



New Trends in
Qualitative
Research



VOLUME 19

Qualitative Research in Social Science

Investigação Qualitativa em Ciências Sociais

Investigación Cualitativa en Ciencias Sociales

DOI:

<https://doi.org/10.36367/ntqr.19.2023.e927>

Julio César Montiel Flores

Juan Pedro Valenzuela Valenzuela

Data de submissão: 03/2023

Data de avaliação: 04/2023

Data de publicação: 09/2023

MÍNERIA DE DATOS PARA UNA ANÁLISIS DEL SENTIMIENTO: ESTUDIO DE CASO

RESUMEN

Los usuarios de internet en todo el mundo pasan un promedio de 2 horas y 25 minutos al día en las redes sociales, por lo que técnicas como el análisis del sentimiento son aplicadas a sus redes sociales, la cual es utilizada para la identificación y clasificación de las emociones empleadas en actividades como investigación de mercados, monitoreo de la reputación de marca, medición de la satisfacción del cliente, entre otros. Esta comunicación pretende determinar la importancia y utilidad del análisis de sentimiento a través de la minería de datos como una herramienta relevante para conocer tendencias de opinión y sentimiento de referentes comunicacionales. La metodología utilizada es de corte cualitativo bajo el método del estudio de caso. Se utilizó Twitter como fuente para la recopilación de información y para su análisis se empleó el software de minería de datos RapidMiner con el operador para análisis de sentimiento. Como parte de las conclusiones, se observa que es posible afirmar que el análisis del sentimiento a través de la minería de datos es una herramienta útil y valiosa para la toma de decisiones sobre las estrategias comunicacionales que pueden llevarse a cabo, no solo en el turismo sino en cualquier actividad que busque posicionarse en la mente del consumidor.

Palabras-Clave

Percepción Social; Análisis de Sentimiento; Minería de datos; Redes Sociales;

DATA MINING FOR SENTIMENT ANALYSIS: A CASE STUDY

Abstract

Internet users around the world spend an average of 2 hours and 25 minutes a day on social networks, so techniques such as sentiment analysis are applied to their social networks, which is used for the identification and classification of the emotions used in activities such as market research, brand reputation monitoring, customer satisfaction measurement, among others. This communication aims to determine the importance and usefulness of sentiment analysis through data mining as a relevant tool to know opinion trends and sentiment of communication referents. The methodology used is qualitative under the case study method. Twitter was used as a source for the collection of information and for its analysis the data mining software RapidMiner was used with the operator for sentiment analysis. As part of the conclusions, it is observed that it is possible to affirm that the analysis of sentiment through data mining is a useful and valuable tool for decision-making on communication strategies that can be carried out, not only in tourism, but in any activity that seeks to position itself in the mind of the consumer.

Keywords

Social Perception; Sentiment Analysis; Data mining; Social Media;

1. Introducción

1.1 La minería de datos en estudios cualitativos

La minería de datos tiene sus raíces en la estadística y la informática, y ha evolucionado a lo largo del tiempo con el surgimiento de nuevas tecnologías y la creciente disponibilidad de datos; el término "minería de datos" se utilizó por primera vez en la década de los 90's por investigadores de IBM (Berry y Linoff, 1997), describiéndola como un proceso de exploración y análisis de grandes cantidades de datos, con el objetivo de descubrir patrones significativos y reglas de asociación. Los avances en tecnología digital han ampliado las opciones de promoción para llegar a los consumidores, utilizando espacios importantes en la actualidad, siendo internet el medio por el que se permite la distribución de información relevante sobre los productos y servicios que ofrecen las empresas para darlos a conocer (Montiel et al., 2017). Con el aumento del uso de Internet y las redes sociales en la década de los 2000, esta se convirtió en una técnica popular para analizar grandes conjuntos de datos en línea, siendo los investigadores quienes la utilizaron para analizar los patrones de comportamiento de los usuarios en internet y con ello personalizar la publicidad en línea (Provost y Fawcett, 2013).

La minería de datos puede ser importante en la investigación cualitativa ya que ayuda a descubrir patrones y tendencias, exploración de datos y validación de hallazgos (Caruana y Niculescu-Mizil, 2006; Kelle y Erzberger, 2010; Marshall y Rossman, 2016), es por ello, que esta es una técnica de gran utilidad para realizar el análisis de grandes conjuntos de datos en la investigación cualitativa, ya que permite a los investigadores encontrar patrones y relaciones en los datos, lo que puede ayudar a hacer inferencias y generar hipótesis para futuras investigaciones.

