



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

**DETECCIÓN DE MOVIMIENTOS DE PIERNAS PERIÓDICOS Y NO
PERIÓDICOS SOBRE EL ELECTROMIOGRAMA UTILIZANDO LA
ENERGÍA DE SHANNON**

MEMORIA PARA OPTAR AL TÍTULO
DE INGENIERO CIVIL ELÉCTRICO

MATIAS EZEQUIEL ARRIOLA ARAVENA

PROFESOR GUÍA:
ALEJANDRO BASSI ACUÑA

MIEMBROS DE LA COMISIÓN:
ÁLVARO SILVA MADRID
ANDRÉS CABA RUTTE

SANTIAGO DE CHILE
2021

DETECCIÓN DE MOVIMIENTOS DE PIERNAS PERIÓDICOS Y NO PERIÓDICOS SOBRE EL ELECTROMIOGRAMA UTILIZANDO LA ENERGÍA DE SHANNON

El objetivo de este trabajo de título fue desarrollar un sistema de detección automática de movimientos de piernas periódicos y no periódicos sobre el electromiograma en registros polisomnográficos pertenecientes al INTA de la Universidad de Chile.

Los movimientos de piernas periódicos han sido asociados a distintas patologías y/o trastornos del sueño, como por ejemplo, el síndrome de las piernas inquietas (RLS), el cual trae consigo la dificultad para conciliar el sueño. Debido a su relación con estos trastornos y/o patologías, es que la detección de movimientos de piernas periódicos sobre el EMG es relevante desde el punto de vista médico.

El sistema de detección desarrollado incorpora características como la Energía de Shannon, y además, algoritmos de clasificación de aprendizaje de máquinas (Random Forest) y técnicas de procesamiento avanzado de señales, tales como la Transformada Wavelet Discreta (DWT) y transformada de Hilbert-Huang (HHT), las que permiten obtener una representación de la señal en un espacio tanto temporal como frecuencial. Además, se obtienen características específicas que permiten filtrar artefactos producidos por el electrocardiograma, y criterios de tiempo para la duración de eventos de movimientos de piernas.

El sistema de detección automático desarrollado corresponde a un sistema modular compuesto por tres módulos con objetivos distintos entre sí: el primer módulo tiene por objetivo detectar la mayor cantidad de eventos de movimiento de piernas utilizando la Energía de Shannon. A su vez, el Módulo 2 tiene el objetivo de filtrar los eventos propuestos por el primer módulo, y de esta manera, filtrar la mayor cantidad de falsos positivos. Para ello, el Módulo 2 aplica la DWT, la HHT, y luego extrae distintas características sobre cada evento propuesto por el Módulo 1. Finalmente, el Módulo 3 tiene por objetivo generar las secuencias de movimientos periódicos a partir de los eventos de movimientos de piernas propuestos por el Módulo 2.

De un total de 27 registros polisomnográficos se construyó una base de datos con 25 registros polisomnográficos de sueño nocturno. Dos de ellos no se consideraron debido a que poseían un alto ruido en cada uno de sus canales. La base de datos fue dividida en tres conjuntos: 11 de entrenamiento, 4 de validación y 10 de prueba.

Los resultados sobre los registros del conjunto de prueba muestran una sensibilidad de 0,87, una precisión de 0,73 y un F1-score del 0,79. En general, el sistema desarrollado presenta resultados bastante buenos, los que son comparables con los mejores resultados mostrados en la literatura.

El algoritmo es parte integral de un software desarrollado como una herramienta de apoyo al trabajo de investigación de los especialistas en sueño.

A Dios y mi familia

Agradecimientos

En primer lugar quisiera agradecer a mi madre y mi padre, por su apoyo incondicional durante estos años y por el esfuerzo realizado para que yo y mis herman@s pudiéramos desarrollarnos en todos los ámbitos de nuestras vidas. Muchas gracias por enseñarnos y guiarnos a intentar siempre ser íntegros y mejores personas.

A mis herman@s por su amistad y apoyo incondicional: A mi hermana Camila sin la cual nada de esto hubiese sido posible. Gracias por acompañarme en todo y ser mi consejera y mejor amiga durante toda mi vida. A mi hermano Alexis que por su espontaneidad y personalidad única, siempre fue un apoyo para aliviar los días donde la carga académica fue ardua. Gracias por alegrarme los días.

Quisiera agradecer a mis amig@s y prim@s por acompañarme durante estos años y brindarme su compañía. A todos los del Club Andino Vomito por vivir momentos de mucha alegría y experiencias inolvidables en el cerro.

Quisiera agradecer al Profesor Alejandro Bassi por su orientación y buena disposición para llevar a cabo la finalización de este proyecto. Al Profesor Álvaro Silva por su buena onda, apoyo y consejos durante este último período. Al Profesor Andrés Caba por su paciencia y orientación entregada durante todo este período.

Finalmente quisiera agradecer al Profesor Marcelo Garrido del INTA de la Universidad de Chile, por la disposición a orientarme y entregarme sus conocimientos.

Tabla de Contenido

1. Introducción	1
1.1. Sobre el estudio del sueño	1
1.2. Sobre la polisomnografía	1
1.3. Sobre los movimientos de piernas periódicos (PLM) y no periódicos (LM) .	3
1.4. Estado del Arte en la detección automática de movimientos de piernas periódicos (PLM) y no periódicos (LM)	5
1.4.1. Métodos basados en umbrales de amplitud del EMG	5
1.4.2. Método basado en Aprendizaje de Máquinas	10
1.5. Objetivos	13
1.5.1. Objetivo General	13
1.5.2. Objetivos específicos	13
1.6. Organización del documento	13
2. Antecedentes generales	15
2.1. Sobre la energía de Shannon	15
2.2. Método Welch	16
2.3. EMA (Exponential Moving Average)	17
2.4. Transformada Wavelet Discreta (DWT)	18
2.5. Transformada de Hilbert-Huang (HHT)	22
2.6. Árboles de decisión y Algoritmo Random Forest	24
2.6.1. Árboles de decisión	24
2.6.2. Algoritmo Random Forest	25
3. Metodología	29
3.1. Sobre los datos utilizados	29
3.2. División de datos en conjuntos de entrenamiento, validación y prueba	30
3.3. Pre-Procesamiento de los datos	32
3.4. Método para la detección automática de movimientos de piernas	34
3.4.1. Primer módulo	35
3.4.2. Segundo módulo	37
3.4.2.1. Sobre las características que se obtienen para lidiar con los Falsos Positivos	38
3.4.2.2. Algoritmo para ajustar umbrales	39
3.4.3. Tercer módulo	42
3.5. Criterios utilizados para evaluar el desempeño de la metodología propuesta .	43
4. Resultados y Análisis	45
4.1. Resultados del módulo 1	45

4.1.1. Sobre los tipos de Falsos Positivos detectados	46
4.2. Resultados del Módulo 2	48
4.3. Comparación de la metodología implementada y estado del arte en detección de LM	51
4.4. Resultados Módulo 3	53
5. Conclusiones	55
6. Glosario de abreviaturas usadas en este documento	57
Bibliografía	59
Anexo A. Ajuste de umbrales en módulo 2 utilizando curva Precisión-Sensibilidad	63
A.1. Curva ROC	63
A.2. Curva Precisión-Sensibilidad	64
Anexo B. Método Out Of Bag (OOB)	68
Anexo C. Resultados obtenidos aplicando EMD (Empirical Mode Decomposition)	69
C.1. EMD (Empirical Mode Decomposition)	69
C.2. Resultados	72
Anexo D. Tablas y resultados	73
D.1. Ajuste de hiperparámetros del algoritmo Random Forest	73
D.2. Importancia de características obtenidas en módulo 2	77

Índice de Tablas

1.1.	Tabla comparativa entre reglas AASM y WASM, para detectar PLM (obtenida de [33]).	4
1.2.	División de datos en conjuntos de entrenamiento, validación y prueba en [12]. .	11
1.3.	Resultados obtenidos por Carvelli et all. [12] sobre el conjunto de prueba (Tabla 1.2). Recall corresponde a la sensibilidad.	11
1.4.	Resultados comparativos obtenidos en [12], aplicando las distintas metodologías sobre un mismo conjunto de datos (60 registros pertenecientes al WSC).	12
3.1.	Número de registros por conjunto.	32
4.1.	Resultados obtenidos por registro en el conjunto de prueba.	45
4.2.	Resultados sobre el conjunto de prueba.	46
4.3.	Resultados de los distintos modelos ajustados y probados sobre el conjunto de validación. Los modelos ajustados siguen la notación de la Tabla D.1 para los distintos conjuntos de hiperparámetros.	48
4.4.	Resultados obtenidos por registro sobre el conjunto de prueba en la salida del Módulo 2.	49
4.5.	Resultados sobre el conjunto de prueba en salida del Módulo 2.	50
4.6.	Resultados comparativos con métodos previos, y la metodología propuesta según cada módulo, esto es, Módulo 1 (Mód 1) y Módulo 2 (Mód 2).	52
4.7.	Resultados comparativos de la metodología implementada en este documento con resultados del experimento desarrollado en Carvelli et al [12], donde replicaron los métodos de [5–8] y los probaron sobre un mismo conjunto de datos, esto es, 60 PSG pertenecientes al WSC marcados por 5 expertos distintos. . .	53
4.8.	PLMI obtenido aplicando regla WASM para la detección de PLM sobre cada registro del conjunto de prueba.	54
C.1.	Resultados sobre el conjunto de prueba en salida del módulo 2 aplicando EMD.	72
D.1.	Resultados de las distintas combinaciones de hiperparámetros para el algoritmo Random Forest y sus respectivos resultados aplicando OOB.	73

Índice de Ilustraciones

1.1.	Laboratorio de Sueño y Neurobiología Funcional del INTA: A: Instalación de electrodos en un niño para la realización de un polisomnograma. B: Sala de polisomnografía (figura obtenida de [32]).	2
1.2.	En la figura se observan 2000 muestras (10 [s]) de las mediciones de un polisomnograma. Entre las distintas señales se muestran el EMG y el ECG, que miden la tensión muscular en las piernas y la actividad cardiaca, respectivamente. Además, se observa el canal MOR (acrónimo para Movimientos Oculares Rápidos) en el EOG y los canales MOV PIERNA y MOV BRAZO, que registran los movimientos corporales en la zona inferior y superior del cuerpo, respectivamente. A su vez, los 5 últimos canales corresponden a canales del EEG.	3
1.3.	Se muestran tres señales del PSG durante una época (30 [s]). Las señales son los EMG de ambas piernas y el ECG. En cada canal de EMG hay un evento de LM marcado, que en conjunto conforman un movimiento bilateral. A: Canal de EMG de la pierna derecha, con una marca de LM del experto. B: Canal de EMG de la pierna izquierda, con una marca de LM del experto. C: Canal de electrocardiograma.	4
1.4.	Se muestran 9 señales del PSG durante una época (30 [s]) con un evento de movimiento periódico (PLM) sobre el EMG de la pierna derecha (Tib-R) y sobre el EMG de la pierna izquierda (Tib-L). Esta figura fue obtenida de [33].	5
1.5.	Muestra ejemplos de detecciones al aplicar distintos filtros pasa altos sobre el EMG con frecuencias de corte de 15 [Hz] y 30 [Hz]. Además, la figura de más a la derecha se aprecia la detección realizada aplicando el filtro adaptativo propuesto en Moore et al [8]. A: Señal original. B: Señal rectificada. C: Señal rectificada con filtro pasa altos ($fc = 15[Hz]$). D: Señal rectificada con filtro pasa altos ($fc = 30[Hz]$). E: Filtro adaptativo. Figura obtenida de [8].	9
1.6.	Esquema del funcionamiento del algoritmo propuesto en Carvelli et all. [12] donde además se observa la arquitectura de la red diseñada. Se puede observar que la entrada a la red neuronal consiste en las 32 ventanas de 0,5[s] concatenadas, y su salida consiste en un vector de dimensiones 1x32 que consiste en la etiqueta -hubo evento o no lo hubo- de cada una de las 32 ventanas de entrada. Figura obtenida de [12].	11
2.1.	Pasos para la obtención de la energía de Shannon sobre un segmento de 30 [s] ($N = 6000$) de señal con presencia de un evento de LM. A: Se observa la implementación de la ecuación 2.1. B: Se aplica ecuación 2.2. C: Se aplica ecuación (2.3).	16
2.2.	Periodograma WM aplicado sobre segmento de señal (señal A) de la Figura 2.1, con una ventana deslizante de 0.16 [s] de ancho.	17

2.3.	Resultado de aplicar el EMA ($S = 20$) sobre una señal ruidosa ($N = 5000$). A: Se observa una señal senoidal contaminada con ruido Gaussiano ($\mu = 0, \sigma = 1$). B: Se observa el resultado de aplicar el EMA sobre la señal ruidosa de A. . . .	18
2.4.	Se pueden observar distintas wavelets, tanto discretas (primera fila) como continuas (segunda fila).	19
2.5.	Wavelet mexican hat con $a_0 = 2^{\frac{1}{2}}$ y $b_0 = 0,5$. A: Tres ubicaciones consecutivas de la wavelet para un índice de escala m e índices de ubicación $n, n + 1, n + 2$. B: Tres ubicaciones consecutivas de la wavelet para un índice de escala $m + 1$ e índices de ubicación $n, n + 1, n + 2$. Se puede observar que al tener un índice de escala mayor, esto ocasiona que la wavelet sea más ancha, y por ende la posición entre dos posiciones consecutivas de la wavelet sea mayor, esto es que b sea más grande. Figura obtenida de [40].	20
2.6.	Se observa el algoritmo que sigue la DWT para descomponer una señal en los distintos niveles. Entran los datos y al aplicar la DWT se obtienen los coeficientes de detalle, los cuales poseen las frecuencias más altas de la señal de entrada, por lo que se interpreta como un filtro pasa altos (PA). A su vez, los llamados coeficientes de aproximación quedan representados en PB, donde se poseen las componentes de la señal de entrada tales que su frecuencia sea menor a la frecuencia mínima de PA en el nivel respectivo. De esta manera, los coeficientes de aproximación se pueden interpretar como la salida de aplicar un filtro pasa bajos (PB) a la señal de entrada. Luego, en cada nivel se separa la señal de entrada en una componente de altas frecuencias (coeficientes de detalle o PA) y componentes de bajas frecuencias (coeficientes de aproximación o PB). Para el primer nivel la señal de entrada corresponde a la señal original, mientras que en niveles sucesivos, la señal de entrada corresponde a los coeficientes de aproximación del nivel anterior.	21
2.7.	En la columna izquierda se observa una señal con frecuencia variable desde una frecuencia baja hacia una frecuencia mayor. A su vez, en la columna derecha se observan los distintos niveles de descomposición de la señal hasta un nivel 5. Se ve claramente que la DWT actúa como una secuencia de filtros pasa altos en cascada, desde una frecuencia de corte mayor (nivel 1) hasta una frecuencia de corte menor (nivel 5). De esta manera, al aplicar la DWT se puede descomponer una señal en sus distintas bandas de frecuencia.	22
2.8.	Ejemplo de funcionamiento de la HHT. En la primera señal se puede observar la señal original, la cual va cambiando tanto su amplitud como su frecuencia a lo largo del tiempo. En la segunda señal se puede observar la frecuencia instantánea estimada. A su vez, en la tercera señal se puede observar la magnitud instantánea estimada. Figura basada en figura 2.7 de [32].	23
2.9.	Ejemplo simple de árbol de decisión para la clasificación de si una persona está en forma o no lo está. Además, se pueden observar las distintas partes de un árbol de decisión: En azul, se representan los distintos nodos del árbol, siendo el nodo superior el nodo raíz, y los inferiores los nodos terminales que poseen las etiquetas (o clases). Finalmente, en verde se pueden apreciar las ramas del árbol de decisión.	24

2.10.	Ejemplo básico de funcionamiento del algoritmo Random Forest. Se puede observar el conjunto de n árboles, el número de nodos por árbol que es igual en todos los árboles en este ejemplo (no necesariamente es igual) y el flujo de decisión (nodos naranjos) sobre cada árbol. Luego sobre el conjunto de todos los resultados de los n árboles, se toma la moda obteniendo la clasificación final.	28
2.11.	Tasa de error sobre conjunto de entrenamiento aplicando método OOB (Out Of Bag). Se observan tres curvas, esto considerando distintos números de variables a la vez para generar los árboles. Se puede observar que considerar un número de variables $m = \sqrt{p}$ o $m = \log_2(p)$ (ambas curvas coinciden), posee mejores resultados que considerar todas las variables a la vez. Además, se puede observar para los tres modelos que el aumentar el número de árboles disminuye el error de clasificación.	28
3.1.	Ejemplo de 30 [s] del EMG de ambas piernas, donde se presentan marcas realizadas por el experto y que poseen un corrimiento en relación al evento real.	30
3.2.	Representación como proceso Gaussiano del conjunto de entrenamiento. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.	31
3.3.	Representación como proceso Gaussiano del conjunto de validación. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.	31
3.4.	Representación como proceso Gaussiano del conjunto de prueba. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.	32
3.5.	Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 10[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.	33
3.6.	Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 10[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.	33
3.7.	Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 10[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.	34
3.8.	Software para la detección de LM y PLM construido en el lenguaje de programación Python versión 3, y módulos que componen la metodología propuesta. La entrada al algoritmo corresponde a la señal del EMG filtrada con un filtro pasa banda (10-100 [Hz]). El primer módulo busca generar un gran número de candidatos a LM. El segundo módulo filtra la mayor cantidad de falsos positivos. El tercer módulo aplica regla de la WASM para detectar PLM. El algoritmo posee dos salidas, una correspondiente a los eventos de LM detectados y otra a los PLM detectados.	35

3.9.	Funcionamiento del Módulo 1, y cada uno de sus pasos. A: Se observa una época (30 [s]) del registro del EMG, la marca realizada por experto y la marca de posible candidato a evento de LM que se obtiene al final del Módulo 1. B: Se puede observar la energía Shannon de la señal visible en A. C: Corresponde a la transformada de Hilbert de la energía de Shannon de la señal obtenida en B. D: Acá se puede observar el último paso del algoritmo, donde se presenta la aplicación del EMA sobre la señal resultante en C, y el umbral correspondiente al segmento de señal de 30 [s] analizado. Además, se puede observar para cada ventana de 0,25 [s] el respectivo valor (media) que se utiliza para comparar con el umbral. Aquellas ventanas cuya media está por sobre el umbral, son consideradas como candidatos a LM. Finalmente las marcas que posean traslape entre ellas, son concatenadas para conformar las marcas finales de candidatos a eventos de LM, las que se pueden observar en A.	37
3.10.	Ejemplo de detección de máximos que están por sobre una amplitud umbral (10[μ V]) en un evento de LM.	38
3.11.	Extracción de características en Módulo 2. En la parte superior se observa un evento de LM sobre el EMG, el contexto que se toma en relación a su centro (en rojo el tiempo posterior y negro tiempo previo al centro del evento), y las características que se obtienen sobre el canal del EMG para distintos contextos dados ($\tau = 3[s]$ o $\tau = 10[s]$). Para un contexto de $\tau = 3[s]$ se obtiene la DWT en el EMG hasta un nivel de 5, y para cada una de las 6 señales obtenidas se extraen cada una de las características indicadas en la figura. En la parte inferior se observa el canal de movimiento de piernas, el contexto dado (que es el mismo al tomado para el EMG), y las características obtenidas. Finalmente, se puede notar que por evento de LM se obtienen 96 características en total.	39
3.12.	Ejemplo de cómo se obtienen zonas del EMG sin eventos de LM, y de esta manera tener ejemplos negativos para entrenar el algoritmo Random Forest. En la parte superior se puede observar que al tomar un punto 20[s] posterior al centro del evento, y luego tomar $\pm 0,5[s]$ (1[s] en total) sobre el punto desplazado existe otro evento de LM, de esta manera dicho segmento no se considera como ejemplo negativo. Por otro lado, en la parte inferior de la figura se observa que al tomar un punto 20[s] posterior al centro del evento, y luego tomar una ventana $\pm 0,5[s]$ sobre ese punto, no existe evento de LM, de esta manera dicho segmento se considera como ejemplo negativo.	40
3.13.	Distribución de ejemplos por clase aplicando el método presentado en la Figura 3.12. Se puede observar una cantidad de ejemplos similares, entre los segmentos del EMG con eventos de LM (Eventos positivos 1815) y entre los segmentos del EMG sin eventos de LM (Eventos negativos 1671).	41
4.1.	Ejemplo de falsos positivos de Tipo I en zonas de baja energía. A: Señal del EMG con falsos eventos detectados. B: Electrocardiograma, que ocasiona pequeños aumentos de amplitud en el EMG.	46
4.2.	Ejemplo de falsos positivos de Tipo II en zonas con artefactos del ECG. A: Señal del EMG con falsos eventos detectados. B: Electrocardiograma, que como se observa, contamina el EMG ocasionando que las amplitudes del EMG crezcan en función a pulsos del ECG.	47

4.3.	Ejemplo donde hay falsos positivos de Tipo III que ocurren en un comienzo del EMG, en las cuales el experto supone que el individuo está despierto, y de esta manera, los eventos que ocurren durante este periodo no son tomados en cuenta.	47
4.4.	Ejemplo de falsos positivos de Tipo IV . La marca de color roja corresponde al evento detectado en el Módulo 1, y la marca negra al evento detectado por el experto médico.	48
4.5.	Se puede observar la importancia de cada señal considerada, según contexto y canal (ver Figura 3.11), en relación a la suma de las importancias de las características obtenidas en cada señal para cada candidato propuesto en el Módulo 1 (ver Figura D.2). EMG y dwt0-5 corresponden al segmento del EMG y las descomposiciones de este segmento hasta un nivel 5 del candidato a evento para un contexto de $\pm 3[s]$ en relación al centro de la detección. A su vez, Contexto EMG corresponde al segmento del EMG para un contexto de $\pm 10[s]$, y Canal movimiento corresponde al canal de movimiento del PSG para un contexto de $\pm 3[s]$. Se puede apreciar que las señales más importantes de acuerdo a las características obtenidas y el algoritmo Random Forest entrenado, corresponden al segmento de $\pm 3[s]$ del EMG y a las descomposiciones mayores de este mismo.	51
A.1.	Curva ROC para un modelo de clasificación utilizando Random Forest. En naranja se obtiene una curva ROC con un $AUC = 0.9$. En azul se observa la clasificación con un $AUC = 0.5$ que corresponde a la peor curva ROC que se podría obtener.	63
A.2.	Curva Precisión-Sensibilidad para distintos umbrales. Mientras la curva esté más cercana a la zona superior derecha, el modelo posee una mejor clasificación, de esta manera el mejor modelo en esta figura viene dado por la curva morada, y el peor por la curva azul. Figura obtenida de [42].	65
A.3.	Curva Precisión-Sensibilidad para distintos umbrales para la característica de número umbrales sobre una magnitud de $10[\mu V]$, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.	65
A.4.	Curva Precisión-Sensibilidad para distintos umbrales para la característica PSD, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.	66
A.5.	Curva Precisión-Sensibilidad para distintos umbrales para la energía y un contexto de $\pm 10[s]$ en relación al centro de la detección, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.	66
C.1.	Funcionamiento de la EMD aplicada sobre una señal compuesta por una señal triangular más una señal senoidal. Se pueden observar los dos primeros modos, donde se comienza desde el modo de mayor frecuencia, hacia el modo de menor frecuencia. A: Señal original. B: Primer modo. C: Segundo modo. Figura basada en figura 2.6 de [32].	70

C.2.	Se observa una época (30 [s]) del EMG, con presencia de un evento de LM y además, el resultado de aplicar EMD y luego transformada de Hilbert sobre cada modo obtenido. Se puede observar hasta el modo 9 que se obtiene de la EMD, y su respectiva transformada de Hilbert para cada modo. En el gráfico de la amplitud se puede observar tanto la amplitud de la señal obtenida de la EMD, como la amplitud instantánea obtenida de la transformada de Hilbert. A su vez, en los gráficos de frecuencia, se puede observar la frecuencia instantánea obtenida de la transformada de Hilbert.	71
C.3.	Extracción de características en módulo 2 aplicando EMD hasta modo 3. En la parte superior se observa un evento de LM sobre el EMG, el contexto que se toma en relación a su centro (en rojo el tiempo posterior y negro tiempo previo al centro del evento), y las características que se obtienen sobre el canal del EMG para distintos contextos dados ($\tau = 3[s]$ o $\tau = 10[s]$). En la parte inferior se observa el canal de movimiento de piernas, el contexto dado (que es el mismo al tomado para el EMG), y las características obtenidas según el contexto τ dado. Finalmente, se puede notar que por evento de LM se obtienen 9 características en total.	71
D.1.	Se puede observar lo obtenido aplicando el método OOB para maximizar la métrica F1-score. Cada curva representa un conjunto diferente perteneciente al conjunto de entrenamiento, que fue dejado como prueba, para distintos valores de hiperparámetros del algoritmo Random Forest. Se desprende de la figura que el número de árboles mayores a una cantidad de 50, no reviste de mejoras al desempeño de la clasificación, además, hay conjuntos en que máximo se alcanza un F1-score de $\sim 0,7$ mientras que en otros se supera el $\sim 0,9$. Los resultados completos para cada una de las combinaciones de los distintos hiperparámetros se pueden observar en la Tabla D.1.	76
D.2.	Importancia de cada una de las características obtenidas para cada marca propuesta como candidato a evento de LM, en la salida del módulo 1.	78

Capítulo 1

Introducción

1.1. Sobre el estudio del sueño

El sueño es un proceso biológico de mucha importancia para los seres humanos. Se caracteriza por una pérdida reversible de la conciencia y de la voluntad, y por un aumento del umbral de respuesta a estímulos externos. Está regulado y controlado por los procesos circadiano y homeostático [25, 26].

El proceso circadiano consiste un proceso biológico generado de manera endógena en el organismo y comprende a aquellos procesos dentro de un sistema biológico que ocurren de manera regular. El ciclo diario de luz y oscuridad gobierna los cambios rítmicos en el comportamiento y/o fisiología de la mayoría de las especies, donde estos cambios están gobernados por un reloj biológico, que en los mamíferos se localizan en dos áreas del cerebro llamadas núcleos supraquiasmáticos [26]. Los ciclos circadianos establecidos por este reloj biológico ocurren en toda la naturaleza y tienen un período de aproximadamente 24 horas. Dicho reloj interno consta de una serie de genes y productos proteicos que codifican y regulan varios procesos en todo el cuerpo [27].

A su vez, el proceso homeostático puede ser entendido como un proceso de autorregulación conducente al mantenimiento de una relativa constancia en las composiciones y las propiedades del medio interno de un organismo.

La importancia del sueño para el desempeño conductual durante la vigilia está establecida desde hace mucho tiempo, pero las razones y los mecanismos subyacentes siguen siendo poco claros [24]. Estudios recientes postulan que el sueño puede desempeñar un papel en la deshabitación, lo que restaura la capacidad de atención y maximiza la disposición del animal para el aprendizaje y el comportamiento dirigido a objetivos durante la vigilia posterior. Además, como resultado de una capacidad atencional reducida se sugiere que la alteración del sueño puede cambiar la naturaleza del comportamiento, convirtiendo al sujeto en alguien más impredecible [24].

Otras hipótesis que cuentan con mayor respaldo sostiene que el sueño estimula el mantenimiento y la mejora de las estructuras sinápticas para, de esta manera, contribuir con los procesos de memoria y plasticidad cerebral [28–31].

1.2. Sobre la polisomnografía

La polisomnografía es un examen que se utiliza para recopilar parámetros fisiológicos durante el sueño. Las mediciones obtenidas por el examen se almacenan en un polisomnograma

(PSG) donde se captan distintas señales tales como el electroencefalograma (EEG), electro-oculograma (EOG), electromiograma (EMG), electrocardiograma (ECG), oximetría de pulso, el flujo de aire y el esfuerzo respiratorio, los cuales son usados regularmente para evaluar las causas subyacentes de las alteraciones del sueño u otro tipo de trastornos. La polisomnografía se utiliza, por ejemplo, para detectar trastornos respiratorios relacionados con el sueño, que incluyen apnea obstructiva del sueño (AOS), apnea central del sueño e hipoventilación-hipoxia relacionada con el sueño. La polisomnografía también se puede utilizar para evaluar otros trastornos del sueño como convulsiones nocturnas, narcolepsia, detección de movimientos periódicos en las extremidades y sueño con movimientos oculares rápidos.

En la Figura 1.1 se puede observar un paciente al cual se le están instalando los electrodos necesarios para obtener el polisomnograma en el Laboratorio de Sueño y Neurobiología Funcional del INTA. Cabe hacer notar que este examen no repercute de manera negativa en los pacientes, existiendo una remota posibilidad de ocasionar irritación de la piel por los adhesivos de los electrodos.

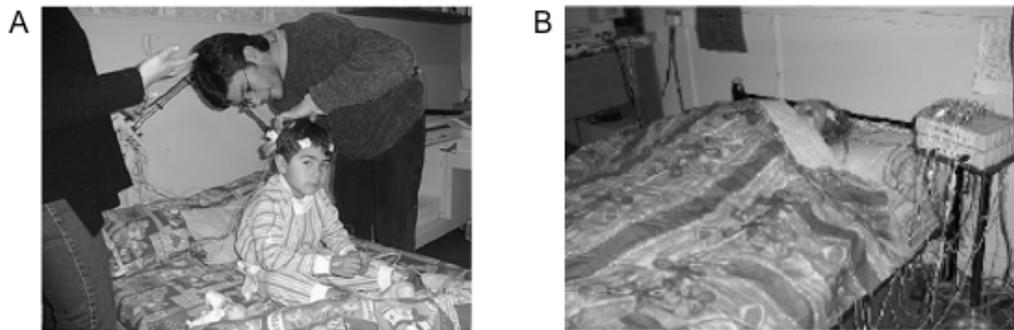


Figura 1.1: Laboratorio de Sueño y Neurobiología Funcional del INTA: A: Instalación de electrodos en un niño para la realización de un polisomnograma. B: Sala de polisomnografía (figura obtenida de [32]).

