

Metodología de Análisis de Emociones para Identificar Riesgo de Cometer Suicidio Generado por el COVID-19¹

Luz Judith Rodríguez-Esparza², Diana Barraza-Barraza³, Jesús Salazar-Ibarra⁴, Rafael G. Vargas-Pasaye⁵

Resumen

Introducción: El comienzo del año 2020 llegó acompañado de una pandemia causada por el virus denominado SARS-CoV-2. Con las medidas de distanciamiento social implementadas para evitar la propagación de este virus, se presentan problemáticas de salud mental, como ansiedad, depresión, etc., trayendo como consecuencia una necesidad de atención a pacientes a distancia. Dadas las cifras alarmantes de incidencias de suicidio en la sociedad actual, aunadas a estas medidas de distanciamiento, son requeridas herramientas de apoyo para identificar individuos en riesgo de cometer suicidio. **Objetivo:** Proponer y evaluar una nueva metodología para calcular riesgo de suicidio en usuarios de Twitter, apoyándose en el análisis de emociones. **Materiales y Métodos:** Usando modelos de aprendizaje estadístico (supervisado y no supervisado),

la metodología propuesta identifica el nivel de riesgo en el texto analizado de 77 tuits de usuarios regulares y de figuras políticas en México y Latinoamérica. **Resultados:** Se encontró que, al comparar los métodos utilizados, el porcentaje de coincidencia en clasificación es cercano al 96 %, siendo los métodos supervisado no paramétrico y no supervisado los que detectaron los niveles extremos de riesgo al suicidio. **Conclusiones:** la metodología propuesta es una herramienta que puede ser de gran apoyo para especialistas del área de salud mental al ayudar a identificar, de manera masiva, la presencia de indicios de enfermedades mentales, para su subsecuente diagnóstico.

Palabras clave: Aprendizaje estadístico; análisis de emociones; modelos supervisados y no supervisados

1 Artículo original, derivado del 1er Encuentro Nacional de Investigación Multidisciplinaria Código 27, ejecutado entre abril de 2019 a abril 2020, financiado por los autores.

2 Doctora en Probabilidad y Estadística. Cátedra CONACYT-Universidad Autónoma de Aguascalientes. ORCID: <https://orcid.org/0000-0003-2241-1102>

3 Doctora en Ciencias de la Ingeniería, Doctora en Administración de Sistemas e Ingeniería. Profesora de Tiempo Completo en la Facultad de Ciencias Exactas. Universidad Juárez del Estado de Durango. ORCID: <https://orcid.org/0000-0002-6041-4657>.

4 Maestro en Estadística Aplicada. Estudiantes de Doctorado en la Universidad Iberoamericana León. ORCID: <http://orcid.org/0000-0001-8864-2662>

5 Maestro en Comunicación. Estudiante de Doctorado en la Universidad Iberoamericana León

Autor para Correspondencia: Diana Barraza-Barraza. E-mail: dbarrazabza@gmail.com

Recibido: 15/04/2020 Aceptado: 06/12/2021

*Los autores declaran que no tienen conflicto de interés

Methodology of Emotion Analysis to Identify Risk of Committing Suicide Generated by COVID-19

Abstract

Introduction: The beginning of 2020 was accompanied by a pandemic caused by the virus called SARS-CoV-2. With social distancing measures implemented to prevent the spread of this virus, mental health problems arose, such as anxiety, depression, etc., resulting in a need for telemedicine. Given the alarming numbers of suicide incidences in today's society, coupled with these distancing measures, support tools are required to identify individuals at risk of committing suicide. **Objective:** To propose and evaluate a new methodology to calculate suicide risk in Twitter users, based

on the analysis of emotions. **Materials and Methods:** Using statistical learning models (supervised and unsupervised), the proposed methodology identifies the level of risk in the analyzed text of 77 tweets from regular users and political figures in Mexico and Latin America. **Results:** It was found that, when comparing the methods used, the percentage of coincidence in classification is close to 96%, being the supervised non-parametric and unsupervised methods those that detected the extreme levels of suicide risk. **Conclusions:** the proposed methodology is a tool that can be of great support for specialists in the mental health area by helping to identify, in a massive way, the presence of signs of mental illness, for its subsequent diagnosis.

Keywords: Statistical learning; emotion analysis; supervised and unsupervised models

Metodologia de Análise de Emoções para Identificar o Risco de Suicídio Gerada por COVID-19

Resumo

Introdução: O início de 2020 foi acompanhado por uma pandemia causada pelo vírus denominado SARS-CoV-2. Com as medidas de distanciamento social implantadas para prevenir a propagação desse vírus, surgem problemas de saúde mental, como ansiedade, depressão, etc., resultando na necessidade de atendimento remoto ao paciente. Dados os números alarmantes de incidentes de suicídio na sociedade atual, juntamente com essas medidas de distanciamento, ferramentas de apoio são necessárias para identificar indivíduos em risco de suicídio. **Objetivo:** propor e avaliar uma nova metodologia para calcular o risco de suicídio em usuários do Twitter, a

partir da análise das emoções. **Materiais e Métodos:** Usando modelos estatísticos de aprendizagem (supervisionados e não supervisionados), a metodologia proposta identifica o nível de risco no texto analisado de 77 tweets de usuários regulares e figuras políticas no México e na América Latina. **Resultados:** Verificou-se que, na comparação dos métodos utilizados, o percentual de coincidência na classificação é próximo a 96%, sendo os métodos não paramétricos supervisionados e não supervisionados aqueles que detectaram os níveis extremos de risco de suicídio. **Conclusões:** A metodologia proposta é uma ferramenta que pode ser de grande apoio aos especialistas da área de saúde mental por ajudar a identificar, de forma massiva, a presença de indícios de doença mental, para seu posterior diagnóstico.

Palavras-chave: Aprendizagem estatística; análise de emoções; modelos supervisionados e não supervisionados

Introducción

En diciembre de 2019 estalló un brote de enfermedades respiratorias generadas por una nueva cepa de la familia coronavirus, designada como SARS-CoV-2. Para el 20 de enero de 2020, la Organización Mundial de la Salud (OMS) declara emergencia de salud pública de interés internacional (OMS, 2020b), sugiriendo medidas para evitar su contagio, tales como distanciamiento social. A partir de este día, diferentes países alrededor del mundo adoptaron el distanciamiento social, cierre de fronteras y cancelación de actividades no esenciales para la vida (clases presenciales, eventos masivos, centros turísticos, etc.).

Al ser declarada como pandemia (OMS, 2020a), la enfermedad COVID-19, generada por el virus SARS-CoV-2, y sus medidas de prevención, diagnóstico y tratamiento han traído diversos problemas de salud mental (Gould, Díaz y Vargas, 2020; Ramírez-Ortín et al., 2020). Fardin (2020) caracteriza estos problemas como comunes a brotes infecciosos, estableciendo que ante una completa cuarentena, se espera que los individuos desarrollen ansiedad de perder a sus seres queridos, ante la posible escasez de alimentos y el exceso de rumores negativos en redes sociales y noticias; miedo a sufrir la enfermedad; o depresión ante la pérdida de amigos y familia.

