

Análisis del lenguaje en grupos de apoyo en Internet de salud mental

Analysing language in Internet support groups for mental health

Gabriela Ferraro¹  <https://orcid.org/0000-0003-3652-9689>

Commonwealth Scientific and Industrial Research Organisation, Australia
Research School of Computer Science, Australian National University
 gabriela.ferraro@csiro.com.au

Luis Salvador-Carulla  <https://orcid.org/0000-0002-5742-9866>

Research School of Population Health, Centre for Mental Health Research,
Australian National University, Canberra, Australia
 luis.salvador-carulla@anu.edu.au

Resumen

Dar asistencia a los moderadores de Grupos de Ayuda en Internet es importante para asegurar su uso de forma segura. Métodos de clasificación de textos que analizan el lenguaje utilizado en estos foros es una de las posibles soluciones. Esta investigación trata de utilizar tecnologías del procesamiento del lenguaje natural y el aprendizaje automático para construir un sistema de clasificación de triaje usando datos del forum de salud mental Reachout.com. Al comparar con el estado de la cuestión, nuestra propuesta alcanza el mejor rendimiento para la clase *crisis* (52%), siendo ésta la de mayor importancia.

Palabras clave: procesamiento del lenguaje natural, aprendizaje automático, clasificación de textos, salud mental.

¹ Recibido: 04.02.2021 | Aceptado: 22.02.2021

Abstract

Assisting moderators to triage critical posts in Internet Support Groups is relevant to ensure its safe use. Automated text classification methods analysing the language expressed in posts of online forums is a promising solution. Natural Language Processing and Machine Learning technologies were used to build a triage post classifier using a dataset from Reachout.com mental health forum. When comparing with the state-of-the-art, our solution achieved the best classification performance for the *crisis* posts (52%), which is the most severe class.

Key words: natural language processing, machine learning, text classification, mental health.

Introducción

Los Grupos de Apoyo en Internet (GAI) son tecnologías importantes y populares para los individuos con problemas de salud mental. Los usuarios de estos grupos buscan recibir apoyo de otros pares que tienen experiencias de vida similares (Islam et al., 2018) y compartir sus historias anónimamente para ayudar a su recuperación (Mikal et al. 2017). En la actualidad, existe un gran rango de GAI. Por ejemplo, aquellos para grupos de personas con enfermedades crónicas como diabetes, cáncer y depresión, entre otras (Islam et al., 2018, Naslund et al. 2018). Algunos estudios sugieren que estos grupos pueden tener un efecto positivo en individuos con problemas de salud mental, aunque también podrían elevar la carga de estrés de los pacientes (Kaplan et al. 2011). Esto sugiere que el uso seguro de GAI como herramienta de apoyo a la salud requiere mayor atención, especialmente el diseño de mecanismos que puedan asistir a mitigar posibles reacciones adversas (Griffiths 2017).

La evaluación y monitorización de GAI es un proceso difícil y costoso debido a que recae en el control manual de las publicaciones (*posts* o *posteos*) en un forum en línea, que es llevado a cabo por moderadores entrenados para tal fin. En consecuencia, la escalabilidad de GAI como potencial sistema de intervención sanitaria digital podría no ser factible. Una alternativa para aliviar esta limitación es el uso de las tecnologías del procesamiento del

lenguaje natural y el aprendizaje automático para construir sistemas de asistencia a los moderadores (Huh et al. 2013). Estas tecnologías pueden utilizarse para detectar y responder a publicaciones potencialmente nocivas, es decir, aquellas que puedan causar estrés a otros usuarios o que describan actos de autolesión.

Los moderadores desempeñan un rol importante en manejar la comunicación entre usuarios en GAI. Estos ofrecen apoyo y consejos a los usuarios, lo que incluye compartir sus propias historias de recuperación, motivar a los usuarios a participar en las discusiones, y promover la adopción de servicios digitales y presenciales de salud mental (Kornfield et al. 2018). Sin embargo, los moderadores pueden no tener las habilidades y competencias necesarias para detectar indicios de casos críticos y guiar adecuadamente en la toma de decisiones relacionadas con seguridad clínica (Hartzler y Pratt 2011).

El triaje de publicaciones en GAI para asistir a los moderadores en la revisión de contenido es una tarea de clasificación de textos diseñada para detectar efectivamente la manifestación de pensamientos, sentimientos, emociones, y posibles comportamientos de los usuarios en un forum (Conway y O'Connor 2016, Tausczik y Pennebaker 2010). Estudios anteriores se han enfocado en evaluar el desempeño de diversos modelos de aprendizaje automático utilizando datos de GAI sobre salud mental, como modelos de regresión (Cohan et al. 2016, Pink et al. 2016, Zirikly et al. 2016), descenso de gradiente estocástico (SGD) (Kim et al. 2016), y análisis discriminante lineal (LDA) (Shickel et al. 2016). De manera similar, la investigación de Islam et al. (2018) utiliza diferentes técnicas de aprendizaje automático para detectar casos de depresión en datos de Facebook. En estos estudios se evaluaron diferentes métodos de clasificación, incluyendo máquina de vectores de soporte (SVM), árboles de decisión, métodos de ensamble y vecinos próximos (KNN). Sin embargo, los métodos actuales no han evaluado el impacto del uso de recursos léxicos y la aplicación de aprendizaje profundo.

Los recursos léxicos pueden considerarse centrales en el modelado de las características lingüísticas de los GAI (Hollingshead et al. 2017, Islam et al. 2018) y en otras publicaciones en línea como Twitter (O’Dea et al. 2015, Jamil et al. 2017, O’Dea et al. 2017). Mientras que el modelado de las características lingüísticas es importante para obtener buenos resultados, otros rasgos como la interacción entre usuarios, la estructura del forum, meta-datos y otros rasgos externos contribuyen también a mejorar la calidad de las predicciones (Carron-Arthur et al. 2015, Mohammad y Turney 2013). Sin embargo, algunos autores han resaltado que utilizar rasgos externos, como por ejemplo la estructura del forum y los meta-datos, podría sesgar la clasificación y, por lo tanto, disminuir la capacidad predictiva de los clasificadores al analizar nuevas publicaciones cuyos rasgos externos sean mínimos o nulos (Carron-Arthur et al. 2015).

Nuestro estudio se centra en el desarrollo de un clasificador automático para triaje de publicaciones utilizando únicamente rasgos textuales derivados de las publicaciones y rasgos extraídos de recursos léxicos. Nuestro objetivo es estudiar el lenguaje en GAI excluyendo la estructura de forum y los hilos de las conversaciones. Al excluir estos rasgos, este estudio se centra en optimizar los aspectos lingüísticos para la clasificación de publicaciones y evitar posibles sesgos causados por usuarios frecuentes y/o más antiguos del forum. Además, dado el alcance de las investigaciones previas en la combinación de diferentes métodos de aprendizaje automático, nuestro estudio se centra en experimentar con diversos conjuntos de rasgos y un número relativamente pequeño de clasificadores lineales y no lineales, incluyendo algunos métodos de aprendizaje profundo.

