

Calidad del sueño, IA y Equinoterapia : Calibración de actígrafos de bajo coste

Omar Teixeira¹ Gian C. Dal Boni¹, Mirko Fañez², Victor M. González³, M.
Dolores Apolo⁴, and Enrique A. de la Cal¹

¹ University of Oviedo, Computer Science Department,
Faculty of Geology, Oviedo, Spain
`{uo281847,fanezmirko,delacal}@uniovi.es`

² Hospital Universitario de Asturias, Unidad del Sueño, Neurología,
Oviedo, Spain
`gianalboni@gmail.com`

³ University of Oviedo, Electrical Engineering Department,
EPI, Gijón, Spain
`vmsuarez@uniovi.es`

⁴ University of Extremadura, 09001, Badajoz-Cáceres, Spain
`mdapolo@unex.es`

Abstract. La calidad del sueño es fundamental para el bienestar general, ya que impacta en la salud física y mental, mejorando el ánimo, la concentración y fortaleciendo el sistema inmunológico. Ensayos clínicos han estudiado tratamientos como la melatonina, el ejercicio físico y la relajación muscular, así como terapias alternativas, entre ellas la equinoterapia, que busca mejorar el sueño y la salud emocional a través de la interacción con caballos.

La Alianza Internacional por la Interacción Saludable con Equinos (AINISE) trabaja para que la equinoterapia sea reconocida como terapia en España. En 2020, realizaron un ensayo piloto con pacientes de artritis, comparando los efectos de la equinoterapia con un programa de ejercicio físico en cuanto a dolor, movilidad y calidad del sueño, monitorizando la actividad diaria mediante el TICWATCH E2. Sin embargo, se encontraron errores en la app del reloj al medir el sueño, lo que motivó un ensayo en el Hospital Universitario Central de Asturias (HUCA) para calibrar los datos con un polisomnógrafo y mejorar la precisión en las fases de sueño. Se emplearon como técnicas machine learning para aprender las fases de sueño 11 técnicas clásicas de las cuales aquí se presentan los resultados obtenidos con RandomForest (RF) y eXtraTrees (XT). En concreto los mejores resultados obtenidos considerando dos fases de sueño (Wake y Sleep) fueron de un macro-F1 de 0,93/0,93 y 0,71/0,70 para un 10CV (Cross-Validation) y LOPO (Leave One Participant Out) respectivamente. Los resultados para 5 fases de sueño fueron sensiblemente peores, lo cual nos hace reflexionar sobre varios aspectos, tales como las técnicas de normalización de las series temporales, la variabilidad de la biomecánica de los pacientes y el desequilibrio en el número de datos de cada fase de sueño.

Keywords: Calidad del sueño, Equinoterapia, Fases de sueño, Machine Learning

1 Introducción

La calidad del sueño es clave para el bienestar general, impactando en la salud física y mental. Dormir bien mejora el ánimo, la concentración, y fortalece el sistema inmunológico. Muchos ensayos clínicos evalúan la calidad del sueño para estudiar tratamientos como la melatonina [11], el ejercicio físico [16], o la relajación muscular [13]. También se han investigado terapias alternativas, como la aromaterapia [18], la musicoterapia [10], la reflexología [10] y la equinoterapia [25], aunque esta última es considerada pseudoterapia en España [2].

La equinoterapia consiste en la intervención de pacientes con condiciones como artritis o traumas psicológicos, mediante la interacción, sobre el caballo o junto al caballo, con caballos bajo supervisión profesional, con el fin de mejorar la salud. Se cree que la conexión emocional y el ejercicio físico pueden contribuir a un mejor sueño. A pesar de las dudas sobre su efectividad, la investigación en esta área sigue creciendo.

