

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias Sociales Y Humanísticas

Evaluación de la gestión de insumos médicos en Ecuador a través de Twitter empleando técnicas de análisis de big data durante la pandemia de la Covid-19

PROYECTO INTEGRADOR

Previo la obtención del Título de:

Licenciado en Auditoría y Control de Gestión

Presentado por:

Ester Rebeca Melo Vargas

Luis Eduardo Quijije Cevallos

GUAYAQUIL – ECUADOR

Año: 2021

DEDICATORIA

Dedico este trabajo de grado a la mujer más valiente y resiliente que he conocido en toda mi vida, quien también es mi mayor inspiración y a quien tengo el honor de llamar “mamá”. Como se lo prometí, este trabajo va dedicado a ella: Lucía del Rocío Vargas Chica.

Ester Melo.

AGRADECIMIENTO

Agradezco a Dios por guiarme y fortalecerme en el camino, esto no sería posible sin su sabiduría y gracia inmerecida. Agradezco a mis padres, Julio Melo y Lucía Vargas, por apoyar mis decisiones, cuidarme en todo momento y desvelarse conmigo cuando fue necesario. Agradezco a mis hermanos, familia y amigos por motivarme a ser mejor cada día y comprender mi ausencia en muchas ocasiones. Agradezco al Pastor Manuel Chilán por sus oraciones y bendiciones hacia mi vida. A mi compañero de tesis por ayudarme a materializar este proyecto. Agradezco a los doctores Mayken Espinoza, Ronald Campoverde y Víctor González, y la doctora Carmen Vaca, por brindarme oportunidades de crecimiento personal y profesional. Así mismo, agradezco a mi tutor y a todos los docentes de la ESPOL que han contribuido en mi formación. Finalmente, agradezco a todas las personas que creyeron en mí.

Ester Melo.

DEDICATORIA

Dedico esta tesis a Dios porque sé que el haber llegado hasta aquí ha sido parte de su plan perfecto en mi vida. A mis padres, Antonio y Margarita, por haberme forjado como persona y por haberme impulsado en el cumplimiento de mis objetivos. A mis hermanos de quienes he aprendido la constancia y perseverancia siendo ellos mi guía desde que era pequeño. A Melany quien ha estado a mi lado en cada momento de mi vida universitaria.

Luis Quijje.

AGRADECIMIENTO

Agradezco a Dios por las innumerables oportunidades académicas otorgadas por su gracia, y por la sabiduría para enfrentar cualquier problema presentado en el camino. Agradezco a mi familia por toda su confianza, amor y apoyo incondicional durante estos años, así como por haber sembrado en mí los valores necesarios para alcanzar mis metas, tales como perseverancia, resiliencia, y constancia; los cuales me han permitido llegar hasta aquí. Agradezco a mi compañera de vida quien ha estado conmigo desde el principio siendo un apoyo importante. A mi compañera de tesis por ayudarme a materializar este proyecto. Agradezco a mis docentes de la ESPOC por inspirarme a ser mejor cada día. Finalmente, agradezco a todas las personas que me apoyaron y creyeron en mí.

Luis Quijje.

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, nos corresponde conforme al reglamento de propiedad intelectual de la institución; *Ester Rebeca Melo Vargas* y *Luis Eduardo Quijije Cevallos* damos nuestro consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”.

Ester Rebeca Melo
Vargas



Luis Eduardo Quijije
Cevallos

EVALUADOR

.....

MSc. Christian Vera Alcívar

TUTOR Y PROFESOR DE LA MATERIA

RESUMEN

La gestión de insumos médicos en Ecuador, desde tiempos remotos, ha representado un obstáculo en el tratamiento oportuno de enfermedades debido a la insuficiencia de estos. Dicha situación se agudizó a principios del 2020 por la pandemia del coronavirus. En consecuencia, las personas optaron por emplear redes sociales, como Twitter, para obtener medicinas y equipo médico. Por ello, el presente estudio tiene como objetivo evaluar la gestión de insumos médicos en Ecuador por parte de la ciudadanía, a través de la red social Twitter empleando técnicas de análisis de big data, para determinar su grado de efectividad durante la pandemia de la Covid-19 desde marzo del 2020 hasta marzo del 2021. Para lograrlo se realizaron entrevistas a médicos para conocer los insumos más demandados. A partir de ello se aplicó web scraping en Twitter para extraer publicaciones sobre la gestión de insumos. Posteriormente se aplicaron técnicas de minería de texto para identificar emociones, clasificar necesidades y validar la información. Por último, se establecieron indicadores de efectividad. Los resultados revelaron que marzo y julio del 2020 y enero del 2021 fueron tiempos críticos, especialmente en provincias como Guayas, Azuay, Imbabura y Santa Elena; siendo consistes con la información reportada por el Comité de Operaciones de Emergencia (COE) de Ecuador. Además, se demostró que una gestión efectiva implicaría un ahorro de \$6.8 millones de dólares en el sector salud y la cobertura del 70% de la demanda de ataúdes. Se concluye que este estudio brinda información de interés para la formulación de políticas públicas relacionadas al abastecimiento de insumos y la atención médica.

Palabras claves: Insumos médicos, Coronavirus, Twitter, Efectividad.

ABSTRACT

The management of medical supplies in Ecuador, since ancient times, has represented an obstacle in the timely treatment of diseases due to their insufficiency. This situation was exacerbated in early 2020 by the coronavirus pandemic. Consequently, people chose to use social networks, such as Twitter, to obtain medicines and medical equipment. Therefore, the present study aims to evaluate the management of medical supplies in Ecuador by citizens, through the social network Twitter using big data analysis techniques, to determine their degree of effectiveness during the Covid-19 pandemic from March 2020 to March 2021. To achieve this, interviews were conducted with doctors to find out the most demanded supplies. From this, web scraping was applied on Twitter to extract publications on the management of supplies. Subsequently, text mining techniques were applied to identify emotions, classify needs, and validate information. Finally, effectiveness indicators were established. The results revealed that March and July 2020 and January 2021 were critical times, especially in provinces such as Guayas, Azuay, Imbabura, and Santa Elena; being consistent with the information reported by the Comité de Operaciones de Emergencia (COE) from Ecuador. In addition, it was shown that effective management would imply a saving of \$ 6.8 million dollars in the health sector and the coverage of 70% of the demand for caskets. It is concluded that this study provides information of interest for the formulation of public policies related to the supply of inputs and medical care.

Keywords: Medical supplies, Coronavirus, Twitter, Effectiveness.

ÍNDICE GENERAL

RESUMEN	I
<i>ABSTRACT</i>	II
ÍNDICE DE FIGURAS	VI
ÍNDICE DE TABLAS	VII
ABREVIATURAS	VIII
CAPÍTULO 1	1
1. Introducción	1
1.1 Descripción del problema	1
1.2 Justificación del problema	1
1.3 Alcance	2
1.4 Objetivos	2
1.4.1 Objetivo General.....	2
1.4.2 Objetivos Específicos.....	2
1.5 Marco teórico	3
1.5.1 Marco conceptual	3
1.5.2 Marco metodológico.....	8
1.5.3 Marco referencial	12
1.5.4 Marco legal	13
CAPÍTULO 2	15
2. METODOLOGÍA	15
2.1 Diseño de la investigación	15
2.2 Entrevista	15
2.2.1 Diseño de la entrevista	15
2.2.2 Validación de entrevistados	16

2.3	Web scraping	17
2.4	Extracción de información	18
2.4.1	Búsqueda de patrones.....	18
2.5	Recuperación de información	19
2.6	Procesamiento de lenguaje natural	20
2.6.1	Análisis de sentimiento	20
2.7	Clustering.....	21
2.7.1	K-medias	21
2.8	Text summarization	21
2.9	Indicadores de Efectividad	22
2.9.1	Tiempos críticos.....	22
2.9.2	Zonas críticas	23
2.9.3	Proyección de pérdidas económicas en el sector de la salud	23
2.9.4	Proyección de la demanda de ataúdes en Guayas	24
CAPÍTULO 3.....		25
3.	Resultados Y ANÁLISIS.....	25
3.1	Análisis de entrevistas.....	25
3.1.1	Entrevista a Dr. Miguel Icaza	25
3.1.2	Entrevista a Dra. Lorena Espinoza.....	25
3.2	Web Scraping.....	26
3.3	Resultados de extracción de información	26
3.3.1	Palabras frecuentes	26
3.3.2	Identificación de patrones	27
3.4	Recuperación de Información.....	31
3.5	Resultados del Análisis de Sentimiento.....	31
3.6	Identificación de la efectividad de las campañas de insumos médicos	34
3.6.1	Método del codo	35

3.6.2	K means	36
3.7	Validación	37
3.8	Indicadores de efectividad	37
3.8.1	Tiempos críticos.....	37
3.8.2	Zonas críticas	39
3.8.3	Proyección de Pérdidas en el Sector Salud	40
3.8.4	Proyección de la demanda de Ataúdes en Guayas.....	41
CAPÍTULO 4		43
4.	Conclusiones Y Recomendaciones	43
4.1	Conclusiones.....	43
4.2	Recomendaciones.....	44
BIBLIOGRAFÍA		46
APÉNDICES		53

ÍNDICE DE FIGURAS

Figura 1.1 Áreas fundamentales de la minería de texto.....	9
Figura 2.1 Diseño de la Investigación.....	15
Figura 2.2 Relación de palabras según su tipo.....	18
Figura 3.1 Top 10 de palabras frecuentes	27
Figura 3.2 Patrones asociados a la palabra "Covid"	28
Figura 3.3 Patrones asociados a la palabra "Coronavirus"	28
Figura 3.4 Muestra del primer patrón asociado a la palabra "covid"	29
Figura 3.5 Muestra del segundo patrón asociado a la palabra "covid".....	29
Figura 3.6 Muestra del primer patrón asociado a la palabra "coronavirus"	30
Figura 3.7 Muestra del segundo patrón asociado a la palabra "coronavirus"	30
Figura 3.8 Resultado del análisis de sentimiento.....	31
Figura 3.9 Porcentaje de tweets negativos por provincias	33
Figura 3.10 Porcentaje de tweets positivos por provincias	34
Figura 3.11 Dispersión de los tweets con base al análisis de sentimiento, corazones y retweets	35
Figura 3.12 Número de clusters mediante el método del codo	35
Figura 3.13 Identificación de campañas de insumo efectivas y no efectivas	36
Figura 3.14 Serie en el tiempo de tweets catalogados negativos	38
Figura 3.15 Curva epidemiológica de casos positivos de covid-19 acumulados por semana epidemiológica.....	39
Figura 3.16 Identificación de la efectividad de las campañas de insumos en Ecuador por provincia.....	39
Figura 3.17 Casos confirmados de coronavirus en Ecuador por provincia	40

ÍNDICE DE TABLAS

Tabla 2.1 Descripción de patrones	19
Tabla 2.2 Estimación de pérdidas económicas en el sector de la salud durante el período de marzo a mayo del 2020	23
Tabla 3.1 Resultado de Análisis de Sentimiento.....	32
Tabla 3.2 Composición de los grupos identificados.....	37

ABREVIATURAS

ESPOL	Escuela Superior Politécnica del Litoral
COE	Comité de Operaciones de Emergencia
OMS	Organización Mundial de la Salud
OIT	Organización Internacional del Trabajo

CAPÍTULO 1

1. INTRODUCCIÓN

1.1 Descripción del problema

La gestión de insumos médicos en Ecuador, desde tiempos remotos, ha representado un obstáculo en el tratamiento oportuno de enfermedades o situaciones de emergencia a nivel de salud. A principios de 2020, dicha falencia se agudizó debido a la pandemia provocada por la Covid-19, la cual ha causado millones de muertes en todo el mundo. Además, durante la pandemia se reportaron casos de corrupción respecto a la distribución de insumos médicos. Por todo aquello, las personas vieron la necesidad de emplear canales de comunicación modernos, como las redes sociales, para obtener medicinas y equipo médico que incluso a nivel hospitalario eran insuficientes.

Sin embargo, a la fecha se desconocen sobre estudios enfocados en la evaluación de la gestión de este tipo de recursos en el territorio ecuatoriano y, sobre todo, de indicadores que permitan medir su impacto en diferentes etapas de la pandemia. Adicionalmente, no se cuenta con datos respecto a la identificación de patrones de comportamiento y zonas críticas o de alta demanda de insumos médicos a través de las redes sociales.

La escasez de esta información imposibilita obtener una visión completa de la realidad ecuatoriana en la lucha contra el aumento de casos positivos y muertes por coronavirus. Así mismo, invisibiliza las prácticas de gestión de insumos médicos ejecutadas por la ciudadanía que podrían guiar la toma de decisiones por parte de las autoridades locales.

1.2 Justificación del problema

En el presente trabajo de investigación se desarrollará un análisis de Big Data a partir de publicaciones extraídas de Twitter relacionadas a la gestión de insumos médicos en el territorio ecuatoriano durante la pandemia de la covid-19. Dicho análisis se basará

en fundamentos teóricos y metodológicos para la construcción de indicadores que permitan monitorear la situación sanitaria y las respuestas de la población.

Esto debido a que, dicha gestión debe satisfacer las necesidades de la sociedad en la mayor medida posible y contribuir a la mitigación de casos positivos de coronavirus. En consecuencia, se presentará un informe que resuma los principales hallazgos respecto al impacto que posee la red social Twitter como medio de difusión de la demanda de dichos insumos y la identificación de zonas críticas.

1.3 Alcance

El desarrollo de este proyecto considera el estudio de las publicaciones de Twitter a nivel nacional durante la pandemia. Por ello, para una comprensión profunda, se recopilarán publicaciones desde marzo de 2020 hasta marzo del presente año. Cabe recalcar que se considerará el uso de esta red social debido a que se conoce que es uno de los medios más empleados para difundir información (Jiang, Huang, & Li, 2021).

1.4 Objetivos

1.4.1 Objetivo General

Evaluar la gestión de insumos médicos en Ecuador por parte de la ciudadanía, a través de la red social Twitter empleando técnicas de análisis de Big data, para determinar su grado de efectividad durante la pandemia de la Covid-19 desde marzo del 2020 hasta marzo del 2021.

1.4.2 Objetivos Específicos

- Recopilar información sobre el tratamiento de la covid-19, mediante entrevistas a expertos, para identificar los insumos médicos más requeridos.
- Crear una base de datos a partir de publicaciones efectuadas en Ecuador y extraídas de Twitter, empleando web scrapping, para analizar la demanda de los insumos identificados.

- Aplicar técnicas de preprocesamiento y minería de datos masivos y de texto, para identificar patrones de comportamiento y zonas críticas o de riesgo.
- Evaluar la gestión de insumos médicos a través de indicadores, para determinar el grado de efectividad de dicha gestión durante la pandemia de la Covid-19.

1.5 Marco teórico

1.5.1 Marco conceptual

Para una mayor comprensión del estudio, se describen los siguientes términos:

1.5.1.1 Coronavirus o Covid-19

La enfermedad por coronavirus 2019, también denominada Covid-19 está causada por el virus SARS-CoV-2, el cual es un nuevo tipo de coronavirus responsable de causar el síndrome respiratorio agudo severo, nombrado así por su similitud con el SARS-CoV que causó similar síndrome respiratorio en China (Cavallo, Rossi, & Chiarelli, 2020).

