CAPÍTULO V

MINERÍA DE TEXTOS Y ANÁLISIS DE REDES SOCIALES EN TWITTER

Natalia Niño Martínez

Profesional en Estadística. Profesional en Negocios Internacionales. Docente investigadora. Fundación Universitaria Unipanamericana. Correo electrónico: nninom@unipanamericana.edu.co.

Carolina Vaca

Profesional en Mercadeo y Publicidad. Especialista en Gerencia de Publicidad. Máster en Gestión y Dirección de Marketing. Docente investigadora. Fundación Universitaria Unipanamericana. Correo electrónico: yadiravaca@unipanamericana.edu.co.

Bibiana Paola Ríos

Profesional en Diseño Visual, Especialista en Gerencia de Diseño, Magíster en Gestión Cultural, Máster en Mercadeo del Arte y Gestión de empresas Relacionadas. Docente investigadora. Fundación Universitaria Unipanamericana. Correo electrónico: bpriosc@unipanamericana.edu.co.

Laura Jimena Rey

Profesional en Comunicación Social. Especialista en Gerencia Integral de Calidad. Especialista en Docencia Universitaria. Máster en Dirección de Marketing. Docente investigadora. Fundación Universitaria Unipanamericana. Correo electrónico: ljrey@unipanamericana.edu.co.

Resumen

En la actualidad, las plataformas digitales y las herramientas tecnológicas se emplean para fortalecer los procesos de relacionamiento externo de las entidades públicas con los medios de comunicación y con la población en general. Es por esto que, se realiza esta investigación; con el objetivo de evaluar el impacto de estas campañas desplegadas en la plataforma Twitter para la cuenta “@Ejercito_CAAID” la cual es administrada por la línea de Asuntos Públicos del Comando de Apoyo de Acción Integral y Desarrollo del Ejército Nacional de Colombia. A través de una metodología exploratoria, de naturaleza cualitativa y cuantitativa en la cual se usarán herramientas de minería de textos y análisis de sentimientos en redes sociales, se pretende identificar tendencias que permitan direccionar las estrategias de mercadeo para fortalecer la marca, y con ello las relaciones públicas de la organización, como por ejemplo el desarrollo de campañas que generen tendencia y permanezcan por periodos importantes de tiempo, y así realmente logran construir concepto y marcar, así como diferenciar en el sentir y visualizar de los usuarios.

Palabras clave: análisis de sentimientos, Fuerzas Militares de Colombia, impacto de marca, minería de textos, redes sociales.

TEXT MINING AND SOCIAL MEDIA ANALYSIS WHIT TWITTER DATA

Abstract

Currently, digital platforms and technological tools are used to strengthen the processes of external relations of public entities with the media and with the general population. This is why this research is carried out; with the aim of evaluating the impact of these campaigns deployed on the Twitter platform for the “@Ejercito_CAAID” account, which is administered by the Public Affairs line of the Support Command for Integral Action and Development of the Colombian National Army. Through an exploratory methodology, of a qualitative and quantitative nature in which text mining tools and sentiment analysis will be used in social networks, the aim is to identify trends that allow directing marketing strategies to strengthen the brand, and with it the public relations of the organization, such as the development of campaigns that generate trends and remain for important periods of time, and thus really manage to build concept and mark, as well as differentiate in the feeling and visualization of users.

Keywords: brand impact, Colombia’s Military Forces, sentiment analysis, social networks, text mining.

Este capítulo es el resultado de un proyecto de investigación del Grupo de Investigación en Ingenierías GIIS, en conjunto con el Grupo EGE titulado “Creación de propuestas de valor para las organizaciones con base en Tecnologías Digitales Emergentes”. Finalizado en el año 2019 y financiado por Unipanamericana Compensar Institución Universitaria, Bogotá, Colombia.

Introducción

La línea de Asuntos Públicos del Comando de Apoyo de Acción Integral y Desarrollo (CAAID), encargada de trabajar y fortalecer todos los procesos de relacionamiento externo con los medios de comunicación y con la población en general, para visibilizar las actividades y el trabajo de Acción Integral, enmarcado en la Política de Defensa y Seguridad del Gobierno Nacional, pone sobre la mesa la utilización de los nuevos medios y herramientas tecnológicas de la era digital, dándole al comando autonomía para el manejo de sus redes sociales propias y la generación de sus propias sinergias comunicativas, contenidos y estructuras de proyección de acuerdo a las necesidades que se vayan presentando.

Dicha propuesta, se enmarca en la evolución y transformación digital; que, durante los últimos años, ha facilitado la comunicación entre individuos y la creación de mecanismos para expresar ideas y opiniones a través de redes sociales, así mismo, se enfoca en la utilización

de la plataforma Twitter, para establecer una interacción entre el público en general y las acciones que lleva a cabo el Comando en pro de la sociedad, buscando visibilizar su accionar y con ello, generar un debate acerca de los asuntos públicos que le conciernen a esta entidad.

A pesar de ser una campaña activa, enfocada en el relacionamiento público de la entidad, no se conoce el efecto que está generando en los usuarios que activamente participan en la discusión acerca de la actuación diaria de los miembros del Comando; lo que lleva a esta investigación, a asumir como objetivo la evaluación del impacto de esta campaña en Twitter, valiéndose de la información contenida en los textos o comentarios relacionados con la cuenta del Comando y usando técnicas modernas de minería de datos, procesamiento del lenguaje natural, análisis de sentimiento y análisis de redes sociales, para así estructurar un estudio descriptivo de marca, de naturaleza cualitativa y cuantitativa.

En virtud de lo anterior, se pretende también, establecer estrategias de proyección desde los mecanismos de posicionamiento de marca y de estrategias publicitarias aterrizadas a un contexto comunicativo, atendiendo a un seguimiento minucioso del comportamiento de la red desde una plataforma de análisis de datos, que se alimenta con 5911 textos extraídos de Twitter, acerca de la cuenta “@Ejercito_CAAID”.

