

Comparación del reconocimiento multimodal de emociones en textos de Twitter

Comparison of multimodal emotion recognition in texts from Twitter

Juan Carlos Garduño-Miralrio^{1*}, David Valle-Cruz² y José Luis Tapia-Fabela²

¹Centro Universitario UAEM Temascaltepec, Universidad Autónoma del Estado de México
Carr. Toluca - Tejupilco Km. 67.5, Barrio de Santiago, 51300 Temascaltepec de González, Méx.
[*jgardunom010@alumno.uaemex.mx](mailto:jgardunom010@alumno.uaemex.mx)

²Unidad Académica Profesional Tianguistenco, Universidad Autónoma del Estado de México
Paraje el Tejocote, San Pedro Tlaltizapan, 52640 Santiago Tianguistenco, Méx.

PALABRAS CLAVE: RESUMEN

PALABRAS CLAVE:

Análisis de emociones, clasificación, CNN, redes sociales, Twitter

El presente artículo plasma los resultados de una investigación basada en la implementación de técnicas de aprendizaje profundo para la detección de emociones en textos, tomando como referencia la recolección de datos de la red social Twitter considerando que, según diversos estudios, es una fuente de información donde muchas personas expresan sus emociones a través de sus publicaciones. El objetivo de esta investigación fue resaltar el grado de precisión que presentan las redes neuronales convolucionales (CNN) para la identificación de emociones frente a otras técnicas de clasificación basadas en el aprendizaje automático, dado que la implementación de este tipo de red neuronal en el rubro del cómputo afectivo es relativamente nueva. A partir de un experimento con datos analizados previamente con otras técnicas de clasificación se obtuvieron resultados que favorecen en tema de precisión a las CNN frente a otros métodos de clasificación como: Conjunto de cadenas de clasificación (ECC), aprendizaje multietiqueta por retropropagación (BP-MLL), Vecino más cercano k de varias etiquetas (ML-kNN), Relevancia binaria (BR), Jerarquía de los clasificadores multietiqueta (HOMER) y otros más.

KEYWORDS: ABSTRACT

Emotion analysis, ratings, CNN, social networks, Twitter

This article presents the results of a research based on the implementation of deep learning techniques for the detection of emotions in texts, taking as a reference the collection of data from the social network Twitter considering that, according to several studies, it is a source of information where many people express their emotions through their publications. The objective of this research was to highlight the degree of accuracy of convolutional neural networks (CNN) for the identification of emotions compared to other classification techniques based on machine learning, given that the implementation of this type of neural network in the field of affective computing is relatively new. From an experiment with data previously analyzed with other classification techniques, results were obtained that favor CNNs over other classification methods such as: Ensemble of Classification Chains (ECC), Backpropagation Multi-Label Learning (BP-MLL), Multi-Label Nearest Neighbor k (ML-kNN), Binary Relevance (BR), Hierarchy of Multi-Label Classifiers (HOMER) and others.

• Recibido: 17 de julio 2021 • Aceptado: 10 de marzo 2022 • Publicado en línea: 28 de febrero 2023

1. INTRODUCCIÓN

El aprendizaje profundo es una técnica de la inteligencia artificial que ha permitido detectar patrones en los datos, realizar predicciones en tendencias y series de tiempo, además de crear artefactos tecnológicos que superan las capacidades de los seres humanos en problemas específicos como la detección de problemas pulmonares en radiografías. La combinación del aprendizaje profundo con el cómputo afectivo brinda posibilidades para detectar y predecir comportamientos, emociones, sentimientos y opiniones de una manera más precisa en comparación con otras técnicas de aprendizaje automático de acuerdo con el estudio de Zatarain-Cabada *et. al.* [1].

Una aplicación del aprendizaje profundo en el big data es la detección de emociones generadas en las publicaciones de redes sociales cuyo objetivo es catalogar los textos con base en las necesidades que se tengan. Sin embargo, existen múltiples técnicas que pueden implementarse en este sentido y es aquí donde surge una disparidad en el rendimiento (en cuestión de precisión) entre todas ellas, dando cabida a una “lucha” por saber cuál técnica es la más viable o precisa para cumplir con esta tarea. Según el estudio de Bollen *et al.* “se ha acumulado evidencia de que las redes sociales están asociadas con niveles elevados de soledad, ansiedad, disgusto e insatisfacción” [2], por lo cual, resulta el medio idóneo para poder implementar un análisis de detección de emociones y sentimientos negativos.

