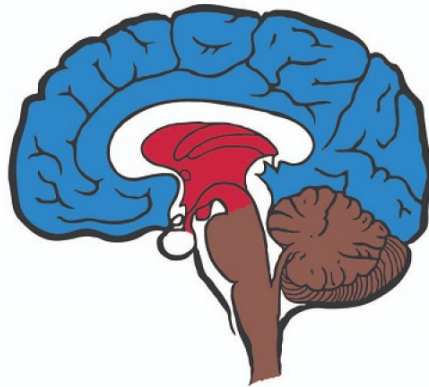


Introducción a la Ciencia de Datos aplicada a escenarios Médico-Biológicos

Análisis de los trastornos del sueño



Cristal Araceli Cardoso Estrada

20/Abril/2025

Índice

1. Introducción	2
1.1. El sueño como actividad biológica	2
1.2. Trastornos del sueño	2
1.3. Factores que pueden desencadenar trastornos del sueño	3
1.4. Trastornos del sueño y sus consecuencias en la salud humana	3
1.5. Abordaje de los trastornos del sueño	4
2. Posibles aplicaciones de Machine Learning para estudios de trastornos del sueño	4
3. Metodología	5
3.1. Análisis Exploratorio de los datos	5
3.2. Etapa de preprocesamiento	6
3.3. Implementación de modelos de entrenamiento	6
4. Resultados y Discusión	6
5. Conclusiones	9
6. Posibles mejoras para el análisis	10

1. Introducción

1.1. El sueño como actividad biológica

Todos los animales, incluidos los unicelulares, tienen periodos de actividad y reposo, existiendo, en este último, una inmovilidad relativa. El sueño es periódico y en general, espontáneo. Además, en el humano, se acompaña de una pérdida de la consciencia vigil.

El sueño tiene distintos grados de profundidad y en cada una de las etapas se presentan modificaciones fisiológicas concretas. Para el estudio de estos cambios se acude a una de las variables que se denominan *indicadores del sueño*: el electroencefalograma (EEG), los movimientos oculares y el tono muscular.[6] Así, la polisomnografía, es el estudio que permite determinar estos indicadores, permitiendo obtener una serie de datos que determinan la calidad del sueño de la persona y un diagnóstico de alguna anormalidad o trastorno.

La relación entre vigilia y sueño debe ser armónica, ya que el desequilibrio entre ambos está considerado como una crisis, situación que además origina trastornos que aminoran el estado de bienestar de una persona y en consecuencia generan problemas de salud pública. La problemática es socialmente relevante debido a que hoy en día las personas están obligadas a cumplir jornadas laborales amplias, lo cual genera que duerman menos horas, deterioro de la calidad de vida, salud y un vacío normativo que promueve la mejora de estas condiciones. Dicho lo anterior, se encuentra documentada la relación que guardan los trastornos del sueño con la obesidad, hipertensión, las enfermedades cardiovasculares y los trastornos psiquiátricos (depresión o ansiedad).[4]

1.2. Trastornos del sueño

Los trastornos del sueño son una serie de alteraciones relacionadas con el proceso de dormir. Existiendo tanto en las etapas de inicio, de mantenimiento, como durante el ciclo sueño-vigilia. Los trastornos del sueño constituyen un grupo numeroso y heterogéneo de procesos, que la mayoría de las veces suelen ser mal diagnosticados y tratados, sus consecuencias en la salud y mortalidad representan una carga de salud pública no reconocida. [5].

El estudio de los trastornos del sueño es un campo de la medicina relativamente nuevo y no fue hasta la década de 1990 que comenzó a implementarse una clasificación de estos trastornos. La primera Clasificación Internacional de los Trastornos del Sueño (ICSD-1) fue elaborada en 1997 y publicada por la Academia Americana de Medicina del Sueño (AASM) y en ella podemos encontrar clasificaciones y descripciones de los trastornos del sueño, así como datos de interés de cada uno.

