RCMOS – Revista Científica Multidisciplinar O
Saber. ISSN: 2675-9128. São Paulo - SP.
Ano VII, v.1 2026 | submissão: 23/07/2026 | aceito: 26/07/2026 | publicação: 29/07/2026

Clasificación Automática de las Etapas del Sueño con Validación por Sujeto: Comparación de Random Forest, XGBoost y SVM en el Conjunto de Datos Sleep-EDFx

Classificação Automática dos Estágios do Sono com Validação por Indivíduo: Comparação entre Random Forest, XGBoost e SVM no Conjunto de Dados Sleep-EDFx

Automatic Sleep Stage Classification with Subject-Wise Validation: Comparison of Random Forest, XGBoost, and SVM on the Sleep-EDFx Dataset

Francisco Hugo César Pereira Vale – Instituto de Informática (INF), Universidad Federal de Goiás (UFG), Goiânia, Brasil, franciscovale@discente.ufg.br
Ronaldo Martins da Costa – Instituto de Informática (INF), Universidad Federal de Goiás (UFG), Goiânia, Brasil, ronaldocosta@ufg.br

Resumen: Este artículo presenta un pipeline reproducible para la clasificación automática de las etapas del sueño (Vigilia W, N1, N2, N3 y REM) utilizando EEG de un solo canal (Fpz-Cz) del conjunto de datos público Sleep-EDFx (PhysioNet), compuesto por 100 sujetos y 175 registros de dos subconjuntos distintos: Sleep Cassette (SC, 78 sujetos sanos, registros domiciliarios) y Sleep Telemetry (ST, 22 sujetos con dificultad leve para conciliar el sueño, únicamente noches con placebo). Se extrajeron 17 características por época de 30 segundos en dos dominios: estadísticas en el dominio temporal (media, varianza, desviación estándar, asimetría, curtosis y parámetros de Hjorth) y características en el dominio de la frecuencia mediante estimación espectral de Welch (potencia absoluta y relativa en las bandas Delta, Theta, Alfa y Beta, además de la razón espectral Delta/Alfa). Se compararon tres clasificadores de naturalezas algorítmicas fundamentalmente diferentes (Random Forest, XGBoost y Máquina de Vectores de Soporte con kernel RBF) mediante validación cruzada Leave-One-Group-Out (LOGO), evitando cualquier fuga de datos entre sujetos. Los experimentos se realizaron en tres escenarios: solo SC, solo ST y el conjunto combinado. Random Forest obtuvo el mejor desempeño general, con Kappa de Cohen de 0,587 en SC y 0,583 en el conjunto combinado. La convergencia de los tres clasificadores hacia un intervalo de Kappa similar (0,53 a 0,59) indica que el techo de desempeño está determinado principalmente por el conjunto de características y la superposición fisiológica inherente entre las etapas, más que por la elección del clasificador.

Palabras clave: Procesamiento de señales EEG. Aprendizaje automático. Polisomnografía. Estadificación del sueño.

Resumo: Este artigo apresenta um pipeline reprodutível para a classificação automática dos estágios do sono (Vigília W, N1, N2, N3 e REM) utilizando EEG de canal único (Fpz-Cz) do conjunto de dados público Sleep-EDFx (PhysioNet), composto por 100 indivíduos e 175 registros de dois subconjuntos distintos: Sleep Cassette (SC, 78 indivíduos saudáveis, registros domiciliares) e Sleep Telemetry (ST, 22 indivíduos com dificuldade leve para iniciar o sono, somente noites com placebo). Foram extraídas 17 características por época de 30 segundos em dois domínios: estatísticas no domínio do tempo (média, variância, desvio padrão, assimetria, curtose e parâmetros de Hjorth) e características no domínio da frequência por meio da estimativa espectral de Welch (potência absoluta e relativa nas bandas Delta, Teta, Alfa e Beta, além da razão espectral Delta/Alfa). Três classificadores de naturezas algorítmicas fundamentalmente diferentes foram comparados (Random Forest, XGBoost e Máquina de Vetores de Suporte com kernel RBF) sob validação cruzada Leave-One-Group-Out (LOGO), impedindo qualquer vazamento de dados entre indivíduos. Os experimentos foram conduzidos em três cenários: somente SC, somente ST e o conjunto combinado. O Random Forest obteve o melhor desempenho geral, com Kappa de Cohen de 0,587 no SC e 0,583 no

conjunto combinado. A convergência dos três classificadores para uma faixa de Kappa semelhante (0,53 a 0,59) indica que o teto de desempenho é determinado principalmente pelo conjunto de características e pela sobreposição fisiológica inerente entre os estágios, e não pela escolha do classificador.

Palavras-chave: Processamento de sinais de EEG. Aprendizado de máquina. Polissonografia. Estadiamento do sono.