La Alianza Internacional por la Interacción Saludable con Equinos (AINISE) busca el reconocimiento de la equinoterapia como terapia en España. En 2020, lograron la beca Ann Kerr-Godal [4] para un ensayo piloto que evaluó su validez en pacientes con artritis, comparando 10 pacientes que recibieron equinoterapia con otros 10 que siguieron un programa de ejercicio físico durante 12 sesiones en seis semanas. Durante el ensayo se recogieron dos grupos de medidas por cada paciente:

1. *Medidas1*: Dolor, movilidad, equilibrio y esfuerzo, evaluados con un goniómetro y sensores corporales.
2. *Medidas2*: Monitorización continua (24 x 7 x 6 semanas) de la actividad diaria y calidad del sueño, mediante con un reloj inteligente de propósito general (no clínico), modelo TICWATCH E2, colocado en cada paciente del piloto.

El primer estudio realizado sobre los datos recolectados correspondiente a las Medidas2 [8], mostró errores de medición de la calidad del sueño debidos a fallos en la app TicSleep incluida en el modelo TICWATCH E2 [1]. Por lo tanto se decidió lanzar un ensayo en la unidad de sueño del Hospital Universitario Central de Asturias (HUCA) para calibrar los datos obtenidos con el modelo TICWATCH. Este ensayo consistiría en colocar el reloj TICWATCH, a otro grupo de pacientes mientras son monitorizados con un polisomnógrafo (PSG), para obtener datos que puedan ser etiquetados con fases de sueño correctas.

1.1 Ensayo clínico en unidad de sueño con wearables

Por lo tanto, el proyecto global en lo que respecta al sueño se basa en dos conjuntos de datos:

- *DatosHUCA*: Datos de 15 pacientes monitorizados en el HUCA con PSG (fase de sueño cada 30 segs), con aceleración y pulso del TICWATCH.
- *DatosPiloto* : Datos de 18 participantes del piloto de equinoterapia, con medidas de aceleración y pulso del TICWATCH.

1.2 Objetivos

En lo que respecta a análisis de DatosPiloto respecto al sueño, el objetivo final de este proyecto radica en comparar evolución de la calidad del sueño de los participantes intervenidos con equinoterapia, respecto a los participantes del grupo de control que fueron intervenidos con ejercicio físico funcional. A continuación se estimará la calidad del sueño entrenando varios modelos machine learning a partir de DatosSueño y se aplicarán estos modelos sobre DatosPiloto.

Pero el presente trabajo solo abarca como objetivo la parte de preparación de los modelos de machine learning para estimar las fases de sueño trabajando con DatosHUCA. En concreto se trata de estimar la fase de sueño en que está una persona en cada uno de los fragmentos de 30 segs del período que la persona ha estado en la "cama".

1.3 Estructura

El documento se organiza en las siguientes secciones, el estado del arte² la propuesta³, material y métodos⁴, resultados ⁵ y por último, las conclusiones ⁶.

2 Estado del arte

La identificación de fases de sueño ⁵ es esencial para el estudio de patrones y el diagnóstico de trastornos del sueño. Aunque la PSG es el estándar de oro, los wearables ofrecen una alternativa interesante, especialmente al combinarse con inteligencia artificial y aprendizaje automático [7,20,26,17].

Las principales alternativas para la clasificación de fases de sueño son la PSG y los dispositivos vestibles (*wearables*).

- **Polisomnografía (PSG)**: Considerada la técnica más precisa, emplea múltiples sensores como EEG, EOG, y EMG para identificar las fases de sueño, aunque su alto costo y necesidad de especialistas limita su uso [7,26].
- **Wearables**: Dispositivos como relojes inteligentes y pulseras de actividad registran parámetros como frecuencia cardíaca, aceleración, y temperatura de la piel, siendo accesibles y cómodos para el usuario. Estos dispositivos han mostrado resultados prometedores en la clasificación de fases del sueño, aunque requieren validación constante para alcanzar la precisión de la PSG [7,21].

Un breve resumen de los sensores empleados en este tipo de estudios se ilustra en la tabla 1

⁵ Las fases del sueño empleadas en este proyecto son las 5 más utilizadas en la literatura, siendo estas "Wake", "REM", "N1", "N2" y "N3", yendo de vigilia a sueño más profundo respectivamente

Sensor	Descripción
EEG	Mide actividad cerebral, útil para todas las fases del sueño.
EOG	Registra movimientos oculares, ayuda en la identificación del sueño REM y vigilia.
ECG	Mide actividad cardíaca, correlacionando con fases de sueño.
Acelerómetro	Detecta movimientos, útil para ciclos de sueño/vigilia.