Para realizar la minería de datos se han generado diferentes herramientas, por ejemplo, RapidMiner la cual es una plataforma de software de minería de datos y aprendizaje automático de código abierto que permite a los usuarios construir y aplicar modelos predictivos a través de una interfaz gráfica de usuario intuitiva, convirtiéndose en una de las herramientas de minería de datos y aprendizaje automático más populares debido a su facilidad de uso y sus capacidades de análisis avanzadas (Kotsiantis et al., 2006a). Dicha herramienta ofrece beneficios a los usuarios como ahorrar tiempo y esfuerzo en el proceso de modelado predictivo, ofreciendo una amplia gama de técnicas de minería de datos, permitiendo elegir la mejor opción para sus necesidades (Kotsiantis et al., 2006b; Li et al., 2017).

1.2 El caso de estudio

El estado de Guanajuato es una de las 32 entidades federativas de México, colinda al norte con los estados de Zacatecas y San Luis Potosí, al sur con Michoacán, al este con el estado de Querétaro y al oeste con el estado de Jalisco; el estado de Guanajuato cuenta con una extensión de 30,491 km², que conforman sus 46 municipios, es un estado que se caracteriza por sus edificios culturales de la época de la Nueva España, recintos religiosos, monumentos históricos, paisajes, riquezas minerales, pueblos fantasmas, productos gastronómicos y de piel;

adicionalmente, es referente de un sinfín de influencias artísticas, desde la indígena hasta la estadounidense, pasando por la española de la época de la conquista, debido a su carácter cosmopolita, en todo el estado se pueden encontrar restaurantes internacionales y algunos especializados en determinados países (MéxicoReal, 2012).

Fernando Olivera Rocha, ex secretario de Turismo de Guanajuato y actual Vicepresidente de Turismo CANIRAC (Cámara Nacional de la Industria de Restaurantes y Alimentos Condimentados), mencionó en una intervención con el tema: Turismo local, destino turístico sin playa de mayor relevancia, comentó: "Guanajuato es un destino sin playa que ofrece grandes ventajas competitivas y una política pública establecida en el Programa Estatal de Turismo y en base a ello, se trabaja a diario para subsistir en la calidad de los servicios y la fisonomía de la Entidad" (Gobierno del Estado de Guanajuato, 2015), afirmaciones que respaldan los datos del observatorio turístico del Estado de Guanajuato quien reporta una inversión privada en el turismo desde el 2013 hasta el 2021 de más de 18, 665 millones de pesos, más de mil millones de dólares, para el cierre del año 2021 se registraron 319,394 visitantes extranjeros en el estado y hasta junio del 2022 se habían registrado 215,600 llegadas de pasajeros al aeropuerto del estado, solo en 2021 se registró un total de 35,481 millones de pesos como derrama económica en dicho sector, casi dos mil millones de dólares (OTEG, 2022).

Además, la Secretaría de Turismo del Estado de Guanajuato en su reporte de Indicadores de la Actividad Turística, en la versión diciembre 2022, comunicaron el registro de un millón 765 mil visitantes al estado, generando en dicho mes una derrama económica de 3 mil 827 millones de pesos, más de 200 millones de dólares, acumulando entre enero y diciembre del 2022, más de 45 mil millones de pesos (más de dos mil millones de dólares); durante el tercer trimestre de 2022, en el Estado de Guanajuato el número de personas trabajando en el rubro Restaurantes y Servicios de Alojamiento fue de 195 mil 830 personas, ocupando el quinto lugar en el Ranking Nacional de Empleo del rubro, siendo a nivel nacional 1 de cada 13 personas que trabajan en la actividad de Restaurantes y Servicios de Alojamiento del estado de Guanajuato (DATATUR, 2023).

1.3 Análisis del sentimiento

De forma general la Real Academia Española define la percepción como la "sensación interior que resulta de una impresión material producida en los sentidos corporales" (Real Academia Española, 2023), desde la visión de la psicología, principal disciplina encargada del estudio de esta, es definida como un proceso cognitivo (mental) sobre la conciencia que reconoce, interpreta y da significado en la generación de juicios sobre las sensaciones que son obtenidas por el ambiente físico y social, dentro del cual se ven involucrados procesos psíquicos como el aprendizaje, la memoria y la simbolización (Vargas, 1994).

Dentro del proceso de percepción ocurren tres momentos los cuales mencionan Bassan y Rodríguez (2020): el primero, de carácter sensorial; el segundo, de carácter selectivo y el tercero, de carácter interpretativo, en el cual se da significado al objeto percibido; dicha percepción, en este caso, trasladada al fenómeno turístico, puede ser percibido a través de una gran cantidad de variables, que por mencionar algunas, puede ser considerada la idiosincrasia de cada pueblo o lugar, su componente étnico-poblacional, el nivel de

relación comercial con el turismo, el aporte económico a la economía vernácula, los preconceptos existentes, las experiencias pasadas o de localidades vecinas, etc.

