



**Colegio de Estudios
Superiores de Administración**

**ESTUDIO DE LA REPUTACIÓN CORPORATIVA DE LA EMPRESA AVIANCA
EN SOCIAL MEDIA, MEDIANTE MINERÍA DE DATOS**

César Ricardo Castro Garzón

Colegio de Estudios Superiores de Administración – CESA

Maestría en Dirección de Marketing

Bogotá

2018

**ESTUDIO DE LA REPUTACIÓN CORPORATIVA DE LA EMPRESA AVIANCA
EN SOCIAL MEDIA, MEDIANTE MINERÍA DE DATOS**

César Ricardo Castro Garzón

Director de tesis:

Ricardo Dicarlo Blanco

Colegio de Estudios Superiores de Administración – CESA

Maestría en Dirección de Marketing

Bogotá

2018

CONTENIDO

Listado de Figuras	4
Listado de Tablas	6
1. Introducción	7
1.1 Antecedentes del problema de investigación	7
1.2 Pregunta de investigación	9
1.3 ¿Por qué es importante la investigación que se propone?.....	10
1.4 Objetivos de la investigación	11
1.5 Objetivos específicos de la investigación	12
1.6 Hipótesis	12
2. Estado del arte	13
3. Marco teórico	19
3.1 Reputación Corporativa	19
3.2 Dimensiones de la Reputación Corporativa.....	21
3.3 Reputación Corporativa en Social Media	23
3.3.1 Percepciones evaluativas de Stakeholders.	25
3.4 Social Media	27
3.5 Twitter como plataforma de análisis de Negocios.....	28
3.6 Análisis de Sentimientos.....	32
3.7 Minería de Datos: Los datos como fuente de valor	36

4.	Marco Referencial	41
4.1	Avianca: Empresa, Huelga de Pilotos y Reputación Empresarial	41
5.	Metodología	49
5.1	Tipo y Diseño de la Investigación	49
5.2	Colección de datos	51
5.3	Extracción de datos	51
5.4	Visualización y análisis de resultados	53
5.4.1	Avianca - Antes de la huelga.....	55
5.4.2	Avianca – Durante la Huelga.	60
5.4.3	Avianca – Después de la Huelga.....	67
6.	Conclusiones y futuros trabajos	72
7.	Bibliografía.....	77

Listado de Figuras

<i>Figura 1 Relación epistémica entre la variable y sus indicadores (Helm,2005)</i>	22
<i>Figura 2 Transformación de datos a conocimiento Accionable (Mohanty, S., Jagadeesh, M., Srivastva, 2013)</i>	36
<i>Figura 3 Procesos de Data Mining (Kumlin, 2016)</i>	38
<i>Figura 4 Técnicas de Análisis de Sentimientos (Kumlin, 2016)</i>	41
<i>Figura 5 Acción Avianca Holding SA - NYSE:AVH (Holding, 2018)</i>	48
<i>Figura 6. Modelo de extracción de datos. Elaboración Propia</i>	52
<i>Figura 7. Timeline - Antes Huelga - Elaboración Propia</i>	56
<i>Figura 8. Keywords - Antes Huelga - Elaboración Propia</i>	57
<i>Figura 9. Polaridad tweets - Antes Huelga - Elaboración Propia</i>	58
<i>Figura 10. Sentiment - Antes Huelga - Elaboración Propia</i>	58
<i>Figura 11. Sentiment Score - Antes Huelga - Elaboración Propia</i>	59
<i>Figura 12. Timeline - Huelga - Pilotos - Elaboración Propia</i>	60
<i>Figura 13. Keywords - Pilotos - Elaboración Propia</i>	61
<i>Figura 14. Polaridad tweets - Huelga - Pilotos - Elaboración Propia</i>	61
<i>Figura 15. Sentiment Score - Huelga - Pilotos - Elaboración Propia</i>	63
<i>Figura 16. Timeline - Huelga - Empresa - Elaboración Propia</i>	64
<i>Figura 17. Keywords - Huelga - Empresa - Elaboración Propia</i>	64
<i>Figura 18. Polaridad tweets - Huelga - Empresa - Elaboración Propia</i>	65
<i>Figura 19. Sentiment Score - Huelga - Empresa - Elaboración Propia</i>	66
<i>Figura 20. Timeline - Después Huelga - Elaboración Propia</i>	67
<i>Figura 21. Keywords - Después Huelga - Elaboración Propia</i>	68

<i>Figura 22. Polaridad tweets - Después Huelga - Elaboración Propia.....</i>	<i>69</i>
<i>Figura 23. Sentiment - Después Huelga - Elaboración Propia</i>	<i>70</i>
<i>Figura 24. Sentiment Score - Después Huelga - Elaboración Propia</i>	<i>70</i>
<i>Figura 25. Relación de las Dimensiones y los Stakeholders (Helm, 2007)</i>	<i>74</i>
<i>Figura 26. Enfoque Dimensional - Reputación Avianca - Análisis de Sentimientos - Elaboración Propia</i>	<i>75</i>

Listado de Tablas

<u>Tabla 1 Periodos Analizados Tweets - Antes, Durante y Después de la Huelga Avianca - Elaboración Propia a partir de #Hashtag y @Mentions</u>	50
<u>Tabla 2. Top posts, Influencer e Impressions - Antes Huelga - Elaboración Propia</u>	57
<u>Tabla 3. Top posts, influencer e Impressions - Pilotos - Huelga - Elaboracion Propia</u>	62
<u>Tabla 4. Top post, Influencer e Impressions - Empresa - Elaboración Propia</u>	66
<u>Tabla 5. Top posts, Influencer e Impressions - Después Huelga - Elaboración Propia</u>	69

1. Introducción

1.1 Antecedentes del problema de investigación

Desde el establecimiento de las estrategias genéricas (Porter, 1989), se ha destacado la importancia de la Reputación Corporativa (CR). Los CEO consideran esta como el activo intangible más importante de una organización, el cual merece toda la atención por parte de la alta gerencia (City, 1993)

Tener una positiva *reputación*, genera un número importante de favorables consecuencias que han sido estudiadas hace tiempo. Estas se relacionan con la intención de compra (Yoon et al., 1993); con la actitud de los compradores hacia los vendedores y productos de una organización (Brown, 1995); con la percepción de la calidad de los productos (Weigelt, 1998); contribuye a la diferenciación entre compañías (Rao, 1994); atrae inversionistas y un bajo costo de capital (Fombrun, 1990) y además establece fuerte identificación entre los empleados de una compañía (Dutton et al, 1994).

Blogging, twitering, facebooking, posting videos, etc., proveen un feedback muy importante a las encuestas, de igual modo a los productos o servicios online. Académicos e investigadores exaltan, cómo estas actividades contribuyen al incremento del *capital social* o *reputación*, lo cual puede verse como una nueva forma de moneda. Estas prácticas podrán generar una nueva área del marketing, de alta generación de valor como el *Branding*, de alto contenido humano y con muchas posibilidades de convertirlo en un activo rentable (Hearn, 2010).

El capital social o reputación empresarial, es el activo intangible más valioso de las compañías y está tomando un valor mucho más significativo para las empresas y sus stakeholders (grupos de interés); superando los balances y/o valoración tradicional de los mismos (Brammer & Pavelin, 2006).

La reputación Corporativa tiene una variedad de definiciones que dependen de la disciplina sobre la cual se desea revisar y los *stakeholder* los cuales va dirigido (Fombrun, Gardberg, & Sever, 2000). Un primer enfoque, eminentemente pragmático, se basa en encuestas hechas a los *stakeholders* involucrados con el mercado (jueces expertos); y sobre estas encuestas se generan índices, como el de la revista *Fortune* (*Fortune AMAC – America's Most Admired Companies*) (Fortune, 2018). En estas encuestas, los expertos analizan a través de la valoración ciertos criterios la reputación de una compañía.

En un segundo enfoque, los académicos han sugerido que, para la determinación de la reputación corporativa, esta se debe basar en una serie más amplia de *dimensiones*. De acuerdo con un concepto más epistemológico sobre el origen y la definición del concepto de reputación corporativa y desde una construcción formativa, se ha desarrollado un concepto más académico. Para definir este *constructo*, se han *refinado* las dimensiones que lo componen, hasta un número de seis mínimas básicas; que son definitivas y muy significativas para la formación del concepto de reputación (Helm, 2005).

Una de las definiciones de reputación, menciona a esta como “*el resultado de un juzgamiento público, que aumenta o disminuye en el tiempo y que es compartida por diferentes stakeholders*” (Colleoni, Arvidsson, Hansen, & Marchesini, 2011). En las redes sociales, los miembros de estas redes se influyen unos a otros, compartiendo sus experiencias, opiniones y modificando constantemente el juicio general de lo discutido

(Bhuiyan, T., Josang, A., Xu, 2010). Por esto, las redes sociales y la web 2.0, actualmente tienen un rol principal y esencial para la construcción o la percepción de la reputación corporativa (RC), dado su alcance, poder de participación y de opinión.

La evaluación de la opinión de los usuarios y sus sentimientos en social media, puede considerarse una excelente aproximación sobre la reputación de las compañías. Utilizando técnicas de minería de datos es posible extraer, monitorear e inclusive predecir las tendencias de esta reputación (Colleoni et al., 2011). Si se logra una aproximación hacia las dimensiones determinantes de la reputación corporativa y se aplica técnicas de minería de datos en un entorno empresarial determinado, se obtendrían unos hallazgos que generan valor, en la medida de poder entender y gestionar aún más la reputación corporativa en la web.

Con base en lo anterior, se considera para este análisis, la compañía de aviación Avianca, buscando como se afectó su reputación corporativa (CR), dados los sucesos de la huelga de Pilotos en septiembre del 2017. Este análisis pretende abarcar una serie de tiempos: antes, durante y después de la huelga, revisando en social media (específicamente en Twitter), la variación en la percepción de esta Empresa y cómo vio afectada su reputación.

1.2 Pregunta de investigación

¿Cómo se representa la Reputación Corporativa en social media, con un enfoque dimensional para la empresa Avianca, mediante la aplicación de técnicas y herramientas de minería de datos?

1.3 ¿Por qué es importante la investigación que se propone?

Desde hace ya algún tiempo, ha tomado relevancia el estudio sobre la construcción de la reputación corporativa (CR) y la responsabilidad social corporativa (CSR), dentro de la literatura de la Gerencia Estratégica. Esto debido a su relación con desempeño financiero (Formbrum, 1990), (Waddock, 1997) y (Dowling, 2002); la percepción de calidad de los productos (Sen, 2001); la productividad de los empleados (Turban, 2003) y las características de administración de los dueños (Johnson and Greening, 1999).

La reputación no puede ser juzgada sin referenciar sus dimensiones. Desde hace varios años, la revista Fortune ha publicado un índice de reputación corporativa, basado en el juicio experto de 10.000 ejecutivos de las más grandes compañías, evaluando varias dimensiones, de las cuales se destaca: la calidad de la administración, la calidad de los productos o servicios, la solvencia financiera, la innovación, la habilidad para atraer o mantener el talento humano, la responsabilidad hacia la comunidad y el medio ambiente, etc. Pero estos índices han sido muy criticados, por responder más a tendencias de evaluación financiera y poco sobre la percepción de la reputación desde el consumidor en general (Caruana & Caruana, 2006).

A partir de las redes sociales y mediante la aplicación de herramientas de minería de datos como: *Opinion mining o Sentimental Analysis (SA)* (Cambria, Schuller, Xia, & Havasi, 2013), podremos explorar y medir la *percepción - reputación* que tienen los usuarios de las redes sociales frente a compañías y sectores de la economía real, en este caso en específico, sobre la industria Aeronáutica, en particular la empresa Avianca.

Estas percepciones son valoradas a través de menciones positivas y negativas (polaridades) en las principales redes sociales, mediante programas de procesamiento del lenguaje natural (*Natural Language Processing* - NLP). Estos programas se desarrollan con base en una lista de palabras definidas, que pueden clasificar las opiniones en favorables o desfavorables. Varios estudios demuestran que las opiniones en las redes sociales pueden ser usadas como predictores de escogencia sobre marcas, productos o servicios. (Kumlin, 2016).

La reputación, además de ser el activo intangible más valioso de una compañía, es una fuente que permite incorporar diferentes modelos de percepción de una empresa. Para los economistas representa el rasgo distintivo de una compañía; para los estrategas, es una barrera para los rivales y una ventaja competitiva; para los contadores es un activo, cuyo valor fluctúa por el mercado y para los mercadólogos es un *poder* que permite atraer más clientes (Fombrun et al., 2000).

Estudiar, configurar y calificar la reputación corporativa de las industrias se justifica, hoy en día más, como una manera de apalancar el camino hacia el éxito de una compañía frente a sus stakeholders; donde estos están cada más integrados mediante las redes sociales.

1.4 Objetivos de la investigación

Describir la Reputación Corporativa en Social Media de la empresa Avianca, mediante minería de datos, para la huelga de Pilotos de septiembre de 2017.

1.5 Objetivos específicos de la investigación

- Analizar el comportamiento de la percepción sobre la Reputación Corporativa de la empresa Avianca; para las series de tiempo correspondientes a: antes, durante y después de la huelga de pilotos de esta empresa en septiembre del 2017.
- Identificar cual(es) dimensión(es) se configura(n) para la Reputación Corporativa en la web 2.0, extraídos de la minería de datos.
- Obtener elementos de información de la extracción de datos, que contribuyan a interpretar la Reputación Corporativa de la empresa Avianca, para los periodos analizados.

1.6 Hipótesis

- Hipótesis 1 – H1: Las variaciones de las opiniones (sentimientos) en social media, determinan cambios en la Reputación Corporativa.
- Hipótesis 2 – H2: Existe una relación directa entre la dimensión de la Reputación Corporativa, *Emotional Appeal* y el resultado del análisis de sentimientos en social media.
- Hipótesis 3 – H3: La Reputación Corporativa de Avianca se favoreció después de ganar el pleito legal de la huelga de sus pilotos en el 2017.

2. Estado del arte

En el mundo digital de hoy, muchas de las organizaciones entienden la importancia de la utilización de nuevas tecnologías y del *big data*. La pregunta es ¿Cómo aprovechar estos avances tecnológicos para generar *valor* y colocar al consumidor como el centro de la investigación, buscando comprender mejor su comportamiento? En la social media se produce una cantidad enorme de datos y estos se pueden utilizar para generar una mejor experiencia al consumidor. Con tecnologías como la web 2.0, ha cambiado la forma como las personas se comunican e intercambian información.

Existen más oportunidades que peligros en la web 2.0, donde se puede advertir ciertos beneficios que solamente la social media puede ofrecer. Estos beneficios se basan en la utilización de la minería de datos en las redes sociales.

Beneficios como el generar *insights* para los negocios, por ejemplo. Igualmente generando valor a través del estudio y la administración de la reputación corporativa en la web y demostrando de esta manera como el análisis de redes sociales puede ayudar a las organizaciones a conocer mejor sus consumidores y contribuir al éxito de las compañías (Kumlin, 2016).

Kumlin (2016), de igual modo aplica un análisis de minería de datos basado en herramientas como *Sentimental Analysis* y *Text mining*, a dos empresas de transporte marítimo, reconocidas en Finlandia, para evaluarlas mediante la polaridad de las opiniones que tienen sus usuarios en twitter, para un periodo de tiempo específico.

Dentro de las conclusiones principales de este trabajo, se tiene que el resultado de las opiniones subjetivas compartidas en las redes sociales, es cada vez más amplia. Esto equivaldría a que sí no se escuchara y se monitoreara la “voz” de los *stakeholders* en la web, la reputación corporativa sería como un *barco*, que viaja a la deriva, sin ninguna influencia (vientos) por parte de las compañías, para orientar su rumbo. Finalmente, concluye que hay muchas ventajas y beneficios por la utilización del análisis de social media, para el manejo de la reputación corporativa (Kumlin, 2016).

El trabajo de Kumlin es un estudio muy reciente sobre reputación corporativa en la web (2016); y donde se exponen las últimas técnicas de análisis de sentimientos (minería de datos), aplicado a compañías locales en Finlandia, para evaluar y comparar su reputación corporativa contra su desempeño bursátil.

Pero, ¿Que es la reputación corporativa y cómo podemos medirla? Esta es una de las preguntas centrales que la autora Maria Barbato se hace sobre la reputación corporativa. Al hacer una revisión de como las diferentes disciplinas (economía, sociología, etc.) definen la reputación corporativa, se encuentra que no hay una definición en común, pero todas apuntan a que la *reputación corporativa es el resultado de las relaciones entre las compañías y sus stakeholders* (Barbato, 2016).

Dentro del trabajo de Barbato, se evaluó si la reputación corporativa puede ser medida mediante técnicas de minería de datos (*sentimental analytics*) y como este análisis puede ser aplicado para mostrar a la variación de la reputación corporativa de la compañía Volkswagen, debido al escándalo del desempeño de sus motores diésel, en el año 2015, y su efecto en la valoración de las acciones de esta compañía.

En las conclusiones que obtuvo Barbato, se demuestra que herramientas como *sentimental analysis*, permite establecer una relación casi directa entre la polaridad de las opiniones en la red (en este caso negativa) y el comportamiento de la acción de Volkswagen (a la baja), durante el periodo analizado. De igual forma menciona que lenguajes de programación como *Python*, permiten de una manera fácil y natural, realizar el análisis de sentimientos sin ahondar en grandes conocimientos de programación (Barbato, 2016).

Al ser también un estudio muy reciente (2016) y con aplicación dirigida a las ciencias de la computación; Barbato hace un mayor énfasis en las herramientas de computación, que permiten realizar esta minería de datos, que en el resultado del análisis de sentimientos en sí mismo.

Por otra parte, autores más académicos han abordado el tema de la reputación corporativa desde su definición epistemológica, donde se plantea que la reputación es una construcción que se *forma* (formativa) y *que no es reflectiva* (conocimiento inductivo). Basada desde la perspectiva del consumidor alemán, Sabrina Helm, conceptualiza que la reputación es una construcción formativa que consiste en diez indicadores (dimensiones) y a partir de estas dimensiones, se puede construir un sólido método para medir la reputación corporativa, convirtiendo estas mediciones en una manera eficiente de calcular y administrar la reputación corporativa (Helm, 2005).

