

Graph Machine Learning (GraphML) para la detección temprana de riesgos para la salud mental: una revisión sistemática de literatura

Graph Machine Learning (GraphML) for early detection of mental health risks: a systematic literature review

Jorge Edwin Delgado Altamirano*
Universidad Nacional Mayor de San Marcos
Lima - Perú
Universidad Nacional de Chimborazo
Riobamba - Ecuador
jdelgado@unach.edu.ec
<https://orcid.org/0000-0002-6782-806X>

Ciro Rodriguez Rodriguez
Universidad Nacional Mayor de San Marcos
Lima - Perú
crodriguezro@unmsm.edu.pe
<https://orcid.org/0000-0003-2112-1349>

*Correspondencia:
jdelgado@unach.edu.ec

Cómo citar este artículo:
Delgado, J., & Rodriguez, C. (2026). Graph Machine Learning (GraphML) para la detección temprana de riesgos para la salud mental: una revisión sistemática de literatura. *Esprint Investigación*, 5(2), 529-547. <https://doi.org/10.61347/10.61347/ei.v5i2.400>

Recibido: 5 de mayo de 2026
Aceptado: 5 de octubre de 2026
Publicado: 26 de octubre de 2026

Copyright: Derechos de autor 2026 Jorge Edwin Delgado Altamirano, Ciro Rodriguez Rodriguez.



Esta obra está bajo una licencia internacional Creative Commons Atribución-NonComercial 4.0.

Resumen: La detección temprana de riesgos en salud mental constituye un desafío debido a la complejidad de las relaciones entre factores clínicos, conductuales y contextuales. El objetivo de esta investigación fue sintetizar y analizar la evidencia científica sobre la aplicación de Graph Machine Learning (GraphML) en la detección temprana de riesgos en salud mental, considerando las arquitecturas utilizadas, los trastornos abordados, las métricas de desempeño, las estrategias de validación y las principales limitaciones. Se realizó una revisión sistemática siguiendo las directrices PRISMA 2020, mediante búsquedas en Scopus y Web of Science de estudios originales publicados entre 2021 y 2026. Tras aplicar los criterios de elegibilidad y evaluar el riesgo de sesgo mediante la herramienta PROBAST, se incluyeron 27 estudios en la síntesis cualitativa. Los resultados evidenciaron un predominio de arquitecturas basadas en redes convolucionales de grafos (GCN), seguidas de redes de atención de grafos (GAT), GraphSAGE y modelos híbridos multimodales, aplicados principalmente a depresión, ansiedad, estrés e ideación suicida. Las métricas más utilizadas fueron exactitud, F1-score y área bajo la curva (AUC), con resultados predictivos favorables, especialmente en modelos que integraron datos clínicos, neuroimágenes, señales fisiológicas e información digital. Sin embargo, se identificaron limitaciones relacionadas con la validación externa, el desequilibrio de clases, la heterogeneidad de los datos y la interpretabilidad. Se concluyó que GraphML presenta potencial para apoyar la detección temprana de riesgos en salud mental, aunque su aplicación clínica requiere mayor validación multicéntrica, explicabilidad y reproducibilidad metodológica.

Palabras clave: Aprendizaje automático sobre grafos, depresión, detección temprana, Graph Machine Learning, redes neuronales sobre grafos.

Abstract: Early detection of mental health risks remains a significant challenge due to the complexity of relationships among clinical, behavioral, and contextual factors. The objective of this study was to synthesize and analyze scientific evidence on the application of Graph Machine Learning (GraphML) for the early detection of mental health risks, considering the architectures employed, disorders addressed, performance metrics, validation strategies, and main limitations. A systematic literature review was conducted following the PRISMA 2020 guidelines, using Scopus and Web of Science to identify original studies published between 2021 and 2026. After applying the eligibility criteria and assessing the risk of bias using the Prediction Model Risk of Bias Assessment Tool (PROBAST), 27 studies were included in the qualitative synthesis. The findings revealed a predominance of Graph Convolutional Network (GCN) architectures, followed by Graph Attention Networks (GAT), GraphSAGE, and multimodal hybrid models, primarily applied to depression, anxiety, stress, and suicidal ideation. The most frequently reported performance metrics were accuracy, F1-score, and area under the curve (AUC), with favorable predictive results, particularly in models integrating clinical data, neuroimaging, physiological signals, and digital information. However, limitations related to external validation, class imbalance, data heterogeneity, and model interpretability were identified. It was concluded that GraphML shows potential to support the early detection of mental health risks, although its clinical application requires stronger multicenter validation, explainability, and methodological reproducibility.

Keywords: Depression, early detection, graph-based machine learning, graph machine learning, graph neural networks.

1. Introducción

La detección temprana de riesgos en salud mental constituye una prioridad científica debido a que los desenlaces clínicos responden a interacciones entre antecedentes médicos, síntomas, factores psicosociales, trayectorias de atención y biomarcadores (Bollen et al., 2022; Mansoor & Ansari, 2024). En este contexto, la inteligencia artificial ha adquirido relevancia en el diagnóstico, monitoreo e intervención de trastornos mentales (Cruz-Gonzalez et al., 2025). Entre sus aplicaciones, el aprendizaje automático basado en grafos destaca por su capacidad para representar relaciones complejas entre variables y estructuras de datos no euclidianos (Paul et al., 2024; Weber, 2025).

Graph Machine Learning (GraphML) comprende un conjunto de métodos de aprendizaje automático que procesan información organizada en grafos, cuyos nodos representan entidades y cuyas aristas describen sus relaciones. Estos métodos utilizan mecanismos de propagación de mensajes, agregación de información vecinal y modelado de dependencias topológicas para identificar patrones complejos (Weber, 2025; Zhou et al., 2020). Su aplicación resulta especialmente útil cuando los fenómenos estudiados involucran interacciones entre pacientes, diagnósticos, eventos clínicos y fuentes heterogéneas de información (Boll et al., 2024). No obstante, su utilidad depende de la arquitectura empleada, la construcción del grafo, la calidad de los datos y la interpretabilidad de los resultados (Boll et al., 2024; Paul et al., 2024).

Los riesgos en salud mental comprenden diversos desenlaces clínicos, entre ellos la ideación suicida, las conductas autolesivas, las recaídas y la evolución crónica de trastornos depresivos, asociados con sufrimiento psicológico y deterioro funcional (Cripps et al., 2020; Knipe et al., 2022). Su naturaleza dinámica y multifactorial dificulta la identificación de patrones mediante modelos predictivos convencionales basados en relaciones estáticas (Nordin et al., 2022). Aunque el aprendizaje automático ha mostrado resultados prometedores en la predicción de estos riesgos, persisten dificultades relacionadas con la disponibilidad de información neurobiológica, la validación externa y la presentación de métricas clínicamente interpretables (Pigoni et al., 2024).

La incorporación de GraphML en la detección temprana de riesgos en salud mental ha despertado interés por su capacidad para representar redes de interacción entre individuos, síntomas y factores contextuales (Dong et al., 2021; Vaida & Huang, 2026). Las arquitecturas neuronales basadas en grafos permiten identificar patrones relacionales en datos de neuroimagen, registros clínicos y plataformas digitales, con aplicaciones en la identificación de trastornos depresivos (Okesanya et al., 2026). Su capacidad para integrar información multimodal favorece el desarrollo de modelos predictivos que consideran simultáneamente diferentes características clínicas y contextuales (Jiang, 2026; Vaida & Huang, 2026). Estas propiedades sustentan su potencial para contribuir a la identificación anticipada de alteraciones en la salud mental.

Diversas investigaciones han examinado las aplicaciones de la inteligencia artificial en la identificación y el manejo de trastornos mentales. Okesanya et al. (2026), mediante una revisión sistemática sobre el trastorno depresivo mayor, identificaron avances en el diagnóstico y la predicción de resultados clínicos mediante aprendizaje profundo y redes neuronales, aunque señalaron heterogeneidad metodológica y diferencias en las métricas de evaluación. Por su parte, Park et al. (2024) analizaron la efectividad de la inteligencia artificial en la detección y el manejo de trastornos depresivos, destacando su potencial clínico. Sin embargo, los autores identificaron desafíos relacionados con la validez externa de los modelos y la presencia de sesgos algorítmicos.