Abstract: This paper presents a reproducible pipeline for automatic classification of sleep stages (Wake W, N1, N2, N3, and REM) using single-channel EEG (Fpz-Cz) from the public Sleep-EDFx dataset (PhysioNet), comprising 100 subjects and 175 recordings from two distinct subsets: Sleep Cassette (SC, 78 healthy subjects, home recordings) and Sleep Telemetry (ST, 22 subjects with mild sleep onset difficulty, placebo nights only). A total of 17 features were extracted per 30-second epoch across two domains: time-domain statistics (mean, variance, standard deviation, skewness, kurtosis, and Hjorth parameters) and frequency-domain features via Welch spectral estimation (absolute and relative power in the Delta, Theta, Alpha, and Beta bands, plus the Delta/Alpha spectral ratio). Three classifiers of fundamentally different algorithmic natures were compared (Random Forest, XGBoost, and Support Vector Machine with RBF kernel) under Leave-One-Group-Out (LOGO) cross-validation, preventing any data leakage between individuals. Experiments were conducted in three scenarios: SC alone, ST alone, and the combined dataset. Random Forest achieved the best overall performance, with Cohen's Kappa of 0.587 on SC and 0.583 on the combined dataset. The convergence of all three classifiers to a similar Kappa range (0.53 to 0.59) indicates that the performance ceiling is primarily determined by the feature set and the inherent physiological overlap between stages, rather than by the choice of classifier.

Keywords: EEG signal processing. Machine learning. Polysomnography. Sleep staging.

1. Introducción

El sueño es un proceso fisiológico fundamental para la homeostasis del organismo, directamente asociado con la consolidación de la memoria [1], la regulación metabólica, la función inmunitaria y la salud cardiovascular [2]. Los trastornos del sueño afectan a una proporción significativa de la población mundial y constituyen un factor de riesgo para afecciones como hipertensión, diabetes tipo 2, depresión y enfermedades neurodegenerativas [2]. El diagnóstico clínico de estas afecciones depende de la identificación precisa de las etapas del sueño (Vigilia W, N1, N2, N3 y REM) a partir de registros polisomnográficos (PSG).

El estándar de referencia actual para este diagnóstico es la polisomnografía (PSG) con estadificación manual realizada por especialistas capacitados, quienes analizan visualmente el electroencefalograma (EEG), el electrooculograma (EOG) y el electromiograma (EMG), segmentados en épocas de puntuación de 30 segundos (en el sentido clínico de la PSG [3], sin confundirlas con épocas de entrenamiento en la optimización de redes neuronales) y clasificados según los criterios de la American Academy of Sleep Medicine (AASM) [3]. Aunque está clínicamente establecido, este proceso presenta limitaciones significativas:

requiere tiempo, es costoso y está sujeto a variabilidad entre evaluadores. Los estudios de metaanálisis indican que el Kappa de Cohen entre especialistas humanos para la estadificación global es aproximadamente 0,76, pero disminuye a solo 0,24 para la etapa N1, lo que muestra que incluso evaluadores experimentados discrepan sustancialmente en las etapas de transición [4]. Esta variabilidad intrínseca representa un techo natural para cualquier sistema automático supervisado, ya que los modelos se entrenan con anotaciones que ya contienen este ruido de concordancia [5].

Ante este escenario, la clasificación automática de las etapas del sueño mediante aprendizaje automático ha surgido como un área de investigación clínica y tecnológicamente relevante. Los enfoques basados en características elaboradas manualmente (extracción manual de características espectrales y estadísticas del EEG), seguidos de clasificadores clásicos como Random Forest, SVM y XGBoost, demuestran un desempeño competitivo y ventajas de interpretabilidad frente a las arquitecturas de aprendizaje profundo, especialmente cuando el número de sujetos disponibles es limitado [6], [7]. Al mismo tiempo, los enfoques de redes neuronales profundas como CNN 1D, LSTM y Transformers han informado valores de Kappa superiores a 0,80 en conjuntos de datos a gran escala con múltiples canales de PSG [8], [9].

Este trabajo evalúa la clasificación automática de las etapas del sueño utilizando el conjunto de datos público Sleep-EDFx de PhysioNet [10], compuesto por los subconjuntos Sleep Cassette (SC) y Sleep Telemetry (ST). Se comparan tres clasificadores de distintas naturalezas algorítmicas (Random Forest, XGBoost y SVM) mediante 17 características extraídas de EEG de un solo canal (Fpz-Cz), con una rigurosa validación Leave-One-Group-Out (LOGO).

