

بسمه تعالی



دانشگاه صنعتی شریف

اصول مهندسی پزشکی

گزارش پروژه

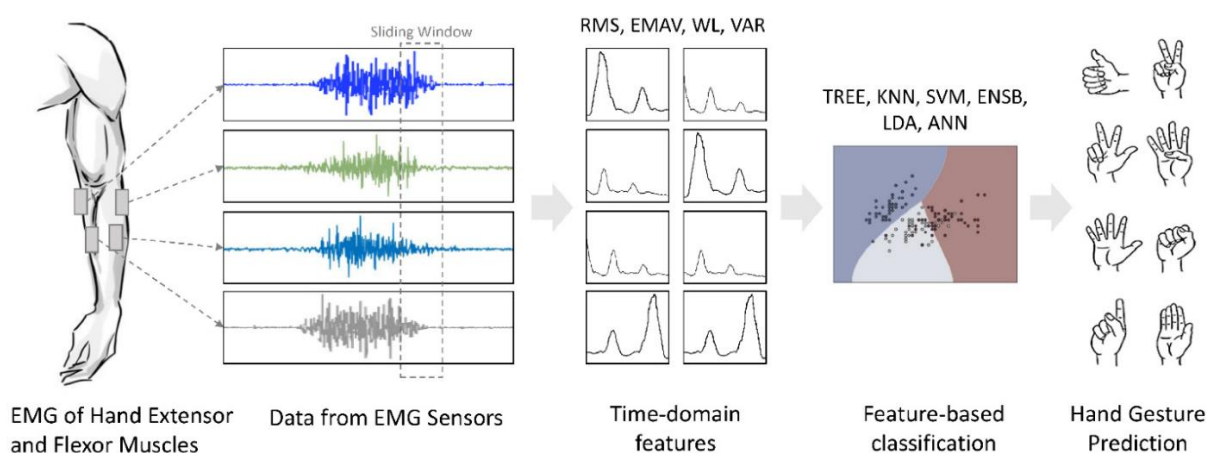
پردازش سیگنال EMG به منظور شناسایی حرکات دست

محمدحسین شفیع زادگان - ۹۹۱۰۴۷۸۱

نوید باقری شورکی - ۹۹۱۰۹۶۵۸

سید حسن نوری - ۴۰۲۲۰۵۴۲۱

هدف از این پروژه پردازش و طبقه بندی سیگنال الکترومایوگرام به منظور تشخیص و شناسایی ژست دست مربوط به هر حرکت است. این کار توسط چهار مرحله اصلی یعنی پیش پردازش سیگنال ها مانند فیلتر کردن آنها، پنجره سازی داده ها، استخراج ویژگی های مهم از روی سیگنال ها الکترومایوگرام و در انتها استفاده از الگوریتم های یادگیری ماشین به منظور طبقه بندی این سیگنال ها انجام می شود. نمای کلی کار نیز در تصویر زیر مشهود است.

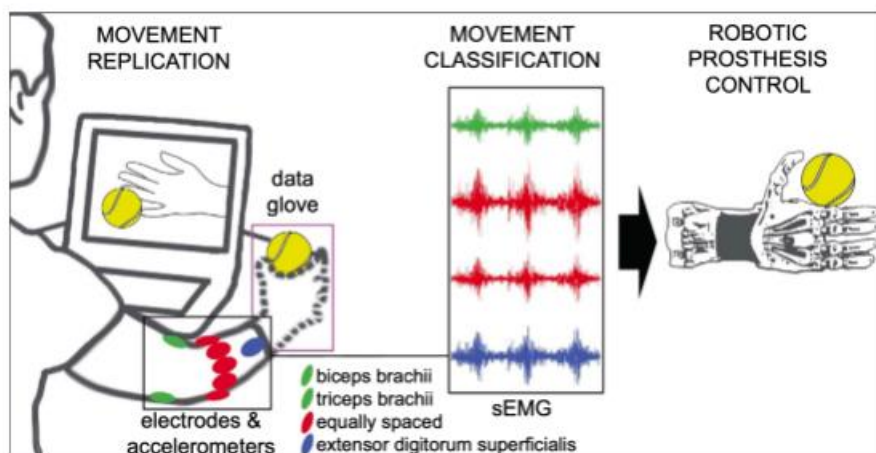


در این پروژه از آنجایی که سیگنال های ضبط شده در آزمایش اصلاً مطلوب نبوده و نتوانستیم دقت کافی را برای آنها داشته باشیم، لذا از دیتاست آماده موجود یعنی Ninapro و دیتاست DB2 استفاده کردیم. ویژگی های افراد داوطلب موجود در این دیتاست مانند قد، وزن و دست مورد استفاده در طول آزمایش در تصویر زیر آورده شده است. در اینجا ما از داده های مربوط به Subject 1, 4, 13 استفاده کردیم.

## DB2 Data

| Subject | Hand   | Handedness | Gender | Age | Height | Weight | zip files               |
|---------|--------|------------|--------|-----|--------|--------|-------------------------|
| 01      | Intact | Right      | Male   | 29  | 187    | 75     | <a href="#">s1.zip</a>  |
| 02      | Intact | Right      | Male   | 29  | 183    | 75     | <a href="#">s2.zip</a>  |
| 03      | Intact | Right      | Male   | 31  | 174    | 69     | <a href="#">s3.zip</a>  |
| 04      | Intact | Left       | Female | 30  | 154    | 50     | <a href="#">s4.zip</a>  |
| 05      | Intact | Right      | Male   | 25  | 175    | 70     | <a href="#">s5.zip</a>  |
| 06      | Intact | Right      | Male   | 35  | 172    | 79     | <a href="#">s6.zip</a>  |
| 07      | Intact | Right      | Male   | 27  | 187    | 92     | <a href="#">s7.zip</a>  |
| 08      | Intact | Right      | Male   | 45  | 173    | 73     | <a href="#">s8.zip</a>  |
| 09      | Intact | Right      | Male   | 23  | 172    | 63     | <a href="#">s9.zip</a>  |
| 10      | Intact | Right      | Male   | 34  | 173    | 84     | <a href="#">s10.zip</a> |
| 11      | Intact | Right      | Female | 32  | 150    | 54     | <a href="#">s11.zip</a> |
| 12      | Intact | Right      | Male   | 29  | 184    | 90     | <a href="#">s12.zip</a> |
| 13      | Intact | Left       | Male   | 30  | 182    | 70     | <a href="#">s13.zip</a> |

این دیتاست شامل اطلاعاتی مانند سیگنال EMG، اطلاعات حرکتی و مقدار نیرو است که از 40 داوطلب هنگامی که 49 حرکت دست به علاوه حالت استراحت را انجام می‌دادند، ثبت شده است. این دیتاست متشکل از 49 حرکت دست بوده که هر کدام 10 بار انجام شده است و از افراد داوطلب خواسته شد بود که حرکاتی که روی مانیتور لپ تاپ مشاهده می‌کنند را انجام دهند. در تصویر زیر شیوه انجام این آزمایشات آورده شده است.



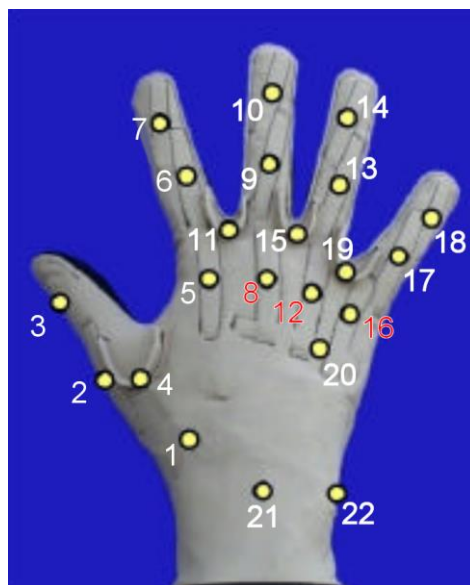
لیست متغیرهای موجود در دیتاست نیز در تصویر زیر آورده شده است که شامل شماره داوطلب، شماره فعالیت، سیگنال EMG ضبط شده توسط 12 الکتروود، مقدار نیروی ضبط شده، سیگنال‌های ضبط شده توسط 12 سنسور موجود روی دستکش و تحریک داده شده است.