En las redes sociales y en la red en general se encuentran multitud de textos, en los cuales deben aplicarse subjetividad y no únicamente clasificarlos según su naturaleza o procedencia. Existen dos formas de enfrentarse al análisis de sentimientos: aplicando un enfoque semántico o aplicando un aprendizaje automático [3]. La implementación de este tipo de algoritmos puede extenderse al análisis de cualquier

tipo de texto dando cabida a la creación de herramientas que ayuden a la prevención o reversión de actitudes negativas en grupos de personas determinados o con algún fin en específico dictado por una organización.

1.1 TRABAJOS RELACIONADOS

Los resultados de experimentos realizados por Konate & Du [4], muestran que la CNN de una capa y la CNN-LSTM han logrado una mayor precisión que los modelos basados en el LSTM/BSLTM. En cuanto a los modelos clásicos, Naive Bayes se formó ligeramente mejor en el carácter tri-gramo en comparación con la palabra TF-IDF. Sin embargo, de manera similar a los modelos de aprendizaje profundo, el SVM logró un nivel de precisión menor en el carácter tri-grama comparado con la palabra TF-IDF.

De acuerdo con un estudio [5], se logró diseñar un algoritmo para el análisis de impresiones que tuvo una efectividad superior al 90% logrando de esta manera identificar impresiones en temas con tendencia en Twitter. En el estudio de Zatarain-Cabada *et. al.* [1] se llevó a cabo el reconocimiento de emociones implementando análisis de sentimientos con diversas técnicas obteniendo los siguientes resultados de acuerdo con las técnicas implementadas que se muestran en la Tabla 1.

En el ámbito psicológico se han estipulado diversas opiniones sobre la clasificación de emociones, principalmente al definir cuáles son las más básicas, destacando 5 trabajos, el primero de ellos donde Tomkins [6], propuso ocho emociones básicas: miedo, ira, angustia, alegría, disgusto, sorpresa, interés y vergüenza, por su parte Plutchik [6], define un pequeño número de emociones básicas o primarias que son la ira, la anticipación, el disgusto, la alegría, el miedo, la tristeza y la sorpresa, además, establece que el resto de las emociones son una mezcla de las

anteriores o derivan de ellas; esto quiere decir, que aparecen como combinación, mezcla o composición de las emociones primarias o emociones básicas. Plutchik afirma que todas las emociones varían en el grado de similitud entre ellas y que cada emoción puede existir variando el grado de intensidad o el nivel de activación.

Tabla 1. Precisión en las técnicas de procesamiento de datos con aprendizaje automático y profundo.

Técnicas implementadas	Eficiencia
Bernoulli NB	76.77%
Multinomial NB	75.31%
SVC	75.79%
Linear SVC	74.69%
SGDC Classifier	76.69%
KNN	68.46%
CNN + LSTM	88.26%

El OCC Model [6], presenta 22 categorías emocionales, agrupadas por pares opuestos: orgullo-vergüenza, admiración-reproche, felicidad-resentimiento, regodeo-compasión, esperanza-miedo, alegría-estrés, satisfacción-miedo-confirmado, alivio-decepción, gratificación-recuerdo, gratitud-ira y amor-odio; a principios de los años 90, Ekman [6], realiza una propuesta con un conjunto de seis emociones básicas: ira, disgusto, miedo, alegría, tristeza y sorpresa las cuales estaban directamente ligadas con las expresiones faciales que las caracterizan mientras que Parrott [6], presenta una lista de emociones más profunda, donde las emociones están categorizadas en una pequeña estructura de árbol, esta estructura tiene tres niveles: emociones primarias, secundarias y terciarias. Parrot toma como emociones primarias: amor, alegría, sorpresa, ira, tristeza y miedo.

El objetivo de esta investigación fue diseñar un algoritmo capaz de identificar las emociones plasmadas en textos con un grado de precisión elevado usando la técnica de aprendizaje profundo denominada redes neuronales convolucionales y contrastarlo ante otras técnicas de aprendizaje automático para determinar cuál técnica presenta mejores resultados de precisión.

2. METODOLOGÍA

Para la comparativa entre las diversas técnicas de clasificación para el análisis de emociones se elaboró una serie de experimentaciones aplicando técnicas de Machine Learning como el lexicón de SenticNet [7] y el clasificador Naive Bayes en conjunto con la Red Neuronal Convolucional con la finalidad de poder comparar de una mejor forma los resultados obtenidos.

