

Evaluación del método DBSCAN para identificar patrones de estrés en estudiantes usando datos no etiquetados

Evaluation of the DBSCAN method to identify student stress patterns using unlabeled data

Reyes-Acosta, J.^{1*} , Rodríguez-Arce, J.²  (*jreyesa838@alumno.uaemex.mx)

Recibido: 2024-11-12 Aceptado: 2024-11-29

RESUMEN

Un desafío clave en la clasificación del estrés es la obtención de datos etiquetados, que requiere la intervención de expertos, como psicólogos, para garantizar precisión, dada la variabilidad individual en la respuesta al estrés. Este estudio evalúa el método DBSCAN (Density-Based Spatial Clustering of Applications with Noise) para identificar respuestas fisiológicas al estrés en datos no etiquetados, destacando su capacidad para detectar patrones en datos ruidosos o complejos. Se analizaron señales de fotopletismografía (PPG) y actividad electrodérmica (EDA) registradas en estudiantes universitarios durante 51 pruebas de inducción de estrés en entornos académicos de bajo, medio y alto estrés. Las señales fueron segmentadas y procesadas para reducir artefactos y asegurar un número homogéneo de muestras por sujeto y escenario. Se extrajeron características de las señales, normalizadas

con z-score, y se aplicó Análisis de Componentes Principales (PCA) para reducir la dimensionalidad y facilitar la visualización, generando conjuntos separados para PPG y EDA. Finalmente, se aplicó DBSCAN al conjunto unificado de datos no etiquetados. Los resultados muestran que DBSCAN no logró formar clústeres claramente definidos con los parámetros seleccionados, evidenciando una limitación del método en este contexto, a pesar de su capacidad para manejar datos complejos. Los hallazgos sugieren la necesidad de explorar métodos más avanzados de procesamiento y clustering para datos no etiquetados en estudios futuros.

Palabras clave: DBSCAN, clustering, aprendizaje no supervisado, identificación de estrés, análisis de señales.

¹Facultad de Ingeniería, Universidad Autónoma del Estado de México, Paseo Universidad S/N, Toluca, Estado de México, México. CP. 50130.

²Facultad de Ingeniería, Universidad Autónoma del Estado de México, Paseo Universidad S/N, Toluca, Estado de México, México. CP. 50130.



ABSTRACT

A critical challenge in stress classification is obtaining labeled data, which requires the intervention of experts, such as psychologists, to ensure accuracy due to the individual variability of stress responses. This study evaluates the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) method to identify physiological stress responses in unlabeled datasets. This method is characterized by its ability to detect patterns in noisy or complex data. Photoplethysmography (PPG) and electrodermal activity (EDA) signals were analyzed from university students during 51 stress induction tests in educational environments representing low, medium, and high-stress levels. The signals were segmented and pre-processed to reduce artifacts, ensuring consistent samples per subject and scenario. Signal features

were extracted and standardized using z-score normalization. Principal Component Analysis (PCA) was used for dimensionality reduction to facilitate visualization, resulting in separate datasets for PPG and EDA. Finally, DBSCAN was applied to the unified set of unlabeled features. The results suggest that DBSCAN failed to form well-defined clusters with the selected parameters, highlighting a limitation of the method in this context despite its capability to handle complex data structures. These findings suggest the need to explore more advanced processing and clustering methods for unlabeled data in future research.

Keywords: DBSCAN, clustering, unsupervised learning, stress identification, signal analysis.

INTRODUCCIÓN

La salud mental, particularmente el estrés, ha ganado relevancia en el ámbito tecnológico. Su impacto en la sociedad es evidente: al menos el 50 % de los adultos experimenta niveles moderados de estrés, mientras que un 25 % presenta niveles altos, según la Asociación Americana de Psicología (APA) [1]. Estas clasificaciones dependen de la escala psicológica empleada, la interpretación de las respuestas y el análisis del experto. En general, puntuaciones altas indican mayor estrés, mientras que las bajas reflejan niveles menores [A].

Este contexto ha motivado múltiples estudios para explorar situaciones, factores y estresores, así como sus efectos en el cuerpo humano. Entre estos, destaca la encuesta anual “Stress in America”, reconocida por evaluar las actitudes y percepciones relacionadas con el estrés en la población de los Estados Unidos [1]. Asimismo, se ha promovido el desarrollo de sistemas tecnológicos para evaluar los niveles de estrés en sujetos, integrando herramientas como informes psicológicos, sistemas de adquisición de señales fisiológicas y algoritmos de aprendizaje automático (Machine Learning, ML) [2].

El estrés es la respuesta psicofisiológica del cuerpo ante estímulos percibidos como amenazas o desafíos, con efectos notables en la salud, la productividad y la calidad de vida. Según la teoría del Síndrome General de Adaptación (GAS) de Selye, el estrés se desarrolla en tres etapas: alarma, resistencia y agotamiento, describiendo cómo el cuerpo enfrenta y se adapta al estrés, hasta llegar al agotamiento si la exposición es prolongada [3].

Una forma efectiva de medir la respuesta somática o fisiológica al estrés es mediante bioseñales, que

son medidas objetivas evaluadas con, que son medidas objetivas evaluadas con precisión. Estas señales proporcionan información crucial sobre cambios metabólicos, morfológicos y funcionales en el cuerpo, permitiendo realizar evaluaciones objetivas de la salud mental de los sujetos [2].

