

<https://doi.org/10.47300/actasidi-unicyt-2022-20>

MINERÍA DE TEXTO APLICADO AL ANÁLISIS DE SENTIMIENTOS PARA LA BÚSQUEDA DE PATRONES DE MOTIVACIÓN EN PROGRAMA DE ACTUALIZACIÓN DOCENTE

Meléndez Gómez, Nelly

Universidad Monteávila, Venezuela
Universidad Internacional de Ciencia y Tecnología, Panamá
nmelendez@uma.edu.ve
ORCID: 0000-0002-2780-2519

Briceño, Magally

Universidad Internacional de Ciencia y Tecnología, Panamá
magally.briceno@unicyt.net
ORCID: 0000-0003-4539-3103

Gibertoni, Jaime

Universidad Monteávila, Venezuela
jgiber197@gmail.com
ORCID: 0000-0002-4763-142X

Lucente, Rosina

Universidad Internacional de Ciencia y Tecnología, Panamá
rosina.lucente@unicyt.net
ORCID: 0000-0002-5041-723X

RESUMEN

El objetivo de esta investigación es analizar los sentimientos expresados por los profesores en torno a su participación en un programa de actualización docente. Para ello se utilizó la minería del texto como un proceso de investigación que permite el manejo automático de datos a partir de lenguaje natural, a fin de identificar patrones, descubrir relaciones, tendencias, desviaciones y obtener nueva información para la toma de decisiones en diferentes ámbitos, entre ellos los programas educativos. El estado del arte de este objeto de estudio partió de una conceptualización de la minería de texto y la aplicación de la inteligencia artificial, debido a los aportes significativos que han brindado a la educación y al proceso de análisis de los sentimientos. Se parte del supuesto que se puede proponer la minería de datos para el estudio de percepciones motivacionales en programas de capacitación. La metodología utilizada fue cuantitativa y cualitativa, el análisis del texto se realizó mediante el método *AnalyzeSentiment* (Google, 2022). Los resultados demuestran que el análisis de las opiniones emitidas por los docentes, atendiendo a preguntas intencionales, puede analizarse para generar un enfoque eficaz en programas de formación. Se concluye que la combinación de métodos de inteligencia artificial para el análisis de lenguaje natural, junto con técnicas de estadística descriptiva y análisis del discurso constituye un enfoque viable para el análisis de datos y sentimientos para mejorar la transmisión de mensajes que se utilizan en el aula de clase.

Palabras clave: Motivación, Minería de texto, Sentimientos, Actualización docente.

ABSTRACT

The objective of this research is to analyze the feelings expressed by teachers about their participation in a teacher update program. For this, text mining was used as a research process that allows the automatic handling of data from natural language, to identify patterns, discover relationships, trends, deviations and obtain new information for decision-making in different areas, including educational programs. The state of the art of this object of study started from a conceptualization of text mining and the application of artificial intelligence, due to the significant contributions they have provided to education and the process of sentiment analysis. It is assumed that data mining can be proposed for the study of motivational perceptions in training programs. The methodology used was quantitative and qualitative, the analysis of the text was conducted using the AnalyzeSentiment method (Google, 2022). The results show that the analysis of the opinions expressed by teachers, attending to intentional questions, can be analyzed to generate an effective approach in training programs. It is concluded that the combination of artificial intelligence methods for natural language analysis, together with descriptive statistics techniques and discourse analysis constitutes a viable approach to the analysis of data and feelings to improve the transmission of messages that are used in the classroom.

Keywords: Motivation, Text mining, Feelings, Teacher update.

1. INTRODUCCIÓN

El objetivo de la investigación fue analizar los sentimientos expresados por los profesores en torno al programa de actualización docente. Para ello se utilizó la minería del texto, que permite el procesamiento automático de datos a partir del texto e identificar patrones, descubrir relaciones, tendencias, desviaciones y obtener nueva información para pasar al descubrimiento de conocimientos y toma de decisiones en diferentes ámbitos uno de ellos son los programas educativos.

Este estudio es relevante en el sistema educativo, específicamente en lo que se refiere al análisis de programas, cursos de capacitación, evaluación de los docentes porque permite que, a partir del texto, las expresiones, imágenes se pueda inferir lo que estos piensan sobre un programa o una situación educativa en particular.

El estado del arte de este objeto de estudio lo constituyó las investigaciones realizadas por investigadores reconocidos en este campo de estudio. Se constató que la minería de datos puede ser aplicada en organizaciones orientadas a la gestión del conocimiento. Tiene grandes aplicaciones por cuanto es una herramienta que identifica patrones, descubre relaciones y hace afirmaciones basadas en patrones determinados.