Fundamentación teórica

Para hablar del Comando de Apoyo de Acción Integral y Desarrollo (CAAID) y de su misionalidad, hay que remontarse al concepto mismo de la Acción Integral desde la historia y el contexto que ha permeado el término en el Ejército Nacional de Colombia. La historia comienza a vislumbrar estrategias que lideradas desde una institución castrense, militar, encargada de portar las armas legales del estado, generaban modelos diferentes de trabajo, ligados a las necesidades específicas de las comunidades y de los contextos que presentaba el estado colombiano. De esta forma, a principios de la década de los setenta, cuando el país se encontraba sumergido en una división partidista, el Ejército, a través de la jurisdicción de sus unidades le correspondía garantizar control territorial y representación total del estado para contrarrestar las acciones violentas, una labor que se consolidó como el primer reto que avizoró los inicios de la Acción Integral, según palabras del General Valencia (2006):

Fueron en aquel tiempo los oficiales subalternos de estas unidades quienes empezaron con una actitud comprensiva y humana a acercarse a los campesinos, tratando de ganar su mente y corazón hasta obtener colaboración de ellos para erradicar el bandolerismo, la delincuencia o cualquier conducta violenta en las zonas rurales. (p.8).

Posteriormente, en los años 62 y 65, el Batallón Colombia No. 1, en la Guerra de Corea, conoció las técnicas del combate irregular dentro de un conflicto armado regular, revisando espacios de cooperación entre países aliados, en busca de fortalecer las entonces llamadas acciones cívico-militares y psicológicas en diversas regiones, lo que dio nacimiento al “Plan Laso”, que buscaba centralizar los apoyos encausados a finalizar la violencia. En palabras del General Valencia Tovar este plan fue en realidad el primer ensayo de acción integral (Valencia, 2006).

Años más tarde y manteniendo vigentes las prácticas que buscaban generar procesos disruptivos para conseguir la finalización de los conflictos, se da creación al “Plan Perla”, puesto en marcha por el entonces comandante del Ejército General Guillermo Pinzón Caicedo, con el que se ordenaba a las “brigadas informar las urgencias no militares en sus jurisdicciones (vías, escuelas, puestos de salud, titulación de tierras, medios de paso sobre corrientes fluviales, ordenamiento y apoyo a la colonización espontánea, necesidades específicas de cada región” (Valencia, 2006).

Más adelante llega a los procesos de planeación estratégica militar el “Plan Andes”, que buscaba a través de las unidades en terreno, prestar asistencia social a las comunidades vulnerables, que manifestaran necesidades básicas, a través de algunos recursos que las mismas unidades tenían, y las gestiones que desde el departamento N°5 creado desde 1963 por el Ejército Nacional, se lograrán ubicar.

En el año 1991, se comienza a robustecer la misionalidad y efectividad de estas actividades enmarcadas dentro de los lineamientos de Acción Integral, y se considera importante la creación de una unidad de formación especializada para estos hombres y mujeres, dando paso a la Escuela de Relaciones Civiles y Militares, “encargada de valorar la naturaleza y la evolución de las relaciones cívico-militares y psicológicas, por lo que un año más tarde, se empiezan a dictar cursos de Operaciones Psicológicas destinados a preparar a los integrantes de las Fuerzas Militares en los escenarios e interacciones posibles con la población civil” (Comando General de las Fuerzas Militares, 2009), atendiendo de la misma forma una visión prospectiva, donde se consideraran las variaciones posibles en estos ejercicios de relacionamiento.

Es de esta forma, como poco a poco y a través del ejercicio directo con las comunidades y con los contextos de violencia diferenciada, propios de cada una de las regiones del país, que la Acción Integral fue adquiriendo un protagonismo en el marco de la Defensa y Seguridad del estado, llegando a conformar unidades estipuladas dentro de los procesos organizativos de la institución, con la Dirección de Acción Integral, posteriormente la Jefatura de Acción Integral y finalmente, mediante Resolución N° 0175 del 2017, expedida por el Ministerio de Defensa Nacional, la creación del Comando de Apoyo de Acción Integral y Desarrollo (CAOID), orgánico de la Jefatura de Estado Mayor de Operaciones (JEMOP), con una estructura que corresponde en el 2019 a más de 2000 hombres y mujeres, ubicados a lo largo y ancho del territorio nacional; distribuidos en dos (2) Brigadas de Acción Integral y Desarrollo, cada una con cuatro (4) Batallones de Acción Integral y Desarrollo, que llegan como unidades de apoyo a soportar los requerimientos de cada una de las Divisiones del Ejército Nacional.

La constitución de la unidad como un comando de apoyo, adscrito a la línea de operaciones del Ejército Nacional, reviste a la Acción Integral de una importancia estratégica que sobrepasa los escenarios de la práctica frente a los temas de prueba y error y lo consolida como una línea de acción que debe trabajarse de manera alterna a los procesos de seguridad y garantía de la democracia a través de las armas. Por lo anterior y de acuerdo con el (Comando General de las Fuerza Militares de Colombia, 2016) la definición de Acción Integral (AIN) es la Integración de las capacidades de las Fuerzas Militares al esfuerzo de Acción Unificada del Estado de manera conjunta, coordinada, combinada e

interinstitucional, en apoyo al progreso de la nación. Este aspecto se soporta actualmente desde el Comando de Apoyo de Acción Integral y Desarrollo desde cuatro (4) líneas de acción estratégica, una de las cuales trabaja la componente de comunicaciones y relaciones públicas, denominada Asuntos Públicos.

La línea estratégica de Asuntos Públicos se encarga de trabajar y fortalecer todos los procesos de relacionamiento externo con los medios de comunicación y con la población en general, buscando llevar información sobre las acciones desarrolladas en los diferentes puntos del país, con el objetivo de visibilizar las actividades y el trabajo que desde Acción Integral se viene desarrollando, enmarcado y en coherencia con la Política de Defensa y Seguridad del Gobierno Nacional, generando procesos de Acción Unificada acordes con las necesidades propias de cada territorio, y de la mano con las demás instituciones del estado, con la empresa privada, con las entidades de cooperación internacional y con la sociedad civil organizada.