El número de canales medidos durante el examen para la obtención del PSG suele variar según el estudio y el centro de investigación. En la Figura 1.2, se puede observar un ejemplo de polisomnograma, donde se presentan 23 canales. Idealmente cada canal del PSG mide actividades fisiológicas distintas, pero normalmente los distintos canales son contaminados por la actividad en otros canales del PSG [5]. De esta manera la actividad en un canal de un proceso distinto al que se está midiendo, son considerados como artefactos, lo que indica que la actividad registrada no corresponde al proceso que se está evaluando, sino a un factor externo. Algunos eventos típicos que producen artefactos en los demás canales del polisomnograma son el electrocardiograma (ECG) y el movimiento rápido de los ojos los que son medidos por el EOG. Electrodo sueltos, o que estén puestos sobre zonas presentes de mucho sudor, también puede producir artefactos en las mediciones de los canales.

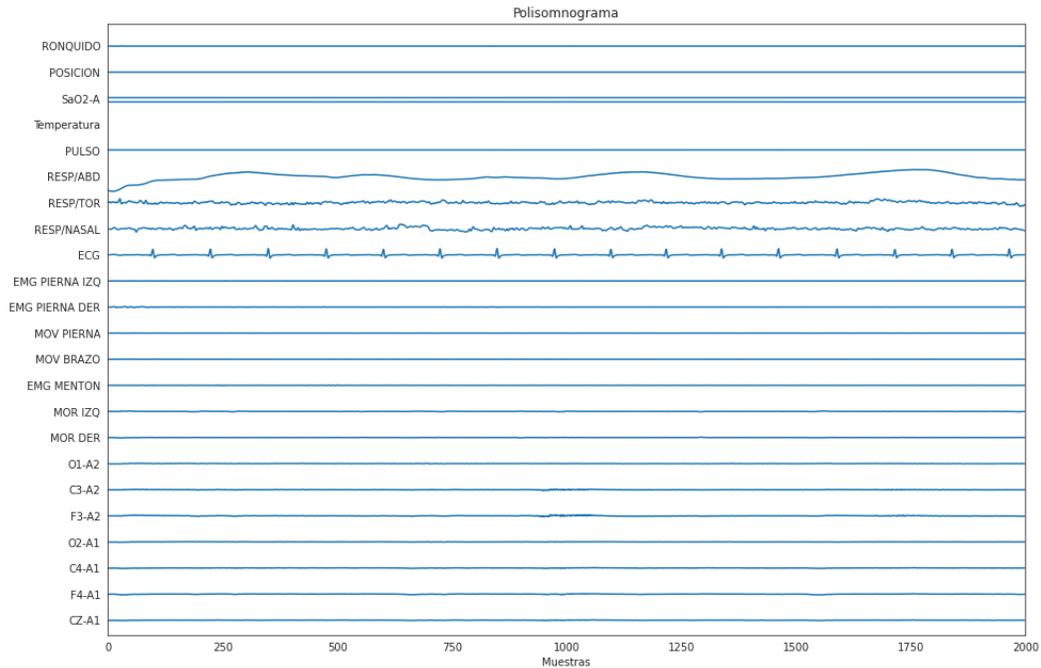


Figura 1.2: En la figura se observan 2000 muestras (10 [s]) de las mediciones de un polisomnograma. Entre las distintas señales se muestran el EMG y el ECG, que miden la tensión muscular en las piernas y la actividad cardiaca, respectivamente. Además, se observa el canal MOR (acrónimo para Movimientos Oculares Rápidos) en el EOG y los canales MOV PIERNA y MOV BRAZO, que registran los movimientos corporales en la zona inferior y superior del cuerpo, respectivamente. A su vez, los 5 últimos canales corresponden a canales del EEG.

1.3. Sobre los movimientos de piernas periódicos (PLM) y no periódicos (LM)

En el polisomnograma, la electromiografía (EMG) de piernas se utiliza para registrar los movimientos de piernas (LM), los cuales se clasifican en periódicos (PLM) o no periódicos, teniendo relevancia el primero mencionado. El trastorno llamado movimiento periódico de las extremidades durante el sueño (PLMD) - que es cuando los movimientos de las piernas se repiten periódicamente durante el sueño- tiene importancia debido a que ha sido asociado a otros trastornos como el RLS (Restless Legs Syndrome)[1], o con trastornos cardiovasculares [2–4]. El RLS es un síndrome frecuente, y se caracteriza por la necesidad de mover las piernas durante la tarde o al comienzo de la noche [26], trayendo consigo la dificultad para conciliar el sueño durante la noche. El conciliar mal el sueño tiene como consecuencia una serie de problemas, tales como la depresión, entre otros [27].

De acuerdo a las definiciones aceptadas por la WASM (World Association of Sleep Medicine) [13] y AASM (American Academy of Sleep Medicine) [14], se considera el inicio de un movimiento de pierna (LM) durante el sueño cuando ocurre un aumento de la amplitud del EMG de $8,0 \mu\text{V}$ o más por encima la línea de base en reposo, y termina cuando la amplitud del EMG disminuye a menos de $2,0 \mu\text{V}$ por encima del nivel de reposo, permaneciendo por debajo de ese umbral durante al menos $0,5 \text{ [s]}$. La duración del evento es el tiempo entre su

inicio y su finalización. Se considera válido cuando dura al menos 0,5 [s] y no más de 10,0 [s].

Los movimientos de piernas pueden ser bilaterales. Esto ocurre cuando se detectan LM en ambas piernas en el mismo instante (traslapadamente) o si la diferencia temporal entre el final de un evento y el comienzo del otro en la otra pierna no supera los 0,5 [s] (ver Figura 1.3).

Si bien tienen definiciones básicas concordantes respecto de que es un movimiento de pierna (LM), existen diferencias entre las definiciones de la WASM y la AASM para determinar si un evento LM pertenece a una secuencia de PLM, las que se observan en la Tabla 1.1.

Tabla 1.1: Tabla comparativa entre reglas AASM y WASM, para detectar PLM (obtenida de [33]).

	WASM	AASM
Dormido o despierto	Todo LM puede formar PLM	Solo LM que ocurren durante el sueño
Movimientos Bilaterales	Lapso Offset-Onset <0.5 [s]	Lapso Onset-Onset <5 [s]
LM relacionados con eventos respiratorios (RRLM)	Se excluyen de PLM	Se excluyen de PLM
Definición de RRLM	Cualquier LM que ocurra dentro de $\pm 0.5[s]$ del final de un evento de apnea hipoapnea	Cualquier LM que ocurra 0.5 [s] después del comienzo hasta 0.5 [s] antes del final de un evento de apnea, hipoapnea, relacionado con la excitación del esfuerzo respiratorio o trastornos respiratorios del sueño

El detectar un evento de movimiento de pierna (ya sea periódico o no periódico), consiste en marcar el inicio y el final de un movimiento de pierna en el EMG. Esto se puede observar en la Figura 1.3, donde se presentan dos eventos de movimientos de piernas marcados y el ECG, que consiste en la señal que causa mayor ruido sobre el EMG, y por ello dificulta la detección de eventos (LM) [5–12, 22].

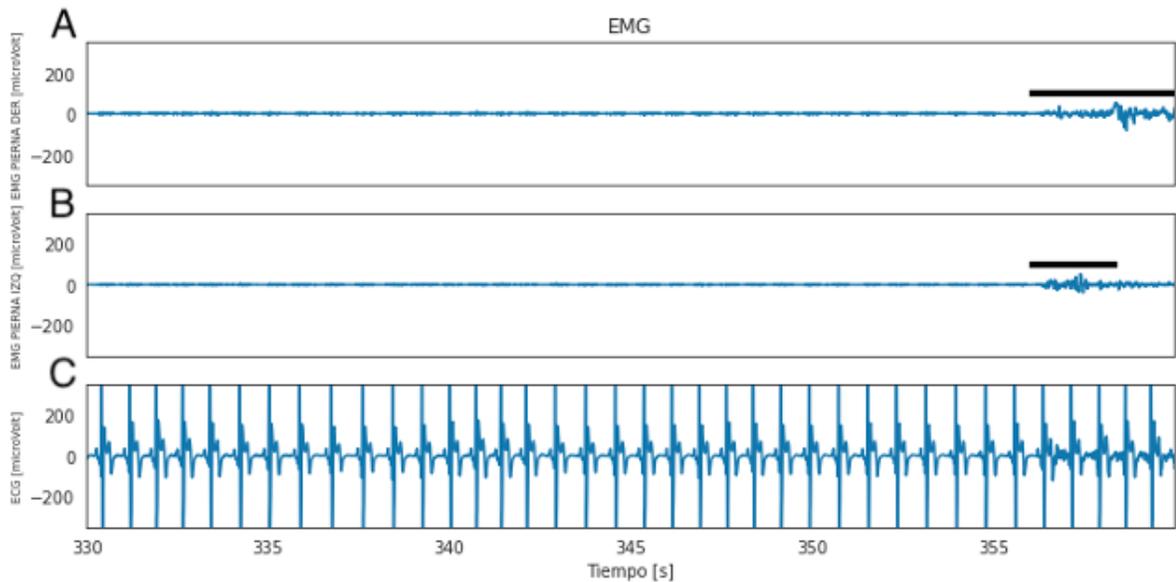


Figura 1.3: Se muestran tres señales del PSG durante una época (30 [s]). Las señales son los EMG de ambas piernas y el ECG. En cada canal de EMG hay un evento de LM marcado, que en conjunto conforman un movimiento bilateral. A: Canal de EMG de la pierna derecha, con una marca de LM del experto. B: Canal de EMG de la pierna izquierda, con una marca de LM del experto. C: Canal de electrocardiograma.

Además, como ya se mencionó previamente, eventos de LM pueden conformar una secuencia de PLM siguiendo alguna de las reglas presentadas en la Tabla 1.1. En la Figura 1.4 se puede observar un ejemplo de un movimiento de pierna periódico a partir de la definición de la WASM. Una vez descritas todas las series de PLM en un registro, se calcula el índice PLM (PLMI) que consiste en dividir el número total de PLM en un registro por la duración total del registro en horas. El parámetro PLMI es relevante, dado que es utilizado para estudios médicos [1, 33].

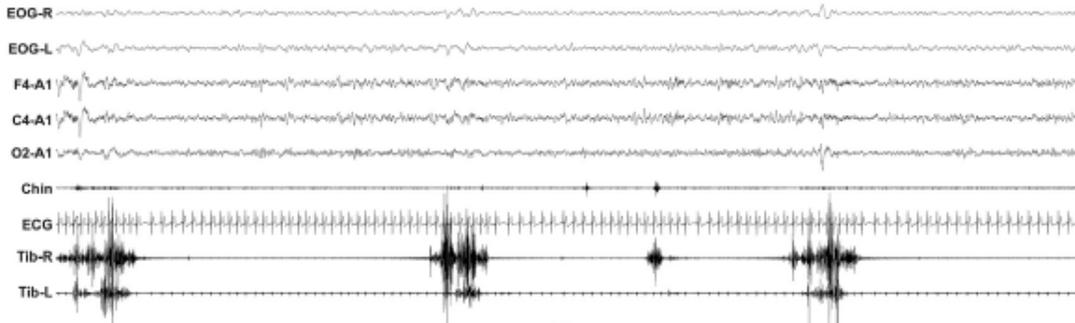


Figura 1.4: Se muestran 9 señales del PSG durante una época (30 [s]) con un evento de movimiento periódico (PLM) sobre el EMG de la pierna derecha (Tib-R) y sobre el EMG de la pierna izquierda (Tib-L). Esta figura fue obtenida de [33].

1.4. Estado del Arte en la detección automática de movimientos de piernas periódicos (PLM) y no periódicos (LM)

Se han desarrollado múltiples algoritmos para la detección automática de movimientos de piernas periódicos y no periódicos. Sin embargo el método más utilizado hasta ahora (Golden Standard) es aún por medio de marcas realizadas por un experto. Las marcas de experto presentan dos principales problemas: primero, existe un sesgo entre las marcas realizadas entre distintos expertos e incluso por un mismo experto (por fatiga) [12], y segundo, se requiere de una gran cantidad de tiempo de experto para realizar esta tarea.

Casi todos los algoritmos de detección automática utilizan métodos de filtrado de la señal, para luego estimar la línea base del EMG y aplicar umbrales para la detección de eventos (LM). Sólo se identificó un trabajo que no utilizó umbrales, que es el trabajo de Carvelli et al. [12], que proponen un modelo basado en redes neuronales.

1.4.1. Métodos basados en umbrales de amplitud del EMG

Polachmer et al. [5] utilizaron un conjunto de 10 grabaciones polisomnográficas con eventos de LM marcados de 5 pacientes con RLS obtenidas durante toda la noche (dos grabaciones por paciente). La media de las edades de los pacientes era de 49 años, y el rango de edades estuvo comprendido entre 37-58 años. Los 10 registros fueron separados en 5 de entrenamiento y 5 de prueba, seleccionando un registro de cada paciente de manera aleatoria para el conjunto de prueba.

Una vez separados los conjuntos, a cada canal del EMG le aplicaron un filtro pasa altos con frecuencia de corte en 16 [Hz] y un filtro notch con frecuencia de 50 [Hz].

Su metodología comienza por rectificar la señal filtrada, y luego aplicar criterios de amplitud y duración para identificar el inicio y el final de un evento: Cada vez que la amplitud de la señal es mayor que una amplitud umbral se marca el inicio del evento, y el final se marca cuando la amplitud de la señal vuelve a ser menor que la amplitud umbral, y se mantiene así por un rango de tiempo.

El algoritmo fue probado para distintos valores de la amplitud umbral (3, 5, 7 y 9 [μ V]), y para distintos rangos de tiempo (0,2, 0,15, 0,1, 0,05[s]) en que la amplitud fuera menor a la amplitud umbral para establecer el fin del episodio. Además se estableció una duración mínima de 0,5 [s] y una distancia mínima de 2[s] entre dos eventos consecutivos, de otra manera las marcas eran concatenadas y se consideraba una sola marca. No se indica como eligieron estos valores. Aplicando este algoritmo sobre el conjunto de prueba obtuvieron como resultados una sensibilidad de 93 % y una especificidad de 92 %.

En este trabajo, a pesar de que los umbrales tanto para amplitud como para la duración se hayan establecido sobre el conjunto de entrenamiento, estos pueden estar sesgados debido a la cantidad de datos, y más aún, debido a que el conjunto de entrenamiento y de prueba posee solo un registro de cada paciente. Por otro lado, pese a los aparentemente buenos resultados obtenidos (sobre el 90 % en ambas métricas) la especificidad reviste de poca importancia al estar trabajando con datos desbalanceados y además, no se reporta alguna métrica que de cuenta de los falsos positivos que pueda tener el modelo, dado que si bien se puede tener una muy buena sensibilidad (muchos verdaderos positivos) el modelo puede tener muchos falsos positivos, y estar detectando zonas del EMG como eventos de LM que en realidad no lo son.

Wetter et al. [6] utilizaron un conjunto de 29 registros polisomnográficos obtenidos durante el sueño del paciente (toda la noche), con eventos de LM marcados por dos expertos del sueño (cada registro marcado por ambos) de 10 pacientes con RLS. La media de las edades de los pacientes fue de 58 años. El conjunto de datos fue dividido en 5 registros de entrenamiento y 24 de prueba. Para el pre-procesamiento del EMG utilizaron un filtro pasa altos con una frecuencia de corte en 16 [Hz], y luego un filtro pasa bajo con frecuencia de corte en 300 [Hz]. Una vez filtrada la señal, la metodología se basa en tres criterios para determinar zonas del EMG que ellos denominaron como de “actividad” en la señal. Estos criterios son la tasa de repetición de máximos en la señal (denominada por ellos como “densidad de actividad”), la media de la amplitud, y la duración. Para determinar la densidad de actividad primero se truncan las amplitudes por sobre 30 [μ V]. Luego se toman ventanas sucesivas de 0,5 [s] en el EMG, para luego recorrer cada ventana (de 0,5 [s]) por medio de sub-ventanas sucesivas de 16 [ms] -sin traslape- donde se capturan “cambios bruscos y rápidos” (nombre dado por ellos) de la señal utilizando la desviación estándar (SD) en dicha ventana de 16 [ms]. Si la SD supera los 0,6 [μ V], la ventana es considerada con cambios bruscos y rápidos en la amplitud (criterio de amplitud). Si el 80 % o más de las ventanas de 16 [ms] de una ventana de 0,5 [s] son clasificadas con cambios bruscos y rápidos de amplitud, entonces la ventana de 0,5 [s] es etiquetada como un evento de LM (criterio de duración). Obtuvieron como resultado una sensibilidad del 92-95 % según marcas hechas por cada uno de los expertos (tomando como ground truth las marcas de un experto a la vez).

En este trabajo lo que más destaca es el uso de sub-ventanas, dado que esto le otorga al modelo la capacidad de ser más preciso a la hora de analizar cada parte del EMG, no obstante, el contexto dado al modelo es bastante pequeño -ventana de 0,5[s]- lo cual no es

adecuado desde el punto de vista de como un experto en sueño realiza el proceso de marcar los eventos de LM sobre el EMG. Además, el no otorgarle un contexto más amplio a cada ventana, esto debe ocasionar la detección de muchos falsos positivos, lo cual puede ser el motivo de que ninguna métrica que de cuenta de los falsos positivos haya sido presentada.

Ferri et al. [7] utilizaron una base de datos de 30 registros polisomnográficos con eventos de LM marcados sobre el EMG. Dentro de estos registros, 15 pertenecían a pacientes con RLS con una media de edad de 50,7 años, y 15 a pacientes sanos con una media de edad de 32,9 años. No se realizó una separación entre conjuntos de entrenamiento, validación y prueba. El pre-procesamiento que utilizaron consistió en aplicar un filtro pasabanda con frecuencias de corte de 10 y 100 [Hz], y un filtro Notch con frecuencia de corte de 50 [Hz]. Su metodología consiste en definir dos umbrales de manera arbitraria para marcar tanto el inicio como el final de un evento de LM en el EMG. Miden la amplitud media de una ventana móvil de 0,5 [s], y prueban con distintos valores de umbrales para esa media para definir el comienzo de un evento (5, 7, 10 y 15 [μ V]), mientras que el umbral para la finalización de un evento es dejada constante en 2 [μ V]. De esta manera detectan un evento marcando el inicio cuando el promedio de la amplitud de la ventana de 0,5 [s] es superior al valor umbral, y la finalización del evento se considera cuando la amplitud media de una ventana móvil de 0,5 [s] es inferior al umbral de término de 2 [μ V]. Aplicando esta metodología obtuvieron una sensibilidad y especificidad mayor al 90% (no se indica específicamente el valor) para todos los valores de umbrales de inicio elegidos.

En este trabajo al haber utilizado umbrales específicos y no haber separado en conjuntos de entrenamiento y prueba, ocurre un error metodológico, debido a que los umbrales que se seleccionaron de alguna manera deben haber sido ajustados mediante un método de prueba y error, en donde luego si se prueba los umbrales seleccionados sobre los mismos datos en que estos umbrales fueron ajustados, el modelo estará sobreajustado. De esta manera, pese a que obtuvieron un buen resultado (sobre el 90% en métricas evaluadas), estas pueden estar claramente sesgadas a los datos con los que trabajaron, en donde además ninguna métrica que mide falsos positivos fue presentada. No obstante, es rescatable el hecho de que las características basadas en la amplitud del EMG, pueden ser utilizadas con buenos resultados para detección de LM, y por implicancia, de PLM.

Huang et al. [9] utilizaron un conjunto de 24 PSG con eventos de LM marcados sobre el canal del EMG, en donde 15 de los pacientes tenía RLS y 9 eran de sujetos sanos. No realizaron separación entre conjuntos de entrenamiento, validación y prueba. Para el pre-procesamiento utilizaron un filtro pasa bajos con una frecuencia de corte de 225 [Hz] y un filtro pasa altos con una frecuencia de corte de 20 [Hz]. Su metodología parte por estimar la línea de base (o línea de reposo) del EMG: Se recorre el EMG desde el inicio buscando el primer intervalo de 10 [s] en los cuales el EMG es estable. De acuerdo a su clasificación, una señal del EMG se considera una señal estable si su amplitud máxima es menor a 20 [μ V] y menor a cinco veces la desviación estándar de la media de las magnitudes de la ventana de 10 [s]. Con esto se busca evitar que la línea base se estime sobre ventanas donde hay eventos de LM. Una vez obtenidos cinco intervalos de 10 [s] (no necesariamente consecutivos) que hayan sido clasificados como estables, la línea base se define tomando la mediana del fragmento de 10 [s] que posea la amplitud más grande.

Una vez estimada la línea base, utilizan la regla de la WASM para marcar el inicio y el final de un evento: Se marca el inicio si la amplitud está por sobre 8 [μ V] de la línea base, y

la finalización si la amplitud baja a no más de 2 [μV] por sobre la línea base, manteniéndose en ese nivel al menos de 0,5[s]. La duración del evento debe ser al menos 0,5 [s] y no más de 10 [s]. Obtuvieron como resultados una sensibilidad del 96 % y una especificidad del 92 %.

El mayor aporte de Huang et al. [9] es la metodología utilizada para estimar la línea base del EMG, no obstante, aunque la metodología es no paramétrica por lo que el ajuste de parámetros (o umbrales) no es necesario -separación en conjuntos de entrenamiento, validación y prueba no es necesario- el método para estimar la línea base debe haber sido diseñado por prueba y error, en donde por ejemplo la necesidad de obtener cinco -y no seis o cuatro por ejemplo- ventanas de 10[s] clasificadas como estables debe haber sido seleccionado de manera de obtener mejores resultados. De esta manera al ajustar dicho número de ventanas sobre los mismos datos que luego se prueba el modelo, hará que el modelo esté sobreajustado a esos datos. Además, pese a obtener una buena sensibilidad como resultado, no se presentan métricas relativas a los falsos positivos, y en cuanto a la especificidad está claramente no reviste de importancia dado que se está trabajando con datos desbalanceados, por lo que dicha métrica se sobre ajusta a los verdaderos negativos, lo cual no tiene ningún interés fisiológico.

Moore et al. [8] utilizaron un conjunto de 78 PSG con eventos de LM marcados sobre el canal de EMG pertenecientes a la base de datos del WSC (Wisconsin Sleep Cohort) y SSC (Stanford Sleep Cohort) -60 del WSC y 18 del SSC-. Similar a lo realizado por Huang et al. [9], su metodología comienza por estimar una línea base (o de reposo). La línea base se estima tomando para cada muestra ($x[n]$) una ventana de 20 [s], 10 [s] antes de la muestra n , y 10 [s] después, para luego obtener la media de dicha ventana. Una vez obtenida la media, a partir de las ecuaciones (1.1), (1.2) y (1.3) se estiman los valores de umbrales que se consideran para el inicio y finalización de un evento. En la ecuación (1.1), m es la media de la amplitud de la ventana de 20 [s], n es la muestra considerada. A su vez, en la ecuación (1.2) U y L son constantes, que se fijan en 8[μV] y 2[μV] respectivamente:

$$a[n] = \begin{cases} m * \ln(m + 1) + U & \text{for } m \leq 50 \\ \infty & \text{for } m > 50 \end{cases} \quad (1.1)$$

$$b[n] = \frac{L}{U} a[n] \quad (1.2)$$

$$c[n] = \frac{a[n] + b[n]}{2} \quad (1.3)$$

Luego, tomando un número de muestras igual a $M=0,15*f$, con f la frecuencia de muestreo (200 [Hz]), se obtiene el RMS de la siguiente forma

$$y[n] = \sqrt{\frac{\sum_{i=-M/2}^{M/2} x(n-i)}{M}} \quad (1.4)$$

Se define el comienzo de un evento como la muestra en que $y(n)$ supera por primera vez el valor de $a(n)$, y la finalización del evento cuando el valor de $y(n)$ en una muestra n decaiga bajo $b(n)$ por al menos 0,05 [s] (10 muestras sucesivas). Notar que si $m>50$, simplemente no se considera dicha muestra como posible inicio de una marca y se prosigue con la muestra $n+1$. En la Figura 1.5 se puede observar el resultado de aplicar este filtro adaptativo y su comparación con filtros pasa altos a frecuencias de corte 15 [Hz] y 30 [Hz].

Aplicando su método obtuvieron como resultados en la detección de PLM, una sensibilidad

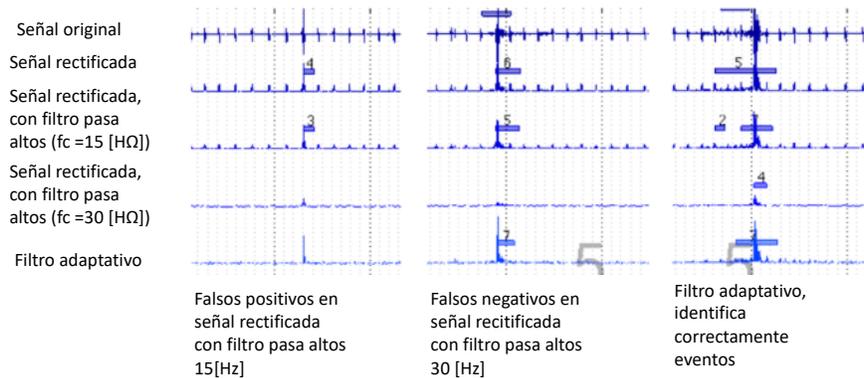


Figura 1.5: Muestra ejemplos de detecciones al aplicar distintos filtros pasa altos sobre el EMG con frecuencias de corte de 15 [Hz] y 30 [Hz]. Además, la figura de más a la derecha se aprecia la detección realizada aplicando el filtro adaptativo propuesto en Moore et al [8]. A: Señal original. B: Señal rectificada. C: Señal rectificada con filtro pasa altos ($f_c = 15[Hz]$). D: Señal rectificada con filtro pasa altos ($f_c = 30[Hz]$). E: Filtro adaptativo. Figura obtenida de [8].

del 0,82 una especificidad del 0,99.

Lo que más destaca de este trabajo es la manera en como estimaron la línea base, en donde el parámetro estimado $a(n)$ se le da rangos en los que puede estar, lo cual es relevante dado que esto permite filtrar aquellas magnitudes que no se deben a eventos de LM, sino que a ruido causado por el ECG. Además, el contexto dado a cada muestra, es suficiente para emular lo que realiza el experto en sueño a la hora de marcar eventos de LM sobre el EMG. Otro punto relevante sobre este trabajo, es lo que se observa en la Figura 1.5, donde se aprecia la necesidad de realizar algún tipo de filtro adaptativo para la detección de eventos LM, y que la utilización de filtros pasa altos para filtrar los artefactos producidos por el ECG, puede ocasionar muchos falsos positivos si la frecuencia de corte es cercana a 10 [Hz] y muchos falsos negativos si la frecuencia de corte es cercana a 30 [Hz]. En cuanto a los resultados, aunque la sensibilidad alcanzada es menor a la de otros trabajos ya analizados, la metodología es mucho más robusta lo cual ha sido demostrado en resultados alcanzados en trabajos posteriores [12], en donde se ha comparado este método con los métodos propuestos en [5–8], obteniendo mejores resultados en comparación a esos. Un punto relevante también es que utilizaron distintas bases de datos para realizar su metodología, lo cual hace que sus resultados sean mucho más robustos al generalizar sobre nuevos datos.

Finalmente los últimos tres trabajos son los de Kayed et al. [22], Pittman et al. [23] y Stefani et al. [11]. Estos trabajos son poco detallados en su metodología por lo que solo se presentan con el fin de realizar la documentación completa sobre lo que se ha realizado en el problema de crear un detector automático de LM, pero no se considerarán para comparar por lo anteriormente mencionado.

Kayed et al. [22] en 1990 presentaron el primer algoritmo para la detección automática de movimientos de piernas. Utilizaron un conjunto de datos de 10 pacientes con RLS, con una media de edad de 54,7 años de los pacientes. Su metodología está poco detallada; sólo se indica que utilizaron criterios de duración de los eventos, y que la amplitud del EMG debía estar por sobre el 50 % de la línea base (o reposo) del EMG. No se indica cómo obtuvieron la línea base. Además, no separaron los registros en conjuntos de entrenamiento, validación y prueba. Obtuvieron como resultados una sensibilidad de 0,94 y una especificidad del 0,85.

La metodología es bastante sencilla y poco detallada, en donde además la base de datos muy pequeña lo cual hace casi imposible sacar grandes conclusiones de este trabajo, no obstante, este trabajo constituye el primer intento de realizar un detector automático de LM y PLM, siendo su mayor aporte la utilización de umbrales sobre la amplitud del EMG, lo cual ha constituido la base para los trabajos posteriores.

Pittmann et al. [23] utilizaron un conjunto de 31 PSG con eventos marcados por tres expertos distintos (cada experto marcó cada PSG). La media de edad de los sujetos fue de 44,3 años. Desarrollaron un algoritmo para detectar no solo PLM sino que también eventos respiratorios, microdespertares y etapas de sueño. Su metodología es poco clara: aunque indican que utilizaron segmentación por bandas de frecuencia y potencia, no indican que umbrales o bandas. Además utilizaron lógica difusa para calcular niveles de membresía y la pertenencia de cada segmento de la amplitud del EMG, a conjuntos previamente definidos, de los que tampoco se indican sus características. Para la detección de eventos de LM utilizaron criterios de energía en segmentos del EMG, para marcar el inicio y el final del evento. No se indica como lo hicieron. Además no realizaron separación en conjuntos de entrenamiento y prueba. Aplicando su metodología presentan dos resultados, según el experto que fuera tomado como ground truth, un “Agreement” del 93,1-93,1-95,6 % y un kappa score del 0,68-0,66-0,7 respectivamente. Kappa score, al igual que F1-score suele ser usada como métrica de detección -o clasificación- cuando los datos están desbalanceados. De esta manera, en comparación a otros trabajos el kappa score es similar al F1-score. En cuanto al “Agreement”, no se indica como calcularon dicha métrica.

