

Análisis del Lenguaje Natural para la Identificación de Alteraciones Mentales en Redes Sociales: Una Revisión Sistemática de Estudios

Mieles Toloza, Ismael^{1,*} ; Delgado Meza, Jesús² ; Acevedo-Suárez, Johana³ 

¹Universidad de Investigación y Desarrollo -UDI, Psicología virtual, Bucaramanga, Colombia

²Universidad de Investigación y Desarrollo -UDI, Psicología, Bucaramanga, Colombia

³Universidad de Investigación y Desarrollo -UDI, Psicología virtual, Bucaramanga, Colombia

Resumen: Las enfermedades mentales constituyen una de las principales causas de angustia en la vida de las personas a nivel individual, y repercuten en la salud y el bienestar de la sociedad. Para captar estas complejas asociaciones, las ciencias computacionales y la comunicación, a través del uso de métodos de procesamiento del lenguaje natural (NLP) en datos recolectados en redes sociales, han aportado prometedores avances para potenciar la atención sanitaria mental proactiva y ayudar al diagnóstico precoz. Por ello, se realizó una revisión sistemática de la literatura acerca de la detección de alteraciones mentales a través de redes sociales, mediante el uso de NLP en los últimos 5 años, que permitió identificar métodos, tendencias y orientaciones futuras, a través del análisis de 73 estudios, de 509 que arrojó la revisión de documentos extraídos de bases de datos científicas. El estudio reveló que, los fenómenos más comúnmente estudiados, correspondieron a Depresión e Ideación suicida, identificados a través del uso de algoritmos como el LIWC, CNN, LSTM, RF y SVM, en datos extraídos principalmente de Reddit y Twitter. Este estudio, finalmente proporciona algunas recomendaciones sobre las metodologías de NLP para la detección de enfermedades mentales, que pueden ser adoptadas en el ejercicio de profesionales interesados en la salud mental, y algunas reflexiones sobre el uso de estas tecnologías.

Palabras clave: redes sociales, salud mental, procesamiento del lenguaje natural, redes neuronales, inteligencia artificial

Natural Language Analysis for the Identification of Mental Disturbances in Social Networks: A Systematic Review of Studies

Abstract: Mental illness is a major cause of distress in people's lives at the individual level and impacts the health and well-being of society. To capture these complex associations, computational science and communication, through the use of natural language processing (NLP) methods on data collected in social networks, have provided promising advances to enhance proactive mental health care and aid in early diagnosis. Therefore, a systematic review of the literature on the detection of mental disorders through social networks, using NLP in the last 5 years, was carried out, which allowed identifying methods, trends and future directions, through the analysis of 73 studies, out of 509 that resulted from the review of documents extracted from scientific databases. The study revealed that the most commonly studied phenomena corresponded to Depression and Suicidal Ideation, identified through the use of algorithms such as LIWC, CNN, LSTM, RF and SVM on data extracted mainly from Reddit and Twitter. This study finally provides some recommendations on NLP methodologies for mental illness detection that can be adopted in the practice of professionals interested in mental health and some reflections on the use of these technologies.

Keywords: social networks, mental health, natural language processing, neural networks, machine learning

1. INTRODUCCIÓN

La salud mental se ha convertido en una temática de interés general a nivel global, dado el impacto que su afectación en la población puede tener en las esferas económicas, políticas sociales y sanitarias. Sin embargo, la salud mental continúa siendo una de las áreas menos atendidas en salud pública, en

donde cerca de 1000 millones de personas viven con un trastorno mental, 3 millones de personas mueren cada año por el consumo nocivo de alcohol, y una persona se suicida cada 40 segundos (Organización Mundial de la Salud, 2020). Adicionalmente, se sabe que una gran parte de los trastornos que padecen las personas son altamente estigmatizados, lo que contribuye significativamente al sufrimiento de los individuos,

*mieles1@udi.edu.co

Recibido: 01/04/2023

Aceptado: 21/11/2023

Publicado en línea: 09/02/2024

10.33333/rp.vol53n1.06

CC 4.0

desincentivando la búsqueda de ayuda profesional y promoviendo el ocultamiento de los síntomas. Ante esta situación, y como alternativa para compartir sus pensamientos y sentimientos relacionados con la alteración en salud mental, las redes sociales se han convertido en un medio para la búsqueda de orientación y ayuda por parte de la comunidad que allí interactúa (Chiong, et al., 2021).

Las redes sociales han tenido un marcado éxito y penetración en la población en general, las cuales son muy populares en adolescentes (Ophir et al., 2019) y frecuentemente usadas por adultos (Mendu et al., 2020). Así mismo, la pandemia del Covid-19 y el aislamiento social trajo consigo que, plataformas como Facebook, Instagram, Twitter y Reddit, fueran mucho más utilizadas, ya que se convirtieron en los canales que permitían mantener la interacción social segura, en medio de la crisis sanitaria (Garg, 2021; Koh & Liew, 2020).

Entre las características de las redes sociales, se identifica que el 80% de los datos en estas plataformas es textual, y crecen diariamente a razón de 1.3 millones de nuevas entradas de contenido (Fernández, 2020). Ante esta gigantesca cantidad de información, surge una serie de estrategias de análisis de datos que hace uso de distintas metodologías (Moessner et al., 2018), algunas básicas como el análisis y frecuencia de palabras clave publicadas por un usuario, y otras más complejas como la interpretación del lenguaje natural (NLP) por medio de la inteligencia artificial (Mori & Haruno, 2021) y el aprendizaje profundo (Wongkoblap et al., 2021; Yao et al., 2020) a través de distintos modelos de redes neuronales (Babvey et al., 2021; Melvin et al., 2019), lo que permite el monitoreo y la perfilación de usuarios, esto resulta muy útil en mercadotecnia, pero también, ventajoso para instituciones educativas, gobiernos y la salud pública en general (Nandhini & Sheeba, 2015).

Estas novedosas posibilidades de análisis de datos han despertado el interés de diversos autores y estudios respecto a la identificación de factores, patrones y modelos que puedan predecir los comportamientos de usuarios (Centola, 2010; Zhang et al., 2013), relacionados con temas de salud pública y salud mental (Camacho et al., 2013; Dos Santos et al., 2019; Mori, & Haruno, 2021; Tan, et al., 2021). Por ejemplo, el estudio realizado por Bae et al. (2021) cuya investigación buscó determinar si el aprendizaje automático podía utilizarse eficazmente para detectar signos de esquizofrenia en usuarios de la red social Reddit, a través de la recopilación de publicaciones y el análisis de textos; o la investigación realizada por Chiong et al. (2021) cuyo objetivo fue determinar si el aprendizaje automático podría usarse de manera efectiva para detectar signos de depresión en usuarios de las redes sociales, a través del análisis de sus publicaciones.

De esta manera, se observa un panorama coherente al formulado por la Organización mundial de la salud (2017), el cual refiere que, las alteraciones en salud mental en los últimos años van aumentando, señalando que 1 de cada 4 personas, a nivel global, sufrirá alguna enfermedad mental durante el transcurso de su vida (Confederación Salud mental de España, 2019) para lo cual resulta interesante tener formas alternativas para su prevención, detección y estudios, que funcionen

paralelamente al ejercicio de la psicología clínica y la psiquiatría tradicional.

Es por ello que, a través del presente estudio de revisión sistemática, se buscó examinar documentos científicos que implementaran modelos de análisis del lenguaje natural de textos publicados en redes sociales, para detectar cuáles de ellos han sido más utilizados, qué fenómenos de salud mental fueron los más abordados y las redes sociales más comúnmente usadas como fuente de información, permitiendo conocer el estado actual del problema en cuestión, identificar nuevas tendencias, y promover en los investigadores interesados, la formulación de estudios con impacto disciplinar y social.

2. METODOLOGÍA

En este trabajo, se realizó un análisis de la literatura (Palmatier et al., 2018), basado en la Declaración PRISMA (Moher et al., 2009; Perestelo-Pérez, 2013), el cual inició con la revisión de documentos, y continuó con el uso de un enfoque analítico para la identificación, principalmente, de modelos computacionales del lenguaje natural más utilizados para la identificación de alteraciones en salud mental a través de redes sociales. Se siguieron unas pautas metodológicas formuladas por expertos en este tipo de investigación (Cobo et al., 2011), que permitieron el análisis de resultados basado en un modelo aceptado por la comunidad científica (López-Belmonte et al., 2020).

2.1 Diseño de la investigación

Partiendo de la revisión sistemática como metodología principal del estudio, se llevó a cabo un diseño de investigación para la búsqueda, registro y análisis de la literatura científica (Martínez et al., 2015), el cual fue acompañado de un análisis de co-palabras (Soler-Costa et al., 2021). Los distintos procesos llevados a cabo en este diseño permitieron obtener los subdominios conceptuales que posibilitaron orientar el proceso de interpretación de resultados.

2.2 Procedimiento

El proceso de búsqueda y selección se compuso de cuatro pasos principales: (1) identificación, (2) cribado, (3) elegibilidad (4) inclusión, con el fin de disminuir sesgos en la selección de documentos, tal y como se requiere en este tipo de estudios. Los pasos seguidos fueron los siguientes:

En primer lugar, se seleccionaron las bases de datos para el proceso de búsqueda, entre ellas, Ebsco, Academic Search, APA PsyArticles, APA Psynfo, ERIC, Fuente Académica, MEDLINE, Engineering Index, Software Reviews on File, Psychology and Behavioral Science Collection, SocINDEX With Full Text, Master File Premier, Medline with Full Text y PsicoDoc, así como también las plataformas Web of Science y Scopus, ya que están consideradas como bases de datos mundiales y contienen un amplio número de estudios de impacto indexados.

En segundo lugar, fueron definidas las palabras a utilizar en la ecuación de búsqueda, en donde inicialmente se revisaron estudios de impacto (Al Asad et al., 2019; Gaur et al., 2019;

Katchapakirin et al., 2018), con el propósito de generar una lista de palabras clave para hacerlas parte del proceso de consulta, entre las que se encontraron, en español: procesamiento del lenguaje natural, salud mental y redes sociales; y en inglés: natural language processing, mental health and social network. Lo anterior permitió detectar los documentos científicos que tenían esos términos en el título, el resumen, y palabras clave de las publicaciones indexadas.

