فصل چهارم طراحی تکلیف، فرایند داده‌گیری، پایگاه‌داده و روش پردازش

# مقدمه

در این فصل ابتدا به شرح چگونگی و کیفیت طراحی تکلیف خواب و داده‌گیری در حین خواب پرداخته شده است. سپس ساختار پایگاه داده موجود که توسط این نرم‌افزار و دستگاه جمع‌آوری‌شده مورد بررسی قرار گرفته شده است. در مرحله بعد نیز به بیان روش‌های پیش‌پردازشی و پردازشی به کار گرفته شده در این پژوهش پرداخته می‌شود. نتایج پیاده‌سازی روش‌های پیش‌پردازش و پردازش در فصل بعد برسی می‌شود.

# طراحی تکلیف و دادگان تجربی

در این زیرفصل پس از معرفی دستگاه ثبت سیگنال‌های الکتروانسفالوگرام استفاده شده در حین خواب به بیان روند تکلیف و چگونگی آن پرداخته می‌شود. همچنین فرایند مرحله‌بندی بصری دادگان خواب توضیح داده می‌شود.

## مشخصات هدبند الکترواسنفالوگرام زی‌مکس

هدبند زی‌مکس شرکت هیپنوداین در سال ۲۰۱۸ به‌منظور ثبت سیگنال‌های مغزی در حین خواب در فضایی خارج از آزمایشگاه، به بازار عرضه شد. زی‌مکس در حال حاضر در دانشگاه‌ها و مؤسسات علمی مختلف در سراسر جهان در حال آزمایش است. این دستگاه از طریق دو حسگر جلویی که فعالیت مغز و حرکات چشمی را ثبت می‌کند، خواب را پایش می‌کند. برخلاف سایر دستگاه‌هایی که از حسگرهای EEG خشک استفاده می‌کنند، زی‌مکس از الکترودهای هیدروژل جامد یکبار مصرف اختصاصی استفاده می‌کند. علاوه بر این، شامل حسگرهای ضربان قلب (که از طریق فتوپلتیسموگرام[[1]](#footnote-1) یا PPG به دست می‌آید)، دما، نور محیط، صدا و حرکات بدن است. در این دستگاه به منظور دریافت دادگان و اعمال آیین‌گان محرک، نیاز به یک رایانه است که از طریق یک دانگل اتصال بی‌سیم به هدبند متصل شود. علی‌رغم اینکه زی‌مکس منبع‌باز نیست، امکان نوشتن توابع مختلف در جاوا اسکریپت برای محرک‌های متفاوت را فراهم می‌کند. زی‌مکس همچنین می‌تواند با زبان‌های برنامه‌نویسی مختلف (MATLAB، Python، PHP، C++، Java، و غیره) از طریق یک سوکت داده TCP/IP ارتباط برقرار کند.

TCP/IP، برای تسهیل ارتباط بین دو دستگاه جداگانه استفاده می‌شود. این رابط دستورالعمل‌هایی را برای مکان ارسال آن ارتباط ارائه می‌دهد. آدرس IP و آدرس پورت موجود در سوکت TCP/IP محل ارسال داده‌های درخواستی را به رایانه گیرنده می‌گوید. همچنین به ایجاد نوعی زبان یکنواخت بین ماشین‌ها کمک می‌کند. هنگامی که یک سوکت ایجاد می‌شود، فرستنده و گیرنده هر دو پورت ارتباطی را آزمایش می‌کنند تا اطمینان حاصل شود که داده‌ها از طریق آن عبور می‌کنند. فرستنده ممکن است درخواستی بفرستد که به گیرنده بگوید هویتش چیست. سپس پاسخی به فرستنده ارسال می‌شود که هویت فرستنده را تأیید می‌کند و درعین‌حال هویت گیرنده را نیز اعلام می‌کند. تا زمانی که ارتباط ادامه پیدا کند، درخواست‌ گیرنده برای دادگان بیشتر ادامه می‌یابد [1].

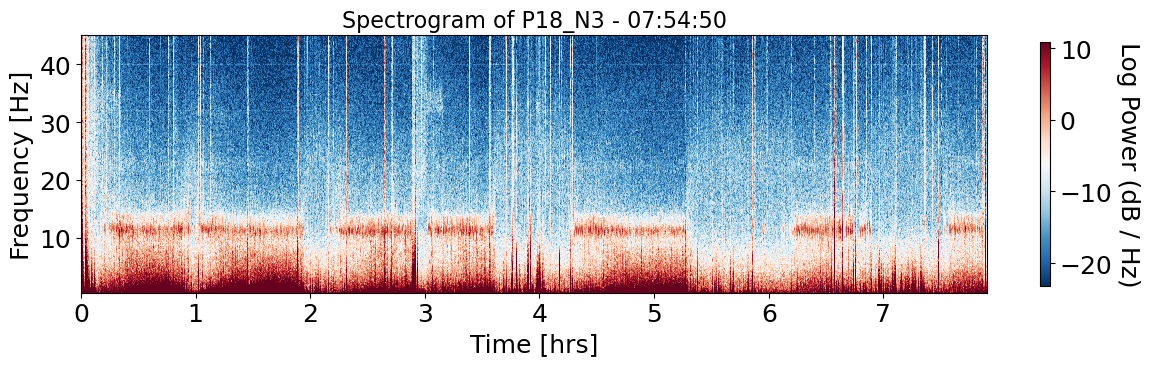
## پایگاه داده‌ی ثبت‌شده

برای دادگان تجربی، ۱۰ شرکت‌کننده راست‌دست سالم (۵ مرد و ۵ زن) با سنین ۱۹ تا ۳۹ سال و برای دو شب متوالی در آزمایشگاه خواب و حافظه داندرز خوابیدند. شرکت‌کنندگان برای اطمینان از برنامه خواب منظم و نداشتن سابقه اختلال خواب، مشکل روانپزشکی یا بیماری عصبی، و یا سابقه مصرف دخانیات یا داروهای تجویزی/تفریحی از قبل غربالگری شدند.

به منظور حذف شرکت‌کنندگان مبتلا به آپنه انسدادی خواب[[2]](#footnote-2) (OSA) یک شب قبل از آزمایش آزمایشگاهی پایش خانگی انجام گرفت. علاوه بر این، یک کارشناس باتجربه داده‌های پلی‌سومنوگرافی تجربی[[3]](#footnote-3) (PSG) را بعد از اولین شب آزمایشی امتیازبندی کرده تا از سلامت خواب شرکت‌کنندگان اطمینان حاصل شود. لازم به ذکر است که در این پژوهش آزمایش ادرار برای تشخیص مصرف دارو قبل از هر شب آزمایش و غربالگری بارداری برای شرکت‌کنندگان زن نیز انجام گرفت.

به‌منظور افزایش کیفیت داده‌گیری و ثبت در ابتدای آزمایش، پیشانی شرکت‌کنندگان توسط پد الکلی ضدعفونی شده و سپس هدبند الکتروانسفالوگرام زی‌مکس بر روی پیشانی آنان قرار داده می‌شود. اندازه و فشار بند دور سری هدبند باتوجه‌به اندازه سر شرکت‌کننده تنظیم می‌شود. به‌طوری که در طول شب فشاری روی پیشانی و دور سر حس نشود و نیز هدبند در اثر حرکات غیرارادی در هنگام خواب جدا نشود. این بدین منظور است که در وهله‌ی اول آرامش شرکت‌کننده در حین خواب مختل نشود و در وهله‌ی دوم کیفیت داده‌گیری در طول شب ثابت باقی بماند.

پس از ثبت داده، طیف‌نگار دادگان ثبت شده توسط برنامه نوشته شده در محیط پایتون نمایش داده شد تا از کیفیت ثبت اطمینان حاصل شود. برای رسم طیف‌نگار از روش تخمین طیف توان چند مخروطی که در فصل سوم شرح داده شد استفاده شده است [2]. نمونه‌ای از طیف‌نگار یک شب کامل خواب شرکت‌کننده ‌‌‌‌‌P8 در شکل 1 آورده شده است.

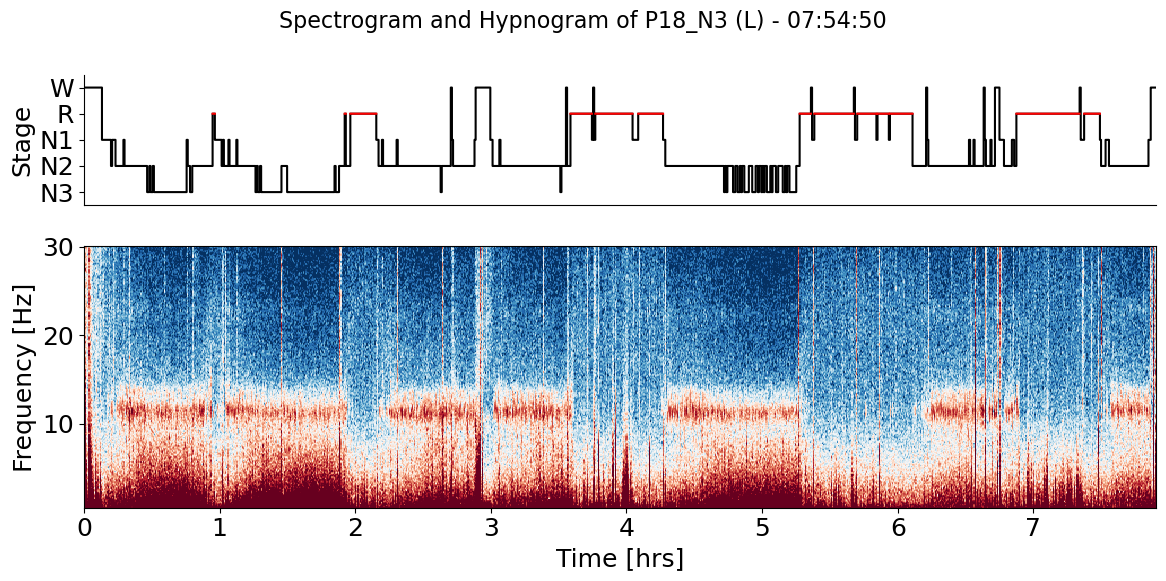


شکل 1 – طیف‌نگار به‌دست آمده از سیگنال الکتروانسفالوگرام یک شب خواب کامل (شرکت‌کننده P8)

## مرحله‌بندی بصری دادگان

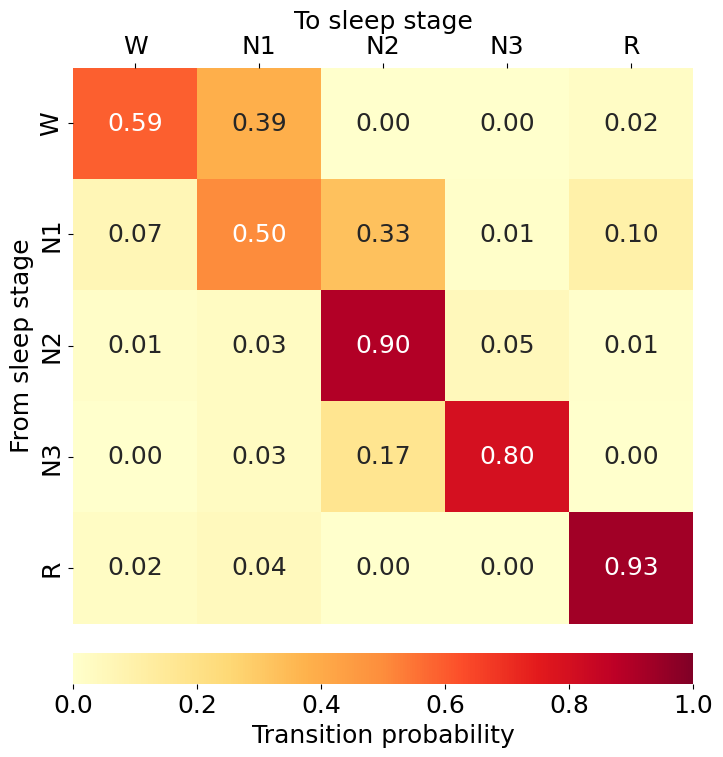
مرحله‌بندی بصری داده‌های خواب قبل از تجزیه‌وتحلیل آماری توسط یک کارشناس خواب بالینی باتجربه با استفاده از دستورالعمل‌های معاصر آکادمی پزشکی خواب آمریکا[[4]](#footnote-4) (AASM) [3] در بیمارستان دانشگاه رادبود هلند انجام شد و سپس نتایج آن با دادگان خواب (مطابق شکل 2) همگام‌سازی شد.

در فرایند مرحله‌بندی بصری خواب، هر ۳۰ ثانیه از دادگان خواب باتوجه‌ به ویژگی‌ها و محتویات فرکانسی آن به یکی از طبقات خواب منصوب می‌شود. این طبقات به‌صورت عددی از -2 تا 4 نمایش داده می‌شوند که به ترتیب بیانگر طبقه‌بندی‌نشده (-2)، نویز/حرکت (-1)، بیداری (0)، N1 (1)، N2 (2)، N3(3) و REM (4) است.