2. Metodología

2.1 Conjunto de Datos

Sleep-EDFx es un conjunto de datos público disponible en PhysioNet [10], ampliamente utilizado como referencia en la literatura sobre estadificación automática del sueño [6], [9]. Se utilizaron dos subconjuntos: Sleep Cassette (SC), con 78 sujetos sanos de 25 a 101 años, sin medicación relacionada con el sueño, que totaliza 153 registros de aproximadamente 20 horas realizados en un entorno domiciliario; y Sleep Telemetry (ST), con 22 sujetos con dificultad leve para conciliar el sueño, registrados en un entorno hospitalario, utilizando exclusivamente las noches de la condición Placebo (22 registros de aproximadamente 9 horas), identificadas en la hoja oficial de metadatos de PhysioNet ST-subjects.xls. Las anotaciones siguen la convención de Rechtschaffen & Kales (R&K), con las etapas N3 y N4 fusionadas en una sola

clase N3 conforme a la convención actual de la AASM [3], lo que da como resultado cinco clases: W, N1, N2, N3 y REM.

Declaración de ética: este estudio utilizó exclusivamente datos secundarios de acceso público del conjunto Sleep-EDFx (PhysioNet). No se realizó ninguna interacción directa con sujetos humanos. La recopilación original de datos se llevó a cabo de acuerdo con la Declaración de Helsinki y fue aprobada por los comités de ética locales de las instituciones responsables. Los datos están completamente anonimizados y los sujetos se identifican únicamente mediante códigos numéricos. Su uso con fines de investigación está permitido por la PhysioNet Credentialed Health Data License. No se requirió aprobación adicional de un comité de ética para este trabajo.

La Figura 1 ilustra las etapas del pipeline propuesto, desde la lectura de los archivos EDF hasta la evaluación final de los clasificadores.

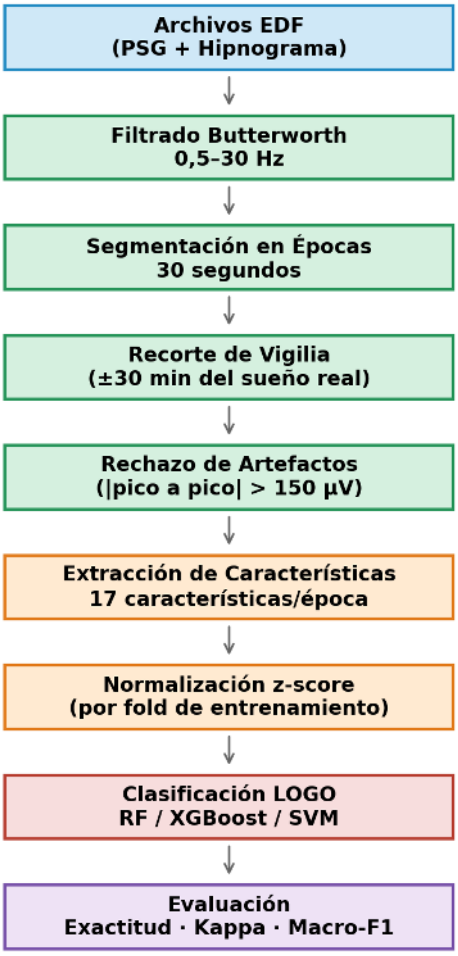


Figura 1. Pipeline de clasificación automática de las etapas del sueño. Las etapas de preprocesamiento se aplican individualmente a cada registro antes de la extracción de características.

2.2 Preprocesamiento

Las señales de EEG del canal Fpz-Cz se filtraron con un filtro pasabanda Butterworth de cuarto orden (0,5 a 30 Hz), preservando las bandas Delta, Theta, Alfa y Beta relevantes para la estadificación del sueño. Las épocas con amplitud pico a pico superior a 150 µV se descartaron mediante rechazo de artefactos. Para reducir el desequilibrio artificial de la clase Vigilia, se aplicó un procedimiento de recorte de vigilia: solo se conservaron las épocas situadas hasta 30 minutos antes de la primera época real de sueño (N1, N2, N3 o REM) y hasta 30 minutos después de la última, descartando los largos períodos de vigilia en los extremos del registro.

La Figura 2 ilustra el hipnograma de un sujeto representativo después del recorte de vigilia, mostrando la estructura cíclica típica del sueño humano, con alternancia entre N1, N2, N3 y REM durante la noche y predominio de la etapa N2 [3].

RCMOS – Revista Científica Multidisciplinar O
Saber. ISSN: 2675-9128. São Paulo - SP.
Ano VII, v.1 2026 | submissão: 23/07/2026 | aceito: 26/07/2026 | publicação: 29/07/2026

