به نام خدا



تمرین سری سوم درس BCI

محمدمهدی حبیبی

شماره دانشجویی:401617033

استاد:دكتر شالچيان

پاییز-زمستان1402

لب	مطا	ست	فم
<del>-</del>			~

سوال 1	3
سوال 2	6
استخراج ویژگی با استفاده از واریانس	6
استفاده از channel selection	12
سوال سوم	13
جواب سوال سوم	17
استفاده از مدل یادگیری ماشین SVM	17
ار زیابی مدل یادگیری ماشین SVM	17
استفاده از مدل یادگیری ماشین KNN	18
ار زیابی مدل یادگیری ماشین KNN	19
سوال چهارم	20
سوال پنجم	
سو ال شيشم 	



### تمرین ۳

- در این تمرین مراحل پیش پردازش، استخراج ویژگی و طبقه بندی سیگنالها مورد ارزیابی قرار می گیرد.
- لطفا با توجه به توضیحات زیر، فایل اسکریپت نوشته شده (.m) و گزارش (pdf) خود را در زمان تعیین شده در سامانه
   LMS ثبت بفرمایید.
- گزارش تمرین باید شامل تصاویر خروجی و توضیحات مربوط به آن و توضیحات مربوط به اسکریپت و فرآیند طی شده در برنامه باشد. (نیازی به بارگذاری دیتاست نیست در فایل زیپ گزارش نیست)
  - تمام شبیه سازی های زیر را بر روی دیتاست تصور حرکت BCICIV\_calib\_ds1b\_100Hz انجام دهید.
    - این مجموعه داده حاوی سیگنالهای EEG برای تصور حرکت ۲ کلاسه است.

۱- در مرحله پیش پردازش، با تحلیل فرکانسی، در صورت لزوم از فیلترهای ناچ، میان گذر و ... استفاده کنید. آیا نیازی به استاندارد (نرمال) نمودن دادهها وجود دارد؟ در این صورت، دادهها را استاندارد کنید.

۲- در مرحله استخراج ویژگی، از ویژگیهای مختلف آماری (واریانس، کورتوسیس)، توان باندهای مختلف (دلتا، تتا، آلفا و بتا)
 استفاده کنید. در این مرحله ویژگیهای دو باند آلفا،دلتا و بتا،تتا را با استفاده از اسکتر پلات ( برای دو جفت باند) رسم کرده و فضای ویژگی رسم شده را تحلیل کنید. آیا در این فضا اختلافی میان ویژگیها قابل رویت است؟

۳- با استفاده از ویژگیهای استخراج شده در مرحله قبل، مسئله طبقه بندی دو کلاسه را برای انواع ویژگیها (ویژگیهای آماری و ویژگیهای فرکانسی) حل نمایید. در این مرحله به دلخواه میتوانید ویژگیها را با یکدیگر ترکیب کنید. راهنما: در این مرحله از دو کلاسیفایر SVM استفاده کنید و نتایج را باهم مقایسه کنید.

۴- با تحلیل نتایج فوق، مناسب ترین نوع ویژگی برای طبقه بندی این مسئله را تعیین نمایید.

۵- با تحلیل نتایج فوق، مناسب ترین نوع کلاسیفایر را برای طبقه بندی این مسئله را تعیین نمایید.

9- از روش ICA استفاده کنید و آرتیفکتها و نویزهای موردنظر را به انتخاب خود حذف کنید. سپس با سیگنالهای جدید مرحله استخراج ویژگی (نوع ویژگی به دلخواه) و طبقه بندی (نوع کلاسیفایر به دلخواه) را تکرار کنید. چه تفاوتی در نتایج ایجاد شده است؟

گام 1: ابتدا دیتا رو بارگذاری میکنیم و در ادامه power spectrum را برای یکی از کانال ها بطور دلخواه رسم میکنیم که در اینجا برای کانال 21 ام اینکار را با دستور pwelch انجام دادیم

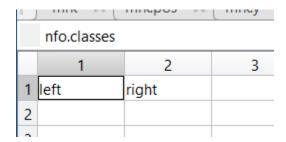
```
load('BCICIV_calib_ds1b_100Hz.mat');

cnt = double(cnt) *0.1;
fs =nfo.fs;
figure
plot(cnt(:,1))

hold on
pwelch(cnt(:,21));figure(gcf)

figure
hold on
pwelch(cnt(:,21),[],[],fs);figure(gcf);% plotting powerspectrum
```

همانطور که از اصلاعاتی که در قسمت nfo.classes هست ،سوژه قراره تصور دست راست و چپ را انجام بده در



نتیجه با یک مسئله ی دوکلاسه روبرو هستیم

#### سوال 1

در مرحله پیش پردازش ،با تحلیل فرکانسی در صورت لزوم از فیلترهای ناچ و میان گذر و . استفاده کنید آیا نیازی به نرمال سازی دیتا وجود دارد؟در این صورت داده ها را استاندارد کنید.