La ICSD-2, publicada en 2005, fue la consecuente clasificación de trastornos del sueño y en ella se obtiene una división de los trastornos en siete grupos en función de su síntoma principal: insomnio, hipersomina, trastorno respiratorio, parasomnia, alteración del ciclo circadiano, movimientos anormales y trastornos aislados. En la ICSD-3 de 2014 (Figura 1) se toma como base la clasificación de 2005, resumiendo en seis categorías los trastornos del sueño:

Grupos de trastornos del sueño en la ICSD-3		
Insomnios	Trastornos respiratorios	Hipersomnias centrales
Insomnio crónico	Apneas obstructivas del sueño	Narcolepsia tipo 1
Insomnio de corta evolución	Del adulto	Narcolepsia tipo 2
Otros insomnios	Del niño	Hipersomnia idiopática
Variantes	Apneas centrales del sueño	Síndrome de Kleine-Levin
Excesivo tiempo en cama	Trastornos de hipoventilación	Síndrome de sueño insuficiente
Dormidor corto	Hipoxemia del sueño	Hipersomnias secundarias
	Síntomas aislados o variantes	Variantes: dormidor largo
	Ronquido	
	Catatrenia	
Trastornos del ritmo sueño-vigilia	Parasomnias	Movimientos anormales
Fase de sueño retrasada	No relacionadas con REM	Piernas inquietas
Fase de sueño avanzada	Relacionadas con REM	Movimientos periódicos de los miembros
Fase de sueño irregular	Otras parasomnias	Calambres del sueño
Ritmo no de 24 horas	Síntomas aislados o variantes	Bruxismo
Trabajo en turnos		Movimientos rítmicos relacionados con el sueño
Jet-lag		Miclonías benignas del lactante
Otros trastornos del ritmo		Miclonías propioespinales
		Otros, secundarios y no especificados

ICSD: International Classification of Sleep Disorders.

Figura 1

1.3. Factores que pueden desencadenar trastornos del sueño

Los trastornos del sueño pueden tener múltiples y diversas causas, que pueden ir desde la edad de una persona, el sexo y también estar relacionados a sus actividades y rutinas diarias.

Algunos de estos factores son:

- **Edad:** El patrón de sueño cambia dependiendo de la edad, ya que los infantes y niños duermen en promedio entre 16 y 20 horas al día, los adultos entre siete y ocho horas, y las personas mayores de 60 años de seis a siete horas diarias en promedio y su sueño —al igual que el de los niños pequeños— está fragmentado; por otro lado, el sueño de ondas lentas delta (el cual es, por cierto, profundo y reparador) también disminuye con la edad.
- **Genéticos:** Algunos estudios han demostrado la existencia de cierta predisposición relacionada con los rasgos del temperamento.
- **Estados funcionales:** Los trastornos del sueño en los adultos, excepto en los adultos mayores, suelen estar relacionados con: estados de nerviosismo, ansiedad, estrés, bajo estado emocional, otras alteraciones emocionales.
- **Internos** Estado de salud: apneas, fiebre, depresión, etc.; Estados fisiológicos: adaptación a los cambios, menopausia, etc.; Cambios en los ritmos circadianos por uso horario (jet-lag), turno laboral nocturno o rotatorio día/noche, por determinadas alteraciones del Sistema Nervioso Central (SNC). Además de los niveles hormonales (cortisol, estrógeno y testosterona), sedentarismo y obesidad.
- **Farmacológicos:** El insomnio puede ser causado por ciertos fármacos como: anfetaminas, corticosteroides, estimulantes adrenérgicos, betabloqueantes, antidepresivos, diuréticos, esteroides anabólicos, teofilina, aminofilina, benzodiacepinas, anticonvulsivantes, levodopa en altas dosis, vitamina B6, ciprofloxacino, levofloxacino, antipsicóticos y antiepilépticos.
- **Externos:** Ambientales, hábitos de vida: consumo de tabaco, alcohol, cafeína, sustancias estimulantes y bebidas energéticas, mala higiene del sueño: falta de horario regular y rutina; uso de dispositivos electrónicos o exposición a luz brillante antes de ir a dormir, ejercicio intenso a última hora de la tarde, etc.)[2]

1.4. Trastornos del sueño y sus consencuencias en la salud humana

Actualmente existe una gran cantidad de evidencia científica en relación a que los trastornos del sueño crónicos son un factor de riesgo significativo para una gran diversidad de enfermedades crónicas, que van desde las enferme-

dades cardiovasculares, hasta los trastornos psiquiátricos, emocionales y cognitivos. Es muy interesante resaltar que dicha evidencia demuestra que la asociación como factor de riesgo no se limita a un tipo de alteración del sueño en particular, sino que se ha asociado a una diversidad de trastornos del sueño o síntomas de los mismos.