Stefani et al. [11] utilizaron 40 PSG con eventos de LM marcados sobre el canal del EMG. Veinte de estos registros pertenecían a pacientes con RLS, mientras que los otros 20 pertenecían a sujetos sanos. El método que desarrollaron forma parte de un software privado, por lo que su algoritmo no se explica. En su trabajo plantean que obtuvieron una sensibilidad del 95 %. No se indica si dividieron el conjunto de datos en conjunto de entrenamiento, validación y prueba.

1.4.2. Método basado en Aprendizaje de Máquinas

Carvelli et al. [12] utilizaron un conjunto de 800 PSG con eventos de LM marcados sobre el canal del EMG. El conjunto de datos se construyó a partir de tres bases de datos distintas: WSC (Wisconsin Sleep Cohort), SSC (Stanford Sleep Cohort) y MrOS (Osteoporotic Fractures in Men), de donde se seleccionaron (aleatoriamente) algunos registros, y posteriormente se separaron en conjuntos de entrenamiento, validación y prueba de acuerdo a la tabla 1.2.

Como pre-procesamiento de la señal aplicaron un filtro pasa banda con frecuencias de

Tabla 1.2: División de datos en conjuntos de entrenamiento, validación y prueba en [12].

	Entrenamiento	Validación	Prueba	Total
WSC	235	10	30	275
SSC	137	10	30	177
MrOS	283	33	32	348

corte de 10 y 100 [Hz]. Luego aplicaron el filtro adaptativo propuesto en Moore et al. [8] que ya fue analizado previamente-hay investigadores que participan en ambos trabajos-. Su metodología consistió en tomar ventanas de 16 [s] sucesivas (sin traslape) en el canal del EMG, para luego tomar ventanas de 0,5 [s] (sin traslape) dentro de la ventana de 16 [s], obteniendo un total de 32 ventanas de 0,5 [s] por cada 16 [s] de señal. Este conjunto de 32 ventanas se utilizó como entrada a una red neuronal. En la figura 2 se puede observar la entrada a la red (INPUT) y la arquitectura de la red propuesta.

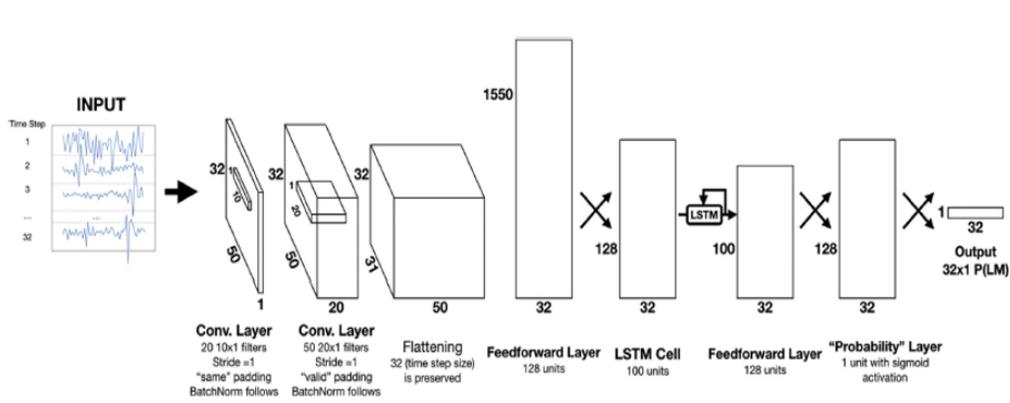


Figura 1.6: Esquema del funcionamiento del algoritmo propuesto en Carvelli et al. [12] donde además se observa la arquitectura de la red diseñada. Se puede observar que la entrada a la red neuronal consiste en las 32 ventanas de 0,5[s] concatenadas, y su salida consiste en un vector de dimensiones 1x32 que consiste en la etiqueta -hubo evento o no lo hubo- de cada una de las 32 ventanas de entrada. Figura obtenida de [12].

Los resultados obtenidos en el conjunto de prueba para la detección de LM, se presentan en la Tabla 1.3.

Tabla 1.3: Resultados obtenidos por Carvelli et al. [12] sobre el conjunto de prueba (Tabla 1.2). Recall corresponde a la sensibilidad.

	WSC	SSC	MrOs	Promedio
Precision	0.79	0.70	0.73	0.74
Recall	0.87	0.72	0.83	0.81
F1 score	0.83	0.71	0.77	0.77

Además, en [12] replicaron algunas de las metodologías ([5–8]) anteriores y procedieron a probarlas, junto con la desarrollada por ellos, en un sub-conjunto de 60 registros pertene-

cientes al WSC. Los resultados obtenidos se pueden observar en la Tabla 1.4.

Uno de los mayores logros de este trabajo es el haber utilizado una gran base de datos, lo cual lo hace un método bastante confiable y generalizable a nuevos datos. Los resultados obtenidos son comparables a los mejores resultados, en donde además se presentan dos métricas relevantes como la precisión y F1-score, donde la primera da cuenta de los falsos positivos y la segunda (F1-score) resume que tan bueno es el modelo tomando en cuenta la sensibilidad y la precisión. Este trabajo corresponde al único que utiliza técnicas de aprendizaje de máquinas, y en específico de aprendizaje profundo. Es de esperarse que la red entrenada por ellos posea resultados similares si se prueba en otros datos, esto dado el gran conjunto de entrenamiento utilizado para entrenar la red. Por otro lado, pese a que este método puede constituir la mejor metodología realizada, en cuanto a resultados robustos y generalizables a más datos, este método es una caja negra, esto es que los resultados obtenidos no son explicables lo cual es un problema desde el punto de vista fisiológico dado que desde este punto de vista, los resultados y su explicación son bastante relevantes.

El mayor logro de este trabajo es haber demostrado que la utilización de redes neuronales para la detección de eventos de LM es posible, en donde además se obtienen buenos resultados en comparación a los métodos anteriores sobre un mismo conjunto de datos (Tabla 1.4).

Tabla 1.4: Resultados comparativos obtenidos en [12], aplicando las distintas metodologías sobre un mismo conjunto de datos (60 registros pertenecientes al WSC).

Métrica	Tauch[5]	Wett[6]	Ferri[7]	Moore[8]	Moore SNR_+ [8]	Carvelli[12]
Precisión	0.79	0.47	0.62	0.88	0.85	0.81
Sensibilidad	0.24	0.96	0.85	0.60	0.73	0.90
F1-score	0.37	0.63	0.72	0.71	0.79	0.85

1.5. Objetivos

1.5.1. Objetivo General

Desarrollar un sistema para la detección automática de eventos de movimientos de piernas sobre el EMG utilizando la energía de Shannon y técnicas de procesamiento de señales, en registros polisomnográficos pertenecientes a datos del INTA de la Universidad de Chile.

1.5.2. Objetivos específicos

- Construir una base de datos de registros polisomnográficos con eventos de LM marcados sobre el canal del EMG.
- Realizar una correcta partición de los datos, buscando que en base a características de la señal, los conjuntos de entrenamiento, validación y prueba queden uniformemente representados.
- Obtener la energía de Shannon en cada registro.
- Construir primer bloque del modelo, el cual tiene por objetivo generar la mayor cantidad de candidatos a ser eventos de LM.
- Generar segundo bloque del modelo, que tiene por objetivo filtrar los eventos generados y generar las detecciones finales del modelo.
- Aplicar regla WASM para generar secuencias de PLM a partir de los LM detectados, entregando el índice PLM (PLMI) por registro.
- Definir las ventajas y desventajas del método y las mejoras que podría recibir en el futuro.

1.6. Organización del documento

A continuación se describe la estructura que presenta la memoria para los capítulos que siguen.

En el capítulo 2 se describen los aspectos teóricos necesarios para el entendimiento del documento. Se presentan las distintas metodologías y técnicas de análisis de señales, y algoritmos de aprendizaje de máquinas que se utilizan en la metodología implementada. Además, se presentan ejemplos de cómo funcionan y la utilización de cada una de estas técnicas y algoritmos, esto para lograr una mayor comprensión del documento.

En el capítulo 3 se describe la metodología utilizada tanto para la construcción de la base de datos, como para el desarrollo del sistema de detección automática de LM. En la primera sección se explica el método utilizado para la obtención de datos, y el tratamiento que se realizó sobre estos.

En la segunda sección se explican los tres bloques que posee el modelo, los cuales se explican de manera separada. Finalmente, se detallan las métricas y elementos que se utilizarán para evaluar el desempeño del modelo propuesto.

En el capítulo 4 se describen y analizan los resultados experimentales obtenidos. En la primera sección se muestran los resultados del sistema de detección automática de LM, y se

comparan con las marcas de LM que realizaron los expertos médicos. En la segunda sección se comparan los resultados con los obtenidos en trabajos previos mencionados en la revisión bibliográfica realizada en el capítulo 1, sobre el estado del arte en detección de eventos LM.

En el capítulo 5 se presentan las conclusiones obtenidas en la memoria. Finalmente se propone el trabajo futuro con el cual se podría mejorar lo realizado.

Capítulo 2

Antecedentes generales

2.1. Sobre la energía de Shannon

La energía de Shannon (ES) fue definida por primera vez en [35]. Para obtener la energía Shannon de una señal se comienza calculando la ecuación (2.1). Esto es, para una señal filtrada f con una secuencia de $n = 0, 1, \dots, N$ muestras se obtiene su señal rectificadas y normalizada $a[n]$

$$a[n] = \frac{|f[n]|}{\max_{n=1}^N |f[n]|} \quad (2.1)$$

Luego, para obtener la energía de Shannon se aplica la ecuación (2.2), sobre $a[n]$ que corresponde a la señal obtenida del paso anterior.

$$s[n] = -a[n]^2 * \log(a[n]^2) \quad (2.2)$$

Notar que estos dos pasos permiten obtener una señal que da cuenta no solo de eventos específicos en una zona de la señal, sino que también el contexto del segmento de señal que se esté analizando, sea también relevante. Finalmente, el último paso consiste en centrar la señal aplicando una normalización estándar dada por la ecuación (2.3), donde μ corresponde a la media de la ventana considerada y σ a la desviación estándar.

$$s[n] = \frac{s[n] - \mu}{\sigma} \quad (2.3)$$

En la Figura 2.1 se puede observar el resultado de aplicar las tres ecuaciones para la obtención de la energía de Shannon sobre un segmento de señal.

Un punto relevante y que diferencia la energía Shannon de la energía discreta dada por la ecuación (2.4), es que la ES da un contexto a la señal, haciendo que cada punto sea relevante por sí mismo, pero a la vez por el contexto en el que esté inmerso.

$$E_N = \sum_{n=0}^N |f[n]|^2 \quad (2.4)$$

Si se observa la ecuación (2.4), al obtener la energía de la señal se obtiene un solo número que representa a todo el segmento evaluado, y aún si se tomara para cada muestra de manera puntual su energía, esto haría que partes de alto ruido sean amplificadas notoriamente, lo cual haría que un modelo basado en umbrales de amplitud detecte eventos donde en realidad

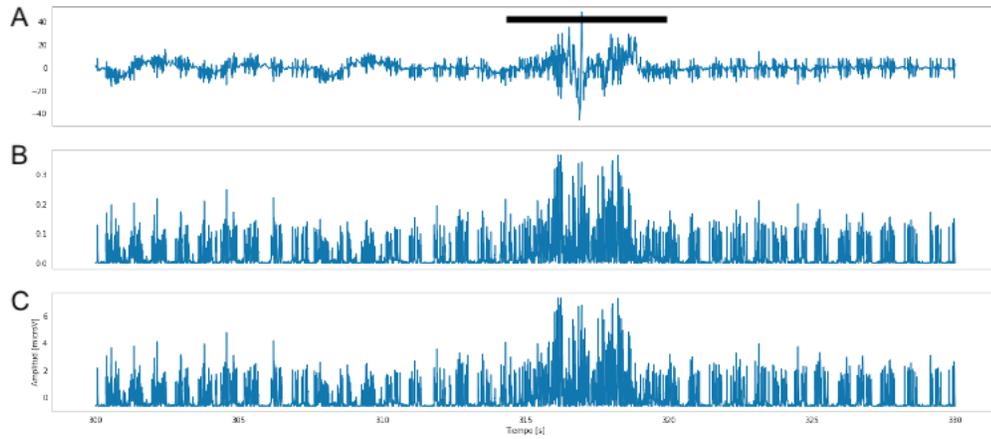


Figura 2.1: Pasos para la obtención de la energía de Shannon sobre un segmento de 30 [s] ($N = 6000$) de señal con presencia de un evento de LM. A: Se observa la implementación de la ecuación 2.1. B: Se aplica ecuación 2.2. C: Se aplica ecuación (2.3).

no hay. De esta manera, se tiene que para la detección de eventos en señales fisiológicas la utilización de la energía de Shannon puede resultar mucho más práctica que la energía discreta (2.4).

2.2. Método Welch

El método Welch (WM) consiste en una metodología para estimar la potencia espectral de la señal. Una de las principales ventajas del WM es que permite estimar la potencia de la señal con un bajo sesgo y varianza, esto al utilizar ventanas deslizantes y el promedio de la potencia estimada por ventanas. La potencia estimada por WM viene dada por la ecuación (2.5).

$$\hat{s}(w) = \frac{1}{KNU} \sum_{i=0}^{K-1} \left| \sum_{n=0}^{N-1} x_i[n]w[n]e^{-jwn} \right|^2 \quad (2.5)$$

Donde

- K : Número de sub-ventanas en el segmento de N muestras.
- N : Número de muestras.
- x_i : Segmento de señal dentro de sub-ventana.
- $w(n)$: Ventana rectangular, hamming u otra.

En la Figura 2.2 se puede observar el periodograma obtenido de aplicar el WM sobre un segmento de 30[s] de señal, que poseía un evento de LM.

De 2.2 se puede observar que para el segmento de señal (señal A) de la Figura 2.1 la mayor potencia espectral se concentra a una frecuencia de $\sim 10[Hz]$.

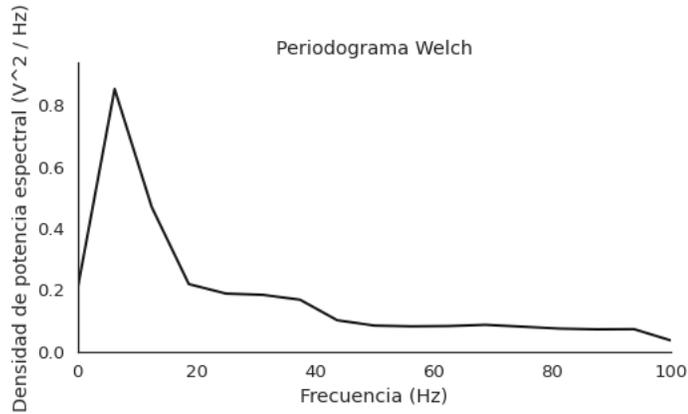


Figura 2.2: Periodograma WM aplicado sobre segmento de señal (señal A) de la Figura 2.1, con una ventana deslizante de 0.16 [s] de ancho.

2.3. EMA (Exponential Moving Average)

El EMA es un tipo de media móvil (MA) que otorga un mayor peso e importancia a los puntos de datos más recientes. El EMA también se conoce como media móvil ponderada exponencialmente, donde una de las características de estas medias, es que reaccionan de manera más significativa a los cambios de amplitudes recientes que una media móvil simple (SMA), la cual aplica la misma ponderación a todas las observaciones del período. El EMA sobre una secuencia de $n = 0, 1, \dots, N$ muestras, se obtiene aplicando la ecuación (2.6).

$$EMA[n] = a[n] \frac{S}{1 + N} + EMA[n - 1] \left(1 - \frac{S}{1 + N}\right) \quad (2.6)$$

Donde S, corresponde a un parámetro que se debe seleccionar y dependerá de que tanto se desea suavizar la señal. En la Figura 2.3 se puede observar el resultado de aplicar el EMA ($S = 20$) sobre una señal ruidosa.

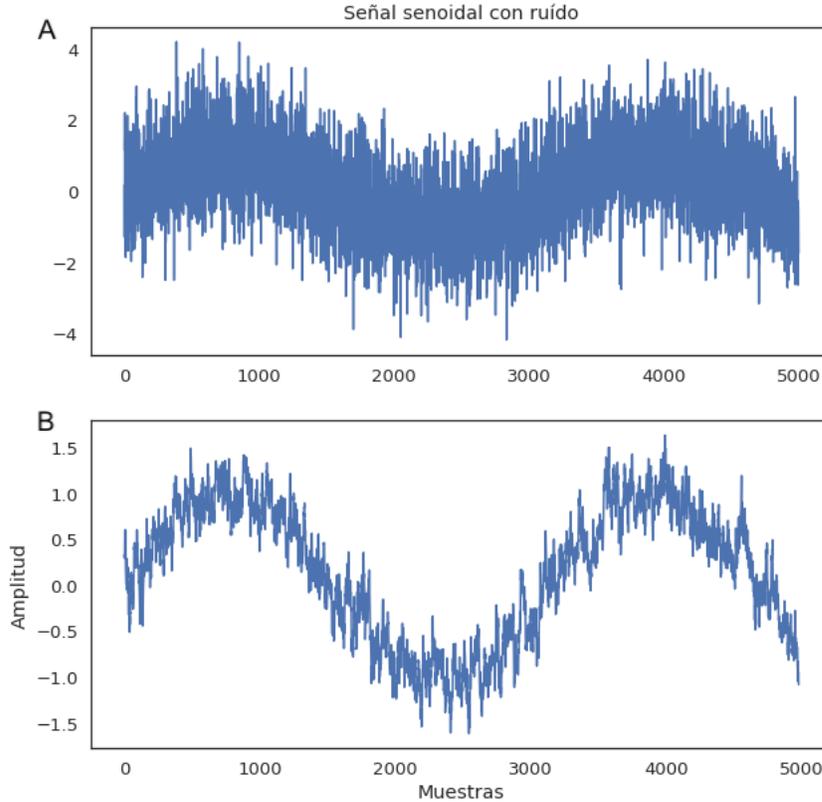


Figura 2.3: Resultado de aplicar el EMA ($S = 20$) sobre una señal ruidosa ($N = 5000$). A: Se observa una señal senoidal contaminada con ruido Gaussiano ($\mu = 0$, $\sigma = 1$). B: Se observa el resultado de aplicar el EMA sobre la señal ruidosa de A.

2.4. Transformada Wavelet Discreta (DWT)

Las wavelets corresponden a ondas con oscilaciones y formas específicas, en donde la forma de cada wavelet puede ser variada a través de un factor de escala y localización. Al obtener la convolución de una wavelet y una señal, se puede obtener información importante de la frecuencia de la señal, y el tiempo específico en que esto ocurre [43].

Una wavelet debe tener energía finita y media cero; esto garantiza que sea integrable y el producto interno entre la wavelet y la señal siempre exista. A su vez, la condición de admisibilidad implica que una wavelet debe tener media cero en el dominio del tiempo, lo cual es necesario para asegurar que sea integrable y que también se pueda calcular su inversa.

Luego, una wavelet ψ está representada por medio de la ecuación (2.7), donde a corresponde al factor de escala y b corresponde a la localización [43].

$$\psi_{a,b} = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.7)$$

Un factor de escala más alto (wavelet más larga) se corresponde con una frecuencia más pequeña, por lo que al escalar la wavelet en el dominio del tiempo, se analizan frecuencias más pequeñas (se logra una resolución más alta) en el dominio de la frecuencia. Y viceversa, al usar una escala más pequeña (a menor), se tiene más detalles en el dominio del tiempo.

Existen variadas familias de wavelets, algunas de las cuales se pueden observar en la Figura 2.4. Cada tipo de wavelets tiene una forma, suavidad y compacidad diferente, lo que les confiere utilidades distintas para distintos propósitos. Dado que solo hay dos condiciones matemáticas que debe satisfacer una wavelet (energía finita y media cero), es fácil generar un nuevo tipo de wavelet.

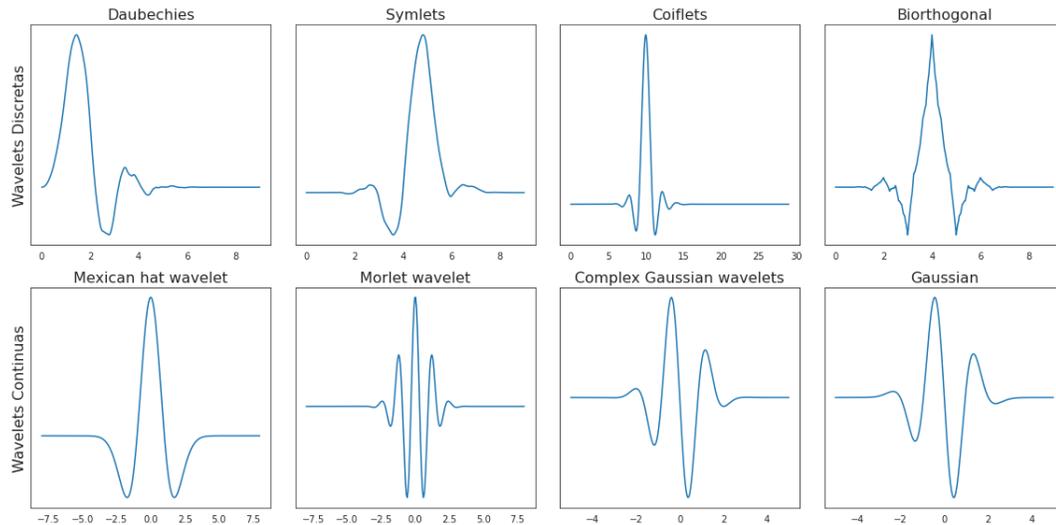


Figura 2.4: Se pueden observar distintas wavelets, tanto discretas (primera fila) como continuas (segunda fila).

A partir del uso de wavelets discretas, se define la convolución de la wavelet con una señal como la Transformada Wavelet Discreta (DWT), la cual permite obtener una representación tanto temporal como frecuencial de la señal convolucionada, lo cual le confiere su mayor ventaja frente a otros métodos de análisis de señales como lo es la FFT, la cual solo posee una interpretación en el espacio de la frecuencia de la señal.

Para obtener la DWT, se debe muestrear los parámetros a y b de la ecuación (2.4). Una forma de muestrear los parámetros a y b , es utilizar una discretización logarítmica de la escala a y vincularla, a su vez, al tamaño de los pasos dados entre b ubicaciones. Para vincular b con a , se debe mover en pasos discretos a cada ubicación b que sea proporcional al factor a ($\Delta b = b_0 a_0^m$). Este tipo de discretización de la wavelet tiene la forma dada por la ecuación (2.8). En la Figura 2.5 se puede observar un ejemplo de una wavelet, y lo que se obtiene al variar los parámetros a y b de acuerdo a la discretización descrita previamente.

$$\psi_{a,b} = \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \quad (2.8)$$

En la ecuación (2.8) los enteros m y n controlan la dilatación y traslación, respectivamente. A su vez, a_0 es un parámetro de paso de dilatación fijo establecido en un valor mayor que 1, y b_0 es el parámetro de ubicación que debe ser mayor que cero (ver Figura 2.5).

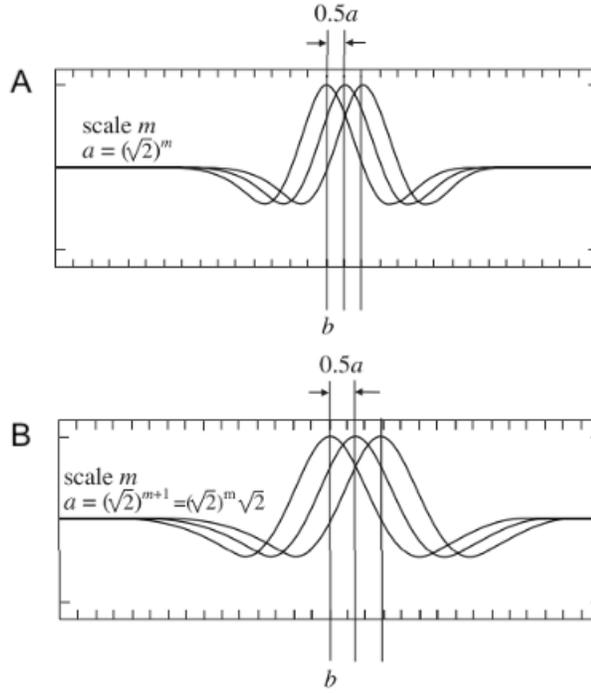


Figura 2.5: Wavelet mexican hat con $a_0 = 2^{\frac{1}{2}}$ y $b_0 = 0, 5$. A: Tres ubicaciones consecutivas de la wavelet para un índice de escala m e índices de ubicación $n, n + 1, n + 2$. B: Tres ubicaciones consecutivas de la wavelet para un índice de escala $m + 1$ e índices de ubicación $n, n + 1, n + 2$. Se puede observar que al tener un índice de escala mayor, esto ocasiona que la wavelet sea más ancha, y por ende la posición entre dos posiciones consecutivas de la wavelet sea mayor, esto es que b sea más grande. Figura obtenida de [40].

Luego, a partir de la ecuación (2.8) se define la DWT en la ecuación (2.9).

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) dt \quad (2.9)$$

Donde $T_{m,n}$ son los valores de la DWT para una grilla de valores m, n de la escala y ubicación, respectivamente. Los valores de $T_{m,n}$ usualmente son denominados como coeficientes de detalle. En cada nivel al aplicar la DWT, se obtienen dos señales, una con componentes de altas frecuencias (coeficientes de detalle) y una con componentes de bajas frecuencias (coeficientes de aproximación). De esta manera, al aplicar la DWT en una señal, se comienza con la escala más pequeña (frecuencias altas). Esto significa que primero se analiza el comportamiento de alta frecuencia. En la segunda etapa, la escala aumenta con un factor de dos (la frecuencia disminuye en un factor de dos), y por ende, se está analizando el comportamiento alrededor de la mitad de la frecuencia máxima. En la tercera etapa, el factor de escala es cuatro y se está analizando el comportamiento de la frecuencia alrededor de un cuarto de la frecuencia máxima. Esto se continúa de manera iterativa, hasta que se alcanza el nivel máximo de descomposición dado por la frecuencia de Nyquist, o por el nivel solicitado a la transformación. Lo anterior, queda representado en la Figura 2.6, donde se observa el algoritmo que sigue la DWT.

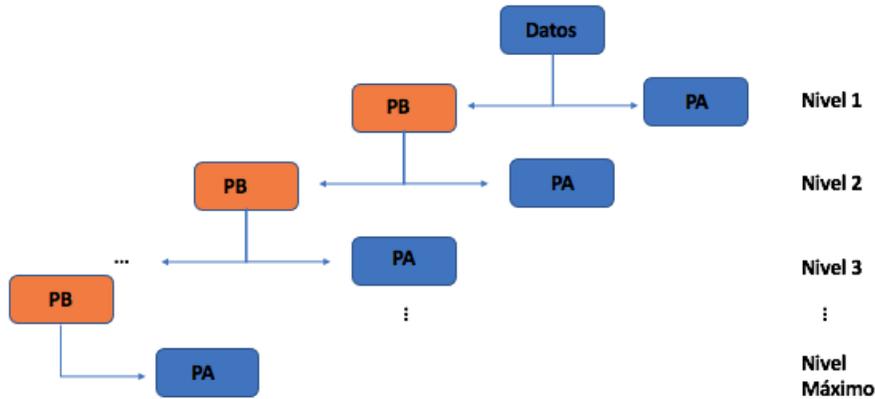


Figura 2.6: Se observa el algoritmo que sigue la DWT para descomponer una señal en los distintos niveles. Entran los datos y al aplicar la DWT se obtienen los coeficientes de detalle, los cuales poseen las frecuencias más altas de la señal de entrada, por lo que se interpreta como un filtro pasa altos (PA). A su vez, los llamados coeficientes de aproximación quedan representados en PB, donde se poseen las componentes de la señal de entrada tales que su frecuencia sea menor a la frecuencia mínima de PA en el nivel respectivo. De esta manera, los coeficientes de aproximación se pueden interpretar como la salida de aplicar un filtro pasa bajos (PB) a la señal de entrada. Luego, en cada nivel se separa la señal de entrada en una componente de altas frecuencias (coeficientes de detalle o PA) y componentes de bajas frecuencias (coeficientes de aproximación o PB). Para el primer nivel la señal de entrada corresponde a la señal original, mientras que en niveles sucesivos, la señal de entrada corresponde a los coeficientes de aproximación del nivel anterior.

A partir de lo anterior, se tiene que la DWT se puede utilizar para separar una señal en distintas bandas de frecuencia. En la Figura 2.7 se puede observar que al obtener distintos niveles se la DWT sobre una señal, permite descomponer dicha señal en diferentes bandas de frecuencia.