En China, epicentro de la epidemia, se han sugerido estrategias para el tratamiento del impacto psicológico de la epidemia COVID-19, entre las cuales se considera el uso de redes sociales para enfrentar la propagación de información falsa y mejorar las formas en que se llevarán a cabo las intervenciones psicológicas (Ho, Chee, & Ho, 2020). Ante esta situación de emergencia mundial, el apoyo de especialistas en salud mental enfrenta un nuevo reto: el trabajo de la telemedicina (Hollander & Carr, 2020).

La atención de pacientes de salud mental en México, en estos tiempos de emergencia mundial, tampoco deberá dejarse de lado, especialmente si se considera el aumento de casos de suicidio en los últimos años en el país (INEGI, 2018), donde los adolescentes conforman el grupo más vulnerable por problemas relacionados con la depresión, la ansiedad, daño neurológico, consumo de sustancias adictivas y problemáticas del tipo económico y familiar, condiciones que se agravan en tiempos de epidemias (Fardin, 2020). Benítez Camacho (2021) ahonda en los aspectos multifactoriales del suicidio, sobre todo la manera en que éstos se agravan después de la pandemia del 2020 en México.

La depresión no es sólo presencia de emociones negativas, sino ausencia de emociones positivas. Las reacciones emocionales que más se han estudiado hasta ahora en relación con la salud han sido la ansiedad, la ira y la tristeza (Abascal & Palmero, 1999). Considerando las relaciones entre las emociones positivas y negativas, es posible plantear la utilidad de las emociones positivas para prevenir enfermedades, para reducir su intensidad, duración y también para alcanzar niveles elevados de bienestar subjetivo (Lyubomirsky, King, & Diener, 2005).

Un modelo teórico ampliamente usado sobre las emociones es el planteado por Plutchik (2001), que habla de ocho emociones básicas (alegría, enojo, miedo, tristeza, aceptación, sorpresa, anticipación y aversión) con distintos niveles de intensidad que generarían otras emociones secundarias (por ejemplo, de la emoción básica de la tristeza, se derivan la pena y la aflicción como emociones secundarias). Este modelo ha sido punto de partida para investigaciones posteriores (Rodríguez-Esparza, Barraza-Barraza, Salazar-Ibarra, & Vargas-Pasaye, 2019).

Así pues, considerando la gravedad de esta problemática en el área de Salud

Pública, las altas tasas de suicidio, así como las consecuencias que éstas conllevan en las familias de quienes terminan con su vida de esta forma, este artículo presenta una metodología que estima un índice de riesgo a cometer suicidio a partir de texto redactado, en la red social Twitter, por el individuo bajo análisis. Aclarando el uso de esta red social por ser la que permite un acceso rápido y libre a las publicaciones de la mayoría de los usuarios, permitiendo así la adquisición de datos masivos. El cálculo del índice se lleva a cabo a través de metodologías del aprendizaje estadístico basándose en apreciaciones de expertos en salud mental.

Este artículo se organiza de la siguiente manera. En la subsección Antecedentes se presentan algunos estudios de aprendizaje estadístico aplicado al estudio de riesgo de suicidio. En la Materiales y métodos se presenta la propuesta metodológica para predecir el índice de riesgo al suicidio de los usuarios del Twitter utilizando aprendizaje estadístico. Los resultados de la investigación se muestran en Resultados. Finalmente, el artículo cierra con discusión y conclusiones.

Antecedentes

El aprendizaje estadístico se refiere a un amplio conjunto de herramientas para comprender los datos. Estas herramientas pueden clasificarse como supervisadas o no supervisadas. En general, el aprendizaje estadístico supervisado implica construir un modelo estadístico para estimar o predecir una salida basada en una o más entradas. Con el aprendizaje estadístico no supervisado, hay entradas pero sin salida que “supervisan” lo acertado de la respuesta; sin embargo podemos aprender relaciones y estructura de tales datos (James, Witten, Hastie, & Tibshirani, 2017).

En la década de 1980, Breiman, Friedman, Olshen y Stone (1984) introdujeron la clasificación y árboles de regresión. Hastie

y Tibshirani (1987) acuñaron el término Modelos Generalizados Aditivos para una clase de extensiones no lineales a los modelos lineales generalizados, proporcionando una implementación práctica de software.

Desde entonces, inspirado en el advenimiento del aprendizaje automático, el aprendizaje estadístico se ha convertido en un nuevo subcampo en estadística, enfocado en modelado y predicción supervisada y no supervisada.

Aprendizaje Estadístico y Redes Sociales: herramientas en la detección del suicidio

Existen actualmente dos tendencias para la predicción de intenciones suicidas: la médica y la social. Esta última realiza las predicciones con base en datos sobre el comportamiento de los usuarios extraídos de sus interacciones en redes sociales o de sus dispositivos inteligentes (Merino, 2019). El uso constante de las redes sociales como una herramienta que permite expresar el sentir y pensar de sus usuarios, ha generado que diversos estudios las consideren como fuente de información sobre el estado anímico de quienes las usan.

Greist, Gustafson, Satuss, Rowse, Laughren y Chiles (1973) fueron de los primeros estudios en incorporar técnicas computacionales para la detección y predicción de riesgo de suicidio, encontrando que el nivel de precisión en las predicciones era más alto cuando se usaban las técnicas computacionales. Adicionalmente, los pacientes preferían contestar entrevistas en computadora que hablar con una persona, al sentirse más cómodos describiendo situaciones o pensamientos suicidas a una entidad que no los juzgaría.

Guntuku, Yaden, Kern, Ungar, y Eichstaedt (2017) encuentran que es Twitter la red social más usada en los estudios para detectar depresión y otros problemas mentales, con

análisis que datan desde el 2013. Christensen, Batterham, y Dea (2014) analizaron los esfuerzos en lo que ellos denominan *e-health*, o salud mental en línea, en tres rubros: búsqueda de usuarios con depresión, efectividad de intervenciones a través de las redes sociales y la naciente intervención cuando a los usuarios se les ha identificado con riesgo de depresión y otros problemas de salud mental.

De Choudhury, Gamon, Counts, y Horvitz (2013), han obtenido escritos públicos en Twitter para generar bases de datos de usuarios que han declarado su diagnóstico depresivo, para comparar su comportamiento con usuarios regulares. Benton, Mitchell, y Hovy (2016) estimaron el riesgo de suicidio en los usuarios de Twitter, tomando como base aquellos que ya han sido diagnosticados con depresión, mientras Shen et al. (2017) optaron por analizar tuits a través de Bayes ingenuos para detectar depresión.

Nadeem (2016) utilizó técnicas de clasificación para el análisis de los textos compartidos por los usuarios a través del tiempo. Jamil, Inkpen, Buddhitha y White (2017) desarrollaron una metodología para la detección de depresión en usuarios de Twitter, a nivel twitter. Debido a la desproporcionalidad que había entre los tuits con contenido depresivo y aquellos que no lo presentaban, optaron por un análisis a nivel usuario, considerando el porcentaje de tuits depresivos, mejorando así la clasificación de usuarios deprimidos. Trabajando también con Twitter, Orabi, Buddhitha, Orabi, e Inkpen (2018) desarrollaron modelos de clasificación con altos niveles de exactitud, al determinar si un usuario presenta o no señales de depresión.