Entre esta relevante evidencia se encuentra el estudio de los trastornos del sueño como factor de riesgo para distintas enfermedades cardiovasculares, así como un incremento en la tasa de mortalidad en aquellos pacientes que experimentan enfermedades cardiovasculares y trastornos del sueño a la vez. Por otra parte también existe abundante evidencia de que los trastornos del sueño producen un incremento en el riesgo de otras enfermedades crónicas (sobre todo metabólicas) que adicionalmente aumentan el riesgo cardiovascular y la mortalidad, entre ellas están la obesidad, la diabetes y el síndrome metabólico. En la esfera neurológica, los trastornos del sueño también son un factor de riesgo independiente para el desarrollo de enfermedad vascular cerebral (infarto cerebral), y además son un factor de pronóstico negativo para la recuperación funcional del mismo. Finalmente, en el área psicoemocional, el insomnio se ha relacionado con el desarrollo de estados de depresión y ansiedad. [1]

1.5. Abordaje de los trastornos del sueño

El estudio y diagnóstico de los trastornos del sueño, requiere de la recabación de distintos estudios para poder llevar a cabo una identificación precisa de la alteración que esté experimentando un paciente. La mayor parte de estos métodos de estudio se centran en registrar la actividad que ocurre durante el periodo de sueño de la persona y también por los síntomas y consecuencias que este experimenta. Como se mencionó al principio, estos estudios contemplan principalmente la polisomnografía (PSG), también conocida como el “estándar de oro” en el área del diagnóstico de los trastornos del sueño. Se realiza en un laboratorio especializado donde se monitorizan simultáneamente múltiples variables biológicas del individuo durante una noche “típica” de sueño, como la estadificación de las etapas del sueño mediante monitoreo continuo del EEG, electrooculograma (EOG), electromiograma submentoniano (EMG), flujo de aire nasal u oral, el esfuerzo respiratorio, oximetría, electrocardiograma (ECG), electromiograma (EMG) del musculo tibial anterior y monitorización de la posición del paciente. [1]

La polisomnografía y todo lo que este estudio abarca, representa un valioso recurso para el diagnóstico de trastornos del sueño, sin embargo, requiere de la utilización de equipos e instalaciones específicas que pueden no estar al alcance de todos. Para estos casos se cuenta con otras alternativas como lo son los cuestionarios y escalas y también otros tipos de estudios como lo son la actigrafía y el estudio de latencias múltiples del sueño.

2. Posibles aplicaciones de Machine Learning para estudios de trastornos del sueño

Estudio sobre las profesiones y la calidad del sueño

Puede realizarse un estudio para observar cómo es la calidad del sueño e personas con distintas profesiones y a través de los resultados se puede trabajar de la mano con las empresas y empleadores para emitir recomendaciones o recursos para que los trabajadores puedan llevar un mejor descanso y disminuir la incidencia de trastornos del sueño.

Estudio sobre el impacto de la actividad física en la mejora de la calidad del sueño

Evaluar cómo impacta la actividad física y el número de pasos diarios en la calidad del sueño de una persona y a partir de ello poder realizar recomendaciones de actividades para personas de acuerdo a su edad y género.

Identificación de los parámetros fisiológicos más comunes para cada tipo de trastorno del sueño

Clasificar qué parámetros influyen mayormente para cada tipo de trastorno del sueño y poder identificar cuál es más probable que padezca una persona de acuerdo a su presión arterial, tasa de latidos del corazón e índice de masa corporal.

3. Metodología

Se trabajó con el dataset *Sleep Disorder* obtenido de Kaggle, el cual se analizó e interpretó en un notebook de Google Colab.

