

# EDUCACIÓN

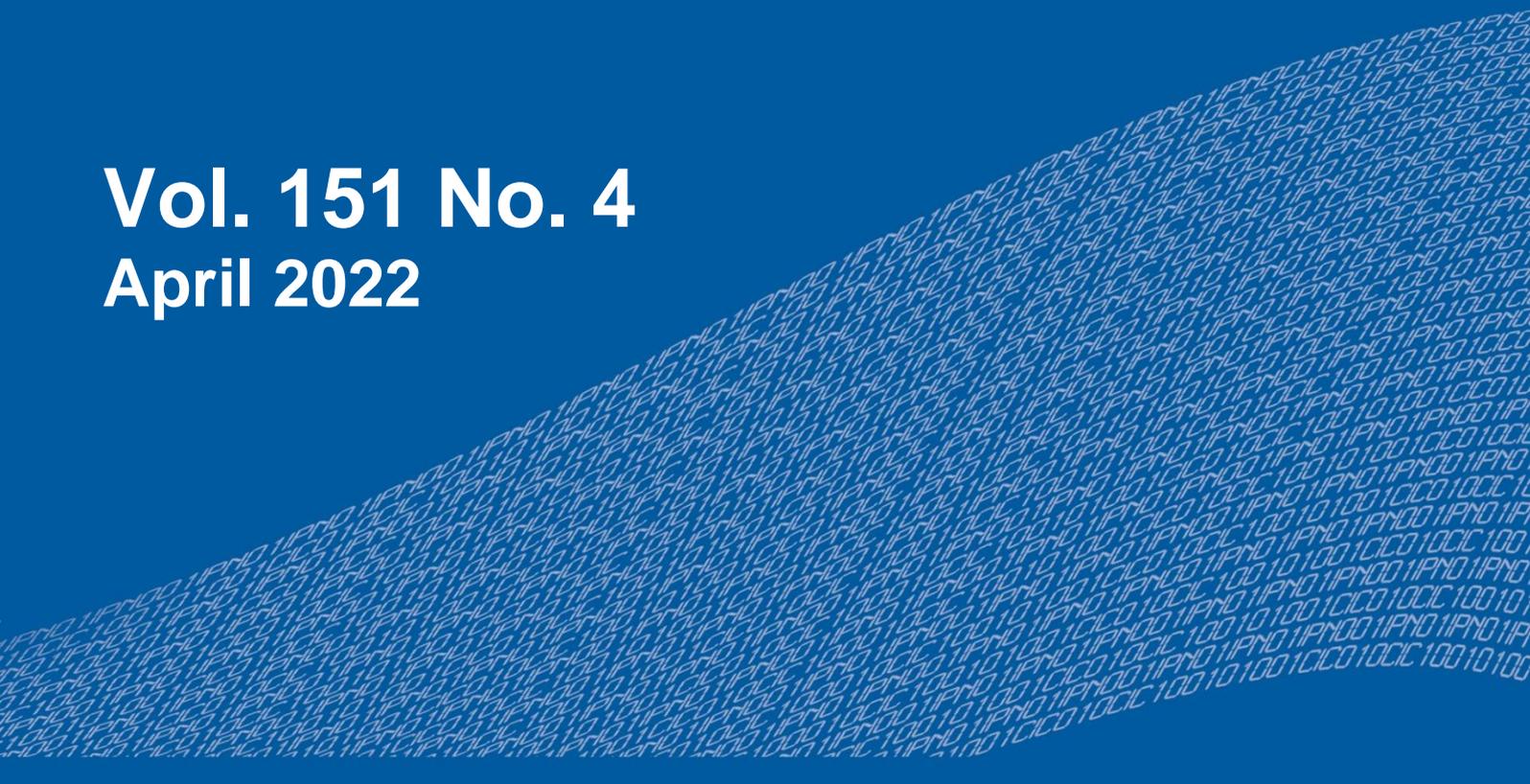
SECRETARÍA DE EDUCACIÓN PÚBLICA



Instituto Politécnico Nacional  
"La Técnica al Servicio de la Patria"

# Research in Computing Science

**Vol. 151 No. 4**  
**April 2022**



# Research in Computing Science

---

## Series Editorial Board

### Editors-in-Chief:

*Grigori Sidorov, CIC-IPN, Mexico*  
*Gerhard X. Ritter, University of Florida, USA*  
*Jean Serra, Ecole des Mines de Paris, France*  
*Ulises Cortés, UPC, Barcelona, Spain*

### Associate Editors:

*Jesús Angulo, Ecole des Mines de Paris, France*  
*Jihad El-Sana, Ben-Gurion Univ. of the Negev, Israel*  
*Alexander Gelbukh, CIC-IPN, Mexico*  
*Ioannis Kakadiaris, University of Houston, USA*  
*Petros Maragos, Nat. Tech. Univ. of Athens, Greece*  
*Julian Padget, University of Bath, UK*  
*Mateo Valero, UPC, Barcelona, Spain*  
*Olga Kolesnikova, ESCOM-IPN, Mexico*  
*Rafael Guzmán, Univ. of Guanajuato, Mexico*  
*Juan Manuel Torres Moreno, U. of Avignon, France*

### Editorial Coordination:

*Alejandra Ramos Porras*

*Research in Computing Science*, Año 21, Volumen 151, No. 4, abril de 2022, es una publicación mensual, editada por el Instituto Politécnico Nacional, a través del Centro de Investigación en Computación. Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, Ciudad de México, Tel. 57 29 60 00, ext. 56571. <https://www.rcs.cic.ipn.mx>. Editor responsable: Dr. Grigori Sidorov. Reserva de Derechos al Uso Exclusivo del Título No. 04-2019-082310242100-203. ISSN: en trámite, ambos otorgados por el Instituto Politécnico Nacional de Derecho de Autor. Responsable de la última actualización de este número: el Centro de Investigación en Computación, Dr. Grigori Sidorov, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738. Fecha de última modificación 01 de abril de 2022.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

Queda estrictamente prohibida la reproducción total o parcial de los contenidos e imágenes de la publicación sin previa autorización del Instituto Politécnico Nacional.

*Research in Computing Science*, year 21, Volume 151, No. 5, April 2022, is published monthly by the Center for Computing Research of IPN.

The opinions expressed by the authors does not necessarily reflect the editor's posture.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research of the IPN.

# Advances in Computing Science and Applications

Abdiel Reyes Vera (ed.)



Instituto Politécnico Nacional  
"La Técnica al Servicio de la Patria"



Instituto Politécnico Nacional, Centro de Investigación en Computación  
México 2022

## ISSN: in process

---

Copyright © Instituto Politécnico Nacional 2022  
Formerly ISSNs: 1870-4069, 1665-9899.

Instituto Politécnico Nacional (IPN)  
Centro de Investigación en Computación (CIC)  
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal  
Unidad Profesional “Adolfo López Mateos”, Zacatenco  
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Electronic edition

## Table of Contents

	Page
Metaphore Detection Using Convolutional Neural Networks .....	5
<i>Ericka Deyanira Ovando-Becerrill, Hiram Calvo-Castro</i>	
Criptoanálisis a sistema de cifrado de imágenes basado en múltiples cajas de sustitución .....	15
<i>Marco T. Ramírez-Torres, César A. Guerra-García, María F. Mendoza Martínez, Jazmín Rangel Saucedo, Luis J. Ontañón</i>	
Comparativa de operadores de reemplazo en algoritmos genéticos aplicado a problemas de optimización .....	27
<i>Carlos Alberto Rolon-Gonzalez, Yenny Villuendas-Rey</i>	
Open Data Analysis to Identify Behavior in Water Quality for Mexico City.....	41
<i>Emmanuel Alejandro Martínez Casillas</i>	
Estado del arte en el reconocimiento de emociones en textos de redes sociales .....	53
<i>Cristian C. Segura, Hiram Calvo</i>	



# Detección de metáforas utilizando redes neuronales convolucionales

Ericka Deyanira Ovando-Becerril, Hiram Calvo-Castro

Instituto Politécnico Nacional,  
Centro de Investigación en Computación,  
Laboratorio de Ciencias Cognitivas Computacionales,  
México

{eovandob2021, hcalvo}@cic.ipn.mx

**Resumen.** Este trabajo presenta una investigación sobre el estado del arte en la interpretación y generación de metáforas dentro del campo de la lingüística computacional. En la segunda sección, se muestran los resultados de un clasificador de lenguaje literal y no literal utilizando redes neuronales convolucionales (CNN) y redes neuronales multicapa (MLP). Finalmente, se presentan y comparan los resultados experimentales, destacando el rendimiento de los enfoques propuestos.

**Palabras clave:** Metáfora, lingüística computacional, procesamiento de lenguaje natural, lenguaje literal.

## Metaphore Detection Using Convolutional Neural Networks

**Abstract.** This work presents an investigation regarding of the state of the art around the interpretation of metaphors and their respective generation in the computational linguistics topics. At second section, this work shows the results of literal and non-literal language classifier using convolutional neural networks and multilayer neural networks and finally the results are presented and compared as an experiment.

**Keywords:** Metaphor, computational linguistics, natural language processing, literal language.

### 1. Introducción

Este trabajo tiene como punto fundamental la detección de la metáfora; en particular, se busca proponer un acercamiento claro al concepto de metáfora, al estado del arte y a la experimentación en torno al tema central de este proyecto. Es relevante para esta investigación identificar claramente conceptos, trabajos y

áreas de oportunidad que conduzcan hacia una propuesta innovadora y funcional que represente una aportación dentro del campo de estudio.

En cuanto a la revisión del concepto de metáfora, se establecen algunos puntos fundamentales, particularmente dentro de las ciencias computacionales cognitivas. En el estado del arte se mencionan las propuesta más relevantes en torno al problema a tratar y se presentan los resultados de: un clasificador mediante redes neuronales convolucionales (CNN) y un clasificador por redes neuronales multicapa. A manera de aclaración se propuso este clasificador por su versatilidad dentro del procesamiento del lenguaje natural y la posibilidad de considerar la información del contexto. Como punto de referencia se considera el auge durante los últimos años del Deep learning en el procesamiento de lenguaje natural para diversas tareas desde traducción hasta análisis contextual. Siendo el Procesamiento del Lenguaje Natural la intersección de Ciencias Computacionales, la Lingüística y la Inteligencia Artificial.

La necesidad del hombre de comunicarse o representar su entorno, mas allá de las imágenes nos remite a la época primigenia. El concepto del lenguaje y de comunicación trasciende entre filósofos, lingüistas y pensadores a través de la historia pero también en épocas más recientes entre computólogos e ingenieros. En este contexto podemos considerar a Platón en el Crátilo o Aristóteles en su Poética como ejemplo de los primeros trabajos sobre esta temática. Entender la metáfora exige comprender [a grandes rasgos] el lenguaje; el lenguaje nos obliga a cuestionar la existencia de los signos y cuestionar la existencia de los signos significa cuestionar todo nuestro entorno, ahora bien, entender todo lo anterior como un complejo sistema a clasificar no es un problema menor dentro del Procesamiento del Lenguaje Natural (PLN).

Pensar en torno a la idea del signo, remite en automático a San Agustín (354-430) y a Occam (1285-1349), de esta forma podemos ir repasando de manera cronológica la concepción de las diversas ideas y conceptos sobre el lenguaje y la metáfora, trabajo que sería muy complicado de llevar a cabo. Por lo anterior, esta primera sección tiene como objetivo identificar las directrices de este proyecto y justificar su relevancia.

A partir del lenguaje como proceso de significación se han acuñado conceptos como: “imagen”, “signo”, “alegoría”, “símbolo”, “emblema”, “parábola”, “mito”, “figura”, “icono” y “metáfora”. A partir de la teoría de los signos de Ricoeur, se puede considerar las ideas de la metáfora muy cercanas al símbolo. La metáfora se concibe como una innovación semántica, un acontecimiento propio del discurso. De acuerdo con Ciceron y Quintiliano la metáfora es una comparación abreviada, y de acuerdo con Aristóteles en su Poética es “la aplicación de una cosa de un nombre ajeno, con lo que efectúa la transferencia de género a especie, de especie a género, de especie a especie, o existe una analogía, esto es una proposición. De manera concisa la analogía es una figura retórica donde la semejanza sirve para sustituir una palabra por otra. Partiendo de este concepto “tradicional” podemos comprender la metáfora como el ejercicio de extrapolar una palabra de una semántica a otra, sin embargo, a partir de este concepto, ya complicado por sí mismo, surge,

de acuerdo con [4] la metáfora conceptual; que no se limita a funcionar en un sistema signado sino que está presente en la manera que nosotros entendemos el mundo, las asociaciones que generamos de manera natural y en los diversos sistemas de signos que nos rodean. A este respecto es importante aclarar que la metáfora no se limita, en este trabajo, a un concepto de una imagen literaria o un recurso lingüístico sino a una estructura compleja que puede tener o no un carácter lingüístico, y que se convierte en objeto de estudio, de interpretación, y también en un objeto a generar.

La metáfora, de acuerdo con [4], representa entonces una idea “origen” y una idea “objetivo” la cual está sujeta al contexto cultural y lingüístico. Por ejemplo, la idea de la vida como un viaje que puede estar presente en frase, en una historia, o una canción, es una metáfora conceptual, que trasciende los límites lingüísticos donde la idea de “origen” es el viaje y la idea “objetivo” es la vida. Otros ejemplos podrían ser la actividad laboral como una competencia, la supervivencia como un juego, o el amor como un viaje.

Por supuesto, el concepto de la metáfora no sólo es cercano al símbolo sino también a la metonimia, donde la designación del significado cambia de acuerdo con el contexto.

La complejidad del propio sistema de la metáfora se robustece con la complejidad del sistema lingüístico, en consecuencia, es un complicado problema dentro de la ingeniería. La metáfora, por si misma, es un reflejo de nuestras habilidades creativas como humanos XVeale16 ¿Cómo imitar o replicar esto desde el cómputo o cómo entenderlo desde la Inteligencia Artificial?

Como se ha analizado de manera breve en este trabajo, la complejidad del problema a abordar es alta, la metáfora es mas que frases o sistemas de palabras, es una forma de pensamiento, la representación del sistema semántico, la identificación y clasificación es un reto, un problema computacional digno de abordar desde el Procesamiento del Lenguaje Natural y fundamental para el desarrollo de la Inteligencia Artificial.

## **2. Estado del arte**

Los primeros proyectos sobre procesamiento del lenguaje natural (PLN) tuvieron un enfoque semántico, en el año 1975 XMason4. La búsqueda de la metáfora parte de los sistemas literatistas que tienen como base la identificación de del significado literal de una frase u oración para poder identificar las metáforas. Fue hasta 1979 que Aarts y Calbert utilizaron un marcador semántico, este sistema funciona a partir de cuatro características que englobaban principalmente el significado literal y el significado secundario. Retomando estos conceptos, Russell en 1992 desarrolla el sistema MAP, con características más específicas y tomando como referencia un diccionario semántico y el desarrollo de estructuras iniciales para expresiones metafóricas XVeale16.

Otra observación importante es que, a diferencia de los primeros algoritmos [6], en las propuestas de [11] y [12], la representación del diccionario semántico

no es mediante listas sino mediante un sistema de diccionarios que representan conceptos conectados por vértices bidireccionales. Aunque existen diversos trabajos realizados en torno a la detección y la generación de metáfora y en este proyecto será complicado mencionarlos todos, será de vital importancia para esta tesis realizar un análisis detallado de cada uno de ellos o en su caso de los más representativos, y poder enriquecer a partir de esto este proyecto. Como ya se ha mencionado, en la actualidad existen diversos autores y trabajos recientes que han realizado valiosas propuestas en torno a la detección de la metáfora, entre los mas representativos para este proyecto se consideran [12], [4], y [2] trabajos que resultan valiosos dentro de la clasificación de la metáfora, y se consideran un punto de partida para su estudio. Es importante mencionar que la interpretación y generación de metáforas que se tiene cómo objetivo final de este proyecto requiere primeramente de la evaluación de los posibles corpus disponibles para la tarea, el preprocesamiento, el modelo del algoritmo primeramente para su clasificación y posteriormente para la generación. De acuerdo a lo anterior es que se ha analizado el estado del arte.

Respecto a los corpus entre los mas mencionado en la literatura son TroFi, MOHnMOH-X y VUA, un referente para el estudio de metáfora en idioma ingles. Así mismo en muchos trabajos como referencia directa del análisis de metáfora se considera el contexto, y se utilizan como herramienta de apoyo WordNet y VerbNet. Un ejemplo es la clasificación de verbos.

Así mismo en torno al Procesamiento del Lenguaje Natural, se consideran a [5] y [1] para identificar las diversas estrategias de vectorización de un texto (word2vec, word vectors, embeddings etc...). Este material presenta el uso de Python para el PLN, y las bibliotecas: Keras, TensorFlow, NLTK y PyTorch cómo recursos importantes dentro de esta investigación.

Algunos otros trabajos no menos relevantes en torno a la tarea de detección de la metáfora son [8], [13], [7], [9], [3], los cuales proponen diversas soluciones, algunas de las mas recientes utilizando redes neuronales. Entre las herramientas que se identifican en los trabajos de los autores mencionados, para la clasificación de texto, se encuentran: redes neuronales, redes neuronales convoluciones, redes neuronales generativas, el algoritmo de Nave Bayes y redes neuronales recurrentes, en sus distintas configuraciones y combinaciones.

El trabajo en torno a la semántica de los diccionarios de metáfora se ha ido complejizando a la velocidad que el avance en los lenguajes de programación lo han permitido, trabajar [hoy en día] entendiendo estos sistemas semánticos metafóricos como sistemas interconectados permite una identificación y generación mas efectivas de la metáfora. Es importante mencionar que la clasificación de la metáfora se relaciona estrechamente con el contexto de la oración o texto que se está analizando.

Aunque existen varios trabajos en torno a la detección de la metáfora las áreas de oportunidad aún son muchas, la generalidad de respuesta de la mayoría de los algoritmos reportados es muy limitada. Los retos de la detección son tan amplios como el numero de idiomas en el mundo; sin

**Tabla 1.** Ejemplos de oraciones literal y no literal de Trofi Dataset del verbo absorber.

Lenguaje Literal	Lenguaje No literal
The Monitor's losses are absorbed by the church's working fund, which has declined during the past two years to about 200million from more than 280 million, largely because of the stock-market crash.	But in the short-term it will absorb a lot of top management's energy and attention, says Philippe Haspeslagh, a business professor at the European management school, Insead, in Paris.
The yellow beta carotene pigment absorbs blue -LRB- not yellow -RRB- laser light.	Mr. Wyss of Data Resources suggested that factories may be having difficulty hiring new workers at a time of such low unemployment or may simply not be able to absorb as many new workers as they would like.
This time, the ground absorbed the shock waves enough to transfer her images to the metal in bas-relief.	Now, do all of them absorb it and grab it?

embargo el aproximarnos a una solución viable representa un gran paso para el Procesamiento del Lenguaje Natural.

### 3. Desarrollo de la solución

#### 3.1. Descripción de la solución al problema planteado

Para esta primera propuesta experimental se implementaron:

1. Una red neuronal convolucional que clasifica lenguaje no literal y lenguaje metafórico
2. Una red neuronal que clasifica lenguaje literal y lenguaje metafórico.

Se desarrolló a partir de Trofi Dataset XGaoGsa, compuesto por 2,146 oraciones metafóricas y 1593 oraciones literales de 51 verbos en inglés.

El preprocesamiento consistió en generar 4 grupos de oraciones:

1. Training data: Literal
2. Training data: No Literal
3. Taste: Literal
4. Taste: No literal

En todas las oraciones: Se eliminaron los signos de puntuación y las stopwords y se convirtió todo el texto a minúsculas. El modelo que se utilizó para la vectorización del corpus fue: bag to words.

Es importante mencionar que para ambas redes se tomaron de referencia los ejercicios presentados en [10]. Así mismo se consideraron como base los modelos

para la clasificación de lenguaje literal y no literal, [1], y [3], respectivamente. Para obtener estos resultados, se trabajó en Python 3.8.

El modelo de la CNN se desarrolló a partir de Pythorch y esta compuesta por tres filtros convolucionales; cada capa convolucional con su respectiva función ReLu y seguida por su función Max Pooling. Para minimizar el error se utilizó el gradiente. La entrada al clasificador está compuesto por vectores  $v$  de  $[1 \times (\text{tamaño del vocabulario})]$  que corresponde a la representación de la oración a clasificar en un modelo vectorial bag to words.

La red neuronal de dos capas, se desarrolló a partir de Numpy. Se utiliza la función sigmoidea como base de las neuronas y se minimiza el error a partir de su derivada. La primera capa de neuronas consta de  $n$  neuronas, recibe la matriz de  $n \times m$ , donde  $m = \text{número de oración del conjunto de datos de entrenamiento}$  y  $n = \text{tamaño del vocabulario}$ . La segunda capa de este experimento se configuró para 30 neuronas.

Para ambos casos únicamente se tienen dos clases: Lenguaje literal y Lenguaje no literal o metafórico.

En este trabajo se presentan los resultados estadísticos de las redes propuestas y se analiza la viabilidad de ambas para solucionar el problema planteado en este documento y en un futuro proponer un algoritmo más robusto.

Estos dos ejercicios de experimentación resultan una propuesta interesante en primer lugar por los resultados preliminares obtenidos. Así mismo permiten plantear posibilidades y repensar el uso de herramientas de redes neuronales multicapa y redes neuronales profundas para proponer una solución más clara.

