

# Uso del Procesamiento de Lenguaje Natural en la Salud Mental

Revisi3n de Alcance

Use of Natural Language Processing in Mental Health: Scoping Review.  
Uso do Processamento de Linguagem Natural em Sa3de Mental: Revis3o de Escopo



Reyk Sayk **Alem3n Acu3a**  
Eider **Pereira Montiel**  
Ever Augusto **Torres Silva**  
David Andr3s **Montoya Arenas**



# 2024

Revista Iberoamericana de

# Psicolog3a

ISSN-I: 2027-1786 | e-ISSN: 2500-6517

Publicaci3n Cuatrimestral

izkes

Photo By/Foto:

**Rip**  
**17<sup>2</sup>**

Volumen 17 #2 may-ago  
| 17 A3os



Planeta Formaci3n y Universidades

ID: [10.33881/2027-1786.rip.17202](https://doi.org/10.33881/2027-1786.rip.17202)

Title: Use of Natural Language Processing in Mental Health

Subtitle: scoping review.

Título: Uso del Procesamiento de lenguaje natural en la Salud Mental

Subtítulo: Revisión de alcance.

Título: Uso do Processamento de Linguagem Natural em Saúde Mental

Subtítulo: Revisão de Escopo.

Alt Title / Título alternativo:

[en]: Use of Natural Language Processing in Mental Health: Scoping Review.

[es]: Uso del Procesamiento de Lenguaje Natural en la Salud Mental: Revisión de Alcance.

[pt]: Uso do Processamento de Linguagem Natural em Saúde Mental: Revisão de Escopo.

Author (s) / Autor (es):

Alemán Acuña, Pereira Montiel, Torres Silva & Montoya Arenas

Keywords / Palabras Clave:

[en]: Anxiety, Depression, Grief, Natural language processing, Mental health, Mental disorders.

[es]: Ansiedad, Depresión, Duelo, Procesamiento de lenguaje natural, Salud mental, Trastornos mentales

[pt]: Ansiedade, Depressão, Luto, Processamento de linguagem natural, Saúde mental, Transtornos mentais.

Financiación / Funding:

Convocatoria del Fondo CTEI del Sistema General de Regalías para el Fortalecimiento en Capacidades de Investigación y Desarrollo Regionales e iniciativas de Desarrollo y Transferencia de Tecnología y Conocimiento para la Innovación, orientadas a atender problemáticas derivadas del COVID 19. BPIN 2020000100726

Submitted: 2023-06-23

Accepted: 2023-10-28

## Resumen.

El propósito de esta revisión es analizar el uso del procesamiento de lenguaje natural en la investigación de trastornos mentales en adultos, como la depresión, la ansiedad y el duelo. Se realizó una búsqueda en cuatro bases de datos relevantes (PubMed, IEEE, ScienceDirect y LILACS) para identificar estudios publicados en español e inglés desde 2017 hasta 2022, sin restricciones en cuanto al país de origen. Se emplearon términos MeSH y de texto libre para identificar investigaciones sobre la aplicación del procesamiento de lenguaje natural en la detección de condiciones de salud mental, como la ansiedad, la depresión y el duelo. En total, se encontraron 136 estudios relacionados, de los cuales se seleccionaron 32 para esta revisión. Se observó un aumento en el uso del procesamiento de lenguaje natural en salud pública, especialmente entre los años 2020 y 2022. Además, se notó que las redes sociales son una fuente de datos frecuentemente utilizada en estos estudios, y que los modelos de aprendizaje automático supervisados son los más comunes para la detección de depresión y ansiedad. El procesamiento de lenguaje natural puede mejorar la detección de problemas de salud mental en la salud pública. Aunque los métodos de aprendizaje supervisado son los más comunes, los algoritmos basados en aprendizaje profundo presentan perspectivas innovadoras, y se espera que esta tecnología continúe creciendo para mejorar la detección y el tratamiento de trastornos mentales. Es crucial seguir investigando y desarrollando estas tecnologías para su aplicación en salud pública.

## Abstract.

The purpose of this review is to analyze the use of natural language processing in the research of mental disorders in adults, such as depression, anxiety, and grief. A search was conducted in four relevant databases (PubMed, IEEE, ScienceDirect, and LILACS) to identify studies published in Spanish and English from 2017 to 2022, with no restrictions on the country of origin. MeSH terms and free-text terms were used to identify research on the application of natural language processing in the detection of mental health conditions, such as anxiety, depression, and grief. In total, 136 related studies were found, of which 32 were selected for this review. An increase in the use of natural language processing in public health was observed, especially between the years 2020 and 2022. Furthermore, it was noted that social media is a frequently used source of data in these studies, and supervised machine learning models are the most common for detecting depression and anxiety. Natural language processing can enhance the detection of mental health issues in public health. While supervised learning methods are the most common, deep learning-based algorithms offer innovative perspectives, and it is expected that this technology will continue to grow to improve the detection and treatment of mental disorders. It is crucial to continue researching and developing these technologies for their application in public health.

## Resumo.

O propósito desta revisão é analisar o uso do processamento de linguagem natural na pesquisa de distúrbios mentais em adultos, como depressão, ansiedade e luto. Foi realizada uma busca em quatro bases de dados relevantes (PubMed, IEEE, ScienceDirect e LILACS) para identificar estudos publicados em espanhol e inglês de 2017 a 2022, sem restrições quanto ao país de origem. Foram utilizados termos MeSH e de texto livre para identificar pesquisas sobre a aplicação do processamento de linguagem natural na detecção de condições de saúde mental, como ansiedade, depressão e luto. No total, foram encontrados 136 estudos relacionados, dos quais 32 foram selecionados para esta revisão. Observou-se um aumento no uso do processamento de linguagem natural na saúde pública, especialmente entre os anos de 2020 e 2022. Além disso, observou-se que as redes sociais são uma fonte de dados frequentemente utilizada nesses estudos e que os modelos de aprendizado de máquina supervisionados são os mais comuns na detecção de depressão e ansiedade. O processamento de linguagem natural pode melhorar a detecção de problemas de saúde mental na saúde pública. Embora os métodos de aprendizado supervisionado sejam os mais comuns, os algoritmos baseados em aprendizado profundo apresentam perspectivas inovadoras, e espera-se que essa tecnologia continue a crescer para melhorar a detecção e o tratamento de distúrbios mentais. É crucial continuar pesquisando e desenvolvendo essas tecnologias para sua aplicação na saúde pública da população e melhorar a adesão às medidas preventivas.

## Citar como:

Alemán Acuña, R. S., Pereira Montiel, E. ., Torres Silva, E. A. & Montoya Arenas, D. A. (2024). Uso del Procesamiento de lenguaje natural en la Salud Mental: Revisión de alcance. Revista Iberoamericana de Psicología, 17 (2), 11-22. Obtenido de: <https://reviberopsicologia.iberu.edu.co/article/view/2731>

Reyk Sayk Alemán Acuña, [MA] Ele  
ORCID: [0000-0003-4789-8498](https://orcid.org/0000-0003-4789-8498)

Source | Filiación:  
Universidad Pontificia Bolivariana

BIO:  
Tecnólogo en Electrónica Industrial.  
Ingeniero Electrónico. Estudiante de  
Maestría en Ingeniería.

City | Ciudad:  
Montería [co]

e-mail:  
[reyk.aleman@upb.edu.co](mailto:reyk.aleman@upb.edu.co)

Eider Pereira Montiel, [MA] Bio  
ORCID: [0000-0002-5959-3870](https://orcid.org/0000-0002-5959-3870)

Source | Filiación:  
Universidad Pontificia Bolivariana,  
Facultad de Medicina, Escuela  
de Ciencias de la Salud, Medellín,  
Antioquia, Colombia

BIO:  
Ingeniero Biomédico. Estudiante de  
Maestría en Ingeniería Biomédica.

City | Ciudad:  
Medellín [co]

e-mail:  
[eider.pereira@upb.edu.co](mailto:eider.pereira@upb.edu.co)

Ever Augusto Torres Silva, MA Bio  
ORCID: [0000-0002-6302-6131](https://orcid.org/0000-0002-6302-6131)

Source | Filiación:  
Universidad Pontificia Bolivariana,  
Facultad de Medicina, Escuela  
de Ciencias de la Salud, Medellín,  
Antioquia, Colombia

BIO:  
Ingeniero Biomédico. Magíster en  
Gestión Tecnológica. Coordinador de  
Proyectos de Innovación en Netux SAS.  
Docente universitario y consultor.