#### Dataset variables

For each exercise, for each subject, the database contains one matlab file with synchronized variables. The variables included in the matlab files are:

- Subject: subject number.
- Exercise: exercise number.
- Emg (12 columns): sEMG signal. Columns 1-8 are the electrodes equally spaced around the forearm at the height of the radio humeral joint. Columns 9 and 10 contain signals from the main activity spot of the muscles flexor and extensor digitorum superficialis, while columns 11 and 12 contain signals from the main activity spot of the muscles biceps brachii and triceps brachii.
- Acc (36 columns): three-axes accelerometers of the 12 electrodes.
- Glove (22 columns): uncalibrated signal from the 22 sensors of the cyberglove.
- *The cyberglove signal corresponds to raw data from the cyberglove sensors, which are expected to be proportional to the angles at the joints.*
- Stimulus (1 column): the movement repeated by the subject, according to the displayed movie.
- Restimulus (1 column): again the movement repeated by the subject, but with the duration of the movement label refined a-posteriori in order to better correspond to the real movement.
- *Read the paper [Gijssberts et al., 2014](#) for more details about relabelling procedure.*
- Repetition (1 column): repetition of the stimulus.
- Rerepetition (1 column): repetition of restimulus.
- Force (6 columns): force recorded during the third exercise
- forcecal (2 x 6 values): the force sensors calibration values, corresponding to the minimal and the maximal force.

شیوه اتصال الکتروودها روی دستکش نیز در تصویر زیر آورده شده است.



همچنین حرکات موجود در آزمایش در هر تمرین نیز همانند تصویر زیر نشان داده شده است. از این دیتاست Exercise B مد نظر ما است.

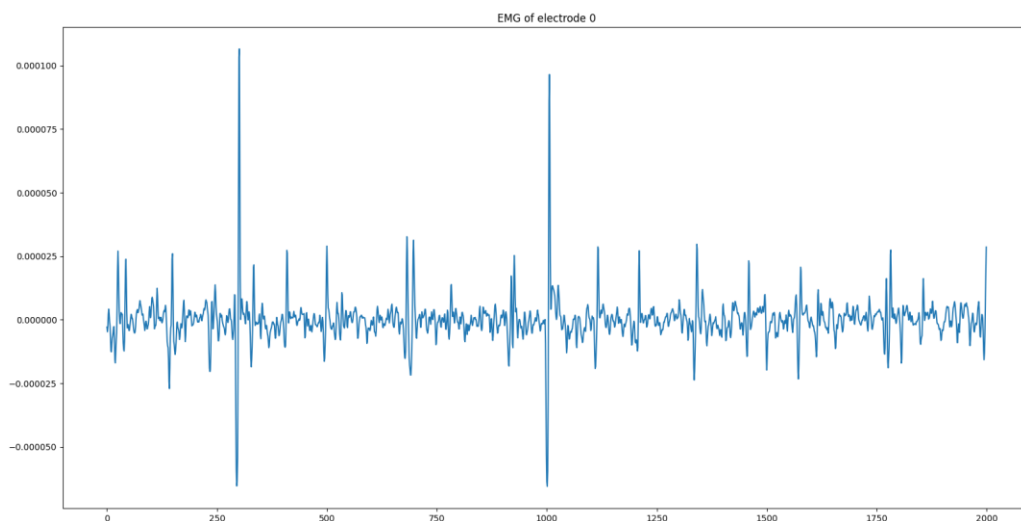
[illegible]

محیط کد زنی مورد استفاده پایتون است. ابتدا فایل دیتاست که یک فایل متلب بوده را همانند یک dataframe با استفاده از دستور `scipy.io.loadmat` در محیط پایتون load می‌کنیم. همچنین متغیرهای موجود در این دیتاست را نیز در cell های مختلف نمایش می‌دهیم. ابتدا نتایج مربوط به داوطلب شماره 1 را بررسی می‌کنیم. فرکانس نمونه برداری را 2 کیلو هرتز تعریف می‌کنیم.

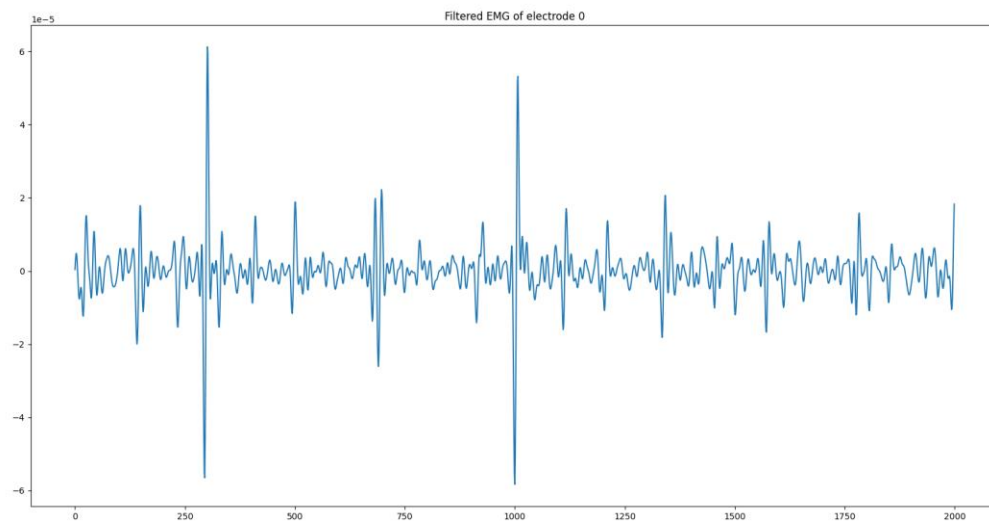
### Subject 1:

(الف)

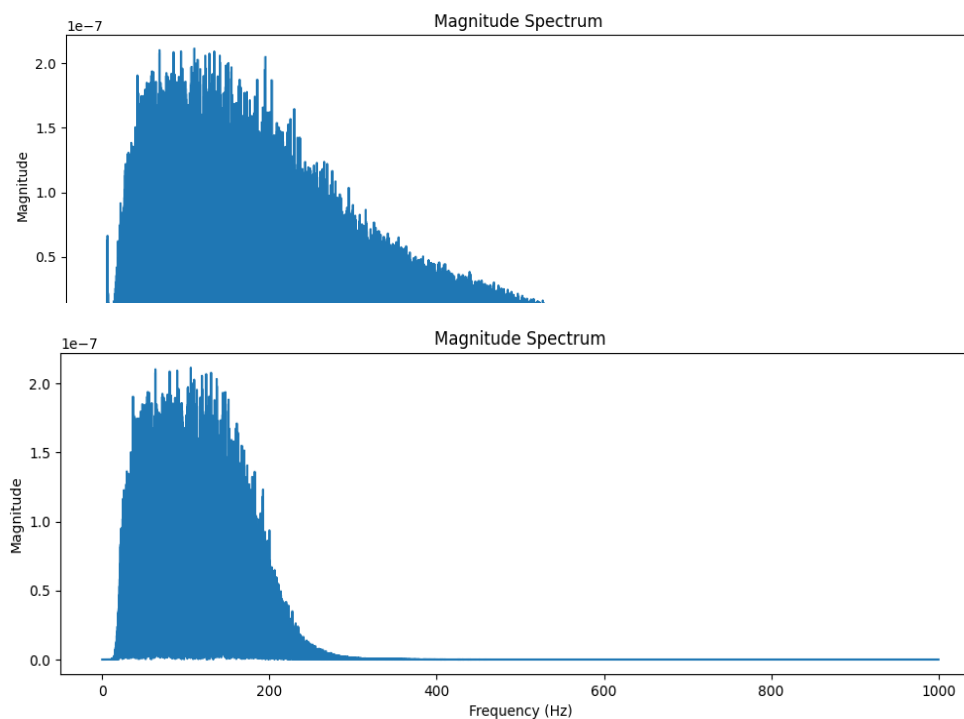
ابتدا سیگنال ضبط شده توسط الکتروود شماره صفر را رسم می‌کنیم.