### 3.2. Evaluación

Para la evaluación de los clasificadores antes mencionados, se presenta la matriz de confusión para ambos casos, a partir de estos datos se consideran los siguientes parámetros:

$$Accuracy = \frac{TrueMetaphore + TrueLiteral}{TrueMetaphore + TrueLiteral + FalseMetaphore + FalseLiteral}, \quad (1)$$

$$Precisión(Metaphora) = \frac{TrueMetaphore}{TrueMetaphore + FalseMetaphore}, \quad (2)$$

$$Precisión(Literal) = \frac{TrueLiteral}{TrueLiteral + FalseLiteral}, \quad (3)$$

$$Recall = \frac{TrueMetaphore}{TrueMetaphore + FalseLiteral}. \quad (4)$$

Estos índices estadísticos nos permiten conocer el desempeño de los clasificadores y determinar mejoras a realizar.

#### 4. Experimentos y resultados

Los experimentos realizados para la clasificación de metáfora se llevaron a cabo mediante dos clasificadores distintos: el primero, utilizando un clasificador de redes neuronales convolucionales, y el segundo, mediante un clasificador de redes neuronales multicapa. A continuación se presentan algunos ejemplos de oraciones y su clasificación en los cuadros 2 y 3.

Los resultados se presentan en los cuadros 4 y 5 respectivamente.

**Tabla 2.** Ejemplo de oraciones clasificadas con redes neuronales convolucionales.

		<i>Clasificación</i>	
		<i>Literal</i>	<i>No Literal</i>
<i>Clase</i>	<i>Literal</i>	Under the proposed merger Bush adds two women aides Wheelabrator Technologies will to address women's issues and absorb Waste Management's plans a big push on child care Tampa Fla incinerator and three stressing a menu of choices for other waste to energy projects parents. the company has under.	
	<i>No Literal</i>	The cooling off period ordered as They do it to cool off the market part of Mr Reagan's naming of a so to speak because they need presidential emergency board in the copper and when they're April to try to resolve the issue. buying they find themselves chasing.	

**Tabla 3.** Ejemplo de oraciones clasificadas por medio de redes neuronales multicapa

		<i>Clasificación</i>	
		<i>Literal</i>	<i>No Literal</i>
<i>Clase</i>	<i>Literal</i>	The Monitor's losses are Operations at metal and absorbed by the church's wood coffin manufacturing working fund which reportedly plants in Missouri Indiana and has declined during the past two Pennsylvania will remain open years to about 200 million from and in fact be expanded with more than 280 million largely additional shifts to absorb some because of the stock market of the workers and production crash. idled elsewhere	
	<i>No Literal</i>	Management of three of the While the rest of the nation five highest yielding general watched with mere curiosity or purpose funds as ranked by the slight concern the nation's Donoghue Organization absorb fourth largest city was a portion of fund expenses. thoroughly absorbed by the approach of Hurricane Gilbert	

**Tabla 4.** Matriz de confusión para el clasificador convolución

		<i>Clasificación</i>	
		Literal	Metafora
<i>Clase</i>	Literal	125	26
	Metáfora	78	31

**Tabla 5.** Matriz de confusión para el clasificador de redes neuronales multicapa

		<i>Clasificación</i>	
		Literal	Metáfora
<i>Clase</i>	Literal	79	20
	Metafora	22	77

**Resultados estadísticos CNN.** Una vez presentadas las matrices de confusión se observa para el caso del clasificador por redes neuronales convolucionales:

$$Accuracy = 0.66$$

$$Precisión(Metáfora) = 0.32$$

$$Precisión(Literal) = 0.82$$

$$Recall = 0.28$$

**Resultados estadísticos redes neuronales multicapa.** Una vez presentadas las matrices de confusión se observa para el caso del clasificador de dos capas de redes neuronales:

$$Accuracy = 0.79$$

$$Precisión(Metáfora) = 0.77$$

$$Precisión(Literal) = 0.79$$

$$Recall = 0.79$$

## 5. Conclusiones y trabajo futuro

Una vez realizada una primera aproximación sobre la detección de metáfora en este trabajo, es evidente la relevancia de la generación y la interpretación

de la metáfora dentro del área del Procesamiento del Lenguaje Natural, que a pesar de la existencia de valiosas propuestas, requiere de nuevos métodos para la obtención de mejores resultados.

Respecto a los resultados de la experimentación reportados en este documento, se ve claramente un mejor desempeño en la red neuronal multicapa. Como trabajo a futuro se considera muy importante evaluar los modelos con un preprocesamiento mas fino y una representación vectorial más robusta cómo la mencionada en [3].

## Referencias

1. Antić, Z.: Python natural language processing cookbook : over 50 recipes to understand, analyze, and generate text for implementing language processing tasks. Birmingham Packt März (2021)
2. Bolognesi, M.: Metaphor and metonymy in the digital age. John Benjamins Publishing Company, (2019)
3. Gao, G., Choi, E., Choi, Y., Zettlemoyer, L.: CorMet: Neural metaphor detection in context. University of Washington, (2018)
4. Kövecses, Z.: Metaphor (2010)
5. Le, D., Thai, M., Nguyen, T.: Multi-task learning for metaphor detection with graph convolutional neural networks and word sense disambiguation. Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 8139–8146 (04 2020) doi: 10.1609/aaai.v34i05.6326
6. Masón, Z.: Cormet: Un sistema de extracción de metáforas convencional computacional y basado en corpus. Computational linguistics, (2004)
7. Maudslay, R., Pimentel, T., Cotterell, R., Teufel, S.: Metaphor detection using context and concreteness. University of Cambridge, (2020)
8. Minh Le, D., Thai, M., Huu, T.: Multi-task learning for metaphor detection with graph convolutional neural networks and word sense disambiguation. University of Florida, USA; University of Oregon, (2020)
9. Ruder, S., Vulić, I., Søgaard, A.: A survey of cross-lingual word embedding models. Journal of Artificial Intelligence Research, vol. 65, pp. 569–631 (08 2019) doi: 10.1613/jair.1.11640
10. Ugik: Text classification using neural networks (10 2017), <http://medium.com/@gk/text-classification-using-neural-networks-f5cd7b8765c6>
11. Veale, T., Keane, M. T.: The competence of sub-optimal theories of structure mapping on hard analogies. In: IJCAI (1). pp. 232–237. Citeseer (1997)
12. Veale, T., O'Donoghue, D.: Computation and blending, (2001)
13. Wilks, Y., Galescu, L., J., A., Dalton, A.: Automatic metaphor detection using large-scale lexical resources and conventional metaphor extraction. Florida Institute for Human and Machine Cognition,



## Criptografía a sistema de cifrado de imágenes basado en múltiples cajas de sustitución

Marco T. Ramírez-Torres, César A. Guerra-García,  
María F. Mendoza Martínez, Jazmín Rangel Saucedo,  
Luis J. Ontañón

Universidad Autónoma de San Luis Potosí,  
Coordinación Académica Región Altiplano Oeste,  
México

tulio.torres@uaslp.mx

**Resumen.** Debido a las necesidades actuales de seguridad, diversos autores han propuesto varios sistemas de cifrado, basados en diferentes principios de funcionalidad y aplicados a diferentes tipos de información. Las cajas de sustitución han demostrado ser una herramienta importante en los sistemas de cifrado simétrico como parte de las operaciones de confusión. Actualmente hay propuestas de sistemas de cifrado que incluyen el uso de múltiples cajas de sustitución, así como las llamadas cajas de sustitución dinámicas. Sin embargo, es importante analizar a profundidad los sistemas de seguridad que se encuentran publicados, antes de ser implementados o tomados como referencia. En este trabajo se analiza un sistema de cifrado para imágenes, propuesto por Majid Khan, donde por estadística y con ataques criptoanalíticos verificaremos la calidad del cifrado, utilizando el ataque chosen-plain image attack. El sistema es quebrado con este ataque, revelando la información original sin conocer la llave de cifrado, utilizando solo dos imágenes arbitrarias. Este trabajo trata de explicar a detalle la forma de realizar un criptoanálisis, para que a futuro las personas que diseñen un sistema de seguridad contemplen estas pruebas.

**Palabras clave:** Criptoanálisis, cajas de sustitución, cifrado de imágenes.

### Cryptanalysis of an Image Encryption System Based on Multiple Substitution Boxes

**Abstract.** Due to current security needs, various authors have proposed several encryption systems based on different functional principles and applied to different types of information. Substitution boxes have proven to be an important tool in symmetric encryption systems as part of confusion operations. Currently, there are proposals for encryption systems that include the use of multiple substitution boxes, as well as so-called dynamic substitution boxes. However, it is important to thoroughly analyze published security systems before implementing them or taking them as a reference. This paper analyzes an image encryption system proposed by Majid Khan. Using statistics and cryptanalytic attacks, we verify the quality of the encryption using the chosen-plain image attack. The system is broken with this attack, revealing the original information without knowing the encryption key, using only two arbitrary images. This paper

attempts to explain in detail how to perform cryptanalysis so that future security system designers can consider these tests.

**Keywords:** Cryptanalysis, substitution boxes, image encryption.

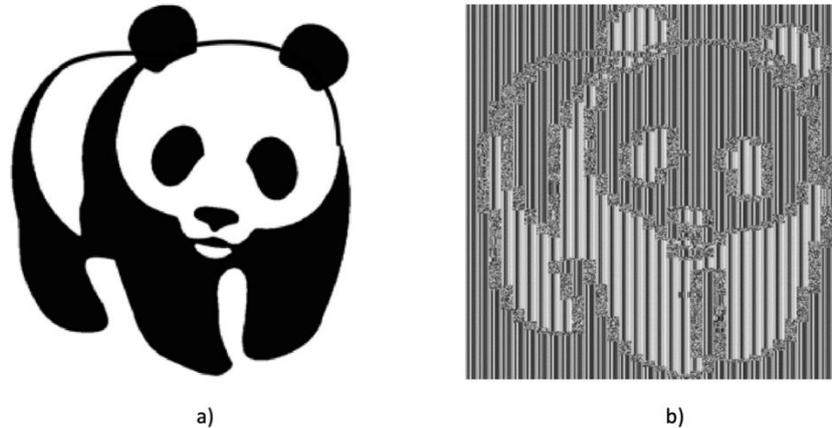
## 1. Introducción

En la actualidad, debido a la demanda de seguridad por todos los procesos que requieren trabajar en línea, y el almacenamiento seguro de información confidencial, se han propuesto diversos algoritmos criptográficos con diferentes enfoques, entre ellos los de enfoque caótico. Esto debido a propiedades que tienen como, ergodicidad y la sensibilidad a condiciones iniciales.

Por otra parte, el cifrado de contenido multimedia ha requerido el desarrollo de nuevos sistemas de cifrado, debido a que esta información requiere seguridad criptográfica y seguridad perceptual [1]. Para explicar esto, nos concentraremos en el caso de las imágenes digitales, las cuales pueden contener una gran cantidad de datos, una alta redundancia y correlación adyacente. Para cifrar una imagen no solo se requiere proteger la información modificando su contenido de una forma segura (seguridad criptográfica), si no que la versión cifrada debe evitar revelar patrones de la imagen original (seguridad perceptual). Un ejemplo de cuando no se consideran estas dos seguridades se puede ver en la Fig. 1.

Sin embargo, en muchas ocasiones estos nuevos sistemas presentan debilidades ante ataques de criptoanálisis y/o criptoanálisis diferencial. Provocando la fuga de información confidencial. Por ejemplo, en [2] utilizan una secuencia pseudoaleatoria generada por un sistema hipercaótico para cifrar las imágenes, utilizando la operación XOR y la suma modular. Este sistema fue analizado en [3] y se encontraron debilidades al momento de aplicar el ataque chosen-plain image attack (CPIA). Por otra parte, en la Ref. [4] se propone un método de cifrado y de generación de cajas de sustitución basado en caos. Sin embargo, en 2018 se publicó en [5], las debilidades del sistema propuesto por Çavuşoğlu et. al, bajo el ataque chosen-plaintext attack (CPA). En [6] los autores proponen un algoritmo de cifrado de imágenes caótico, basado en la entropía de la información. En el mismo año, Li et. al, en [7] revelan los problemas de seguridad que presenta dicho sistema contra ataques diferenciales. Por lo que podemos ver, es necesario un análisis mas profundo con diferentes escenarios para poder validar un nuevo sistema de cifrado.

En este trabajo analizamos una propuesta hecha por Majid Khan [8], la cual presenta una debilidad ante el ataque CPIA. El sistema utiliza los parámetros iniciales de un sistema caótico como llave secreta para generar cajas de sustitución. Estas cajas son también llamadas S-box, en este sistema las emplean para cifrar imágenes sustituyendo los valores de los coeficientes de los píxeles. En el artículo original presentan una validación del sistema, principalmente basada en pruebas estadísticas y con poca variedad de imágenes. Por lo que en esta investigación trataremos de describir a detalle el ataque, su justificación y referencia, para que en futuras propuestas contemplen una mayor variedad de pruebas de forma justificada.



**Fig. 1.** a) Imagen original de dos tonos, b) versión cifrada de a) con el sistema AES en modo ECB.

Este artículo tiene la siguiente estructura, en la sección 2 se describe el sistema propuesto por Khan, tanto la generación de las cajas de sustitución como el algoritmo de cifrado. La sección 3 muestra el criptoanálisis aplicado paso a paso. Y finalmente, en la sección 4 se incluyen las conclusiones.

## 2. Descripción del sistema de cifrado

El sistema de cifrado propuesto por Khan [8] utiliza sistemas caóticos que suelen ser empleados en criptografía debido a su ergodicidad y a que son sensibles a condiciones iniciales. Dicho esquema está basado en un sistema caótico multi-paramétrico, utilizando la combinación de los sistemas de Lorenz y Rossler. El autor afirma que este sistema es altamente no lineal y tiene comportamiento caótico. El conjunto de ecuaciones de este sistema se muestra a continuación:

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x) - y - z, \\ \frac{dy}{dt} &= rx - y - 20xz + x + ay, \\ \frac{dz}{dt} &= 5x - bz + \delta + x(x - c).\end{aligned}\tag{1}$$

Para el sistema de ecuaciones (1) se establecen los parámetros  $\delta$ ,  $r$ ,  $a$ ,  $b$  y  $c$  que son los que gobiernan las características de salida y los valores generados por el sistema.

El algoritmo de cifrado de imágenes es descrito a continuación:

1. Las trayectorias son generadas en el sistema caótico multi-paramétrico, seleccionando las condiciones iniciales.
2. Se colectan las muestras de la trayectoria de datos caóticos.

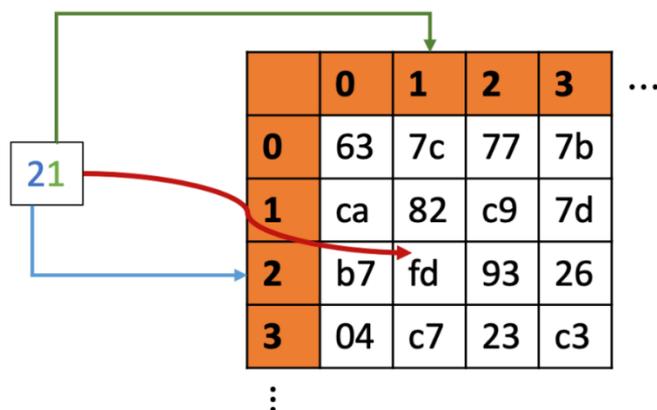
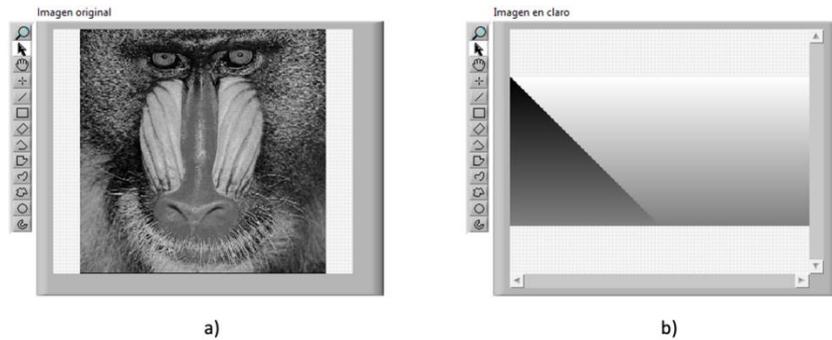


Fig. 2. Sustitución del valor 21 (hex) utilizando una caja de sustitución.

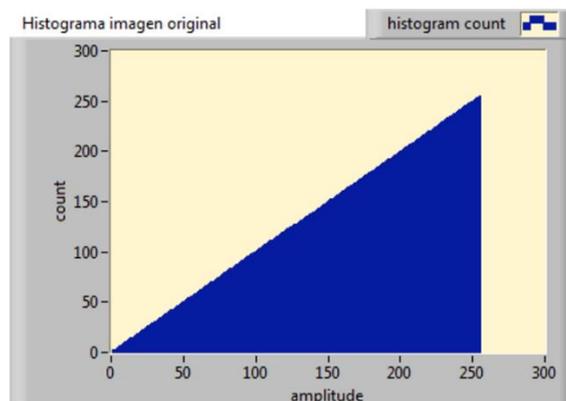
3. Aplicando una transformación afín, las muestras de la salida son ajustadas en un rango de 0 a 255.
4. Se conforma la S-box como una matriz de  $16 \times 16$ , utilizando las muestras ajustadas en el paso anterior.
5. Los píxeles de las imágenes que serán los datos de entrada a la S-box, son convertidos a números binarios de ocho bits.
6. Dicho número binario debe ser separado en dos bloques de cuatro bits, también llamados nibbles y convertidos nuevamente a base decimal.
7. La imagen cifrada se obtiene al usar los dos números (de cuatro bits) del paso 6, para sustituir el valor del píxel. Estos dos números serán usados como coordenadas de renglón y columna en la S-box. El número de los cuatro bits más significativos señala en la S-box, el renglón del valor de sustitución. Mientras que el número de los cuatro bits menos significativos señala la columna. El número localizado en la S-box se utiliza para sustituir el coeficiente en la imagen original. Este paso se repite hasta cifrar toda la imagen.

Para explicar con más detalle el paso 7 del algoritmo, suponga el caso de un píxel cuyo coeficiente es 21 (hex). En la Fig. 2 se puede ver una sección de la S-box. Los bits más significativos del valor de entrada 21 (hex), señalan el renglón 2, mientras que los menos significativos la columna 1. Por lo tanto, el valor 21 (hex) se sustituye por fd (hex).

Estos son todos los pasos que abarca el esquema de cifrado diseñado por Khan. Consta únicamente de una función de confusión que intercambia los valores de entrada (los coeficientes de píxel) por los valores de la caja de sustitución.



**Fig. 3.** Imágenes en claro. a) Imagen de mandril, ampliamente utilizada en el procesamiento de imágenes. b) Imagen hecha con todos los colores de la escala de grises de 8 bits.



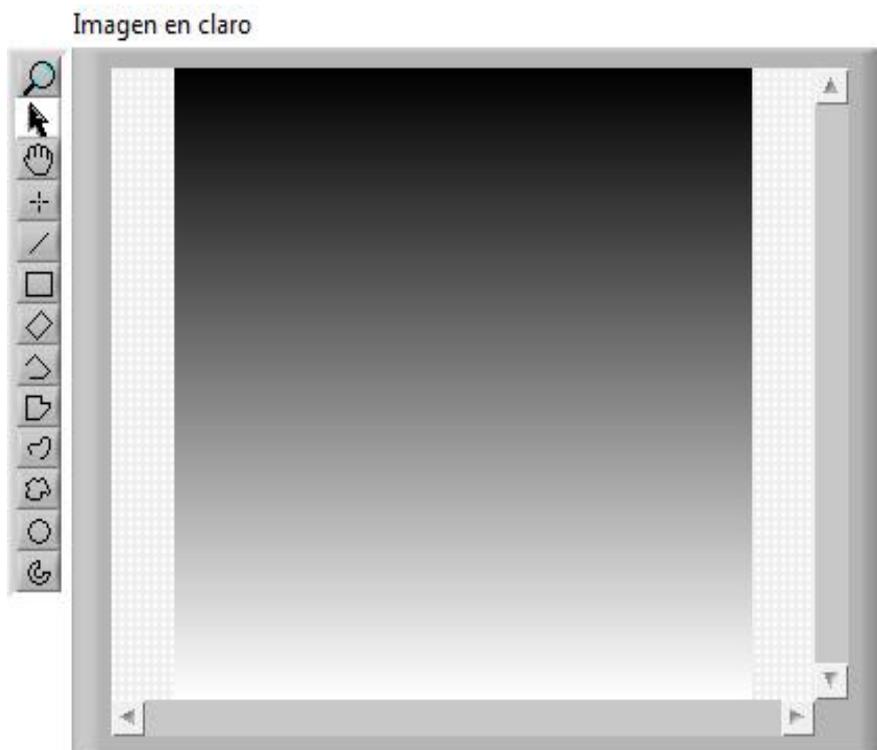
**Fig. 4.** Histograma de los coeficientes de los píxeles de la Fig. 3b).

Esta es toda la descripción que se da en el artículo, en el resto se dedican a validar el cifrado con pruebas estadísticas y criptoanálisis diferencial.

### 3. Criptoanálisis propuesto

Antes de comenzar a describir el ataque que se emplea en esta investigación, es importante definir lo que es una debilidad o ruptura en un sistema de cifrado. En [9] el autor define romper un cifrado, como encontrar una debilidad en el cifrado que puede ser explotada con menor complejidad que por fuerza bruta. Por lo tanto, es necesario aclarar que quizás se requieran cantidades irreales de conjuntos de textos conocidos o escogidos por el atacante. Así como considerar que el atacante tiene habilidades que tal vez no sean posibles en el mundo real.