شکل 2 – طیف‌نگار همگام‌سازی شده با طبقات مختلف خواب (شرکت‌کننده P8)

برای اطمینان از مرحله‌بندی صحیح مراحل خواب توسط کارشناس‌، ماتریس انتقال خواب طراحی و پیاده‌سازی شد. این ماتریس، احتمال رخداد مراحل مختلف خواب را پس از مرحله‌ی کنونی نشان می‌دهد. پیاده‌سازی این ماتریس بر پایه مرحله‌بندی کارشناس است. یک نمونه از این ماتریس مربوط به شرکت‌کننده‌ی P18 در شکل 3 نشان داده شده است. در این ماتریس محور عمودی مرحله‌ی خواب اولیه و محور افقی مرحله‌ی خواب ثانویه است. مجموع مقادیر عناصر هر سطر برابر ۱ است. به‌صورت بدیهی عناصر روی قطر اصلی این ماتریس که نشان‌دهنده رخداد دوباره هر مرحله خواب است، دارای بیشترین مقادیر هستند. پس از آن نیز خط درایه‌هایی که بلافاصله بالای قطر اصلی قرار دارد دارای بیشترین مقادیر هستند. به‌این‌ترتیب با استناد بر این حقیقت فیزیولوژیکی که مراحل خواب به‌صورت متوالی رخ می‌دهند، می‌توان از صحت مرحله‌بندی خواب اطمینان حاصل کرد.



شکل 3 – ماتریس انتقال مراحل مختلف خواب برای شرکت‌کننده‌ی P8

# روش پیش‌پردازش

ثبت سیگنال EEG همواره با نویز و آرتیفکت همراه است. به این معنی که سیگنال ابتدایی ثبت‌شده از شرکت‌کنندگان که دادة خام نیز نام دارد، در وهلة اول به‌جز سیگنال حاصل از فعالیت مغز حاوی نویزها و آرتیفکت‌هایی با دو منشأ فیزیولوژی و غیرفیزیولوژی است. از جمله نویزهای غیرفیزیولوژی می‌توان به آثار برق شهر در سیگنال ثبت‌شده و از جمله آرتیفکت‌های غیر فیزیولوژی می‌توان به آثار سیگنال فعالیت الکتریکی چشم در سیگنال ثبت‌شده اشاره کرد [4].

از جمله اهداف پیش‌پردازش دادة خام، حذف نویزها و آرتیفکت‌های مذکور در کنار اهداف دیگری همچون کاهش هزینة محاسباتی در مرحلة پردازش است تا پس از گذراندن این مراحل، سیگنال آمادة پردازش نهایی شود. در ادامه مراحل و نحوة انجام پیش‌پردازش سیگنال‌ها بررسی می‌شود.

## اصلاح خط پایه

مرحله اول اصلاح خط پایه از سیگنال است. خط پایه یا بیس‌لاین را می‌توان یک آفست در نظر گرفت که با سیگنال هر کانال جمع شده است. ازآنجایی‌که سیگنال EEG یک سیگنال تعیین‌کننده زمان است، ممکن است اغلب دارای انحرافات زمانی باشد که با سوال تجربی ما ارتباطی ندارد. منابع داخلی و خارجی مختلف ممکن است باعث رانش‌های زمانی شوند که در طول زمان و همچنین در الکترودها تغییر می‌کنند. برای کاهش تأثیر این‌گونه دریفت‌ها، به‌اصطلاح اصلاح خط پایه انجام می‌شود. برای اصلاح خط پایه معمولاً مقادیر متوسط ولتاژ هر الکترود در یک بازه زمانی محاسبه می‌شود و سپس این میانگین از آن بازه زمانی سیگنال کم می‌شود [5].

## اعمال فیلتر روی سیگنال

فیلترینگ دیجیتال یک مرحله پیش‌پردازش رایج در هنگام تجزیه‌وتحلیل داده‌های EEG است. روش معمول در پردازش سیگنال EEG استفاده از یک فیلتر بالاگذر برای فیلترکردن فرکانس‌های آهسته کمتر از 0.1 هرتز یا اغلب حتی 1 هرتز و یک فیلتر پایین گذر برای فیلترکردن فرکانس‌های بالای 40 یا 50 هرتز است [2].

فیلترها را می‌توان به‌گونه‌ای طراحی کرد که دارای پاسخ ضربه محدود[[5]](#footnote-5) (FIR) یا پاسخ ضربه نامحدود[[6]](#footnote-6) (IIR) باشد. پاسخ ضربه خروجی (یا پاسخ) یک سیستم است هنگامی که ورودی سیستم، تابع دلتای دیراک است. از آنجا که تابع ضربه همه فرکانس‌ها را داراست، پاسخ ضربه تعیین‌کننده پاسخ یک سیستم خطی تغییرناپذیر با زمان به همه فرکانس‌ها است. تبدیل فوریه آن به‌عنوان پاسخ فرکانسی شناخته می‌شود. دانستن پاسخ ضربه می‌تواند در مشخص‌کردن رفتار فیلتر به محقق کمک کند.

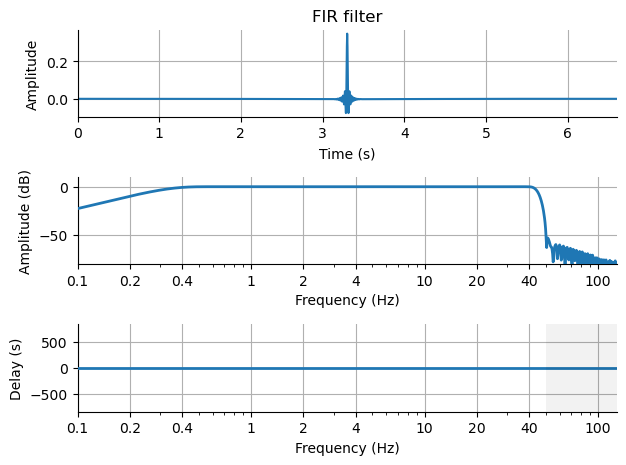
یک فیلتر FIR، همان‌طور که از نامش پیداست، یک پاسخ ضربه‌ای برای مدت‌زمان محدودی دارد، پس از آن خروجی به صفر می‌رسد و تاخیر‌های مساوی در تمام فرکانس‌ها ایجاد می‌کند (که به‌عنوان پاسخ فاز خطی نیز شناخته می‌شود). در مقابل، فیلترهای IIR (همچنین به‌عنوان فیلترهای بازگشتی شناخته می‌شوند) یک پاسخ ضربه‌ای نامحدود دارند که در آن بخشی از خروجی فیلتر به‌عنوان بازخورد استفاده می‌شود. این باعث ایجاد تأخیرهای نابرابر در فرکانس‌های مختلف می‌شود (مشخصات فاز غیرخطی). این بدان معناست که سیگنال خروجی باتوجه به ورودی با جابجایی بیشتر برخی از اجزای فرکانس نسبت به سایرین در زمان جابه‌جا می‌شود. بااین‌حال، مزیت اصلی فیلترهای IIR این است که از نظر محاسباتی کارآمدتر هستند.

یکی دیگر از جنبه‌های طراحی فیلتر جهت سیگنال استفاده شده به‌عنوان ورودی است. فیلترهایی که فقط شامل اطلاعات گذشته و حال هستند به‌عنوان فیلترهای علّی شناخته می‌شوند درحالی‌که فیلترهایی که به ورودی گذشته و آینده وابسته هستند به‌عنوان فیلتر غیرعلی شناخته می‌شوند. یک فیلتر علّی پس از شروع پاسخ تولید می‌کند (t=0). فیلترهای غیرعلّی حتی قبل از شروع به دلیل فیلترکردن به سمت عقب پاسخ می‌دهند و همچنین لوب‌های جانبی بزرگ‌تری تولید می‌کنند. فیلتر غیرعلّی در عمل با فیلترکردن داده‌ها یک‌بار به جلو و سپس دوباره به عقب به دست می‌آید و می‌تواند در به‌حداقل‌رساندن تأخیر فاز در سیگنال کمک کند. ازآنجایی‌که فیلتر غیرعلّی به داده‌های کامل نیاز دارد، این کار را نمی‌توان به‌صورت بلا‌درنگ (مانند فیلتر علی) انجام داد؛ لیکن یک عملیات برون‌خط است که پس از جمع‌آوری کامل داده انجام می‌شود.

در این پژوهش از یک فیلتر FIR با مشخصات زیر برای پیش‌پردازش سیگنال‌های EEG استفاده شده است.

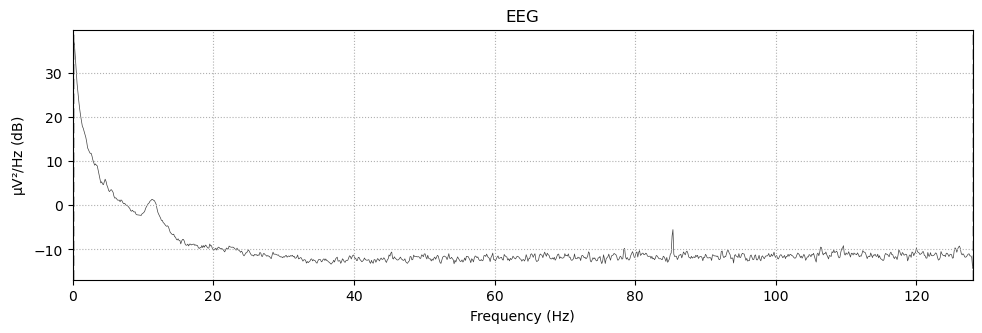
* یک فیلتر FIR با فاز صفر، غیرعلی و میان‌گذر
* طراحی پنجره حوزه زمانی با روش firwin
* پنجره همینگ با ریپل باند عبور 0.0194 و تضعیف باند توقف 53 دسی‌بل
* لبه پایین باند عبور (فرکانس قطع پایین): 0.50 هرتز
* لبه بالای باند عبور (فرکانس قطع بالا): 45 هرتز
* پهنای باند انتقال پایین: 0.50 هرتز (فرکانس قطع ۶- دسی‌بل: 0.25 هرتز)
* پهنای باند انتقال بالا: 11.25 هرتز (فرکانس قطع ۶- دسی‌بل: 50.62 هرتز)
* طول (مرتبه) فیلتر: 1691 نمونه (6.605 ثانیه)

مشخصات زمانی و فرکانسی پاسخ ضربه این فیلتر را می‌توان در شکل 4 مشاهده کرد.

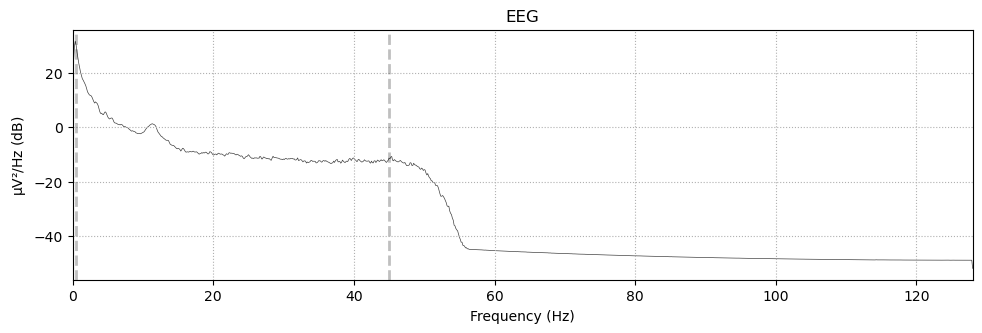


شکل 4 - مشخصات زمانی و فرکانسی پاسخ ضربه فیلتر اعمال شده بر روی دادگان الکتروانسفالوگرام

طیف توان برای یک سری زمانی توصیف کننده توزیع توان به مولفه‌های فرکانسی تشکیل‌دهنده سیگنال است. بر اساس آنالیز فوریه، هر سیگنال فیزیکی را می‌توان به تعدادی از فرکانس‌های گسسته، یا طیفی از فرکانس‌ها در یک محدوده پیوسته تجزیه کرد. میانگین آماری یک سیگنال معین یا نوعی از سیگنال (شامل نویز) وقتی که از دیدگاه محتوای فرکانس آن تحلیل می‌شوند، طیف آن سیگنال نامیده می‌شوند. یک نمونه از طیف توان سیگنال مربوط به شرکت‌کننده‌ی ۱۸ قبل و بعد از اعمال فیلتر در شکل ۵ و ۶ نشان داده شده است.



شکل 5 – طیف توان سیگنال پیش از اعمال فیلتر میانگذر ۰.۵ تا ۴۵ هرتز



شکل 6 - طیف توان سیگنال پس از اعمال فیلتر میانگذر ۰.۵ تا ۴۵ هرتز

## دوره‌بندی زمانی[[7]](#footnote-7) (ایپاک کردن) سیگنال

از زمان اولین ثبت الکتروانسفالوگرام (EEG)، دانشمندان به دنبال توصیف الگوهای تکرارشونده پیچیده فعالیت عصبی مشاهده شده در طی ساعات طولانی در طول خواب بوده‌اند. در درجه اول، این الگوها از طریق فرایند مرحله‌بندی خواب سازمان‌دهی شده‌اند که با استفاده از طبقه‌بندی مبتنی بر قانون خواب، زنجیره خواب را به مجموعه‌ای از مراحل مجزا تقسیم می‌کند. این مراحل از طریق بازرسی بصری EEG خواب و سایر شکل‌های موج فیزیولوژیکی در دوره‌های مجزای 30 ثانیه یا پنجره‌های زمانی تعریف می‌شوند [2].

کتابچه راهنمای آکادمی پزشکی خواب آمریکا همچنین امتیازدهی به خواب را در دوره‌های متوالی 30 ثانیه‌ای توصیه می‌کند. این مدت‌زمان برای امتیازدهی دستی مقرون‌به‌صرفه است و به طور کاملاً دقیق، دوره زمانی و ساختار کلان خواب معمولی را بدون خطر جاب‌جایی‌های بیش از حد مرحله خواب منعکس می‌کند [3].