Dentro de esta línea se conciben los nuevos medios y las herramientas tecnológicas que la era digital pone sobre la mesa; por lo anterior, y teniendo en cuenta la importancia del trabajo de Acción Integral, al Comando se le da autonomía y autorización para el manejo de sus redes sociales propias, es decir, que de forma independiente tiene los espacios de trabajo y publicación exclusivos para los temas de Acción Integral (AIN), atendiendo a los lineamientos de la oficina de Comunicaciones Estratégicas del Ejército, pero con la autonomía de poder generar sus propias sinergias comunicativas, contenidos y estructuras de proyección de acuerdo a las necesidades que se vayan presentando. Dentro de las redes sociales que maneja el Comando de Apoyo de Acción Integral (CAAID), se encuentra Facebook, Instagram y Twitter, entre otras.

Esta última, es una de las que más se utiliza, sin embargo, no deja de ser un reto el generar movimiento a través de esta red social, que se encuentre dentro de los parámetros y las líneas de comunicación que el usuario promedio de esta red social exige, sin entrar en controversias, discusiones, y procesos de opinión fuera de los marcos institucionales establecidos y acorde a los parámetros propios de la red.

Las organizaciones y sus marcas

La comunicación corporativa es un pilar fundamental en la gestión de las organizaciones, debe incluir un plan estratégico en el que se planteen tácticas concretas, acciones asociadas a las relaciones públicas, y con ello, toda una propuesta publicitaria que apunte a la generación de una imagen de marca positiva, esto último, para referirse a los valores y a la filosofía de una organización, entendidas como el concepto fundamental de la marca (De Toro, 2009).

En ese sentido, las marcas deben desarrollar un concepto único y diferenciador, que les permita enfrentarse a un escenario actual de consumo y de producción, mediado por el uso de la tecnología; para responder a esto Cannata (2013), propone una comunicación de marca que vincule todos los sentidos, donde se cuenten historias de forma visual, táctil, auditiva, gustativa y olfativa, intentando crear para el consumidor un nuevo mundo de referencia, que les permita diferenciar y comprender el territorio de cada marca.

Para que la organización pueda responder efectivamente a estos nuevos escenarios de consumo, no debe hablarse únicamente de una comunicación corporativa, deben vincularse aspectos que tengan que ver con el concepto de la marca, y como esta logra proyectarse en diferentes soportes y escenarios.

Para Costa (2012) la marca es una realidad global, consta de aspectos económicos, ideales, culturales, sociales, pero sobre todo, de un conjunto de decisiones y estrategias creativas que contribuyen significativamente en una gestión de marca exitosa. Es así, como una gestión de marca debe responder a cuestionamientos sobre ¿cómo debe construirse la identidad?, ¿cómo influir positivamente sobre la imagen que tienen los públicos sobre la organización?, y entre otras cosas ¿cómo reflejar en una imagen de marca, todo lo que es la organización?

Costa (2012) confirma que el paradigma del siglo XXI en lo que tiene que ver con la gestión de las organizaciones, debe partir de cinco vectores: la identidad, la cultura, la acción, la comunicación y la imagen. Todo esto en función de comunicar asertivamente la respuesta a preguntas como: ¿quién la empresa?, ¿qué hace?, ¿qué dice?, ¿qué es para mí? y para responder todos estos cuestionamientos ¿cómo debería hacerlo?

Para Hernández, Figueroa y Correa (2018), todas las organizaciones que buscan permanecer en el mercado deben gestionarse estratégicamente, pues aquellas que no desarrollan, entre otras cosas, su marca bajo una línea de administración estratégica tienen garantizado su fracaso. Siguiendo su teoría, los elementos que potencian una estrategia de marca, están relacionados con la notoriedad, el posicionamiento, la comunidad, la permanencia y el ahorro, este último, se refiere al uso eficiente el presupuesto para posicionar una marca.

Es entonces, que muchas estrategias se diseñan para gestionar una marca; algunas apuntan a aspectos corporativos y otras, por ejemplo, a partir del uso de la tecnología y los lenguajes que vienen con ella. Belén (2014) afirma que el consumidor actual, cada vez es más demandante y personalizado, busca respuestas inmediatas y por tal motivo necesita una marca que se adapte a sus necesidades: las marcas deben proveer información veraz y de primera mano evitando información falsa.

Redes Sociales y Twitter

Hoy en día las redes sociales son una extensión de las marcas a nivel digital, lo cual ha dado a las organizaciones la oportunidad de generar nuevas maneras de relacionarse con los clientes, estableciendo una comunicación más rápida y directa en comparación con algunos medios tradicionales, de igual forma presenta nuevas formas de pauta a nivel digital, lo que conlleva a la generación de nuevas estrategias de publicidad o marketing.

De manera que, desde la llegada del internet y las nuevas tecnologías de la información, se ha transformado la forma en que las personas se comunican, se relacionan e interactúan a nivel digital, las redes sociales se pueden definir como un conjunto de individuos dentro de una estructura de relaciones. Esa estructura es un soporte casi “virtual”, un tejido, un net que se origina entre los individuos interrelacionados por vínculos de amistad, de colaboración, o de otro tipo de interés. (Fresno, Marqués y Paunero, 2014).

Con estos cambios las comunidades comienzan a tener un nuevo significado para las organizaciones, estas al ser creadas por individuos con intereses, gustos o necesidades en común, pretenden generar y fortalecer las relaciones entre usuarios para intercambiar información, productos, artículos, entre otros; reflejando una transformación en la forma de comunicar, siendo considerado un medio de comunicación más.

La digitalización ha permitido que los medios de comunicación tradicionales migren a las plataformas digitales, al igual que los consumidores y las organizaciones, es decir que el marketing digital pretende fortalecer la comunicación entre los diferentes actores y así mismo comercializar.

