج اسپایک‌های شناسایی شده
..... ۷ شکل ۵- نتایج خوش‌بندی kmeans با فضای کاهش بعد یافته PCA
..... ۸ شکل ۶- ماتریس درهم‌یختگی اسپایک‌های شناسایی شده
..... ۸ شکل ۷- شکل موج اسپایک‌های بدست آمده از حد آستانه جدید
..... ۹ شکل ۸- نتایج خوش‌بندی tsne با فضای کاهش بعد یافته
..... ۱۰ شکل ۹- raster plot نورون پنجم
..... ۱۰ شکل ۱۰- raster plot نورون هشتم
..... ۱۱ شکل ۱۱- PSTH نورون پنجم برای زوایای مختلف
..... ۱۱ شکل ۱۲- PSTH نورون هشتم برای زوایای مختلف
..... ۱۲ شکل ۱۳- نرخ شلیک نورون پنجم به ازای هر زاویه
..... ۱۲ شکل ۱۴- نرخ شلیک نورون پنجم در زاویه های ۱۳۵ و منفی ۴۵
..... ۱۴ شکل ۱۵- معیار پوشش طبقه بندها برای زوایای مختلف در پنجره های زمانی
..... ۱۴ شکل ۱۶- معیار پوشش طبقه بندها برای شعاع‌های مختلف در پنجره های زمانی

سوال اول – پیش‌پردازش اسپایک‌ها: مرتب‌سازی

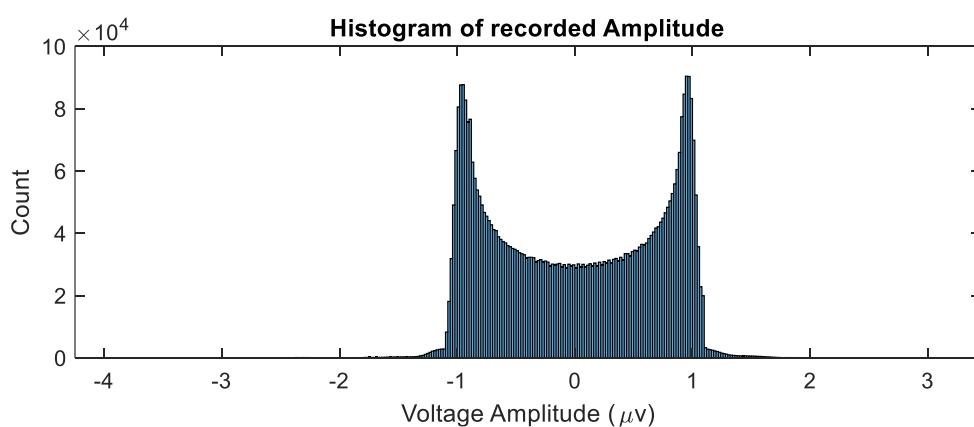
۱.۱ - شروع

در این قسمت ابتدا دادگان خام اولیه را بارگزاری کرده و نمودار آن را بر اساس زمان ترسیم می‌کنیم. نمودار حاصل در شکل ۱ قابل مشاهده است. لازم به ذکر است محور زمان با تقسیم تعداد دادگان بر نرخ نمونه‌برداری^۱ بدست آمده است.



شکل ۱- نمودار دامنه ولتاژ سیگنال زمانی بر اساس زمان

در ادامه هیستوگرام^۲ ولتاژها را نیز رسم می‌کنیم. چنانچه در شکل ۲ قابل مشاهده است، به نظر می‌رسد در بازه‌های کوچک‌تر از منفی ۱ و بزرگ‌تر از مثبت ۱ تعداد کمی داده وجود دارد که می‌توانند نشان‌دهنده نویز باشند. همچنین نمودار حاصل یک توزیع U شکل را حاصل شده است، در صورتی که یک سیگنال بدون نویز توزیعی گوسی خواهد داشت. یعنی تراکم بیشتر تعداد ولتاژها در مقادیر مثبت و منفی ۱ میکروولت نشان‌دهنده نویز در این نقاط است.