City | Ciudad:  
Medellín [co]

e-mail:  
[ever.torres@upb.edu.co](mailto:ever.torres@upb.edu.co)

Dr David Andrés Montoya Arenas,  
Dr MA Psi  
ORCID: [0000-0001-6647-4696](https://orcid.org/0000-0001-6647-4696)

Source | Filiación:  
Universidad Pontificia Bolivariana

BIO:  
Psicólogo. Magíster en neuropsicología.  
Doctor en psicología con orientación  
en neurociencia cognitiva aplicada.  
Docente interno facultad de psicología.  
Investigador Asociado Minciencias

City | Ciudad:  
Medellín [co]

e-mail:  
[david.montoyaarenas@upb.edu.co](mailto:david.montoyaarenas@upb.edu.co)

# Uso del Procesamiento de Lenguaje Natural en la Salud Mental

## Revisión de Alcance

Use of Natural Language Processing in Mental Health: Scoping Review  
Uso do Processamento de Linguagem Natural em Saúde Mental: Revisão de Escopo

Reyk Sayk **Alemán Acuña**

Eider **Pereira Montiel**

Ever Augusto **Torres Silva**

David Andrés **Montoya Arenas**

## Introducción

Actualmente, los problemas psicológicos, como la depresión, la ansiedad y los trastornos mentales relacionados con el duelo, representan una carga significativa para la salud pública y pueden tener un impacto negativo en la calidad de vida de las personas (**OPS, 2022a**). Según la Organización Mundial de la Salud (OMS), se estima que 1 de cada 4 personas en todo el mundo experimentará algún tipo de trastorno mental a lo largo de sus vidas, lo cual incluye trastornos como la depresión, la ansiedad, de la personalidad, entre otros (**OMS, 2022**).

De acuerdo con un informe elaborado por la Organización de las Naciones Unidas (**ONU, 2022**), se observó un aumento del 27% en la prevalencia de trastornos depresivos y de ansiedad a nivel mundial entre los años 2019 y 2022 debido a la aparición de la pandemia del COVID-19. El confinamiento ha sido uno de los factores que ha conllevado varios desafíos para la salud mental, incluyendo el impacto económico, la preocupación por contagiar a familiares, conflictos con los seres queridos y cambios en las rutinas laborales habituales (**Kuliukas et al., 2021; Mahalingasivam et al., 2021; OPS, 2022b**).

Además, investigaciones recientes han comprobado que el coronavirus (COVID-19) y su familia de virus tienen una notable tendencia a afectar el sistema nervioso central (Carod Artal, 2020). Esta afectación del sistema nervioso central se ha relacionado con la aparición de diversos síntomas neurológicos, neuropsiquiátricos y neuropsicológicos en los pacientes, que incluyen manifestaciones depresivas y de ansiedad (Hermoso Contreras et al., 2022; Simón Martínez, 2022). Se cree que estos síntomas pueden persistir a largo plazo y afectar la capacidad funcional de los individuos, lo que podría ser uno de los posibles factores del aumento de estas condiciones de salud mental post-COVID-19 (Elixmahir Dávila-Marrero et al., 2021; Valdivieso Jimenez, 2021).

Otro aspecto relevante para considerar en el aumento de las condiciones de salud mental durante la pandemia del COVID-19 son las variables sociodemográficas y sociocognitivas de la población afectada. Iturbide Fernández (2021), a través de un estudio transversal realizado en la población adulta mexicana, se comparó estas variables con las condiciones de salud mental relacionadas con el COVID-19. Los resultados revelaron una correlación negativa del 11.7% con la resiliencia a la ansiedad y del 23.7% con la depresión, lo que pone de manifiesto la importancia de comprender cómo factores como el contexto social y las percepciones cognitivas pueden influir en la salud mental de las personas durante situaciones de crisis (Ávila-Toscano et al., 2022).

Los trastornos mentales, que se manifiestan en diversas formas, como la persistente sensación de tristeza y apatía característica de la depresión, así como la respuesta excesiva al estrés y las preocupaciones constantes asociadas a la ansiedad, ejercen un profundo impacto en el bienestar emocional y el funcionamiento diario de las personas. Estos trastornos pueden dejar huellas o signos característicos en diversos tipos de registros, desde notas clínicas hasta publicaciones en redes sociales (Mayo Clinic, 2021; Medlineplus, 2018; NIMH, 2021).

En este contexto, el Procesamiento de Lenguaje Natural (NLP, por sus siglas en inglés) emerge como una herramienta invaluable para analizar el lenguaje y el discurso escrito de las personas, lo que permite la detección temprana, el diagnóstico y la provisión de apoyo terapéutico en el ámbito de la salud mental (Yusufov et al., 2022). La contribución del NLP en la investigación y tratamiento de dichos trastornos es de singular relevancia, ya que favorece una atención más precisa y oportuna para aquellos que se ven afectados por estas condiciones. (DeSouza et al., 2021; Preoțiu-Pietro et al., 2015).

En la actualidad, el NLP es una de las técnicas más utilizadas para el análisis de datos no estructurados en el procesamiento de datos clínicos (Mehta & Pandit, 2018). El NLP, como rama de la inteligencia artificial, se enfoca en permitir que las máquinas comprendan y procesen el lenguaje humano (Enzyme, 2020; Hays DG, 1967; IBM, 2021; Lior Rokach & Oded Maimon, 2008). Al aplicar esta técnica al análisis de datos clínicos, se logra una mayor eficiencia y rapidez en el procesamiento de grandes cantidades de información textual, como registros médicos y notas clínicas. Esto, a su vez, permite a los profesionales de la salud tomar decisiones más informadas y precisas en lo que respecta al diagnóstico y tratamiento de los pacientes (Mishra & Garg, 2018; The MathWorks, 2020).

A diferencia de los enfoques convencionales empleados por psicólogos durante un largo período, como la entrevista clínica, la observación del comportamiento y la aplicación de pruebas psicológicas estandarizadas, como el Inventario de Depresión de Beck y la Escala de Depresión de Hamilton, para diagnosticar la depresión y la ansiedad en los pacientes, estos métodos se fundamentan en la valoración subjetiva del paciente y en la interpretación del profesional de la salud mental

(Benitez Molina & Caballero Badillo, 2017; Moral de la Rubia, 2013). La entrevista clínica se convierte en una herramienta crucial para recopilar información sobre los síntomas del paciente, su historial médico y sus antecedentes familiares.

De igual manera, la observación del comportamiento resulta valiosa para identificar señales de depresión y ansiedad, como la pérdida de interés en actividades cotidianas, la disminución de energía y la expresión de tristeza (Chávez-Negrete et al., 2021). Además, las pruebas psicológicas estandarizadas, como el Inventario de Depresión de Beck y la Escala de Depresión de Hamilton, representan herramientas efectivas para evaluar la intensidad de los síntomas de depresión y ansiedad (Purriños, 2013; Sanz et al., 2003).

Sin embargo, estos métodos presentan restricciones, ya que su eficacia depende de la habilidad del paciente para comunicar sus síntomas con precisión y de la capacidad del profesional de la salud para interpretarlos adecuadamente. Además, su implementación puede ser costosa y requerir una inversión considerable en tiempo y recursos tanto en su aplicación como en su análisis (Ibáñez & Echeburúa, 2015).

En contraste, los enfoques basados en el NLP se centran en el análisis de patrones lingüísticos y emocionales presentes en el discurso y la escritura del paciente, que se encuentren presentes en el texto escrito por el paciente, con el objetivo de detectar señales indicativas de depresión y ansiedad (Córdoba Cortés, 2020).

Además, en el campo de la salud mental, los modelos basados en NLP son herramientas importantes debido a que muchos síntomas de problemas de salud mental, como la depresión y la ansiedad, se reflejan en el lenguaje escrito no estructurado proporcionado por los pacientes (Cohen et al., 2014). A través de este lenguaje, los pacientes pueden brindar información relevante sobre sus emociones y su estado mental, incluyendo el duelo u otras condiciones emocionales. Esto hace que los modelos basados en NLP sean útiles para identificar patrones y comportamientos que pueden ayudar en el diagnóstico e identificación de estas condiciones de salud mental (Saffar et al., 2023).

De acuerdo con lo anterior, el lenguaje se convierte en una pieza clave para el estudio del comportamiento de los pacientes en diversas condiciones de salud mental, y el Procesamiento de Lenguaje Natural (NLP) utiliza los recursos lingüísticos disponibles para analizar el texto de manera cualitativa y cuantitativa, proporcionando una visión más profunda de los datos (Le Glaz et al., 2021). Los métodos de NLP pueden centrarse en diversas características, como las opciones léxicas, la sintaxis y la semántica, para llevar a cabo tareas como el modelado de temas, la agrupación en clústeres y la clasificación (Sharma et al., 2022).

Este tipo de herramientas puede contribuir a apoyar el proceso diagnóstico que realiza el personal de salud. Al permitir un análisis más detallado y preciso de la información lingüística de los pacientes, el NLP puede ayudar a los profesionales de la salud mental en la identificación temprana de trastornos mentales y proporcionar un apoyo valioso en el proceso de diagnóstico.

En este sentido, el objetivo de este artículo es analizar el uso del procesamiento del lenguaje natural en el estudio de trastornos mentales en la población adulta, incluyendo depresión, ansiedad y duelo. Se revisan investigaciones de los últimos cinco años relacionadas con la detección e identificación de estos problemas psicológicos. En última instancia, se busca identificar los avances más destacados en el uso del Procesamiento de Lenguaje Natural en este campo de investigación y proporcionar una visión general de las tendencias emergentes para orientar futuras investigaciones en el ámbito de la salud mental.

Además, se examinan los diferentes tipos de técnicas de procesamiento del lenguaje natural utilizadas para identificar y caracterizar estos trastornos mentales. También se consideran otros factores importantes, como los diversos tipos de datos utilizados en el análisis, como textos de redes sociales, notas clínicas, registros electrónicos de salud, entre otros. Del mismo modo, se analiza el tipo de estructura de los datos y el modelo de clasificación utilizado, todo ello con el fin de comprender las tendencias en el uso de NLP en el campo de la salud mental.