De igual manera, Cennimo (2021) define al Covid-19 como:

La enfermedad causada por un nuevo coronavirus llamado coronavirus del síndrome respiratorio agudo severo 2 (SAR- CoV-2,), el cual fue identificado por primera vez en medio de un brote de casos de enfermedad respiratoria en la ciudad de Wuhan.

Por lo cual, debido al consenso encontrado en las definiciones, se categorizará al covid-19 como la enfermedad causada por el virus SARS-CoV-2 en el desarrollo de esta investigación.

1.5.1.2 Pandemia de Covid-19

La enfermedad por coronavirus fue caracterizada como pandemia el 11 de marzo del 2020 por la Organización Mundial de la Salud (OMS), debido a la extensión de la epidemia por todos los continentes. Para esa fecha, Ecuador ya registraba 17 casos positivos de coronavirus y, de hecho, ocupaba el tercer lugar en Latinoamérica por ello, después de Brasil y Argentina (Pérez, 2020).

En consecuencia, el Ministerio de Salud Pública del Ecuador emitió el Acuerdo Ministerial No. 00126-2020 y estableció el Estado de Emergencia Sanitaria. A partir de este acuerdo se estipularon medidas para contener la propagación del coronavirus, entre las cuales se destacan las siguientes:

- Adopción de medidas de prevención en el transporte público por parte de los Gobiernos Autónomos Descentralizados (GAD).
- Cuidado especial a los adultos mayores, personas con discapacidad y con enfermedades crónicas y catastróficas.
- Fortalecimiento de las medidas de bioseguridad al personal de salud.
- Prevención y control en los puntos de ingreso al país.
- Restricciones para eventos masivos y de concentración de gente, considerando la zona y el número de casos.
- Uso de plataformas tecnológicas en telemedicina, educación en línea y teletrabajo para aquellos casos que lo ameriten.
- Otras (Andramuño, 2020).

Sin embargo, dichas medidas estuvieron acompañadas de una inadecuada gestión, resultando inefectivas y conllevando a la agudización de la situación sanitaria del país (Acosta, 2020). De tal manera que, el 13 de marzo el presidente de turno decretó el Estado de Excepción, dando lugar al Comité de Operaciones de Emergencia (COE Nacional) para la coordinación de la crisis sanitaria (ONU, 2020).

1.5.1.3 Insumos médicos para la prevención y el tratamiento de la Covid-19

La prevención y el tratamiento de la Covid-19 aún sigue siendo un tema de discusión e investigación. Lo que sí es seguro, es que prácticas como el uso de mascarillas de tela o quirúrgicas, constante lavado y desinfección de manos con jabón, alcohol y gel, desinfección de espacios, uso de gafas anti-salpicaduras y/o visores faciales, entre otras; disminuyen las probabilidades de contagio (Moncada, et al., 2020). Por lo tanto, los materiales empleados en tales prácticas constituyen elementos claves para prevenir la covid-19.

Por otra parte, entre muchos insumos médicos asociados a la mitigación de los efectos de esta enfermedad, la literatura destaca el uso del fosfato de clorina para tratar

la neumonía provocada por el virus. A su vez, sugiere la aplicación de fármacos antiinflamatorios y antivíricos, así como como altas dosis de vitamina C (Nadeem, y otros, 2020).

Por último, se han desarrollado diversas vacunas cuya efectividad ha sido comprobada para la inmunización de la covid-19. Entre ellas, la AstraZeneca, Pfizer y Sinovac se encuentran disponibles en Ecuador dentro del plan de vacunación llevado a cabo por este (La Hora, 2021).

Para efectos de este estudio se considerarán todos los insumos médicos mencionados, además de los que sugieran los expertos en el área. Estos, en conjunto, serán de utilidad en el desarrollo y ejecución de la metodología planteada en el capítulo siguiente.

1.5.1.4 **Big data**

De acuerdo con Castillo (2018), el big data se define como un enorme conjunto de datos que sobrepasa la capacidad del software tradicional para poder ser capturados, gestionados y procesados; los cuales poseen tres características principales: volumen, variedad y velocidad (págs. 18-19).

Por su parte Kitchin (2014) manifiesta que, a pesar de que no hay una única definición para el big data, la literatura señala algunas características esenciales. Es así como, el big data son grandes conjuntos de datos, dinámicos, variados, detallados y de bajo costo que están interrelacionados de muchas formas diferentes. Además, acota que el crecimiento explosivo de los datos se basa en diferentes tecnologías, técnicas, infraestructura y procesos; y su inclusión en las prácticas cotidianas, como el uso de dispositivos móviles y redes sociales, que permite que muchos usuarios tengan acceso al mismo tiempo desde distintas ubicaciones.

Por otro lado, Oracle (2011) sostiene que el big data es la obtención de valor a partir de la toma de decisiones empresariales tradicionales basadas en bases de datos que se relacionan entre sí, aumentada con nuevas fuentes de datos no estructurados como blogs, medios de comunicación social, redes de sensores, datos de imágenes,

entre otros; que difieren en tamaño, estructura, y formato. Es decir, el big data es la inclusión de fuentes de datos adicionales para incrementar las operaciones existentes.

Para este caso de estudio se usará la definición de Kitchin debido a que toma en cuenta las diferentes tecnologías que se usan, así como su inclusión en las prácticas cotidianas a través de las redes sociales.

1.5.1.5 Redes sociales

Para Estévez (2013) las redes sociales son “plataformas alojadas en la web con servicios y herramientas que posibilitan a las personas interactuar con otros usuarios” (p. 67). En esa misma línea, Boyd y Ellison (2007) definen a las redes sociales como:

Servicios basados en la web que permiten a las personas construir un perfil público o semipúblico dentro de un sistema delimitado, articular una lista de otros usuarios con los que comparten una conexión y; ver y recorrer su lista de conexiones y las realizadas por otros dentro del sistema. (pág. 210)

De igual manera, para Celaya (2009) “las redes sociales son lugares en Internet donde las personas publican y comparten todo tipo de información, personal y profesional, con terceras personas, conocidos y absolutamente desconocidos” (págs. 106-108).

Para este caso de estudio se toma como definición de redes sociales la propuesta por Celaya, ya que es la que de manera general representa mejor el concepto a utilizarse.

1.5.1.6 Twitter

Gómez y García (2011) definen Twitter como “una herramienta que permite a los usuarios enviar y recibir de manera instantánea mensajes denominados tweets” (p. 539). Una definición más extensa la proporciona Falcón (2011) quien acota que Twitter es un servicio catalogado como una red social y un blog a la vez, ya que permite compartir pensamientos, información, enlaces, entre otros; con cualquier usuario web, así como comunicarse directamente con otros usuarios de la red ya sea de manera pública o privada. Como red social, Twitter permite crear comunidades de usuarios interconectados que comparten intereses comunes; mientras que como blog permite que el contenido creado por sus usuarios pueda ser visto por cualquier usuario web.

En este estudio se considerará la definición de Falcón al ser más representativa y guardar mayor relación con el objetivo de estudio.

1.5.1.7 **Web Scraping**

El web scraping o raspado web es una técnica que consiste en la extracción de información de uno o varios sitios web, a fin de procesarla en estructuras sencillas que luego se almacenan en bases de datos, archivos CSV u hojas de cálculo (Saurkar, Pathare, & Gode, 2018). Para la ejecución de esta técnica se hace uso de lenguajes de programación o aplicaciones especializadas en la obtención de información a través de sitios webs (Martínez, Rodríguez, Vera, & Parkinson, 2019).

En otras palabras, el raspado web constituye una herramienta capaz de transferir los datos no estructurados de sitios webs a medios de almacenamiento, de forma estructurada y automatizada; valiéndose de lenguajes de programación o softwares. De esta manera, los sitios webs se convierten en fuentes de datos valiosas para la investigación, la industria y demás agentes interesados en la información que proporcionan dichos sitios.

Dentro de dicho mundo virtual, las redes sociales corresponden a las más ricas fuentes de información sobre la idiosincrasia de las personas. Estas, a pesar de sus políticas de seguridad y privacidad, no son ajenas a ser rastreadas o exploradas. Una de esas redes sociales es Twitter, de la cual se puede obtener, principalmente, la opinión pública respecto a diversas temáticas.

Dada la facilidad que brinda el raspado web para la obtención de información de diversas fuentes, especialmente de la red social de interés en este estudio; es necesario tener presente la definición y características descritas previamente, a lo largo de este documento.

1.5.2 Marco metodológico

1.5.2.1 Entrevistas

La obtención de información a través de entrevistas es una estrategia común en el ámbito investigativo; por ello, la literatura relacionada a la realización y el análisis de estas es muy amplia (Azevedo, et al., 2017). La finalidad de las entrevistas es “es obtener información de forma oral y personalizada sobre acontecimientos, experiencias y opiniones de personas” (Folgueiras, 2005).

De esta manera, la entrevista figura como una herramienta idónea para la comprensión profunda del tema de estudio, desde diferentes perspectivas. Precisamente, en este caso se requiere de la experticia de los profesionales en el área de salud o afines, para la identificación de elementos claves en lo que respecta a la gestión de insumos médicos durante la pandemia de covid-19.

1.5.2.1.1 Entrevista Semiestructurada

Entre los diversos tipos de entrevistas, la entrevista semiestructurada es la más adecuada para este caso, ya que combina la precisión con la flexibilidad. En este tipo de entrevista es posible contar con un guion que sirva de referencia para realizar las preguntas de forma secuencial. No obstante, el entrevistador puede agregar preguntas en la medida en que lo considere pertinente durante la entrevista. Así mismo, el entrevistado puede proporcionar respuestas más extensas, permitiendo enriquecer el estudio (Folgueiras, 2005).

1.5.2.2 Minería de texto

La minería de texto, también conocida como “descubrimiento de conocimiento”, se puede describir como un subproceso de la minería de datos que consiste en la identificación y extracción de datos significativos a partir de datos textuales (Talib, Kashif, Ayesha, & Fatima, 2016). Este subproceso comprende cinco áreas fundamentales para la obtención de información relevante, las cuales se ilustran en la figura 1.1 y se describen en las siguientes secciones.

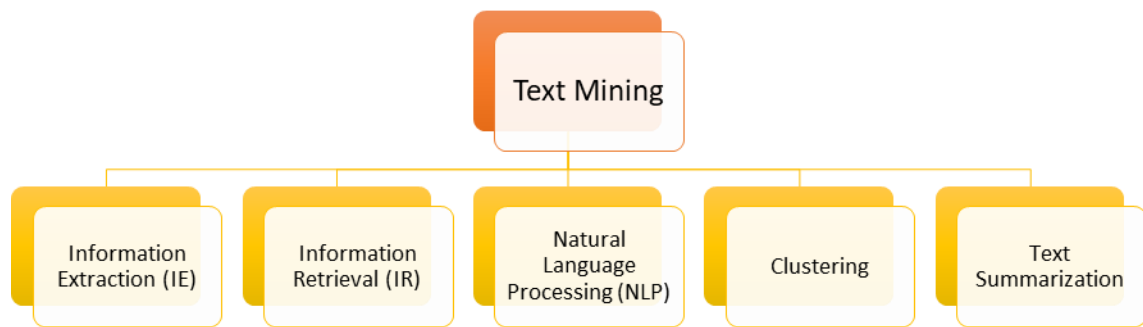


Figura 1.1 Áreas fundamentales de la minería de texto

Fuente: Talib, Kashif, Ayesha, & Fatima (2016)

1.5.2.2.1 Extracción de información

De acuerdo con Sarawagi (2008) es una técnica de minería de datos que permite extraer información importante de una gran cantidad de texto, en donde pueden ser categorizadas de acuerdo con las necesidades del usuario.

Por otro lado, Howard & Ruder (2018) consideran que la tarea de extraer información de documentos requiere demasiado esfuerzo y conocimiento del tema, puesto que existe una carga exagerada de texto donde la información de utilidad puede que se encuentre dispersa por todo el documento.

Gralinski, y otros (2020) mencionan que es necesario construir un sistema más sofisticado que posea precisión y un proceso de recuperación, logrando así extraer los datos de interés, y a su vez crear una base de datos con resultados favorables, debido a que buscar de manera secuencial la información mediante palabras exactas en todo el documento puede no ser lo más eficaz, ya que se estaría utilizando valores estandarizados.

1.5.2.2.2 Recuperación de información

La recuperación de información o, en inglés, “information retrieval” es un proceso que, como su nombre lo indica, recupera o devuelve información relevante de un conjunto de datos textuales. El proceso inicia con palabras claves que sirven como base

para la obtención de datos significativos asociados a dichos términos (Talib, Kashif, Ayesha, & Fatima, 2016).

Como se ha mencionado previamente, este estudio requiere de información que agregue valor a la comprensión de la gestión de insumos médicos. Por lo tanto, se establecerán palabras claves que serán empleadas en el algoritmo de recuperación de información para obtener los aspectos más relevantes asociados a dichas palabras. De tal manera que, esta herramienta reduce el tiempo de trabajo que implicaría hacerlo de forma manual, siendo útil y eficiente.

1.5.2.2.3 Procesamiento de lenguaje natural

Para Sierra (2014), el Procesamiento de Lenguaje Natural (PLN) es “una disciplina de la IA que se ocupa de la formulación e investigación de mecanismos computacionales para la comunicación entre personas y máquinas mediante el uso de Lenguajes Naturales” (pág. 3).

De manera similar, el PLN también es definido como “un subcampo de la inteligencia artificial que tiene como objetivo aprender, comprender, reconocer y producir contenido de lenguaje humano” (Zeroual & Lakhouaja, 2018, pág. 83); a través de sus áreas de desarrollo que incluyen la “recuperación y extracción de información, generación de resúmenes automáticos, minería de datos, análisis de sentimientos”, entre otras. (Hernández & Gómez, 2013, pág. 87)

1.5.2.2.4 Agrupación

La agrupación o “clustering” en inglés, refiere a una técnica que, en minería de texto, cumple la función de clasificar documentos de textos mediante la utilización de técnicas de agrupación no supervisada (Kaushik & Naithani, 2016). Es decir, para este caso de estudio, el algoritmo permitiría agrupar la información obtenida basándose en alguna característica o similitud, manteniendo así una mejor organización de los textos.

Uno de los algoritmos más utilizadas para aplicar este tipo de técnica es el K-media. Este refiere a una agrupación que consiste en establecer el centro de los grupos, los cuales se atribuirán a cada punto de datos. Luego se establecerá el lugar de cada

grupo en el valor medio de todos los puntos de datos que encajen en ese grupo (Talib, Kashif, Ayesha, & Fatima, 2016).

1.5.2.2.5 Síntesis de texto

La Síntesis de Texto, también conocida como Text Summarization, es una “técnica que permite resumir un texto de forma computacional y automática, basada en el resultado analítico de la idea principal o la esencia clave de los documentos” (Maylawati, Kumar, Y., Kasmin, & Ramdhani, 2019, pág. 1).

