

Análisis de sentimiento en Instagram: Poder Ejecutivo de Honduras.

Adela Meza
Facultad de Ciencias, Departamento de Matemáticas
Universidad Nacional Autónoma de Honduras
e-mail: adela.meza@unah.hn

ÍNDICE

I. Introducción	1
II. Definición del Problema	1
III. Preliminares y Notación	1
III-A. Antecedentes	1
III-B. Impacto de las redes sociales en la política	2
III-C. Análisis de sentimiento	2
III-D. Análisis de contenido	2
III-E. Materiales	2
IV. RESULTADOS PRELIMINARES	2
V. RESULTADOS	3
V-A. Muestra	3
V-B. Recolección de datos.	4
V-C. Resultados del análisis de sentimiento.	4
V-D. Resultados del análisis de contenido.	5
VI. ANÁLISIS, INTERPRETACIÓN Y CONCLUSIONES	5
VII. LIMITACIONES Y RECOMENDACIONES	6
VIII. TRABAJO FUTURO	6
IX. AGRADECIMIENTOS	6
X. DISPONIBILIDAD DE DATOS	6
Referencias	6

ÍNDICE DE FIGURAS

1. Algoritmo solución del problema ejemplo.	3
2. Subjetividad y polaridad de la muestra.	5
3. Nube de adjetivos frecuentes en cuentas con bajo engagement (arriba) y engagement promedio (abajo).	5

ÍNDICE DE CUADROS

I. Código para realizar análisis de sentimiento.	3
II. Muestra de la población.	3
III. Muestra de la población. (Porcentaje bajo de Engagement)	4
IV. Polaridad en cuentas con engagement promedio.	4
V. Polaridad en cuentas con engagement bajo.	4
VI. Subjetividad en cuentas con engagement promedio.	5
VII. Subjetividad en cuentas con engagement bajo.	5
VIII. Adjetivos en cuentas con bajo engagement	5
IX. Adjetivos en cuentas con engagement promedio	5

Análisis de sentimiento en Instagram: Poder Ejecutivo de Honduras.

Resumen—Para la política mediática, es de suma importancia contar con documentos con el suficiente rigor matemático que puedan respaldar conclusiones respecto al nivel de aceptación que tiene un gobierno. En este escenario, analizamos la polaridad y subjetividad de 900 entradas de texto de cuentas de noticieros/periódicos hondureños en Instagram mediante procesamiento del lenguaje natural con machine learning y análisis de contenido. Los resultados revelaron una tendencia a la neutralidad y objetividad en los comentarios de cuatro cuentas con engagement promedio y cuatro con bajo engagement, con el reiterado empleo de los adjetivos bueno, excelente, mejor, gran, etc. En suma, las cuentas expresan una tendencia a ser comentarios informativos e interrogativos. Resultados que se pueden precisar en investigaciones futuras mediante la expansión del campo y objeto de estudio, automatización en la corrección de errores ortográficos/gramaticales y la elaboración de criterios de referencia para interpretar los resultados.

Palabras clave: Procesamiento del lenguaje natural (PLN), análisis/minería de textos, análisis de contenido, polaridad, subjetividad.

I. INTRODUCCIÓN

En Honduras, actualmente, son pocos los estudios o reportes rigurosos en los que se miden preferencias o estudian comentarios, opiniones y emociones frente a las acciones a nivel de gobierno (en este primer semestre del año 2022). Una de las herramientas utilizadas frecuentemente para medir el nivel de popularidad o aceptación que tiene un candidato o las acciones que el gobierno ha desempeñado, se basa solamente en encuestas emitidas por noticieros populares u organismos dedicados a la recopilación de este tipo de datos.

Para las ciencias sociales, y la política mediática, es de suma importancia contar con documentos con el suficiente rigor matemático para respaldar conclusiones respecto al nivel de apoyo o aceptación que tienen aquellos que organizan, gestionan y poseen las instituciones del Estado; O que ofrezcan un panorama general de la opinión pública en relación a este tema.