Desde una perspectiva poblacional, Ruan et al. (2026) desarrollaron una revisión sobre aplicaciones del aprendizaje automático en la promoción de la salud mental de adultos mayores, identificando

como principales líneas de investigación la predicción de riesgos y la detección de estados mentales. Sus hallazgos destacaron la importancia de mejorar la calidad de los datos y la interpretabilidad de los modelos para favorecer su aplicación clínica. Por otra parte, Razavi et al. (2024) examinaron técnicas de aprendizaje automático y profundo para detectar, predecir y monitorear el estrés, identificando aplicaciones de máquinas de vectores de soporte, bosques aleatorios y redes neuronales. Los autores señalaron resultados prometedores, aunque reconocieron desafíos relacionados con la personalización de los modelos y el procesamiento de información en tiempo real.

A pesar de los avances descritos, la evidencia científica sobre GraphML aplicada a la detección temprana de riesgos en salud mental permanece distribuida entre estudios con diferentes arquitecturas, fuentes de datos, poblaciones y objetivos predictivos. Esta diversidad dificulta establecer comparaciones sistemáticas sobre el desempeño de los modelos y determinar su aplicabilidad en distintos contextos clínicos. Aunque las revisiones existentes examinan aplicaciones generales de inteligencia artificial y aprendizaje automático, resulta necesario integrar específicamente los hallazgos relacionados con las arquitecturas basadas en grafos. En consecuencia, se requiere una síntesis que permita identificar los enfoques utilizados, los riesgos abordados, las estrategias de validación y las limitaciones reportadas.

En este contexto, el objetivo de la presente revisión sistemática fue sintetizar y analizar la evidencia científica disponible sobre la aplicación de Graph Machine Learning en la detección temprana de riesgos en salud mental, considerando las arquitecturas empleadas, los trastornos abordados, las métricas de desempeño, los enfoques de validación y las principales limitaciones identificadas. Para orientar el análisis, se formularon cuatro preguntas de investigación: (1) ¿Qué arquitecturas de GraphML se han utilizado para la detección temprana de riesgos en salud mental?; (2) ¿Qué trastornos y riesgos de salud mental se han abordado mediante estas arquitecturas?; (3) ¿Qué métricas de desempeño y enfoques de validación se han empleado para evaluar los modelos propuestos?; y (4) ¿Cuáles son las principales limitaciones para la aplicación de GraphML en salud mental?

La investigación busca proporcionar una visión integrada del estado de la evidencia científica sobre el uso de GraphML en la identificación temprana de riesgos en salud mental. Sus resultados permitirán reconocer las principales tendencias metodológicas, los campos de aplicación y las características de evaluación de los modelos desarrollados. De esta manera, se pretende aportar fundamentos científicos que orienten futuras investigaciones hacia diseños transparentes, reproducibles y metodológicamente consistentes, con potencial para apoyar la detección temprana y la toma de decisiones en salud mental.

2. Metodología

La presente revisión sistemática se desarrolló conforme a las directrices de PRISMA 2020, con el propósito de favorecer la transparencia y reproducibilidad del proceso de identificación, selección, evaluación y síntesis de la evidencia científica (Page et al., 2021). El procedimiento metodológico comprendió la definición de los criterios de elegibilidad, la formulación de las estrategias de búsqueda, la selección de estudios, la evaluación del riesgo de sesgo y la extracción de información. Finalmente, se realizó una síntesis narrativa de los hallazgos relacionados con la aplicación de Graph Machine Learning (GraphML) en la detección temprana de riesgos en salud mental.

Criterios de inclusión y exclusión

Se incluyeron estudios originales publicados entre 2021 y 2026, disponibles hasta el 6 de mayo de 2026, que abordaron aplicaciones de Graph Machine Learning, redes neuronales de grafos o arquitecturas basadas en grafos relacionadas con la detección temprana, predicción de riesgos, evaluación o

diagnóstico de condiciones de salud mental. Se consideraron investigaciones que presentaron resultados empíricos y reportaron métricas de desempeño, procedimientos de validación o información metodológica suficiente para evaluar los modelos desarrollados. No se establecieron restricciones de idioma durante la identificación y selección de las publicaciones.

Se excluyeron los registros duplicados, editoriales, cartas al editor, resúmenes de congresos, capítulos de libro, tesis y artículos de revisión. También se descartaron investigaciones centradas exclusivamente en técnicas convencionales de aprendizaje automático o aprendizaje profundo que no incorporaron estructuras basadas en grafos. Se excluyeron estudios relacionados con enfermedades neurológicas o biomédicas sin vinculación explícita con riesgos en salud mental. Finalmente, se eliminaron las publicaciones que no presentaron resultados empíricos, métricas de desempeño o información metodológica suficiente para su análisis.

Fuentes de información y estrategia de búsqueda

La identificación de los estudios se realizó mediante consultas en las bases de datos Scopus y Web of Science, seleccionadas por su cobertura multidisciplinaria y disponibilidad de literatura científica especializada. Estas fuentes permitieron recuperar investigaciones relacionadas con inteligencia artificial, aprendizaje automático basado en grafos y aplicaciones orientadas a la salud mental. La búsqueda se limitó a artículos científicos originales, de acuerdo con los criterios de elegibilidad establecidos. Los registros recuperados constituyeron la base documental para el posterior proceso de selección y evaluación de los estudios.

La estrategia de búsqueda se estructuró mediante tres categorías conceptuales: aprendizaje automático basado en grafos, detección temprana y salud mental. Se utilizaron términos en inglés, operadores booleanos y truncamientos para identificar publicaciones relacionadas con las dimensiones del estudio. Las consultas se realizaron en los campos de título, resumen y palabras clave, con adaptaciones específicas a la sintaxis de cada base de datos. Las cadenas de búsqueda empleadas y el número de registros recuperados se presentan en la tabla 1.

Tabla 1

Estrategia de búsqueda

Base de datos	Cadena de búsqueda	Estudios
Scopus	TITLE-ABS-KEY("graph machine learning" OR "GraphML" OR "graph neural network*" OR "GNN" OR "graph convolutional network*" OR "GCN" OR "graph attention network*" OR "GAT" OR "GraphSAGE") AND TITLE-ABS-KEY("early detection" OR "early diagnosis" OR "risk prediction" OR "predictive model*" OR "risk assessment" OR screening) AND TITLE-ABS-KEY("mental health" OR depression OR anxiety OR stress OR suicid* OR "psychological disorder*" OR "psychiatric disorder*" OR "mental disorder*" OR "mental illness" OR "emotional disorder*") AND PUBYEAR > 2020 AND PUBYEAR < 2027 AND LIMIT-TO(DOCTYPE, "ar")	83
Web of Science	TS=("graph machine learning" OR "graphql" OR "graph neural network*" OR "GNN" OR "graph convolutional network*" OR "GCN" OR "graph attention network*" OR "GAT" OR "GraphSAGE") AND TS=("early detection" OR "early diagnosis" OR "risk prediction" OR "predictive model*" OR "risk assessment" OR screening) AND TS=("mental health" OR depression OR anxiety OR stress OR suicid* OR "psychological disorder*" OR "psychiatric disorder*" OR "mental disorder*" OR "mental illness" OR "emotional disorder*")	57
Total		140

Nota. Las cadenas se transcribieron a partir de la información proporcionada. En Web of Science se mantuvo el término "graphql" para no alterar una estrategia cuya ejecución original no se ha comprobado. El filtro de tipo documental fue «Article» y debe verificarse que también se aplicó el período 2021–2026. Los 140 resultados corresponden a registros recuperados antes de la eliminación de duplicados.

Proceso de selección de estudios

El proceso de selección se desarrolló por etapas, de acuerdo con las recomendaciones de PRISMA 2020. Inicialmente, los registros identificados en Scopus y Web of Science se exportaron y organizaron para detectar y eliminar duplicados. Posteriormente, dos revisores examinaron de manera independiente los títulos y resúmenes, aplicando los criterios de inclusión y exclusión establecidos. Los artículos potencialmente elegibles se evaluaron mediante la lectura de sus textos completos para determinar su incorporación definitiva a la revisión.