La adquisición, procesamiento y análisis de bioseñales como el electrocardiograma (ECG) la electromiografía (EMG), la actividad electrodérmica (EDA), la electroencefalografía (EEG) y la fotopletismografía (PPG) permiten crear biomarcadores útiles para detectar el estrés. Estas señales, relacionadas con el sistema nervioso simpático (SNS), aportan información clave para identificar respuestas de estrés en el organismo [2].

En entornos controlados, el monitoreo continuo de señales fisiológicas, como la variabilidad cardíaca y la conductancia de la piel, permite detectar la activación del sistema nervioso simpático ante situaciones de estrés [2]. Integrar un análisis de señales de ECG, EMG y EDA con cuestionarios y entrevistas psicológicas minimiza el sesgo asociado a la subjetividad en la percepción individual [4].

Adicionalmente, la integración de herramientas como el aprendizaje automático, una rama de la inteligencia artificial que utiliza algoritmos para analizar datos, identificar patrones complejos y automatizar decisiones, ha transformado el diagnóstico de enfermedades y la predicción de resultados clínicos. Combinado con técnicas de procesamiento de señales, ha mejorado notablemente la atención médica [5].

En el ámbito de la psicología, el aprendizaje automático se emplea para facilitar el análisis de señales fisiológicas, como PPG, que refleja los cambios en el volumen sanguíneo asociados con la actividad cardíaca, y EDA, que mide la conductancia de la piel como un indicador de la respuesta autonómica al estrés. Algoritmos como SVM (Support Vector Machine), MLP (Multilayer Perceptron), KNN (K-Nearest Neighbors) y DT (Decision Tree) permiten identificar estados y niveles de estrés (bajo, medio, alto). Estas herramientas mejoran la evaluación precisa del estrés, potenciando las intervenciones en salud mental y la calidad de vida [2].

En la literatura predominan las técnicas de aprendizaje supervisado para identificar estados de estrés, las cuales dependen de datos etiquetados para entrenar modelos, evaluar su desempeño y ajustar parámetros [2]. Sin embargo, la obtención de etiquetas precisas durante la evaluación del estrés es un proceso lento, costoso y que requiere la intervención de expertos, como psicólogos, especialmente en contextos reales donde debe considerarse la variabilidad individual de las respuestas somáticas al estrés. Generalmente, estas etiquetas se asignan en función de la actividad y los estresores utilizados, y, en ocasiones, se complementan con autoinformes, como la Escala de Estrés Percibido (PSS), lo que permite reflejar estados binarios (estrés/no estrés) o niveles discretos (bajo, medio, alto) [6]. Si bien estos enfoques son efectivos en escenarios controlados, su dependencia de etiquetas limita su escalabilidad y aplicabilidad en contextos reales con datos fisiológicos complejos y dinámicos.

En contraste, los algoritmos de aprendizaje no supervisado, al no requerir etiquetas, los convierte en una alternativa factible para identificar patrones en las respuestas somáticas al estrés. Ofreciendo una forma autónoma de detectar estados de estrés en datos fisiológicos no etiquetados, lo que sugiere una mejor adaptación a problemas como la variabilidad individual y la identificación de respuestas fisiológicas de estrés en entornos reales.

En este contexto, el algoritmo DBSCAN puede ser un algoritmo auxiliar debido a su utilidad para identificar patrones en conjuntos de datos que, como en el caso del estrés, pueden contener ruido o estructuras complejas. A diferencia de métodos como K-means, DBSCAN no requiere definir previamente el número de clústers, ya que los determina automáticamente según la densidad de los puntos [7]. Los

puntos se clasifican en tres categorías: Core (núcleo), Frontera y Ruido [8, 9]. Las ventajas de DBSCAN son notables, destacándose su capacidad para identificar clústers de formas arbitrarias, manejar eficazmente el ruido y operar sin la necesidad de un número predefinido de clústers [9]. Estas características resultan especialmente útiles en la clasificación del estrés, donde las relaciones entre los datos suelen ser diversas y complejas [2].

De este modo, este estudio evalúa la eficacia del método DBSCAN en la identificación de respuestas fisiológicas al estrés en estudiantes universitarios, utilizando datos no etiquetados. El conjunto de datos incluye señales de PPG y EDA obtenidas durante pruebas controladas de inducción de estrés. El objetivo es identificar patrones fisiológicos que permitan clasificar tres niveles de estrés (bajo, medio y alto), asociados a los diferentes escenarios de la prueba de inducción. Estas pruebas están diseñadas para generar un incremento gradual del estrés en los participantes. A pesar de la ausencia de etiquetas, se espera que DBSCAN sea capaz de detectar y diferenciar patrones en las señales fisiológicas, asociando estos niveles de estrés a los escenarios de la prueba.