El presente estudio utiliza métodos del estado de la cuestión para extraer rasgos derivados del forum en línea Reachout (Hollingshead et al. 2017). Nuestro objetivo es obtener la mejor clasificación para las publicaciones etiquetadas como *crisis* y resultados competitivos para las otras etiquetas, descritas en la Sección 1. Además, hemos realizado un análisis cualitativo

de las publicaciones que requieren la atención inmediata de los moderadores. Los objetivos de esta investigación son los siguientes:

- ✓ Proporcionar pistas de las características lingüísticas de las publicaciones críticas.
- ✓ Estudiar la eficacia de recursos léxicos en los sistemas clasificación basados en aprendizaje automático para triaje en los datos del forum Reachout.

1. Materiales y Métodos

1.1 Los datos

Este estudio utiliza una colección de datos del forum Australiano de salud mental Reachout. Estos datos fueron publicados por la Tarea Conjunta Lingüística Computacional y Psicología Clínica (CLPsych: Computational Linguistics y Clinical Psychology Shared Task) (Hollingshead et al., 2017). Los participantes del forum Reachout tienen entre 18 y 25 años y todas sus publicaciones están escritas en inglés. Cada publicación en la colección de datos está etiquetada con un patrón de semáforo para indicar la severidad de la publicación, y consecuentemente la necesidad de atención de un moderador (Tabla 1). La Tabla 2 muestra la distribución entre los datos de entrenamiento y evaluación.

Tabla 1 Etiquetas, descripciones y ejemplos de las clases de severidad en los datos de Reachout

Etiqueta	Descripción	Ejemplo
verde	Intervención del moderador no necesaria. Se puede dejar para que responda la comunidad más amplia de compañeros.	<i>"Estoy orgulloso de haber podido llamar y mantener una conversación telefónica con mi madre".</i>
amarillo	Un moderador debe abordar la publicación en algún momento, pero no es necesario que lo haga de inmediato.	<i>"Hay tantas cosas en las que estoy pensando, pero mis medicamentos ralentizan mis pensamientos y los hacen más manejables".</i>
rojo	Un moderador debe responder a la publicación lo antes posible.	<i>"Me siento impotente y las cosas parecen inútiles. Odio sentirme tan mal."</i>

crisis	El autor, o alguien que conozca, corre un riesgo inminente de sufrir daños o de hacerse daño a sí mismo o a otros. Estas publicaciones deben tener prioridad sobre todos las demás.	<i>“Estoy teniendo fuertes pensamientos sobre acabar con mi vida, nada ayuda.”</i>
---------------	---	--

Tabla 2 Distribución de etiquetas en los sets de entrenamiento y evaluación en los datos de Reachout

	Entrenamiento	%	Evaluación	%
crisis	40	3.36	42	10.5
rojo	137	11.53	48	12
amarillo	296	24.91	94	23.5
verde	715	60.18	216	54
Total	1188	-	400	-

Precisión, cobertura y valor-F (denominado también F-score o medida-F) son las métricas de evaluación usadas para medir el desempeño de los clasificadores. Precisión se define como la proporción de elementos asignados a una clase en particular que son correctos; cobertura es la proporción de elementos a los que se asigna correctamente una categoría y valor-F es la media armónica entre la precisión y la cobertura. La métrica macro valor-F es la preferida, dado que otorga más peso a las clases con menos instancias y, en este caso, más importantes, como en los casos de *rojo* y *crisis*. Como en Altszyler (2018), calculamos también el valor-F para las clases *crisis* versus *non-crisis*, ya que esta métrica mide la capacidad del clasificador para detectar las clases más severas. Las métricas de evaluación oficiales de Hollingshead et al. (2017) se describen a continuación:

Macro Valor-F: esta métrica se calcula entre las clases *crisis*, *rojo* y *amarillo*, después de excluir la clase *verde*.

Valor-F entre *flagged* vs. *non-flagged*: esta métrica separa las publicaciones que requieren la atención de un moderador (p. ej., *crisis*, *rojo*, *amarillo*) en comparación con las publicaciones que pueden ser

ignoradas (p. ej., *verde*). Esta es la métrica más importante en CLPsych ya que mide la capacidad del clasificador para identificar las publicaciones que requieren la atención de un moderador.

Valor-F entre *urgente* vs. *no-urgente*: esta métrica es el promedio Valor-F entre las clases *urgente* (*crisis + rojo*) y *no-urgente* (*amarillo + verde*).

La Tabla 3 muestra los métodos y rasgos utilizados en estudios previos utilizando los datos de Reachout. Para más detalles, consultar Milne (2016) y Hollingshead (2017).

Tabla 3 Ejemplos de rasgos usados para triaje con los datos de Reachout. *Usado por* se refiere a estudios previos donde los rasgos fueron utilizados

Rasgos de lexicones	Utilizado por
Lexicón LIWC (Pennebaker. et al., 2015)	Cohan et al. (2016), Malmasi et al. (2016)
Lexicón MPQA (Kroenke et al., 2001)	Cohan et al. (2016), Altszyler et al. (2018)
Lexicón PERMA (Schwartz et al., 2016)	Altszyler et al. (2018)
Lexicón Emolex (Mohammad y Turney, 2013)	Altszyler et al. (2018)
Lexicón DepecheMood (Staiano y Guerini, 2014)	Cohan et al. (2016), Altszyler et al. (2018)
Otros rasgos	
Diversidad léxica	Altszyler et al. (2018)
Topic modeling	Cohan et al. (2016)
TF-IDF ponderado	Kim et al., 2016, Brew, (2016)
Incrustaciones de caracteres	Malmasi et al. (2016)
Incrustaciones de palabras	Altszyler et al. (2018), Brew, (2016), Kim et al. (2016), Malmasi et al. (2016)