در بخش مربوط به پیش پردازش سیگنال، ابتدا سیگنال را از یک فیلتر میان گذر با فرکانس قطع 20 و 200 هرتز عبور می‌دهیم. این فیلتر را به صورت یک فیلتر `butterworth` با مرتبه 4 تعریف کرده‌ایم و با استفاده از دستور `filtfilt` روی سیگنال اعمال نمودیم. نمودار سیگنال فیلتر شده در تصویر زیر آمده است که همانطور که مشاهده می‌کنیم مقدار نویز به صورت قابل توجهی حذف شده است. برای استفاده از فیلتر نیز از دستور آماده `butter` استفاده شده است.

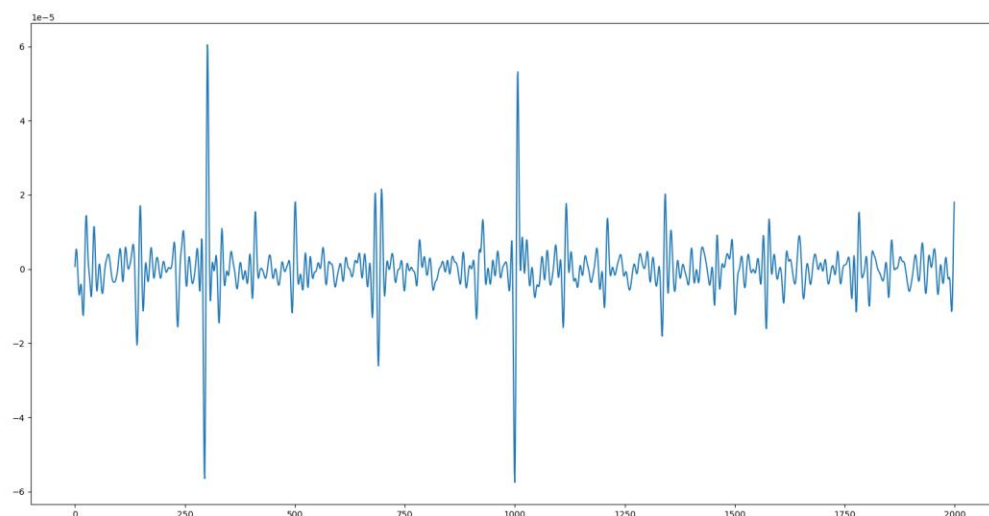


حال اسپکتروم سیگنال را پیش از فیلتر کردن و بعد از فیلتر کردن نمایش می‌دهیم. همانطور که مشاهده می‌کنیم پس از اعمال فیلتر، فرکانس‌های بیرون از بازه 20 تا 200 هرتز به میزان خوبی فیلتر شده‌اند.

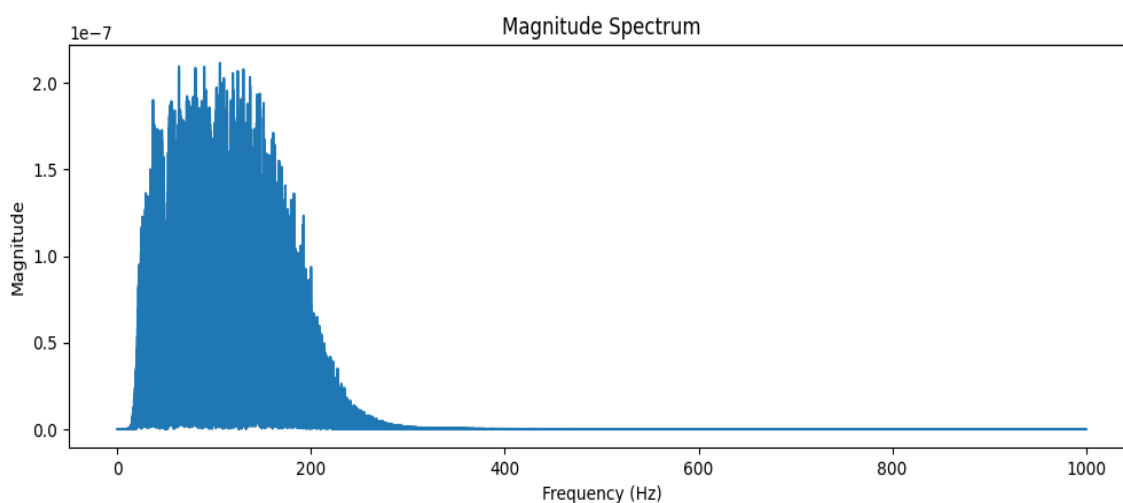




حال از یک Notch filter برای حذف برق شهر 50 هرتز استفاده می کنیم.



حال اسپکتروم سیگنال را پس اعمال این فیلتر نیز رسم می کنیم. همانطور که می بینیم دره بوجود آمده در فرکانس 50 هرتز نشان از حذف شدن نویز این فرکانس توسط فیلتر است.





(ب)

در این بخش داده‌ها را بر اساس کانال نرمالیزه می‌کنیم. با استفاده از مفهوم Z-score سیگنال را به صورت میانگین صفر و واریانس یک نرمالیزه می‌کنیم. این کار توسط دستور StandardScaler انجام می‌شود. این کار را در یک cell جدا گانه همراه با کامنت گذاری درون کد انجام دادیم.

(ج)

از آنجایی که سیگنال EMG تصادفی و غیر خطی است، برای استخراج الگوهای متفاوت از آن نیازمند پنجره بندی مناسب هستیم تا بتوانیم ویژگی‌های مربوطه را از هر پنجره استخراج کنیم. حال به سراغ پنجره بندی سیگنال می‌رویم. برای این کار یک تابع به اسم window\_signal تعریف کرده‌ایم که سیگنال‌ها را به پنجره‌های 200 میلی ثانیه‌ای و با همپوشانی 190 میلی ثانیه تقسیم بندی می‌کنیم. از آنجایی که این پنجره‌ها همپوشانی دارند، لیبل مربوطه را به حرکتی که در پنجره مورد نظر طولانی‌تر بوده اختصاص می‌دهیم. این کار را در یک cell جدا گانه همراه با کامنت گذاری درون کد انجام داده‌ایم.

(د)

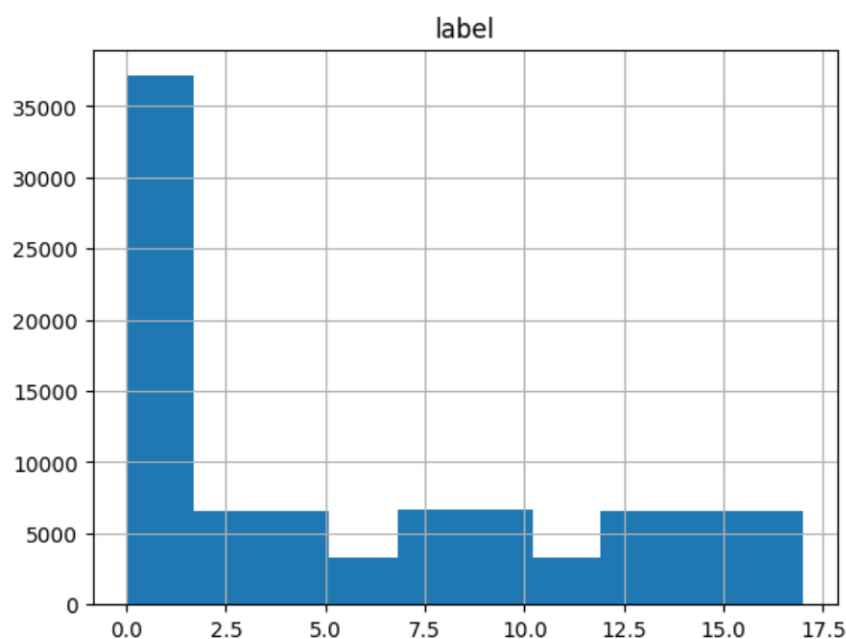
حال نوبت به استخراج ویژگی‌های مورد نظر از این پنجره‌های سیگنال است. در این بخش ما از تمامی 10 ویژگی مطرح شده در استفاده کردیم. لیست ویژگی‌ها نیز در پایین آورده شده است. هر کدام از این ویژگی‌ها را نیز توسط توابع موجود در numpy و همانند فرمول مطرح شده در صورت پروژه تعریف می‌کنیم.