El turismo es un fenómeno social, cultural y económico que supone el desplazamiento de personas a países o lugares fuera de su entorno habitual por motivos personales, profesionales o de negocios (OMT, 2023). El turismo es un fenómeno muy susceptible a lo que sucede en su entorno, por ejemplo, Rodríguez-Toubes y Álvarez (2023) mencionan que el descenso de las llegadas de turistas está relacionado con la reacción de los turistas a la percepción de amenazas a su salud, su seguridad personal o, simplemente, a su disfrute en las vacaciones. Riveros y Gamarra (2018) enfatizan como las actitudes y percepciones de los grupos o individuos frente al proceso turístico y al visitante, pueden ser positivas o negativas, activas o pasivas, de manera que las reacciones pueden tomar la forma de una de estas reacciones, o bien, coexistir una combinación de estas.

Bo Pang y Lillian Lee son autores relevantes en el campo del análisis del sentimiento y han contribuido en gran manera al desarrollo de técnicas de procesamiento de lenguaje natural para la identificación de sentimientos, su trabajo destaca la importancia del análisis del sentimiento en una amplia variedad de aplicaciones y la necesidad de técnicas precisas y eficaces para la identificación de emociones expresadas en el lenguaje humano, definiendo el análisis del sentimiento como una técnica utilizada para la identificación y clasificación de las emociones expresadas en un texto en tres posturas: positivas, negativas o neutrales, dicha técnica ha sido empleada en una gran variedad de aplicaciones, en las que se incluye la investigación de mercado, el monitoreo de la reputación de la marca, la medición de la satisfacción del cliente y la identificación de tendencias en las opiniones y actitudes de los consumidores (Pang y Lee, 2008)

Según el informe de We Are Social UK (2021), los usuarios de internet en todo el mundo pasan un promedio de 2 horas y 25 minutos al día en las redes sociales, por ello no es de extrañarse que el análisis de sentimientos en redes sociales sea una técnica cada vez más utilizada para comprender cómo las personas se sienten acerca de marcas, productos, un evento o un tema específico en internet, ya que dicha técnica está basada en el procesamiento del lenguaje natural y la minería de datos para identificar y clasificar las emociones expresadas en los textos de las redes sociales.

Es importante resaltar que los contenidos que se publican en las redes sociales son muestra de las apropiaciones que las personas realizan a través de sus experiencias y constituye un conjunto muy diverso de formas de participar con opiniones en temas de interés.

2. Metodología

Esta investigación se basa en la tradición de la metodología cualitativa que como Quecedo y Castaño (2003) la justifican, representa una investigación que genera datos descriptivos a partir del análisis de las propias palabras de las personas, habladas o escritas y de la conducta observable.

La investigación cualitativa es de carácter inductivo (Taylor y Bogdan en Quecedo y Castaño, 2003) pues permite comprender y desarrollar conceptos tomando en cuenta las pautas de los datos; utilizar un diseño flexible y partir de un conjunto de interrogantes que pueden o no sean planteadas vagamente. La investigación cualitativa por tanto se enfoca en la comprensión de las personas en su contexto y así observar cómo experimentan la realidad.

El método de investigación utilizado para esta investigación fue el estudio de caso que Yin (2014) explica es aquel que busca entender el fenómeno en su complejidad a través de la recolección rigurosa y el análisis de una amplia variedad de fuentes de datos, siendo una estrategia de investigación empírica valiosa para entender la complejidad y la riqueza de los fenómenos en su contexto, útil para investigar situaciones en las que no es posible controlar las variables de interés o en las que el investigador tiene acceso limitado a la situación, reconociendo su potencial para generar teorías, más que para probar hipótesis generalizables.

Para la recolección de la información obtenida se utilizó la plataforma de Twitter que Restrepo, García y Arboleda (2020) la define como una red social de alcance mundial que permite suministrar información en tiempo real y la interacción entre personas en relación con una temática específica. Según estos autores, corresponde a una red enfocada en microblogging que opera bajo publicaciones limitadas a 280 caracteres; fue creada en Estados Unidos por Jack Dorsey en el año de 2006.

Aunque Twitter por número de usuarios no es una red con el alcance de otras como Facebook, Whatsapp e Instagram, Twitter es utilizada por 556 millones de usuarios en el mundo (Statista, 2023). Por otro lado, Restrepo et al. (2020) destacan la importancia de Twitter como una herramienta de alta utilidad y que emplean con mayor interés personas del campo científico ya que esta red proporciona una comunicación más concreta entre distintos participantes.