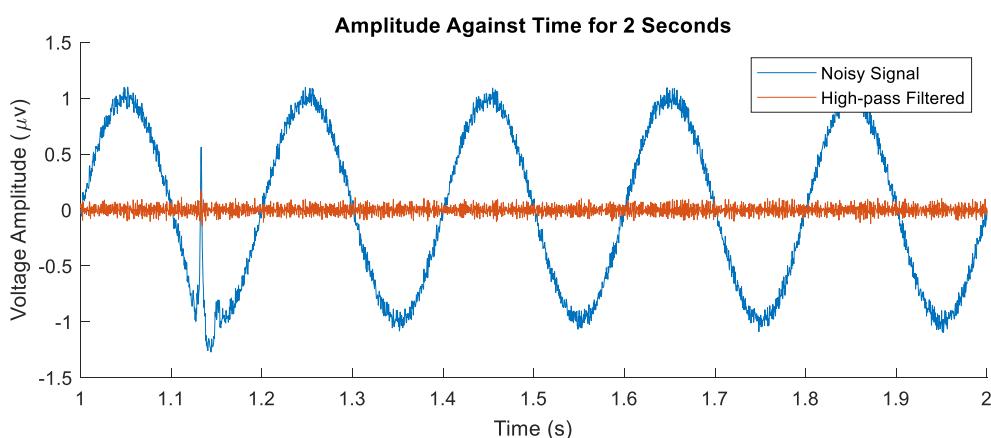
شکل ۲- هیستوگرام ولتاژ سیگنال ضبط شده

Sampling rate^۱
Histogram^۲

۲.۱ – فیلتر کردن داده‌ها

در این بخش دادگان را از یک فیلتر بالا گذر از نوع Butterworth و با فرکانس قطع ۳۰۰ هرتز عبور می‌دهیم تا فرکانس‌های زیر این مقدار فیلتر شوند. طراحی این فیلتر با دستور butter و پیاده‌سازی آن با دستور `filtfilt` انجام شده است. لازم به ذکر است که ورودی تابع butter فرکانس نایکوئیست^۱ می‌باشد و با تقسیم فرکانس قطع بر نصف نرخ نمونه‌برداری، این مقدار محاسبه شده است.

در نهایت سیگنال فیلتر شده و سیگنال اولیه در شکل ۳ قابل مشاهده‌اند. برای مشاهده بهتر روند سیگنال این مقایسه در یک پنجره به اندازه ۲ ثانیه انجام شده است.



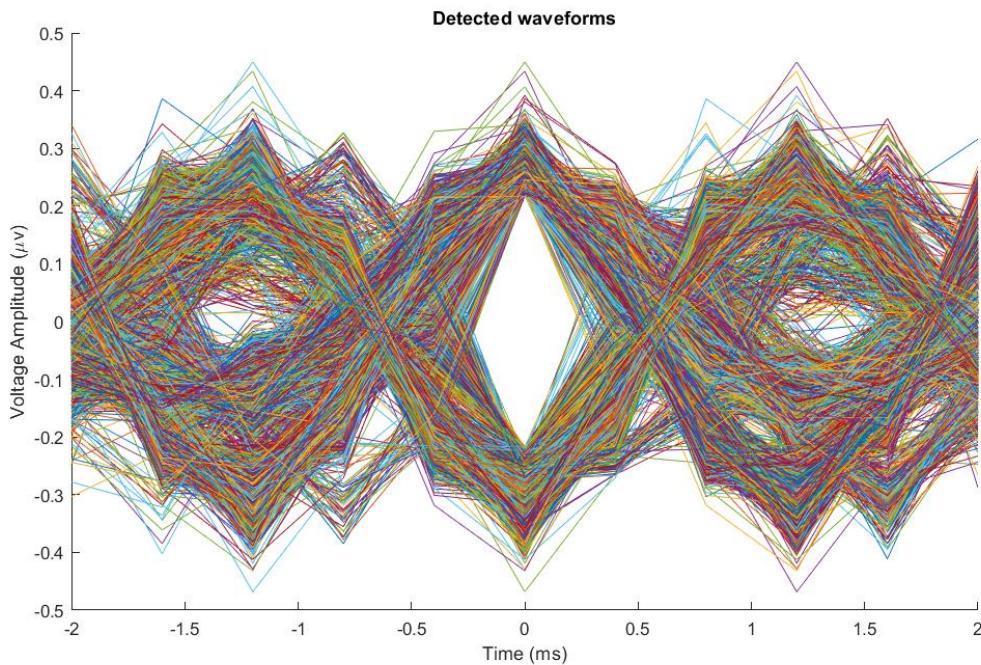
شکل ۳- مقایسه سیگنال نویزی و فیلتر شده (پنجره ۲ ثانیه‌ای)

۳.۱ – شناسایی اسپایک‌ها

در این قسمت هدف شناسایی پیک‌های سیگنال و یافتن اسپایک‌ها از میان آن‌ها می‌باشد. به این منظور ابندا با دستور `findpeaks` نقاط پیک را پیدا می‌کنیم، سپس با استفاده از حد آستانه^۲ معروفی شده در صورت سوال و محاسبه آن، پیک‌هایی که مقدار بیشتری نسبت به حد آستانه دارند را به عنوان اسپایک شناسایی می‌کنیم. سپس در قدم بعد شکل موج^۳ این اسپایک‌ها را در یه پنجره زمانی ۲ میلی‌ثانیه‌ای قبل و بعد هر اسپایک بدست می‌آوریم. این شکل موج‌ها دادگان ما برای خوش بندی را تشکیل می‌دهند.

Nyquist ^۱
Peak ^۲
Spike ^۳
Threshold ^۴
Waveform ^۵

در شکل ۴ تمام شکل موج‌ها در یک نمودار قابل مشاهده است. چنانچه مشاهده می‌شود برخی از اسپایک‌ها مثبت و به صورت قله و برخی به صورت دره و منفی هستند. این امر می‌توان نشان‌دهنده اختلاف جهت جریان این اسپایک‌ها باشد.



شکل ۴- شکل موج اسپایک‌های شناسایی شده

۴.۱- استخراج ویژگی‌ها

در بخش چهارم با اعمال الگوریتم PCA روی شکل موج‌ها، به کاهش ابعاد فضای ویژگی^۱ پرداختیم.