En tercer lugar, se estableció la siguiente ecuación de búsqueda en español: “procesamiento del lenguaje natural” AND “salud mental” AND “redes sociales” y para el inglés, “natural language processing” AND “mental health” AND “social network”. Esta ecuación se utilizó para la búsqueda de los títulos de las publicaciones de las bases de datos anteriormente mencionadas. Además, para la selección de categorías de búsqueda en las bases de datos, se escogieron todas aquellas relacionadas con el ámbito de la medicina, ciencias computacionales y psicología.

El proceso de búsqueda y reporte se llevó a cabo entre octubre de 2021 a junio de 2023. En el informe se decidió elegir todo tipo de artículos científicos en idiomas inglés y español, y se intentó abarcar una buena parte de la literatura sobre el estado del problema. Estas acciones dieron como resultado un total de 509 publicaciones, de las cuales 170 correspondieron a documentos duplicados. A continuación, para refinar la búsqueda se establecieron varios criterios; como criterio de inclusión se tomaron los años de publicación de los artículos en el periodo 2018 a 2022, y como criterios de exclusión, se separaron aquellos documentos anteriores a 2018 (n=87), documentos mal indexados en las bases de datos (n=90) y estudios que no reportaran fenómenos de salud mental, datos analizados no provenientes de redes sociales, y que no expresaran uso del procesamiento de lenguaje natural (n=89). La aplicación de estos criterios produjo una unidad de análisis de 73 documentos científicos. Las acciones seguidas se muestran en el diagrama de flujo según la declaración PRISMA (Figura 1), en donde se establece el informe de documentos sobre los constructos analizados y las reducciones sufridas por el volumen inicial, tras la aplicación de los criterios establecidos, hasta llegar a la unidad de análisis final establecida en este estudio.

2.3 Análisis de los datos

Todos los artículos finalmente seleccionados (73 en total) fueron analizados y procesados por los investigadores. Se hizo uso de una matriz de sistematización que contempló información referida a: datos de identificación del artículo, estructura teórica y metodológica, y, principales hallazgos. De esta manera, el proceso de análisis permitió la detección de información relacionada con las categorías deductivas definidas: a) redes sociales analizadas, idiomas de publicación y campos de estudio de las publicaciones académicas, b) fenómenos de salud mental más abordados en los estudios y c) modelos de procesamiento del lenguaje natural de textos publicados en redes sociales para la identificación de alteraciones en salud mental; dicha información se obtuvo producto de la concatenación de expresiones comunes, palabras clave, redes de co-palabras (Marín-Marín et al., 2021), extraída de los textos de análisis seleccionados, en las que tres investigadores del presente estudio coincidieron, y

cuyos resultados son expresados en las secciones Resultados y Discusión, de este manuscrito.

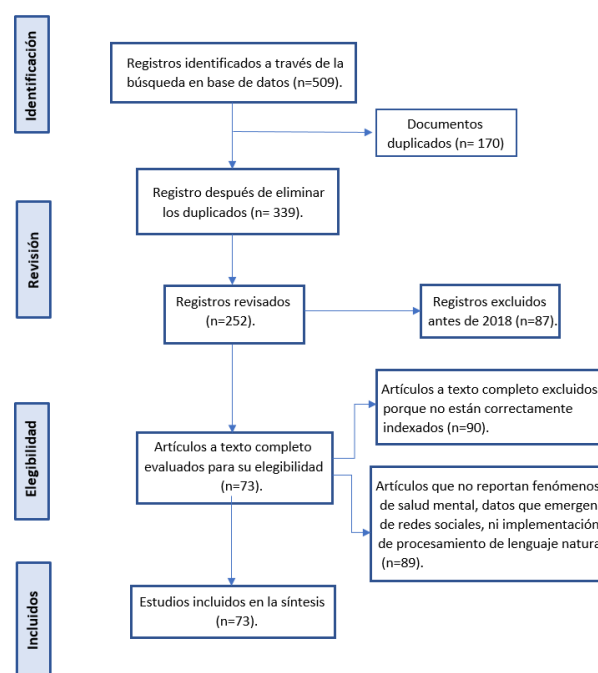


Figura 1. Diagrama de flujo de la información PRISMA a través de las diferentes etapas de la revisión sistemática

3. RESULTADOS

Este apartado informa acerca de los resultados obtenidos y relacionados con las categorías deductivas establecidas, que se vinculan a los objetivos definidos para este estudio, otorgando valor en la investigación sobre la detección de enfermedades mentales a través de NLP, cuyo interés, de acuerdo con lo expresado en las investigaciones consultadas, ha ido en crecimiento en los últimos años.

Se iniciará con la presentación de resultados relacionados con metadatos de interés para esta investigación como, año de publicación de artículos, localización e idioma en el que se desarrolló el estudio, áreas del conocimiento de las revistas en donde fueron publicados las investigaciones, y las redes sociales utilizadas para la extracción de información en dichos estudios; posteriormente, se presentarán los algoritmos usados para el NLP y los fenómenos psicológicos más abordados.

3.1 Años de publicación de los artículos, país de origen e idiomas de publicación de los estudios

En la búsqueda inicial de documentos de investigación publicados entre los años 2018 y 2022, se observa una tendencia dentro del número de estudios publicados sobre el tema por año (Figura 2), y se identifica un aumento sustancial en el número de publicaciones entre el 2021 y el 2022, siendo este último el año en el que se publicó el 46.5% del total de los documentos seleccionados.

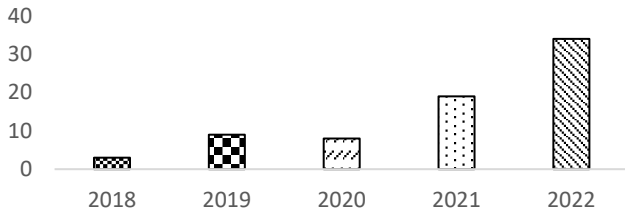


Figura 2. Numero de documentos agrupados por año de publicación

De otro lado, varios países han demostrado su interés en estudiar las alteraciones en salud mental en sus territorios. De los datos extraídos en la presente revisión, se identifica un predominio en el número de publicaciones por parte de EE.UU. (21,9 %), India (21,9 %) y China (8,2 %) (Figura 3); así mismo España destaca como el único país hispano hablante que reporta publicaciones.

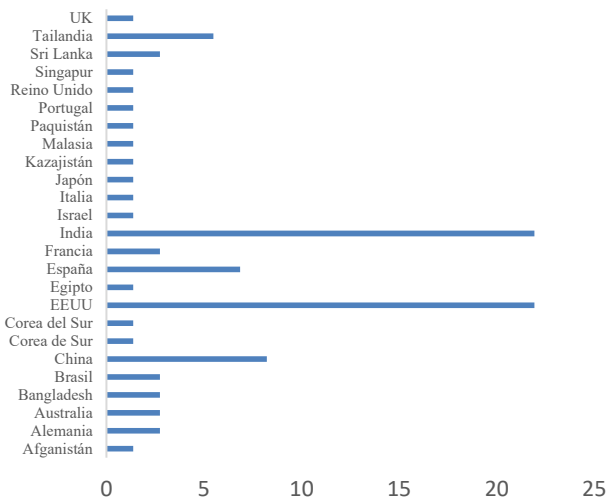


Figura 3. Porcentaje por país de origen de los documentos analizados

El 100% de los documentos incluidos en este estudio se encuentra publicado en inglés. Por su parte, para el análisis de datos, algunas investigaciones utilizaron textos en su lengua materna (Katchapakirin et al., 2018, Eldin et al., 2019; Priya Sri et al., 2021), otros usan textos en su lengua materna traducidos al inglés (López Úbeda et al., 2019; Mehedy et al., 2021; Chatrinan et al., 2021), y en otros, este dato no se indicaban específicamente en los estudios (Wang et al., 2020; Ricard & Hassanpour, 2021; Sabina et al., 2021; Ragheb et al., 2021) (Figura 4).

3.2 Áreas del conocimiento de las publicaciones y redes sociales utilizadas en los estudios

Referente a las áreas del conocimiento y/o temáticas de las publicaciones, se identifica el predominio de revistas relacionadas con Ciencias aplicadas, ingeniería y tecnología, seguida por Psicología, comportamiento humano y salud mental (Figura 5).

3.3 Red social, tipo de dato, alteración de salud mental y algoritmos

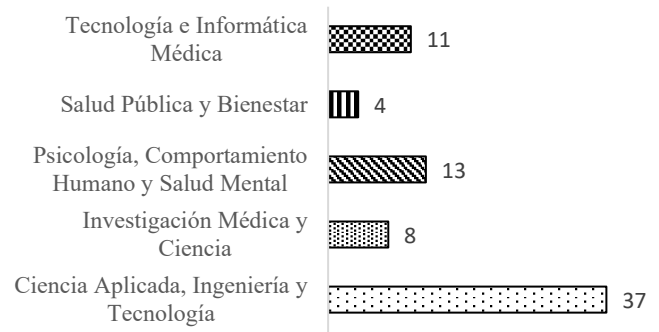


Figura 5. Tema general de la revista y número de artículos relacionados con dicha temática

En lo que tiene que ver con las redes sociales más utilizadas en los estudios, y que permitieron la recopilación de datos para su posterior análisis, en su orden correspondieron a, Twitter (48,4 %), Reddit (25,8 %) y Facebook (12,9%) (Figura 6).

El tipo de dato analizado en el 100% de los estudios corresponde a textos, sin embargo, el 6,3% de los estudios revisados incluyó, de forma adicional, análisis de imágenes como fotografías, emoticones o ilustraciones.

Respecto a las tecnologías implementadas para el NLP, se observó un uso frecuente de algoritmos de machine learning, deep learning y softwares o librerías diseñadas para el análisis del lenguaje natural; entre los más frecuentemente implementados destacan para la etapa de preprocesamiento el Word2Vec con el 25% de los estudios, TF-IDF (25%) y el GloVe (18%), para la etapa de clasificación el LIWC (39%), CNN (32%), LSTM (33%), RF (32%), SVM (31%). Es conveniente señalar que muchos estudios planteaban su modelo con más de un algoritmo en sus metodologías de análisis (Figura 6).