Durante la evaluación de elegibilidad, se verificó la relación de cada estudio con las aplicaciones de GraphML y su pertinencia para la detección temprana de riesgos en salud mental. Se consideraron las características metodológicas, la disponibilidad de resultados empíricos y la información presentada sobre el desempeño de los modelos. Las discrepancias identificadas entre los dos revisores se sometieron a la intervención de un tercer evaluador. La decisión final sobre la inclusión de los estudios se estableció mediante consenso entre los evaluadores.

Evaluación del riesgo de sesgo

La evaluación del riesgo de sesgo se realizó mediante la herramienta *Prediction Model Risk of Bias Assessment Tool* (PROBAST), aplicada a los estudios de desarrollo y validación de modelos predictivos. Esta herramienta permitió examinar aspectos metodológicos relacionados con cuatro dominios: participantes, predictores, resultados y análisis. La valoración consideró las características de las poblaciones estudiadas, la selección de variables predictoras, la definición de los desenlaces y los procedimientos estadísticos empleados. Los resultados de esta evaluación se utilizaron para interpretar la solidez metodológica de la evidencia incluida en la revisión.

Extracción y síntesis de datos

La extracción de información se realizó mediante una matriz diseñada para organizar sistemáticamente las características de los estudios incluidos. Se registraron el autor, año de publicación, país, diseño metodológico, tamaño y características de la población, fuentes de datos, trastornos o riesgos de salud mental y arquitecturas de GraphML utilizadas. También se recopilieron las técnicas de preprocesamiento, los procedimientos de validación, las métricas de desempeño y los principales resultados obtenidos. Finalmente, se incorporaron las limitaciones y conclusiones relevantes para facilitar la comparación e interpretación de la evidencia.

Posteriormente, se desarrolló una síntesis narrativa orientada a describir, comparar e interpretar los hallazgos de los estudios seleccionados. La información se organizó mediante categorías temáticas relacionadas con las preguntas de investigación, considerando las arquitecturas de GraphML, los trastornos y riesgos analizados, las métricas de desempeño y los enfoques de validación. También se examinaron las limitaciones metodológicas y las dificultades reportadas para la aplicación de estos modelos en contextos de salud mental. Esta organización permitió identificar tendencias y diferencias entre los estudios, así como integrar la evidencia disponible sobre su potencial para la detección temprana de riesgos.

3. Resultados

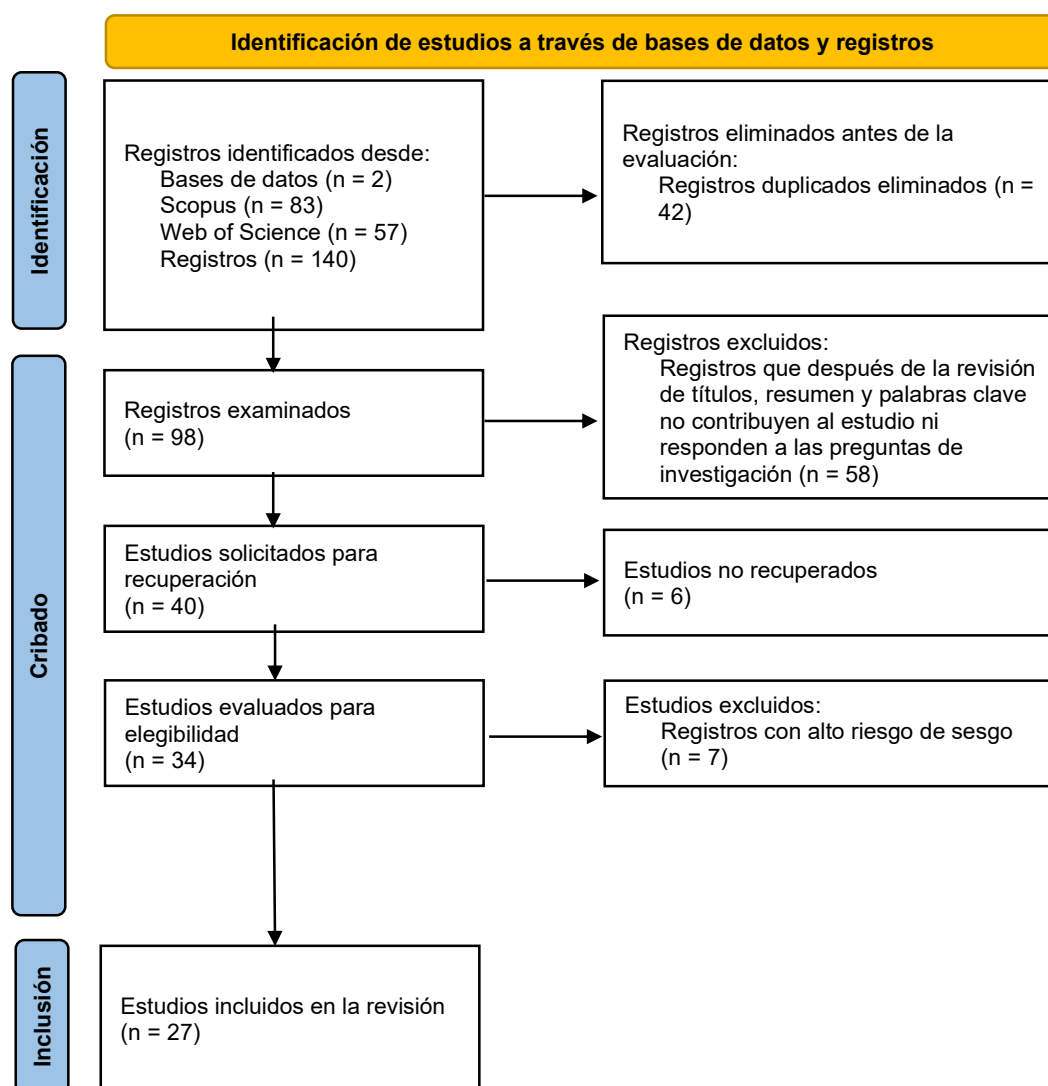
El proceso de selección se desarrolló conforme a las directrices de PRISMA 2020. En la fase de identificación se recuperaron 140 registros, de los cuales 83 procedieron de Scopus y 57 de Web of Science. Tras eliminar 42 duplicados, se conservaron 98 registros para el cribado de títulos y

resúmenes. De estos, 58 fueron excluidos por incumplir los criterios de elegibilidad, mientras que 40 se consideraron potencialmente relevantes.

De los 40 documentos seleccionados para recuperación, seis no pudieron obtenerse a texto completo. En consecuencia, se evaluó la elegibilidad de 34 artículos, de los cuales siete fueron excluidos por presentar un riesgo de sesgo alto, según la valoración metodológica reportada. Finalmente, 27 estudios se incorporaron a la síntesis cualitativa de la revisión sistemática. La Figura 1 presenta las etapas de identificación, cribado, recuperación y selección, junto con el número de registros correspondiente a cada fase.

Figura 1

Diagrama de flujo PRISMA

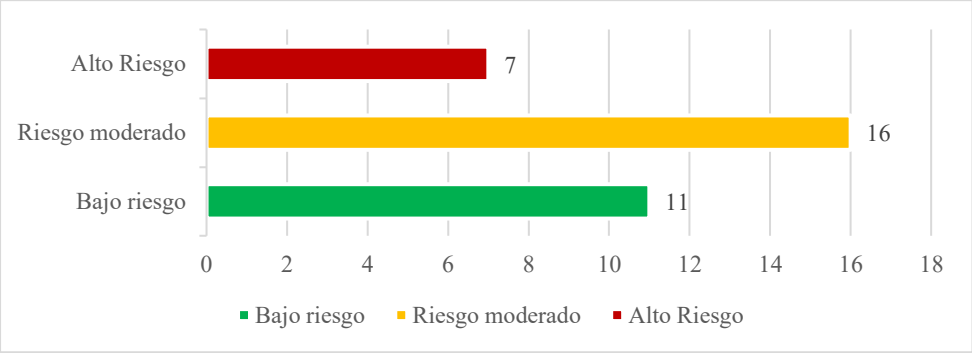


Evaluación del riesgo de sesgo

La evaluación metodológica comprendió 34 estudios potencialmente elegibles mediante la herramienta PROBAST. Según la clasificación presentada en la revisión, 11 investigaciones fueron consideradas de riesgo bajo, 16 de riesgo moderado y siete de riesgo alto. Estas últimas fueron excluidas de la síntesis final, por lo que se conservaron 27 estudios para el análisis de los resultados.