الگوریتم PCA به وسیله کاهش ابعاد، با مسئله نحسی ابعاد^۲ مقابله می‌کند، همچنین به تصویرسازی دوبعدی داده‌ها نیز کمک خواهد کرد. لازم به ذکر است که ویژگی‌های جدید استخراج شده لزوماً دارای معنی و مفهوم نمی‌باشند.

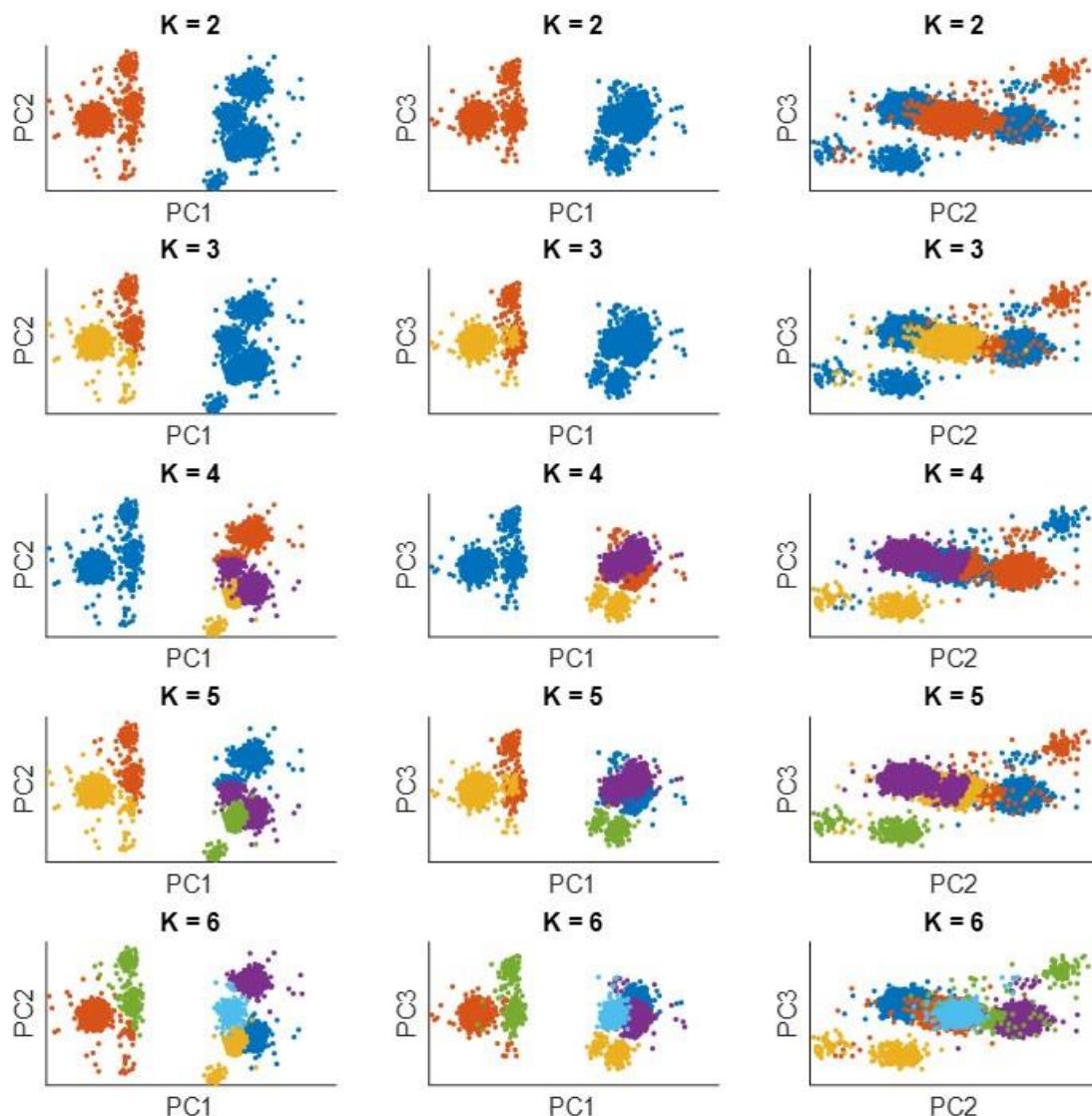
دادگان اصلی دارای ۱۱ ویژگی به ازای هر میلی ثانیه بوده که نشان دهنده دامنه ولتاژ سیگنال‌ها است. در این قسمت با اعمال PCA سه عدد از ویژگی‌هایی که بیشترین واریانس داده‌ها را پوشش می‌دهند انتخاب می‌کنیم.

Feature reduction^۱
Curse of dimensionality^۲

۵.۱ - خوشبندی اسپایک ها

در قسمت پنجم با استفاده از الگوریتم kmeans عملیات خوشبندی^۱ را انجام می‌دهیم. برای این کار از فضای با ابعاد کاهش یافته استفاده خواهیم کرد. در ادامه الگوریتم را تا تعداد ۶ خوشبندی تکرار کرده و هر بار نمودار مربوط به هر جفت از ویژگی‌ها را رسم می‌کنیم. نتایج حاصله در شکل ۵ قابل مشاهده است. با بررسی نتایج می‌توان نتیجه گرفت که با انتخاب ۵ خوشبندی، داده‌ها تقریباً به خوبی خوشبندی شده و ۵ نورون متمایز شناسایی می‌شوند.