## Feature Extraction

Recommended features:

- Mean absolute value (MAV)
- Standard deviation (STD)
- Variance (VAR)
- Waveform length (WL)
- Zero crossing (ZC)
- Root mean square (RMS)
- Number of peaks (NP)
- Slope sign change (SSC)
- Correlation coefficient (Cor)
- Integrated absolute value (IAV)

حال با اعمال هر یک از این ویژگی‌ها بر روی پنجره‌های زمانی جدا شده، dataframe مورد نظر را به همراه لیبل آنها ذخیره می‌کنیم. همچنین از آنجایی که exercise B مد نظر است بنابراین 17 لیبل داریم. اما اگر نمودار هیستوگرام تعداد پنجره‌های سیگنال را بر حسب لیبل‌ها رسم کنیم خواهیم دید که دیتای مربوط به حالت استراحت rest باعث شده تا توازن داده‌های ما بهم بخورد و به اصطلاح دیتاست را imbalance کرده است. همچنین از آنجایی که طبقه بندی این حالت برای ما مهم نیست لذا این حالت را حذف خواهیم کرد.

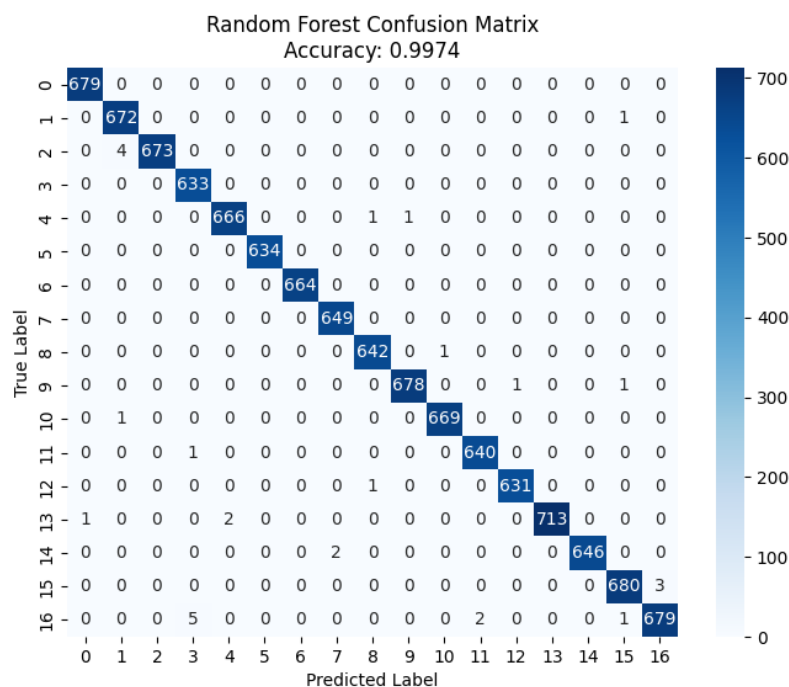
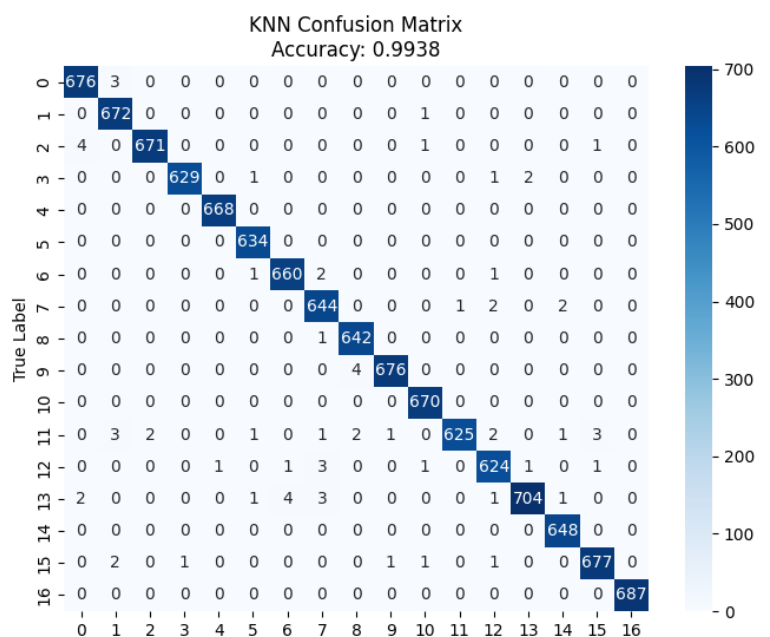


|       | Channel1_MAV | Channel1_STD | Channel1_VAR | Channel1_WL | Channel1_ZC | Channel1_RMS | Channel1_I |
|-------|--------------|--------------|--------------|-------------|-------------|--------------|------------|
| 331   | 0.515060     | 0.390652     | 0.471868     | 0.310114    | 0.410045    | 0.696874     | 0.7522     |
| 332   | 0.488845     | 0.398306     | 0.468998     | 0.304301    | 0.406603    | 0.717662     | 0.7829     |
| 333   | 0.478391     | 0.411778     | 0.462300     | 0.318238    | 0.396446    | 0.775304     | 0.7877     |
| 334   | 0.478142     | 0.423685     | 0.480526     | 0.329182    | 0.411051    | 0.796140     | 0.7530     |
| 335   | 0.502756     | 0.415550     | 0.479357     | 0.333576    | 0.422473    | 0.827958     | 0.7309     |
| ...   | ...          | ...          | ...          | ...         | ...         | ...          | ...        |
| 90045 | 0.297823     | 0.285804     | 0.316806     | 0.284572    | 0.298882    | 0.391661     | 0.3581     |
| 90046 | 0.299552     | 0.288627     | 0.318003     | 0.286064    | 0.298690    | 0.393943     | 0.3466     |
| 90047 | 0.300013     | 0.288636     | 0.317375     | 0.283789    | 0.299756    | 0.389949     | 0.3465     |
| 90048 | 0.300321     | 0.290637     | 0.320436     | 0.284609    | 0.301363    | 0.393934     | 0.3436     |
| 90049 | 0.296492     | 0.289605     | 0.316936     | 0.284395    | 0.298754    | 0.387136     | 0.3408     |

56239 rows × 121 columns

(۵)

حال داده‌های dataframe را به داده‌های آموزشی و تست با استفاده از دستور train\_test\_split تقسیم می‌کنیم و 80 درصد داده‌ها را به عنوان داده آموزشی انتخاب می‌کنیم. ما از دو طبقه بند KNN و Random Forest استفاده کرده‌ایم.



همانطور که مشاهده می‌کنیم هر دو الگوریتم به مقدار بسیار خوبی توانسته‌اند تا داده‌های مربوط به حرکات دست را طبقه بندی کنند. این اتفاق نیز متاثر از خود دیتاست بوده که دیتاست مورد استفاده بسیار خوب بوده و ما نیز از تمامی ویژگی‌های مطرح شده در صورت جزوه برای طبقه بندی استفاده کردیم.