El ataque utilizado en este criptoanálisis es el chosen plain-image attack (CPIA), es una versión específica para imágenes del ataque chosen plaintext attack, que ha sido



**Fig. 5.** Imagen en escala de grises hecha con los valores de 0 a 255.

utilizado en otros criptoanálisis, por ejemplo [10 y 11]. En este ataque el adversario es capaz de escoger las imágenes en claro a la entrada del sistema, y obtener las respectivas imágenes cifradas. En este ataque el adversario tratará de encontrar alguna debilidad en el sistema sin conocer la llave secreta, utilizando únicamente las imágenes en claro y cifradas.

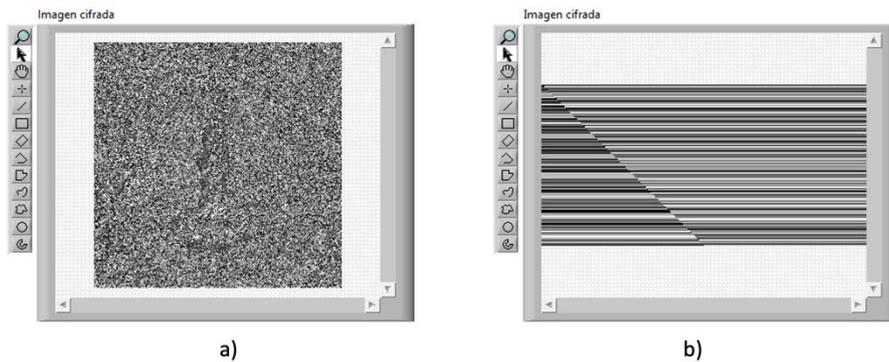
Debido a que es parte del propósito de este artículo servir como una explicación detallada de las pruebas de criptoanálisis, queremos aclarar que una vez especificadas las consideraciones del ataque, la forma en que se llevará a cabo no es única. Se pueden hacer diferentes usos con la información disponible, diferentes pruebas y análisis.

Así como la selección de las imágenes en claro pueden tener diferentes propósitos. Por lo tanto, repetir un ataque tal cual como es reportado en algún artículo, para un nuevo sistema puede no ser concluyente para medir la calidad del cifrado. Antes de replicar algo, se debe analizar el porqué de las condiciones y si cumplen con los supuestos del ataque.

Una vez explicado esto, se procede a realizar el ataque. Las imágenes seleccionadas por el atacante son de 8 bits y se muestran en la Fig. 3.

**Tabla 1.** Caja de sustitución de valores de 8 bits sin signo.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	74	3	93	16	37	8B	5B	E1	A2	94	F2	14	77	2A	2B	F9
1	7D	C	9C	1C	32	7C	25	C2	A8	9F	D4	1A	7F	B	F7	52
2	85	15	A5	22	2D	6E	13	B9	AE	AA	D2	34	8F	6	72	66
3	8E	1E	AD	2C	20	5F	1	B0	B4	AF	D0	3A	90	23	69	DC
4	97	27	B6	31	19	4F	EF	A7	C1	BD	C9	48	89	33	58	9D
5	A0	30	BE	35	11	3F	DD	8C	CF	CA	C0	56	87	64	50	7E
6	A9	39	C6	3D	0	2F	CC	86	E4	D1	B7	5D	62	96	F	45
7	B2	43	CE	41	F6	1F	BA	81	F3	D5	B5	65	5C	FB	51	26
8	BB	4C	DE	44	EC	E	88	76	18	D8	B1	91	4E	D	6B	7A
9	C4	55	E6	46	E2	FD	78	73	2E	DB	AC	B8	47	17	95	FF
A	CD	5E	ED	49	D7	EB	68	71	36	E3	AB	4	FE	29	F5	63
B	D6	67	F4	4A	CB	DA	59	6F	53	E5	A3	21	D3	9A	9B	83
C	DF	70	FC	4D	BF	C8	3B	6D	5A	E7	A1	28	BC	57	D9	24
D	E8	79	2	4B	B3	A4	12	6C	61	E9	C3	42	98	38	F3	1D
E	F1	82	9	40	A6	92	5	84	75	EE	C7	54	1B	A	E0	9E
F	FA	8A	10	3C	99	80	F8	8D	7B	F0	8	6A	3E	60	C5	7



**Fig. 6.** Imágenes cifradas. a) Imagen del mandril cifrada. b) Imagen 3b) cifrada.

La Fig. 3b) es en escala de grises a 8 bits (0 a 255), fue diseñada para aportar dos tipos de información en este ataque. Se sabe la localización y cantidad de píxeles con un valor de coeficiente específico. La imagen tiene dimensiones de  $128 \times 256$ , bajo el siguiente patrón:

$$\begin{bmatrix}
 0 & 255 & \dots & \dots & \dots & 255 \\
 1 & 1 & 254 & \dots & \dots & 254 \\
 2 & 2 & 2 & 253 & \dots & 253 \\
 \vdots & & \ddots & & & \vdots \\
 \vdots & & & \ddots & & \vdots \\
 127 & 127 & \dots & 128 & \dots & 128
 \end{bmatrix}, \tag{2}$$

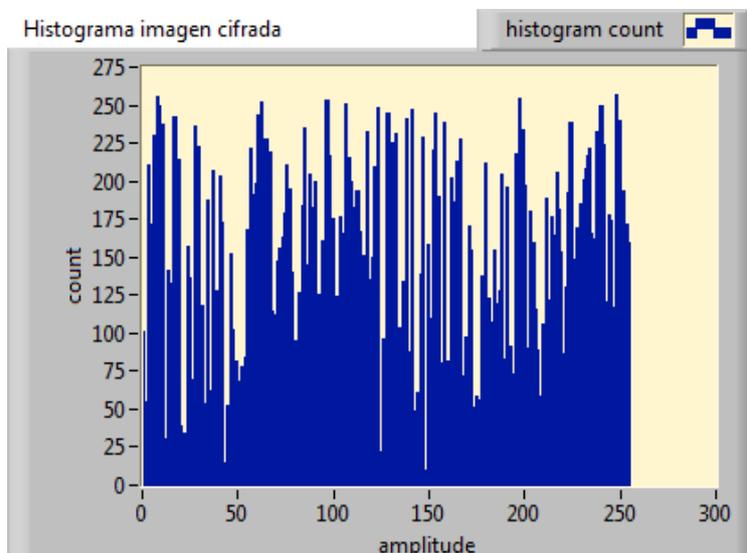


Fig. 7. Histograma de la Fig. 6b).

donde el primer renglón está formado por un píxel con el valor de 0 y 255 píxeles con ese mismo valor. El segundo renglón contiene dos píxeles con el valor de 1 y 254 píxeles con el valor de 254, y así sucesivamente un lado aumenta un píxel, mientras que el otro lo disminuye. El histograma de esta imagen, que es un gráfico para representar distribuciones de frecuencias, hecho con los valores de los píxeles, se ve de la siguiente manera.

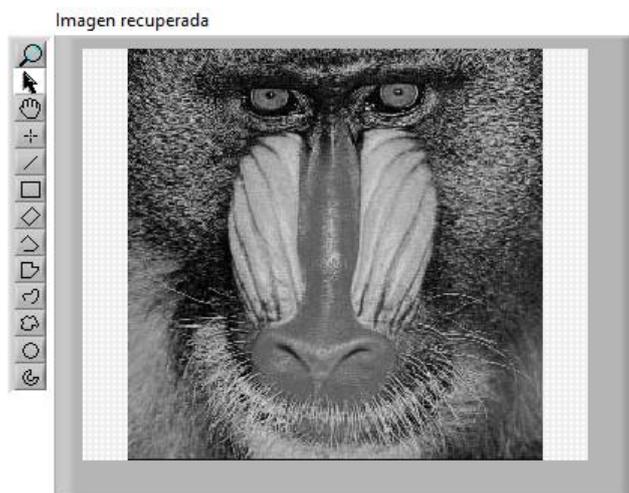
Se pudo haber utilizado una imagen más sencilla, como la de la Fig. 5, que es igual a la imagen seleccionada en los criptoanálisis [12]. Es un barrido de todos los valores de 0 a 255. Sin embargo, pensando en más casos, nuestra propuesta serviría para sistemas de cifrado similares, incluso si estos tuvieran una operación de difusión. Las operaciones de difusión sirven para cambiar de posición los píxeles, como por ejemplo las funciones de mapeo. Si el sistema propuesto por Khan o el sistema propuesto en [13], agregaran alguna función de mapeo, el criptoanálisis propuesto en [12] con la Fig. 5, no funcionaría.

Siguiendo con el ataque se seleccionan los parámetros iniciales del sistema caótico para poder generar la caja de sustitución. Para una mejor descripción utilizaremos la caja de sustitución reportada en el artículo por el autor, la cual se muestra a continuación.

Ambas imágenes en claro son cifradas con la misma caja de sustitución. El resultado se muestra en la Fig. 6.

Si además analizamos el histograma de la Fig. 6b), obtendríamos la siguiente gráfica:

Como se puede ver, el sistema de cifrado no es capaz de ocultar la redundancia de la imagen original, solo intercambia los valores. Un buen sistema de cifrado arrojaría un histograma uniforme, sin importar la imagen de entrada. Esto nos permite romper el sistema al menos de dos formas, que se podrían extender a calcular una caja de sustitución inversa. 1) Acorde a su posición en la imagen original. En la imagen cifrada



**Fig. 8.** Imagen recuperada con el ataque chosen plain-image attack.

podemos analizar los píxeles y acorde a su posición podemos deducir utilizando la imagen original, el valor por el cual se intercambiaron los valores 0, 1, 2, hasta 255. 2) Haciendo un análisis de histograma. Utilizando el histograma de la imagen cifrada Fig. 6b) es posible deducir los valores por los cuales se intercambiaron los coeficientes de píxel. Porque sabemos el número de píxeles de cada valor en la imagen original. Por ejemplo, el valor que solo tenga un píxel originalmente era el valor de 0, y el valor que tenga 2 píxeles originalmente era el valor 1.

Utilizando las imágenes de las Fig. 3b) y 6b), y en caso de que se requiera el histograma de la Fig. 7, podemos recuperar la Fig. 3a) con un porcentaje del 100% de igualdad. El resultado es la Fig. 8.

Como se puede ver, la imagen original del mandril se puede recuperar sin conocer la llave secreta. Se necesita solo una imagen conocida que contenga todos los valores de la codificación y se conozcan sus posiciones. Incluso el sistema de cifrado también sería vulnerable al ataque chosen ciphertext attack, al pasar la imagen de la Fig. 3b) por la operación de descifrado.

Si se analizan sistemas de cifrado que están bien establecidos como el AES (Advanced Encryption Standard), se puede ver que utilizan varias rondas con operaciones de confusión y difusión, incluyendo cajas de sustitución. Por lo tanto, la recomendación en el caso de diseño de sistemas de cifrado de imágenes, es hacer sistemas más robustos y combinar diversas operaciones para poder romper la alta correlación y la correlación adyacente.

Condiciones como que ambas imágenes se cifraron con la misma caja de sustitución, o dicho de forma más general, que ambas imágenes fueron cifradas con la misma condición inicial, son parte de los aspectos que, si el sistema de cifrado no limita en sus pasos, es válido aplicar en el criptoanálisis. Como se mencionó anteriormente, en los criptoanálisis las reglas son un poco más flexibles, permitiendo considerar que el adversario puede hacer este tipo de acciones de cifrar con la misma condición inicial, como seleccionar las imágenes en claro, etc. Prueba de la utilidad de éste, es que los

sistemas de cifrado avanzados pueden pasar estas pruebas, bajo las mismas consideraciones en los ataques.

#### 4. Conclusiones

En esta investigación se presenta un criptoanálisis a un sistema de cifrado para imágenes, basado en la generación de múltiples de cajas de sustitución a partir de un sistema caótico. El sistema de cifrado presenta una debilidad ante el ataque chosen plain-image attack, lo que implicaría que presenta debilidades ante otro tipo de ataques. Si el sistema hubiera sido evaluado con imágenes más simples, no hubiera pasado las pruebas estadísticas que reporta el autor en el artículo, desde la inspección visual. Otro problema que se puede señalar es que las cajas de sustitución generadas por el sistema caótico no son evaluadas. Cuando se propone un método de generación de cajas de sustitución, éstas se deben pasar por diversas pruebas, para confirmar que el método funciona y que las cajas cumplen con ciertas propiedades y estadísticas. En la ref. [14] se puede consultar un trabajo de esa área donde se propone un método y se explican las pruebas a las que se someten las cajas.

Una recomendación que se puede dar con este trabajo es que, al momento de diseñar sistemas de cifrado, se usen imágenes solidas que son el caso más simple. Si el sistema es bueno, será capaz de ocultar la alta redundancia de las imágenes en el cifrado. Esto se puede analizar calculando los histogramas.

Esperamos que este trabajo sirva de referencia para diseñadores de criptosistemas, y de esta manera contemplen aspectos de seguridad necesarios para poder pasar pruebas y ataques. Es muy importante evitar simplemente repetir punto a punto pruebas que se encuentran en la literatura, sin interpretar su porqué y si es aplicable al nuevo sistema que se esté diseñando.

#### Referencias

1. Lian, S.: *Multimedia Content Encryption: Techniques and Applications*. CRC press (2008) doi: 10.1201/9781420065282.
2. Zhu, C.: A Novel Image Encryption Scheme Based on Improved Hyperchaotic Sequences. *Optics Communications*, 285(1), pp. 29–37 (2012)
3. Li, C., Liu, Y., Xie, T.: Breaking a Novel Image Encryption Scheme Based on Improved Hyperchaotic Sequences. *Nonlinear Dynamics*, 73(3), pp. 2083–2089 (2013) doi: 10.1007/s11071-013-0924-6.
4. Çavuşoğlu, Ü., Kaçar, S., Pehlivan, I.: Secure Image Encryption Algorithm Design Using a Novel Chaos Based S-Box. *Chaos, Solitons & Fractals*, Elsevier, 95, pp. 92–101 (2017) doi: 10.1016/j.chaos.2016.12.018.
5. Zhu, C., Wang, G., Sun, K.: Cryptanalysis and Improvement on an Image Encryption Algorithm Design Using a Novel Chaos Based S-box, *Symmetry*, 10(9), 399 (2018) doi: 10.3390/sym10090399.
6. Ye, G., Pan, C., Huang, X.: A Chaotic Image Encryption Algorithm Based on Information Entropy. *International Journal of Bifurcation and Chaos*, 28(01), 1850010 (2018) doi: 10.1142/S0218127418500104.

7. Li, C., Lin, D., Feng, B.: Cryptanalysis of a Chaotic Image Encryption Algorithm Based on Information Entropy. In: IEEE Access, 6, pp. 75834–75842 (2018) doi: 10.1109/ACCESS.2018.2883690.
8. Khan, M.: A Novel Image Encryption Scheme Based on Multiple Chaotic S-boxes. Nonlinear Dynamics, 82(1), pp. 527–533 (2015) doi: 10.1007/s11071-015-2173-3.
9. Schneier, B.: A Self-Study Course in Block-Cipher Cryptanalysis. Cryptologia, 24(1), pp. 18–33 (2000) doi: 10.1080/0161-110091888754..
10. Chen, Y., Tang, C., Ye, R.: Cryptanalysis and Improvement of Medical Image Encryption Using High-Speed Scrambling and Pixel Adaptive Diffusion. Signal Processing, 167, 107286 (2020) doi: 10.1016/j.sigpro.2019.107286.
11. Li, C., Li, S., Chen, G.: Cryptanalysis of an Image Encryption Scheme Based on a Compound Chaotic Sequence. Image and Vision Computing, 27(8), pp. 1035–1039 (2020) doi: 10.1016/j.imavis.2008.09.004.
12. Ahmad, M., Doja, M.N., Beg, M.S.: Cryptanalysis and Improvement of an Image Encryption Scheme Using Fourier Series. 3D Research, 8(4), pp. 1–11 (2017) doi: 10.1007/s13319-017-0150-y.
13. Khan, M.: A Novel Image Encryption Using Fourier series. Journal of Vibration and Control, 21(16), pp. 3450–3455 (2014) doi: 10.1177/1077546314523029.
14. Aboytes-González, J.A., Murguía, J.S., Mejía-Carlos, M.: Design of a Strong S-box Based on a Matrix Approach. Nonlinear Dynamics, 94(3), pp. 2003–2012 (2018) doi: 10.1007/s11071-018-4471-z.



# Comparativa de operadores de reemplazo en algoritmos genéticos aplicado a problemas de optimización

Carlos Alberto Rolon-Gonzalez<sup>1</sup>, Yenny Villuendas-Rey<sup>2</sup>

<sup>1</sup> Instituto Politécnico Nacional,  
Centro en Investigación en Computación,  
México

<sup>2</sup> Instituto Politécnico Nacional,  
Centro de Innovación y Desarrollo Tecnológico en Cómputo,  
México

carlosagrolon@gmail.com, yvilluendasr@ipn.mx

**Resumen.** Los algoritmos genéticos han demostrado ser metodologías viables para resolver problemas de optimización, búsqueda y aprendizaje en las máquinas. El algoritmo genético clásico da lugar a una gran variedad de modelos, tomando en consideración los operadores genéticos que se implementen. Uno de los operadores genéticos con mayor impacto entre iteraciones es el operador de reemplazo. En este trabajo se presenta una comparativa y análisis de los resultados obtenidos por distintos operadores de reemplazo del estado del arte aplicado a funciones de minimización.

**Palabras clave:** Algoritmos genéticos, operadores de reemplazo, algoritmos evolutivos, optimización de funciones.

## Comparison of Replacement Operators in Genetic Algorithms Applied to Optimization Problems

**Abstract.** Genetic algorithms have proven to be viable methodologies for solving optimization, search, and machine learning problems. The classical genetic algorithm gives rise to a wide variety of models, taking into account the genetic operators implemented. One of the genetic operators with the greatest impact between iterations is the replacement operator. This paper presents a comparison and analysis of the results obtained by different state-of-the-art replacement operators applied to minimization functions.

**Keywords:** Genetic algorithms, replacement operators, evolutionary algorithms, function optimization.

## 1. Introducción

En los últimos años existe un creciente interés por el análisis de experimentos en el ámbito de los algoritmos evolutivos y las metaheurísticas [1]. Entre las técnicas evolutivas destacan los algoritmos genéticos (AG) por su adaptabilidad a los problemas y robustez.

Los AG son una técnica de búsqueda que se basa en la teoría de la evolución y selección de sistemas biológicos de Darwin [2]. Son métodos adaptativos y generalmente son utilizados en problemas de búsqueda y optimización de parámetros.

El desarrollo de los algoritmos genéticos se debe en gran parte a los trabajos de John Holland y sus colaboradores. En la década de 1960, Holland establece las teorías que imitaban el funcionamiento de selección natural y su adaptación en sistemas robustos [3]. Con base en las estrategias del cómputo evolutivo de Rechenberg [4, 5], Holland propone en [6] una heurística basada en los principios genéticos y evolución biológica como metodología para resolver problemas llamada algoritmo genético.

El algoritmo genético clásico toma una representación de un problema (codificación) y genera un grupo de posibles soluciones (llamada población). La idea de AG es que durante la ejecución del algoritmo la población evolucione, para que esto ocurra son aplicados los operadores genéticos: selección, cruce, mutación y reemplazo.

En el proceso de selección se eligen aquellas soluciones que presenten los mejores valores en su función de ajuste o *fitness*. En el proceso de cruce los individuos seleccionados producen soluciones descendientes de manera que contengan información parcial de cada uno de los progenitores. La mutación es un mecanismo que modifica de forma aleatoria algunos valores de los descendientes, de esta forma se desplazan valores numéricos hacia zonas del espacio de búsqueda que no pueden ser alcanzadas por los otros operadores genéticos. Por último, la técnica de reemplazo indica que miembros formarán parte de la siguiente población.

La capacidad de optimización y adaptabilidad de los AG permite que los investigadores puedan aplicarlos en áreas como robótica [7], economía [8], energía [9], medicina [10], geografía [11], entre otros. Dentro de los problemas de optimización encontramos las funciones de minimización. El problema de minimizar funciones [12] consiste en obtener el mínimo valor de una función real seleccionando sistemáticamente los valores de entrada, sujeto a un conjunto permitido, y computando el valor de la función. El desafío de estos problemas es obtener la mejor solución con la metodología más eficiente posible, es decir, aquella que utilice menos recursos y pueda generar un mayor rendimiento.

El algoritmo genético clásico da lugar a una gran variedad de modelos, tomando en consideración los operadores genéticos que se implementen. El operador de reemplazo juega un papel importante en el desarrollo de los AG; es el procedimiento encargado de decidir qué miembros de la población se preservarán entre generaciones. Bajo el teorema no-free-lunch [13] se puede aseverar que no existe el algoritmo ideal para todas las funciones posibles, por lo cual se analiza en este documento el impacto que tienen los operadores de reemplazo más utilizados en el estado del arte aplicados al problema de minimización de funciones.

## 2. Trabajos relacionados

### 2.1. Cómputo evolutivo

La computación evolutiva (CE) [14] es la rama de la computación que engloba a las técnicas inspiradas en la evolución biológica para la resolución de problemas, generalmente métodos de búsqueda estocásticos basados en ideas evolutivas de la selección natural y la genética. Los algoritmos de CE destacan por su fuerte adaptabilidad y autoorganización.