از انجا که فرکانس نمونه برداری دیتای مورد نظرمون برابر با 100 هرتز می باشد در نتیجه طبق قضیه نایکویست حداکثر اطلاعاتی که داریم تا 50 هرتز می باشد و همین طور در ادامه نیز چون می دونیم بیشتر اطلاعات تصور حرکتی بین بازه 8 تا 30هرتز می باشد برای همین از یک فیلتر میان گذر با فرکانس 8 تا 30 هرتز استفاده میکنیم بنابراین اصلا نیازی به طراحی فیلتر notch در این پروژه نمی باشد .

#### نرمال ساز<u>ی</u>

```
bands=[1,4 ;4,8; 8,12; 12,30];
for band=1:4
    for ch=1:59
disp(['Working on Bands:',num2str(band),'channel:',num2str(ch)]);

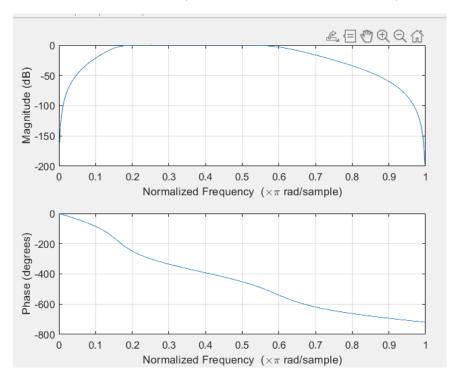
%[b,a] =buuter(order,wn,ftype)
    if band ==1
        [b,a] = butter(4,bands(band,2)/(fs/2));%make IIR filter (band pass filter )
        else
        [b,a] = butter(4,bands(band,:)/(fs/2));%make IIR filter (band pass filter )

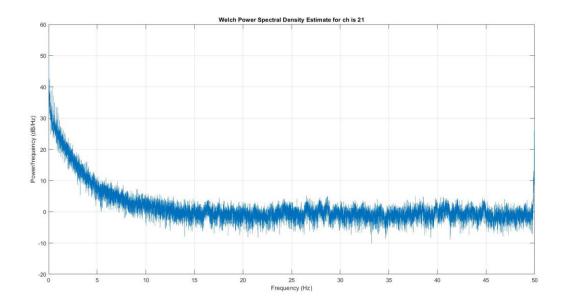
end
temp_cnt = filtfilt(b,a,cnt(:,ch));%we use filtfilt for problem of filt that make phase shift

temp_cnt = (temp_cnt-mean(temp_cnt))/std(temp_cnt);%Normalize
end
end
```

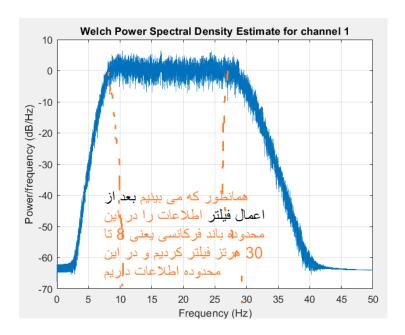
از اونجایی که می دانیم اطلاعات تصور حرکتی بیشتر در باند میو و بتا (hz30-8) می باشد ،سیگنال مون را در این بازه فیلتر میکنیم تا صرفا از اطلاعات محدوده ی 13 هر تز استفاده کنیم .

\*از دستور filtfilt نيز استفاده ميكنيم براى اعمال فيلتر تا دچار اختلاف فاز نشويم \*





در شکل بالا نمایشPower Spectral رو قبل از اعمال فیلتر میان گذر می بینیم و بعد از اعمال فیلتر میان گذر 8 تا 30 هرتز بصورت پایین میشود البته برای کانال های مختلف رسم کردیم ولی نتایج برای کانال های مشابه نیز همان می باشد



### سوال 2

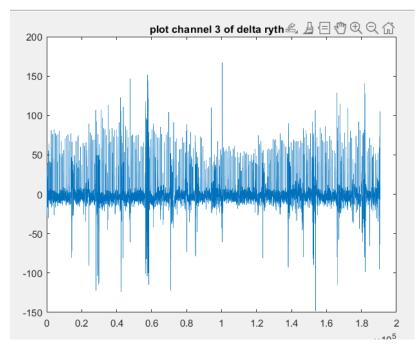
۲- در مرحله استخراج ویژگی، از ویژگیهای مختلف آماری (واریانس، کورتوسیس)، توان باندهای مختلف (دلتا، تتا، آلفا و بتا)
 استفاده کنید. در این مرحله ویژگیهای دو باند آلفا،دلتا و بتا،تتا را با استفاده از اسکتر پلات ( برای دو جفت باند) رسم کرده و فضای ویژگی رسم شده را تحلیل کنید. آیا در این فضا اختلافی میان ویژگیها قابل رویت است؟