## Método

En esta sección, se seleccionaron y analizaron los artículos siguiendo las pautas de la declaración **PRISMA-ScR** (Tricco et al., 2018). La revisión sigue los 3 pasos fundamentales para este tipo de revisiones (Peters et al., 2015): (i) planificación de la revisión; (ii) identificación de la literatura relevante y selección; (iii) realización de la revisión y registro de los resultados.

## Planificación de la Revisión: Criterios de Elegibilidad

Los estudios que se relacionan con el uso del procesamiento del NLP en la salud mental abarcan un amplio campo de investigación. Por consiguiente, se establecieron criterios específicos para seleccionar los artículos relevantes para este estudio.

### Criterios de selección:

1. Artículos publicados en los últimos 5 años (2017- 2022).
2. Artículos que se centren en el uso de procesamiento de lenguaje natural en la salud mental.
3. Artículos que aborden temas relacionados con la ansiedad, la depresión, el duelo o los síntomas depresivos.
4. Artículos que se basen en investigaciones cualitativas o cuantitativas.
5. Artículos que estén disponibles en español o inglés.

### Artículos excluidos:

6. Artículos centrados en poblaciones con edades menores a 18 años.
7. Artículos sin disponibilidad de texto completo debido a la falta de licenciamiento institucional.
8. Publicaciones como revisiones, conferencias, poster, preprint, cartas al editor, capítulos de libros y editoriales.

## Identificación y Selección de la Literatura Relevante

Se llevó a cabo una búsqueda exhaustiva de la literatura utilizando cuatro bases de datos: Pubmed, IEEE, ScienceDirect y LILACS. La búsqueda se realizó el 18/11/2022, abarcando publicaciones desde el 1/01/2017 y sin restricciones de país de origen. Los resultados de la búsqueda se importaron a Rayyan, una herramienta en línea que facilita la revisión sistemática y metaanálisis de forma eficiente y **colaborativa** (Ouzzani et al., 2016). Los términos MeSH utilizados incluyeron “natural language processing”, “anxiety”, “depression”, “depressive Symp-

toms” y “mourning” para identificar estudios relacionados con la implementación de NLP en la detección de estas condiciones de salud mental. Además, se incluyeron los términos de texto libre como “text analysis” y “text Mining”, comunes en estudios relacionados con el NLP en la literatura científica. Como resultado, se obtuvieron las siguientes ecuaciones de búsqueda:

**Pubmed:** (“Natural Language Processing”[MeSH Terms] OR “Text Analysis”[Title/Abstract] OR “Text Mining”[Title/Abstract]) AND (Anxiety[MeSH Terms] OR Depression[MeSH Terms] OR Mourning[MeSH Terms] OR “Depressive Symptoms”[Title/Abstract])

**IEEE:** (“Mesh\_Terms”:Natural Language Processing” OR “Full Text Only”:”Text Mining” OR “Full Text Only”:”Text Analysis”) AND (“Mesh\_Terms”: Anxiety OR “Mesh\_Terms”:Depression OR “Full Text Only”:”Depressive Symptoms” OR “Mesh\_Terms”:Mourning)

**ScienceDirect:** (“Natural Language Processing” OR “Text Analysis” OR “Text Mining”) AND (Anxiety OR Depression OR Mourning OR “Depressive Symptoms” )

**LILACS:** (mh:(“Procesamiento de Lenguaje Natural” OR mh:(“Análisis de texto”) OR “ mh:Minería de datos”)) AND ( mh:(Ansiedad) OR mh:(Depresión) OR mh:(duelo) OR mh:(“Síntoma de depresión”))

En las bases de datos que lo permitieron, se aplicaron filtros para restringir los resultados por rango de fechas, idioma inglés y español y disponibilidad de artículo completo.

## Realización de la revisión y registro de los resultados.

La selección de artículos se realizó en dos fases, siguiendo los criterios de elegibilidad predefinidos. En primer lugar, se llevó a cabo una revisión del título y resumen por parte de dos revisores de manera ciega. Cualquier conflicto se resolvió a través de consenso sin necesidad de intervención de un tercer evaluador. Una vez seleccionados los artículos, se procedió a una lectura minuciosa de cada uno de ellos, extrayendo los datos relevantes relacionados con el uso de NLP en la salud mental. Para facilitar este proceso, se creó una matriz de dominio de análisis y se elaboraron resúmenes de cada artículo destacando los resultados y conclusiones principales.

Finalmente, se llevaron a cabo comparaciones y análisis temáticos entre los estudios revisados para identificar patrones y tendencias en las investigaciones relacionadas con el uso del procesamiento de lenguaje natural en el área de la salud mental.

## Resultados

El propósito fundamental de esta revisión fue llevar a cabo un análisis exhaustivo de la utilización de técnicas y métodos clásicos del Procesamiento de Lenguaje Natural en la investigación de trastornos mentales en la población adulta, con un enfoque específico en la depresión, ansiedad y los sentimientos de duelo. Para ilustrar de manera esquemática el proceso de búsqueda y selección de estudios en cuatro bases de datos científicas, se utilizó el método PRISMA para la búsqueda y selección de artículos, como se muestra en la Figura 1. En dicha figura, se puede observar que la búsqueda inicial arrojó un total de 136

estudios, de los cuales se identificaron 5 duplicados. Posteriormente, se realizó una revisión exhaustiva de los títulos y resúmenes de los 131 artículos que permanecieron después de la primera fase de selección. La exclusión de estos estudios se basó en una serie de criterios previamente definidos. De este conjunto, un total de 90 artículos (68.7%) se excluyeron debido a que no se ajustaban a los objetivos específicos de la investigación.

Estos criterios de exclusión se establecieron cuidadosamente para garantizar la consistencia y relevancia de los estudios seleccionados. En primer lugar, se descartaron 8 artículos correspondiente a revisiones, ya que su enfoque no se alineaba con el análisis original de la presente investigación.

En segundo lugar, se excluyeron 12 artículos que no aplicaban el procesamiento de lenguaje natural, una técnica esencial en nuestra metodología de análisis. Además, se eliminaron 69 artículos cuyo objetivo principal no estaba relacionado con la detección o clasificación de condiciones de salud mental, como la depresión, ansiedad y los sentimientos de duelo, que constituían el núcleo central de nuestra revisión.

Finalmente, se identificó un único artículo que incluía una población menor de 18 años en su estudio, lo que resultó en su exclusión del análisis, ya que nuestra investigación se centraba exclusivamente en adultos.

De los 131 artículos seleccionados, se obtuvieron un total de 41 artículos (31.3%) que se sometieron a una revisión exhaustiva de texto completo. Durante esta fase, se identificó un artículo que no se encontraba disponible en las bases de datos de la biblioteca virtual UPB, lo que resultó en su exclusión del análisis. Además, se excluyeron otros 4 artículos debido a que no abordaban específicamente la detección de trastornos como la depresión, ansiedad o el duelo en la población adulta.

Adicionalmente, se excluyó un artículo de conferencia y un estudio que hacía uso de fotos de perfil de usuarios sin aplicar técnicas de procesamiento de lenguaje natural. También se excluyeron dos estudios adicionales, uno de los cuales era un estudio de caso y el otro una carta al editor, por no estar alineados con los objetivos de nuestra revisión.

Después de este riguroso proceso de selección, se identificaron y conservaron un total de 32 artículos que cumplían con los criterios de inclusión para llevar a cabo la revisión y el análisis en profundidad.

## Características Generales de los Artículos Incluidos

El análisis de los artículos seleccionados aportó valiosa información al presente estudio. Se observó que la mayoría de los estudios incluidos (62,5%) emplearon corpus y bases de datos en inglés como principales fuentes de información para la identificación de los trastornos de interés. Sin embargo, también se identificaron investigaciones que utilizaron corpus y bases de datos en otros idiomas para el entrenamiento de los modelos clasificación, como alemán (6,2%), chino mandarín (15,6%), español (6,2%), francés (3,1%), italiano (3,1%) y tailandés (3,1%).