Table 1: Sensores utilizados en estudios de sueño [14]

2.1 Aplicación de la Inteligencia Artificial y Machine Learning

Los algoritmos de aprendizaje automático y las redes neuronales han mostrado avances en la clasificación automática de fases de sueño usando datos de wearables. A pesar de ser menos precisos que la PSG, estos dispositivos, integrados con IA, permiten un monitoreo continuo y accesible de las fases de sueño [6,12,19].

Estudios recientes han implementado modelos avanzados como redes neuronales Long Short-Term Memory (LSTM) y DeepSleepNet, los cuales combinan CNN y LSTM bidireccionales para mejorar la precisión en la clasificación de las etapas de sueño [26,23,24]. Sin embargo, la precisión de estos modelos depende de la calidad de los datos de entrenamiento, lo que puede ser un desafío en entornos no clínicos.

La integración de señales multicanal y la combinación de datos de wearables y PSG han demostrado mejoras significativas en la precisión de clasificación, indicando que estos enfoques combinados podrían ser efectivos para aplicaciones en monitoreo de sueño en entornos clínicos y no clínicos [5,22].

3 Propuesta

La propuesta de clasificación de fases de sueño se desarrolla mediante la recopilación y procesamiento de datos de 15 pacientes, con el objetivo de crear un modelo de aprendizaje automático capaz de identificar las fases de sueño. Este flujo de trabajo incluye tres etapas principales: obtención de datos, preprocesamiento y entrenamiento de modelos de aprendizaje superficial (SL). Cada una de estas etapas se desglosa en varios pasos específicos, según se detalla a continuación (ver Figura 1).



Fig. 1: Esquema general de la propuesta

3.1 Obtención de datos

Para este estudio, se obtuvieron datos de 15 pacientes del Hospital Universitario Central de Asturias (HUCA) a lo largo de 10 meses. Cada paciente utilizó dispositivos portátiles (relojes TICWATCH y teléfonos BQ) que capturaron datos biométricos, como aceleración tridimensional (*3DACC*) y frecuencia cardíaca (*HR*). Además, se realizó un registro estándar de fases de sueño (PSG) para cada paciente, lo cual permite utilizar estos datos como referencia [8].

3.2 Preprocesamiento de datos

El preprocesamiento es fundamental para estructurar y limpiar los datos antes de construir los modelos. Esta fase incluye:

1. **Consolidación.**

Se combinan los datos de aceleración y frecuencia cardíaca con las fases de sueño obtenidas a partir del PSG. Dado que las frecuencias de los relojes (10 Hz) y de la PSG (1 muestra cada 30 segundos, conocida como *epoch*) son distintas, se sincronizan antes de proceder.

2. **Limpieza.**

Se aplican filtros para eliminar artefactos y datos irrelevantes. Primero, se aplica un filtro horario que excluye datos fuera del rango de sueño (20:00-12:00). Luego, un filtro de artefactos de movimiento elimina series temporales en las que las variaciones del pulso superan un umbral determinado, lo que indica actividad o movimiento inesperado [15].

3. **Escalado.**

Se ajusta la media de las series temporales y se aplican técnicas de escalado robusto [3] para minimizar el impacto de valores atípicos. Esta transformación asegura que todas las características de los datos tengan valores comparables.

4. **Extracción de *features*.**

Se crean matrices de características (*features*) que incluyen medidas calculadas cada 30 segundos, como aceleración (*3DACC*), *HR*, magnitud de aceleración ("MAGACC"), amplitud de movimiento ("AoM") y ángulo del brazo ("ArmAngle"). Estas medidas proporcionan información relevante sobre la actividad y posición del cuerpo durante el sueño.