به این منظور تابعی نوشته شد که سیگنال خواب ثبت‌شده، فرکانس نمونه‌برداری، و طول ایپاک به واحد ثانیه را به صورت ورودی دریافت کند و آن را به ماتریسی تبدیل کند که هر سطر بیان‌گر یک ایپاک می‌باشد. برای مثال یک ساعت داده خواب که به‌صورت یک ماتریس ۱ در 921600 درایه‌ای بود را به ماتریسی 120 در 7680 تبدیل می‌کند. این بدان معنی است که یک ساعت خواب دارای ۱۲۰ ایپاک ۳۰ ثانیه بوده و هر ایپاک دارای 7680 نمونه داده الکتروانسفالوگرام است.

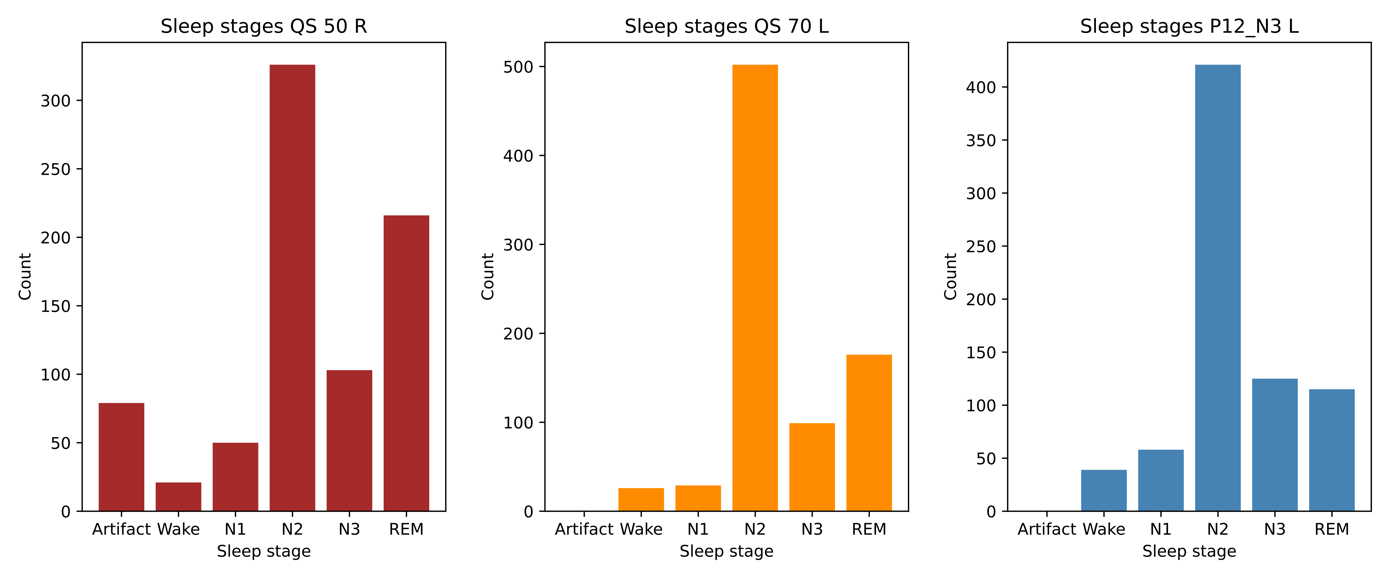
## حذف ایپاک های بد

عواملی مانند حرکت شرکت کننده یا قطع و وصل شدن رابط دستگاه ثبت و رایانه باعث اعوجاج سیگنال های ثبت شده الکتروانسفالوگرام می‌شود. این سیگنال ها مصنوع یا آرتیفکت نامیده می‌شوند. این مصنوعات با برسی چشمی ایپاک های ثبت شده قابل تشخیص هستند. پس از اعمال فیلتر و دوره بندی زمانی سیگنال، ایپاک هایی که دارای آرتیفکت هستند تشخیص داده شده و به آنها برچسب (کلاس) ۱- نسبت داده می‌شود. به این ترتیب ایپاک هایی که دارای برچسب ۱- هستند در مراحل بعدی پردازش و طبقه‌بندی اعمال نمی‌شوند.

## افزایش دادگان

هر کدام از مراحل مختلف خواب دارای مدت زمان متفاوتی است که باعث بوجود آمدن چالشی در طبقه‌بندی خودکار خواب می‌شود. شکل زیر نمونه ای از توزیع مراحل خواب در سه جلسه ثبت مختلف را نشان می‌دهد. عدم تعادل توزیع برچسب ها یک مشکل رایج هنگام کار با داده های واقعی است و زمانی رخ می‌دهد که نمونه هایی از یک کلاس یا چندین کلاس بیش از حد در یک مجموعه داده حضور داشته باشند.

تصور کنید برچسب ایپاک های بدست آمده از دادگان ثبت شامل بیش از ۸۰ درصد طبقه N2 است. در این مورد، ما می‌توانیم با پیش‌بینی برچسب اکثریت (طبقه N2 خواب) برای همه نمونه‌ها، بدون کمک الگوریتم یادگیری ماشین نظارت شده، به دقت بیش از ۸۰ درصدی در مجموعه داده‌های آزمایشی دست یابیم. بنابراین، آموزش مدلی بر روی چنین مجموعه داده‌ای که تقریباً ۸۰ درصد دقت آزمون را به دست می‌آورد به این معنی است که مدل ما هیچ چیز مفیدی از ویژگی‌های ارائه‌شده در این مجموعه‌داده نیاموخته است.

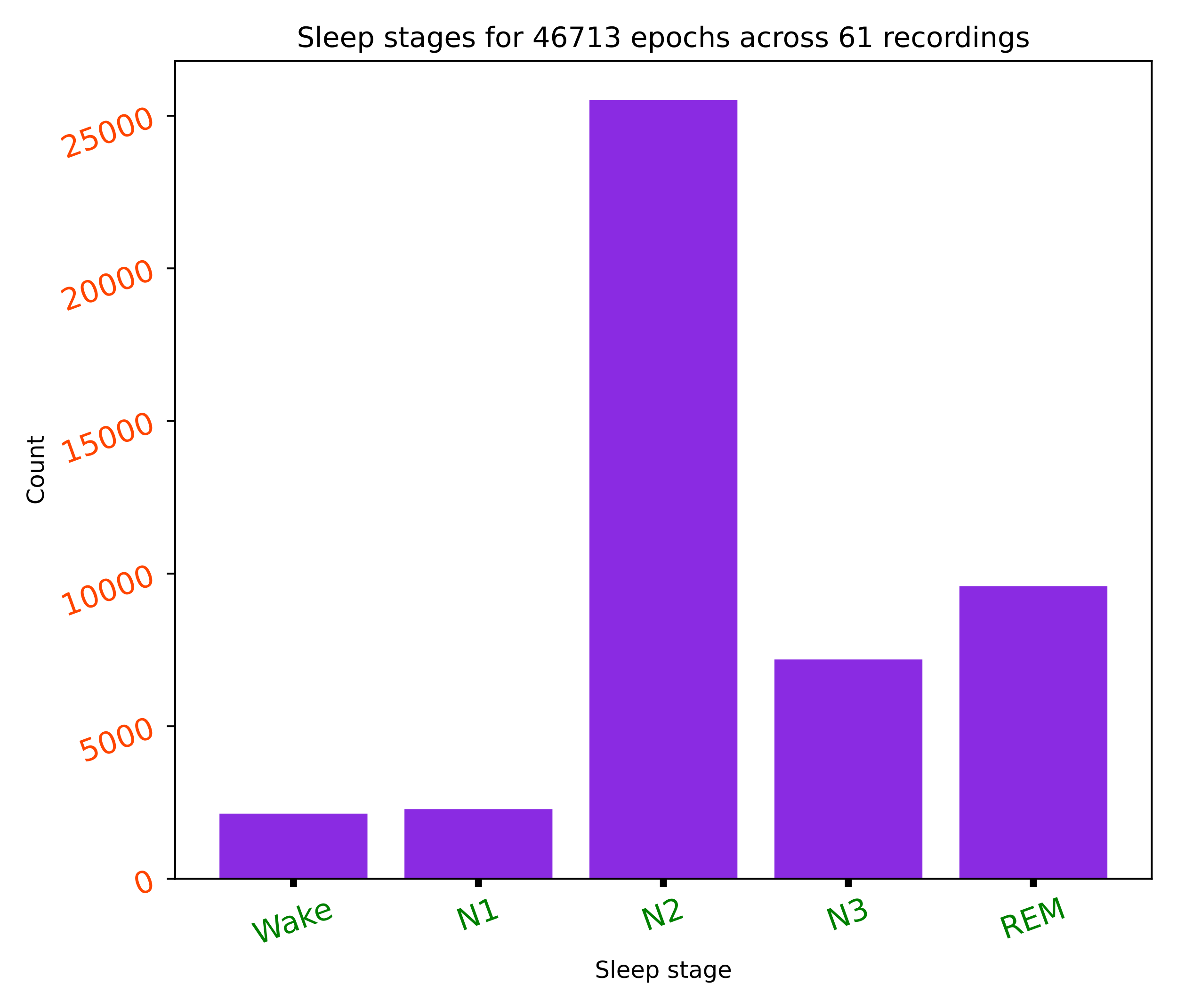


شکل 7 – توزیه مراحل مختلف خواب در سه جلسه ثبت متفاوت

نمودار جعبه ای در شکل زیر نشان دهنده وجود این ناهماهنگی در میان تمام جلسات ثبت است. نقاط یک‌رنگ نشان دهنده دفعات تکرار برچسب ها در طول یک شب ثبت است. در مجموع ۶۱ رنگ موجود است و هر رنگ نماینده یک جلسه ثبت است. میانه دفعات تکرار هر برچسب در طول جلسات مختلف ثبت توسط خط زرد رنگ در وسط نمودار جعبه ای نمایش داده شده است. پر واضح است که برچسب N2 دارای بیشترین تکرار بین مراحل خواب در اکثر جلسات است. در نهایت تعداد دفعات تکرار هر مرحله خواب در مجموع ۶۱ جلسه ثبت در شکل زیر نمایش داده شده است. این نمودار نیز تکرار بیش از حد مرحله خواب N2 را تایید میکند.