Al lograr un mayor nivel de interacción entre los públicos y las marcas por medio de las plataformas digitales, se fortalece el posicionamiento y credibilidad de la marca, lo que conlleva a lograr un alto nivel de compromiso (*engagement*). En efecto el Social Media o las Redes Sociales hacen parte de las estrategias de marketing digital, la cual se centra en generar acciones que permitan generar, incrementar o convertir cada interacción una fortaleza para la marca y la organización.

En definitiva, el marketing digital permite gestionar los medios digitales en tres categorías: los medios propios (*owned media*); los medios ganados (*earned media*) y finalmente los medios pagos (*paid media*) para lograr el alcance deseado.

Otro factor a tener en cuenta es el público que utiliza las redes sociales, el poder identificar y entender ¿quién es?, ¿dónde se encuentra?, ¿cómo distribuye su tiempo?, ¿cuánto tiempo destina a nivel online?, ¿qué redes sociales utiliza?, ¿en qué ocasiones?, ¿para qué las utiliza?, ¿qué necesidades tienen?, ¿qué esperan de la marca?, son solo algunas de las preguntas que toda organización y marca deben hacerse, para poder direccionar el uso de cada una de las redes sociales en las que está presente la marca.

Uno de los medios utilizados por los consumidores, son las publicaciones, estas independientemente de la red social, facilitan la creación de nuevos canales de comunicación, sino a su vez propenden por la generación y transmisión de información entre múltiples usuarios, gracias a esto se identificó una oportunidad en el que las marcas, son conscientes y estrategias al momento de generar contenidos en los diferentes medios digitales, con el fin de motivar a los usuarios a producir contenidos e interactuar entre sí en torno a las marcas.

De esta manera la valoración de los comentarios y opiniones han generado que el discurso publicitario tradicional ha ido perdiendo eficacia. Hoy los usuarios prefieren leer las opiniones de usuarios...y se les atribuye más credibilidad a ellos que a todo lo que las empresas y marcas puedan decir de sí mismas (Maciá y Santoja, 2018).

Twitter

La red social Twitter desde su creación en el 2006, se caracterizó por su facilidad, en donde los usuarios a través de 140 caracteres generan mensajes puntuales sobre distintas temáticas, los tuits permiten la publicación y recepción de pequeños mensajes, pero dada la cantidad de mensajes generados en un lapso, no es posible leer todos los contenidos publicados por los diferentes usuarios; a través de los seguidores se puede interactuar tanto con empresas, como marca y otros usuarios.

A nivel organizacional los elementos clave dentro de la red social permiten, identificar a los usuarios, filtrar, agrupar y gestionar los contenidos (Maciá y Santoja, 2018); entre estos elementos clave, se destacan Hastags a través del uso del numeral # que establece la relación entre la etiqueta y el contenido del tuit; esto facilita la búsqueda de publicaciones; las tendencias cuando varios usuarios utilizan el mismo Hashtag se convierte en Trending Topic, como reconocimiento por ser el más utilizado en un periodo de tiempo; los Acortadores de URL permiten maximizar el uso del mensaje reduciendo las URL; las listas permiten organizar las cuentas con el fin de perder los contenidos a los cuales se les quiere hacer seguimiento, facilita el acceso cuando se tiene un gran número de cuentas en seguimiento y finalmente los buscadores por medio de criterios de selección Twitter ayuda a la búsqueda de contenido por medio de palabras clave utilizadas en la cuenta o en los tuits o etiquetas, todo esto debido a los grandes flujos de información que se generan en un día.

Convenio de información obtenida a través de la red social Twitter

Para el procesamiento de la información obtenida a través de las redes sociales, es pertinente identificar si el sector a estudiar está interesado en el aprovechamiento de la información obtenida a través de las diferentes plataformas y de esta manera potencializar los estudios de marca y posicionamiento de la empresa.

Debido a lo anterior se realiza un convenio con el Comando de Acción Integral y Desarrollo del Ejército Nacional, el cual brindó el apoyo para gestión de la información y el permiso para el estudio y publicación de los datos obtenidos a través de su cuenta de @Ejercito_CAAID de la plataforma Twitter.

A través de este acuerdo de información la investigación va a estar orientada al análisis de la polaridad de los comentarios obtenidos a través de los tuits, y de esta forma estudiar la imagen de la marca frente a las opiniones positivas o negativas de los usuarios. De igual manera la descarga y procesamiento de la información permitirá realizar estudios de marca y posicionamiento involucrando varias áreas del conocimiento.

Procesamiento del Lenguaje Natural

El Procesamiento del Lenguaje Natural puede ser definido como el procesamiento automático (o semiautomático) del lenguaje humano. El termino NPL por sus siglas en inglés, se utiliza algunas veces de forma más restringida, excluyendo la recuperación de información y a veces incluso excluyendo la traducción automática (Copestake, 2004).

El Procesamiento del Lenguaje Natural es contrastado con la “lingüística computacional”, pensado como un campo más aplicado e incluso los términos alternativos como “tecnología del lenguaje” o “ingeniería del idioma” son a menudo preferidos.

Copestake (2004), afirma que NPL es un campo multidisciplinario, es decir que está estrechamente relacionada con la lingüística, sin embargo, se debe tener en cuenta que su grado de relación con la teoría varía considerablemente. De igual forma el procesamiento

del lenguaje natural tiene relación con la investigación en ciencia cognitiva, Psicología, Filosofía y Matemáticas (especialmente la lógica) y, por último, se relaciona con la teoría del lenguaje formal, las técnicas de compilación, la prueba de teoremas, el aprendizaje de maquina e interacción humano-ordenador.

Gracias a su multidisciplinariedad, la NPL ha sido muy bien acogida en la industria lo cual ha hecho que se revolucione el tema, y ha permitido que nuevos autores como Beyosolow II (2018) hablen del lenguaje natural como un subcampo de la ciencia computacional, el cual está enfocado a enseñar a los computadores a entender el lenguaje tal cual como un humano lo entendería, de igual forma diseñar mecanismos en donde se puedan resolver tareas relacionadas con el entendimiento del sentimiento del texto y la generación de respuestas automatizadas a preguntas generadas por los usuarios.