La segunda red social elegida por los hondureños, aunque notablemente menos popular que Facebook, es Instagram, la cual cuenta con 940,000 usuarios en todo el país [1]. Por lo que, a pesar de ser una plataforma por medio de la cual se intercambia información principalmente de forma visual, es conveniente determinar el sentimiento general que los usuarios tienen respecto a ésta temática.

II. DEFINICIÓN DEL PROBLEMA

Utilizando una aplicación de la minería de datos, conocida como análisis de sentimiento, se analizará la polaridad y subjetividad de comentarios realizados en publicaciones (de

Instagram) relacionadas a las acciones que el poder ejecutivo ha desempeñado en sus primeros 6 meses de gobierno. Posteriormente, se realizará un análisis de contenido determinando los adjetivos más frecuentes en las entradas de texto analizadas. Esto, con el objetivo de obtener un reporte que incluya un análisis detallado, con elementos matemáticos, en relación a la opinión pública respecto a este tema. Es decir, el enunciado del problema propuesto es una predicción de sentimientos relacionados con el poder ejecutivo de Honduras con la ayuda del aprendizaje automático o machine learning.

III. PRELIMINARES Y NOTACIÓN

III-A. Antecedentes

El objetivo general de ésta investigación es amplificar los estudios relativos al análisis de sentimiento semántico en Instagram, en particular, determinando la actitud de los usuarios en esta red social con respecto a los primeros 6 meses de gobierno del Poder Ejecutivo en Honduras. Por ésta razón, en primer lugar, se examinará brevemente el trabajo previo sobre la aplicación que ha tenido el análisis de sentimientos en redes sociales.

Dentro de las investigaciones más recientes, tenemos la tesis de pregrado [2], realizada en el año 2018, cuyo producto final es una aplicación web que interactúa con twitter y puede ser utilizada para analizar la respuesta en las redes sociales (mediante gráficos estadísticos) a un producto o servicio de determinada empresa o entidad. El siguiente año, el Ingeniero Tomás Garcés, realizó el proyecto de título [3], con diferentes contribuciones al campo de AS. Dentro de ellas encontramos: la creación de un software que permite la limpieza, estandarización, y tratamiento del texto en idioma español. Así como el desarrollo de un software de machine learning, escalable, que permita extrapolar su aplicabilidad a distintas temáticas (política, TV, juegos, etc) y con un mayor trabajo, a distintos idiomas.

Estas dos investigaciones utilizan como plataforma principal, Twitter; dado que las publicaciones contienen, en su mayoría, texto y permite acceder libremente a gran parte de la información publicada. Tratándose de Instagram, en el año 2020, se publica la investigación [4], que pretende estudiar y profundizar en la expresión del yo infantil a través de la escritura. Identificando el grado de polaridad y subjetividad expresado a nivel léxico en cuentas infantiles, haciendo uso del machine learning, en concreto, el análisis de sentimiento.

Con respecto al tema de aplicar el AS para determinar la aprobación que tiene el poder ejecutivo Hondureño en sus

primeros meses de gobierno, no hay una investigación directa. Pero si se cuenta con herramientas que lo emplean y se podrían utilizar para este propósito, como ser el caso de las plataformas: Hootsuite Insights, RapidMiner, Social mention, Brand24. Que permiten minar texto para revelar percepciones de influencers, historias, tendencias y sentimientos.

III-B. Impacto de las redes sociales en la política

En numerosas investigaciones relacionadas a la influencia de las redes sociales en la sociedad, se concluye que estas se han convertido en un espacio de transformación, divulgación y expresión social. En [5], Blasco Fontecilla, un reconocido médico psiquiatra, plantea que las redes sociales tienen dos funciones primordiales: 1) Facilitar la comunicación y el intercambio de información; y 2) Facilitar la creación de comunidades.

De forma general, estas se pueden ver como un servicio del mundo informático, que ofrece a los usuarios una plataforma de comunicación a través de Internet en la que pueden configurar un perfil y de esta manera facilitar la interacción con otros usuarios en la red.

En particular, no se puede negar su protagonismo en el discurso político digital. Basta imaginar un contexto en el que no se utilicen las redes sociales como medio de transmisión de la información política. Aún se contaría con la TV, radio y periódicos, sin embargo, es claro que no se tendría el mismo alcance de personas; dado que muchas utilizan las redes sociales como medio principal para obtener información.