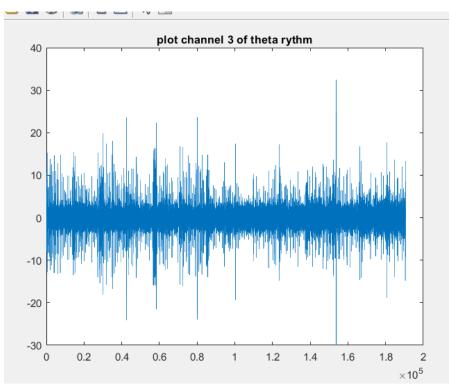
### استخراج ویژگی با استفاده از واریانس

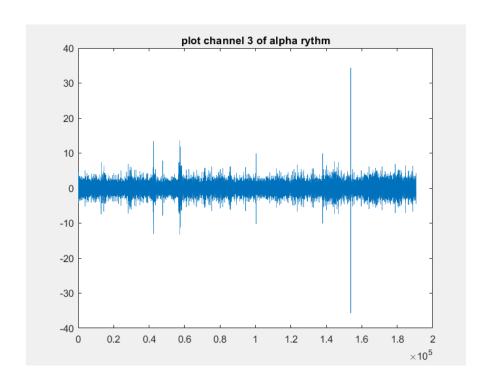
یکی از ویژگی های مهمی که می تونیم از سیگنال های EEG استخراج کنیم ،واریانس سیگنال هامون می باشد که با استفاده از دستور (var(signal می تونیم واریانس سیگنال مورد نظرمون رو بدست بیاریم.

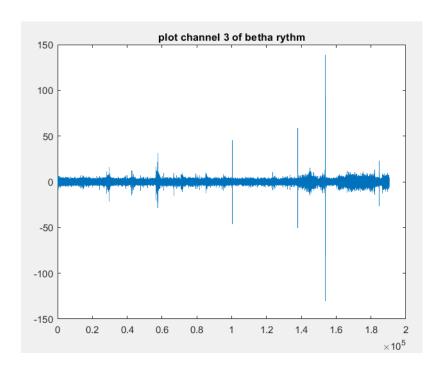
```
%% feature extraction
% create feature using variance
for i =1:length(class2)

class1_Feature (i,:) = var(class1{i});
class2_Feature (i,:) = var(class2{i});
end
```

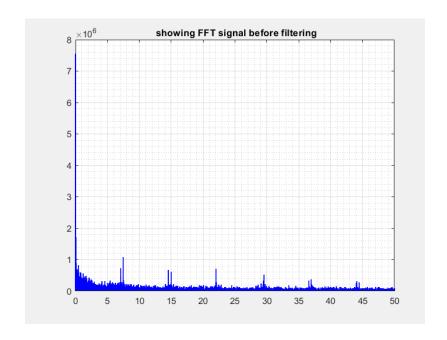


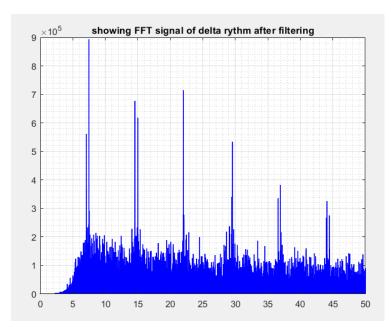






#### استفاده از تبدیل فوریه





همین طور که در دو شکل صفحه قبل دیدیم وقتی سیگنال را در حوزه فوریه می بریم و محدوده فرکانسی انرا محدود میکنیم ،میتوانیم اطلاعات فرکانسی مهمی را از ان استخراج کنیم

استخراج ویژگی های آماری برای باند های مختلف فرکانسی دلتا ،تتا،میو و بتا:

```
for i = 1:100
    var1_delta(i, :) = var(delta1{i});
    var1_theta(i, :) = var(theta1{i});
    var1_mu(i, :) = var(mu1{i});
    var1 betha(i, :) = var(beta1{i});
    kur1_delta(i, :) = kurtosis(delta1{i});
    kur1 theta(i, :) = kurtosis(theta1{i});
    kur1_mu(i, :) = kurtosis(mu1{i});
    kur1_beta(i, :) = kurtosis(beta1{i});
    power1_delta(i, :) = sum(abs(fft(delta1{i})) .^ 2);
    power1 theta(i, :) = sum(abs(fft(theta1{i})) .^ 2);
    power1_mu(i, :) = sum(abs(fft(mu1{i})) .^ 2);
    power1 beta(i, :) = sum(abs(fft(beta1{i})) .^ 2);
end
for i = 1:n2
    var2_delta(i, :) = var(delta2{i});
    var2_theta(i, :) = var(theta2{i});
    var2_mu(i, :) = var(mu2{i});
    var2 beta(i, :) = var(beta2{i});
    kur2 delta(i, :) = kurtosis(delta2{i});
    kur2_theta(i, :) = kurtosis(theta2{i});
    kur2_mu(i, :) = kurtosis(mu2{i});
    kur2_beta(i, :) = kurtosis(beta2{i});
    power2 delta(i, :) = sum(abs(fft(delta2{i})) .^ 2);
    power2_theta(i, :) = sum(abs(fft(theta2{i})) .^ 2);
    power2 mu(i, :) = sum(abs(fft(mu2{i})) .^ 2);
    power2_beta(i, :) = sum(abs(fft(beta2{i})) .^ 2);
end
```