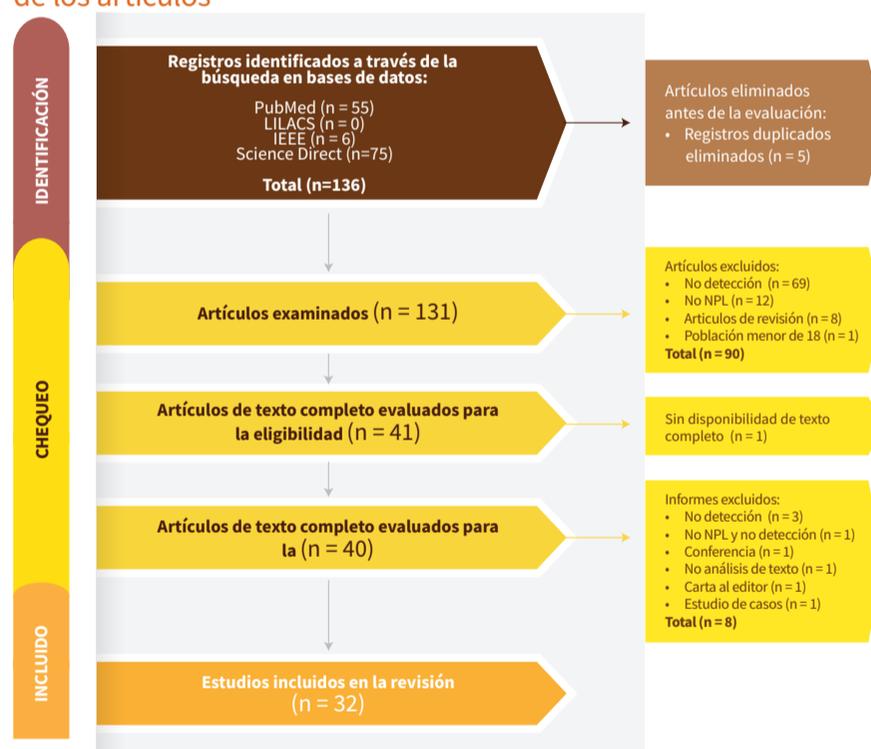
En cuanto a los trastornos relacionados con la salud mental, se destacó que la mayoría de los estudios incluidos se centraron en la depresión, representando un total de 24 artículos. A pesar de esta predominancia, se identificaron investigaciones que abordaron otras cate-

gorías de trastornos de salud mental, como la ansiedad, con un total de 8 artículos. Es relevante mencionar que ninguno de los estudios revisados se enfocó específicamente en la identificación de sentimientos de duelo o trastornos relacionados.

En relación con las fuentes de datos utilizadas en los estudios, se encontró una variabilidad significativa. La mayoría de ellos emplearon redes sociales como fuente principal de información para sus investigaciones, con un 32,5% de los estudios revisados utilizando esta plataforma. Esto refleja la creciente disponibilidad de datos en redes sociales y su utilidad para la investigación en salud mental. Sin embargo, también se identificaron otras fuentes de datos, como bases de datos sobre trastornos mentales (18,6%), cuestionarios (4,6%), encuestas (9,3%), chats de usuarios (4,6%), ensayos escritos por pacientes (2,3%), mensajes de texto (4,6%), notas clínicas (4,6%), diarios electrónicos (2,3%), blogs (2,3%) y registros clínicos electrónicos (13,9%).

Estos resultados resaltan la importancia del idioma de los corpus utilizados y la elección de las fuentes de datos en la investigación de procesamiento de lenguaje natural relacionada con trastornos de salud mental en adultos. La diversidad de idiomas de los corpus refleja la internacionalización de esta área de estudio. Asimismo, la preferencia por las redes sociales como fuente de datos se justifica por la riqueza de información emocional y cognitiva que estas plataformas proporcionan.

Figura 1. Diagrama de Flujo PRISMA para la búsqueda y selección de los artículos

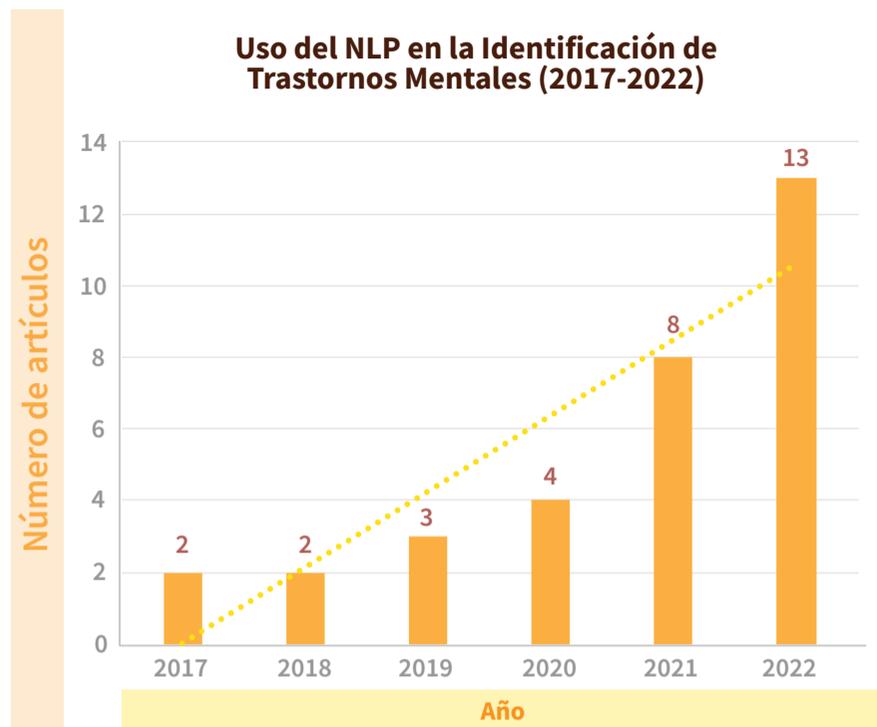


Nota. Elaboración de los autores.

## Evolución de las Publicaciones

Durante los últimos cinco años (2017-2022), se ha observado una tendencia creciente en el número de publicaciones que abordan el uso del Procesamiento de Lenguaje Natural como herramienta para la identificación de condiciones de salud mental, como se muestra en la Figura 2. En 2017 se publicaron muy pocos trabajos en este campo, pero gradualmente comenzaron a ganar mayor atención e interés.

**Figura 2.** Número de artículos publicados de 2017 a 2022 que usan el Procesamiento de Lenguaje Natural para la identificación de trastornos mentales.



Nota. El gráfico representa el crecimiento de los artículos publicados en los últimos años en el uso del NLP en la salud mental. Elaboración de los autores.

A continuación, se expondrán los diversos aspectos identificados en los artículos seleccionados para esta revisión. Se explorarán los hallazgos relevantes en relación con las técnicas clásicas de NLP, así como las diferencias entre enfoques supervisados y no supervisados. Esto permitirá obtener una comprensión completa de las tendencias y enfoques en la investigación sobre trastornos de salud mental en adultos.

Técnicas y Herramientas del Procesamiento de Lenguaje Natural (NLP)

En el contexto de esta investigación, las técnicas y herramientas empleadas en el NLP se pueden categorizar en dos grupos distintos: supervisadas y no supervisadas, dependiendo del método de extracción de información utilizado en los conjuntos de datos (Microsoft, 2022; Tran & Kavuluru, 2017).

En esta revisión, se identificaron un total de 22 estudios que hicieron uso de técnicas o herramientas supervisadas de NLP para la identificación de trastornos, como la depresión y la ansiedad, en los datos. Estas técnicas requieren etiquetas previas para llevar a cabo dicha identificación (Chiong et al., 2021; Haug & Kurpicz-Briki, 2022; Leis et al., 2019; Shehmir Javid, 2022; Tadesse et al., 2019). Además, se encontraron 10 estudios que implementaron técnicas o herramientas de NLP no supervisadas, las cuales operan de manera autónoma, buscando patrones y tendencias dentro del texto sin la supervisión humana directa, lo que las hace especialmente útiles en la identificación de información relevante en grandes conjuntos de datos de texto. (Charles Roe, 2017; Resnik et al., 2015; Vasudha Rani & Sandhya Rani, 2016).

En el análisis de las técnicas de NLP supervisado identificadas en esta revisión, se observó que el 55,5% de los artículos implementaban la herramienta LIWC (Linguistic Inquiry and Word Count). Además, se encontraron diversas técnicas supervisadas de NLP, incluyendo el Reconocimiento de Entidades Nombradas (NER) con un 22,2%, el uso de N-grams con un 11,1%, la representación de Bag of Words (BOW) con un 5,5%, y el método TF-IDF (Frecuencia de Término-Inversa de la Frecuencia en Documentos) también con un 5,5%.

Por otro lado, al analizar los enfoques de NLP no supervisados utilizados en los artículos revisados, se constató que el 40% de ellos empleaban la técnica de LDA (Latent Dirichlet Allocation). Además, se identificaron otras técnicas, como el análisis de sentimiento utilizando Senticon, con un 20% de frecuencia, el uso de Word2Vec con un 10%, la métrica de Valencia-Arousal-Dominancia (VAD) también con un 10%, y la herramienta NimbleMiner con un 20%.