En una investigación sobre el discurso político en las redes sociales, se menciona que la importancia de éstas radica en que resulta imposible plantear cualquier estrategia de comunicación política sin planificar acciones en el soporte digital. Y, por el contrario, difícilmente un mensaje, debate o campaña política puede instalarse en la opinión pública [9]. Dada esta naturaleza inherente de ser un medio de difusión masiva, se puede apreciar el impacto que tienen en la realidad socio política prevaleciente en el país.

III-C. Análisis de sentimiento

El análisis de sentimientos o minería de opiniones emplea el procesamiento del lenguaje natural para extraer un patrón significativo de conocimiento de una gran cantidad de datos textuales. Examina comentarios, opiniones, emociones, creencias, puntos de vista, preguntas, preferencias, actitudes y solicitudes comunicadas por el escritor en una cadena de texto. [6]

Esta tarea puede realizarse examinando la polaridad, subjetividad y las emociones presentes en un texto. Entendiéndose por polaridad el grado de positividad o negatividad del texto, es decir, son positivos si presentan un estado de satisfacción, dicha o felicidad; negativos si presentan un estado de abatimiento, desilusión o tristeza; o neutrales si presentan un estado que no es ni negativo ni positivo. Y, por subjetividad, se entiende si el texto representa

una opinión o un hecho.

Como expresa el autor Yong Shi en [8], el análisis de sentimiento se puede aplicar ampliamente a una gran cantidad de escenarios de aplicación, como mejorar el servicio al cliente y analizar las redes sociales. Dado que en una sola oración, y utilizando los parámetros involucrados en el análisis de sentimiento, se puede extraer la opinión sobre algún producto, idea, marca, entre otras.

III-D. Análisis de contenido

Como se discutía en la sección anterior, el análisis de sentimiento también involucra extraer del texto analizado un resumen de opiniones o sentimientos y emociones del autor. Justamente este es uno de los objetivos del análisis de contenido, una técnica de investigación social para analizar un mensaje.

Existen diversos tipos y métodos para analizar un texto, pero nos centraremos en el *análisis de contenido temático*, que recurre a la lógica para resumir, definir categorías y verificar la validez de los argumentos y de las conclusiones. Llevan al analista a actuar principalmente como verificador, clasificador y eventualmente estadístico [10]. La técnica más simple de análisis temático consiste en crear una lista con la frecuencia de aparición de todas las palabras de un texto. Técnica que se ajusta para ser implementada computacionalmente.

III-E. Materiales

Para realizar el análisis de sentimiento en el texto extraído de las publicaciones en Instagram, se utilizará una de las librerías de Python más famosas y fáciles de utilizar, TextBlob. Esta es una biblioteca para procesar datos textuales que proporciona una API simple para sumergirse en tareas comunes de procesamiento de lenguaje natural (PLN), como el etiquetado de partes del discurso, la extracción de frases nominales, el análisis de sentimientos, la clasificación, la traducción, entre otras. [11]

Cada entrada de texto produce un resultado de polaridad (muy negativo, neutro o muy positivo) y de subjetividad (muy objetivo o muy subjetivo). Es decir, para cada comentario, con la librería TextBlob, podemos obtener resultados en el intervalo $[-1, 1]$ para medir la polaridad, donde -1 significa muy negativa, 0 significa neutral y +1 significa muy positiva. Y resultados en el intervalo de $[0, 1]$ para medir la subjetividad, donde 0 significa muy objetivo y 1 significa muy subjetivo.

Para el análisis de contenido se utilizará el lector inmersivo de Outlook en la identificación de adjetivos, y el software gratuito en línea WordCounter para la localización de los ítems más frecuentes.

IV. RESULTADOS PRELIMINARES

Consideremos el siguiente problema: *Realizar un análisis de sentimiento a un comentario extraído de una publicación*

de instagram relacionada al Poder ejecutivo de Honduras, utilizando la librería de Python: TextBlob.

