

34

Fecha de presentación: Septiembre, 2021

Fecha de aceptación: Noviembre, 2021

Fecha de publicación: Diciembre, 2021

ANÁLISIS DE SENTIMIENTOS:

HERRAMIENTA PARA ESTUDIAR DATOS CUALITATIVOS EN LA INVESTIGACIÓN JURÍDICA

SENTIMENT ANALYSIS: A TOOL FOR STUDYING QUALITATIVE DATA IN LEGAL RESEARCH

Maikel Yelandi Leyva Vázquez¹

E-mail: ub.c.investigacion@uniandes.edu.ec

ORCID: <https://orcid.org/0000-0001-7911-5879>

Remigio Edmundo Hernández Cevallos¹

E-mail: ub.remigiohernandez@uniandes.edu.ec

ORCID: <https://orcid.org/0000-0003-3782-8031>

Jesus Estupiñan Ricardo¹

E-mail: ua.jesusestupinan@uniandes.edu.ec

ORCID: <https://orcid.org/0000-0002-1595-6174>

¹ Universidad Regional Autónoma de Los Andes. Ecuador.

Cita sugerida (APA, séptima edición)

Leyva Vázquez, M. Y., Hernández Cevallos, R. E., & Estupiñan Ricardo, J. (2021). Análisis de sentimientos: herramienta para estudiar datos cualitativos en la investigación jurídica. *Revista Universidad y Sociedad*, 13(S3), 262-266.

RESUMEN

La investigación jurídica posee gran importancia por el papel que juega en la resolución de problemas legales. Los datos cualitativos proporcionados por herramientas de investigación como las entrevistas individuales constituyen un método para extraer conclusiones amplias. Sin embargo, por su subjetividad, se basan en gran medida en los prejuicios inconscientes de los autores o investigadores. Debido a lo cual el objetivo del presente trabajo radica en aplicar el análisis de sentimientos en la investigación jurídica mediante un método de procesamiento de lenguaje natural y los Números Neutrosóficos de Valor Único. La aplicación del método VADER neutrosófico mediante la herramienta Orange permitió el cumplimiento del objetivo propuesto de la investigación demostrándose su aplicabilidad mediante estudios de casos.

Palabras clave: Investigación jurídica, análisis de sentimientos, números neutrosóficos de valor único, Vader, Orange.

ABSTRACT

Legal research is of great importance because of the role it plays in solving legal problems. Qualitative data provided by research tools such as individual interviews provide a method for drawing broad conclusions. However, because of their subjectivity, they rely heavily on the unconscious biases of the authors or researchers. Due to which the aim of the present work lies in applying sentiment analysis in legal research using a natural language processing method and the Neutrosophic Single Value Neutrosophic Numbers. The application of the neutrosophic VADER method by means of the orange tool allowed the fulfillment of the proposed objective of the investigation demonstrating its applicability by means of case studies.

Keywords: Legal research, sentiment analysis, neutrosophic single-valued numbers, Vader, Orange.

INTRODUCCIÓN

En las ciencias sociales, una metodología de investigación primaria como las entrevistas individuales constituye una técnica muy utilizada para obtener conocimientos significativos y extraer conclusiones amplias las cuales ayudan a realizar un análisis cualitativo (Sasangohar et al., 2021).

Sin embargo, estos análisis son subjetivos y se basan en gran medida en los prejuicios inconscientes de los autores o investigadores. Además, el aprendizaje de cada nueva entrevista disminuye a un ritmo elevado y no es un uso eficiente del valioso tiempo de los investigadores. Lo cual es una de las limitantes del uso de los datos cualitativos en la investigación jurídica. (Estupiñán Ricardo et al., 2020; Álvarez et al. 2021).

Por tanto, los autores de la presente investigación consideran pertinente el uso de la Neutrosofía como complemento por el tratamiento a las indeterminaciones que suelen acompañar este tipo de estudios; debido a que esta ciencia, se caracteriza por el concepto de indeterminación. A pesar de no haberse utilizado ampliamente en el análisis de sentimientos de las entrevistas. Los conjuntos neutrosóficos se utilizan para el análisis de sentimientos de las entrevistas como herramienta de investigación cualitativa (Kandasamy et al., 2020). Este estudio es el primer paso de una investigación que señala las indeterminaciones en el análisis discursivo (Smarandache et al., 2018; Palacios et al. 2021).