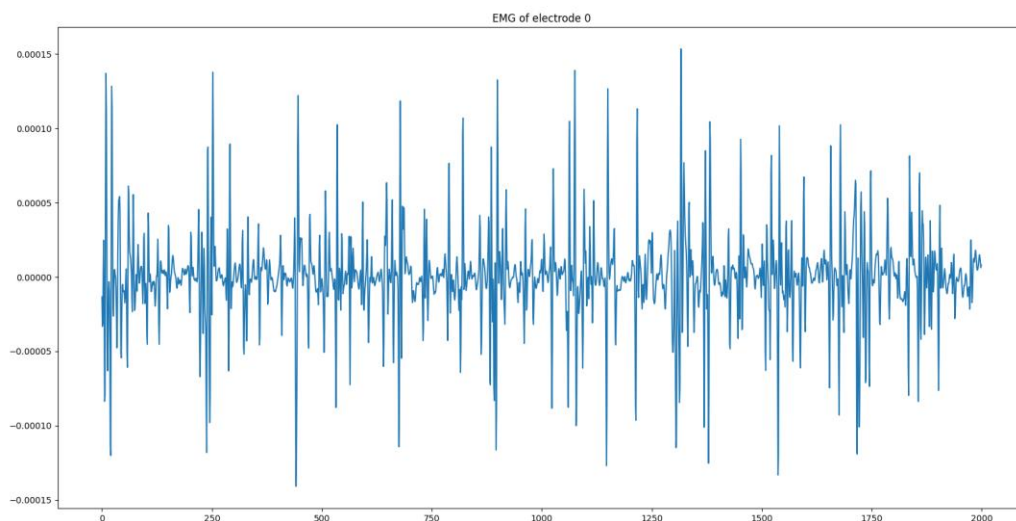
(ی)

دیتاست مورد استفاده اهمیت ویژه‌ای دارد به گونه‌ای اگر حرکات انجام شده توسط فرد به خوبی صورت گرفته نباشد و نویزها و آریتمیک‌های گوناگون مثل لرزش دست و برق شهر زیاد باشد، در ادامه نمی‌توان سیگنال‌های مربوط به هر حرکت را به درستی طبقه بندی کرد. داده‌های ثبت شده در آزمایشگاه این مشکل را داشتند و به گونه‌ای سیگنال حرکات مختلف به درستی از هم تمیز داده نشده بود برای همین از دیتاست آماده استفاده کردیم که این داده‌های ضبط شده در آن بسیار مناسب و مطلوب بود. همچنین ویژگی‌های مورد استفاده نیز اهمیت دارند و با افزایش آنها می‌توان دقت را افزایش داد و به گونه‌ای با افزایش بعدهای بیشتر بهتر می‌توانیم سیگنال‌ها را از هم تمیز دهیم. قبل از استفاده از سیگنال‌ها نیز باید پیش پردازش به صورت مناسب انجام شود. برای اینکار باید از فیلتر مناسب استفاده کرده و بازه فرکانسی که محتوای اصلی سیگنال در آن قرار دارد را استخراج کرده و همچنین با استفاده از فیلتر مناسب مثل notch نویز برق شهر را نیز حذف کنیم.

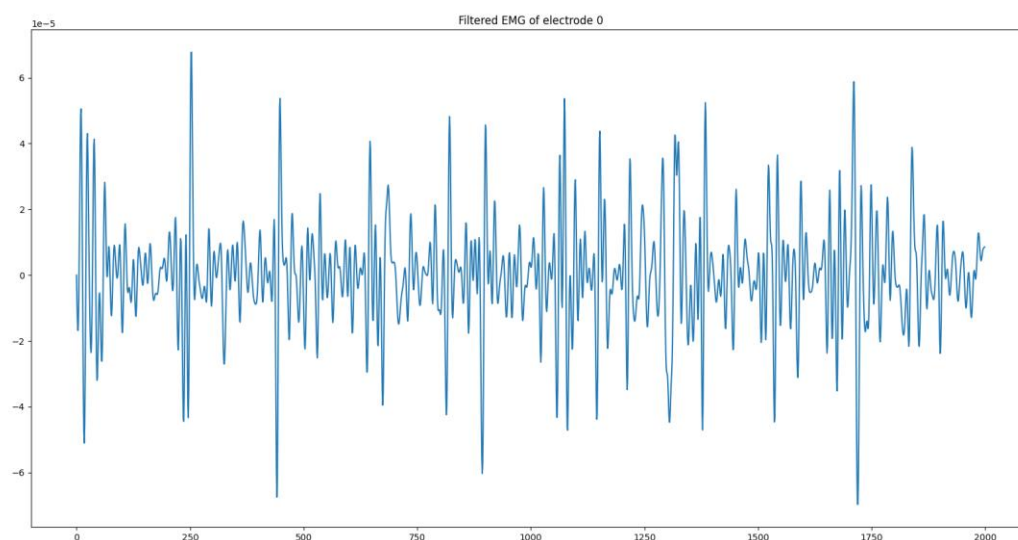
حال داده‌های مربوط به دو داوطلب دیگر را بررسی می‌کنیم.

### **Subject 4:**

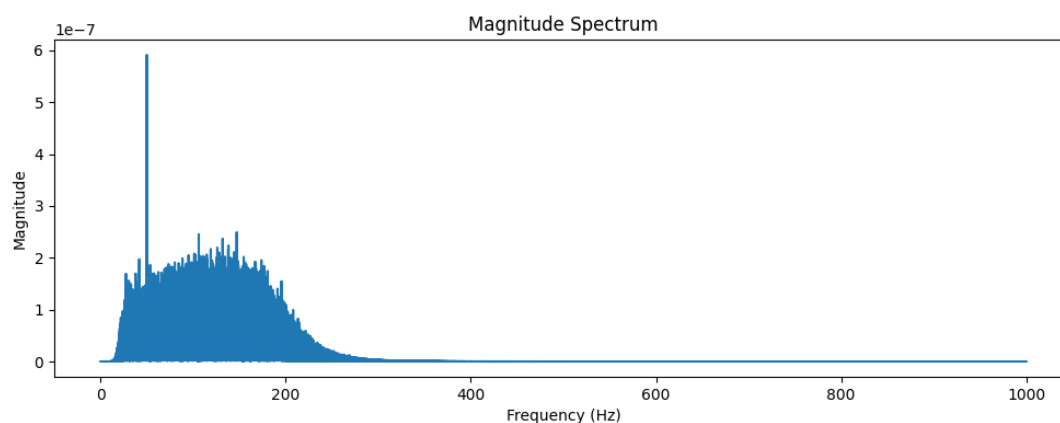
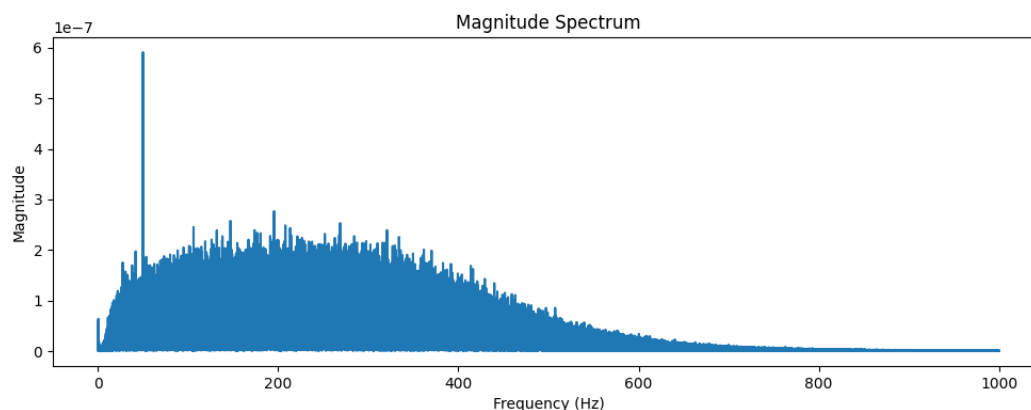
ابتدا سیگنال ضبط شده توسط الکترود شماره صفر را رسم می‌کنیم.