La técnica aplicada por esta autora para validar sus resultados se basa en la determinación de las covarianzas entre dimensiones encontradas; determinando cuáles de estas dimensiones son significativas, por sus relaciones de valor numérico positivo y de alta significancia estadística. Todo esto con base en unas encuestas aplicadas a 1681

consumidores en Alemania. En el estudio se revisa, desde el punto de vista académico, como se construye el concepto reputación corporativa, desde una estructura formativa; y con base en este modelo, poder determinar al menos seis dimensiones, que tienen una correlación positiva y una significancia muy alta, para la determinación de la reputación corporativa, esto aplicado a los consumidores en general. El estudio se basó en una encuesta aleatoria, con una muestra significativa de consumidores no expertos, como los que hoy conforman las redes sociales (Helm, 2005).

Dentro de los estudios clásicos y de mayor citación sobre la reputación corporativa, se destaca el realizado por Charles Fombrun, de la Universidad de Nueva York. Fombrun se planteó el desarrollar el más completo instrumento para medir la reputación corporativa. Junto con otros colegas revisó los rankings que se tenían en la época y observó sus debilidades. A partir de allí, desarrolló un instrumento de marca registrada: *The reputation Quotient* (RQ); que se convirtió en un referente para la medición de la reputación corporativa, según la literatura revisada. Al final, mediante técnicas como *focus group* y estudios pilotos en diferentes industrias, creo el instrumento de mayor cobertura que, en ese momento, podía medir la reputación corporativa, mediante varias categorías – dimensiones, para diferentes industrias (Fombrun et al., 2000).

La técnica que aplicó Fombrun fue a través de un desarrollo Nomológico-Deductivo y mediante un diseño a escala, para la definición del constructo (Reputación), Este desarrollo tuvo las siguientes etapas: Definición del constructo – Diseño de la escala de medición – Tests Pilotos – Administración de la escala y tamaño de los ítems – y finalmente una Validación y Normalización de los resultados.

Este trabajo de referencia clásico, antes de la llegada del internet y las redes sociales, se consideró como el instrumento de mayor cobertura para medir la reputación corporativa en diferentes *stakeholders* y para toda clase de industrias. *Tiene básicamente seis categorías-dimensiones para evaluar la reputación corporativa: Atractivo emocional, Productos y Servicios, Visión y liderazgo, ambiente de trabajo, responsabilidad social y ambiental y finalmente desempeño financiero* (Fombrun et al., 2000).

Para determinar cómo ha evolucionado la reputación, desde los principios de la era industrial hasta la implementación de rankings y rating en la web 2.0; Alison Hearn, muestra como ahora, los Blogs, twitter, posting de video en Youtube o las evaluaciones con sus respectivos rankings en tripadvisor; generan en la esfera digital un *capital social o reputación*.

El trabajo Hearn, examina cómo esta reputación ha sido históricamente definida y entendida; y como actualmente ha sido agregada, medida, estructurada y orientada hacia los intereses de las compañías en la web. Considera como el uso de estos rankings, listas o scores, van construyendo reputación y sí estos realmente constituyen acciones libres de los intereses de las compañías (Hearn, 2010).

El proceso que aplicaba Hearn, se basa en un discurso argumentativo, de carácter inductivo, que pretende desde una recopilación histórica de las definiciones clásicas de la reputación, llegar hasta como hoy en día, se puede gestionar una administración y una medición de la reputación corporativa en la web 2.0, para beneficio de las empresas.

En el artículo, Hearn pretende elevar el estudio de la reputación corporativa en la web, a una nueva *disciplina* dentro del mercadeo, semejante al *branding*; que puede generar mayor valor sí se desarrolla y se gestiona como el concepto de *marca*; base sobre la cual se desarrollan varios estudios de reputación corporativa recientes (Hearn, 2010).

Dentro de los estudios más recientes encontrados, tenemos una tesis de doctorado que analiza las redes sociales en la industria petrolera y como a través de minería de datos (*sentimental analysis*) en Twitter, se puede obtener el *sentimiento* del público en general sobre compañías *controversiales* como son las Petroleras. Se pretende también determinar cuál es la precisión de la aplicación de métodos de evaluación como *sentimental analysis* y también como los ambientes culturales diferentes (Petroleras de Occidente contra Petroleras de Oriente) pueden generar *polaridades* diferentes frente a un mismo tema (Aldahawi, 2015).

Adicionalmente tenemos otra tesis que estudia la reputación corporativa a través de social media y la aplica al escándalo sucedido en Europa respecto a la falsedad de la información de los motores de la automotriz Volkswagen (VW) para sus motores diésel. (Barbato, 2016)

Estos dos últimos estudios mencionados, al ser realizados para obtener títulos en ciencias de la computación aplicadas a la administración, su base fundamental está en el desarrollo y aplicación de programas de cómputo, que buscan encontrar la mayor exactitud posible en la polaridad de las opiniones en redes sociales (Twitter); pero son escasos en su análisis y asociación al concepto de Reputación Corporativa (CR) y como esta extracción de información pueden generar indicadores o insights que den *valor* a la extracción y seguimiento de la Reputación Corporativa en redes sociales para las Compañías.

En el caso colombiano, el estudio que determina la Reputación Corporativa de las empresas más influyentes en el País, se denomina *MERCO (Monitor Empresarial de Reputación Corporativa)* que se realiza anualmente, que es pre-configurado por altos ejecutivos colombianos y expertos en temas, especialmente financieros y establece un

ranking de las 100 principales empresas colombianas en diferentes sectores, utilizando índices, como el que presenta la revista Fortune, mediante encuestas (Merco, 2017).

3. Marco teórico

3.1 Reputación Corporativa

La reputación Corporativa ha sido una gran cuestión para el *management* de las compañías. Rindova et al. (2005) revisaron más de 60 artículos sobre el tema y encontraron que la reputación corporativa y su medición tiene un sin número de definiciones y metodologías de medición. Las compañías han entendido que capitalizar este activo intangible genera una gran rentabilidad y posiciona de la mejor manera ante sus stakeholders.

Dependiendo de la disciplina que la aborde, para el constructo (constructo puede entenderse como una entidad hipotética que resulta difícil de definir en el marco de una teoría científica) de Reputación Corporativa, se destaca una primera definición desde la disciplina del *mercadeo*, donde se describe como *la asociación que los individuos establecen a partir del nombre de una compañía* (Barbato, 2016). Revisando una definición más amplia se puede mencionar que es una *construcción colectiva, que describe múltiples percepciones agregadas, de los diferentes stakeholders sobre el desempeño de una compañía* (Fombrun et al., 2000). Si el desempeño corporativo es una construcción multidimensional, se espera que la reputación corporativa también lo sea. Finalmente, todo esto se traduce en el resultado de la relación entre la compañía y sus stakeholders (Barbato, 2016).

Existen dos grandes tendencias que abordan la reputación corporativa. La primera se basa en la percepción de como los *stakeholders* ven a la corporación (imagen) y como la corporación se presenta a sí misma (identidad). La combinación de estos dos resultados, imagen e identidad, genera el concepto de reputación corporativa. Este modo de definir y de encontrar la reputación corporativa, se basa en el concepto de marca (Barbato, 2016).

Existe otra tendencia, recopilada históricamente y sobre la cual se mueven los índices de tradiciones de medición de reputación corporativa (*Fortune Index*, GMAC, etc.). Aquí se determinan ciertas dimensiones que explican la percepción de la reputación y se aplican *encuestas* a los *stakeholders* para medir la percepción que tienen en función de cada una de las dimensiones presentadas. Esta ha sido el método de mayor uso para medir la reputación corporativa (Fombrun et al., 2000).

La era digital y particularmente social media como Twitter y Facebook han creado una situación diferente para el manejo de la reputación corporativa. *Social Media se ha caracterizado por su fácil búsqueda, abierta participación, publicidad limitada, dialogo, redes, amplia y rápida difusión de la información y su respectivo feedback* (Aula, 2010). El impacto de estas tecnologías expone un riesgo en términos de la reputación a las compañías. Es muy mencionado el caso de un músico, cuya guitarra fue dañada en un viaje de la compañía *United Airlines* y al no ser atendida su reclamación, posteo un video en Youtube en el 2009 titulado “*United breaks Guitars*” y fue visto por millones de personas, haciendo que la compañía tuviera que compensar al músico y disculparse públicamente por este impase (Aula, 2010).

No hay mucha información respecto a las nuevas posibilidades que ofrece la disponibilidad de las opiniones de los clientes/usuarios en la web y su relación con la

Reputación Corporativa. Con la masiva producción de datos en social media, es posible extraer, monitorear e inclusive predecir las tendencias de la Reputación Corporativa, por la agregación de opiniones subjetivas, usando técnicas de minería de datos como *sentiment analysis* (Colleoni et al., 2011).

3.2 Dimensiones de la Reputación Corporativa

Hay un gran interés entre los investigadores respecto a lo referente sobre la Reputación y su aplicación en el ámbito empresarial. No hay consenso sobre las dimensiones de la reputación y por eso los investigadores han recurrido al modelo de ecuaciones estructurales (*structural equation modeling*), muy aplicado a las ciencias sociales; donde a un constructo o variable latente le es asignado un set de indicadores que, en el caso de la reputación, se asocia a un modelo *formativo* de indicadores. Este acercamiento epistemológico respecto a la reputación permite concluir que el modelo formativo (no reflectivo), es el que mejor explica el concepto de reputación, ya que son las dimensiones la *causa* que forma el concepto y que la agregación de las diferentes dimensiones configuran el constructo. Ver Figura 1 (Helm, 2005).

Figure 1: Epistemic relationships between the latent variable and its indicators

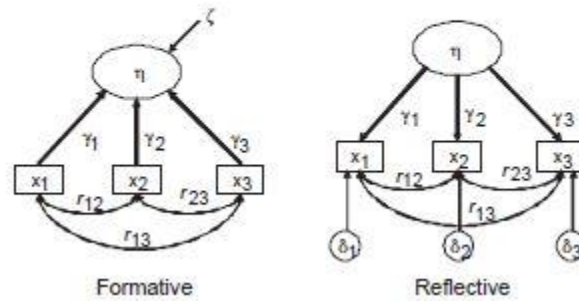


Figura 1 Relación epistémica entre la variable y sus indicadores (Helm,2005)

En cuanto a las dimensiones de la reputación, según la investigación realizada, de acuerdo con su naturaleza y cantidad, también han sido motivo de debate. Asumiendo lo que la mayoría de investigaciones mencionan como válido (Caruana & Caruana, 2006), (Etter, Colleoni, Illia, Meggiorin, & D'Eugenio, 2017), (Colleoni et al., 2011), (Helm, 2005), (Helm, 2007), (Fombrun et al., 2000), las dimensiones que mejor explican el constructo o la variable latente correspondiente a la Reputación Corporativa son:

- *Emotional Appel* (¿Tengo buenos sentimientos hacia la firma?, ¿Admiro la compañía?, ¿Confío en la compañía?)
- *Products and Services* (¿Respalda sus productos o servicios?, ¿Desarrolla productos o servicios innovadores?, ¿Ofrece productos o servicios de alta calidad?, ¿Ofrece productos o servicios acordes con el valor pagado?)
- *Vision and Leadership* (¿Tienen un excelente liderazgo dentro de su categoría?, ¿Tienen una visión clara de su futuro?, ¿Reconoce y toma ventajas de las oportunidades del mercado?)

- *Workplace Environment* (¿Es bien administrada?, ¿Parece una buena compañía para trabajar?, ¿Parece que tiene buenos empleados?)
- *Social and Environmental Responsibility* (¿Soporta buenas causas?, ¿Es una compañía responsable ambientalmente?, ¿Mantiene altos estándares en el trato con la gente?)
- *Financial Performance* (¿Tiene buena rentabilidad?, ¿Aparenta bajo riesgo para invertir?, ¿Tiene mejor desempeño que sus competidores?, ¿Luce como una compañía con alta promesa de futuro crecimiento?)

3.3 Reputación Corporativa en Social Media

Con el advenimiento de la web 2.0, las redes han empoderado a los consumidores, cambiando los hábitos y comportamientos de los usuarios, dado el intercambio de información y experiencias; influenciando a sus pares en los procesos de decisión de productos y servicios (He, Zha, & Li, 2013).

La pérdida de control sobre el flujo de la información, puede afectar significativamente la reputación corporativa, pero las compañías pueden influenciar activamente esta reputación a través del *engagement o implicación*. Aquí aparece el concepto de la reputación online, el cual representa la reputación que las compañías establecen en internet, principalmente fomentando la participación de la comunidad y el refinamiento de búsquedas, con herramientas como *google search*. Esta participación de la comunidad consiste en una comunicación de doble vía entre la compañía y sus *stakeholders*.

Pero lo que *construye* reputación son las experiencias directas o indirectas que reciben estos grupos de interés y que se comparten a través de *word-of-mouth* (WOM), la web u otras plataformas (Etter et al., 2017).

Estudiosos del tema de la reputación coinciden en su importancia como un activo intangible muy valioso, ya que crea barreras en el mercado y aumenta las ventajas competitivas. En otras palabras, la habilidad para crear, manejar y explotar estos intangibles, permiten a las compañías manejar los mercados y no que los mercados manejen a las compañías.

El resultado de la definición de la reputación corporativa en la web, está comprendida por dos dimensiones principalmente: la calidad percibida por los *stakeholders* y la importancia de la organización en la mente de los mismos (Rindova, Williamson, Petkova, & Sever, 2005). Las evaluaciones de los stakeholders pueden ser descritas como la opinión acerca de las compañías por parte de los clientes, inversores, empleados y el público en general, que expresa aptitudes subjetivas, emocionales y cognitivas. La importancia de las organizaciones se refiere al grado en el largo plazo de la atención colectiva y su reconocimiento. Es decir, *la reputación en la web es el resultado del juzgamiento público, que puede incrementarse o disminuir en el tiempo y que es compartido por los diferentes stakeholders.*

Tradicionalmente, la opinión del público fue manejada a través de diferentes canales tradicionales como prensa, radio o televisión, y la comunicación corporativa utilizaba estos canales para enviar el mensaje deseado a su público objetivo; pero actualmente la opinión pública tiene más independencia en relación a los canales y las compañías. La

comunicación en social media es participativa y bidireccional, por lo tanto, la opinión acerca de las compañías es más difícil de controlar.

El mecanismo de reputación Online está directamente relacionado con los grados de *co-producción* y el *engagement* de los stakeholders para la creación y circulación de contenidos. De aquí, parte el concepto del *Prosumer*, ya que no hay distinción entre productores y consumidores de contenido en digital. El ambiente de participación en social media hace que la construcción de la reputación sea menos controlada por las empresas y su cocreación sea muy dinámica entre los stakeholders.

La bidireccionalidad de la social media representa un espacio para expresar opiniones e ideas, donde la gente se comunica de manera horizontal. Estas nuevas tecnologías han permitido una comunicación directa entre las organizaciones y sus clientes y entre ellos mismos (clientes), permitiendo también que las opiniones personales y los pensamientos sean accesibles a una comunidad global de internautas.

La reputación y su confianza son el resultado de los comportamientos previos de la compañía hacia los stakeholders y de la interacción entre las comunidades de los stakeholders. Esto hace que el manejo de la reputación este cambiando hacia valores más éticos y en el establecer unas relaciones más afectivas con las comunidades de stakeholders digitales, más que persuadirlas al corto plazo (Verhoef, Peter; Kooge, 2016)

3.3.1 Percepciones evaluativas de Stakeholders.

Las percepciones de los Stakeholders respecto a los atributos corporativos, se refieren a la opinión general acerca de compañía, que se definen como una actitud. *Actitud entendida*

como tendencia psicológica que esta expresada por la evaluación particular de alguna entidad, con algún grado de favorecimiento o desfavorecimiento y que haciendo referencia a la reputación corporativa, tiene que ver con las respuestas afectivas y cognitivas (Einwiller, Carroll, & Korn, 2010).

El rol de las respuestas afectivas (emociones y sentimientos), es reconocido cada vez más como un factor relevante que dirige las evaluaciones de los consumidores. Las actitudes hacia las compañías están definidas por las opiniones y los sentimientos que circulan en las redes digitales. Ahora, la gente obtiene el conocimiento de las compañías buscando e interpretando las señales online. Estas señales no solamente están basadas en las comparaciones entre compañías, sino también como las redes sociales perciben el desempeño de los productos y/o servicios de las mismas. Una vez las personas construyen su percepción, comparten la misma con otros y *la verdad subjetiva, se convierte en la verdad colectiva sobre una compañía (Aula, 2010).*

Esto es particularmente verdad en plataformas como Facebook y Twitter, que están específicamente diseñadas para compartir emociones, sentimientos y opiniones entre sus usuarios. De acuerdo con Aula, *Social Media tiene el efecto de presentar la verdad colectiva, y esta se basa en las emociones y los sentimientos de los usuarios.* En consecuencia, *la Reputación Corporativa Online puede ser vista como los sentimientos y las emociones alrededor de una compañía, que si es positiva, tendrá la capacidad de atraer más y más opiniones y sentimientos positivos de sus stakeholders; pero a la vez, lo contrario es igualmente posible (Arvidsson, 2011).*

3.4 Social Media

La iniciación del concepto de social media puede ser rastreado a finales de 1970, cuando comenzó como un Boletín digital (*Buletin Board Systems – BBS*). Este era una plataforma donde se podía cargar anuncios, leer noticias e inclusive intercambiar mensajes entre los usuarios. Sobre esta base y los fórums, se iniciaron los primeros Social media *networks*, que permitían tener chat entre varios usuarios, crear grupos de interés y seleccionar amigos. *Six Degree* fue lanzado en 1997, considerándose la primera forma moderna de Social media. En 2002, *Friendster* apareció para el público en general de los Estados Unidos, mientras *MySpace* y *LinkedIn* fueron lanzadas en el 2003. Facebook y Twitter, estuvieron disponibles para los usuarios en 2004 y 2006 respectivamente. Las bases de datos disponibles de estas redes sociales son de gran importancia para extraer y utilizar su información con propósitos de análisis y predicción.