Atendiendo a lo anterior se establece como objetivo del presente documento aplicar el análisis de sentimientos en la investigación jurídica mediante un método de procesamiento de lenguaje natural y los Números Neutrosóficos de Valor Único (SVN). Con el objetivo de humanizar el trabajo del procesamiento de los datos cualitativos y proveer a los juristas de una forma rápida y eficiente de generar conclusiones en sus investigaciones. Lo cual está dado por la importancia que engloba el enfoque cualitativo como forma del jurista de resolver los problemas de la realidad socio jurídica a las que se enfrenta.

Preliminares

Neutrosofía en el análisis de sentimientos conceptos básicos

La neutrosofía es una teoría matemática desarrollada por el erudito rumano Florentin Smarandache para lidiar con la indeterminación (Smarandache, 2002). Ha sido la base para el desarrollo de nuevos métodos para lidiar con información indeterminada e inconsistente como neutrosóficos conjuntos de lógica neutrosófica y, especialmente,

en problemas de toma de decisiones (Sleem et al., 2020). El valor de verdad en el conjunto neutrosófico es el siguiente (Ricardo et al. 2018; Smarandache et al., 2020):

Sea, sea una evaluación neutrosófica de un mapeo de un grupo de fórmulas proposicionales para y para cada oración:

$$v(p)=(T,I,F) \tag{1}$$

Facilitar la aplicación práctica en problemas del mundo real. (Smarandache et al., 2021), se propuso el uso de Conjuntos neutrosóficos de valor único (SVNS), a través de los cuales es probable que se utilicen términos lingüísticos para obtener una mayor interpretabilidad de los resultados (Saqlain et al., 2021).

Sea X un universo de discurso, un SVNS A sobre X tiene la siguiente forma (Cadena et al., 2020):

$$A = \{ \langle x, u_a(x), r_a(x), v_a(x) \rangle : x \in X \} \tag{2}$$

Dónde

$$\langle x \rangle : X \rightarrow [0,1], r_a(x) : X \rightarrow [0,1] \text{ y } v_a(x) : X \rightarrow [0,1] \tag{3}$$

con $0 \leq u_a(x), r_a(x), v_a(x) \leq 3, \forall x \in X$

Los intervalos denotan las pertenencias relacionadas con verdadero, indeterminado y falso de x en A, respectivamente (Villafuerte et al., 2020). Por razones de conveniencia, un número neutrosófico de valor único (SVN) se expresa como A = (a, b, c), donde a, b, c ∈ [0, 1] y $0 \leq a + b + c \leq 3$.

Sea A = (a, b, c) un número neutrosófico de valor único, una función de puntuación S relacionada con un valor neutrosófico de valor único, basada en el grado de pertenencia a la verdad, el grado de pertenencia a indeterminación y el grado de pertenencia a la falsedad se define por (Leyva-Vázquez et al., 2020):

$$s(V_i) = 2 + T_i - F_i - I_i \tag{4}$$

Tabla 1: Términos lingüísticos asociados a los números neutrosóficos de valor único.

Términos lingüísticos	Números SVNN
Extremadamente buena (EB)	(1,0,0)
Muy muy buena (MMB)	(0.9, 0.1, 0.1)
Muy buena (MB)	(0.8,0,15,0.20)
Buena(B)	(0.70,0.25,0.30)
Medianamente buena (MDB)	(0.60,0.35,0.40)
Media(M)	(0.50,0.50,0.50)

Medianamente mala (MDM)	(0.40,0.65,0.60)
Mala (MA)	(0.30,0.75,0.70)
Muy mala (MM)	(0.20,0.85,0.80)
Muy muy mala (MMM)	(0.10,0.90,0.90)
Extremadamente mala (EM)	(0,1,1)

Fuente: (Şahin & Yiğider, 2014)

Se propone la función de puntuación para conjuntos neutrosóficos de un solo valor para hacer la distinción entre números (tabla 1). Donde los mismos se fusionarán en la asociación de los términos o vocablos que indiquen polaridad. La clasificación de la polaridad se hará atendiendo positiva (EB, MMB, MB, B, MDB), neutral (M) y negativa (MDM, MA, MM, MMM, EM). Para su de-neutrosificación aplicar ecuación 4 para determinar en un rango de , el nivel de sentimiento y evaluar así el proceso para arribar a conclusiones.