#### از دستور روبرو نیز برای نمایش scatterدو جفت باند (الفا و بتا ) و (دلتا و تتا )استفاده میکنیم

```
figure
scatter(var_delta, var_alpha)
xlabel('Delta Variance')
ylabel('Alpha Variance')
title('Scatter Alpha Var vs Delta Var')

figure
scatter(kur_delta, kur_alpha)
xlabel('Delta Kur')
ylabel('Alpha Kur')
title('Scatter of Alpha Kur vs Delta Kur')

figure
scatter(power_delta, power_alpha)
xlabel('Delta Power')
ylabel('Alpha Power')
title('Scatter of Alpha Power vs Delta Power')
```

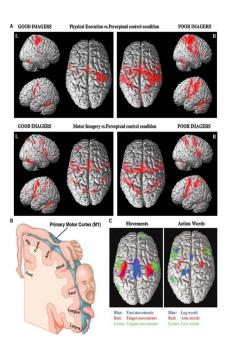
#### استفاده از channel selection

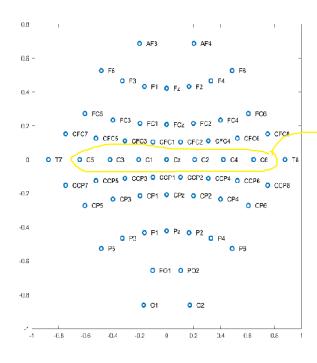
نیز بزنیم

#### %% channel selection

```
channel_index = 26:32; % select channels :c5 c3 c1 cz c2 c4 c6
channel_index(4) = []; % ignore information of Cz
cnt_ch = new_cnt(:,channel_index);
```

همانطور که انتظار داریم بیشتر اطلاعات تصور حرکتی مربوط به ناحیه primary motor areaیا همان ناحیه M1 می باشد ،بنابراین می تونیم بجای اینکه از اطلاعات تمامی 59 کانال ثبت شده استفاده کنیم ،از اطلاعات کانال های مربوط به این task استفاده کنیم ،در نتیجه یک عملیات channel selection نیاز داریم که بصورت زیر عمل میکنیم:





12

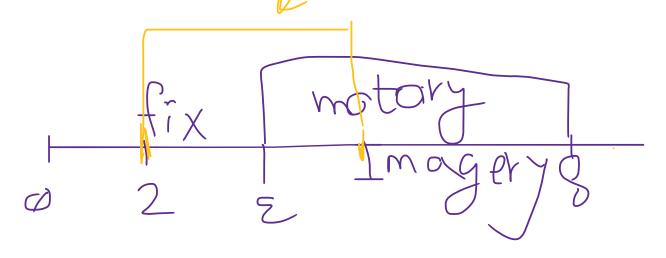
### سوال سوم

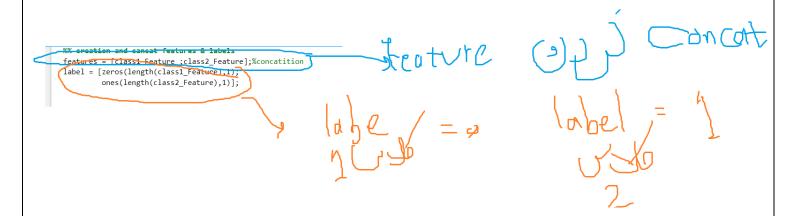
۳- با استفاده از ویژگیهای استخراج شده در مرحله قبل، مسئله طبقه بندی دو کلاسه را برای انواع ویژگیها (ویژگیهای آماری و ویژگیهای فرکانسی) حل نمایید. در این مرحله به دلخواه میتوانید ویژگیها را با یکدیگر ترکیب کنید.
 راهنما: در این مرحله از دو کلاسیفایر SVM و KNN استفاده کنید و نتایج را باهم مقایسه کنید.

```
number_class1 = sum(mrk.y == 1); %right hand
number_class2 = sum(mrk.y == -1); % left hand

class1 = cell(1,number_class1);
class2 = cell(1 , number_class2);
```

```
79
          %% state 2 when using channel selection
80
          for i =1:length(mrk.pos)
              if mrk.y(i) ==1
81
                  class1{i}= cnt_ch(mrk pos(i)+200 | mrk.pd*(i)+600,:
82
              elseif mrk.y(i) ==-1
83
                  class2{i}= cnt_ch(mrk.pos(i)+200 : mrk.pos(i)+600,:);
84
              end
85
86
87
          end
88
```