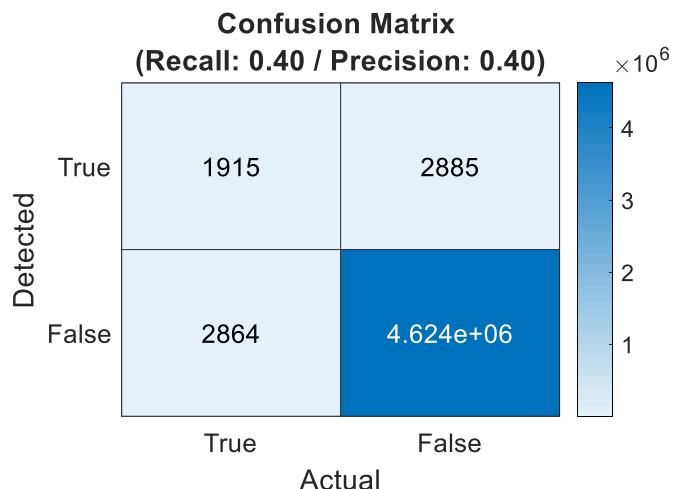
Kmeans Clustering with PCA Feature Extraction



شکل ۵ - نتایج خوشبندی kmeans با فضای کاهش یافته PCA

N پرسش

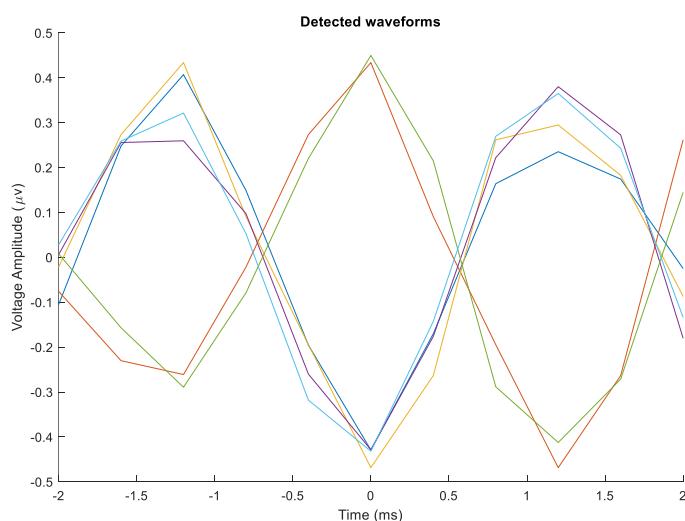
در اینجا به مقایسه اسپایک‌های اصلی و اسپایک‌های شناسایی شده، می‌پردازیم. با محاسبه مواردی چون خطای نوع یک و نوع دو و همچنین مقادیر صحت و پوشش این کار را انجام می‌دهیم. موارد گفته شده در ماتریس درهم‌ریختگی شکل ۶ قابل مشاهده است.