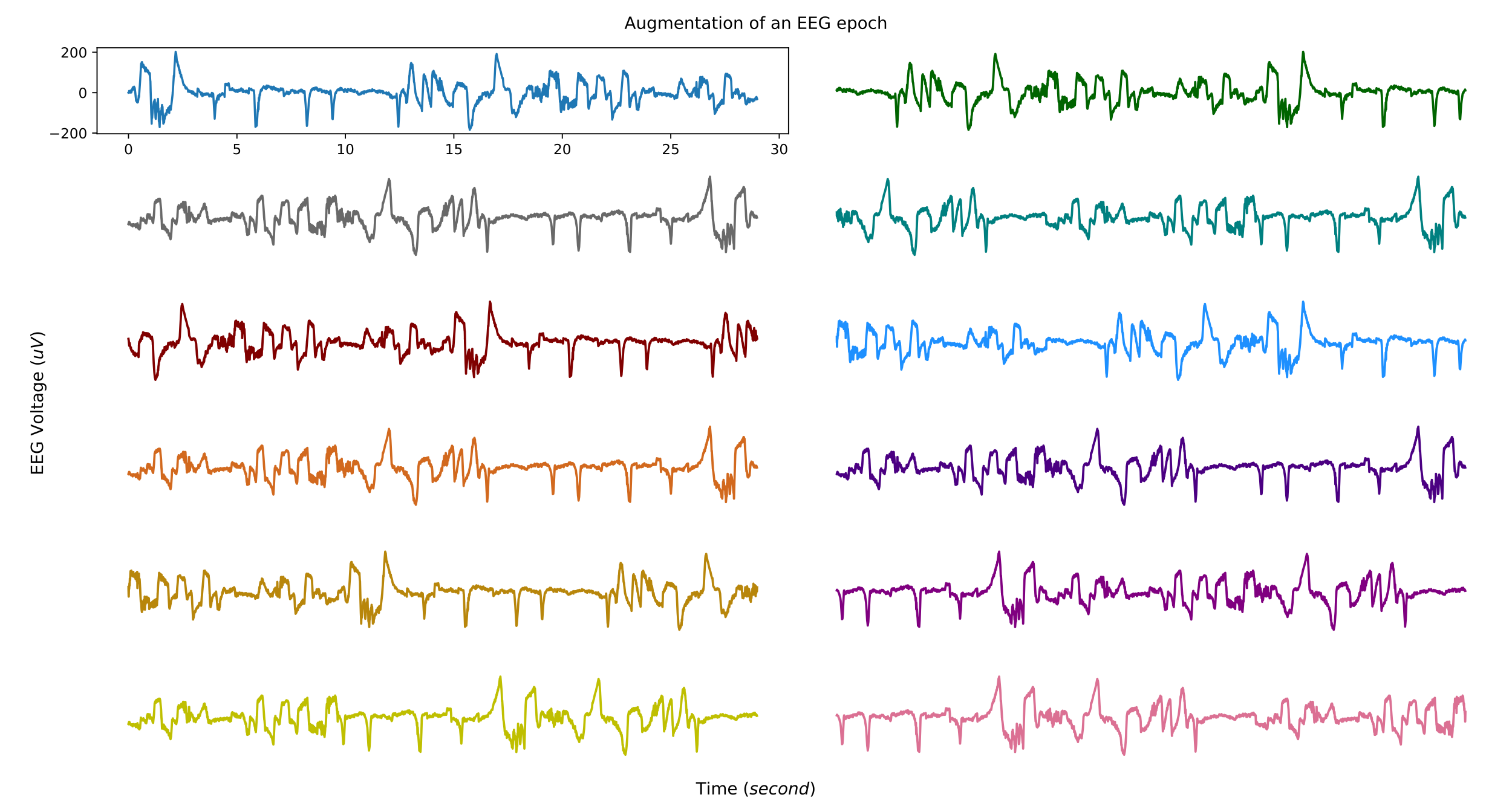


شکل 8 - نمودار جعبه ای تعداد دفعات تکرار هر مرحله خواب در هر ۶۱ جلسه ثبت



شکل 9 - نمودار هیستوگرام تعداد دفعات تکرار هر مرحله خواب در هر ۶۱ جلسه ثبت

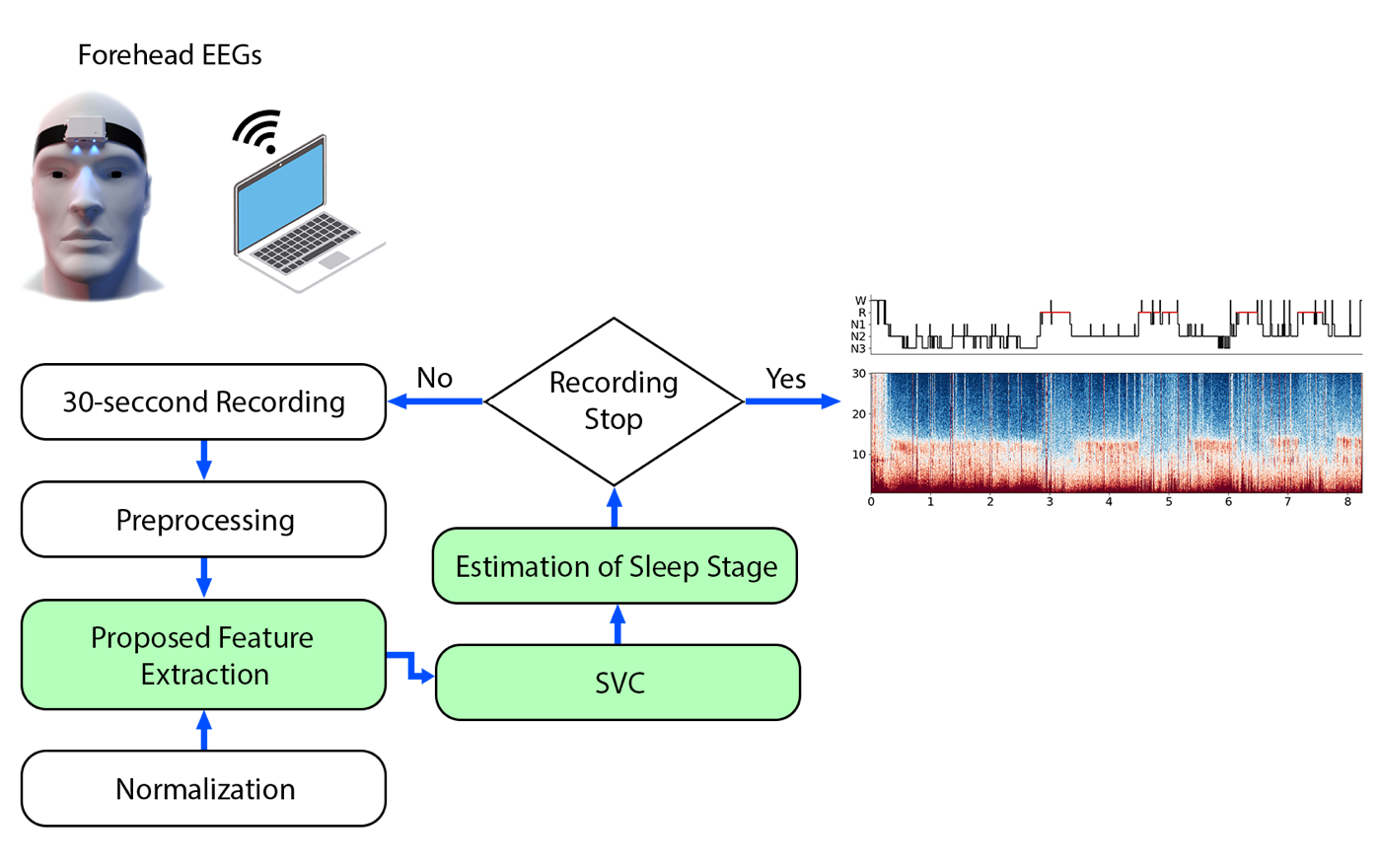
وجود این مشکل باعث شد تا از روش‌های افزایش دادگان[[8]](#footnote-8) برای تولید داده های خواب جدید برای ایپاک هایی با برچسب کم تکرار استفاده کنیم. روش های متفاوتی برای افزایش دادگان وجود دارد. یکی از این روش ها که برای دادگان الکتروانسفالوگرام خواب مناسب است روش **برش و جابه‌جایی** است. این روش با تقسیم سیگنال به قسمت های جدا از هم و جا‌به‌جایی این قسمت ها سیگنال جدیدی با برچسبی مشابه برچسب اولیه ایجاد می‌کند. به این ترتیب تمام جایگشت های موجود و معکوس آن ها (از انتها به ابتدا) نیز قابل استفاده هستند. در این روش ابتدا سیگنال الکتروانسفالوگرام هر ایپاک به دو قسمت مساوی تقسیم می‌شود. سپس ترتیب قسمت دوم و اول جابه‌جا شده و سیگنال جدیدی تولید می‌شود. همچنین با معکوس کردن سیگنال اولیه یا قسمت های بریده شده از انتها به ابتدا و تولید جایگشت های موجود ایپاک های جدیدی حاوی سیگنال های الکتروانسفالوگرام جدید با برچسب مورد نیاز تولید می‌شود. شکل زیر نشان دهنده تعدادی از رویکرد های مختلف افزایش داده است. سیگنال موجود در جعبه بالا سمت چپ سیگنال ایپاک الکتروانسفالوگرام اولیه است و سیگنال های دیگر از روی آن با روش برش و جابه‌جایی تولید شده اند.

****

شکل 10 - رویکرد های مختلف افزایش داده توسط روش برش و جا‌به‌جایی

# مراحل پردازش و استخراج ویژگی

در مرحلة پردازش با استخراج ۴۵ ویژگی از سیگنال که در ادامه به شرح هرکدام پرداخته خواهد شد، و پیاده‌سازی یک آزمون آماری برای پی‌بردن به معناداری آماری تفاوت میان این ویژگی‌ها در مراحل مختلف خواب مورد بررسی، امکان دسته‌بندی نهایی هر ایپاک با استفاده از یک روش یادگیری ماشین و مبتنی بر این ویژگی‌ها، بر اساس ۵ برچسب مراحل خواب فراهم می‌شود. خلاصه این مراحل را می‌توان در شکل زیر مشاهده کرد. قسمت های مشخص شده با رنگ سبز مراحل استخراج ویژگی، الگوریتم مرحله‌بندی و تعیین مرحله خواب می‌باشد.



شکل 11 – بلوک دیاگرام مراحل طبقه‌بندی خواب. قسمت های مشخص شده با رنگ سبز مراحل استخراج ویژگی، الگوریتم مرحله‌بندی و تعیین مرحله خواب می‌باشد.

شبکه عصبی مغز یک سیستم غیرخطی است، یعنی یک سیستم غیرثابت با رابطه غیرخطی بین علت و معلول. یکی از راه‌های تحلیل شبکه عصبی مغز، استخراج ویژگی از سیگنال EEG است. بیشترین تکنیک‌های استفاده شده برای استخراج ویژگی‌ها مانند تبدیل فوریه، خطی هستند. اگرچه این یک روش خطی است، اما تبدیل فوریه همچنین فرض می‌کند که دامنه همه اجزای فرکانس در طول زمان ثابت است که در مورد نوسانات مغزی صدق نمی‌کند، زیرا آن‌ها غیرایستا هستند. به دلیل پیچیدگی پویایی مغز، نیاز به روش‌های استخراج ویژگی وجود دارد که بتواند غیرخطی بودن و غیرثابت بودن پویایی مغز را به‌درستی در نظر بگیرد. با افزایش قدرت محاسباتی در سال‌های اخیر، بسیاری از محققان بر روی بهبود روش‌های استخراج ویژگی کار می‌کنند و تعداد فزاینده‌ای از ویژگی‌های مختلف استخراج‌شده از سیگنال EEG وجود دارد. در ادامه این بخش، از حروف پررنگ برای اولین‌بار استفاده از نام یک ویژگی خاص و حروف کج برای اولین بار استفاده از نام روش استخراج ویژگی استفاده خواهد شد. جدول زیر فهرستی از تمام ویژگی‌های معرفی شده را نشان می‌دهد.

جدول ۱ - فهرستی از ویژگی‌های حوزه زمان، فرکانس، زمان-فرکانس و ویژگی‌های غیرخطی مورد استفاده در این پژوهش

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| گروه | نام ویژگی | مخفف | گروه | نام ویژگی | مخفف |
| حوزه فرکانس | میانگین | Mean\_psd | حوزه زمان | میانگین | Mean |
| میانه | Median\_psd | میانه | Median |
| واریانس | Var\_psd | واریانس | Var |
| انحراف معیار | Std\_psd | انحراف معیار | Std |
| دامنه بین چالگی | Iqr\_psd | دامنه بین چالگی | Iqr |
| چولگی | Kurt\_psd | چولگی | Kurt |
| کشیدگی | Skew\_psd | کشیدگی | Skew |
| فعالیت هجورت | Hact\_psd | نرخ عبور از صفر | Nzc |
| پیچیدگی هجورت | Hcomp\_psd | فعالیت هجورت | Hact |
| تحرک هجورت | Hmob\_psd | پیچیدگی هجورت | Hcomp |
| دلتا | Delta | تحرک هجورت | hmob |
| تتا | Theta | انرژی | E |
| آلفا | Alpha | آنتروپی‌ها | آنتروپی شانون | shEnt |
| بتا | Beta | آنتروپی رنی | Renyi |
| گاما | Gamma | آنتروپی تالیس | tsallisEnt |
| سیگما | Sigma | آنتروپی طیفی | Spec\_entropy |
| ۳۶ نسبت باند ها | ... | آنتروپی پاسخ | Resp\_entropy |
|  | Ta\_b | آنتروپی حالت | Sate\_entropy |
|  | Ta\_ab | آنتروپی موجک | Wavelet\_entropy |
|  | Gb\_da | آنتروپی تقریبی | App\_entropy |
| غیرخطی | پیچیدگی لمپل-زیو | Lziv | آنتروپی نمونه | Samp\_entropy |
| تجزیه و تحلیل نوسانات بدون روند | Dfa | آنتروپی حبابی | bubbleEnt |
| بعد فراکتال پتروسین | Petrosian | آنتروپی دیفرانسیلی | diffEnt |
| بعد فرکتال هیگوچی | Higuchi | آنتروپی جایگشت | Perm\_entropy |
| بعد فرکتال کاتز | Katz | آنتروپی پراکندگی | Disp\_entropy |
| اندازه‌گیری گرایش مرکزی | CTM | آنتروپی فازی | FuzzEnt |
| ضرایب هرست | Hurt\_Xi | آنتروپی حبابی | bubbleEnt |

## ویژگی‌های حوزه‌ی زمان، فرکانس و زمان - فرکانس

### ویژگی‌های حوزه زمان

ساده‌ترین ویژگی‌های سیگنال EEG ویژگی‌های آماری مانند **میانگین**[[9]](#footnote-9)، **میانه**[[10]](#footnote-10)، **واریانس**[[11]](#footnote-11)، **انحراف معیار**[[12]](#footnote-12)، **چولگی**[[13]](#footnote-13)، **کشیدگی**[[14]](#footnote-14) و موارد مشابه است [6]. **نرخ عبور از صفر**[[15]](#footnote-15) **(ZCR)** [7] یک ویژگی آماری نیست، اما یک ویژگی ساده است که تعداد دفعاتی که سیگنال از محور x عبور می‌کند را نشان می‌دهد. *تجزیه و تحلیل دامنه-پریود* بر اساس تجزیه و تحلیل امواج نیمه، یعنی سیگنال‌های بین دو تقاطع صفر است. با تجزیه و تحلیل دامنه-پریود، می توان **تعداد امواج**[[16]](#footnote-16)، **مدت زمان موج**[[17]](#footnote-17)، **دامنه قله**[[18]](#footnote-18) و **فرکانس لحظه‌ای**[[19]](#footnote-19) را استخراج کرد (فقط بر اساس نیم موج مشاهده شده منفرد) [8].

پارامترهای هجورت[[20]](#footnote-20) ویژگی‌هایی هستند که بر اساس واریانس مشتقات سیگنال EEG به دست می‌آیند. **تحرک**[[21]](#footnote-21)، **فعالیت**[[22]](#footnote-22) و **پیچیدگی**[[23]](#footnote-23) [9] سه مشتق اول سیگنال و پرکاربردترین پارامترهای هجورت هستند. **میانگین قدر مطلق تحرک**، **فعالیت** و **پیچیدگی** نیز می‌تواند به عنوان یک ویژگی استفاده شود [10].