La figura 2 presenta la distribución de las categorías reportadas y permite identificar la composición metodológica de la evidencia seleccionada.

Figura 2
Evaluación del riesgo de sesgo mediante PROBAST



Arquitecturas GraphML y riesgos de salud mental abordados

Los 27 estudios incluidos examinaron diferentes arquitecturas de aprendizaje automático basado en grafos aplicadas a problemas de clasificación, detección y predicción relacionados con la salud mental. Las investigaciones utilizaron datos de neuroimagen, señales fisiológicas, registros clínicos, información académica y plataformas digitales. Entre las arquitecturas identificadas destacaron las redes convolucionales de grafos (GCN), las redes de atención de grafos (GAT), GraphSAGE y los modelos híbridos. La tabla 2 presenta las características de los datos, los riesgos analizados, los modelos utilizados y los principales resultados de cada investigación.

Tabla 2
Trastornos de salud mental y arquitecturas GraphML

Estudio	Riesgo o trastorno	Arquitectura GraphML	Principal resultado reportado
Shi et al. (2025)	Ansiedad, depresión y estrés	PsyGraph-SSL	Superó los modelos comparativos en precisión, F1 y AUC.
Mo et al. (2025)	Ansiedad	ASF-MDGNC	Exactitud del 93,48 % y AUC del 94,58 %.
Patro et al. (2026)	Estrés	Att-MSR-GraphSAGE	Exactitud del 96,10 % y F1 del 94,21 %.
Chen et al. (2024)	Trastorno del espectro autista	DeepASD	Mejoras reportadas del 13,25 % en exactitud y del 7,69 % en AUC.
Liu (2024)	Depresión	WGRNN	F1-micro del 67,6 % y F1-macro del 65,2 %.
He et al. (2024)	Depresión	DPD Net	Coeficiente de concordancia de 0,682.
Li et al. (2023)	Autismo, depresión y trastorno bipolar	FSL-BrainNet	Mejoras de exactitud y AUC frente a los modelos comparativos.
Fu et al. (2025)	Depresión	GAT multimodal	MAE de 5,26 y CCC de 0,555.
Kanakala et al. (2025)	Diversos riesgos de salud mental	MindFusionNet	Exactitud del 98,68 %.
Priyadarshana et al. (2025)	Depresión y gravedad depresiva	GEM	Mejoras reportadas en F1 y clasificación de gravedad.
Liu et al. (2025)	Ideación suicida	GCN semántico	Superó varios clasificadores convencionales.
Kim y Sohn (2022)	Depresión en adultos mayores	GAT	Identificación de anomalías conductuales.
Chen et al. (2026)	Depresión mayor y recurrente	MAF-GNN	Mejoras reportadas del 2,8 % y 4,1 % en

			exactitud.
Wang et al. (2026)	Depresión ansiosa y no ansiosa	MAMF-GCN	Exactitud del 98,12 % y AUC de 0,9867.
Polowczyk et al. (2026)	Diversos trastornos mentales	Apriori-GAT y BERT	Mejor rendimiento global frente a los modelos individuales.
Kim et al. (2025)	Depresión mayor	MSGNN, GIN y GAT	Precisión promedio reportada de 0,882 ± 0,021.
Jena et al. (2026)	Depresión, ansiedad y otros riesgos	CNSGL	Exactitud de 0,925 y F1 de 0,923.
Liu et al. (2025)	Respuesta al tratamiento antidepresivo	LGCIF-GNN	AUC de 0,78 en validación interna y 0,72 en externa.
Ibitoye et al. (2025)	Adicción digital, ansiedad y depresión	GCN	Exactitud del 80,3 % y ROC-AUC de 0,89.
Yokotani et al. (2025)	Ansiedad social	GraphSAGE	F1 de 0,717.
Lin et al. (2025)	Crisis psicológica	GCN	Exactitud de 0,987; se reportaron dificultades de sensibilidad.
Du (2023)	Riesgos psicológicos estudiantiles	GCN espectral	Exactitud del 78,24 % y F1 de 0,5048.
Li et al. (2022)	Trastorno del espectro autista	TE-HI-GCN	Mejoras reportadas en exactitud y AUC.
Zhu et al. (2023)	Depresión mayor	DGCNN	Exactitud del 72,1 % y AUC de 0,788.
Cao et al. (2021)	Trastorno del espectro autista	DeepGCN	Exactitud del 73,7 %.
Zhou et al. (2025)	Depresión mayor	DGCN	Exactitud del 82,5 % y AUC de 0,869.
Kavitha y Shakkeera (2026)	Estrés fisiológico y psicológico	2D-CNN-GACL-ECGNet	F1 entre el 92,8 % y el 98,7 %, según el conjunto de datos.

Nota. Se presenta una síntesis de los resultados declarados. Para la versión definitiva deben conservarse las características de los datos y los tamaños muestrales, y verificarse las atribuciones bibliográficas de cada estudio.

Los resultados evidenciaron una presencia destacada de arquitecturas GCN y sus variantes, utilizadas principalmente para analizar conectividad cerebral, datos multimodales y relaciones entre características clínicas. Las arquitecturas GAT incorporaron mecanismos de atención para identificar relaciones relevantes, mientras que GraphSAGE permitió representar interacciones entre usuarios y estructuras de vecindad. También se identificaron modelos híbridos que combinaron grafos con transformadores, mecanismos de atención y técnicas de aprendizaje multimodal. Estas diferencias reflejaron la diversidad de enfoques utilizados para representar información relacional en problemas de salud mental.

En cuanto a los trastornos estudiados, la depresión y sus diferentes manifestaciones constituyeron el principal campo de aplicación de GraphML. También se identificaron investigaciones relacionadas con ansiedad, estrés, ideación suicida, autolesión, trastorno bipolar y trastornos del neurodesarrollo. Los modelos se aplicaron tanto a datos clínicos y neurobiológicos como a información conductual obtenida mediante plataformas digitales. Esta distribución mostró una concentración de investigaciones en problemas depresivos, acompañada de aplicaciones orientadas a otros riesgos psicológicos y psiquiátricos.

Métricas de desempeño y enfoques de validación

Las investigaciones utilizaron diferentes indicadores para evaluar el desempeño de las arquitecturas GraphML. Las métricas se organizaron en siete familias conceptuales, entre las que destacaron las medidas de clasificación, las curvas ROC y precisión-recall, los indicadores de regresión y las medidas de calibración. Según las frecuencias declaradas en el documento, las métricas de clasificación estuvieron presentes en el 63 % de los estudios y las medidas basadas en curvas ROC/PR en el 59 %. La tabla 3 resume las familias identificadas, sus principales indicadores y la frecuencia porcentual reportada.

Tabla 3*Métricas de desempeño*

Familia conceptual	Métricas específicas	Frecuencia reportada
Clasificación binaria y multiclase	Accuracy, precision, recall, F1-score, sensibilidad y especificidad	63 %
Curvas ROC/PR	AUC y AUPRC	59 %
Regresión y variables continuas	MAE, RMSE, CCC y R ²	15 %
Correlación y acuerdo	Kappa, MCC y balanced accuracy	11 %
Ranking y recuperación	Hits@k, MR, MRR y Top-N	11 %
Calibración	Brier score, ECE y log loss	7 %
Otros indicadores	G-mean, FPR y AVP	7 %

Nota. Se conservaron los porcentajes declarados originalmente. Antes de publicar, deben verificarse las frecuencias absolutas sobre los 27 estudios. Las categorías no son mutuamente excluyentes.