METODOLOGÍA

Las pruebas de inducción de estrés se utilizan para generar respuestas psicológicas y fisiológicas al recrear situaciones estresantes, siendo clave para desarrollar tecnologías de detección de estrés mediante señales fisiológicas. Sin embargo, su capacidad para diferenciar de manera eficaz entre estados de estrés y no estrés sigue siendo incierta. Múltiples investigaciones suponen que estas pruebas provocan cambios fisiológicos significativos y, bajo esta premisa, han desarrollado clasificadores de estrés con alta precisión en condiciones controladas [2]. Sin embargo, su aplicabilidad en entornos reales es limitada, destacando la necesidad de un análisis más detallado de señales fisiológicas, especialmente en contextos de alta demanda cognitiva. Además, la respuesta al estrés varía según el entorno o contexto, incluso al realizar actividades idénticas, como se propone en la prueba de inducción de estrés empleada en este estudio.

Para llevar a cabo esta evaluación, se analizaron señales de PPG y EDA de 51 estudiantes licenciatura del área de ingeniería de la Universidad Autónoma del Estado de México, incluyendo 16 mujeres y 35 hombres, con edades de 19 a 30 años (edad promedio de 22 años). Cada participante realizó una única prueba de inducción de estrés, durante la cual se registraron las señales fisiológicas.

La prueba consistió en el desarrollo de tres exámenes en tres escenarios distintos, cada participante completó la prueba una única vez, resolviendo exámenes de comprensión lectora o de razonamiento lógico-matemático. Los escenarios se diferenciaron por los estresores aplicados, aunque actividad era la misma (ver Figura 1):

- Primer escenario: Sin estresores, se utilizó para establecer la línea base de las respuestas fisiológicas.
- Segundo escenario: Incluyó dos estresores psicológicos: restricción de tiempo (15 minutos) y ruido ambiental de conversaciones, reproducidas en una computadora portátil. Estos fueron seleccionados para inducir un aumento en la respuesta fisiológica al estrés. [10].
- Tercer escenario: Se añadió un estresor físico a los estresores psicológicos del segundo escenario: pequeñas descargas eléctricas inferiores a 5 mA y con una duración menor a 5 segundos al final de la prueba. Antes de este escenario, el dispositivo empleado fue calibrado para determinar el

umbral de estimulación de cada participante, aunque las descargas nunca fueron aplicadas. Estas se utilizaron únicamente como amenaza para intensificar la respuesta fisiológica, sin causar daño a los sujetos [10, 11].

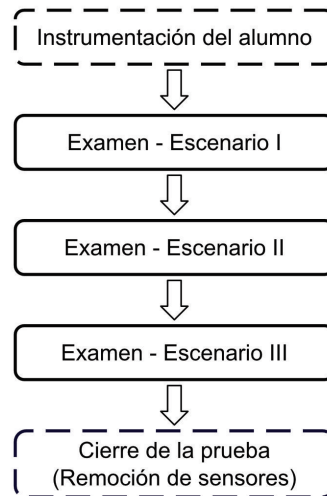


Figura 1: *Distribución general de la prueba de inducción de estrés académico.*

Durante la prueba, cada participante utilizó una banda con sensores de PPG y EDA en la frente, conectada a una computadora para registrar las señales de forma continua. Esta ubicación evitó limitar el movimiento de las manos, donde típicamente se realiza la instrumentación para adquirir estas señales, y redujo la susceptibilidad a artefactos de movimiento [2].

En cada escenario, los participantes realizaron un examen de comprensión lectora o razonamiento lógico-matemático, con un registro continuo de las señales a una frecuencia de 100 Hz para monitorear la respuesta fisiológica. Al finalizar los tres exámenes, se retiró la banda, concluyendo el registro de datos.

Tras la adquisición, las señales se analizaron en una segunda etapa que incluyó el preprocesamiento, la extracción y procesamiento de características, y la aplicación del método DBSCAN. La metodología utilizada se presenta en la Figura 2.

Durante el preprocesamiento, las señales se segmentaron y clasificaron según el escenario de adquisición, asociándose a los tres niveles de estrés esperados:

- Estrés bajo: Primer escenario sin estresores.
- Estrés medio: Segundo escenario con dos estresores psicológicos.
- Estrés alto: Tercer escenario con estresores psicológicos y físicos.

Debido a la variabilidad en la longitud y al ruido por artefactos de movimiento en las señales crudas, se aplicó un filtrado específico según el espectro de cada señal: pasa-bajas a 15 Hz para PPG [12] y a 0.15 Hz para EDA [13], eliminando componentes no deseados que afectarían la extracción de características. Las

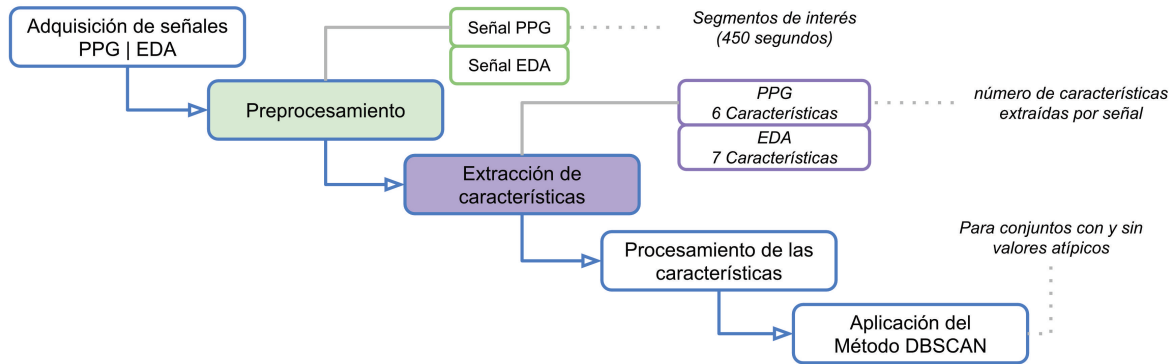


Figura 2: Diagrama a bloques general para la aplicación del método DBSCAN en señales de PPG y EDA.