**K-Complex**[11] شکل موج مشخصی از سیگنال EEG است که در مرحله دوم از مراحل خواب حرکت غیرسریع چشم رخ می‌دهد. **انرژی** **سیگنال (E)** مجموع مربع‌های دامنه سیگنال است.

### ویژگی‌های حوزه فرکانس

چگالی طیفی توان[[24]](#footnote-24) (PSD) سیگنال که مبنایی برای محاسبه ویژگی‌های حوزه فرکانس است، با چندین روش پارامتری و غیرپارامتری قابل محاسبه است. روش‌های غیرپارامتری بیشتر مورداستفاده قرار می‌گیرند و شامل روش‌هایی مانند تبدیل فوریه (معمولاً با الگوریتم تبدیل فوریه سریع [12] محاسبه می‌شود)، روش ولش[[25]](#footnote-25) [13]، یا روش چندمخروطی تامپسون[[26]](#footnote-26) [14] است. نمونه‌هایی از روش‌های پارامتری برای تخمین PSD، مدل‌های خودبازگشتی [15]، مدل‌های خودبازگشتی چند متغیره [16] یا مدل‌های میانگین متحرک خودبازگشتی[[27]](#footnote-27) (ARMA) هستند. مدل‌های غیرپارامتری کاربرد گسترده‌تری دارند، زیرا نیازی به انتخاب پارامترهایی مانند ترتیب مدل، که در مدل‌های خودبازگشتی وجود دارد، نیست.

ویژگی‌های آماری مانند **میانگین**، **میانه**، **واریانس**، **انحراف معیار**، **چولگی**، **کشیدگی** و موارد مشابه نیز در حوزه فرکانس استفاده می‌شود. **توان‌های نسبی باندهای فرکانسی خاص**، پراستفاده‌ترین ویژگی‌های حوزه فرکانس در تمام زمینه‌های تجزیه‌وتحلیل سیگنال‌های EEG هستند. متداول‌ترین باندهای فرکانسی **دلتا** (δ, 0.5-4 هرتز)، **تتا** (θ، 4-8 هرتز)، **آلفا** (α، 8-12 هرتز)، **بتا** (β، 16-30 هرتز) و **گاما** (γ، 30-40 هرتز) هستند. همچنین یک باند موسوم به **نوار سیگما** (σ، 12-16 هرتز) وجود دارد که گاهی اوقات **دوک‌های خواب** نامیده می‌شود. چندین **نسبت بین باندهای فرکانس** به طور گسترده به‌عنوان ویژگی در تجزیه‌وتحلیل سیگنال EEG استفاده می‌شود، به‌عنوان‌مثال ، ، ، ، ، و [17]–[19]**.** همچنین تمامی ۳۰ جای‌گشت موجود از نسبت دو باند فرکانسی متفاوت به عنوان ویژگی مستقل در نظر گرفته شده است.

دامنه حوزه فرکانس سیگنال را نیز می‌توان با استفاده از *روش‌های تجزیه موجک* [16] به دست آورد. برخلاف تبدیل فوریه، که سیگنال را به سینوسی تجزیه می‌کند، تجزیه موجک از یک تابع موجک مادر برای تجزیه استفاده می‌کند.

از مدل‌های خودبازگشتی می‌توان ویژگی‌هایی مانند **ضرایب بازتاب**[[28]](#footnote-28) یا **ضرایب همبستگی** **جزئی**[[29]](#footnote-29) را استخراج کرد. **ضرایب موجک** به‌دست‌آمده پس از اعمال تجزیه موجک نیز می‌تواند به‌عنوان ویژگی استفاده شود. PSD معمولاً برای به‌دست‌آوردن آماره‌های مرتبه دوم سیگنال EEG استفاده می‌شود. بااین‌حال، می‌توان طیف‌های مرتبه بالاتر را نیز در نظر گرفت. به‌عنوان‌مثال، **جفت‌شدگی فاز** [20] اجزای فرکانسی مختلف را می توان با تحلیل طیفی مرتبه بالاتر به دست آورد.

### ویژگی‌های زمان - فرکانس

تجزیه‌وتحلیل سیگنال EEG در حوزه زمان و فرکانس به طور هم‌زمان یک ابزار قدرتمند است، زیرا سیگنال EEG یک سیگنال غیرایستا است [21], [22]. مهم‌ترین مؤلفه تحلیل حوزه زمان-فرکانس، امکان مشاهده تغییرات فرکانس در طول زمان است. تبدیل فوریه کوتاه مدت (STFT) ساده‌ترین ابزاری است که از جداسازی یکنواخت سیگنال مشاهده شده استفاده و اجزای فرکانس آن را محاسبه می‌کند. یک *طیف نگار* [21] را می توان با استفاده از STFT به دست آورد. تبدیل موجک [23] روش جایگزین معمولی برای طیف‌نگار است که ضرایبی را نیز به‌عنوان ویژگی‌هایی از حوزه زمان-فرکانس ارائه می‌کند. مزیت اصلی تبدیل موجک در مقایسه با طیف‌نگار، اندازه پنجره متغیر است که به فرکانس‌های طیف بستگی دارد.

## ویژگی‌های غیرخطی

پویایی مغز یک سیستم پیچیده را تشکیل می‌دهد. یک سیستم، زمانی پیچیده است که از زیرسیستم‌های غیرخطی بسیاری ساخته شده باشد که نمی‌توان آن‌ها را بدون تغییر ویژگی‌های دینامیکی‌شان به زیرسیستم‌های کوچک‌تر تقسیم کرد. سیستم‌های فراکتال اغلب برای توصیف پویایی و دینامیک مغزی اندازه‌گیری شده با سیگنال EEG استفاده می‌شوند. برای توضیح سیستم‌های فراکتال ابتدا باید قانون مقیاس‌بندی را معرفی کرد. قانون مقیاس‌بندی یک تابع خود مشابه را به‌عنوان تابعی از پارامتر مقیاس (به‌طور بدون علامت) توصیف می‌کند، به عنوان مثال، . هنگامی که به یک سیگنال خود وابسته اعمال می شود، هر محور باید با یک ضریب توان متفاوت مقیاس شود تا تغییرات آماری معادل در هر دو جهت به دست آید. اگر در جهت محور استفاده شود، پس باید در جهت محور استفاده شود. ضریب توان را **توان هرست**[[30]](#footnote-30) می نامند. توان هرست یک ابزار برای اندازه‌گیری حافظه بلندمدت سیگنال است و با بعد فراکتال با معادله برای سری‌های زمانی خود‌متشابه مرتبط است.

سری زمانی اگر به صورت خطی با **توان مقیاس بندی رنی**[[31]](#footnote-31) خود وابسته باشد، منوفرکتال است، در غیر این صورت، چندفراکتال است. بعد **تعمیم‌یافته رنی چند فراکتال**[[32]](#footnote-32) به صورت رابطه‌ی ۱ تعریف می‌شود. برای توضیحات بیشتر در مورد فرکتال بودن و چندفراکتال بودن یک سری زمانی، به [24]–[26] رجوع شود.

|  |  |
| --- | --- |
| (رابطه ۱) |  |

در تجزیه‌وتحلیل سیگنال EEG، تمام ابعاد فراکتال بر اساس جاذبه زیرین (ساختار هندسی که سیستم اتلاف ساکن در فضای حالت خود به سمت آن می‌کشد) سیگنال تخمین زده می‌شود [27]. در یک مفهوم دقیق ریاضی، بیشتر سری‌های زمانی دارای بعد **فراکتال پشتیبان یک‌بعدی**  (یا بعد ظرفیت یا بعد هاوس‌دورف[[33]](#footnote-34)) هستند، اگر مقادیر گم‌شده‌ای وجود نداشته باشد. صرف نظر از مقدار ، بعد اطلاعات و بعد همبستگی قابل محاسبه است [27], [28]. بعد همبستگی D2 را می توان با هر دو روش تک فراکتال و چند فرکتال محاسبه کرد. **بعد فراکتال کاتز**[[34]](#footnote-35) **(KFD)** [28]، **بعد فراکتال پتروسین**[[35]](#footnote-36) (PFD) [25]، و **بعد فراکتال هیگوچی**[[36]](#footnote-37) **(HFD)** رویکردهای متفاوتی برای تخمین بعد فراکتالی هستند. با تجزیه و تحلیل سری های زمانی چندفرکتالی، یک **طیف فراکتالی** متشکل از ابعاد فراکتالی متعدد را می‌توان به دست آورد [29], [30].

روش‌های تجزیه‌وتحلیل سری‌های زمانی فراکتال را می‌توان به روش‌های تحلیل ایستا (مانند *تحلیل نوسانات*[[37]](#footnote-38) [31]، *آنالیز دامنه مقیاس‌گذاری‌شده هرست*[[38]](#footnote-39) [32] و موارد مشابه)، تجزیه و تحلیل غیر ایستا (مانند *تحلیل نوسانات کاهش‌یافته*[[39]](#footnote-40) [33]، *تجزیه و تحلیل میانگین متحرک مرکزی*[[40]](#footnote-41) [34]، *مجموع مساحت مثلث*[[41]](#footnote-42) [35] و موارد مشابه)، و تجزیه و تحلیل چندفرکتالی (مانند *مدول تبدیل موجک ماکسیما*[[42]](#footnote-43) [30]، *تحلیل نوسانی چندفراکتالی بدون روند*[[43]](#footnote-44) [31] و موارد مشابه) طبقه‌بندی کرد. هر یک از این روش‌ها تخمین خود را از ابعاد فراکتال یا ویژگی‌های توان مقیاس‌بندی ارائه می‌دهد.

**نماهای لیاپانوف**[[44]](#footnote-45) (LE) [36] معیارهای پیچیدگی جاذبه هستند. اگر یک سیستم حداقل یک توان لیاپانوف مثبت داشته باشد، آنگاه سیستم را می‌توان به عنوان یک سیستم پویای آشوبی توصیف کرد. یک توان مثبت لیاپانوف به واگرایی نمایی دو مسیر مجاور در جاذب در طول زمان اشاره می‌کند [37]. **پیچیدگی لمپل-زیو**[[45]](#footnote-46) **(LZC) [38]** اندازه‌گیری پیچیدگی است که سری‌های زمانی را باینری می‌کند و سپس وقوع کاراکترهای باینری متوالی یا «کلمات» را جستجو می‌کند، سپس تعداد دفعاتی را که یک «کلمه» جدید است شمارش می‌کند. **اندازه‌گیری گرایش مرکزی**[[46]](#footnote-47) **(CTM)** [39] اندازه‌گیری تغییرپذیری سری‌های زمانی مشاهده شده است و نشان‌دهنده درصد نقاطی در نمودار پراکندگی است که در یک شعاع معین قرار می‌گیرند. **اطلاعات متقابل خودکار**[[47]](#footnote-48) **(AMI)** [40]یک اندازه‌گیری اطلاعات متقابل است که برای نسخه‌های با تأخیر زمانی همان سری‌های زمانی EEG اعمال می‌شود. **برگشت‌ناپذیری زمانی**[[48]](#footnote-49) [41] یک سری زمانی بر تأثیر پویایی غیرخطی، نویز غیرگاوسی یا هر دو دلالت دارد. این یک ویژگی آماری است که بر اساس جهتی که زمان در آن طی می‌شود متفاوت است؛ به عنوان مثال، هر دنباله‍ای از اندازه‌گیری‌ها احتمال وقوع متفاوتی نسبت به زمان معکوس آن دارد.

*نمودار بازرخداد*[[49]](#footnote-50) [42] یک روش گرافیکی برای تشخیص الگوهای تکرارشونده در سری‌های زمانی است. *تحلیل کمیت تکراری*[[50]](#footnote-51) (RQA) گروهی از الگوریتم‌ها برای تعیین کمیت خودکار نمودارهای بازرخداد است. RQA یک روش مقاوم در برابر نویز است، به این معنی که حتی زمانی که نسبت سیگنال به نویز سیگنال‌های در نظر گرفته شده نامطلوب باشد، نتایج خوبی به دست می‌دهد [43]. **نرخ بازگشت**[[51]](#footnote-52) **(RR)** احتمال وقوع مجدد حالت خاصی از یک سری زمانی است. **جبر**[[52]](#footnote-53) **(Det)** درصد نقاطی است که خطوط مورب را در نمودار بازرخداد تشکیل می‌دهند و **لمیناریته**[[53]](#footnote-54) **(Lam)** درصد نقاط تشکیل خطوط عمودی در نمودار بازرخداد است. **میانگین طول خط مورب**[[54]](#footnote-55) **(L)**، **حداکثر طول مورب**[[55]](#footnote-56) **(Lmax)** و **حداکثر طول خطوط عمودی**[[56]](#footnote-57) **(Vmax)** نیز به عنوان ویژگی‌های مبتنی بر RQA استفاده می‌شود. **زمان به دام انداختن**[[57]](#footnote-58) **(TT)** میانگین طول خط عمودی است و به قابلیت پیش‌بینی سری‌های زمانی مربوط می‌شود. **واگرایی**[[58]](#footnote-59) **(Div)** مقدار متقابل حداکثر طول خط مورب است و می‌تواند روندی شبیه به شارهای لیاپانوف مثبت داشته باشد. **آنتروپی نمودار بازرخداد**[[59]](#footnote-60) **(ENTR)** پیچیدگی ساختار قطعی[[60]](#footnote-61) سیستم را منعکس می‌کند.