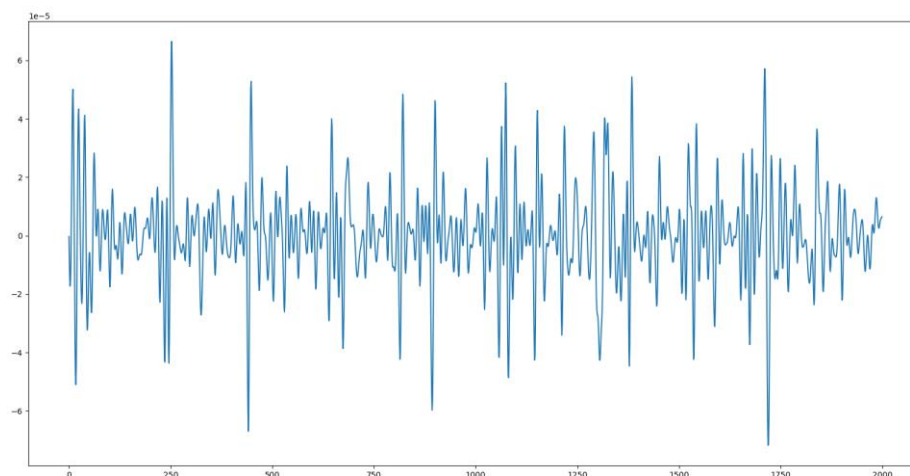
حال سیگنال را پس از اعمال فیلتر میان گذر Buterworth رسم می‌کنیم.

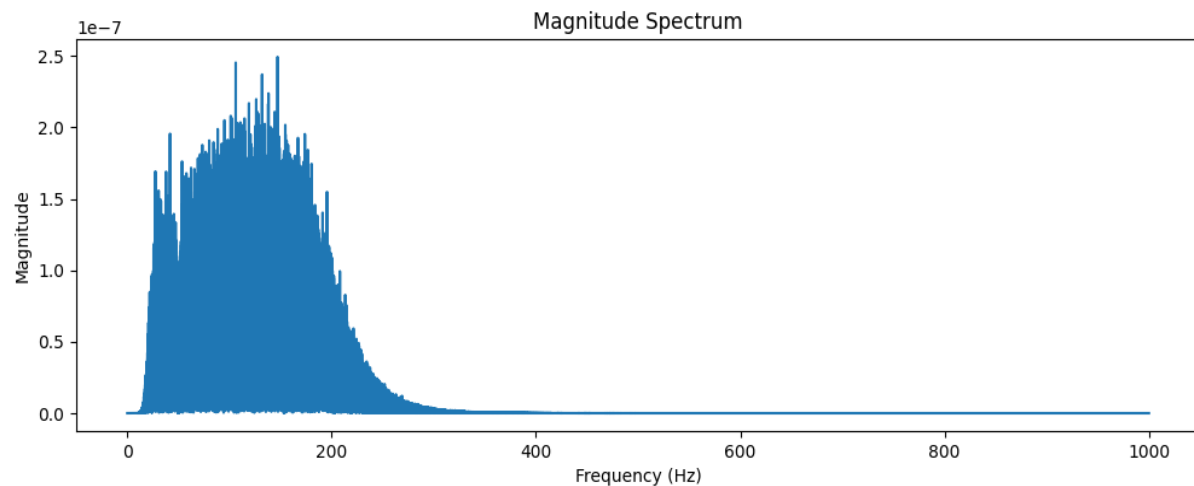


حال اسپکتروم سیگنال قبل و بعد از فیلتر کردن رسم می کنیم.

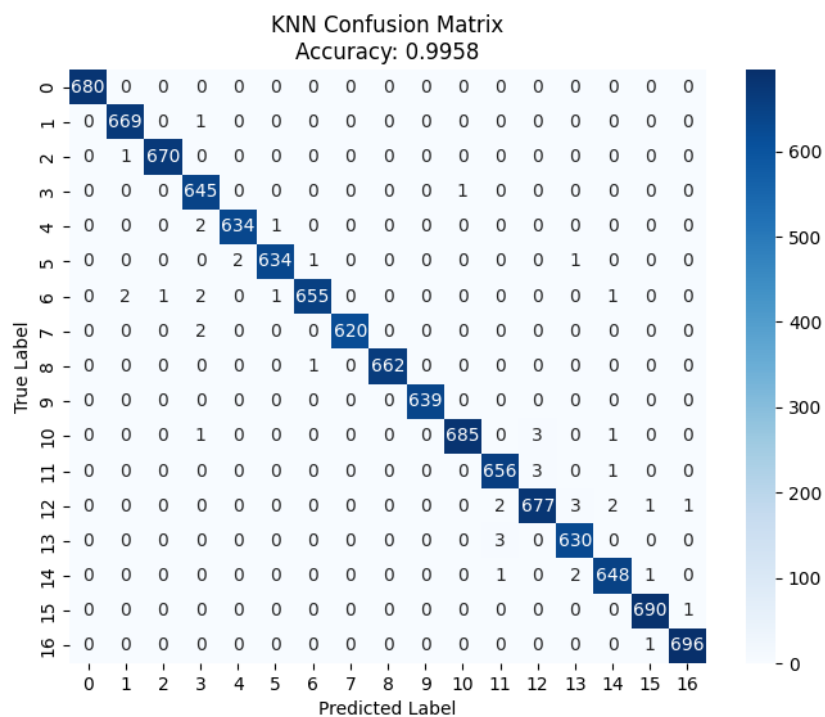


حال به منظور حذف نویز برق شهر در فرکانس 50 هرتز از notch filter استفاده کرده و نمودار سیگنال و اسپکتروم آن را پس اعمال این فیلتر نیز رسم می کنیم.

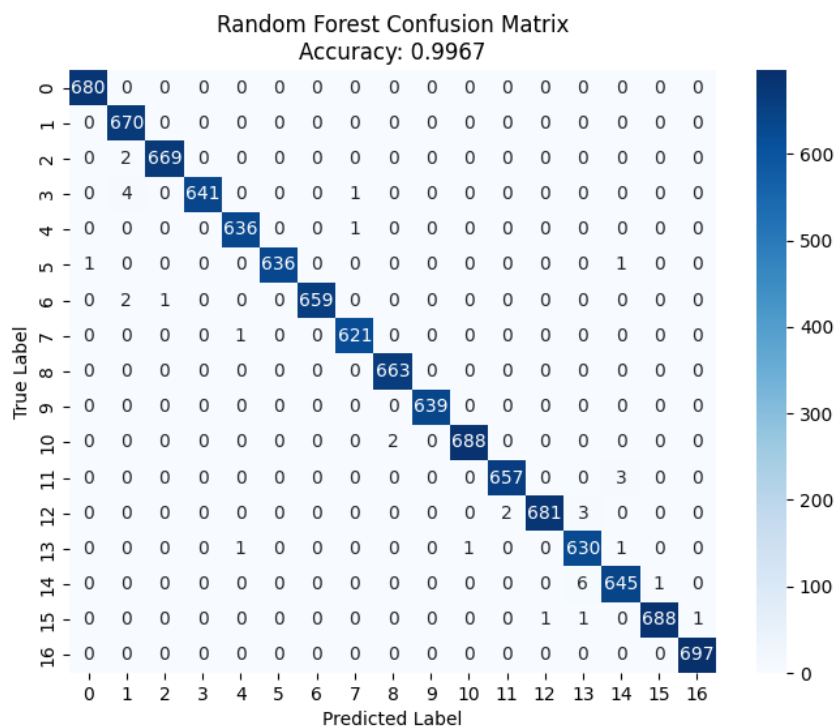




حال با استفاده از دو الگوریتم KNN و Random Forest و استفاده از تمامی ویژگی‌ها، هر 17 حرکت را طبقه بندی می‌کنیم.