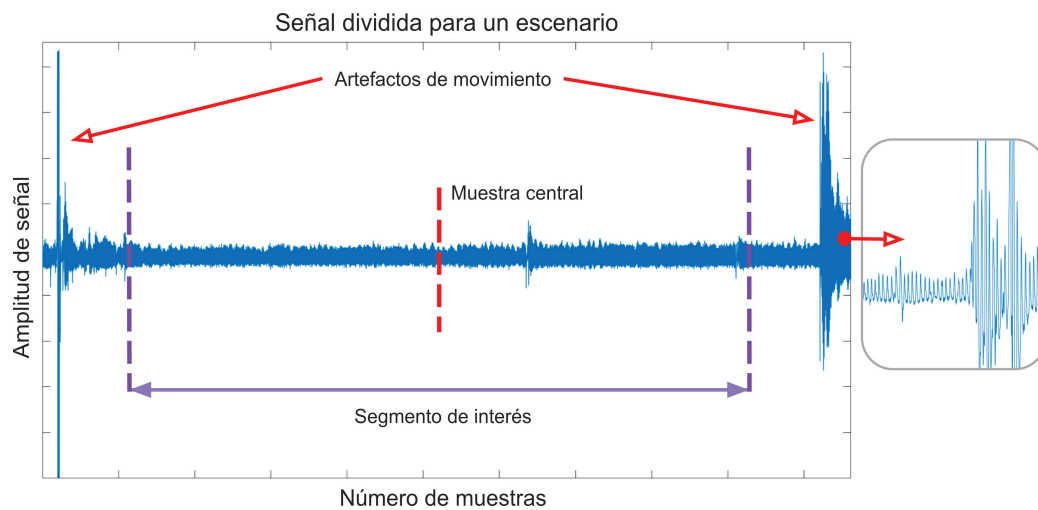


Figura 3: Selección de los intervalos de interés en las señales para reducir los artefactos de movimiento inicial y final en cada escenario de la prueba.

señales se segmentaron eliminando muestras iniciales y finales afectadas por ruido, obteniendo intervalos centrales con una duración uniforme de 450 segundos (ver Figura 3). Finalmente, las señales filtradas se agruparon en dos conjuntos, PPG y EDA, para los tres escenarios.

Para cada conjunto de señales, se realizó una extracción de características usando ventanas móviles ajustadas al tipo de señal. Para PPG, se emplearon ventanas de 90 segundos con un traslape de 45 segundos por segmento de interés en los tres escenarios [12]. Para EDA, las ventanas fueron de 60 segundos con un traslape de 30 segundos por segmento en cada escenario. [13].

En las señales PPG, se extrajeron seis características por ventana: promedio (C1), desviación estándar (C2) y valor cuadrático medio (RMS, por sus siglas en inglés) de los intervalos pulso-pulso (C3) así como el valor promedio (C4), desviación estándar (C5) y RMS de la frecuencia cardíaca (C6) [12]. Para las señales de EDA, se calcularon siete características por ventana: pendiente de la regresión lineal (C1), área entre la señal y su regresión lineal (C2), valor promedio (C3), desviación estándar (C4), RMS (C5), promedio de variaciones (C6) y cambio máximo (C7) [14]. Este proceso se aplicó a cada señal filtrada y en los tres escenarios estudiados. Este proceso se aplicó a cada señal filtrada y en los tres escenarios estudiados, seleccionando las características más utilizadas en la literatura.

Realizado el procesamiento de las señales y la extracción de características, se estructuraron dos conjuntos de datos que abarcan los tres escenarios analizados (ver Figura 4). El primer conjunto incluye las características extraídas de la señal de PPG, asociadas a los niveles de estrés en cada escenario, mientras que el segundo contiene las características derivadas de la señal de EDA, también vinculadas a estos niveles de estrés. Los datos fueron normalizados mediante z-score (media 0, desviación estándar 1) para garantizar una contribución uniforme de las características al análisis (ver Figura 5).