En el dataset de Kaggle se destacan algunas secciones relevantes, se inició importando las librerías pertinentes y leyendo el archivo .csv en el cual al desplegar las primeras filas mostró las siguientes columnas:

- Person ID
- Gender
- Age
- Occupation
- Sleep Duration
- Quality of sleep
- Physical activity level
- Stress level
- BMI Category
- Blood pressure
- Heart rate
- Daily steps
- Sleep Disorder

Posterior a esto se realizaron distintas fases para el análisis de los datos.

3.1. Análisis Exploratorio de los datos

Se realizó un Análisis Exploratorio de los Datos o EDA. En esta sección se muestran diferentes funciones que nos permiten conocer la cantidad de datos que contiene nuestro dataframe, así como los tipos de datos que se manejan. También, se indentifican los valores nulos y se reemplazan algunos valores que se consideren por otros términos que sean más claros o que nos permitan interpretarlo de forma más eficiente. Asimismo, aquí se obtuvieron los valores estadísticos relevantes de los datos numéricos como lo son la media, mediana, desviación estándar, etc. Esto mediante la función `.describe()`.

Después, se desplegaron algunas gráficas que muestran información relevante respecto a las columnas del dataframe.

Posteriormente se utilizó *label encoding* para asignar valores numéricos a las tres categorías de la columna *Sleep Disorder* (No Disorder, Insomina, Sleep Apnea), donde:

- 0 = Insomnia
- 1 = No Disorder
- 2 = Sleep Apnea

3.2. Etapa de preprocesamiento

En esta etapa se definieron dos tipos de factores: numéricos y categóricos, y se realizó un proceso de escalamiento para los valores numéricos. El proceso de preprocesamiento es crucial para que a la hora de aplicar los modelos de entrenamiento, estos puedan procesar de mejor manera los datos y sean más eficientes y con ello los resultados a obtener sean más precisos.

Cabe destacar que en el código base también se implementó la técnica SMOTE para los datos desbalanceados, esto es importante para que nuestros modelos de entrenamiento no estén sesgados a la hora de desempeñarse.

3.3. Implementación de modelos de entrenamiento

Se implementaron tres modelos de entrenamiento diferentes, el primero de ellos fue el modelo de Regresión Logística; después un modelo XGBoost Classifier, modelo Gradient Boosting Machine y finalmente el modelo K-Neighbors Classifier. Para cada uno de estos métodos se obtuvieron las métricas de exactitud, precisión y recuperación, así como el valor F1 que nos permite determinar qué tan bien se desempeñó cada uno de los modelos.

4. Resultados y Discusión

Modelo de Regresión Logística

Al implementar el modelo de regresión logística y obtener los valores de las métricas generales¹ (que nos permiten evaluar qué tan bien se desempeñó nuestro modelo) se obtuvieron los siguientes datos:

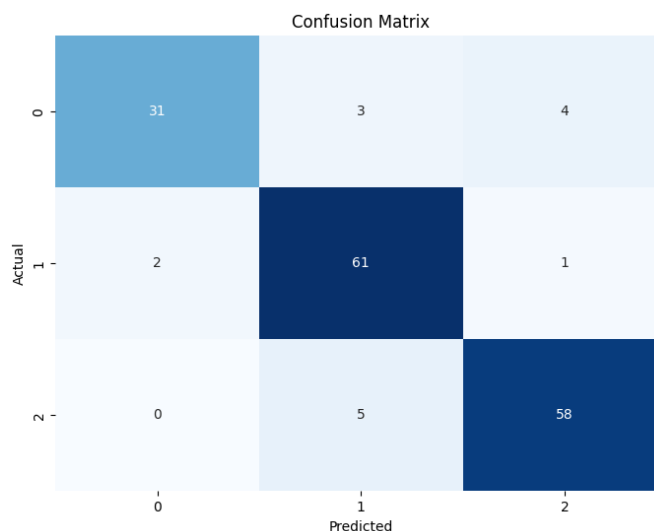
Accuracy: 0.9090909090909091

Precision: 0.910767756617559

Recall: 0.9090909090909091

F1-score: 0.9084234273263183

Y la matriz de confusión:



Que en su diagonal principal nos permite determinar que este modelo obtuvo un mejor desempeño con la categoría 1 (No Disorder) y en sí, nos muestra una puntuación de exactitud del 90 % al trabajar con los diferentes parámetros para determinar si hay o no un trastorno del sueño.