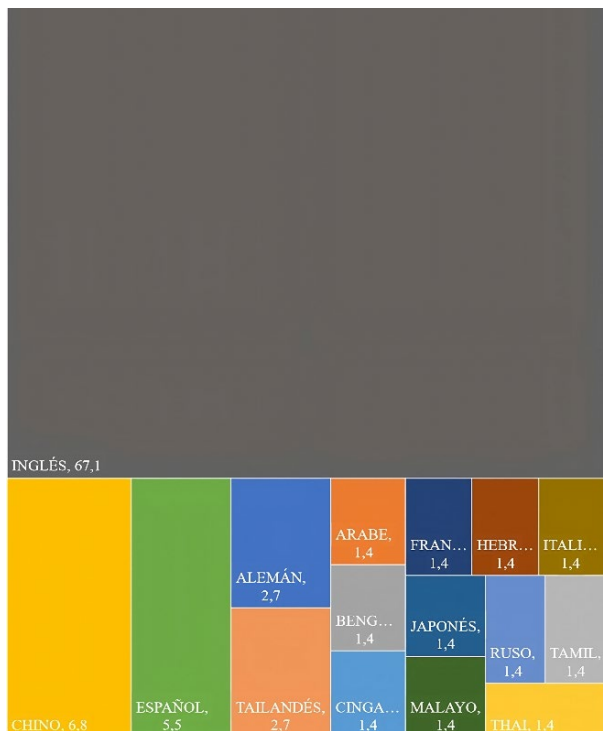


Figura 4. Porcentaje del idioma de los textos analizados en los estudios incluidos

Tabla 1. Glosario de tecnologías usadas			Cont Tabla 1. Glosario de tecnologías usadas		
GRUPO	SIGLA	ALGORÍTMO	GRUPO	SIGLA	ALGORÍTMO
Preprocesamiento Tokenizado	Bow	Bag of Words		RF	Random Forests
	CBoW	Continuous Bag of Words		SGD	Stochastic Gradient Descent
Preprocesamiento Vectorizado	GloVe	Global Vectors for Word Representation		SUN	Semantic Understanding Network
	Word2Vec	Algoritmo Word2vec		SVM	Support Vector Machine
TPreprocesamiento Vectorizado	XLM-R	Cross-lingual Language Model Pretraining		XLNet	Generalized Autoregressive Pretraining for Language Understanding
	TF-IDF	Term frequency – Inverse document frequency			
Categorización	CNN	Convolutional Neural Network	Categorización		
	GRU	Gated Recurrent Unit		LIWC	Linguistic Inquiry and Word Count
	HAN	Hierarchical Attention Network		NLTK	Natural Language Toolkit
	LSTM	Long Short-Term Memory			
	MLP	Multilayer Perceptron		SMOTE	Synthetic Minority Oversampling Technique
	NCNL	Negatively Correlated Noisy Learners		VADER	Valence Aware Dictionary and Sentiment Reasoner
	RMN	Relationship Modeling Network			
	RNN	Recurrent Neural Network			
	AVNNET	Neural Networks Using Model Averaging			
	BERT	Bidirectional Encoder Representations from Transformers			
	CGB	Powell-Beale Algorithm			
	DT	Decision Tree			
	EUN	Emotional Understanding Network			
	FP-Tree	Frequent Pattern Tree			
	KNN	K-Nearest-Neighbor			
LDA	Latent Dirichlet Allocation				
LR	Logistic Regression				
MIL	Multiple Instance Learning				
NB	Naive Bayes				

El uso de distintos algoritmos para el NLP configura el modelo implementado en los análisis de las alteraciones de salud mental, en donde a su vez, destacan las unidades de análisis, comprendidas por las redes sociales contempladas en el estudio como fuente de información, la forma de extracción de la data (Figura 6), siendo en muchos casos grupales (71,6%) a través de un foro temático, corpus preexistentes o un hashtag y/o individuales (28,4%), de lo cual se consiguió la información publicada por usuarios específicos (Tabla 2).

3.4 Fenómenos de Salud mental estudiados en las investigaciones

Finalmente, en lo que corresponde a los fenómenos psicológicos, se identificó un notable interés, por parte de los estudios, en las alteraciones de salud mental relacionadas con factores emocionales (Figura 6), siendo la Depresión (33,9%), la Ideación Suicida (16,1%) y la Ansiedad (9,8%), las temáticas que con mayor frecuencia son reportadas como objeto de análisis en los estudios seleccionadas. Cabe mencionar que, fenómenos como Estrés (4,5%), la Bipolaridad (3,6%), la Soledad (3,6%) los Desórdenes Alimenticios (3,6%) y el Cyberbullying (3,6%), son temáticas que han venido posicionándose como fenómenos de interés y pueden resultar mucho más interesantes en el desarrollo de estudios posteriores (Figura 6).

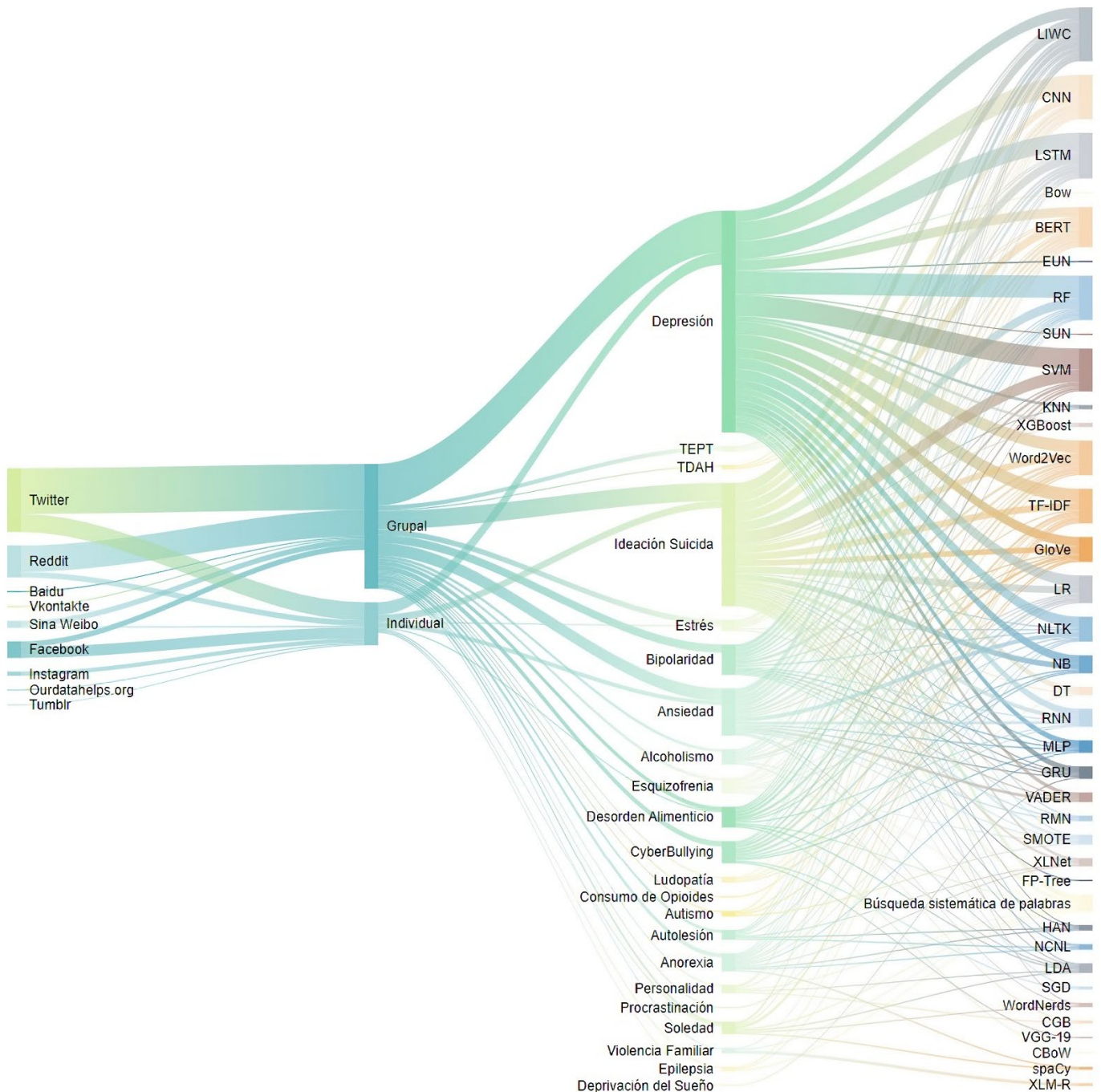


Figura 6. Frecuencia en la red social, tipo de dato, alteración de salud mental y algoritmos implementados en los documentos analizados

Tabla 2. Autores por uso de algoritmo y alteración de salud mental más frecuentes

ALGORITMO	ANSIEDAD	BIPOLARIDAD	DEPRESIÓN	IDEACIÓN SUICIDA	SOLEDAD
Bow			Ferreira et al., 2022		
Búsqueda sistemática de palabras			Ophir et al.,2019; Ren et al., 2021; Ramirez-Cifuentes et al., 2021; Narynov et al.,2020	Ramírez-Cifuentes et al., 2021; Ambalavan et al., 2019	Guntuku et al.,2019
GloVe	Silveira et al., 2018	Silveira et al., 2018	Ren et al., 2021; Wongkoblap et al., 2021; Ramirez-Cifuentes et al., 2021; Silveira et al., 2018; Uban et al., 2021; Ferreira et al., 2022; Nadeem et al., 2022	Yao et al., 2020; Ramirez-Cifuentes et al., 2021; Silveira et al., 2018; Uban et al., 2021	
spaCy					Koh y Liew, 2020;

Cont. **Tabla 2.** Autores por uso de algoritmo y alteración de salud mental más frecuentes.