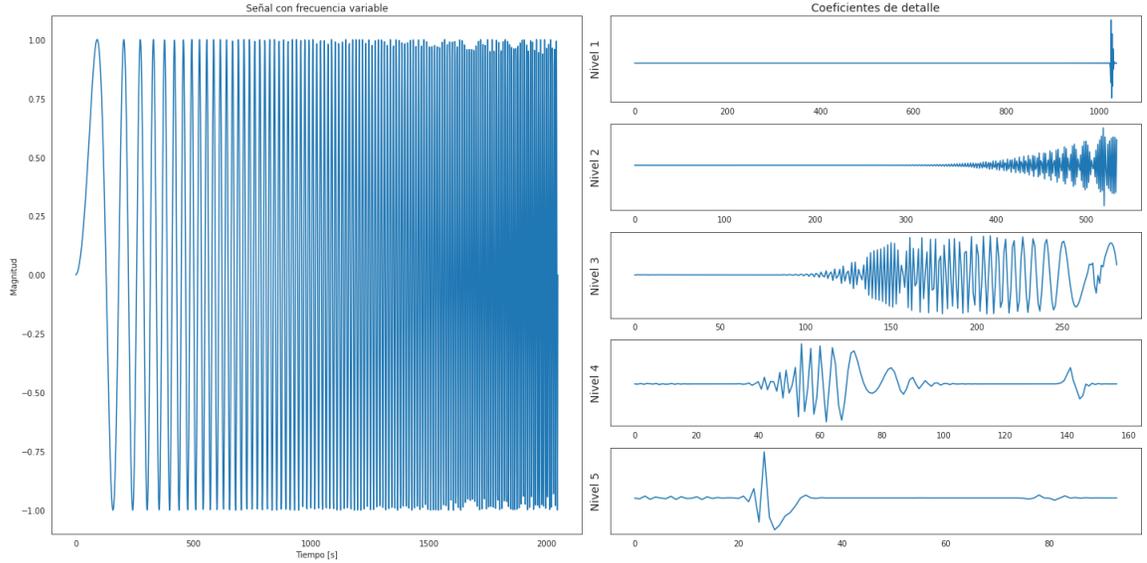


Figura 2.7: En la columna izquierda se observa una señal con frecuencia variable desde una frecuencia baja hacia una frecuencia mayor. A su vez, en la columna derecha se observan los distintos niveles de descomposición de la señal hasta un nivel 5. Se ve claramente que la DWT actúa como una secuencia de filtros pasa altos en cascada, desde una frecuencia de corte mayor (nivel 1) hasta una frecuencia de corte menor (nivel 5). De esta manera, al aplicar la DWT se puede descomponer una señal en sus distintas bandas de frecuencia.

2.5. Transformada de Hilbert-Huang (HHT)

La transformada de Hilbert (HT) corresponde a la convolución entre una señal $x(t)$ y la función $\frac{1}{\pi t}$ (ver ecuación (2.10)) [45]. A diferencia de la FFT, al aplicar la transformada de Hilbert se obtiene una representación de la señal tanto en frecuencia como en amplitud, las que se denominan como frecuencia y amplitud instantánea, respectivamente. La transformada de Hilbert se define en la ecuación (2.10).

$$h(t) = \frac{P}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (2.10)$$

Donde P es el valor principal de Cauchy de la integral singular [44]. La función $x(t)$ y su transformada de Hilbert $h(t)$, forman un par conjugado complejo, por lo que es posible definir una señal analítica $z(t)$ tal que:

$$z(t) = x(t) + h(t)i = a(t)e^{i\theta(t)} \quad (2.11)$$

$$a(t) = \sqrt{x(t)^2 + h(t)^2} \quad (2.12)$$

$$\theta(t) = \tan^{-1}\left(\frac{h(t)}{x(t)}\right) \quad (2.13)$$

Donde la $a(t)$ representa la amplitud instantánea y $\theta(t)$ la función de fase instantánea,

de la cual se obtiene la frecuencia instantánea w (ecuación (2.14)), la que es única para un tiempo dado.

$$w = \frac{d\theta(t)}{dt} \quad (2.14)$$

Uno de los problemas que surge de aplicar directamente la HT a una señal, es que esta debe cumplir ciertos requisitos con el propósito de que su frecuencia instantánea tenga significado físico [44, 45]. De esta manera, en [43] demostraron que descomponiendo la señal de entrada en señales oscilatorias en torno a cero, permiten obtener una frecuencia instantánea con interpretación física. Para ello, descomponen la señal utilizando EMD (Empirical Mode Decomposition) la cual descompone la serie temporal en modos base derivados de la señal de entrada (Funciones de Modos Intrínsecos, FMI). Luego, una vez separada la señal en las FMI, se obtiene la HT sobre cada uno de ellos. Esta metodología se conoce como transformada de Hilbert-Huang (HHT).

En la Figura 2.8 se puede observar la aplicación de la HHT sobre una señal, obteniendo su amplitud y frecuencia instantánea.

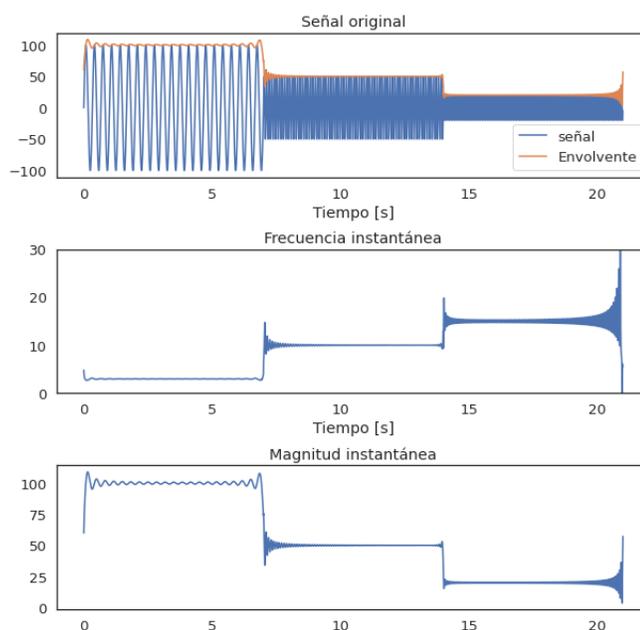


Figura 2.8: Ejemplo de funcionamiento de la HHT. En la primera señal se puede observar la señal original, la cual va cambiando tanto su amplitud como su frecuencia a lo largo del tiempo. En la segunda señal se puede observar la frecuencia instantánea estimada. A su vez, en la tercera señal se puede observar la magnitud instantánea estimada. Figura basada en figura 2.7 de [32].

Se puede observar en la Figura 2.8 que la HHT va estimando correctamente tanto la amplitud como la frecuencia que la señal original va teniendo a lo largo del tiempo.

2.6. Árboles de decisión y Algoritmo Random Forest

2.6.1. Árboles de decisión

Los árboles de decisión corresponden a una metodología de aprendizaje supervisado (o predictivo), de esta manera el objetivo es aprender un mapeo de las entradas (características) de un modelo a las salidas (etiquetas) [41].

Un árbol de decisión se representa como un diagrama de flujo con una estructura de árbol, donde cada nodo interno denota una prueba en un atributo, cada rama representa un resultado de la prueba y cada nodo hoja (nodo terminal) tiene una etiqueta de clase. El algoritmo comienza por el nodo raíz, y se prosigue a través del árbol por medio de las ramas que vayan cumpliendo con la pregunta planteada en los nodos respectivos. En la Figura 2.9 se puede observar un ejemplo sencillo de un árbol de decisión para la clasificación de si una persona está en forma o no.

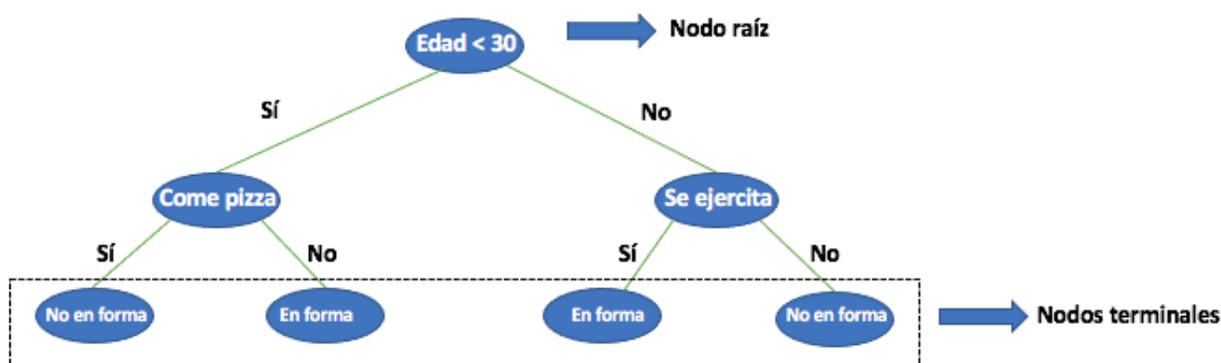


Figura 2.9: Ejemplo simple de árbol de decisión para la clasificación de si una persona está en forma o no lo está. Además, se pueden observar las distintas partes de un árbol de decisión: En azul, se representan los distintos nodos del árbol, siendo el nodo superior el nodo raíz, y los inferiores los nodos terminales que poseen las etiquetas (o clases). Finalmente, en verde se pueden apreciar las ramas del árbol de decisión.

El árbol se entrena buscando la división óptima de las características. Para ello existen diferentes criterios que se utilizan para medir la calidad de una división. Estos criterios son el índice de impureza Gini y la entropía. La impureza Gini mide la frecuencia con la que cualquier elemento del conjunto de datos se etiqueta incorrectamente cuando se etiqueta al azar. La impureza (o índice) Gini se obtiene por medio de la ecuación (2.15).

$$Gini = 1 - \sum_i p_i^2 \quad (2.15)$$

donde p_i es la probabilidad de la clase i . El valor mínimo del índice de Gini es 0. Esto sucede cuando el nodo es puro, esto significa que todos los elementos contenidos en el nodo son de una clase única. Por lo tanto, este nodo no se volverá a dividir. De esta manera, la división óptima está dada por las características que permiten obtener el valor más bajo en su índice Gini. Por otro lado, el valor máximo del índice Gini se obtiene cuando la probabilidad

de las distintas clases es la misma, esto es que el modelo no logra separar las clases a partir de las características ingresadas. La importancia de Gini o la disminución media de la impureza (MDI) calcula la importancia de cada característica como la suma del número de divisiones (en todos los árboles) que incluyen la característica, proporcionalmente al número de muestras que divide.

Similarmente al índice de Gini, para la entropía la división óptima es elegida por la característica con menos entropía. Obtiene su valor máximo cuando la probabilidad de las distintas clases es la misma y un nodo es puro cuando la entropía tiene su valor mínimo, que es 0. La entropía se calcula a partir de la ecuación (2.16).

$$\text{Entropía} = - \sum_i p_i \log_2(p_i) \quad (2.16)$$

Dentro de las ventajas que poseen los árboles de decisión se tiene que son simples y rápidos de construir, son robustos frente a la inclusión de variables irrelevantes e invariantes al escalar el problema. Además son de fácil interpretación, esto debido a que uno puede seguir el transcurso que se siguió en el árbol hasta el nodo terminal (clasificación final). De esta manera, tomando el ejemplo de la Figura 2.9, se puede inferir fácilmente que una persona es clasificada como estar en forma si su edad es menor a 30 años y no come pizza, o si es mayor a 30 años y realiza ejercicio.

Por otro lado, dentro de las desventajas de los árboles de decisión se puede mencionar su inestabilidad, esto es que cualquier pequeño cambio en los datos de entrada puede suponer un árbol de decisión completamente diferente. Además no se puede garantizar que el árbol generado sea el óptimo, donde hay conceptos que no son fácilmente aprendidos por el árbol, siendo fácil crear árboles sesgados. Y por último se puede mencionar que un árbol de decisión puede llegar a ser demasiado complejo con facilidad, perdiendo su utilidad (que sea fácil de interpretar).

2.6.2. Algoritmo Random Forest

Generalmente los árboles pequeños (pocas ramificaciones) tienen poca varianza, pero poseen un bias alto no representando bien la relación entre las entradas y las clases. Por otro lado, los árboles grandes se sobreajustan al conjunto de entrenamiento, por lo que tienen muy poco bias pero mucha varianza. Una forma de solucionar este problema son los métodos de ensamble, los cuales pueden ser separados en dos grandes grupos:

- **Bagging:** Es el diminutivo de bootstrap aggregation. Este método consiste en ajustar varios modelos sobre subconjuntos de datos tomados del conjunto de entrenamiento. Finalmente la clasificación (o predicción) viene dada como la media (variables continuas) de la salida de cada modelo ajustado, o como la moda (variables categóricas).
- **Boosting:** En este método se ajustan secuencialmente múltiples modelos, llamados weak learners, de forma que cada modelo aprende de los errores del anterior. Finalmente, para la clasificación (o predicción) se toma la media de todas las predicciones (variables continuas) o la moda (variables categóricas).

El algoritmo de Random Forest fue propuesto en [39], el cual se plantea como solución a la baja precisión y tendencia al sobreajuste de los árboles de decisión, y corresponde a una modificación del proceso de bagging que consigue mejorar los resultados gracias a que decorrelaciona aún más los árboles generados en el proceso. Para esto en vez de considerar todos

los predictores (características) como en bagging, toma un número que debe ser ajustado de manera tal de disminuir el error de predicción. Usualmente el número de predictores que se elige es $m = \sqrt{p}$, o $m = \log_2(p)$, con p el número total de predictores (ver Figura 2.11).

El algoritmo Random Forest crea una serie de árboles de decisión (bosque) sobre subconjuntos de datos del conjunto de entrenamiento, obteniendo una salida para cada árbol perteneciente al bosque (ver Figura 2.10). Finalmente para obtener la clasificación final se obtiene la moda (o la media si las variables son continuas) sobre el conjunto de salidas de cada árbol. Esto queda detallado de mejor manera en el algoritmo 1.

1. **para** $b=1$ en B **hacer**

- Tomar un conjunto Z^* de tamaño N del conjunto de entrenamiento.
- Crear un árbol sobre el conjunto Z^* , repitiendo recursivamente los siguientes pasos para cada nodo terminal del árbol, hasta que se alcance el tamaño mínimo de nodo n_{min} .
 - I) Seleccione m variables al azar de las p variables.
 - II) Elija la mejor variable / punto de división entre los m .
 - III) Divida el nodo en dos nodos secundarios.

fin

2. La salida corresponde al conjunto de árboles $\{T_b\}_1^B$.

3. Para clasificación: Sea $C_b(x)$ la predicción de clase del b -ésimo árbol. Entonces la clase obtenida del algoritmo es $C^B(x) = \text{Moda}(C_b(x)_1^B)$

Algoritmo 1: Algoritmo Random Forest, donde B : Número de árboles, Z^* : Sub-conjunto de datos tomados del conjunto de entrenamiento, p : Número de variables o características, $C_b(x)$: Clase obtenida por el árbol b -ésimo, $\{T_b\}_1^B$: Conjunto de árboles. Obtenido de [40].

Uno de los problemas que poseen los árboles de decisión es su inestabilidad, en donde dependiendo de los datos de entrada se pueden generar árboles completamente distintos, y por ende, clasificaciones distintas sobre nuevos datos. El algoritmo Random Forest, puede superar esto aumentando el número de árboles del bosque generando una reducción en la varianza de la clasificación.

Por otro lado, si bien un número de árboles mayor puede reducir el error de clasificación (Figura 2.11), con esto también aumenta la dificultad para interpretar los resultados. No obstante, existen distintos métodos para evaluar la importancia de las distintas características [40]. Uno de estos métodos consiste en utilizar el índice de pureza Gini. Esto se hace obteniendo para cada nodo de cada árbol cuánto se reduce la impureza al dividir los datos en las ramas y nodos sucesivos. Finalmente, todos los valores de reducción en la impureza se suman y se determina en promedio cuánto aporta cada variable a la disminución de la impureza.

Además del número de árboles, el algoritmo Random Forest posee una serie de hiperparámetros tales como número de nodos por árbol, número de características a tomar, entre otras. Una de las metodologías que se utiliza para ajustar los hiperparámetros del Random Forest, consiste en el OOB (Out Of Bag). Este método consiste en generar muestras por bootstrapping sobre el conjunto de entrenamiento, lo cual conlleva que, en promedio, cada ajuste use sólo aproximadamente dos tercios de las observaciones originales (Ver anexo B). Al tercio restante en el cual se prueba el modelo ajustado se le llama OOB.

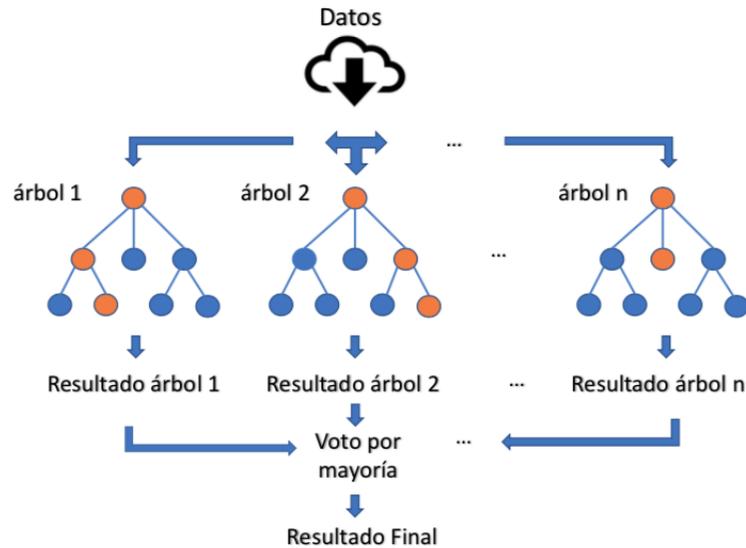


Figura 2.10: Ejemplo básico de funcionamiento del algoritmo Random Forest. Se puede observar el conjunto de n árboles, el número de nodos por árbol que es igual en todos los árboles en este ejemplo (no necesariamente es igual) y el flujo de decisión (nodos naranjos) sobre cada árbol. Luego sobre el conjunto de todos los resultados de los n árboles, se toma la moda obteniendo la clasificación final.

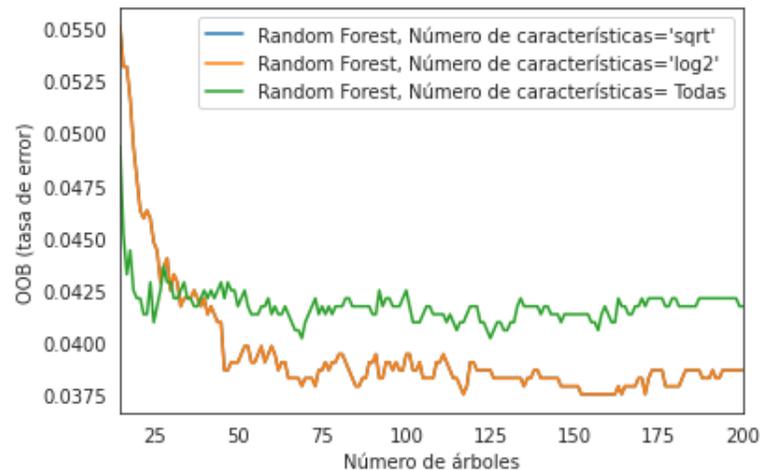


Figura 2.11: Tasa de error sobre conjunto de entrenamiento aplicando método OOB (Out Of Bag). Se observan tres curvas, esto considerando distintos números de variables a la vez para generar los árboles. Se puede observar que considerar un número de variables $m = \sqrt{p}$ o $m = \log_2(p)$ (ambas curvas coinciden), posee mejores resultados que considerar todas las variables a la vez. Además, se puede observar para los tres modelos que el aumentar el número de árboles disminuye el error de clasificación.

Capítulo 3

Metodología

3.1. Sobre los datos utilizados

Los registros que conforman la base de datos utilizada en este trabajo fueron obtenidos de sujetos sanos de ambos sexos en el Laboratorio de Sueño y Neurobiología Funcional, perteneciente al Instituto de Nutrición y Tecnología de los Alimentos (INTA) de la Universidad de Chile.

Se obtuvieron 27 registros polisomnográficos, de donde se seleccionaron 25 de ellos. Los registros que fueron dejados fuera de la base de datos, poseían una gran cantidad de ruido en el EMG y además muy pocos eventos marcados en base a la duración del registro.

Los registros de PSG se adquirieron durante el periodo de sueño nocturno en 27 pacientes distintos (un registro para cada paciente) en una habitación especialmente acondicionada para ello en el Laboratorio de Sueño y Neurobiología Funcional del INTA. El procedimiento se estandarizó para limitar las potenciales influencias del ambiente, los ritmos circadianos y/o de la ingesta de alimentos en los registros. Los pacientes fueron transportados desde sus casas al laboratorio, arribando una a dos horas antes de su horario normal de inicio de sueño, periodo durante el cual se les alimentaba, realizando además sus rutinas diarias previas a acostarse (lavarse los dientes, ponerse pijama, etc.); una vez en la cama, se procedía a realizar la colocación de los electrodos, y luego el paciente se disponía a dormir. La temperatura se mantenía constante (20-22°C) a lo largo de la sesión de registro. Durante la adquisición un experto permanecía en la habitación monitoreando todo el proceso. El protocolo de toma de muestras considera la aplicación de la PSG por dos noches consecutivas en cada sujeto [34].

Uno de los puntos relevantes sobre los datos utilizados es que solo están presentes las marcas de eventos relacionados a eventos LM, esto es, no existen marcas de experto para los PLM. Además de las marcas realizadas por el experto para eventos de LM, en muchas ocasiones estas presentaban un corrimiento en relación al evento real. Esto se representa en la Figura 3.1.

Esto hace patente la necesidad de lidiar de algún modo con aquellas marcas que están corridas, dado que el algoritmo se podría estar ajustando a zonas donde realmente no hubo evento de LM.

Otro punto relevante, es que en los eventos de LM detectados por el experto, no se realizó una distinción si estos tenían correlación con eventos respiratorios, de esta manera eventos RRLM (ver Tabla 1.1) no fueron filtrados por experto.

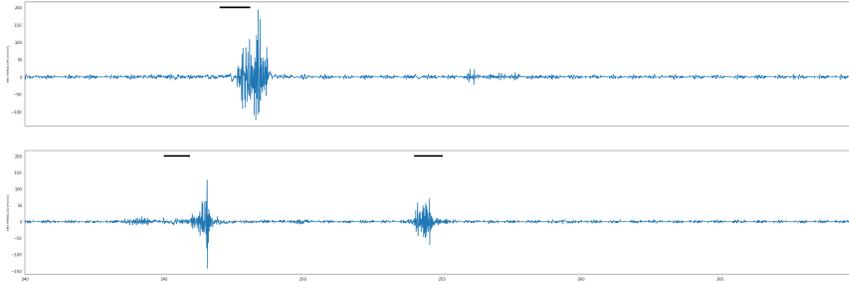


Figura 3.1: Ejemplo de 30 [s] del EMG de ambas piernas, donde se presentan marcas realizadas por el experto y que poseen un corrimiento en relación al evento real.

3.2. División de datos en conjuntos de entrenamiento, validación y prueba

Una de las necesidades básicas a la hora de desarrollar un algoritmo para detectar eventos de manera automática, es la correcta división de los datos en los conjuntos de entrenamiento, validación y prueba, en donde es imperativo que la variabilidad presente en los datos esté fielmente representada en cada uno de los conjuntos.

Para obtener una división pertinente de los datos, primero se aplica el siguiente algoritmo de caracterización de registros:

- Se comienza recorriendo cada uno de los registros por épocas -ventanas de 30 [s]- con traslape de 15 [s] entre épocas sucesivas -la mitad-.
- Luego, para cada época primero se estima su potencia espectral máxima utilizando el método Welch: Para ello, se recorre la ventana de 30 [s] por medio de sub-ventanas deslizantes de 0,16 [s] con traslape de la mitad entre ventanas sucesivas.
- Para cada época, se estima la energía de Shannon, y luego se toma la media de toda la ventana de 30 [s].
- Se guarda para cada época de 30 [s], la energía de Shannon media y la potencia espectral máxima.
- Se itera sobre cada época -con traslape de la mitad- en cada registro, realizando esto para todos los registros.

Una vez obtenida para cada época de cada registro estas dos características -energía de Shannon media y máxima potencia espectral- se procede a graficar cada registro como un proceso Gaussiano, donde cada punto en el plano representa a la energía Shannon media obtenida -eje x- y a la potencia espectral máxima -eje y-. Para obtener una mejor visualización para cada registro solo se gráfica su elipse correspondiente (proceso Gaussiano) y su media (punto central de la elipse).

Finalmente se prueban distintas combinaciones (de manera experimental) tales que permitan obtener una representación similar en los tres conjuntos (entrenamiento, validación y prueba).

Aplicando el método anterior se obtienen las siguientes representaciones. Para el conjunto de entrenamiento se obtiene la Figura 3.2.

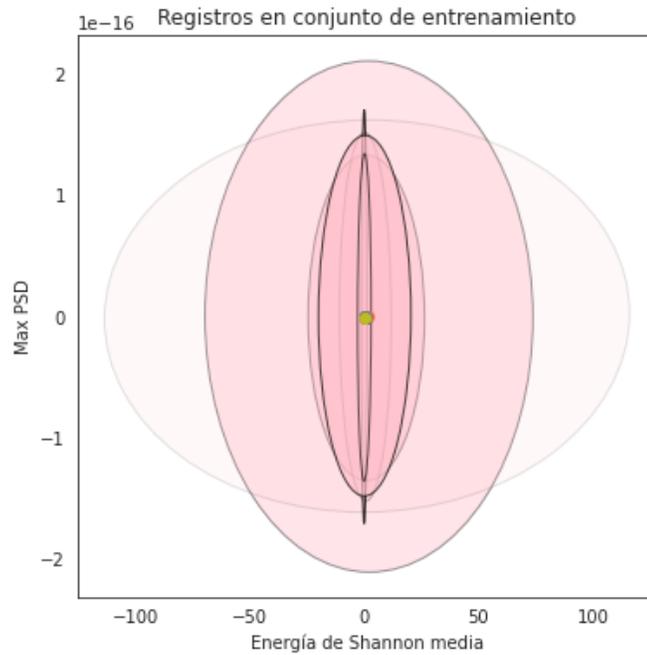


Figura 3.2: Representación como proceso Gaussiano del conjunto de entrenamiento. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.

A su vez, para el conjunto de validación se obtiene la Figura 3.3.

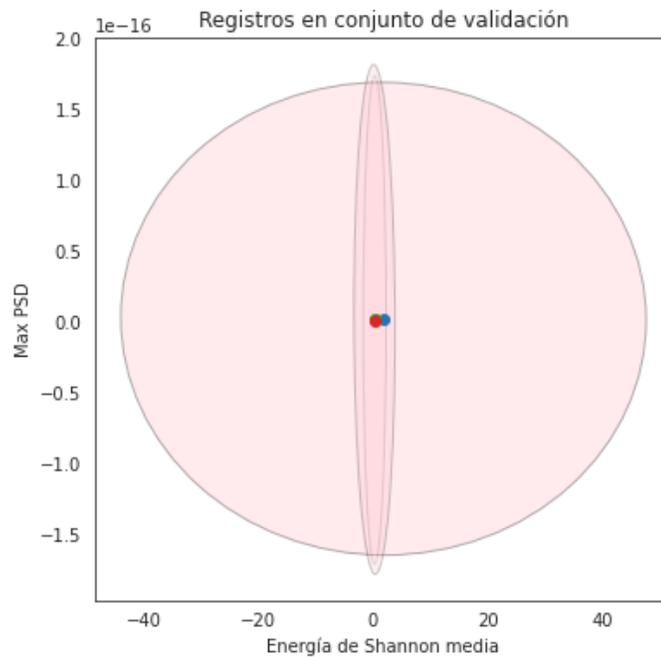


Figura 3.3: Representación como proceso Gaussiano del conjunto de validación. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.

Finalmente, para el conjunto de prueba se obtiene la Figura 3.4.

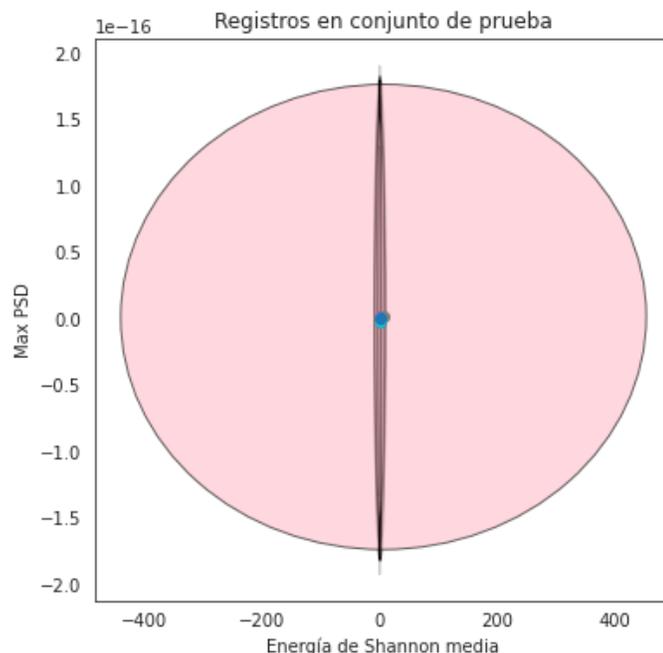


Figura 3.4: Representación como proceso Gaussiano del conjunto de prueba. Cada elipse corresponde a un registro completo, y el centro corresponde a la media de las características tomadas por época (30 [s]) de todo el registro.

Finalmente, en la Tabla 3.1 se puede observar el número de registros seleccionados para cada uno de los conjuntos.

Tabla 3.1: Número de registros por conjunto.

Conjunto	Número de registros
Entrenamiento	11
Validación	4
Prueba	10

3.3. Pre-Procesamiento de los datos

El principal artefacto que se produce sobre el EMG corresponde al que ocasiona el ECG. Normalmente se sugiere ([38]) que el ECG se filtre por un filtro pasa altos de frecuencia de corte $f_c = 30[Hz]$. No obstante, al igual que lo desarrollado en [8], se determina que al probar con distintos tipos de filtros y frecuencias de corte, un filtro pasa altos con frecuencia de corte de 30 [Hz], si bien filtra, también atenúa demasiado la señal del EMG.

En la Figura 3.5 se puede observar que un filtro pasa altos en $f_c = 10[Hz]$ permite atenuar el efecto producido por ECG sobre el EMG, no obstante, al tener frecuencias solapadas, el efecto del ECG aún es visible sobre el EMG.