¹Accuracy = exactitud, Precision = precisión, Recall = Sensibilidad, F1-score = Puntuación F1

Modelo XGBoost Classifier

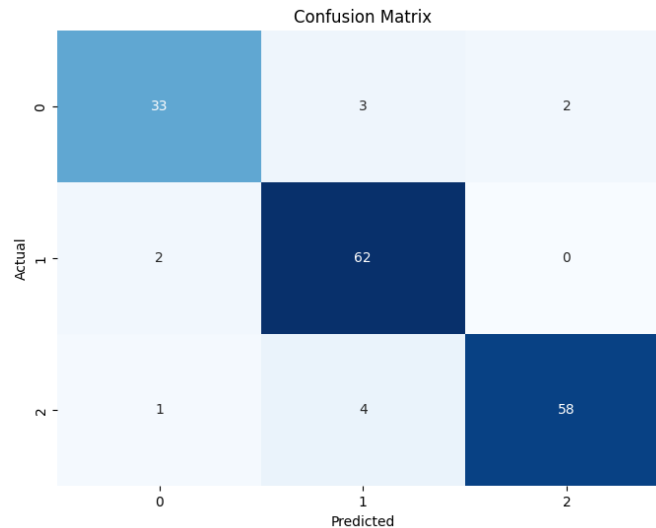
Para el modelo XGBoost Classifier se obtuvieron las siguientes métricas y matriz de confusión:

Accuracy: 0.9272727272727272

Precision: 0.9287307861220904

Recall: 0.9272727272727272

F1-score: 0.9271254483064495



Para el caso de este modelo, podemos observar que el valor de exactitud aumentó al 92 % en comparación con el modelo de Regresión Logística, y en general el valor de sus métricas está muy bien balanceado. Además, en la matriz de confusión se observa el aumento en las predicciones realizadas correctamente.

Modelo Gradient Boosting Machine

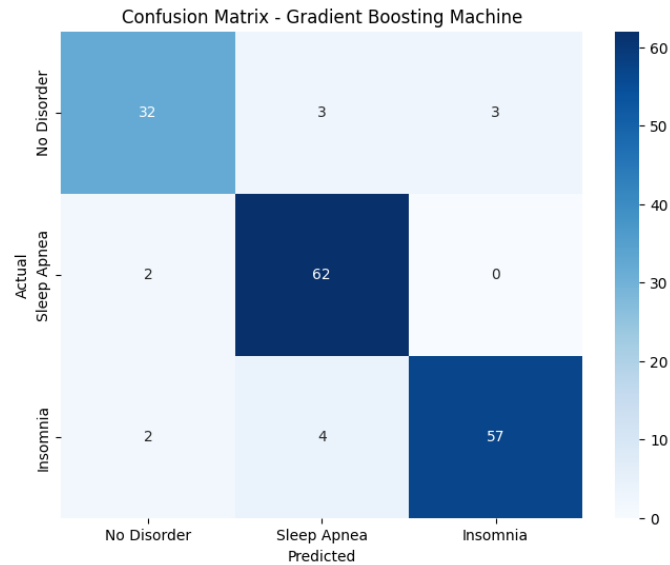
A partir de la implementación de esre modelo se obtuvieron las siguientes métricas y matriz de confusión:

Accuracy: 0.9152

Precision: 0.9160

Recall: 0.9152

F1 Score: 0.9147



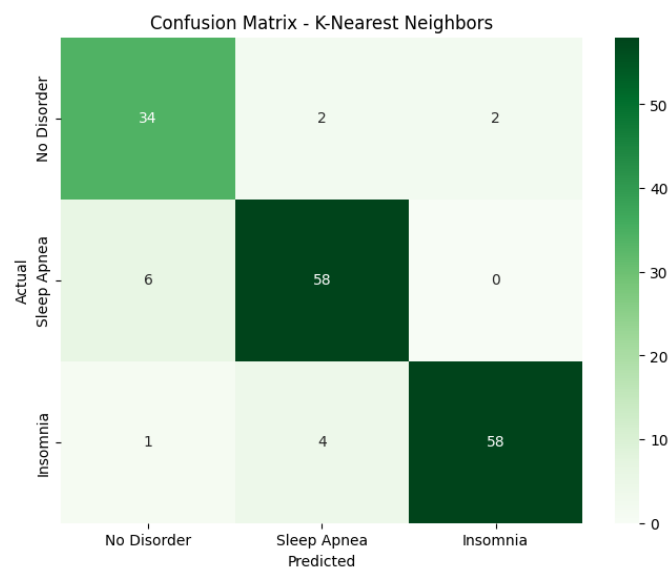
A partir de las cuales se puede observar que este modelo obtuvo un menor desempeño en comparación con el modelo XGB Classifier, pero a la vez tiene mejor puntuación en cuanto a exactitud, precisión y puntuación F1 en relación al modelo de Regresión Logística. De igual forma, en la matriz de confusión observamos el mejor desempeño para predicciones de Sleep Apnea.

Modelo K-Neighbors Classifier

Al probar con el modelo de K-Neighbors Classifier, se muestran las siguientes métricas de evaluación para este modelo:

Accuracy: 0.9212
Precision: 0.9231
Recall: 0.9212
F1 Score: 0.9213

Y su respectiva matriz de confusión

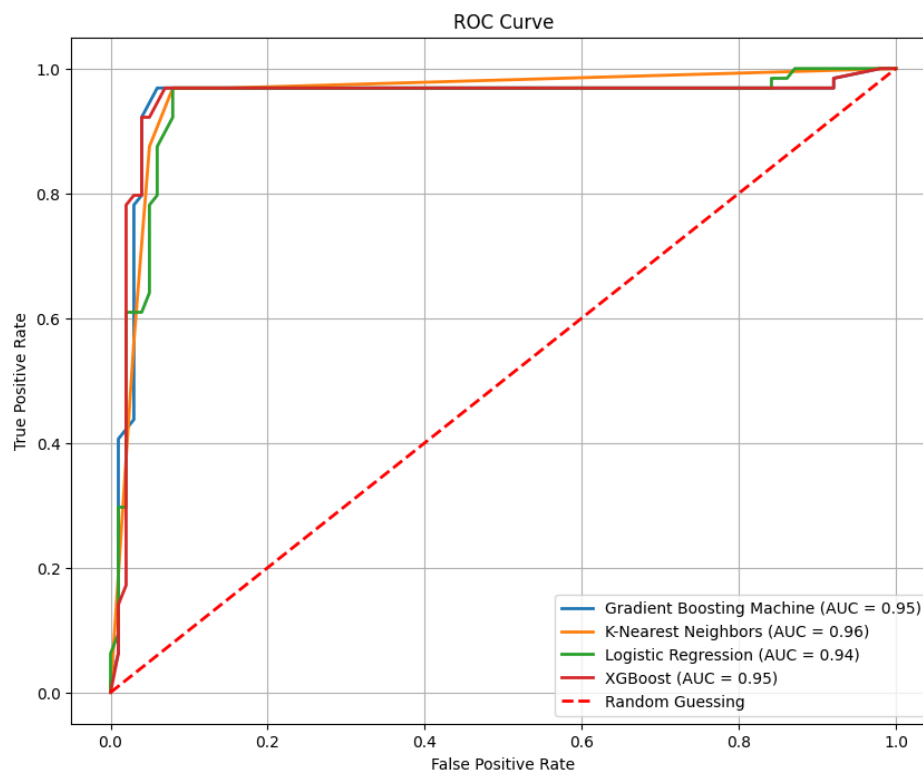


Podemos observar que los valores de las métricas de precisión para el modelo K-Neighbors son muy cercanas a las obtenidas con el modelo XGB Classifier, sin embargo este último es el que ha obtenido el mejor desempeño, aunque por muy poca diferencia. En relación a los modelos de Regresión Logística y Gradient Boosting Machine, vemos que funciona mejor.