Esta importante tarea del Procesamiento del Lenguaje Natural tiene dos amplias categorías de enfoque, la extracción y la abstracción. La primera elabora un resumen a partir de un subconjunto de palabras o frases del texto original; mientras que la segunda construye primero una representación semántica interna, para luego crear un resumen a través de técnicas de generación de lenguaje natural (Gudivada, 2018).

Es decir, la síntesis de texto “identifica y extrae las frases clave de los documentos de entrada para producir resúmenes de texto automáticos”, “sin comprometer el contenido y significado del texto original” (Mohdab, Jan, & Shah, 2020, pág. 1).

1.5.2.3 Análisis de texto

El análisis de texto a menudo suele confundirse con la minería de texto, pero su diferencia radica los resultados que proporcionan. Como se ha mencionado previamente, la minería de texto tiene como objetivo devolver información relevante, lo cual implica que sus resultados tienden a ser más cualitativos. Mientras que el análisis de texto ofrece información cuantitativa, ya que utiliza los resultados de las técnicas de minería de texto, para encontrar patrones y tendencias (Welbers, Atteveldt, & Benoit, 2017).

Dado lo expuesto, y en congruencia con los objetivos de este estudio, se acogerán únicamente las técnicas asociadas a la minería de texto. Esto debido a que se requiere analizar la opinión pública en cuanto a la gestión de insumos médicos. Además, por factores de tiempo y recursos computacionales, no será posible abarcar la aplicación de técnicas de análisis de texto.

1.5.2.4 Indicadores de Efectividad

En los últimos años, las investigaciones de mercado y estudios afines se han concentrado en analizar la interacción de los usuarios en las redes sociales, como Twitter, para determinar la acogida de un determinado producto o servicio (Agarwal & Damle, 2020). En consecuencia, se han definido algunos indicadores clave de rendimiento, también conocidos como KPI. Los más empleados en Twitter son: número de likes o corazones, número de retweets, número de seguidores, engagement o compromiso, lealtad y rendimiento (Latorre-Martínez, Orive-Serrano, & Íñiguez-Dieste, 2018).

Aunque los KPI anteriores aportan información valiosa sobre el grado de aceptación por parte de la audiencia, estos son cuantitativos y es posible que no proporcionen información más profunda como las emociones. Es decir, si un tweet relacionado a un producto o servicio posee “más likes o corazones”, no implica que este será el más demandado o vendido (Agarwal & Damle, 2020).

Por ello, otro indicador frecuentemente utilizado en la literatura para este tipo de estudios es el análisis de sentimientos (Ferrer-Serrano, Latorre-Martínez, & Lozano-Blasco, 2020), ya que permite recoger información cualitativa importante como las emociones y sentimientos en una medida cuantitativa denominada "puntaje de sentimiento". Este puntaje puede sacar conclusiones sobre la efectividad y el éxito de un producto o servicio ofertado (Agarwal & Damle, 2020).

1.5.3 Marco referencial

García y Berton (2021) exploran una amplia cantidad de tweets relacionados al covid-19 en Brasil y Estados Unidos a través de la utilización de la identificación de temas y el análisis de sentimientos; donde concluyen que, la mayoría de los mensajes analizados son negativos dado que son datos sobre la pandemia del coronavirus; sin embargo, al observar las cifras, se evidencia diferencias significativas entre algunos temas en los que fueron clasificados los tweets previamente; ya que por un lado el cuidado de la proliferación fue el tema con mayor porcentaje de mensajes negativos,

alrededor de 60% en ambos países; y por otro lado, los tratamientos fueron el tema con mayor cantidad de mensajes positivos, el 33% de los mensajes en total (págs. 1-15).

Por otro lado, Khan et. al (2020) realizaron una investigación con el fin de brindar un análisis de sentimiento sobre la reacción de las personas hacia las decisiones relacionadas al covid-19 tomadas por el gobierno o autoridades locales, mediante un sistema que automatiza el análisis de los tweets y los clasifica como positivos, negativos o neutros. Los resultados demostraron que el número máximo de personas pensó y percibió como positivas dichas decisiones a pesar del aumento de personas infectadas y fallecidas. También se realizó un análisis de 3 meses correspondiente al subcontinente indio, el cual mostró que la variación de los sentimientos positivos, negativos y neutros se mantuvo constante con el número de casos que incrementaban día a día (págs. 2761-2774).

1.5.4 Marco legal

La Constitución de la República del Ecuador, es la Norma Suprema, en la cual se instauran las normas esenciales que defienden los derechos, libertades y obligaciones del Estado, las instituciones y de todos los ciudadanos; es así que en la misma se establece en el Art 363 que será el Estado el responsable de asegurar la disposición de los insumos médicos de calidad, garantizar el control de la comercialización y producción del mismo, y de otorgar el acceso a estos medicamentos de acuerdo a las necesidades epidemiológicas del pueblo ecuatoriano. (República, 2018).

Adicional a eso, se han creado reglamentos como el Reglamento de Buenas Prácticas de Almacenamiento, Distribución y Transporte para establecimientos farmacéuticos, el cual según el Art. 1 y el Art.2 establece las normas e implementa los instrumentos adecuados a fin de conceder la certificación que avale dichas buenas prácticas; todo esto, con el objetivo de que se mantengan las propiedades, características y calidad de los medicamentos, dispositivos médicos, reactivos, bioquímicos de diagnóstico, productos para la industria farmacéutica, entre otros.

Para el cumplimiento de aquello, es importante un correcto proceso de distribución, por lo cual los Art. 60 y Art, 64 hacen referencia al aval de la correcta conservación de los productos; es decir, que los recipientes, cajas o métodos de embalaje utilizados en el proceso afecten la calidad del producto, sino que más bien deben protegerlo de factores externos; lo cual es muy importante mencionar debido a que forma parte de la responsabilidad del establecimiento.

Asimismo, respecto a la transportación de los productos los Art. 66 y Art. 68 establece que el establecimiento debe asegurar la integridad de los productos transportados, tomar las medidas necesarias para evitar derrames, rupturas, contaminación por otros materiales y asegurar el mantenimiento de la cadena de frío en los productos que requieran conservación a temperatura controlada. (República del Ecuador, 2014).

Por otra parte, debido a que en este proyecto se utilizarán datos provenientes de la red social Twitter es necesario considerar la Ley Orgánica de Protección de Datos Personales, cuyo propósito es procurar el apropiado tratamiento y flujo de datos personales a fin de proteger los derechos esenciales de los ciudadanos como el de las libertades individuales (Art. 2); sin embargo, al tratarse de datos anonimizados, es decir que no es posible identificar al titular, la ley no tiene aplicativo, quedando exentos de ella. (Asamblea Nacional, 2021)

No obstante, es necesario tener presente el proyecto de ley del Uso Responsable de Redes Sociales planteado en el 2019, ya que en caso de ser aprobado y de constar en el Registro Oficial podría convertirse en una limitante de este proyecto. La razón se encuentra en el Art, 7 del mismo, el cual planea que será necesaria la autorización expresa de los usuarios para capturar y almacenar su información personal, así como la transparencia en la utilización de esas bases de datos, caso contrario tendrá una sanción civil o penal de acuerdo con el caso.

CAPÍTULO 2

2. METODOLOGÍA

2.1 Diseño de la investigación

El presente estudio tiene un enfoque cualitativo de tipo observacional, prospectivo, transversal y descriptivo; lo cual implica la recopilación y representación sistemática de datos a fin de caracterizar una determinada situación. Por ello, en esta investigación se empleó la entrevista y el web scraping o raspado web como instrumentos de recolección de datos, para luego realizar un análisis aplicando técnicas de minería de texto. El detalle de los instrumentos y procedimientos ejecutados se describe en las siguientes secciones. La metodología general se resume en la figura 2.1.



Figura 2.1 Diseño de la Investigación

2.2 Entrevista

2.2.1 Diseño de la entrevista

Para la entrevista se optó por un estilo semiestructurado que constó de tres partes: experiencias durante la pandemia, insumos médicos y perspectiva sobre la población; estas cuentan con seis, cinco y tres preguntas, respectivamente. La entrevista fue dirigida a profesionales de la salud, quienes estuvieron en la primera línea de defensa contra el coronavirus, tales como: médicos generales, internistas y especialistas en áreas afines a inmunología, cardiología, virología, entre otros.

La ejecución de entrevistas se llevó a cabo del 10 al 15 de julio del presente año, de forma virtual a través de la plataforma Zoom con una duración aproximada de dos horas por entrevista. Para ello, primero se estableció el contacto con los entrevistados por medio de una invitación formal vía correo electrónico. En esta invitación se proporcionó un enlace de Doodle, en donde se dio a elegir a los entrevistados los horarios de su preferencia para el desarrollo de la entrevista.

Una vez confirmada la participación de los entrevistados, se procedió a enviarles un correo de agradecimiento como respuesta y, adicionalmente, el enlace de acceso a la entrevista a través de Zoom según el horario más votado por el entrevistado en Doodle. Posteriormente, el día de la entrevista, se envió un correo como recordatorio 15 minutos antes de que esta inicie.

2.2.2 Validación de entrevistados

Dado que la profesión médica es altamente demandante y ahora más por la pandemia, únicamente fue posible entrevistar a dos profesionales en medicina. No obstante, para efectos de este estudio, fue fundamental que estos cumplan con aspectos que proporcionen mayor credibilidad a su opinión o perspectiva y estos fueron: que su experiencia laboral sea mayor a 5 años y que hayan trabajado durante la pandemia. Considerando estas condiciones, en las siguientes subsecciones se presenta una breve validación de los entrevistados.

2.2.2.1 Dr. Miguel Icaza

El doctor Miguel Icaza ejerce su profesión en el Hospital de Especialidades “Dr. Abel Gilbert Pontón” de Guayaquil. Cuenta con más de 10 años de experiencia en el campo y, dentro de su experiencia, se destaca su intervención en pacientes con condiciones críticas de salud, dada su especialidad de médico interno. Por ello, durante la pandemia ha atendido principalmente a pacientes con coronavirus en estado grave o crónico, lo cual lo convierte en un testigo fiel de la realidad sanitaria de la época.

2.2.2.2 Dra. Lorena Espinoza

La doctora Lorena Espinoza labora como pediatra en un centro de salud tipo C de Esmeraldas, perteneciente a la red del Ministerio de Salud Pública. Aunque durante la pandemia también tuvo la oportunidad de laborar en varios hospitales de Guayaquil debido a la demanda de médicos. Su experiencia comprende más de una década como médico general y cerca de 5 años como especialista en Pediatría. De manera que, a lo largo de la crisis sanitaria ha podido evidenciar los efectos del covid-19 tanto en adultos como en niños y mujeres embarazadas.

2.3 Web scraping

Para poder extraer información de la red social Twitter, se utilizó la librería Twint de Python, la cual es de código abierto y permite obtener tweets de los usuarios sin la necesidad de tener las limitaciones de la API, cabe destacar que la red social Twitter tiene limitado el raspado de información de aproximadamente 3200 tweets por mes buscado.

Para la extracción de la información mediante Twint se delimitó la zona geográfica en Ecuador con coordenadas de latitud, longitud y un radio en kilómetros, donde se genera una circunferencia para una búsqueda más exacta. Adicionalmente, se establecieron parámetros de tiempo desde el 1 de marzo del 2020 hasta el 31 de marzo del 2021, siendo así que se genera un archivo csv detallando los siguientes datos:

- **Id_username:** Nombre del usuario del que realiza el tweet.
- **Retweets_count:** Cuántos retweets tuvo la publicación.
- **Likes_count:** Cuántos corazones tuvo la publicación.
- **Hashtags:** Cuáles son los #hashtags en la publicación.
- **Place:** Coordenadas de latitud y longitud del usuario.
- **Date:** Día en el que se ha realizado la publicación.
- **Language:** En qué idioma se encuentra el tweet.

Por otro lado, para el proceso y lectura de la base de datos se utilizó el lenguaje de programación Python junto con su IDE Pycharm con el fin de

realizar un mapeo para identificar a qué provincia pertenecen las coordenadas de cada usuario de twitter. Para ello, se hizo uso de un mapa en formato geojson realizado por Marin (2017), el cual permite visualizar las 24 provincias del Ecuador y sus zonas no delimitadas.

Como primer paso se extrajo las coordenadas de los usuarios por tweet, con sus respectivos ejes de longitud y latitud, lo cual permite que este sea catalogado como un punto en el Ecuador. Después, mediante la librería Shapely se buscaba a qué provincia pertenece la persona que realizó el tweet, con el objetivo de generar un mapa de colores, en donde se muestre gráficamente los resultados de forma cuantitativa.

2.4 Extracción de información

2.4.1 Búsqueda de patrones

Como se mencionó en el capítulo anterior, la extracción de información consiste en la obtención de categorías del texto, es decir, determinar el tipo de palabra de acuerdo con su función en la oración. Los tipos de palabras identificadas por esta técnica son: sustantivo, pronombre, verbo, adjetivo, adverbio, preposición, conjunción e intersección (Jurafsky & Martin, 2021); como se muestra en la figura 2. El conocimiento de la categoría de la palabra facilita la identificación de patrones y obtención de palabras claves que permitan contar con información relevante, según los objetivos del investigador (Ahmad & Varma, 2018).

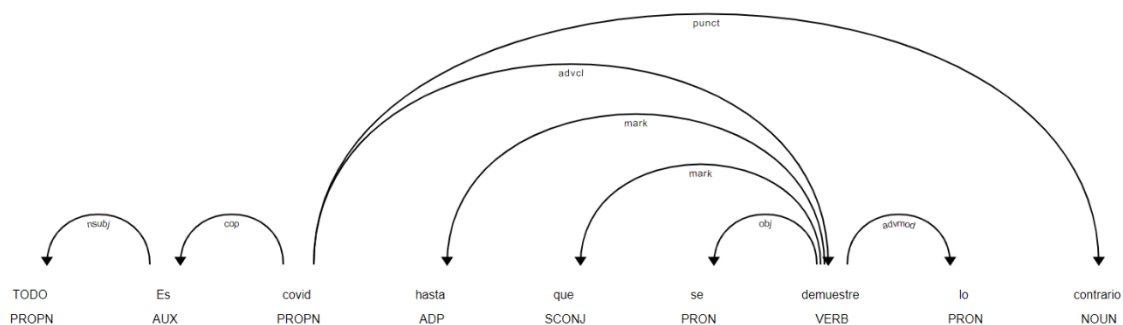


Figura 2.2 Relación de palabras según su tipo

Fuente: Autores

En este caso, se requería conocer las características atribuidas a los dos sustantivos que se presentaban con mayor frecuencia en todas las publicaciones, a los cuales se denominó “palabras bases”. Para ello, se optó por extraer palabras frecuentes y una vez obtenido el par de términos más frecuentes, se establecieron reglas o patrones que permitieran asociar dichos sustantivos con uno o varios adjetivos. En la determinación de las reglas se consideraron todas las palabras derivadas, similares o compuestas respecto a los dos términos más frecuentes; por ejemplo: covid, covid-19, covid19, COVID-19, etc. Los patrones establecidos se detallan en la tabla 2.1.