El procedimiento para el tratamiento de los datos es el siguiente:

a) Recogida de datos:

El análisis de datos se realizó con los 2000 textos recolectados mediante el uso del servidor IFTTT o en español SIEEE que es un tipo de servicio web que permite crear y programar acciones para automatizar diferentes tareas y acciones en Internet, desde su sitio web y también desde su aplicación móvil [8].

b) Generación del corpus:

Se creo un corpus, seleccionando 1000 textos al azar de los datos obtenidos. Cada uno de ellos fue leído de forma independiente por dos personas especialistas en el reconocimiento de emociones en textos para asignar una etiqueta a cada uno, con los valores de las emociones identificadas en ellos, en aquellos casos en los que las etiquetas de ambas personas no coinciden en un texto una tercera persona se encargó de

etiquetar los textos para desempatar las mismas.

c) Preprocesamiento:

El preprocesamiento es una fase previa a la extracción de características para el análisis de la emoción, pero de suma importancia que implica los siguientes pasos:

1) Eliminar las palabras de parada.

2) Sustituir las vocales acentuadas por sus equivalentes no acentuadas.

3) Eliminar URLs, nombres de usuario, emoticones, números, espacios y signos de puntuación.

4) Aplicar el stemming o lematización: que es relacionar una palabra flexionada o derivada con su forma canónica o lema [9].

Todos los pasos anteriores de preprocesamiento se realizan automáticamente utilizando el lenguaje de programación Python.

d) Extracción de características:

Una vez etiquetados y preprocesados los textos, fue necesario extraer características útiles para aplicar métodos de aprendizaje profundo para el análisis de emociones.

e) Entrenamiento de clasificadores para el análisis de emociones Valle-Cruz et. al. [5]

Debido al objetivo de esta experimentación se decide clasificar los textos en dos grupos, el primero dedicado para los textos considerados como positivos y uno más para aquellos que han sido etiquetados como negativos. Para el clasificador NB se establecen los siguientes parámetros de probabilidad para las palabras tomando como ejemplo la oración “Me encanta este hermoso lugar”.

Tabla 2. Ejemplificación de la ponderación asignada a las palabras para su análisis.

w	P(w +)	P(w -)
Me	0.1	0.2
Encanta	0.1	0.001
Este	0.01	0.01
Hermoso	0.05	0.005
Lugar	0.1	0.1

Si $P(s|c) = \prod_{i \in \text{posiciones}} P(w_i|c)$ entonces:

$$P('Me encanta este hermoso lugar'|+) = 0.1+0.1+0.01+0.05+0.1=0.36$$

$$P('Me encanta este hermoso lugar'|-) = 0.2+0.001+0.01+0.005+0.1=0.316$$

$P(s|pos) > P(s|neg)$ entonces el texto se clasifica como positivo.

Las características, muy a menudo se clasifican de acuerdo con la cantidad de datos que se utilizan para su clasificación, la ganancia de información resulta ser una métrica bastante utilizada ya que indica cuantos bits de información resultan de la existencia de una palabra.

$$G(w) = - \sum_{i=1}^c P(c_i) \log P(c_i) + P(w) \sum_{i=1}^c P(c_i|w) \log P(c_i|w) + P(\bar{w}) \sum_{i=1}^c P(c_i|\bar{w}) \log P(c_i|\bar{w})$$

En el caso del lexicón de SenticNet se propone lo siguientes:

Basados en el modelo Houghlass de emociones [10], se determinaron las emociones positivas y negativas para llevar a cabo una clasificación similar a la elaborada con NB, obteniendo como resultado lo siguiente:

- Emociones positivas:
 - Interés, atracción, satisfacción, placer, amor, alegría, deseo, entusiasmo, gratitud y esperanza.
- Emociones negativas:
 - Impotencia, frustración, odio, rencor, aburrimiento, decepción, inquietud, miedo, apatía, ira, asco, tristeza y vergüenza.

Para la red neuronal convolucional se establece el siguiente modelo con base en pruebas de distribución de los datos:

- Entrada secuencial con dimensión 45x1 donde 45 es el número máximo de palabras por tweet.
- Matriz de *embedding* que representa la capa del mismo nombre.
- Capa de *bigramas* con su respectiva capa de *convolución* y *pooling*, la función de activación es *relu*.
- Capa de *trigramas* que conserva las características de la capa anterior.
- Capa de *cuatrigramas* con las mismas características.
- Capa oculta con 256 neuronas activadas con función *relu*.
- Capa de salida con una neurona activada por una función *sigmoid*

En conjunto con el modelo se establece la métrica *accuracy* y un optimizador *Adam* [11].

3. RESULTADOS

Una vez realizado el tratamiento de los datos se comienza con la fase de experimentación, con 4 épocas de

entrenamiento para la CNN obteniendo los siguientes resultados:

Tabla 3. Resultados obtenidos con la CNN.