Aunque las redes sociales ya hacen parte integral de nuestras vidas, es necesario dar una definición formal de las mismas. Una red social puede ser definida como un grupo basado en aplicaciones de internet, que sobre la base de las propiedades de la web 2.0, permiten la creación y el intercambio de contenido generado de manera propia o importada. (Kaplan & Haenlein, 2010).

Otros autores se definen *Social Media como un set de herramientas y tecnologías que permiten a la gente conectarse con otros y construir relaciones de manera eficiente*, es decir, que también se puede describir Social Media a través de la misma red: *Set de nodos conectados por relaciones. Estos nodos pueden ser individuos, grupos de individuos o páginas web* (Fielding, Lee, & Blank, 2008).

Las redes sociales pueden dividirse en dos grupos: Contenido-céntricas y las Usuario-céntricas.

- **Contenido-Céntricas:** Estas corresponden a plataformas donde el usuario postea sobre una variedad de temas. Estos post pueden ser comentados y reenviados a otras personas. Las más comunes son Twitter, Tumblr y Wordpress. Twitter es una plataforma de *microblogs*, que permiten a los usuarios interactuar y comunicarse con otros. Los usuarios pueden transmitir en tiempo real mensajes llamados Tweets de hasta 140 caracteres de longitud. Estos tweets pueden tomar formas diferentes, como texto, imágenes o videos, en los cuales los usuarios expresan sus opiniones, para informar a otros sobre actividades, eventos o nuevas noticias. Otros ejemplos de este tipo de redes sociales de contenido son YouTube o Flickr, donde los usuarios comparten videos, imágenes y se discute los contenidos con otras personas.
- **Usuario-Céntricas:** Estas plataformas se enfatizan en la identidad de los usuarios, suministrando perfiles o compartiendo intereses personales. Los usuarios actualizan sus estatus, comparten fotografías, videos y links. La más popular es Facebook y se estima que 57% de los adultos en Estados Unidos usan Facebook y el 64% de estos adultos revisan su perfil diariamente, para el año 2012.

3.5 Twitter como plataforma de análisis de Negocios

Se ha estimado que más de 316 millones de usuarios activos de twitter postean alrededor de 500 millones de tweets por día en 35 lenguas diferentes. Twitter puede ser clasificado como una red social usuario-céntrica (Twitter, 2018).

Las organizaciones se pueden *amplificar* con los diálogos online, para intensificar su prosperidad, usando plataformas de social media como Twitter. Un estudio revela como las 500 empresas más ricas de Fortune 500, se han beneficiado por la construcción de una relación Online con sus *stakeholders*, cuando empiezan a utilizar de manera regular social media como Twitter. Esto a través de un dialogo continuo entre la compañía y su público, presentando información relevante de la organización (Rybalko & Seltzer, 2010).

En otro estudio se revisa como se puede comprometer más a los stakeholders con las compañías por medio de Twitter. Se demuestra que más post positivos, fueron recibidos por las compañías, que tienen principalmente una buena comunicación y un buen *enganche* con sus stakeholders a través de Twitter. En el caso contrario, compañías con menor *enganche*, reciben mayores tweets negativos por sus deficiencias en la comunicación (Wigley & Lewis, 2012).

Las tendencias actuales de la comunicación en la web y de redes sociales están muy relacionadas con el concepto de co-creación. Este concepto se focaliza en la generación consistente de valor compartido entre consumidores y las compañías. Social Media tambien ofrece beneficios a los negocios, basándose en el *Know how* de la minería de datos, debido a la basta cantidad de datos disponible en las redes sociales. Obtener información relevante es muy útil para ayudar a la toma de decisiones y para el aprendizaje que conlleva a la acción inteligente.

Tambien se puede medir en las redes sociales, la eficiencia y la efectividad sobre el conocimiento de los clientes, que influye en la estrategia de la compañía. Además, se puede generar conocimiento de nuevos productos o servicios y monitorear el grado de aceptación de estas innovaciones.

En el estudio de Wigley y Lewis, se muestra como las compañías en crisis pueden mantener su reputación, si enganchan a sus stakeholders y al público mediante una comunicación de doble vía. Cuando esto sucede, la compañía debe explicar cuáles fueron las causas raíces del problema y que acciones se tomarán para corregir y prevenir cualquier situación de nueva ocurrencia o repetición del problema. Se ha demostrado que social media *viaja* más rápido que los canales tradicionales de comunicación y es más económico (Wigley & Lewis, 2012).

Según la información de la compañía Twitter en su sitio web (Twitter, 2018), los tweets permiten enviar mensajes cortos de 140 caracteres. Es uno de los 10 sitios web más visitados del mundo y fue descrito como “el SMS de la internet”.

Los usuarios pueden agrupar tweets por tema o por tipo, utilizando *hashtags*, los cuales son palabras o frases con el prefijo “#”. El símbolo “@” acompañado del nombre de un usuario es utilizado para mencionar o responder a dicho usuario.

Estos *hashtags*, cuyas palabras o las frases sin espacios son precedidas por el símbolo hash “#”, son usados como etiquetas o marcas de palabras claves en los tweets. Las compañías usan los *hashtags* para categorizar sus tweets y atraer a usuarios que, pueden no ser seguidores (*followers*) de sus páginas, marcas, productos, servicios o eventos. Cuando se busca por *hashtags*, podemos encontrar tweets de un específico tema, más fácilmente que, buscando por todo el texto del tweet. Los *hashtags* también permiten reconocer cuáles son los temas más recurrentes (*trending topics*), para ser fácilmente analizados.

Los hipervínculos (*Hyperlinks*), en redes sociales se refieren especialmente a páginas web. En el caso de twitter, todos los hipervínculos (URLs) posteados en los tweets

son automáticamente reducidos a máximo 22 caracteres por el servicio de twitter t.co. Este servicio permite compartir largos URL sin exceder el número de caracteres límite, a la vez que revisa estas URL de listas peligrosas, para proteger a los usuarios. Las compañías usan estos hipervínculos para guiar a los usuarios a sus sitios web, blogs u otros sitios de su interés.

Los *Retweets* (RT), son tweets que son re-posteados de un tweet original, por otro usuario y aparece como (RT@usuario). Los usuarios de twitter usan esta característica para compartir información con sus *followers*. Cualquier tweet puede ser reenviado (*retweet*), así no sea seguidor y esos tweets serán vistos por otros como un tweet normal. Las compañías utilizan los *retweets* para distribuir positivos e interesantes tweets de su compañía, posteados por diferentes usuarios. Es una manera muy efectiva de conseguir seguidores.

Las menciones (*Mentions*), corresponde a la parte de los tweets que contiene nombre del usuario, precedido por el símbolo “@”. La mención es útil si se quiere seguir a un usuario, que puede estar enviando información a sus seguidores o no seguidores. Las compañías pueden enganchar directamente a los usuarios, mencionándolos en sus tweets. Monitorear las menciones, permite a las compañías resolver los problemas, antes que escalen demasiado, conservando su reputación y dando una respuesta personalizada a los problemas (Becerra, 2016).

Algunas compañías han tomado la iniciativa de solicitar a sus consumidores un feedback para obtener una evaluación objetiva y cuantitativa de la calidad de sus productos o servicios. Este feedback puede ser *minado* y evaluado respecto a los sentimientos relacionados con los productos o servicios. Estos sentimientos pueden ser clasificados como positivos, negativos o neutros. Esta evaluación sobre los sentimientos (*Data mining*)

es muy efectiva para la toma de decisiones, ya que puede customizar los productos o servicios analizados.

Los tweets alcanzan millones de usuarios más rápido que las canales tradicionales, por lo tanto, esto es una ventaja para social media marketing; ya que es de muy bajo costo y de rápidos y amplios resultados. Adicionalmente, la efectividad de las campañas puede ser monitoreadas en tiempo real, con herramientas como el análisis de sentimientos (*Sentiment Analysis- SA*).

3.6 Análisis de Sentimientos

El análisis de sentimientos (SA) tiene por objetivo determinar las opiniones, emociones y actitudes reportadas en fuentes como documentos, textos cortos, blogs, entre otros (Pang, Lee, & Vaithyanathan, 2002). Los consumidores y usuarios han incrementado sus voces para expresar sus sentimientos, los cuales pueden ser positivos, negativos o neutros, en la forma textual de los post en social media, para un público en general. El análisis de sentimiento ha sido aplicado para predecir los destinos más apetecidos, las opiniones sobre debates políticos y su subsecuente elección entre los votantes, ventas de productos, etc. (Zhu, 2010).

Debido al volumen cada vez mayor de datos, es necesario utilizar soluciones como *machine learning* (aprendizaje automático), que han propuesto análisis de sentimientos en línea, sobre datos de texto. El Análisis de sentimientos intenta hacer de manera automática, la determinación de las opiniones desde el lenguaje natural escrito. Es hacer parte de un nuevo paradigma del cómputo, conocido como *Tecnologías Cognitivas*.

El análisis de sentimientos se refiere a la categorización de documentos no estructurados, donde se extrae la actitud emocional de las personas, según las expresiones de sus respuestas. Este análisis se focaliza en la expresión de las opiniones en el texto más que en los temas tratados (Bhuiyan, Josang, & Xu, 2010).

Hay dos enfoques para la realización del análisis de sentimientos:

- *Machine Learning*: La evaluación de los sentimientos en las máquinas de aprendizaje automático (*machine learning*) usan las características lingüísticas y la sintaxis de los textos. *Machine Learning* envuelve dos variables: Parámetro y valor. El parámetro representa la palabra como tal. El valor representa su frecuencia. Cada palabra es reseñada usando API (*Application Programming Interface*), donde a la vez se anota su frecuencia. Este enfoque puede dividirse en dos métodos: Supervisados y No Supervisados. Los datos etiquetados son usados en el método supervisado, mientras en los No Supervisados, no hay datos etiquetados. La diferencia entre los dos métodos es que, para el caso de los supervisados, es necesario una clasificación previa de los datos, mientras los no supervisados esta clasificación no se realiza. En ambos existe un algoritmo que detecta y clasifica de manera automática los parámetros (palabras) acordes a sus sentimientos (Colleoni et al., 2011).
- *Lexicon-based approach*: Involucra el cálculo de la orientación de los documentos a partir de la polaridad y la fuerza de las palabras, las frases o los textos. En este caso, se asume que las palabras tienen polaridad propia, es decir que las palabras tienen una polaridad inherente, independiente del contexto donde se aplica. La clasificación implica tener previamente definido instancias de textos o frases. Estas instancias las hacen unos diccionarios, los cuales se crean manual o

automáticamente a través del uso de palabras semillas (set de palabras que tienen fuertes asociaciones positivas o negativas). Generalmente la orientación de los documentos está determinada, por la agregación de los valores de las palabras y como estas están ranqueadas en los diccionarios. Los más populares diccionarios son *WordNet*, *General Enquirer*, *ANEW (Affective Norms for English Words)* (Aldahawi, 2015).

En el caso de la clasificación de sentimientos con *machine learning*, primero un set de datos es creado manualmente, codificando unas frases de acuerdo con su sentimiento, entonces un algoritmo es entrenado de acuerdo con esta clasificación previa. El resultado es que el algoritmo es ahora capaz de clasificar un nuevo objeto (opinión) de acuerdo con la polaridad o valencia. En el caso del enfoque por *lexicon-based*, la creación de conocimiento empieza con listas de palabras que son pre-codificadas por su polaridad y algunas veces por su fuerza emocional y ocurrencia, permitiendo que se clasifique un texto por su polaridad. El análisis más común de sentimientos es por su polaridad y en algunos casos por la fuerza de la misma.

En el caso de las listas de palabras con clasificación de valencia emocional, *ANEW*, esta es considerada como la referencia más importante y consiste en un set de material verbal, que clasifica las palabras, en términos de placer, excitación y dominancia, para crear un estándar en el uso para las emociones. En el caso de twitter, Nielsen ha propuesto una nueva lista de palabras, que contiene 2477 palabras únicas y un ranking de -5 (muy negativo) a +5 (muy positivo), y no utiliza placer o dominancia. Nielsen ha usado esta lista de palabras en más de 1000 tweets, para compararla con *ANEW* y ha observado un mejor

desempeño en su set de palabras para la valoración de las polaridades del sentimiento (Oakleaf, 2009)

Ahora, dentro de los problemas que podemos tener al hacer un análisis de sentimientos en twitter, que impactan de manera negativa en la precisión del análisis, tenemos:

- Contexto: Cuando un texto expresa una opinión, el contexto juega un amplio papel para determinar la verdadera polaridad del texto. Un texto se puede considerar positivo o negativo dependiendo del escenario.
- Presencia de términos ambiguos: Los tweets actualmente contienen muchos términos ambiguos, así como palabras mal escritas, abreviaciones o *slangs*, que son irregularidades que no pueden ser analizadas bajo el lenguaje estándar. Por lo tanto, pueden afectar significativamente la precisión del procesamiento de lenguaje natural (*Natural Language Processing – NLP*), utilizado por la herramienta de análisis de sentimientos en twitter.
- Diversidad en el uso de expresiones: Hay muchas maneras de expresar los sentimientos, pero si estos se hacen de manera irónica, lo positivo puede quedar como negativo y viceversa. Lo cual genera un conflicto para la precisión del análisis. Ahora, según el *ajuste* que se haga del programa y del algoritmo, puede que el análisis sea demasiado *sensitivo*, para descartar contenido relevante o demasiado general, haciendo que el análisis asuma información que no es relevante para el mismo (Boumaiza, 2015).

3.7 Minería de Datos: Los datos como fuente de valor

La consecuencia natural de la internet y de la web 2.0 es la explosión de datos. Un apalancamiento apropiado viene de la administración de los datos. Pero los datos solos no nos pueden brindar los elementos necesarios para tomar decisiones. Los profesores Mohanty, Jagadeesh y Srivatsa, (figura 2) describen la transformación de los datos en un modelo piramidal, que va desde los datos bases hasta un conocimiento accionable (Mohanty, S., Jagadeesh, M., Srivatsa, 2013).

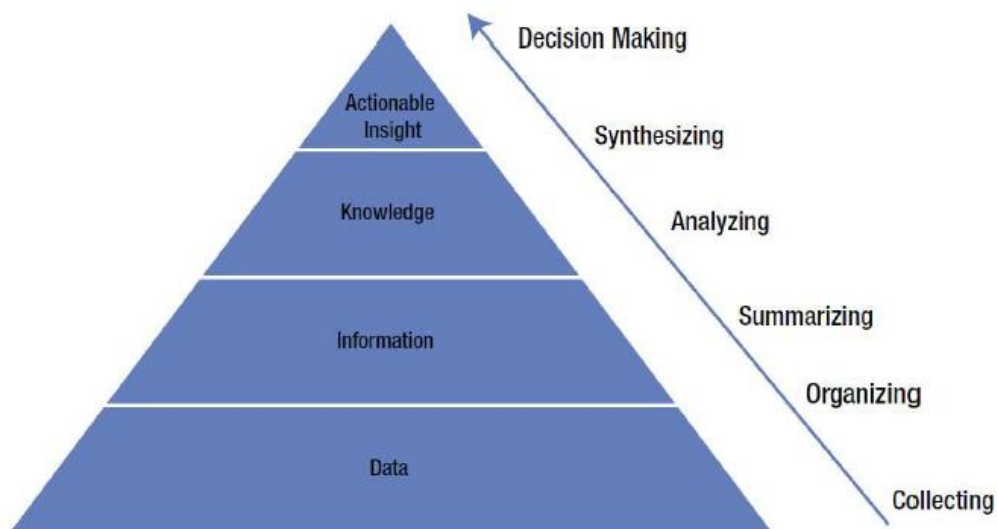


Figura 2 Transformación de datos a conocimiento Accionable (Mohanty, S., Jagadeesh, M., Srivatsa, 2013)

Porque los datos y la creación de valor para las compañías, son dos cosas distintas, las compañías necesitan extraer conocimientos útiles de los datos. Los negocios hoy requieren estar orientados hacia los datos para ser exitosos en el tiempo. Por lo tanto, es necesario pensar como los datos disponibles se pueden aplicarse a los problemas

específicos de las compañías. El mayor desafío y obstáculo es implementar este concepto a la práctica (Etter et al., 2017)

El proceso de la extracción del conocimiento a partir de los datos se conoce como Minería de datos o *Data mining* (Kumlin, 2016). En la minería de datos son muy importante los procesos y en forma práctica sin minería de datos no es posible extraer conocimiento de los mismos. En este orden de ideas, la minería de datos consta de los siguientes procesos (figura 3):

1. Integración de datos: Se conoce como ETL (*Extraction, Transform and Load data*)
2. Limpieza de datos: Selección de datos relevantes para el análisis
3. Consolidación de datos: Transformación de la información en formatos apropiados para la realización de operaciones básicas de resumen o agregación
4. Minería: Donde se aplica programas para extraer patrones de información
5. Presentación y Evaluación: Representación y conocimiento de los patrones extraídos

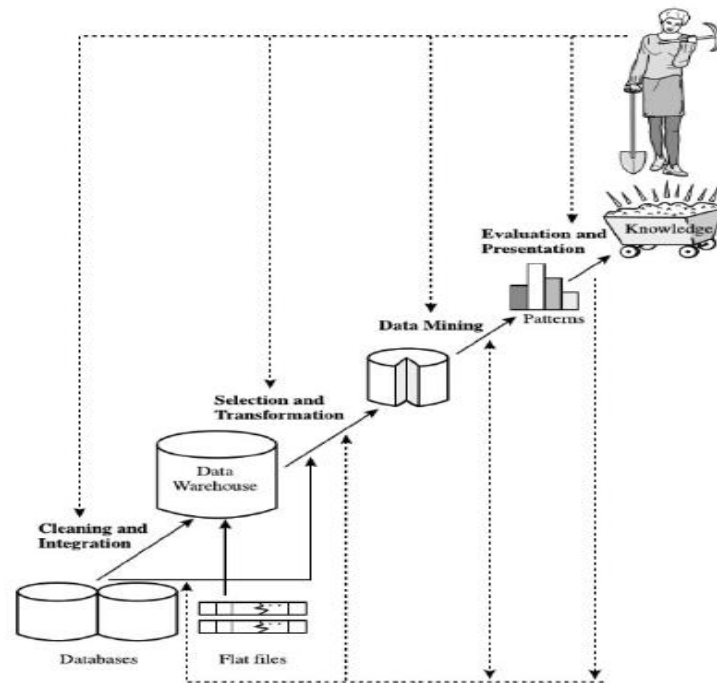


Figura 3 Procesos de Data Mining (Kumlin, 2016)

Los datos en redes sociales ayudan a obtener en tiempo real *insights* sobre el desempeño de la compañía, las experiencias del consumidor y la reputación corporativa. La equivocación más común de las empresas es focalizar su gestión en social media, en las herramientas y en las plataformas; y no en sus objetivos de negocio. El principal reto es identificar las metas y las métricas que le dan sentido a las empresas, ya que medir por medir no es una meta. El siguiente paso es encontrar las herramientas y los canales de interés para las compañías. Uno de los medios para capturar los datos en social media y que tiene mayor relevancia es el análisis de sentimientos o *sentimental analysis* (Kumlin, 2016).