Método VADER

El método VADER (Valence Aware Dictionary for Sentiment Reasoning text sentiment analysis en inglés) (Borg & Boldt, 2020) resulta una de las herramientas para el análisis de sentimientos. La misma ostenta uno de los mayores éxitos porque no solo habla de la puntuación positiva o negativa, sino también, sobre qué tan positivo o negativo puede ser un sentimiento, lo que resulta conveniente para esta investigación (Gilbert, 2018).

Se basa en reglas que clasifican las polaridades de las emociones de las opiniones utilizando un listado de palabras llamado lexicón. Este contiene un total de 7,517 palabras incluidos emoticones, abreviaciones, acrónimos e iniciales etiquetados mediante una valencia de -4 a 4, el lexicón de VADER se obtuvo aplicando la metodología de Machine Learning, Wisdom-of-the-Crowd (la sabiduría de los grupos) y el uso de lexicones (LIWC, ANEW, GI). La clasificación de la polaridad (positiva, negativa o neutra) de una opinión se realiza mediante los valores de cada palabra en el lexicón del método.

MATERIALES Y MÉTODOS

Se empleó la estadística neutrosófica para el cálculo de la muestra. La población total calcula la mediante la siguiente expresión:

$$n = \frac{N Z^2 pq}{d^2(N - 1) + Z^2 pq} \quad (5)$$

Donde:

p = proporción aproximada del fenómeno en estudio en la población de referencia

q = proporción de la población de referencia que no presenta el fenómeno en estudio ($1 - p$).

El nivel de confianza deseado (Z). Indica el grado de confianza que tendrá el valor verdadero del parámetro en la población se encontrará en la muestra calculada.

La precisión absoluta (d). Es la amplitud deseada del intervalo de confianza a ambos lados del valor real de la diferencia entre las dos proporciones (en puntos porcentuales). N es tamaño de la población

En este caso de desea un nivel de confianza de un 90%, $z = 1.645$, $d = 0.1$, $p = 0.4$ y $N = [10,20]$ El resultado al que denominamos muestra neutrosófica $n = [4, 16]$ indica que la muestra debe estar en valores entre 4 y 16 individuos.

Para el procesamiento de la información se utilizó Orange Data Mining (Demšar et al., 2013) para analizar el sentimiento en las entrevistas (Figura 1).

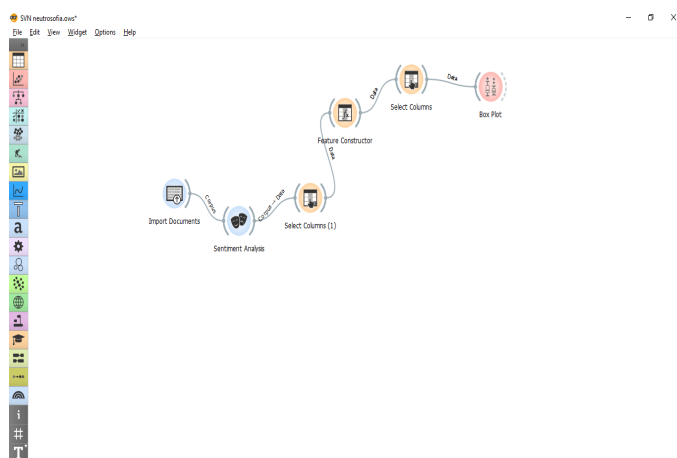


Figura 1. Flujo de datos en Orange

En este software el componente de análisis de opinión predice la opinión de cada documento en un corpus, el cual fue programado acorde a la metodología propuesta.