شکل ۶- ماتریس درهم‌ریختگی اسپایک‌های شناسایی شده

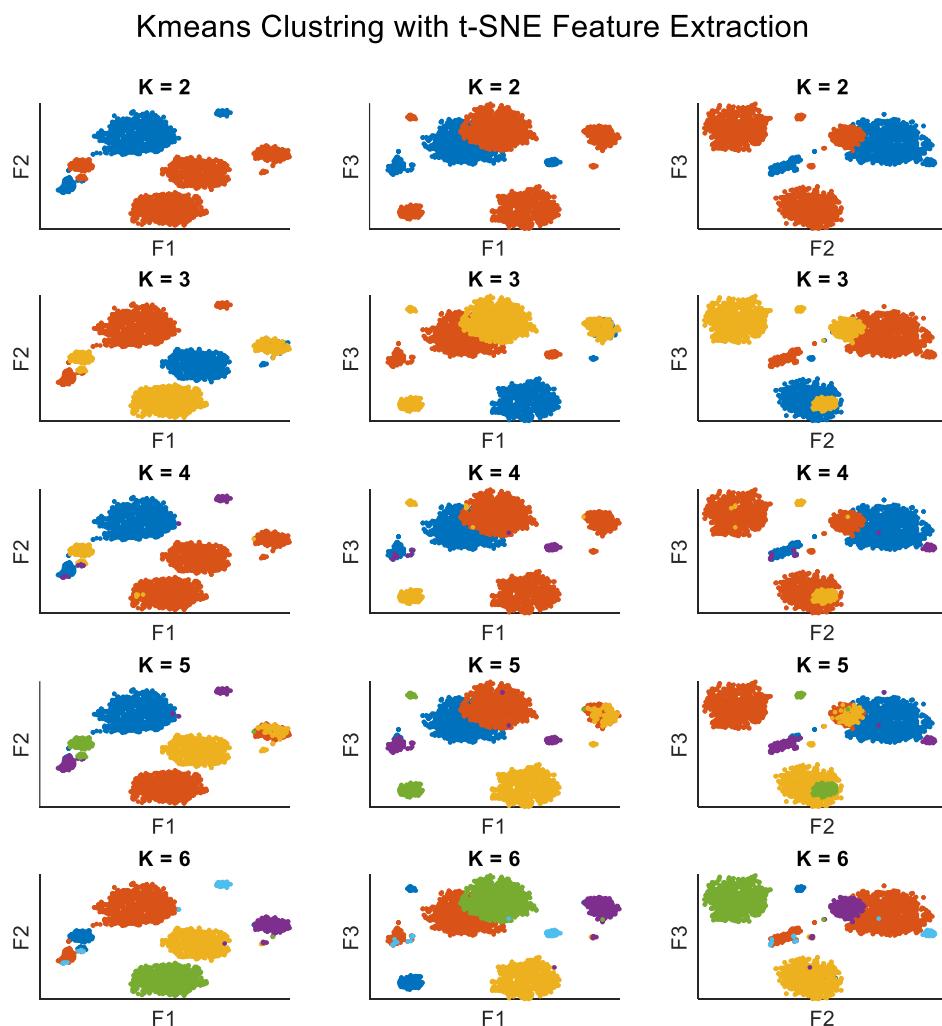
O پرسش

در این قسمت با استفاده از حد آستانه معرفی شده جدید، عملیات شناسایی اسپایک‌ها را تکرار می‌کنیم. در نهایت چنانچه در شکل ۷ پیداست تنها ۶ اسپایک شناسایی می‌شود. این تعداد بسیار کم بوده و برای عملیات خوش‌بندی مناسب نمی‌باشد. لذا می‌توان نتیجه گرفت که انتخاب حد آستانه مناسب امری بسیار مهم می‌باشد.



شکل ۷- شکل موج اسپایک‌های بدست آمده از حد آستانه جدید

با تکرار مراحل قبلی برای خوشبندی اما این بار با استفاده از الگوریتم t-SNE برای کاهش ویژگی‌ها به نتایج جدیدی خواهیم رسید. نمودارهای مربوط به خوشبندی‌ها تا ۶ عدد، در شکل ۸ قابل مشاهده است. با توجه در شکل به نظر می‌رسد این بار با تعداد ۶ خوشبندی بهترین عملکرد را داشته‌ایم، و ۶ نورون شناسایی شده است. همچنین ممکن است کمی افزایش تعداد خوشبندی‌ها همچنان بهبود عملکرد را به دنبال داشته باشد.