El análisis de sentimientos hace parte de las herramientas de la minería de datos y consiste en, de manera analítica, determinar las emociones detrás de los textos en las redes.

También es llamado *opinión mining*. El análisis de sentimientos detecta de manera subjetiva la polaridad de un texto, clasificándolo en opinión negativa o positiva (Boumaiza, 2015).

Hay diferentes maneras de clasificar la polaridad de un texto. Se puede realizar a partir de un programa de aprendizaje automático supervisado (*machine learning*) o mediante programas de léxico no supervisado. Este último es el más utilizado para el desarrollar análisis de sentimientos (figura 4).

De acuerdo con un artículo escrito por el experto en tecnología en El País de España, Enrique Martín, el aprendizaje automático o *machine learning* cambiará la forma como vivimos y trabajamos. Esta tecnología cognitiva ya está presente en nuestras vidas diarias: teléfonos, software de mapas, incluso en el algoritmo de Facebook.

Según Google, *machine learning*, se deriva de la Inteligencia artificial y consiste en desarrollar códigos o procesos que le permitan a las máquinas aprender a través de un conjunto de datos que se le entregan y que, conforme al desarrollo de la máquina, se van corrigiendo. En otras palabras, es moldear la experiencia y la manera en que un cerebro (en este caso, una máquina) la procesa. El aprendizaje automático reduce el tiempo en el que se crean pronósticos, cada vez más precisos y fiables y pueden tener un impacto en la capacidad de las empresas para planificar, hacer presupuestos, identificar anomalías en grandes conjuntos de datos, etc. (Espectador, 2017).

La inteligencia Artificial no es una ciencia ficción, ya está incorporada en los productos que usamos a diario. La asistente Siri de Apple, las recomendaciones de libros de Amazon, el suministro de noticias de Facebook y la lista de descubrimiento semanal de

música de Spotify, son ejemplos de servicios impulsados por algoritmos de aprendizaje automático.

Esta ciencia que ha estado disponible durante décadas, está disfrutando de un renacimiento, impulsado por la avalancha de datos creados por teléfonos inteligentes y por sensores, y debido al poder de supercomputación disponible para procesar los datos. Según la firma de investigación tecnológica Tractica, el mercado de Inteligencia Artificial crecerá de U\$D 643,7 millones en 2016 a U\$D 36.800 millones para el 2025 (Espectador, 2017).

Las técnicas como el aprendizaje profundo y las redes neuronales supuestamente imitan al cerebro humano: ellas detectan patrones amplios en enorme conjunto de datos con el fin de catalogar imágenes, de reconocer voces y de tomar decisiones (Han, J., Pei, J., Kamber, 2006).

El siguiente paso de la inteligencia artificial será generar un algoritmo al que no se le tendrá que *enseñar* una habilidad específica, como un juego de ajedrez, sino que lo adquirirá por ensayo y error, precisamente cómo lo hace un niño. Ciertas compañías como *DeepMind* de Google, con sede en Londres, están trabajando para que esto se haga realidad (Espectador, 2017).

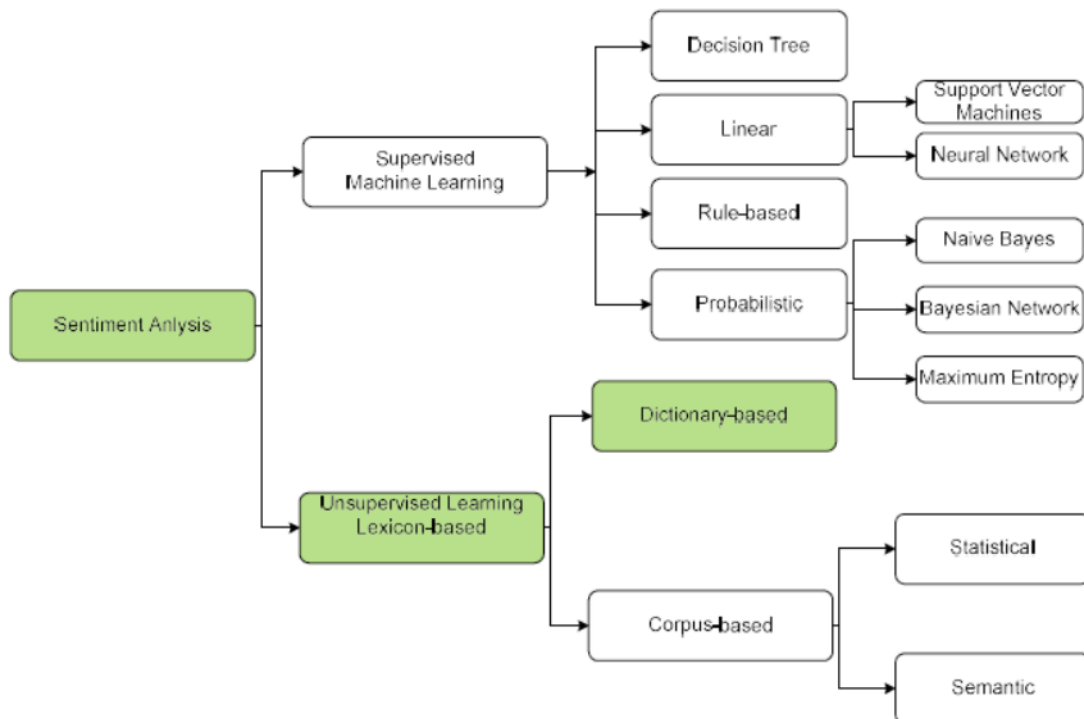


Figura 4 Técnicas de Análisis de Sentimientos (Kumlin, 2016)

4. Marco Referencial

4.1 Avianca: Empresa, Huelga de Pilotos y Reputación Empresarial

Avianca (Aerovías del Continente Americano), es la marca comercial que representa a las aerolíneas latinoamericanas integradas en Avianca Holdings S.A. El equipo humano está integrado por más de 19.000 colaboradores. Con las aerolíneas de Avianca, especializadas en el transporte de pasajeros y carga, se atienden en forma directa más de 108 destinos en 26 países de América y Europa, a bordo de una flota de 189 aeronaves de corto, mediano y largo alcance.

Tiene membresía con Star Alliance y ofrece conectividad a más de 1.300 destinos en 191 países de los cinco continentes. Tiene un programa de lealtad LifeMiles, que cuenta con 7.2 millones de socios, con múltiples beneficios y ventajas exclusivas que ofrece este esquema de fidelidad. Tiene presencia en comunidades oficiales de Facebook, Twitter, YouTube, Pinterest y Foursquare.

Su Misión es: Conectar al mundo con Latinoamérica, buscando dar siempre más. Su Visión a 2020 es: Con la mejor gente y tecnología para una experiencia excepcional, seremos la aerolínea latinoamericana preferida en el mundo; y sus principales valores son: Seguridad, servir con pasión, disfrutar lo que se hace e innovar con impacto sostenible (Avianca, 2018).

German Efromovich, principal socio de Avianca, compro la compañía en el 2004, casi quebrada (llevaba casi dos años acogida a la ley de quiebras de Estados Unidos) y sin revisión del estado de la empresa (conocido como *due diligence*), se comprometió a inyectarle 65 millones de dólares a través de su holding *Synergy*, a cambio que el grupo Santo Domingo le diera el 73% de sus acciones y que ellos asumieran directamente la mitad del pasivo pensional de la aerolínea (unos 120 millones de dólares).

Así Efromovich se quedó con Avianca, le quito cargas financieras y no pago por las acciones. En cambio, uso el dinero prometido para renovarla y ponerla a crecer. Con el capital que se le inyectó, se renovó la flota y la marca e incremento sus frecuencias. A la vez Efromovich compro las acciones restantes a la Federación Nacional de Cafeteros; y nombro al ex ministro Fabio Villegas como presidente, logrando así que, en el 2006, la empresa diera utilidades por primera vez en 12 años. En el 2010 se fusionó con la salvadoreña Taca, lo que le permitió competir con Copa-Continental y con LAN y dejó a

Avianca muy fuerte en el mercado centroamericano. Efromovich mantuvo el 78% de Avianca Taca y los dueños de Taca, la familia Kriete, retuvo el 22 por ciento restante a través de *Kingsland Holding*.

Pero en los últimos años, la situación se ha deteriorado, tanto para Avianca, como para sus socios: AviancaTaca; ya que se han quedado pequeños frente a la fusión de Lan y Tam (hoy Latam), que tiene como accionista a Qatar Airways, con una intención de operar conjuntamente con American Airlines. De otra parte, Aeroméxico busco asociarse con Delta e Interjet con American o United. A nivel local, han perdido mercado interno, dado el auge de las aerolíneas de bajo costo como Viva Colombia o Easy Fly; además de la férrea competencia para los vuelos internacionales.

En el 2015, los títulos de Avianca Holdings (dueña de las aerolíneas, excepto Avianca Brasil), costaban una cuarta parte de lo que valían dos años atrás, y aunque en 2016 se recuperaron, solo doblaron su precio. En el 2016 de igual modo, un informe de gestión mostro que la posición financiera de Avianca se desmejoró frente al 2015, disminuyendo su capital de trabajo, lo que significa que, aunque funcionaba con ganancias, las deudas se estaban convirtiendo en una carga que no la dejaba crecer, ni tener dinero para invertir.

En medio de estas turbulencias, en abril de 2016, la empresa anunció que iniciaba la búsqueda de un socio estratégico y mencionó que *Synergy*, había puesto como garantía de un préstamo la mayoría de sus acciones de la empresa. En julio de ese año, *Bloomberg* reveló que el beneficiario del préstamo era el fondo inversionista *Elliot Management* y en septiembre del mismo año, varios medios internacionales mencionaron que el préstamo era de 800 millones de dólares.

Luego, tras acercamientos con United Airlines, Delta y Copa, un reporte de la calificadora Fitch, mostraba que Avianca tenía problemas de liquidez, y entre el 4 de enero y 3 de febrero de 2017, los títulos cayeron un 30% en Wall Street. Con este golpe, Efromovich detuvo el anuncio de la alianza con United Airlines en febrero de 2017; lo que también llevó a una demanda por parte de Kingsland, socio de Avianca, alegando que Efromovich, como principal accionista, eligió a United por su interés personal y no por la aerolínea; ya que supuestamente United era compañía que menor valor le daba a Avianca en el negocio, pero facilitaba el pago a Elliot, fondo que le había prestado dinero.

Después de esto, tuvo dificultades en la entrada a Argentina, en marzo de 2017, pero logro un avance laboral al firmar un acuerdo colectivo con el sindicato ODEAA (Organización de aviadores de Avianca), creado en el 2004, y que evitaría cualquier problema laboral con la mitad de los pilotos, por lo menos hasta 2021; pero faltaba ACDAC (Asociación de Aviadores Civiles).

La relación entre ACDAC y Efromovich nunca fue buena. Este sindicato nació en 1949 y tenía una participación poderosa en Avianca. En el 2013, venció la convención colectiva de trabajo, que habían firmado en el 2009, y las negociaciones no llegaron a un acuerdo y tras agotarlas, Avianca ofreció un plan de beneficios y bonos especiales a los Pilotos, con la condición que se salieran de la convención, es decir del sindicato de ACDAC. En el 2015 y luego de muchas batallas legales, ACDAC, logro que todos sus miembros recibieran los beneficios, ya entregados, como si fuera una convención colectiva.

Pero en el 2014, por efectos de la reforma tributaria de 2012, a todos los pilotos les tocaba pagar más impuestos, a través de su declaración de renta, lo cual llevo a que se redujeran sus ingresos mensuales en un 15%. El golpe fue tan fuerte que ambos sindicatos,

ODEAA (entre 600 y 800 Pilotos) y ACDAC (entre 250 y 500), se pronunciaron; y en conjunto con la empresa, buscaron revivir la gabela que los escindiera de este impuesto, que habían tenido por 30 años y que el Estado había subsidiado. Algunos congresistas los apoyaron, pero finalmente fracasaron en este intento.

En 2017, se tenía que negociar de nuevo la convención y dado un pronto cierre de negociaciones con ODEAA, la gente de ACDAC debía tomar la decisión de irse a paro o no. La decisión que finalmente tomaron fue decretar la huelga el 15 de septiembre de 2017, iniciando el cese de actividades el día 19 de septiembre. Las exigencias de los pilotos iban desde una reducción de la jornada laboral a 40 horas mensuales, hasta un incremento salarial del 50%, para quedar nivelados con el sueldo de los pilotos de otros países en los que opera la aerolínea. Para Avianca, este pliego de peticiones sumaba más de USD 200 millones, una suma imposible para una compañía, cuyas utilidades netas alcanzaron USD 22 millones de dólares en primer semestre de 2017.

Durante la huelga Avianca, a través de su dueño, expresó que la huelga era ilegal e inmoral; ya que según explicó era ilegal porque la aviación es un servicio público esencial e inmoral con los usuarios de Avianca. La compañía desarrollo una fuerte campaña de comunicación que busco que la opinión apoyará a la empresa y no a los pilotos. En redes sociales difundió un extracto de las reivindicaciones de ACDAC, que suscitaron indignación, ya que la mayoría de los colombianos, ganan muy por debajo de los sueldos de un piloto, que puede llegar a 30 millones mensuales. A través el hashtag #YoNoParo, la compañía quería demostrar el compromiso de sus otros trabajadores con la operación de la misma, que incluyo comerciales en televisión y abrió una convocatoria para enrolar pilotos extranjeros, para mitigar la crisis.

Por otra parte, ACDAC respondió con mensajes de los familiares de los pilotos y sus explicaciones de cómo se habían deteriorado sus ingresos, explicando que trabajan igual que los pilotos de Avianca de otros países, pero su salario era inferior. Buscaron el apoyo de organizaciones sindicales, congresistas de izquierda y en redes sociales impulsaron el hashtag #TodosSomosPilotos (Lewin, 2017).

Finalmente, y luego de 51 días, los pilotos de ACDAC votaron a favor de levantar la huelga, el 10 de noviembre de 2017. La legalidad de la huelga fue siempre un tema de debate durante y después de la huelga. El tribunal de Cundinamarca la declaró ilegal el 6 de octubre de 2017 y el tribunal de arbitramento nunca pudo operar por las acciones legales llevadas a cabo por la empresa y el sindicato. La Corte Suprema de Justicia ratificó el fallo de ilegalidad el 29 de noviembre de 2017, dando como razones que la huelga fue confirmada por una minoría de la empresa y que esta recae sobre un servicio público esencial.

Esta huelga fue uno de hechos económicos más importante en el 2017 en Colombia. Avianca había despedido casi 50 pilotos a febrero de 2018. De acuerdo a una carta enviada a los accionistas, el presidente de Avianca, Hernan Rincon, en febrero de 2018, mencionó que el año 2017, pudo haber sido el mejor de la historia de la compañía y a pesar de los problemas relacionados con la huelga, obtuvo un aumento del 5.13% en su EBITDA (USD 885 millones) comparado con el 2016.

Aunque no hay consenso, se menciona que durante la huelga hubo pérdidas por entre USD 800.000 y USD 1 millón diarios, más de 12.300 vuelos se cancelaron y 400.000 pasajeros se vieron afectados, durante un paro que duro 51 días; el más largo conocido hasta ahora de una aerolínea comercial.

De otra parte, en octubre de 2017, se presentó el informe *Monitor Empresarial de Reputación Corporativa (Merco)*, que evalúa más de 33.000 encuestas, con 18 fuentes de información, donde empresarios y diferentes stakeholders, valoran las empresas en las variables de las que son expertos. Esta evaluación establece un ranking de las 100 empresas con mejor valoración de su reputación en Colombia, encontrándose a Avianca dentro de top 10 de las compañías mejor valoradas (puesto 7 entre 100 – año 2017) y siendo la primera en el segmento de transporte de viajeros.

Este ranking se hace en Colombia en alianza con el Centro Nacional de Consultoría, donde ejecutivos, CEO de dirección de empresas que facturan más de USD 30 millones, eligen las diez empresas de mejor reputación corporativa en Colombia para cada sector. De allí, surgen casi 800 empresas para un ranking provisional, que pasan una valoración de expertos, donde diferentes stakeholders valoran de 0 a 100 a cada una de las empresas del ranking provisional en las variables donde son expertos. Analistas financieros, gobierno/administración, ONG, sindicatos, asociaciones de consumidores, influenciadores y catedráticos de universidad hacen sus análisis.

Posteriormente, los técnicos de la Compañía que publica el análisis Merco, evalúan los principales indicadores (200) de las empresas incluidas en el ranking provisional; donde se incluye el estudio de Merco de Consumo (en el que la población/consumidores valoran la reputación comercial de las empresas en función de diez variables), al igual que Merco talento (las empresas con más capacidad de atraer y retener Talento), para dar con el ranking Corporativo definitivo. El seguimiento de la metodología establecida por Merco para la elaboración del ranking de empresas con mejor reputación en Colombia ha sido

objeto de la revisión independiente por parte de KPMG en Colombia. En total, se evaluaron 33.685 encuestas con 18 fuentes de información (Dinero, 2017).

Es muy importante, de la misma forma, ver el comportamiento de la acción de Avianca en la bolsa de New York, antes, durante y después de la huelga. A continuación, se anexa la figura 5, que muestra este comportamiento.