Reece y Danforth (2017) se enfocaron en el análisis de fotografías subidas a Instagram para detectar signos de depresión en los usuarios. Por medio de métodos estadísticos, los autores encontraron que es posible identificar usuarios depresivos a través de las fotografías

que comparten, siendo capaces de detectar signos de depresión incluso antes de que los usuarios sean diagnosticados por parte de un especialista de salud mental.

Más recientemente, Ryu et al. (2018), analizaron un modelo de aprendizaje automático para predecir ideación suicida, alcanzando una precisión del 78.3 % con los datos de prueba. Mientras que Oh et al. (2017) experimentaron la posibilidad de predecir intentos suicidas con un algoritmo de clasificación alcanzando un 77.9 % de detección, en muestras grandes.

Tanto Facebook como Twitter han implementado rutinas considerando palabras y/o frases con tintes suicidas. En 2017, Facebook comenzó a utilizar aprendizaje automático para ampliar la capacidad de brindar ayuda oportuna, identificando publicaciones de personas que podrían estar en riesgo, como frases en publicaciones y comentarios preocupados de amigos y familiares (Card, 2018).

La herramienta Tree Hole Rescue (Want, 2019), desarrollada en China para evitar el suicidio, rastrea ciertas palabras en la red social Weibo (equivalente a Twitter) y estima un riesgo de las personas a cometer suicidio. Desde abril de 2018 a noviembre de 2019, el programa ha salvado a 700 personas. No es un software abierto ni con artículos científicos publicados, por lo que su replicación es imposible de llevar a cabo.

En México, si bien la telemedicina ha existido desde principios de la década de 2010's (Secretaría de Salud, 2011), los esfuerzos por atender la salud mental, especialmente después del aislamiento social por la enfermedad COVID-19 aún no han sido publicados. Mercado (2020) presenta sugerencias de trabajo en línea para servicios de psicoterapia, mientras Nicolini (2020) explica las incidencias de depresión y ansiedad debidos a la pandemia; sin describir alguno de ellos herramientas tecnológicas para la

detección de riesgos de suicidio o depresión en redes sociales.

Aun cuando la tecnología no puede reemplazar a los especialistas de salud mental, puede ser una ayuda para conectar a más personas necesitadas con la ayuda apropiada. De momento, los esfuerzos en redes sociales han estado encaminados a detectar signos de depresión, sin adentrarse en el estudio del riesgo de suicidio en los usuarios de dichas redes.

Materiales y Métodos

Se proponen algoritmos de aprendizaje estadístico de regresión y clasificación para calcular un Índice de Riesgo de Suicidio (IRS) a través de minería de emociones. Se busca que, con estos algoritmos, los expertos en salud mental tengan nuevas herramientas para determinar si un paciente, posible paciente o usuario de redes sociales que no ha recibido atención especializada, requiere atención focalizada, dado su nivel de riesgo a cometer acciones suicidas. La metodología consiste en

- **Entrada de emociones:** frases redactadas por el usuario, en este caso, tuits.

- **Minería de emociones:** identificación de emociones expresadas por el usuario en el tuit, a través de NRC Word-Emotion Association Lexicon, que es una lista de palabras y sus asociaciones a las ocho emociones básicas de Plutchik (2001) y a dos sentimientos (positivo o negativo) (Mohammad & Turney, 2013).
- **Cálculo de índice:** las emociones identificadas se alimentan a los métodos de aprendizaje estadístico para calcular el IRS del usuario. Dicho IRS puede tomar uno de cuatro niveles: Leve, Moderado, Severo y Extremo (Gómez, 2012).

Minería de emociones

La frase proporcionada al software generará un perfil de emociones identificando la proporción de éstas en la frase. Por ejemplo, la frase “Su orfandad es una manera terrible de violentar, pero más vale un final terrible que un terror sin final” (Vega-Gil, 2019) genera un perfil de emociones como el que se muestra en la Figura 1, donde se aprecia el porcentaje de cada emoción identificado en la frase. Este perfil de emociones es la entrada de los modelos de aprendizaje estadístico discutidos en la siguiente sección.

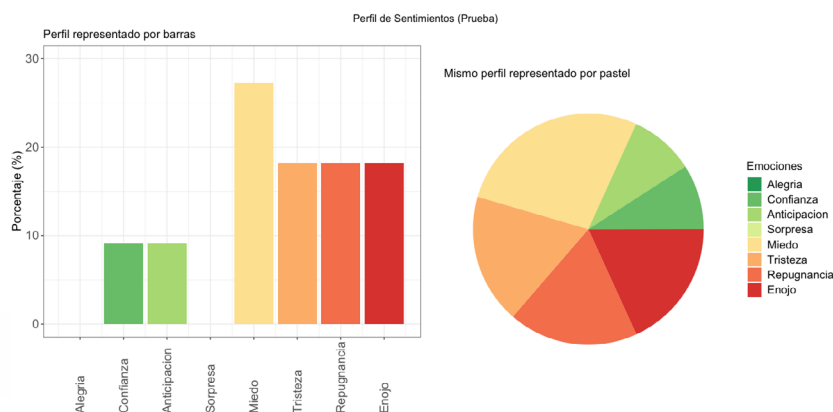


Figura 1. Perfil de emociones generado a través de minería de emociones

Fuente: Elaborado por los autores

A partir de este punto, en toda notación matemática, es Tristeza, es Enojo, es Miedo, es Aversión, es Anticipación, es Sorpresa, es Confianza y es Alegría, las ocho emociones básicas presentadas por Plutchik (2001), y Es el conjunto de las ocho emociones por perfil, expresado en forma de vector.

Métodos del Aprendizaje Estadístico

Consideremos el siguiente modelo general

$$IRS = f(X_1, X_2, \dots, X_8) + \epsilon,$$

donde las X son las emociones de acuerdo con la sección anterior, y IRS es la variable de respuesta. La función f es una función fija pero desconocida, que dependerá de los modelos de aprendizaje utilizados en este estudio, y el término ϵ representa el término aleatorio presente en todo modelo estadístico.

El conjunto de entrenamiento para la modelación y para estimar f está dado por $\{(x_1, IRS_1), (x_2, IRS_2), \dots, (x_n, IRS_n)\}$, donde n es el número de expresiones que componen el conjunto de entrenamiento, y $x_i, i = 1, 2, \dots, n$, es el perfil de emociones como se definió en la sección Minería de emociones.

Dependiendo de la existencia de una variable de respuesta, aprendizaje estadístico considera dos tipos: **Supervisado** (donde para cada $x_i, i = 1, 2, \dots, n$, hay una respuesta asociada IRS_i) y el **No Supervisado** (para cada $x_i, i = 1, 2, \dots, n$ no existe una respuesta asociada IRS_i). Dentro de los métodos supervisados, se encuentran dos métodos: el **paramétrico**, que supone la forma funcional de f y el **no paramétrico**, que no supone una forma funcional de f . Dependiendo el tipo de respuesta, se puede realizar **Regresión** (IRS en términos numéricos) o bien **Clasificación** (IRS como uno de los cuatro niveles definidos). Considerando esta información, las propuestas de ajuste y predicción para el IRS , que se explican en las siguientes subsecciones, son:

1. Supervisado paramétrico: GAM.
2. Supervisado No paramétrico: Splines.
3. No Supervisado: Análisis Discriminante Lineal (ADL).