Incrustaciones de oraciones	Le y Mikolov (2014)
Etiquetas morfo-sintácticas (POS-Tags)	Malmasi et al. (2016)
Pronombres	Altszyler et al. (2018)
Análisis de sentimientos	Shickel et al. (2016), Zirikly et al. (2016)
Autor de una publicación	Altszyler et al. (2018)
Historia de una publicación	Pink et al. (2016)
Cadena de respuesta de una publicación	Altszyler et al. (2018)
Hora de la publicación	Altszyler et al. (2018)
Tiempo entre publicaciones	Altszyler et al. (2018)
Día de la semana de la publicación	Altszyler et al. (2018)
Referencia a consejeros	Altszyler et al. (2018)
Referencias de auto-flagelación	Altszyler et al. (2018)
Referencia a teléfonos de ayuda	Altszyler et al. (2018)
Algoritmos	
LDA: topic modeling no supervisado	Shickel et al. (2016)
SGD: clasificación supervisada	Kim et al. (2016)
Máquinas de vectores de soporte (SVM): clasificación supervisada	Altszyler et al. (2018), Brew, (2016), Malmasi et al. (2016), Zirikly et al. (2016)
Regresión lógica: clasificación supervisada	Cohan et al. (2016), Pink et al. (2016), Zirikly et al. (2016)

1.2 Predicción de alertas

Los datos de Reachout $D = \{x_i, y_i\}_{i=1}^n$ consisten en n instancias de entrenamiento, donde la instancia i_{th} es un vector de rasgos x_i y una

etiqueta y_i . La tarea de clasificación es predecir la etiqueta y_i dado el vector de rasgos x_i para cada publicación en el forum tal que:

$$\hat{y}_i = \arg \max_{y_i} P_{\theta}(y_i|x_i)$$

En esta investigación entrenamos un modelo SVM multiclase con un kernel lineal (Vapnik, 1963). SVM es un método de aprendizaje automático supervisado ampliamente usado en clasificación de textos. Los hiperparámetros¹ fueron seleccionados con el algoritmo de búsqueda de cuadrículas² utilizando la técnica de validación cruzada de 5 veces con los datos de entrenamiento. El hiperparámetro C^3 es 1 con $L1$ como tipo de regularización y la función de pérdida⁴ *hinge*, con un máximo de 2000 iteraciones.

Para comparar SVM con otros clasificadores basados en aprendizaje automático, también entrenamos modelos utilizando el método de vecinos próximos y el de naive Bayes. Dado que los métodos de aprendizaje automático profundo son el estado del arte en muchas tareas del procesamiento del lenguaje natural, también entrenamos dos clasificadores de este tipo de redes neuronales: perceptrón multi-capas (MLP) y redes neuronales recurrentes (RNN) con memoria de corto a largo plazo (LSTM).

El set de rasgos se detalla en la Tabla 4. Todos los rasgos fueron derivados de las publicaciones de los datos de Reachout. Tal como ya se indicó, no

-
- ¹ En aprendizaje automático, un hiperparámetro es un parámetro cuyo valor se define antes del proceso de aprendizaje, mientras que el valor de otros parámetros se deriva del proceso de aprendizaje.
 - ² Búsqueda de cuadrículas es una técnica para encontrar los mejores hiperparámetros que consiste en buscar exhaustivamente entre los valores de cada posición en una cuadrícula.
 - ³ El hiperparámetro C refiere al valor de regularización que sirve como grado de importancia que se determina para las instancias mal clasificadas. Mientras mayor es el valor, menos instancias mal clasificadas son toleradas.
 - ⁴ Una función de pérdida o coste mide el error entre el valor estimado y el valor real, con el fin de optimizar los parámetros de un modelo predictivo.

utilizamos rasgos derivados de las estructuras del forum o de las interacciones entre los usuarios. Como se detalla en la Tabla 4, hemos incluido también rasgos adicionales derivados de recursos léxicos como MPQA, lenguaje ofensivo, y lexicones de salud mental. Durante la extracción de rasgos, las negaciones fueron modeladas como en Cimino (2014). Es decir, cada vez que un término de una publicación es encontrado en un lexicón, su negación es analizada inspeccionando el *término_{i-1}*. Como en Cimino (2014), utilizamos la siguiente lista de negaciones:

- ✓ *no, nadie, nada, ninguno, nunca, ni, en ninguna parte, difícilmente, escasamente, apenas, no es, no estaba, no, no puedo, no quiero, no debería, no podría, no ha, no he (no, nobody, nothing, none, never, neither, nor, nowhere, hardly, scarcely, barely, don't, isn't, wasn't, doesn't, ain't, can't, won't, wouldn't, shouldn't, couldn't, hasn't, haven't, didn't)*

Si una negación es encontrada, la polaridad del término se invierte en caso de que el lexicón difiera entre términos positivos y negativos (p.ej., lexicón PERMA). En caso contrario, se ignora y no se incluye como rasgo.

Tabla 4 Set de rasgos de clasificación de triaje con los datos de Reachout, '*' indica lexicones utilizados en trabajos anteriores (ver Tabla 3)

Rasgos léxicos	Descripción
MPQA lexicón*	Número de palabras con polaridad MPQA en cada publicación.
DepecheMood lexicón*	La cantidad de palabras que se superponen entre cada categoría en DepechMood y una publicación.
Emolex lexicón*	Número de palabras que se superponen entre cada categoría en el léxico NRC-Emotion-Lexicon-v0.92 y una publicación
Mental Disorder lexicón 5	Número de palabras que se superponen entre el lexicón y una publicación.
PHQ 9 lexicón	Número de palabras que se superponen entre el PHQ 9 y una publicación.

PERMA lexicón (1)*	Número de bi-gramas y tri-gramas que se superponen entre PERMA y una publicación.
PERMA lexicón (2)	Número de bi-gramas y tri-gramas que se superponen entre las categorías de negativos PERMA y una publicación.
PERMA lexicón (3)	La suma de pesos de los bi-gramas y tri-gramas se superponen entre PERMA y una publicación.
Lexicón Palabras ofensivas	El número de palabras que se superponen entre la lista de palabras ofensivas del lexicón.
Otros rasgos	
TF-IDF ponderad	Representación de n-gramos TF-IDF de cada publicación con las características máximas superiores elegidas por Scikit-learn en función de la frecuencia del término
Pronombres	El número de pronombres utilizados en cada publicación, incluidos yo, tú, él, ella, eso, nosotros, ellos, ellos (<i>I, me, you, he, him, she, her, it, we, us, they, them</i>)
Mean word length	La longitud promedio de las palabras en una publicación.
Incrustaciones de oraciones	Representación de oraciones calculada promediando las incrustaciones de palabras FastText previamente entrenadas y ajustadas con el conjunto de datos Reachout
Incrustaciones de oraciones de la última oración de una publicación	Representación de oraciones de la última oración en cada publicación calculada promediando de incrustaciones de palabras FastText entrenadas con el conjunto de datos Reachout
Análisis de sentimiento	El sentimiento de cada publicación clasificada por un clasificador de sentimiento entrenado por nosotros con representaciones de palabras incrustadas con GloVe (Pennington et al., 2014) y incrustaciones de emoticonos.
Ranking del usuario	El título de la publicación.
Número de enlaces web	Número total de enlaces web en una publicación.
Número de referencias a servicios de ayuda	<i>mental health, australia, general practitioner, doctor, psychologist, counsellor, gp (general practitioner), emergency, 000, lifeline, 131114, 13 11 14, kids help line, 1800 55 1800,</i>