Aunque existe una variedad de modelos computacionales evolutivos, estos comparten una base conceptual común: la simulación de la evolución de estructuras individuales a través de procesos biológicos (llamados operadores genéticos). Estos procesos dependen del desempeño percibido de las estructuras individuales definidas por un entorno.

Los algoritmos evolutivos mantienen una población de estructuras que evolucionan de acuerdo con reglas de selección y otros operadores genéticos, como el cruce y la mutación. Cada individuo de la población es medido cuantitativamente por su aptitud al medio ambiente. La selección centra la atención en los miembros con mejor adaptación al ambiente. El cruce y la mutación perturban a esos individuos, proporcionando heurísticas generales para la exploración.

Entre los algoritmos más conocidos de la computación evolutiva están la programación evolutiva [15], estrategias evolutivas [16], programación genética [17] y los algoritmos genéticos [2]. Es en estos últimos se centra este artículo.

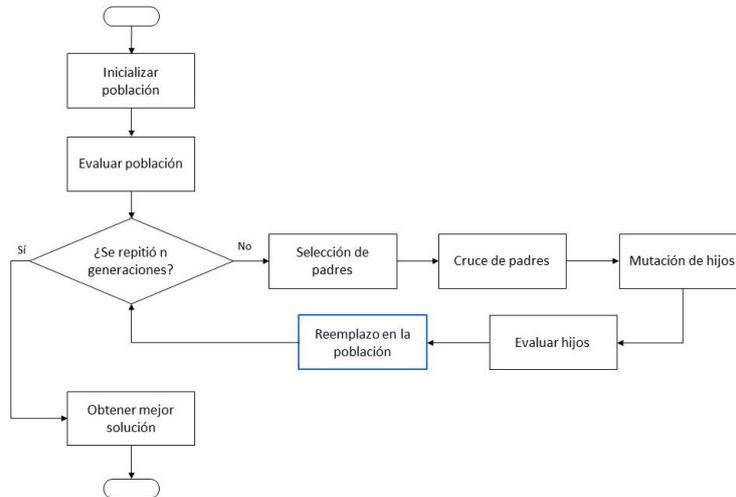
### 2.2. Algoritmos genéticos

Los algoritmos genéticos son técnicas que simulan la selección natural y el proceso evolutivo de los organismos vivos para resolver problemas de optimización, búsqueda y aprendizaje en las máquinas.

Esta estrategia pertenece a las técnicas basadas en poblaciones. Dado un problema particular a resolver, la entrada de AG es un conjunto de posibles soluciones al problema (llamado conjunto de individuos) codificadas de alguna manera y que, asociada una función llamada *fitness*, evalúa cuantitativamente cada solución candidata.

En cada iteración del algoritmo (llamado generación) se le aplican a la población los cuatro operadores de los AG: cruce, selección, mutación y reemplazo, en donde las soluciones candidatas son sometidas a acciones aleatorias semejantes a la evolución biológica.

Durante la ejecución del algoritmo el tamaño de la población se mantiene constante. Después de aplicar un número determinado de generaciones, la salida del AG es el mejor miembro de la población, es decir, la solución con el mejor valor *fitness*. El proceso de los AG se observa en la Fig. 1 y se explica a continuación:



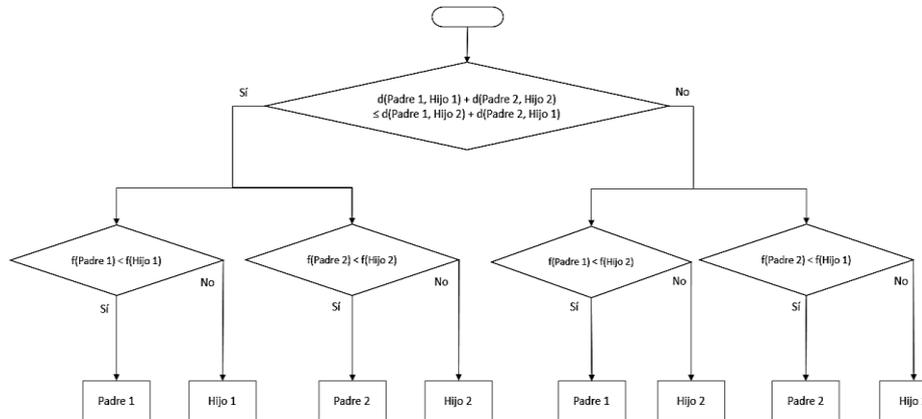
**Fig. 1.** Proceso de un algoritmo genético. La etapa de reemplazo en la población (remarcada en color azul) es el proceso de estudio de este trabajo.

- Inicializar población: Se genera una población inicial de soluciones candidatas, usualmente de forma uniformemente aleatoria
- Evaluación. Paso donde se calcula el valor *fitness* de cada solución candidata. Este proceso se realiza después de obtener la población inicial o cuando se obtiene una nueva solución descendiente.
- Selección: Se escogen qué individuos van a disponer de oportunidades de reproducirse. Las soluciones elegidas son llamadas *padres*.
- Cruce: Se genera soluciones descendientes (llamados *hijos*) a partir de las soluciones *padres*. La descendencia resulta de la recombinación de los *padres*.
- Mutación: Produce variaciones de modo aleatorio en cada una de las soluciones *hijo*, usualmente mediante un cambio aleatorio en la vecindad de la solución. La mutación es aplicada (con una tasa muy baja) a cada gen de la solución.
- Reemplazo: La población descendiente que se crea mediante la selección, cruce y mutación compite para reemplazar a algún miembro de la población inicial siguiendo algún criterio determinado. Estos son encargados de determinar qué soluciones continuarán en la siguiente generación.

En la siguiente sección se describirán algunos de los métodos de reemplazo más utilizados en el estado del arte.

### 2.3. Teorema de no free-lunch

Al estudiar los algoritmos genéticos se puede observar la gran gama de métodos que existen en cada operador de los que lo componen. Este trabajo analiza al operador de



**Fig. 2.** Proceso del reemplazo crowding. En primera instancia, se emparejan los padres con los hijos de acuerdo con una métrica de similitud (función  $d$ ). En una segunda fase, toma la decisión sobre cuál de ellos sobrevive en la población considerando el fitness (función  $f$ ).

reemplazo; estos algoritmos son encargados de elegir, mediante un criterio, los individuos de la población que se preservarán entre generaciones. La importancia de los algoritmos de reemplazo es que estos orientan al espacio de búsqueda preservando do soluciones con mayor adaptación observada hasta el momento, y manteniendo la diversidad genética.

En este contexto, cada investigador de cómputo evolutivo diseña y crea algoritmos genéticos con funciones de reemplazo esperando que la salida sea el mínimo global del problema, sin embargo, la demostración del teorema de no-free-lunch [13] excluye la existencia de un algoritmo o modelo ideal. El teorema de no-free-lunch indica que no existe un algoritmo que sea universalmente mejor que los demás, siempre existirán escenarios con una distribución de probabilidad en la que falle. Este teorema es muy importante y gobierna la efectividad de los algoritmos de optimización [18].

Ahora, teniendo en cuenta lo demostrado en el teorema de no-free-lunch, se presenta esta investigación donde se implementan una serie de algoritmos de reemplazo más empleados en el estado del arte con el objetivo de analizar su comportamiento en funciones de minimización.

### 3. Materiales y métodos

#### 3.1. Operadores de reemplazo

Los operadores de reemplazo utilizados en este trabajo son de los principales en el estado del arte [19], y se encuentran descritos a continuación.

- Reemplazo elitista: En este método consideran las soluciones *hijos* y toda la población para la competencia, la nueva población es compuesta por las mejores soluciones.

- Reemplazo de padres: Se obtiene espacio para la nueva descendencia (*hijos*) liberando el espacio ocupado por las soluciones *padres*.
- Reemplazo de padres por competencia: Los *hijos* compiten con los *padres* por ocupar su espacio en la población, entre este grupo son elegidas las soluciones con mejor *fitness*.
- Reemplazo proporcional. A cada miembro de la población e *hijos* se le asigna una probabilidad de ser elegido teniendo en cuenta su valor *fitness*. Los mejores individuos recibirán una mayor probabilidad de ser elegidos para la nueva población, pero no la certeza.
- Reemplazo crowding: Introducido por De Jong [20], es una técnica utilizada en algoritmos genéticos para preservar la diversidad en la población y evitar la convergencia prematura a óptimos locales. Consiste en emparejar cada descendencia con un individuo similar en la población actual y decidir cuál permanecerá en la población. En la Fig. 2 se puede observar este proceso.

### 3.2. Funciones de prueba

Las funciones utilizadas para probar los algoritmos genéticos fueron obtenidas de [21, 22, 23] y se muestran en la Tabla 1, estas son utilizadas como función *fitness* de cada algoritmo genético. Todas las funciones de prueba son problemas de minimización definidos como:

$$\text{Min } f(\mathbf{x}), \mathbf{x} = [x_1, x_2, \dots, x_D],$$

donde  $D$  es la dimensión de la función. Las funciones empleadas en este experimento presentan su valor de mínimo global cuando todas las componentes del vector solución son iguales a cero.

### 3.3. Parámetros del experimento

Los parámetros considerados para la ejecución de los algoritmos genéticos se muestran en la Tabla 2. Para cada función, se realizaron 10 ejecuciones independientes. La dimensión  $D$  se definió como  $D=15$  y los valores de cada dimensión se definieron como  $x_i \in [-100, 100]$ .

La metodología del operador de selección proporcional está definida en [24, 25], el operador de cruce Blx- $\alpha$  en [26] y el operador de mutación Gaussiana en [27]. Los valores de probabilidad de cada operador se definieron considerando [28]. Los algoritmos genéticos fueron desarrollados en el lenguaje de programación Python (versión 3.6) e implementados en Google Colaboratory [29].

**Tabla 1.** Funciones utilizadas.

Nombre	Función
Sphere	$f(\mathbf{x}) = \sum_{i=1}^D x_i^2,$ (1)
Rastrigin	$f(\mathbf{x}) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10),$ (2)
Schwefel	$f(\mathbf{x}) = \sum_{i=1}^D \left( \sum_{j=1}^i x_j^2 \right)^2,$ (3)
Rosenbrock	$f(\mathbf{x}) = \sum_{i=1}^{D-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2),$ (4)
Zakharov	$f(\mathbf{x}) = \sum_{i=1}^D x_i^2 + \left( \sum_{i=1}^D 0.5x_i \right)^2 + \left( \sum_{i=1}^D 0.5x_i \right)^4,$ (5)
Discus	$f(\mathbf{x}) = 10^6 x_1^2 + \sum_{i=2}^D x_i^2,$ (6)
HGBat	$f(\mathbf{x}) = \left  \left( \sum_{i=1}^D x_i^2 \right)^2 - \left( \sum_{i=1}^D x_i \right)^2 \right ^{1/2} + \left( .5 \sum_{i=1}^D x_i^2 + \sum_{i=1}^D x_i \right) / D + .5,$ (7)
Different Powers	$f(\mathbf{x}) = \sqrt{\sum_{i=1}^D  x_i ^{2+4\frac{i-1}{D-1}}},$ (8)
Happy Cat	$f(\mathbf{x}) = \left  \sum_{i=1}^D x_i^2 - D \right ^{1/4} + \left( .5 \sum_{i=1}^D x_i^2 + \sum_{i=1}^D x_i \right) / D + .5,$ (9)
Griewank	$f(\mathbf{x}) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1.$ (10)

### 3.4. Métricas de desempeño

Existen diversas métricas para medir el desempeño de algoritmos, la más utilizadas para los problemas de minimización son las medidas de tendencia central [30]: mínimo, máximo, promedio y desviación estándar. A continuación, se muestra cómo se calculan estas métricas:

$$\text{Min}(Y) = \min\{y_1, y_2, \dots, y_n\}, \quad (11)$$

$$\text{Max}(Y) = \max\{y_1, y_2, \dots, y_n\}, \quad (12)$$

$$\text{Promedio}(Y) = \bar{Y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (13)$$

$$\text{Desviación Estandar}(Y) = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{Y})^2}{n}}, \quad (14)$$

**Tabla 2.** Parámetros de los algoritmos genéticos.

Parámetro	Valor
Tamaño de población	10
Iteraciones	250
Máxima cantidad evaluaciones	500
Operador de selección	Proporcional
Operador de cruce	Blx- $\alpha$
Probabilidad de cruce	0.9
Operador de mutación	Gaussiana
Probabilidad de mutación (por gen)	0.08
Dimensión de la función	15
Valor posible en cada dimensión	$x_i \in [-100, 100]$
Ejecuciones independientes	10

**Tabla 3.** Mejores valores de los algoritmos genéticos con distintos métodos de reemplazo en funciones de minimización.

Función	Reemplazo elitista	Reemplazo proporcional	Reemplazo de padres	Reemplazo de padres por competencia	Reemplazo Crowding
Sphere	0.49	14.06	5.94E+04	<b>0.03</b>	1.26
Rastrigin	<b>67.67</b>	161.23	6.11E+04	97.12	122.29
Schwefel	938.43	1.16E+03	2.48E+05	836.42	<b>679.59</b>
Rosenbrock	36.39	103.03	7.80E+09	<b>5.85</b>	129.15
Zakharov	1.80E+03	3.41E+03	8.51E+04	<b>180.99</b>	626.70
Discus	25.23	31.62	1.24E+06	<b>0.19</b>	0.52
HGBat	0.46	34.46	7.28E+04	<b>0.41</b>	1.23
Different Powers	<b>1.07</b>	69.98	2.18E+05	7.58	1.80
Happy Cat	<b>0.44</b>	2.59	2.24E+03	0.69	0.65
Griewank	<b>0.54</b>	1.00	1.00	0.82	0.79
Total Ganados	4	0	0	<b>5</b>	1

donde  $n = 10$  y  $Y = [y_1, y_2, \dots, y_n]$  son los valores obtenidos de las 10 ejecuciones independientes del algoritmo genético.

#### 4. Resultados y discusión

En este trabajo se evalúan algoritmos genéticos con distintos operadores de reemplazo con el objetivo de optimizar funciones de minimización. Las métricas

**Tabla 4.** Peores valores de los algoritmos genéticos con distintos métodos de reemplazo en funciones de minimización.

Función	Reemplazo elitista	Reemplazo proporcional	Reemplazo de padres	Reemplazo de padres por competencia	Reemplazo Crowding
Sphere	16.23	1.47E+03	3.49E+05	<b>10.89</b>	215.84
Rastrigin	<b>131.83</b>	2.84E+03	1.83E+05	223.27	309.65
Schwefel	<b>9.14E+03</b>	1.26E+04	3.43E+06	1.24E+04	9.45E+03
Rosenbrock	7.79E+05	2.59E+06	5.88E+11	<b>2.19E+05</b>	1.33E+07
Zakharov	1.18E+04	2.50E+04	4.02E+09	<b>1.53E+04</b>	4.07E+04
Discus	5.43E+05	3.79E+07	7.23E+09	1.29E+05	<b>1.98E+03</b>
HGBat	<b>4.26</b>	880.17	4.32E+05	103.40	315.17
Different Powers	<b>131.91</b>	2.29E+03	1.86E+06	307.39	606.71
Happy Cat	<b>1.66</b>	73.54	1.25E+04	2.29	6.52
Griewank	<b>0.96</b>	1.28	3.16	0.98	0.97
Total Ganados	<b>6</b>	0	0	3	1

utilizadas para evaluar y comparar el desempeño de cada algoritmo fueron las medidas de tendencia central.

Los resultados del experimento se pueden observar en las Tablas 3-6, cada tabla corresponde a una métrica; en la Tabla 3 se observa el mejor valor *fitness* (el más cercano al mínimo global) dentro de las ejecuciones de cada algoritmo genético, la Tabla 4 muestra el peor valor *fitness* (el más lejano al mínimo global), en la Tabla 5 se presentan un promedio de valores *fitness* obtenidos por cada algoritmo y en la Tabla 6 la desviación estándar de estos valores. Cada columna representa el operador de reemplazo probado, mientras que cada fila corresponde a la función de minimización empleada. Los mejores resultados para cada función de minimización en particular se muestran en negrita.

Los resultados muestran que el operador de reemplazo de padres por competencia logró el mejor valores *fitness* (Tabla 3) en cinco de las diez funciones de minimización, seguido del reemplazo elitista logrando el mejor valor en cuatro funciones y el método Crowding en una función. Los algoritmos de reemplazo proporcional y reemplazo de padres no logran mejores resultados en ninguna de las funciones comparadas.

En los resultados de los peores valores *fitness* (Tabla 4) destaca el operador de reemplazo elitista al obtener el menor valor en seis funciones. El algoritmo de reemplazo de padres con competencia logra el menor valor entre los peores valores en tres funciones y, como en el caso anterior, el método Crowding en una función. Se identifica que los peores valores *fitness* para todas las funciones son obtenidos por el reemplazo proporcional.

El promedio de los valores obtenido en las ejecuciones, mostrado en la Tabla 5, muestran como mínimo valor de promedio el reemplazo elitista para cinco funciones,

**Tabla 5.** Promedio de los resultados de algoritmos genéticos con distintos métodos de reemplazo en funciones de minimización.

<b>Función</b>	<b>Reemplazo elitista</b>	<b>Reemplazo proporcional</b>	<b>Reemplazo de padres</b>	<b>Reemplazo de padres por competencia</b>	<b>Reemplazo Crowding</b>
Sphere	4.49	232.49	1.66E+05	<b>2.83</b>	58.63
Rastrigin	<b>93.01</b>	726.77	1.09E+05	141.21	188.14
Schwefel	4.35E+03	6.08E+03	9.27E+05	5.07E+03	<b>3.85E+03</b>
Rosenbrock	7.99E+04	3.06E+05	1.81E+11	<b>2.65E+04</b>	1.34E+06
Zakharov	1.16E+04	1.54E+04	1.01E+09	<b>7.01E+03</b>	1.06E+04
Discus	6.95E+04	4.57E+06	2.43E+09	1.50E+04	<b>8.03E+03</b>
HGBat	<b>1.58</b>	261.77	1.52E+05	12.37	55.20
Different Powers	<b>50.44</b>	725.93	6.55E+05	120.70	105.73
Happy Cat	<b>0.98</b>	20.79	5.34E+03	1.44	2.68
Griewank	<b>0.88</b>	1.03	1.54	0.95	0.92
Total Ganados	<b>5</b>	0	0	3	2

**Tabla 6.** Desviación estándar de los resultados de algoritmos genéticos con distintos métodos de reemplazo en funciones de minimización.

<b>Función</b>	<b>Reemplazo elitista</b>	<b>Reemplazo proporcional</b>	<b>Reemplazo de padres</b>	<b>Reemplazo de padres por competencia</b>	<b>Reemplazo Crowding</b>
Sphere	4.95	430.20	8.95E+04	<b>3.13</b>	65.46
Rastrigin	<b>20.26</b>	778.10	3.75E+04	37.53	65.01
Schwefel	2.45E+03	3.33E+03	8.71E+05	3.77E+03	<b>2.44E+03</b>
Rosenbrock	2.33E+05	7.66E+05	1.83E+11	<b>6.50E+04</b>	3.97E+06
Zakharov	<b>5.93E+03</b>	6.49E+03	1.36E+09	6.21E+03	1.07E+04
Discus	1.62E+05	1.13E+07	2.10E+09	3.81E+04	<b>714.77</b>
HGBat	<b>1.08</b>	236.79	9.96E+04	30.44	92.35
Different Powers	<b>38.62</b>	575.71	5.20E+05	99.75	186.62
Happy Cat	<b>0.40</b>	20.49	2.58E+03	0.53	1.83
Griewank	0.12	0.08	0.68	<b>0.04</b>	0.05
Total Ganados	<b>5</b>	0	0	3	2

seguido del reemplazo de padres por competencia, en tres funciones, y el reemplazo Crowding (en dos funciones).

En los valores de la desviación estándar de los valores *fitness* (Tabla 6) presentan menor dispersión en los resultados obtenidos por el operador de reemplazo elitista en seis funciones. Escenario similar expuesto anteriormente, también logra destacar por su

menor valor de desviación estándar el reemplazo de padres por competencia (en dos funciones) y el reemplazo Crodwing (en una función).

Bajo el esquema del teorema de no-free-lunch, de que no existe el algoritmo genético ideal, no se presenta el escenario donde un operador de reemplazo muestre los mejores resultados para todas las funciones de minimización empleadas. Al analizar los resultados en conjunto, los tres operadores de reemplazo que destacan, independiente de la métrica, son el reemplazo de padres por competencia, el reemplazo elitista y el reemplazo Crodwing. Se reafirma el operador de reemplazo de padres por competencia como una mejora al método de padres, esto al mostrar mejores resultados.

## **5. Conclusiones**

En esta investigación se presenta un informe detallado sobre los experimentos realizados en algoritmos genéticos con distintos operadores de reemplazo conocidos en el estado del arte.

Entre los diversos operadores de reemplazo revisados y probados destacan el reemplazo elitista, el reemplazo de padres por competencia y el reemplazo Crodwing al generar soluciones con los mejores valores mínimos. Los algoritmos genéticos desarrollados en este trabajo representan una solución viable para optimizar funciones de búsqueda de soluciones.

Para futuras investigaciones, sugerimos modificar los parámetros de los algoritmos genéticos y la dimensión y dominio en las funciones de minimización. En este sentido creemos que el uso de algoritmos genéticos puede ser de gran beneficio en la búsqueda de resultados en problemas de optimización.