Modelos de Supervisados Paramétricos y no Paramétricos

Para el modelo paramétrico, se ha seleccionado un modelo *GAM* que permite relaciones no lineales entre la respuesta IRS y el conjunto de emociones, tomando la forma

$$IRS = \beta_0 + g_1(x_1) + g_2(x_2) + \dots + g_8(x_8) + \epsilon.$$

donde $g_j(x_j), j = 1, 2, \dots, 8$, es una función suave no lineal que se estima por separado para cada una de las emociones, y después se suman sus contribuciones para dar como resultado un IRS numérico (0 a 100 en este ejercicio).

Por otro lado, de los modelos no paramétricos, se considera el modelo *spline*, una curva que se ajusta por regiones, es decir, dependiendo de los valores que tome el perfil de emociones. Estos modelos están dados por:

$$IRS = \beta_0 + \beta_1 b_1(x_1) + \beta_2 b_2(x_2) + \dots + \beta_8 b_8(x_8) + \epsilon,$$

donde $b_j(x_j), j = 1, 2, \dots, 8$ son funciones fijas, polinómicas de x_j .

Modelos no supervisados: Análisis Discriminante Lineal

La predicción de una respuesta cualitativa para una observación se puede denominar clasificación de la observación, ya que implica asignar a la observación a una categoría, o clase. Para este estudio, se asume que el perfil de emociones, $X = (X_1, \dots, X_8)$, tiene una distribución Gaussiana multivariada (o normal multivariada), con vector de medias μ_k ($k = 4$ el número de niveles de IRS , es decir 4: Leve,

Moderado, Severo y Extremo) y una matriz de covarianza común en todas las clases Σ . La densidad X de está dada por

$$f_X(x) = \frac{1}{(2\pi)^4 |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_k)' \Sigma^{-1}(x - \mu_k)\right).$$

Usando esta información, el ADL determinará si un perfil de emociones pertenece a cada uno de los cuatro niveles de IRS usando la ecuación:

$$\delta_k(x) = x' \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k' \Sigma^{-1} \mu_k + \log(\pi_k),$$

donde π_k es la proporción de las observaciones de entrenamiento que pertenecen al nivel k de IRS. En otras palabras, $\pi_k = \frac{n_k}{n}$, siendo n_k el número de elementos en el nivel k de ISR. Estos números se determinan usando la información del conjunto de entrenamiento.

Resultados

Generación de Base de Datos de Entrenamiento

En 2019, se llevó a cabo el 1er. Encuentro Nacional de Investigación Interdisciplinaria

Código 27 en la ciudad de Durango, México, que tuvo como objetivo conjuntar un grupo de expertos en salud mental para comentar la problemática del suicidio en México. Con la participación de expertos en la materia se llevó a cabo el siguiente ejercicio:

1. Se entregó a cada mesa de trabajo siete perfiles de emociones diferentes como el mostrado en la Figura 2. Note que ambas gráficas representan la misma información. Se generaron 100 perfiles de manera aleatoria, cuya representación numérica (porcentaje por emoción) se encuentra en Tabla 1.
2. En cada mesa de trabajo, los expertos en salud mental llegaron a un acuerdo sobre el nivel de riesgo de suicidio (0 a 100) en que se encontraba la persona que haya dado como resultado tal perfil de emociones.
3. Se registró este nivel de riesgo otorgado a cada perfil. Ver por ejemplo la Tabla 1.

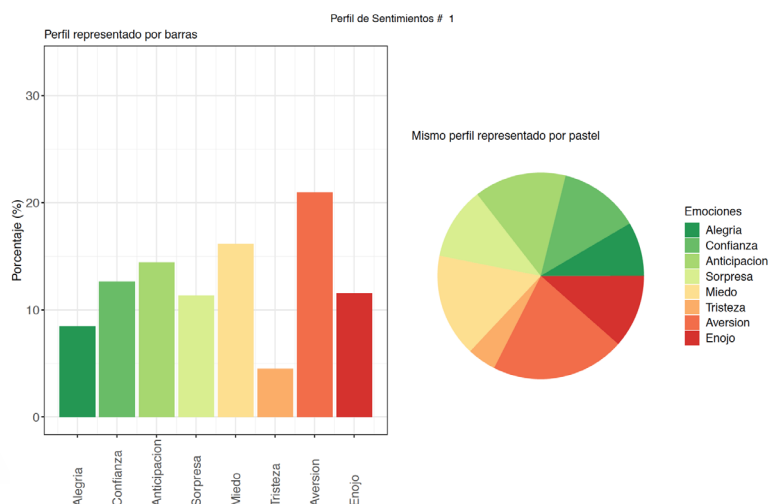


Figura 2. Perfil de emociones presentado a los expertos en salud mental.

Fuente: Elaborado por los autores

Tabla 1. Propuesta de Riesgo de Suicidio para perfil de emociones.

Perfil	Emoción en cada perfil (%)								Riesgo de Suicidio (%)
	Alegría	Confianza	Anticipación	Sorpresa	Miedo	Tristeza	Aversión	Enojo	
1	8.44	12.66	14.42	11.34	16.16	4.50	20.97	11.52	70
2	15.13	11.41	18.51	15.64	8.24	14.65	6.91	9.51	40
3	20.56	1.28	17.36	22.92	17.69	6.66	7.16	6.35	60
4	7.85	9.67	0.98	25.91	26.82	0.06	22.24	6.48	30
5	1.62	27.70	2.67	2.78	27.80	28.15	0.51	8.77	60

Fuente: Elaborado por los autores.

La información generada por los expertos en salud mental se usó como conjunto de datos de entrenamiento para los algoritmos de aprendizaje estadístico presentados en la metodología. Dado que los expertos proporcionan información numérica y no nivel de riesgo (leve, moderado, severo o extremo), se realizó un análisis de conglomerados con distancia promedio. Los resultados se

presentan en la Figura 2, identificando cada nivel con colores: rojo (leve), azul (moderado), verde (severo) y cian (extremo). Las líneas horizontales marcan la clasificación propuesta del IRS en este trabajo. Así pues, los rangos propuestos para los niveles del IRS son: Leve 0-45, Moderado 46-70, Severo 71-90 y Extremo de 91-100.

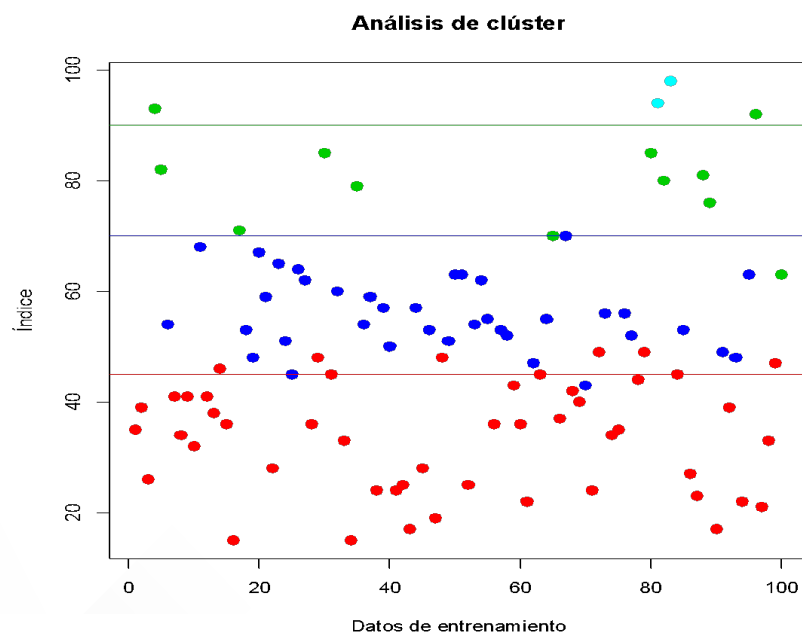


Figura 3. Análisis de clúster para los datos de entrenamiento.