En la Figura 3.6, se puede observar que a seleccionar una frecuencia de corte mayor

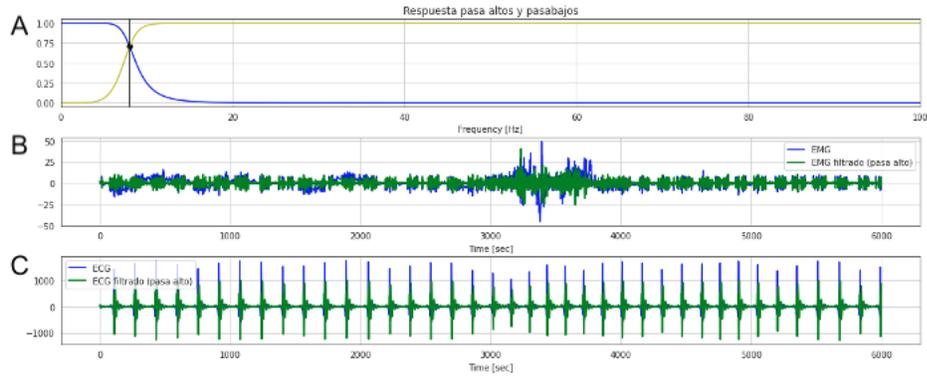


Figura 3.5: Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 10[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.

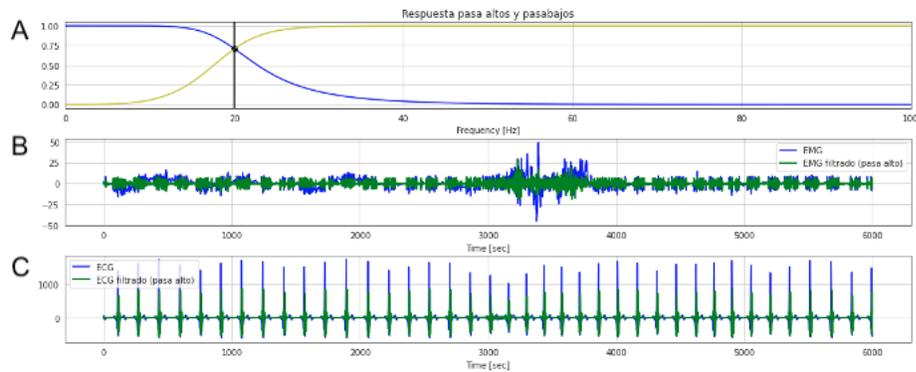


Figura 3.6: Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 20[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.

($f_c = 20[Hz]$) el efecto del ECG se filtra mucho mejor que con una frecuencia de corte de $f_c = 10[Hz]$, no obstante también se atenúa demasiado la señal del EMG.

En la Figura 3.7, se puede observar que al seleccionar una frecuencia de corte mayor ($f_c = 30[Hz]$) al igual que con una frecuencia de corte $f_c = 20[Hz]$ el efecto del ECG se filtra mucho mejor que en con una frecuencia de corte de $f_c = 10[Hz]$, pero la señal del EMG se vuelve demasiado pequeña

De esta manera, se tiene que al igual que en trabajos anteriores [8], se selecciona un filtro pasa banda de orden 4 con frecuencias de corte 10-100[Hz], esto dado que si bien frecuencias de corte mayores filtran mejor el ruido del ECG, también hacen disminuir demasiado la señal del EMG.

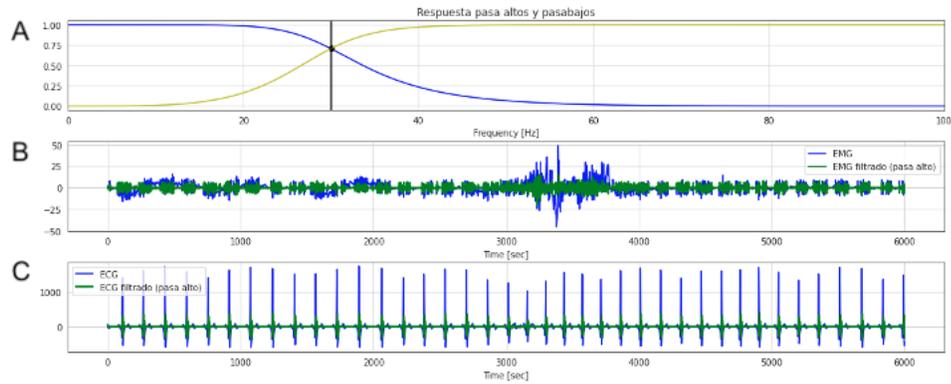


Figura 3.7: Representación del efecto de un filtro pasa altos con frecuencia de corte $f_c = 10[Hz]$ en 30[s] del EMG y ECG. A: Respuesta en frecuencia del filtro. B: Segmento del EMG real y filtrado. C: Segmento del ECG filtrado.

3.4. Método para la detección automática de movimientos de piernas

La metodología que se propone, está basada en tres módulos: El primer módulo tiene por objetivo generar la mayor cantidad de candidatos a eventos de LM. El segundo módulo tiene por objetivo filtrar la mayor cantidad de falsos positivos posibles. Finalmente, en el tercer módulo se aplica la regla de la WASM (ver Tabla 1.1) para generar candidatos a ser partes de secuencias de PLM. Cada uno de los módulos del método propuesto se construyeron en el lenguaje de programación Python versión 3, y forman parte de un software para la detección de eventos de movimientos de piernas sobre el EMG, en el cual se ingresa el polisomnograma y como salida se obtienen los eventos de LM y PLM detectados sobre el EMG. Esto queda ejemplificado en el siguiente diagrama (Figura 3.8).

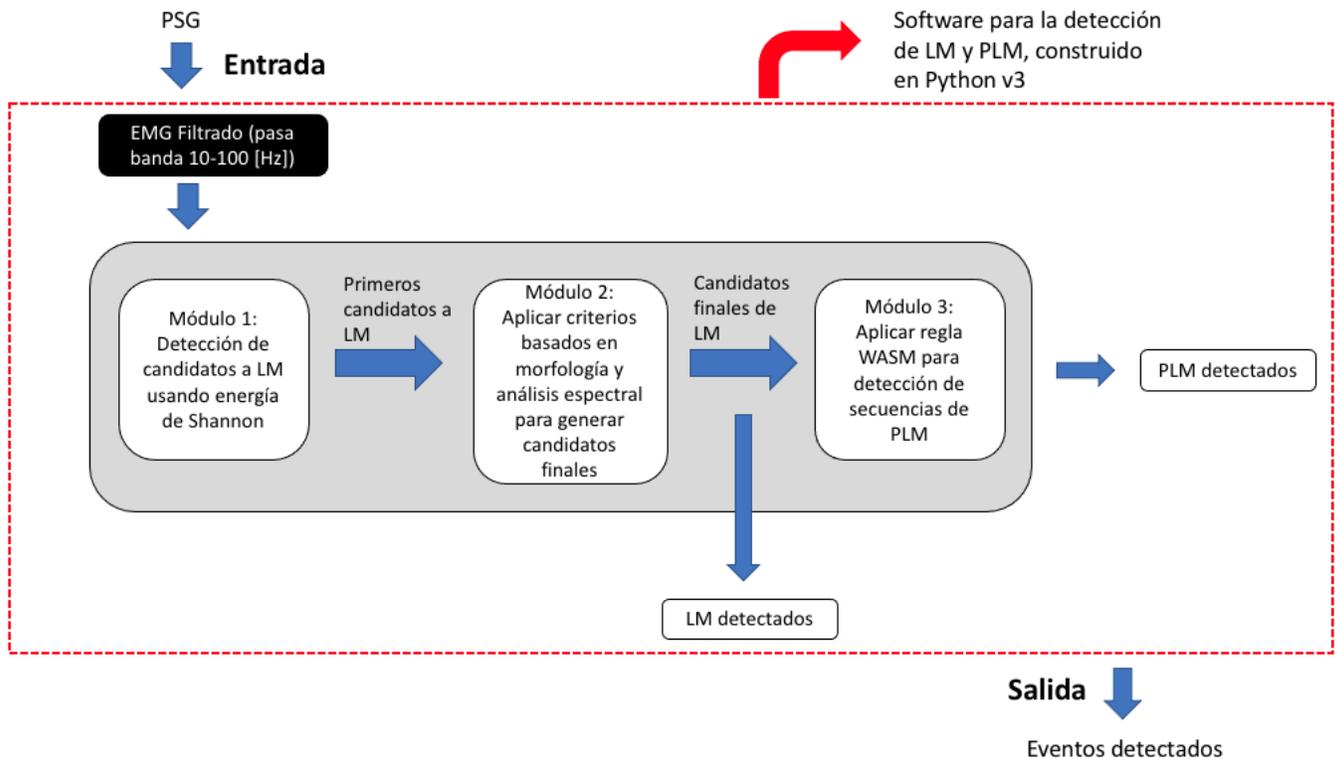


Figura 3.8: Software para la detección de LM y PLM construido en el lenguaje de programación Python versión 3, y módulos que componen la metodología propuesta. La entrada al algoritmo corresponde a la señal del EMG filtrada con un filtro pasa banda (10-100 [Hz]). El primer módulo busca generar un gran número de candidatos a LM. El segundo módulo filtra la mayor cantidad de falsos positivos. El tercer módulo aplica regla de la WASM para detectar PLM. El algoritmo posee dos salidas, una correspondiente a los eventos de LM detectados y otra a los PLM detectados.

En la subsección 3.1 se mencionó sobre el hecho de que las marcas de eventos de PLM no están presentes en los datos. De esta manera, la aplicación de métricas de desempeño sobre los eventos de PLM detectados no será posible, a diferencia de los eventos de LM, donde será posible realizar un análisis más acabado. Además, dado que en el proceso de realizar la detección de LM sobre el EMG en los datos con los que se prueba el modelo, el experto no realizó distinción de si el evento estaba relacionado con eventos respiratorios (RRLM), en este trabajo tampoco se tomará en cuenta si eventos de LM están correlacionados a eventos respiratorios.

3.4.1. Primer módulo

Este módulo busca generar la mayor cantidad de candidatos a eventos de LM. El algoritmo para generar candidatos que se sigue en este módulo es el siguiente:

1. Se recorre cada registro por épocas sucesivas (ventanas de 30 [s] sin traslape).
2. Se obtiene la energía de Shannon de toda la ventana de 30[s].
3. Se aplica la transformada de Hilbert sobre la señal resultante del punto anterior.

4. Se obtiene el EMA de la señal resultante del punto anterior.
5. Se obtiene el valor del umbral para detectar eventos. El umbral se seleccionó como $Umbral = \mu + \sigma$, que corresponde a la media de la ventana (μ) más la desviación estándar (σ) de la señal obtenida en el punto 4).
6. Una vez obtenido el umbral, se aplica una ventana deslizante de 0,25 [s] con traslape de la mitad (0,125[s]) sobre la señal obtenida en 4). Finalmente si la media de la amplitud de la ventana deslizante está por sobre el umbral del punto 5), se genera un candidato a LM. De otra manera no se considera esa ventana y se prosigue a la siguiente ventana de 0,25[s].

El funcionamiento del Módulo 1, y cada uno de sus pasos se puede observar en la Figura 3.9.

Finalmente, la salida del Módulo 1 corresponde a todas aquellas ventanas de 0,25 [s] que fueron consideradas como posibles eventos de LM. Aquellas marcas sucesivas que poseían traslape entre ellas, se concatenan y se consideran como un solo evento. De esta manera, todas las marcas que fueron concatenadas y que cumplan con el criterio de duración de la WASM, esto es que un evento de LM no puede durar más de 10[s] ni menos de 0,5 [s], corresponden a la salida del Módulo 1, y son los candidatos primarios a ser eventos de LM.

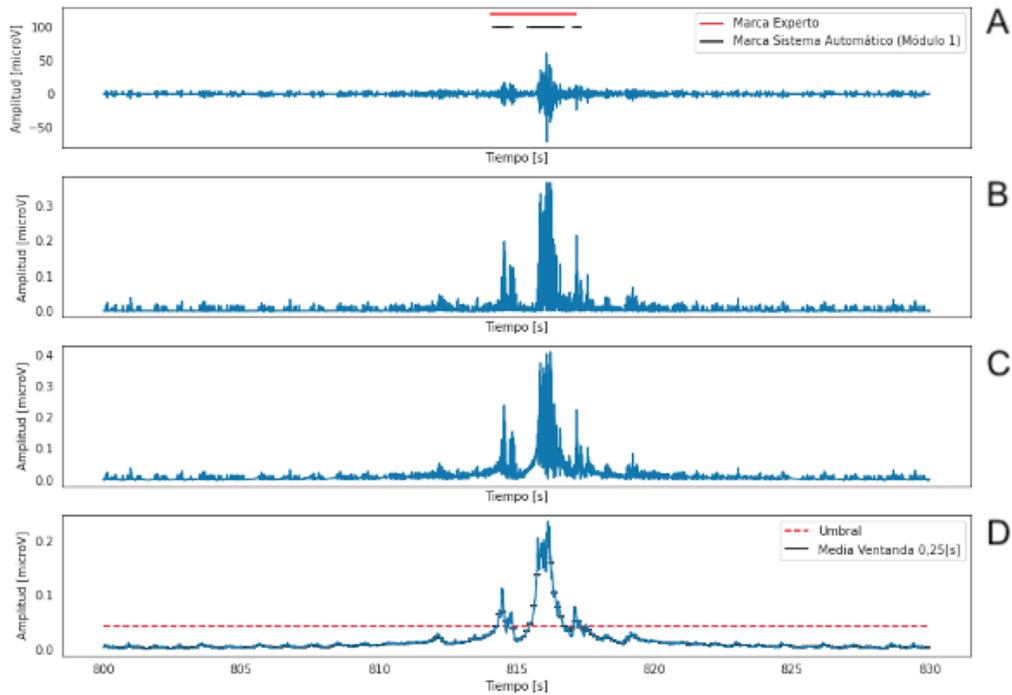


Figura 3.9: Funcionamiento del Módulo 1, y cada uno de sus pasos. A: Se observa una época (30 [s]) del registro del EMG, la marca realizada por experto y la marca de posible candidato a evento de LM que se obtiene al final del Módulo 1. B: Se puede observar la energía Shannon de la señal visible en A. C: Corresponde a la transformada de Hilbert de la energía de Shannon de la señal obtenida en B. D: Aquí se puede observar el último paso del algoritmo, donde se presenta la aplicación del EMA sobre la señal resultante en C, y el umbral correspondiente al segmento de señal de 30 [s] analizado. Además, se puede observar para cada ventana de 0,25 [s] el respectivo valor (media) que se utiliza para comparar con el umbral. Aquellas ventanas cuya media está por sobre el umbral, son consideradas como candidatos a LM. Finalmente las marcas que posean traslape entre ellas, son concatenadas para conformar las marcas finales de candidatos a eventos de LM, las que se pueden observar en A.

3.4.2. Segundo módulo

Este módulo tiene por objetivo filtrar la mayor cantidad de falsos positivos. Para ello se recorre cada candidato primario obtenido del Módulo 1, y se extraen distintas características las cuales permiten determinar si dicho candidato corresponde a un evento LM. La salida de este módulo corresponde a los eventos finales de LM detectados.

A continuación, primero se explican qué características son obtenidas para cada evento (marcas de experto) o candidato a evento del Módulo 1, y luego el algoritmo utilizado para seleccionar los umbrales sobre estas características que permitan filtrar la mayor cantidad de falsos positivos.

3.4.2.1. Sobre las características que se obtienen para lidiar con los Falsos Positivos

Filtrar falsos positivos relacionados a artefactos producidos por el ECG: Para cada uno de los candidatos obtenidos del primer módulo, se obtiene el centro de la marca del evento candidato a LM y se considera un contexto de $\pm 3[s]$ en relación a su centro. Luego, se obtiene el número de máximos que posee dicho segmento que están por sobre una magnitud de $10[\mu V]$, la cual se tomó en base a lo indicado por la WASM y AASM [13, 14] para la detección de eventos de LM, que requiere que la amplitud debe ser de al menos $8[\mu V]$ por sobre la línea base. Luego, utilizando el conjunto de validación y entrenamiento, se selecciona empíricamente un umbral cercano, que en este caso, se fija en un umbral de $10[\mu V]$. En la Figura 3.10 se puede observar un ejemplo de lo anterior.

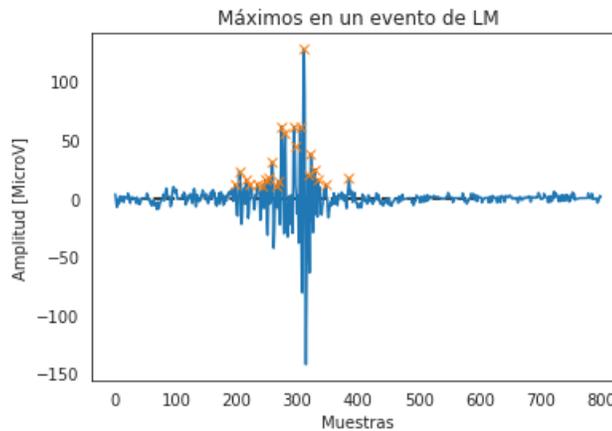


Figura 3.10: Ejemplo de detección de máximos que están por sobre una amplitud umbral ($10[\mu V]$) en un evento de LM.

Finalmente, el número de máximos que superan dicho umbral es ajustado utilizando el algoritmo Random Forest.

Filtrar falsos positivos en zonas de baja energía y potencia espectral:

Para filtrar este tipo de falsos positivos, se obtienen características basadas en la potencia espectral, la energía y la magnitud de la señal. Para ello, para cada marca propuesta en el Módulo 1, se consideran dos contextos de tiempo distintos, tomando como referencia el centro de la marca propuesta: un contexto de $\pm 3[s]$ y otro de $\pm 10[s]$. Esto se hace en base al análisis visual que realiza el experto médico para detectar eventos de LM sobre el EMG. Además, para un contexto de $\pm 3[s]$ se considera también el canal de movimiento de piernas, el cual consiste en un actígrafo que por medio de un acelerómetro mide movimiento, diferenciándose del EMG en que este último mide actividad eléctrica muscular. Luego, según el contexto de tiempo considerado (ver Figura 3.11), se extraen distintas características, las que se enuncian a continuación:

- Energía: Se obtiene la energía del segmento utilizando la ecuación (2.4).
- PSD: Se obtiene la PSD utilizando el Método Welch, y luego se obtiene el máximo de la PSD.
- Max frecuencia: La frecuencia donde se concentra el máximo de la PSD.

- Número de máximos sobre $10[\mu V]$.
- Estadísticos: De las magnitudes del segmento considerado, se obtiene la desviación estándar, la media, mediana, varianza, RMS y percentiles 5, 25, 75 y 95.

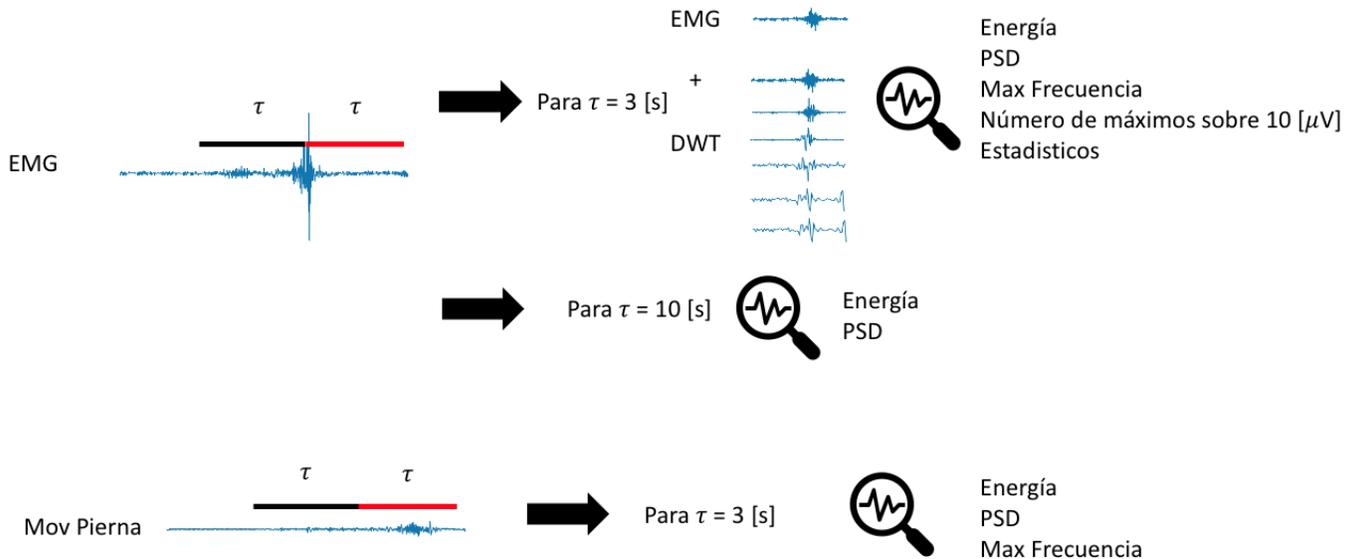


Figura 3.11: Extracción de características en Módulo 2. En la parte superior se observa un evento de LM sobre el EMG, el contexto que se toma en relación a su centro (en rojo el tiempo posterior y negro tiempo previo al centro del evento), y las características que se obtienen sobre el canal del EMG para distintos contextos dados ($\tau = 3[s]$ o $\tau = 10[s]$). Para un contexto de $\tau = 3[s]$ se obtiene la DWT en el EMG hasta un nivel de 5, y para cada una de las 6 señales obtenidas se extraen cada una de las características indicadas en la figura. En la parte inferior se observa el canal de movimiento de piernas, el contexto dado (que es el mismo al tomado para el EMG), y las características obtenidas. Finalmente, se puede notar que por evento de LM se obtienen 96 características en total.

De la Figura 3.11 se puede observar que por evento de LM propuesto en el Módulo 1, se extraen 9 señales: El segmento del EMG de $3[s]$ y las descomposiciones de este segmento hasta un nivel 5 (comenzando desde nivel cero), el segmento del EMG de $\pm 10[s]$ en relación al centro de la detección y el segmento de $\pm 3[s]$ del canal de movimiento (Mov Pierna en Figura 3.11) del PSG.

Análogamente para las características, se tiene que por evento de LM propuesto por el Módulo 1 se extraen 96 características: 2 para un contexto de $\pm 10[s]$ en relación al centro de la detección en el EMG, 3 para el canal de movimiento y 8 para cada señal que se obtiene de un contexto de $\pm 3[s]$ en el EMG (7 señales).

3.4.2.2. Algoritmo para ajustar umbrales

Del punto anterior se tiene que por evento de LM, ya sea candidato (validación y prueba) o un evento real marcado por el experto (entrenamiento), se obtienen 96 características. Esto

queda ejemplificado en la Figura 3.11, donde se observa que según contexto considerado del evento (3 [s] o 10 [s]) en relación a su centro, se extraen distintas características sobre el canal del EMG y el canal de movimiento de piernas.

Para ajustar los umbrales con los que debe cumplir cada candidato, se extraen las mismas características sobre el conjunto de entrenamiento, validación y prueba. Luego se entrenan distintos Random Forest (modelos), con distintos hiperparámetros entre sí. Como el algoritmo Random Forest requiere para su entrenamiento tener tanto ejemplos positivos (eventos de LM) como ejemplos negativos (zonas del EMG sin eventos de LM), los ejemplos positivos se obtienen utilizando las características obtenidas sobre todas las zonas del EMG con eventos detectados por el experto y que están en el conjunto de entrenamiento. Por otro lado, los ejemplos negativos son obtenidos sobre zonas del EMG sin eventos de LM. Para esto, para cada marca realizada por el experto se toma el centro de la marca y se desplaza a un punto 20[s] posterior sobre el EMG en relación al centro de la marca. Luego, una vez desplazado en 20[s] en relación al centro de la marca del experto, se toma un segmento $\pm 0,5$ [s] (1[s] de ancho) sobre dicho punto, y de esta manera si el segmento de 1[s] posee un IOU (ecuación (3.1)) de cero sobre todas las marcas realizadas por el experto, entonces las características obtenidas sobre dicho segmento son consideradas como ejemplo negativo, de otra manera ($IOU \neq 0$) se pasa a la marca siguiente realizada por el experto. Esto se puede apreciar de mejor forma en la Figura 3.12.

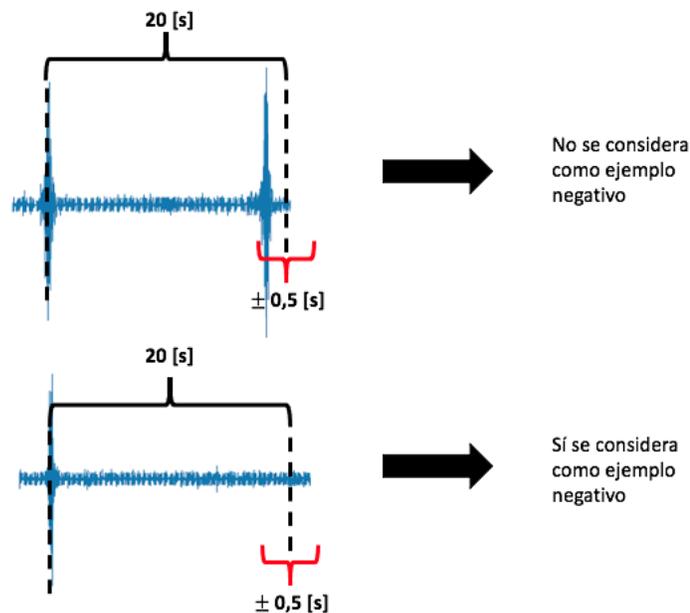


Figura 3.12: Ejemplo de cómo se obtienen zonas del EMG sin eventos de LM, y de esta manera tener ejemplos negativos para entrenar el algoritmo Random Forest. En la parte superior se puede observar que al tomar un punto 20[s] posterior al centro del evento, y luego tomar $\pm 0,5$ [s] (1[s] en total) sobre el punto desplazado existe otro evento de LM, de esta manera dicho segmento no se considera como ejemplo negativo. Por otro lado, en la parte inferior de la figura se observa que al tomar un punto 20[s] posterior al centro del evento, y luego tomar una ventana $\pm 0,5$ [s] sobre ese punto, no existe evento de LM, de esta manera dicho segmento se considera como ejemplo negativo.

Aplicando el método de la Figura 3.12 se obtiene una cantidad de ejemplos por clase presentados en la Figura 3.13, donde se aprecia una cantidad de ejemplos similares para ambas clases.

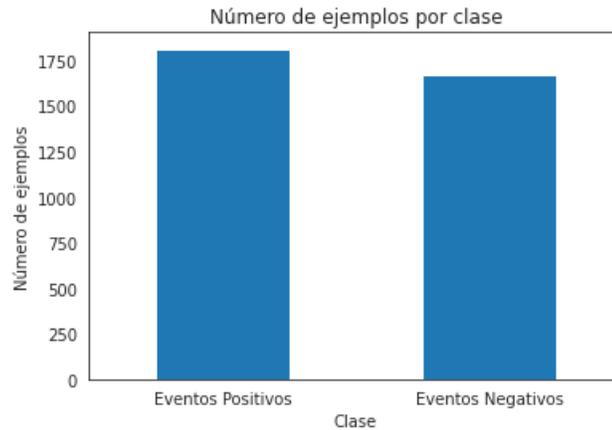


Figura 3.13: Distribución de ejemplos por clase aplicando el método presentado en la Figura 3.12. Se puede observar una cantidad de ejemplos similares, entre los segmentos del EMG con eventos de LM (Eventos positivos 1815) y entre los segmentos del EMG sin eventos de LM (Eventos negativos 1671).

A su vez, para ajustar los hiperparámetros del algoritmo Random Forest, se utilizan las características obtenidas sobre los candidatos a eventos obtenidos del Módulo 1 en aquellos registros del conjunto de validación. Finalmente, se selecciona el modelo (Random Forest) que posea el mejor F1-score sobre el conjunto de validación como modelo final, el cual es aplicado sobre el conjunto de prueba.

En la sección de Anexo A, se detalla otra metodología para ajustar los umbrales de las características obtenidas en el Módulo 2. La conclusión que se obtuvo de aquello es que no existe uniformidad en las distintas características obtenidas sobre los eventos posibles, esto es que no existen eventos que posean, por ejemplo, un número elevado de máximos por sobre una magnitud de $10[\mu V]$ y una alta energía. De esta manera, en vez de utilizar un ajuste de umbrales utilizando la curva Precisión-Sensibilidad, se reemplaza por el algoritmo Random Forest, el cual debido a sus características permite obtener una discriminación mucho más precisa para cada evento propuesto en el Módulo 1.

3.4.3. Tercer módulo

El tercer módulo consiste en aplicar la regla WASM (ver Tabla 1.1) para la detección de PLM. Para esto se recorren todas las detecciones finales de LM obtenidas en el Módulo 2, y se aplica regla de detección de PLM. Esto queda clarificado en el algoritmo 2

```
N=Número de eventos detectados
s,i,j=0,0,1
mientras  $i < N$  and  $j < N$  hacer
  si  $2 < |evento[i] - evento[j]|$  entonces
    |  $j += 1$ 
    |  $i += 1$ 
  fin
  en otro caso
    | si  $|evento[s] - evento[i]| < 90$  entonces
    | | PLM=evento[s] hasta evento[i]
    | fin
    |  $s = j$ 
    |  $i = j$ 
    |  $j += 1$ 
  fin
fin
```

Algoritmo 2: Algoritmo para determinar si un LM pertenece a una secuencia de PLM.

3.5. Criterios utilizados para evaluar el desempeño de la metodología propuesta

Para evaluar el desempeño del modelo propuesto, se utiliza la metodología planteada en [37], donde se propone una métrica basada en el parámetro IOU (Intersection Over Union). El IOU se define según la ecuación (3.1).

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (3.1)$$

Donde

- A: Corresponde al número de muestras que están sobre la marca realizada por la metodología propuesta (sistema automático).
- B: Corresponde al número de muestras que están dentro de la marca realizada por experto médico.