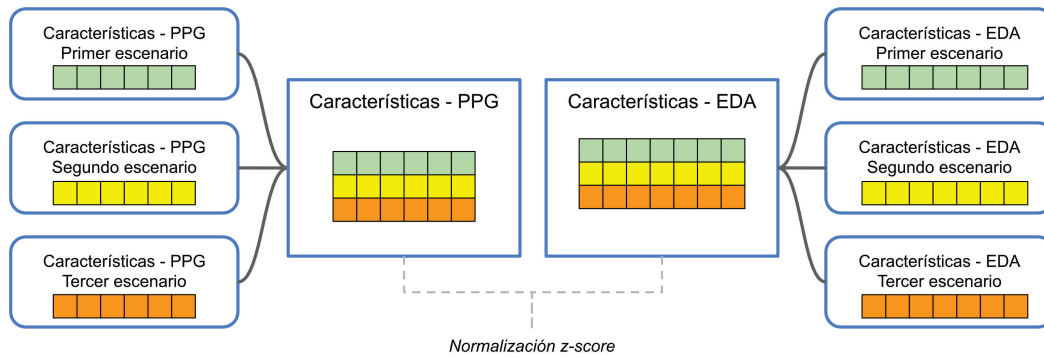


Figura 4: Estructuración de las características de las señales de PPG y EDA para los tres escenarios de la prueba de inducción de estrés.

Después, se utilizaron gráficos de caja para identificar datos atípicos en las características de EDA y PPG. En EDA, se detectaron valores atípicos en las características C1, C2, C4, C6 y C7, mientras que en PPG solo se observaron en C2 y C5.

Dada la dispersión de los datos, se realizaron dos pruebas con DBSCAN. En la primera, se incluyeron todas las características, abarcando tanto datos típicos como atípicos, para obtener una visión general de los patrones en las características de las señales. En la segunda, se excluyeron los valores atípicos, enfo-

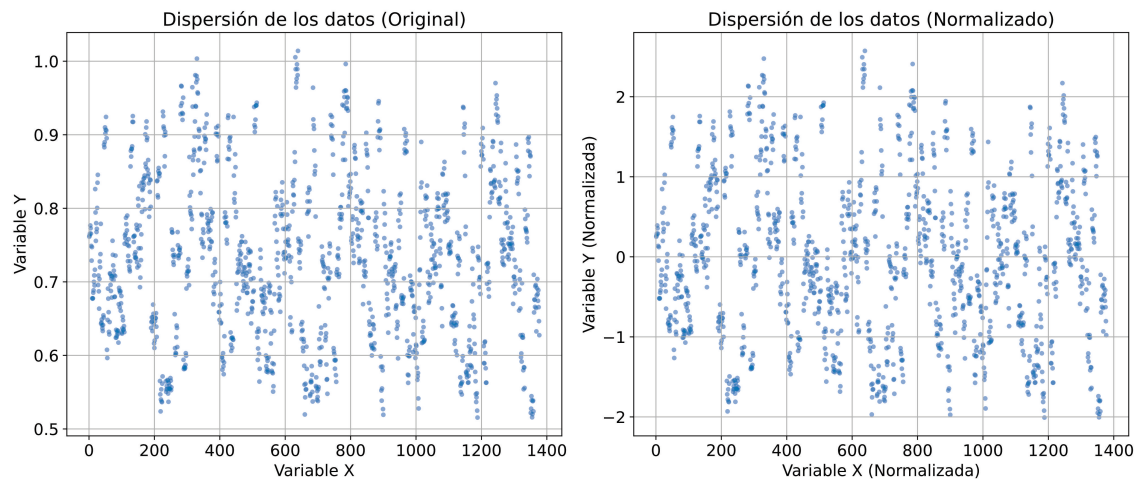


Figura 5: *Dispersión representativa de los datos antes y después de la normalización (media 0, desviación estándar 1).*

cándose en patrones más consistentes y representativos. Esto permitió comparar los clústeres generados en ambos conjuntos de datos.

Para facilitar la visualización de los resultados, se aplicó la técnica de PCA para reducir la dimensionalidad a dos componentes principales en ambos conjuntos de datos (seis características para PPG y siete para EDA). Aunque efectiva, esta técnica puede conllevar cierta pérdida de información. Se incorporó un identificador visual para los niveles de estrés esperados: 1 (bajo), 2 (medio) y 3 (alto), como se ilustra en la Figura 6. Es importante señalar que este etiquetado se empleó exclusivamente con fines visuales y no influyó en la aplicación del método DBSCAN, permitiendo observar la distribución inicial de las características según el escenario.

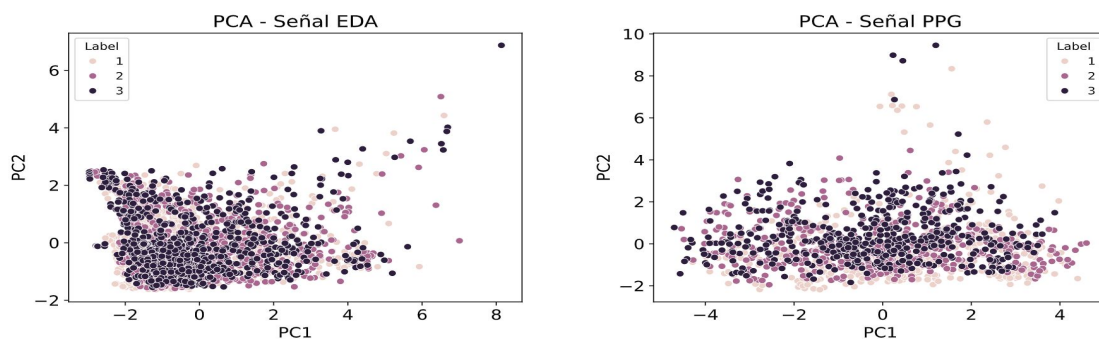


Figura 6: *Visualización de la aplicación de PCA para visualizar la dispersión de los niveles de estrés en el conjunto de características extraídas de las señales de EDA y PPG.*

Posteriormente, se implementó el método DBSCAN para identificar patrones de respuestas fisiológi-

cas asociados con los niveles de estrés en cada escenario. Para ello, fue necesario determinar los valores de los hiper-parámetros del modelo, como la distancia máxima entre dos muestras (eps) para formar un clúster o región densa, y el número mínimo de muestras requeridas para considerar un dato como punto central [7, 8].