	1800551800, <i>salvation army care line</i> , 1300 36 36 22, 1300363622, <i>e-couch</i> , <i>moodgym</i> , <i>bluepages</i> , <i>black dog institute</i> , <i>reachout</i> , <i>beyondblue</i> , www.moodgym.anu.edu.au , www.ecouch.anu.edu.au , www.bluepages.anu.edu.au , www.researchout.org.au , www.blackdoginstitute.org.au
Número de referencias de expresiones de auto-flagelación	suicidarme, suicidarme, suicidarme, cortarme, cortarme, lastimarme, lastimarme, lastimarme, lastimarme, quiero morir, no quiero vivir, terminar con mi vida, matar, lastimar, cortar, quiero morir, no quiero vivir (<i>suicide, kill myself, kill my self, cut myself, cut my self, hurt myself, hurt my self, harm myself, harm my self, I want to die, don't want to live, end my life, kill, hurt, cut, want to die, I don't want to live</i>)
Número de referencias a supervisores y/o consejeros	<i>supervisor, supervisores, mentor, gerente, tutor, administrador de casos, gerentes, gerente, psicólogo, psiquiatra, gp (médico general), gps, consejero, consejero (supervisor, supervisors, mentor, manager, tutor, case-manager, managers, manager, psych, psychiatrist, gp (general practitioner), gps, counsellor, counselor)</i>

2. Resultados

Esta sección incluye los resultados de los experimentos de triaje automático y el análisis cualitativo de la clase *crisis*

2.1 Resultados experimentales de la clasificación de triaje

Los resultados de los experimentos de clasificación de triaje utilizando diferentes rasgos, incluyendo recursos léxicos y tratando la negación, se presentan en la Tabla 5. Los mejores resultados están señalados en negrita. El uso exclusivo de recursos léxicos resultó tener un desempeño bajo para todas las clases. La inclusión de los rasgos TF-IDF¹ contribuyó a mejorar la clasificación en todas las clases. Los mejores resultados se consiguieron con una combinación de rasgos TF-IDF + lexicones con negación, con un valor-F de 0.44. El set de rasgos más complejos, incluyendo todos los rasgos en la Tabla 4, muestran resultados competitivos con la mayoría de los sistemas

¹ TF-IDF (Spärck Jones 1972) es la cantidad de veces que una palabra aparece en un documento ponderado por el número de palabras frecuentes en una colección de documentos.

del estado del arte relevados por Altszyler (2018), y con los clasificadores de base presentados por Milne (2016).

Tabla 5 Clasificación con diferentes sets de rasgos

	Macro Valor-F	Flagged	Urgente	Crisis
Solo lexicones	0.24	0.38	0.38	0.2
Lexicones + negaciones	0.19	0.43	0.37	0.04
TF-IDF + lexicones	0.38	0.71	0.53	0.44
TF-IDF + lexicones + negaciones	0.44	0.74	0.63	0.52
SVM + rasgos Tabla 4	0.43	0.77	0.59	0.51

Además de los mencionados experimentos, llevamos a cabo pruebas con clasificadores lineales y no lineales. El clasificador Naive Bayes fue entrenado con los mismos rasgos que nuestro clasificador SVM. Naive Bayes es un clasificador lineal que se caracteriza por ser simple y rápido. De manera similar, KNN fue entrenado con el mismo set de rasgos. Los hiperparámetros, como el número de vecinos, fueron seleccionados con un esquema de búsqueda de cuadrículas utilizando un rango de 1 a 25. La Tabla 6 muestra los resultados de Naive Bayes y KNN, los cuales dan peores resultados que SVM y otros sistemas del estado de la cuestión.

De manera similar, MLP fue entrenado utilizando el mismo set de rasgos que nuestro SVM. El tamaño de las capas ocultas varía entre 100 y 300, dependiendo del conjunto de datos de desarrollo. El modelo RNN + LSTM fue entrenado con representaciones de palabras incrustadas pre-entrenadas, sin añadir ningún otro rasgo, ya que una de las ventajas de este método es su capacidad de aprender representaciones de rasgos automáticamente. Hiperparámetros importantes, como el número de *epochs*, el tamaño de las capas ocultas y el tamaño del lote (*batch size*)

fueron escogidos utilizando una porción del set de entrenamiento como set de desarrollo. Los modelos de aprendizaje profundo tuvieron un peor desempeño que los modelos evaluados anteriormente. Esto no resulta sorprendente, ya que los métodos profundos requieren muchos datos de entrenamiento y el tamaño de los datos de Reachout es insuficiente, especialmente para algunas clases que tienen pocas instancias, como por ejemplo *rojo* y *crisis* (ver Tabla 2).

Tabla 6 Comparación del Valor-F de clasificadores utilizando el set de evaluación

Sistema	Macro Valor-F	Flagged	Urgente	Crisis
Base	0.3	0.61	0.44	-
Naïve Bayes	0.28	0.67	0.42	0.39
KNN	0.14	0.39	0.08	0
MLP	0.38	0.71	0.58	0.39
RNN+LSTM	0.28	0.44	0.01	0
Altszyler (2018)	0.44	0.9	0.68	0.48
TF-IDF + lexicons + negaciones	0.44	0.74	0.63	0.52
SVM + rasgos en Tabla 4	0.43	0.77	0.59	0.51

2.2 Análisis cualitativo de la clase crisis

Para llevar a cabo un análisis cualitativo de la clase *crisis* se seleccionaron 40 publicaciones de esta clase aleatoriamente. Algunas de las publicaciones contienen errores ortográficos y/o gramaticales que no han sido corregidos por los autores de este artículo. Hemos marcado estas publicaciones con la etiqueta [SIC]. Mientras que dichos errores podrían impactar negativamente en la calidad del procesamiento automático,

queda fuera del alcance de este estudio medir dicho impacto. En la muestra seleccionada se encontraron seis patrones lingüísticos a partir de los cuales sugerimos algunas recomendaciones para futuros modelos. Los patrones encontrados se detallan a continuación:

Expresar desesperación en crisis

Muchas de las publicaciones utilizan palabras que describen sentimientos de desesperación inmediata. Desesperación e impotencia pueden estar asociadas con un aumento de riesgo de suicidio (Cash et al. 2013). El sentimiento de impotencia puede tener su origen en la creencia de que las situaciones incómodas no tienen escapatoria, como por ejemplo en el caso de la siguiente publicación: *“Intenté hacer esto para mi ansiedad, pero acabé enfrentándome a estos desafíos”* (“*I tried doing this for my anxiety, but I ended up faced with these challenges*”). La desesperación es la combinación de los sentimientos de impotencia y depresión que resultan de una experiencia negativa (Cash et al. 2013, Liu et al. 2015). Un ejemplo de este tipo de sentimiento se encuentra en el siguiente testimonio: *“Ya tuve bastante de mi amigo enojo! No me molestaré en seguir intentando porque estoy frustrado”* (“*I am fed up with my friend anger! I can’t bother trying anymore because I am frustrated*”).