Nº	Patrón	Descripción	Ejemplo
1	Palabra base + adjetivo	Obtener una caracterización simple respecto a la palabra base.	Mascarilla efectiva
2	Palabra base + adjetivo + adjetivo o sustantivo complementario	Obtener una caracterización compuesta respecto a la palabra base.	Mascarilla efectiva KN95

Tabla 2.1 Descripción de patrones

Fuente: Autores

A partir de dichos patrones, se realizó la búsqueda de tweets que cumplieran con estos. La búsqueda y configuración de reglas fueron llevadas a cabo mediante la librería Spacy en Python. Posteriormente, los resultados fueron cuantificados y visualizados para comprender la percepción de la población respecto a los términos bases y conservar la información relevante.

2.5 Recuperación de información

Seguido de la extracción de la información, fue necesario aplicar la recuperación de la información como última fase de limpieza previo a la aplicación del procesamiento de lenguaje natural. Como se mencionó en el capítulo anterior, esta técnica se basa en un conjunto de palabras claves previamente definidas que permiten recuperar información útil y relevante de la base de datos textuales conformado por los tweets (IBM, 2020). Por lo tanto, este

paso es fundamental para mejorar la precisión de los tweets asociados a los objetivos de la investigación.

Anteriormente twitter restringía las publicaciones hasta 140 caracteres por tweet, por lo que el usuario ocupaba diferentes métodos tales como: abreviaturas, supresión de letras, creación de hilos de tweets, entre otras cosas; lo cual limitaba la obtención de la información importante. Sin embargo, en el año 2017 la red social duplicó la cantidad de caracteres proporcionándole a sus usuarios mayor amplitud para expresar sus sentimientos en un mismo tweet, logrando a su vez fomentar mejores opiniones y, por ende, aumentado la probabilidad de recuperar información valiosa (Rosen, 2017).

2.6 Procesamiento de lenguaje natural

2.6.1 Análisis de sentimiento

Para el análisis de sentimiento se ocupó la librería classifier proveniente de SentimentClassifier, el cual realiza un preprocesamiento de la información eliminando lo que no es relevante, como signos de puntuación que hayan sido mal utilizados; además de remover páginas web, eliminar caracteres especiales, y realizar una modificación en las letras que se encuentren en mayúsculas y minúsculas permitiendo tener un mejor procesamiento de los datos.

Sin embargo, no convierte todas las letras en minúsculas, debido a que trabaja con Word Embeddings, el cual es una técnica del procesamiento natural del lenguaje y la modificación completa de la palabra puede cambiar el contexto gramatical para determinar si el tweet es bueno, neutro o malo. Esto genera una probabilidad que se clasifica de la siguiente manera:

- Cuando la probabilidad se acerca a 0 es considerado un tweet malo.
- Cuando la probabilidad se acerca a 0.5 es considerado un tweet neutral.
- Cuando la probabilidad se acerca a 1 es considerado un tweet bueno.

Lo cual permite identificar las emociones de los tweets y clasificarlas de acuerdo con su sentimiento ya sea bueno, malo o neutral.

2.7 Clustering

Como se mencionó previamente, el análisis de sentimientos proporciona información cualitativa de gran valor para comprender de manera profunda lo que la información cuantitativa por sí sola no puede explicar. Sin embargo, al tratarse de registros de todo el Ecuador, se contaba con una perspectiva general que limitaba la interpretación específica de los resultados. En ese sentido, para mejorar la caracterización del problema, se optó por segmentar o agrupar la información mediante el algoritmo K-medias.

2.7.1 K-medias

El algoritmo K-medias es uno de los más empleados en la literatura por su simplicidad, sobre todo en problemas de aprendizaje no supervisado (Adewole, Han, Wu, Song, & Sangaiah, 2018), como en este caso. No obstante, la ejecución de este algoritmo depende un número K que representa la cantidad de grupos a formar. Para obtener este número se pueden emplear diferentes métodos, entre ellos el método del codo, el cual se utilizó en este estudio debido a su bajo costo computacional y el poco espacio de memoria que consume (Ahuja & Dubey, 2017).

Una vez obtenido el número K correspondiente, se ejecutó el algoritmo K-medias considerando como variables el número de likes o corazones, el número de retweets y el valor del sentimiento. Posteriormente, los resultados de este análisis fueron segmentados e ilustrados a nivel provincial para lograr una caracterización más precisa.

2.8 Text summarization

Para poder validar la información recuperada, se utilizó la librería NLTK de Python, la cual es "un conjunto de módulos de programa de código abierto, tutoriales y conjuntos de problemas, que proporciona material didáctico de lingüística computacional" (Loper & Bird, 2020). La importancia de utilizar esta herramienta radica en que reduce la probabilidad de tener datos aberrantes los cuales podrían sesgar los resultados de la investigación; es decir, si no se realiza

la validación, se estarían procesando en exceso datos que no proporcionen información relevante, tales como palabras vacías o stopwords.

Para ello, una vez realizado el pre-procesamiento de información se generó un score utilizando la técnica *cosine similarity* con el fin de calcular la importancia de cada vector. El cosine similarity es un modelo que no necesita ser entrenado debido a que trabaja con un algoritmo interno, en donde convierte las oraciones en vectores calculando el ángulo de coseno de acuerdo con la frecuencia de palabras que se encuentre inmerso en cada uno de los vectores. Luego de calcular la importancia, toma los vectores que tengan la puntuación más alta para realizar un resumen que tenga estructura gramatical de acuerdo con ese orden.

2.9 Indicadores de Efectividad

En congruencia con lo reportado por estudios similares, los indicadores de efectividad corresponden a medidas de carácter cuantitativo y cualitativo que se pueden extraer de los tweets a fin de conseguir información sobre la interacción y capacidad de respuestas de los usuarios. En este caso, con los resultados de las técnicas aplicadas y los reportes emitidos por el Comité de Operaciones de Emergencia (COE) nacional e instituciones afines, se establecieron cuatro indicadores: tiempos críticos, zonas críticas, proyección de pérdidas económicas en el sector salud y proyección en la demanda de ataúdes.

2.9.1 Tiempos críticos

Para definir los tiempos críticos, se graficó una serie de tiempo de todos los tweets obtenidos, identificando los picos o puntos destacados que indicaban un comportamiento diferente o significativo. Luego, se compararon los resultados obtenidos con los tiempos críticos declarados por el COE Nacional, a fin de identificar semejanzas o diferencias y validar los resultados del análisis ejecutado.

2.9.2 Zonas críticas

Las zonas críticas se definieron a partir los grupos identificados mediante la técnica de agrupación, los cuales fueron ilustrados en el mapa de Ecuador. Los resultados obtenidos se compararon con la información reportada por el COE Nacional, respecto a las zonas más afectadas durante la pandemia, considerando el período de estudio.

2.9.3 Proyección de pérdidas económicas en el sector de la salud

Se realizó una proyección de pérdidas económicas en el sector de la salud de marzo a mayo del 2020, a partir de los resultados de un análisis realizado por el Ministerio de Salud y la Secretaría Técnica de Planificación del Ecuador, y la Organización Internacional del Trabajo (OIT). En la tabla 2.2 se detalla la composición de las pérdidas económicas en este sector que equivalen a 870 millones aproximadamente, según el análisis realizado por dichas entidades.

Componente	Total (millones USD)
Producción y distribución de bienes y servicios de salud	724.69
Recursos, infraestructura y activos físicos del sector salud	83.37
Acceso de la población a bienes y servicios de salud	59.80
Gobernabilidad del sector salud	1.60
Incremento de la vulnerabilidad y riesgos del sector salud	0.53
Total de pérdidas del sector salud	869.99

Tabla 2.2 Estimación de pérdidas económicas en el sector de la salud durante el período de marzo a mayo del 2020

Fuente: Organización Internacional del Trabajo (2021)

Para calcular la proyección de pérdidas del sector salud, primero se obtuvo la proporción de pérdidas del sector salud con respecto a la población ecuatoriana, es decir, la representación de la pérdida económica por habitante (ecuación 2.1). Luego, la proporción obtenida se multiplicó por el total de tweets equivalentes a los meses de marzo a mayo de 2020, considerando que cada tweet provenía una cuenta diferente (ecuación 2.2). En consecuencia, se obtuvo

la proyección de la pérdida con relación a los ciudadanos que emplearon sus cuentas de Twitter para gestionar insumos.

$$(2.1) \textit{Proporción de pérdidas del sector salud} = \frac{\textit{Total de pérdidas del sector salud}}{\textit{Total de la población ecuatoriana}}$$

$$(2.2) \textit{Proyección de pérdidas del sector salud} \\ = \textit{Proporción de pérdidas del sector salud} * \textit{Total de tweets equivalentes}$$

2.9.4 Proyección de la demanda de ataúdes en Guayas

De la misma manera, se estableció una proyección de la demanda de ataúdes en Guayas durante los inicios de la pandemia. Según España & otros (2020), durante los meses de marzo y abril del 2020, se emitieron 17000 actas de defunción en total, siendo 8.5 veces más que las actas emitidas mensualmente en un período normal. Es decir, hubo exceso considerable en la demanda total de ataúdes entre esos meses, para la cual los productores u ofertantes no se encontraban preparados.

Con esta referencia se procedió a calcular la proyección de la demanda de ataúdes en Guayas durante ese lapso. Para ello, se dividió el total de actas de defunción emitidas entre el total de tweets recopilados en Guayas (ecuación 2.3), obteniendo así el valor equivalente a la demanda de ataúdes. Luego, se comparó lo obtenido con la proporción de la demanda de ataúdes, esta última se calculó a partir de la relación entre el total de actas de defunción emitidas y el total de la población guayasense (ecuación 2.4).

$$(2.3) \textit{Proyección de demanda de ataúdes} = \frac{\textit{Total de actas de defunción emitidas}}{\textit{Total de tweets en Guayas}}$$

$$(2.4) \textit{Proporción de demanda de ataúdes} = \frac{\textit{Total de actas de defunción emitidas}}{\textit{Total de la población guayasense}}$$

CAPÍTULO 3

3. RESULTADOS Y ANÁLISIS

3.1 Análisis de entrevistas

3.1.1 Entrevista a Dr. Miguel Icaza

De la entrevista con el Dr. Icaza se destacó que, tanto a nivel público como privado, se llevaron a cabo las medidas necesarias para afrontar la emergencia sanitaria y que, en todo tiempo, se priorizó el bienestar y la seguridad del personal de salud. Los insumos necesarios como: mascarillas, alcohol, gel antibacterial, entre otros.

Sin embargo, a pesar de que los hospitales cuentan con bombas de oxígeno, este fue cada vez más escaso a medida que avanzaba la pandemia y debido a la poca colaboración de la población al seguir los cuidados recomendados. Según el doctor, esta situación fue más crítica en la segunda ola de la pandemia, es decir, de abril a junio de 2020.

3.1.2 Entrevista a Dra. Lorena Espinoza

La entrevista con la Dra. Espinoza brindó una perspectiva distinta de la obtenida con el Dr. Icaza, ya que, por el contrario, el lugar donde desempeñó sus labores padeció de una gran escasez de mascarillas, específicamente, las N-95. A partir de esta escasez se tuvo que establecer un protocolo a fin de extender la vida útil de este insumo y este consistió en reutilizar las mascarillas tres guardias seguidas, es decir, a la semana solo empleaban dos mascarillas de ese tipo.

A partir de las entrevistas realizadas, se determinaron los insumos más mencionados o comunes y estos fueron:

1. Mascarilla
2. N-95
3. Nitazoxanida

4. Paracetamol
5. Oxígeno
6. Azitromicina
7. Amoxicilina
8. Ivermectina
9. Jengibre
10. Canela
11. Eucalipto
12. Árbol de neem
13. Colufase
14. Alcohol
15. Gel antibacterial

Con este conjunto de palabras, se procedió a realizar la extracción de datos y a aplicar las técnicas de minería de texto establecidas en el capítulo anterior.

3.2 Web Scraping

En la búsqueda del archivo se ocuparon las palabras claves que se obtuvieron directamente de las entrevistas, las cuales son: mascarilla, N-95, nitazoxadina, paracetamol, oxígeno, axitromicina, amoxicilina, ivermectina, jengibre, canela, eucalipto, árbol de neem, colufase, alcohol, gel antibacterial; lo cual dio como resultado una base de datos pura de 131098 tweets en la que estaban incluidos los tweets de todo el país. Además, se obtuvieron las características de los tweets, tal como retweets y likes o corazones. La cantidad de retweets fue de 556151 y la cantidad de likes fue de 1276567, estos posteriormente fueron de utilidad para aplicar las técnicas de minería de texto, cuyos resultados se describen en las siguientes secciones.

3.3 Resultados de extracción de información

3.3.1 Palabras frecuentes

Previo a la identificación de patrones, se obtuvo un top 10 de palabras frecuentes para seleccionar el par de términos que servirían como palabras bases para la determinación de patrones. Como se esperaba, en congruencia

con el estudio en cuestión, los resultados indicaron que los dos sustantivos más frecuentes entre las publicaciones de Twitter eran “covid” y “coronavirus”, siendo mencionados más de 30000 y 10000 veces, respectivamente. Seguido de estos términos, se encontraron otras palabras comunes en el contexto pandémico, las cuales se muestran en la figura 3.1.

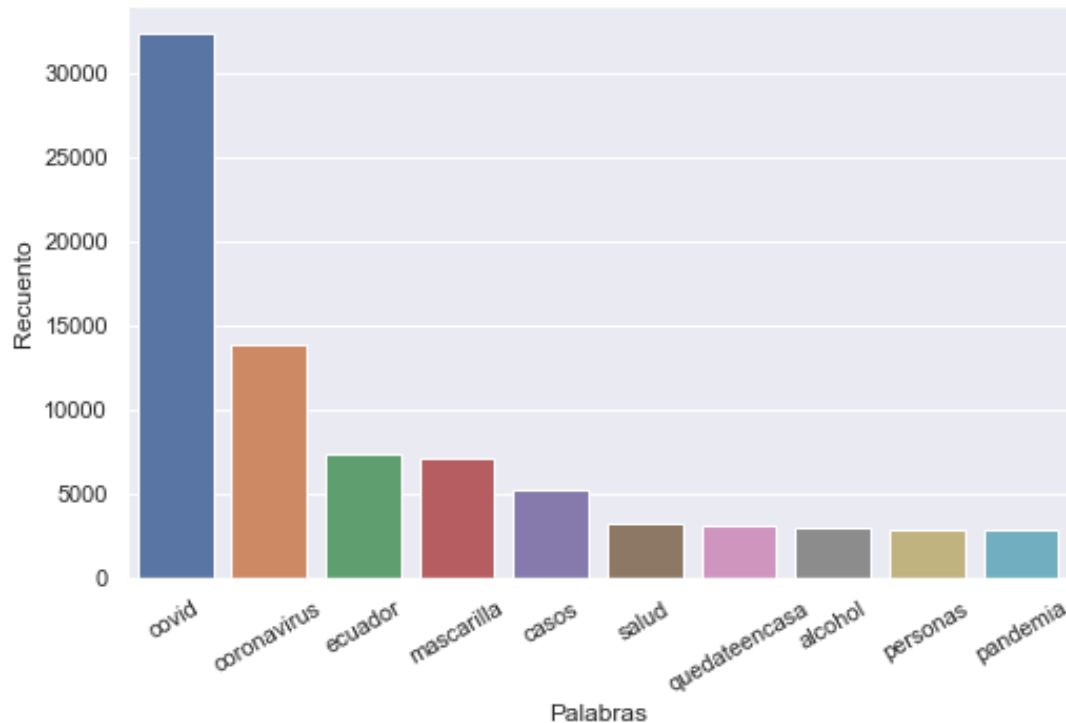


Figura 3.1 Top 10 de palabras frecuentes

Fuente: Autores.