3.3 Entrenamiento de modelos SL

Con las matrices de *features* creadas anteriormente se procede al entrenamiento de los modelos seleccionados. Para ello, el flujo se divide en las siguientes 4 partes:

1. **Particionado.**

Se dividen los datos en conjuntos de entrenamiento y validación. Se utilizan dos técnicas de validación: "Stratified KFold Cross-Validation" y "Leave-One-Participant-Out" (LOPO) [8], que aseguran que las particiones reflejen la variabilidad de las clases (fases de sueño).

2. Sobremuestreo (*Oversampling*).

Para equilibrar el conjunto de datos, se aplica la técnica SMOTE [9], que genera nuevas muestras sintéticas para las clases minoritarias. Esto evita que haya desbalance entre las distintas fases de sueño y mejora la capacidad de generalización del modelo.

3. Entrenamiento de modelos.

Se entrenan 11 modelos de clasificación, incluyendo métodos como *SVC*, *RidgeClassifier*, *RandomForest*, *AdaBoost* y *KNeighbors*, entre otros. Estos modelos se optimizan a través de una búsqueda de parámetros (GridSearch) para obtener el mejor rendimiento posible en cada partición de datos.

4. Análisis de resultados.

Finalmente, se realizan comparaciones mediante boxplots y matrices de confusión para evaluar el rendimiento de cada modelo. Estos análisis permiten identificar los modelos que mejor clasifican las fases de sueño y entender en qué clases se producen más errores.

4 Materiales y métodos

Esta sección describe los pasos y herramientas utilizadas para crear modelos que pueden identificar las fases del sueño usando datos obtenidos de dispositivos portátiles. A continuación, se detallan los datos utilizados, las técnicas para mejorar su calidad, los métodos para validar los modelos, y los ajustes para optimizar su precisión.

4.1 Conjunto de datos y procesamiento

Se trabajó con datos de 15 pacientes del Hospital Universitario Central de Asturias (HUCA), que incluyen mediciones de movimiento y ritmo cardíaco obtenidos con dispositivos portátiles, junto con las fases del sueño registradas por un sistema clínico estándar (PSG). Cada paciente tiene un registro de aproximadamente 9 horas de sueño. Los datos se clasificaron en tres niveles de calidad (*Alta*, *Media*, *Baja*), considerando factores como la duración del sueño y la estabilidad de las señales. La mayoría de los datos pertenecen a 9 pacientes de alta calidad.

Para preparar estos datos, se aplicaron varias técnicas de ajuste que buscan hacer los valores más uniformes y comparables. Esto ayuda a mejorar la precisión de los modelos, asegurando que los datos sean consistentes y minimizando los efectos de diferencias individuales entre pacientes.

4.2 Técnicas de validación

Para probar la efectividad de los modelos, se utilizaron dos métodos de validación. El primero, llamado "Stratified 10-Fold Cross-Validation", divide los datos en 10 grupos y prueba el modelo en cada grupo, asegurando que cada fase del sueño esté equilibradamente representada. El segundo método, "Leave One

Participant Out" (LOPO), usa los datos de un paciente a la vez como prueba, mientras que los datos de los demás se emplean para entrenar el modelo. Este método permite comprobar cómo funcionaría el modelo con datos de un paciente completamente nuevo. Ambos métodos de validación se pueden ver en la figura 2.

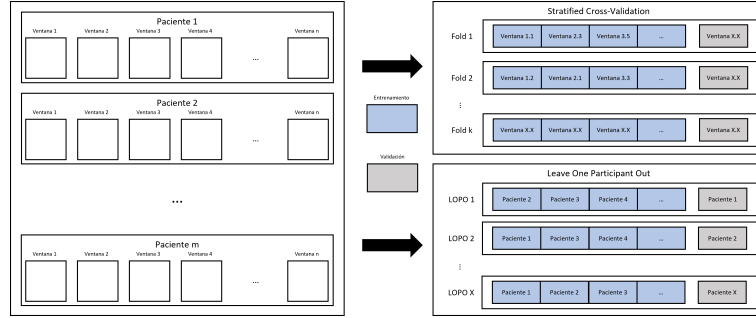


Fig. 2: Particionado en el entrenamiento

4.3 Métricas de evaluación

Para medir el desempeño de los modelos se utilizaron dos indicadores principales: *Precisión* (la proporción de predicciones correctas) y *Puntaje F1* (que mide el equilibrio entre predicciones correctas y errores). Estas métricas ayudan a entender cuán confiables son los modelos al clasificar las fases del sueño.