Las siguientes frases, extraídas de la clase *crisis*, expresan desesperación:

- ✓ *A veces también puedo sentirme bastante desesperado. Empiezo a preguntarme si alguna vez podré mejorar. Ya es bastante difícil vivir.”* (“*I can feel pretty hopeless at times too. I start questioning if I can ever get better. It’s hard enough to live.*”) (Post ID: 136600)
- ✓ *“No, pero estoy harto de toda mi vida en este momento y de mi existencia...”* (“*No, but I am pretty friggin sick of my entire life at this point and my existence...*”) (Post ID: 135818)

Recomendaciones: rasgos categoriales pueden modelarse utilizando las siguientes palabras clave: feel tired (sentirse cansado), fed up (harto), better dead (mejor muerto), give up life (renunciar a la vida), the end is near (el final está cerca), sick of life (harto de la vida), sick of existence (harto de la existencia), holding on (aguantar), hopeless times (tiempos desesperados), hope (esperanza), trying help or talking (intentar ayudar o hablar), hard to try or do (difícil intentar o hacer).

Publicaciones cortas de crisis y respuesta emocional

Las publicaciones cortas contienen descripciones concisas de los sentimientos negativos de una persona. En contraste con las publicaciones largas, las publicaciones cortas contienen mayor variación de sentimientos positivos y negativos. Como destacaron O’Dea et al. (2017), recursos léxicos como listas de palabras no son suficientes para detectar expresiones irónicas, sarcásticas y metafóricas. Por lo tanto, cualquier texto con menos de 50 palabras debe ser analizado con cautela. Otras limitaciones a la hora de interpretar publicaciones cortas incluyen el uso de la negación (Gkotsis et al. 2016).

Las siguientes son frases extraídas de la clase *crisis* de publicaciones cortas que describen sentimientos negativos:

- ✓ *Me estoy sofocando. Ya no sé si puedo hacer esto".*
"I'm suffocating. I don't if I can do this anymore." [SIC] (Post ID: 138064)
- ✓ *@redhead No sé cuánto tiempo podré mantener la compostura antes de estar jodido" ("@redhead I don't know how long I can even keep myself together before I'm screwed.") (Post ID: 138067)*

Recomendaciones: puede crearse un rasgo para denominar como cortas a las publicaciones con menos de 50 palabras, o publicaciones que no contienen más de 2 oraciones. Además, puede considerarse la posibilidad de detectar solamente sentimientos negativos y el tratamiento de todas las negaciones.

Publicaciones largas de la clase *crisis* y afrontamiento emocional

En las publicaciones largas de la clase *crisis* se detectó que los usuarios comienzan expresando emociones negativas, seguido de emociones positivas relacionadas con su capacidad para recuperarse. Esto podría verse como un signo positivo ya que la persona intenta superar una emoción negativa (Naslund et al. 2014). Sin embargo, los usuarios de estas publicaciones suelen expresar emociones negativas luego de demostrar signos de positividad.

Las siguientes son frases extraídas de publicaciones de la clase *crisis* que expresan patrones de insatisfacción con los servicios de salud:

- ✓ *(Negativo) Me siento extremadamente cansado cada mañana. Está llegando al punto en que contemplo llamar para avisar que estoy enfermo para no levantarme. (Positivo) A pesar del cansancio, me he levantado y me voy a trabajar, porque sé que necesito enfrenar el mundo. (Negativo) Sigo teniendo pensamientos para acabar con todo .. “(Negative) (Feeling extremely tired each morning. It’s getting to the point I’m contemplating ringing in sick to avoid getting up. (Positive) **Despite the tiredness, I’ve bee [SIC] getting up and going to work, because I know I need to face the world. (Negative) Keep having thoughts to end it all..”** (Post ID: 135898)*
- ✓ *(Negativo) No dormí hasta pasadas las 2 de anoche. Fue muy difícil dormir y luego despertarme esta mañana. Solo quería ignorar el mundo hoy. (Positivo) Finalmente me quedé dormido. Finalmente me levanté y me puse a trabajar. Me enfrenté al mundo y sonreí un poco. (Negativo) Realmente lo pasé mal durante mi turno de hoy ...” (Negative) (“Didn’t sleep til way after 2 last night. It was super hard to sleep, then wake this morning. I just wanted to ignore the world today. (Positive) **I eventually fell asleep. I eventually got up and went to work. I faced the world y smiled a little. (Negative) Really struggled through my shift today...”** (Post ID: 137919)*

Recomendaciones: puede crearse un rasgo para denominar como largas a las publicaciones que contienen más de 50 palabras. La identificación de emociones positivas y negativas en la misma publicación puede modelarse

también como un rasgo. Esto podría hacerse, por ejemplo, identificando una emoción negativa seguida de una serie de emociones positivas. Para detectar emociones positivas relacionadas con salud mental pueden usarse las siguientes palabras clave: *getting there (llegando)*, *I faced the world (me enfrenté al mundo)*, *getting up (me levanté)*, y *working (trabajando)*.

Insatisfacción con los servicios de salud

Uno de los patrones de comportamiento encontrados es la insatisfacción con los servicios de salud. Personas que buscan servicios de ayuda (p. ej., servicios de salud, consejeros, o tratamiento) algunas veces se sienten sin esperanzas, rechazadas o frustradas. Este patrón se encuentra en usuarios que parecen estar utilizando el forum para desahogar su insatisfacción o frustración con los servicios locales de apoyo a la salud mental, o de fracasos con respecto a su tratamiento.