Análisis de sentimientos

Haddaway (2015) define el análisis de sentimientos como un subcampo del procesamiento del lenguaje natural (NPL) que trata de identificar y extraer opiniones de un texto. El objetivo del análisis de sentimientos es medir las actitudes, evaluaciones y emociones basándose en el tratamiento computacional de la subjetividad contenida en los textos.

La minería de opinión se refiere al análisis de los comentarios o expresiones de un autor, la cual tiene varios desafíos debido a que, debe determinarse si un documento o sección del mismo es subjetivo u objetivo y si la opinión expresada es positiva o negativa. Debido a la complejidad del lenguaje humano, su gran expresividad y ambigüedades, el problema de la clasificación de sentimientos no es trivial, dado que implica extraer información semántica de manera automática.

Para lograr esta tarea, desde los noventas; y trabajando de la mano con modelos estadísticos avanzados se han venido desarrollando técnicas para resolver este y otros problemas relacionados con la extracción automática de conocimiento léxico, desambiguación léxica y estructural, inferencia gramatical y etiquetado robusto de la estructura sintáctica de un texto o frase, analizando grandes cantidades de texto (Márques, 2000).

De este desarrollo se generó un gran progreso en la adaptación de los modelos estadísticos actuales para la creación de modelos especializados en la extracción automática de información proveniente un texto, basados en diccionarios léxicos que buscan con la menor cantidad de información disponible lograr una muy buena inferencia del contenido semántico de un texto incluyendo su polaridad.

Un ejemplo de dichos avances es VADER (Valence Aware Dictionary and Sentiment Reasoner) o Diccionario Consciente de Valencia y Razonador de Sentimientos. Este modelo, de sencilla implementación en diferentes lenguajes de programación como Python o JavaScript, está diseñado específicamente para calcular el sentimiento de los textos publicados en las redes sociales usando reglas definidas y un diccionario lexicográfico (C.J. y Gilbert, 2014).

Este modelo, no requiere información adicional dado que se encuentra construido sobre un diccionario léxico de sentimientos en inglés, el cual es generalizable y sus reglas de asignación de intensidad y sentimiento de un texto fueron cuidadosamente validadas para puntuar la valencia o sentimiento en una escala entre -1 y 1, donde -1 significa extremadamente negativo y 1 extremadamente positivo (Calderon, 2017).

Metodología

Como se ha venido mencionando, esta investigación se enmarca en una metodología de naturaleza exploratoria, y se vale de herramientas de análisis cualitativo y cuantitativo modernas, como la minería de datos, procesamiento del lenguaje natural, y análisis de redes sociales para extraer hallazgos relevantes de la información recolectada. En este apartado se resume cómo fueron aplicadas herramientas para la obtención y estructuración de los datos y para la inferencia de la polaridad o sentimiento de los textos descargados, en aras de generar información adicional que permita cuantificar el impacto de la campaña del comando sobre sus usuarios más activos en Twitter.

Obtención de datos

Uno de los retos que representa el análisis de la información presente en las redes sociales, es el hecho de que esta no se encuentra estructurada. Para realizar cualquier análisis descriptivo o estadístico es necesario contar con una base de datos previamente validada y acondicionada, para evitar errores de inferencia acerca de los datos.

La extracción y estructuración de información disponible en las redes sociales, se realiza mediante una técnica conocida como el “Web Scrapping”; diseñada específicamente para extraer información del código HTML de una página web usando un programa o software (Haddaway, 2015).

Dado que la información de las redes sociales puede ser sensible, las diferentes plataformas como Twitter usan programas denominados *-Interfaces de Programación de Aplicaciones-* (API's por sus siglas en inglés) para que los usuarios accedan de manera segura y controlada al contenido que los usuarios consideran como público.

En el portal de desarrolladores de Twitter (2019), se encuentra una documentación clara y concisa de los procedimientos que los interesados deben llevar a cabo para registrar su propia API y así obtener acceso a dicha información a través de lenguajes de programación como R o Python.

Para este trabajo, se realizó todo el proceso de registro de una API con permisos de descarga de información pública presente en la plataforma. A través de esta, y valiéndose del lenguaje de programación Python; se obtuvieron un total de 5911 textos cortos o tweets, en los cuales se hace mención específica a la cuenta “@Ejercito_CAAID”. Los textos fueron descargados en el periodo comprendido entre el 5 de mayo y 7 de julio de 2019, obteniendo, además; información adicional como fecha y hora exactas de publicación, usuario que lo publicó, número de favoritos y *retweets* (ver ilustración 1).

Ilustración 1. Proceso de descarga de información.

Fuente: elaboración propia (2019).

Asignación de polaridad a los textos

Como se menciona en un apartado previo, VADER es un modelo para asignar la polaridad a un texto valiéndose de diccionarios lexicográficos en inglés, lo que presupone una limitación al análisis de sentimientos de textos en lenguas distintas. Sin embargo, la literatura demuestra que se puede sobrepasar esta restricción sin sacrificar la calidad del análisis semántico de los textos, autores como Denecke, (2008) y Khurshid, Cheng y Almas (2006), proponen metodologías de identificación de idioma y traducción para lograr análisis de sentimientos en múltiples lenguas con buenas tasas de precisión.

Para esta investigación se usará la metodología propuesta por Denecke, (2008), la cual se basa en la estandarización de los textos en diferentes lenguas, valiéndose de herramientas de traducción; para luego usar modelos basados en diccionarios lexicográficos construidos en idioma inglés y obtener la polaridad o nivel de negatividad, positivismo o neutralidad del comentario.

Los tweets generalmente se componen de la frase u opinión acompañada de contenido multimedia como, menciones, hashtags, emojis y links que podrían entorpecer la tarea de traducción de los textos. Para que la traducción guarde la mayor fidelidad al texto en español, es necesario que este conserve su construcción semántica y de sintaxis original, por lo que se deben limpiar conservando elementos como hashtags, menciones y emojis; los cuales generalmente dan sentido al comentario dado que se tratan de sujetos dentro de la oración.