طول بازه مربوط به تصور حرکتی را در واقع باید از 400تا 800 بگیریم چراکه فرکانس نمونه برداری مون برابر با 100هرتز می باشد و سوژه در بازه بین 4 تا 8 ثانیه تصور حرکت را انجام داده است ،در این پروژه برای این که بررسی کنیم ببینیم ایا با تغییر طول بازه مورد نظر دقت بهتر میشود یا خیر برای همین کمی از قسمت ترایالی که به سوژه حتی تسک را نشان داده اند را نیز بررسی کردیم

طبق این کد نیز برای کلاس یک label، را برابر با 0 گذاشتیم و برای کلاس دوم یا همان تصور دست چپ label، را برابر 1 گذاشتیم

# number\_class1 = sum(mrk.y == 1);

با دستور هایلایت شده روبرو می توانیم تعداد ترایال هایی که سوژه تصور دست راست یا کلاس اول را کرده است را بدست بیاریم که برابر با 100 می شود .

برای تعداد ترایال های تصور دست چپ نیز برابر با 100 می شود .

```
%% state 2 when using channel selection
for i =1:length(mrk.pos)
   if mrk.y(i) ==1
        class1{i}= cnt_ch(mrk.pos(i)+200 : mrk.pos(i)+600,:);
   elseif mrk.y(i) ==-1
        class2{i}= cnt_ch(mrk.pos(i)+200 : mrk.pos(i)+600,:);
   end
```

end

بعد در اینجا نیز ترایال هایی که سوژه تصور دست راست کرده را در کلاس یک و ترایال هایی که سوژه تصور دست چپ کرده را در کلاس دوم می ریزیم

class1 ×												
Ix200 cell												
1	2	3	4	5	6	7	8	9	10	11	12	13
01x6 doub	[]	401x6 doub	[]	401x6 doub	. 401x6 doub	. 401x6 doub	. 401x6 doub	401x6 doub	[]	401x6 doub		[]

حالا وقتی class1 را باز کنیم یک سلول 1\*200 می بینیم که ترایال هایی که در کلاس یک تصور دست راست کرده مقدار پیدا کردند و مابقی تهی شدند برای اینکه این مقادیر تهی را نیز از بین ببریم از دستور پایین استفاده میکنیم :

```
class1 = class1(~cellfun ('isempty' , class1));
class2 = class2(~cellfun ('isempty' , class2));
```

حالا ديگه ابعاد class1 و class2 بصورت 1\*100 مي باشد كه ديگه مقادير تهي نداريم .

در گام بعد نوبت این هست از ویژگی های اماری که استخراج کردیم استفاده کنیم برای همین انها را در یک ماتریس features می ریزیم که بصورت زیر اینکار را میکنیم:

```
%% feature selection
37
38
          mean cnt = mean(new cnt);
39
          ZeroMeanCnt = new_cnt - mean_cnt ;
          CovMatrix = cov(ZeroMeanCnt);
40
41
          Number EigVec = 35; % we choose only 35 principal components
42
43
          [EigVec , EigVal] = eig(CovMatrix);
          [Sort_EigVec , Sort_Index] = sort(diag(EigVal), 'descend');
44
          chooseEigVec = Sort Index(1:Number EigVec);%only 10 principal components
45
          cnt_pca = cnt * EigVec(:,chooseEigVec);
46
```

همین طور label گذاری را نیز بصورت بالا انجام می دهیم

از اونجایی که label ها رو خودمون درست کردیم در نتیجه برای صدتای اول label ای که گذاشتیم برابر با 0 و برای صدتای دوم نیز label برابر با 1هست ،برای اینکه کار یادگیری ماشین مون درست باشه بهتره ترتیب این label گذاری مون بصورت رندوم باشه ،برای اینکار نیز از دستور روبرو استفاده میکنیم:

استفاده از PCA

در قسمت بالا نیز همانطور که می بینیم از PCAاستفاده کردیم تا ابعاد فیچر مون رو کاهش بدیم تا بتونیم عملکرد مدل را بهبود بدیم

```
ر ابندا باید میانگین رو صفر کنیم و بعدش از ماتریس COV: ابندا باید میانگین رو صفر کنیم و بعدش از ماتریس Principal component اولیه استفاده میکنیم و در اینجا ما از 35تا principal component اولیه استفاده کردیم در (۱۹۵۳ تنیجه ابعاد فیچرمون از 19000*59 کاهش پیدا کرده است. (۱۹۵۳ تنیجه ابعاد فیچرمون از 19000*59 کاهش پیدا کرده است. (۱۹۵۳ تنیجه ابعاد فیچرمون از 19000*59 کاهش پیدا کرده است. (۱۹۵۳ تنیجه ابعاد فیچرمون از 19000*59 کاهش ایک العشان از ۱۹۵۳ تنیجه ابعاد فیچرمون از ایک العشان العشا
```

```
جواب سوال سوم
استفاده از مدل یادگیری ماشین SVM
```

```
%% % Classification using SVM
SVMModel = fitcsvm(features , label , 'KernelFunction', 'gaussian', 'ClassNames',[0;1]);
```

با دستور روبرو یک object می سازیم که مدل svm رو روی دیتامون fit کنیم ،نوع کرنل ماشین بردار پشتیبان مون رو نیز برابر gaussian قرار دادیم

### ارزیابی مدل یادگیری ماشین SVM

از اونجایی که دیتامون یک دیتای نوروساینسی می باشد و حجم دیتا بسیار کم هست ،بهتره از kfold cross validation یا random subsampling

همین طور که از کد بالا نیز می بینیم ،k را برابر 5 قرار دادیم و یک PartitionModel ایجاد کردیم

برای بدست اور دن صحت یا همون Accuracy نیز از کد پایین استفاده کردیم:

ACC=1-kfoldLoss(PartitionModel , 'LossFun', 'classiferror');

#### نتيجه

ACC for using SVM model is:0.53

#### استفاده از مدل یادگیری ماشین KNN

```
%% Classification using KNN

KnnModel = fitcknn(features , label , "NumNeighbors" (5)

ACC=1-kfoldLoss(PartitionModel , 'LossFun', 'classiferror');

* validation

PartitionModel = crossval(KnnModel , "NumNeighbors",5);

* validation

PartitionModel = crossval(KnnModel , "NumNeighbors",5);

* validation

PartitionModel = crossval(KnnModel , 'kfold' , 5);

ACC=1-kfoldLoss(PartitionModel , 'LossFun', 'classiferror');

* validation using k-fold cross validation

PartitionModel = crossval(SVMModel , 'kfold' , 5);

* validation using k-fold cross validation

PartitionModel = crossval(SVMModel , 'kfold' , 5);
```

k fold cross validation=5 استفاده کر دیم که در ادامه نیز بطور خلاصه این متد

ارزیابی را در صفه بعد توضیح دادیم

#### نتيجه

# ACC for using KNN model is :0.585

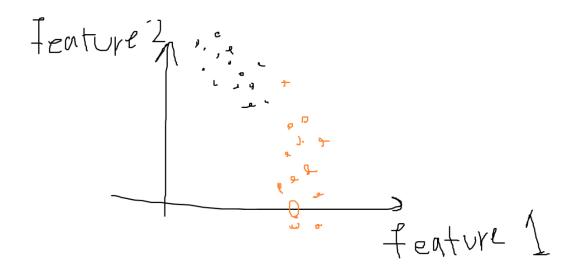
#### ارزیابی مدل یادگیری ماشین KNN

برای اینکه دقتی که بدست می اوریم قابل اعتماد باشد و به این صورت نباشد که بصورت شانسی مدل مون در قسمت خوب دیتا فیت شده باشد ،دیتا را به k تا fold می شکنیم و یک از این قسمت ها را برای test و مابقی را برای train می زاریم و هر سری یک ACC بدست می اوریم و مطابق شکل بالا این کار را تکرار میکنیم و دراخر یک میانگین گیری از تمامی دقت های بدست امده انجام می دهیم تا به صحت و عملکرد با اطمینان بیشتری دست پیدا کنیم.

## سوال چهارم

۴- با تحلیل نتایج فوق، مناسب ترین نوع ویژگی برای طبقه بندی این مسئله را تعیین نمایید.

مناسب ترین ویژگی با توجه به تحلیل هایی که انجام دادیم این می باشد که ابتدا سیگنال را از حوزه زمان به حوزه فوریه ببریم و سپس از باند های الفا و بتا استفاده کنیم در واقع یک فیلتر میانگذر 8 تا 30 هرتز بزنیم و بعدش از power سیگنال و واریانس سیگنال استفاده کنیم و در اخر از یک فیلتر فضایی CSP استفاده کنیم و این گونه می شود که فضای ویژگی بصورت شکل پایین می شود:



### سوال ينجم

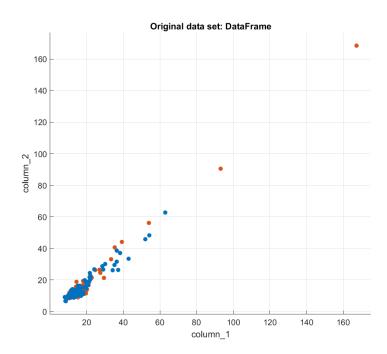
۵- با تحلیل نتایج فوق، مناسب ترین نوع کلاسیفایر را برای طبقه بندی این مسئله را تعیین نمایید.