Los datos se obtuvieron directamente de los foros de un diplomado b-learning dirigido al desarrollo de competencias digitales multicanal en docentes de distintos niveles educativos. Para ello, se hizo una pregunta abierta que permitiera extraer el sentimiento de los docentes. Esta quedó planteada así ¿Qué te motiva para hacer este diplomado?".

La investigación se organiza en las siguientes partes: Resumen, Introducción, Marco Conceptual, Materiales y Métodos, Resultados y Discusión y Conclusiones. Finalmente se presentan las referencias bibliográficas.

2. MARCO CONCEPTUAL

Este estudio es de enfoque interdisciplinario debido a que se aplica en diversos ámbitos del conocimiento y como método permite generar nueva información y conocimientos. En la literatura encontramos autores e investigaciones que han venido trabajando sobre el tema.

Al respecto mencionaremos a Justicia de la Torre (2017) plantea que la minería de textos es un nuevo paradigma basada en el desarrollo del conocimiento diferente al concepto tradicional referente a bases de datos.

Witten et. al (2004), la minería de texto es el proceso de analizar escritos o conjuntos de enunciados para extraer información que resulta útil para propósitos particulares.

Sánchez y Martín-Bautista (2014) plantean que la minería de texto hace referencia al descubrimiento no trivial potencialmente útil de conocimiento partiendo de una colección de documentos de texto no estructurado. Y puede realizarse una analogía con la minería de datos, encargada de descubrir conocimiento en bases de datos.

Sukanya y Biruntha (2012) plantean que la minería de texto es un campo interdisciplinario joven el cual se basa en la recuperación de información, minería de datos, aprendizaje de máquina, estadística y lingüística computacional.

Los conceptos señalados permiten indicar que la minería de textos es un campo reciente en análisis del texto y extracción de conocimiento. No existe un concepto único sobre este objeto de estudio y menos aún acerca de las técnicas o instrumentos que pueden utilizarse para la extracción del conocimiento previamente desconocido a partir del texto escrito.

Es importante además reconocer la importancia que tiene el uso de la inteligencia artificial en la minería de texto por cuanto se nutre de datos y los emplea para desarrollar algoritmos y para constituir la 'lógica' de las máquinas. En definitiva, utiliza los datos para obtener información del entorno e interactuar con él en consecuencia.

Saint-Pierre (2018) indica que la inteligencia artificial (IA) es aquella que se puede desarrollar a partir de la programación de computadoras o máquinas para que imiten las funciones "cognitivas", como «aprender» y «resolver problemas» propias de los seres humanos, por ejemplo, reconocer el habla humana (Siri), apoyar la conducción de autos autónomos (Tesla, Google car) y reconocer objetos y personas en imágenes y vídeos (Facebook).

Señala el autor antes indicado, que una de las grandes aplicaciones de la Inteligencia Artificial a la minería es la automatización de sus procesos y la toma de decisiones inteligente en base a la data histórica.

3. MATERIALES Y MÉTODOS

El tipo de investigación aplicada fue no experimental, dado que no hubo manipulación de las variables durante el proceso de investigación (Hernández et. al, 2006). Los datos usados para análisis se tomaron directamente de los foros de un diplomado e-learning dirigido al desarrollo de competencias digitales multicanal en docentes de distintos niveles educativos.

El diseño de la investigación se considera mixto, porque combina elementos cualitativos y cuantitativos derivados de la aplicación de técnicas de minería de datos y el análisis del contenido semántico, para lo cual se utilizó *Cloud Natural Language*, perteneciente a la empresa Google y específicamente se aplicó la técnica *AI Natural Language*, para obtener desde el motor de aprendizaje automático estadísticas del texto no estructurado.

El total de los participantes fue de 23 docentes (20 mujeres y 3 hombres; edad promedio = 45 años; media de desempeño docente = 10,28 años), que laboran en cinco (5) instituciones educativas venezolanas ubicadas en los estados Zulia, Carabobo y Distrito Capital. Los niveles educativos de trabajo fueron: preescolar (3), primaria (7), secundaria (7) y universitarios (3).