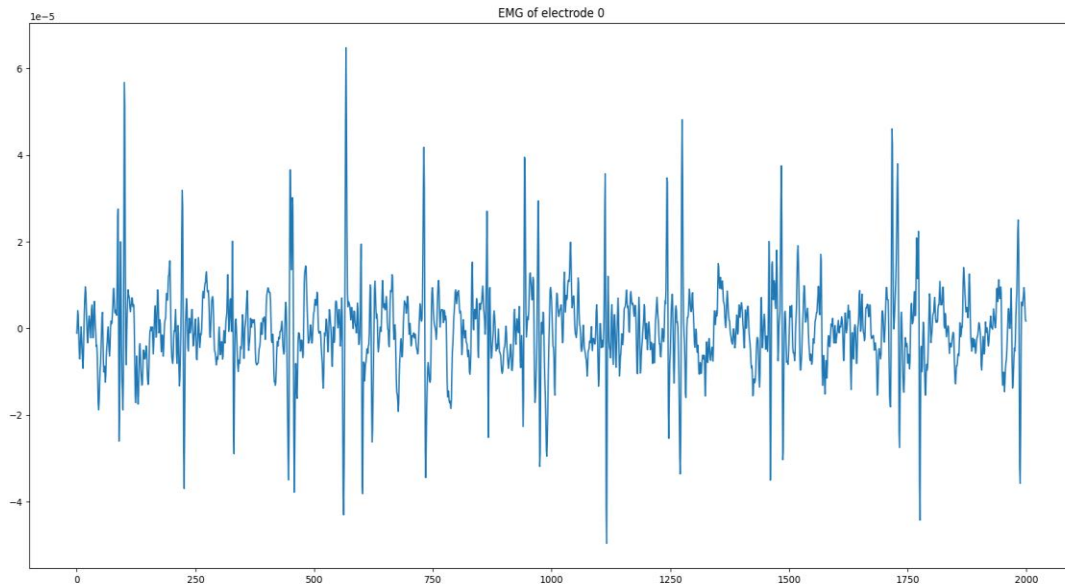




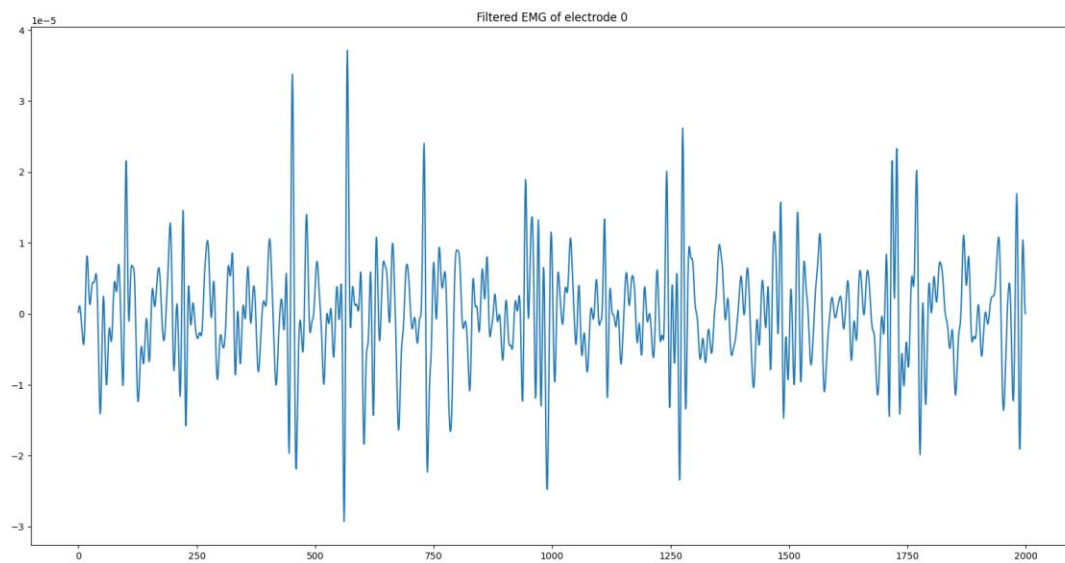
همانطور که مشاهده می‌کنیم برای داوطلب شماره 4 نیز توانسته‌ایم به دقت بسیار خوبی دست پیدا کنیم و تقریباً تمامی داده‌ها توانسته‌اند به درستی طبقه بندی شوند.

### Subject 13:

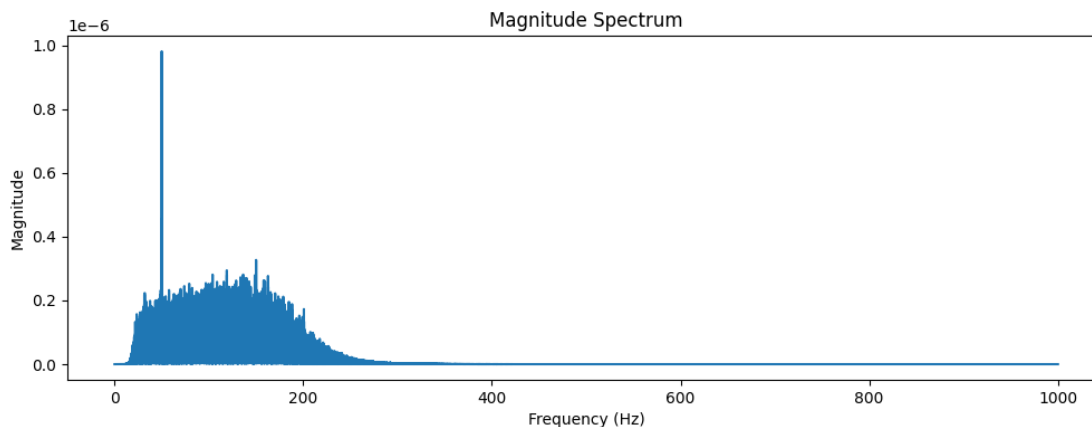
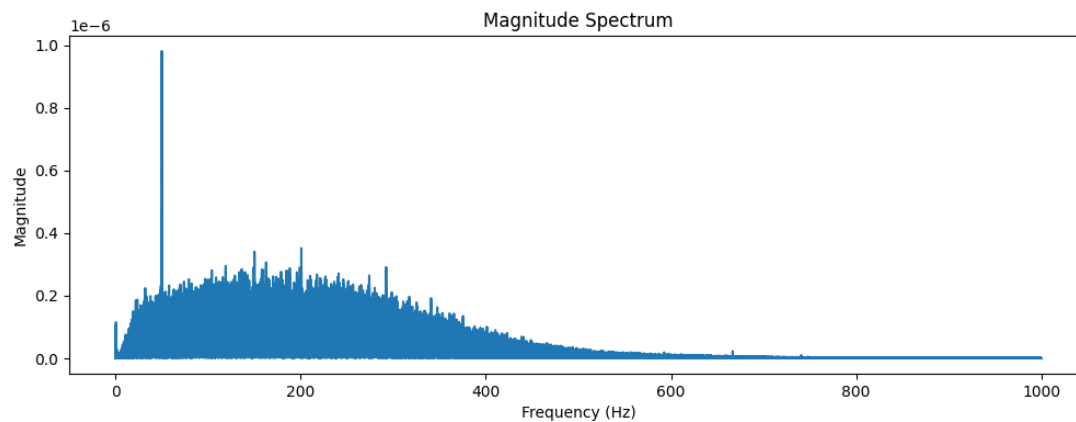
ابتدا سیگنال ضبط شده توسط الکترود شماره صفر را رسم می‌کنیم.



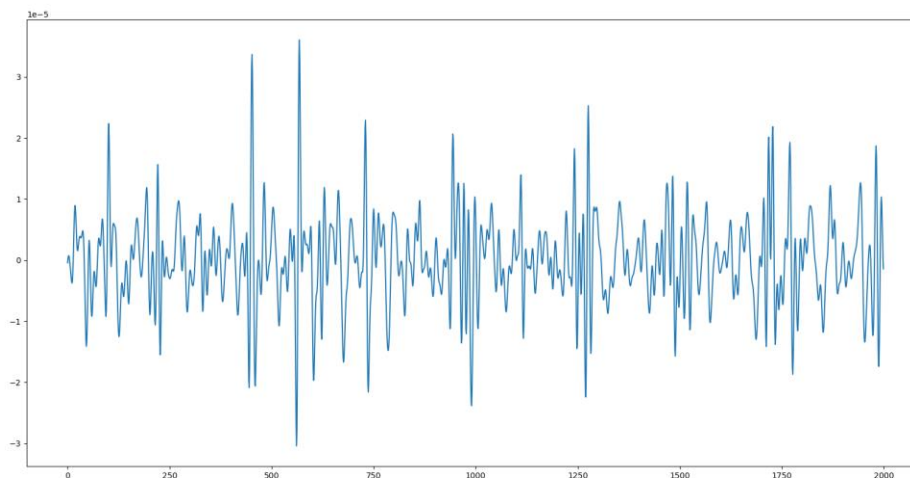
حال سیگنال را پس از اعمال فیلتر میان گذر Buterworth رسم می‌کنیم.