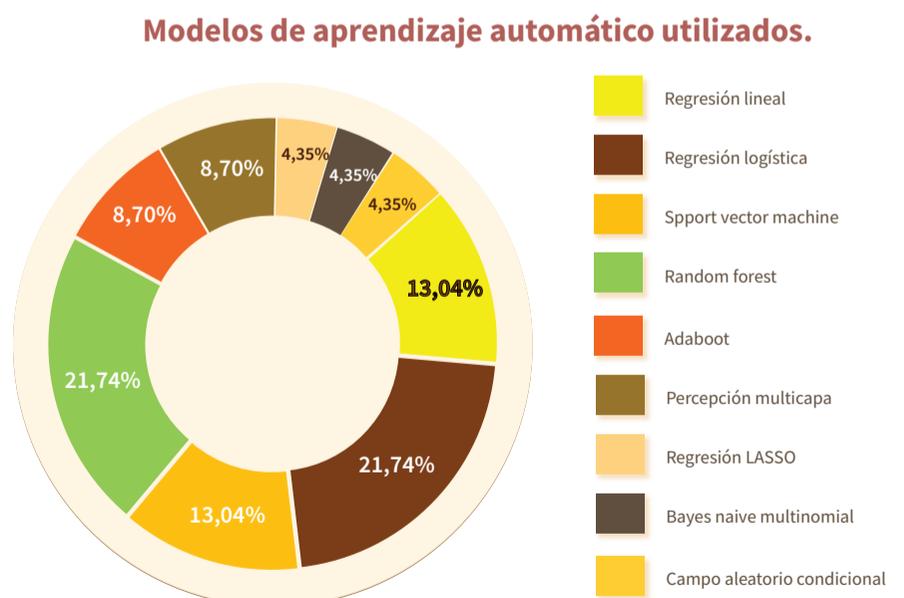
Los hallazgos encontrados en las técnicas del NLP revelan un interés destacado por parte de los investigadores y profesiones en área en el uso de herramientas como LIWC, un programa de computadora que analiza textos y permite el análisis automatizado del contenido lingüístico y emocional en textos relacionados con la salud mental (Chiong et al., 2021). Este interés puede atribuirse a su facilidad de uso y a la capacidad de obtener resultados rápidos y coherentes.

Además, se observa un interés en la implementación de técnicas más avanzadas como LDA y VAD, para la identificación de patrones y comportamientos más profundos mediante un análisis de modelados de temas para dimensionar las emociones presentes en los datos textuales relacionado con las condiciones de salud mental de los pacientes (Resnik et al., 2015; Vasudha Rani & Sandhya Rani, 2016).

## Algoritmos y Modelos de Clasificación de Texto

La identificación precisa de trastornos a partir de datos es un componente esencial en la investigación relacionada con la salud mental. En este contexto, la etapa de clasificación se erige como un pilar fundamental para categorizar y diferenciar diversos tipos de trastornos, aprovechando los datos recolectados por profesionales de la salud (Wiering & Schmidhuber, 1998; Yang et al., 2022). Esta revisión pone de manifiesto que, en la práctica, los métodos de clasificación más ampliamente empleados por los expertos en esta área son los algoritmos de aprendizaje supervisado y los modelos más usados se ilustra en la Figura 3.

**Figura 3.** Variedad de Modelos de Aprendizaje Automático Aplicados en la Investigación.



Nota. La gráfica muestra el porcentaje de los modelos de aprendizaje automático supervisados más comúnmente empleados en la revisión. Elaboración de los autores.

Según se muestra en la Figura 3 los modelos de clasificación más utilizados por los investigadores son la regresión logística y el Random Forest, ambos con una frecuencia del 21,74% en los artículos revisados.

Esto se debe a su capacidad para predecir resultados clínicos y su capacidad para identificar factores de riesgo y factores de prevención asociados con diferentes trastornos mentales. Además, estos modelos son relativamente fáciles de interpretar y proporcionan información útil para la toma de decisiones clínicas (Burkhardt et al., 2021; Guntuku et al., 2019; Wang et al., 2022).

Por otro lado, esta revisión ha revelado varios estudios que indican un creciente uso de los algoritmos y modelos de aprendizaje profundo en la clasificación e identificación de trastornos mentales en el ámbito de la salud pública, como se ilustra en la Figura 4. Estos estudios demuestran resultados superiores en comparación con los algoritmos y métodos de clasificación tradicionales (Rizwan et al., 2022).

**Figura 4.** Uso de algoritmos de aprendizaje profundo en las publicaciones entre los años 2017 a 2022.



Nota. La gráfica representa la tendencia a usar nuevos métodos de clasificación de texto, como el uso del aprendizaje profundo o Deep learning en inglés, a partir de modelos preentrenados del lenguaje para el NLP o Transformers. Elaboración de los autores.

La ilustra la tendencia en el uso de algoritmos de aprendizaje profundo en el contexto de la salud mental a lo largo de los años. Aunque el número de publicaciones que aplican estos algoritmos es limitado, se ha observado un incremento en los últimos años. Se espera que esta tendencia continúe en el futuro (Chiong et al., 2021; Noraset et al., 2022; Tran & Kavuluru, 2017). Este hallazgo refleja un creciente interés por parte de los profesionales de la salud mental en la adopción de estos algoritmos como herramientas para la identificación de trastornos mentales (Patel et al., 2022).

## Detección de Trastornos Mentales Mediante Diversas Técnicas y Herramientas de Procesamiento del Lenguaje Natural (NLP)

La implementación de NLP en la identificación de trastornos mentales, como la depresión y la ansiedad, ha resultado en investigaciones reveladoras. Un estudio realizado por Leis et al. (2019) arrojó resultados de gran relevancia en este contexto. Descubrieron que alrededor del

80% de los tweets publicados por individuos que padecen depresión muestran un uso más frecuente de pronombres singulares en primera persona. Por otro lado, se observó una disminución significativa en la frecuencia de pronombres plurales en primera y segunda persona en usuarios afectados por la depresión. Estos hallazgos sugieren un patrón lingüístico distintivo asociado a la depresión en el discurso en línea, caracterizado por una mayor individualización y una menor referencia a interacciones sociales.

Este estudio destaca la importancia de la detección temprana de trastornos mentales a través del análisis del lenguaje en plataformas de redes sociales y cómo los patrones lingüísticos pueden proporcionar insights valiosos en el ámbito de la salud mental

Por otra parte, Tadesse et al. (2019) llevaron a cabo un estudio que hizo uso de técnicas PLN con un enfoque supervisado. En su metodología, utilizaron una combinación de herramientas, incluyendo el Análisis del Lenguaje a través del Diccionario (LIWC), Modelado de Tópicos Latentes (LDA) y el análisis de N-gramas, para extraer características relevantes de los textos analizados. No obstante, el aspecto más sobresaliente de su investigación radica en la implementación de un modelo de clasificación de aprendizaje automático de tipo supervisado, en concreto un Perceptrón Multicapa (MLP). Este enfoque demostró ser altamente eficaz, logrando una precisión del 91% en la detección de la depresión y ansiedad en las publicaciones de los usuarios de Reddit.

Este resultado resalta la eficacia de un enfoque supervisado en la identificación de patrones de lenguaje asociados con el estado emocional de los individuos en la mencionada plataforma en línea, en este caso, Reddit. La combinación de herramientas de PLN y un modelo de aprendizaje automático supervisado puede proporcionar una comprensión más profunda de la detección de trastornos de salud mental en el discurso en línea o blog.

De manera similar, Vaci et al. (2020) emplearon el NLP para organizar datos clínicos sobre la depresión. En su estudio, utilizaron una combinación de métodos para extraer información relevante de los registros electrónicos de salud. Los resultados obtenidos demostraron un elevado nivel de precisión en la extracción de datos relacionados con pacientes que padecen depresión, presentes en dichos registros electrónicos. Estos hallazgos respaldan la eficacia del procesamiento del lenguaje natural en la identificación y análisis de trastornos mentales, tanto en el ámbito del discurso en línea como en la gestión de datos clínicos.

Además, es relevante destacar que investigaciones recientes, como la realizada por Sharma et al. (2021), han logrado avances significativos en el campo de la identificación de trastornos mentales al explorar con éxito el uso de tecnologías avanzadas. Estos estudios han empleado arquitecturas de redes neuronales profundas en conjunto con el Procesamiento del Lenguaje Natural (NLP) y técnicas de Latent Dirichlet Allocation (LDA). Un aspecto destacado de estas investigaciones es la utilización de blockchain como una herramienta para garantizar la seguridad y la integridad de los datos clínicos de los pacientes. Gracias a estos enfoques tecnológicos más sofisticados, los investigadores lograron obtener una precisión superior al 90% en la detección de trastornos como la depresión y la ansiedad. Estos avances subrayan la eficacia de la combinación de tecnologías avanzadas en la identificación y clasificación de trastornos mentales a partir de datos clínicos.

Otro estudio similar, realizado por Rizwan et al. (2022), se enfocó en la evaluación de la intensidad de la depresión utilizando datos de Twitter. Este estudio llevó a cabo experimentos con cuatro modelos de lenguaje basados en Transformers. Estos modelos fueron sometidos a

una exhaustiva evaluación mediante el aprendizaje por transferencia a partir del Procesamiento del Lenguaje Natural (NLP) con el objetivo de clasificar múltiples clases de intensidad de la depresión, que comprenden tres niveles: 'severo', 'moderado' y 'leve'.

Los resultados obtenidos demostraron que el modelo implementado es efectivo y confiable, logrando una impresionante precisión del 92% en la clasificación de tweets relacionados con la depresión. Estos hallazgos subrayan el potencial de las tecnologías basadas en el procesamiento del lenguaje natural, como los modelos de lenguaje basados en Transformers y el aprendizaje por transferencia, para evaluar la intensidad de la depresión y clasificar eficazmente diferentes niveles de gravedad. Estas herramientas tecnológicas avanzadas desempeñan un papel prometedor en la detección temprana y el monitoreo de la depresión a partir del análisis de texto en línea.

Por último, en otro estudio realizado por Noraset et al. (2022), presentó LAPoMM (Learning Automata-based Protocol for Mobile Multihop Networks), una innovadora red móvil multisalvo basada en autómatas de aprendizaje de tipo supervisado. Este marco revolucionario permite el monitoreo en tiempo real de indicadores de salud mental a partir de datos de redes sociales, sin la necesidad de utilizar conjuntos de datos etiquetados en lenguajes específicos.