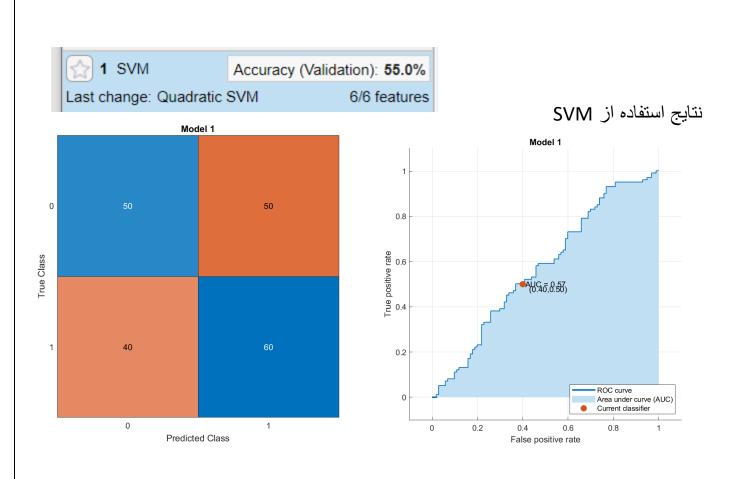
برای اینکه به نتیجه های دقیق تری برسیم از محیط گرافیکی متلب نیز حتی کمک گرفتیم تا بتونیم مدل های یادگیری ماشین بیشتری را نیز با پارامتر های مختلف امتحان کنیم ،که در ادامه دوتا از مدل ها ینی SVMو KNN با پارامتر های مختلف را نتایج شون را اور دیم

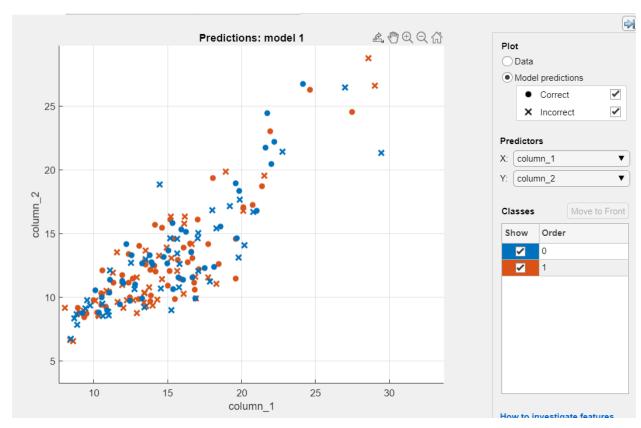
در این قسمت از محیط گرافیکی متل یعنی "classificationLearner

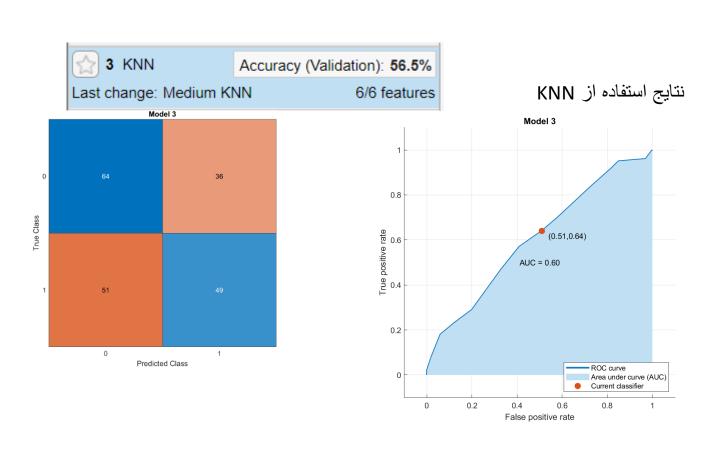
استفاده کردیم و Dataframe مون را وارد این محیط میکنیم New Session from Workspace Validation Data set Data Set Variable Cross-Validation DataFrame 200x7 double Protects against overfitting by partitioning t data set into folds and estimating accuracy Use columns as variables on each fold. Use rows as variables 3 Cross-validation folds: Response From data set variable From workspace O Holdout Validation column\_7 double 0 .. 1 ▼ Recommended for large data sets. 25 🗅 **Predictors** Name Туре Range ✓ column 1 double 8.06101 .. 167.22 Resubstitution Validation ✓ column 2 double 6.56858 .. 168.538 No protection against overfitting. The app uses all the data for both training and validation. column\_3 double 6.65096 .. 157.359 ✓ column\_4 double 7.03207 .. 160.761 Remove All Add All How to prepare data Read about validation Start Session Cancel Response variable is numeric. Distinct values will be interpreted as class labels.