Para la extracción de datos desde la red social de Twitter y su análisis se utilizó el software de Minería de Datos de RapidMiner, ya que este puede procesar datos de diferentes fuentes, como archivos de texto, bases de datos, hojas de cálculo y servicios web, entre otros; la minería de datos es una tecnología y estrategia basada en un modelado matemático que permite comprender el contenido de una base de datos y así transformar esos datos en información útil (Gutiérrez y Molina, 2014); por otro lado, la minería de datos comprende un conjunto de metodologías, aplicaciones y tecnologías que facilitan la colección, depuración y transformación de datos que no se encuentran estructurados, a fin de que faciliten su análisis y conversión en conocimiento (Marcano y Talavera, 2007).

Para lograr la conexión con Twitter desde RapidMiner se gestionó el acceso mediante una cuenta de usuario y un código que asegura la autenticación para su uso.

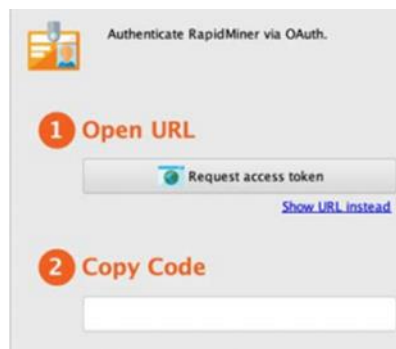


Figura 1. Autorización para la conexión a Twitter desde RapidMiner



Figura 2. Autorización de cuenta de Twitter desde RapidMiner

Una vez hecha la conexión directa desde ahí; se procedió a la extracción de tweets con la cantidad de 10 mil de ellos bajo el texto clave: "Turismo Guanajuato" como se aprecia en la figura número 3.

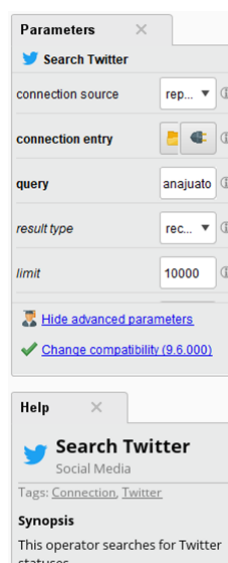


Figura 3. Parámetros para la búsqueda de Tweets desde RapidMiner

Se obtuvo una base de datos en Excel con los resultados de la búsqueda realizada. La búsqueda realizada en Twitter arrojó de 678 tweets bajo la categoría de "recientes o populares", esto representa que no se alcanzaron a captar los 10 mil tweets de la búsqueda inicial. Se procedió a limpiar la base de datos extrayendo caracteres que no proporcionan información para este análisis. Posteriormente se realizó el análisis de sentimiento sobre la base de datos resultante con el software de RapidMiner, tal como se muestra en el modelo de proceso de la Figura 4.

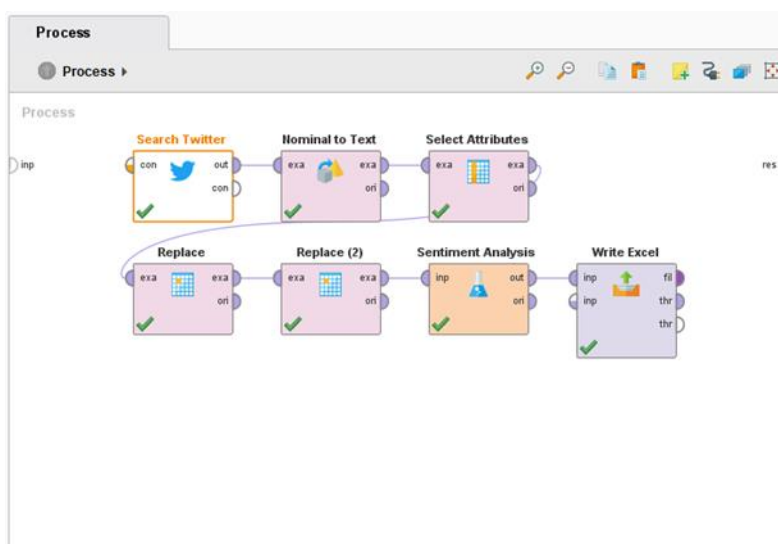


Figura 4. Modelo de proceso para análisis de sentimiento en RapidMiner

Se extrajeron determinados atributos de la información original arrojada en Twitter como: Geo-Location-Latitude, Geo-Location-Longitude, Id, Retweet-Count y Text. Los primeros dos parámetros refieren a la ubicación geográfica de donde se genera el tweet; luego se tiene el número que identifica al texto; Retweet-Count refiere al número total de veces que se volvió a publicar el tweet original; y finalmente el texto del tweet. La extracción de parámetros se muestra en la figura 5.