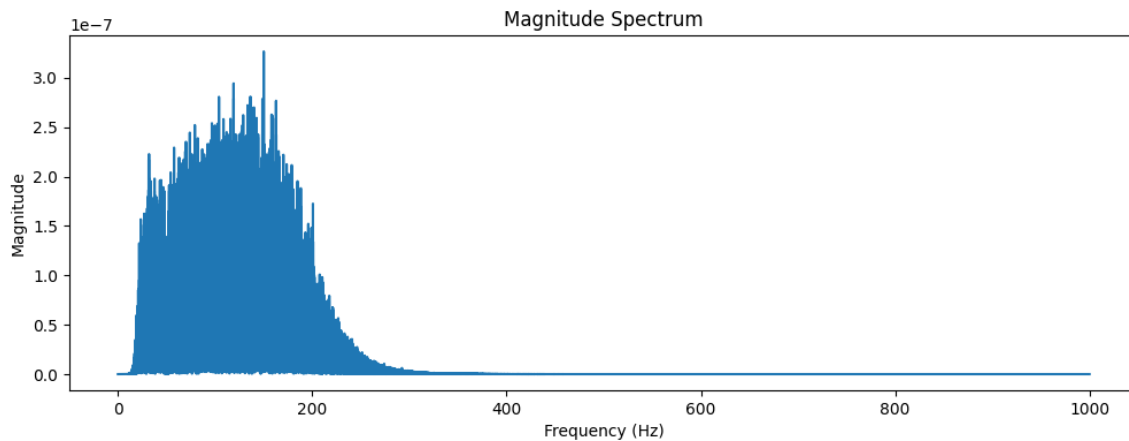


حال اسپکتروم سیگنال قبل و بعد از فیلتر کردن رسم می کنیم.

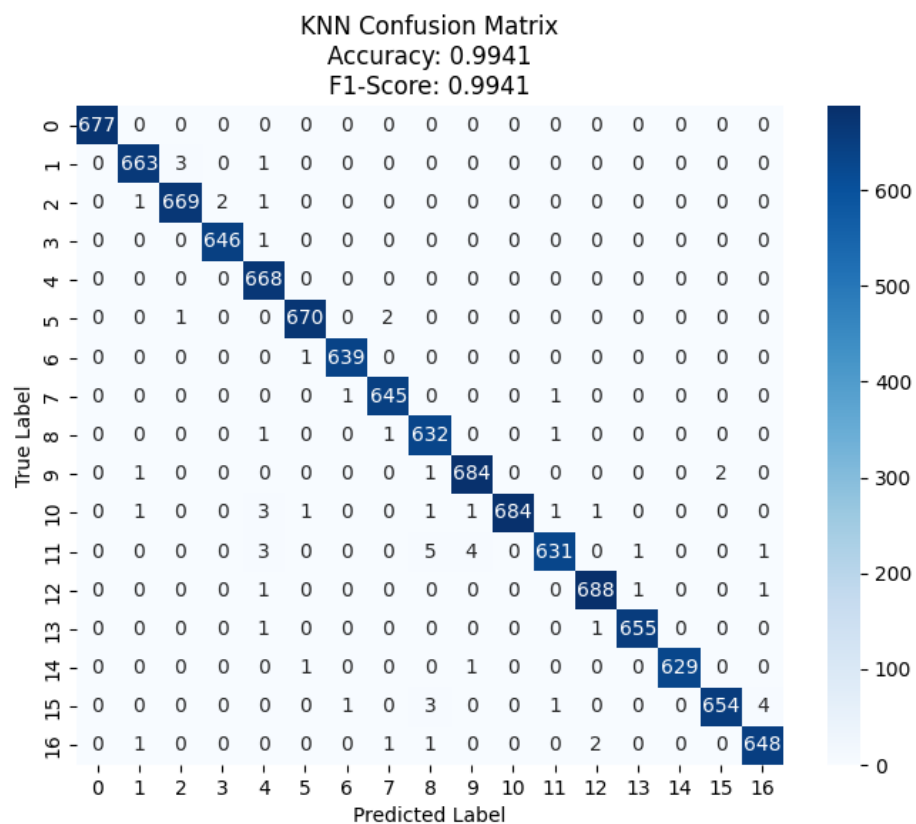


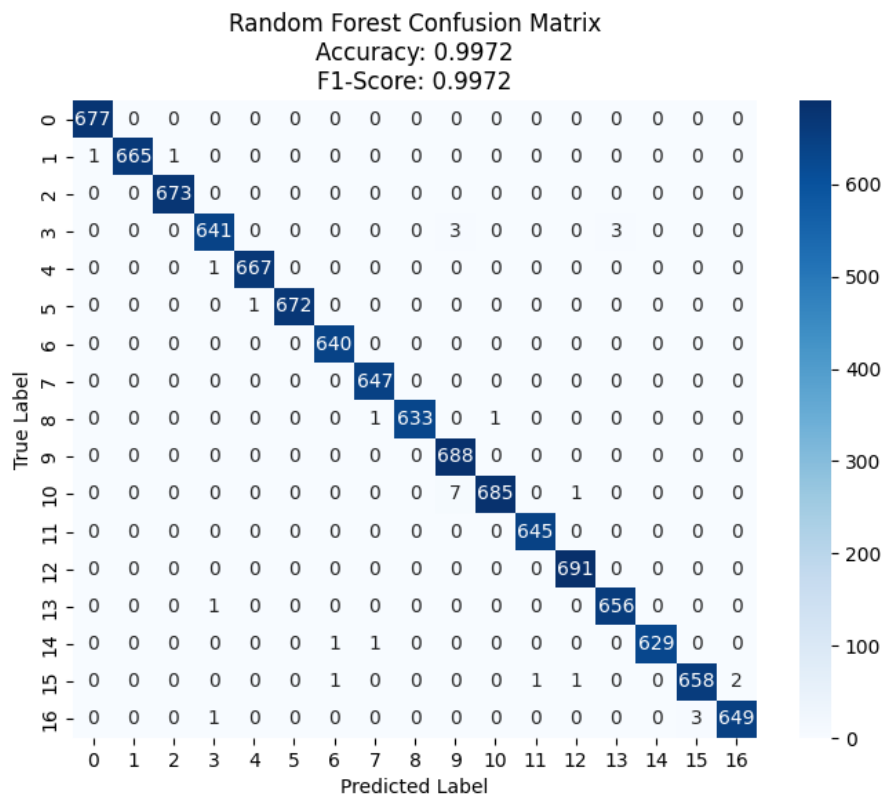
حال به منظور حذف نویز برق شهر در فرکانس 50 هرتز از notch filter استفاده کرده و نمودار سیگنال و اسپکتروم آن را پس اعمال این فیلتر نیز رسم می کنیم.





حال با استفاده از دو الگوریتم KNN و Random Forest و استفاده از تمامی ویژگی‌ها، هر 17 حرکت را طبقه بندی می‌کنیم.





همانطور که مشاهده می‌کنیم برای داوطلب شماره 13 نیز توانسته‌ایم به دقت بسیار خوبی دست پیدا کنیم و تقریباً تمامی داده‌ها توانسته‌اند به درستی طبقه بندی شوند.

کد همراه با این فایل آپلود شده است. همچنین لینک آن نیز در [این صفحه گیت‌هاب](#) موجود است. در این گزارش سعی کردیم تا همه افراد در بخش‌های مختلف دخیل باشند و به طور عمده بخش کد زنی توسط آقای شفیع زادگان و بخش گزارش توسط آقایان باقری و نوری انجام شده است.