Las siguientes son frases extraídas de publicaciones de la clase *crisis* que expresan insatisfacción con los servicios sanitarios:

- ✓ *(Servicio) “Hoy **mi médico de cabecera** se retrasó, lo que aumentó mi ansiedad. Al principio, no se dio cuenta de que era mi cita de seguimiento de 3 meses. Ella tampoco tenía idea de que el psicólogo debía escribir una carta, pero se le acabó el tiempo o se olvidó. Meh. Solo quería huir y esconderme. Fue MUY difícil no cerrarme y correr. Ni siquiera la escuché llamar mi nombre la primera vez. Blergh ...” (Service) (“**My gp** was running late today, which heightened my anxiety. At first she didn’t realise it was my 3month follow up apt. She also had no idea that the psych was meant to write a letter, so the psych either ran out of time or forgot. Meh. I just wanted to run away and hide. It was SO VERY hard not to close off and run. I didn’t even hear her call my name the first time. Blergh...”)* [SIC] (Post ID: 137919)
- ✓ *“estoy pensando mal en acabar con mi vida, nada ayuda, ni siquiera (Servicio) **mi consejero**” (“im having bad thought about ending my life, nothing helps not even (Service) **my counsellor**”)* [SIC] (Post ID: 136895)
- ✓ *“Sé que mirar atrás a las experiencias de terapia que no funcionaron solo me desanimará. Estoy muy impaciente y molesto. Estoy tratando de*

*encontrar los profesionales (de servicio) adecuados para mí, es un proceso muy frustrante. A veces también puedo sentirme bastante desesperado. Empiezo a preguntarme si alguna vez podré mejorar. Ya es bastante difícil vivir” (“I know looking back at therapy experiences that didn’t work out will only discourage me. I’m highly impatient and annoyed. I’m trying to find the right (Service) **professionals for me**, its [SIC] a very frustrating process. I can feel pretty hopeless at times too. I start questioning if I can ever get better. It’s hard enough to live”)* (Post ID: 136600)

Recomendaciones: rasgos categóricos pueden usarse para modelar menciones a los servicios de salud.

Contar historias para expresar una crisis

De acuerdo con Smithson et al. (2011), hay dos tipos de comportamiento en los usuarios de GAI que buscan ayuda. El primer comportamiento es el de una persona que quiere comunicar su historia, y el otro el de una persona que busca ayuda. Llegar a tiempo para ofrecer ayuda es crucial, pero si los consejos son dados muy tempranamente, es probable que sean rechazados. Se encontró que algunas personas se unen al forum para recibir consejos acerca de una situación realizando publicaciones donde se abren a contar sus problemas.

Las siguientes son algunas historias extraídas de la clase *crisis*:

- ✓ *(Evento) Así que hoy fui a los médicos y me dijeron que la quimioterapia en la que estoy no está funcionando, que mi cuerpo no está reaccionando como debería, lo que significa que ahora necesito comenzar este nuevo tratamiento. eso me va a golpear mucho más que la quimioterapia ast... (Event) (So today I went to the doctors and they told me that the **chemotherapy that I am on is not working**, my body isnt [SIC] reacting to it the way it should be, which means that I now need to start this new treatment that is going to knock me around a lot more then [SIC] the ast chemotherapy...” [SIC])* (Post ID: 136116)
- ✓ *(Evento) “Me mudé de casa a una relación de facto hace aproximadamente un año, y a pesar de tener problemas con mi madre, con quien solía vivir (madre soltera), tengo la sensación de que ella está*

muy sola y a menudo se pone llorosa por eso. (Evento) Mencionó hoy que también podría suicidarse porque siente que ya no vale la pena” (Event) (*“I moved out of home into a defacto relationship about a year ago now, and despite having troubles with my mum, who I used to live with (single parent), I have the feeling that she is very lonely and she often gets teary about that. (Event). She mentioned today that she may as well just kill herself because she feels like she’s not really worth it anymore.”*) (Post ID: 137384)

Recomendaciones: rasgos que reflejen la detección de secuencia de eventos personales. Los eventos personales pueden contener rasgos déicticos como *hoy*, *ayer*, y *mañana*. Además, pueden detectarse respuestas positivas y negativas relativas a cada evento.

Buscar ayuda de pares durante una *crisis*

Las publicaciones etiquetadas como *crisis* tratan más acerca de buscar ayuda que de dar apoyo a otros usuarios. Conseguir apoyo de otros usuarios en línea es un comportamiento habitual en personas con problemas de salud mental (Naslund et al. 2014]). Carron-Arthur et al. (2015) diferencia entre las publicaciones que tratan de proveer ayuda a otros usuarios y las que buscan ayuda de otros usuarios en GAI. Las publicaciones de apoyo se caracterizan por usuarios que proveen ayuda emocional e información sobre servicios de ayuda tales como, por ejemplo, ofrecer enlaces para encontrar ayuda. Por otro lado, las publicaciones de los usuarios que buscan ayuda se caracterizan por buscar compañerismo.

Las siguientes son frases extraídas de la clase *crisis* de usuarios que buscan ayuda:

- ✓ “... Sufriendo de ansiedad y depresión, este tipo de relación me está retrasando significativamente. (Búsqueda de consejos) ¿Alguna vez alguien más ha tenido una madre depresiva que les preocupa cuando se han mudado de casa? He ido a su casa a menudo y no estoy seguro de qué más puedo hacer para ayudarla realmente ...” (“... *Suffering from anxiety and depression [SIC] myself, this kind of relationship is setting me back quite significantly. (Advice Seeking) Has anyone else ever had a*

depressed parent that they are worried about when they move out of home? *I have been going to her place often and not sure what else I can do to really help her...*) (Post ID: 137384)

- ✓ “Ella está muy deprimida y siempre quiere morir. Estoy muy asustada y trato de ayudar, pero en el fondo soy bastante inútil para ayudar. (Buscando consejo) ¿Algún buen consejo? Porque a ella le gusta hablar conmigo porque soy agradable con ella y no la juzgo” (“*She’svery [SIC] depressed always wnat [SIC] to die. I’mpretty [SIC] scared and I try to help but deep down I’m pretty useless for helping.(Advice Seeking) Any good tips? Because she likes to talk to me because i’m nice to her and doesn’t [SIC] judge her.*”) (Post ID: 135748)

Recomendaciones: extraer rasgos sobre la detección de preguntas relacionadas con respuestas emocionales y de apoyo emocional.

3. Comparación con otras investigaciones

Nuestro mejor clasificador consigue resultados comparables con el estado del arte en la clasificación de publicaciones con los datos de Reachout y otros clasificadores de base. El sistema de base de Milne et al. (2016) utiliza rasgos de unigramas y bigramas y el clasificador de regresión lógica de la herramienta *scikit-learn* con sus parámetros por defecto (Pedregosa et al. 2011). El mejor sistema de clasificación de CLPsych 2017 también utiliza SVM. En este caso, utiliza un set de rasgos más complejos, incluyendo rasgos de la estructura del forum e interacciones entre las publicaciones, por ejemplo, hilos de conversaciones y hora de las publicaciones, entre otros. Este sistema consigue mejores resultados que nuestro mejor clasificador para las clases *flagged* y *urgente*. Interesantemente, nuestro mejor clasificador consigue mejores resultados para la clase *crisis*, que es la clase más importante ya que los moderadores deben responder a este tipo de publicaciones de forma inmediata. Al igual que en otros sistemas, encontramos que utilizar TF-IDF y lexicones ayuda a mejorar la clasificación. El mejor clasificador de la edición de CLPsych 2016 utilizó enigramas ponderados por TF-IDF e incrustaciones de palabras (*word embeddings*) de las publicaciones usando la herramienta *sent3vec* y un

clasificador SGD. El sistema de Kim et al. (2016) también utilizó enigramas ponderados por TF-IDF y consiguió resultados similares.