4.4 Optimización de modelos

Se aplicó un proceso llamado *GridSearch* para encontrar los mejores parámetros en cada modelo. Este proceso prueba diferentes configuraciones hasta encontrar la que ofrece mejores resultados. Los modelos evaluados incluyen algoritmos como *RandomForest*, *SVC* y *XGBoost*, entre otros. Los parámetros se basaron en ejemplos de plataformas como *Kaggle* y *Medium*, donde otros investigadores han probado configuraciones similares.

4.5 Configuración de experimentos

Para la configuración experimental de este estudio, se llevaron a cabo varias iteraciones ajustando técnicas de normalización, particionado de datos y etiquetas a clasificar para optimizar el rendimiento de los modelos y comprender mejor su comportamiento. En el **Experimento 3**, se aplicó un ajuste de media y escalado robusto global, seguido de normalización gaussiana en las matrices de

features y selección de características mediante ANOVA; además, se implementó un sobremuestreo para balancear las clases. En el **Experimento 6**, se utilizó el conjunto completo de pacientes y se ajustó la clasificación a un problema binario, distinguiendo únicamente entre "Wake" y "Sleep". Finalmente, en el **Experimento 7**, también con el conjunto completo de datos, se simplificaron las etiquetas a tres clases: "Wake", "REM" y "NREM" (agrupando "N1", "N2" y "N3"), manteniendo las mismas técnicas de preprocesamiento y balanceo de clases del experimento anterior.

5 Resultados

Se A continuación se presentan los resultados de los experimentos 3, 6 y 7, los cuales implementaron las diferentes configuraciones descritas en 4.5 para mejorar la clasificación de fases del sueño.

5.1 Experimento 3

En el Experimento 3 se mantuvieron cinco fases de sueño y se emplearon técnicas de normalización avanzadas y selección de *features*. Los modelos *RandomForest* (RF) y *eXtraTrees* (XT) obtuvieron los mejores resultados en la validación *Stratified 10-Fold Cross-Validation* (SCV), alcanzando un *F1 Score* promedio de 0.90. Sin embargo, en *Leave-One-Participant-Out* (LOPO), el rendimiento fue significativamente menor (alrededor de 0.57), lo que indica una dificultad para generalizar entre pacientes, dichos resultados se pueden ver en la Tabla 2.

Modelo	F1 Score (SCV)	F1 Score (LOPO)
RandomForest (RF)	0.90	0.57
eXtraTrees (XT)	0.89	0.56

Table 2: Resumen de resultados para Experimento 3

5.2 Experimento 6

En este experimento, se simplificó el problema a una clasificación binaria entre "Wake" y "Sleep", lo que mejoró notablemente los resultados. En SCV, los modelos RF y XT lograron un *F1 Score* promedio de aproximadamente 0.93, mientras que en LOPO también se observó una mejora, alcanzando un promedio de 0.71. Esta reducción de clases facilitó la captura de patrones generales, aunque la generalización en LOPO sigue siendo un desafío. Un resumen de estos resultados se ilustra en la Tabla 3

Modelo	F1 Score (SCV)	F1 Score (LOPO)
RandomForest (RF)	0.93	0.71
eXtraTrees (XT)	0.93	0.70

Table 3: Resumen de resultados para Experimento 6

5.3 Experimento 7

Finalmente, en el Experimento 7 se probó una clasificación de tres fases: "Wake", "NREM" y "REM". En SCV, RF y XT nuevamente lideraron con un *F1 Score* cercano a 0.93. En LOPO, los modelos alcanzaron un *F1 Score* promedio de 0.61, ligeramente inferior al experimento 6 pero mejor que en el Experimento 3. Esta configuración permitió una mejor distinción entre las fases principales, aunque sigue habiendo dificultades para generalizar entre pacientes. Estos resultados se muestran en la Tabla 4.