De esta manera, se definen los falsos negativos, verdaderos positivos y falsos positivos como :

- **Falsos positivos (FP)**: Son considerados como falsos positivos todos aquellos eventos detectados por el sistema automático y que para cada marca realizada por el experto médico poseen un IOU de cero ($IOU = 0$).
- **Verdaderos positivos (VP)**: Son considerados como verdaderos positivos todos aquellos eventos detectados por el sistema automático y que para cada marca realizada por el experto médico poseen un IOU distinto de cero ($IOU \neq 0$).
- **Falsos negativos (FN)**: Son considerados como falsos negativos todos aquellos eventos marcados por el experto médico y que para cada marca detectada por el sistema automático poseen un IOU de cero ($IOU = 0$).

Luego, a partir de las definiciones anteriores se utilizan las siguientes métricas para evaluar el desempeño de la metodología propuesta.

1. Sensibilidad: Corresponde a la proporción de casos correctamente detectados en relación al total de eventos de LM marcados por el experto médico.

$$Sensibilidad = \frac{VP}{VP + FN} \quad (3.2)$$

2. Precisión: Corresponde a la proporción de casos correctamente detectados en relación al total de eventos de LM detectados por el sistema.

$$Precisión = \frac{VP}{VP + FP} \quad (3.3)$$

3. $F1_{score}$: Corresponde a una ponderación de las métricas 1-2), lo cual es útil dado que permite comparar en una única métrica tanto la probabilidad de que el modelo detecte un evento (sensibilidad) y la probabilidad de que este sea una detección correcta (precisión).

$$F1_{score} = 2 * \frac{Sensibilidad * Precisión}{Sensibilidad + Precisión} \quad (3.4)$$

En la metodología implementada se buscará aumentar el valor de la métrica F1-score, dado que esta métrica permite que el modelo no se sobreajuste a la Sensibilidad o a la Precisión, y permite que en un solo valor quede representada que tan alta es la probabilidad de que el modelo detecte un evento (Sensibilidad) y la probabilidad de que los eventos detectados sean realmente eventos de LM (Precisión).

Capítulo 4

Resultados y Análisis

En este capítulo se presentan los resultados y análisis de manera separada para cada módulo de la metodología propuesta, esto dado que cada módulo tiene objetivos distintos, y su desempeño se debe evaluar de manera distinta en base a los objetivos que posea el módulo.

4.1. Resultados del módulo 1

El Módulo 1 tiene por objetivo generar la mayor cantidad de candidatos a eventos de LM. Luego, se busca que este módulo posea una alta sensibilidad (ecuación (3.2)).

Al probar el sistema sobre el conjunto de prueba, se obtienen los siguientes resultados.

Tabla 4.1: Resultados obtenidos por registro en el conjunto de prueba.

Id Registro	Verdaderos positivos	Falsos negativos	Falsos positivos
2	192	9	31907
3	80	9	27848
4	87	9	27092
6	131	6	30519
7	150	6	41448
8	201	18	21943
9	69	9	29966
20	68	1	32724
21	125	16	27262
26	165	19	19084
Total	1268	102	289793

Se puede observar que el Módulo 1 cumple con poder detectar casi todos los eventos de LM, donde se obtiene una buena robustez en el sentido de que independiente del número de eventos que posea el registro, el Módulo 1 logra detectar casi todos los eventos.

Claramente, se puede observar que el número de falsos positivos es muy alto, esto para cada uno de los registros del conjunto de prueba. Esto último también queda reflejado en las métricas obtenidas (ver Tabla 4.2).

De lo anterior se tiene que el primer módulo es efectivo para detectar eventos de LM, no

Tabla 4.2: Resultados sobre el conjunto de prueba.

Métrica	Valor
Sensibilidad	0,93
Precisión	0,004
$F1_{score}$	0,01

obstante, se requiere un filtro (Módulo 2) que permita borrar de los posibles candidatos a eventos de LM aquellos que no lo son. Además, se tiene que si bien la Energía de Shannon es una buena metodología para detectar eventos de LM, es sensible a variaciones de amplitud y a zonas del EMG en que la energía es baja. La Energía de Shannon al tomar en cuenta el contexto de un segmento de señal, aquellas zonas que poseen una baja energía pero presentan un aumento de amplitud en relación a su contexto, son propuestos como eventos de LM por el primer módulo, siendo que realmente no lo son. Este, y los otros tipos de falsos positivos, son analizados en la siguiente subsección.

4.1.1. Sobre los tipos de Falsos Positivos detectados

A partir de los eventos propuestos por el Módulo 1, se tienen los siguientes tipos de falsos positivos.

- **Tipo I:** Son aquellos relacionados a zonas de una baja energía y una baja potencia espectral. Esto se produce debido a que la Energía de Shannon toma en cuenta el contexto del segmento de señal que se está analizando en cada iteración, lo que ocasiona que breves aumentos de energía en relación a su contexto sean detectados como posibles eventos de LM. Esto se ejemplifica en la Figura 4.1.

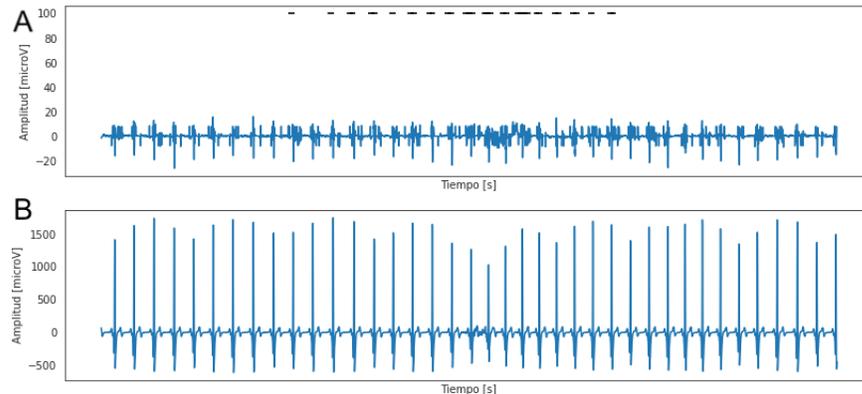


Figura 4.1: Ejemplo de falsos positivos de **Tipo I** en zonas de baja energía. A: Señal del EMG con falsos eventos detectados. B: Electrocardiograma, que ocasiona pequeños aumentos de amplitud en el EMG.

- **Tipo II:** Son aquellos producidos por artefactos ocasionados por el ECG. El ECG al contaminar la señal del EMG, produce cambios de amplitud en este último, los cuales no están relacionados a eventos de LM. Esto se puede observar en la Figura 4.2.

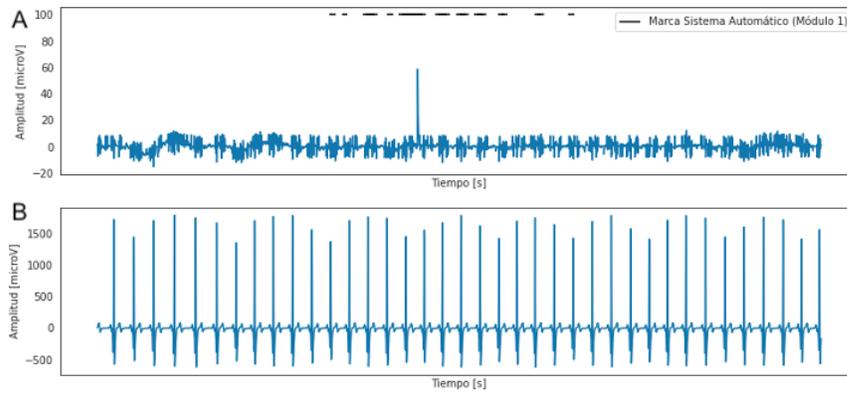


Figura 4.2: Ejemplo de falsos positivos de **Tipo II** en zonas con artefactos del ECG. A: Señal del EMG con falsos eventos detectados. B: Electrocardiograma, que como se observa, contamina el EMG ocasionando que las amplitudes del EMG crezcan en función a pulsos del ECG.

Para eliminar los tipos de falsos positivos I y II, se extraen las características descritas en el capítulo de Metodología, y se aplica el algoritmo Ranfom Forest para seleccionar los umbrales.

- **Tipo III:** Son todos aquellos candidatos a eventos que ocurren durante el inicio del registro ($\sim 300[s]$) y el final del registro (últimos $\sim 1000[s]$). Dado que se busca contabilizar aquellos eventos de movimientos de piernas que ocurren durante el sueño, el experto médico no suele marcar aquellos eventos que ocurren en un comienzo dado que usualmente a una persona le toma un tiempo quedarse dormida, y a su vez, el final dado que en el proceso de examen en que se obtiene el PSG, puede ocurrir que la persona despierte antes, pero de igual manera debe esperar a que el examen finalice. La Figura 4.3 da cuenta de dos falsos positivos que ocurren durante los primeros 300[s] del EMG para un determinado registro.

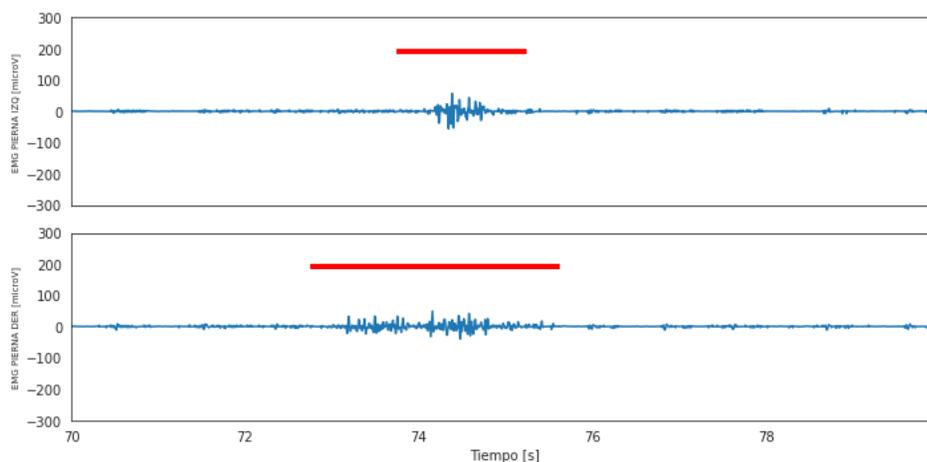


Figura 4.3: Ejemplo donde hay falsos positivos de **Tipo III** que ocurren en un comienzo del EMG, en las cuales el experto supone que el individuo está despierto, y de esta manera, los eventos que ocurren durante este periodo no son tomados en cuenta.

Para eliminar este tipo de falsos positivos, se eliminan de los candidatos a eventos de LM a aquellos eventos que ocurren durante los primeros 300 [s] del EMG, y durante los últimos 1000 [s] del EMG.

- **Tipo IV:** Son todos aquellos candidatos a eventos que fueron clasificados como falsos positivos (o falsos negativos) que se producen debido al corrimiento de las marcas realizadas por el experto. Esto se puede observar en la Figura 4.4, donde se presentan distintos eventos con corrimiento de marcas.

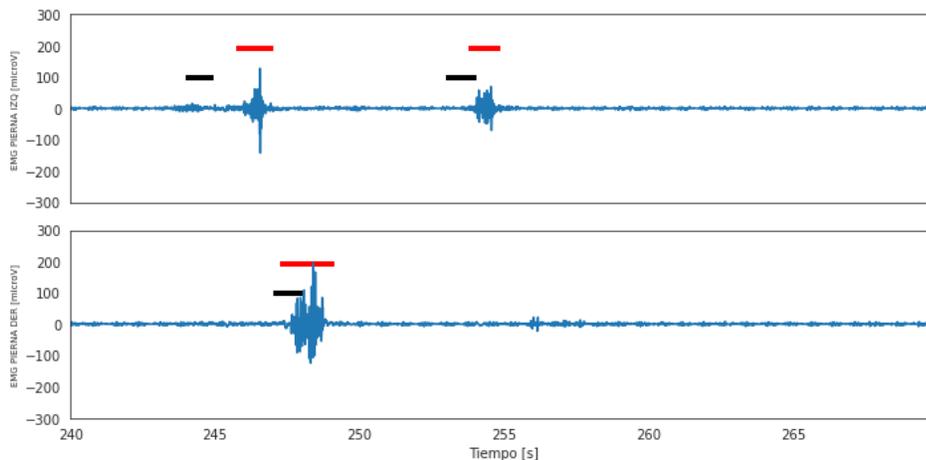


Figura 4.4: Ejemplo de falsos positivos de **Tipo IV**. La marca de color roja corresponde al evento detectado en el Módulo 1, y la marca negra al evento detectado por el experto médico.

Para eliminar este tipo de falsos positivos, cada marca realizada por el sistema automático se amplía en $\pm 2[s]$, siempre y cuando dicha ampliación no haga que la marca que se está ampliando se traslape con otra marca.

4.2. Resultados del Módulo 2

Para el ajuste del algoritmo Random Forest, primero se obtienen aquellos hiperparámetros que permiten obtener un mejor desempeño en cuanto a la métrica F1-score utilizando el método OOB (ver Anexo B y D.1). Luego, utilizando las 7 mejores combinaciones (Tabla D.1) de los hiperparámetros que maximizan el F1-score por el método OOB, se entrenan 7 Random Forest distintos y se prueban sobre el conjunto de validación. Los resultados obtenidos para cada modelo sobre el conjunto de validación están presentes en la Tabla 4.3.

Tabla 4.3: Resultados de los distintos modelos ajustados y probados sobre el conjunto de validación. Los modelos ajustados siguen la notación de la Tabla D.1 para los distintos conjuntos de hiperparámetros.

Métrica Modelos	3	9	12	48	30	21	0
Precisión	0,6043	0,6047	0,6050	0,6047	0,6047	0,6047	0,5903
Sensibilidad	0,8737	0,8737	0,8737	0,8737	0,8737	0,8737	0,8754
F1-score	0,7145	0,7147	0,7149	0,7147	0,7147	0,7147	0,7051

Se puede observar en la Tabla 4.3 que las diferencias en los resultados para las primeras 7 combinaciones no presentan muchas diferencias sobre el conjunto de validación, por lo que es de esperarse que cualquiera de estas combinaciones de hiperparámetros para el algoritmo Random Forest posea resultados similares en el conjunto de prueba. De esta manera, se selecciona el modelo de hiperparámetros 3 de la Tabla 4.3. Una vez entrenado el algoritmo Random Forest con aquellos hiperparámetros que maximizan la métrica F1-score en el conjunto de validación, se realiza la detección final sobre los eventos propuestos por el Módulo 1.

En la Tabla 4.4 se pueden observar los resultados de la detección final sobre cada uno de los registros del conjunto de prueba, donde se aprecia una gran mejora en cuanto al número de falsos positivos que se tenía en la salida del primer módulo (Tabla 4.1). No obstante, el número de falsos negativos aumenta, esto es, el número de eventos de LM presentes en el EMG y que no fueron detectados por el algoritmo aumentan.

Tabla 4.4: Resultados obtenidos por registro sobre el conjunto de prueba en la salida del Módulo 2.

Id Registro	Verdaderos positivos	Falsos negativos	Falsos positivos
2	189	12	71
3	62	27	44
4	84	12	7
6	118	19	19
7	154	2	87
8	204	15	83
9	73	5	84
20	43	26	2
21	114	27	4
26	152	32	43
Total	1193	177	444

Se puede notar que pese a que el número de falsos negativos aumenta en relación al primer módulo, aún el número total de ellos es bastante inferior al de verdaderos positivos alcanzado. Además, en cuanto a los falsos positivos, se aprecia claramente una mejora importante, disminuyendo de un total de 289793 obtenidos en el primer módulo a solo 444 en el segundo.

A su vez, en la Tabla 4.5 se puede observar el resultado en cuanto a las métricas evaluadas. Se tiene que las métricas de precisión y F1-score mejoran bastante en relación a las obtenidas en el primer módulo. De esta manera, el Módulo 2 cumple con el objetivo de filtrar la mayor cantidad de falsos positivos, sin perjudicar demasiado la Sensibilidad del modelo propuesto, la cual solo disminuye en un 6% en relación a lo obtenido en el primer módulo.

Tabla 4.5: Resultados sobre el conjunto de prueba en salida del Módulo 2.

Métrica	Valor
Sensibilidad	0,87
Precisión	0.73
$F1_{score}$	0,79

En el Anexo C, se puede observar el resultado de la metodología aplicando EMD para la descomposición en bandas de frecuencia en el segundo módulo. Este algoritmo (EMD) fue reemplazado por la DWT, dado que la DWT permite una descomposición mayor y debido a que su tiempo de procesamiento es menor que el que requiere la EMD.

En cuanto a la interpretabilidad de la metodología propuesta, primero es necesario recordar que por evento de LM propuesto por el Módulo 1 se obtienen 9 señales distintas y 96 características en total. De esta manera para obtener la importancia de cada características se obtiene su porcentaje de importancia MDI, y luego, para obtener la importancia de cada señal en la detección final, se toma la suma de las MDI de las características extraídas sobre cada señal respectiva (observar Figura 4.5). Luego, analizando por separado las importancias de las características y señales:

- En cuanto a las características, si se observa la Figura D.2 se tiene que dentro de las características más importantes son aquellas que se obtienen sobre el EMG, siendo la más importante el RMS, seguidos por la energía y PSD del segmento. Es notorio además que características como los percentiles que dan cuenta de la distribución que poseen las magnitudes de las señales consideradas, cobran relevancia para la clasificación, en conjunto con la energía y PSD del segmento. A su vez, entre las características con menor relevancia está la frecuencia en que se ubica el máximo de la PSD, la media y la mediana.
- En cuanto a la importancia de cada señal considerada: Si se observa la Figura 4.5, se tiene que las señales más importantes para realizar la detección corresponden al EMG para un contexto de $\pm 3[s]$ en relación al centro del evento propuesto por el Módulo 1, y además, las descomposiciones desde un nivel de 2 hasta un nivel de 5 de este segmento. De esta manera, en base a las características obtenidas zonas de una frecuencia media y baja son más relevantes para la correcta detección de eventos de LM, mientras que zonas de alta frecuencia (dwt0-1) son menos relevantes. Esto se condice con un uso común que posee la DWT, que es filtrar frecuencias altas sobre señales, dado que usualmente las frecuencias más altas están asociadas a ruido [43]. Además de las frecuencias altas del EMG, en base a las características obtenidas, el contexto de $\pm 10[s]$ desde el centro de la detección (Contexto EMG) no posee relevancia para realizar una correcta clasificación, siendo incluso la señal con menos importancia. Por último, se puede observar que el canal de movimiento posee una importancia ínfima en relación al de otras señales, en donde pese a que sirve para filtrar algunos falsos positivos ($MDI \neq 0$) esto no es generalizable a todos los candidatos a evento.

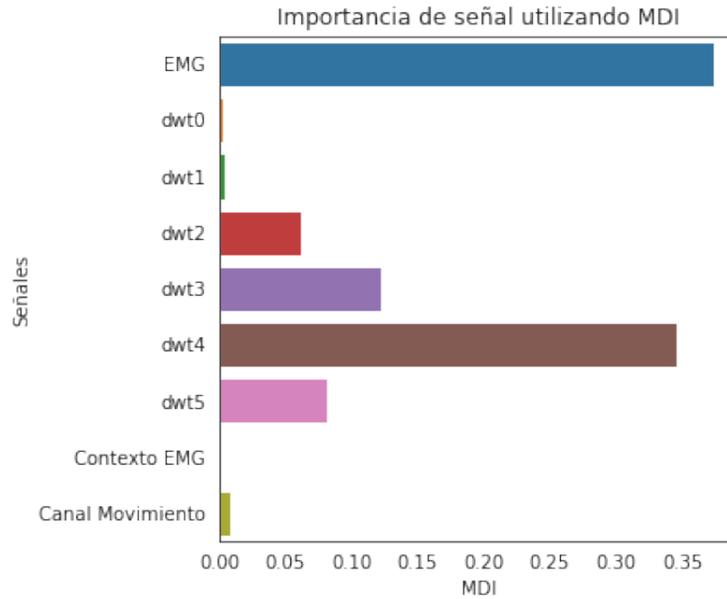


Figura 4.5: Se puede observar la importancia de cada señal considerada, según contexto y canal (ver Figura 3.11), en relación a la suma de las importancias de las características obtenidas en cada señal para cada candidato propuesto en el Módulo 1 (ver Figura D.2). EMG y dwt0-5 corresponden al segmento del EMG y las descomposiciones de este segmento hasta un nivel 5 del candidato a evento para un contexto de $\pm 3[s]$ en relación al centro de la detección. A su vez, Contexto EMG corresponde al segmento del EMG para un contexto de $\pm 10[s]$, y Canal movimiento corresponde al canal de movimiento del PSG para un contexto de $\pm 3[s]$. Se puede apreciar que las señales más importantes de acuerdo a las características obtenidas y el algoritmo Random Forest entrenado, corresponden al segmento de $\pm 3[s]$ del EMG y a las descomposiciones mayores de este mismo.

4.3. Comparación de la metodología implementada y estado del arte en detección de LM

En la Tabla 4.6 se pueden observar los resultados obtenidos en los trabajos anteriores para las métricas utilizadas en este trabajo, esto es, Precisión, Sensibilidad y F1-score. Como ya se vio en la presentación del estado del arte, la mayoría de los trabajos previos publicaron sólo las métricas de Sensibilidad y Especificidad. No obstante, el uso de la Especificidad en la detección de eventos de LM y PLM no reviste de interés, esto, dado que la Especificidad toma en cuenta los verdaderos negativos de un modelo (ver ecuación (A.1)), los cuales no tienen interés desde el punto de vista fisiológico, dado que un verdadero negativo se interpreta como una zona del EMG en la cual el modelo detecta correctamente que ahí no hubo un evento de LM. Además, se tiene que como los eventos de LM son escasos en relación a la duración total del registro polisomnográfico, es muy probable que la especificidad alcanzada sea muy alta, lo cual se puede fácilmente malinterpretar como que el modelo es mejor al tener una Especificidad mayor, cosa que no es cierta dado que el modelo simplemente estará sobreajustado a zonas del EMG en que no ocurren eventos de interés fisiológico.

De las métricas anteriores, la más importante corresponde al F1-score debido a que logra

capturar qué tan bueno es el modelo en cuanto a la probabilidad de detectar un evento de LM (Sensibilidad) y en cuanto a la probabilidad de que un evento detectado sea realmente un evento de LM (Precisión). El F1-score al ser una combinación de estas métricas (Precisión y Sensibilidad) permite que la metodología no se sobreajuste a una métrica en específico. Esto sucede por ejemplo en los trabajos de [5] y [6], donde para [5] se obtiene una alta precisión (ver Tabla 4.7) pero una baja Sensibilidad, lo cual ocasiona que se obtenga un F1-score bajo (0,37). Similarmente para [6] se obtiene una alta Sensibilidad, pero una baja Precisión lo que ocasiona que se obtenga un bajo F1-score (0,63). Luego, se tiene que la métrica F1-score permite determinar qué tan bueno es un modelo.

Ahora, observando la Tabla 4.6 se tiene que el Módulo 1 de la metodología desarrollada, alcanza la segunda mejor Sensibilidad (en conjunto con [5] y [6]), siendo superada solo por el trabajo de Huang et al [9]. A su vez, en cuanto a las métricas de Precisión y F1-score, estas solo fueron publicadas en el trabajo de Carvelli et al. [12]. Se puede observar que el Módulo 2 de la metodología propuesta, logra una mejor Sensibilidad y F1-score que los alcanzados en [12]. Esto último es relevante considerando que el método desarrollado en [12] corresponde a una metodología basada en aprendizaje profundo, donde se utilizó una gran cantidad de datos de entrenamiento. No obstante, el número de registros del conjunto de prueba en [12], también es más grande que la utilizada en la metodología propuesta en este documento, por lo que es probable que lo realizado en [12] sea más generalizable, no solo que la metodología propuesta en este documento, sino que también a cualquier otro trabajo realizado hasta ahora, debido a la gran cantidad de datos que utilizaron (Tabla 1.2) en cada uno de sus conjuntos de entrenamiento, prueba y validación.

Tabla 4.6: Resultados comparativos con métodos previos, y la metodología propuesta según cada módulo, esto es, Módulo 1 (Mód 1) y Módulo 2 (Mód 2).

Métrica	Tauch[5]	Wett[6]	Ferri[7]	Huang [9]	Moore[8]	Moore SNR_+ [8]	Carvelli [12]	Mód 1	Mód 2
Precisión	-	-	-	-	-	-	0,74	0,004	0,73
Sensibilidad	0,93	0,93	>0,9	0,96	0,82	0,88	0,81	0,93	0,87
F1-score	-	-	-	-	-	-	0,77	0,01	0,79

Pese a esto, otro punto muy relevante es que desde el punto de vista fisiológico la posibilidad de poder interpretar los resultados es importante. Esto último le confiere una gran desventaja a lo realizado en [12] frente a todas las otras metodologías realizadas, debido a que es un método basado en aprendizaje profundo y sus resultados no poseen explicación fisiológica. De esta manera, si bien por un lado puede ser el método más generalizable, al mismo tiempo es el menos interpretable lo cual representa un gran problema desde el punto de vista médico.

Por otro lado, en [12] desarrollaron un experimento en el que replicaron las metodologías de [5–8], y junto con la suya, las probaron sobre un mismo conjunto de 60 registros pertenecientes al WSC con eventos de LM y PLM marcados sobre el EMG por 5 expertos distintos. En la Tabla 4.7 se presentan los resultados obtenidos en la metodología propuesta en este documento, y los del experimento desarrollado en [12]. En cuanto a la Precisión, la metodología propuesta logra la cuarta mejor Precisión, siendo superada por los trabajos de [5], [8] y [12]. A su vez, en cuanto a la Sensibilidad, la metodología propuesta, para el Módulo 1 y Módulo 2 logra la segunda y tercera mejor Sensibilidad respectivamente, siendo superada

por el trabajo de [6] (Módulo 1 y 2) y por [12] (Módulo 2).

Finalmente, se puede observar que la metodología propuesta logra obtener el segundo mejor F1-score en conjunto con el trabajo de Moore et al [8], en donde solo son superados por Carvelli et al. [12].

Tabla 4.7: Resultados comparativos de la metodología implementada en este documento con resultados del experimento desarrollado en Carvelli et al [12], donde replicaron los métodos de [5–8] y los probaron sobre un mismo conjunto de datos, esto es, 60 PSG pertenecientes al WSC marcados por 5 expertos distintos.

Métrica	Tauch[5]	Wett[6]	Ferri[7]	Moore[8]	Moore SNR_+ [8]	Carvelli [12]	Mód 1	Mód 2
Precisión	0,79	0,47	0,62	0,88	0,85	0,81	0,004	0,73
Sensibilidad	0,24	0,96	0,85	0,60	0,73	0,90	0,93	0,87
F1-score	0,37	0,63	0,72	0,71	0,79	0,85	0,01	0,79

Si se comparan los resultados obtenidos por [12] en las Tablas 4.6 y 4.7, se puede observar que sus resultados mejoran al probar el modelo en el subconjunto de datos del WSC. De esta manera, es muy probable que los 60 registros seleccionados para esa prueba estén algo más idealizados de lo real, esto quiere decir que los 60 registros deben poseer un bajo nivel de ruido, las detecciones realizadas por los expertos hayan sido revisadas varias veces por cada uno de los expertos (5 expertos en total), y las detecciones sean claras en el sentido de no presentar corrimiento en relación al evento real y también en relación al medio se note un claro aumento de amplitud. Pese a esto, si se comparan los resultados de la Tabla 4.7 se puede observar que la metodología propuesta está dentro de los tres mejores resultados en cuanto al F1-score, igualando con Moore et al. [8] y siendo superados solo por Carvelli et al. [12], el cual como ya se mencionó previamente, corresponde a un modelo no interpretable.

De esta manera, al ser el F1-score la métrica que captura qué tan bueno es un modelo, se tiene que en base a esta métrica la metodología propuesta en este documento, está dentro de los mejores trabajos realizados.

4.4. Resultados Módulo 3

El Módulo 3 tiene por objetivo generar las secuencias de PLM a partir de los eventos de LM detectados en el Módulo 2. Luego, aplicando la regla de la WASM (Tabla 1.1) se detectan las secuencias de PLM sobre cada registro del conjunto de prueba, y luego se calcula su PLMI, el cual se obtiene al dividir el número total de PLM en un registro por la duración total de este en horas. En la Tabla 4.8 se puede observar el número de PLM obtenidos sobre cada registro del conjunto de prueba, la duración de cada registro en horas y su respectivo PLMI.

Cabe mencionar que de acuerdo a la regla WASM, las secuencias de PLM deben estar conformadas por al menos 4 eventos de LM y tener un intervalo de entre 5 – 90[s] entre eventos. De esta manera, todos aquellos LM que no cumplen con lo anterior se catalogan como eventos aislados, y no se consideran para el conteo de PLM, y en consecuencia, para el PLMI.

Tabla 4.8: PLMI obtenido aplicando regla WASM para la detección de PLM sobre cada registro del conjunto de prueba.

Id Registro	PLMs	Duración en horas	PLMI
2	49	7,54	6,50
3	18	7,42	2,43
4	23	6,94	3,31
6	36	8,31	4,33
7	49	8,23	5,95
8	59	6,57	8,98
9	17	8,39	2,03
20	15	6,95	2,16
21	37	8,01	4,62
26	39	7,57	5,15

El dar una explicación fisiológica del índice de PLM está fuera del alcance de esta memoria, ya que para ello se requiere de expertos médicos. No obstante, de la Tabla 4.8 se puede observar que los registros poseen una duración en horas similar, y aquellos pacientes con un PLMI más alto corresponden a aquellos con id 2, 6, 7, 8 21 y 26, y más bajo los pacientes con id 3, 4 y 9.