3.3.2 Identificación de patrones

Una vez obtenidas las palabras bases “covid” y “coronavirus”, se procedió a establecer los patrones descritos en la metodología, por cada uno de los términos mencionados y, luego, se realizó la búsqueda de tweets que cumplieran con las reglas. De la palabra base “covid” se obtuvieron 314 patrones, de los cuales 296 patrones correspondían al primer patrón y 18, al segundo. Mientras que del término “coronavirus” se obtuvieron 399 patrones, de los cuales 371 correspondían al primer patrón y 28, al segundo. Gráficamente los resultados se muestran en las figuras 3.2 y 3.3.

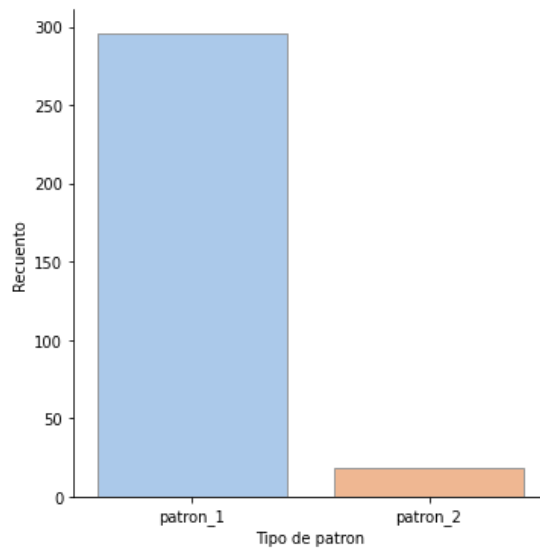


Figura 3.2 Patrones asociados a la palabra "Covid"

Fuente: Autores

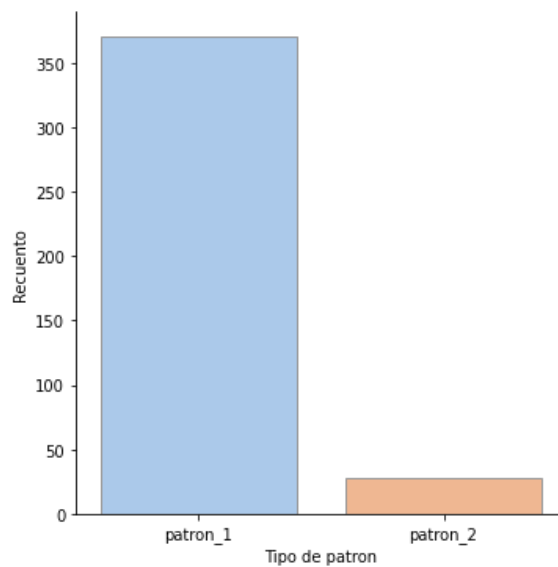


Figura 3.3 Patrones asociados a la palabra "Coronavirus"

Fuente: Autores.

Dada la cantidad de patrones identificados, mostrar cada uno de ellos resultaría complejo. Por ello, a continuación, se presenta una muestra de los patrones identificados por cada una de las palabras bases.

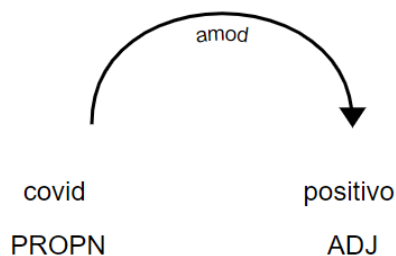


Figura 3.4 Muestra del primer patrón asociado a la palabra "covid"

Fuente: Autores

Como se observa en la figura 3.4, para la palabra “covid”, el primer patrón identificado asocia esta con el adjetivo “positivo”. En primera instancia se denota que, en efecto, se cumple adecuadamente la regla y se conserva el sentido de este resultado con el contexto estudiado, es decir, la pandemia. Esta relación identificada hace referencia a los casos positivos del virus que, hasta ahora, continúan siendo controversiales.

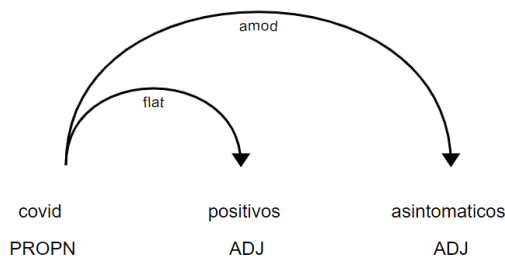


Figura 3.5 Muestra del segundo patrón asociado a la palabra "covid"

Fuente: Autores.

Tal como se presenta en la figura 3.5, para el segundo patrón de la palabra “covid”, relaciona esta con los adjetivos “positivos” y “asintomáticos”, haciendo referencia los casos positivos del virus y una de las características de este. Nuevamente, se comprueba que el tweet asociado al patrón cumple adecuadamente con sus especificaciones y mantiene una relación coherente entre estas.

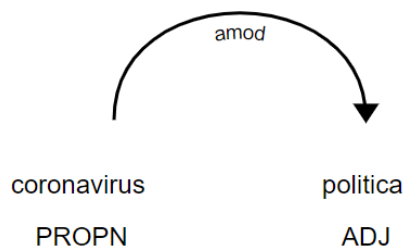


Figura 3.6 Muestra del primer patrón asociado a la palabra "coronavirus"

Fuente: Autores

La figura 3.6 representa la relación identificada entre la palabra base “coronavirus” y el adjetivo “política” en respuesta al primer patrón. Dicho adjetivo indica que la relación debe interpretarse desde una perspectiva política, lo cual es congruente con uno de los sucesos que intensificó el impacto negativo de la pandemia en Ecuador durante el periodo de estudio y esta situación fue la gestión de la emergencia sanitaria por parte de las autoridades mediante políticas y normas.

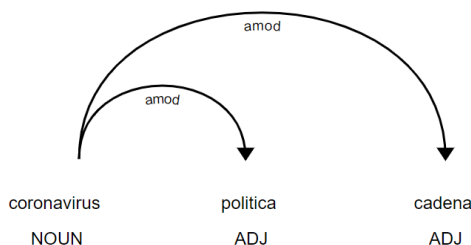


Figura 3.7 Muestra del segundo patrón asociado a la palabra "coronavirus"

Fuente: Autores

El segundo patrón asociado a la palabra coronavirus complementa la comprensión del primer patrón, haciendo referencia a las cadenas políticas que se intensificaron por ese tiempo debido a algunos casos de corrupción asociados a la distribución de mascarillas KN95 y otros insumos. De manera que el patrón es consistente con el tema en cuestión.

Por lo tanto, este procedimiento permitió verificar que los tweets sean relevantes y así conservar aquellos que realmente aportaban información a las campañas de insumos médicos. Los tweets verificados y conservados con este

procedimiento fueron alrededor de 131000, confirmando la pertinencia de los tweets recopilados mediante web scraping.

3.4 Recuperación de Información

De la base de 131098 tweets obtenidos, se realizó una limpieza de datos aberrantes que no pertenecían al territorio ecuatoriano empleando la técnica de recuperación de información. Para ello, se utilizó el código del tweet junto a una condición que consistía en que, si el tweet pertenecía a una provincia del Ecuador, este era considerado un dato válido, caso contrario era eliminado. Esto permitió obtener una base de datos completos y validos de 112578 tweets. Por último, a partir de las palabras claves se comprobó que cada uno de los tweets estaban relacionados a la gestión de insumos médicos, conservando información relevante.

3.5 Resultados del Análisis de Sentimiento

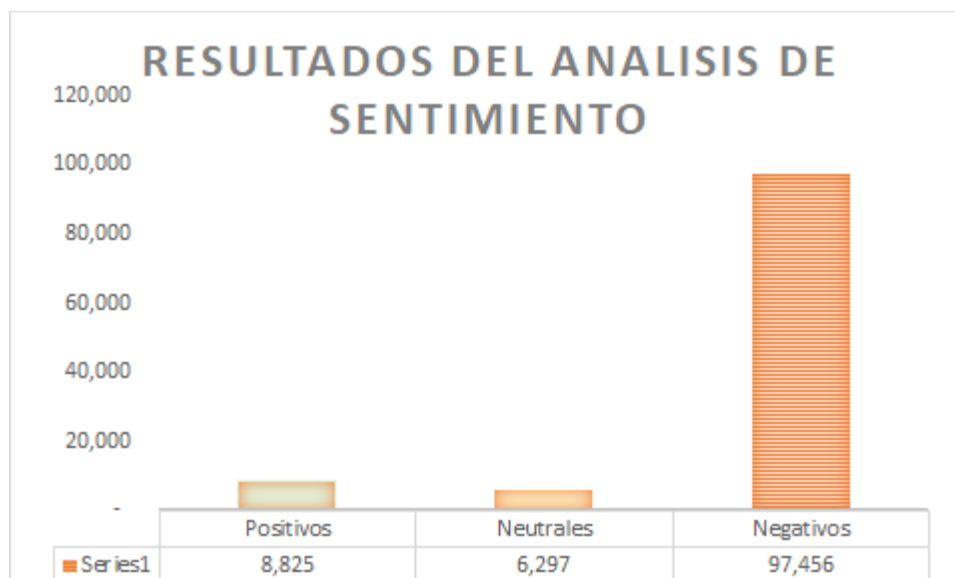


Figura 3.8 Resultado del análisis de sentimiento

Fuente: Autores

Como se puede apreciar en la figura 3.8, los resultados que ha otorgado el análisis de sentimiento determinan que se han analizado un total de 112578 tweets, de los cuales 8825 son positivos y corresponden el 7.84% del total de la muestra, 6297 son neutrales y corresponden el 5.59% de la muestra para lo cual

se consideraron los valores que oscilan con un margen ± 0.05 , y finalmente 97456 son negativos y representan el 86.57%.

Tabla 3.1 Resultado de Análisis de Sentimiento

Provincias	#Tweets	# Positivos	#Neutrales	#Negativos
Azuay	17513	1386	961	15166
Bolivar	676	49	37	590
Cañar	2171	107	86	1978
Carchi	12	1	1	10
Cotopaxi	10459	716	558	9185
Chimborazo	434	25	20	389
El oro	2213	189	109	1915
Esmeralda	1376	117	62	1197
Guayas	24349	2056	1495	20798
Imbabura	19072	1319	1025	16728
Loja	994	76	40	878
Los Ríos	2328	187	140	2001
Manabí	1856	174	116	1566
Morona Santiago	330	30	18	282
Napo	9571	662	516	8393
Pastaza	154	35	7	112
Tungurahua	1100	109	50	941
Pichincha	2470	189	144	2137
Zamora Chinchipe	1359	108	69	1182
Galapagos	133	14	7	112
Sucumbios	291	23	12	256
Orellana	151	37	6	108
Santo Domingo de los Tsachilas	117	8	10	99
Santa Elena	13403	1203	800	11400
Zona no delimitada	46	5	8	33

Fuente: Autores

Por otro lado, en la Tabla 3.1 se puede apreciar la cantidad total por provincia de los tweets considerados como negativos, positivos y neutrales. La mayor cantidad de tweets negativos provienen de Guayas (21%) seguidos de Imbabura (17%), Azuay (16%), Santa Elena (12%) y Cotopaxi (9,4% de acuerdo con lo que se puede apreciar en la figura 4, lo cual está relacionado con el número total de tweets, ya que justamente son las provincias de las cuales se registró un mayor número de tweets durante el tiempo considerado en el estudio.

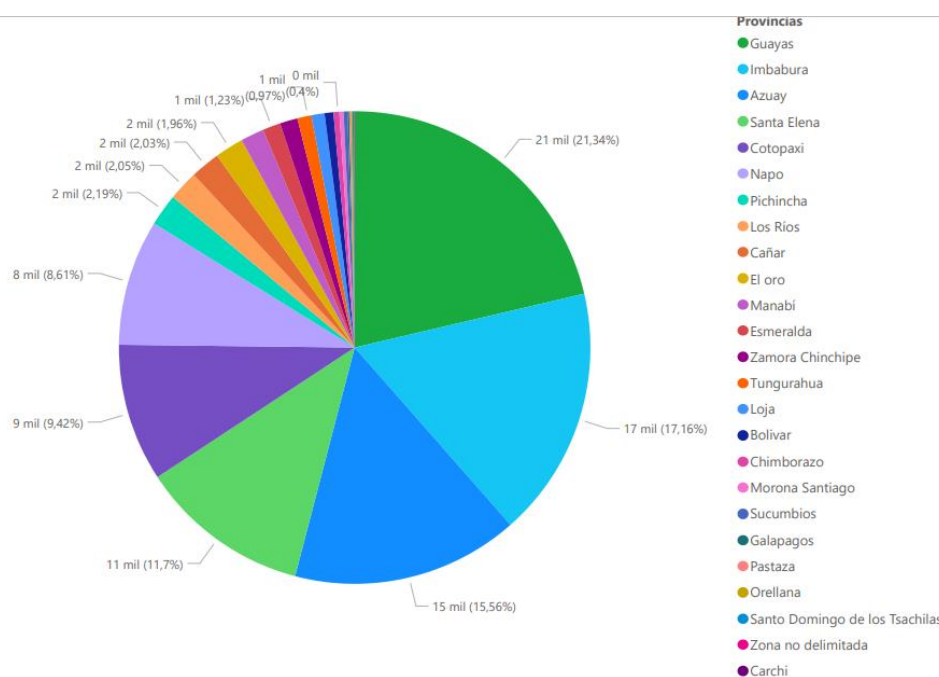


Figura 3.9 Porcentaje de tweets negativos por provincias

Fuente: Autores

De igual manera, la figura 3.10 muestra que Guayas (23%), Imbabura (16%), Azuay (15%), Santa Elena (14%) y Cotopaxi (8%) encabezan las listas de las provincias que más emitieron tweets catalogados como positivos durante el período de estudio del proyecto, lo cual también es congruente con la cantidad de tweets que provienen de cada una de las provincias mencionadas.