## آنتروپی‌ها

آنتروپی اولین‌بار توسط شانون در سال ۱۹۴۸ [44], [45] به حوزه نظریه اطلاعات معرفی شد. آنتروپی اطلاعات شانون بر اساس رابطه ۲ محاسبه می‌شود، که در آن توزیع احتمال داده‌های مشاهده‌شده است.

|  |  |
| --- | --- |
| (رابطه ۲) |  |

این برای اندازه‌گیری عدم قطعیت یا تصادفی بودن در سری های زمانی مشاهده شده استفاده می شود. بسیاری از مدل های مشتق شده از آنتروپی اطلاعات در تجزیه و تحلیل EEG استفاده می شود. آنتروپی ها ممکن است به عنوان ویژگی های غیر خطی در نظر گرفته شوند، اما در این پژوهش به دلیل نوع خاص محاسبه در یک زیربخش جداگانه توضیح داده می‌شوند.

**آنتروپی رنی**[[61]](#footnote-62) [46] با رابطه ۳ تعریف می شود:

|  |  |
| --- | --- |
| (رابطه ۳) |  |

که در آن و . این تعمیم آنتروپی شانون در مورد مقدار محدود است. **آنتروپی درجه دوم رنی** (یا فقط آنتروپی رنی) موردی است که . **آنتروپی *تسلیس***[[62]](#footnote-63) ***(آنروپی کیو***[[63]](#footnote-64)***)*** [47] تعمیم آنتروپی بولتزمن-گیبس[[64]](#footnote-65) از ترمودینامیک آماری است و با رابطه ۴ تعریف می شود:

|  |  |
| --- | --- |
| (رابطه ۴) |  |

که در آن k یک ثابت مثبت و q پارامتر غیر وسعت[[65]](#footnote-66) است. برای ، آنتروپی واکنش معنی دارتری به رویدادهایی دارد که اغلب رخ می دهند، در حالی که برای ، آنتروپی واکنش معنی دارتری به رویدادهای نادر دارد.

سه آنتروپی فوق را می‌توان از سیگنال EEG خام محاسبه کرد. علاوه‌براین، آن‌ها مبنایی برای محاسبه چندین آنتروپی دیگر در زمینه آنالیز EEG هستند. **آنتروپی کراسکوف**[[66]](#footnote-67) **(KE)** [6] یک برآوردگر بی‌طرفانه از آنتروپی شانون برای یک نمونه تصادفی d بعدی است. **آنتروپی طیفی**[[67]](#footnote-68) **(SEN)** [48] با رابطه آنتروپی شانون بر اساس PSD نرمال شده سیگنال EEG محاسبه می شود. **آنتروپی طیفی رنی درجه دوم**[[68]](#footnote-69) **(QRSEN)** [49] با استفاده از رابطه آنتروپی رنی محاسبه می‌شود و تفاوتش در مقایسه با آنتروپی طیفی این است که وزن‌های بالاتری را به فرکانس‌های پایین‌تر می‌دهد. مدول تجاری M-Entropy[[69]](#footnote-70) [50] از دو جزء مختلف آنتروپی طیفی یعنی **آنتروپی پاسخ**[[70]](#footnote-71) **(RE)** و **آنتروپی حالت**[[71]](#footnote-72) **(SE)** استفاده می کند. آنتروپی حالت شامل طیف بین 0.8 تا 32 هرتز است، در حالی که آنتروپی پاسخ شامل طیف بین 0.8 تا 47 هرتز است.

**آنتروپی موجک (WE)** [51], [52] تا حدودی شبیه به آنتروپی طیفی است، با این تفاوت که بر اساس ضرایب تجزیه موجک سری زمانی داده شده محاسبه می شود. دو تعمیم آنتروپی موجک وجود دارد - **آنتروپی موجک *تسلیس***[[72]](#footnote-73) (TWE) و **آنتروپی موجک رنی**[[73]](#footnote-74)(RWE). **آنتروپی طیفی هیلبرت-هوانگ**[[74]](#footnote-75) (HHSE) آنتروپی شانون را برای طیف هیلبرت-هوانگ اعمال می کند که با *تبدیل هیلبرت-هوانگ* به دست می آید. **آنتروپی انرژی لاگ**[[75]](#footnote-76) (LogEn) مشابه آنتروپی موجک است، اما فقط از جمع لگاریتم احتمالات استفاده می کند. **آنتروپی چند رزولوشن**[[76]](#footnote-77)از ترکیب پنجره سازی و تبدیل موجک برای تشخیص تغییرات در پارامترهایی که فرآیند مشاهده شده را تعریف می کنند (یعنی پارامترهای پویایی مغزی) استفاده می کند.

**آنتروپی تقریبی (ApEn)** [53] از آنتروپی کولموگروف مشتق شده است و استفاده از آن در تجزیه و تحلیل سیگنال EEG (و سایر سیگنال های فیزیولوژیکی) گسترده است. این ویژگی به کمّی‌سازی بی نظمی یک سری زمانی می‌پردازد. سری‌های زمانی قابل پیش‌بینی، یعنی سری‌های زمانی با الگوهای تکراری زیاد، مقدار کمی آنتروپی تقریبی دارند. **آنتروپی نمونه (SampEn)** [54] به عنوان بهبودی برای آنتروپی تقریبی معرفی شده که خطای آنتروپی تقریبی را با حذف دو ایراد آن کاهش می دهد - (1) خود انطباقی[[77]](#footnote-78) و (2) وابستگی به طول سری زمانی. آنتروپی نمونه نیز تقریبی از پیچیدگی سیگنال است. **آنتروپی نمونه درجه دوم** **(QSE)** [55] یک آنتروپی نمونه است که به پارامتر شعاع داده r حساس نیست. این اجازه می دهد تا r در صورت نیاز برای دستیابی به برآوردهای مطمئن از احتمال شرطی تغییر کند. **آنتروپی چند مقیاسی** **(MSE)** [56] تعمیم یک اندازه گیری آنتروپی (مانند آنتروپی نمونه) به مقیاس های زمانی مختلف است. **آنتروپی چند مقیاسی اصلاح شده (MMSE)** [57] از همان روش MSE استفاده می کند، اما روش میانگین متحرک را جایگزین دانه بندی درشت[[78]](#footnote-79) می کند. **آنتروپی چند مقیاسی مرکب (CMSE)** اصلاحی از MSE است که با مشکل افزایش واریانس و تخمین خطا برای سری‌های زمانی کوتاه مقابله می‌کند.

**آنتروپی جای‌گشت (PE)** [58] تنوع سیگنال را بر اساس تکرار الگوهای ترتیبی تخمین می زند. این الگوریتم برای بدست آوردن الگوهای ترتیبی و احتمال وقوع آنها به پارامتر (ترتیب جایگشت) نیاز دارد. سپس این احتمالات در بیان آنتروپی شانون اعمال می شود. علاوه بر این، **آنتروپی جایگشت رنی (RPE) [59]**، **جایگشت آنتروپی رنی (PEr) [60]**، **آنتروپی جایگشت چند متغیره (MvPE) [61]**، و **آنتروپی جایگشت تسلیس (TPE) [62]** را می توان برای الگوهای ترتیبی محاسبه کرد. **آنتروپی پراکندگی (DisE) [63]** اصلاحی در آنتروپی جایگشت است که با مشکل از دست دادن اطلاعات دامنه مقابله می کند (زیرا آنتروپی جایگشت فقط ترتیب مقادیر دامنه را در نظر می گیرد اما خود مقادیر را در نظر نمی گیرد). **آنتروپی جایگشت آگاه از دامنه (AAPE) [64]** بر اساس ایده مشابه استفاده از مقدار سیگنال با آنتروپی جایگشت است. **آنتروپی حباب (BE) [65]** مشابه آنتروپی جایگشت با تفاوت اصلی در روش مورد استفاده برای رتبه بندی بردارها در فضای جاسازی[[79]](#footnote-80) است. یعنی، آنتروپی جایگشت از تکرار الگوهای ترتیبی و آنتروپی حبابی از تعداد مراحل مورد نیاز برای مرتب‌سازی یک بردار با الگوریتم مرتب‌سازی حبابی استفاده می‌کند. محاسبه **آنتروپی دیفرانسیل (DifEn) [66]** بر اساس بیان آنتروپی شانون و تخمین تابع چگالی احتمال زیربنایی سری‌های زمانی است. **آنتروپی فازی (FuzzyEn) [67]** مبتنی بر مفهوم مجموعه های فازی است که اولین بار توسط دکتر زاده [68] معرفی شد. این شبیه به آنتروپی نمونه است، اما به جای استفاده از تابع هوی‌ساید[[80]](#footnote-81) برای محاسبه فاصله، از تابع عضویت فازی استفاده می کند. **آنتروپی انتقال (TrEn) [69]** از مفاهیمی مشابه اطلاعات متقابل برای کمی کردن مبادله اطلاعات بین دو سیستم استفاده می کند. این یک اندازه گیری نامتقارن برای انتقال اطلاعات از فرآیند X به فرآیند Y است که تأثیر مقادیر گذشته فرآیندهای X و Y را بر ارزش فعلی فرآیند Y اندازه گیری می کند.

## بررسی برخی خصوصیات آماری ویژگی‌های به‌دست‌آمده

قدم اول پیش از پیاده‌سازی هرگونه آزمون و محک آماری بر روی داده‌های حاصل از ویژگی‌های استخراج شده از ۲ کانال پیشانی این است که با رسم هیستوگرام[[81]](#footnote-82) داده‌ها توزیع آماری آن‌ها بررسی شده و آزمون مناسب برای مقایسه‌ی گروه‌های تحت مطالعه انتخاب شود. چرا که آزمون‌های آماری خود به دو دستة پارامتری و غیرپارامتری تقسیم شده و آزمون‌های پارامتری از یک سری پیش‌فرض در مورد جامعه‌ای که به آن اعمال می‌شود استفاده می‌کند. هدف آزمون‌های آماری این است که بدون قربانی کردن قدرت پیش‌بینی، ویژگی‌هایی که اضافی یا نامربوط هستند (در واقع می‌توانند بر عملکرد مدل تأثیر منفی بگذارند) را شناسایی و در صورت نیاز حذف کند.

مزیت اصلی انتخاب ویژگی این است که برازش بیش از حد[[82]](#footnote-83) را کاهش می‌دهد. همچنین با حذف داده‌های اضافی، این امکان را به مدل می‌دهد تا فقط بر ویژگی‌های مهم داده‌ها تمرکز کند و از پردازش ویژگی‌هایی که تاثیر ناچیزی در صحت طبقه‌بندی نهایی دارند جلوگیری می‌شود. یکی دیگر از مزایای حذف اطلاعات نامربوط این است که دقت پیش‌بینی‌های مدل را بهبود می‌بخشد. همچنین زمان محاسبات موردنیاز برای به‌دست‌آوردن مدل را کاهش می‌دهد. در نهایت، داشتن تعداد کمتری از ویژگی‌ها، مدل را قابل‌تفسیرتر کرده و درک آن را آسان‌تر می‌کند. به‌طورکلی، انتخاب ویژگی خوب کلیدی برای پیش‌بینی مقادیر با دقت بالا است.

دو مورد از متداول‌ترین روش انتخاب ویژگی برای داده‌های ورودی عددی زمانی که متغیر هدف دسته‌ای[[83]](#footnote-84) است (مثلاً مدل‌سازی پیش‌بینی طبقه‌بندی)، آماره‌ی F آزمون تحلیل واریانس[[84]](#footnote-85) (ANOVA) و معیار اطلاعات متقابل[[85]](#footnote-86) است.

شرط استفاده از آزمون تحلیل واریانس یا به‌اختصار ANOVA، نرمال بودن توزیع آماری داده و نیز برقراری شرط کروی بودن[[86]](#footnote-87) است.

پس از رسم هیستوگرام اگر به‌صورت چشمی واضح بود که داده از توزیع نرمال پیروی نمی‌کند از آزمون‌های غیرپارامتری استفاده خواهد شد. در صورت شباهت هیستوگرام به توزیع نرمال از آزمون‌های سنجش نرمالیتی مانند آزمون کولموگروف - اسمیرنوف برای اطمینان از نرمال بودن توزیع استفاده خواهد شد.

### انتخاب ویژگی بر اساس آزمون آماری تحلیل واریانس F (ANOVA F-test)

ANOVA مخفف "تحلیل واریانس" است و یک آزمون فرضیه آماری پارامتری برای تعیین اینکه آیا میانگین دو یا چند نمونه از داده‌ها (اغلب سه یا بیشتر) از توزیع یکسان می‌آیند یا خیر است. آماره F یا آزمون F، دسته‌ای از آزمون‌های آماری است که نسبت بین مقادیر واریانس‌ها را محاسبه می‌کند،

### انتخاب ویژگی بر اساس آزمون آماری مربع کای[[87]](#footnote-88)

آزمون مربع کای آزمون آماری استقلال برای تعیین وابستگی دو متغیر است. شباهت‌هایی با ضریب تعیین R2 دارد. بااین‌حال، آزمون مربع کای فقط برای داده‌های طبقه‌ای یا اسمی قابل‌استفاده است.

از تعریف مربع کای به‌راحتی می‌توانیم کاربرد این تکنیک در انتخاب ویژگی را استنباط کنیم. فرض کنید یک متغیر هدف (به‌عنوان‌مثال، برچسب طبقه خواب) و برخی ویژگی‌های دیگر (متغیرهای ویژگی) دارید که هر نمونه از داده‌ها را توصیف می‌کند. حال، آمار مربع کای را بین هر متغیر ویژگی و متغیر هدف محاسبه می‌کنیم و وجود رابطه بین متغیرها و متغیر هدف را مشاهده می‌کنیم. اگر متغیر هدف مستقل از متغیر ویژگی باشد، می‌توانیم آن متغیر ویژگی را کنار بگذاریم. اگر آنها وابسته باشند، متغیر ویژگی بسیار مهم است.