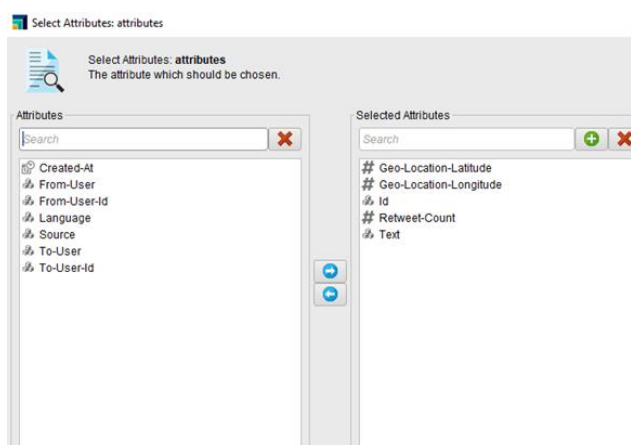


Figura 5. Atributos extraídos de la información original de Twitter para análisis de sentimiento en RapidMiner

Para este caso en particular se utiliza un modelo muy sencillo de análisis de sentimiento que puede ser escalado, pero los límites exigidos de la extensión de esta comunicación han impedido un análisis más extenso; sin embargo, se subraya la importancia de la minería de datos como una herramienta muy práctica para realizar aproximaciones a la comprensión de la opinión pública sobre determinados temas.

3. Resultados y Discusión

Como ya se ha mencionado, la búsqueda arrojó un total de 678 tweets como recientes o populares. En la figura 6 se muestra un extracto de la base de datos resultante.

	A	B	C	D	E	F	G	H	I	J
1	Text	Geo-Location-Latitude	Geo-Location-Longitude	Retweets	Id	polarity(Text)	confidence	agreement	subjectivity	irony(Text)
2	@AnaLiliaHerrera EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SAL			.0	163138711	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
3	@m_ebrard EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALMANT			.0	163138702	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
4	@PartidoMorenaMx EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO S			.0	163138688	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
5	@alitomorenoc @INEMexico @PRI_Nacional EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GU			.0	163138677	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
6	@alitomorenoc EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163138586	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
7	@MexicoenseTV @alfredodelmazo EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUA			.0	163138305	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
8	@Edomex EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALMANTIN			.0	163138272	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
9	@Sebas_RM @mario_delgado @LourdesPaz24 @A_QuinteroMX @AreliaCastilla2 EL FUNDADOR DEL			.0	163138272	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
10	@PRI_Nacional EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163138263	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
11	@LuisBNava @Cisco EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO S			.0	163138216	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
12	@kenialopez EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALMA			.0	163138186	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
13	@WeraReynoso EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163137292	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
14	@LauraBarreraF @melissavargasmx @Eduardo_Zarzosa @JazminJaimesA11 @acryd_50 EL FUNDADO			.0	163137015	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
15	@LauraBarreraF @melissavargasmx @Eduardo_Zarzosa @JazminJaimesA11 @acryd_50 EL FUNDADO			.0	163137004	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
16	@LauraBarreraF EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163136974	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
17	@LauraBarreraF @melissavargasmx @Eduardo_Zarzosa @JazminJaimesA11 @acryd_50 EL FUNDADO			.0	163136946	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
18	@LauraBarreraF EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163136913	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
19	@LauraBarreraF EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLO GUANAJUATO Y ORGULLO SALM			.0	163136875	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
20	@LauraBarreraF @PRI_Nacional @CristinaRuiz_5 EL FUNDADOR DEL ORGULLO MEXICANO, ORGULLC			.0	163136867	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC
21	@LauraBarreraF @melissavargasmx @Eduardo_Zarzosa @JazminJaimesA11 @acryd_50 EL FUNDADO			.0	163136812	P	100.0	AGREEMENT	OBJECTIVE	NONIRONIC

Figura 6. Extracto de la base de datos resultante

Adicionalmente, el resultado del proceso diseñado en RapidMiner genera otros campos para su análisis: Polarity (polaridad); confidence (confianza); agreement (acuerdo); subjectivity (subjetividad), irony (ironía). El campo al que se enfoca esta investigación corresponde al de polaridad ya que refiere a la evaluación del sentimiento que recibe valores como: Muy positivo, positivo, neutro, negativo, muy negativo y ninguno. Esto significa que cualquier texto expresado en el tweet puede tener una connotación negativa, positiva o neutra con algunas intensidades, o bien estar ausente.