El estudio cuantitativo llevado a cabo en esta investigación se alinea con estudios previos que encontraron patrones parecidos en las interacciones en línea de personas con problemas de salud mental que utilizan redes sociales. Como se ha destacado en trabajos anteriores, las interacciones en línea entre pares (*peer-to-peer*) pueden mejorar la salud y conductas psicosociales al facilitar una variedad de comportamientos positivos que empoderen a los participantes, como la búsqueda de información y el apoyo emocional (Naslund et al. 2014). Sin embargo, estas redes sociales pueden al mismo tiempo resultar dañinas si, por ejemplo, el contenido promueve la autoflagelación, el suicidio o comportamientos relacionados con trastornos alimenticios (Dyson et al. 2016; Gerrard 2018). Particularmente, las redes sociales pueden promover contenido problemático que puede ser difícil de identificar (Gerrard 2018).

Este estudio se limita al uso de un solo set de datos. Más datos serían necesarios para conseguir mayor generalización y evitar la sobrevaloración de la capacidad predictiva de los modelos obtenidos. En segundo lugar, el tamaño del set de entrenamiento debería ser mayor. Por último, no se llevó a cabo un análisis de errores que examine el porqué de la confusión entre clases.

Conclusión

Este estudio demuestra una solución que utiliza una variedad de recursos léxicos y técnicas de aprendizaje automático para asistir a moderadores en GAI. En comparación con otras investigaciones similares, el presente trabajo extrae rasgos léxicos del contenido textual de las publicaciones con el fin de evitar posibles sesgos en la clasificación. Uno de nuestros clasificadores consiguió los mejores resultados para la clase de publicaciones etiquetadas como *crisis* (valor-F 0.52) y resultados competitivos en otras clases (*no-verde*, *flagged* y *urgente*). Estos resultados sugieren que sería posible construir un clasificador potente que sólo

necesite procesar el contenido textual de las publicaciones individuales. Además, nuestro estudio demuestra las limitaciones que existen al utilizar solamente rasgos léxicos a nivel de palabra.

La mayoría de los resultados de nuestro análisis cuantitativo podría traducirse en rasgos que potencialmente mejoren los resultados de clasificación. Como ya hemos mencionado, este análisis podría utilizarse para distinguir las características lingüísticas en los mensajes urgentes. Por ejemplo, rasgos específicos para caracterizar desesperación podrían mejorar la clasificación de la clase *crisis*. Sugerencias para trabajo futuro incluyen también la inclusión de rasgos para detectar a usuarios que buscan ayuda y para aquellos que la ofrecen, y la identificación de los roles de los participantes, por ejemplo, líderes, influyentes y opinadores (Carron-Arthur 2015).

Referencias bibliográficas

Altszyler, E., Berenstein, A. J., Milne, D. N., Calvo, R. A. y Slezak, D. F. (2018). Using contextual information for automatic triage of posts in a peer-support forum. En Loveys, K., Niederhoffer, K., Prud'hommeaux, E., Resnik, R., y Resnik, P., editors, *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, CLPsych@NAACL-HTL, New Orleans, LA, USA, June 2018*, pp. 57–68. Association for Computational Linguistics.

Carron-Arthur B., Ali K., Cunningham JA. y Griffiths KM (2015). *From help-seekers to influential users: A systematic review of participation styles in online health communities*. Journal of Medical Internet Research.

Brew, C. (2016). Classifying reachout posts with a radial basis function svm. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, págs. 138–142. Association for Computational Linguistics.

Cash, S.J., Thelwall, M., Peck, S.N., Ferrell, J.Z. y Bridge, J. A. (2013). Adolescent suicide statements on myspace. *Cyberpsychology, Behavior, and Social Networking*, 16(3):166–174. PMID: 23374167.

Cimino, A., Cresci, S., Dell'Orletta, F. y Tesconi, M. (2014). Linguistically-motivated y lexicon features for sentiment analysis of italian tweets. *4th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA 2014)*, pp. 81–86.

Cohan, A., Young, S. y Goharian, N. (2016). Triageing mental health forum posts. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical*

Psychology, pp. 143–147, San Diego, CA, USA. Association for Computational Linguistics.

Conway, M. y O'Connor, D. (2016). Social media, big data, y mental health: current advances y ethical implications. *Current Opinion in Psychology*, 9:77–82. Social media and applications to health behavior.

Gerrard, Y. (2018). Beyond the hashtag: Circumventing content moderation on social media. *New Media & Society*, 20(12):4492–4511.

Gkotsis, G., Velupillai, S., Oellrich, A., Dean, H., Liakata, M. y Dutta, R. (2016). Don't let notes be misunderstood: A negation detection method for assessing risk of suicide in mental health records. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, pp. 95–105, San Diego, CA, USA. Association for Computational Linguistics.

Griffiths, K. M. (2017). Mental health internet support groups: just a lot of talk or a valuable intervention? *World Psychiatry*, 16(3):247–248.

Hartzler, A. y Pratt, W. (2011). Managing the personal side of health: How patient expertise differs from the expertise of clinicians. *J Med Internet Res*, 13(3):e62.

Hollingshead, K., Irely, M. E. y Loveys, K., (eds.) (2017). *Proceedings of the Fourth Workshop on Computational Linguistics y Clinical Psychology — From Linguistic Signal to Clinical Reality*. Association for Computational Linguistics, Vancouver, BC.

Huh, J., Yetisgen-Yildiz, M. y Pratt, W. (2013). Text classification for assisting moderators in online health communities. *Journal of Biomedical Informatics*, 46(6):998–1005. Special Section: Social Media Environments.

Islam, M.R., Kabir, M.A., Ahmed, A., Kamal, A. R. M., Wang, H. y Ulhaq, A. (2018). Depression detection from social network data using machine learning techniques. *Health Information Science y Systems*, 6(1):8.

Kaplan, K., Salzer, M., Solomon, P., Brusilovskiy, E. y Cousounis, P. (2011). Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. 72:54– 62.

Kim, S. M., Wang, Y., Wan, S. y Paris, C. (2016). Data61-Csiro systems at the Clpsych 2016 shared task. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, pp. 128–132, San Diego, CA, USA. Association for Computational Linguistics.