Figura 5 Acción Avianca Holding SA - NYSE:AVH (Holding, 2018)

5. Metodología

5.1 Tipo y Diseño de la Investigación

El tipo de investigación propuesta es una investigación descriptiva, ya que se pretende dar una descripción de la Reputación Corporativa en social media de la empresa Avianca, mediante minería de datos para unos periodos de tiempo específicos (Bernal, 2006)

Por otra parte, el Diseño de la investigación es no experimental (cuasi experimental), de tipo longitudinal, que busca comparar la Reputación Corporativa de Avianca en social media, específicamente en Twitter; antes, durante y después de la huelga de Pilotos en septiembre de 2017 (Bernal, 2006)

El tamaño de muestra de datos analizados es de 74.027 tweets, durante los periodos de análisis determinados.

Los tweets analizados corresponden a una serie de tiempos que van desde: Antes de la huelga de Avianca (agosto 28/17 a septiembre 15/17), durante la huelga (septiembre 16/17 a noviembre 15/17) y después de la huelga (noviembre 16/17 a febrero 28/18).

En la siguiente tabla se mencionan los periodos, las menciones y los hashtags utilizados para clasificar los tweets en los diferentes periodos analizados:

Periodos	Antes de la Huelga Agosto 28/17 – Septiembre 15/17	Durante la Huelga Septiembre 16/17 – Noviembre 15/17	Después de la Huelga Noviembre 16/17 a Febrero 28/18
<i>Mención @</i>	@AVIANCA	Pilotos: @ACDAC1	@AVIANCA
		Empresa: @YoSoyAvianca	
<i>Hashtag #</i>		Pilotos: #PilotoSomosTodos	
		Empresa: #YoNoParo	

Tabla 1 Periodos Analizados Tweets - Antes, Durante y Después de la Huelga Avianca - Elaboración Propia a partir de #Hashtag y @Mentions

Los periodos se establecieron, según juicio del autor, para que sirvieran de puntos de inflexión, en el análisis de la reputación de Avianca. El periodo inicial tomo quince días antes de la huelga (agosto 28/17 a septiembre 15/17), que busca establecer un punto inicial del estado de la reputación de Avianca.

De igual manera, para el periodo concerniente a la huelga (septiembre 16/17 a noviembre 15/17), el análisis se hizo tanto para los pilotos, como para la Empresa. Esto permite contraponer y visualizar de una mejor manera, cuáles fueron las polaridades (sentimientos) que tuvieron los usuarios de twitter frente a estos dos grupos enfrentados

durante la huelga. Además de evaluar si la gestión hecha en redes sociales, tanto los pilotos, como la empresa, tuvo buena acogida y repercusión.

Finalmente, el periodo correspondiente a después de la huelga, es el más amplio (noviembre 16/17 a febrero 28/18), ya que Avianca tuvo serios inconvenientes en ese momento, e inclusive más allá del periodo analizado, para restablecer de manera oportuna sus servicios, razón por la cual era necesario tener un mayor tiempo de análisis.

Según el problema de investigación planteado y los objetivos de este análisis, se escogió a Twitter como la plataforma de social media sobre la cual se hará la minería de datos, respecto a la empresa seleccionada (Avianca), ya que permite un análisis de mayor profundidad por su gran contenido de opiniones.

5.2 Colección de datos

Por el tamaño de la muestra de datos encontrados (más de 70 mil), fue necesario *correr* los datos en un servidor de computo dedicado a esta tarea durante tres semanas, esto mediante la utilización de un programa en lenguaje *Python*, que analizo el contenido de los tweets y obtuvo unos datos de salida, que van a ser analizados más adelante.

5.3 Extracción de datos

Estos se hicieron a través de una revisión y elaboración de API (*Application Programming Interface*) de la red social Twitter, mediante el siguiente esquema (figura 6):

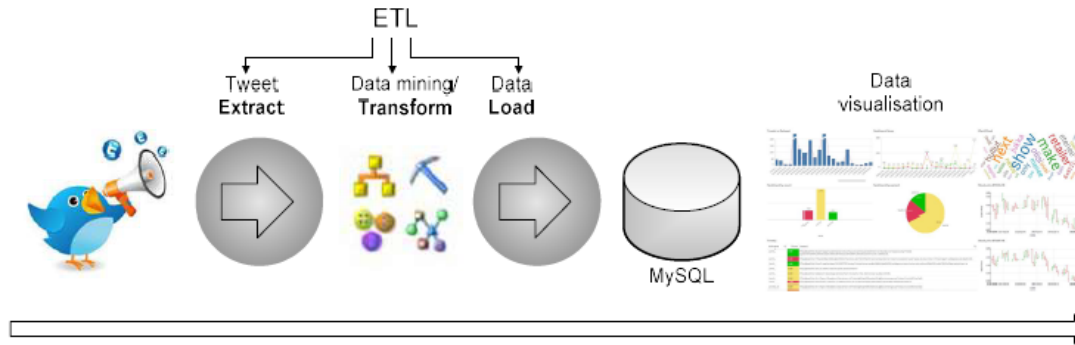


Figura 6. Modelo de extracción de datos. Elaboración Propia

Dentro de este proceso se utilizaron herramientas del programa *Python* (Bibliotecas) que permitieron hacer la comparación de los textos de los tweets contra datos y textos, ya etiquetados, con valores de polaridad positiva como negativa.

Se realizó una autenticación para la aplicación API en twitter y luego se implementó la misma, en servidores dedicados a esta tarea para extraer las *entidades* (información), solicitadas en la API, para que más tarde esta información pudiera ser almacenada en archivos de extensión .csv, para un posterior análisis y visualización.

Dentro de las *entidades* solicitadas en la aplicación se extrajeron las siguientes:

- *Date*: Fecha y hora del tweet
- *Retweets*: Numero de retweets correspondientes al tweet extraído
- *Favorites*: Numero de “likes” correspondiente al tweet extraído
- *Text*: Correspondiente al texto del tweet extraído
- *Mentions*: Corresponde a los @usuario mencionados en el tweet
- *Hashtags*: Corresponde a los #hashtags mencionados en el tweet
- *ID*: Corresponde a un número de identificación del tweet

- *Permalink*: Corresponde a la URL del tweet extraído

La aplicación de la API nos dio como resultado la siguiente información para cada uno de los periodos analizados:

- *Keywords*: Palabras con mayor presencia en los tweets
- *Cantidad*: Corresponde al número encontrado para cada Keyword
- *Tweets Positivos*: Numero de tweets catalogados como positivos
- *Tweets Negativos*: Numero de tweets catalogados como negativos
- *Tweets Neutrales*: Numero de tweets catalogados como neutros

Con los archivos documentados (archivos tipo .csv), que contienen la información extraída y que pueden ser visualizados y analizados, en archivos tipo Excel, se extrajo lo siguiente:

- *Top Posts*: Corresponden a los tweets con mayor número de re-tweet, donde se les analizó:
 - *Polaridad*: Si es negativa, positiva o neutral
 - *Influencer*: Nombre del usuario @username con mayores *retweet*
 - *Número de Followers*: Cantidad de seguidores que tiene el *influencer*
 - *Número de Impression* del tweet: Numero de *retweet*, multiplicado por el Numero de *followers*.

5.4 Visualización y análisis de resultados

Luego de extraídos los datos, es necesario mostrarlos y analizarlos para generar *valor* de la información encontrada y así pasar a la etapa del desarrollo de conocimiento.

El número de Tweets analizados totales fue de *74.027* y sus resultados están representados en las siguientes tablas y figuras, a través de las cuales se describe el análisis de los tweets, para cada uno de los periodos mencionados, bajo el siguiente esquema:

- *Timeline*: Muestra el número de tweets, retweets y likes en el periodo de tiempo.
- *Keywords*: Muestra las palabras predominantes y en forma gráfica muestra su relación de tamaño respecto a las otras palabras recurrentes, en el periodo de tiempo.
- *Top posts, Influencer e Impressions*: Tablas que muestran los tweets de mayor penetración (*retweets*), su polaridad, contenido y el número de impresiones alcanzado. Este número se calcula por la multiplicación del número de *retweets* por el número de seguidores (*followers*) del influenciador (influencer) que generó el tweet original.
- Polaridad tweets: Muestra la relación entre los tweets positivos encontrados y los tweets negativos, en el periodo de tiempo.
- *Sentiment Analisis y Sentiment Score*: Muestra el porcentaje de tweets positivos, negativos y neutros, encontrados en el periodo de tiempo. El *Sentiment Score* se refiere a:

$$\sum \textit{Tweets Positivos} / \sum (\textit{Tweets Positivos} + \textit{Tweets Negativos})$$

Sentiment Score > 50% - Reputación Positiva

Sentiment Score < 50% - Reputación Negativa

Sentimental Score = 50% - Reputación Neutra

Este Score nos indica cual es el sentimiento o la polaridad global de las opiniones que tienen los usuarios de Twitter en el periodo de tiempo, para los *#hashtags* y las menciones *@mencion*, del periodo analizado.

Se considera tambien como la medida de la reputación de la empresa, ya que Reputación Corporativa Online puede ser vista como los sentimientos y las emociones alrededor de una compañía, que si es positiva, tendrá la capacidad de atraer más y más opiniones y sentimientos positivos de sus stakeholders; pero a la vez, lo contrario es igualmente valido (Arvidsson, 2011).

Para representar estas polaridades, se proponen tres tipos de *Emoji* para su representación:



Representa una polaridad o emoción positiva



Representa una polaridad o emoción negativa



Representa una polaridad o emoción neutra

5.4.1 Avianca - Antes de la huelga.

Este periodo corresponde a antes de la huelga de Pilotos de septiembre de 2017. Se estableció como un periodo de referencia, para establecer el estado de la reputación

corporativa antes del punto de inflexión que corresponde a la huelga y su longitud es de 15 días antes del hecho.

El vínculo de análisis fue toda la información (tweets) que tuvieran como mención a @Avianca, dentro de su texto, ya que este representa el usuario principal y válido de la empresa dentro de la red Twitter.

Antes de la Huelga: Agosto 28 a Septiembre 15/17 - Etiqueta: @Avianca

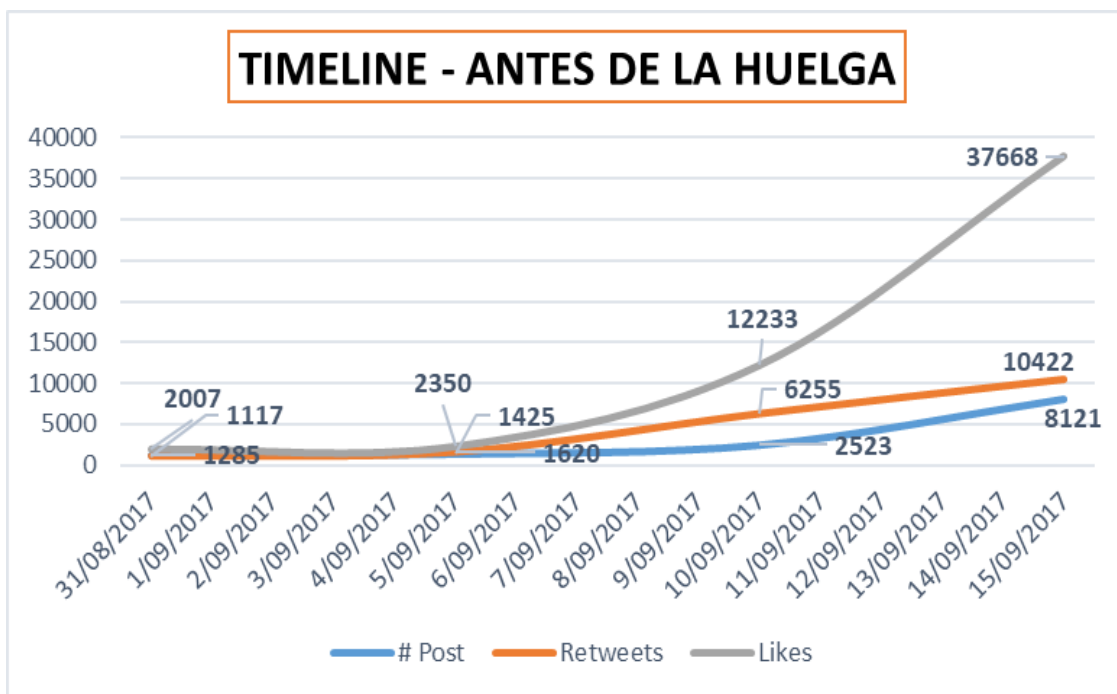


Figura 7. Timeline - Antes Huelga - Elaboración Propia

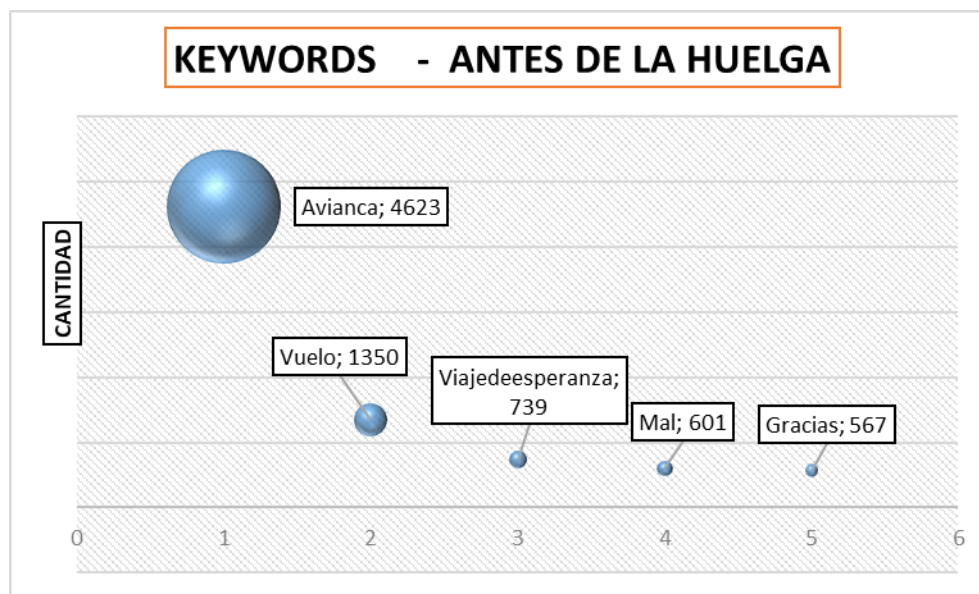


Figura 8. Keywords - Antes Huelga - Elaboración Propia



Polaridad	Fecha	Retweets	Likes	Text	Influencer	Followers	Impression
	28/08/2017	112	43	Confirmamos que este mensaje que circula en las redes sociales es falso, Recomendamos no enviar datos personales. #ProtegeTusDatos pic.twitter.com/8vFMGzxcwCR	@Avianca	850.000	95.200.000
	10/09/2017	681	1408	Después de tantas sonrisas despedimos a @Pontifex_es en este #ViajeDeEsperanza . Mira junto a @ sus mejores momentos abordo de nuestro #B787 pic.twitter.com/ZrzPldiiRJ	@Avianca	850.000	578.850.000
							674.050.000