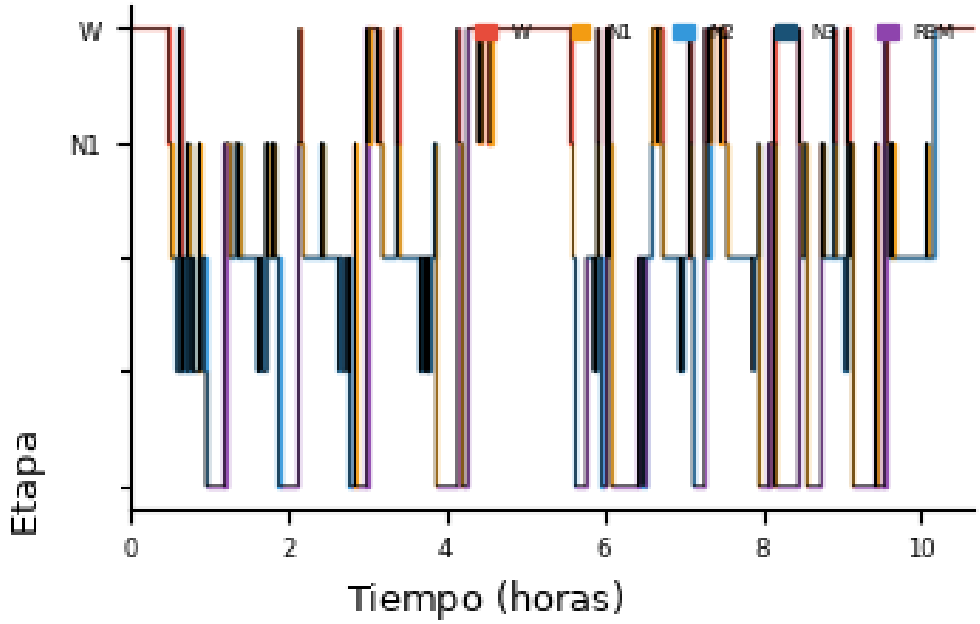


Figura 2. Ejemplo de hipnograma después del recorte de vigilia. Cada color representa una etapa del sueño: Vigilia (W), N1, N2, N3 y REM.

2.3 Extracción de Características

Se extrajeron 17 características por época de 30 segundos del canal de EEG Fpz-Cz, organizadas en dos dominios. En el dominio temporal: media, varianza, desviación estándar, asimetría, curtosis y los tres parámetros de Hjorth (actividad, movilidad y complejidad). En el dominio de la frecuencia, mediante estimación espectral de Welch: potencia absoluta y relativa en las bandas Delta (0,5 a 4 Hz), Theta (4 a 8 Hz), Alfa (8 a 12 Hz) y Beta (12 a 30 Hz), además de la razón espectral Delta/Alfa, un indicador relevante para distinguir el sueño profundo de la vigilia.

2.4 Clasificadores y Validación

Se evaluaron tres clasificadores de naturalezas algorítmicas fundamentalmente diferentes: Random Forest (ensemble por bagging, 300 estimadores, pesos de clase balanceados), XGBoost (ensemble por boosting, 300 estimadores, profundidad máxima 6) y Máquina de Vectores de Soporte con kernel RBF, con balanceo de clases aplicado en todos los modelos. Para la SVM, debido a la complejidad computacional cuadrática del kernel RBF en grandes conjuntos de datos, cada fold se entrenó con una submuestra estratificada de hasta 8.000 épocas, conservando las proporciones originales de las clases, una práctica estándar en la literatura para clasificadores basados en kernel a escala [7]. Las características se normalizaron mediante

estandarización z-score ajustada exclusivamente en el fold de entrenamiento, evitando la fuga de información estadística hacia el conjunto de prueba.

La validación siguió la estrategia Leave-One-Group-Out (LOGO), en la que cada fold excluye completamente a un sujeto para la prueba y se entrena con todos los demás, garantizando que ninguna época del mismo sujeto aparezca simultáneamente en los conjuntos de entrenamiento y prueba. Los tres clasificadores se evaluaron en tres escenarios independientes: solo Sleep Cassette (78 folds), solo Sleep Telemetry Placebo (22 folds) y conjunto combinado (100 folds).

3. Resultados

La Tabla 1 presenta los resultados consolidados de los tres clasificadores en los tres escenarios de evaluación, expresados en Exactitud, Macro-F1 y Kappa de Cohen.

Tabla 1. Desempeño de los Clasificadores por Escenario (LOGO)

Escenario	Clasificador	Exactitud	Macro-F1	Kappa
Cassette (n=78)	Random Forest	0.705	0.625	0.587
Cassette (n=78)	XGBoost	0.675	0.630	0.568
Cassette (n=78)	SVM (RBF)*	0.642	0.602	0.531
Telemetry (n=22)	Random Forest	0.687	0.596	0.548
Telemetry (n=22)	XGBoost	0.704	0.629	0.582
Telemetry (n=22)	SVM (RBF)*	0.673	0.605	0.554
Combinado (n=100)	Random Forest	0.700	0.625	0.583
Combinado (n=100)	XGBoost	0.673	0.631	0.567
Combinado (n=100)	SVM (RBF)*	0.646	0.606	0.536

**SVM: submuestra estratificada de 8.000 épocas/fold.*

Random Forest obtuvo el mejor desempeño general, con Kappa de Cohen de 0,587 en el escenario Cassette y 0,583 en el conjunto combinado. XGBoost presentó resultados competitivos y superó a Random Forest en el escenario Telemetry (Kappa de 0,582 frente a 0,548). La SVM con kernel RBF, evaluada con una submuestra de entrenamiento, mostró menor exactitud pero un Macro-F1 comparable, lo que destaca su sensibilidad al volumen de datos de entrenamiento disponible.