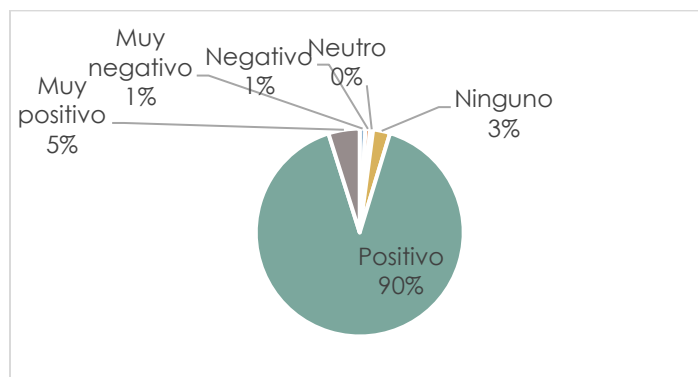


Figura 7. Polaridad resultante de análisis de tweets

De la figura anterior puede decirse que la percepción del turismo en el estado de Guanajuato, México, genera comentarios positivos en un 90% de acuerdo con el análisis de sentimiento aplicado a los tweets referidos en la categoría de “recientes o populares”, que en un sentido más general correspondería al 95% de ellos si se toma adicionalmente el valor de “muy positivo”.

La confianza refleja el grado en que es confiable el modelo de análisis de construido, para el caso que se presenta se obtuvo un nivel de confianza mayor e igual al 92% lo que significa que es confiable el modelo. Por otro lado, el resultado representa el acuerdo que existe entre los sentimientos encontrados en el texto analizado; el 99.4% de los sentimientos encontrados concuerdan en la polaridad.

En lo que refiere a la ironía, entendida como el hecho de comunicar algo cuando en realidad quiere decir otra cosa; el análisis del sentimiento arrojó que en el 100% de los tweets analizados no refleja ironía alguna. En tanto a la variable subjetividad referida a la valoración personal de lo expresado, los resultados mostrados identifican que el 94% de los textos son objetivos, es decir, refieren particularmente a expresiones inherentes al objeto de la comunicación. Lo anterior se refleja en la figura 8.

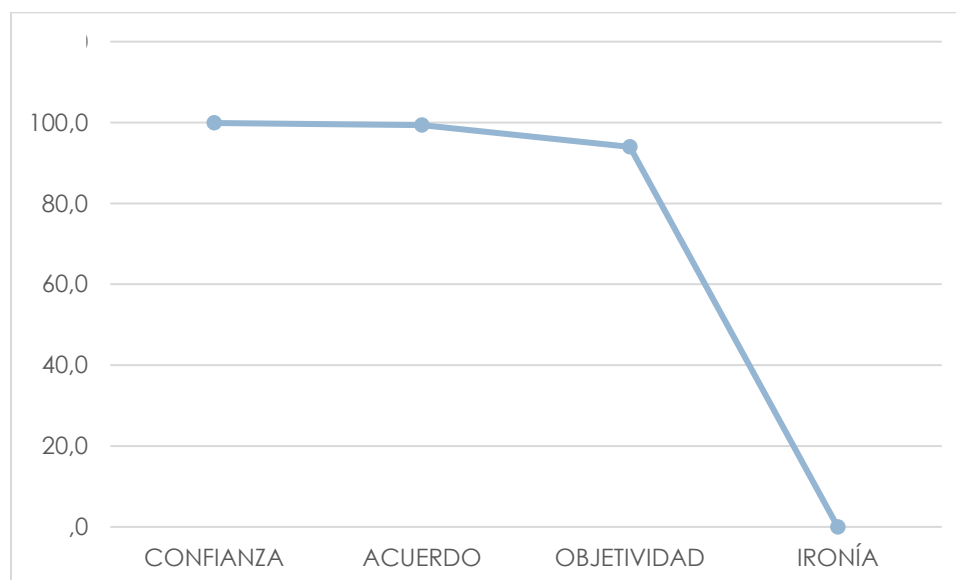


Figura 8. Otros resultados obtenidos, expresados en porcentaje

Entre los retweets con mayor frecuencia registrada en este análisis de sentimiento corresponden a los siguientes tweets:

Tabla 1. Publicaciones con mayor frecuencia en retweets.

Tweet	Retweets
#Guanajuato es la ciudad más bonita de México ??? Ven a conocerla y enamórate ❤️ ...	28.0
¿Ya viste la hora? 🕒 Es hora de que me invites un cantarito en Guanajuato Capital	10.0
para la promoción del #turismo 🦋, logramos reinventarnos ??y encontrar nues...	10.0
Hoy nos reconocen por nuestra labor y esfuerzo ??para la promoción del #turismo 🦋, logramos reinventarnos ??y encontrar nuestras	10.0