Tabla 2. Top posts, Influencer e Impressions - Antes Huelga - Elaboración Propia

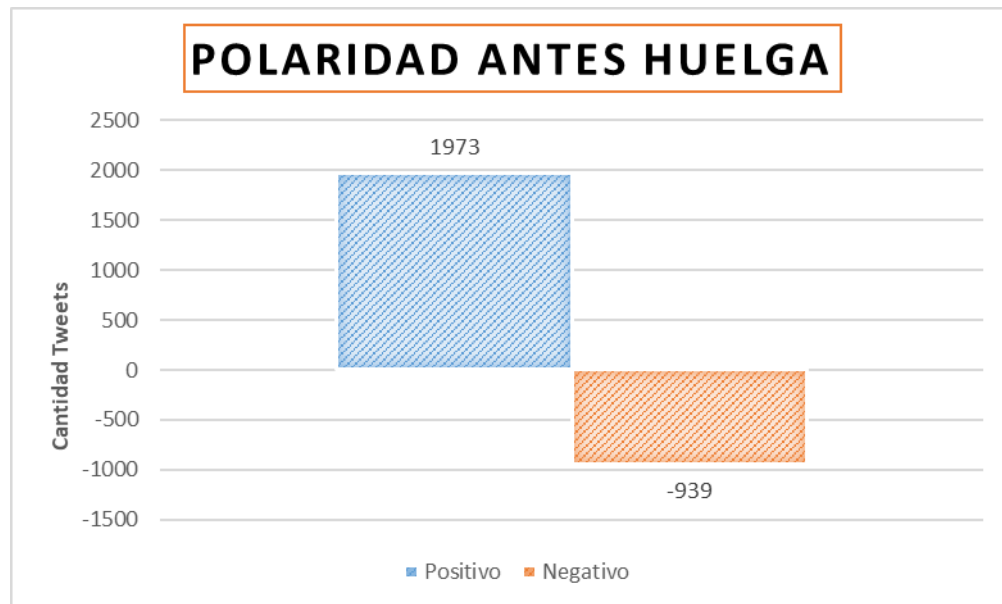


Figura 9. Polaridad tweets - Antes Huelga - Elaboración Propia

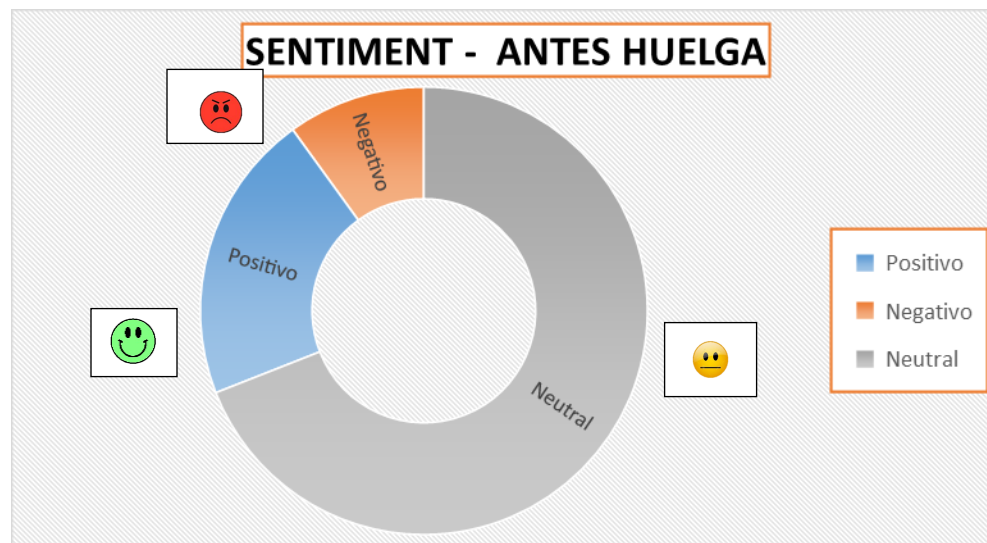


Figura 10. Sentiment - Antes Huelga - Elaboración Propia

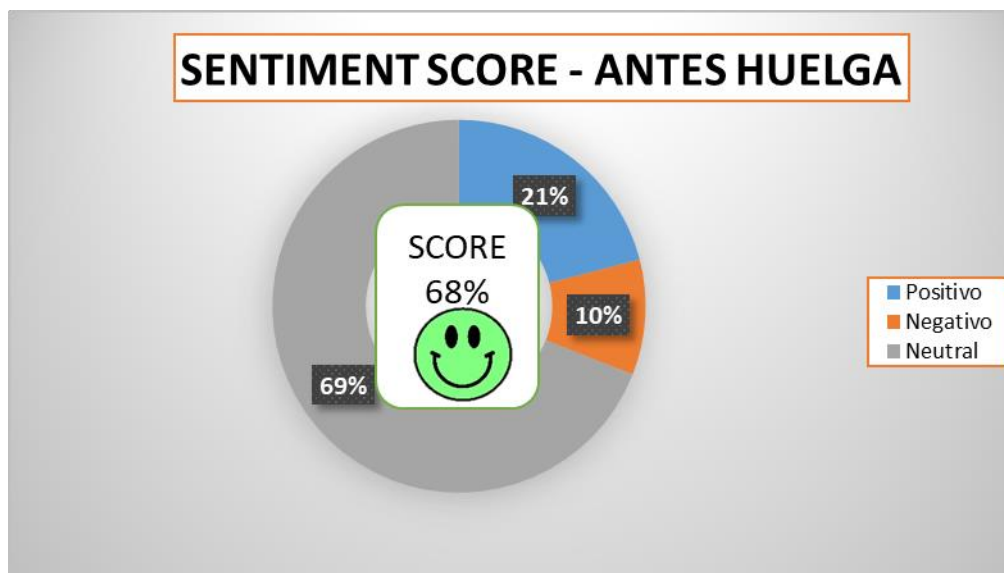


Figura 11. Sentiment Score - Antes Huelga - Elaboración Propia

De acuerdo con las figuras mostradas, como resultado de la extracción de la información para el periodo antes de la huelga, se encontró lo siguiente

- Al final de periodo se observa un alto incremento de la actividad en twitter debido al anuncio de la huelga, aumentando los indicadores de Post, retweets y likes.
- No se observa ninguna mención de huelga en las keywords, lo cual podría significar que la posibilidad de huelga no estaba contemplada para los usuarios de twitter
- Se puede mencionar que el sentimiento hacia la marca Avianca era positivo en términos generales, ya que las menciones/polaridades positivas, superan las negativas, dando un Score del 68%, lo cual significa que, en twitter la reputación se movía en terreno positivo antes de la huelga.

5.4.2 Avianca – Durante la Huelga.

- **PILOTOS**

Septiembre 16/17 a noviembre 15/17

Etiqueta: @ACDC1 y/o Hashtag #PilotosSomosTodos

Durante la huelga, se enfrentaron dos posiciones: Los Pilotos pertenecientes al sindicato ACDAC y la empresa (Avianca). En esta parte se va a mostrar lo extraído, correspondiente a los Pilotos:

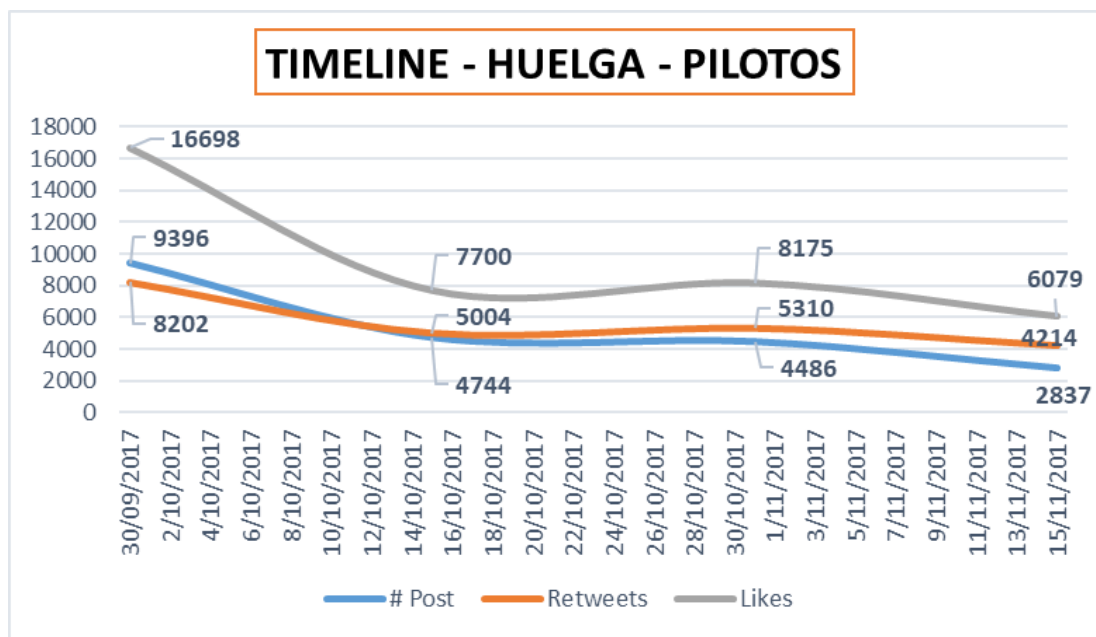


Figura 12. Timeline - Huelga - Pilotos - Elaboración Propia

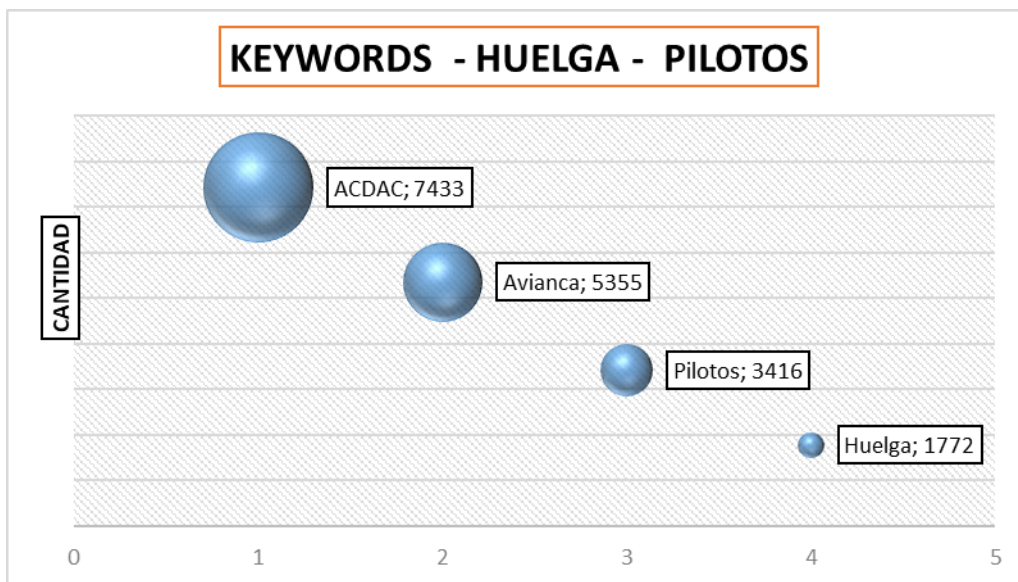


Figura 13. Keywords - Pilotos - Elaboración Propia

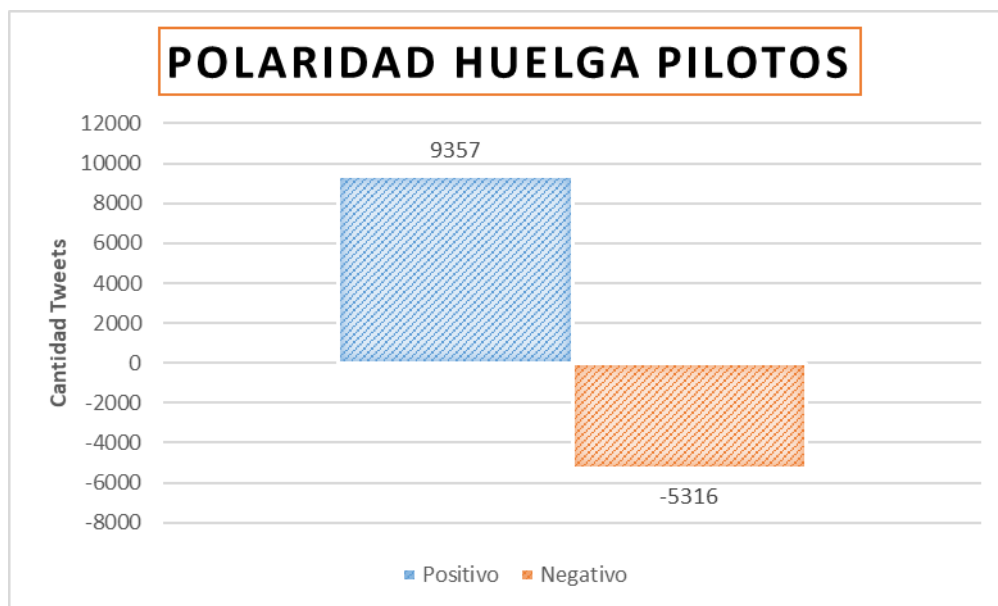


Figura 14. Polaridad tweets - Huelga - Pilotos - Elaboración Propia





Polaridad	Fecha	Retweets	Likes	Text	Mentions	Influencer	Followers	Impressions
	27/09/2017	669	724	Lamentamos que nuestra propuesta final no haya sido aceptada por parte de @Acdac1 en el plazo establecido. #YoNoParo pic.twitter.com/kyP4dbjbSj	@Acdac1	@Avianca	850.000	568.650.000
	7/10/2017	144	131	La unión y Fuerza corean los pilotos de @Acdac1 pic.twitter.com/xBxKKhgm49"	@Acdac1	@AurelioSuarez	58.800	8.467.200
	17/10/2017	328	397	Mensaje de los pilotos de @Acdac1 para German Efromovich y @Avianca pic.twitter.com/UrfMhjlruk	@Acdac1 @Avianca	@HassNassar	677.000	222.056.000
	8/11/2017	203	235	Celebro que el juez haya fallado a favor de @Acdac1 - Suspenden despidos injustificados de pilotos en huelga de #Avianca y solicita reintegro de los trabajadores ya despedidos. Con la movilización se puede! pic.twitter.com/5qyX8zcPRO	@Acdac1	@VictorJCorrea	22.600	4.587.800
								803.761.000

Tabla 3. Top posts, influencer e Impressions - Pilotos - Huelga - Elaboracion Propia

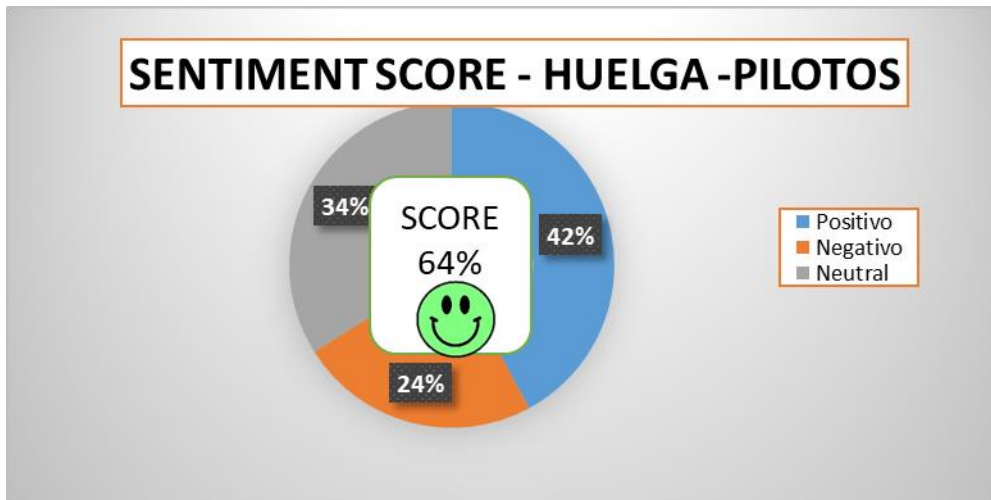


Figura 15. Sentiment Score - Huelga - Pilotos - Elaboración Propia

- **Empresa**

Septiembre 15/17 a noviembre 15/17

Etiqueta: @YoSoyAvianca y/o Hashtag #YoNoParo

Durante la huelga, se enfrentaron dos posiciones: Los Pilotos pertenecientes al sindicato ACDAC y la empresa (Avianca). En esta parte se va a mostrar lo extraído, correspondiente a la empresa:

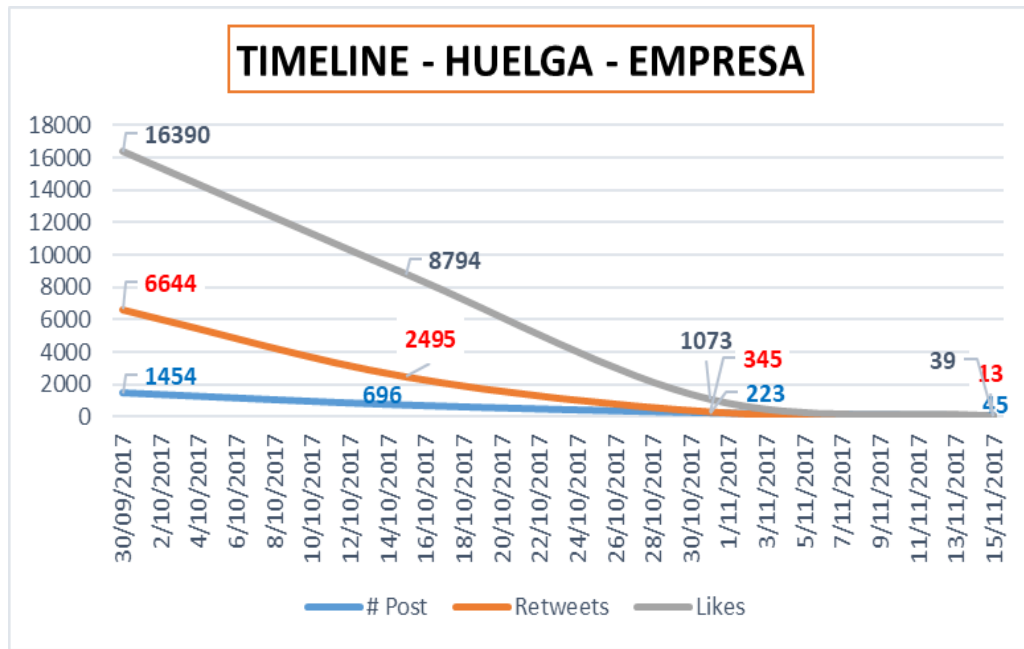


Figura 16. Timeline - Huelga - Empresa - Elaboración Propia

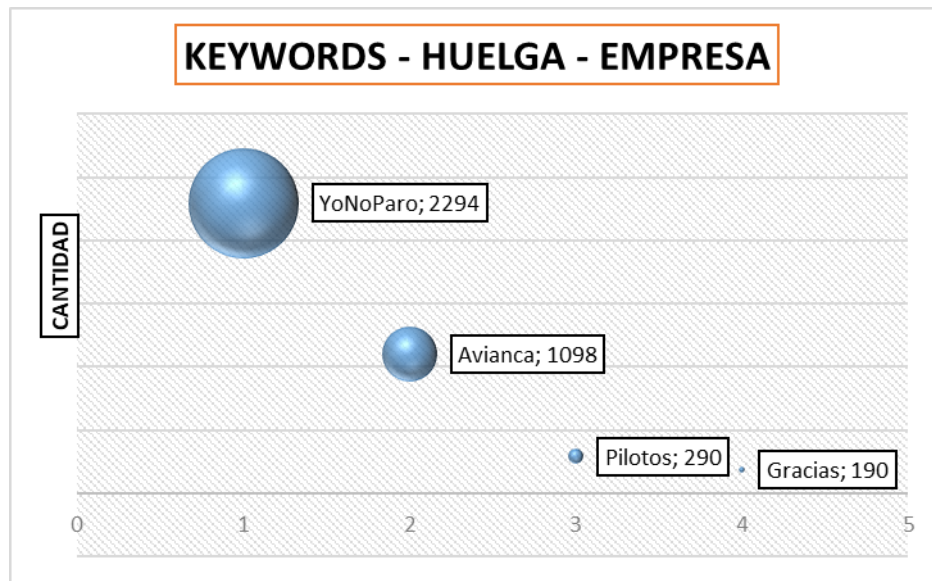


Figura 17. Keywords - Huelga - Empresa - Elaboración Propia

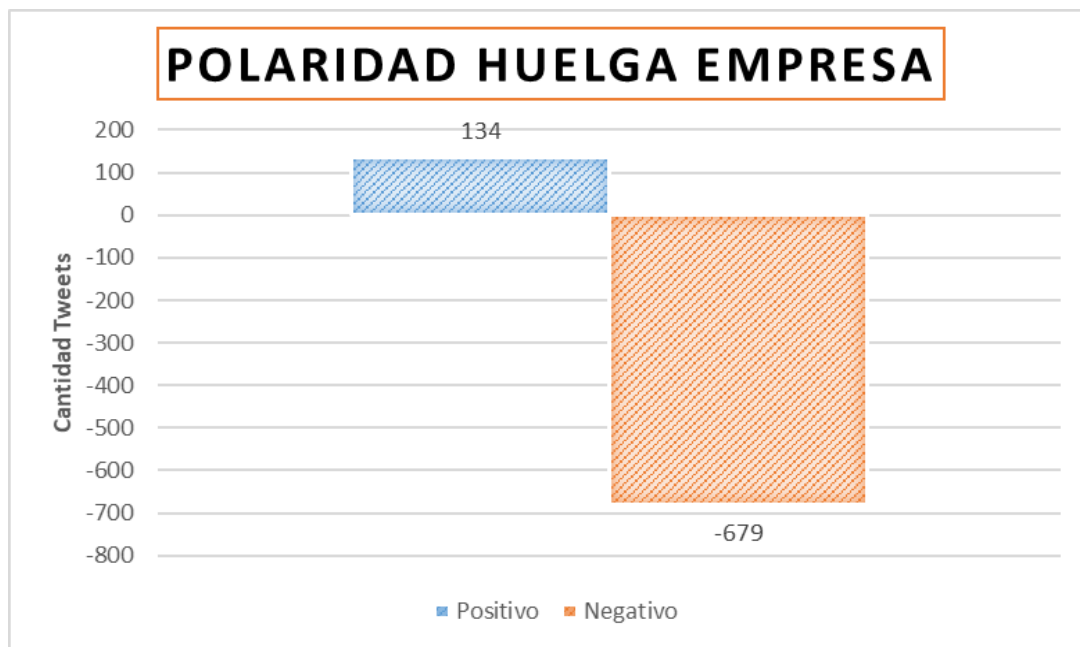






Figura 18. Polaridad tweets - Huelga - Empresa - Elaboración Propia

Polaridad	Fecha	Retweets	Likes	Text	Hashtag	Influencer	Followers	Impression
	22/09/2017	1348	4314	#YoNoParo porque sabemos que nuestros viajeros necesitan llegar a su destino y estamos aquí para ayudarlos. pic.twitter.com/CUTUL6jIV	#YoNoParo	@Avianca	850.000	1.145.800.000
	3/10/2017	429	935	Implementamos de forma inmediata aprobación de @AerocivilCol para incorporar Pilotos extranjeros y fortalecer nuestra operación. #YoNoParo pic.twitter.com/9ddmtKtv5E	#YoNoParo	@Avianca	850.000	364.650.000
	22/10/2017	50	143	Continúan los esfuerzos por hacer frente a la contingencia operacional desde hoy contamos con 20 vuelos diarios adicionales. #YoNoParo pic.twitter.com/PG76F5iRh0	#YoNoParo	@Avianca	850.000	42.500.000
	1/11/2017	3	3	En #WETRADE2017 nuestro gerente @cariasann muestra como buen ejemplo de manejo de reputación el paro de @Avianca Recuerdan el #YoNoParo ?	#WETRADE2017 #YoNoParo	@Avianca	850.000	2.550.000

				pic.twitter.com/hwDPAtvm FK				
								1.555.500.00 0

Tabla 4. Top post, Influencer e Impressions - Empresa - Elaboración Propia

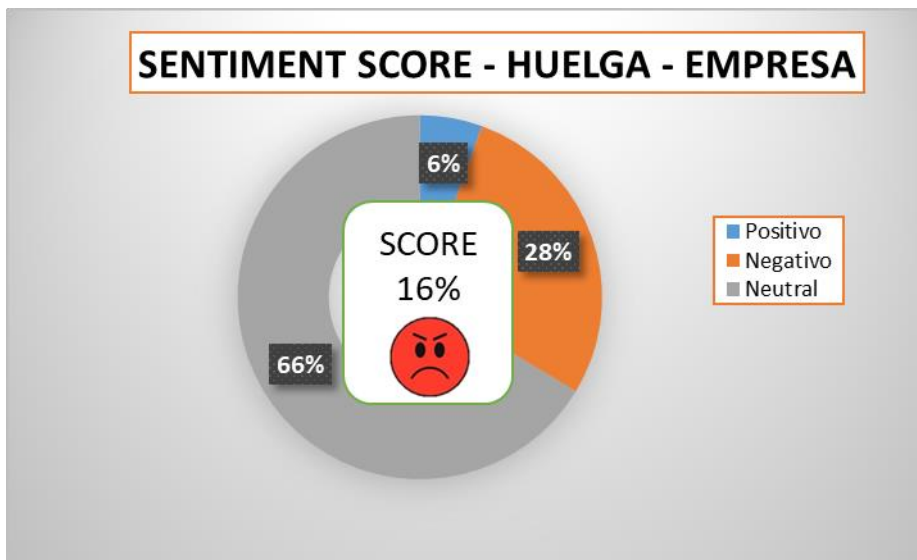


Figura 19. Sentiment Score - Huelga - Empresa - Elaboración Propia

Durante la huelga se observa de manera muy clara un cambio de tendencia respecto a la Reputación (sentimientos) y a la vez una diferencia en la percepción (polaridad) entre las partes enfrentadas: Pilotos y Empresa.

En la línea de tiempo vemos como los Pilotos reciben mucha mayor interacción (retweets) de sus comentarios en twitter, respecto a los comentarios de Avianca, pero la estrategia digital utilizada por Avianca, a través de su hashtag #YoNoParo, tuvo mucha profundidad (más de 1140 millones de impresiones).

Finalmente, esta estrategia digital no logro generar un ambiente positivo hacia Avianca, ya que la medición de la polaridad de los comentarios hacia la empresa, marco una tendencia muy negativa, con un score de apenas el 16%. Caso muy contrario respecto a la polaridad hacia los Pilotos que fue positiva en un 64%.

En la realidad (Offline), Avianca gano esta disputa desde lo legal, pero dejo una profunda marca en cuanto a su reputación corporativa en la red (Online).

5.4.3 Avianca – Después de la Huelga.

Finalmente vamos a analizar el periodo después de la huelga que, dada su implicación, se evaluó para un tiempo de tres meses y medio.

Después de la Huelga: noviembre 16/17 a febrero 28/18

Etiqueta: @Avianca

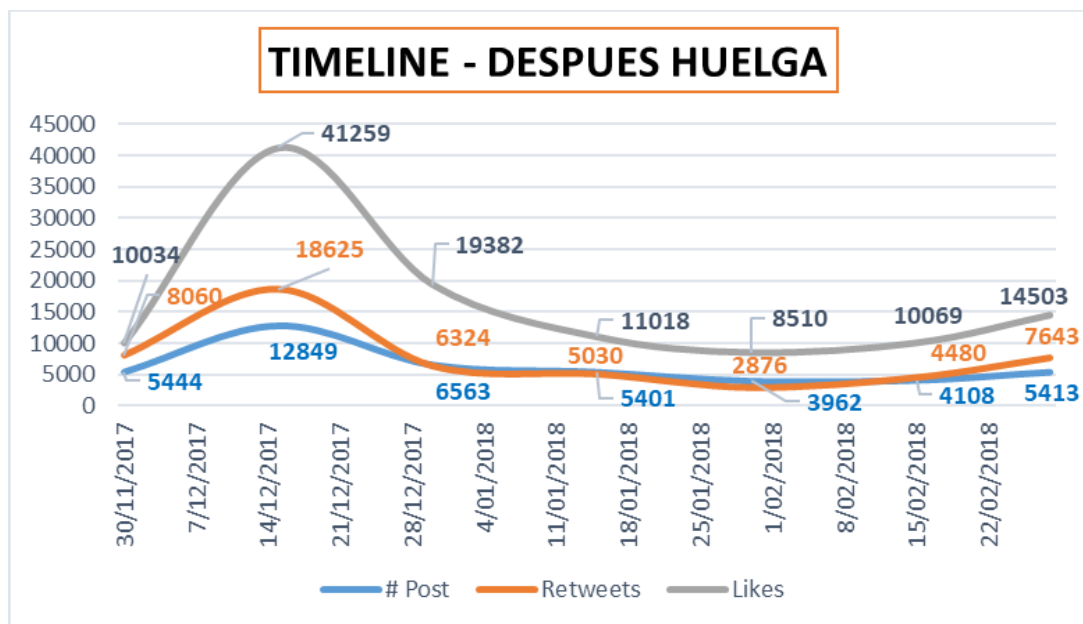


Figura 20. Timeline - Después Huelga - Elaboración Propia

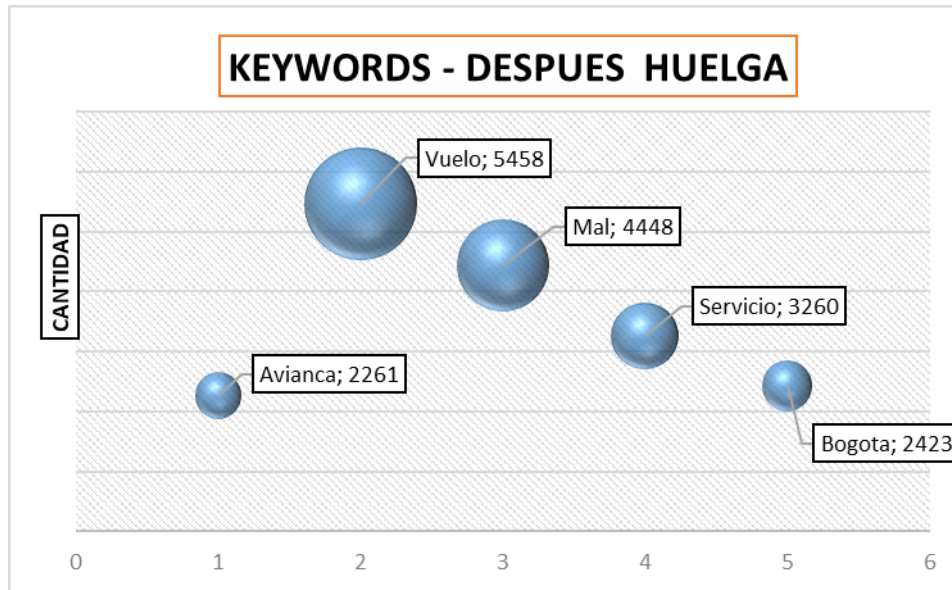






Figura 21. Keywords - Después Huelga - Elaboración Propia

Polaridad	Fecha	Retweets	Likes	Text	Influencer	Followers	Impression
	29/11/2017	1117	2532	700 mil pesos un vuelo Bogotá-Cali en un fin de semana sin festivo. Vengan, amigos de @Avianca ¿Seguro que quieren ser la aerolínea de los colombianos? ¡Es un abuso!	@Rmayorga	596.000	665.732.000
	6/12/2017	1257	2617	La mayoría de los usuarios se pusieron del lado de @Avianca en el paro de pilotos, y la muestra de gratitud de la empresa fue incrementar el valor de los tiquetes al doble y al triple. Avianca dejó de ser colombiana hace mucho, aunque opere aquí	@stevenarce	185.000	232.545.000
	16/12/2017	812	1853	Hola @avianca , estoy muy triste porque siento que por culpa de ustedes no voy a llegar a los premios @UnivisionSports a los que estoy nominada como Mejor Deportista Extremo de Latinoamérica. ¿Sera que hablamos?	@SofiGomez	140.000	113.680.000
	10/01/2018	669	1321	El icopor jamás se biodegrada. Invito a @avianca a que deje de lado el icopor en 2018.	@CadenaGaitan	10.600	7.091.400




				iHay muchas opciones y lo agradecerás! pic.twitter.com/T7jjZEPOMP			
	25/01/2018	260	378	El tiquete aereo de @Avianca Bogotá-Neiva es el más caro del mundo. El vuelo es de 25 minutos y vale \$833.000. Vamos a legislar para ponerle topes y evitar estos abusos.	@ernestomaciast	91.800	23.868.000
	2/02/2018	485	1143	De verdad hay que hacer algo para que @Avianca respete! Lo voy a organizar y espero me apoyen. No más mamadera de gallo !	@CancinoAbog	31.800	15.423.000
	25/02/2018	827	743	A ver si entendí la ecuacion: Para @petrogustavo el petróleo es maldito y lo va a erradicar, tambien va a erradicar a @Avianca y va a repotenciar a @AerolineaSatena , va a tanquear los aviones con gasolina? un derivado del petroleo, o con los paneles solares de @GustavoBolívar ?	@Lepuyana	1.459	1.206.593
							1.059.545.993

Tabla 5. Top posts, Influencer e Impressions - Después Huelga - Elaboración Propia

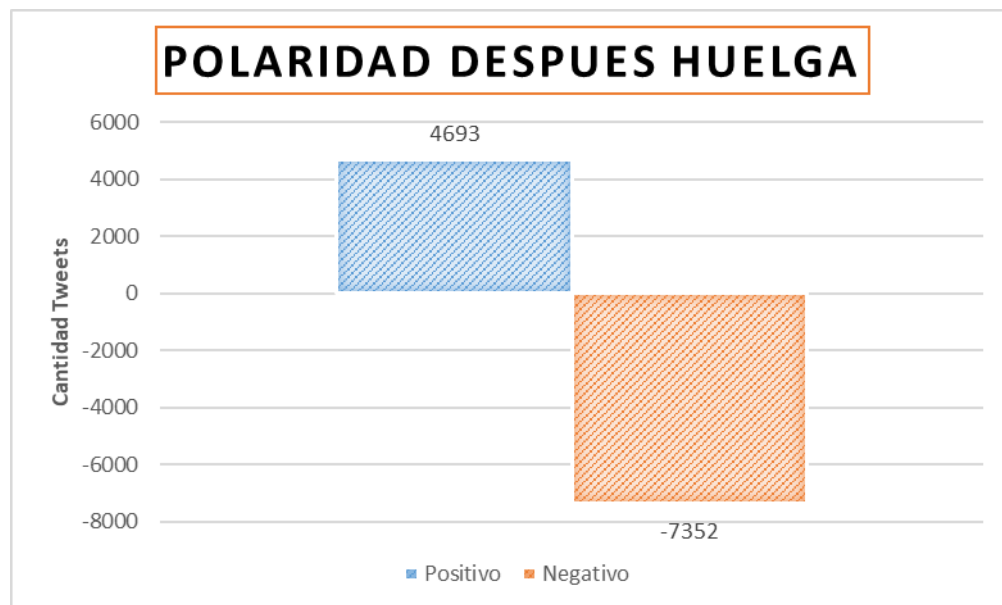


Figura 22. Polaridad tweets - Después Huelga - Elaboración Propia

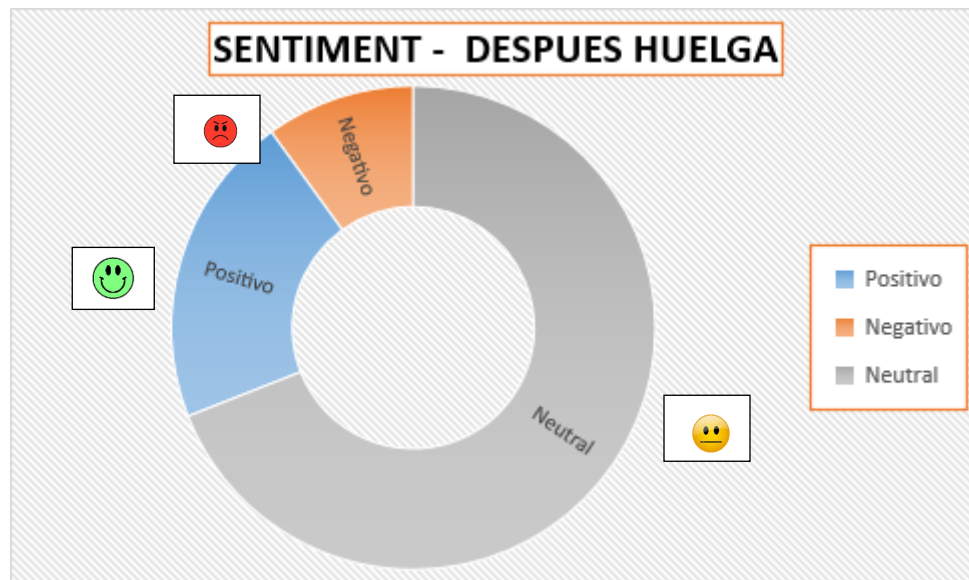


Figura 23. Sentiment - Después Huelga - Elaboración Propia

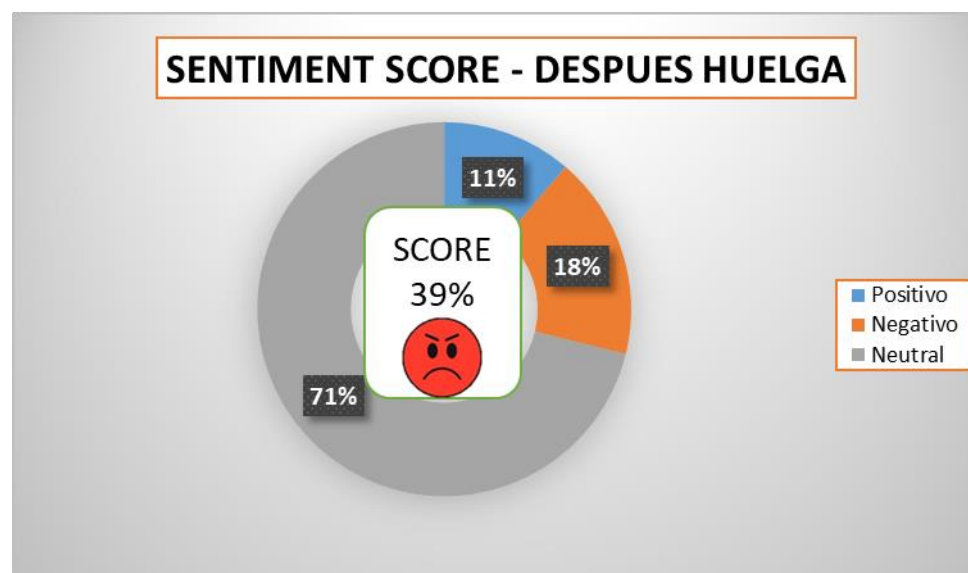


Figura 24. Sentiment Score - Después Huelga - Elaboración Propia

Como se puede observar, hubo bastante actividad después de terminada la huelga, en el mes de diciembre de 2017, seguramente por inicio de la época de vacaciones y donde se empezaron a sentir los primeros efectos de la huelga. Es importante destacar que las palabras que más se utilizaron fueron: vuelo, mal y servicio; esto a raíz de las múltiples cancelaciones de vuelos que se dieron como consecuencia de la huelga y por el despido de bastantes pilotos durante esta época, sin plan de contingencia en el corto plazo.