Los investigadores descubrieron una alta correlación entre los casos reales de depresión y las señales de sentimiento negativo predichas, así como entre los intentos de suicidio y las señales negativas (por ejemplo, miedo, tristeza y disgusto) y la tendencia suicida. Es relevante señalar que, en esta investigación, utilizaron modelos de lenguaje preentrenados de gran tamaño, como RoBERTa, BERT y LaBSE, como parte de su metodología. Estos modelos avanzados de procesamiento de lenguaje natural han demostrado ser eficaces en la identificación y análisis de patrones lingüísticos y emocionales relacionados con la salud mental en datos de redes sociales, fortaleciendo aún más la base tecnológica para la investigación en este campo.

En conjunto, todos estos métodos de Procesamiento del Lenguaje Natural (NLP) ofrecen una valiosa contribución para mejorar la detección y el diagnóstico de trastornos mentales. La aplicación de estos métodos en la práctica clínica proporciona a los profesionales de la salud una herramienta con un alto nivel de rendimiento para tomar decisiones fundamentadas en el ámbito de los trastornos mentales. Estos avances, respaldados por la investigación reciente y la tecnología, están transformando la atención de la salud mental al permitir una detección temprana, un tratamiento más efectivo y un mejor seguimiento de los pacientes. Esta evolución no solo beneficia a los profesionales de la salud, sino que también tiene un impacto significativo en la vida de las personas que padecen trastornos mentales y sus familias, mejorando la calidad de la atención y la calidad de vida en general.

## Discusión

Los resultados presentados previamente subrayan la importancia del Procesamiento de Lenguaje Natural como una herramienta valiosa en el análisis de trastornos mentales en la población adulta, en particular aquellos relacionados con la depresión y la ansiedad. Este análisis se basa en datos no estructurados, que incluyen una amplia gama de fuentes, como cuestionarios, encuestas, chats de usuarios, ensayos escritos por pacientes, mensajes de texto, notas clínicas, diarios electrónicos, blogs, registros clínicos electrónicos y publicaciones en redes sociales.

Los datos no estructurados ofrecen una perspectiva única de la

experiencia de las personas con trastornos mentales, permitiendo una comprensión más completa y profunda de sus pensamientos, emociones y comportamientos. Sin embargo, el uso de NLP en este contexto también conlleva desafíos, como la necesidad de abordar cuestiones de privacidad de datos y garantizar que los algoritmos sean justos y no sesgados.

Es importante destacar que el avance en esta área de investigación requiere una colaboración interdisciplinaria entre profesionales de la salud mental, lingüistas computacionales y expertos en NLP. Esta cooperación permitirá aprovechar al máximo el potencial de las herramientas de NLP para mejorar la detección temprana, el diagnóstico y el tratamiento de los trastornos mentales en la población adulta.

Es esencial subrayar que, si bien los métodos de aprendizaje automático y NLP ofrecen valiosas perspectivas en la investigación de la salud mental, es crucial considerarlos como herramientas de apoyo para la práctica clínica en lugar de reemplazar la evaluación y el juicio clínico. Estos enfoques pueden enriquecer el diagnóstico y tratamiento de trastornos mentales, pero deben utilizarse con precaución y de manera complementaria a los expertos clínicos en el cuidado de la salud mental.

En este contexto, un estudio similar realizado por Glaz et al. (2021) señaló que las técnicas de NLP y los modelos de aprendizaje automático proporcionan información a partir de datos inexplorados y sobre los hábitos diarios de los pacientes que generalmente son inaccesibles para los psicólogos. Considerando estas técnicas como una herramienta adicional en cada paso de la atención de la salud mental: diagnóstico, pronóstico, eficacia del tratamiento y monitoreo, los profesionales de la salud mental pueden tomar decisiones más informadas y personalizadas en beneficio de sus pacientes.

Por otro lado, en consonancia con la importancia de considerar el NLP y el aprendizaje automático como herramientas complementarias en la atención de la salud mental, Malhotra A & Jindal R (2022) destacaron la necesidad apremiante de desarrollar sistemas de evaluación y vigilancia de la salud mental en tiempo real. Estos sistemas se perfilan como valiosos aliados en la identificación predictiva y no invasiva de condiciones críticas como la depresión, la autolesión y el riesgo de suicidio.

La investigación se centra en la revisión sistemática de diversas arquitecturas de aprendizaje profundo basadas en redes neuronales, destinadas a abordar esta compleja tarea de procesamiento de lenguaje computacional. Es importante destacar que, hasta la fecha, gran parte de las investigaciones previas en este ámbito se han centrado exclusivamente en el contenido en inglés de plataformas como Twitter y Reddit. Entre las arquitecturas de aprendizaje profundo más destacadas y utilizadas para este propósito se encuentran CNN, LSTM, BERT y sus variantes, que han demostrado ser preferidas en la resolución de este desafío.

Además, es relevante mencionar que, en su mayoría, se han empleado conjuntos de datos textuales unimodales para el entrenamiento de modelos supervisados en esta área de investigación. Sin embargo, se deben abordar desafíos adicionales, como la necesidad de conjuntos de datos multilingües y el potencial sesgo en los datos, para garantizar la eficacia y la aplicabilidad de estos sistemas en diversos contextos culturales y lingüísticos.

De igual forma, en continuidad con la exploración de las herramientas de NLP en el ámbito de la salud mental, Greco et al. (2023) se enfocaron en la aplicación de modelos de lenguaje basados en Transformer, como Bi-LSTM, BERT y GPT, para abordar cuestiones de salud mental. Estos modelos han demostrado ser herramientas poderosas para la identificación y predicción de problemas de salud mental a par-

tir de datos de texto.

En su investigación, se empleó un enfoque de procesamiento de texto, utilizando tanto Bi-LSTM como BERT, y se observó un éxito significativo en la predicción de problemas de salud mental. Además, se introdujo un modelo preentrenado denominado MentalBERT, diseñado específicamente para el ámbito de la salud mental, que demostró utilidad en la detección de emociones y la clasificación de textos relacionados con salud mental. Asimismo, se destacó que los modelos basados en GPT se mostraron efectivos en la generación de contenido relacionado con la salud mental, como respuestas a preguntas sobre depresión y ansiedad.

No obstante, se subrayó la necesidad de explorar la aplicación de estos modelos en idiomas con recursos limitados y considerar enfoques multimodales que integren no solo texto, sino también imágenes y datos multimedia. Esta dirección de investigación promete enriquecer aún más la capacidad de estas herramientas para abordar las complejidades de la salud mental en una variedad de contextos culturales y lingüísticos.

## Limitaciones de la revisión

Existen varias limitaciones asociadas con esta revisión. En primer lugar, se utilizó un conjunto limitado de bases de datos para la búsqueda de estudios, lo que podría afectar la diversidad de los estudios presentados y la representatividad de la investigación en el campo de la salud mental. En segundo lugar, la revisión se limitó a un período específico entre 2017 y 2022, lo que podría haber pasado por alto desarrollos más recientes en el uso de Procesamiento de Lenguaje Natural en este campo en constante evolución.

Además, se evidenció en la revisión que la mayor parte de los estudios en los que se usó NLP en salud mental se llevaron a cabo en inglés, con escasa documentación en español y otros idiomas, lo que restringe el acceso a la información por parte de investigadores y profesionales hispanohablantes en este campo. Esto resalta la necesidad de fomentar investigaciones en español y facilitar la disponibilidad de recursos en lenguas no predominantes en la comunidad científica.

Adicionalmente, se han identificado otras limitaciones metodológicas y posibles fuentes de sesgo en esta revisión. Como se detalla en la sección de resultados, al implementar las ecuaciones de búsqueda en diversas bases de datos, se observó una ausencia notable de artículos o estudios relacionados con la implementación de los modelos GPT (Generative Pre-trained Transformer). Estos modelos han sido pioneros en la revolución del campo del procesamiento del lenguaje, tanto en el análisis como en la generación de texto en los últimos años. La ausencia de estos estudios se atribuye a la estructura de las ecuaciones utilizadas en las bases de datos seleccionadas. Se recomienda encarecidamente que futuras revisiones consideren estrategias para garantizar la inclusión de investigaciones sobre estos modelos revolucionarios, lo que enriquecerá significativamente el panorama de la investigación en este campo en constante evolución.

## Conclusión

Tras revisar las metodologías y resultados de los artículos seleccionados para este estudio sobre el uso del NLP en la identificación de condiciones mentales como la depresión y la ansiedad, se puede concluir que esta tecnología tiene un potencial significativo para mejorar

la detección de trastornos de salud mental. Los métodos de NLP en conjunto con los algoritmos de aprendizaje supervisado son los más comunes entre los investigadores debido a su facilidad de implementación, según lo observado en esta revisión.

Por otra parte, los métodos de NLP en conjuntos con los algoritmos de aprendizaje profundo fueron los menos utilizados en los estudios revisados, pero presentaron una perspectiva innovadora en el análisis de grandes cantidades de datos. Además, en los resultados se puede observar una tendencia a la implementación de estos nuevos métodos, gracias a su capacidad para identificar patrones con mayor precisión que otros métodos de clasificación tradicionales.