Modelo	F1 Score (SCV)	F1 Score (LOPO)
RandomForest (RF)	0.93	0.61
eXtraTrees (XT)	0.92	0.60

Table 4: Resumen de resultados para Experimento 7

En resumen, SCV mostró consistentemente un rendimiento superior en todos los experimentos debido a la inclusión de datos de entrenamiento y validación del mismo paciente, mientras que LOPO presentó dificultades para generalizar a pacientes nuevos. Los mejores modelos en todas las configuraciones fueron *RandomForest* y *eXtraTrees*.

6 Conclusiones

En este estudio se evaluaron diversas configuraciones y técnicas de preprocesamiento para la clasificación de fases del sueño a partir de datos de acelerometría (ACC) y fotopletismografía (PPG) en pacientes. Se emplearon dos enfoques de validación: *Stratified 10-Fold Cross-Validation* (SCV) y *Leave-One-Participant-Out* (LOPO), para comparar su efectividad en la generalización de los modelos entre diferentes pacientes.

Los resultados mostraron que SCV proporcionó consistentemente un *F1 Score* superior al de LOPO en todos los experimentos, debido a la inclusión de datos del mismo paciente en los conjuntos de entrenamiento y validación, lo cual facilita la captura de patrones específicos de cada individuo. Sin embargo, este enfoque no permite evaluar la capacidad de los modelos para generalizar a nuevos pacientes. Por otro lado, LOPO presentó un rendimiento considerablemente menor, destacando las dificultades de los modelos para adaptarse a las diferencias individuales

entre pacientes. Estas limitaciones de generalización son especialmente evidentes en la clasificación de múltiples fases del sueño, donde los datos presentan una mayor variabilidad.

Entre los modelos evaluados, *RandomForest* y *eXtraTrees* destacaron por su rendimiento en la mayoría de los experimentos, tanto en SCV como en LOPO. Estos modelos obtuvieron los mejores *F1 Scores* al reducir el problema a una clasificación binaria (fases "Wake" y "Sleep") o de tres clases ("Wake", "NREM", "REM"), lo que simplificó el desafío de la clasificación y permitió capturar patrones generales de sueño. Este enfoque resultó en una mejora significativa en el rendimiento de LOPO, alcanzando un *F1 Score* de hasta 0.71 en el Experimento 6 y 0.61 en el Experimento 7.

A pesar de los avances logrados, este estudio pone de manifiesto los desafíos asociados con la clasificación de fases del sueño en un contexto de datos heterogéneos y con variabilidad interindividual significativa. La reducción del número de clases mostró ser una estrategia efectiva para mejorar la precisión, pero limita la capacidad de los modelos para clasificar todas las fases de sueño de forma detallada. Esto sugiere que, para aplicaciones clínicas o de monitorización personalizada, es necesario explorar métodos adicionales que mejoren la capacidad de generalización, tales como técnicas de adaptación de dominio, modelos específicos por paciente o el uso de arquitecturas de redes neuronales profundas preentrenadas en datos similares.

En conclusión, la técnica de validación SCV es adecuada para la evaluación preliminar de modelos en estudios donde los datos provienen de un mismo grupo de pacientes. Sin embargo, para aplicaciones en contextos reales, donde es fundamental la generalización a nuevos individuos, la técnica LOPO sigue siendo un estándar más representativo. Futuros trabajos deberán centrarse en mejorar esta capacidad de generalización y en evaluar estos modelos con datos adicionales y pacientes externos al conjunto de entrenamiento.