Fuente: Elaborado por los autores.

Entrenamiento de Modelos

Resultados para Modelo Supervisado Paramétrico, usando GAM.

Utilizando el paquete estadístico R, el análisis de varianza correspondiente al modelo GAM se presenta en la Tabla 2, siendo significativas para el IRS las variables Tristeza, Enojo, Miedo, Aversión y Anticipación (marcadas con ***).

Tabla 2. ANOVA para el modelo GAM para estimar el IRS.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Tristeza	1	22736.9	22736.9	986.8619	>2.20E-16	***
Enojo	1	9356.8	9356.8	406.1193	>2.20E-16	***
poly(Miedo,2)	2	1069.3	534.6	23.2051	8.06E-09	***
Aversión	1	643.2	643.2	27.9188	9.08E-07	***
Anticipación	1	305.4	305.4	13.2574	0.0004579	***
Sorpresa	1	43.4	43.4	1.8853	0.1732215	
poly(Confianza,2)	2	108.3	54.1	2.3498	0.1013454	
poly(Alegria,2)	2	32.5	16.3	0.7063	0.4962501	
Residuals	88	2027.5	23			

Fuente: Elaborado por los autores.

Resultados para Modelo Supervisado No Paramétrico, usando Splines.

Para este modelo se utilizaron splines naturales (ns, por sus siglas en inglés) con distintos grados de libertad (James et al., 2017).

El análisis de varianza correspondiente a este modelo se presenta en la Tabla 3, siendo significativas para el IRS las variables Tristeza, Enojo, Miedo, Aversión y Anticipación (marcadas con *** y **).

Tabla 3. ANOVA para el modelo IRS utilizando splines naturales.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
ns(Tristeza, 3)	3	23845.2	7948.4	356.8235	>2.20E-16	***
ns(Enojo, 3)	3	8351.7	2783.9	124.9763	>2.20E-16	***
ns(Miedo, 2)	2	1180.4	590.2	26.4951	1.40E-09	***
ns(Aversion, 2)	2	775.2	387.6	17.4009	5.17E-07	***
ns(Anticipación, 2)	2	254.4	127.2	5.7093	0.004791	**
ns(Sorpresa, 2)	2	38.3	19.2	0.8605	0.426796	

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ns(Confianza, 2)	2	44.3	22.1	0.9943	0.37447
ns(Alegría, 2)	2	29.6	14.8	0.6642	0.51748
Residuals	81	1804.3	22.3		

Fuente: Elaborado por los autores.

Resultados para Modelo No Supervisado, usando Análisis de Discriminante Lineal.

Este modelo arroja las probabilidades de pertenecer a cada uno de los niveles de IRS que se presentan en la Tabla 4. En esta tabla se puede apreciar que, de los datos de entrenamiento, las probabilidades de pertenecer a los niveles de IRS leve o moderado son mayores.

Tabla 4. Probabilidades de pertenencia a cada nivel de IRS

Leve	Moderado	Severo	Extremo
0.45	0.43	0.08	0.04

Fuente: Elaborado por los autores.

El Análisis de Discriminante Lineal (ADL) describe los grupos con un vector de medias (renglones en la Tabla 5). Recordando que se trata de porcentajes de emociones identificadas en cada perfil, es posible ver que en el nivel de riesgo leve está representado por bajos porcentajes de emociones negativas como tristeza, enojo, miedo o aversión, y altos niveles de confianza y alegría. Por otro lado, el nivel de riesgo de suicidio extremo está identificado por altos porcentajes de emociones negativas (tristeza, enojo, miedo y aversión), mientras que las emociones positivas se presentan en muy bajos niveles o nulos (alegría, sorpresa y anticipación).

Tabla 5. Medias de emociones por grupo de IRS

	Tristeza	Enojo	Miedo	Aversión	Anticipación	Sorpresa	Confianza	Alegría
Leve	6.234	2.578	7.009	4.399	16.965	7.704	31.976	23.132
Moderado	14.636	10.277	15.604	8.544	14.978	6.753	18.487	10.719
Severo	30.565	14.018	16.585	17.345	6.637	3.750	6.013	5.085
Extremo	32.6175	20.737	24.947	20.382	0.000	0.000	1.315	0.000

Fuente: Elaborado por los autores.

Se presentaron tres modelos con metodologías diferentes, realizando una comparación del comportamiento de estos modelos con datos de prueba. Se consideraron 77 nuevas frases para predecir cuál es su nivel de IRS. La Figura 4 presenta la comparación

de estimaciones de IRS para los modelos paramétrico (GAM) y no paramétrico (splines); estos modelos asignan un valor numérico al IRS. Estos modelos coinciden en estimaciones de IRS en un 96 % de las frases estudiadas.

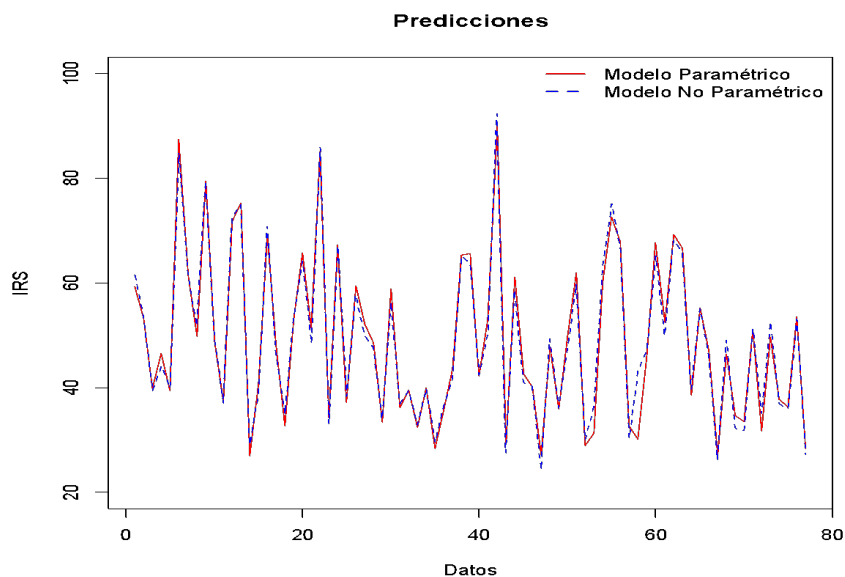


Figura 4. Predicción del IRS para datos de prueba. Comparación de modelo paramétrico y no paramétrico.

Fuente: Elaborado por los autores.

Al realizar la comparación de las predicciones entre los tres métodos propuestos, presentando los porcentajes de coincidencia entre los modelos (modelo supervisado vs modelo

paramétrico o vs modelo no paramétrico) en la Tabla 4, se observa que los tres modelos tienen altos porcentajes de coincidencia en las estimaciones.

Tabla 6. Porcentaje de coincidencias de predicciones entre los métodos analizados.