Las métricas de exactitud, precisión, sensibilidad y F1 constituyeron indicadores frecuentes para evaluar la capacidad de clasificación de los modelos. Por su parte, el AUC permitió examinar la discriminación entre categorías, mientras que los indicadores de regresión se utilizaron en estudios orientados a estimar variables continuas. Las medidas de calibración y otros indicadores especializados presentaron una menor frecuencia de uso. La diversidad de métricas dificultó la comparación directa entre modelos evaluados con poblaciones, conjuntos de datos y procedimientos metodológicos diferentes.

Las estrategias de validación incluyeron procedimientos de validación cruzada, particiones de entrenamiento y prueba, evaluación externa y experimentos complementarios de generalización. Estas técnicas se emplearon para estimar el desempeño predictivo de las arquitecturas y examinar su comportamiento frente a diferentes configuraciones de datos. También se identificaron investigaciones que utilizaron estudios de ablación, optimización de hiperparámetros y comparaciones con modelos de referencia. La tabla 4 presenta las estrategias identificadas y los estudios asociados a cada procedimiento.

Tabla 4*Estrategias de validación*

Estrategia	Características	Estudios reportados
Validación cruzada estratificada de 10 pliegues	Conservó las proporciones de las clases durante las particiones.	Chen et al. (2026); Wang et al. (2026); Kim et al. (2025); Cao et al. (2021); Zhou et al. (2025).
Validación cruzada de 10 pliegues	Utilizó diez particiones rotativas para entrenamiento y evaluación.	Li et al. (2023); Polowczyk et al. (2026); Lin et al. (2025); Zhu et al. (2023).
Validación cruzada de 5 pliegues	Aplicó cinco particiones rotativas.	He et al. (2024); Kanakala et al. (2025); Chen et al. (2026).
Entrenamiento, validación y prueba	Separó los datos para ajustar y evaluar los modelos.	Patro et al. (2026); He et al. (2024); Fu et al. (2025).
División entrenamiento-prueba	Empleó subconjuntos independientes para entrenamiento y evaluación.	Mo et al. (2025); Liu et al. (2025); Du (2023); Jena et al. (2026).
Validación interna y externa	Evaluó el desempeño con datos internos y externos.	Liu et al. (2025).
Leave-One-Site-Out	Evaluó la generalización utilizando centros excluidos del entrenamiento.	Kim et al. (2025); Zhou et al. (2025).
Generalización y ablación	Examinó el aporte de componentes y configuraciones del modelo.	Patro et al. (2026); Li et al. (2023); Kavitha y Shakkeera (2026).
Optimización de hiperparámetros	Ajustó parámetros y configuraciones del modelo.	Jena et al. (2026); Yokotani et al. (2025).
Evaluación comparativa	Contrastó los resultados con modelos de referencia.	Chen et al. (2026); Kim y Sohn (2022); Kavitha y Shakkeera (2026).
Sin validación claramente especificada	No se identificó una descripción suficiente del procedimiento.	Shi et al. (2025); Liu (2024); Ibitoye et al. (2025).

Nota. Se eliminaron referencias repetidas dentro de una misma categoría. Un estudio pudo utilizar varias estrategias, por lo que las categorías no fueron excluyentes. Las atribuciones deben comprobarse con los artículos originales.

Los resultados mostraron una variedad de procedimientos para estimar el desempeño de los modelos GraphML. La validación cruzada y las particiones independientes de entrenamiento y prueba se identificaron entre las estrategias utilizadas, mientras que algunos estudios incorporaron evaluaciones multicéntricas o conjuntos externos. Los experimentos de ablación permitieron examinar componentes específicos de las arquitecturas, aunque no sustituyeron la validación predictiva independiente. Estas diferencias metodológicas deben considerarse al interpretar la capacidad de generalización y la posible aplicabilidad clínica de los resultados.

Principales limitaciones para la aplicación de GraphML en salud mental

Las investigaciones incluidas reportaron limitaciones relacionadas con la disponibilidad de datos, la construcción de los grafos, la interpretabilidad de los modelos y los procedimientos de validación. Estas dificultades se agruparon en siete categorías temáticas para facilitar su análisis y comparación. Las limitaciones identificadas comprendieron aspectos metodológicos, tecnológicos y clínicos que condicionaron la interpretación del desempeño de las arquitecturas. La tabla 5 presenta cada categoría, sus principales características y sus implicaciones para la aplicación de GraphML en salud mental.

Tabla 5

Limitaciones de los modelos GraphML

Limitación	Descripción	Implicaciones
Generalización externa limitada	El desempeño obtenido en determinados conjuntos de datos no necesariamente se mantuvo en otras poblaciones.	Restringió la transferibilidad de los modelos a contextos clínicos diferentes.
Tamaño muestral y representatividad	Se identificaron muestras pequeñas, poblaciones específicas y desequilibrios entre clases.	Incrementaron el riesgo de sobreajuste y dificultaron la detección de casos minoritarios.
Dependencia de datos multimodales	Algunos modelos necesitaron integrar simultáneamente neuroimágenes, texto, señales fisiológicas o información conductual.	Las modalidades ausentes, incompletas o ruidosas pudieron afectar su funcionamiento.
Interpretabilidad limitada	Determinadas arquitecturas presentaron dificultades para explicar las variables y relaciones asociadas con sus predicciones.	Limitaron la comprensión y evaluación clínica de las decisiones automatizadas.
Complejidad computacional	Los modelos incorporaron múltiples capas, mecanismos de atención y procedimientos de fusión.	Incrementaron los requisitos computacionales y las necesidades de infraestructura.
Sensibilidad a la construcción del grafo	El rendimiento dependió de la definición de nodos, aristas, umbrales y relaciones de conectividad.	Introdujo variabilidad metodológica y dificultades de reproducción.
Limitaciones clínicas, éticas y regulatorias	Se identificaron restricciones relacionadas con la validación clínica y el tratamiento de información sensible.	Condicionaron las posibilidades de implementación en entornos asistenciales.

Nota. Las categorías se sintetizaron a partir de la información proporcionada. Las referencias correspondientes deben reincorporarse tras verificar su concordancia con los 27 estudios incluidos.

La generalización externa limitada y las características de los conjuntos de datos constituyeron obstáculos relevantes para interpretar los resultados de los modelos. Las muestras reducidas, el desequilibrio entre categorías y la heterogeneidad de las fuentes de información condicionaron las evaluaciones predictivas. Por otra parte, la construcción de grafos y la selección de relaciones entre nodos introdujeron diferencias metodológicas entre investigaciones. Estas condiciones dificultaron la comparación directa del desempeño y la extrapolación de resultados a poblaciones distintas de las utilizadas durante el desarrollo de los modelos.

Las limitaciones tecnológicas estuvieron relacionadas con la complejidad computacional, la integración de múltiples fuentes y las dificultades de interpretación de determinadas arquitecturas. En el ámbito clínico, las investigaciones identificaron desafíos relacionados con la validación

independiente, la explicación de las predicciones y la utilización de información sensible. Estos factores evidenciaron diferencias entre el rendimiento obtenido durante la experimentación y las condiciones necesarias para su aplicación asistencial. En conjunto, los resultados permitieron reconocer los principales aspectos metodológicos y técnicos que condicionaron el uso de GraphML en la detección temprana de riesgos en salud mental.

4. Discusión

La presente revisión sistemática permitió sintetizar y analizar la evidencia científica sobre la aplicación de Graph Machine Learning (GraphML) en la detección temprana de riesgos en salud mental. Los resultados evidenciaron un predominio de arquitecturas basadas en GCN, GAT y modelos híbridos multimodales, aplicados principalmente a la depresión, ansiedad, estrés e ideación suicida. Estas arquitecturas mostraron resultados predictivos favorables mediante indicadores como exactitud, F1-score y AUC, especialmente en investigaciones que integraron diferentes fuentes de información. Sin embargo, las diferencias entre poblaciones, conjuntos de datos y procedimientos de validación impidieron establecer la superioridad general de una arquitectura sobre las demás.