References

1. Mobvoi (formerly TicWatch) - Aplicaciones en Google Play. URL <https://play.google.com/store/apps/details?id=com.mobvoi.companion.aw{%&}hl=es>
2. Observatorio contra las Pseudociencias, Pseudoterapias, Intrusismo y Sectas Sanitarias | CGCOM. URL <https://www.cgcom.es/observatorios/oppiss>
3. RobustScaler — scikit-learn 1.5.1 documentation. URL <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>
4. AINISE, O.: Clinical trial and data collection, multi- center, randomised trial, with experimental group, control group, and blind evaluation by third parties, in order to determine the effect of Therapeutic riding, on a group of people with Arthritis. Ann Kern-Godal's Horse-Assisted Therapy Grants (2020)
5. Almutairi, H., Hassan, G.M., Datta, A.: Machine-Learning-Based-Approaches for Sleep Stage Classification Utilising a Combination of Physiological Signals: A Systematic Review. Appl. Sci. 2023, Vol. 13, Page 13280 **13**(24), 13,280 (2023). DOI 10.3390/APP132413280. URL <https://www.mdpi.com/2076-3417/13/24/13280/htmlhttps://www.mdpi.com/2076-3417/13/24/13280>

6. Birrer, V., Elgendi, M., Lambercy, O., Menon, C.: Evaluating reliability in wearable devices for sleep staging. *npj Digit. Med.* 2024 71 **7**(1), 1–14 (2024). DOI 10.1038/s41746-024-01016-9. URL <https://www.nature.com/articles/s41746-024-01016-9>
7. Boe, A.J., McGee Koch, L.L., O'Brien, M.K., Shawen, N., Rogers, J.A., Lieber, R.L., Reid, K.J., Zee, P.C., Jayaraman, A.: Automating sleep stage classification using wireless, wearable sensors. *npj Digit. Med.* 2019 21 **2**(1), 1–9 (2019). DOI 10.1038/s41746-019-0210-1. URL <https://www.nature.com/articles/s41746-019-0210-1>
8. de la Cal, E.A., Fañez, M., Apolo, M.D., García-Gómez, A., González, V.M.: Smartwatch Sleep-Tracking Services Precision Evaluation Using Supervised Domain Adaptation. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13469 LNAI, pp. 14–26 (2022). DOI 10.1007/978-3-031-15471-3_2
9. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002). DOI 10.1613/JAIR.953. URL <https://www.jair.org/index.php/jair/article/view/10302>
10. Cooke, M., Ritmala-Castrén, M., Dwan, T., Mitchell, M.: Effectiveness of complementary and alternative medicine interventions for sleep quality in adult intensive care patients: A systematic review. *Int. J. Nurs. Stud.* **107**, 103,582 (2020). DOI 10.1016/j.ijnurstu.2020.103582. URL <https://doi.org/10.1016/j.ijnurstu.2020.103582>
11. Fatemeh, G., Sajjad, M., Niloufar, R., Neda, S., Leila, S., Khadijeh, M.: Effect of melatonin supplementation on sleep quality: a systematic review and meta-analysis of randomized controlled trials. *J. Neurol.* **269**(1), 205–216 (2022). DOI 10.1007/S00415-020-10381-W/FIGURES/4. URL <https://link.springer.com/article/10.1007/s00415-020-10381-w>
12. Grandner, M.A., Lujan, M.R., Ghani, S.B.: Sleep-tracking technology in scientific research: Looking to the future (2021). DOI 10.1093/sleep/zsab071
13. Harorani, M., Davodabady, F., Masmouei, B., Barati, N.: The effect of progressive muscle relaxation on anxiety and sleep quality in burn patients: A randomized clinical trial. *Burns* **46**(5), 1107–1113 (2020)
14. Imtiaz, S.A.: A Systematic Review of Sensing Technologies for Wearable Sleep Staging. *Sensors* 2021, Vol. 21, Page 1562 **21**(5), 1562 (2021). DOI 10.3390/S21051562. URL <https://www.mdpi.com/1424-8220/21/5/1562/htm><https://www.mdpi.com/1424-8220/21/5/1562>
15. Ismail, S., Akram, U., Siddiqi, I.: Heart rate tracking in photoplethysmography signals affected by motion artifacts: a review. *EURASIP J. Adv. Signal Process.* **2021**(1), 1–27 (2021). DOI 10.1186/S13634-020-00714-2/FIGURES/7. URL <https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-020-00714-2>
16. Kirmizigil, B., Demiralp, C.: Effectiveness of functional exercises on pain and sleep quality in patients with primary dysmenorrhea: a randomized clinical trial. *Arch. Gynecol. Obstet.* **302**(1), 153–163 (2020)
17. Långkvist, M., Karlsson, L., Loutfi, A.: Sleep Stage Classification Using Unsupervised Feature Learning. *Adv. Artif. Neural Syst.* **2012**(1), 107,046 (2012). DOI 10.1155/2012/107046. URL <https://onlinelibrary.wiley.com/doi/full/10.1155/2012/107046><https://onlinelibrary.wiley.com/doi/abs/10.1155/2012/107046><https://onlinelibrary.wiley.com/doi/10.1155/2012/107046>
18. Mohammadi, F., Moradi, M., Niazi, A., Jamali, J.: The Impact of Aromatherapy with Citrus Aurantium Essential Oil on Sleep Quality in Pregnant Women with Sleep Disorders: A Randomized Controlled Clinical Trial. *Int. J. Community Based Nurs. Midwifery* **10**(3), 160–171 (2022). DOI 10.30476/IJCBNM.2022.92696.1900