ALGORITMO	ANSIEDAD	BIPOLARIDAD	DEPRESIÓN	IDEACIÓN SUICIDA	SOLEDAD
TF-IDF	Mendu et al., 2020;		Guntuku et al., 2019; Al Asad et al., 2019; Li et al., 2022; Sun y Luo, 2022; Kour y Gupta, 2022; Kumar et al., 2022; Sarkar et al., 2022; Kumar y Nisha, 2022; Ferreira et al., 2022	Yao et al., 2020; Sun y Luo, 2022; Kour y Gupta, 2022; Chatterjee et al., 2022; Noraset et al., 2022	Guntuku et al., 2019; Mendu et al., 2020;
Word2Vec	Kim et al., 2020; Nasrullah y Jalali, 2022	Kim et al., 2020; Nasrullah y Jalali, 2022	Joshi y Patwardhan, 2020; Wongkoblap et al., 2021; Ramírez-Cifuentes et al., 2021; Narynov et al., 2020; Eldin et al., 2019; Kour y Gupta, 2022; Tejaswini et al., 2022; Chadha y Kaushik, 2022; Kour y Gupta, 2022	Yao et al., 2020; Ramírez-Cifuentes et al., 2021; Glaser et al., 2020; Kour y Gupta, 2022; Liu et al., 2022	
BERT	Silveira et al., 2021; Viviani et al., 2021; Zanwar, et al., 2022	Silveira et al., 2021; Zanwar, et al., 2022	Silveira et al., 2021; Ramírez-Cifuentes et al., 2021; Chatrinan et al., 2021; Ragheb et al., 2021; Prince y Srinivas, 2022; Yang et al., 2022; Kumar y Nisha, 2022;	Silveira et al., 2021; Ramírez-Cifuentes et al., 2021; Chatrinan et al., 2021; Ragheb et al., 2021; Prince y Srinivas, 2022; Noraset et al., 2022; Schoene et al., 2022	
CNN	Kim et al., 2020; Wang et al., 2020; Nasrullah y Jalali, 2022	Kim et al., 2020; Wang et al., 2020; Nasrullah y Jalali, 2022	Wongkoblap et al., 2021; Ramírez-Cifuentes et al., 2021; Chatrinan et al., 2021; Wang et al., 2020; Tejaswini et al., 2022; Prince y Srinivas, 2022; Rego et al., 2021; Chadha y Kaushik, 2022; Kour y Gupta, 2022; Yang et al., 2022; Kumar y Nisha, 2022; Meena y Thulasi Bai, 2022; Saini et al., 2022; Nadeem et al., 2022	Yao et al., 2020; Ramírez-Cifuentes et al., 2021; Chatrinan et al., 2021; Sun y Luo, 2022; Prince y Srinivas, 2022; Schoene et al., 2022	
DT			Chiong et al., 2021; Li et al., 2022; Chanda et al., 2022; Kumar y Nisha, 2022	Liu et al., 2022	
EUN			Ren et al., 2021		
FP-Tree				Glaser et al., 2020	
GRU	Silveira et al., 2021	Silveira et al., 2021	Silveira et al., 2021; Wongkoblap et al., 2021; Yang et al., 2022; Nadeem et al., 2022	Silveira et al., 2021	
HAN			Uban et al., 2021	Uban et al., 2021	
KNN			Li et al., 2022; Chanda et al., 2022	Lekkas et al., 2021	
LDA	Mendu et al., 2020;		Kumar et al., 2022	Chatterjee et al., 2022	Mendu et al., 2020;
LIWC	Mendu et al., 2020; Silveira et al., 2021; Viviani et al., 2021; Wang et al., 2020; Zanwar, et al., 2022	Silveira et al., 2021; Wang et al., 2020; Zanwar, et al., 2022	Guntuku et al., 2019; Silveira et al., 2021; Wongkoblap et al., 2021; Ramírez-Cifuentes et al., 2021; Uban et al., 2021; Wang et al., 2020; Prince y Srinivas, 2022; Chanda et al., 2022	Lekkas et al., 2021; Silveira et al., 2021; Ramírez-Cifuentes et al., 2021; Uban et al., 2021; Arilla et al., 2022; Prince y Srinivas, 2022; Liu et al., 2022; Schoene et al., 2022	Guntuku et al., 2019; Mendu et al., 2020;
LR			Chiong et al., 2021; Ramírez-Cifuentes et al., 2021; Sun y Luo, 2022; Kumar et al., 2022; Yang et al., 2022; Sarkar et al., 2022; Kumar y Nisha, 2022; Ferreira et al., 2022	Yao et al., 2020; Ramírez-Cifuentes et al., 2021; Ambalavan et al., 2019; Sun y Luo, 2022; Chatterjee et al., 2022; Liu et al., 2022	
LSTM	Nasrullah y Jalali, 2022; Zanwar, et al., 2022; Chen et al., 2022	Nasrullah y Jalali, 2022; Zanwar, et al., 2022	Ren et al., 2021; Narynov et al., 2020; Priya Sri et al., 2021; Chatrinan et al., 2021; Uban et al., 2021; Eldin et al., 2019; Tejaswini et al., 2022; Chadha y Kaushik, 2022; Kour y Gupta, 2022; Yang et al., 2022; Chen et al., 2022; Meena y Thulasi Bai, 2022; Nadeem et al., 2022	Coppersmith et al., 2018; Chatrinan et al., 2021; Uban et al., 2021; Arilla et al., 2022; Noraset et al., 2022; Chen et al., 2022; Schoene et al., 2022	

Cont. Tabla 2. Autores por uso de algoritmo y alteración de salud mental más frecuentes

ALGORITMO	ANSIEDAD	BIPOLARIDAD	DEPRESIÓN	IDEACIÓN SUICIDA	SOLEDAD
MLP	Wang et al., 2020	Wang et al., 2020	Chiong et al.,2021; Wongkoblap et al., 2021; Mehedy et al., 2022; Wang et al., 2020	Ambalavan et al., 2019	
NB			Mehedy et al., 2022; Al Asad et al., 2019; Li et al., 2022; Sun y Luo, 2022; Kour y Gupta, 2022; Ferreira et al., 2022	Lekkas et al., 2021; Sun y Luo, 2022; Noraset et al., 2022; Liu et al., 2022	
NCNL			Ragheb et al., 2021	Ragheb et al., 2021	
NLTK	Kim et al., 2020: Nasrullah y Jalali, 2022; Chen et al, 2022	Kim et al., 2020; Nasrullah y Jalali, 2022	Chiong et al.,2021; Joshi y Patwardhan, 2020; Katchapakirin er al., 2018; Al Asad et al., 2019; Tejaswini et al., 2022; Rego et al., 2021; Yang et al., 2022; Chen et al, 2022	Ambalavan et al., 2019; Chen et al, 2022	
RF	Wang et al., 2020	Wang et al., 2020	Chiong et al.,2021; Guntuku et al.,2019; Ramirez-Cifuentes et al., 2021; Katchapakirin er al., 2018; Mehedy et al., 2022; Chatrinan et al., 2021; Wang et al., 2020; Sun y Luo, 2022; Kour y Gupta, 2022; Chanda et al.,2022; Kumar et al., 2022; Kour y Gupta, 2022; Sarkar et al., 2022; Kumar y Nisha, 2022; Ferreira et al., 2022; Saini et al., 2022	Yao et al., 2020; Ramírez-Cifuentes et al., 2021; Chatrinan et al., 2021; Sun y Luo, 2022; Sun y Luo, 2022; Kour y Gupta, 2022; Liu et al., 2022	Guntuku et al.,2019
RMN	Silveira et al., 2018	Silveira et al., 2018	Silveira et al., 2018	Silveira et al., 2018	
RNN	Silveira et al., 2021; Chen et al, 2022	Silveira et al., 2021	Silveira et al., 2021; Priya Sri et al. 2021; ElDin et al., 2019; Chen et al, 2022; Kumar y Nisha, 2022	Silveira et al., 2021; Yao et al., 2020; Chen et al, 2022	
SGD				Ambalavan et al., 2019	
SMOTE	Kim et al., 2020	Kim et al., 2020	Chatrinan et al., 2021	Chatrinan et al., 2021; Noraset et al., 2022	
SUN			Ren et al., 2021		
SVM	Wang et al., 2020; Chen et al, 2022	Wang et al., 2020	Chiong et al.,2021; Katchapakirin er al., 2018; Mehedy et al., 2022; Chatrinan et al., 2021; Wang et al., 2020; Al Asad et al., 2019; Kour y Gupta, 2022; Prince y Srinivas, 2022; Chanda et al.,2022; Kumar et al., 2022; Sarkar et al., 2022; Chen et al, 2022; Kumar y Nisha, 2022; Ferreira et al., 2022; Saini et al., 2022	Yao et al., 2020; Ambalavan et al., 2019; Chatrinan et al., 2021; Kour y Gupta, 2022; Prince y Srinivas, 2022; Chatterjee et al., 2022; Noraset et al., 2022; Liu et al., 2022; Chen et al, 2022	
VADER	Silveira et al., 2021; Viviani et al., 2021	Silveira et al., 2021	Silveira et al., 2021; Li et al., 2022; Ferreira et al., 2022	Silveira et al., 2021	
WordNerds	Marshall et al., 2022		Marshall et al., 2022		Marshall et al., 2022
XLNet			Ragheb et al., 2021; Prince y Srinivas, 2022	Ragheb et al., 2021; Prince y Srinivas, 2022	
XGBoost			Lekkas et al., 2021; Kim et al., 2020; Kumar et al., 2022; Sarkar et al., 2022; Saini et al., 2022	Chatterjee et al., 2022	

4. DISCUSIÓN DE RESULTADOS

De acuerdo con los resultados obtenidos en la revisión, en primer lugar, se destaca el número de publicaciones sobre procesamiento del lenguaje natural para el análisis de fenómenos de salud mental a partir de datos en redes sociales durante de los años de observación, siendo los años 2022 y 2021, los de mayor producción científica sobre el tema. Dicho aumento en los últimos dos años se atribuye como producto de

la pandemia de Covid-19 y pos pandemia, no sólo porque algunos estudios indagaban sobre afectaciones emocionales por el Covid a través de redes sociales (Babvey et al. 2021), sino que, ante la dificultad de realizar estudios de forma presencial, se dio prioridad al análisis del comportamiento por medios digitales como redes sociales, ya que esto no implicaba un contacto directo con los participantes (Hlatshwako et al., 2021).