Época (epoch)	Perdida (loss)	Precisión (accuracy)
1	0.3653	0.8401
2	0.3383	0.8544
3	0.3090	0.8691
4	0.2782	0.8838

Los resultados obtenidos tras 40 ejecuciones de la técnica en comparación con las dos técnicas adicionales que se implementaron se plasman en la Tabla 4 donde se muestran los promedios de la precisión alcanzada por cada una de estas.

Tabla 4. Resultados de la precisión obtenida en la clasificación de emociones con diversas técnicas de Machine Learning tras 40 ejecuciones.

Técnica	Perdida (loss)	Precisión (accuracy)
Lexicón SenicNet	0.2922	0.7523
Emociones positivas		0.8076
Emociones negativas		0.697
Clasificador NB	0.2853	0.7245
Emociones positivas		0.736
Emociones negativas		0.713
CNN	0.2782	0.8838
Emociones positivas		0.919
Emociones negativas		0.8486

4. DISCUSIÓN DE RESULTADOS

De acuerdo con la teoría de Freud, inferimos a partir de los tuits que las impresiones generadas por los usuarios de la

red social son una mezcla de diferentes emociones identificadas en la literatura, de este modo los marcos emocionales son útiles para determinar con mayor precisión el sentimiento en las redes sociales. Ekman, Tomkins y Parrot proponen emociones que son útiles para clasificar documentos, expresiones faciales y expresiones corporales [5].

Las redes neuronales convolucionales han presentado una precisión mayor a las dos técnicas con las que se ha contrastado, obteniendo una superioridad del 18%, considerando como la mayor precisión el 88.38% obtenido por las mismas, sin embargo, se presenta un grado de pérdida de datos lo que da cabida al mejoramiento del algoritmo. Cabe destacar que la implementación de las tres técnicas de aprendizaje automático ha permitido identificar que para los tres casos se ha tenido una mayor precisión en la clasificación de textos con emociones positivas que en el caso de las emociones negativas.

6. TRABAJO FUTURO

La presente investigación es el comienzo de un largo camino en el rubro del cómputo afectivo, como siguiente paso dentro del mismo se plantea llevar a cabo la experimentación con diversas arquitecturas de redes neuronales convolucionales y distintos optimizadores con la finalidad de poder mejorar el algoritmo presentado en este artículo. Posteriormente se implementará el resultado en el diseño de un nuevo algoritmo que sirva como herramienta para la detección de violencia de género en las publicaciones en redes sociales.

6. CONCLUSIONES

En este trabajo, se propuso una investigación seminal para examinar el

potencial de las redes neuronales convolucionales en el análisis de emociones, concretamente mediante el procesamiento de lenguaje natural para explorar las emociones que los usuarios de las redes sociales expresan en sus publicaciones.

Se implementó un enfoque de análisis de carga emocional y comparamos los resultados de los cálculos a partir del cálculo del promedio de precisión para cada una de las técnicas. Con base en el marco del reloj de arena de las emociones, se determinó que las emociones predominantes en los textos analizados fueron, el interés, la atracción, la satisfacción, el placer, el amor, la alegría, el deseo, el entusiasmo, la gratitud y la esperanza, ya que estas emociones alcanzaron los niveles más altos; esto significa que hubo altos niveles de emociones positivas en las publicaciones.

Los resultados demuestran que las redes neuronales convolucionales es una técnica más eficiente en cuanto a precisión en la detección de emociones a pesar de que la principal aplicación que se les da es la de reconocimiento de patrones y no la del procesamiento de lenguaje natural.