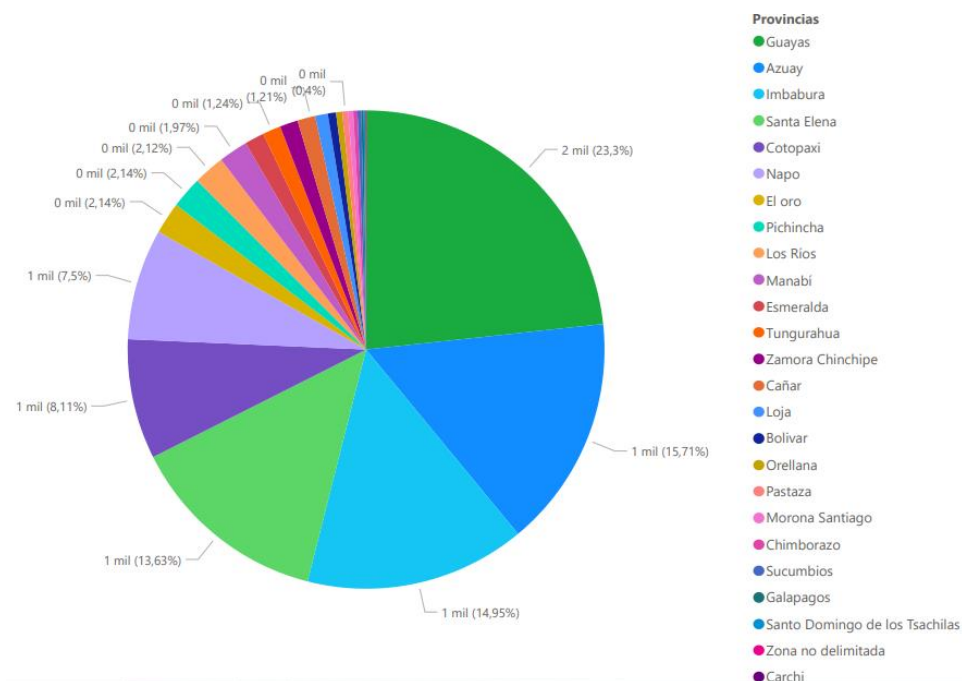


Figura 3.10 Porcentaje de tweets positivos por provincias

Fuente: Autores

A través de este análisis descriptivo se determina que la mayor cantidad de tweets fueron negativos y que las provincias que más emitieron tweets tanto positivos como negativos son Guayas, Imbabura, Azuay, Santa Elena y Cotopaxi, las cual conforman aproximadamente el 75% de la población estudiada.

3.6 Identificación de la efectividad de las campañas de insumos médicos

Como se mencionó en el capítulo anterior, mediante esta técnica se relacionaron tres variables o características: los sentimientos o emociones, los likes o corazones y la cantidad de retweets. Estas variables, en otras palabras, representan indicadores que en conjunto posibilitan la comprensión de la efectividad de las campañas de insumos (Agarwal & Damle, 2020). La dispersión de los tweets con respecto a dichas variables se muestra en la figura 3.11.

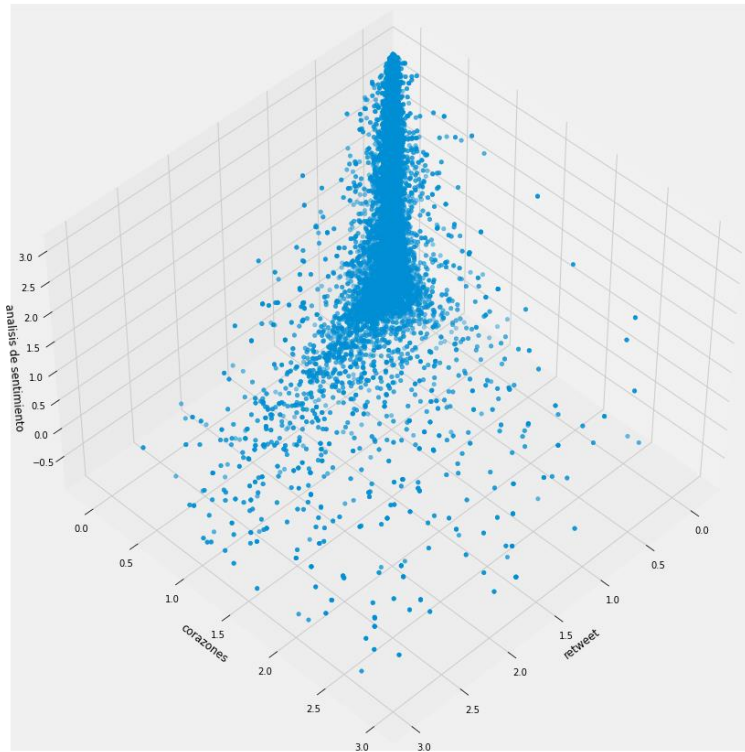


Figura 3.11 Dispersión de los tweets con base al análisis de sentimiento, corazones y retweets

Fuente: Autores

3.6.1 Método del codo

Los resultados del método del codo aplicado para obtener el número K apropiado de grupos a formar, mostraron que este era 2. Como se observa en la figura 3.12, la curva se aplana significativamente en el cluster 2, lo que se traduce en el número de grupos (Tibshirani, Walther, & Hastie, 2001).

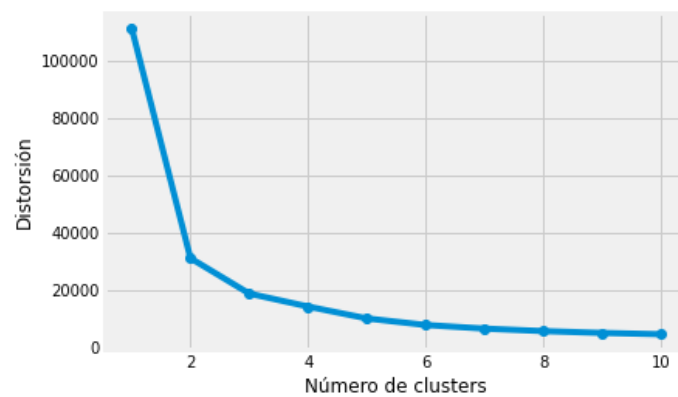


Figura 3.12 Número de clusters mediante el método del codo

Fuente: Autores

3.6.2 K means

Como resultado del método del codo, se identificaron 2 grupos claramente diferenciados, como se observa en la figura 3.13. La composición de los grupos obtenidos a partir del K-medias muestra que el primer cluster contenía en promedio los sentimientos negativos, ya que el valor es cercano a cero. No obstante, su capacidad de respuesta era mayor, dado que el valor de los likes o corazones y los retweets era superior al segundo grupo. Mientras que el segundo grupo se componía de sentimientos neutrales, debido a que su valor era cercano a 0.5 y poseía una menor capacidad de respuesta, reflejada en su bajo promedio en likes y retweets.

En otras palabras, el primer grupo representaba las “campañas efectivas”, dado que contenían los sentimientos más negativos y la mayor capacidad de respuesta por parte de la audiencia; lo cual es congruente considerando que los tweets que exponían necesidades de insumos poseían mayor atención. A diferencia del segundo grupo cuyos tweets referían a normas o resoluciones por parte de las autoridades y, por ende, tenían menor atención, ya que no se relacionaban con la necesidad de insumos. Por ello, el segundo grupo representaba las “campañas no efectivas”. En la tabla 3.2 se detallan la composición por grupo.

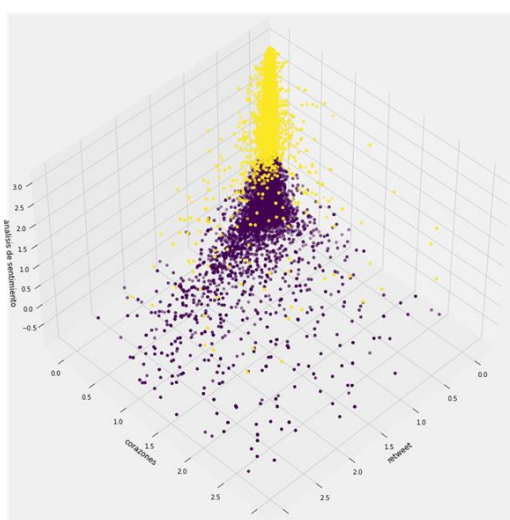


Figura 3.13 Identificación de campañas de insumo efectivas y no efectivas

Fuente: Autores

Cluster	Tipo	Retweet	Corazones o Likes	Análisis de Sentimiento
1	Efectivo	2.437176	5.254753	0.072025
2	No efectivo	1.917113	5.160046	0.485975

Tabla 3.2 Composición de los grupos identificados

Fuente: Autores

3.7 Validación

De acuerdo con la segmentación de información mediante el clustering de los datos catalogados como relevantes, y a su vez, agrupados por palabras se obtuvo una síntesis de texto o “text summarization”, la cual permitió validar la información condensada, es decir verificar que los tweets mantenían el mismo sentimiento.

Por un lado, se tiene que de los 112578 tweets recopilados la mayor parte fueron los negativos, los cuales corresponden a un valor total de 97456, mostrando así que son los más relevantes y que cuya información no se pierde, sino que más bien reflejaba emociones negativas. Mientras que los tweets positivos y neutrales no fueron tan significativos ya que del total a penas 8825 y 6297 respectivamente, corresponden a dichos sentimientos.

3.8 Indicadores de efectividad

3.8.1 Tiempos críticos

Los resultados de la serie de tiempo sobre la evolución de los tweets catalogados como negativos muestran que en el mes de marzo se registró la cifra más elevada, lo cual coincide con la declaración de emergencia sanitaria del 12 de marzo por parte del gobierno ecuatoriano. A partir de esa fecha los tweets negativos fueron decreciendo hasta llegar a julio, en donde se evidencia un punto crítico o de crecimiento; dado que en ese tiempo también se establecieron restricciones a la movilidad. Así mismo, el sentimiento negativo reflejado en los tweets muestra un crecimiento para diciembre, alcanzando un máximo relativo en enero de 2021 y decreciendo nuevamente para los siguientes

meses. Cabe mencionar que el gobierno anunció nuevamente estado de excepción junto con toque de queda a finales de diciembre con una duración de 30 días, lo cual coincide con la evolución de los tweets negativos en esos meses.

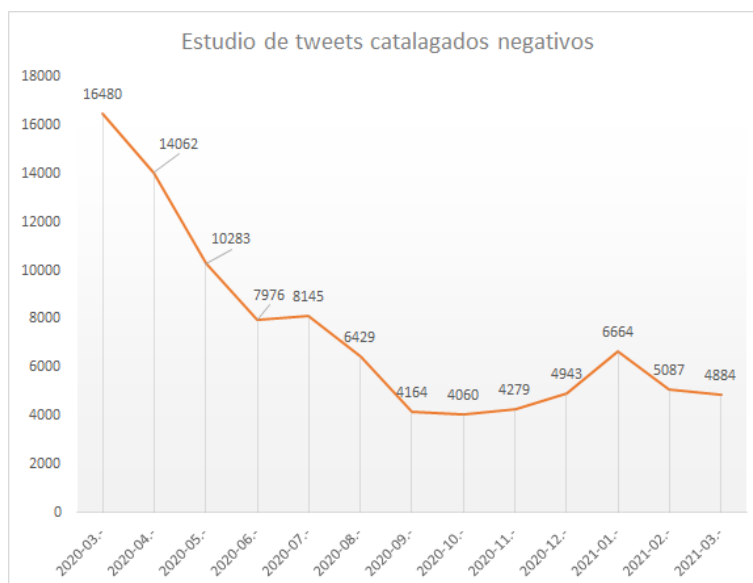


Figura 3.14 Serie en el tiempo de tweets catalogados negativos

Fuente: Autores

Para validar los resultados obtenidos, estos se compararon con los casos positivos de coronavirus en el país registrados semanalmente desde el 13 de marzo del 2020 hasta el 31 de marzo del 2021, los cuales se ilustran en la figura 3.15. Como se observa, desde la semana 11 a la 15 del 2020, se reportó un incremento significativo de los casos positivos de covid-19, correspondiendo a los meses de marzo y abril. Así mismo, de la semana 26 a la 31 del mismo año, se mantuvo una curva de crecimiento de los casos positivos de coronavirus, correspondiendo a los meses de julio y agosto.

Por último, durante las tres primeras semanas del 2021, se evidenció un crecimiento constante de coronavirus, correspondientes a enero de ese año. Por lo tanto, se confirma la consistencia de los tiempos críticos obtenidos a partir de los tweets negativos que coinciden con los casos de coronavirus reportados por el COE nacional (2021).

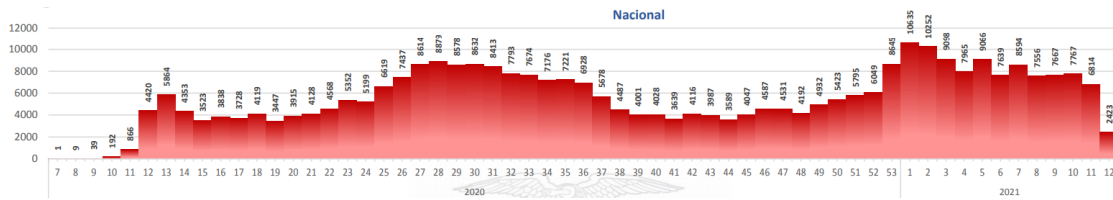


Figura 3.15 Curva epidemiológica de casos positivos de covid-19 acumulados por semana epidemiológica

Fuente: Autores

3.8.2 Zonas críticas

Los resultados de la efectividad de las campañas de insumos en el país por provincia se ilustran en la figura 3.16. Como se puede observar, las provincias con campañas no efectivas están representadas con un color morado oscuro el cual va tornándose más claro a medida que aumenta la cantidad de campañas efectivas.

Se evidencia entonces que las provincias en donde se concentraron las campañas efectivas fueron Guayas, Imbabura, Azuay y Santa Elena. Es decir, en estas zonas hubo una mayor urgencia o necesidad de insumos médicos, por lo que la ciudadanía optó por gestionar tales recursos a través de twitter a fin de conseguirlos. Dada la crisis que se reflejaba en los tweets, estos alcanzaron una mayor capacidad de respuesta y, por ende, resultaron efectivos.

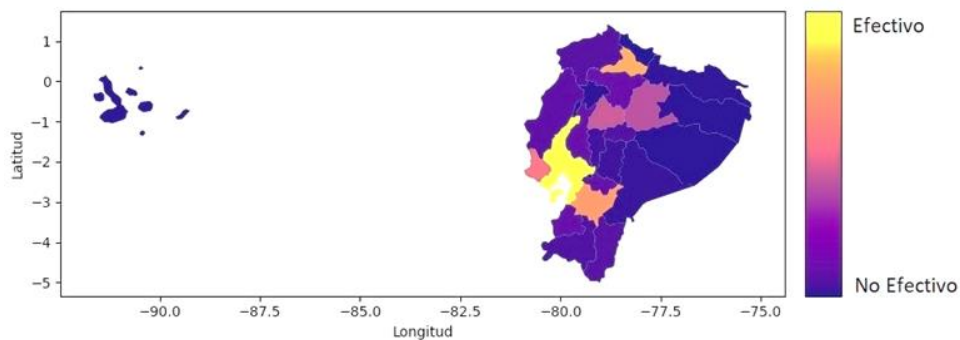


Figura 3.16 Identificación de la efectividad de las campañas de insumos en Ecuador por provincia

Fuente: Autores

Los resultados obtenidos fueron validados con la información de casos positivos de coronavirus por provincia reportados por el COE nacional (2021), los cuales se ilustran en la figura 3.17. De esto se evidencia una importante coincidencia entre los resultados de las campañas efectivas y las zonas críticas reportadas por el COE, siendo Guayas, Azuay, Imbabura y Santa Elena las provincias que reportaron más del 1.2% de casos confirmados de coronavirus acumulados por provincia referente al total nacional. Es decir, debido a que estas zonas eran las más críticas, hubo mayor movilidad por parte de las personas para realizar campañas de insumos a través de Twitter.