TextBlob utiliza el algoritmo Naive Bayes de aprendizaje supervisado que se basa en el teorema de Bayes. Toma cada clase (texto) como variable independiente y predice la probabilidad de que pertenezca a un grupo, sea este positivo, negativo o neutro.

[12]

El enfoque que aplica este paquete al análisis de sentimiento se basa en reglas y, por lo tanto, requiere un conjunto predefinido de palabras categorizadas, es decir, Textblob tiene almacenado un conjunto de palabras que reflejan o describen emociones en un texto. Estas palabras pueden, por ejemplo, cargarse desde la base de datos NLTK (Natural Language Toolkit), que calcula el sentimiento a partir de reglas creadas con anterioridad, como ser, contar la cantidad de apariciones de una determinada palabra o analizar la estructura de la oración. Además, los sentimientos se definen en función de las relaciones semánticas y la frecuencia de cada palabra en una oración de entrada que permite obtener como resultado una salida más precisa. [13]

Así, dada una oración, cada palabra se califica según subjetividad, polaridad o si es un modificador (también conocidos como intensificadores), que intensifican el significado del texto en función de su patrón, resultando en una polaridad combinada para textos más grandes. TextBlob ignora la polaridad y la subjetividad cuando se incluye una palabra modificadora, en lugar de confiar únicamente en la intensidad para calcular el sentimiento del texto.

Teniendo esto en cuenta, planteamos el siguiente algoritmo de solución para el problema enunciado anteriormente:

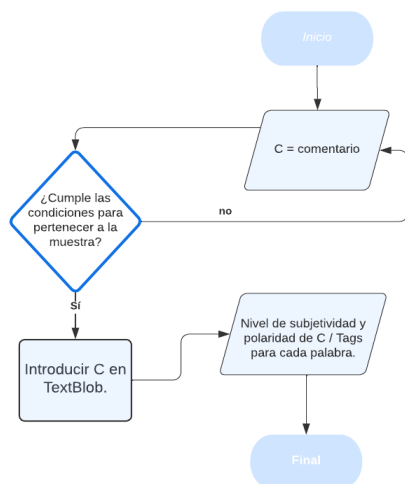


Figura 1. Algoritmo solución del problema ejemplo.

Siguiendo el diagrama solución, el primer paso es considerar una línea de texto que se encuentre dentro de la

Código en TextBlob

```

c = TextBlob("Está excelente, ella fue una heroína,
lo que duele fue que en vida nunca se le hizo un reconocimiento
por parte de nosotros como Hondureños y Hondureñas, y está bien,
pero, también se necesita que trabajen en asuntos de verdadera
necesidad, de urgente necesidad. ")
comentario = c.translate(from_lang="es",to="en")
print(comentario," \n")
print(comentario.sentiment)
print(comentario.tags)

```

Cuadro I
CÓDIGO PARA REALIZAR ANÁLISIS DE SENTIMIENTO.

Canal	Engagement	Seguidores	Imágenes observadas	Entradas de texto extraídas
1	0.10 %	57,570	15	200
2	0.05 %	115,420	12	200
3	0.02 %	187,300	12	200
4	0.78 %	277,630	12	200

Cuadro II
MUESTRA DE LA POBLACIÓN.

muestra, es decir, un comentario extraído de una publicación de instagram relacionada al poder ejecutivo en Honduras. Una vez se tiene el texto, se procede a realizar el AS en TextBlob, para ello, consideramos el código de la tabla anterior. Obtenemos como resultado los siguientes niveles de polaridad y subjetividad: Sentiment(polarity=0.44166666666666665, subjectivity=0.7875).

Si bien el comentario seleccionado tiene ciertos errores gramaticales y ortográficos, estos no alteran el significado del texto en función de su patrón. Dado que al aplicar las respectivas correcciones a la línea de texto, el resultado del AS es el mismo. Sin embargo, es importante mencionar que al realizar esta misma prueba en otros textos que tenían errores ortográficos/gramaticales más notorios, se presentaba una diferencia (poco significativa) en el valor de la polaridad y la subjetividad, más no en el signo de dicho valor. En [14] podemos encontrar el código y texto utilizados.