## **Referencias**

1. Wang, Z., Qin, C., Wan, B.: A Comparative Study of Common Nature-Inspired Algorithms for Continuous Function Optimization. *Entropy*, 23(7), 874 (2021) doi: 10.3390/e23070874.
2. Buontempo, F.: *Genetic Algorithms and Machine Learning for Programmers: Create AI Models and Evolve Solutions*. Pragmatic Bookshelf (2019)
3. Holland, J.H.: Outline for a Logical Theory of Adaptive Systems. *Journal of the ACM*, 9(3), pp. 297–314 (1962) doi: 10.1145/321127.321128.
4. Rechenberg, I.: *Cybernetic Solution Path of an Experimental Problem*. Royal Aircraft Establishment Library Translation, 1122 (1965)
5. Rechenberg, I.: *Evolutionsstrategien. Simulationsmethoden in der Medizin und Biologie*, pp. 83–114 (1973)
6. Holland, J. H.: *Adaptation in Natural and Artificial Systems*. University of Michigan Press (1975)
7. López-González, A., Campaña, J.M., Martínez, E.H.: Multi Robot Distance Based Formation Using Parallel Genetic Algorithm. *Applied Soft Computing*, 86, 105929 (2020) doi: 10.1016/j.asoc.2019.105929.
8. NoParast, M., Hematian, M., Ashrafian, A.: Development of a Non-Dominated Sorting Genetic Algorithm for Implementing Circular Economy Strategies in the Concrete Industry. *Sustainable Production and Consumption*, 27, pp. 933–946 (2021) doi: 10.1016/j.spc.2021.02.009.

9. Elsoragaby, S., Yahya, A., Mahadi, M.R.: Applying Multi-Objective Genetic Algorithm (MOGA) to Optimize the Energy inputs and Greenhouse Gas Emissions (GHG) in Wetland Rice Production. *Energy Reports*, 6, pp. 2988–2998 (2020) doi: 10.1016/j.egy.2020.10.010.
10. Książek, W., Gandor, M. Pławiak, P.: Comparison of Various Approaches to Combine Logistic Regression with Genetic Algorithms in Survival Prediction of Hepatocellular Carcinoma. *Computers in Biology and Medicine*, 134, 104431 (2021) doi: 10.1016/j.compbiomed.2021.104431.
11. Parimi, P., Rout, R.R.: Genetic Algorithm Based Rumor Mitigation in Online Social Networks Through Counter-Rumors: A Multi-Objective Optimization. *Information Processing & Management*, 58(5), 102669 (2021) doi: 10.1016/j.ipm.2021.102669.
12. Hussain, K., Salleh, M.N.M., Cheng, S.: Common Benchmark Functions for Metaheuristic Evaluation: A Review. *JOIV: International Journal on Informatics Visualization*, 1(4-2), pp. 218–223 (2017) doi: 10.30630/joiv.1.4-2.65.
13. Wolpert, D.H., Macready, W.G.: No Free Lunch Theorems for Optimization. In: *IEEE Transactions on Evolutionary Computation*, 1(1), pp. 67–82 (1997) doi: 10.1109/4235.585893.
14. Fogel, D.B.: *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 1, John Wiley & Sons (2006)
15. Yao X.: *Evolutionary Computation. Evolutionary. Optimization. International Series in Operations Research & Management Science*, 48. Springer, Boston, MA (2003) doi: 10.1007/0-306-48041-7\_2.
16. Goldberg, D.E.: *Genetic Algorithms in Search. Optimization and Machine Learning*, Addison Wesley (1989)
17. Koza, J.R.: *Genetic Programming*. MIT Press (1992)
18. Joyce, T., Herrmann, J.M.: A Review of no Free Lunch Theorems, and their Implications for Metaheuristic Optimization. *Nature-Inspired Algorithms and Applied Optimization*, pp. 27–51 (2018) doi: 10.1007/978-3-319-67669-2\_2.
19. Lozano, M., Herrera, F., Cano, J.R.: Replacement Strategies to Preserve Useful Diversity in Steady-State Genetic Algorithms. *Information Sciences*, 178(23), pp. 4421–4433 (2008) doi: 10.1016/j.ins.2008.07.031.
20. De Jong, K.A.: *Analysis of the Behavior of a Class of Genetic Adaptive Systems*. University of Michigan (1975)
21. Li, X., Tang, K., Omidvar, M.N.: Benchmark Functions for the CEC 2013 Special Session and Competition on Large-Scale Optimization. Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University (2013)
22. C.T., Yue, K.V., Price, P.N.: Problem Definitions and Evaluation Criteria for the CEC 2020 Special Session and Competition on Single Objective Bound Constrained Numerical Optimization. Technical Report. Rep., Zhengzhou University and Nanyang Technological University (2019)
23. Awad, M.Z., Ali, P.N., Suganthan, J.J.: Problem Definitions and Evaluation Criteria for the CEC 2017 Special Session and Competition on Single Objective Real-Parameter Numerical Optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore (2016)
24. Johnson, J.M., Rahmat-Samii, V.: Genetic Algorithms in Engineering Electromagnetics. *IEEE Antennas and Propagation Magazine*, 39(4), pp. 7–21 (1997) doi: 10.1109/74.632992.
25. Back, T.: *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press (1996)

26. Herrera, F., Lozano, M., Sánchez, A.M.: A Taxonomy for the Crossover Operator for Real-Coded Genetic Algorithms: An Experimental Study. *International Journal of Intelligent Systems*, 18(3), pp. 309–338 (2003) doi: 10.1002/int.10091.
27. Singh, P., Dwivedi, P., Kant, V.: A Hybrid Method Based on Neural Network and Improved Environmental Adaptation Method Using Controlled Gaussian Mutation with Real Parameter for Short-Term Load Forecasting. *Energy*, 174(C), pp. 460–477 (2019) doi: 10.1016/j.energy.2019.02.141.
28. Srinivas, M., Patnaik, L.M.: Genetic Algorithms: A Survey. *Computer*, 27(6), pp. 17–26 (1994) doi: 10.1109/2.294849.
29. Alphabet Google Colaboratory: <https://colab.research.google.com/> (2021)
30. Zhang, Q.L.: Metrics for Meta-Heuristic Algorithm Evaluation. In: *IEEE Computer Society*. 2, pp. 1241–1242 (2003) doi: 10.1109/ICMLC.2003.1259677.



# Open Data Analysis to Identify Behavior in Water Quality for Mexico City

Emmanuel Alejandro Martínez Casillas

Instituto Politécnico Nacional,  
Mexico

ecasillas.emc@gmail.com

**Abstract.** This manuscript is based on the project “Prototype of monitoring network and forecasting model to know water quality in Mexico City”, with register number “TTT-2021/1-15”, the project is made up of three modules: 1) a data collection module, 2) a data processing module, 3) a data presentation module. Module 1) is made up of IoT sensors, which will collect data on water quality, considering the chemical, physical and biological characteristics of water, which are considered suitable for human use in accordance with NOM-012- SSA1-1993. Once the information is collected, it will be stored in 2) a data repository with a data manager in the cloud, the communication mechanisms between the modules will be web services. Module 2) will oversee calculating the water quality forecast for a specific date or period, for this, machine learning algorithms will be used for getting the predictions and for detecting the anomalies (Outlier values), the detection of outliers in the geographic analysis, will represent the origin of a problem. Finally, module 3) will display the graphs and maps using the dashboard technique (dashboard in a responsive web interface).

**Keywords:** Exploratory, water quality, clustering.

## 1 Introduction

Currently many refer to the current era, as the data era [1], this since people currently consider information as the “new gold” [6], so it can be understood, the great importance of data in our times, as well as what can be done and achieved with them is mentioned, of course, for that we must first go through the real challenge, which is to be able to analyze and extract the information from said data.

In Mexico, the culture about data has been growing, however, it is still very little, since science in our country there is an endless amount of data, which has not yet been treated and its potential has not been discovered, since Mexico is in the 76th place within a list of 140 countries, in the index of adoption of information technologies favorable to economic development.

For example, one of the sectors of which there is a lot of information data still unprocessed on the water system, day by day many people within the CDMX consume water with a quality, which in some cases is harmful, if it is consumed or used in a certain amount [10]. Then people question the National Water Commission

(CONAGUA), the reason why the quality of the water they receive has been deteriorating over the years. In many cases, the detection and correction of this problem is very slow, and it is until the water quality has decreased by half that a true notion that there is a problem is obtained, so we can conclude that It is necessary to have a system that predicts or is capable of recognizing according to the successive behavior of water quality, what problem could be the cause of said decrease in water quality and combat it before it spreads and reaches a critical point at which the water quality is lower than that recommended for human consumption and use.

The problem of calculating or knowing the quality of water in CDMX began a long time ago, serious consequences have already occurred, as happened in the 90's, where a large part of the population suffered from a disease called cholera, which is a bacterial disease, spread through contaminated water, causing severe diarrhea and dehydration in people, which can be fatal within hours.

Based on articles and reliable sources such as scientific studies by universities and government reports [9], the idea arises of being able to generate a system that helps to solve this problem, so important, since we are talking about water, which is an invaluable resource for the daily life of all human beings and whose conservation and care, also within the factors that define what quality of life is, is the basic service of water supply, in times like the pandemic, having good quality water becomes a basic requirement to be able to face diseases.

The principle of the first law of geography is understood, which says that the objects closest to each other in space have more relationship [7], and this is important when we talk about a network such as the CDMX, the which connects all municipalities, but does not obtain the same water quality results in each municipality, This concept is further deepened, but it is also seen that within a network there is also a relationship through time [8], this is very helpful for understanding more, how this problem has advanced within the network through space and time, so that thanks to concepts we can better understand how the problem of water quality is a problem that grows over the years and spreads across the network.

## **1.1 Problem Statement**

At present only 33.2 percent of the water quality monitoring sites operated by the National Water Commission (CONAGUA) meet the acceptable limits of water quality, this means that 66.8 percent of the sites have low water quality. This represents a severe problem, since there are people who can be intoxicated without being informed of what they are really consuming, in the water of the CDMX, because the information is not easy to obtain and the statistics found in the government reports, they are statewide, so, in smaller regions, the information is not accurate. Currently Mexico has a national water quality measurement network, but this only monitors the main bodies of water in the country, that is, it monitors very large ranges, so it is more difficult to obtain accurate information from smaller communities, such as the case of the CDMX mayors, where it is not known whether the affected area is the entire network or only a part of it.

Within the network that CONAGUA currently manages, in CDMX there are few monitoring sites and most of these do not comply with the acceptable limits of water quality, this shows that in CDMX the quality is poor, but the current network does not consider a model that forecasts for a specific period. With these data on a dashboard, you can support decision-making and know or infer causes of the problem or be able to monitor the data, to know how the water quality has improved or worsened over time and in certain areas.

This problem is a reality in Mexico, where cases of diarrhea have been increasing, this because the presence of contaminants in the water has increased, between 7 and 12 percent, this is serious and means that the risk of finding contaminants in water is comparable to the risk in European countries and some parts of China.

## **2 State of the Art**

In this space there will be a comparison of scientific articles which have the same objective as the proposed project, but they use different techniques, also at the bottom of this section is the table 1, with a deeper comparison.

### **2.1 Spatial Variation Impact of Landscape Patterns and Land Use on Water Quality Across an Urbanized Watershed in Bentong, Malaysia**

This research aimed to quantify and illustrate the effects of land use and landscape configuration on water quality in Bentong River, Malaysia. The study sampled 22 sites during the normal and wet season in 2018. FRAGSTATS was used to analyze the spatial change of landscape metrics. The results showed that water quality was closely associated with landscape configurations and land cover proportions. It also indicated that the susceptibility of water to degradation increased with a great interrelation of different land uses [2].

### **2.2 Quantifying the Contributions of Structural Factors on Runoff Water Quality from Green Roofs and Optimizing Assembled Combinations Using Taguchi Method**

In this study, runoff plots of extensive green roofs with Taguchi designed structural factors and levels were constructed and simulated rainfall experiments were conducted. Influences of structural factors on outflow water quality of green roofs were statistically assessed and quantified. Runoff water quality of green roofs with assembled combinations at specific levels were optimized and predicted by using the Taguchi method [3].

**Table 1.** Comparison of projects.

Projects	Sensor type	Type of data analysis	Presentation of the results	Applied on
P1	Different	Different	Different	Malaysia
P2	Different	Equal	Different	China
P3	Different	Equal	Equal	China
P4	Different	Equal	Equal	

### 2.3 A Holistic Assessment of Water Quality Condition and Spatiotemporal Patterns in Impounded Lakes Along the Eastern Route of China's South-to-North Water Diversion Project

Water quality is one of the key determinants for assessing effectiveness and success of water diversions, but rarely studied at a spatial scale that crosses large river basins. Multiple statistical methods and the water quality index (WQI) were used to assess overall condition and detect spatiotemporal patterns of water quality in a series of impounded lakes along the Eastern Route of China's South-to-North Water Diversion Project [4].

### 2.4 Water Quality Related to Conservation Reserve Program (CRP) and Cropland Areas: Evidence from Multi-Temporal Remote Sensing

Therefore, aiming to quantify the relationship between CRP enrollment, cropland area, and the downstream water quality, we propose an approach that combines archived survey data, water quality monitoring data (total nitrogen content, TN), and remote sensing observations. By constructing the long-term datasets (1999–2014 annually) in Google Earth Engine and conducting multiple linear regression, they explained 79% variation in TN by the area of total CRP enrollment (CRP\_all), area of corn and soybeans croplands, and discharge [5].

## 3 Methodology

A prototype was built to know the quality of the water that is delivered to a house every day, as well as the quality that the water will have in the future, through a model that accurately predicts the quality of the water, in a way that people can more quickly and easily know what they consume at that time and in the near future.

The first step for the development of this system will consist of choosing the sensors that will be used for data collection, these sensors must have at least the possibility of measuring concentrations of total dissolved solids (TDS), pH level, and turbidity, once the sensors have been chosen, they will be placed, one will be placed in each city hall of the CDMX, giving a total of 16 sensors, said sensors will be found within the government sites of each city hall, the installation within these sites will be thanks to an agreement with the CDMX secretariat of science and technology, they will be installed within these sites for two reasons; The first is the security of the

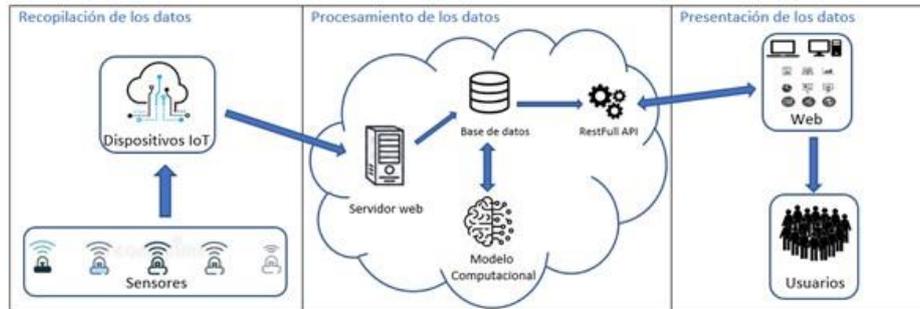


Fig. 1. Data sequence.

sensors and the second is to take advantage of the internet within said sites to send the data captured by each sensor to the database, through IoT devices, said the database will store the information for then be able to process it through a model, which will be helpful for predicting [15], according to the data of each sensor, how the water quality may improve or worsen in the future, in each municipality, then through geographic information systems, the results will be shown, occupying an Outlier type analysis, this analysis will be occupied since as we are talking about a network it is important to know what is the relationship of the water quality in one municipality with respect to the water quality in other municipalities, in order to be able to use the system, the user needs a connection that allows access to the internet, the system will not ask the user for permissions, but will first show a map of the sensor network s and according to the element on which the user clicks, the respective graphs corresponding to that municipality will appear, showing information on both the current quality of water that is being consumed, as well as the quality that is predicted. It is expected that with this system users will have more information about the water service they are receiving, but also that with this it will be easier and faster to detect where the quality of water is declining and could put the quality of life at risk. the users who consume it, figure 1 illustrate the connection between the modules that make up the system.

### 3.1 Data and Water Quality

Once each one of the properties has been analyzed, take into account the area to be measured, it is time to characterize all this to the CDMX water system, which is the area that the project will cover, for this we must take into account that the CDMX has an area of land of 1485 square km, as well as the number of 8.92 million inhabitants.

The CDMX has a drinking water coverage of 98%: It is important to highlight that 7.5% of the country's population resides in CDMX, as well as that 17% of the country's economic activity passes through here. Considering the previous data and in accordance with the SACMEX.

As reported by SACMEX, every year in CDMX the quality of the water suffers a deterioration since each year, the water level of the aquifer decreases by one meter and as the water is extracted from a greater depth the water undergoes certain

changes, in addition to the fact that this also encourages the city to sink 30 cm each year.

In 2019, several actions were undertaken to improve the service of the CDMX water system, one of these actions has been the sectorization of the secondary network in all municipalities to improve the service to users with pressure control schemes, hydraulic balances, and increased efficiency, also considering the CDMX wells that are active, restored or that continue to have problems.

The concept of water quality focuses on the use for a city or for the use of people daily in their homes, for this CONAGUA is based on the following parameters, to say that the quality of the water is, the qualities of the water that CONAGUA has defined are:

- Green: for those that correctly meet all the parameters for water quality.
- Yellow: for those that violate one or more of the following parameters E\_COLI, CF, SST and OD%.
- Red: for those that fail to comply with one or more of the following parameters BOD5, COD, TOX and ENTEROC.

The water quality limits, so that it can be considered as potable for human use and consumption, are established in the current Official Mexican Standard NOM-127-SSA1, said standard defines limits such as those exemplified in table 2.

The data was obtained directly from reports by CONAGUA and SACMEX of wells and aquifers that feed the CDMX, this data is public online and include data from 2011 to 2020, it contains information on parameters such as E\_COLI, CF, SST, OD%, pH, turbidity, TDS, among others.

This project will focus on parameters such as turbidity, pH, as well as a total of all the solids dissolved in the water, following the CONAGUA color code, if only one parameter is met, the point will be displayed in red, in case of meeting two parameters of yellow and in case of three of green color.

In figure 2, it can be seen the data taken from the source of the government base of SACMEX, which contains data by year of each measure for water quality considered in CDMX, these records are divided by municipalities, the measurements of the sensors and the year; latitude and longitude data were added for space exploration.

From this table, queries were generated, where it was divided only by mayor's office to facilitate analysis of the data as shown in figure 3, the possibility that there was a pattern between the growth or decrease of some measures throughout the period was analyzed as can be seen, in the municipalities that are closest to each other, such as the Gustavo A. Madero mayor's office and the Azcapotzalco mayor's office, despite their proximity, there is no similar pattern of variation of the measures.

Queries were made showing the processes of inserting information to the table of concentrates and the calculation of the error, as can be seen, the insertion to the table of concentrates will occur each time information from all monitoring points arrives, while the calculation of the error will be given once the date of the insertion of data from the monitoring points coincides with the date of the prediction made by means of the forecast model, where the database manager itself by means of a trigger that It

**Table 2.** Parameters and their limit allowable at Official Mexican Standard NOM-127-SSA1.

Parameters	Limit Allowable
Total coliform organisms	2 MPN / 100 ml
Fecal coliform organisms	Not detectable MPN / 100 m
Color	20 true color units on the platinum-cobalt scale.
Odor and Taste	Pleasant (those that are tolerant for most of the consumers)
Turbidity	5 nephelometric turbidity units (NTU) or its equivalent in another method.
Aluminum	0.20
Free residual chloride	0.2-1.5
Copper	2.00
Total hardness as (CaCO3)	500
pH	6.5-8.5
Pesticides in microorganisms	0.03
Global alpha radioactivity	0.1
Zinc	5.00
Sodium	200.00
Global beta radioactivity	1.0
Mercury	0.001

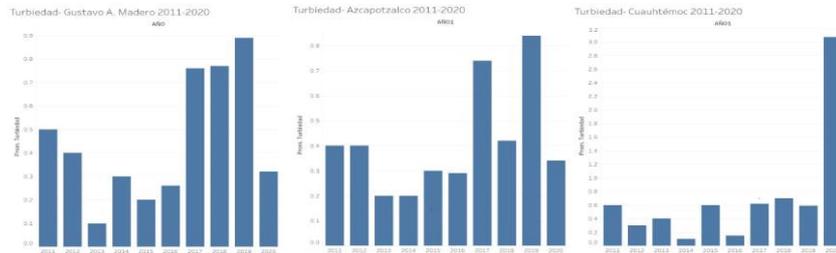
INDICE	ALCALDIA	NOMBRE_ALCALDIA	MUESTRAS	PROMEDIO_CLORO	pH	Turbiedad	Dureza_Total	Cloruro	Hierro	Manganeso	LATITUD	LONGITUD	AÑO
1	1	Ávaro Obregón	438	1.94	3.64	0.35	72.97	27.89	0.05	0.057	19.35867	-99.20329	2016
2	2	Azcapotzalco	589	0.88	2.53	0.29	47.33	15.03	0.038	0.035	19.48596	-99.18594	2016
3	3	Benito Juárez	407	0.83	2.64	0.3	21.54	4.58	0.034	0.037	19.3984	-99.15766	2016
4	4	Coyoacán	439	0.98	1.99	0.31	41.37	10.82	0.04	0.055	19.3467	-99.16174	2016
5	5	Cuajimalpa	171	0.9	2.46	0.24	44.94	29.4	0.036	0.034	19.359306	-99.29388055555556	2016
6	6	Cuauhtémoc	213	1.02	2.08	0.15	34.31	11.06	0.027	0.027	19.4450667	-99.14611666666667	2016
7	7	Gustavo A. Madero	737	0.94	2.32	0.26	48.85	19.13	0.032	0.039	19.49392	-99.11075	2016
8	8	Itacaicolco	252	1.68	2.98	0.3	43.74	20.16	0.068	0.038	19.3952778	-99.09777777777778	2016
9	9	Itzapalapa	1606	1.17	2.76	0.4	56.1	17.53	0.039	0.038	19.35529	-99.06224	2016
10	10	Magdalena Contreras	168	0.85	3.17	0.23	44.44	16.78	0.045	0.048	19.33212	-99.21118	2016
11	11	Miguel Hidalgo	311	0.88	3.53	0.3	47.33	14.99	0.046	0.043	19.43411	-99.20024	2016
12	12	Mixco Alta	117	0.76	2.96	0.27	87.99	39.42	0.03	0.053	19.19251	-99.02317	2016
13	13	Tlalhuac	414	0.89	2.35	0.29	32.7	16.03	0.033	0.029	19.28689	-99.00507	2016

**Fig 2.** Basic Exploration of the Concentrate Table.

will calculate the difference between the result of the prediction and that measured by the IoT device [14].