Modelo 1	Modelo 2	Porcentaje de Coincidencia
No Supervisado	Supervisado Paramétrico	90.90%
No Supervisado	Supervisado No Paramétrico	92.20%
Supervisado Paramétrico	Supervisado No Paramétrico	96.00 %

Fuente: Elaborado por los autores

En la Tabla 7 se presentan explícitamente cuántos usuarios se clasificaron en cada categoría para los tres métodos usados. Se agrega además la predicción obtenida utilizando el Proceso Jerárquico Analítico (PJA) presentado en Rodríguez-Esparza et al. (2019).

La misma información se presenta en la Figura 5. En ambos elementos se puede observar que las metodologías coinciden en la predicción de niveles leve, moderado y severo. Sin embargo, sólo el método no supervisado clasifica frases en niveles extremos de IRS.

Tabla 7. Clasificación del IRS de los datos de prueba dada por los métodos analizados.

	Leve	Moderado	Severo	Extremo
Supervisado Paramétrico	34	36	7	0
Supervisado No Paramétrico	35	34	7	1
No supervisado	33	38	3	3
PJA	38	31	8	0

Fuente: Elaborado por los autores.

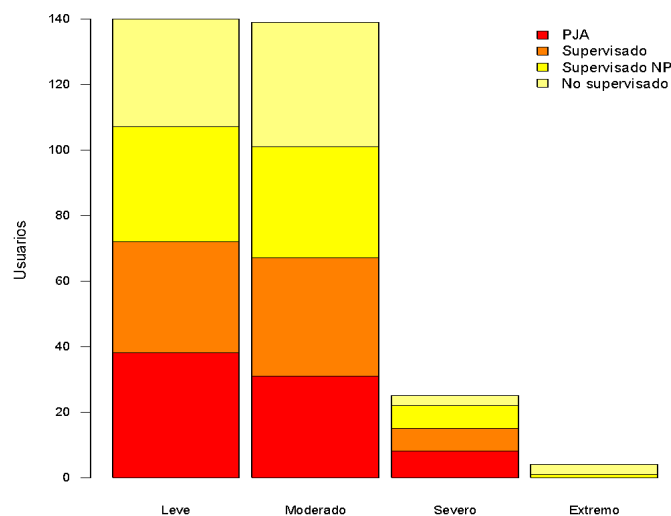


Figura 5. Clasificación de los datos de prueba obtenida por los cuatro modelos considerados.

Fuente: Elaborado por los autores.

Prueba de Modelos con Nuevo Conjunto de Datos: COVID-19

Se decidió usar esta metodología para analizar las emociones de diversos gobernadores y jefes de estado, de México y Latinoamérica respectivamente, expresados en sus tuits entre los días 15 de marzo a 12 de abril del año 2020, analizando 2,377 palabras para gobernadores de la República Mexicana, y 1,371 palabras para los mandatarios de los 18 países latinoamericanos. La Figura 6 muestra los resultados obtenidos aplicando los modelos supervisados (paramétrico y no paramétrico) para los tuits de los gobernadores de México,

se puede observar que la mayoría de éstos muestran un IRS por debajo del 60. El único gobernador con IRS sobre 70 es el gobernador de Hidalgo, quien el 28 de marzo de 2020 dio positivo a COVID-19 (Milenio, 2020), el primer gobernador mexicano en tener tal diagnóstico. En la Figura 7, se comparan los resultados de predicción de ISR con modelos supervisados, para jefes de estado en los países de América Latina, para quienes este índice reporta niveles menores a 60, con el presidente cubano reportando un índice ligeramente superior a 60. Para ninguno de los gobernadores o presidentes se encuentra un ISR que emitan preocupación por alguno de ellos.

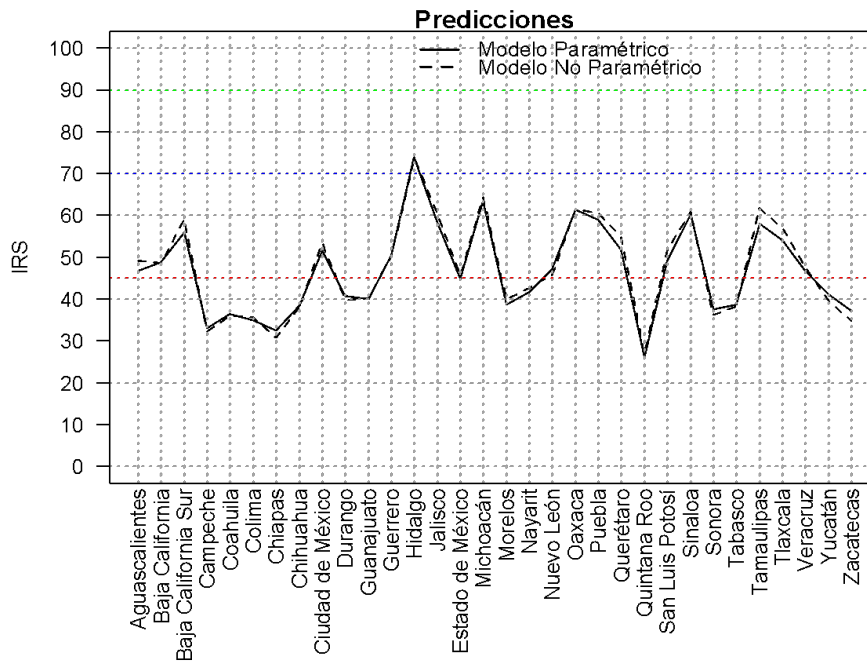


Figura 6. Predicciones utilizando los modelos supervisados para los gobernadores de México.

Fuente: Elaborado por los autores.

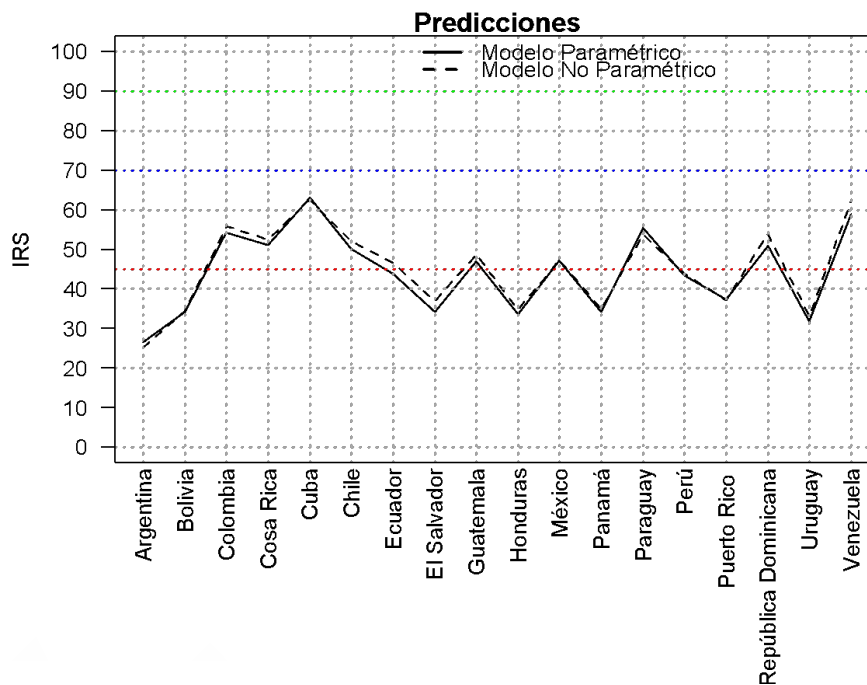


Figura 7. Predicciones utilizando los modelos supervisados para algunos presidentes.