Los hallazgos coincidieron con investigaciones que destacaron el avance de las redes neuronales basadas en grafos en aplicaciones biomédicas y de salud mental. Vaida y Huang (2026) señalaron que estas arquitecturas favorecieron la integración de información clínica multimodal mediante la representación de relaciones complejas entre pacientes, síntomas, biomarcadores y factores contextuales. Esta perspectiva fue consistente con los estudios analizados, en los que la combinación de neuroimágenes, señales fisiológicas, textos y datos conductuales permitió representar diferentes dimensiones de los problemas de salud mental. En consecuencia, el principal aporte de GraphML se relacionó con su capacidad para integrar información heterogénea y modelar dependencias que podrían resultar difíciles de representar mediante enfoques convencionales.

En relación con los trastornos abordados, se identificó una concentración de investigaciones orientadas a la depresión, seguida de aplicaciones relacionadas con ansiedad, estrés y otros riesgos psicológicos. Fisher et al. (2026) destacaron las aplicaciones del procesamiento del lenguaje natural y el aprendizaje automático en la detección de indicadores asociados con la depresión. Por su parte, Pinge et al. (2024) identificaron el potencial de los dispositivos portátiles y sensores fisiológicos para la detección y el monitoreo del estrés. Estos antecedentes coincidieron con la diversidad de fuentes de información encontrada en la revisión y evidenciaron que las aplicaciones de GraphML se extendieron desde el análisis de señales fisiológicas hasta la identificación de patrones conductuales y lingüísticos relacionados con la salud mental.

El análisis de las métricas de desempeño evidenció una utilización frecuente de indicadores de clasificación, principalmente exactitud, sensibilidad, precisión, F1-score y AUC, lo que reflejó el predominio de modelos orientados a diferenciar categorías diagnósticas o niveles de riesgo. No obstante, los resultados obtenidos en diferentes investigaciones no fueron directamente comparables debido a las variaciones en el tamaño muestral, la distribución de clases y las características de los conjuntos de datos. La menor presencia de indicadores de calibración limitó la información disponible sobre la correspondencia entre las probabilidades predichas y los riesgos observados. Por tanto, la interpretación del rendimiento predictivo requirió considerar tanto la capacidad discriminativa de los modelos como las condiciones metodológicas utilizadas para evaluarlos.

Otro hallazgo relevante fue la incorporación de mecanismos de explicabilidad y modelado de relaciones causales en determinadas arquitecturas GraphML. Chan et al. (2025) identificaron

variabilidad en los biomarcadores neurofuncionales asociados con trastornos psiquiátricos, lo que evidenció desafíos para establecer patrones reproducibles mediante redes neuronales basadas en grafos. Por su parte, Moser et al. (2024) examinaron estrategias de aprendizaje profundo explicable aplicadas a la detección del estrés mediante sensores portátiles. Estos aportes permitieron reconocer que la interpretabilidad constituye una dimensión relevante del desarrollo de modelos predictivos, especialmente cuando sus resultados se utilizan para identificar factores asociados con alteraciones psicológicas o fundamentar decisiones clínicas.

En cuanto a los procedimientos de validación, los estudios incluidos utilizaron principalmente validación cruzada y particiones de entrenamiento y prueba, mientras que la evaluación con conjuntos de datos externos fue menos frecuente. Esta distribución evidenció que una parte importante del desempeño reportado procedió de evaluaciones internas, cuyas condiciones no necesariamente reprodujeron la variabilidad existente en contextos asistenciales. Aunque algunos modelos incorporaron validaciones multicéntricas y experimentos de generalización, persistieron diferencias en los criterios empleados para evaluar su estabilidad. En este sentido, la validación externa representó un elemento fundamental para interpretar la capacidad de transferencia de las arquitecturas GraphML entre poblaciones, instituciones y escenarios clínicos.

Los resultados también permitieron identificar limitaciones metodológicas y éticas relacionadas con la implementación de inteligencia artificial en salud mental. Vidal et al. (2025) señalaron que los sistemas de inteligencia artificial podían reproducir sesgos asociados con datos desequilibrados y poblaciones insuficientemente representadas, comprometiendo la equidad de sus predicciones. Por su parte, Ali et al. (2023) identificaron desafíos generales vinculados con la calidad de los datos, la transparencia y la implementación de herramientas de inteligencia artificial en el sector sanitario. Estas observaciones coincidieron con las limitaciones encontradas en la revisión, particularmente la generalización externa restringida, la dependencia de datos multimodales y la sensibilidad del desempeño a los procedimientos utilizados para construir los grafos.

La revisión evidenció que la construcción del grafo constituyó un componente metodológico determinante, debido a que la selección de nodos, aristas, umbrales de similitud y mecanismos de conectividad influyó en la representación de las relaciones analizadas. Estas decisiones condicionaron la información disponible para el aprendizaje y pudieron generar diferencias entre modelos desarrollados con fuentes de datos semejantes. De igual manera, las arquitecturas que integraron mecanismos de atención, múltiples capas o información multimodal presentaron mayores exigencias de procesamiento y configuración. Por consiguiente, la complejidad arquitectónica debió interpretarse conjuntamente con los beneficios predictivos reportados y las condiciones necesarias para reproducir los resultados.

La presente revisión presentó limitaciones que debieron considerarse al interpretar sus hallazgos. La búsqueda bibliográfica se restringió a Scopus y Web of Science, por lo que pudieron omitirse investigaciones disponibles en otras fuentes de información científica. La heterogeneidad de los diseños, poblaciones, arquitecturas y métricas dificultó la comparación directa de los resultados, mientras que la disponibilidad limitada de validaciones externas restringió las conclusiones sobre aplicabilidad clínica. Además, la información incompleta sobre calibración, procedimientos de validación y explicabilidad en determinados estudios condicionó la evaluación conjunta de su solidez metodológica.

A partir de los hallazgos identificados, futuras investigaciones podrían profundizar en el desarrollo y evaluación de grafos dinámicos capaces de representar cambios temporales en síntomas, comportamientos y factores psicosociales. También resultó pertinente fortalecer la validación externa, la transparencia en la construcción de grafos y la evaluación de indicadores clínicamente

interpretables. En aplicaciones relacionadas con ideación suicida, monitoreo conductual y análisis de información digital, se reconoció la importancia de considerar la privacidad, la protección de datos y los posibles sesgos algorítmicos. Estas líneas de investigación permitirían profundizar en la evaluación de GraphML como herramienta complementaria para la detección temprana y el apoyo a la toma de decisiones en salud mental.

5. Conclusiones

La presente revisión sistemática evidenció que las arquitecturas basadas en GCN y sus variantes fueron las más utilizadas para la detección temprana de riesgos en salud mental, seguidas de GAT, GraphSAGE y modelos híbridos multimodales. Estas tecnologías permitieron representar relaciones complejas entre variables clínicas, conductuales, fisiológicas y contextuales, favoreciendo la identificación de patrones asociados con diferentes condiciones psicológicas. Sus aplicaciones se concentraron principalmente en depresión, ansiedad, estrés e ideación suicida, con menor presencia de investigaciones relacionadas con trastornos del espectro autista, trastorno bipolar y aislamiento social.

Los modelos analizados presentaron resultados predictivos favorables, evaluados principalmente mediante exactitud, F1-score y AUC, complementados con sensibilidad, especificidad y, en menor proporción, indicadores de regresión y calibración. Las estrategias de evaluación más utilizadas comprendieron validación cruzada y divisiones de entrenamiento y prueba, mientras que la validación externa multicéntrica fue menos frecuente. Aunque los resultados evidenciaron el potencial de GraphML para identificar patrones relacionados con riesgos en salud mental, la heterogeneidad de las métricas, poblaciones y procedimientos de validación limitó la comparación directa entre arquitecturas y la interpretación de su aplicabilidad clínica.

Las principales limitaciones identificadas comprendieron la escasa validación externa, los tamaños muestrales reducidos, el desequilibrio entre clases, la dependencia de información multimodal y las dificultades de interpretabilidad de determinadas arquitecturas. Estos factores condicionaron la generalización de los resultados y evidenciaron diferencias entre el desempeño experimental y las exigencias de su aplicación en entornos asistenciales. En consecuencia, la consolidación de GraphML como herramienta complementaria para la detección temprana de riesgos en salud mental requiere fortalecer la validación prospectiva y multicéntrica, la explicabilidad de los modelos y la estandarización de procedimientos metodológicos reproducibles, considerando también la protección de datos y las implicaciones éticas de su utilización.