Kornfield, R., Sarma, P. K., Shah, D. V., Mc-Tavish, F., Lyucci, G., Pe-Romashko, K. y Gustafson, D. H. (2018). Detecting recovery problems just in time: Application of automated linguistic analysis and supervised machine learning to an online substance abuse forum. *J Med Internet Res*, 20(6):e10136.

Kroenke, K., Spitzer, R. L. y Williams, J. B. W. (2001). The phq-9. *Journal of General Internal Medicine*, 16(9):606–613.

Le, Q. y Mikolov, T. (2014). Distributed representations of sentences and documents. En *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32, ICML'14*, págs. II–1188–II–1196. JMLR.org.

Liu, R.T., Kleiman, E.M., Nestor, B.A. y Cheek, S. M. (2015). The hopelessness theory of depression: A quarter-century in review. *Clinical Psychology: Science and Practice*, 22(4):345–365.

Malmasi, S., Zampieri, M. y Dras, M. (2016). Predicting post severity in mental health forums. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, pp. 133–137. The Association for Computational Linguistics.

Mikal, J., Hurst, S. y Conway, M. (2017). Investigating patient attitudes towards the use of social media data to augment depression diagnosis and treatment: a qualitative study. En *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality*, pp. 41–47, Vancouver, BC. Association for Computational Linguistics.

Milne, D. N., Pink, G., Hachey, B. y Calvo, R. A. (2016). Clpsych 2016 shared task: Triaging content in online peer-support forums. En *Proceedings of the Third Workshop on Computational Linguistics and Clinical Psychology*, pp. 118–127, San Diego, CA, USA. Association for Computational Linguistics.

Mohammad, S. y Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29:436–465.

Naslund, J. A., Grye, S. W., Aschbrenner, K. A. y Elwyn, G. (2014). Naturally occurring peer support through social media: The experiences of individuals with severe mental illness using Youtube. *PLoS One*, 9(10).

O’Dea, B., Larsen, M. E., Batterham, P. J., Calear, A. L. y Christensen, H. (2017). A linguistic analysis of suicide-related twitter posts. *Crisis*, 38(5):319–329. PMID: 28228065.

O’Dea, B., Wan, S., Batterham, P.J., Calear, A.L., Paris, C. y Christensen, H. (2015). Detecting suicidality on twitter. *Internet Interventions*, 2(2):183–188.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vyerplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. y Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pink, G., Radford, W. y Hachey, B. (2016). Classification of mental health forum posts. En *Proceedings of the 3rd Workshop on Computational Linguistics y Clinical Psychology*:

From Linguistic Signal to Clinical Reality, CLPsych@NAACL-HLT 2016, June 16, 2016, San Diego, California, USA, pp. 180–182.

Schwartz, H.A., Sap, M., Kern, M.L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., Blanco, E., Dziurzynski, L., Park, G., Stillwell, D., Kosinski, M., Seligman, M. E. y Ungar, L. H. (2016). Predicting individual well-being through the language of social media, pp. 516–527.

Shickel, B., Heesacker, M., Benton, S., Ebadi, A., Nickerson, P. y Rashidi, P. (2016). Self-reflective sentiment analysis. En *Proceedings of the Third Workshop on Computational Linguistics y Clinical Psychology*, págs. 23–32. Association for Computational Linguistics.

Smithson, J., Sharkey, S., Hewis, E., Jones, R., Emmens, T., Ford, T. y Owens, C. (2011). Problem presentation and responses on an online forum for young people who self-harm. *Discourse Studies*, 13(4):487–501.

Spärck Jones, K. (1972). "A Statistical Interpretation of Term Specificity and Its Application in Retrieval". *Journal of Documentation*. 28: 11–21.

Staiano, J. y Guerini, M. (2014). Depeche mood: a lexicon for emotion analysis from crowd annotated news. En *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, págs. 427–433, Baltimore, Maryland. Association for Computational Linguistics.

Tausczik, Y. R. y Pennebaker, J. W. (2010). The psychological meaning of words: Liwc y computerized text analysis methods. *Journal of Language y Social Psychology*, 29(1):24–54.

Vapnik, A. Y. L. (1963). Recognition of patterns with help of generalized portraits. volume 24, págs. 774–780.

van Genderen, M. y Vlaker, J. (2018). Virtual healthcare; use of virtual, augmented and mixed reality. *Nederlands tijdschrift voor geneeskunde*, 162:D3229.

Pennebaker, J., Boyd, R., Jordan, K. y Blackburn, . (2015). The Development and Psychometric Properties of LIWC2015. 10.15781/T29G6Z.

Zirikly, A., Kumar, V. y Resnik, P. (2016). The gw/umd Clpsych 2016 shared task system. En *Proceedings of the Third Workshop on Computational Linguistics y Clinical Psychology*, págs. 166–170. Association for Computational Linguistics.

Notas biográficas

Gabriela Ferraro

Gabriela Ferraro es investigadora en procesamiento del lenguaje natural y lingüística computacional en Commonwealth Scientific and Industrial Research Organisation, Australia y profesora honoraria en Australian National University. Su investigación trata sobre el uso del aprendizaje automático para enseñar a las computadoras a leer y escribir texto. Durante los últimos 15 años ha participado en numerosos proyectos relacionados con aplicaciones de tecnología del lenguaje como simplificación automática de textos, corrección de errores gramaticales, tecnologías del lenguaje para la salud mental, entre muchas otras. Le gusta enseñar y trabajar con estudiantes, y le apasiona trabajar y supervisar a jóvenes investigadoras. En 2018 recibió el premio Woman in Science Career Award en reconocimiento a su carrera.

Luis Salvador-Carulla

Luis Salvador-Carulla es el director del Centro de Investigación en Salud Mental de la Escuela de Investigación en Salud de la Población de la Universidad Nacional de Australia (ANU) en Canberra (Australia). Ha sido asesor de la Generalitat de Catalunya (España), el Ministerio de Sanidad español, la Comisión Europea (CE) y la Organización Mundial de la Salud (OMS). Su investigación se ha centrado en el desarrollo de sistemas de apoyo a la toma de decisiones en salud y política social, incluyendo herramientas para el análisis de la eficiencia técnica y la evaluación comparativa, indicadores para el análisis de políticas de salud y establecimiento de prioridades en salud mental y discapacidad. Ha coordinado el Proyecto Atlas Integrado de Salud Mental para mapear los servicios de salud mental en más de 30 áreas locales alrededor del mundo. Recibió el premio Leon Eisenberg de la Escuela de Medicina de Harvard en 2012 por sus contribuciones en el campo de los trastornos del desarrollo