propias fortalezas?? como el destino favorito. ??? ¿Felicidades a todo el sector y al Consejo de Promoción Turística del Municipio de #GuANajuato! #Guanajuato consolida su presencia en @AnatoNacional con el mercado colombiano ??.	6.0
Año con año, la #SECTURGTO se presenta en... #GuANajuato empieza con G de Grandeza ?? y también de #Google porque muy pronto estaremos digitalizando el turismo ?? con el apoyo de los grandes gigantes de la tecnología. ??? Pendientes, porque les platicaré otras cosas que ando gestionado. Hasta aquí mi reporte. ?????	5.0
@cardonamex @lopezobrador_ @RicardoMonrealA @DavidMonrealA Rascale por igual, al estado mas violento, Guanajuato con fiscal...	4.0
@JJGueroAlvarez habla #AlAire del Turismo en #Guanajuato con @AltavozTurismo ?? @ElCastorViajes Escucha la entrevista ??	3.0
ó Expo Boda Capital en #Guanajuato ??.	3.0
Este año celebramos la 3ra. edición de este gran evento de Turismo de Romanc... Hoy inició Expo Boda Capital en #Guanajuato ??.	3.0
Este año celebramos la 3ra. edición de este gran evento de Turismo de Romance -que concluye el 26 de febrero- en una de nuestras Ciudades Patrimonio de la Humanidad. San Miguel de Allende es una ciudad del estado mexicano de Guanajuato (México). Es considerada la ciudad más bella de ese país y una de las mejores del mundo para el turismo. ¡Nos inspiran las ciudades!	3.0
@SECTURGTO fortalece alianzas estratégicas en la vitrina más importante de turismo: #ANATO2023 en #Colombia. Dicha...	3.0
alianzas estratégicas que promuevan la oferta turística de #Guanajuato, la #SECTURGTO partic...	3.0
Con el objetivo de fortalecer ?? alianzas estratégicas que promuevan la oferta turística de #Guanajuato, la #SECTURGTO participa en #ANATO2023, Feria de Turismo más importante de Colombia. @AnatoNacional es el punto de encuentro de sectores del turismo en #Colombia.	3.0

4. Consideraciones finales

El análisis de sentimiento es una herramienta que permite observar la actitud que asumen las personas que publican en internet sobre un tema en particular, para llevar a cabo este análisis se requiere de extraer la información pertinente; en este caso se realizó sobre la red social Twitter, una red que se caracteriza por su comunicación directa y concreta en pocos caracteres.

La polaridad de las publicaciones indica la actitud sobre el tema en referencia en valores como positivo, negativo, neutro o ausente, lo que permite identificar la empatía de las personas sobre los referentes de análisis. Esta información puede ser tomada en cuenta para el desarrollo de estrategias comunicacionales relevantes que coadyuven al posicionamiento

en la mente del consumidor de marcas, productos etcétera, pero también para identificar de qué lugar proviene la percepción.

El análisis de sentimiento utiliza sistemas para el procesamiento de opiniones y que a través de la minería de datos puede facilitar la comprensión de la actitud de los usuarios de internet sobre determinados temas y conocer así las tendencias para la toma de decisiones en diferentes campos de la comunicación.

5. Referências

Bassan, C. E., & Rodríguez, A. J. (2020). Estudio de percepción del turismo por parte de los residentes de la localidad de Barreal, departamento de Calingasta (San Juan). *Revista de Turismo e Identidad*, 1(1), Article 1. <https://revistas.uncu.edu.ar/ojs3/index.php/turismoeidentidad/article/view/4354>

Berry, M. J., & Linoff, G. (1997). *Data mining techniques: For marketing, sales, and customer support*. John Wiley & Sons.

[https://books.google.es/books?hl=es&lr=&id=Ni5nMDO1OfEC&oi=fnd&pg=PR19&dq=Berry,+M.+J.,+%26+Linoff,+G.+\(1997\).+Data+mining+techniques:+For+marketing,+sales,+and+customer+support.+John+Wiley+%26+Sons.&ots=vac3rpEMIM&sig=ndDs4y6j2MM_gzvXL2xAxGBKLEo](https://books.google.es/books?hl=es&lr=&id=Ni5nMDO1OfEC&oi=fnd&pg=PR19&dq=Berry,+M.+J.,+%26+Linoff,+G.+(1997).+Data+mining+techniques:+For+marketing,+sales,+and+customer+support.+John+Wiley+%26+Sons.&ots=vac3rpEMIM&sig=ndDs4y6j2MM_gzvXL2xAxGBKLEo)

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. *En Proceedings of the 23rd international conference on Machine learning* (pp. 161-168).

DATATUR. (2023). *Indicadores de la Actividad Turística en el Estado de Guanajuato-diciembre 2022* (pp. 1-8). Secretaria de Turismo de Guanajuato. http://www.observatorioturistico.org/cenDoc/8ad47-Boletin-Actividad-Turistica12_DIC2022.pdf

Gobierno del Estado de Guanajuato. (2015, septiembre 14). Guanajuato, mejor destino turístico sin playa": secretario de Turismo. *Boletines Dependencias*.