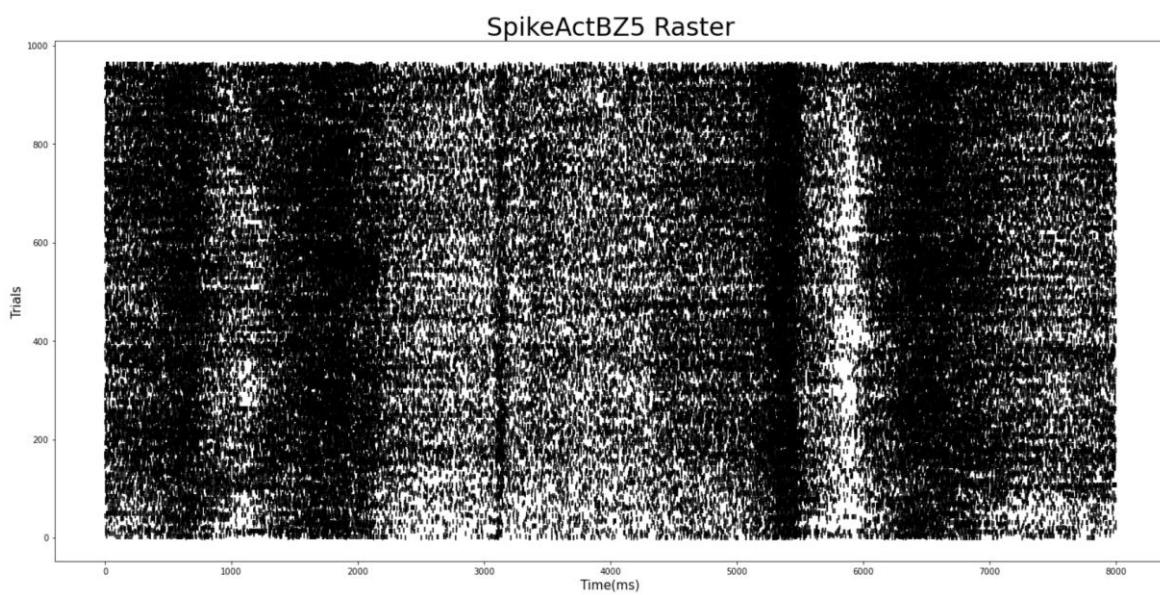


شکل ۸- نتایج خوشبندی kmeans با فضای کاهش بعد یافته tsne

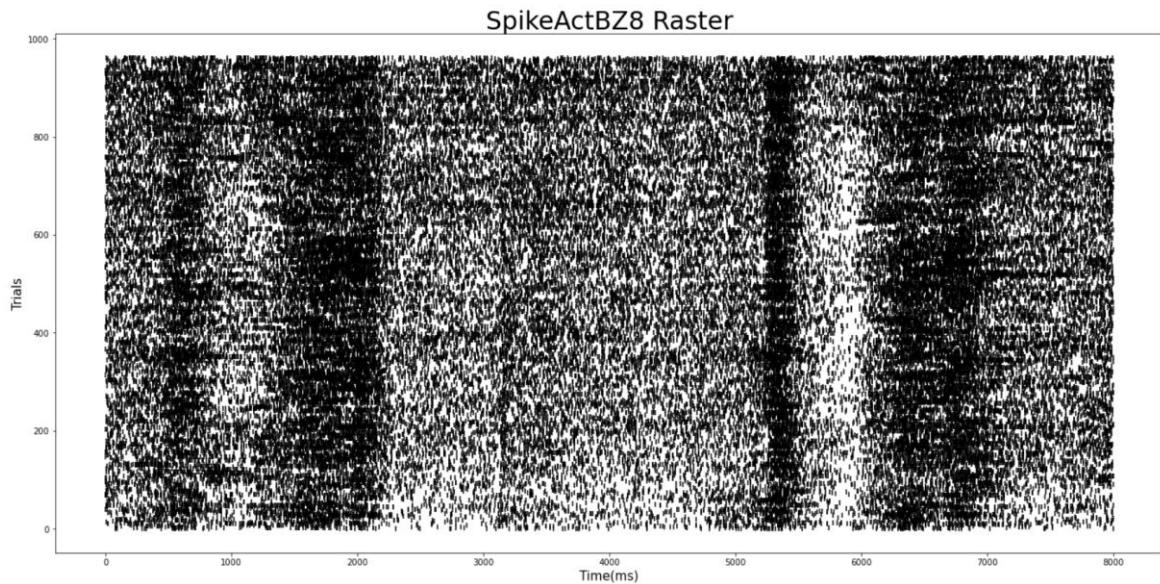
سوال دوم – کاوش داده‌ها: سیگنال‌های زمانی گسسته

RASTER PLOT

در این قسمت به ازای هر کانال اسپایک‌های شناسایی شده در واحد زمان را رسم می‌کنیم. دو عدد از بهترین نمودارهای به دست آمده، در شکل‌های ۹ و ۱۰ قابل مشاهده است. چنانچه مشخص است این دو نمودار مربوط به نورون‌های ۵ و ۸ می‌باشند. با توجه به task مسئله (نمایش محرک در زمان ۳۰۰۰ و به خاطرسپاری آن در زمان ۵۵۰۰)، به نظر می‌رسد این دو نورون به این وقایع به خوبی پاسخ داده‌اند زیرا تعداد اسپایک‌ها در این زمان‌ها نسبت به زمان‌های دیگر افزایش یافته‌است.



شکل ۹ نورون پنجم raster plot

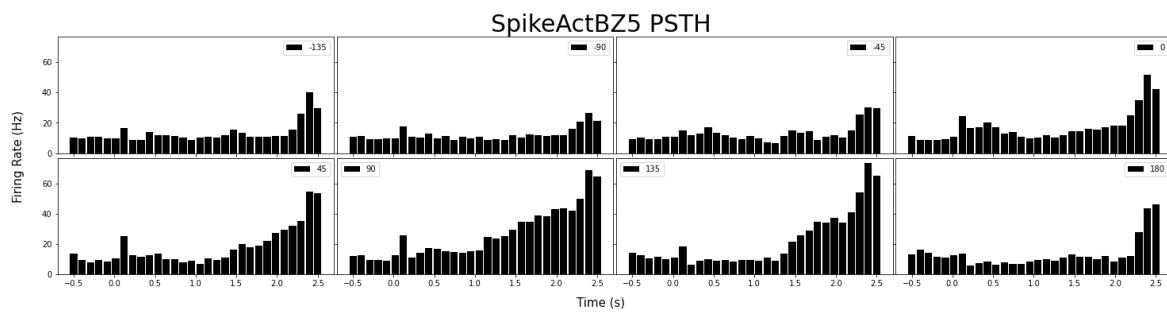


شکل ۱۰ نورون هشتم raster plot

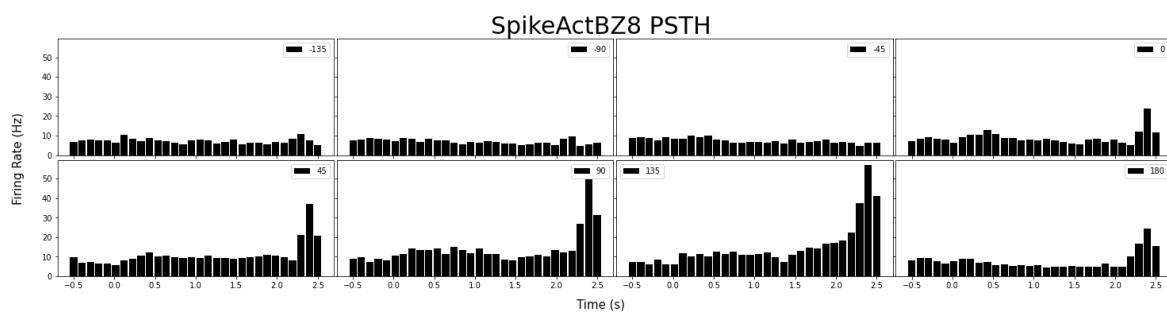
PERI-STIMULUS TIME HISTOGRAM (PSTH)

از این نمودار برای توصیف پاسخ نورون استفاده می‌شود. در این نمودار در واقع میانگین پاسخ نورون‌ها به ازای پنجره‌های زمانی تعریف شده، محاسبه می‌شود، همچنین با ضرب نتیجه حاصل در نرخ نمونه‌برداری، فرکانس پاسخ نورون بدست خواهد آمد. در اینجا این کار را به ازای زاویه‌های مختلف برای هر نورون انجام می‌دهیم. در شکل‌های ۱۱ و ۱۲ نمودارهای PSTH مربوط به نورون‌های بخش قبلی که بیشترین پاسخ را داشتند، به ازای زوایای مختلف قابل مشاهده است.