Por su parte, en lo referente al idioma de publicación, todos los documentos consultados fueron publicados originalmente en inglés, a pesar que los orígenes de las investigaciones fueron de países no angloparlantes como Alemania (Fazekas et al., 2021), Francia (Ambalavan et al., 2019), Italia (Viviani et al., 2021) o España (Sabina et al., 2021; Ardila et al., 2022) en Europa; Bangladesh (Mehedy et al., 2021), China (Li et al., 2021; Sun & Luo, 2022), Corea del sur (Kim et al., 2020), India (Joshi & Patwardhan, 2020; Kour & Gupta, 2022; Tejaswini et al., 2022), Israel (Ophir et al., 2019), Japón (Mori & Haruno, 2021), Kazajistán (Narynov et al., 2020), Malasia (Priya et al., 2021), Singapur (Koh & Liew, 2020), Sri Lanka (Perera, & Fernando, 2021) y Tailandia (Katchapakirin et al., 2018) en Asia; Brasil (Silverira et al., 2021) en Latinoamérica o Egipto (ElDin et al., 2019) en África.

De esta manera, la mayoría de los estudios realizaron sus análisis sobre textos en inglés, siendo algunos de ellos extraídos en su idioma original y luego traducidos al inglés para el análisis (Chatrinan et al., 2021; ElDin et al., 2019; Katchapakirin et al., 2018; López et al., 2019; Narynov et al. 2020; Priya Sri et al., 2021). Todo ello se debe a que, gran parte de las tecnologías para el análisis del lenguaje natural fue desarrollada para la interpretación de textos en inglés como, por ejemplo, el Word2Vec, BoW o el NLTK, por lo cual muchas de estas herramientas no están adaptadas a las particularidades de otros idiomas (Chatrinan et al., 2021).

En cuanto a las temáticas de las publicaciones, se identificó que, la ciencia aplicada, la ingeniería y la tecnología, son el interés principal de las revistas que publicaron los documentos seleccionados, aunque también, con menor frecuencia, se encontraron las revistas con temática de psicología, comportamiento humano y salud mental, campos que deberían resultar más notables, tomando en cuenta que la salud mental es tópico fundamental en su área de estudio. Este fenómeno puede atribuirse al conocimiento especializado que se requiere para la extracción, la depuración y el procesamiento del lenguaje natural de los datos, tareas mucho más cotidianas en las ciencias informáticas que en las ciencias sociales y de la salud, donde pueden parecer distantes, poco frecuentes o desconocidas (Urban & Gates, 2021).

En lo referente a las fuentes de información en los estudios analizados, se identifica que las redes sociales más usadas fueron Twitter y Reddit, seguida de Facebook, cuyo resultado es similar al obtenido por Zhang et al. (2022), quienes consideran que estas redes sociales, además de ser de las más estudiadas y populares (Harrigian et al., 2020), se caracterizan por la predominancia de textos, organizando la información a través de palabras clave, hashtags, hilos, etc. (Ricard & Hassanpour, 2021). Adicionalmente, Twitter permite la recopilación de información por medio de su API (Interfaz de programación de aplicaciones) y Reddit organiza las temáticas por medio de foros denominados subreddits, lo que facilita la extracción de los datos relevantes para cada investigación, tarea mucho más compleja en redes sociales con predominancia audiovisual como Facebook, Instagram o Tumblr, las cuales son infrutilizadas en investigaciones de salud mental (Harrigian et al., 2020).

Los fenómenos de salud mental abordados en los documentos seleccionados evidencian un especial interés por las alteraciones emocionales, siendo la depresión, la ideación suicida y la ansiedad, las temáticas con mayor frecuencia en las investigaciones, y cuyo resultado es coherente con el estudio realizado por Chancellor & De Choudhury (2020), en donde la depresión y el suicidio también resultan ser los fenómenos más recurrentes en su análisis. Así mismo, las alteraciones emocionales presentan una alta prevalencia en la población mundial, de acuerdo al informe de la OMS (2017) y cuya situación se vio exacerbada en el periodo de confinamiento y distanciamiento social por Covid-19, lapso en el cual se reporta una mayor prevalencia de sintomatología depresiva, ansiedad, comportamiento suicida, entre otras alteraciones de salud mental en la población a nivel global (Huarcaya-Victoria, 2020; Kumar & Nayar, 2021; Mac-Ginty et al., 2021).

Respecto al procesamiento de los datos, se identifican dos grupos de algoritmos fundamentales para el NLP; en primer lugar los de ordenamiento, en donde se limpia la data eliminando la información irrelevante como los emoticones o signos de puntuación, se tokeniza la data, se divide el contenido en palabras individuales y se convierte ("No, es triste") en un arreglo ["No", ",", "es", "triste"] y se vectoriza, proceso en el que se convierten los tokens en vectores numéricos comprensibles para los algoritmos de aprendizaje automático, modelos que generalmente requieren datos de entrada numérica. Entre los más frecuentemente usados se detectaron el TF-IDF, Word2Vec, GloVe.

En segundo lugar, entre los algoritmos de clasificación usados para la identificación de alteraciones psicológicas, en los documentos revisados, se identifica el SVM, RF, CNN, LIWC y el LSTM, siendo este último el que mayor precisión evidenció en el análisis de las alteraciones emocionales como la depresión y la ideación suicida (Zhang et al., 2017; ElDin et al., 2019; Ren et al., 2021; Chadha & Kaushik, 2022; Kour & Gupta, 2022; Yang et al., 2022). Por su parte Gong et al. (2018), indica que el uso de LIWC puede superar la dificultad de obtener datos a gran escala en el campo de la salud, esto es especialmente relevante en el entrenamiento de modelos de Deep Learning, ya que un diccionario creado por humanos puede proporcionar una alternativa efectiva y concisa al aprendizaje automático para convertir texto en representaciones vectoriales.

A pesar de las grandes ventajas que el análisis de textos de redes sociales para la identificación de alteraciones psicológicas a través de NLP representa, es importante ejercer prudencia en su uso y destacar la necesidad de contar con el acompañamiento de profesionales en salud mental para un diagnóstico adecuado (Calvo et al., 2017). Si bien esta herramienta puede ser útil para la detección temprana de alteraciones psicológicas como lo promueve el proyecto eRisk (Crestani et al., 2022), se debe tener en cuenta que los datos analizados pueden estar sujetos a sesgos, por ejemplo, al considerar variables de personalidad como la "triada oscura", que con frecuencia tiene una presencia más activa en las redes sociales (Preotiuc-Pietro et al., 2016), lo que puede generar una mayor representación en los datos recolectados. Así mismo, el exacerbar el control de la información digital para

mantener la privacidad, debe ser un compromiso implícito establecido en los estudios, ya que estos pueden ser utilizados para ser analizados y comercializados en mercados secundarios, generando preocupaciones sobre las implicaciones éticas del uso de datos públicos, la calidad de estos, las consecuencias de una mala interpretación y la vulneración de la privacidad (Bauer et al., 2017).

Particularmente en los aspectos éticos, Wongkoblap et al. (2017) en su estudio, destacan la inexistencia de una ruta clara en el uso de la data pública proveniente de redes sociales, ya que, se evidencian prácticas adoptadas de forma autónoma en cada estudio, que varían de un investigador a otro, lo que resulta similar a lo encontrado en la presente investigación, en donde pocas publicaciones, principalmente aquellas que analizaron los datos de usuarios individuales, contaron algunas con avales éticos (Fazekas et al., 2021; Guntuku et al., 2019; Katchapakirin et al., 2018; Kim et al., 2020; Mori & Haruno, 2021; Ambalavan et al., 2019; Ramírez-Cifuentes et al., 2021; Schoene et al., 2022; Wongkoblap et al., 2021), mientras que la gran mayoría, interesados en el análisis de forma grupal, no detallan ninguna consideración ética en sus estudios.

Lo anterior, conduce a reflexionar acerca de la importancia de tener en cuenta, para este tipo de investigaciones, aspectos como, las fuentes de sesgo, las implicaciones éticas, y el acompañamiento de profesional del área de la salud, en el análisis automatizado, y obtener con ello, una comprensión completa, adecuada y responsable del fenómeno.

5. CONCLUSIONES

El interés de esta investigación se encaminó hacia la detección de modelos de lenguaje natural para la identificación de alteraciones mentales en redes sociales. Dicha detección precoz de síntomas de trastornos mentales es una estrategia importante y eficaz para otorgar datos que aportan al diagnóstico de la salud mental en la población., a partir de los hallazgos se puede concluir que:

Existe un aumento significativo en la cantidad de publicaciones sobre el procesamiento del lenguaje natural para el análisis de fenómenos de salud mental en datos de redes sociales, especialmente en los años 2021 y 2022 durante la pandemia de COVID-19.

La mayoría de los estudios consultados fueron publicados originalmente en inglés, a pesar de que las investigaciones se realizaron en diversos países no angloparlantes de Europa, Asia, Latinoamérica, Oceanía y África.

Las revistas con temáticas de ciencia aplicada, ingeniería y tecnología, son las que predominan en la publicación de los documentos, en comparación con las revistas especializadas en psicología, comportamiento humano y salud mental.

Las redes sociales más utilizadas para la recopilación de datos en los estudios son Twitter y Reddit, debido a su predominancia de texto y facilidad para la extracción de información relevante. En contraste, redes sociales audiovisuales como Facebook, Instagram y Tumblr, son menos utilizadas en investigaciones de salud mental. Así

mismo, la extracción de datos de forma grupal es el común denominador de los estudios.

Los trastornos emocionales como la depresión, la ideación suicida y la ansiedad, son los temas más frecuentemente abordados en los estudios consultados, lo cual es coherente con la alta prevalencia de estos trastornos a nivel mundial, especialmente durante la pandemia de COVID-19.

En cuanto al procesamiento de datos, se identifican dos grupos principales de algoritmos utilizados en el procesamiento del lenguaje natural: los algoritmos de ordenamiento para limpiar y preparar los datos, y los algoritmos de clasificación, siendo los algoritmos SVM, RF, CNN, LIWC y LSTM los más usados, para la identificación de alteraciones psicológicas. Cabe mencionar que los modelos propuestos por los estudios contemplan una combinación de ellos y de forma general todos reportaban efectividades elevadas.