De otra parte, los post más destacados, tuvieron una relación amplia respecto a las tarifas y al mal servicio prestado por Avianca después de la huelga, con connotaciones claramente negativas, lo cual hace que el sentimiento en general sea ampliamente negativo y que se obtenga un score de sentimientos del 39%. Todo esto nos indica, que después de la huelga, la reputación de Avianca se vio seriamente comprometida y que, durante el periodo analizado, no pudo salir del terreno negativo.

6. Conclusiones y futuros trabajos

Dentro de la literatura revisada a la fecha, se encontraron muy pocos estudios respecto a la reputación corporativa y su desarrollo en la web. No se encontró ninguno que trabajara a partir de un enfoque dimensional la reputación corporativa, con aplicación en social media.

Por lo cual el problema, las hipótesis y los objetivos planteados en esta investigación son bastante novedosos; no solo por la forma de cómo se pretendió abordar el estudio con un enfoque dimensional sobre la reputación, sino que además se hizo a través de minería de datos, donde se analizaron más de 70.000 tweets. La minería de datos es algo relativamente nuevo y de muy alto potencial, ya que corresponde a todo lo relacionado con Inteligencia Artificial, *Machine learning*, *Big Data*, etc., que el profesor Thomas Davenport (Davenport, 2018), investigador del MIT, llama *Tecnologías Cognitivas*; y que ya están presentes en muchas de nuestras actividades cotidianas, que están cambiando nuestras vidas y potenciando algo que está a nuestro alcance: Datos.

Desde la revisión de la literatura referenciada sobre la Reputación Corporativa, se ha encontrado una serie de abordajes respecto a cómo definir este *constructo*. Todos los enfoques tienen sus ventajas, pero a la vez sus falencias. Ahora, con el desarrollo de la web, el empoderamiento y la democratización de las opiniones en social media; la reputación y sus dimensiones deberán plantearse el reto de hacer corresponder lo Offline con lo Online; ya que lo que actualmente muestra estas primeras aproximaciones, es que los formatos y ranquin, definidos a través de encuestas a expertos; se aleja bastante de la opinión y el sentimiento general que puede manejarse en una red social como Twitter, para una empresa específica como Avianca.

Si observamos evaluaciones tan elaboradas con la de Merco, para la evaluación de la Reputación Corporativa en Colombia, encontramos que Avianca goza de una Reputación extraordinaria dentro del ámbito empresarial y en el juicio de expertos; y que hechos tan importantes como la huelga sucedida el año pasado, parecería no impactar lo suficiente, tanto a los expertos, como a los inversionistas, ya que la acción de Avianca no sufrió mayor modificación durante este periodo de huelga y la empresa se clasificó en el top diez de las empresas de mayor reputación en Colombia para el 2017.

Ahora, lo que percibió el principal stakeholders de Avianca, que son sus usuarios, sí se vio reflejado en el cambio del posicionamiento de la reputación entre los mismos y que se manifestó en las opiniones de la red social Twitter, donde se extrajeron los datos, afirmando la Hipótesis 1.

La profesora Sabrina Helm, que es una experta en investigación sobre la reputación empresarial y su relación con los *stakeholders*, muestra que todos stakeholders tienen una relación directa con la dimensión de la reputación corporativa correspondiente a *Emotional appeal*. Figura 25.

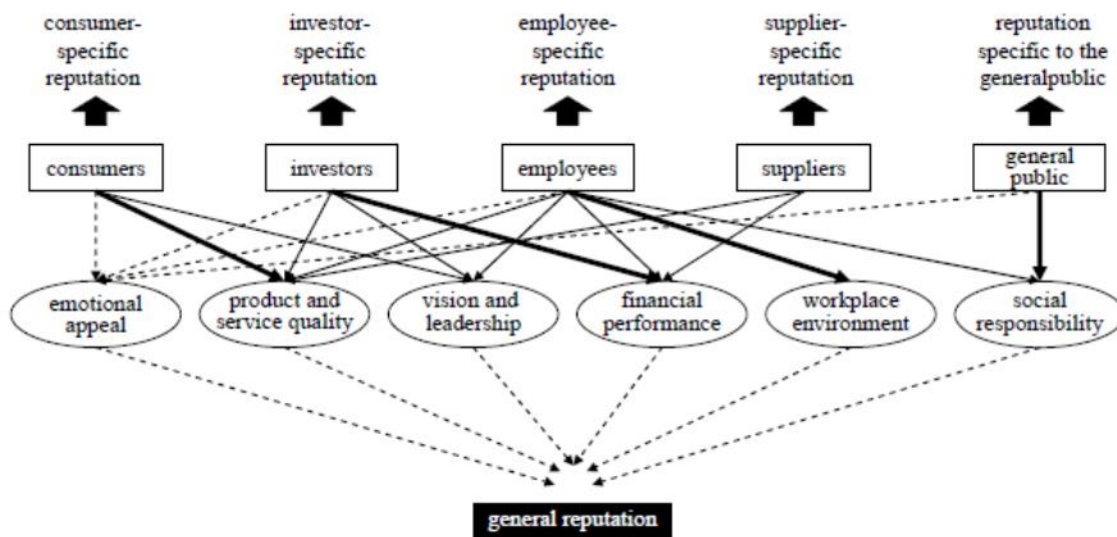


Figura 25. Relación de las Dimensiones y los Stakeholders (Helm, 2007)

Este *Emotional appeal*, donde se relacionan todos los *stakeholders*, es el que se postea, se discute, se comparte y sobre el cual se generan polaridades en las redes sociales. Es aquí, donde el análisis de sentimientos se puede convertir en un *predictor* de la reputación corporativa y este análisis está al alcance de las empresas en tiempo real; aprovechando la *sabiduría* que nos dan los datos. Aquí se demuestra, como una técnica como el análisis de sentimientos puede generar *valor* a las empresas, que es el fin último del aprovechamiento de los datos mediante las tecnologías cognitivas.

Para el caso analizado de Avianca, en el periodo total de tiempo de esta investigación, podemos concluir que, si bien su Reputación a nivel de los expertos (Reputación Merco – Dimensiones: *Vision & Leadership and Social Responsibility*) y agentes financieros (Reputación Bursátil – Dimensión: *Financial Performance*), ha salido bien librada; para el caso del público en general y principalmente entre sus usuarios, sí se percibió un cambio en la percepción de Reputación hacia el lado negativo, enfatizado

principalmente en la dimensión: *Emotional Appeal*, confirmando así Hipótesis 2, como lo muestra la figura 26:

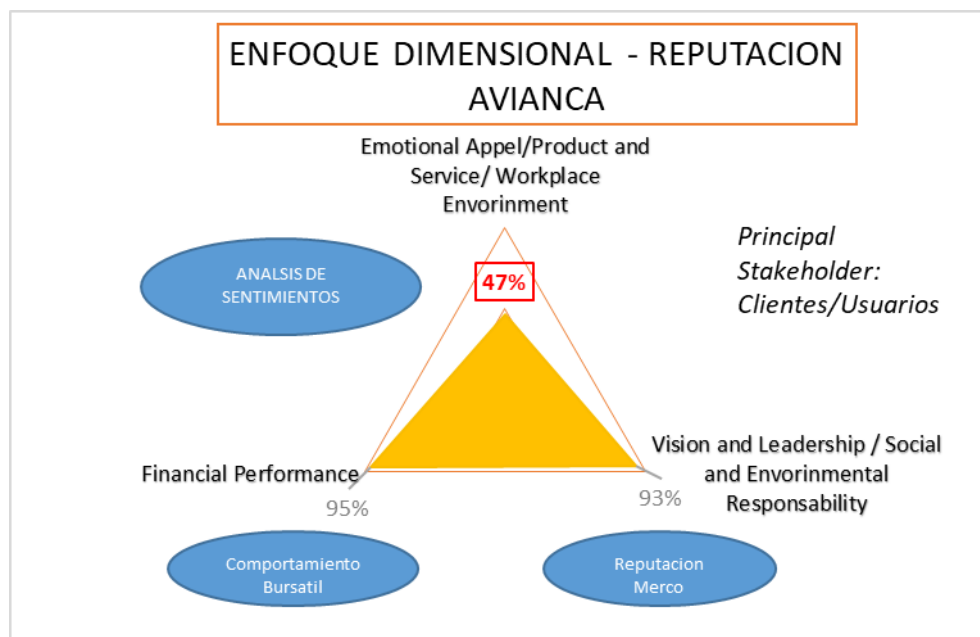


Figura 26. Enfoque Dimensional - Reputación Avianca - Análisis de Sentimientos - Elaboración Propia

A lo largo de la investigación se observó como la reputación corporativa de Avianca se vio afectada desde el momento del anuncio de la huelga de pilotos en septiembre de 2017 y hasta el final de periodo analizado, ya finalizada la huelga. En social media y específicamente en Twitter, la reputación de Avianca, paso de un terreno positivo a una posición negativa, que no les fue posible cambiar, a pesar de sus estrategias el terreno digital, negando de esta manera la Hipótesis 3.

Temas como el servicio, las tarifas y las cancelaciones de vuelos, fueron muy repetitivos en los post colocados, contribuyendo así a apalancar ese sentimiento negativo que se mantuvo durante y después de la huelga hacia la compañía. Si bien la estrategia

digital de la empresa en postear los esfuerzos que estaban realizando para mantener el nivel de servicio durante la huelga tuvo una alta penetración y difusión; no fueron suficientes para mitigar las carencias en prestación del servicio en el mundo real (offline), que tuvieron y han tenido, durante la huelga y después de la misma. Todo esto se vio reflejado en los *post*, *keywords*, *polaridades* de las opiniones y en el *score* final de reputación obtenido, que se extrajo y configuro a través de la minería de datos, en cada uno de los momentos analizados.

Espero con esta investigación, poder abrir un nuevo camino para posibles aplicaciones de la *minería de datos* y la *social analítica web* para el marketing, que puedan crear *valor* al sector empresarial y de servicios. En la literatura se reseña una serie de limitaciones que actualmente tiene el análisis de sentimientos, por eso se considera necesario el desarrollo de otros trabajos que, utilizando aplicaciones más avanzadas, puedan dar mayor alcance y precisión a las conclusiones que sobre trabajos similares, se presenten a futuro. También se considera necesario y además se sugiere, una mayor profundización respecto a los grupos de interés que opinan en las redes sociales, para generar procesos de agrupación (clúster) y segmentación de poblaciones en la web; que puedan generar acciones de marketing más específicas sobre los intereses de estos segmentos; que es un límite todavía no muy explorado y explotado en la investigación del marketing.

7. Bibliografía

- Aldahawi, H. A. (2015). Mining and Analysing Social Network in the Oil Business : Twitter Sentiment Analysis and Prediction Approaches, 158.
- Arvidsson, A. (2011). General Sentiment: How value and affect converge in the information economy. *Sociological Review*, 59(SUPPL. 2), 39–59. <https://doi.org/10.1111/j.1467-954X.2012.02052.x>
- Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy and Leadership*, 38(6), 43–49. <https://doi.org/10.1108/10878571011088069>
- Avianca. (2018). Avianca.com. Retrieved from <https://www.avianca.com/co/es/nuestra-compania/informacion-institucional/perfil-corporativo>
- Barbato, M.-M. (2016). Measuring Corporate Reputation through online Social Media, (December). Retrieved from http://amslaurea.unibo.it/12434/1/Tesi_Barbato.pdf
- Becerra, C. M. (2016). Analisis de Sentimientos en Twitter : El bueno , el malo y el > :(.
- Bernal, C. A. (2006). *Metodología de la investigación*.
- Bhuiyan, T., Josang, A., Xu, Y. (2010). *Trust and reputation management in web-based social network*.
- Bhuiyan, T., Josang, A., & Xu, Y. (2010). Trust and reputation management in web-based social network. *Web Intelligence and Intelligent ...*, 207–232. <https://doi.org/10.5772/8375>
- Boumaiza, A. (2015). A Survey on Sentiment Analysis and Visualization. *Journal of Emerging Technologies in Web Intelligence*, 7(1), 35–43. <https://doi.org/10.12720/jetwi.7.1.35-43>

- Brammer, S., & Pavelin, S. (2006). Corporate reputation and social performance: The importance of fit. *The Journal of Management Studies*, 43(3), 435–456.
- Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, 28(2), 15–21.
<https://doi.org/10.1109/MIS.2013.30>
- Caruana, A., & Caruana, A. (2006). Corporate reputation : concept and measurement, 6(2), 109–118.
- City, D. (1993). A Framework linking intangible resources and capabilities to Sustainable competitive. *Management*, 14(July 1992).
- Colleoni, E., Arvidsson, A., Hansen, L. K., & Marchesini, A. (2011). Measuring corporate reputation using sentiment analysis. *The 15th International Conference on Corporate Reputation: Navigating the Reputation Economy, New Orleans, USA, May 18. - 20. 2011.*
- Davenport, T. H. (2018). What 's Your Cognitive Strategy ?, 59(4).
- Dinero. (2017). Las empresas colombianas con mejor reputacion corporativa en 2017. Retrieved from <https://www.dinero.com/edicion-impresa/negocios/articulo/monitor-empresarial-de-reputacion-corporativa-merco-2017/251533>
- Einwiller, S. A., Carroll, C. E., & Korn, K. (2010). Under what conditions do the news media influence corporate reputation the roles of media dependency and need for orientation. *Corporate Reputation Review*, 12(4), 299–315. <https://doi.org/10.1057/crr.2009.28>
- Espectador, E. (2017). Qué es machine learning y porqué será tan grande como la llegada de internet. *Elespectador.Com*. Retrieved from <https://www.elespectador.com/tecnologia/que-es-machine-learning-y-porque-sera-tan-grande-como-la-llegada-de-internet-articulo-728101>
- Etter, M., Colleoni, E., Illia, L., Meggiorin, K., & D'Eugenio, A. (2017). Measuring organizational

- legitimacy in social media: Assessing citizens' judgments with sentiment analysis. *Business & Society*, 1–38. <https://doi.org/10.1177/0007650316683926>
- Fielding, N., Lee, R., & Blank, G. (2008). *The SAGE Handbook of Online Research Methods*, (January 2008). <https://doi.org/10.4135/9780857020055>
- Fombrun, C. J., Gardberg, N. A., & Sever, J. M. (2000). The Reputation QuotientSM: A multi-stakeholder measure of corporate reputation. *Journal of Brand Management*, 7(4), 241–255. <https://doi.org/10.1057/bm.2000.10>
- Fortune. (2018). fortune.com. Retrieved from <http://fortune.com/worlds-most-admired-companies/>
- Han, J., Pei, J., Kamber, M. (2006). *Data Mining: Concepts and Techniques*.
- He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, 33, 464–472. <https://doi.org/10.1016/j.ijinfomgt.2013.01.001>
- Hearn, A. (2010). 10-3Hearn, 421–438.
- Helm, S. (2005). Designing a Formative Measure for Corporate Reputation. *Corporate Reputation Review*, 8(2), 95–109. <https://doi.org/10.1057/palgrave.crr.1540242>
- Helm, S. (2007). One reputation or many? *Corporate Communications: An International Journal*, 12(3), 238–254. <https://doi.org/10.1108/13563280710776842>
- Holding, C. (2018). Avianca. Retrieved from <https://www.avianca.com/co/es/nuestra-compania/informacion-institucional/perfil-corporativo>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>

- Kumlin, M. (2016). Social media as a tool for corporate reputation management.
- Lewin, J. E. (2017). El paro de los pilotos es la batalla final por Avianca. *La Silla Vacía*, (Septiembre 29). Retrieved from <https://lasillavacia.com/el-paro-de-los-pilotos-es-la-batalla-final-por-avianca-62741>
- Merco. (2017). Merco.com. Retrieved from <http://www.merco.info/co/>
- Mohanty, S., Jagadeesh, M., Srivatsa, H. (2013). *Extracting Value from Big Data: In-Memory Solutions, Real Time Analytics, and Recommendation Systems*.
https://doi.org/https://doi.org/10.1007/978-1-4302-4873-6_8
- Oakleaf, M. (2009). Writing information literacy assessment plans: A guide to best practice. *Communications in Information Literacy*, 3(2), 80–90. <https://doi.org/10.1002/asi>
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - EMNLP '02*, 10(July), 79–86.
<https://doi.org/10.3115/1118693.1118704>
- Porter, M. (1989). *How Competitive Forces Shape Strategy*.
https://doi.org/https://doi.org/10.1007/978-1-349-20317-8_10
- Rindova, V. P., Williamson, I. O., Petkova, A. P., & Sever, J. M. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of Management Journal*.
<https://doi.org/10.5465/AMJ.2005.19573108>
- Rybalko, S., & Seltzer, T. (2010). Dialogic communication in 140 characters or less: How Fortune 500 companies engage stakeholders using Twitter. *Public Relations Review*, 36(4), 336–341.
<https://doi.org/10.1016/j.pubrev.2010.08.004>

Twitter. (2018). Twitter.com. Retrieved from <https://twitter.com/?lang=es>

Verhoef, Peter; Kooge, E. W. N. (2016). *Creating Value with Big Data Analytics*.

Wigley, S., & Lewis, B. K. (2012). Rules of engagement: Practice what you tweet. *Public Relations Review*, 38(1), 165–167. <https://doi.org/10.1016/j.pubrev.2011.08.020>

Zhu, F. (2010). Impact of Online Consumer Reviews on Sales : The Moderating Role of Product and Consumer. *Journal of Marketing*, 74(March), 133–148.
<https://doi.org/10.1509/jmkg.74.2.133>