## طبقه‌بندی

پس از استخراج ویژگی و بررسی آماری تفاوت‌های موجود میان گروه‌های تحت مطالعه، طبقه‌بندی پنج گروه بیداری (0)، N1 (1)، N2 (2)، N3(3) و REM (4) با استفاده از الگوریتم ماشین بردار پشتیبان چند کلاسه انجام گرفت که در ادامه جزئیات این روش‌ها آورده شده است.

### دسته‌بندی با ماشین بردار پشتیبان (SVM)

طریقة کار SVM به این صورت است که برای هر برچسب ابتدا بردارهای پشتیبان را می‌یابد. بردار پشتیبان از داده‌هایی می‌گذرد که بیشترین حاشیه را ایجاد می‌کنند. سپس با استفاده از این بردارهای پشتیبان ابرصفحة جداکنندة گروه‌ها را با میانگین‌گیری از بردارهای موجود مشخص می‌کند که این ابرصفحه بالاترین قدرت تفکیک (بیشترین دقت) را دارد.

در ماشین‌های بردار پشتیبان، سعی می‌شود که یک مدل خطی بر داده‌ها برازش شود که بیشترین حاشیه را داشته باشد. در اینجا حاشیه فاصله نزدیک‌ترین نقطه داده‌های ویژگی آموزش به خط جداکننده است (شکل 12).



شکل 12 – نمایش عملکرد SVM در یک فضای دو بعدی با دو برچسب [70]

طبقه‌بندهای SVM در ساده‌ترین نوعشان در طبقه‌بندی باینری اعمال می‌شوند و نقاط داده را به دو کلاس تقسیم می‌کند. برای طبقه‌بندی چند کلاسه نیز، از همان اصل استفاده می‌شود و مسئله چند کلاسه به چند مورد طبقه‌بندی باینری تقسیم می‌شود.

### ارزیابی صحت الگوریتم طبقه‌بندی

در این پژوهش برای ارزیابی طبقه‌بند از روش ارزیابی دسته‌های k تایی[[88]](#footnote-89) استفاده شد. در این روش ارزیابی، پارامتر k برابر ۱۰ در نظر گرفته شد. یکی از الزامات این روش آن است که داده پیش از شروع آموزش به‌صورت تصادفی بر زده[[89]](#footnote-90) شود. در طی این روش مقدار k روی ۱۰ ثابت شده است، مقداری که از طریق آزمایش به‌طورکلی به یک تخمین مهارت مدل با بایاس کم و واریانس متوسط می‌انجامد.

روال کلی به شرح زیر است:

1. مجموعه‌داده را به‌صورت تصادفی مخلوط می‌کنیم.
2. مجموعه‌داده را به k گروه تقسیم می‌کنیم.
3. برای هر گروه منحصربه‌فرد:
4. گروهی به‌عنوان مجموعه ‌داده‌های نگهدارنده یا آزمون در نظر گرفته می‌شود.
5. گروه‌های باقی‌مانده به‌عنوان مجموعه ‌داده‌های آموزشی در نظر گرفته می‌شود.
6. مدل روی داده‌ی آموزش برازش شده و در مجموعه آزمون ارزیابی می‌شود.
7. امتیاز ارزیابی ذخیره و مدل را کنار گذاشته می‌شود.
8. عملکرد مدل را با استفاده از معیار های صحت، دقت، یادآوری و امتیاز F1 ارزیابی می‌کنیم.

# جمع‌بندی

در این فصل پس از معرفی دستگاه ثبت دادگان الکتروانسفالوگرام و معرفی دادگان تجربی ثبت شده در پژوهش به بررسی ساختار پایگاه دادة موجود پرداخته شد. سپس فرایند طبقه‌بندی بصری این دادگان موردبحث قرار گرفت و نحوه دسته‌بندی و نمایش این دادگان شرح داده شد. در ادامه روش پیش‌پردازش شامل حذف خط پایه، اعمال یک فیلتر میان‌گذر بر دادگان و دوره‌بندی زمانی، بیان و سپس مراحل پردازش و استخراج ویژگی توضیح داده شد. این ویژگی‌ها شامل ویژگی‌های حوزه زمان، فرکانس، زمان - فرکانس، ویژگی‌های غیرخطی و معیارهای آنتروپی متعدد می‌شوند. پس از آن به بررسی روش‌های مختلف انتخاب ویژگی پرداخته شد. در پایان نیز روش طبقه‌بندی استفاده شده یعنی SVM تشریح شد. در فصل بعد به ارایه نتایج پیاده‌سازی و عملکرد نرم‌افزار و پیش‌پردازش و پردازش دادگان پرداخته می‌شود. همچنین نتایج انتخاب ویژگی های برتر و عملکرد مدل SVM بر روی دادگان مورد برسی قرار خواهد گرفت.

# مراجع

[1] I. Suzuki, K. Yamada, T. Yamakawa, M. Hashiba, and K. Akazawa, “Delivery of medical multimedia contents through the TCP/IP network using RealSystem,” *Comput Methods Programs Biomed*, vol. 70, no. 3, pp. 253–258, Mar. 2003, doi: 10.1016/s0169-2607(02)00012-3.

[2] M. J. Prerau, R. E. Brown, M. T. Bianchi, J. M. Ellenbogen, and P. L. Purdon, “Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis,” *Physiology*, vol. 32, no. 1, pp. 60–92, Jan. 2017, doi: 10.1152/physiol.00062.2015.

[3] I. C, “The AASM Manual for the Scoring of Sleep and Associated Events : Rules,” *Terminology and Technical Specification*, 2007, Accessed: Apr. 23, 2022. [Online]. Available: https://cir.nii.ac.jp/crid/1573668925445260928

[4] P. Anderer *et al.*, “Artifact processing in computerized analysis of sleep EEG - a review,” *Neuropsychobiology*, vol. 40, no. 3, pp. 150–157, Sep. 1999, doi: 10.1159/000026613.

[5] B. Maess, E. Schröger, and A. Widmann, “High-pass filters and baseline correction in M/EEG analysis. Commentary on: ‘How inappropriate high-pass filters can produce artefacts and incorrect conclusions in ERP studies of language and cognition,’” *J Neurosci Methods*, vol. 266, pp. 164–165, Jun. 2016, doi: 10.1016/j.jneumeth.2015.12.003.

[6] D. Jiang, Y. Lu, Y. Ma, and Y. Wang, “Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement,” *Expert Systems with Applications*, vol. 121, pp. 188–203, May 2019, doi: 10.1016/j.eswa.2018.12.023.

[7] N. Michielli, U. R. Acharya, and F. Molinari, “Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals,” *Comput Biol Med*, vol. 106, pp. 71–81, Mar. 2019, doi: 10.1016/j.compbiomed.2019.01.013.

[8] null Geering, null Achermann, null Eggimann, and null Borbély, “Period-amplitude analysis and power spectral analysis: a comparison based on all-night sleep EEG recordings,” *J Sleep Res*, vol. 2, no. 3, pp. 121–129, Sep. 1993, doi: 10.1111/j.1365-2869.1993.tb00074.x.

[9] B. Hjorth, “EEG analysis based on time domain properties,” *Electroencephalogr Clin Neurophysiol*, vol. 29, no. 3, pp. 306–310, Sep. 1970, doi: 10.1016/0013-4694(70)90143-4.

[10] X.-W. Wang, D. Nie, and B.-L. Lu, “EEG-Based Emotion Recognition Using Frequency Domain Features and Support Vector Machines,” in *Neural Information Processing*, Berlin, Heidelberg, 2011, pp. 734–743. doi: 10.1007/978-3-642-24955-6\_87.

[11] G. H. B. S. Oliveira *et al.*, “Multitaper-based method for automatic k-complex detection in human sleep EEG,” *Expert Systems with Applications*, vol. 151, p. 113331, Aug. 2020, doi: 10.1016/j.eswa.2020.113331.

[12] H. J. Nussbaumer, “The Fast Fourier Transform,” in *Fast Fourier Transform and Convolution Algorithms*, H. J. Nussbaumer, Ed. Berlin, Heidelberg: Springer, 1981, pp. 80–111. doi: 10.1007/978-3-662-00551-4\_4.

[13] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, Jun. 1967, doi: 10.1109/TAU.1967.1161901.

[14] D. J. Thomson, “Spectrum estimation and harmonic analysis,” *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982, doi: 10.1109/PROC.1982.12433.

[15] H. Akaike, “Fitting autoregressive models for prediction,” *Ann Inst Stat Math*, vol. 21, no. 1, pp. 243–247, Dec. 1969, doi: 10.1007/BF02532251.

[16] A. Neumaier and T. Schneider, “Estimation of parameters and eigenmodes of multivariate autoregressive models,” *ACM Trans. Math. Softw.*, vol. 27, no. 1, pp. 27–57, Mar. 2001, doi: 10.1145/382043.382304.

[17] H. J. Eoh, M. K. Chung, and S.-H. Kim, “Electroencephalographic study of drowsiness in simulated driving with sleep deprivation,” *International Journal of Industrial Ergonomics*, vol. 35, no. 4, pp. 307–320, Apr. 2005, doi: 10.1016/j.ergon.2004.09.006.

[18] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, “Using EEG spectral components to assess algorithms for detecting fatigue,” *Expert Systems with Applications*, vol. 36, no. 2, Part 1, pp. 2352–2359, Mar. 2009, doi: 10.1016/j.eswa.2007.12.043.

[19] T. L. T. da Silveira, A. J. Kozakevicius, and C. R. Rodrigues, “Automated drowsiness detection through wavelet packet analysis of a single EEG channel,” *Expert Systems with Applications*, vol. 55, pp. 559–565, Aug. 2016, doi: 10.1016/j.eswa.2016.02.041.

[20] K. Schwab, M. Eiselt, C. Schelenz, and H. Witte, “Time-variant parametric estimation of transient quadratic phase couplings during electroencephalographic burst activity,” *Methods Inf Med*, vol. 44, no. 3, pp. 374–383, 2005.

[21] L. Debnath, *Wavelets and Signal Processing*. Springer Science & Business Media, 2012.

[22] A. Abbate, C. DeCusatis, and P. K. Das, *Wavelets and Subbands: Fundamentals and Applications*. Springer Science & Business Media, 2012.

[23] Y. Meyer, *Wavelets and Operators*, vol. 1. Cambridge: Cambridge University Press, 1993. doi: 10.1017/CBO9780511623820.

[24] J. W. Kantelhardt, “Fractal and Multifractal Time Series,” in *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, Ed. New York, NY: Springer, 2009, pp. 3754–3779. doi: 10.1007/978-0-387-30440-3\_221.

[25] M. F. Barnsley, “Chapter V - Fractal Dimension,” in *Fractals Everywhere (Second Edition)*, M. F. Barnsley, Ed. Academic Press, 1993, pp. 171–204. doi: 10.1016/B978-0-12-079061-6.50010-3.

[26] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, 2004.

[27] W. S. Pritchard and D. W. Duke, “Measuring ‘chaos’ in the brain: a tutorial review of EEG dimension estimation,” *Brain Cogn*, vol. 27, no. 3, pp. 353–397, Apr. 1995, doi: 10.1006/brcg.1995.1027.

[28] P. Grassberger and I. Procaccia, “Estimation of the Kolmogorov entropy from a chaotic signal,” *Phys. Rev. A*, vol. 28, no. 4, pp. 2591–2593, Oct. 1983, doi: 10.1103/PhysRevA.28.2591.

[29] T. Zorick and M. A. Mandelkern, “Multifractal Detrended Fluctuation Analysis of Human EEG: Preliminary Investigation and Comparison with the Wavelet Transform Modulus Maxima Technique,” *PLoS One*, vol. 8, no. 7, p. e68360, Jul. 2013, doi: 10.1371/journal.pone.0068360.

[30] null Muzy, null Bacry, and null Arneodo, “Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method,” *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, vol. 47, no. 2, pp. 875–884, Feb. 1993, doi: 10.1103/physreve.47.875.

[31] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, “Multifractal detrended fluctuation analysis of nonstationary time series,” *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1, pp. 87–114, Dec. 2002, doi: 10.1016/S0378-4371(02)01383-3.

[32] H. E. Hurst, “The Problem of Long-Term Storage in Reservoirs,” *International Association of Scientific Hydrology. Bulletin*, vol. 1, no. 3, pp. 13–27, Jan. 1956, doi: 10.1080/02626665609493644.

[33] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, “Mosaic organization of DNA nucleotides,” *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, vol. 49, no. 2, pp. 1685–1689, Feb. 1994, doi: 10.1103/physreve.49.1685.

[34] J. Alvarez-Ramirez, E. Rodriguez, and J. Carlos Echeverría, “Detrending fluctuation analysis based on moving average filtering,” *Physica A: Statistical Mechanics and its Applications*, vol. 354, pp. 199–219, Aug. 2005, doi: 10.1016/j.physa.2005.02.020.

[35] H. Lotfalinezhad and A. Maleki, “TTA, a new approach to estimate Hurst exponent with less estimation error and computational time,” *Physica A: Statistical Mechanics and its Applications*, vol. 553, p. 124093, Sep. 2020, doi: 10.1016/j.physa.2019.124093.

