دسته‌بندی مراحل خواب از نوار قلب با استفاده از مدل‌های یادگیری ماشین

دانشجو: پرنیان فاضل

استاد راهنما: دکتر سیامک محمدی

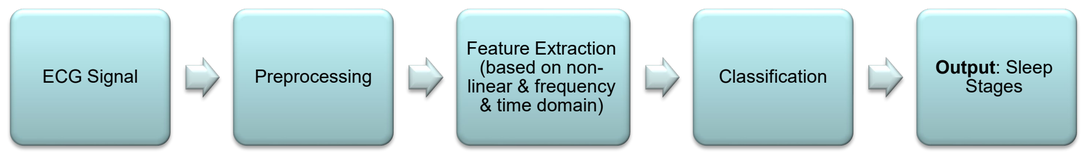
دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران

خلاصه

خواب یک حالت فیزیولوژیکی پیچیده است که بر عملکرد مغز، سیستم عصبی خودمختار و ریتم‌های شبانه‌روزی تأثیر می‌گذارد. روش‌های سنتی ارزیابی خواب مانند پلی‌سومنوگرافی یا بررسی نوار مغز، هزینه‌بر و در شرایطی غیر عملی هستند. سیگنال‌های نوار قلب به عنوان جایگزینی ارزان‌تر ارائه می‌شوند و درک بهتری از تغییرات ضربان قلب را فراهم می‌کنند. هدف این پژوهش، اجرا و مقایسه عملکرد مدل‌های یادگیری ماشین در دسته‌بندی مراحل خواب با استفاده از سیگنال‌های نوار قلب است. پیش‌پردازش داده، استخراج ویژگی و آموزش مدل‌ها برای پیش‌بینی مراحل خواب از سیگنال‌های ضربان قلب انجام می‌شود. طبق نتایج این پژوهش، دسته‌بندی مراحل خواب بر اساس سیگنال‌های قلب دارای پتانسیل بالقوه‌ای است و رویکردی دسترس‌پذیرتر و اقتصادی‌تر برای تشخیص و نظارت بر اختلالات خواب ارائه می‌دهد. در این پژوهش برای اولین بار به دسته‌بندی مراحل خواب به ۴ دسته بیداری، REM، Low و Deep بر روی سیگنال نوار قلب از مجموعه داده دانشگاه دابلین با استفاده از مدل‌های یادگیری ماشین پرداخته شده است.

روش پیشنهادی

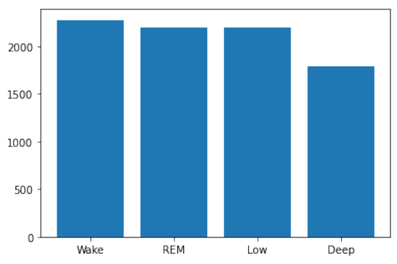
جهت انجام پروژه و بررسی عملکرد مدل‌های مختلف یادگیری ماشین برای دسته‌بندی مراحل مختلف خواب، از مجموعه‌داده‌ی نوار قلب جمع‌آوری شده از بیماران مراجعه کننده به یک بیمارستان (مجموعه داده UCD) استفاده شده است. مراحل پروژه به صورت کلی در نمودار زیر نشان داده شده است:



برای استفاده از این سیگنال‌ها ابتدا پیش‌پردازش انجام شد. با توجه به اینکه نمی‌توان یک سیگنال نوار قلب که مدت زمان طولانی ظبط شده است را جهت یادگیری به مدل یادگیری ماشین ارائه داد، باید آن را به قطعات کوچکتر تقسیم کرده و سپس برچسب‌گذاری کرد؛ سیگنال‌ها به بازه‌های کوچکتر ۳۰ ثانیه‌ای تقسیم شدند.

الگوریتم هایی که برای یادگیری مدل‌ها استفاده شدند مبتنی بر ویژگی هستند، بنابراین سه دسته ویژگی استخراج شد: ویژگی‌های حوزه زمان، ویژگی‌های حوزه فرکانس و ویژگی‌های غیر خطی. ویژگی‌های سیگنال‌های نوار قلب بر اساس تغییرات ضربان قلب استخراج می‌شوند؛ بنابراین، بایستی ابتدا سیگنال‌های نوار قلب به سیگنال‌های تغییرات ضربان قلب (HRV) تبدیل شوند و سپس ویژگی‌ها استخراج شوند. در مجموع ۷۲ ویژگی استخراج شد که این ویژگی‌ها برای آموزش مدل‌های یادگیری ماشین استفاده شد.