چنانچه در نتایج مشخص است، به عنوان مثال برای نورون پنجم در زاویه‌های ۰ و ۴۵ و ۹۰ پاسخ خوبی (به ازای زمان‌های مربوط به نمایش محرک و یادآوری آن) دیده می‌شود یا در زاویه‌های منفی ۱۳۵ و منفی ۹۰ و منفی ۴۵ اسپایک‌های کمتری رخ داده و نورون مربوطه به محرک‌ها پاسخ کمتری می‌دهد.



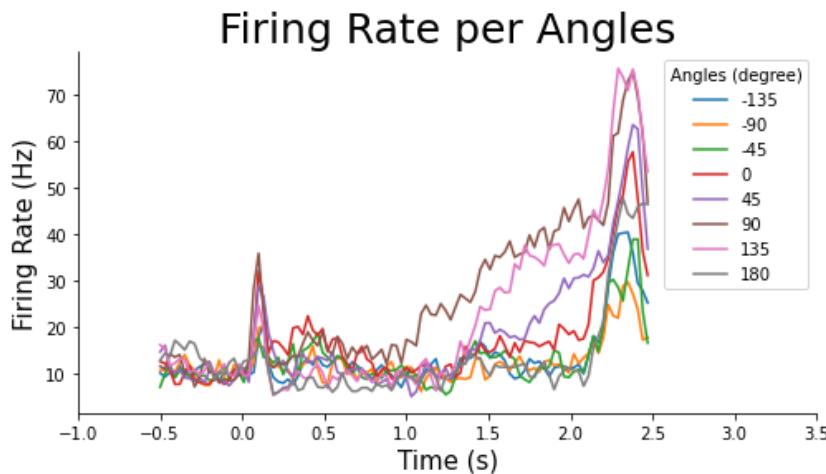
شکل ۱۱ - نورون پنجم برای زوایای مختلف



شکل ۱۲ - نورون هشتم برای زوایای مختلف

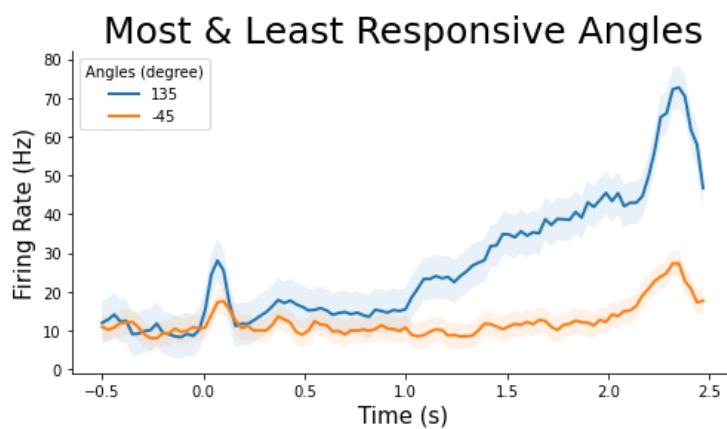
FIRING RATE

در این بخش بهترین نورون از نظر پاسخدهی به محرک‌ها را انتخاب کرده (نورون پنجم) و به ازای زوایای مختلف، نرخ شلیک این نورون را در زمان رسم می‌کنیم. نتیجه در شکل ۱۳ قابل مشاهده است.



شکل ۱۳- نرخ شلیک نورون پنجم به ازای هر زاویه

در ادامه دو زاویه که دارای کمترین و بیشترین پاسخدهی هستند را انتخاب کرده و به بررسی تفاوت پاسخ در آن‌ها می‌پردازیم. نمودار نرخ شلیک این دو زاویه به همراه بازه‌های اطمینان در شکل ۱۴ قابل مشاهده است.



شکل ۱۴- نرخ شلیک نورون پنجم در زاویه‌های ۱۳۵ و منفی ۴۵

چون این دو نمودار، میانگین آزمایش‌ها بوده و به هم نیز مرتبط هستند، در اینجا برای بررسی اینکه تفاوت قابل توجهی با هم دارند یا خیر، از آزمون آماری paired t-test استفاده می‌کنیم. در نهایت، این آزمون مقدار p-value برابر با $3e-18$ را نتیجه می‌دهد که از 0.05 کمتر بوده و نتیجه می‌گیریم اختلاف این دو نمودار معنادار و قابل توجه است.