Estudio de casos

Se utilizó un grupo de 5 entrevistas, donde por conveniencia, el número neutrosófico de valor único asociado al análisis de sentimiento se expresará como $A = (\text{pos}, \text{net}, \text{neg})$, donde pos, net y neg positivo son puntuaciones compuestas positivas, neutrales y negativas, respectivamente acorde a la tabla 1. Los resultados para cada caso se exponen en la tabla 2.

Tabla 2. Asociación Vader-SVN para cada caso de estudio.

Caso	Número SVN
Caso 1	(0.052,0.909,0.04)
Caso 2	(0.336,0.622,0.042)
Caso 3	(0.044,0.814,0.142)
Caso 4	(0,1,0)
Caso 5	(0.075,0.746,0.179)

Las puntuaciones de cada entrevista se calcularon utilizando un componente de *Feature Constructor* (Figura 2).

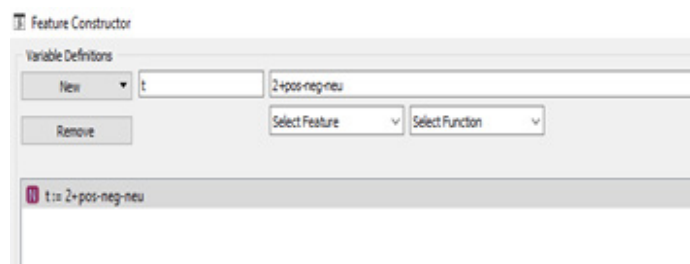


Figura 2. Cálculo de puntaje con componente Feature Constructor

Fuente: Salida de datos de Orange

Los resultados del cálculo utilizando la ecuación 4 como método de de-neutrosificación se muestran en la Figura 3.

id	path	seter	name	t
2	C:/Users/.../Pie...	Case2		1.672
5	C:/Users/.../Joh...	Case5		1.15
1	C:/Users/.../Du...	Case1		1.103
3	C:/Users/.../Ca...	Case3		1.088
4	C:/Users/.../Ih...	Case4		1

Figura 3. Cálculo de puntuación

Fuente: Salida de Orange

Los métodos muestran un puntaje de sentimiento para cada entrevista, esa función de puntaje permite clasificar números neutrosóficos de un solo valor y da un solo valor numérico que posibilita la comparación y análisis con respecto a otros casos.

Para los casos estudiados los resultados arrojaron que las opiniones se encuentran segmentadas de la siguiente manera:

- Polaridad media: Caso 2

- Polaridad medianamente negativa: Caso 5
- Polaridad muy negativa: Casos 1, 3 y 4

Por lo que la situación se califica de media baja, debido a la puntuación mostrada y su clasificación de polaridad.

CONCLUSIONES

El análisis de sentimiento resulta una herramienta útil para el jurista sobre todo para el procesamiento rápido y eficaz de un gran cúmulo de datos cualitativos sometidos a la subjetividad de los expertos que los evalúan. Lo cual puede condicionar a favor o en contra del proceso las evidencias obtenidas y llegar a conclusiones fallidas.

La aplicación del método VADER y su fusión con los números neutrosóficos de valor único permitió el cumplimiento del objetivo propuesto de la investigación. Ya que se pudo aplicar una herramienta al análisis de datos cualitativos que produjo resultados rápidos y veraces acorde al pensamiento humanos

Se confirmó la importancia que engloba el enfoque cualitativo como forma del jurista de resolver los problemas de la realidad socio jurídica a las que se enfrenta. Ya que se pudieron obtener resultados vitales para el arribo de conclusiones en la investigación.

REFERENCIAS BIBLIOGRÁFICAS

- Álvarez Gómez, S. D., Romero Fernández, A. J., Estupiñán Ricardo, J., & Ponce Ruiz, D. V. (2021). Selección del docente tutor basado en la calidad de la docencia en metodología de la investigación. *Conrado*, 17(80), 88-94.
- Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2020.113746>
- Cadena, M. T., Burgos, M. F. J., & Montenegro, M. J. J. (2020). Neutrosophic Case-Based Reasoning Method to Determine the Profitability of the Tourism Sector in the City of Riobamba. *Neutrosophic Sets and Systems*, 37(1), 1-13.
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., ... & Zupan, B. (2013). Orange: data mining toolbox in Python. *the Journal of machine Learning research*, 14(1), 2349-2353.