En el futuro, el desarrollo de nuevos métodos que incluyan diferentes estrategias de aprendizaje, nuevos paradigmas de procesamiento de lenguaje natural, modelos y métodos multimodales, validados y aplicados en distintos contextos, apoyarán la detección de enfermedades mentales con un énfasis en la interpretabilidad, lo que resulta crucial para la aceptación de las aplicaciones de detección por parte de profesionales de la salud.

Así mismo, es importante destacar algunas preocupaciones sobre el componente ético implementado en estos estudios, dado que no se identificó una regulación clara sobre la recopilación y el manejo de datos, sino más bien, una serie heterogénea de estrategias adelantadas por iniciativa de los mismos autores. Esto genera interrogantes sobre las consecuencias que esta información pueda tener en contextos sanitarios, políticos y sociales, llevando a sugerir, para futuras investigaciones, con similares características, la consulta de publicaciones como la de Arigo et al. (2018) quien propone algunas directrices éticas relevantes para el desarrollo de estos estudios.

Como limitaciones en este estudio, se destaca que la selección de artículos se llevó a cabo a través de unas bases de datos, dejando de lado algunas otras que bien pueden brindar información complementaria. Sólo se seleccionaron los artículos relacionados con los datos extraídos de redes sociales, y aquellos que utilizan NLP para la detección de alteraciones en salud mental.

Los retos que deja esta investigación podrán abordarse en investigaciones futuras sobre la detección de problemas de salud mental. Aspectos como, los modelos, ubicaciones de los estudios, el idioma, la diversidad de fuentes de datos, el tipo de extracción de datos, no sólo de textos, sino también, imágenes, audios y videos, se convertirán en áreas potenciales de exploración, ya que, como bien se ha revelado en este estudio, la utilización de algoritmos o híbridos para la detección de enfermedades mentales o emocionales, se podrán convertir en fuentes de información que aporten a la formulación de alternativas de prevención, promoción y control, en políticas públicas.

Por último, se debe manifestar que la presente investigación ofrece información sobre una serie de herramientas de aprendizaje automático que emplean muchas de las técnicas de NLP, capaces de ofrecer a las instituciones u organizaciones de salud competentes una ayuda sistemática en el rastreo de la actividad de los usuarios en las redes sociales. Se destaca las herramientas acá expuestas, las cuales demuestran la capacidad de generar alertas tempranas ante evidencia de padecimientos asociados a algún tipo de trastorno mental, aspecto que potencia una intervención anticipada y oportuna, además de dar cabida al desarrollo de alternativas en el tratamiento temprano de trastornos psicológicos.

Los resultados son prometedores y arrojan luz sobre muchas aplicaciones en el mundo real para la detección precoz de alteraciones mentales a partir de los datos de las redes sociales.

Declaración de intereses concurrentes

Los autores declaran que no tienen intereses contrapuestos.

Agradecimientos

Los autores agradecen a la Universidad de Investigación y Desarrollo -UDI por su apoyo administrativo durante este estudio. Las interpretaciones, conclusiones y análisis son de exclusiva responsabilidad de los autores y no comprometen ni vinculan de ninguna manera a la Universidad de Investigación y Desarrollo -UDI o a las instituciones a las que los autores están afiliados.

Detalles de la financiación

Este trabajo fue financiado por la Universidad de Investigación y Desarrollo -UDI, Bucaramanga, Colombia.

REFERENCIAS

- Al Asad, N., Pranto, M., Afreen, S., & Islam, M. (2019). Depression detection by analyzing social media posts of user. *International Conference on Signal Processing, Information, Communication & Systems*, 13-17. <http://dx.doi.org/10.1109/SPICSCON48833.2019.9065101>
- Ambalavan, A. K., Moulahi, B., Azé, J. & Bringay, S. (2019). Unveiling online suicide behavior: What can we learn about mental health from suicide survivors of Reddit?. *MedInfo*, 264(1), 50-54. <https://doi.org/10.3233/SHTI190181>
- Arigo, D., Pagoto, S., Carter-Harris, L., Lillie, S., & Nebeker, C. (2018). Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery. *Digital health*, 4, 1-15. <https://doi.org/10.1177/2055207618771757>
- Arilla-Andrés, S., García-Martínez, C., & Hoyo, Y. L. Del. (2022). Detection of Suicide Risk Through Social Media: Pilot Study. *Revista Internacional de Tecnología, Ciencia y Sociedad*, 11. <https://doi.org/10.37467/revtechno.v11.4384>
- Babvey, P., Capela, F., Cappa, C., Lipizzi, C., Petrowski, N. & Ramirez-Marquez, J. (2021). Using social media data for assessing children's exposure to violence during the COVID-19 pandemic. *Child Abuse & Neglect*, 116(2), 1-14. <https://doi.org/10.1016/j.chiabu.2020.104747>
- Bae, Y., Shim, M. & Lee, W. (2021). Schizophrenia Detection Using Machine Learning Approach from Social Media Content. *Sensors*, 21(17), 1-18. <https://doi.org/10.3390/s21175924>
- Bauer, M., Glenn, T., Monteith, S., Bauer, R., Whybrow, P. C., & Geddes, J. (2017). Ethical perspectives on recommending digital technology for patients with mental illness. *International journal of bipolar disorders*, 5(1), 1-14. <https://doi.org/10.1186/s40345-017-0073-9>
- Calvo, R., Milne, D., Hussain, M., & Christensen, H. (2017). Natural language processing in mental health applications using non-clinical texts. *Natural Language Engineering*, 23(5), 649-685. <https://doi.org/10.1017/S1351324916000383>
- Camacho, J., Moreno, S., Suarez-Obando, F., Puyana, J., & Gómez-Restrepo, C. (2013). El procesamiento de lenguaje natural y su relación con la investigación en salud mental. *Revista Colombiana de Psiquiatría*, 42(2), 227-233. [https://doi.org/10.1016/S0034-7450\(13\)70011-8](https://doi.org/10.1016/S0034-7450(13)70011-8)
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329(5996), 1194-1197. <https://doi.org/10.1126/science.1185231>
- Chadha, A., & Kaushik, B. (2022). A Hybrid Deep Learning Model Using Grid Search and Cross-Validation for Effective Classification and Prediction of Suicidal Ideation from Social Network Data. *New Generation Computing*, 40(4), 889-914. <https://doi.org/10.1007/s00354-022-00191-1>
- Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3, 1-11. <https://doi.org/10.1038/s41746-020-0233-7>
- Chanda, K., Roy, S., Mondal, H., & Bose, R. (2022). To Judge Depression and Mental Illness on Social Media Using Twitter. *Universal Journal of Public Health*, 10(1), 116-129. <https://doi.org/10.13189/ujph.2022.100113>
- Chatrinan, K., Kangpanich, A., Wichit, T., Noraset, T., Tuarob, S., & Tawichsri, T. (2021). Towards Approximating Population-Level Mental Health in Thailand Using Large-Scale Social Media Data. In *International Conference on Asian Digital Libraries*, 13133, 334-343. https://doi.org/10.1007/978-3-030-91669-5_26
- Chatterjee, M., Samanta, P., Kumar, P., & Sarkar, D. (2022). Suicide Ideation Detection using Multiple Feature Analysis from Twitter Data. *2022 IEEE Delhi Section Conference, DELCON 2022, February*. <https://doi.org/10.1109/DELCON54057.2022.9753295>
- Chen, Z., Zhang, R., Xu, T., Yang, Y., Wang, J., & Feng, T. (2020). Emotional attitudes towards procrastination in people: A large-scale sentiment-focused crawling analysis. *Computers in Human Behavior*, 110, 1-11. <https://doi.org/10.1016/j.chb.2020.106391>
- Chiong, R., Budhi, G., 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>
- Cobo, M., Lopez-Herrera, A. Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of*