سوال سوم – تک نورون‌ها: اندازه‌گیری اطلاعات

۱

برای انتقال اطلاعات مربوط به تعدادی از محرک‌ها، پاسخ نورون‌ها به ازای هر محرک باید متفاوت باشد. آنتروپی یک معیار اندازه‌گیری تفاوت پاسخ‌های است، ولی اطلاعاتی راجب منشاء این تفاوت به دست نمی‌دهد. یک نورون فقط زمانی می‌تواند اطلاعاتی از محرک بدهد که تفاوت پاسخ‌های آن با تغییرات محرک مرتبط باشد. یک راه تشخیص ارتباط بین تفاوت محرک‌ها و تفاوت پاسخ نورون، مقایسه پاسخ‌های به دست آمده از محرک‌های مختلف در هر آزمایش با پاسخ‌های اندازه‌گیری شده در آزمایش‌هایی با تکرار نمایش همان محرک، می‌باشد. پاسخ‌هایی که راجب هویت محرک، دارای اطلاعات هستند، باید تفاوت بیشتری برای محرک‌های مختلف نتیجه بدهند. اطلاعات متقابل^۱، تفاوت بین آنتروپی پاسخ کلی و آنتروپی پاسخ میانگین را برای آزمایش‌هایی که شامل نمایش تکراری یک محرک هستند، نتیجه می‌دهد.

۲

محاسبه اطلاعات متقابل می‌تواند دشوار باشد، زیرا باید فرکانس رخداد تعداد زیادی از اسپایک‌ها مختلف را مشخص شود، و این امر حجم زیادی از داده‌ها را نیاز دارد.

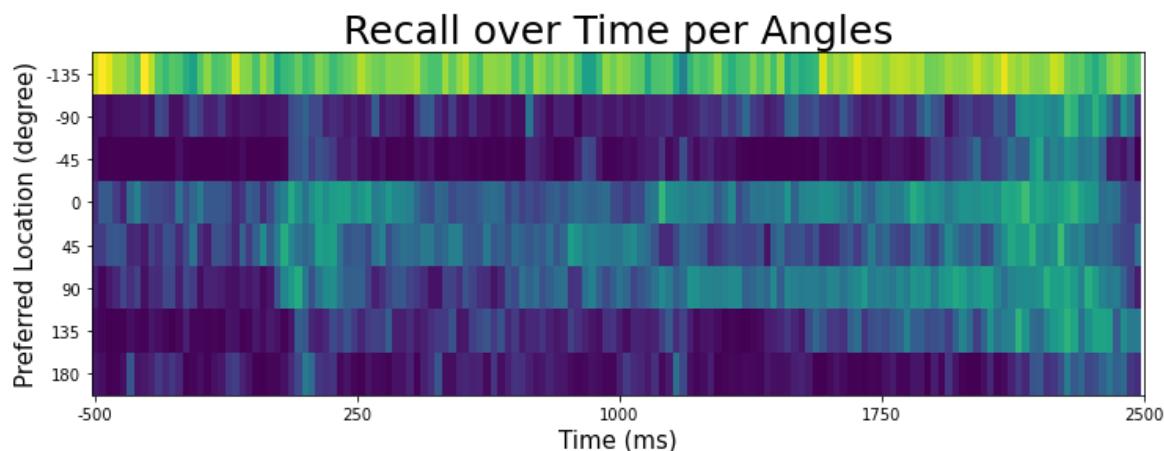
۳

۴

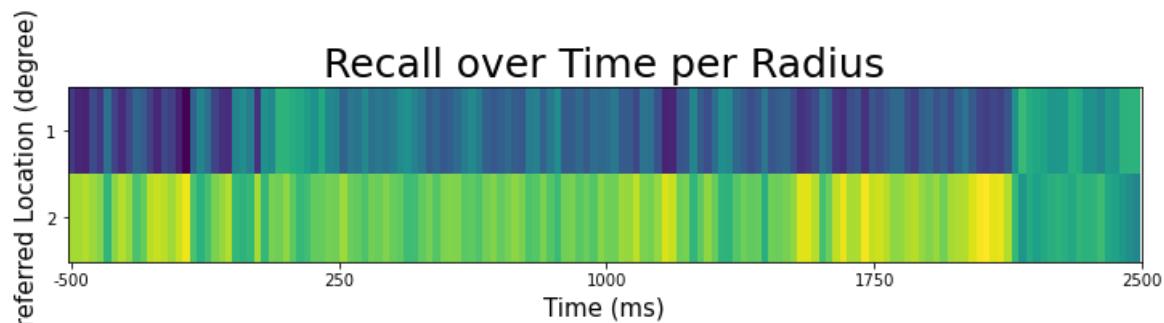
۵

سوال چهارم – جمعیت نورون‌ها: اندازه‌گیری اطلاعات

روش اطلاعات متقابل، معیار ساده‌ای برای کار کردن با جمعیت نورون‌ها نمی‌باشد. یک روش جایگزین استفاده از دقت طبقه‌بندها می‌باشد. در اینجا به طور خاص از طبقه‌بند SVM استفاده می‌کنیم. در نهایت معیار پوشش زوایای مختلف در شکل ۱۵ و شعاع‌های مختلف در شکل ۱۶ قابل مشاهده است. بنظر می‌رسد در پنجره زمانی‌ای که محرک در آن به نمایش در می‌آید(یعنی زمان صفر)، افزایشی نسبی در مقدار پوشش طبقه‌بندها اتفاق می‌افتد.



شکل ۱۵ - معیار پوشش طبقه‌بندها برای زوایای مختلف در پنجره‌های زمانی



شکل ۱۶ - معیار پوشش طبقه‌بندها برای شعاع‌های مختلف در پنجره‌های زمانی