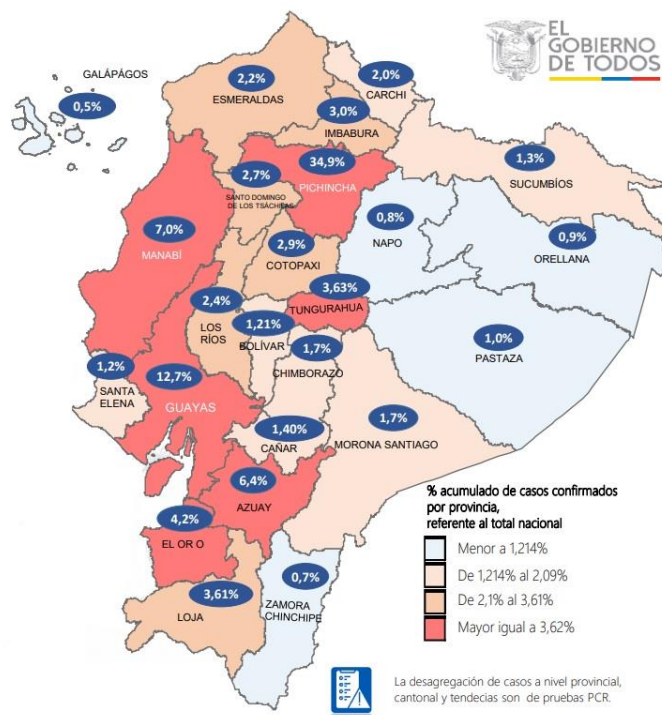


Figura 3.17 Casos confirmados de coronavirus en Ecuador por provincia

Fuente: Autores

3.8.3 Proyección de Pérdidas en el Sector Salud

De acuerdo con lo establecido en el capítulo anterior, se calculó la proporción de pérdidas económicas en el sector salud (ecuación 3.1) considerando las pérdidas estimadas en \$870 millones de dólares y la población ecuatoriana estimada en 17.2 millones de habitantes (ecuación 3.2). Para conocer la proyección de pérdidas (ecuación 3.4), la proporción obtenida

(ecuación 3.3) se multiplicó por el total de tweets equivalentes que correspondieron a 131 mil (ecuación 3.5), resultando la proyección en aproximadamente \$6.63 millones de dólares (ecuación 3.6). Esto quiere decir que, si se hubiese considerado la información de las campañas de insumos médicos en Twitter, dicha proyección en pérdida significaría un ahorro. Este ahorro pudo dirigirse a mejorar la atención en los hospitales o ampliar la infraestructura de los mimos, evitando así el prematuro fallecimiento por coronavirus.

$$(3.1) \textit{Proporción de pérdidas del sector salud} = \frac{\textit{Total de pérdidas del sector salud}}{\textit{Total de la población ecuatoriana}}$$

$$(3.2) \textit{Proporción de pérdidas del sector salud} = \frac{870000000}{17200000}$$

$$(3.3) \textit{Proporción de pérdidas del sector salud} = 50.58$$

$$(3.4) \textit{Proyección de pérdidas del sector salud} \\ = \textit{Proporción de pérdidas del sector salud} * \textit{Total de tweets equivalentes}$$

$$(3.5) \textit{Proyección de pérdidas del sector salud} = 50.58 * 131000$$

$$(3.6) \textit{Proyección de pérdidas del sector salud} = \$ 6.63 \textit{ millones (USD)}$$

3.8.4 Proyección de la demanda de Ataúdes en Guayas

Así mismo, se calculó la proyección de la demanda de ataúdes en Guayas (ecuación 3.7), considerando 17000 actas de defunción y 24349 tweets en Guayas (ecuación 3.8), obteniendo una proyección de 0.70 (ecuación 3.9). Para comparar esta proyección, se calculó también la proporción de demanda de ataúdes (ecuación 3.10), considerando el mismo número de actas de defunción y la población estimada del Guayas, la cual correspondió a 3.6 millones de habitantes (ecuación 3.11). En consecuencia, se obtuvo una proporción de 0.0047 (ecuación 3.12). Esto quiere decir que, si se hubiese considerado las campañas de insumos a través de Twitter la demanda de ataúdes se hubiera atendido en un 70%, ya que la necesidad de insumos hubiese orientado a los ofertantes de ataúdes a incrementar su producción.

$$(3.7) \textit{Proyección de demanda de ataúdes} = \frac{\textit{Total de actas de defunción emitidas}}{\textit{Total de tweets en Guayas}}$$

$$(3.8) \textit{Proyección de demanda de ataúdes} = \frac{17000}{24349}$$

$$(3.9) \textit{Proyección de demanda de ataúdes} = 0.70$$

$$(3.10) \textit{Proporción de demanda de ataúdes} = \frac{\textit{Total de actas de defunción emitidas}}{\textit{Total de la población guayasense}}$$

$$(3.11) \textit{Proporción de demanda de ataúdes} = \frac{17000}{3600000}$$

$$(3.12) \textit{Proporción de demanda de ataúdes} = 0.0047$$

CAPÍTULO 4

4. CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones

En el presente estudio se recopiló información sobre el tratamiento de la covid-19, mediante entrevistas a expertos, para identificar los insumos médicos más requeridos. A partir de ello se obtuvo una base de 15 palabras, entre las cuales se destacaron recursos como: mascarilla, nitaxozanida, paracetamol, oxígeno, entre otras.

La identificación de insumos más demandados sirvió de base para la aplicación de web scraping en Twitter, obteniendo como resultado alrededor de 131000 tweets, 556151 retweets y 1276567 likes o corazones, desde el 11 de marzo del 2020 hasta el 31 de marzo del 2021. Con esta información se creó una base de datos que facilitó la ejecución de técnicas de preprocesamiento y minería de texto.

Como resultado de la aplicación de técnicas de preprocesamiento y minería de texto se obtuvieron 112578 tweets completos y válidos, que conservaban información relevante referente a la gestión de insumos llevada a cabo por la ciudadanía a través de esta plataforma. Así mismo, se pudo identificar las emociones de los ciudadanos, siendo mayormente negativas debido a la emergencia sanitaria y a la necesidad de conseguir recursos médicos. A su vez, se evidenciaron dos grupos de campañas claramente diferenciadas y estas correspondían a las campañas efectivas y no efectivas.

En resumen, se logró evaluar la gestión de insumos médicos a través de indicadores de efectividad, a partir de los cuales se concluye que dicha gestión fue efectiva, ya que revelaron que con una gestión efectiva podría haberse ahorrado \$6.63 millones de dólares en el sector de la salud y podría haberse abastecido la demanda de ataúdes en un 70%. Además, se identificó una importante coincidencia con los informes proporcionados por el COE nacional,

siendo marzo y julio del 2020 y enero del 2021 los meses más críticos, especialmente en zonas como Guayas, Imbabura, Azuay y Santa Elena. Estos resultados representan información de interés para la formulación de políticas públicas relacionadas al abastecimiento de insumos y atención médica.

Por lo tanto, esta investigación presenta un marco metodológico de utilidad para el manejo de la información a través de plataformas modernas de comunicación, como Twitter; el cual puede ser aplicado tanto en situaciones de emergencia sanitaria como en otro tipo de circunstancias de riesgo, a corto y largo plazo.

Cabe destacar que este trabajo es el primer estudio en Ecuador enfocado en la evaluación de la gestión de insumos médicos a través de Twitter, posibilitando una visión completa de la realidad ecuatoriana en la lucha contra el coronavirus. A su vez, sus hallazgos mantienen una relación directa con la mejora del sistema de salud, contribuyendo así la salud en general, siendo esta el tercer objetivo de desarrollo sostenible para la agenda 2030.

4.2 Recomendaciones

Considerando las limitaciones de este estudio asociadas al tiempo y disponibilidad de recursos, se recomienda ampliar la muestra de entrevistados enfatizando en su diversidad geográfica. Es decir, contar con la experiencia de profesionales de la salud provenientes de distintas partes del país, podría mejorar la caracterización de las zonas a fin de poseer una perspectiva más completa de la realidad de estas durante la pandemia.

Así mismo, como trabajo futuro se espera ampliar el estudio a nivel latinoamericano a fin de comprender las similitudes o diferencias en las medidas de acción llevadas a cabo por los habitantes de cada país de esta región, frente a la afectación en la gestión de insumos médicos y la atención hospitalaria. Además de comprender las ventajas y desventajas de las decisiones

establecidas por las autoridades locales e internacionales como respuesta a la pandemia del coronavirus.

Por último, se espera definir un modelo que permita establecer los factores socioeconómicos que inciden en la efectividad de la gestión de insumo médicos; ya que se ha demostrado que incluso en provincias altamente urbanizadas como el Guayas, se evidenció una escasez de insumos. Es decir, una zona con facilidades de acceso a insumos en condiciones normales no implica una adecuada gestión de estos en momentos de crisis. Conocer esta relación a través de un modelo, podría mejorar el establecimiento de políticas de gestión.

BIBLIOGRAFÍA

- Acosta, A. (abril de 2020). El Coronavirus en los tiempos del Ecuador. *Análisis Carolina*, 1-18. doi:10.33960/AC_23.2020
- Adewole, K., Han, T., Wu, W., Song, H., & Sangaiah, A. (2018). Twitter spam account detection based on clustering and classification methods. *The Journal of Supercomputing*. doi:10.1007/s11227-018-2641-x
- Agarwal, S., & Damle, M. (2020). Sentiment Analysis to Evaluate Influencer Marketing: Exploring To Identify the Parameters of Influence. *Palarch's Journal Of Archaeology Of Egypt/Egyptology*, 17(6), 4784-4800.
- Ahmad, S., & Varma, R. (2018). Information extraction from text messages using data mining techniques. *Malaya Journal of Matematik*, S(1), 26-29. doi:10.26637/MJM0S01/05
- Ahuja, S., & Dubey, G. (2017). Clustering and Sentiment Analysis on Twitter Data. *2nd International Conference on Telecommunication and Networks (TEL-NET 2017)*. doi:10.1109/tel-net.2017.8343568
- Andramuño, C. (2020). *Acuerdo N° 00126 - 2020*. Quito: Ministerio de Salud Pública. Recuperado el 25 de junio de 2021, de https://www.salud.gob.ec/wp-content/uploads/2020/03/SRO160_2020_03_12.pdf
- Asamblea Nacional. (26 de mayo de 2021). *Asamblea Nacional*. Obtenido de <https://www.asambleanacional.gob.ec/es/leyes-aprobadas?leyes-aprobadas=All&title=&fecha=>
- Azevedo, V., Carvalho, M., Fernandes-Costa, F., Mesquita, S., Soares, J., Teixeira, F., & Maia, Â. (2017). Transcrever entrevistas: questões conceituais, orientações práticas e desafios. *Revista de Enfermagem Referência*, 4(14), 160. doi:10.12707/RIV17018
- Boyd, D. M., & Ellison, N. B. (1 de Octubre de 2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230. doi:10.1111/j.1083-6101.2007.00393.x
- Castillo Romer, J. A. (2018). El termino big data y sus definiciones. En J. A. Castillo Romer, *Big Data. IFCT128PO* (págs. 17-19). Antequera: IC Editorial.

- Cavallo, F., Rossi, N., & Chiarelli, F. (2020). Novel coronavirus infection and children. *Acta Biomedica*, 91(2), 172-176.
- Cavallo, F., Rossi, N., & Chiarelli, F. (11 de Mayo de 2020). Novel coronavirus infection and children. *Acta Biomedica*, 91(2), 172-176. doi:10.23750 / abm.v91i2.9586
- Celaya, J. (2009). La empresa en las redes sociales. En J. Celaya, *La empresa en la web 2.0* (págs. 106-108). Madrid: Grupo Planeta.
- Cennimo, D. J. (14 de Mayo de 2021). *Medscape*. Obtenido de <https://www.medscape.com/answers/2500114-197401/what-is-covid-19>
- España, S., Montes, R., Fowks, J., & Rivas, F. (junio de 2020). *Ataúdes de cartón contra la falta de recursos en América Latina*. Obtenido de El País: <https://elpais.com/sociedad/2020-06-30/ataudes-de-carton-contra-la-falta-de-recursos-en-america-latina.html>
- Estévez, N. G. (Mayo de 2013). Obtenido de https://idus.us.es/bitstream/handle/11441/15976/P_T.D._PROV26.pdf;jsessionid=9534DA560A0348127ECF35BD10616512?sequence=1
- Falcón, J. (2011). *TWITTER. Marketing Personal y Profesional*. Madrid: RC Libros.
- Ferrer-Serrano, M., Latorre-Martínez, M.-P., & Lozano-Blasco, R. (2020). Universidades y comunicación. Papel de Twitter durante el inicio de la crisis sanitaria de la Covid-19. *Profesional de la información*, 29(6), 1-18. doi:10.3145/epi.2020.nov.12
- Folgueiras, P. (2005). *La Entrevista*. Recuperado el 23 de junio de 2021, de CORE: <https://core.ac.uk/download/pdf/43554789.pdf>
- García, K., & Berton, L. (2021). Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing*, 101, 1-15. doi:doi.org/10.1016/j.asoc.2020.107057
- Gelbukh, A. (2010). *Komputer Sapiens*, 6(32). Obtenido de https://d1wqtxts1xzle7.cloudfront.net/30768432/Procesamiento_de_lenguaje_natural_y_sus_aplicaciones-with-cover-page-v2.pdf?Expires=1624632505&Signature=JyO3dchXnyNcNle2j5PV2GUy-YG5Ayry5OYdaAJCLMdlCDXVLBQDht3gvZuOvp48oTRzIKyXs8tXYihHvOdpN~9gffxZ0~UugF8ZRAs