Se observa que la combinación de los dos subconjuntos (SC + ST, 100 sujetos) no produjo una degradación del desempeño respecto de los escenarios aislados, manteniendo el Kappa entre 0,53 y 0,58 para todos los clasificadores. Esto sugiere que los patrones de sueño en el EEG de sujetos sanos (Cassette) y de sujetos con dificultad leve para conciliar el sueño bajo placebo

(Telemetry) son suficientemente homogéneos para que un clasificador entrenado conjuntamente generalice en ambas poblaciones.

Entre las 17 características extraídas, las del dominio espectral tienden a ser fisiológicamente más informativas para la estadificación del sueño, ya que cada etapa presenta una firma de banda característica: las ondas delta predominan en el sueño profundo (N3), los husos del sueño y los complejos K son marcadores de N2, y las actividades theta y alfa caracterizan N1 y vigilia, respectivamente [3]. Se prevé un análisis cuantitativo de la importancia relativa de las características como etapa posterior de este trabajo.

La Figura 3 presenta la matriz de confusión normalizada del clasificador Random Forest en el escenario combinado, agregando todas las predicciones fuera de muestra de los 100 folds LOGO.

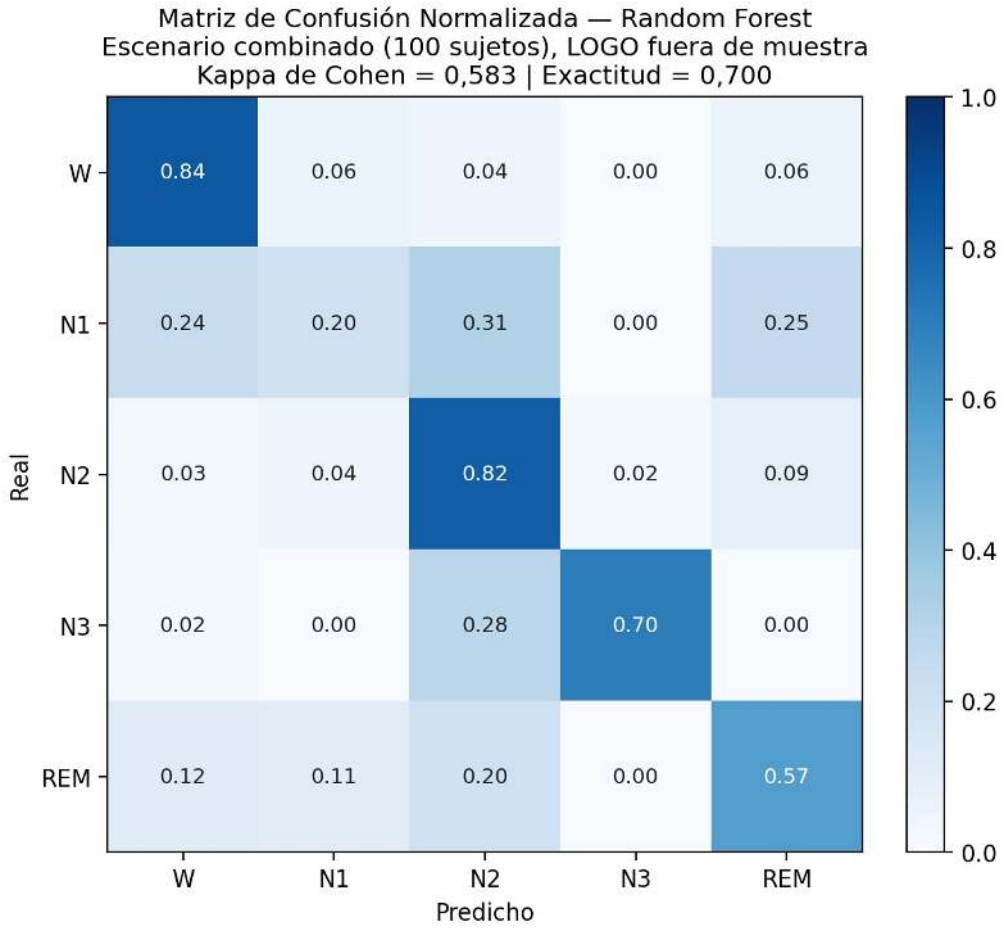


Figura 3. Matriz de confusión normalizada de Random Forest, escenario combinado (100 sujetos), que agrega todas las predicciones LOGO fuera de muestra. Los valores representan la proporción de épocas reales clasificadas en cada etapa predicha.