Referencias

- Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M., & Dwivedi, Y. (2023). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge*, 8(1), 100333. <https://doi.org/10.1016/j.jik.2023.100333>
- Boll, H., Amirahmadi, A., Ghazani, M., de Moraes, W., de Freitas, E., Soliman, A., Etminani, F., Byttner, S., & Recamonde-Mendoza, M. (2024). Graph neural networks for clinical risk prediction based on electronic health records: A survey. *Journal of Biomedical Informatics*, 151, 104616. <https://doi.org/10.1016/j.jbi.2024.104616>
- Bollen, J., ten Thij, M., Lorenzo-Luaces, L., & Rutter, L. (2022). Beyond risk: Individual mental health trajectories from large-scale social media data. In F. Crestani, D. E. Losada, & J. Parapar (Eds.), *Early detection of mental health disorders by social media monitoring: The first five years of the eRisk project* (pp. 265–287). Springer International Publishing. https://doi.org/10.1007/978-3-031-04431-1_12
- Cao, M., Yang, M., Qin, C., Zhu, X., Chen, Y., Wang, J., & Liu, T. (2021). Using DeepGCN to identify the autism spectrum disorder from multi-site resting-state data. *Biomedical Signal Processing and Control*, 70, 103015. <https://doi.org/10.1016/j.bspc.2021.103015>
- Chan, Y. H., Girish, D., Gupta, S., Xia, J., Kasi, C., He, Y., Wang, C., & Rajapakse, J. C. (2025). Discovering robust biomarkers of psychiatric disorders from resting-state functional MRI via graph neural networks: A systematic review. *NeuroImage*, 319, 121422. <https://doi.org/10.1016/j.neuroimage.2025.121422>
- Chen, W., Yang, J., Sun, Z., Zhang, X., Tao, G., Ding, Y., Gu, J., Bu, J., & Wang, H. (2024). DeepASD: A deep adversarial-regularized graph learning method for ASD diagnosis with multimodal data. *Translational Psychiatry*, 14(1), 375. <https://doi.org/10.1038/s41398-024-02972-2>
- Chen, Z., Wu, J., Pu, L., Li, S., Zhang, Y., Yao, D., & Yan, H. (2026). MAF-GNN: Graph neural network-based multi-atlas brain functional information fusion for major depressive disorder diagnosis with rs-fMRI. *Information Processing & Management*, 63(3), 104524. <https://doi.org/10.1016/j.ipm.2025.104524>
- Cripps, R., Hayes, J., Pitman, A., Osborn, D., & Werbeloff, N. (2020). Characteristics and risk of repeat suicidal ideation and self-harm in patients who present to emergency departments with suicidal ideation or self-harm: A prospective cohort study. *Journal of Affective Disorders*, 273, 358–363. <https://doi.org/10.1016/j.jad.2020.03.130>
- Cruz-Gonzalez, P., He, A., Lam, E., Ng, I., Li, M., Hou, R., Chan, J., Sahni, Y., Guasch, N., Miller, T., Lau, B., & Vidaña, D. (2025). Artificial intelligence in mental health care: A systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine*, 55, e18. <https://doi.org/10.1017/S0033291724003295>
- Dong, G., Tang, M., Cai, L., Barnes, L., & Boukhechba, M. (2021). Semi-supervised graph instance transformer for mental health inference. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1221–1228). IEEE. <https://doi.org/10.1109/ICMLA52953.2021.00198>
- Du, Z. (2024). Research on the synergistic nurturing mechanism of ideological work and student

- management in colleges and universities under the binary structure. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–18. <https://doi.org/10.2478/amns.2023.2.01485>
- Fisher, H., Jaffe, N., Pidvirny, K., Tierney, A., Vaidean, M., Dongre, P., & Webb, C. (2026). Language-based detection of depression with machine learning: Systematic review and meta-analysis. *npj Digital Medicine*, 9(1), 273. <https://doi.org/10.1038/s41746-026-02448-1>
- Fu, C., Qian, F., Su, Y., Su, K., Song, S., Niu, M., Shi, J., Liu, Z., Liu, C., Ishi, C., & Ishiguro, H. (2025). Facial action units guided graph representation learning for multimodal depression detection. *Neurocomputing*, 619, 129106. <https://doi.org/10.1016/j.neucom.2024.129106>
- He, M., Bakker, E., & Lew, M. (2024). DPD (DePression Detection) Net: A deep neural network for multimodal depression detection. *Health Information Science and Systems*, 12(1), 53. <https://doi.org/10.1007/s13755-024-00311-9>
- Ibitoye, A., Ravindran, S., Afe, O., & Abiodun, A. (2025). Predicting digital addiction patterns with machine learning for personalised mental health support. *Journal of Social Computing*, 6(4), 359–377. <https://doi.org/10.23919/JSC.2025.0020>
- Jena, M., Khan, N., Lee, M., & Rho, S. (2026). Neuro-symbolic graph learning for causal inference and continual learning in mental-health risk assessment. *Computer Modeling in Engineering & Sciences*, 146(1), 44. <https://doi.org/10.32604/cmescs.2025.075119>
- Jiang, D. (2026). Predicting student mental health through entropy-based features and interpretable cross-attention transformer networks. *PLOS ONE*, 21(4), e0347294. <https://doi.org/10.1371/journal.pone.0347294>
- Kanakala, S., Prashanthi, V., Chinnaiah, V., Sharada, K., & Sumalatha, M. (2025). FMindMonitorAI: An AI-driven framework for social media-based mental health analysis using MindFusionNet. *International Research Journal of Multidisciplinary Scope*, 6(3), 1587–1610. <https://doi.org/10.47857/irjms.2025.v06i03.04620>
- Kavitha, P., & Shakkeera, L. (2026). 2D-CNN-GACL-ECGNet graph attention: A robust framework for electrocardiogram-based stress detection. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 15(2), 1529–1538. <https://doi.org/10.11591/ijai.v15.i2.pp1529-1538>
- Kim, J., & Sohn, M. (2022). Graph representation learning-based early depression detection framework in smart home environments. *Sensors*, 22(4), 1545. <https://doi.org/10.3390/s22041545>
- Kim, S., Bong, S., Yun, S., Kim, D., Yoo, J., Choi, K., Park, H., Jeon, H., Kim, J., Jang, J., & Jeong, B. (2025). Neurobiologically interpretable causal connectome for predicting young adult depression: A graph neural network study. *Journal of Affective Disorders*, 377, 225–234. <https://doi.org/10.1016/j.jad.2025.02.076>
- Knipe, D., Padmanathan, P., Newton-Howes, G., Chan, L., & Kapur, N. (2022). Suicide and self-harm. *The Lancet*, 399(10338), 1903–1916. [https://doi.org/10.1016/S0140-6736\(22\)00173-8](https://doi.org/10.1016/S0140-6736(22)00173-8)
- Li, L., Jiang, H., Wen, G., Cao, P., Xu, M., Liu, X., Yang, J., & Zaiane, O. (2022). TE-HI-GCN: An ensemble of transfer hierarchical graph convolutional networks for disorder diagnosis. *Neuroinformatics*, 20(2), 353–375. <https://doi.org/10.1007/s12021-021-09548-1>
- Li, L., Wen, G., Cao, P., Liu, X., Zaiane, O. R., & Yang, J. (2023). Exploring interpretable graph convolutional networks for autism spectrum disorder diagnosis. *International Journal of Computer Assisted Radiology and Surgery*, 18(4), 663–673. <https://doi.org/10.1007/s11548-022->