- Gómez, L., & García, C. (2011). Twitter. *Colombian Journal of Anesthesiology*, 38(4), 539-540. Obtenido de <https://www.revcolanest.com.co/index.php/rca/article/view/809>
- Gralinski, F., Stanisawek, T., Wróblewska, A., Lipiński, D., Kaliska, A., Rosalska, P., . . . Biecek, P. (2020). *Kleister: A novel task for Information Extraction involving Long Documents with Complex Layout*.
- Gudivada, V. N. (2018). Chapter 12 - Natural Language Core Tasks and Applications. En V. N. Gudivada, V. N. Gudivada, & C.R. (Edits.), *Handbook of Statistics* (págs. 403-428). doi:<https://doi.org/10.1016/bs.host.2018.07.010>
- Hernández, M., & Gómez, J. (Julio de 2013). Aplicaciones del Procesamiento del Lenguaje Natural. *Revista Politécnica*, 32(1), 87-96. Obtenido de https://revistapolitecnica.epn.edu.ec/ojs2/index.php/revista_politecnica2/article/view/32
- Howard, J., & Ruder, S. (Julio de 2018). *Universal Language Model Fine-tuning for Text Classification*. Melbourne: Association for Computational Linguistics. Obtenido de <https://www.aclweb.org/anthology/P18-1031.pdf>
- IBM. (16 de Noviembre de 2020). *IBM*. Obtenido de <https://www.ibm.com/cloud/learn/text-mining>
- Jiang, Y., Huang, X., & Li, Z. (mayo de 2021). Spatiotemporal Patterns of Human Mobility and Its Association with Land Use Types during COVID-19 in New York City. (W. Kainz, Ed.) *International Journal of Geo-information*, 10(5), 344. doi:10.3390/ijgi10050344
- Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing* (3era Edición ed.). New Jersey: Prentice Hall. Obtenido de <https://web.stanford.edu/~jurafsky/slp3/17.pdf>
- Kaur, H., Talluri, M., & He, J. (. (2015). Get Twitter Information : A Collaborative Android Application for Big Data Analysis. *Get Twitter information: A collaborative Android applInternational Conference on Collaboration Technologies and Systems (CTS)*, 483-483. doi:10.1109/cts.2015.7210475
- Kaushik, A., & Naithani, S. (2016). A Comprehensive Study of Text Mining Approach. *IJCSNS International Journal of Computer Science and Network Security*, 16(2), 69-76. Recuperado el 17 de junio de 2021, de

https://d1wqtxts1xzle7.cloudfront.net/56626954/A_Comprehensive_Study_of_Text_Mining_Approach.pdf?1526988620=&response-content-disposition=inline%3B+filename%3DA_Comprehensive_Study_of_Text_Mining_App.pdf&Expires=1624020644&Signature=CBVcr78Hcd3XcH8NVI2~JK

- Khan, R., Shrivastava, P., Kapoor, A., Tiwari, A., & Mittal, A. (Agosto de 2020). Social media analysis with AI: sentiment analysis techniques for the analysis of twitter covid-19 data. *Journal of Critical Reviews*, 7(9), 2761-2774. doi:10.31838/jcr.07.09.437
- Kitchin, R. (2014). The Real-time City? Big Data and Smart Urbanism. *GeoJournal*, 79, 1-14. doi:10.1007/s10708-013-9516-8
- La Hora, E. d. (21 de junio de 2021). *Más vacunas llegaron al Ecuador*. Recuperado el 23 de junio de 2021, de La Hora: <https://www.lahora.com.ec/mas-vacunas-llegaron-al-ecuador/>
- Latorre-Martínez, P., Orive-Serrano, V., & Íñiguez-Dieste, D. (2018). Medición y análisis de la audiencia social de las televisiones autonómicas en Facebook y Twitter. *El profesional de la información*, 27(5), 1061-1070. doi:10.3145/epi.2018.sep.10
- Loper, E., & Bird, S. (2020). *University of Pennsylvania*. Obtenido de University of Pennsylvania: <https://arxiv.org/pdf/cs/0205028.pdf>
- Marín Díaz, J. P. (22 de Julio de 2017). *Github*. Obtenido de https://github.com/jpmarindiaz/geo-collection/blob/master/ecu/ecuador.geojson?short_path=9b67de6
- Martínez, R., Rodríguez, R., Vera, P., & Parkinson, C. (2019). Análisis de técnicas de raspado de datos en la web aplicado al Portal del Estado Nacional Argentino. *XXV Congreso Argentino de Ciencias de la Computación (CACIC)*, (págs. 457-466). Río Cuarto. Recuperado el 17 de junio de 2021, de <http://sedici.unlp.edu.ar/handle/10915/91026>
- Maylawati, D., Kumar, Y., Kasmin, F., & Ramdhani, M. (2019). An idea based on sequential pattern mining and deep learning for text. *Journal of Physics: Conference Series*, 1402(7), 1-10. doi:<https://doi.org/10.1088/1742-6596/1402/7/077013>

- Mohdab, M., Jan, R., & Shah, M. (1 de Abril de 2020). Text document summarization using word embedding. *Expert Systems with Applications*, 143, 1-10. doi:<https://doi.org/10.1016/j.eswa.2019.112958>
- Moncada, D., Hernández, C., Gil, V., Cabrales, J., Cure, O., Hanna, F., . . . Rendón, I. (2020). Recomendaciones para prevención y manejo de los casos de coronavirus en el servicio de Cardiología Intervencionista y hemodinamia. *Revista Colombiana de Cardiología*, 1-5. Recuperado el 20 de junio de 2021, de <http://revcolcard.org/wp-content/uploads/2020/04/recomendaciones-Colegio-de-hemodinamia-1.pdf>
- Nadeem, M., Zamzami, M., Choudhry, H., Murtaza, B., Kazmi, I., Ahmad, H., & Shakoori, A. (2020). Origin, Potential Therapeutic Targets and Treatment for Coronavirus Disease (COVID-19). *PathogenS*, 9(4), 307. doi:10.3390/pathogens9040307
- Nicks, B. (21 de febrero de 2020). *Medscape*. (O. Wong, Editor) Recuperado el junio de 2021, de <https://espanol.medscape.com/diapositivas/59000118>
- OMS, O. M. (11 de marzo de 2020). *Organización Panamericana de la Salud*. Recuperado el junio de 2021, de <https://www.paho.org/es/noticias/11-3-2020-oms-caracteriza-covid-19-como-pandemia>
- ONU, O. d. (2020). *Plan de Respuesta Humanitaria COVID-19 Ecuador*. Equipo Humanitario de País. Organización de las Naciones Unidas. Recuperado el 20 de junio de 2021, de <https://reliefweb.int/sites/reliefweb.int/files/resources/20200430-EHP-ECUADOR-COVID-19.pdf>
- Oracle. (Octubre de 2011). *Oracle:Big Data for the Enterprise*. Oracle. Obtenido de <https://www.oracle.com/technetwork/database/bi-datawarehousing/wp-big-data-with-oracle-521209.pdf>
- Organización Internacional del Trabajo, O. (2021). *El sistema de salud ecuatoriano y la COVID-19*. Organización Internacional del Trabajo. Obtenido de https://www.ilo.org/wcmsp5/groups/public/---americas/---ro-lima/---sro-lima/documents/publication/wcms_799790.pdf
- Patel, A., & Thakral, P. (2016). The Best Clustering Algorithms in Data Mining. *International Conference on Communication and Signal Processing*, (págs. 2042-2046). doi:10.1109/icccsp.2016.7754534

- Pérez, N. (11 de marzo de 2020). *El Universo*. Recuperado el 20 de junio de 2021, de <https://www.eluniverso.com/noticias/2020/03/11/nota/7776414/ecuador-tercero-latinoamerica-casos-positivos-coronavirus/>
- República del Ecuador. (2014). *REGLAMENTO DE BUENAS PRÁCTICAS DE ALMACENAMIENTO, DISTRIBUCIÓN Y TRANSPORTE*. Quito: Decreto Ejecutivo 1272.
- República, M. d. (2018). *Normativa de buenas practicas para laboratorios farmaceuticos*. Quito: Resolución 8. Obtenido de <https://www.controlsanitario.gob.ec/wp-content/uploads/downloads/2016/06/Acuerdo-Ministerial-676.pdf>
- Rosen, A. (7 de Noviembre de 2017). *Blog Twitter*. Obtenido de https://blog.twitter.com/es_la/topics/product/2017/Hoy-es-mas-facil-Twittear
- Sarawagi, S. (2008). *Information extraction. Foundations and Trends in Databases*. Boston: Now Publishers.
- Saurkar, A., Pathare, K., & Gode, S. (2018). An Overview On Web Scraping Techniques And Tools. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4(4), 363-367. Recuperado el 17 de junio de 2021, de <http://www.ijfrcsce.org/index.php/ijfrcsce/article/view/1529/1529>
- Servicio Nacional de Gestión de Riesgos y Emergencias, S. (2021). *Informe de Situación (SITREP) e Infografía*. Servicio Nacional de Gestión de Riesgos y Emergencias (SNGRE), Quito. Obtenido de <https://www.gestionderiesgos.gob.ec/wp-content/uploads/2021/03/INFOGRAFIA-NACIONALCOVID19-COE-NACIONAL-08h00-3132021.pdf>
- Sierra, J. (20 de Diciembre de 2014). *SIAN-Lab*. Obtenido de <http://www.scian.cl/archivos/uploads/1419931797.0517>
- Sistema Nacional Descentralizado de Gestión de Riesgos, S. (2021). *Informes de Situación del Coronavirus en Ecuador*. Servicio Nacional de Gestión de Riesgos y Emergencias (SNGRE), Quito. Obtenido de <https://www.gestionderiesgos.gob.ec/wp->


content/uploads/2021/01/INFOGRAFIA-NACIONALCOVID19-COE-
NACIONAL-Y-ANEXO-08h00-31012021.pdf

- Talib, R., Kashif, M., Ayesha, S., & Fatima, F. (2016). Text Mining: Techniques, Applications and Issues. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 7(11), 414-418. Recuperado el 17 de junio de 2021, de https://www.researchgate.net/publication/311394659_Text_Mining_Techniques_Applications_and_Issues/fulltext/5844091808ae8e63e625730a/Text-Mining-Techniques-Applications-and-Issues.pdf
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society Series B*, 63(2), 411-423. doi:10.1111/1467-9868.00293
- Welbers, K., Atteveldt, W. V., & Benoit, K. (2017). Text Analysis in R. *Communication Methods and Measures*, 11(4), 245-265. doi:10.1080/19312458.2017.1387238
- Zeroual, I., & Lakhouaja, A. (2018). Data science in light of natural language. *Procedia Computer Science*, 127, 82-91. doi:<https://doi.org/10.1016/j.procs.2018.01.101>

APÉNDICES

APÉNDICE A

FORMATO DE ENTREVISTA

 <p>Escuela Superior Politécnica del Litoral</p>
<p>FACULTAD DE CIENCIAS SOCIALES Y HUMANÍSTICAS ENTREVISTA PARA EL PROYECTO INTEGRADOR DE LA CARRERA DE LICENCIATURA EN AUDITORÍA Y CONTROL DE GESTIÓN</p>
<p>Entrevistado: Ocupación: Entrevistadora: Ester Rebeca Melo Vargas.</p>
<p>Experiencias durante la pandemia</p> <ol style="list-style-type: none">1. ¿Podría describir su experiencia en el ámbito profesional a inicios de la pandemia en Ecuador?2. Según su experiencia, ¿cuál fue momento más crítico de la pandemia en Ecuador? ¿por qué?3. Desde su perspectiva, ¿cree que se pudo responder de forma oportuna o eficaz a dicho momento crítico?4. ¿Qué medidas cree que debieron seguirse para afrontar la crisis en el sector hospitalario?5. ¿Considera que la situación del sector hospitalario ha mejorado?
<p>Insumos médicos</p> <ol style="list-style-type: none">1. ¿Podría indicar qué tipo de insumos médicos se demandaron a inicios de la pandemia?2. ¿Dichos insumos se mantuvieron constantes durante la pandemia u otros fueron demandados mientras la pandemia progresaba?3. Actualmente, ¿cuáles son los insumos que se emplean con mayor frecuencia para mitigar o tratar la covid-19?

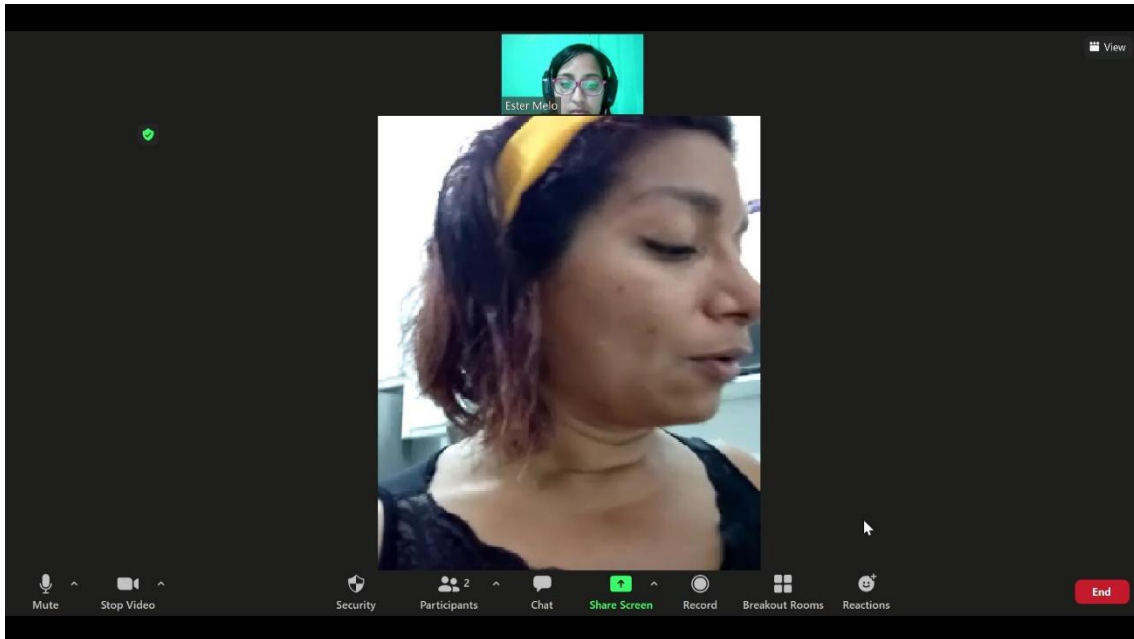
4. ¿Conoce de antibióticos, antivirales, fármacos u otros insumos médicos que hayan sido suministrados por las personas para mitigar o tratar la covid-19, aunque estos no tenían relación con la enfermedad?
5. ¿Conoce de algún insumo o método basado en la medicina natural que también haya sido suministrado por las personas?
6. ¿Conoce sobre la demanda de vacunas para la covid-19 en Ecuador?
7. ¿Cuáles son las vacunas más demandadas por la población?

Perspectiva sobre la población

1. ¿Qué opina respecto a que las personas se automediquen para mitigar o tratar la covid-19?
2. ¿Cree que la gente está dispuesta a vacunarse? ¿Por qué?
3. Si pudiera enviar un mensaje a la población ecuatoriana, ¿qué le diría?

APÉNDICE B

EVIDENCIA DE ENTREVISTA A LA DRA. LORENA ESPINOZA



APÉNDICE C

EVIDENCIA DE ENTREVISTA AL DR. MIGUEL ICAZA