Las Figuras 4, 5 y 6 presentan, respectivamente, la Exactitud, el Kappa de Cohen y el Macro-F1 de los tres clasificadores en los tres escenarios evaluados, lo que permite una comparación visual directa entre los algoritmos.

RCMOS – Revista Científica Multidisciplinar O
Saber. ISSN: 2675-9128. São Paulo - SP.
Ano VII, v.1 2026 | submissão: 23/07/2026 | aceito: 26/07/2026 | publicação: 29/07/2026

En cuanto a la Exactitud (Figura 4), Random Forest obtuvo los mejores resultados en los escenarios Cassette (0,705) y Combinado (0,700), mientras que XGBoost superó a los demás en el escenario Telemetry (0,704). La SVM con kernel RBF mostró una exactitud consistentemente menor, lo que refleja el submuestreo de entrenamiento adoptado debido a restricciones computacionales.

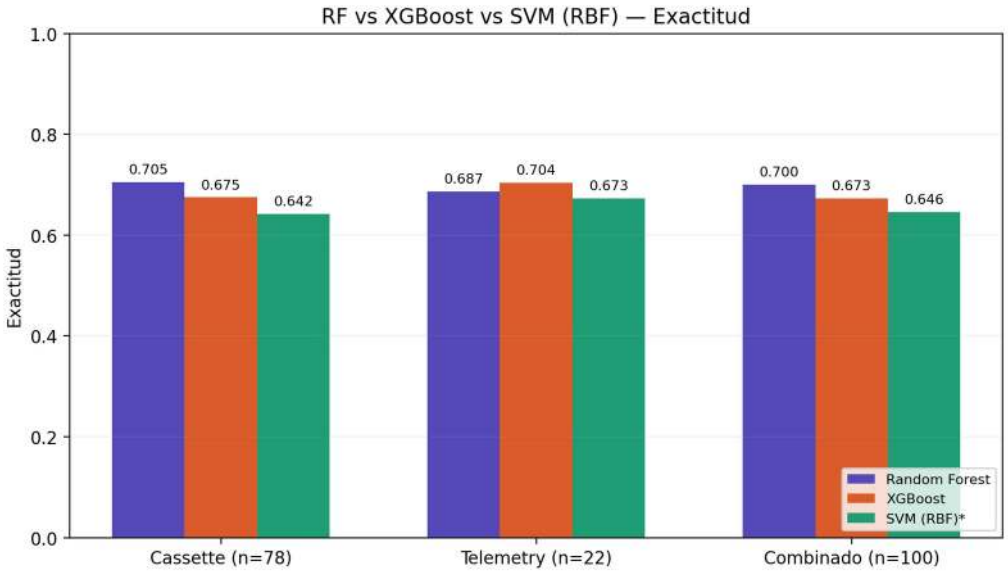


Figura 4. Exactitud de los clasificadores por escenario (LOGO). *SVM entrenada con una submuestra de 8.000 épocas/fold.

El Kappa de Cohen (Figura 5), métrica estándar en la literatura sobre estadificación del sueño porque descuenta el efecto de la concordancia por azar, confirma el mismo patrón: Random Forest lidera en Cassette (0,587) y Combinado (0,583), XGBoost en Telemetry (0,582), y la SVM mantiene un desempeño competitivo dada la restricción de entrenamiento (0,531 a 0,554). Todos los valores se sitúan en el intervalo de moderado a sustancial según la escala de Landis & Koch, comparables con el límite inferior de concordancia humana entre evaluadores para etapas individuales informado en la literatura [4].

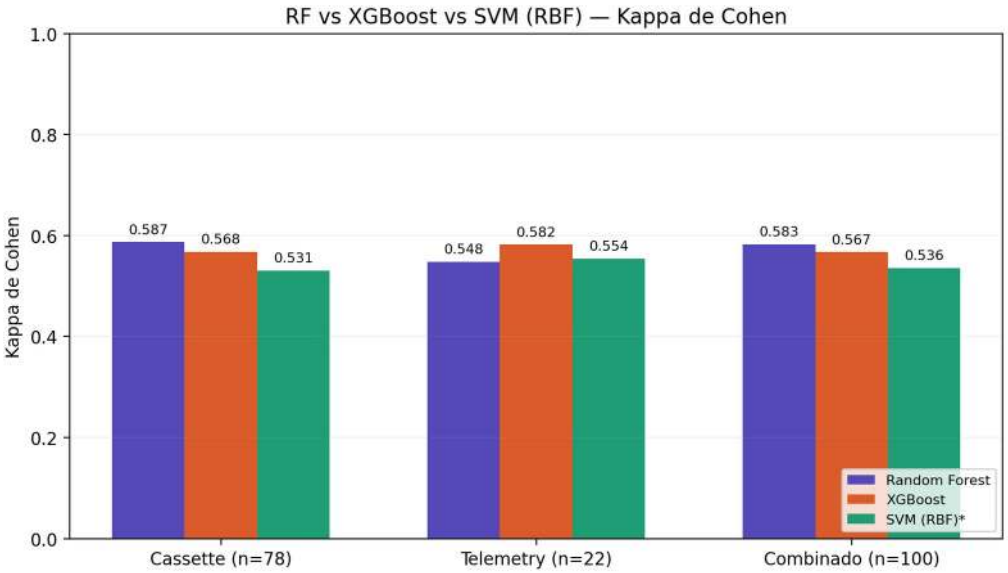


Figura 5. Kappa de Cohen de los clasificadores por escenario (LOGO).