Para determinar la distancia máxima entre dos puntos vecinos, se empleó el método de los k-ésimos vecinos más cercanos (KNN) con un valor típico de $k=5$, que representa menos del 0.5 % de los datos. Esto permitió calcular iterativamente la distancia máxima entre cada punto y sus cinco vecinos más cercanos. Las distancias obtenidas generaron una curva característica de crecimiento exponencial, cuyo punto de inflexión se utilizó para determinar el valor óptimo de eps, diferenciando regiones densas de puntos dispersos o ruido. De esta forma, para las características de la señal EDA, el valor eps fue 4 con todas las características y 0.2 sin datos atípicos. Para la señal PPG, los valores fueron 4.5 y 0.2 en las mismas condiciones.

RESULTADOS Y DISCUSIÓN

En la Figura 7 se observa cómo el método DBSCAN aplicado al conjunto de datos completo (incluidos los valores atípicos) identificó solo una región densa, usando un mínimo de 5 vecinos y valores de eps de 4 y 4.5 para cada señal, respectivamente. Este resultado indica una diferenciación limitada en los datos, con la identificación de un único grupo en ambos casos. Esto contrasta con los grupos inicialmente identificados según los escenarios (ver Figura 6), donde se esperaba una mejor segmentación de los datos.

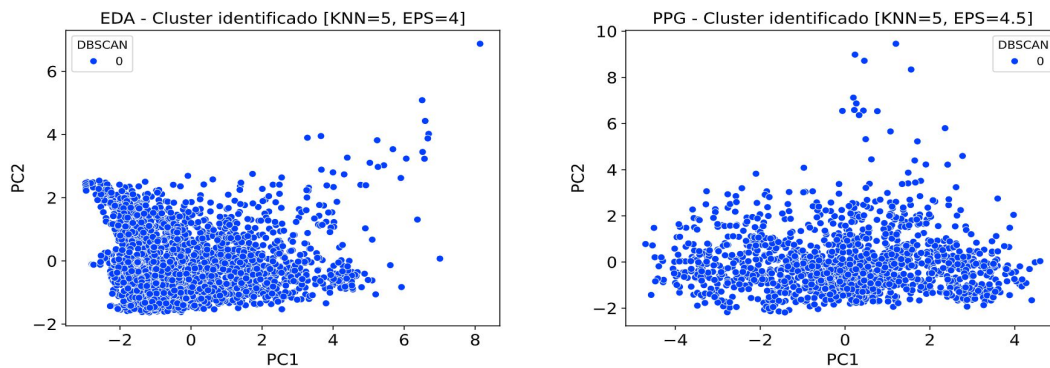


Figura 7: Visualización de los clústeres obtenidos mediante DBSCAN aplicado a las características de las señales de EDA y PPG, considerando valores atípicos.

Al eliminar los valores atípicos y aplicar nuevamente DBSCAN, esta vez con un valor de eps de 0.2 para ambos conjuntos y manteniendo el número de vecinos, se identificaron cuatro grupos en las características de EDA y tres en las de PPG. Sin embargo, como muestra la Figura 8 las regiones adicionales contienen pocos elementos y ruido, mientras que el volumen principal de datos permanece asociado al clúster inicial. Esto sugiere que DBSCAN, con estos hiper-parámetros, no es eficaz para formar grupos en datos asociados a respuestas fisiológicas de estrés, en comparación con otros métodos como ANOVA o técnicas de aprendizaje automático supervisado.

Sin embargo, estos resultados no implican que algoritmos de aprendizaje no supervisado no sean capaces de identificar grupos en este tipo de datos. Un ejemplo es el método Gaussian Mixture, aplicado al conjunto de datos sin valores atípicos, que mostró mayor flexibilidad al detectar tres grupos. Los clústeres generados (ver Figura 9) evidencian parcialmente la segmentación esperada de los datos fisiológicos, aunque no coinciden completamente con los datos originales, destacando la capacidad del método para adaptarse a grupos dispersos o concentrados [15]. Así, este modelo se presenta como una alternativa prometedora para detectar respuestas de estrés en las señales de PPG y EDA.