Fuente: Elaborado por los autores.

Al comparar los resultados de los modelos supervisados con el modelo no supervisado, como se muestra en la Tabla 8 y el resumen en la Tabla 9, se aprecia que todos los gobernadores de México se encuentran en niveles que no preocupan, manteniendo una actitud activa positiva tratando de transmitir los riesgos de la pandemia, no así el gobernador de

Hidalgo, en quien sí se aprecian niveles severos (modelos supervisados) y extremo (modelo no supervisado). En un análisis de las emociones identificadas en los tuits de este gobernador, se aprecian emociones orientadas a la tristeza y el miedo. En esta situación, se recomienda prestar atención a la salud mental de este gobernador, pues su ISR es preocupante.

Tabla 8. IRS calculados a partir de tuits de gobernadores en México usando modelos supervisados y no supervisado

Estado	Modelos			Estado	Modelos		
	Paramétrico	No Paramétrico	No Supervisado		Paramétrico	No Paramétrico	No Supervisado
Aguascalientes	Moderado	Moderado	Moderado	Morelos	Leve	Leve	Leve
Baja California	Moderado	Moderado	Moderado	Nayarit	Leve	Leve	Leve
Baja California Sur	Moderado	Moderado	Moderado	Nuevo León	Moderado	Moderado	Moderado
Campeche	Leve	Leve	Leve	Oaxaca	Moderado	Moderado	Moderado
Coahuila	Leve	Leve	Leve	Puebla	Moderado	Moderado	Moderado
Colima	Leve	Leve	Leve	Querétaro	Moderado	Moderado	Moderado
Chiapas	Leve	Leve	Leve	Quintana Roo	Leve	Leve	Leve
Chihuahua	Leve	Leve	Leve	San Luis Potosí	Moderado	Moderado	Moderado
Ciudad de México	Moderado	Moderado	Moderado	Sinaloa	Moderado	Moderado	Moderado
Durango	Leve	Leve	Leve	Sonora	Leve	Leve	Leve
Guanajuato	Leve	Leve	Leve	Tabasco	Leve	Leve	Leve
Guerrero	Moderado	Moderado	Moderado	Tamaulipas	Moderado	Moderado	Moderado
Hidalgo	Severo	Severo	Extremo	Tlaxcala	Moderado	Moderado	Moderado
Jalisco	Moderado	Moderado	Moderado	Veracruz	Moderado	Moderado	Moderado
Estado de México	Leve	Moderado	Moderado	Yucatán	Leve	Leve	Leve
Michoacán	Moderado	Moderado	Moderado	Zacatecas	Leve	Leve	Leve

Fuente: Elaborado por los autores.

Tabla 9. Resumen de ISR para los tuits analizados de los gobernadores de México.

Modelo	ISR			
	Leve	Moderado	Severo	Extremo
Paramétrico	15	16	1	0
No Paramétrico	14	17	1	0
No supervisado	14	17	0	1

Fuente: Elaborado por los autores.

En el análisis realizado a los tuits de presidentes de países de América Latina, presentados en las Tabla 10 y Tabla 11, se aprecia que todos ellos se encuentran en

niveles leve o moderado, es decir, ninguno de ellos expresa emociones en sus tuits que levanten alarma sobre intenciones suicidas.

Tabla 10. IRS calculados a partir de tuits de presidentes de América Latina usando modelos supervisados y no supervisado

País	Modelos			País	Modelos		
	Paramétrico	No Paramétrico	No Supervisado		Paramétrico	No Paramétrico	No Supervisado
Argentina	Leve	Leve	Leve	Honduras	Leve	Leve	Leve
Bolivia	Leve	Leve	Leve	México	Moderado	Moderado	Moderado
Colombia	Moderado	Moderado	Moderado	Panamá	Leve	Leve	Leve
Costa Rica	Moderado	Moderado	Moderado	Paraguay	Moderado	Moderado	Moderado
Cuba	Moderado	Moderado	Moderado	Perú	Leve	Leve	Moderado
Chile	Moderado	Moderado	Moderado	Puerto Rico	Leve	Leve	Leve
Ecuador	Leve	Moderado	Leve	República Dominicana	Moderado	Moderado	Moderado
El Salvador	Leve	Leve	Leve	Uruguay	Leve	Leve	Leve
Guatemala	Moderado	Moderado	Moderado	Venezuela	Moderado	Moderado	Moderado

Fuente: Elaborado por los autores.

Tabla 11. Resumen de ISR para los tuits analizados de los presidentes de América Latina

Modelo	ISR			
	Leve	Moderado	Severo	Extremo
Paramétrico	9	9	0	0
No Paramétrico	8	10	0	0
No supervisado	8	10	0	0

Fuente: Elaborado por los autores.

Discusión

El análisis aquí presentado, da pie a una amplia discusión respecto al uso de las tecnologías y el uso de aprendizaje estadístico para la detección de riesgos suicidas o problemas de salud mental en análisis de texto.

La clasificación obtenida para el conjunto de datos de prueba, presenta niveles de coincidencia arriba del 90 %. El modelo GAM no identificó ninguna frase (usuario) con clasificación Extrema, el modelo con Splines identificó a un usuario, mientras que el ADL identificó a 3 usuarios. Sin embargo, es importante señalar que tanto los GAM como los Splines identificaron a 7 usuarios con clasificación Severa, tres de ellos son los que aparecen como Extremos en el ADL. Así pues, concluimos que los métodos supervisados se relajan más en la clasificación, no así para el método no supervisado.

Haciendo la comparación con las predicciones del PJA con los métodos propuestos observamos que el número de usuarios en cada categoría es similar, sin embargo y realizando un análisis más exhaustivo, la clasificación de usuarios no coincide per se. Además de que con el PJA tampoco se logró identificar a ningún usuario con intenciones extremas suicidas.

Se usó la metodología propuesta en este artículo para el análisis de emociones y

predicción de IRS en los tuits de autoridades de México y América Latina. Ninguno de los modelos encontró que alguno de los personajes analizados mostrara riesgos de suicidio, a excepción del gobernador del estado de Hidalgo, en México, cuyos tuits reflejan tristeza y miedo. Cabe resaltar que este gobernador fue el primero en ser diagnosticado positivo a COVID-19 a finales de marzo de 2020, generando así, una alarma sobre la salud mental de dicho gobernador.

Si bien es entendible que, después de una enfermedad desconocida, las personas sientan tristeza y/o miedo, se sugiere voltear a ver a quienes desarrollan estas emociones para un monitoreo más cercano y evitar que les lleven a estados de salud mental más complejos que puedan desencadenar en episodios de depresión severos y riesgo de suicidio. Se presentaron evidencias de usuarios de Twitter que, si bien no establecen explícitamente su deseo de suicidarse, sus mensajes en esta red social previos a cometer dicho acto, sí dan indicaciones de perfiles de emociones con altos niveles de tristeza, miedo y/o enojo.