با توجه به اینکه مجموعه داده نامتوازن بود، مجموعه داده متوازن شد. تعداد و توازن هر مرحله از خواب در مجموعه داده در شکل روبرو مشخص شده است:



در نهایت با استفاده از مدل‌های یادگیری ماشین، دسته‌بندی مراحل خواب به ۴ دسته بیداری، REM، Low، Deep انجام شد. مدل‌های استفاده شده، رگرسیون لجستیک، درخت تصمیم، جنگل تصادفی، درخت تقویت شده با گرادیان، ماشین بردار پشتیبان و K نزدیک‌ترین همسایه هستند. برای پیدا کردن فراپارامترهای بهینه مربوط به هر مدل، از روش جستجوی بیزی استفاده شد.

نتایج

برای آموزش مدل‌های یادگیری ماشین، ترکیب‌های مختلف فراپارامترها مورد آزمون قرار گرفت و همچنین از سرچ بیزی استفاده شد تا بهترین نتیجه‌ی این مدل‌ها برای دسته‌بندی مراحل خواب به ۴ دسته بیداری، REM، Low و Deep برای مجموعه داده‌ی مورد نظر به دست آید.

برای ارزیابی مدل‌ها از متریک دقت (Accuracy) و F1-score استفاده شد. دقت بیانگر نسبت پیش‌بینی‌های درست انجام شده توسط مدل به کل پیش‌بینی‌های انجام شده است. F1-score نشان‌دهنده توازن و سازگاری میان صحت و جامعیت مدل در پیش‌بینی‌هاست و تصویر جامعی از عملکرد مدل ارائه می‌دهد. نتیجه هر یک از مدل‌های رگرسیون لجستیک (LR)، درخت تصمیم (DT)، جنگل تصادفی (RF)، درخت تقویت شده با گرادیان (GBC)، ماشین بردار پشتیبان (SVM) و K نزدیک‌ترین همسایه (KNN) با معیارهای F1-score و دقت در جدول زیر نشان داده شده است.

معیار / مدل	LR	RF	DT	SVM	GBC	KNN
Accuracy	۵۳٪	۶۰٪	۵۳٪	۵۶٪	۵۸٪	۵۵٪
F1-score	۴۲٪	۴۸٪	۴۴٪	۴۷٪	۵۲٪	۴۶٪

بیشترین دقت مربوط به مدل جنگل تصادفی (RF) برابر ۶۰٪ است. جنگل تصادفی با ترکیب تعداد زیادی درخت تصمیم‌گیری و استفاده از روش‌هایی مثل تصادفی‌سازی، از واریانس و بیش‌برازش کمتری نسبت به مدل‌های دیگر برخوردار است. همین موارد باعث شده دقت بالاتری در این مسئله دسته‌بندی داشته باشد.

بیشترین F1-score مربوط به مدل درخت تقویت شده با گرادیان (GBC) برابر ۵۲٪ است. GBC با تمرکز بر نمونه‌های دشوار، توانایی مدل‌سازی روابط پیچیده بین داده‌ها را دارد و با تنظیم پارامترهای مناسب، به نسبت بقیه مدل‌ها تعادل بهتری میان صحت و جامعیت پیش‌بینی ایجاد می‌کند. همین موارد باعث شده F1-score بالاتری نسبت به بقیه مدل‌ها در این مسئله دسته‌بندی داشته باشد.

جمع بندی

در این پژوهش با استفاده از سیگنال‌های نوار قلب و مدل‌های یادگیری ماشین، امکان طبقه‌بندی مراحل خواب به عنوان جایگزینی برای روش‌های پرهزینه مانند پلی‌سومنوگرافی یا بررسی نوار مغز مورد بررسی قرار گرفت. پس از پیش‌پردازش‌های لازم، الگوریتم‌های مختلف یادگیری ماشین مورد ارزیابی قرار گرفتند. نتایج نشان داد سیگنال نوار قلب می‌تواند در طبقه‌بندی مراحل خواب، به عنوان راهی به صرفه‌تر از روش‌های سنتی، استفاده شود.

کاربردهای صنعتی:

با توجه به نتایج این پژوهش، می‌توان دستگاه‌های ارزان‌تر و قابل استفاده در منزل برای غربالگری اختلالات خواب طراحی کرد. عرضه این محصولات به بازار توسط شرکت‌های تجهیزات پزشکی، علاوه بر منافع مالی، با افزایش دسترسی به خدمات، موجب ارتقای سلامت عمومی نیز می‌شود.

مراجع اصلی

1. W. Hayet and Y. Slim, “Sleep-wake stages classification based on heart rate variability,” 2012 5th International Conference on BioMedical Engineering and Informatics. IEEE, Oct. 2012.
2. W. McNicholas, L. Doherty, S. Ryan, J. Garvey, P. Boyle, and E. Chua, “St. Vincent’s University Hospital / University College Dublin Sleep Apnea Database.” physionet.org, 2004.
3. F. Bozkurt, M. K. Uçar, C. Bilgin, and A. Zengin, “Sleep–wake stage detection with single channel ECG and hybrid machine learning model in patients with obstructive sleep apnea,” Physical and Engineering Sciences in Medicine, vol. 44, no. 1, pp. 63–77, Jan. 2021.