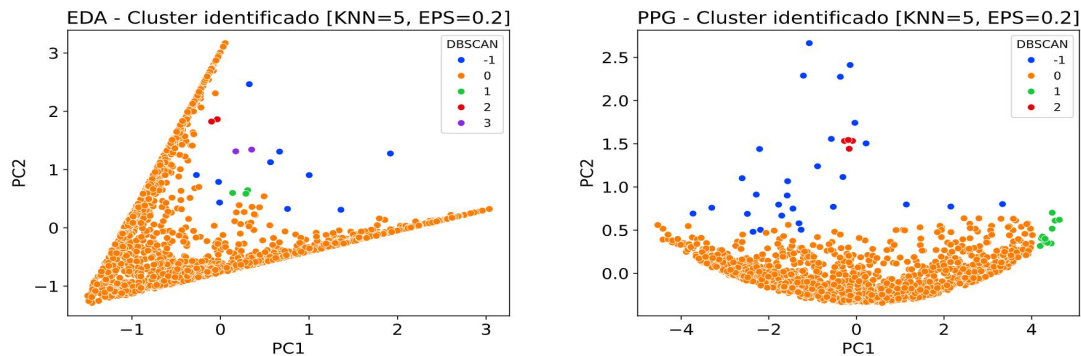


Figura 8: Visualización de los clústeres obtenidos mediante DBSCAN aplicado a las características de las señales de EDA y PPG, no considerando valores atípicos.

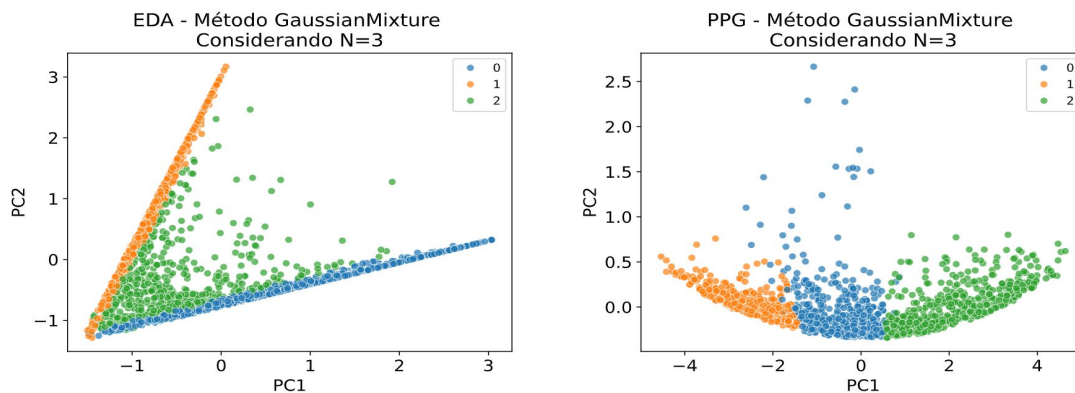


Figura 9: Visualización de los clústeres obtenidos mediante el método GaussianMixture aplicado a las características de las señales de EDA y PPG, no considerando valores atípicos.

Los resultados obtenidos con DBSCAN evidencian su capacidad limitada para distinguir entre diferentes niveles de respuesta al estrés utilizando datos no etiquetados. La falta de diferenciación en los clústeres sugiere que los hiper-parámetros y características seleccionados no fueron suficientes para capturar la complejidad de las respuestas fisiológicas. Aunque la eliminación de valores atípicos mejoró

parcialmente la segmentación, la prevalencia de ruido y la dispersión de los datos en clústeres con pocos elementos indican que DBSCAN no alcanzó la precisión necesaria para este tipo de análisis.

Estos resultados resaltan los desafíos del aprendizaje no supervisado en el análisis de estrés y la importancia de explorar enfoques alternativos. Por ejemplo, el modelo Gaussian Mixture, al permitir que cada grupo tenga su propia varianza, demostró ser una alternativa prometedora para analizar datos con distribuciones complejas y variadas [15]. Esta característica le proporciona la flexibilidad necesaria para adaptarse a grupos dispersos o concentrados, superando las limitaciones de DBSCAN en el análisis de señales fisiológicas.

De esta forma, los resultados obtenidos abren oportunidades para futuras investigaciones, que podrían centrarse en el ajuste de hiper-parámetros, la integración de DBSCAN con otros métodos de agrupamiento o la aplicación de técnicas de preprocesamiento para mejorar los resultados. Aunque DBSCAN no identificó los tres niveles de estrés como se esperaba, el estudio resalta la complejidad de las respuestas fisiológicas al estrés y la necesidad de enfoques más robustos, que permitan seguir explorando y perfeccionando técnicas de análisis de señales, especialmente cuando el etiquetado no es factible, en el contexto de la salud mental.

CONCLUSIONES

Este estudio aporta evidencia sobre las limitaciones del método DBSCAN para la segmentación de señales fisiológicas para la identificación de estrés. Aunque la combinación de la reducción de dimensionalidad mediante PCA y DBSCAN es útil para el análisis exploratorio de datos no etiquetados, los resultados muestran que los clústeres identificados no son representativos de las etiquetas del conjunto de datos original, identificando un solo grupo o grupos con pocos elementos. Esto sugiere que, la hipótesis de que el método DBSCAN podría detectar patrones y diferenciar niveles de estrés en señales fisiológicas no etiquetadas no es válida.

Los hallazgos obtenidos subrayan la complejidad inherente en los datos fisiológicos, donde factores como la correcta selección de parámetros y la variabilidad individual en la respuesta al estrés dificultan el uso de métodos de aprendizaje automático. Esto es especialmente relevante en situaciones reales, donde no siempre es posible contar con un experto o realizar un etiquetado constante de los datos.