<https://boletines.guanajuato.gob.mx/2015/09/14/guanajuato-mejor-destino-turistico-sin-playa-secretario-de-turismo/>

Gutiérrez, J., & Molina, B. (2014). Identificación de técnicas de minería de datos para apoyar la toma de decisiones en la solución de problemas empresariales. *Revista Ontare*, 3(2), 33-51. <https://doi.org/10.21158/23823399.v3.n2.2015.1440>

Kelle, U., & Erzberger, C. (2010). *Qualitative and quantitative methods in research on gender and education*. Sense Publishers.

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006a). Predictive modeling of student performance in distance education using machine learning techniques. *Applied Artificial Intelligence*, 20, 411-426. <https://doi.org/10.1080/08839510600782643>

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006b). Predictive modeling of student performance in distance education using machine learning techniques. *Expert Systems with Applications*, 31(3), 820-829. <https://doi.org/10.1016/j.eswa.2005.10.029>

Li, J., Lu, R., & Chen, G. (2017). Distributed data processing with RapidMiner. *IEEE Transactions on Knowledge and Data Engineering*, 29(2), 464-477. <https://doi.org/10.1109/TKDE.2016.2619814>

Marcano, A., & Talavera, R. (2007). Minería de Datos como soporte a la toma de decisiones empresariales. *Opinión*, 23(52), 104-118. <https://www.redalyc.org/pdf/310/31005208.pdf>

Marshall, C., & Rossman, G. B. (2016). Marshall, C., & Rossman, G. B. (2016). Designing qualitative research. Sage publications. Sage Publications.

MéxicoReal. (2012, abril 18). Estado de Guanajuato en la República Mexicana. Mexico Real. <https://mr.travelbymexico.com/685-estado-de-guanajuato/>

Montiel, J. C., Lozano, M. C., & Díaz, I. A. (2017). Las prácticas de promoción digital de las vinícolas en el Valle de Guadalupe, B.C., México y su contribución en la construcción de la marca destino, una perspectiva desde las ciencias sociales y administrativas. TURYDES: Revista sobre Turismo y Desarrollo local sostenible, 10(22), 1-14. <https://www.eumed.net/rev/turydes/22/promocion-digital-bodegas.html>

OMT. (2023). Glosario de términos de turismo | OMT. <https://www.unwto.org/es/glosario-terminos-turisticos>

OTEG. (2022). Observatorio Turístico del Estado de Guanajuato. <http://observatorioturistico.org/>

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 1-2, 1-135. <https://doi.org/10.1561/1500000011>

Provost, F., & Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.

Quecedo, R., & Castaño, C. (2003). Introducción a la metodología de investigación cualitativa. Revista de Psicodidáctica, 14, 5-40. <https://www.redalyc.org/pdf/175/17501402.pdf>

Real Academia Española. (2023). Percepción [Versión 23.6 en línea]. Diccionario de la lengua española, 23.a ed. <https://dle.rae.es/percepcion>

Restrepo, L. F., García, G., & Arboleda, E. (2020). El poder de Twitter en la comunicación investigativa en las áreas de la educación, marketing y política. 13(2). <https://doi.org/10.12804/revistas.urosario.edu.co/disertaciones/a.8100>

Riveros, T. I., & Gamarra, F. K. (2018). Imagen de la actividad turística: Una percepción de los residentes. Arandu UTIC, 5(2). <http://www.utic.edu.py/revista.ojs/index.php/revistas/article/view/67>

Rodríguez-Toubes, D., & Álvarez, J. (2023). Vista de Vulnerabilidad del turismo y comunicación institucional ante desastres: Estudio de casos. REDMARKA, 6(137-161). <https://doi.org/10.17979/redma.2013.01.011.4810>

Statista. (2023). Usuarios mundiales de las redes sociales líderes en 2023. Statista. <https://es.statista.com/estadisticas/600712/ranking-mundial-de-redes-sociales-por-numero-de-usuarios/>

Vargas, L. M. (1994). Sobre el concepto de percepción. Alteridades, 4(8), Article 8. <https://doi.org/10.24275/uam/izt/dcsh/alteridades>

We Are Social UK. (2021, enero 27). Digital 2021. We Are Social UK. <https://wearesocial.com/uk/blog/2021/01/digital-2021-uk/>

Yin, R. K. (2014). Case study research: Design and methods (5th ed.). Sage Publications.

Julio César Montiel Flores

Universidad de Guanajuato, México

 <https://orcid.org/0000-0001-6891-1431>✉ jc.montiel@ugto.mx**Juan Pedro Valenzuela Valenzuela**

Universidad de Guanajuato, México

 <https://orcid.org/0000-0002-0744-0127>✉ jp.valenzuelavalenzuela@ugto.mx