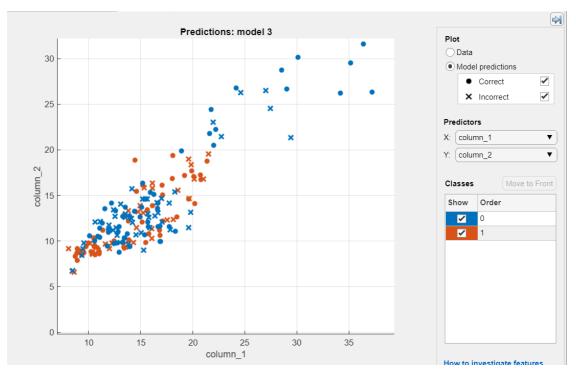
# توزیع دیتا وقتی که از channel selection استفاده کردیم بصورت زیر می باشد







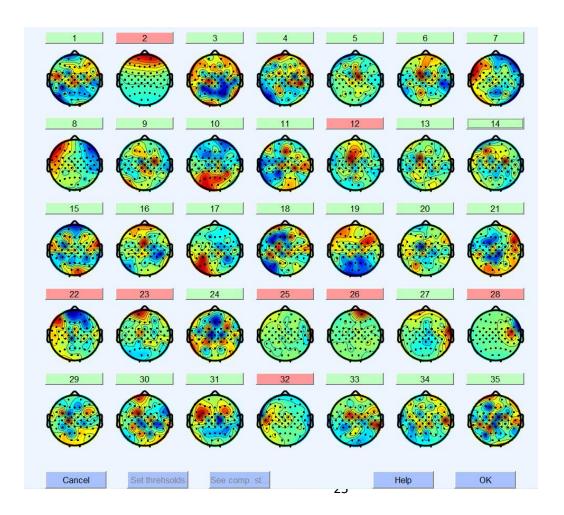


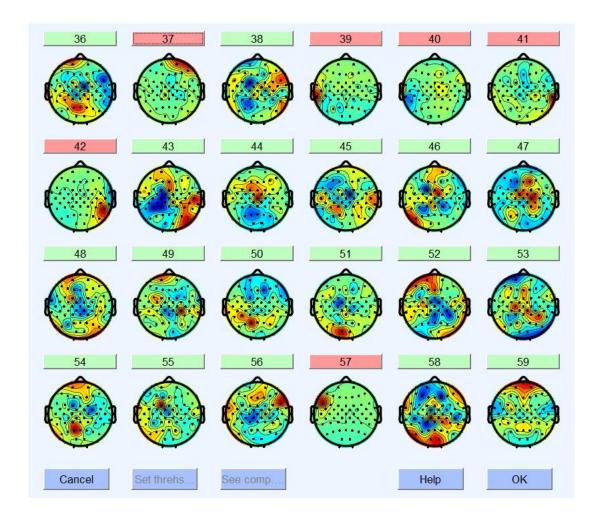


### سوال شيشم

۶- از روش ICA استفاده کنید و آرتیفکتها و نویزهای موردنظر را به انتخاب خود حذف کنید. سپس با سیگنالهای جدید مرحله استخراج ویژگی (نوع ویژگی به دلخواه) و طبقه بندی (نوع کلاسیفایر به دلخواه) را تکرار کنید.
چه تفاوتی در نتایج ایجاد شده است؟

اول که دیتا را مانند تمرین سری دوم وارد محیط eeglab میکنیم بعدش موقعیت کانال ها را مشخص میکنیم و در ادامه رفرنس گذاری را بر مبنای Common Average Refrence می زاریم ،حالا از ICA استفاده میکنیم و بعدش در ادامه Componentهایی که نویزی هستند و حاوی اطلاعات مفیدی نمی باشند را حذف میکنیم که در ادامه Componentهایی که میخاهیم حذف بکنیم را علامت گذاری کردیم و بصورت زیر می باشد





حالا که componentهایی که نویزی بودند را مشخص کردیم انها را حذف میکنیم

Components [2,12,22,23,25,26,28,32,37,39,40,41,42,57] flagged for rejection.

Do you want to remove these components?

Note: we recommend removing components in STUDY instead

Cancel

Manual rej.

Yes

بعدش سیگنال تمیز شده را از قسمت export می گیریم و نام سیگنال مون را برابر با cntRemoveBadComponents می گذاریم