### 3.2 Data Exploration

In addition, a simulation of the Outlier analysis was carried out, it should be remembered that as already mentioned earlier in this document, the Outlier analysis consists of not only considering the information of the place itself but also of the neighbors and although in the statistics these points are neglected in geospatial analysis is highly relevant [13], since they can indicate possible sources of failures or



**Fig. 3.** Graphs on the distribution of turbidity in municipalities (GAM, Azcapotzalco and Cuauhtémoc) from 2011 to 2020.

```

DECLARE @YEAR INT
DECLARE @CONT INT
SET @YEAR=2011
SET @CONT = 0
WHILE (@YEAR<=2021)
BEGIN
PRINT @YEAR
IF (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Gustavo A. Madero' ) = @YEAR) (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Azcapotzalco' ) = @YEAR)
SET @CONT=@CONT+1
IF (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Gustavo A. Madero' ) = @YEAR) (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Cuauhtémoc' ) = @YEAR)
SET @CONT=@CONT+1
IF (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Gustavo A. Madero' ) = @YEAR) (SELECT ph FROM CLOUD WHERE NOMBRE_ALCALDEA='Benito Juárez' ) = @YEAR)
SET @CONT=@CONT+1
IF @CONT=3
PRINT 'OUTLIER RED'
IF @CONT=2
PRINT 'OUTLIER RED'
IF @CONT=1
PRINT 'OUTLIER BLUE'
IF @CONT=0
PRINT 'OUTLIER BLUE'
PRINT @CONT
PRINT @YEAR
SET @YEAR=@YEAR+1
SET @CONT=0
END
    
```

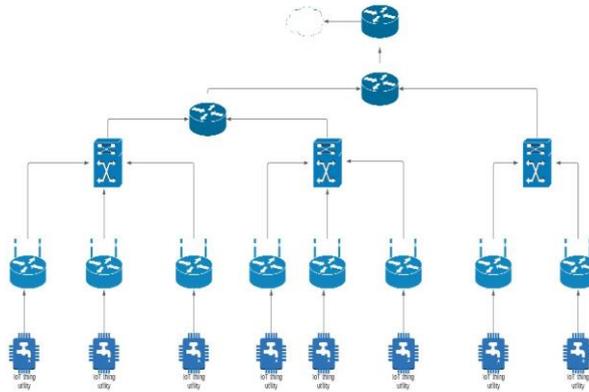
**Fig. 4.** Example and result of an Outlier analysis simulation.

also points where the sensors are not sending the correct information. For this simulation, several exercises were carried out and an example and its result are shown below in figure 4.

## 4 Tests and Results

For the test scenario, the homes of people close to the developer were taken as monitoring points. If the minimum concentrations of chlorine in the detections made at the monitoring points want to be guaranteed, consider that the chlorine concentration decays once the water leaves the purification plant, in places far from the plant, residual chlorine may be absent, thus allowing the increase in bacterial levels. During the tests, the parameters of pH level, turbidity and total dissolved solids (TDS) will be measured.

As can be seen in figure 5, the apparatus in the bottom line was used to represent the modems of the houses where the sensors are located, as the communication between the sensors and the cloud system or even between them is not observed. direct, it requires several routers that communicate to the entire network, the main node was also added above since being the coordinator it was easier to exemplify it like this, although this does not mean that it is in a network space superior to that of the other nodes.



**Fig. 5.** Exemplification of the CDMX network to which the sensors are connected.

In the second part of the tests, the section of the model and its predictions will be seen, the clustering technique will be used, it is a technique whose purpose is to be able to group the objects by similarity, in groups or sets so that the members of the same group have similar characteristics, the geospatial data will be the variable on which the grouping criterion will be based.

To apply an Outlier type analysis, it is important to define a numerical characteristic, which will help to make the correct division of characterization by zone, for example the level of pH, turbidity, and the number of total dissolved solids can be taken as a variable.

As the Outlier analysis can be seen to be more complete, finally, figures 6 and 7 are shown, some Dashboards generated with the Tableau tool, which serve to further explore the measures.

The web application where the results will be projected is a responsive web dashboard, giving the user options to choose the type of map on which he wants to see the sensor network, what parameter he wants to observe on his screen to measure water quality, view the graphs of the data of each sensor, the options offered by the web application are:

- Location: highlighted points that are the locations of the sensors. Each of these points will have two suboptions, which will be if you want to see the data from the sensors or from the forecast model.
- Charts: the following options bars, pie, histogram, and scatter.
- Report: the application will generate a PDF report of the information, whether it is general or a specific point that the user has selected, adding both the data from the sensors and the data from the predictions made by the model.

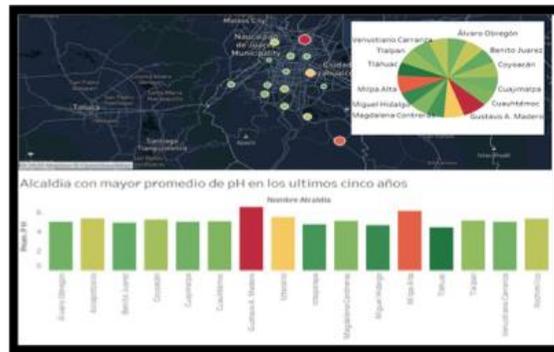


Fig. 6. Dashboard of the pH measurements in the last 5 years in each of the municipalities of the CDMX.

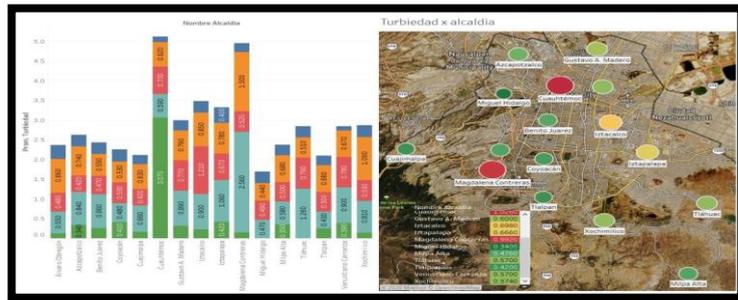


Fig. 7. Dashboard of the turbidity measurements in the last 5 years in each of the municipalities of the CDMX.



Fig. 8. Visualization of the web application on a computer.

## 5 Conclusions and Future Work

This project has been interesting from the first stage of collecting information, consulting the reports of both SACMEX and CONAGUA, in this way it was possible

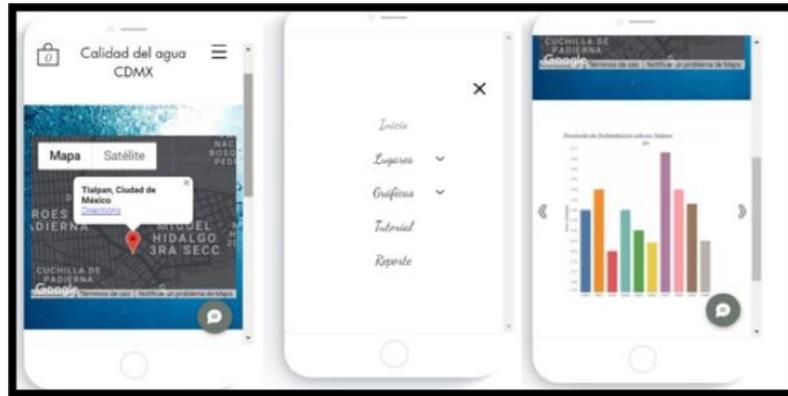


Fig. 9. Visualization of the web application on a cell phone.

to observe that none of the municipalities at the time delegations took continuous measures that allowed to have a good management on the control of water quality, there are also data that show a series of consequences due to this carelessness of control, going from being close to the permissible limit to being almost double, thus exceeding what is adequate for consumption or human use, but without informing the population of these major problems.

On the other hand, it could be observed thanks to the geospatial analysis that over the years no concrete and rapid solution has been given to water quality problems, which has led to a spread of problems around large parts of the network of Mexico City, However, this error has been reduced to only one unit in the pH parameter and for the other parameters the error is even smaller, so it will soon be a model with a fairly high percentage of reliability.

For future work we plan to reach a final agreement with the corresponding authorities to be able to use the system on their behalf and thus serve as a support to improve the quality of life of the inhabitants of Mexico City; in future work the idea is to expand the study area to other large cities in Mexico.

**Acknowledgments.** Special thanks to M. in C. Roberto Eswart Zagal Flores and Dr. Miguel Felix Mata Rivera for their counseling in the development of the project “Prototype of monitoring network and forecasting model to know water quality in Mexico City”.

## References

1. Caldentey, F.: The Age of Data: What Benefits and Risks are there in Big Data. the ‘Big Brother’ of the 21st Century| UNIR (2020)
2. Zakariya-Nafi'Shehab, F., Nor Rohaizah-Jamil, S., Ahmad Zaharin-Aris, T.: Spatial Variation Impact of Landscape Patterns and Land Use on Water Quality Across an

- Urbanized Watershed in Bentong, Malaysia. *Ecological Indicators*, 122, (2021) doi: 10.1016/j.ecolind.2020.107254.
3. WenLiu, F., Bernard A., Engel, S., Weiping Chen, T.: Quantifying the Contributions of Structural Factors on Runoff Water Quality from Green Roofs and Optimizing Assembled Combinations Using Taguchi Method. In: *Journal of Hydrology*, 593 (2021)
  4. Xiao Qu, F., Yushun Chen, S., Han Liu, T.: A Holistic Assessment of Water Quality Condition and Spatiotemporal Patterns in Impounded Lakes Along the Eastern Route of China's South-to-North Water Diversion Project. *Water Research*, 185 (2020) doi: 10.1016/j.watres.2020.116275.
  5. Dameng Yin, F., Le Wang, S., Zhenduo Zhu, T.: Water Quality Related to Conservation Reserve Program (CRP) and Cropland Areas: Evidence from Multi-Temporal Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, 96, (2021) doi: 10.1016/j.jag.2020.102272.
  6. Anonymous, F.: Data is the New Gold. Deloitte (2020)
  7. Goodchild, F.: First Law of Geography. *International Encyclopedia of Human Geography*. Elsevier, pp. 179–182 (2009) doi: 10.1016/B978-008044910-4.00438-7.
  8. Costa Rica, F.: Towards a Network Geography: A New Paradigm of Space Analysis Alternative to the Regional Approach. Redalyc.org (2020)
  9. Agua, F.: Water Quality in Mexico. Government of Mexico (2020)
  10. Villanueva, F.: Contaminated Water in Mexico Increased Hospitalization Due to Diarrhea by 12%. *La Jornada* (2020)
  11. CONAGUA, F.: Water Quality Report. Government of Mexico (2020)
  12. SACMEX, F.: Water Quality Reports. Government of Mexico City (2020)
  13. Blasco Fernández, I: Outlier Detection Methodologies in Spatial, Temporal and Spatio-Temporal Data. *Revista Cartográfica*, 96, pp. 139–157 (2018)
  14. Rose, K., Eldridge, S., Chapin, F.: The Internet of Things a Brief Review. *Internet Society*, (2015)
  15. Sandoval, L.J.: Automatic Learning Algorithms for Analysis and Data Prediction. *ITCA FEPADE*, 11 (2018)

# Estado del arte en el reconocimiento de emociones en textos de redes sociales

Cristian C. Segura, Hiram Calvo

Instituto Politécnico Nacional,  
Laboratorio de Ciencias Cognitivas Computacionales,  
Centro de Investigación en Computación,  
Mexico

{cseguram2021, hcalvo}@cic.ipn.mx

**Resumen** El reconocimiento de emociones es una tarea de gran importancia en las tareas de procesamiento de lenguaje natural con diferentes aplicaciones en varias áreas del conocimiento. La detección implícita de emociones en los textos principalmente en aquellas interacciones en las redes sociales es posiblemente uno de los problemas más desafiante a resolver debido a que las emociones a transmitir y expresar se encuentran ocultas en el texto y encontrarlas requiere comprender el contexto de los diferentes tipos de texto. En el presente documento se revisará el estado del arte de los métodos actuales de reconocimiento de emociones con diferentes fuentes y en diferentes contextos identificando los diferentes enfoques utilizados en el proceso para así proceder con la comparación entre la literatura identificada. Finalmente se presenta una propuesta basada en el procesamiento de textos distorsionados, debido a que éste no ha sido abordado a profundidad en los trabajos presentados anteriormente.

**Palabras clave:** Reconocimiento de emociones, procesamiento de lenguaje natural, detección implícita, redes sociales, análisis de texto, contexto.

## A Survey of State of the Art in the Recognition of Emotions in Texts in Social Networks

**Abstract.** Emotion detection is a task of great importance in natural language processing tasks with different applications in various topics of knowledge. The implicit detection of emotions in the texts, mainly in those interactions in social networks, is possibly one of the most challenging problem to solve because the emotions to be transmitted and expressed are hidden in the text and finding them requires understanding the context of the different types of text . This document will review the state of the art of current emotion detection methods using different sources and in different contexts, identifying the different approaches used in the process and comparing them with the identified literature.

Finally, a proposal based on the text processing of distorted texts is presented, because this has not been researched in depth in the works previously presented.

**Keywords:** Emotion recognition, natural language processing, implicit detection, social media, text analysis, context.

## 1. Introducción

El uso de las redes sociales ha crecido de una manera imponente en los últimos años Leong2019. Las redes sociales son definidas como los sitios y las herramientas en línea que facilitan la interacción entre los diferentes usuarios permitiéndoles la oportunidad de intercambiar información, opiniones, intereses e incluso emociones. Dentro de los principales usos de las redes sociales se encuentran el entretenimiento, comunicación, intercambio de opiniones, búsqueda de información entre otros. Con el auge de estas plataformas sociales, una gran cantidad de horas diarias son invertidas de manera virtual en herramientas como Instagram, Facebook, Twitter, YouTube, WhatsApp entre muchas otras. Las generaciones actuales viven en un mundo donde lo primero que se ve en la mañana y lo último en la noche es un teléfono celular, por lo que de acuerdo con algunos autores sugieren que las redes sociales han cambiado la forma de la interacción entre individuos y el comportamiento de los usuarios de manera individual y colectiva en todo el mundo Dhir2018,Tateno2019.

El mayor uso de plataformas de redes sociales ayuda a incrementar el número de “conexiones” pueden generar diferentes tipos de emociones afectivas pudiendo llegar al punto de generar cierto tipo de confusión emocional. Existen dos tipos de lazos relacionales en las redes sociales, los denominados lazos fuertes, que son establecidos con personas allegadas de una manera muy personal e íntima, como parejas románticas, mejores amigos, familiares cercanos entre otros. [2] se refiere al segundo tipo de lazos denominados débiles que se encuentran en los círculos sociales con allegados, compañeros o conocidos. Según de la fortaleza de las relaciones entre los individuos, el uso de las redes sociales puede ser saludable o disfuncional. En ambos casos, estas relaciones afectan el bienestar de la persona brindando felicidad o placer al usuario en el mejor de los escenarios.

Existen diferentes investigaciones realizadas que indican que cuando se experimentan una fuerte cantidad de sentimientos positivos o negativos, hay una gran motivación al querer compartirlos con los demás Rime2009. De acuerdo con la teoría del intercambio social de emociones de Rimé, compartir las emociones ayuda en dos objetivos: la satisfacción de las necesidades socioemocionales que consiste en obtener apoyo de terceros para revalidar los sentimientos, normalizar las experiencias y sentir estados emocionales positivos. Y como segundo objetivo, se encuentran las necesidades cognitivas, que implica recibir consejos de los demás para ayudar a dar sentido a las experiencias vividas.

Las redes sociales les brindan a las personas oportunidades nunca antes vistas para lograr estos dos objetivos Brady2020,Rime2020 permitiendo una

conexión inmediata y en tiempo real con las personas cercanas, pero también con aquellas que se encuentran a una distancia geográfica considerable. Este tipo de interacciones generan al usuario un sistema de retroalimentación que satisface las necesidades socioemocionales y cognitivas incrementando los niveles de apoyo social percibido y recibido. En algunos casos, el apoyo social que se consigue en una red social es percibido de manera ampliamente superior en el mundo virtual que en el mundo real. Algunos estudios han evidenciado que ciertas personas que sufren de depresión reciben menos apoyo de otras personas que no sufren de esta condición en el mundo fuera de línea y han recibido más apoyo de las personas mediante la interacción en redes sociales Park2016, Starr2008.

De la misma manera que las redes sociales brindan nuevas oportunidades para buscar y brindar apoyo, también permite expresar los sentimientos de formas que pueden lastimar a otras personas. Las redes sociales son plataformas que facilitan en gran medida el ciberacoso y el matoneo, comportamientos antisociales que impactan de manera negativa al bienestar de las demás personas. Aunque las diferencias juegan un papel importante en quienes se involucran en este tipo de actitudes, hay características de las redes sociales que promueven este comportamiento; por ejemplo, al no contar con un contacto cara a cara, es difícil determinar mediante un medio electrónico la reacción física de una persona ante un comentario o acción y que permita limitar un comportamiento que pueda llegar a ser agresivo Lapidot-Lefler2012, Lieberman2020a, Suler2004. De igual manera la facilidad y accesibilidad a los dispositivos para acceder a una red social, permiten que sea más rápido compartir las emociones que se sienten cuando están en un punto máximo y son comunicadas sin detenerse a pensar en las posibles consecuencias de los comentarios precipitados.

Las redes sociales también desempeñan un papel vital en la difusión de sentimientos de indignación moral, lo cual contribuye a la deshumanización de otras personas y reducir acciones que se realicen de manera colectiva y aumentar las brechas existentes entre diversos puntos de opinión o características sociales. Se evidenció en estudios que ciertas personas son más propensas a realizar comentarios de manera intensa sobre indignación moral en línea que estando fuera de ella Turel2018.

Estas redes en general demuestran ser una plataforma que permite satisfacer el deseo de compartir emociones con las demás personas algunas veces con resultados positivos y de bienestar que generan nuevas oportunidades para que las personas puedan tener un apoyo emocional y superar ciertas circunstancias en las que se puedan encontrar; pero también pueden ser promotoras de acciones negativas y ser una plataforma masificadora de ciberacoso, *trolleo* y difusión masiva de odio.

A continuación presentaremos los trabajos actuales que se han realizado, partiendo de modelos para representar emociones (sección 2), recursos utilizados en el reconocimiento de emociones (sección 3), posteriormente presentamos las técnicas utilizadas en el reconocimiento de emociones en textos (sección 4). Para concluir presentamos un esbozo de lo que hemos encontrado como uno de los múltiples casos que no han sido atacados dentro de este problema en la sección 6.

## 2. Modelos para representar emociones

Los modelos de emociones son fundamentales para los sistemas detectores de las mismas ya que definen cómo pueden ser representadas. Estos modelos afirman que las emociones pueden presentarse en varios estados por lo cual nace la necesidad de distinguir cada uno de ellos; por ello cuando se realiza cualquier tipo de actividad concerniente a la detección de emociones es importante realizar la definición del modelo a utilizar. Existen diferentes modelos que representan emociones, pero se hará énfasis en los modelos de emociones discretos y dimensionales.

### 2.1. Modelos de emociones discretos (*Discrete emotion Models - DEMs*)

Clasifican las emociones en diferentes categorías, donde se toman los siguientes tres modelos a consideración:

Modelo de Paul Ekman Kowalska2017. Realiza diferenciación de emociones con un conjunto de seis categorías básicas. En su modelo afirma que existen seis emociones originadas desde diferentes sistemas neuronales como el resultado de cómo se percibe una situación, por ello son independientes. Estas emociones son ira, felicidad, sorpresa, tristeza, temor y desagrado. Adicionalmente, estas emociones en conjunto pueden generar otras más complejas como la vergüenza, orgullo, lujuria, culpa, avaricia, deseo, celos entre otras.

Modelo de Plutchik y Kellerman Kellerman1989. Afirma también que existen emociones primarias las cuales ocurren en pares opuestos y se encargan de producir emociones más complejas con sus combinaciones. Nombra un total de ocho emociones primarias en pares fundamentales las cuales son alegría frente a tristeza, confianza frente a disgusto, ira frente a miedo y sorpresa frente a anticipación. De acuerdo con Plutchik, para cada una de las emociones existen varios grados de intensidad que ocurren de acuerdo como los eventos son contruidos por un individuo.