[36] I. I. Shevchenko, “Lyapunov exponents in resonance multiplets,” *Physics Letters A*, vol. 378, no. 1, pp. 34–42, Jan. 2014, doi: 10.1016/j.physleta.2013.10.035.

[37] C. Shen, S. Yu, J. Lü, and G. Chen, “Designing Hyperchaotic Systems With Any Desired Number of Positive Lyapunov Exponents via A Simple Model,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 8, pp. 2380–2389, Aug. 2014, doi: 10.1109/TCSI.2014.2304655.

[38] A. Lempel and J. Ziv, “On the Complexity of Finite Sequences,” *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75–81, Jan. 1976, doi: 10.1109/TIT.1976.1055501.

[39] M. E. Cohen, D. L. Hudson, and P. C. Deedwania, “Applying continuous chaotic modeling to cardiac signal analysis,” *IEEE Engineering in Medicine and Biology Magazine*, vol. 15, no. 5, pp. 97–102, Sep. 1996, doi: 10.1109/51.537065.

[40] D. Abásolo, J. Escudero, R. Hornero, C. Gómez, and P. Espino, “Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer’s disease patients,” *Med Biol Eng Comput*, vol. 46, no. 10, pp. 1019–1028, Oct. 2008, doi: 10.1007/s11517-008-0392-1.

[41] null Daw, null Finney, and null Kennel, “Symbolic approach for measuring temporal ‘irreversibility,’” *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, vol. 62, no. 2 Pt A, pp. 1912–1921, Aug. 2000, doi: 10.1103/physreve.62.1912.

[42] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, “Recurrence Plots of Dynamical Systems,” *EPL*, vol. 4, no. 9, pp. 973–977, Nov. 1987, doi: 10.1209/0295-5075/4/9/004.

[43] J. P. Zbilut, A. Giuliani, and C. L. Webber, “Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification,” *Physics Letters A*, vol. 246, no. 1, pp. 122–128, Sep. 1998, doi: 10.1016/S0375-9601(98)00457-5.

[44] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.

[45] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, Oct. 1948, doi: 10.1002/j.1538-7305.1948.tb00917.x.

[46] A. Renyi, *Probability Theory*. Courier Corporation, 2007.

[47] C. Tsallis, RenioS. Mendes, and A. R. Plastino, “The role of constraints within generalized nonextensive statistics,” *Physica A: Statistical Mechanics and its Applications*, vol. 261, no. 3, pp. 534–554, Dec. 1998, doi: 10.1016/S0378-4371(98)00437-3.

[48] J. Fell, J. Röschke, K. Mann, and C. Schäffner, “Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures,” *Electroencephalogr Clin Neurophysiol*, vol. 98, no. 5, pp. 401–410, May 1996, doi: 10.1016/0013-4694(96)95636-9.

[49] N. Kannathal, M. L. Choo, U. R. Acharya, and P. K. Sadasivan, “Entropies for detection of epilepsy in EEG,” *Comput Methods Programs Biomed*, vol. 80, no. 3, pp. 187–194, Dec. 2005, doi: 10.1016/j.cmpb.2005.06.012.

[50] H. Viertiö-Oja *et al.*, “Description of the Entropy algorithm as applied in the Datex-Ohmeda S/5 Entropy Module,” *Acta Anaesthesiol Scand*, vol. 48, no. 2, pp. 154–161, Feb. 2004, doi: 10.1111/j.0001-5172.2004.00322.x.

[51] S. Mo, E. Mj, van G. Mj, Y.-H. Am, J. Vh, and V. Ap, “Quantification of epileptiform electroencephalographic activity during sevoflurane mask induction,” *Anesthesiology*, vol. 107, no. 6, Dec. 2007, doi: 10.1097/01.anes.0000291444.68894.ee.

[52] C. Zhang, H. Wang, and R. Fu, “Automated Detection of Driver Fatigue Based on Entropy and Complexity Measures,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 168–177, Feb. 2014, doi: 10.1109/TITS.2013.2275192.

[53] S. M. Pincus, I. M. Gladstone, and R. A. Ehrenkranz, “A regularity statistic for medical data analysis,” *J Clin Monit*, vol. 7, no. 4, pp. 335–345, Oct. 1991, doi: 10.1007/BF01619355.

[54] J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” *Am J Physiol Heart Circ Physiol*, vol. 278, no. 6, pp. H2039-2049, Jun. 2000, doi: 10.1152/ajpheart.2000.278.6.H2039.

[55] B. García-Martínez, A. Martínez-Rodrigo, R. Zangróniz Cantabrana, J. M. Pastor García, and R. Alcaraz, “Application of Entropy-Based Metrics to Identify Emotional Distress from Electroencephalographic Recordings,” *Entropy*, vol. 18, no. 6, Art. no. 6, Jun. 2016, doi: 10.3390/e18060221.

[56] J. Courtiol *et al.*, “The multiscale entropy: Guidelines for use and interpretation in brain signal analysis,” *J Neurosci Methods*, vol. 273, pp. 175–190, Nov. 2016, doi: 10.1016/j.jneumeth.2016.09.004.

[57] S.-D. Wu, C.-W. Wu, K.-Y. Lee, and S.-G. Lin, “Modified multiscale entropy for short-term time series analysis,” *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 23, pp. 5865–5873, Dec. 2013, doi: 10.1016/j.physa.2013.07.075.

[58] D. Cuesta-Frau, J. P. Murillo-Escobar, D. A. Orrego, and E. Delgado-Trejos, “Embedded Dimension and Time Series Length. Practical Influence on Permutation Entropy and Its Applications,” *Entropy (Basel)*, vol. 21, no. 4, p. 385, Apr. 2019, doi: 10.3390/e21040385.

[59] X. Zhao, P. Shang, and J. Huang, “Permutation complexity and dependence measures of time series,” *EPL*, vol. 102, no. 4, p. 40005, May 2013, doi: 10.1209/0295-5075/102/40005.

[60] N. Mammone, J. Duun-Henriksen, T. W. Kjaer, and F. C. Morabito, “Differentiating Interictal and Ictal States in Childhood Absence Epilepsy through Permutation Rényi Entropy,” *Entropy*, vol. 17, no. 7, Art. no. 7, Jul. 2015, doi: 10.3390/e17074627.

[61] S. He, K. Sun, and H. Wang, “Multivariate permutation entropy and its application for complexity analysis of chaotic systems,” *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 812–823, Nov. 2016, doi: 10.1016/j.physa.2016.06.012.

[62] Z. Liang *et al.*, “EEG entropy measures in anesthesia,” *Front Comput Neurosci*, vol. 9, p. 16, Feb. 2015, doi: 10.3389/fncom.2015.00016.

[63] M. Rostaghi and H. Azami, “Dispersion Entropy: A Measure for Time-Series Analysis,” *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 610–614, May 2016, doi: 10.1109/LSP.2016.2542881.

[64] A. H and E. J, “Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation,” *Computer methods and programs in biomedicine*, vol. 128, May 2016, doi: 10.1016/j.cmpb.2016.02.008.

[65] G. Manis, M. Aktaruzzaman, and R. Sassi, “Bubble Entropy: An Entropy Almost Free of Parameters,” *IEEE Trans Biomed Eng*, vol. 64, no. 11, pp. 2711–2718, Nov. 2017, doi: 10.1109/TBME.2017.2664105.

[66] “Cover: Elements of information theory - Google Scholar.” https://scholar.google.com/scholar\_lookup?title=Elements+of+Information+Theory&author=T.M.+Cover&author=J.A.+Thomas&publication\_year=2005& (accessed May 02, 2022).

[67] W. Chen, Z. Wang, H. Xie, and W. Yu, “Characterization of surface EMG signal based on fuzzy entropy,” *IEEE Trans Neural Syst Rehabil Eng*, vol. 15, no. 2, pp. 266–272, Jun. 2007, doi: 10.1109/TNSRE.2007.897025.

[68] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.

[69] null Schreiber, “Measuring information transfer,” *Phys Rev Lett*, vol. 85, no. 2, pp. 461–464, Jul. 2000, doi: 10.1103/PhysRevLett.85.461.

[70] S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow*. Packt Publishing, 2017.

1. Photoplethysmogram [↑](#footnote-ref-1)
2. Obstructive Sleep Apnea [↑](#footnote-ref-2)
3. Polysomnography [↑](#footnote-ref-3)
4. American Academy of Sleep Medicine [↑](#footnote-ref-4)
5. Finite Impulse Response [↑](#footnote-ref-5)
6. Infinite Impulse Response [↑](#footnote-ref-6)
7. Epoch [↑](#footnote-ref-7)
8. Data Augmentation [↑](#footnote-ref-8)
9. mean [↑](#footnote-ref-9)
10. median [↑](#footnote-ref-10)
11. variance [↑](#footnote-ref-11)
12. standard deviation [↑](#footnote-ref-12)
13. skewness [↑](#footnote-ref-13)
14. kurtosis [↑](#footnote-ref-14)
15. Zero-crossing rate [↑](#footnote-ref-15)
16. number of waves [↑](#footnote-ref-16)
17. wave duration [↑](#footnote-ref-17)
18. peak amplitude [↑](#footnote-ref-18)
19. instantaneous frequency [↑](#footnote-ref-19)
20. Hjorth [↑](#footnote-ref-20)
21. Hjorth Mobility [↑](#footnote-ref-21)
22. Hjorth Activity [↑](#footnote-ref-22)
23. Hjorth Complexity [↑](#footnote-ref-23)
24. Power Spectral Density [↑](#footnote-ref-24)
25. Welch Method [↑](#footnote-ref-25)
26. Thompson multitaper method [↑](#footnote-ref-26)
27. AutoRegressive-Moving Average [↑](#footnote-ref-27)
28. Reflection Coefficients [↑](#footnote-ref-28)
29. Partial Correlation Coefficients [↑](#footnote-ref-29)
30. Hurst exponent [↑](#footnote-ref-30)
31. Renyi scaling exponent [↑](#footnote-ref-31)
32. Renyi generalized dimension of multifractals [↑](#footnote-ref-32)
33. Hausdorff dimension [↑](#footnote-ref-34)
34. Katz fractal dimension [↑](#footnote-ref-35)
35. Petrosian fractal dimension [↑](#footnote-ref-36)
36. Higuchi fractal dimension [↑](#footnote-ref-37)
37. Fluctuation Analysis [↑](#footnote-ref-38)
38. Hurst’s Rescaled-Range Analysis [↑](#footnote-ref-39)
39. Detrended Fluctuation Analysis [↑](#footnote-ref-40)
40. Centered Moving Average Analysis [↑](#footnote-ref-41)
41. Triangle Total Areas [↑](#footnote-ref-42)
42. Wavelet Transform Modulus Maxima [↑](#footnote-ref-43)
43. Multifractal Detrended Fluctuation Analysis [↑](#footnote-ref-44)
44. Lyapunov exponents [↑](#footnote-ref-45)
45. Lempel-Ziv complexity [↑](#footnote-ref-46)
46. Central tendency measure [↑](#footnote-ref-47)
47. Auto-mutual information [↑](#footnote-ref-48)
48. Temporal irreversibility [↑](#footnote-ref-49)
49. recurrence plot [↑](#footnote-ref-50)
50. Recurrence quantification analysis [↑](#footnote-ref-51)
51. recurrence rate [↑](#footnote-ref-52)
52. Determinism [↑](#footnote-ref-53)
53. laminarity [↑](#footnote-ref-54)
54. average diagonal line length [↑](#footnote-ref-55)
55. maximum length of diagonal [↑](#footnote-ref-56)
56. maximum length of vertical lines [↑](#footnote-ref-57)
57. Trapping time [↑](#footnote-ref-58)
58. Divergence [↑](#footnote-ref-59)
59. Entropy of the recurrence plot [↑](#footnote-ref-60)
60. deterministic [↑](#footnote-ref-61)
61. Rényi [↑](#footnote-ref-62)
62. Tsallis Entropy [↑](#footnote-ref-63)
63. q-entropy [↑](#footnote-ref-64)
64. Boltzman–Gibbs entropy [↑](#footnote-ref-65)
65. non-extensity parameter [↑](#footnote-ref-66)
66. Kraskov entropy [↑](#footnote-ref-67)
67. Spectral entropy [↑](#footnote-ref-68)
68. Quadratic Renyi’s spectral entropy [↑](#footnote-ref-69)
69. Commercial M-Entropy Module [↑](#footnote-ref-70)
70. response entropy [↑](#footnote-ref-71)
71. state entropy [↑](#footnote-ref-72)
72. Tsallis wavelet entropy [↑](#footnote-ref-73)
73. Rényi’s wavelet entropy [↑](#footnote-ref-74)
74. Hilbert–Huang spectral entropy [↑](#footnote-ref-75)
75. Log energy entropy [↑](#footnote-ref-76)
76. Multiresolution entropy [↑](#footnote-ref-77)
77. self-matches [↑](#footnote-ref-78)
78. coarse-graining [↑](#footnote-ref-79)
79. embedding space [↑](#footnote-ref-80)
80. Heaviside [↑](#footnote-ref-81)
81. histogram [↑](#footnote-ref-82)
82. overfitting [↑](#footnote-ref-83)
83. categorical [↑](#footnote-ref-84)
84. Analysis of Variance [↑](#footnote-ref-85)
85. mutual information statistic [↑](#footnote-ref-86)
86. sphericity [↑](#footnote-ref-87)
87. chi-square [↑](#footnote-ref-88)
88. K-fold cross validation [↑](#footnote-ref-89)
89. Shuffle [↑](#footnote-ref-90)