Conclusiones

Este artículo presentó una nueva metodología para predecir intenciones suicidas de los usuarios de Twitter, basados en lo que éstos expresan en la plataforma, mediante

métodos de aprendizaje estadístico, usando tres métodos: GAM, Splines y ADL. Se utilizó una base de datos de entrenamiento generada en una reunión nacional de expertos en salud mental. Se probaron los modelos con tuits recuperados de usuarios genéricos. Adicionalmente, se analizaron tuits de gobernadores de México y jefes de estado de países de Latinoamérica emitidos después de la declaratoria de pandemia sobre el COVID-19.

Se encontró que los modelos de aprendizaje estadístico empleados presentan altos niveles de coincidencia en la clasificación de los usuarios de Twitter de acuerdo a su índice de riesgo al suicidio, detectando alarmas en usuarios que estudios previos no han detectado. Es importante establecer que será mejor una clasificación de usuarios en riesgos mayores

de los que realmente se encuentran, es decir, es preferible un falso positivo que indique a especialistas de salud mental que existen usuarios que requieren atención a un falso negativo con pérdidas irreparables. Recalcando que las herramientas aquí diseñadas no tienen como objetivo sustituir a los especialistas en salud mental, sino que buscan servir como ayuda a éstos para identificar a individuos con alto riesgo de atentar contra sus vidas. Como trabajo a futuro se deja el monitoreo en el tiempo del Índice de Riesgo al Suicidio para observar las tendencias en el mismo.

Agradecimientos

Se agradece el apoyo por parte del CONACYT a través de una Cátedra en la Universidad Autónoma de Aguascalientes.

Referencias

- Abascal, E. G. F., & Palmero, F. (1999). *Emociones y salud*. Ariel.
- Benítez Camacho, E. (2021). Suicidio: el impacto del Covid-19 en la salud mental. *Revista de Medicina y Ética*, 32(1), 15-39.
- Benton, A., Mitchell, M., & Hovy, D. (2016). Multi-task learning for mental health using social media text. *ArXiv Preprint ArXiv:1712.03538*.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Card, C. (2018). How Facebook AI Helps Suicide Prevention. *Facebook Newsroom*.
- Christensen, H., Batterham, P. J., & Dea, B. O. (2014). E-Health Interventions for Suicide Prevention. *International Journal of Environmental Research and Public Health*, 11(8), 8193-8212. <https://doi.org/10.3390/ijerph110808193>
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting Depression via Social Media. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media Predicting* (Vol. 2, pp. 128-137).
- Fardin, M. A. (2020). COVID-19 and Anxiety: A Review of Psychological Impacts of Infectious Disease Outbreaks. *Archives of Clinical Infectious Diseases*, (In Press).
- Gómez, G. A. (2012). Evaluación del riesgo de suicidio: enfoque actualizado. *Revista Médica Clínica Las Condes*, 23(5), 607-615.
- Gould, M. S. F., Diaz, G. C., & Vargas, M. A. R. (2021). Impacto sobre la salud mental durante la pandemia COVID 19 en

- Paraguay. *Revista Virtual de la Sociedad Paraguaya de Medicina Interna*, 61-68.
- Greist, J. H., Gustafson, D. H., Satuss, F., Rowse, G. L., Laughren, T. P., & Chiles, J. A. (1973). A computer interview for suicide-risk prediction. *American Journal of Psychiatry*, 130(12), 1327-1332.
- Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18, 43-49.
- Hastie, T., & Tibshirani, R. (1987). Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398), 371-386.
- Ho, C. S., Chee, C. Y., & Ho, R. C. (2020). Mental Health Strategies to Combat the Psychological Impact of COVID-19 Beyond Paranoia and Panic. *Annals of the Academy of Medicine, Singapore*, 49(1), 1-3.
- Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? Telemedicine for covid-19. *New England Journal of Medicine*.
- INEGI. (2018). Estadísticas a propósito del día mundial para la prevención del suicidio. Recuperado de https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2018/suicidios2018_Nal.pdf
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning: with Applications in R*, 7th Editio. New York: Springer.
- Jamil, Z., Inkpen, D., Buddhitha, P., & White, K. (2017). Monitoring Tweets for Depression to Detect At-risk Users. *Université d'Ottawa/University of Ottawa*.
- Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, 131(6), 803.
- Mercado, R. G. (2020). Propuesta de atención para los servicios de psicoterapia en línea (tele psicoterapia) derivados del Covid-19 en México. *Psicología y Salud*, 30(1), 133-136.
- Merino, M. (2019). Los dos tipos de predicción del suicidio con Inteligencia Artificial y cómo se consiguen. Retrieved September 20, 2019, from <https://www.xataka.com/robotica-e-ia/dos-tipos-prediccion-suicidio-inteligencia-artificial-como-se-consiguen>
- Milenio. (2020). Gobernador de Hidalgo da positivo a coronavirus. *Milenio*. Retrieved from <https://www.milenio.com/estados/coronavirus-hidalgo-omar-fayad-positivo-covid-19>
- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436-465.
- Nadeem, M. (2016). Identifying depression on Twitter. *ArXiv Preprint ArXiv:1607.07384*.
- Oh, J., Yun, K., Hwang, J.-H., & Chae, J.-H. (2017). Classification of suicide attempts through a machine learning algorithm based on multiple systemic psychiatric scales. *Frontiers in Psychiatry*, 8, 192.
- OMS. (2020a). WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020. Retrieved from <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--11-march-2020>

- OMS. (2020b). *WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV)*. Geneva: World Health Organization. Retrieved from [https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihremergency-committee-on-novel-coronavirus-\(2019-ncov\)](https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihremergency-committee-on-novel-coronavirus-(2019-ncov))
- Orabi, A. H., Buddhitha, P., Orabi, M. H., & Inkpen, D. (2018). Deep Learning for Depression Detection of Twitter Users. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic* (pp. 88-97).
- Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4), 344-350.
- Ramírez-Ortiz, J., Castro-Quintero, D., Lerma-Córdoba, C., Yela-Ceballos, F., & Escobar-Córdoba, F. (2020). Consecuencias de la pandemia COVID-19 en la Salud Mental asociadas al aislamiento social. *SciELO Preprints*, 1, 1-21.
- Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6(1), 15.
- Rodríguez-Esparza, L. J., Barraza-Barraza, D., Salazar-Ibarra, J., & Vargas-Pasaye, R. G. (2019). SALAZAR-IBARRA, Jesús y VARGAS-PASAYE, Rafael Gerardo. Index of Suicide Risk in Mexico using Twitter. *Journal of Social Researches*, 5-15.
- Ryu, S., Lee, H., Lee, D.-K., & Park, K. (2018). Use of a machine learning algorithm to predict individuals with suicide ideation in the general population. *Psychiatry Investigation*, 15(11), 1030.
- Secretaría de Salud (2011). 4 Experiencias de Telemedicina en México. *Sistema Nacional de Telesalud*.
- Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., ... Zhu, W. (2017). Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence* (pp. 3838-3844).
- Vega-Gil, A. (2019). Nota Suicida. Retrieved April 5, 2019, from <https://twitter.com/ArmandoVegaGil/status/1112666222951391233>
- Want, Y. (2019). The Chinese suicides prevented by AI from afar. *BBC News*. Retrieved from <https://www.bbc.com/news/technology-50314819>