Comparación de modelos

Una vez que se implementaron los distintos modelos de entrenamiento, se procedió a realizar una comparación del desempeño de cada uno de los modelos, esto mediante la generación de curvas ROC. La curva ROC nos dice qué tan bueno puede distinguir el modelo entre dos cosas o características. El AUC es el área bajo la curva ROC. Este puntaje nos da una buena idea de qué tan bien funciona el modelo. Según se desplaza la curva hacia la esquina superior izquierda del gráfico, la calidad del modelo va aumentando.[3]

A continuación se muestra la gráfica obtenida a partir de la comparación de los modelos utilizados:



Podemos ver de forma gráfica cuál fue el rendimiento de cada uno de los modelos y el valor del AUC, que para los modelos utilizados es muy cercano a 1, lo cual indica un muy buen desempeño.

5. Conclusiones

El proyecto de análisis del sueño muestra cómo se relacionan distintas variables del estado de salud y actividad de una persona, así como su ocupación en cuanto a si una persona puede padecer trastorno de insomnio, apnea del sueño o bien, si puede estar exluido de padecer cualquiera de estas anomalías del sueño.

Referente al tema de los modelos de entrenamiento utilizados, algunas características pueden tener más peso en cuanto a los resultados de predicción de cada modelo. Basándonos en la literatura, la edad, el género, edad y actividad física, que en este caso puede medirse por la cantidad de pasos diarios, pueden ser factores clave para el rendimiento y precisión de estos modelos.

Sería interesante probar cómo se desempeñan los modelos de entrenamiento añadiendo características adicionales para cada persona, así como diferentes tipos de trastornos del sueño más allá del insomnio y la apnea del sueño.

6. Posibles mejoras para el análisis

Para este estudio en particular, según la bibliografía consultada, existen diversos factores que pueden afectar o propiciar el desarrollo de trastornos del sueño, algunos de estos factores y variables fueron expuestas en el dataset con el que se trabajó, sin embargo sería interesante recabar datos adicionales que permitan expandir el modelado y trabajar con valores adicionales que permitan realizar una clasificación más precisa de los trastornos del sueño, o bien, identificar factores adicionales que pueden intervenir para este tipo de anomalías.

Otros valores que pueden adicionarse para el estudio son:

- Hora a la que la persona suele dormir
- Tipo de actividad física que realiza
- Si padece depresión o ansiedad
- Horario de trabajo (matutino, vespertino, nocturno)
- Consumo de fármacos y/o sustancias

Al obtener estos datos, que pueden influir en los resultados positivos o negativos para trastornos del sueño, se pueden implementar modelos que, por ejemplo nos puedan indicar de qué manera influye el horario para dormir y los turnos de trabajo en los trastornos del sueño o en la ausencia de ellos.

También se podría determinar cómo influyen los diferentes tipos de actividades físicas en la calidad del sueño o realizar una comparación entre las personas que realizan actividades físicas y los parámetros de presión arterial e índice de masa corporal.

Referencias

- [1] Paul Carrillo-Mora et al. «Trastornos del sueño: ¿qué son y cuáles son sus consecuencias?» En: *Revista de la Facultad de Medicina, Universidad Nacional Autónoma de México* 61 (feb. de 2018), págs. 6-20.
- [2] Miguel Collado et al. «Epidemiología de los trastornos del sueño en población mexicana: seis años de experiencia en un centro de tercer nivel». En: *Anales médicos (México, D.F.)* (2016).
- [3] Ligdi Gonzalez. *Curvas Roc y área bajo la curva (AUC)*. Mayo de 2023. URL: <https://aprendeia.com/curvas-roc-y-area-bajo-la-curva-auc-machine-learning/>.
- [4] G. Miranda-Nava. «Trastornos del sueño». En: *Medicina e Investigación Universidad Autónoma del Estado de México* 6.1 (2018), págs. 68-78. ISSN: 2594-0600. URL: <https://medicinainvestigacion.uaemex.mx/article/view/18989>.
- [5] Instituto Mexicano del Seguro Social. *Diagóstico y Tratamiento de los Trastornos del Sueño*. México, 2010.
- [6] J.L. Velayos. *Medicina del Sueño: Enfoque multidisciplinario*. Editorial Médica Panamericana S.A., 2009. ISBN: 9788479039936. URL: <https://books.google.com.mx/books?id=Ov7NoAEACAAJ>.