Capítulo 5

Conclusiones

De los resultados obtenidos, se tiene que el Módulo 1 cumple su objetivo, donde logra capturar prácticamente todos los eventos de piernas presentes en los distintos registros del conjunto de prueba. De esta manera, la energía de Shannon es una buena característica para poder detectar eventos de LM sobre el EMG. No obstante, el umbral debe ser ajustado y puede variar en distintos conjuntos de datos pertenecientes a otros estudios.

De los resultados obtenidos en el Módulo 2, se obtiene que para la detección de eventos de LM, en base al porcentaje de importancia MDI y las características obtenidas, aquellas frecuencias más bajas de la señal del EMG son más relevantes que las frecuencias altas, lo cual se condice con un uso común que posee la DWT en señales que es descomponer la señal en bandas y eliminar las frecuencias altas, dado que usualmente son asociadas a ruido [43]. Además, de las características importantes para la detección de eventos, cobran relevancia la energía y la PSD del segmento, y además características estadísticas básicas como percentiles, que dan cuenta de la distribución de las magnitudes del segmento. A su vez, dentro de las características menos relevantes está por ejemplo la frecuencia en la que se concentra la máxima PSD.

En cuanto a los resultados en el conjunto de prueba, el Módulo 2 logra filtrar un gran número de falsos positivos, sin disminuir demasiado la sensibilidad de la metodología implementada. De esta manera, el Módulo 2 cumple con su objetivo.

Luego, de los eventos de LM propuestos por el Módulo 2 se obtuvieron las respectivas secuencias de PLM sobre cada registro del conjunto de prueba, y a partir de esto su índice PLMI que corresponde al índice utilizado para realizar un análisis médico de los PLM. De esta manera, el sistema cumple con el objetivo de aplicar la regla de la WASM para la detección de eventos de PLM, y luego calcular los respectivos PLMI.

Por otro lado, comparando los resultados obtenidos con el estado del arte en detección automática de LM y PLM, en base a la métricas evaluadas, la metodología propuesta está dentro de los mejores tres modelos estando en segundo lugar (F1-score) con lo realizado en Moore et al. [8] y superado por lo propuesto en Carvelli et al [12]. No obstante, este último trabajo corresponde a una metodología basada en aprendizaje profundo cuyos resultados no poseen explicación fisiológica, lo cual es una gran desventaja frente a lo realizado en este documento, o frente a las otras metodologías implementadas previamente en [5–8].

En relación a los objetivos de la memoria, primero en relación a los objetivos específicos se tiene que cada uno de ellos fue logrado:

- Se construyó una base de datos de registros polisomnográficos con eventos de LM marcados.

- En base a la obtención de dos características (PSD y Energía de Shannon) se realizó una correcta partición en los conjuntos de entrenamiento, validación y prueba.
- Se obtuvo la Energía de Shannon de cada registro.
- Se generó el primer bloque del modelo el cual cumple con obtener un gran número de verdaderos positivos (alta Sensibilidad).
- Se generó el segundo bloque del modelo el cual logra filtrar un gran número de falsos positivos.
- Se implementó el tercer bloque del modelo, el cual a partir de los LM detectados, y regla de la WASM genera secuencias periódicas, o en su defecto determina si un evento es aislado.

A partir de los puntos anteriores, se tiene que el objetivo general de la memoria de construir un sistema para la detección automática de movimientos de piernas sobre el EMG utilizando la Energía de Shannon para registros polisomnográficos del INTA de la Universidad de Chile fue logrado.

Finalmente, como trabajo futuro se proponen los siguientes puntos:

- Evaluar si la extracción de alguna característica, ya sea alguna que se extrae en el segundo módulo o una completamente nueva, permite filtrar eventos en el primer módulo y con esto disminuir la gran cantidad de falsos positivos en la salida del primer módulo.
- Evaluar si es posible expresar la extracción de características en ambos módulos de manera matricial, y de esta manera disminuir el tiempo de procesamiento al utilizar una GPU para realizar los cálculos.
- Extracción de otras características y evaluar si esto logra mejorar los resultados.
- Incluir algún algoritmo para la detección de eventos RRLM y así poder filtrarlos.
- Evaluar el desempeño de la metodología propuesta en esta memoria sobre nuevos registros pertenecientes a otras bases de datos (WSC o SSC).
- Implementar una interfaz interactiva para el Software, donde el experto no solo visualice las detecciones sobre el EMG, sino que también pueda modificarlas.

Capítulo 6

Glosario de abreviaturas usadas en este documento

AASM: American Academy of Sleep Medicine.
AOS: Apnea obstructiva del sueño.
ES: Energía de Shannon.
EMD: Descomposición modal empírica (empirical mode decomposition).
EEG: Electroencefalograma.
ECG: Electrocardiograma
EMG: Electromiograma.
EMA: Exponential Moving Average.
EOC: Electrooculograma.
EOG: Electrooculografía.
FP: Falso positivo.
FN: Falso negativo.
Fc: Frecuencia de corte.
FFT: Transformada rápida de Fourier.
HHT: Transformada de Hilbert-Huang.
INTA: Instituto de nutrición y tecnología de los alimentos.
IOU: Intersection Over Union.
IMF: Intrinsic mode functions.
LM: Movimiento de pierna.
MrOS: Osteoporotic Fractures in Men Study.
MOR: Movimientos Oculares Rápido.
MOV PIERNA: Movimiento de piernas.
MOV BRAZO: Movimiento de brazo.
MDI: Mean decrease in impurity
MA: Media móvil.
Mód 1: Módulo 1.
Mód 2: Módulo 2.
OOB: Out Of Bag.
PLM: Movimiento de pierna periódico.
PLMD: Movimiento de pierna periódico durante el sueño.
PLMI: Índice PLM.
PSG: Polisomnograma.

ROC: Receiver operating characteristic.

RLS: Síndrome de Piernas Inquietas (Restless Legs Syndrome).

RRLM: Movimientos de piernas (LM) relacionados a eventos respiratorios.

PSD: Densidad de potencia espectral.

SSC: Stanford Sleep Cohort.

SD: Desviación estándar.

SMA: Media móvil simple.

VP: Verdadero positivo.

WSC: Wisconsin sleep cohort.

WASM: World Association of Sleep Medicine.

WM: Método Welch.

Bibliografia

- [1] Ferri, R., Koo, B. B., Picchietti, D. L., Fulda, S. Periodic leg movements during sleep: phenotype, neurophysiology, and clinical significance, *Sleep medicine*, vol. 31, pp. 29–38, 2017. [<https://doi.org/10.1016/j.sleep.2016.05.014>].
- [2] May AM, Blackwell T, Stone KL, et al. Osteoporotic Fractures in Men (MrOS) study group. Longitudinal relationships of periodic limb movements during sleep and incident atrial fibrillation. *Sleep medicine* vol. 25. pp. 78-86, 2016. [<https://doi.org/10.1016/j.sleep.2016.08.009>].
- [3] Yumino, D., Wang, H., Floras, J., Newton, G., Mak, S., Ruttanaumpawan, P., Parker, J. and Bradley, T., 2011. Relation of Periodic Leg Movements During Sleep and Mortality in Patients With Systolic Heart Failure. *The American Journal of Cardiology*, vol. 107, no. 3, pp.447-451. [<https://doi.org/10.1016/j.amjcard.2010.09.037>].
- [4] Skomro R, Silva R, Alves R, et al. The prevalence and significance of periodic leg movements during sleep in patients with congestive heart failure. *Sleep Breath*, vol.13, pp. 43-7, 2009. [<https://doi.org/10.1007/s11325-008-0207-5>].
- [5] Tauchmann N, Pollmacher T. Automatic detection of periodic leg movements (PLM). *Journal of Sleep Research*, vol. 5, pp 273-275, 1996. [<https://doi.org/10.1111/j.1365-2869.1996.00273.x>].
- [6] Thomas C. Wetter, MD, Gerhard Dirlich, PhD, Jonathan Streit, Claudia Trenkwaldler, MD, Andreas Schuld, MD, Thomas Pollmächer, MD, An Automatic Method for Scoring Leg Movements in Polygraphic Sleep Recordings and Its Validity in Comparison to Visual Scoring, *Sleep*, vol. 27, no. 2, pp. 324–328, 2004. [<https://doi.org/10.1093/sleep/27.2.324>].
- [7] Ferri R, Zucconi M, Manconi M, et al. Computer-assisted detection of nocturnal leg motor activity in patients with restless legs Syndrome and periodic leg movements during sleep. *Sleep Medicine* vol.28, pp. 998-1004, 2004. [<https://doi.org/10.1093/sleep/28.8.998>].
- [8] Moore H, Leary E, Lee SY, et al. Design and validation of a periodic leg movement detector. *PloS One*, vol. 12, no. 9, pp. 1-30, 2014. [<https://doi.org/10.1371/journal.pone.0114565>].
- [9] Huang AS, Skeba P, Yang MS, et al. MATPLM1, A MATLAB script for scoring of periodic limb movements: preliminary validation with visual scoring. *Sleep Medicine*, vol. 16, no. 12, pp. 1541-1549, 2015. [<https://doi.org/10.1016/j.sleep.2015.03.008>].
- [10] D. Alvarez-Estevéz, D. Wahid, R.M. Rijsman, Validation of an automatic scoring algorithm for the analysis of periodic limb movements according to the WASM2016 guidelines, *Sleep Medicine*, vol. 40, no. 1, pp. e13-e14, 2017. [<https://doi.org/10.1016/j.sleep.2017.11.032>].

- [11] Stefani A, Heidebreder A, Hackner H, et al. Validation of a leg movements count and periodic leg movements analysis in a custom polysomnography system. *BMC Neurology*, vol. 17, no. 42, 2017. [<https://doi.org/10.1186/s12883-017-0821-6>].
- [12] Lorenzo Carvelli, Alexander N. Olesen, Andreas Brink-Kjær, Eileen B. Leary, Paul E. Peppard, Emmanuel Mignot, Helge B.D. Sørensen, Poul Jennum, Design of a deep learning model for automatic scoring of periodic and non-periodic leg movements during sleep validated against multiple human experts, *Sleep Medicine*, vol. 69, pp. 109-119, 2020. [<https://doi.org/10.1016/j.sleep.2019.12.032>].
- [13] Ferri R, Fulda S, Allen R, et al. The official World Association of Sleep Medicine (WASM) 2016 standards for recording and scoring leg movements in polysomnograms developed by a joint task force from the International and the European Restless Legs Syndrome Study Group (IRLSSG and EURLSSG). *Sleep Medicine*, vol. 26, pp. 86-95, 2016. [<https://doi.org/10.1016/j.sleep.2016.10.010>].
- [14] Berry RB, Brooks R, Gamaldo CE, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, Ver. 2.3. Darien (IL): American Academy of Sleep Medicine, 2016.
- [15] J. Lin and L. Qu, “Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis,” *Journal of sound and vibration*, vol. 234, no. 1, pp. 135–148, 2000. [<https://doi.org/10.1006/jsvi.2000.2864>].
- [16] Young T, Finn L, Peppard PE, et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. *Sleep*, vol. 38, no. 8, pp.1071-1078, 2008. [<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2542952/>].
- [17] Blackwell, T., Yaffe, K., Ancoli Israel, S., Redline, S., Ensrud, K.E., Stefanick, M.L., Laffan, A., Stone, K.L. and, Associations Between Sleep Architecture and Sleep Disordered Breathing and Cognition in Older Community-Dwelling Men: The Osteoporotic Fractures in Men Sleep Study. *Journal of American Geriatrics Society*, vol. 59, pp.2217-2225, 2009. [<https://doi.org/10.1111/j.1532-5415.2011.03731.x>].
- [18] Janet Babich Blank, Peggy Mannen Cawthon, Mary Lou Carrion-Petersen, Loretta Harper, J. Phillip Johnson, Eileen Mitson, Romelia Ramírez Delay, Overview of recruitment for the osteoporotic fractures in men study (MrOS), *Contemporary Clinical Trials*, vol. 26, no. 5, pp. 557-568, 2005. [<https://doi.org/10.1016/j.cct.2005.05.005>].
- [19] Guo-Qiang Zhang, Licong Cui, Remo Mueller, Shiqiang Tao, Matthew Kim, Michael Rueschman, Sara Mariani, Daniel Mobley, Susan Redline, The National Sleep Research Resource: towards a sleep data commons, *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1351–1358, 2018. [<https://doi.org/10.1093/jamia/ocy064>].
- [20] Chen X, Wang R, Zee P, Lutsey PL, Javaheri S, Alcántara C, Jackson CL, Williams MA, Redline S. Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA). *Sleep*. vol. 38, no. 6, pp. 877-888, 2015. [<https://doi.org/10.5665/sleep.4732>].
- [21] Klingelhofer L, Bhattacharya K, Reichmann H. Restless legs syndrome. *Clin Med (Lond)*, vol. 14, no. 4, pp. 379-382, 2016. [<https://doi.org/10.7861/clinmedicine.16-4-379>].
- [22] Medic G, Wille M, Hemels ME. Short- and long-term health consequences of sleep disruption. *Nat Sci Sleep*, vol. 9, pp. 151-161, 2017. [<https://doi.org/10.2147/NSS.S134864>].

- [23] Pittman SD, MacDonald MM, Fogel RB, Malhotra A, Todros K, Levy B, Geva AB, White DP. Assessment of automated scoring of polysomnographic recordings in a population with suspected sleep-disordered breathing. *Sleep*. Vol 27 no. 7, pp. 1394-403, 2004. [<https://doi.org/10.1093/sleep/27.7.1394>].
- [24] Vladyslav V Vyazovskiy, Mark E Walton, Stuart N Peirson, David M Bannerman, Sleep homeostasis, habits and habituation, *Current Opinion in Neurobiology*, Volume 44, 2017, pp. 202-211, ISSN 0959-4388, [<https://doi.org/10.1016/j.conb.2017.05.002>].
- [25] A. A. Borbely, "A two process model of sleep regulation," *Human Neurobiology*, vol. 1, pp. 195-204, 1982. [<https://doi.org/10.1111/jsr.12371>].
- [26] M. A. Carskadon and W. C. Dement, "Normal human sleep: an overview", *Principles and Practice of Sleep Medicine*: M. H. Kryger, T. Roth, W. C. Dement (eds.), Elsevier Inc., 4th edition, pp 13-23, Philadelphia, USA, 2005. [<https://doi.org/10.1016/j.mcna.2004.01.001>].
- [27] M. Steriade, "Grouping of brain rhythms in corticothalamic systems", *Neuroscience*, vol. 137, no. 4, pp. 1087-1106, 2006. [<https://doi.org/10.1016/j.neuroscience.2005.10.029>].
- [28] P. Peirano and C. Algarín, "Sleep in brain development", *Biological Research*, vol. 40, no. 4, pp. 471-478, 2007. [<https://doi.org/10.1159/000508055>].
- [29] M. P. Walker, T. Brakefield, A. Morgan, A. J. Hobson and R. Stickgold, "Practice with sleep makes perfect: sleep-dependent motor skill learning", *Neuron*, vol. 35, no. 1, pp. 205-211, 2002. [[https://doi.org/10.1016/s0896-6273\(02\)00746-8](https://doi.org/10.1016/s0896-6273(02)00746-8)].
- [30] M. P. Walker and R. Stickgold, "Sleep, memory, and plasticity", *Annual Review of Psychology*, vol. 57, no. 1, pp. 139-166, 2006. [<https://doi.org/10.1146/annurev.psych.56.091103.070307>].
- [31] SS. Yoo, P. T. Hu, N. Gujar, F. A. Jolesz and M. P. Walker, "A deficit in the ability to form new human memories without sleep", *Nature Neuroscience*, vol. 10, no. 3, pp. 385-392, 2007. [<https://doi.org/10.1038/nn1851>].
- [32] L. Causa, "Detección automática y caracterización de patrones de husos sigma en registros polisomnográficos de niños," Tesis de magister, Universidad de Chile, Enero 2012. [http://repositorio.uchile.cl/bitstream/handle/2250/102747/cf-causa_lm.pdf?sequence=3&isAllowed=y].
- [33] Ferri R, Fulda S "Quantifying Leg Movement Activity During Sleep". *Sleep Med Clin* Vol 11, no. 4, pp. 403-420, 2016. [<https://doi.org/10.1016/j.jsmc.2016.08.005>].
- [34] G. Orellana, "Clasificación de sueño mediante medición de actividad motora", Memoria de pregrado, Universidad de Chile, 2014. [<http://repositorio.uchile.cl/bitstream/handle/2250/131049/Clasificaci%c3%b3n-de-sue%c3%bl0-mediante-medi%c3%b3n-de-la-actividad-motora.pdf?sequence=1&isAllowed=y>].
- [35] Hamed Beyramienanlou, Nasser Lotfivand, "Shannon's Energy Based Algorithm in ECG Signal Processing", *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 8081361, 16 pages, 2017. [<https://doi.org/10.1155/2017/8081361>].
- [36] G. Rilling, P. Flandrin and P. Gonçalves, "On empirical mode decomposition and its algorithms", *Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP)*, Italy, vol. 1, 2003. [<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.586.812&rep=rep1&type=pdf>].

- [37] S. Warby, S. Wendt, P. Welinder, E. Munk, O. Carrillo, H. Sorensen, P. Jennum, P. Pppard, P. Perona, and E. Mignot, “Sleep-spindle detection: Crowdsourcing and evaluating performance of experts, non-experts and automated methods”, *Nat. Methods*, vol. 11, no. 4, pp. 385–392, 2014. [<https://doi.org/10.1038/nmeth.2855>].
- [38] Redfern M, Hughes R, Chaffin D. High-pass filtering to remove electrocardiographic interference from torso EMG recordings. *Clin Biomech (Bristol, Avon)*, vol. 8, n0.1 , pp. 44-48, [[https://doi.org/10.1016/s0268-0033\(05\)80009-9](https://doi.org/10.1016/s0268-0033(05)80009-9)].
- [39] L. Breiman, «Random Forests,» *Machine Learning*, vol. 45, n° 1, pp. 5-32, 2001. [<https://doi.org/10.1023/A:1010933404324>].
- [40] T. Hastie, R. Tibshirani y J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition, New York: Springer, 2009. ch. 15, pp. 587–603.
- [41] Murphy, K. P. . *Machine learning: a probabilistic perspective*. Cambridge, MA: MIT Press, 2012. pp. 2-3.
- [42] Simon L., Webster R. y Rabin J., Revisiting Precision and Recall Definition for Generative Model Evaluation, *Proceedings of the 36 th International Conference on Machine Learning*, Long Beach, California, PMLR 97, 2019. [<https://arxiv.org/pdf/1905.05441.pdf>].
- [43] P. Addison, *The illustrated wavelet transform handbook: Introductory theory and applications in science, engineering, medicine and finance*. CRC press, 2017.
- [44] Huang, N. E., Shen, Z., Long, S. R., Wu, M., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., y Liu, H. H., The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. London, Ser. A*, 454, pp.903-993. 1998. [<http://doi.org/10.1098/rspa.1998.0193>].
- [45] Huang, N. E, Wu, Z., A Review on Hilbert-Huang Transform: Method and its applications to geophysical studies. *Rev. Geophys.* 46. RG2006. 2008. [<https://doi.org/10.1029/2007RG000228>].

Anexo A

Ajuste de umbrales en módulo 2 utilizando curva Precisión-Sensibilidad

A.1. Curva ROC

La curva AUC - ROC es una medida de rendimiento en problemas de clasificación. Consiste en obtener la tasa de verdaderos positivos y falsos positivos para varios valores de umbrales. ROC es una curva de probabilidad y AUC representa el grado o medida de separabilidad. Indica cuánto es capaz el modelo de distinguir entre clases. Cuanto mayor sea el AUC, mejor será el modelo para clasificar correctamente las distintas clases del modelo.

Un modelo excelente tiene un AUC cercano al 1, lo que significa que tiene una buena medida de separabilidad. Un modelo deficiente tiene un AUC cercano a 0, lo que significa que tiene una mala medida de separabilidad. De hecho, significa que está correspondiendo el resultado. Por otro lado, cuando el AUC es 0,5, significa que el modelo no tiene capacidad de separación de clases en absoluto. En la figura A.1 se puede observar un ejemplo de curva ROC.

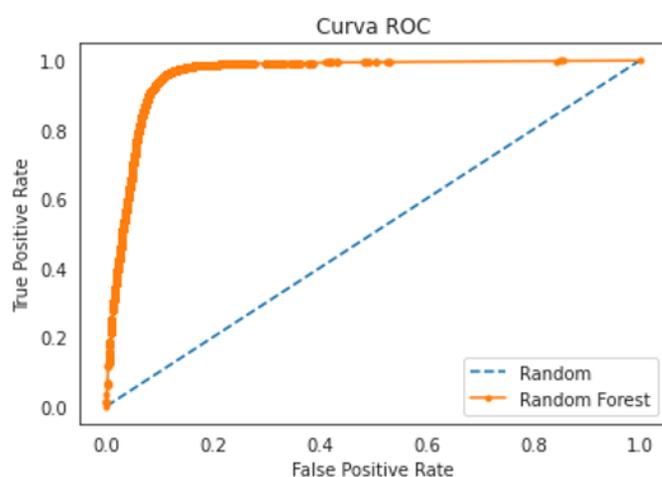


Figura A.1: Curva ROC para un modelo de clasificación utilizando Random Forest. En naranja se obtiene una curva ROC con un $AUC = 0.9$. En azul se observa la clasificación con un $AUC = 0.5$ que corresponde a la peor curva ROC que se podría obtener.

De la figura A.1 se puede ver claramente que a medida que la curva ROC está más cercana al borde superior izquierdo, el modelo de clasificación es mejor.

Notar además que el eje y (tasa de verdaderos positivos) de la curva ROC corresponde a la sensibilidad (3.2), y que el eje x (tasa de falsos positivos) corresponde a la especificidad, dada por la ecuación (A.1).

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (\text{A.1})$$

De la ecuación (A.1) se puede observar que el modelo puede tener una gran medida de especificidad si su tasa de verdaderos negativos es alta. No obstante, para la detección de LM, esto desde el punto de vista fisiológico no reviste de mayor interés dado que el clasificar correctamente zonas del EMG en no hay eventos como que ahí no hubo evento (verdadero negativo) no tiene ningún aporte fisiológico. Además, si se considera el caso de clases desbalanceadas (que es el caso de detectar eventos LM) donde los eventos negativos (ausencia de eventos LM) es mucho mayor a eventos positivos (presencia de LM), ocasiona que se puede obtener una especificidad muy alta, pero el modelo estará sobreajustado a la clase mayoritaria de eventos negativos. De esta manera, usualmente se prefiere la curva precisión-sensibilidad cuando se trabaja con clases desbalanceadas [42].

A.2. Curva Precisión-Sensibilidad

Otra forma que se probó para ajustar los umbrales de las características en el módulo 2 y evaluar el trade-off que existe entre la sensibilidad y la precisión, es la curva precisión-sensibilidad (precision-recall en inglés). La curva precisión-sensibilidad a diferencia de la curva ROC, es más útil al trabajar con modelos cuyas clases están desbalanceadas, dado que utiliza métricas que dan cuenta del desbalance entre las clases, esto es, la sensibilidad y la precisión (ver ecuaciones (3.3) y (3.2)), mientras que la curva ROC utiliza la sensibilidad y la especificidad sobre ajustándose a la clase mayoritaria, que en el caso de detección de LM es la ausencia de eventos de LM. En la figura A.2 se puede observar un ejemplo de la curva Precisión-Sensibilidad para el ajuste de distintos umbrales en un modelo.

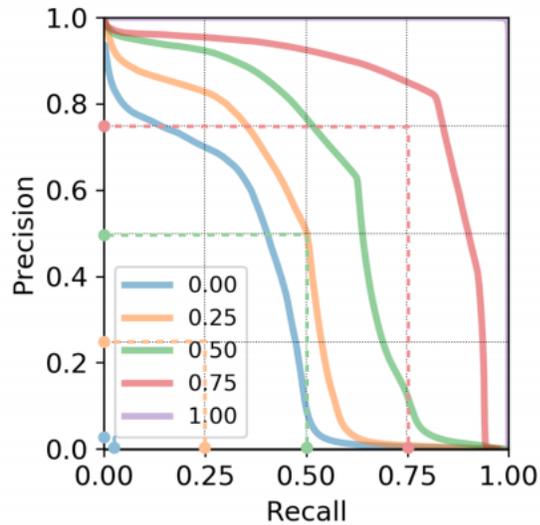


Figura A.2: Curva Precisión-Sensibilidad para distintos umbrales. Mientras la curva esté más cercana a la zona superior derecha, el modelo posee una mejor clasificación, de esta manera el mejor modelo en esta figura viene dado por la curva morada, y el peor por la curva azul. Figura obtenida de [42].

Luego, tomando las características que se obtienen para un contexto de $\pm 3[s]$ en relación al centro del evento en el EMG, se seleccionaron una serie de umbrales para alguna de sus características y se obtuvo su curva Precisión-Sensibilidad.

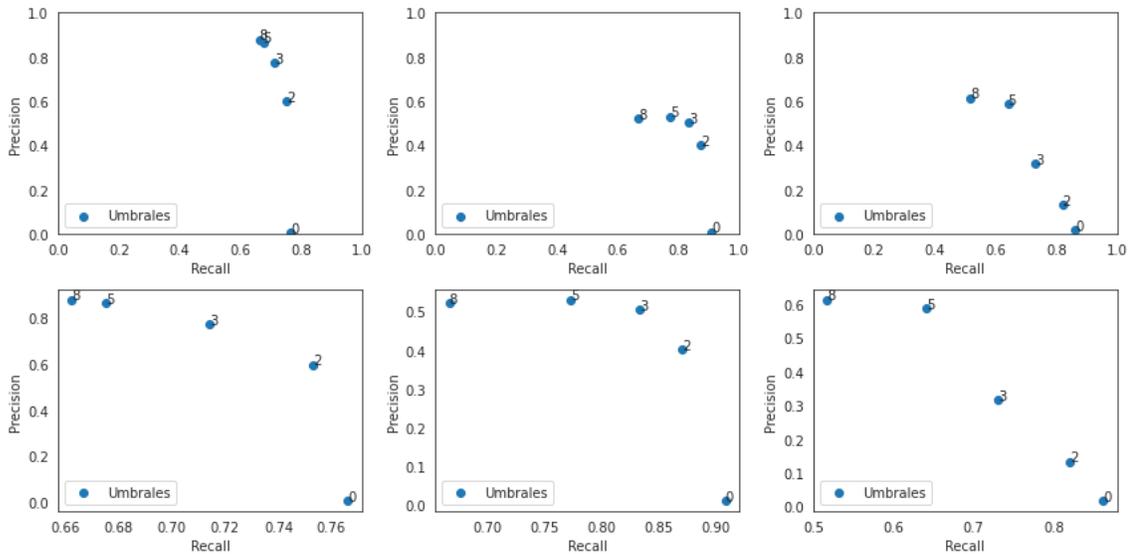


Figura A.3: Curva Precisión-Sensibilidad para distintos umbrales para la característica de número umbrales sobre una magnitud de $10[\mu V]$, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.

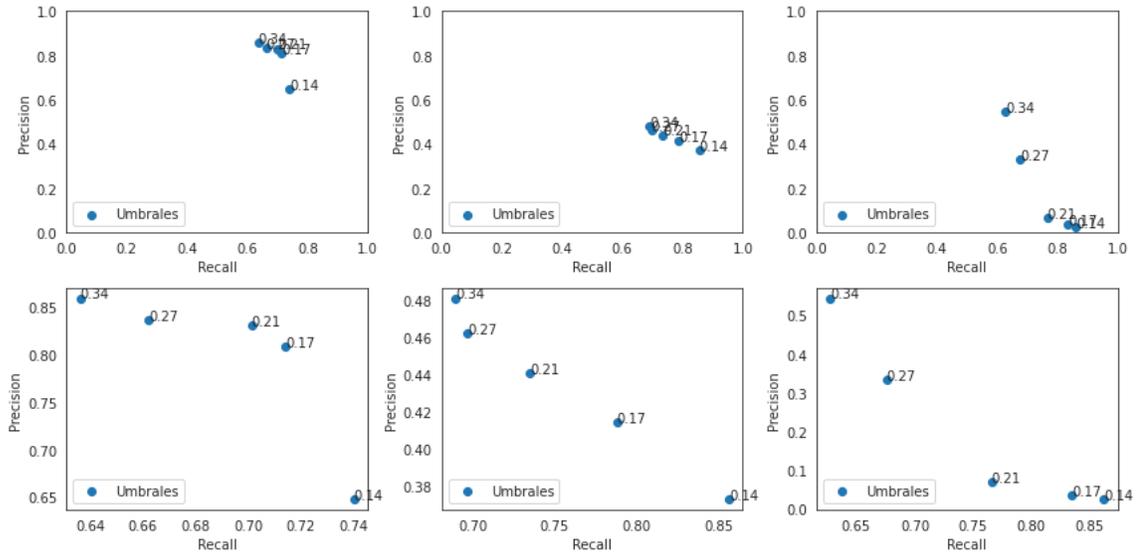


Figura A.4: Curva Precisión-Sensibilidad para distintos umbrales para la característica PSD, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.

Finalmente para un contexto de $\pm 10[s]$ se obtuvo su energía, y luego para distintos umbrales se obtuvo su curva Precisión-Sensibilidad (figura A.5).