El análisis de sentimiento se llevó a cabo utilizando el modelo preentrenado proporcionado por Google en forma de prueba gratuita. Tras realizar el análisis de sentimientos, se obtiene un valor para cada respuesta que va de cero a uno; cuanto más cercano a uno, más positivo es el sentimiento. La pregunta abierta de la que se extrajo el sentimiento de los docentes fue: "¿Qué te motiva para hacer este diplomado?". Las respuestas a esta pregunta son sometidas a técnicas de *Machine Learning* (ML) para lenguaje natural, donde el software "inspecciona el texto dado e identifica la opinión emocional predominante, clasificándola como positiva, negativa o neutra. El análisis de sentimientos se ejecuta mediante el método *analyzeSentiment*" (Google, 2022), que presenta en cuadros identificados con distintos colores según grado y la magnitud. Este rango o

escala tiene los siguientes valores: 0.25 a 1.00 (verde), 0.25 a -0.25 (amarillo) y -0.25 a -1.0 (rojo). (<https://cloud.google.com/natural-language>)

Como criterios para la inclusión o exclusión de expresiones se tomó en cuenta el valor en términos numéricos de manifestación de positiva de sentimientos, haciendo énfasis en los valores más altos de 0.7 en adelante.

El análisis de resultados se realizó desde el punto de vista cuantitativo y cualitativo.

4. RESULTADOS Y DISCUSIÓN

Los resultados del análisis aplicando la herramienta *analyzeSentiment*, arrojan expresiones valoraciones de carga emotiva en la escala con puntajes desde -0.25 hasta 1.0, tal como se indicó con anterioridad.

El procesamiento cuantitativo es corroborado con el análisis cualitativo de las expresiones que surgen de surgen cuatro (4) categorías derivadas del texto. Las categorías resultantes son: Autoexpresión, Percepciones del futuro, Ser docente y Facilitador, esta última corresponde a los altos grados de expresiones positivas del docente.

Análisis cuantitativo

En la Tabla 1 puede observarse el análisis estadístico realizado a las expresiones de mayor carga de sentimiento positivo de los participantes.

Tabla 1

Análisis cuantitativo de los datos por categoría

Categoría	N (cantidad)	Promedio	Moda	Desviación
Autoexpresión	19	0.863	0.90	0.054
Percepciones del futuro	11	0.800	0.90	0.109
Ser docente	14	0.857	0.90	0.067
Facilitador	10	0.870	0.90	0.048

Fuente: elaboración propia (2022).

En el análisis de estos resultados se observó que los docentes, en la mayor parte de las expresiones positivas de alto nivel corresponden a las categorías de Autoexpresión y la retroalimentación del Facilitador, a partir de las participaciones. Por tal razón, los promedios en estas categorías son mayores, con una desviación estándar más baja y poco significativa.

La categoría Ser docente ocupa un tercer lugar en el promedio de manifestación de expresiones positivas, se incluyen aquí la percepción del rol docente, consideraciones acerca de ese trabajo, hábitos de trabajo y la resiliencia como característica presente, en el desempeño cotidiano.

En cuanto a la categoría Percepciones del futuro, se evidencia una carga de sentimientos positivos de más bajo nivel con relación a las otras categorías. Esto se relaciona positivamente con el evidente deterioro de la educación en Venezuela y las dificultades para ejercer la profesión docente (Reto País, 2021; ONU, 2022). No obstante, los resultados se ubican en 0.8, lo cual es altamente positivo.

Análisis Cualitativo

La sustentación cualitativa se basa en el análisis de las expresiones de los participantes con relación a su participación en el programa de capacitación. Se utilizó el software Atlas Ti (versión 22) para realizar el análisis de los datos.

En la Tabla 2 puede observarse las familias de códigos derivados del análisis de expresiones de los docentes.

Tabla 2.

Resumen de la codificación derivada del análisis discursivo de los docentes

Códigos	N (cantidad)	Promedio (%)
Competencias blandas	9	16.07
Cualidades personales	10	17.86
Deseos de aprender	7	12.50
Efectos de la pandemia	3	5.36
Percepción del futuro	6	10.71
Vocación docente	11	19.64
Sentimientos	3	5.36
Tecnologías	7	12.50

Fuente: elaboración propia (2022).

Entre las expresiones verbales, por citar un ejemplo, se encuentran aspectos relacionados con la vocación docente: "Me encanta aprender, enseñar y cada día me doy cuenta de que nací para esta vocación", "aprender muchísimo y así poder aplicarlo a mi día a día", "me ha tocado aprender cosas que nunca pensé hacer y me encanta" y "Siempre me ha gustado compartir mis conocimientos".

5. CONCLUSIONES

La herramienta analyzeSentiment, arrojó expresiones y valoraciones de carga emotiva en la escala con puntajes desde -0.25 hasta 1.0. Este procesamiento cuantitativo es corroborado con el análisis cualitativo de las opiniones que surgen de cuatro (4) categorías derivadas del texto: Autoexpresión, Percepciones del futuro, Ser docente y Facilitador.