- the American Society for Information Science and Technology*, 62, 1382–1402. <https://doi.org/10.1002/asi.21525>
- Confederación Salud Mental España (2019). La salud mental en cifras. <https://comunicalasaludmental.org/guiadeestilo/la-salud-mental-en-cifras/>
- Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural language processing of social media as screening for suicide risk. *Biomed Inform Insights*, 10, 1-11 <https://doi.org/10.1177/1178222618792860>
- Crestani, F., Losada, D., & Parapar, J. (Ed.). (2022). *Early Detection of Mental Health Disorders by Social Media Monitoring: The First Five Years of the ERisk Project*. Springer Nature.
- Dos Santos, B., Steiner, M., Fenerich, A., & Lima, R. (2019). Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. *Computers & Industrial Engineering*, 138. <https://doi.org/10.1016/j.cie.2019.106120>
- ElDin, D., Taha, M., & Khalifa, N. (2019). SentiNeural: A Depression Clustering Technique for Egyptian Women Sentiments. *International Journal of Advanced Computer Science and Applications*, 10(5), 550-555. <https://doi.org/10.14569/IJACSA.2019.0100572>
- Fazekas, B., Megaw, B., Eade, D. & Kronfeld, N. (2021). Insights into the real-life experiences of people living with epilepsy: A qualitative ethnographic study. *Epilepsy & Behavior*, 116, 1-8. <https://doi.org/10.1016/j.yebeh.2020.107729>
- Fernández, R. (2020). Panorama mundial de las redes sociales. Statista. <https://es.statista.com/estudio/32777/panorama-mundial-delas-redes-sociales-dossier-statista>.
- Ferreira, R., Trifan, A., & Oliveira, J. L. (2022). Early risk detection of mental illnesses using various types of textual features. *CEUR Workshop Proceedings*, 3180, 905–920. <https://ceur-ws.org/Vol-3180/paper-72.pdf>
- Garg, M. (2021). A survey on different dimensions for graphical keyword extraction techniques. *Artificial Intelligence Review*, 54, 4731–4770. <https://doi.org/10.1007/s10462-021-10010-6>
- Gaur, M., et al., (2019). Knowledge-aware assessment of severity of suicide risk for early intervention. *The world wide web conference*, 514-525. <https://doi.org/10.1145/3308558.3313698>
- Glaser, E., Morain, A., Gemmell, J. & Raicu, D. (2020). Comparing automatically extracted topics from online suicidal ideation and the responses they invoke. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 1818-1825. <https://doi.org/10.1145/3341105.3373902>
- Gong, Y., Shin, K., & Poellabauer, C. (2018, August). Improving LIWC using soft word matching. In *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*, 18, 523-523. <https://doi.org/10.1145/3233547.3233632>
- Guntuku, S., et al., (2019). Studying expressions of loneliness in individuals using twitter: an observational study. *BMJ open*, 9(11), 1-8. <http://dx.doi.org/10.1136/bmjopen-2019-030355>
- Harrigian, K., Aguirre, C., & Dredze, M. (2020). On the state of social media data for mental health research. *arXiv preprint, 1*. <https://doi.org/10.48550/arXiv.2011.05233>
- Hlatshwako, T., Shah, S., Kosana, P., Adebayo, E., Hendriks, J., Larsson, E. C., ... Tucker, J. (2021). Online health survey research during COVID-19. *The Lancet Digital Health*, 3(2), Article e76–e77. [https://doi.org/10.1016/s2589-7500\(21\)00002-9](https://doi.org/10.1016/s2589-7500(21)00002-9)
- Huarcaya-Victoria, J. (2020). Consideraciones sobre la salud mental en la pandemia por COVID 19. *Revista peruana de medicina experimental y salud pública*. 3(2), 327-334. <https://doi.org/10.17843/rpmesp.2020.372.5419>
- Joshi, D., & Patwardhan, M. (2020). An analysis of mental health of social media users using unsupervised approach. *Computers in Human Behavior Reports*, 2, 1-9. <https://doi.org/10.1016/j.chbr.2020.100036>
- Katchapakirin, K., Wongpatikaseree, K., Yomaboot, P., & Kaewpitakkun, Y. (2018). Facebook social media for depression detection in the Thai community. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 1-6. <https://doi.org/10.1109/JCSSE.2018.8457362>
- Kim, J., Lee, J., Park, E., & Han, J. (2020). A deep learning model for detecting mental illness from user content on social media. *Scientific reports*, 10(1), 1-6. <https://doi.org/10.1038/s41598-020-68764-y>
- Koh, J., & Liew, T. (2020). How loneliness is talked about in social media during COVID-19 pandemic: text mining of 4,492 Twitter feeds. *Journal of psychiatric research*. 1-19. <https://doi.org/10.1016/j.jpsychires.2020.11.015>
- Kour, H., & Gupta, M. K. (2022). Depression and Suicide Prediction Using Natural Language Processing and Machine Learning. *Lecture Notes in Networks and Systems*, 370, 117–128. https://doi.org/10.1007/978-981-16-8664-1_11
- Kour, H., & Gupta, M. K. (2022). Predicting the language of depression from multivariate twitter data using a feature-rich hybrid deep learning model. *Concurrency and Computation: Practice and Experience*, 34(24). <https://doi.org/10.1002/cpe.7224>
- Kumar, A., & Nayar, K. (2021). COVID 19 and its mental health consequences. *Journal of Mental Health*, 30(1), 1-2. <https://doi.org/10.1080/09638237.2020.1757052>
- Kumar, P., Samanta, P., Dutta, S., Chatterjee, M., & Sarkar, D. (2022). Feature Based Depression Detection from Twitter Data Using Machine Learning Techniques. *Journal of Scientific Research*, 66(02), 220–228. <https://doi.org/10.37398/jsr.2022.660229>
- Kumar, S., & Nisha, Z. (2022). Does Social Media Feed Tell about Your Mental State? A Deep Randomised Neural Network Approach. *Proceedings of the International Joint Conference on Neural Networks, 2022-July*, 1–8. <https://doi.org/10.1109/IJCNN55064.2022.9892210>
- Lekkas, D., Klein, R., & Jacobson, N. (2021). Predicting acute suicidal ideation on Instagram using ensemble machine learning models. *Internet interventions*, 25, 1-9. <https://doi.org/10.1016/j.invent.2021.100424>
- Li, C., Liu, H., Yin, B., & Yang, J. (2022). Weibo Depression Posts Detection by Natural Language Processing. *Highlights in Science, Engineering and Technology*, 16, 430–437. <https://doi.org/10.54097/hset.v16i.2605>

- Li, Q., Zhao, L., Xue, Y., & Feng, L. (2021). Stress-buffering pattern of positive events on adolescents: An exploratory study based on social networks. *Computers in Human Behavior, 114*, 1-14. <https://doi.org/10.1016/j.chb.2020.106565>
- Liu, J., Shi, M., & Jiang, H. (2022). Detecting Suicidal Ideation in Social Media: An Ensemble Method Based on Feature Fusion. *International Journal of Environmental Research and Public Health, 19*(13), 8197. <https://doi.org/10.3390/ijerph19138197>
- López-Belmonte, J., Marín-Marín, J.-A., Soler-Costa, R. & Moreno-Guerrero, A. (2020). Arduino advances in web of science. A scientific mapping of literary production. *IEEE Access, 8*, 128674–128682. <https://doi.org/10.1109/ACCESS.2020.3008572>
- López-Úbeda, P., Plaza-del-Arco, F., Díaz-Galiano, M., Lopez, L., & Martín-Valdivia, M. (2019). Detecting anorexia in Spanish tweets. *Proceedings of the International Conference on Recent Advances in Natural Language Processing*. 655-663. https://doi.org/10.26615/978-954-452-056-4_077
- Mac-Ginty, S., Jiménez-Molina, A. & Martínez, V. (2021). Impacto de la pandemia por COVID 19 en la salud mental de estudiantes universitarios de Chile. *Revista Chilena de Psiquiatría y neurología de la infancia y la adolescencia, 32*(1), 23-37. <https://psicologia.udp.cl/cms/wp-content/uploads/2021/04/Rev-SOPNIA-2021-23-37.pdf>
- Marín-Marín, J. A., Moreno-Guerrero, A. J., Dúo-Terrón, P., & López-Belmonte, J. (2021). STEAM in education: a bibliometric analysis of performance and co-words in Web of Science. *International Journal of STEM Education, 8*(1), 1-21. <https://doi.org/10.1186/s40594-021-00296-x>
- Marshall, C., Lanyi, K., Green, R., Wilkins, G. C., Pearson, F., & Craig, D. (2022). Using Natural Language Processing to Explore Mental Health Insights From UK Tweets During the COVID-19 Pandemic: Infodemiology Study. *JMIR Infodemiology, 2*(1), 1–14. <https://doi.org/10.2196/32449>
- Martínez, M. A., Cobo, M. J., Herrera, M., & Herrera-Viedma, E. (2015). Analyzing the scientific evolution of social work using science mapping. *Research on Social Work Practice, 25*(2), 257–277. <https://doi.org/10.1177/1049731514522101>
- Meena, R., & Thulasi Bai, V. (2022). Depression Detection on COVID 19 Tweets Using Chimp Optimization Algorithm. *Intelligent Automation and Soft Computing, 34*(3), 1643–1658. <https://doi.org/10.32604/iasc.2022.025305>
- Mehedy, M., Nanda, U. & Faruque, O. (2021). Ranking Mental Illness among Social Media Users. *International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering*, 1-4. <https://doi.org/10.1109/IC4ME253898.2021.9768617>
- Melvin, S., Jamal, A., Hill, K., Wang, W. & Young, S. (2019). Identifying Sleep-Deprived Authors of Tweets: Prospective Study. *JMIR mental health, 6*(12), 1-9. <https://doi.org/10.2196/13076>
- Mendu, S., Baglione, A., Bacc, S., Wu, C., Ng, B., Shaked, A., Clore, G., Boukhechba, M., & Barnes, L. (2020). A framework for understanding the relationship between social media discourse and mental health. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW2), 1-23. <https://doi.org/10.1145/3415215>
- Moessner, M., Feldhege, J., Wolf, M., & Bauer, S. (2018). Analyzing big data in social media: Text and network analyses of an eating disorder forum. *International Journal of Eating Disorders, 51*(7), 656-667. <https://doi.org/10.1002/eat.22878>
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(7), Article e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Mori, K., & Haruno, M. (2021). Differential ability of network and natural language information on social media to predict interpersonal and mental health traits. *Journal of personality, 89*(2), 228-243. <https://doi.org/10.1111/jopy.12578>
- Nadeem, A., Naveed, M., Islam Satti, M., Afzal, H., Ahmad, T., & Kim, K. Il. (2022). Depression Detection Based on Hybrid Deep Learning SSCL Framework Using Self-Attention Mechanism: An Application to Social Networking Data. *Sensors, 22*(24), 1–28. <https://doi.org/10.3390/s22249775>
- Nandhini, B., & Sheeba, J. (2015). Online social network bullying detection using intelligence techniques. *Procedia Computer Science, 45*, 485-492. <https://doi.org/10.1016/j.procs.2015.03.085>
- Narynov, S., Mukhtarkhanuly, D., & Omarov, B. (2020). Dataset of depressive posts in Russian language collected from social media. *Data in Brief, 29*, 105195. <https://doi.org/10.1016/j.dib.2020.105195>
- Nasrullah, S., & Jalali, A. (2022). Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model. *Computational Intelligence and Neuroscience, 2022*. <https://doi.org/10.1155/2022/9404242>
- 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>
- Ophir, Y., Asterhan, C., & Schwarz, B. (2019). The digital footprints of adolescent depression, social rejection and victimization of bullying on Facebook. *Computers in Human Behavior, 91*, 62-71. <https://doi.org/10.1016/j.chb.2018.09.025>
- Organización Mundial de la Salud (2020). Día Mundial de la Salud Mental: una oportunidad para impulsar un aumento a gran escala de la inversión en salud mental. <https://www.who.int/es/news/item/27-08-2020-world-mental-health-day-an-opportunity-to-kick-start-a-massive-scale-up-in-investment-in-mental-health#:~:text=La%20salud%20mental%20es%20una,s e%20suicida%20cada%2040%20segundos>
- Organización Mundial de la Salud. (2017). *Depression and Other Common Mental Disorders*. <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf>
- Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: Purpose, process, and structure. *Journal*