- Estupiñán Ricardo, J., Vaca Rosado., V. M., Piedra Fernández, J., & Mantilla Martínez., S. (2020). Importancia de la investigación jurídica para la formación de los profesionales del Derecho en Ecuador. *Revista Dilemas Contemporáneos: Educación, Política y Valores.*, 7(Edición Especial), 1–25.
- Gilbert, E. (2018). VADER: A Parsimonious Rulebased Model for Sentiment Analysis of social media Text. *Georgia Institute of Technology, Atlanta*. https://www.researchgate.net/publication/275828927_VADER
- Kandasamy, I., Vasantha, W. B., Obbineni, J. M., & Smarandache, F. (2020). Sentiment analysis of tweets using refined neutrosophic sets. *Computers in Industry*, 115, 103180.
- Leyva-Vázquez, M., Quiroz-Martínez, M. A., Portilla-Castell, Y., Hechavarría-Hernández, J. R., & González-Caballero, E. (2020). A New Model for the Selection of Information Technology Project in a Neutrosophic Environment. *Neutrosophic Sets and Systems*, 32(1), 344-360.
- Palacios, A. J. P., Ricardo, J. E., Piza, I. A. C., & Herrería, M. E. E. (2021). Phenomenological hermeneutical method and neutrosophic cognitive maps in the causal analysis of transgressions against the homeless. *Neutrosophic sets and systems*, 44, 147-156.
- Ricardo, J. E., Peña, R. M., Zumba, G. R., & Fernández, I. I. O. (2018). *La Pedagogía como Instrumento de Gestión Social: Nuevos Caminos para la Aplicación de la Neutrosofía a la Pedagogía*. Infinite Study.
- Şahin, R., & Yiğider, M. (2014). A Multi-criteria neutrosophic group decision making method based TOPSIS for supplier selection. ArXiv Preprint ArXiv:1412.5077.
- Saqlain, M., Jafar, N., Moin, S., Saeed, M., & Broumi, S. (2020). Single and Multi-valued Neutrosophic Hypersoft set and Tangent Similarity Measure of Single valued Neutrosophic Hypersoft Sets. *Neutrosophic Sets and Systems*, 32, 317–329.
- Sasangohar, F., Dhala, A., Zheng, F., Ahmadi, N., Kash, B., & Masud, F. (2021). Use of telecritical care for family visitation to ICU during the COVID-19 pandemic: an interview study and sentiment analysis. *BMJ quality & safety*, 30(9), 715-721.
- Sleem, A., Abdel-Baset, M., & El-henawy, I. (2020). PyIVNS: A python-based tool for Interval-valued neutrosophic operations and normalization. *SoftwareX*.
- Smarandache, F. (2002). *Neutrosophy, a new Branch of Philosophy*. Infinite Study.
- Smarandache, F., Abdel-Basset, M., & Broumi, S. (2021). *Neutrosophic Sets and Systems*, vol. 40, 2021. *Neutrosophic Sets and Systems*, 40(1), 1-32.
- Smarandache, F., Quiroz-Martínez, M. A., Ricardo, J. E., Hernández, N. B., & Leyva Vázquez, M. Y. (2020). Application of neutrosophic offsets for digital image processing. *Investigacion Operacional*, 41(5), 603-611
- Smarandache, F., Teodorescu, B., & Teodorescu, M. (2018). Uncertainty Communication Solution in Neutrosophic Key. *SSRN Electronic Journal*.
- Villafuerte, J. L. R., Torres, L. D. T., & Jimenez, L. T. (2020). Neutrosophic Hypothesis to validate a modification for Article 630 of the Integral Organic Criminal Code of Ecuador. *Neutrosophic Sets and Systems*, 37(1), 31.