Se pudo corroborar que la mayor parte de las muestras positivas en alto nivel corresponde a las categorías de Autoexpresión y la retroalimentación del Facilitador partir de las participaciones, mientras que la categoría Ser docente ocupa un tercer lugar en el promedio de manifestación de expresiones positivas, en ello se incluyen la percepción del rol docente, consideraciones acerca de ese trabajo, hábitos de trabajo y la resiliencia como características presente sen el desempeño cotidiano.

En cuanto a la categoría Percepciones del futuro, se evidencia una carga de sentimientos positivos de más bajo nivel de sentimientos con relación a las otras categorías. Los resultados se ubican en 0.8, lo cual es altamente positivo.

Se evidenció además que el procesamiento automático de datos a partir del texto permite identificar patrones, descubrir relaciones, tendencias, desviaciones y obtener nueva información para pasar al descubrimiento de conocimientos y toma de decisiones en diferentes ámbitos de la educación.

En consecuencia, se demostró el valor académico de este tipo de análisis por cuanto permite a los docentes poseer indicadores cuantitativos acerca de las opiniones emocionales de los grupos de estudiantes y, por ende, seleccionar las estrategias pedagógicas idóneas para la conducción de los procesos de enseñanza y de aprendizaje

REFERENCIAS

- Google. (2022). *Google Cloud*. Obtenido de <https://cloud.google.com/natural-language/docs/analyzing-sentiment>
- Hernández et al., (2006). *Metodologías de la investigación* (5ta. edición ed.). México, D.F.: McGraw-Hill Interamericana.
- Justicia de la Torre, M. (2017). Nuevas técnicas de minería de texto. Universidad de Granada. Disponible en: <https://digibug.ugr.es/bitstream/handle/10481/46975/26606203.pdf?sequence=1>
- Moreno-Marcos PM, et al., (2019). A Learning Analytics Methodology for Understanding Social Interactions in MOOCs. *IEEE Trans Learn Technol.* 2019;12(4):442-455. Disponible en: https://e-archivo.uc3m.es/bitstream/handle/10016/31149/learning_ITLT_ps_2019.pdf;jsessionid=54760D640107DF1E07BFD88BC8B178DE?sequence=1
- Obtenido de <https://reports.unocha.org/es/country/venezuela-bolivarian-republic-of/#cf-YIFCv0f4IMbfPUZ40b4OW>
- ONU. (2022, October 13). Venezuela - Mayo-Junio 2022. Estatus del cluster: Educación. <https://reports.unocha.org/es/country/venezuela-bolivarian-republic-of/card/YIFCv0f4IM/>
- Reto País. (2021 de Noviembre de 2021). *10 datos que revelan la crisis de la educación básica y media en Venezuela*. Obtenido de <https://retopaisvenezuela.org/10-datos-que-revelan-la-crisis-de-la-educacion-basica-y-media-en-venezuela/>
- Saint-Pierre, T. (2018). ¿Qué es la inteligencia artificial y cómo se puede aplicar en la Minería?. Fundación Chile. Disponible en: <https://fch.cl/noticianoticia-destacadanoticia-antigua/que-es-la-inteligencia-artificial-y-como-se-puede-aplicar-en-la-mineria/>
- Sánchez, D., Martín-Bautista, M. Un enfoque deductivo para la minería de texto [en línea]. <<http://www.softcomputing.es/estylf08/es/2006-XIII%20Congreso/articulos/40.pdf>> [06 febrero 2014].
- Sukanya, M. y - Biruntha, S. (2012) Techniques on text mining. *IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)* Disponible en: https://www.researchgate.net/publication/261317364_Techniques_on_text_mining/citation/download
- Witten, I.H. et al., (2004). Text mining in a digital library. *International Journal on Digital Libraries*, 4(1), 56-59. Disponible en: <https://researchcommons.waikato.ac.nz/handle/10289/1298>

i

ⁱ Los autores del trabajo autorizan a la Universidad Internacional de Ciencia y Tecnología (UNICYT) a publicar este resumen en extenso en las Actas del Congreso IDI-UNICYT 2022 en Acceso Abierto (Open Access) en formato digital (PDF) e integrarlos en diversas plataformas online bajo la licencia CC: Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

La Universidad Internacional de Ciencia y Tecnología y los miembros del Comité Organizador del Congreso IDI-UNICYT 2022 no son responsables del contenido ni de las implicaciones de lo expresado en este artículo.