19. Nguyen, Q.N., Le, T., Huynh, Q.B., Setty, A., Vo, T.V., Le, T.Q.: Validation Framework for Sleep Stage Scoring in Wearable Sleep Trackers and Monitors with Polysomnography Ground Truth. *Clocks Sleep* 2021, Vol. 3, Pages 274-288 **3**(2), 274–288 (2021). DOI 10.3390/CLOCKSSLEEP3020017. URL <https://www.mdpi.com/2624-5175/3/2/17/html><https://www.mdpi.com/2624-5175/3/2/17>
20. Radha, M., Fonseca, P., Moreau, A., Ross, M., Cerny, A., Anderer, P., Long, X., Aarts, R.M.: ARTICLE OPEN A deep transfer learning approach for wearable sleep stage classification with photoplethysmography DOI 10.1038/s41746-021-00510-8. URL <https://doi.org/10.1038/s41746-021-00510-8>
21. Roberts, D.M., Schade, M.M., Mathew, G.M., Gartenberg, D., Buxton, O.M.: Detecting sleep using heart rate and motion data from multisensor consumer-grade wearables, relative to wrist actigraphy and polysomnography. *Sleep* **43**(7), 1–19 (2020). DOI 10.1093/SLEEP/ZSAA045. URL <https://pubmed.ncbi.nlm.nih.gov/355403/><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7355403/?report=abstract><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7355403/>
22. Smith, A., Anand, H., Milosavljevic, S., Rentschler, K.M., Pocivavsek, A., Valafar, H.: Application of Machine Learning to Sleep Stage Classification. *Proc. - 2021 Int. Conf. Comput. Sci. Comput. Intell. CSCI* 2021 pp. 349–354 (2021). DOI 10.1109/CSCI54926.2021.00130
23. Supratak, A., Dong, H., Wu, C., Guo, Y.: DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* **25**(11), 1998–2008 (2017). DOI 10.1109/TNSRE.2017.2721116
24. Supratak, A., Guo, Y.: TinySleepNet: An Efficient Deep Learning Model for Sleep Stage Scoring based on Raw Single-Channel EEG. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2020-July**, 641–644 (2020). DOI 10.1109/EMBC44109.2020.9176741
25. White-Lewis, S., Johnson, R., Ye, S., Russell, C.: An equine-assisted therapy intervention to improve pain, range of motion, and quality of life in adults and older adults with arthritis: A randomized controlled trial. *Appl. Nurs. Res.* **49**, 5–12 (2019)
26. Zhang, Y., Yang, Z., Lan, K., Liu, X., Zhang, Z., Li, P., Cao, D., Zheng, J., Pan, J.: Sleep Stage Classification Using Bidirectional LSTM in Wearable Multi-sensor Systems. *INFOCOM 2019 - IEEE Conf. Comput. Commun. Work. INFOCOM WKSHPs* 2019 pp. 443–448 (2019). DOI 10.1109/INFOCOMW.2019.8845115