Por otro lado, los resultados obtenidos cuestionan la efectividad de DBSCAN en este contexto, destacando la necesidad de explorar métodos alternativos, como K-Medoids o Gaussian Mixture, siendo este último utilizado en el estudio, donde se lograron identificar clústeres definidos, aunque no coincidentes con los datos originales. Además, es necesario optimizar la selección de los hiper-parámetros de DBSCAN, así como los modelos de aprendizaje empleados, realizando más iteraciones para determinar adecuadamente estos valores, con el fin de formar agrupamientos más consistentes.

Futuras investigaciones deben centrarse en técnicas de preprocesamiento más avanzadas, como la extracción de características en el dominio de la frecuencia, y explorar otros métodos de clustering o enfoques híbridos que combinen varios algoritmos, como aquellos que incluyan la exclusión de valores atípicos o poda, especialmente en contextos de datos no etiquetados o experimentos donde realizar un etiquetado riguroso no sea posible. En resumen, este estudio sugiere que, aunque DBSCAN y otros métodos de aprendizaje no supervisado pueden servir como punto de partida, son necesarios enfoques más robustos y flexibles para clasificar eficazmente las respuestas fisiológicas al estrés.

REFERENCIAS

- [1] American Psychological Association. Stress in America. [Online]. Available: <https://www.apa.org/news/press/releases/stress>. [Accessed: Nov. 28, 2024].
- [2] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis. Review on Psychological Stress Detection Using Biosignals. *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, Jan.–Mar. 2022. DOI: 10.1109/TAFFC.2019.2927337.
- [3] L. M. Manosso, C. R. Gasparini, and G. Z. Réus. Definitions and Concepts of Stress. in *Glutamate and Neuropsychiatric Disorders: Current and Emerging Treatments*, Z. M. Pavlovic, Ed. Cham: Springer International Publishing, 2022, pp. 27–63. DOI: 10.1007/978-3-030-87480-3_2.
- [4] M. Z. Baig and M. Kavakli. A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. *Multimodal Technologies and Interaction*, vol. 3, no. 2, Article 37, 2019. DOI: 10.3390/mti3020037.
- [5] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons. Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*, vol. 19, no. 1, p. 64, Mar. 2019. DOI: 10.1186/s12874-019-0681-4.
- [6] M. U. Mustafa, S. A. Buzdar, A. Ikhlq, M. Nisa, S. Malik, S. Saeed, M. S. Khan, and A. Javid. Machine learning approach for multi-class stress assessment with electroencephalography signals. *Journal of Population Therapeutics and Clinical Pharmacology*, vol. 30, no. 19, pp. 1464–1478, 2023. DOI: 10.53555/jptcp.v30i19.3852.
- [7] Scikit-Learn. Clustering. 2024. [Online]. Available: <https://scikit-learn/stable/modules/clustering.html>. [Accessed: Nov. 28, 2024].
- [8] MathWorks. DBSCAN Clustering. 2024. [Online]. Available: https://la.mathworks.com/help/stats/dbscan-clustering.html#mw_0ad71796-4ae1-4620-a0e5-9ea5dcded2c6. [Accessed: Nov. 28, 2024].
- [9] O. Kulkarni and A. Burhanpurwala. A Survey of Advancements in DBSCAN Clustering Algorithms for Big Data. in *2024 3rd International Conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC)*, 2024, pp. 106–111. DOI: 10.1109/PARC59193.2024.10486339.
- [10] D. H. Lamb. On the distinction between physical and psychological stressors. *Motivation and Emotion*, vol. 3, pp. 51–61, Mar. 1979. DOI: 10.1007/BF00994160.
- [11] J. Rodríguez-Arce, L. Lara-Flores, O. Portillo-Rodríguez, and R. Martínez-Méndez. Towards an anxiety and stress recognition system for academic environments based on physiological features. *Computer Methods and Programs in Biomedicine*, vol. 190, p. 105408, 2020. DOI: 10.1016/j.cmpb.2020.105408.

- [12] E. Mejía-Mejía and P. A. Kyriacou. Duration of photoplethysmographic signals for the extraction of Pulse Rate Variability Indices. *Biomedical Signal Processing and Control*, vol. 80, p. 104214, 2023. DOI: 10.1016/j.bspc.2022.104214.
- [13] D. Leiner, A. Fahr, and H. Früh. EDA Positive Change: A Simple Algorithm for Electrodermal Activity to Measure General Audience Arousal During Media Exposure. *Communication Methods and Measures*, vol. 6, no. 4, pp. 237–250, 2012. DOI: 10.1080/19312458.2012.732627.
- [14] R. Martinez, A. Salazar-Ramirez, A. Arruti, E. Irigoyen, J. I. Martin, and J. Muguerza. A Self-Paced Relaxation Response Detection System Based on Galvanic Skin Response Analysis. *IEEE Access*, vol. 7, pp. 43730–43741, 2019. DOI: 10.1109/ACCESS.2019.2908445.
- [15] Scikit-Learn. Gaussian Mixture Models. 2024. [Online]. Available: <https://scikit-learn.org/1.5/modules/mixture.html>. [Accessed: Nov. 28, 2024].