En conclusión, los hallazgos de esta revisión de alcance apuntan a que el uso del NLP en la identificación de trastornos de salud mental está actualmente enfocado en la implementación de algoritmos de aprendizaje supervisado, aunque se espera que se investigue y se desarrollen tecnologías basadas en algoritmos de aprendizaje profundo mediante modelos preentrenados como los Transformers en el futuro, ya que estas tecnologías presentan un gran potencial para mejorar la detección e identificación de trastornos de salud mental. Por consiguiente, es importante continuar investigando y desarrollando estas tecnologías para poder aplicarlas eficazmente en el campo de la salud mental.

## Agradecimientos

Este artículo es parte integral del proyecto “Desarrollo de Estrategias para la Atención Integral de la Salud Mental de las Personas durante y después del período de confinamiento, que fue financiado con recursos del Sistema General de Regalías de Colombia. Esta iniciativa se enmarca en la Convocatoria del Fondo CTEI del Sistema General de Regalías para el Fortalecimiento de las Capacidades de Investigación y Desarrollo Regionales, así como en las Iniciativas de Desarrollo y Transferencia de Tecnología y Conocimiento para la Innovación, con un enfoque específico en la resolución de problemáticas derivadas del COVID-19. El proyecto está registrado bajo el BPIN 2020000100726. Agradecemos el apoyo y la financiación brindados para llevar a cabo esta importante labor en el departamento de Córdoba.

## Referencias

- Ávila-Toscano, J. H., Hernández-Chang, E. A., Marengo-Escuderos, A., Herrera-Bravo, M., & Rambal-Rivaldo, L. (2022). Procesos cognitivos implicados en conductas de protección ante COVID-19 (SARS-CoV-2). *Revista Iberoamericana de Psicología*, 15(1), 1–10. <https://doi.org/10.33881/2027-1786.rip.15101>
- Benitez Molina, A., & Caballero Badillo, M. C. (2017). Estudio psicométrico de las escalas de depresión, ansiedad y funcionalidad familiar en estudiantes de la Universidad Industrial de Santander. *Acta Colombiana de Psicología*, 20(1), 221–241. <https://doi.org/10.14718/ACP.2017.20.1.11>
- Carod Artal, F. J. (2020). Complicaciones neurológicas por coronavirus y COVID-19. *Revista de Neurología*, 70(09), 311. <https://doi.org/10.33588/rn.7009.2020179>
- Charles Roe. (2017, December 6). *Identify Data Patterns with Natural Language Processing and Machine Learning*. DATAVERSITY. <https://www.dataversity.net/identify-data-patterns-natural-language-processing-machine-learning/#>
- Chávez-Negrete, L. A., Olivares-Luna, A. M., Rivera-Rodríguez, J. M., &

- Pedraza-Avilés, A. G. (2021). Relación entre depresión y ansiedad con el aislamiento social debido al confinamiento en adultos mayores. *Revista Peruana de Investigación En Salud*, 5(4), 273–278. <https://doi.org/10.35839/repis.5.4.1067>
- Chiong, R., Budhi, G. S., Dhakal, S., & Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. *Computers in Biology and Medicine*, 135. <https://doi.org/10.1016/j.compbiomed.2021.104499>
- Cohen, A. S., Mitchell, K. R., & Elvevåg, B. (2014). What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. *Schizophrenia Research*, 159(2–3), 533–538. <https://doi.org/10.1016/j.schres.2014.09.013>
- Córdoba Cortés, M. N. (2020). *Análisis Lingüístico Comparativo de Narrativas Autobiográficas entre Pacientes con Depresión y Trastorno Obsesivo Compulsivo* [Universitat Ramon Llull]. <http://hdl.handle.net/10803/668271>
- DeSouza, D. D., Robin, J., Gumus, M., & Yeung, A. (2021). Natural Language Processing as an Emerging Tool to Detect Late-Life Depression. *Frontiers in Psychiatry*, 12. <https://doi.org/10.3389/fpsy.2021.719125>
- Elixmahir Dávila-Marrero, Gladiliz Rivera-Delpín, Ashley Rodríguez-Mercado, Raúl Olivo-Arroyo, & Jorge A. Montijo. (2021, May 26). Persistent Cognitive Manifestations Related to COVID-19. *Universidad de Puerto Rico*. <https://revistas.upr.edu/index.php/psicologias/article/view/18973/16394>
- Enzyme. (2020). *Natural Language Processing: ¿Cómo es la técnica Word Embeddings?* <https://enzyme.biz/blog/natural-language-processing#:~:text=Sobre%20estos%20fundamentos%2C%20un%20ejemplo,entonces%20el%20vector%20%E2%80%9CReina%E2%80%9D>
- Greco, C. M., Simeri, A., Tagarelli, A., & Zumpano, E. (2023). Transformer-based language models for mental health issues: A survey. *Pattern Recognition Letters*, 167, 204–211. <https://doi.org/10.1016/j.patrec.2023.02.016>
- Haug, S., & Kurpicz-Briki, M. (2022). Burnout and Depression Detection Using Affective Word List Ratings. *Studies in Health Technology and Informatics*, 292, 43–48. <https://doi.org/10.3233/SHTI220318>
- Hays DG. (1967). Introduction to Computational Linguistics, Mathematical Linguistics and Automatic Language Processing. *Cambridge: American Elsevier Publishing Co*. <https://www.cambridge.org/core/journals/journal-of-linguistics/article/abs/david-g-hays-introduction-to-computational-linguistics-mathematical-linguistics-and-automatic-language-processing-2-new-york-american-elsevier-publishing-co-1967-pp-xvi-231-si-figures/F97978F4A8D49D04DAD542C8894308E1>
- Hermoso Contreras, Cristina Andrea Pelegrín Valero, Carmelo Mariano, Olivera Pueyo, & Francisco Javier. (2022). *Detección de síntomas psiquiátricos y trastornos del comportamiento en pacientes con demencia. Utilidad de la versión española del Cambridge Behavioural Inventory - Revised (CBI - R)*. Universidad de Zaragoza. <https://dialnet.unirioja.es/servlet/tesis?codigo=310720>
- Ibáñez, C., & Echeburúa, E. (2015). Función y limitaciones del pronóstico en la evaluación diagnóstica en el ámbito de la psicología clínica. *Clínica y Salud*, 26(1), 17–22. <https://doi.org/10.1016/j.clysa.2014.10.008>
- IBM. (2021). *Aprendizaje supervisado*. <https://www.ibm.com/mx-es/topics/supervised-learning>
- Iturbide Fernández, P. (2021). Impacto del COVID-19 en población adulta mexicana: conductas de salud pública y salud mental. *Revista Iberoamericana de Psicología*, 14(3), 1–12. <https://doi.org/10.33881/2027-1786.rip.14301>
- Kuliukas, L., Hauck, Y., Sweet, L., Vasilevski, V., Homer, C., Wynter, K., Wilson, A., Szabo, R., & Bradfield, Z. (2021). A cross sectional study of midwifery students' experiences of COVID-19: Uncertainty and expendability. *Nurse Education in Practice*, 51, 102988. <https://doi.org/10.1016/j.nepr.2021.102988>
- Le Glaz, A., Haralambous, Y., Kim-Dufor, D.-H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., DeVlyder, J., Walter, M., Berrouguet, S., & Lemey, C. (2021). Machine Learning and Natural Language Processing in Mental Health: Systematic Review. *Journal of Medical Internet Research*, 23(5), e15708. <https://doi.org/10.2196/15708>
- Leis, A., Ronzano, F., Mayer, M. A., Furlong, L. I., & Sanz, F. (2019). Detecting signs of depression in tweets in Spanish: Behavioral and linguistic analysis. *Journal of Medical Internet Research*, 21(6). <https://doi.org/10.2196/14199>
- Lior Rokach, & Oded Maimon. (2008). Data mining with decision trees: theory and applications. *World Scientific*. [https://doc.lagout.org/Others/Data%20Mining/Data%20Mining%20with%20Decision%20Trees\\_%20Theory%20and%20Applications%20%282nd%20ed.%29%20%5BRokach%20%26%20Maimon%202014-10-23%5D.pdf](https://doc.lagout.org/Others/Data%20Mining/Data%20Mining%20with%20Decision%20Trees_%20Theory%20and%20Applications%20%282nd%20ed.%29%20%5BRokach%20%26%20Maimon%202014-10-23%5D.pdf)
- Mahalingasivam, V., Craik, A., Tomlinson, L. A., Ge, L., Hou, L., Wang, Q., Yang, K., Fogarty, D. G., & Keenan, C. (2021). A Systematic Review of COVID-19 and Kidney Transplantation. *Kidney International Reports*, 6(1), 24–45. <https://doi.org/https://doi.org/10.1016/j.ekir.2020.10.023>
- Malhotra, A., & Jindal, R. (2022). Deep learning techniques for suicide and depression detection from online social media: A scoping review. *Applied Soft Computing*, 130, 109713. <https://doi.org/10.1016/j.asoc.2022.109713>
- Mayo Clinic. (2021). *Trastorno de ansiedad*. MayoClinic.Org. <https://www.mayoclinic.org/>
- Medlineplus. (2018). *Depresión*. <https://medlineplus.gov/spanish/ency/article/003213.htm>
- Mehta, N., & Pandit, A. (2018). Concurrence of big data analytics and healthcare: A systematic review. *International Journal of Medical Informatics*, 114, 57–65. <https://doi.org/10.1016/j.ijmedinf.2018.03.013>
- Microsoft. (2022). *Large-scale custom natural language processing*. <https://learn.microsoft.com/en-us/azure/architecture/ai-ml/idea/large-scale-custom-natural-language-processing>
- Mishra, V., & Garg, T. (2018). A systematic study on predicting depression using text analytics. *Journal of Fundamental and Applied Sciences*, 10(2). <https://www.ajol.info/index.php/jfas/article/view/172123>
- Moral de la Rubia, J. (2013). Validación de un formato simplificado del Inventario de Depresión de Beck (BDI-2). *Psicología Iberoamericana*, 21(1), 42–52. <https://doi.org/10.48102/pi.v21i1.165>
- NIMH. (2021). *Depresión*. Instituto Nacional de La Salud Mental. <https://www.nimh.nih.gov/health/publications/espanol/depression-sp>
- Noraset, T., Chatrinan, K., Tawichsri, T., Thaipisutikul, T., & Tuarob, S. (2022). Language-agnostic deep learning framework for automatic monitoring of population-level mental health from social networks. *Journal of Biomedical Informatics*, 133. <https://doi.org/10.1016/j.jbi.2022.104145>
- OMS. (2022, September 28). *La OMS y la OIT piden nuevas medidas para abordar los problemas de salud mental en el trabajo*. Organización Mundial de La Salud. <https://www.who.int/es/news/item/28-09-2022-who-and-ilo-call-for-new-measures-to-tackle-mental-health-issues-at-work>
- ONU. (2022). *Estado de la salud mental tras la pandemia del COVID-19 y progreso de la Iniciativa Especial para la Salud Mental (2019-2023) de la OMS*. Naciones Unidas. <https://www.un.org/es/cr%C3%B3nica-onu/estado-de-la-salud-mental-tras-la-pandemia-del-covid-19-y-progreso-de-la-iniciativa#:~:text=Se%20calcula%20que%20la%20pandemia,las%20personas%20de%20buscar%20ayuda>
- OPS. (2022a, January 13). *Estudio advierte sobre elevados niveles de depresión y pensamientos suicidas en personal de salud de América Latina durante la pandemia*. Organización Panamericana de Salud (OPS). <https://www.paho.org/es/noticias/13-1-2022-estudio-advierte-sobre-elevados-niveles-depresion-pensamientos-suicidas-personal>
- OPS. (2022b, January 13). *Estudio advierte sobre elevados niveles de depresión y pensamientos suicidas en personal de salud de América Latina durante la pandemia*. Organización Panamericana de La Salud. <https://www.paho.org/es/noticias/13-1-2022-estudio-advierte-sobre-elevados-niveles-depresion-pensamientos-suicidas-personal>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic*