Modelo de Orthony, Clore y Collins (*OCC Model*) Colby1989. Afirman que las emociones son el resultado de cómo un individuo percibe los eventos y esas emociones varían de acuerdo con el grado de intensidad en el que son percibidas. Realizan la diferenciación de 22 emociones definiendo su modelo de la siguiente manera: ira, felicidad, sorpresa, tristeza, temor, desagrado, envidia, desagrado, agrado, dolor, alivio, envidia, reproche, autorreproche, aprecio, vergüenza, lástima, admiración, decepción, esperanza, miedos reafirmados, agradecimiento.

### 2.2. Modelos de emociones dimensionales (*Dimensional emotion models DiEMs*)

El modelo dimensional supone que las emociones no son independientes unas de las otras y que existe una relación entre ellas, por lo que surge la necesidad de colocarlas de manera conjunta en un espacio dimensional que represente qué tan relacionadas están las emociones y realizando una representación de los dos



Fig. 1. Modelo de emociones de Russell. adaptado de Russell1980.

estados conductuales básicos fundamentales: bien y mal. Los siguientes son los dos modelos básicos dimensionales que se utilizarán:

El modelo de Russell (1980) consiste en una estructura de dos dimensiones denominada circunferencia de afecto. El modelo define dos valores de medida, *valence* que diferencia las emociones en un rango de positiva a negativa, y *arousal* las diferencia en un contexto de emoción o apatía. Esta relación se muestra en la figura 1.

El modelo de Plutchik presenta las emociones en un disco bidimensional con *valence* en el eje vertical y *arousal* en el horizontal. La figura muestra las emociones en capas, donde las más internas corresponden a derivaciones de las ocho emociones primarias, seguidas precisamente por estas emociones primarias y en la parte exterior combinaciones de las emociones primarias. Según la posición en la figura, se determina qué tan relacionadas se encuentran las emociones. En la figura 2 se presenta el esquema de Plutchik.

Según el problema a resolver se realiza la selección del modelo emocional a trabajar. Los modelos de emoción discretos usualmente son los seleccionados para los problemas de clasificación de emociones debido a la simplicidad, pero carecen de una gama más amplia de clase de emociones, la intensidad y grado de ocurrencia en contraste con los modelos dimensionales. Pero los modelos dimensionales son recomendables cuando se desea categorizar una mayor cantidad de emociones y se quiere diferenciar entre emociones que son muy similares entre sí.

### 3. Recursos utilizados en el reconocimiento de emociones

Comúnmente en el reconocimiento de emociones se utilizan dos tipos de conjunto de datos, los *corpora* que corresponden al conjunto de textos completos clasificados según la emoción a la que se encuentra asociado cada uno de ellos,

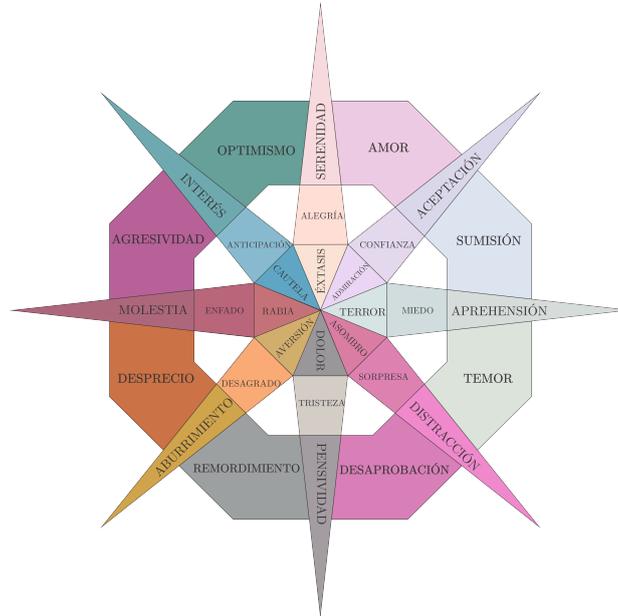


Fig. 2. Modelo de emociones de Plutchik. Adaptado de Kellerman1989.

y los *lexicons* que corresponden a un conjunto de listas ordenadas con palabras asociadas cada una de ellas a una emoción.

### 3.1. Corpora

[1]<sup>1</sup> presentan un total de 185 historias infantiles donde la emoción se clasifica a nivel de las oraciones que se encuentran en cada una de las historias y se asigna una etiqueta a cada una de las frases de acuerdo con el cuadro 1.

Para la clasificación se realiza la asignación en pares de las mismas historias y con el objetivo de evitar sesgos cada uno de los evaluadores trabaja de manera separada y realiza la asignación de una sigla de manera independiente. Cuando existe un desacuerdo en la etiqueta asignada, el evaluador principal elige una de las etiquetas disputantes.

El corpus de Aman2007 cuenta con un conjunto de publicaciones en diferentes blogs que son obtenidas mediante el uso de palabras semilla para asignar a cada emoción. Por ejemplo, palabras como miedo, susto, o pánico son clasificados en la categoría de temor. La clasificación se realiza a nivel de oración en una de las ocho categorías clasificadoras (felicidad, tristeza, enojo, desagrado, sorpresa, temor, emociones combinadas y sin emoción).

<sup>1</sup> [people.rc.rit.edu/~coagla/affectdata/index.html](http://people.rc.rit.edu/~coagla/affectdata/index.html)

**Tabla 1.** Emociones básicas usadas en la clasificación.

Sigla	Clase de emoción
A	Enojo
D	Disgusto
F	Temor
H	Felicidad
Sa	Tristeza
Su+	Sorpresa positiva
Su-	Sorpresa negativa

[3]<sup>2</sup> describen el resultado de la aplicación de una encuesta internacional de antecedentes de emociones y reacciones (Internations Survey on Emotion Antecedents and Reactions – ISEAR); donde un selecto grupo de sicólogos de todo el mundo trabajaron en este proyecto y hubo una participación aproximada de 3000 estudiantes indicando situaciones del día a día en las que han experimentado emociones de alegría, miedo, temor, tristeza, desagrado, vergüenza y culpa.

En SemEval-2007 Strapparava2007<sup>3</sup> se cuenta con titulares de noticias obtenidos de diferentes ejemplares, como BBC, CNN, New York Times y el motor de búsqueda de noticias de Google. La estructura con la que cuenta un titular de noticias permite realizar la clasificación a nivel de oraciones. Cada una es clasificada de acuerdo con una o más de las siguientes emociones: enojo, disgusto, temor, alegría, tristeza y sorpresa. La implementación cuenta con dos conjuntos de datos uno para entrenamiento con 250 titulares, y uno de prueba con un total de 1000 titulares.

SemEval-2018 Mohammad2018<sup>4</sup> cuenta con la recopilación de tuits, donde se clasifican de tipo neutral o si expresan una o más emociones dentro de las que se encuentran enojo, desagrado, temor, alegría, amor, optimismo, pesimismo, tristeza, sorpresa y confianza. De igual manera existe diferenciación entre los conjuntos de entrenamiento. Los tuits fueron obtenidos en tres idiomas: inglés, árabe y español.

SemEval-2019 Chatterjee2019<sup>5</sup> presenta diálogos textuales entre dos personas donde el primero de ellos es quien inicia las conversaciones, luego el segundo responde y de nuevo el primero vuelve conversar en un esquema de turnos. Cada una de las conversaciones es etiquetada de acuerdo con la emoción que expresa; alegría, enojo, tristeza entre otras. La clasificación de la emoción es realizada con base en la interacción en el tercer turno de la conversación.

SemEval-2020 Patwa2020<sup>6</sup> aborda la clasificación de tuits en tres grupos principales de emociones: positivos que expresan felicidad, exaltar a una persona,

<sup>2</sup> [github.com/PoorvaRane/Emotion-Detector/blob/master/ISEAR.csv](https://github.com/PoorvaRane/Emotion-Detector/blob/master/ISEAR.csv)

<sup>3</sup> [web.eecs.umich.edu/~mihalcea/affectivetext/\#resources](http://web.eecs.umich.edu/~mihalcea/affectivetext/\#resources)

<sup>4</sup> [competitions.codalab.org/competitions/17751](http://competitions.codalab.org/competitions/17751)

<sup>5</sup> [www.humanizing-ai.com/emocontext.html](http://www.humanizing-ai.com/emocontext.html)

<sup>6</sup> [paperswithcode.com/paper/semEval-2020-task-9-overview-of-sentiment](https://paperswithcode.com/paper/semEval-2020-task-9-overview-of-sentiment)

grupo, país o celebrar algo. Negativos que expresan ataques a alguna persona, grupo, producto, o país, también disgusto, o tristeza hacia algo y crítica. Y neutrales que corresponden a noticias o publicidad.

En el cuadro 4 se realiza la clasificación de los conjuntos de datos especificados indicando el número de emociones según el tipo de emoción detectado y el número de ocurrencias de cada una de las emociones encontradas.

### 3.2. Lexicones

Existen diversos lexicones, también conocidos como diccionarios, que son utilizados comúnmente para el análisis de emociones. A continuación se describen los recursos existentes.

WordNet<sup>7</sup> se trata de una base de datos léxica en inglés en línea. Agrupa verbos, sustantivos, adjetivos, y adverbios en conjuntos de sinónimos denominados *synsets*.

WordNet-Affect<sup>8</sup> es una extensión de WordNet; un subconjunto de los synsets de éste con palabras que expresan emociones directas o indirectas, se utilizan etiquetas semánticas para su asignación.

SentiWordNet<sup>9</sup> asigna una de los tres posibles valores de clasificación positivos, negativos, u objetivos a cada uno de los synsets de WordNet

AFINN<sup>10</sup> consta de palabras calificadas de manera manual para el valor valence con un número entero entre menos cinco (negativo) y más cinco (positivo).

NRC Word-Emotion Association Lexicon<sup>11</sup> fue creado de manera manual utilizando el sistema de turcos mecánicos de Amazon. Ocho emociones, que son enojo, anticipación, disgusto, miedo, alegría, tristeza, sorpresa y confianza, y las categorías de sentimientos, positivos y negativos están incluidos.

NRC Affect Intensity Lexicon<sup>12</sup> proporciona puntuaciones de intensidad con valores reales por las emociones enojo, miedo, tristeza y alegría.

NCR Valence, Arousal y Dominance Lexicon<sup>13</sup> incluye una lista de más de 20.000 palabras y sus valores de *valence*, *arousal* y *dominance*. Las puntuaciones van de 0 a 1.

NRC Hashtag Emotion Lexicon<sup>14</sup> es un lexicón generado de manera automática a partir de tuits que incluyen *etiquetas* de palabras de emoción, como #happy. Asocia las palabras con las emociones ira, disgusto, miedo, tristeza, anticipación, sorpresa, alegría y confianza.

<sup>7</sup> [wordnet.princeton.edu/download/current-version](http://wordnet.princeton.edu/download/current-version)

<sup>8</sup> [multiwordnet.fbk.eu/english/home.php](http://multiwordnet.fbk.eu/english/home.php)

<sup>9</sup> [github.com/aesuli/SentiWordNet](https://github.com/aesuli/SentiWordNet)

<sup>10</sup> [github.com/fnielsen/afinn](https://github.com/fnielsen/afinn)

<sup>11</sup> [saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm](http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm)

<sup>12</sup> [www.saifmohammad.com/WebPages/AffectIntensity.htm](http://www.saifmohammad.com/WebPages/AffectIntensity.htm)

<sup>13</sup> [saifmohammad.com/WebPages/nrc-vad.html](http://saifmohammad.com/WebPages/nrc-vad.html)

<sup>14</sup> [saifmohammad.com/WebPages/lexicons.html\#NRCTwitter](http://saifmohammad.com/WebPages/lexicons.html\#NRCTwitter)

NRC Hashtag Sentiment Lexicon<sup>15</sup> fue creado de manera automática a partir de tuits que incluyen etiquetas de palabras de sentimiento como **#amazing**. Asocia palabras con un sentimiento positivo o negativo.

Sentiment140 Lexicon<sup>16</sup> fue generado también automáticamente a partir de tuits con emoticonos.

Finalmente, Spanish Emotion Lexicon (SEL)Sidorov2013<sup>17</sup> Contiene un total de 2036 palabras las cuales se encuentran asociadas con el factor de probabilidad de uso afectivo (FPA) con respecto a por lo menos una de las siguientes emociones básicas: alegría, ira, miedo, tristeza, sorpresa y disgusto. Rangel2014

#### 4. Técnicas utilizadas en el reconocimiento de emociones en textos

En trabajos previos realizados normalmente se realiza la distinción entre cuatro enfoques utilizados para reconocer las emociones en el texto que son los basados en palabras clave, basados en reglas, basados en el aprendizaje y basados en aprendizaje profundo. Los artículos se clasifican con base en el enfoque propuesto, el uso de dicha clasificación ayudará a evaluar estos enfoques en función de su desempeño, fortalezas y limitaciones y establecer una comparación entre ellos. Las emociones explícitas son reconocidas principalmente con enfoques basados en palabras clave. Los otros tres se utilizan principalmente para reconocer las emociones implícitas en el texto, a pesar de que también han sido utilizados para el reconocimiento de emociones explícitas.

##### 4.1. Enfoque basado en palabras clave

Este enfoque realiza la búsqueda de ocurrencias de palabras claves en un texto dado y realiza la asignación de una etiqueta de emoción de acuerdo con la palabra clave que ha sido detectada. La aproximación más utilizada es la técnica denominada *keyword-spotting*. Como inicio en el proceso, se realiza la definición de una lista de palabras para el etiquetado de las clases mediante el uso de los léxicos. Luego al texto se le realiza un preprocesamiento, que consiste en realizar la tokenización, eliminación de palabras gramaticales y su posterior lematización. Luego se procede a detectar las palabras clave de la emoción que se encuentran en el texto y compararlas contra el léxico utilizado. Posteriormente se analiza la intensidad de la emoción, se verifica si la emoción detectada corresponde a una negación y finalmente se realiza la asignación de la etiqueta de emoción para cada una de las oraciones en el texto. En la figura 3 se presenta el esquema general del enfoque en palabras clave.

<sup>15</sup> [saifmohammad.com/WebPages/lexicons.html\#NRCTwitter](https://saifmohammad.com/WebPages/lexicons.html\#NRCTwitter)

<sup>16</sup> [help.sentiment140.com/for-students](https://help.sentiment140.com/for-students)

<sup>17</sup> <https://www.cic.ipn.mx/~sidorov/>

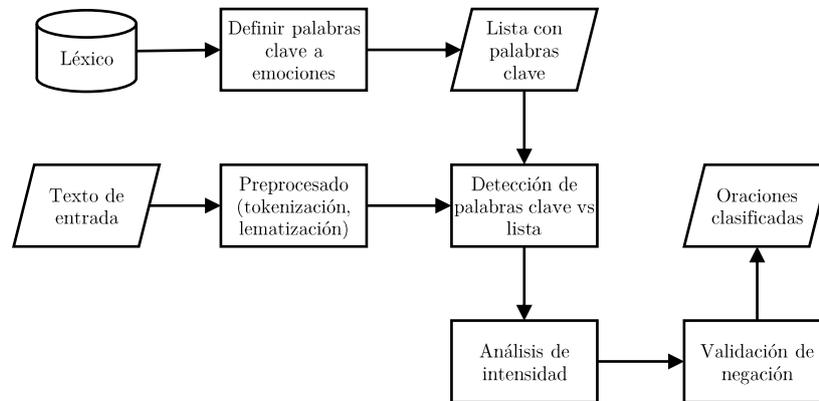


Fig. 3. Enfoque basado en palabras clave.

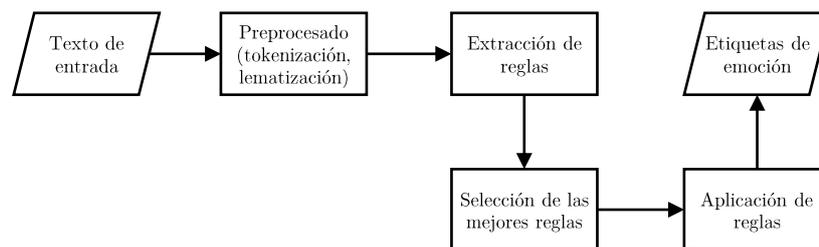


Fig. 4. Enfoque basado en reglas.

#### 4.2. Enfoque basado en reglas

Se basa en la manipulación de diferentes fuentes de conocimiento para realizar la interpretación de la información de entrada y realizar su correspondiente clasificación. En primer lugar, se realiza el preprocesamiento del texto de entrada, donde se aplica tokenización, eliminación de palabras gramaticales, lematización, POS tagging y análisis de dependencias. Luego, las reglas de emoción son extraídas utilizando conceptos lingüísticos, estadísticos y computacionales. Las mejores reglas son seleccionadas y finalmente son aplicadas al texto de entrada para determinar cada una de las etiquetas de emoción. En la figura 3 se representa el enfoque basado en reglas.

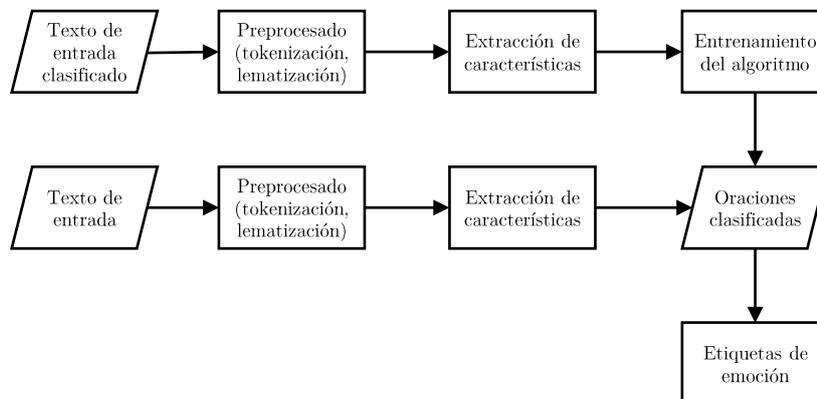


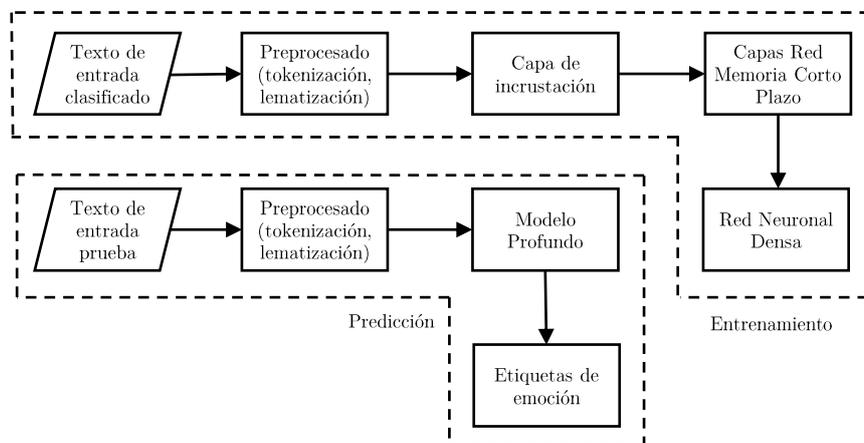
Fig. 5. Enfoque basado en aprendizaje.

#### 4.3. Enfoque basado en aprendizaje

Este enfoque proporciona al sistema clasificador la habilidad de aprender de manera automática y mejorar su desempeño con base en la experiencia adquirida en el proceso. Los algoritmos de *Machine Learning* son usualmente clasificados con tipo de aprendizaje supervisado y no supervisado. Uno de los algoritmos más utilizados en la clasificación de emociones en textos es el algoritmo de Máquinas de Vectores de Soporte por sus siglas en inglés SVM el cual es un algoritmo del tipo supervisado. En este algoritmo el funcionamiento consiste en igual manera de iniciar con el preprocesamiento del texto donde se realiza la tokenización, remoción de palabras gramaticales, lematización y POS tagging. El siguiente paso es extraer funciones útiles y se seleccionan aquellas con la mayor ganancia de información. Dado el conjunto de características y las etiquetas de emoción, el algoritmo SVM genera un hiperplano óptimo. Finalmente, el modelo SVM entrenado se utiliza para clasificar emociones en el texto de entrada para realizar la clasificación. En la figura 5 se representa el enfoque basado en aprendizaje.

#### 4.4. Enfoque basado en aprendizaje profundo

El aprendizaje profundo es una rama del *Machine Learning* en donde los diferentes programas implementados manejan el concepto de aprender de la experiencia donde el mundo es interpretado en conceptos de jerarquía de conceptos; realizando la definición de cada uno de los conceptos en términos de la relación con conceptos más sencillos. Esta aproximación permite que un programa adquiera conceptos complicados realizando su construcción a partir de conceptos más simples Goodfellow2016. Dentro de los modelos de aprendizaje profundo más utilizados se encuentra el de *Long Short Term Memory* (LSTM), el cual consiste en una forma especial de red neuronal recurrente con la capacidad



**Fig. 6.** Enfoque basado en aprendizaje profundo.

de manejar dependencias a largo plazo. Este método cuenta con la ventaja que supera el problema de desvanecimiento del gradiente. En este procedimiento se inicia con el preprocesamiento del texto, donde se realiza la tokenización, eliminación de palabras vacías y lematización. Después de eso, la capa de incrustación se construye y se alimenta a una o más capas de LSTM. Luego, la salida se alimenta a una red neuronal densa (DNN) con unidades iguales al número de etiquetas de emoción y una función de activación sigmoide para realizar la clasificación correspondiente. En la figura 6 se representa el enfoque basado en aprendizaje profundo.

#### 4.5. Medidas de evaluación

Cada uno de los trabajos identificados en el estado del arte realizan la medición de su desempeño mediante diferentes técnicas, para así realizar la medición de la similitud entre los conjuntos pronosticados contra los datos reales y también identificar el nivel de precisión en el pronóstico. Algunas medidas identificadas en los documentos son las siguientes.