El Macro-F1 (Figura 6) revela un patrón distinto: XGBoost presenta el mayor Macro-F1 en los tres escenarios (0,630, 0,629 y 0,631), superando a Random Forest, que lidera en Exactitud y Kappa. Esta inversión indica que XGBoost clasifica mejor las clases minoritarias (especialmente N1 y N3), mientras que Random Forest tiende a concentrar las predicciones correctas en las clases mayoritarias (N2 y Vigilia). Para aplicaciones clínicas en las que la identificación correcta de todas las etapas es igualmente relevante, Macro-F1 constituye una métrica complementaria indispensable para la Exactitud y el Kappa de Cohen.

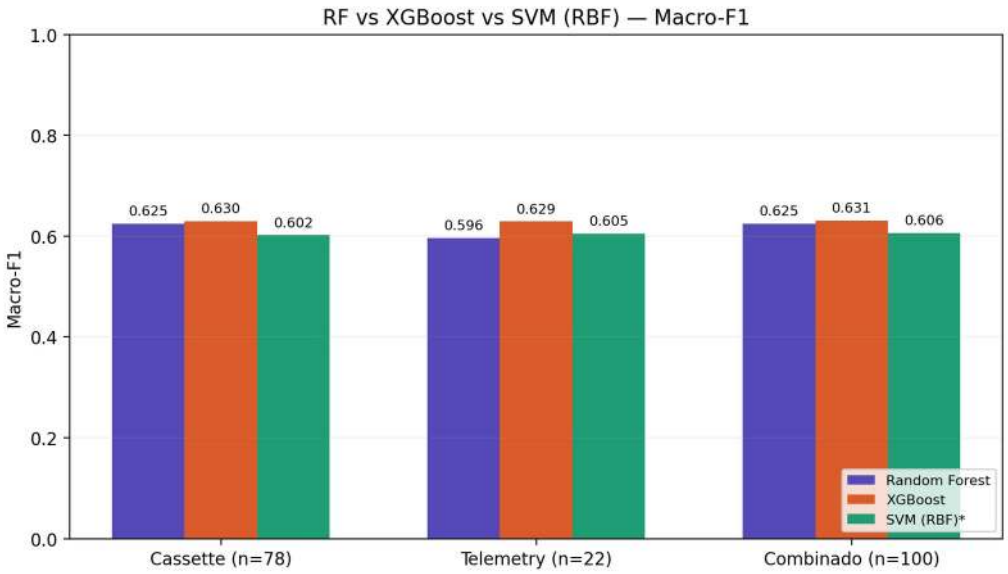


Figura 6. Macro-F1 de los clasificadores por escenario (LOGO).

4. Discusión

Los valores de Kappa de Cohen obtenidos (0,53 a 0,59) se sitúan en el intervalo de moderado a sustancial según la escala de Landis & Koch, comparables con el límite inferior de concordancia humana entre evaluadores informado para etapas individuales en la literatura. En particular, el Kappa humano para N1 es de solo 0,24 [4], etapa que también constituye el principal desafío para los clasificadores automáticos de este trabajo. El hecho de que tres algoritmos de naturalezas fundamentalmente diferentes (bagging, boosting y frontera geométrica) converjan en el mismo intervalo de desempeño sugiere que el techo observado es una propiedad del conjunto de características y de la superposición fisiológica inherente entre las etapas, más que un artefacto de la elección del clasificador.

La estrategia de validación LOGO adoptada es más rigurosa que el k-fold aleatorio empleado con frecuencia en la literatura, ya que elimina cualquier posibilidad de fuga de información entre épocas del mismo sujeto, una fuente común de optimismo artificial en trabajos que dividen por época en lugar de por sujeto [7].

La matriz de confusión normalizada de Random Forest en el escenario combinado (Figura 3) destaca patrones de error coherentes con la literatura sobre estadificación automática del sueño. La etapa N2 alcanza la mayor tasa de reconocimiento (0,83), reflejo de su alta prevalencia en el conjunto de datos y de sus marcadores electrofisiológicos característicos. La etapa REM muestra una buena discriminación (0,70), con confusión principalmente con N1 (0,14) y N2 (0,13). La etapa N3 se confunde frecuentemente con N2 (0,57), una dificultad esperada para características de bandas agregadas en ventanas de 30 segundos. La etapa N1 es la más desafiante (0,31), con errores distribuidos entre W, N2 y REM [4].