- Reviews*, 5(1), 210. <https://doi.org/10.1186/s13643-016-0384-4>
- P. Resnik, W. Armstrong, L. Claudino, T. Nguyen, V.-A. Nguyen, & J. Boyd-Graber. (2015). Beyond LDA: Exploring supervised topic modeling for depression-related language in twitter. *Proc. 2nd Workshop Comput. Lingüística Clin. Psicología.*, 99–107. <https://aclanthology.org/W15-1212/>
- Patel, R., Irving, J., Brinn, A., Taylor, M., Shetty, H., Pritchard, M., Stewart, R., Fusar-Poli, P., & McGuire, P. (2022). Associations of presenting symptoms and subsequent adverse clinical outcomes in people with unipolar depression: A prospective natural language processing (NLP), transdiagnostic, network analysis of electronic health record (EHR) data. *BMJ Open*, 12(4). <https://doi.org/10.1136/bmjopen-2021-056541>
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141–146. <https://doi.org/10.1097/XEB.0000000000000050>
- Preoțiu-Pietro, Daniel, Johannes Eichstaedt, Gregory Park, Maarten Sap, Laura Smith, Victoria Tobolsky, H. Andrew Schwartz, & Lyle Ungar. (2015). The role of personality age and gender in tweeting about mental illness. *Proc. 2nd Workshop Comput. Lingüística Clin. Psicología.*, 21–30. <https://aclanthology.org/W15-1203/>
- Purriños, M. J. (2013). Escala de hamilton-hamilton depression rating scale (HDRS). *Serv. Epidemiol*, 2, 1–4. [https://medsol.co/informacion/medixen/depresion\\_escala\\_hamilton.pdf](https://medsol.co/informacion/medixen/depresion_escala_hamilton.pdf)
- Rizwan, M., Mushtaq, M. F., Akram, U., Mehmood, A., Ashraf, I., & Sahelices, B. (2022). Depression Classification From Tweets Using Small Deep Transfer Learning Language Models. *IEEE Access*, 10, 129176–129189. <https://doi.org/10.1109/ACCESS.2022.3223049>
- Saffar, A. H., Mann, T. K., & Ofoghi, B. (2023). Textual emotion detection in health: Advances and applications. *Journal of Biomedical Informatics*, 137, 104258. <https://doi.org/10.1016/j.jbi.2022.104258>
- Sanz, J., Luis, A., Carmelo, P., & Resumen, V. (2003). *CLÍNICA Y SALUD 249 ARTÍCULOS Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 2. Propiedades psicométricas en población general The spanish adaptation of Beck's Depression Inventory-II (BDI-II): 2. Psychometric properties in the general population* (Vol. 14). <https://www.redalyc.org/pdf/1806/180617972001.pdf>
- Sharma, C., Sharma, S., & Sakshi. (2022). Latent DIRICHLET allocation (LDA) based information modelling on BLOCKCHAIN technology: a review of trends and research patterns used in integration. *Multimedia Tools and Applications*, 81(25), 36805–36831. <https://doi.org/10.1007/s11042-022-13500-z>
- Sharma, S., Kalra, V., & Agrawal, R. (2021). Depression discovery in cancer communities using deep learning. In *Handbook of Deep Learning in Biomedical Engineering* (pp. 123–154). Elsevier. <https://doi.org/10.1016/B978-0-12-823014-5.00004-1>
- Shehmir Javaid. (2022, March 31). *Data Labeling For Natural Language Processing (NLP)*. AIMultiple. <https://research.aimultiple.com/nlp-data-labeling/>
- Simón Martínez, V. (2022). *Alteraciones neuropsicológicas en el trastorno obsesivo compulsivo refractario al tratamiento. UNIVERSIDAD COMPLUTENSE DE MADRID*. <https://dialnet.unirioja.es/servlet/tesis?codigo=313006>
- Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. *IEEE Access*, 7, 44883–44893. <https://doi.org/10.1109/ACCESS.2019.2909180>
- The MathWorks, Inc. (2020). *Cree modelos lingüísticos multipalabra y análcelos con Machine Learning*. <https://la.mathworks.com/discovery/ngram.html#:~:text=El%20modelado%20de%20n%2Dgramas,de%20word%20embedding%2C%20como%20word2vec.>
- Tran, T., & Kavuluru, R. (2017). Predicting mental conditions based on “history of present illness” in psychiatric notes with deep neural networks. *Journal of Biomedical Informatics*, 75, S138–S148. <https://doi.org/10.1016/j.jbi.2017.06.010>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Vaci, N., Liu, Q., Kormilitzin, A., De Crescenzo, F., Kurtulmus, A., Harvey, J., O'Dell, B., Innocent, S., Tomlinson, A., Cipriani, A., & Nevado-Holgado, A. (2020). Natural language processing for structuring clinical text data on depression using UK-CRIS. *Evidence-Based Mental Health*, 23(1), 21–26. <https://doi.org/10.1136/ebmental-2019-300134>
- Valdivieso Jimenez, G. (2021). Uso de psicofármacos para síntomas neuropsiquiátricos en pacientes hospitalizados con COVID-19. *Horizonte Médico (Lima)*, 21(2), e1272. <https://doi.org/10.24265/horizmed.2021.v21n2.13>
- Vasudha Rani, V., & Sandhya Rani, K. (2016). Twitter Streaming and Analysis through R. *Indian Journal of Science and Technology*, 9(45). <https://doi.org/10.17485/ijst/2016/v9i45/97914>
- Wiering, M., & Schmidhuber, J. (1998). Aprendizaje automático. *Machine Learning*, 33(1), 105–115. <https://doi.org/10.1023/A:1007562800292>
- Yang, K., Zhang, T., & Ananiadou, S. (2022). A mental state Knowledge-aware and Contrastive Network for early stress and depression detection on social media. *Information Processing and Management*, 59(4). <https://doi.org/10.1016/j.ipm.2022.102961>
- Yusufov, M., Pirl, W. F., Braun, I., Tulsy, J. A., & Lindvall, C. (2022). Natural Language Processing for Computer-Assisted Chart Review to Assess Documentation of Substance use and Psychopathology in Heart Failure Patients Awaiting Cardiac Resynchronization Therapy. *Journal of Pain and Symptom Management*, 64(4), 400–409. <https://doi.org/10.1016/j.jpainsymman.2022.06.007>