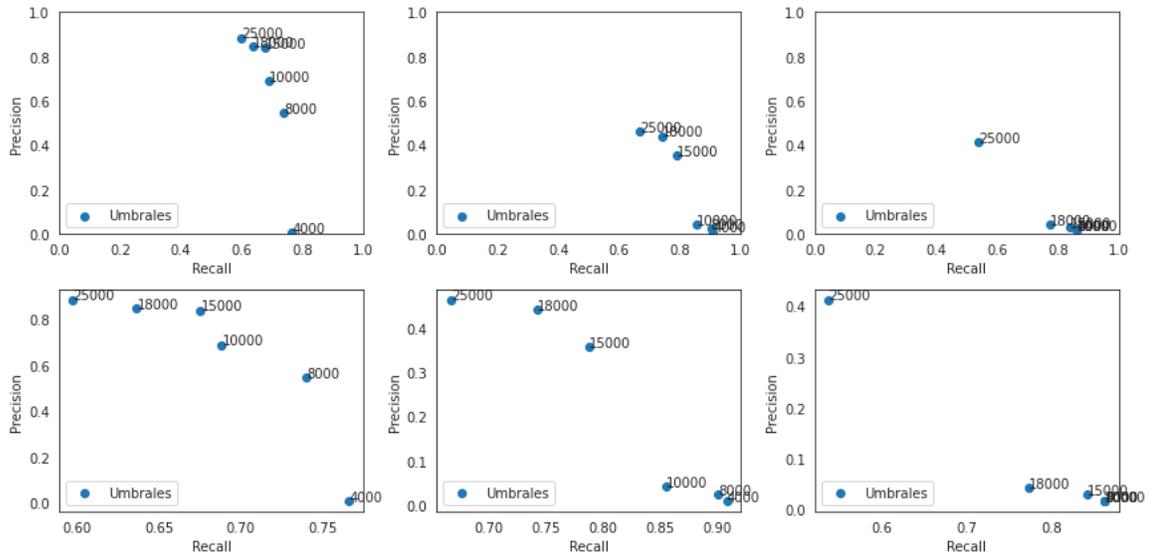


Figura A.5: Curva Precisión-Sensibilidad para distintos umbrales para la energía y un contexto de $\pm 10[s]$ en relación al centro de la detección, sobre el conjunto de validación. La parte superior corresponde a la curva Precisión-Sensibilidad y la inferior se realiza un acercamiento sobre ciertas magnitudes para una mejor visualización.

De las figuras A.3, A.4 y A.5 se obtiene que existe un fuerte trade-off en cuanto a la Precisión y Sensibilidad mirando cada una de las características por separado. De lo anterior, se infiere que no existen eventos que cumplan con tener una uniformidad entre las tres

características, esto es por ejemplo tener un determinado número de máximos por sobre un umbral de 10[s] y una alta energía a la vez. De esta manera, se debe realizar una selección de umbrales más específica y que dé cuenta de las distintas diversidades de eventos, en base a las características que se obtienen.

Anexo B

Método Out Of Bag (OOB)

El método Out Of Bag (OOB) es una forma de validar el modelo de Random Forest. En este, solo se utiliza un subconjunto de árboles de decisión para determinar la puntuación OOB. Esto conduce a reducir el efecto de agregación general en el bagging. Por lo tanto, en general, la validación en un conjunto completo de árboles de decisión es mejor que un subconjunto de árboles de decisión para estimar la puntuación. Sin embargo, en ocasiones, el conjunto de datos no es lo suficientemente grande y, por lo tanto, dejar de lado una parte para la validación no es asequible.

En consecuencia, en los casos en los que no se tiene un gran conjunto de datos y se requiere utilizar todo el dataset como el conjunto de datos de entrenamiento, la puntuación OOB proporciona una buena compensación. Además, se puede demostrar que al realizar el método OOB, al ir generando muestras por bootstrapping, en cada ajuste se utilizan dos tercios de las muestras del conjunto de entrenamiento, denominando al tercio restante sobre el que se prueba el algoritmo ajustado OOB.

Sea N el número de ejemplos en el conjunto de entrenamiento. Luego, la probabilidad de no elegir una muestra en un sampleo aleatorio es

$$\frac{N-1}{N} \tag{B.1}$$

De esta manera, utilizando un muestreo con reemplazo, la probabilidad de no elegir N muestras es

$$\left(\frac{N-1}{N}\right)^N \tag{B.2}$$

Finalmente, tomando el límite de (B.2)

$$\lim_{N \rightarrow \infty} \left(\frac{N-1}{N}\right)^N = \lim_{N \rightarrow \infty} \left(1 - \frac{1}{N}\right)^N = 0.368 \tag{B.3}$$

De esta manera, se prueba que cerca del 36,8% de las muestras del conjunto de entrenamiento están presentes como OOB.

Anexo C

Resultados obtenidos aplicando EMD (Empirical Mode Decomposition)

C.1. EMD (Empirical Mode Decomposition)

EMD [36] es un método no lineal que permite descomponer señales no estacionarias, de forma adaptativa, en una colección de señales componentes denominadas funciones modales intrínsecas (intrinsic mode functions, IMFs), las que incluyen sólo componentes de frecuencias similares. La EMD opera en forma iterativa, separando una serie temporal (señal original) en una parte de alta frecuencia (IMF) y una parte con frecuencias menores o residuales. La primera iteración sobre la señal que se analiza crea el IMF1, la segunda iteración genera el IMF2 a través de la aplicación de la EMD sobre el residuo 1, obteniendo un residuo 2. La descomposición EMD se da por finalizada una vez que la señal de residuo es cero (o suficientemente cercana a cero) [32]. Notar que la EMD va descomponiendo la señal desde frecuencias mayores a frecuencias menores de la señal, donde se pueden obtener todos los modos que hayan sido requeridos, o que hayan sido obtenidos de manera iterativa sobre la señal hasta que su residuo sea cero. Si se suman todos los modos obtenidos, se recupera la señal original.

En la figura C.1 se puede observar el funcionamiento de la EMD sobre una señal artificial compuesta por distintas frecuencias.

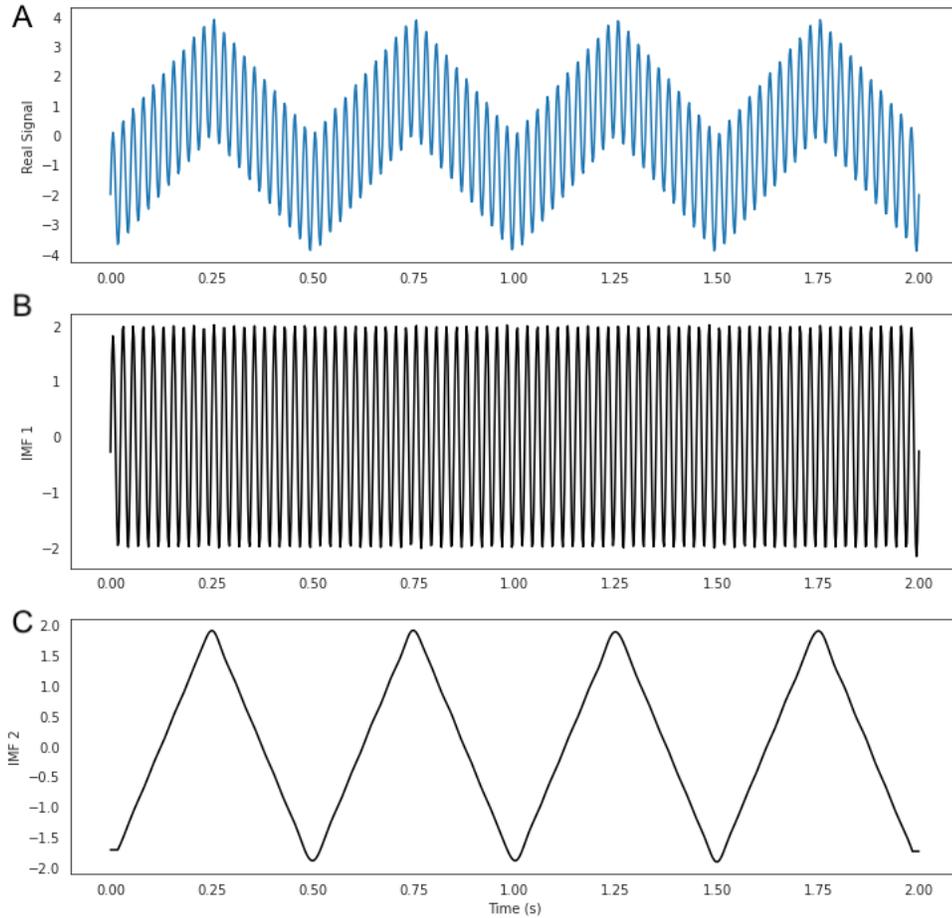


Figura C.1: Funcionamiento de la EMD aplicada sobre una señal compuesta por una señal triangular más una señal senoidal. Se pueden observar los dos primeros modos, donde se comienza desde el modo de mayor frecuencia, hacia el modo de menor frecuencia. A: Señal original. B: Primer modo. C: Segundo modo. Figura basada en figura 2.6 de [32].

En la sección 2.1 se observaron aspectos más teóricos concernientes a la EMD y transformada de Hilbert. Estas metodologías han sido ampliamente usadas en la literatura de detección de eventos fisiológicos y han demostrado ser una buena herramienta para ello [32]. Si se observa la figura C.2 se puede observar el resultado de aplicar EMD sobre el EMG, donde se obtienen 9 modos de la señal, y su respectiva transformada de Hilbert para cada modo.

De esta manera, al aplicar ambas metodologías en conjunto (EMD y transformada de Hilbert) se puede ir descomponiendo la señal en sus distintas bandas de frecuencia (EMD) y su respectiva envolvente de amplitud (transformada de Hilbert), lo cual reduce el ruido de la señal, y facilita la extracción de características.

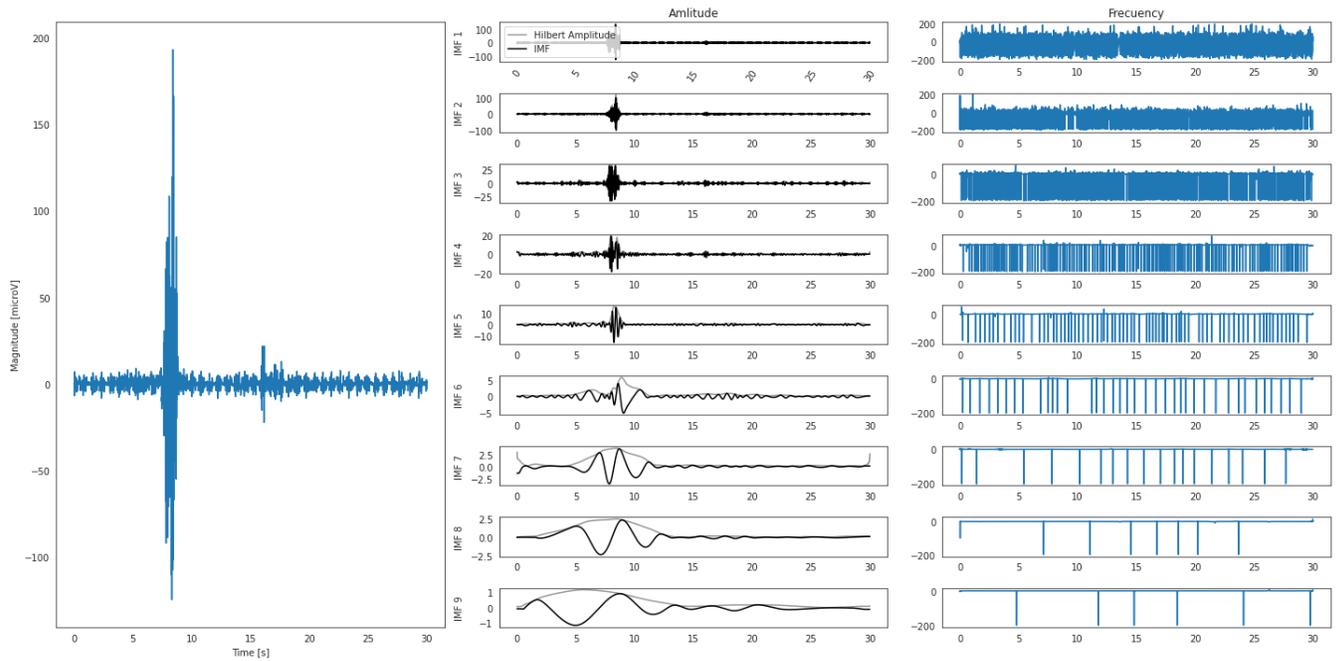


Figura C.2: Se observa una época (30 [s]) del EMG, con presencia de un evento de LM y además, el resultado de aplicar EMD y luego transformada de Hilbert sobre cada modo obtenido. Se puede observar hasta el modo 9 que se obtiene de la EMD, y su respectiva transformada de Hilbert para cada modo. En el gráfico de la amplitud se puede observar tanto la amplitud de la señal obtenida de la EMD, como la amplitud instantánea obtenida de la transformada de Hilbert. A su vez, en los gráficos de frecuencia, se puede observar la frecuencia instantánea obtenida de la transformada de Hilbert.

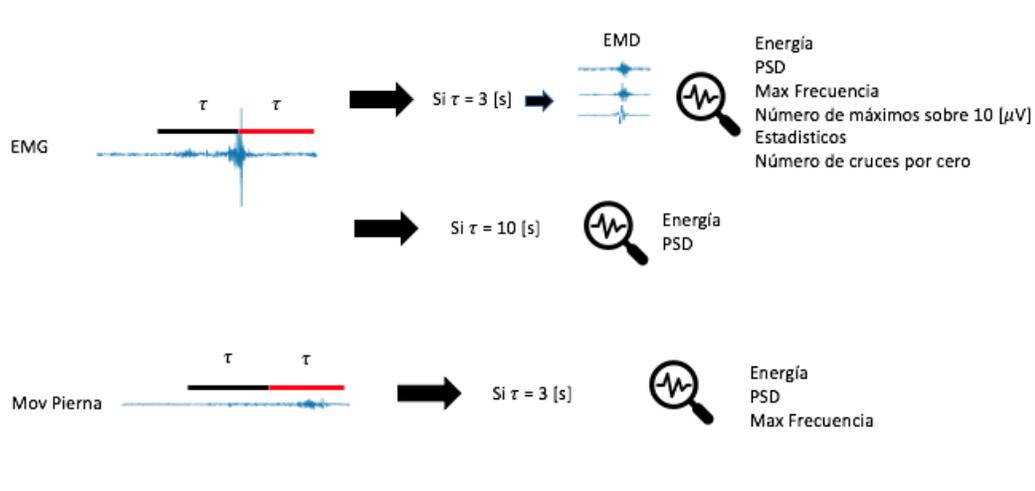


Figura C.3: Extracción de características en módulo 2 aplicando EMD hasta modo 3. En la parte superior se observa un evento de LM sobre el EMG, el contexto que se toma en relación a su centro (en rojo el tiempo posterior y negro tiempo previo al centro del evento), y las características que se obtienen sobre el canal del EMG para distintos contextos dados ($\tau = 3[s]$ o $\tau = 10[s]$). En la parte inferior se observa el canal de movimiento de piernas, el contexto dado (que es el mismo al tomado para el EMG), y las características obtenidas según el contexto τ dado. Finalmente, se puede notar que por evento de LM se obtienen 9 características en total.

C.2. Resultados

Reemplazando la DWT (nível 5) por la EMD (modo 3) para la descomposición en bandas de frecuencia aplicada en módulo 2, se obtienen los siguientes resultados sobre el conjunto de prueba.

Tabla C.1: Resultados sobre el conjunto de prueba en salida del módulo 2 aplicando EMD.

Métrica	Valor
Sensibilidad	0.88
Precisión	0,58
$F1_{score}$	0,70

Notar que los resultados son un tanto peores en cuanto a F1-score y Precisión, lo cual se debe a que solo es posible obtener hasta el modo 3 de descomposición para los eventos propuestos, esto debido a la aproximación lineal necesaria en la EMD, la cual diverge para alguna de las marcas obtenidas a la salida del primer módulo (módulo 1). Además, en cuanto a tiempo de procesamiento, la EMD al ser un algoritmo iterativo, es mucho más costoso que la DWT, lo cual también le confiere una ventaja adicional a la DWT por sobre la EMD.

Anexo D

Tablas y resultados

D.1. Ajuste de hiperparámetros del algoritmo Random Forest

Se aplicó el método OOB para la siguiente grilla de valores de hiperparámetros:

- Profundidad máxima que pueden alcanzar los árboles.: 3, 5, 10, 20,100 y 250.
- Número mínimo de observaciones que debe de tener cada uno de los nodos hijos para que se produzca la división (un valor decimal se interpreta como fracción del total de observaciones de entrenamiento): 0,05, 0,1 y 0,5.
- Número de predictores considerados en cada división: sqrt y log2.

De lo anterior, se tiene que en total se prueban 216 combinaciones distintas de hiperparámetros del algoritmo Random Forest. En la tabla D.1 se pueden observar cada uno de los resultados obtenidos para cada una de las combinaciones, así como también en la figura D.1 la convergencia de cada OOB al variar el número de árboles y el respectivo F1-score alcanzado.

Tabla D.1: Resultados de las distintas combinaciones de hiperparámetros para el algoritmo Random Forest y sus respectivos resultados aplicando OOB.

Nº	<i>oob_{f1}</i>	<i>ccp_{alpha}</i>	<i>max_{depth}</i>	<i>max_{features}</i>	<i>Leaf</i>	<i>Árboles</i>
3	0.9315	0.0	3.0	log2	0.05	250.0
9	0.9314	0.0	100.0	log2	0.05	250.0
12	0.93144	0.0	5.0	log2	0.05	250.0
48	0.93144	0.0	250.0	log2	0.05	250.0
30	0.93144	0.0	20.0	log2	0.05	250.0
21	0.931440	0.0	10.0	log2	0.05	250.0
0	0.93112	0.0	3.0	sqrt	0.05	250.0
45	0.93023	0.0	250.0	sqrt	0.05	250.0
57	0.93023	0.01	3.0	log2	0.05	250.0

Sigue en la página siguiente.

Nº	<i>oob_{f1}</i>	<i>ccP_{alpha}</i>	<i>max_{depth}</i>	<i>max_{features}</i>	<i>samples_{leaf}</i>	<i>n_{estimators}</i>
9	0.93023	0.0	5.0	sqrt	0.05	250.0
27	0.9302	0.0	20.0	sqrt	0.05	250.0
36	0.9302	0.0	100.0	sqrt	0.05	250.0
18	0.93023	0.0	10.0	sqrt	0.05	250.0
102	0.92998	0.01	250.0	log2	0.05	250.0
93	0.92998	0.01	100.0	log2	0.05	250.0
66	0.9299	0.01	5.0	log2	0.05	250.0
84	0.929983	0.01	20.0	log2	0.05	250.0
75	0.9299	0.01	10.0	log2	0.05	250.0
120	0.9296	0.05	5.0	log2	0.05	250.0
129	0.929637	0.05	10.0	log2	0.05	250.0
147	0.929637	0.05	100.0	log2	0.05	250.0
138	0.9296375	0.05	20.0	log2	0.05	250.0
156	0.9296375	0.05	250.0	log2	0.05	250.0
7	0.9294583	0.0	3.0	0.1	250.0	
97	0.929458	0.01	100.0	0.1	250.0	
43	0.929458	0.0	100.0	0.1	250.0	
106	0.92945	0.01	250.0	0.1	250.0	
52	0.92945	0.0	250.0	0.1	250.0	
61	0.92945		0.01	3.0	0.1	250.0
88	0.929458	0.01	20.0	0.1	250.0	
70	0.929458	0.01	5.0	0.1	250.0	
34	0.9294583	0.0	20.0	0.1	250.0	
79	0.9294583	0.01	10.0	0.1	250.0	
16	0.9294583	0.0	5.0	0.1	250.0	
25	0.9294	0.0	10.0	0.1	250.0	
174	0.9293	0.1	5.0	log2	0.05	250.0
175	0.929314	0.1	5.0	log2	0.1	250.0
166	0.92931	0.1	3.0	log2	0.1	250.0
183	0.929314	0.1	10.0	log2	0.05	250.0
202	0.92931	0.1	100.0	log2	0.1	250.0
184	0.92931	0.1	10.0	log2	0.1	250.0
211	0.9293144	0.1	250.0	log2	0.1	250.0
165	0.92931	0.1	3.0	log2	0.05	250.0
111	0.92931	0.05	3.0	log2	0.05	250.0
192	0.92931	0.1	20.0	log2	0.05	250.0
201	0.929314	0.1	100.0	log2	0.05	250.0
193	0.929314	0.1	20.0	log2	0.1	250.0

Sigue en la página siguiente.

Nº	<i>oob_{f1}</i>	<i>ccP_{alpha}</i>	<i>max_{depth}</i>	<i>max_{features}</i>	<i>samples_{leaf}</i>	<i>n_{estimators}</i>
210	0.92931	0.1	250.0	log2	0.05	250.0
157	0.92906	0.05	250.0	log2	0.1	250.0
112	0.92906	0.05	3.0	log2	0.1	250.0
121	0.92906	0.05	5.0	log2	0.1	250.0
148	0.92906	0.05	100.0	log2	0.1	250.0
130	0.929066	0.05	10.0	log2	0.1	250.0
139	0.92906	0.05	20.0	log2	0.1	250.0
6	0.92896	0.0	3.0	0.05	250.0	
15	0.92896	0.0	5.0	0.05	250.0	
33	0.928967	0.0	20.0	0.05	250.0	
42	0.92896	0.0	100.0	0.05	250.0	
51	0.92896	0.0	250.0	0.05	250.0	
24	0.9289	0.0	10.0	0.05	250.0	
81	0.92891	0.01	20.0	sqrt	0.05	250.0
54	0.928915	0.01	3.0	sqrt	0.05	250.0
99	0.92891	0.01	250.0	sqrt	0.05	250.0
63	0.928915	0.01	5.0	sqrt	0.05	250.0
90	0.92891	0.01	100.0	sqrt	0.05	250.0
72	0.928915	0.01	10.0	sqrt	0.05	250.0
85	0.92887	0.01	20.0	log2	0.1	250.0
76	0.928877	0.01	10.0	log2	0.1 250.0	
67	0.928877	0.01	5.0	log2	0.1	250.0
58	0.928877	0.01	3.0	log2	0.1	250.0
103	0.928877	0.01	250.0	log2	0.1	250.0
94	0.928877	0.01	100.0	log2	0.1	250.0
69	0.9287794	0.01	5.0	0.05	250.0	
105	0.9287794	0.01	250.0	0.05	250.0	
87	0.928779	0.01	20.0	0.05	250.0	
60	0.928779	0.01	3.0	0.05	250.0	
78	0.92877	0.01	10.0	0.05	250.0	
96	0.92877945	0.01	100.0	0.05	250.0	
108	0.928609	0.05	3.0	sqrt	0.05	250.0
31	0.928590	0.0	20.0	log2	0.1	250.0
40	0.9285905	0.0	100.0	log2	0.1	250.0
4	0.9285905	0.0	3.0	log2	0.1	250.0
49	0.9285905	0.0	250.0	log2	0.1	250.0
22	0.9285905	0.0	10.0	log2	0.1	250.0
13	0.92859053	0.0	5.0	log2	0.1	250.0

Sigue en la página siguiente.

Nº	oob_{f1}	ccp_{alpha}	max_{depth}	$max_{features}$	$samplesleaf$	$n_{estimators}$
118	0.92828579	0.05	5.0	sqrt	0.1	250.0
154	0.92828579	0.05	250.0	sqrt	0.1	250.0
145	0.9282857	0.05	100.0	sqrt	0.1	250.0
127	0.9282857	0.05	10.0	sqrt	0.1	250.0
109	0.9282	0.05	3.0	sqrt	0.1	250.0
136	0.9282857	0.05	20.0	sqrt	0.1	250.0
115	0.9281942	0.05	3.0	0.1	250.0	
114	0.9281942	0.05	3.0	0.05	250.0	
132	0.928194	0.05	10.0	0.05	250.0	
124	0.9281942	0.05	5.0	0.1	250.0	
187	0.92819429	0.1	10.0	0.1	250.0	
186	0.9281942	0.1	10.0	0.05	250.0	
123	0.9281942	0.05	5.0	0.05	250.0	
133	0.928194	0.05	10.0	0.1	250.0	
141	0.928194	0.05	20.0	0.05	250.0	

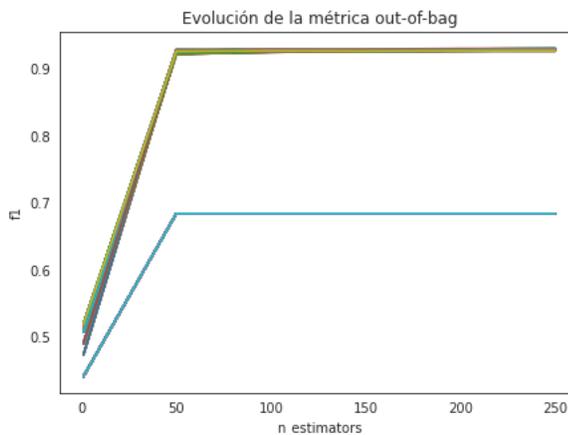


Figura D.1: Se puede observar lo obtenido aplicando el método OOB para maximizar la métrica F1-score. Cada curva representa un conjunto diferente perteneciente al conjunto de entrenamiento, que fue dejado como prueba, para distintos valores de hiperparametros del algoritmo Random Forest. Se desprende de la figura que el número de árboles mayores a una cantidad de 50, no reviste de mejoras al desempeño de la clasificación, además, hay conjuntos en que máximo se alcanza un F1-score de $\sim 0,7$ mientras que en otros se supera el $\sim 0,9$. Los resultados completos para cada una de las combinaciones de los distintos hiperparámetros se pueden observar en la Tabla D.1.

D.2. Importancia de características obtenidas en módulo 2

En la Figura D.2 se puede observar la importancia de cada una de las características obtenidas para cada marca propuesta en el módulo 1, como candidata a evento de LM. En la Figura D.2, se sigue la siguiente notación para las características:

- Nombre característica + Nivel de descomposición (0-5) para las características: percentil 5, percentil 25, percentil 75, percentil 95, Mediana, Media, Desviación estándar (std), Varianza, RMS ,Número de máximos (sobre $10[\mu V]$), Energía, PSD, Frecuencia máxima PSD.
- Por su nombre en relación a la señal considerada: Para el contexto de $\pm 10[s]$ del candidato a evento, se tienen las características energía (Energía Contexto) y PSD (PSD contexto). A su vez, para el canal de movimiento, se tienen la energía (Energía canal Mov), PSD (PSD canal Mov) y frecuencia donde se concentra la máxima PSD (Frecuencia máxima PSD canal Mov).

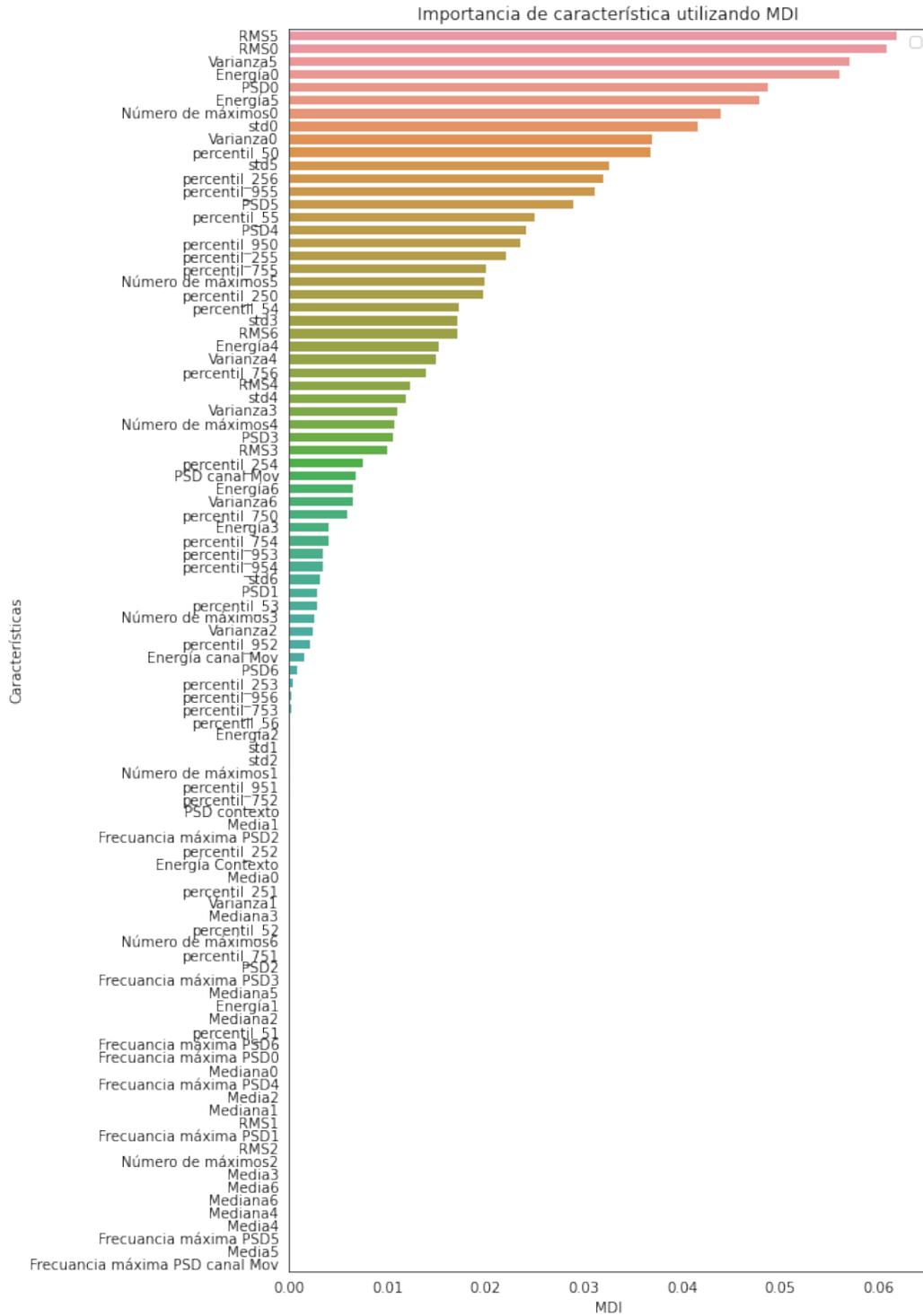


Figura D.2: Importancia de cada una de las características obtenidas para cada marca propuesta como candidato a evento de LM, en la salida del módulo 1.