02780-3

- Lin, J., Tian, J., Wang, T., Li, D., & Jiang, X. (2025). Psychological crisis detection based on behavioral data: A new approach to non-contact measurement. *BMC Psychology*, 13(1), 1355. <https://doi.org/10.1186/s40359-025-03604-0>
- Liu, R., Hou, X., Liu, S., Zhou, Y., Zhou, J., Qiao, K., Qi, H., Li, R., Yang, Z., Zhang, L., Cui, J., Jin, C., Yu, A., & Wang, G. (2025). Predicting antidepressant response via local-global graph neural network and neuroimaging biomarkers. *npj Digital Medicine*, 8(1), 515. <https://doi.org/10.1038/s41746-025-01912-8>
- Liu, Y. (2024). Depression detection via a Chinese social media platform: A novel causal relation-aware deep learning approach. *The Journal of Supercomputing*, 80(8), 10327–10356. <https://doi.org/10.1007/s11227-023-05830-y>
- Liu, Y., Zheng, J., Zeng, Y., Luo, F., & Tian, X. (2025). Graph learning based suicidal ideation detection via tree-drawing test. *Frontiers in Psychiatry*, 16, 1617650. <https://doi.org/10.3389/fpsy.2025.1617650>
- Mansoor, M., & Ansari, K. (2024). Early detection of mental health crises through artificial-intelligence-powered social media analysis: A prospective observational study. *Journal of Personalized Medicine*, 14(9), 958. <https://doi.org/10.3390/jpm14090958>
- Mo, H., Wu, H., Rong, Q., Hu, Z., Yi, M., & Chen, P. (2025). An anxiety screening framework integrating multimodal data and graph node correlation. *Artificial Intelligence in Medicine*, 167, 103189. <https://doi.org/10.1016/j.artmed.2025.103189>
- Moser, M., Ehrhart, M., & Resch, B. (2024). An explainable deep learning approach for stress detection in wearable sensor measurements. *Sensors*, 24(16), 5085. <https://doi.org/10.3390/s24165085>
- Nordin, N., Zainol, Z., Mohd Noor, M., & Chan, L. (2022). Suicidal behaviour prediction models using machine learning techniques: A systematic review. *Artificial Intelligence in Medicine*, 132, 102395. <https://doi.org/10.1016/j.artmed.2022.102395>
- Okesanya, O., Oso, T., Hassan, M., Adebayo, U., Vandy, A., Othman, Z., Ali, I., Ahmed, M., Musa, S., Eshun, G., & Lucero-Prisno, D., III. (2026). Applications of artificial intelligence and machine learning in the diagnosis and management of major depressive disorder: A systematic review. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery*, 62(1), 89. <https://doi.org/10.1186/s41983-026-01156-7>
- Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Park, Y., Park, S., & Lee, M. (2024). Effectiveness of artificial intelligence in detecting and managing depressive disorders: Systematic review. *Journal of Affective Disorders*, 361, 445–456. <https://doi.org/10.1016/j.jad.2024.06.035>
- Patro, K., Routray, S., Madan Kumar, C., Kranthi Kumar, L., Rao, M., & Kottala, R. (2026). Att-MSR-GraphSAGE: An attention-guided multi-scale residual graph neural network for human-centric stress-factor modeling. *IEEE Access*, 14, 37435–37450. <https://doi.org/10.1109/ACCESS.2026.3671315>

- Paul, S., Saha, A., Hasan, M., Noori, S., & Moustafa, A. (2024). A systematic review of graph neural network in healthcare-based applications: Recent advances, trends, and future directions. *IEEE Access*, 12, 15145–15170. <https://doi.org/10.1109/ACCESS.2024.3354809>
- Pigoni, A., Delvecchio, G., Turtulici, N., Madonna, D., Pietrini, P., Cecchetti, L., & Brambilla, P. (2024). Machine learning and the prediction of suicide in psychiatric populations: A systematic review. *Translational Psychiatry*, 14(1), 140. <https://doi.org/10.1038/s41398-024-02852-9>
- Pinge, A., Gad, V., Jaisighani, D., Ghosh, S., & Sen, S. (2024). Detection and monitoring of stress using wearables: A systematic review. *Frontiers in Computer Science*, 6, 1478851. <https://doi.org/10.3389/fcomp.2024.1478851>
- Polowczyk, A., Polowczyk, A., & Woźniak, M. (2026). Mental illness text classification based on late fusion combining BERT with Apriori-informed graph attention network. *Computer Standards & Interfaces*, 98, 104157. <https://doi.org/10.1016/j.csi.2026.104157>
- Priyadarshana, Y., Liang, Z., & Piumarta, I. (2025). GEM: Graph attention encoder for multi-task depression severity detection in multi-party conversations. *Social Network Analysis and Mining*, 15(1), 76. <https://doi.org/10.1007/s13278-025-01495-y>
- Razavi, M., Ziyadidegan, S., Mahmoudzadeh, A., Kazeminasab, S., Baharlouei, E., Janfaza, V., Jahromi, R., & Sasangohar, F. (2024). Machine learning, deep learning, and data preprocessing techniques for detecting, predicting, and monitoring stress and stress-related mental disorders: Scoping review. *JMIR Mental Health*, 11, e53714. <https://doi.org/10.2196/53714>
- Ruan, Y., Liang, H., Yamamoto, S., & Lin, S. (2026). Machine learning in mental health promotion for older adults: A scoping review. *BMC Geriatrics*, 26, 821. <https://doi.org/10.1186/s12877-026-07543-2>
- Shi, X., Pan, J., Yuan, D., Li, M., & Pan, Y. (2025). Advanced data analysis and prediction model for student mental health risk assessment. *Frontiers in Psychology*, 16, 1682083. <https://doi.org/10.3389/fpsyg.2025.1682083>
- Vaida, M., & Huang, Z. (2026). Multimodal graph neural networks in healthcare: A review of fusion strategies across biomedical domains. *Frontiers in Artificial Intelligence*, 8, 1716706. <https://doi.org/10.3389/frai.2025.1716706>
- Vidal, J., Iglesias, C., & Montenegro, I. (2025). Sesgos en la inteligencia artificial en el sector salud: Una revisión sistemática en el contexto iberoamericano. *Revista Iberoamericana de la Propiedad Intelectual*, (22), 123–154. <https://doi.org/10.26422/RIPI.2025.2200.bar>
- Wang, S., Zhang, J., Zhang, R., Li, Z., Zhang, C., Zhao, Z., Pang, T., Gao, Z., Wang, Z., Yu, Y., Ren, W., & Wang, C. (2026). MAMF-GCN model for anxious and non-anxious depression classification and neuroimaging marker recognition. *International Journal of Psychophysiology*, 219, 113301. <https://doi.org/10.1016/j.ijpsycho.2025.113301>
- Weber, M. (2025). Geometric machine learning. *AI Magazine*, 46(1), e12210. <https://doi.org/10.1002/aaai.12210>
- Yokotani, K., Takano, M., Abe, N., & Kato, T. (2025). Predicting social anxiety disorder based on communication logs and social network data from a massively multiplayer online game: Using a graph neural network. *Psychiatry and Clinical Neurosciences*, 79(5), 274–281. <https://doi.org/10.1111/pcn.13804>

- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81. <https://doi.org/10.1016/j.aiopen.2021.01.001>
- Zhou, N., Yuan, Z., Zhou, H., Lyu, D., Wang, F., Wang, M., Lu, Z., Huang, Q., Chen, Y., Huang, H., Cao, T., Wu, C., Yang, W., & Hong, W. (2025). Using dynamic graph convolutional network to identify individuals with major depression disorder. *Journal of Affective Disorders*, 371, 188–195. <https://doi.org/10.1016/j.jad.2024.11.035>
- Zhu, M., Quan, Y., & He, X. (2023). The classification of brain network for major depressive disorder patients based on deep graph convolutional neural network. *Frontiers in Human Neuroscience*, 17, 1094592. <https://doi.org/10.3389/fnhum.2023.1094592>

Transparencia

Conflicto de interés

Los autores declaran que no existen conflictos de interés de naturaleza alguna como parte de la presente investigación.

Fuente de financiamiento

Los autores financiaron completamente la investigación.

Contribución de autoría

Jorge Edwin Delgado Altamirano: Conceptualización, metodología, software, análisis formal, investigación, gestión de datos, visualización, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, recursos,

Ciro Rodriguez Rodriguez: Validación, análisis formal, investigación, supervisión.

Los autores contribuyeron activamente en el análisis de los resultados, revisión y aprobación del manuscrito final.