Teniendo en cuenta lo anterior, para la calificación de los textos descargados sobre la cuenta @Ejercito_CAAID se debe realizar primero una limpieza de los textos, eliminando únicamente enlaces y entidades de retweet y conservando hashtags, emojis y menciones las cuales dan sentido al texto y pueden marcar la diferencia en el entendimiento del sentimiento de un texto. Luego del pre-procesamiento se realiza el proceso de traducción usando el módulo TextBlob y la respectiva identificación de la polaridad usando VADER. El proceso de asignación de polaridad a los textos se resume en la ilustración 2.

Ilustración 2. Proceso de calificación de los textos.

Fuente: elaboración propia (2019).

Resultados

Como se menciona en el apartado de obtención de datos, con el proceso de descarga se obtiene acceso a información adicional como número de favoritos o retweets de los textos. Esta información permite realizar un análisis exploratorio de la visibilidad de la cuenta en el periodo estudiado y además posibilita indagar sobre la aceptación general de los comentarios relacionados con la misma.

La primera aproximación al análisis de la actividad de la cuenta se puede realizar a través de los hashtags, que son un mecanismo de la red social Twitter que permite a los usuarios unirse a la opinión en temas específicos y hacer seguimiento a tendencias importantes acerca de un tópico en general. A pesar de que tienen un propósito dentro de la plataforma, en el ámbito de la minería de textos; representan un elemento valioso para poder indagar acerca de las temáticas relevantes y discusiones en las que se ve involucrada la cuenta.

La tabla 1, resume los temas o Hashtags que implican mayor relevancia dentro del periodo estudiado en función del número de favoritos que reciben los comentarios que los contienen (los Hashtags se estandarizan previamente convirtiendo todo su contenido en minúscula debido a la gran variedad de combinaciones de escritura que puede digitar un usuario durante su uso).

En la tabla 1, se observa como el hashtag #HeroesBicentenarios ocupa el primer puesto en cuanto a número de favoritos recibidos, esta etiqueta se enmarca dentro de la Campaña Bicentenario; llevada a cabo por el Ejército Nacional y diseñada para fortalecer la imagen y reputación de la marca Ejército usando canales de comunicación digitales entre otros (Ejército Nacional, 2019). La aceptación de los usuarios reflejada en los favoritos muestra indicios del impacto que esta ha tenido en el periodo estudiado, sin embargo; se debe evaluar si se debe a un hecho temporal aislado (día específico en el que el hashtag fue tendencia) o si la campaña logra un impacto positivo prolongado en el periodo estudiado. La gráfica 1, muestra el comportamiento temporal del hashtag #HeroesBicentenarios incluyendo su frecuencia de uso y el número de favoritos recibidos a diario.

Tabla 1. Top 10 Hashtags usados en función de los favoritos que recibe el comentario.

Posición	Hashtag	Conteo de favoritos	%
1	Heroesbicentenarios	1009	15,6%
2	Eldato	395	6,1%
3	batallonaccionintegral1	348	5,4%
4	batallonaccionintegral5	336	5,2%
5	Paradestacar	303	4,7%
6	Hoy	303	4,7%
7	Esnoticia	254	3,9%
8	Decimabrigada	179	2,8%
9	feencolombia	169	2,6%
10	envideo	163	2,5%

Fuente: elaboración propia (2019).

Gráfica 1. Comportamiento temporal impacto del hashtag #Héroes Bicentenarios

Fuente: elaboración propia (2019).

En la gráfica 1, se observa cómo el nivel de uso del hashtag crece durante los fines de semana, indicando que los usuarios que participan de la discusión o los eventos relacionados con esta etiqueta logran un mayor impacto en los días mencionados. Se observa también cómo el número de favoritos se encuentra relacionado con la frecuencia de uso del hashtag mostrando picos importantes de aceptación y en general la impresión de los usuarios, relacionada con la Campaña Bicentenarios en la plataforma Twitter es positiva y se mantiene durante el periodo estudiado.

Word cloud visualization of the top 100 terms from the 'herb' dataset. The most prominent words are 'herb', 'centenar', 'bataillon', 'nacion', 'integral', 'colombia', 'paradesta', 'caia', 'informativo', 'ahora', 'servicios', 'el', 'dato', 'noticia', 'fuen', 'via', 'aest', 'taha', 'bataillon', 'nacion', 'integral', 'colombia', 'paradesta', 'caia', 'informativo', 'ahora', 'servicios', 'el', 'dato', 'noticia', 'fuen', 'via', 'aest', 'taha'.

Fuente: elaboración propia (2019).

Dado que Twitter es una red social, también es relevante analizar sobre la red de usuarios que impulsan la actividad de la marca en la plataforma. Una forma de aproximarse a esto es analizar qué cuentas hacen retuit a los textos relacionados con la marca, dado que estos se convierten en fuertes aliados en el momento de posicionar la cuenta y de generar discusión sobre los temas positivos o negativos que la involucran. La gráfica 2, resume el top 10 de cuentas que generan más actividad alrededor de la cuenta @Ejercito_CAAID en función del porcentaje de retuits que hicieron durante el periodo estudiado.

98

Gráfica 2. Cuentas con mayor porcentaje de retuits a la cuenta @Ejercito_CAAID.

Fuente: elaboración propia (2019).

Análisis del sentimiento de los textos

La identificación de polaridad o carga de sentimientos relacionada a los textos, permite realizar un análisis diferenciado de los comentarios realizados acerca de la marca. Dado que los análisis realizados en el apartado anterior se centran en la totalidad de la información recolectada, sin distinción alguna, este pretende abrir un nuevo enfoque de análisis de la información recolectada.