Consideramos que este tipo de investigaciones puede ser un medio para impulsar el uso de las redes neuronales convolucionales en el ámbito del procesamiento del lenguaje natural, para mejorar las herramientas existentes y ampliar el área de oportunidad para la creación de nuevos instrumentos orientados a este campo.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Zatarain Cabada, R., Barrón Estrada, M. L., Cárdenas López, H. M. Reconocimiento multimodal de emociones orientadas al aprendizaje, *Research in Computing Science*. 2019, 148 (7), 153-165.
- [2] Bollen, J., Gonçalves, B., van de Leemput, I., Ruan, G. The happiness paradox: your friends are happier than you. *EPJ Data Science*. 2017, 6 (4), doi: [10.1140/epjds/s13688-017-0100-1](https://doi.org/10.1140/epjds/s13688-017-0100-1).
- [3] ITelligent, Análisis de sentimiento, ¿qué es, cómo funciona y para qué sirve?, [en línea] Itelligent, Blog, [recuperado el 19 de julio de 2017] de <https://itelligent.es/es/analisis-de-sentimiento>. 2017
- [4] Konate, A., Du, R. Sentiment Analysis of Code-Mixed Bambara-French Social Media Text Using Deep Learning Techniques. *Wuhan University Journal of Natural Sciences*. 2018, 23, 237-243. doi: [10.1007/s11859-018-1316-z](https://doi.org/10.1007/s11859-018-1316-z).
- [5] Valle-Cruz, D., Lopez-Chau, A., Sandoval-Almazan, R. Impression analysis of trending topics in Twitter with classification algorithms. Impression analysis of trending topics in Twitter with classification algorithms. In Charalabidis, Y. Et al. (Eds.). *Proc. of the 13th Int. Conf. on Theory and Practice of Electronic Governance (ICEGOV '20)*. New York, ACM, 2020, 430-441. doi: [10.1145/3428502.3428570](https://doi.org/10.1145/3428502.3428570).
- [6] V. Francisco, *Identificación Automática del Contenido Afectivo de un Texto y su Papel en la Presentación de Información*, Tesis, Madrid, España: Universidad Complutense de Madrid, 2008.
- [7] Sentic Team, Talking about SenticNet, SenticNet, 2009. <https://sentic.net>
- [8] Fernández, Y. Qué es IFTTT y cómo lo puedes utilizar para crear automatismos en tus aplicaciones, Xataka, 2019. <https://www.xataka.com/basics/que-ifttt-como-puedes-utilizar-para-crear-automatismos-tus-aplicaciones>
- [9] Urdaneta-Fernández, L.A. Reducir el número de palabras de un texto: lematización y radicalización (stemming) con Python, Medium, 2019. <https://medium.com/qu4nt/reducir-el-número-de-palabras-de-un-texto-lematización-y-radicalización-stemming-con-python-965bfd0c69fa>
- [10] Cambria, E., Hupont, I., Hussain, A., Cerezo, E., Baldassarri, S. Sentic Avatar: Multimodal Affective Conversational Agent with Common Sense. In: Esposito, A., et al. (eds). *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues*. LNCS, 6456. Berlin: Springer, 2011, 81-95, doi: [10.1007/978-3-642-18184-9_8](https://doi.org/10.1007/978-3-642-18184-9_8).
- [11] Pérez-López, R. A. *Uso de redes neuronales convolucionales aplicado a sentiment analysis*, Tesis, Santiago, Chile: Universidad de Chile, 2019. <https://repositorio.uchile.cl/handle/2250/174536>

ACERCA DE LOS AUTORES



Juan Carlos Garduño Miralrio. Egresado de la Ingeniería en Sistemas Computacionales por el Tecnológico de Estudios Superiores de Valle de Bravo (TESVB) en el año 2016. Profe-

sor investigador del TESVB desde el 2017, acreedor al Estimulo al Desempeño Docente en 2019. Actualmente es estudiante de la Maestría en Ciencias de la Computación en el Centro Universitario UAEM Temascaltepec y colaborador del cuerpo académico “Procesamiento de datos con industria 4.0” de la Ingeniería en Sistemas Computacionales formado en el TESVB así como miembro del cuerpo académico “Tecnologías aplicadas a la conservación de especies y producción animal en entornos rurales” perteneciente al Centro Universitario UAEM Temascaltepec.



David Valle Cruz. Profesor en la Universidad Autónoma del Estado de México y miembro del Sistema Nacional de Investigadores de México. Ingeniero en Computación, tiene una Maestría

en Informática y un Doctorado en Economía y Administración. Ha sido investigador visitante en el Center for Technology in Government (CTG), SUNY Albany, NY, y en el Laboratorio de Ciencias de la Computación y Sistemas Multiagentes del CINVESTAV, Guadalajara, México. Sus artículos han sido publicados en importantes revistas, como Government Information Quarterly, Cognitive Computation, First Monday, Information Polity, e International Journal of Public Sector Management (entre otras). Sus intereses de investigación están relacionados con la inteligencia artificial aplicada, los medios sociales y las tecnologías emergentes en el sector público.



José Luis Tapia Fabela

Nació en la ciudad de México en 1977. En 2008 recibió el título de Doctor en Ciencias en Ingeniería Electrónica otorgado por el Instituto Tecnológico de Toluca (ITT). Actualmente se encuentra laborando como profesor adscrito en la Universidad Autónoma del Estado de México y como profesor de la Maestría en Ciencias de la Computación, sus áreas de interés son las siguientes: modelado de descargas de alta intensidad, cómputo aplicado a la resolución de problemas científicos y electrónica de potencia., ha realizado publicaciones en la Mexican International Conference on Artificial Intelligence y en IEEE transactions on plasma science y algunas otras revistas.