- of the *Academy of Marketing Science*, 46(1), 1-5. <https://doi.org/10.1007/s11747-017-0563-4>
- Perera, A., & Fernando, P. (2021). Accurate Cyberbullying detection and prevention on social media. *Procedia Computer Science*, 181, 605-611. <https://doi.org/10.1016/j.procs.2021.01.207>
- Perestelo-Pérez, L. (2013). Estándares sobre cómo desarrollar y reportar revisiones sistemáticas en psicología y salud. *Revista Internacional de Psicología Clínica y de la Salud*, 13(1), 49-57. [http://dx.doi.org/10.1016/S1697-2600\(13\)70007-3](http://dx.doi.org/10.1016/S1697-2600(13)70007-3)
- Preotiuc-Pietro, D., Carpenter, J., Giorgi, S., & Ungar, L. (2016). Studying the Dark Triad of personality through Twitter behavior. In *Proceedings of the 25th ACM international on conference on information and knowledge management*, 761-770. <http://wwwbp.org/papers/darktriad16cikm.pdf>
- Prince, M. C., & Srinivas, L. N. B. (2022). A Review and Design of Depression and Suicide Detection Model Through Social Media Analytics. 443-455. https://doi.org/10.1007/978-981-16-5652-1_40
- Priya, E., Savita, K., & Zaffar, M. (2021). Depression Detection in Tweets from Urban Cities of Malaysia using Deep Learning. *International Conference on Research and Innovation in Information Systems (ICRIIS)*, 1-6. <https://doi.org/10.1109/ICRIIS53035.2021.9617079>
- Ragheb, W., Aze, J., Bringay, S., & Servajean, M. (2021). Negatively Correlated Noisy Learners for At-risk User Detection on Social Networks: A Study on Depression, Anorexia, Self-harm and Suicide. *IEEE Transactions on Knowledge and Data Engineering*, 35(1), 770-783 <https://doi.org/10.1109/TKDE.2021.3078898>
- Ramirez-Cifuentes, D., Largeron, C., Tissier, J., Baeza-Yates, R., & Freire, A. (2021). Enhanced Word Embedding Variations for the Detection of Substance Abuse and Mental Health Issues on Social Media Writings. *IEEE Access*, 9, 130449-130471. <https://doi.org/10.1109/ACCESS.2021.3112102>
- Rego, B., Rego, N., & Kunder, M. (2021). Social Media Analysis for Mental Health Evaluation. *International Journal for Research in Applied Science and Engineering Technology*, 9(4), 1453-1460. <https://doi.org/10.22214/ijraset.2021.33962>
- Ren, L., Lin, H., Xu, B., Zhang, S., Yang, L., & Sun, S. (2021). Depression detection on reddit with an emotion-based attention network: algorithm development and validation. *JMIR Medical Informatics*, 9(7), Article e28754. <https://doi.org/10.2196/28754>
- Ricard, B., & Hassanpour, S. (2021). Deep learning for identification of alcohol-related content on social media (Reddit and Twitter): Exploratory analysis of alcohol-related outcomes. *Journal of medical internet research*, 23(9), <https://doi.org/10.2196/27314>
- Sabina, A. Chulvi, B., & Rosso, P. (2021). On the explainability of automatic predictions of mental disorders from social media data. *International Conference on Applications of Natural Language to Information Systems. Lecture Notes in Computer Science*, 12801, 301-314. https://doi.org/10.1007/978-3-030-80599-9_27
- Saini, G., Yadav, N., & Kamath S, S. (2022). Ensemble Neural Models for Depressive Tendency Prediction Based on Social Media Activity of Twitter Users. *Lecture Notes in Electrical Engineering*, 848, 211-226. https://doi.org/10.1007/978-981-16-9089-1_18
- Sarkar, D., Kumar, P., Samanta, P., Dutta, S., & Chatterjee, M. (2022). A Two-Level Multi-Modal Analysis for Depression Detection From Online Social Media. *International Journal of Software Innovation*, 10(1), 1-22. <https://doi.org/10.4018/IJSI.309114>
- Schoene, A. M., Bojanic, L., Nghiem, M. Q., Hunt, I. M., & Ananiadou, S. (2022). Classifying suicide-related content and emotions on Twitter using Graph Convolutional Neural Networks. *IEEE Transactions on Affective Computing*, XX(X), 1-12. <https://doi.org/10.1109/TAFFC.2022.3221683>
- Silveira, B., Couto, A., & Murai, F. (2018). Online social networks in health care: a study of mental disorders on Reddit. *IEEE/WIC/ACM International Conference on Web Intelligence (WI)* 568-573. <https://doi.org/10.1109/WI.2018.00-36>
- Silveira, B., Silva, H., Murai, F., & da Silva, A. (2021). Predicting user emotional tone in mental disorder online communities. *Future Generation Computer Systems*, 125, 641-651. <https://doi.org/10.1016/j.future.2021.07.014>
- Soler-Costa, R., Moreno-Guerrero, A. J., Lopez-Belmonte, J., & Marín-Marín, J. (2021). Co-word analysis and academic performance of the term TPACK in web of science. *Sustainability*, 13(3), 2-20. <https://doi.org/10.3390/su13031481>
- Sun, L., & Luo, Y. (2022). Identification and analysis of depression and suicidal tendency of Sina Weibo users based on machine learning. *Advances in Educational Technology and Psychology*, 6(9), 108-117. <https://doi.org/10.23977/aetp.2022.060916>
- Tan, H., Peng, S., Zhu, C., You, Z., Miao, M., & Kuai, S. (2021). Long-term Effects of the COVID-19 Pandemic on Public Sentiments in Mainland China: Sentiment Analysis of Social Media Posts. *Journal of Medical Internet Research*, 23(8), 1-12, Article e29150. <https://doi.org/10.2196/29150>
- Tejaswini, V., Babu, K., & Sahoo, B. (2022). Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model. *ACM Transactions on Asian and Low-Resource Language Information Processing*. <https://doi.org/10.1145/3569580>
- Urban, A. S., Chulvi, B., & Rosso, P. (2021). On the Explainability of Automatic Predictions of Mental Disorders from Social Media Data. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 12801 LNCS*. Springer International Publishing. https://doi.org/10.1007/978-3-030-80599-9_27
- Urban, C., & Gates, K. (2021). Deep learning: A primer for psychologists. *Psychological Methods*, 26(6), 743-773. <https://doi.org/10.1037/met0000374>
- Viviani, M., Crocamo, C., Mazzola, M., Bartoli, F., Carrà, G., & Pasi, G. (2021). Assessing vulnerability to psychological distress during the COVID-19 pandemic through the analysis of microblogging content. *Future*

BIOGRAFÍAS

- Generation Computer Systems*, 125, 446-459. <https://doi.org/10.1016/j.future.2021.06.044>
- Wang, Y., Zhao, Y., Zhang, J., Bian, J., & Zhang, R. (2020). Detecting associations between dietary supplement intake and sentiments within mental disorder tweets. *Health informatics journal*, 26(2), 803-815. <https://doi.org/10.1177/1460458219867231>
- Wongkoblapp, A., Vadillo, M. A., & Curcin, V. (2017). Researching mental health disorders in the era of social media: systematic review. *Journal of medical Internet research*, 19(6), 228. <https://www.jmir.org/2017/6/e228/>
- Wongkoblapp, A., Vadillo, M. A., & Curcin, V. (2021). Deep Learning With Anaphora Resolution for the Detection of Tweepers With Depression: Algorithm Development and Validation Study. *JMIR Mental Health*, 8(8), Article e19824. <https://doi.org/10.2196/19824>
- 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), 102961. <https://doi.org/10.1016/j.ipm.2022.102961>
- Yao, H., Rashidian, S., Dong, X., Duanmu, H., Rosenthal, R. N., & Wang, F. (2020). Detection of suicidality among opioid users on reddit: Machine learning-based approach. *Journal of medical internet research*, 22(11), Article e15293. <https://doi.org/10.2196/15293>
- Zanwar, S., Wiechmann, D., Qiao, Y., & Kerz, E. (2022). *Exploring Hybrid and Ensemble Models for Multiclass Prediction of Mental Health Status on Social Media*. 184-196, <http://arxiv.org/abs/2212.09839>
- Zhang, M., Lu, S., Li, M., Zhai, Q., Zhou, J., Lu, X., Xu, J., Xue, J., & Zhong, N. (2017). SVM classification model in depression recognition based on mutation PSO parameter. *EDP Sciences*, 8(01037), 1-8. <https://doi.org/10.1051/bioconf/20170801037>
- Zhang, T., Schoene, A., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: a narrative review. *NPJ digital medicine*, (46), 1-13. <https://doi.org/10.1038/s41746-022-00589-7>
- Zhang, W., Seltzer, T., & Bichard, S. (2013). Two sides of the coin: Assessing the influence of social network site use during the 2012 US presidential campaign. *Social Science Computer Review*, 31(5), 542-551. <https://doi.org/10.1177/0894439313489962>



Ismael Leonardo Mieles Toloza, Magíster en psicología y programador, ha trabajado en proyectos desde las neurociencias, neuropsicología y las nuevas tecnologías, orientadas a la investigación en cognición social y la alteración emocional en pacientes con daño cerebral, ha participado como consultor en la propuesta y desarrollo de proyectos sociales en comunidades indígenas con entidades nacionales colombianas, Becario Colciencias 2016 y docente investigador de la facultad de psicología de la UDI.



Jesús Armando Delgado Meza, Estudiante de doctorado en Ciencias de la Educación, Magíster en Administración de Empresas, Especialista en Psicología Clínica y Psicólogo. Docente de la Universidad e Investigación y Desarrollo – UDI. Líder del grupo de Investigación UDIPSI de la UDI. Ha participado en

proyectos de investigación Minciencias de Colombia y desarrollado investigaciones en áreas sociales, clínicas y educativas de la Psicología, en más de 20 años como profesional.



Johana Acevedo Suárez. Magíster en psicología UPB-BGA. Docente en Educación Superior para Pregrado de Psicología y líder de proyectos estratégicos en primera infancia en la Universidad de Santander, Docente Cátedra en la Universidad Pontificia Bolivariana y actualmente

Docente Investigadora de la Universidad de Investigación y Desarrollo. Reconocida como docente virtual UDI (2019) y líder de semillero Connecting People UDI (2022). Con experiencia en el área de la Psicología Clínica y de la Salud, dominio de baterías neuropsicológicas y rehabilitación cognitiva.