*Índice de Jaccard:*

$$\mathbf{IJ} = \frac{1}{|S|} \sum_{s \in S} \frac{|G_s \cap P_s|}{|G_s \cup P_s|}, \quad (1)$$

donde  $S$  es el conjunto de frases a clasificar,  $G_s$  el conjunto de todas las etiquetas disponibles para clasificación y  $P_s$  el conjunto de todas las etiquetas predecidas para la frase.

*Precisión:*

$$\mathbf{P}_e = \frac{VP_e}{VP_e + FP_e}, \quad (2)$$

donde  $e$  corresponde a la etiqueta de emoción,  $VP$  a Verdadero Positivo y  $FP$  a Falso Positivo.

*Recall* :

$$\mathbf{R}_e = \frac{VP_e}{VP_e + FN_e}, \quad (3)$$

donde  $e$  corresponde a la etiqueta de emoción,  $VP$  a Verdadero Positivo y  $FN$  a Falso Negativo.

*F<sub>score</sub>* :

$$\mathbf{F}_1 = 2 \times \frac{P_e \times R_e}{P_e + R_e}, \quad (4)$$

$P_e$  corresponde al valor de la precisión y  $R_e$  al valor del recall calculados previamente.

*Accuracy* :

$$\mathbf{A} = \frac{\sum_{e \in E} VP + \sum_{e \in E} VN}{\sum_{e \in E} VP + \sum_{e \in E} VN + \sum_{e \in E} FP + \sum_{e \in E} FN}, \quad (5)$$

donde  $E$  corresponde al conjunto de etiquetas de emoción,  $VP$  corresponde al número de verdaderos positivos,  $VN$  es el número de verdaderos negativos,  $FP$  es el número de Falsos Positivos y  $FN$  es el número de falsos negativos.

#### 4.6. Resumen del estado del arte

En el cuadro 5 se presenta el resumen del estado del arte revisado realizando una comparación entre la bibliografía mediante una tabla donde se indica el artículo mencionado, el idioma de los conjuntos de datos utilizados, el enfoque utilizado, la herramienta y/o algoritmo utilizados en la implementación, el nombre del corpus utilizado y el desempeño del modelo implementado.

### 5. Trabajo propuesto

Hemos identificado diversas soluciones que se han presentado para resolver el problema de reconocimiento de emociones en distintos conjuntos de datos. Una parte importante que identificamos es que no se ha estudiado a profundidad la posible distorsión en las formas de escribir texto. Como ejemplo de ello, podemos considerar la escritura *Leet* o *1337 5p34k*. Ésta consiste en la combinación de caracteres alfabéticos y numéricos con el objetivo de reemplazar letras por dígitos e incluso símbolos que tienen parecido con la letra que se desea sustituir. Usualmente este tipo de escritura es utilizada para omitir los filtros de moderación de lenguaje en las redes sociales y principalmente expresar emociones

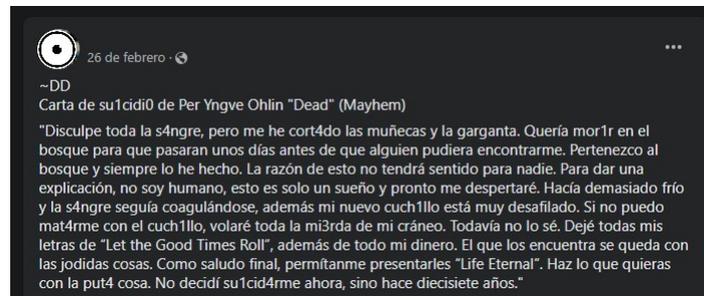


Fig. 7. Ejemplo de escritura *leet* en mensajes depresivos.

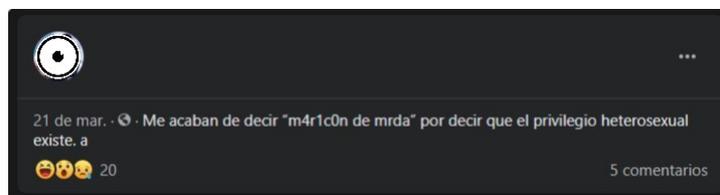


Fig. 8. Ejemplo de escritura *leet* en mensajes de odio.

radicales de odio o tristeza profunda como se puede evidenciar en las figuras 7 y 8. La identidad de los autores de las publicaciones fue eliminada de manera intencional para su protección.

En la documentación revisada y verificando en los conjuntos de datos utilizados, todos los textos tienen en común que son frases o escritos completos donde su gramática se encuentra configurada completamente con palabras y no tienen en cuenta este tipo de cambios en las letras, por lo que en los preprocesamientos de los textos antes de los métodos de entrenamiento, estas palabras pueden ser descartadas en el proceso de limpieza y cambiar la emoción que se transmite. Si se retoma la frase de la figura 8 y se descartan las palabras escritas en forma leet y las palabras con letras faltantes, la frase resultante es la siguiente:

*“Me acaban de decir de por decir que el privilegio heterosexual existe”*

Con esta simple observación se puede evidenciar que, del mensaje de odio agresivo identificado en el escrito original se convierte en uno que carece de emoción y podría ser clasificado con la etiqueta de emoción neutra lo cual obviamente es una mala asignación.

Varios de los usuarios de las redes sociales han identificado que escribir mensajes con este tipo de escritura no genera repercusiones en sus perfiles que usualmente consisten en el bloqueo de sus cuentas, lo cual ha incentivado que esta técnica se haya masificado en más de una de las redes sociales actuales. Realizando una búsqueda sencilla con la palabra “suicidio” en Google, arroja un aproximado de 950,000 resultados y revisando algunos de ellos son bastante perturbadores. Por ello realizar el análisis de emociones en textos incluyendo

esta nueva variable, podría generar resultados más precisos en la clasificación de emociones extremistas que usualmente suelen ser las más peligrosas y que desencadenan en acciones que atentan contra la integridad física de las personas.

### 5.1. Algunas investigaciones en la detección de emociones negativas

Realizando la consulta a trabajos de investigación centrados en el análisis de emociones negativas, se realizó un análisis de las tecnologías implementadas para dicha tarea. En el trabajo de Villatoro-Tello2017 el objetivo principal consiste en realizar una predicción de manera anticipada de casos de depresión en las publicaciones realizadas por los usuarios. Como preprocesamiento del texto deciden por abandonar las técnicas comunes como las bolsas de palabras y realizar el modelamiento de los datos al igual que los modelos de predicción mediante el uso de grafos y clasificar las publicaciones en dos clases de resultados, si la persona que la escribió se encuentra deprimida o no deprimida. Las ejecuciones del modelo se realizaron con 5 configuraciones diferentes: LyRA donde el modelo es construido mediante los términos como palabras independientes. LyRB construye el modelo mediante modelos de 3-gramas. LyRC es una configuración híbrida generando un vector de  $2 \times 4 \times N$  extraídas de las características de las dos configuraciones anteriores. LyRD es un método de ensamblaje conservador que utiliza las salidas de las tres configuraciones anteriores y realiza la asignación de deprimido si las tres primeras configuraciones coinciden en asignar dicha categoría. Finalmente, LyRE es similar a la configuración anterior, pero donde la clasificación se realiza con base en la clasificación de la mayoría de los tres modelos iniciales en un esquema 2 de 3.

En la investigación de Hosseini-Saravani2020 la meta consiste en medir manifestaciones conductuales asociadas a la depresión con un total de 21 categorías de síntomas y actitudes que de acuerdo con estudios clínicos son características que poseen aquellas personas con algún nivel de depresión. Cada categoría es graduada con 4 o 5 frases para validar con cual se puede sentir identificado el paciente y se asigna un rango de 0 a 3 para reflejar el rango de la severidad asociada a cada una de las categorías. Como método de clasificación se utilizó el método de *Naïve Bayes*.

Los siguientes dos trabajos ya se enfocan en tendencias más fuertes que corresponden a la identificación de tendencias suicidas en personas que ya pueden contar con un nivel bastante profundo de depresión. Tadesse2020 realiza la identificación de publicaciones con tendencias suicidas en la red social Reddit. La metodología implementada consistió en el diseño e implementación de un modelo híbrido de Redes Neuronales Convolucionales (CNN) y redes recurrentes del tipo *Long Short Term Memory* (LSTM), donde las entradas al sistema son procesadas por la capa LSTM y sus salidas son el insumo para la sección CNN. Como resultado del estudio se obtuvieron aquellas palabras que son muy frecuentes y presentan una alta ocurrencia en aquellas publicaciones con tendencias suicidas como los son "suicidio", "quiero morir", "morir maldita sea", "deseo suicida", "deseo morir", "me quiero ir." entre otras. La medida de *accuracy* para esta implementación fue del 91.7%.

**Tabla 2.** Tabla de equivalencia propuesta para vocales y algunas consonantes.

Letra	Alfanumérica	Alfabéticos	Caracteres especiales
A	4	ÁĂÄÅ	@
E	3	ÊË	€
I	1	Ï	i
O	0	ÔÕÖ	Ø
U	-	ÛÜ	-
L	1	Łł	!£
R	-	Ŕŕ	®
S	5	ŚśŞş	§
T	7	Ŧ	-

Sawhney2020 de igual manera realiza el análisis de tendencias suicidas en la red social Twitter mediante un conjunto de tuits en idioma inglés, pero basando la recopilación de datos en un usuario en específico y considerando su comportamiento de publicaciones en un intervalo de tiempo para así poder detectar con mayor precisión si existe un patrón en cada una de las publicaciones y poder llegar así a considerar un punto de alerta máxima en la que el usuario pueda llegar a intentar a hacer algo en contra de su vida. Como metodología implementada se utilizaron transformadores entrenados previamente, en particular el denominado BERT. En una primera capa de este modelo realiza la tokenización de todas las publicaciones históricas para así realizar la creación de un vector emocional inicial. Para la capa subsiguiente se realizó la implementación de capas *Time-Aware Long-Short Term Memory* (T-LSTM) debido a que precisamente el comportamiento del usuario a analizar puede ser cambiante en los intervalos de tiempo por lo que este sistema resulta más eficaz para comparar datos anteriores en el tiempo con comportamientos actuales. La medida de *accuracy* para esta implementación fue del 85.1 %.

## 5.2. Propuesta inicial de trabajo

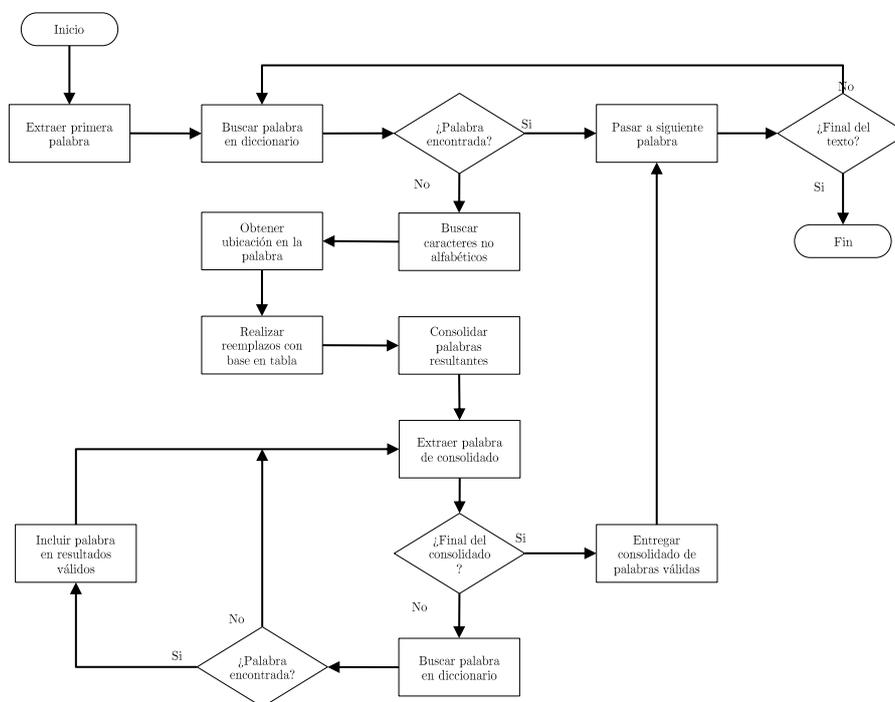
De acuerdo con el trabajo realizado por Tundis2021 en el que se pueden evidenciar varias formas en las que se realiza la representación de frases como combinación entre diferentes tipos de caracteres e idiomas. Para la presente propuesta se utilizará la categoría denominada *single row - single language* y como subcategorías se realizará uso de aquellas implementaciones alfanuméricas, mezcla de diferentes caracteres alfabéticos y el uso de caracteres especiales en la representación de letras.

Por ello, en el cuadro 2 se realiza una definición inicial para las vocales en cada una de las tres subcategorías y para algunas consonantes, la cual posteriormente debe ser modificada para incluir todas las letras del alfabeto.

Se realizó un ejercicio de prueba de concepto con las palabras "deprimirse", "d3pr1m1rse", "dep®imirse", "d3p®imirse" como insumo de entrada para que

**Tabla 3.** Resultado de evaluación de palabras con Stanza.

Palabra	Etiqueta		Lema 1	Etiqueta		Etiqueta
	Gramatical	Lematización		Gramatical	Lema 2	
				Lema 1	Lema 2	Lema 2
deprimirse	-	Si	Deprimir	Verbo	Él	Pronombre
d3pr1m1rse	Numeral	-	-	-	-	-
depⓂimirse	-	Si	dep<UNK>imir	Sustantivo	Él	Pronombre
d3pⓂimirse	Verbo	No	-	-	-	-



**Fig. 9.** Diagrama de flujo reglas de preprocesamiento.

fueran analizados por la librería de procesamiento de lenguaje natural Stanza para verificar cómo realiza la clasificación de estas palabras. El resultado se presenta en el cuadro 3 y se puede evidenciar que cada una de las cuatro palabras fue categorizada de manera diferente y para aquella que pudo ser lematizada la librería no pudo reconocer el carácter especial Ⓜ detectándolo como algo desconocido. Y en el último caso detectó la palabra como un verbo el cual no lematizó, lo cual también es inválido.

**Tabla 4.** Corpora para reconocimiento de emociones en textos.

Emoción	Alm et al	Aman	ISEAR	SemEval 2007		SemEval 2018		SemEval 2019		SemEval 2020	
				Entr	Prueba	Ent	Prueba	Ent	Prueba	Ent	Prueba
Alegría	445	536	1094	35	113	2477	1442	4243	284	-	-
Amor	-	-	-	-	-	700	516	-	-	-	-
Anticipación	-	-	-	-	-	978	425	-	-	-	-
Confianza	-	-	-	-	-	357	153	-	-	-	-
Culpa	-	-	1093	-	-	-	-	-	-	-	-
Disgusto	-	172	1096	9	12	2602	1099	-	-	-	-
Enojo	-	179	1096	20	21	2544	1101	5506	298	-	-
Enojo-Disgusto	218	-	-	-	-	-	-	-	-	-	-
Miedo	166	115	1095	32	92	1242	485	-	-	-	-
Neutral	-	2800	-	146	650	2207	75	-	-	3974	206
Optimismo	-	-	-	-	-	1984	1143	-	-	6005	3061
Otros	-	-	-	-	-	-	-	14948	4677	-	-
Pesimismo	-	-	-	-	-	795	375	-	-	2023	522
Sorpresa	114	115	-	8	42	361	170	-	-	-	-
Tristeza	264	173	1096	41	104	2008	960	5463	250	-	-
Vergüenza	-	-	1096	-	-	-	-	-	-	-	-
Total	1207	4090	7666	291	1034	18255	7944	30160	5509	12002	3789

Con base en el análisis realizado, se realiza como propuesta inicial la secuencia de reglas presentada en la figura 9 como un proceso previo para el preprocesamiento del texto con las características mencionadas y cuyo resultado ya podría pasar a la etapa de lematización para el proceso posterior de la detección de emociones en el texto.

En el mejor de los escenarios como palabras resultantes del proceso se va a generar una sola palabra válida. para el ejemplo de la palabra "d3p®imirse" se realizaría el reemplazo del valor 3 por E y ® por R de acuerdo con lo definido en el cuadro 2. Pero si se realiza una modificación y ahora la palabra a procesar es cop|a, y al realizar la transformación es posible tener dos resultados en donde el caracter | puede ser reemplazado por L o I obteniendo como resultados las palabras copia y copla las cuales son perfectamente válidas y seguramente van a existir en los diccionarios. Por ello el análisis del contexto será un factor determinante al momento de seleccionar la palabra final que en este caso reemplazará a cop|a.

## 6. Conclusiones

Los usuarios de las redes sociales se han dado cuenta que las publicaciones realizadas en las diferentes redes sociales empiezan a contar con una mayor restricción donde estas son revisadas y con base en la detección de palabras de alerta pueden advertir que la publicación puede atentar contra las normas de la comunidad de la plataforma y que la red social se encuentra en la potestad de tomar las medidas necesarias para evitar este tipo de comportamientos. Debido a esto los usuarios empiezan a recurrir a nuevas técnicas para realizar sus publicaciones con métodos de ofuscación en las palabras con varios niveles de complejidad según los conocimientos con los que posea el usuario para que así los filtros de restricción no generen ningún tipo de alertamiento y la publicación sea aceptada sin ningún inconveniente. Como se pudo evidenciar en la prueba de

**Tabla 5.** Resumen de trabajos sobre reconocimiento de emociones en textos en las redes sociales.

Referencia	Idioma	Método	Herramienta - Algoritmo	Dataset	Desempeño	
					Medida	Valor (%)
Bandhakavi2017	Inglés	Aprendizaje clásico	Máquinas de vectores de soporte	SemEval-07	F-Score	40.99
				Twitter		49.55
				Blogs		58.32
				ISEAR		31.21
Yasmina2016	Inglés	Aprendizaje clásico	Máquinas de vectores de soporte WEKA	Comentarios en YouTube	Presición	92.75
					Recall	72.72
					Accuracy	68.82
Wu2018	Inglés	Aprendizaje Profundo	Redes Neuronales Convolucionales	SemEval2018	Pearson	81.00
Tao2020	Chino	Aprendizaje Profundo	Redes de Memoria a corto plazo	Reseñas Hoteles	Presición	88.47
			LDA (Latent Dirichlet Allocation)		Recall	79.88
Ghanbari-Adivi2019	Inglés	Híbrido: Reglas y Aprendizaje clásico	Máquinas de vectores de soporte Vecino más cercano Perceptrón multicapa Árboles de decisión	ISEAR	Accuracy	87.51
					Presición	82.00
					Recall	92.88
					Accuracy	88.49
					F-Score	81.00
Sailunaz2019	Inglés	Aprendizaje clásico	Naive Bayes	ISEAR	Presición	82.00
			Máquinas de vectores de soporte Bosques aleatorios	Twitter	Accuracy	66.86
Ibrahiem2020	Inglés	Aprendizaje clásico	Clasificación binaria	SemEval2018	Presición	63.00
			Redes Neuronales Convolucionales	NRC Emotion NRC Hashtag	Recall	82.00
Baziotis2018	Inglés	Aprendizaje Profundo	Clasificación binaria	SemEval2017	F-Score	69.00
			Redes Neuronales Recurrentes		Jaccard	57.88
Ge2019	Inglés	Aprendizaje Profundo	Redes de Memoria a Corto Plazo	SemEval2019	Presición	75.98
			Redes Neuronales Convolucionales		Recall	74.88
Patwa2020	Hinglish Spanglish	Aprendizaje Profundo	Redes Neuronales Convolucionales	SemEval2020	F-Score	75.42
						F-Score

concepto, realizar la modificación de una palabra que puede generar un alto nivel de alerta cambia la forma en que esta es interpretada por las herramientas de procesamiento de lenguaje natural actuales, lo cual ocasiona que la emoción real a transmitir llegue a ser clasificada de manera errónea con el sentimiento real que se desea transmitir y esto conlleva a que no se puedan prevenir situaciones de riesgo en la vida de las personas, por lo que el estudio de investigación propuesto presenta un punto de vista antes no investigado e innovador y que está adquiriendo relevancia debido a que las redes sociales en los últimos años realizar un mayor control sobre aquello que publican los usuarios en sus plataformas.

## 7. Trabajo futuro

Como trabajo futuro, se debe realizar una muy amplia, cuidadosa y profunda definición de la tabla de equivalencias en la sustitución de cada uno de los diferentes caracteres especiales debido a que su definición y la aplicación de las reglas sobre ésta consistirá en el factor principal en la cantidad de palabras válidas que saldrán como resultado del algoritmo de preprocesamiento en los textos. Adicionalmente, también de acuerdo con esos resultados, se deben analizar y finalmente seleccionar la mejor herramienta tecnológica para realizar la clasificación de cada uno de los textos.

## **Referencias**

1. Alm, C. O., Roth, D., Sproat, R.: Emotions from text: Machine learning for text-based emotion prediction. HLT/EMNLP 2005 - Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, , no. October, pp. 579–586 (2005)
2. Haythornthwaite, C.: Social networks and internet connectivity effects. *Information, Communication & Society*, vol. 8, no. 2, pp. 37–41 (2005)
3. Scherer, K. R., Wallbott, H. G.: Evidence for universality and Cultrual Variation, vol. 66, no. 2, pp. 310–328 (1994)

Electronic edition  
Available online: <http://www.rcs.cic.ipn.mx>



<http://rsc.cic.ipn.mx>



Centro de Investigación  
en Computación