5. Conclusión

Este trabajo presentó un pipeline reproducible para la clasificación automática de las etapas del sueño utilizando el conjunto Sleep-EDFx de PhysioNet, con una rigurosa validación por sujeto (LOGO) y la comparación de tres clasificadores de naturalezas distintas. Los resultados demuestran que los enfoques basados en características de EEG de un solo canal elaboradas manualmente son viables e interpretables, y alcanzan un Kappa de Cohen de hasta 0,587 con Random Forest, compatible con la literatura para configuraciones similares de un solo canal y validación LOGO.

La elección de clasificadores clásicos frente a arquitecturas de aprendizaje profundo fue deliberada: modelos como Random Forest, XGBoost y SVM presentan un menor costo

computacional de inferencia, requisito fundamental para su implementación en dispositivos móviles y sistemas embebidos con recursos de procesamiento limitados. Esto posiciona a los clasificadores evaluados como candidatos naturales para su integración en aplicaciones de monitorización del sueño en tiempo real, donde la eficiencia computacional es tan relevante como la exactitud del modelo.

Como trabajo futuro, se propone: (i) desarrollar una aplicación web que ofrezca el pipeline propuesto como servicio, permitiendo a los usuarios cargar archivos EDF y obtener la clasificación automática de las etapas del sueño con generación interactiva de hipnogramas, democratizando el acceso a la estadificación automatizada sin requerir infraestructura local; (ii) ampliar las características mediante la incorporación de canales EOG y EMG y contexto temporal entre épocas consecutivas; (iii) realizar un análisis sistemático de la importancia de las características, identificando las bandas espectrales y los parámetros estadísticos más discriminativos por etapa, con el potencial de reducir el conjunto de características y optimizar el costo computacional para la implementación móvil; y (iv) evaluar la SVM con kernel RBF en el conjunto completo por fold, utilizando implementaciones escalables como SGDClassifier con aproximación de kernel mediante el método de Nyström, eliminando la necesidad de submuestreo.

Referencias

DAULATABAD, V. et al. Systematic review: the impact of poor sleep quality on psychological and physical health. Journal of Heart Valve Disease, v. 30, p. 79–85, 2025.

FIORILLO, L.; MONACHINO, A.; VAN DER MEER, M.; WARNCKE, V.; SCHMIDT, M.; WENZ, F.; GRUN, R.; FARACI, F. Bridging AI and clinical practice: Integrating automated sleep scoring algorithm with uncertainty-guided physician review. arXiv preprint arXiv:2312.14996, 2023.

GOLDBERGER, A. L.; AMARAL, L. A. N.; GLASS, L.; HAUSDORFF, J. M.; IVANOV, P. Ch.; MARK, R. G.; MIETUS, J. E.; MOODY, G. B.; PENG, C.-K.; STANLEY, H. E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, v. 101, n. 23, p. e215–e220, 2000.

IBER, C.; ANCOLI-ISRAEL, S.; CHESSON, A.; QUAN, S. F. The AASM Manual for the Scoring of Sleep and Associated Events. Westchester, IL: American Academy of Sleep Medicine, 2007.

LEE, Y. J.; LEE, J. Y.; CHO, J. H.; CHOI, J. H. Interrater reliability of sleep stage scoring: a meta-analysis. Journal of Clinical Sleep Medicine, v. 18, n. 1, p. 193–202, 2022.

MASAD, I. S.; ALQUDAH, A.; QAZAN, S. Automatic classification of sleep stages using EEG signals and convolutional neural networks. PLOS ONE, v. 19, n. 1, p. e0297582, Jan. 2024.

SEKKAL, R. N.; BEREKSI-REGUIG, F.; RUIZ-FERNANDEZ, D.; DIB, N.; SEKKAL, S. Automatic sleep stage classification: From classical machine learning methods to deep learning. Biomedical Signal Processing and Control, v. 77, p. 103751, 2022.

VAN DER AAR, J. F.; VAN DEN ENDE, D. A.; VAN MEULEN, F. B.; VAN GILST, M. M. Deep transfer learning for automated single-lead EEG sleep staging with channel and population mismatches. Frontiers in Physiology, v. 14, Jan. 2024.

WEIGHALL, A.; KELLAR, I. Sleep and memory consolidation in healthy, neurotypical children and adults: a summary of systematic reviews and meta-analyses. Emerging Topics in Life Sciences, v. 7, n. 5, p. 513–524, Dec. 2023.

XU, X.; ZHANG, B.; XU, T.; TANG, J. An effective and interpretable sleep stage classification approach using multi-domain EEG and EOG features. Bioengineering, v. 12, n. 3, p. 286, 2025.