Como se mencionó anteriormente las nubes de palabras son herramientas útiles para la identificación de palabras claves y de contexto general. Sin embargo, estas pueden variar significativamente dependiendo de la polaridad o sentimiento que refleje el texto.

En el análisis de la polaridad de los sentimientos expresados a través de los tuits, podemos ver que en la nube de palabras de los usuarios que realizan comentarios positivos (ver ilustración 4) acerca del Comando Integral del Ejército Nacional, se tocan temas relacionados con Colombia, paz, fe, amor, esfuerzos, patria, familia, honor entre otros; reflejando la confianza y la admiración de los usuarios al trabajo que realiza el Ejército por el bienestar del país y de las personas.

Ilustración 4. Nube de palabras de comentarios positivos acerca de la cuenta @Ejercito CAAID.

Fuente: elaboración propia (2019).

De igual forma se destacan palabras como cooperación, vida, gracias y comunidad, por lo que se puede inferir que el Ejército Nacional a través del comando especial está generando impacto importante en las comunidades y sus actividades y esfuerzos se ven destacados en las redes sociales.

Dentro del análisis de la polaridad de los sentimientos, también se pueden encontrar comentarios negativos (ilustración 5), en donde palabras como guerra, vulnerable y víctima, muestran como los diálogos negativos alrededor de la cuenta se dan acerca de temas como conflicto armado y sus consecuencias; lo que permite afirmar que, a pesar de la aceptación de los usuarios al trabajo del Ejército Nacional, se siguen generando casos en donde los comentarios generan rechazo a la guerra y a las afectaciones a poblaciones más vulnerables como los niños.

Ilustración 5. Nube de palabras de comentarios negativos acerca de la cuenta @Ejercito CAAID.

Fuente: elaboración propia (2019).

La ilustración 6, da cuenta de los comentarios neutros, es decir, opiniones sin carga de sentimiento; las palabras como comunidades, desarrollo, niños, hombres entre otras reflejan que los usuarios también centran la discusión en temas sociales y resaltan las actividades realizadas en la comunidad por el Ejército Nacional a través del comando especial (ver gráfica 3).

Ilustración 6. Nube de palabras de comentarios neutros acerca de la cuenta @Ejercito_CAAID.

Fuente: elaboración propia (2019).

Gráfica 3. Comportamiento temporal del volumen de comentarios según su polaridad @ejercito_caaid.

Fuente: elaboración propia (2019).

El cálculo de la polaridad o sentimiento implícito en los textos permite también un análisis de impacto más específico, pues al contrastar el comportamiento temporal del volumen de *retuits* y de favoritos con el comportamiento temporal del volumen de textos positivos y negativos se pueden identificar si el impacto que genera en la red social está causando un efecto positivo o negativo entre los usuarios. En las gráfica 3 y 4, se observa cómo un pico de volumen de comentarios positivos, implica un aumento de los *retuits* y favoritos en los días siguientes y, un pico de comentarios negativos no genera una perturbación evidente en el comportamiento del volumen de *retuits* y favoritos. Lo anterior permite afirmar, que dentro del periodo evaluado; las campañas realizadas en la red social generaron un impacto positivo, incrementando la actividad de la cuenta y los usuarios.

Gráfica 4. Comportamiento temporal del volumen de comentarios según su número de Retweets y Favoritos @ejercito_caaid.

Fuente: elaboración propia (2019).

Conclusiones

La vigencia y permanencia de una marca, como aspecto fundamental en su relación con el usuario, está relacionado entre otros aspectos, con la claridad en el concepto y en el discurso utilizado, así como en el uso de estrategias que les permita tener una relación a largo plazo, o como respuesta a un plan estratégico definido para la marca; para el caso de la marca del Comando Acción Integral en Twitter, el hashtag #HéroesBicentenarios es un ejemplo de continuidad dentro de las redes sociales, pues los datos muestran cómo el volumen de menciones y de favoritos se mantienen en rangos elevados durante los fines de semana y volúmenes aceptables entre semana. Esto quiere decir, que la relación del usuario con la marca, se consolida durante los fines de semana, aspecto que debe ser aprovechado por la organización para definir propuestas de fidelización y permanencia con la marca a través de la estrategia de marketing de contenidos adecuadas para este tipo de plataformas.

La etiqueta anterior no es la única que genera relevancia dentro del periodo estudiado, en la nube de Hashtags se observa cómo las etiquetas #ElDato y #ParaDestacar implican un alto volumen de aceptación por la cantidad de favoritos, mostrando el potencial que tienen estas etiquetas dentro de la marca para mostrar las acciones paralelas que desarrolla el Comando, esto confirma que el uso de estas etiquetas permite a la marca lograr un mayor posicionamiento a nivel digital y paralelamente fortalecer el relacionamiento con los usuarios, los cuales a través de los diferentes mensajes pueden identificar nuevos atributos, ventajas y competencias de la marca.

Además de la continuidad, se debe tener en cuenta, la forma en que la marca genera impacto en ciertas redes de usuarios, durante el periodo estudiado se genera una red de usuarios, en su mayoría institucionales; que fomentan la difusión y la discusión de los temas relacionados con el Comando mediante el retuit de los textos, lo cual es relevante debido a que, el apoyo y voz a voz positivo de la marca por parte de otros usuarios implica una mayor y más rápida difusión de la información.

En ese sentido, en las redes sociales no solo se mide la participación de los usuarios con la marca, sino también, la relación y el uso que debe darse en estrategias de comunicación publicitaria a las nubes de palabras encontradas, esto quiere decir, que en este ejercicio es fundamental tener como referencia las palabras generadas empáticamente por los usuarios para enriquecer el discurso mediático y con ello los mensajes emitidos por la marca en los diferentes contenidos publicitarios.

Además del análisis del volumen de comentarios, retuits y favoritos, usando modelos de procesamiento del lenguaje natural se obtuvo el sentimiento asociado a los textos publicados durante el periodo estudiado. Esto permitió ampliar el análisis e identificar si las discusiones generadas alrededor de la cuenta generan impacto positivo o negativo dentro de la red.

Con las nubes de palabras se puede evidenciar que los comentarios positivos guardan una fuerte relación con el impacto que el *Comando de Acción Integral* genera en las comunidades, por medio de sus actividades e iniciativas, las cuales destacan en redes sociales, a diferencia de los comentarios negativos, dentro de los cuales las discusiones se generan alrededor de temas relacionados con la guerra, el conflicto armado y sus consecuencias. Siguiendo esto, una marca como esta, debe desarrollar estrategias para escuchar y entender a sus usuarios, compartiendo con ellos emociones, necesidades, aspiraciones e intereses.

Adicionalmente de entender el contexto sobre el cual se generan las discusiones dependiendo de su polaridad o sentimiento, se analiza el impacto de los comentarios dentro de la red, encontrando que; cuando se generan altos volúmenes de comentarios positivos, la red responde elevando el volumen de actividad (*retuits* y favoritos), mientras que los altos niveles de negativismo en los comentarios se generan por hechos aislados, los cuales no generan un impacto visible en la actividad de la red alrededor de la cuenta del comando.

El uso de redes sociales (Twitter), como extensiones para ofrecer información y generar conexión con los usuarios a partir de las mismas, desde un ejercicio de experimentación de emociones y comentarios que van generando un termómetro en la opinión pública sobre el concepto institucional hacia un lado u otro, debe obedecer a una estrategia desde nivel de comunicación, que obedezca a la naturaleza de los segmentos que interactúan, y a las

preferencias de los mismos frente a palabras, horarios y días de interacción, lo anterior con el ánimo de generarle sentido organizacional, más allá de la presencia en entornos digitales, al uso de estas herramientas, y a los objetivos que quieren lograrse en las audiencias que interactúan con los mismos.

El desarrollo de campañas que generen tendencia y permanezcan por periodos importantes de tiempo, realmente logran construir concepto y marcar diferencia en el sentir y visualizar de los usuarios, adicional a generar arrastre de nuevos segmentos, determinados o seducidos por las diferentes estrategias y temáticas que engloba la campaña y que hacen que poco a poco por ejemplo o por intención individual, los usuarios logren encontrar una temática o actividad que le dé sentido a sumarse a la campaña. El seguimiento permanente de estos medios y el comportamiento de sus audiencias es estratégico para determinar las líneas de seguimiento y los momentos de incursión de nuevas tendencias.

Referencias bibliográficas

- Belén, M. (2014). *Su presencia en la Web 2.0 Análisis y evaluación de la presencia de la marca en Internet, redes sociales y blogs* Marina Belén Rodríguez. Centro de Estudios en Diseño y Comunicación. Facultad de Diseño y Comunicación. Consultado en: https://fido.palermo.edu/servicios_dyc/publicacionesdc/archivos/501_libro.pdf
- Beyosolow II, T. (2018). *Applied Natural Language Processing with Python*. San Francisco, California: Apress.
- Calderón, P. (10 de Abril de 2017). *VADER Sentiment Analysis Explained*. Obtenido de Data Meets Media: www.datameetsmedia.com/vader-sentiment-analysis-explained/
- Cannata, J. (2013). *Carlos Ávalos. La marca: identidad y estrategia*. Austral Comunicación, 2(1), 138-139.
- C.J., H., y Gilbert, E. (2014). *VADER: A Parsimonious Rule-based Model For Sentiment Analysis of Social Media Text*. Georgia Institute of Technology, Atlanta.
- Comando General de las Fuerzas Militares. (2009). *Instrucciones Generales sobre la Doctrina de la Acción Integral. Cartilla de Trabajo*. Bogotá: Imprenta Militar.
- Comando General de las Fuerza Militares de Colombia (2016) “*Manual de Acción Integral Conjunta FF. MM 5-1*”. Imprenta Comando General de las Fuerzas Militares.
- Copestake, A. (2004). *Natural Language Processing Lecture Synopsis*. Department of Computer Science and Technology - University of Cambridge.
- Costa, J. (2012). Construcción y gestión estratégica de la marca: Modelo MasterBrand. *Revista Luciérnaga-Comunicación*, 4(8), 20-25.
- Denecke, K. (2008). *Using SentiWordNet for Multilingual Sentiment Analysis*. Hanover, Germany: Research CenterL3S.
- De Toro, J. (2009). *La marca y sus circunstancias*. Ediciones Deusto. Barcelona.

- Ejército Nacional. (12 de 07 de 2019). *Héroes Bicentenarios*. Obtenido de Ejército de Colombia EJC: <https://www.ejercito.mil.co/index.php?idcategoria=459101>
- Fresno, M., Marqués, P, y Paunero, D. (2014). *Conectados por redes sociales: Introducción al análisis de redes sociales y casos prácticos*. Barcelona: Editorial UOC. Recuperado el 27 de Julio de 2019.
- Haddaway, N. (2015). The Use of Web-scraping Software in Searching for Grey Literature. *The Grey Journal*, 90-186.
- Hernández, C., Figueroa, E., y Correa, L. (2018). *Reposicionamiento de marca: el camino hacia la competitividad de las pequeñas y medianas empresas*. *Rev.investig. desarro.innov*, 9(1), 33-46.
- Khurshid, A., Cheng, D., y Almas, Y. (2006). *Multi-lingual Sentiment Analysis of Financial News Streams*. Palermo: Grid Technology for Financial Modeling and Simulation.
- Maciá, F., y Santoja, M. (2018). *Marketing en redes sociales*. Madrid: Difusora Larousse - Anaya Multimedia. Recuperado el 04 de agosto de 2019.
- Márques, L. (2000). *Machine Learning and Natural Language Processing*. Universitat Politècnica de Catalunya. Barcelona: Departament de Llenguatges i Sistemes Informàtics.
- Twitter Developer. Obtenido de <https://developer.twitter.com/en/docs>
- Valencia, A. (2006). *Teoría y Práctica de la Acción Integral*. *Revista Fuerzas Armadas*, 60(198), 06-11.