V. RESULTADOS

V-A. Muestra

Dadas las capacidades de análisis del texto en español en TextBlob, el lector inmersivo de Outlook y WordCounter, se seleccionó una muestra de cuatro cuentas de noticieros/periódicos hondureños (800 entradas de texto) de Instagram. Estas se delimitaron en base al número de seguidores y utilizando la plataforma analítica HypeAuditor, que cuenta con una calculadora de *Engagement*, mediante la cual se puede medir el nivel de vínculo entre la cuenta y su comunidad. A través de ese vínculo, los seguidores se animan a interactuar con la cuenta, comentando una publicación, compartiendo un contenido o dando un me gusta en una publicación. Estas cuentas presentarían una media de engagement del 0,24 %, con un rango de seguidores de entre 57,000 a 280,000 usuarios.

Canal	Engagement	Seguidores	Imágenes observadas	Entradas de texto extraídas
5	0.05 %	15,830	11	25
6	0.12 %	18,570	07	25
7	0.08 %	40,470	13	25
8	0.05 %	55,450	10	25

Cuadro III
MUESTRA DE LA POBLACIÓN. (PORCENTAJE BAJO DE ENGAGEMENT)

V-B. Recolección de datos.

Las características de la muestra se acogen a los siguientes aspectos: 1) texto de entrada en español; 2) texto en publicaciones relacionadas al poder ejecutivo; 3) texto en publicaciones realizadas por noticieros/periódicos hondureños y 4) texto dirigido al actual gobierno hondureño o a las acciones desempeñadas por el mismo. Las unidades de análisis se presentan en el cuadro II, donde se anonimizan los datos personales e identificativos de las cuentas.

Para fines investigativos se consideraron 4 cuentas más, de las que se extrajo una menor cantidad de entradas de texto. Esto debido a que la media de engagement que presentaron fue de 0,07 %, con un rango de seguidores de entre 15,000 a 56,000 usuarios. Razón por la cual las entradas de texto por publicación eran limitadas. Estas unidades de análisis se presentan en el cuadro III.

Para evitar el sesgo político se tomaron ciertas medidas en cuenta:

- Selección de cuentas: se realizó un listado de todos los noticieros y periódicos locales en vigencia, posteriormente se seleccionaron las cuentas con mayor número de seguidores y mayor porcentaje de engagement, esto para garantizar la cantidad necesaria de entradas de texto.
- Selección de publicaciones: para cada mes (desde enero, hasta junio del 2022), se realizó un listado con las publicaciones disponibles relacionadas al poder ejecutivo y que tuvieran mayor participación por parte de los seguidores. Luego, se seleccionaron de forma aleatoria 2 de ellas (en el caso del canal 1, para tres de los meses se escogieron 3 publicaciones, esto por la disponibilidad de datos).
- Selección de comentarios: Se seleccionaron los primeros comentarios de cada publicación en base a la cantidad necesaria para completar la muestra y el formato aceptado por TextBlob.

En el caso de las cuentas con bajo engagement no se recopilan los datos de un número específico de publicaciones por la cantidad de texto disponible.

V-C. Resultados del análisis de sentimiento.

Los resultados obtenidos a raíz del registro y procesamiento de las 900 entradas de texto en TextBlob se encuentran en los cuadros IV y V. Se analizó cada entrada individualmente y en correlación al grupo que pertenecen en materia de engagement y polaridad.

Canal	Número de Entradas positivas	Número de Entradas negativas	Media de polaridad positiva	Media de polaridad negativa	Media de polaridad de la cuenta
1	75	29	0.3784	-0.1971	0.1133
2	78	36	0.3529	-0.1921	0.1031
3	62	34	0.3715	-0.2884	0.661
4	105	31	0.4717	-0.2992	0.2013
Total	320	130	0.3937	-0.2442	0.1209

Cuadro IV
POLARIDAD EN CUENTAS CON ENGAGEMENT PROMEDIO.

Canal	Número de Entradas positivas	Número de Entradas negativas	Media de polaridad positiva	Media de polaridad negativa	Media de polaridad de la cuenta
5	10	3	0.3293	-0.4000	0.6916
6	13	3	0.3897	-0.3964	0.1551
7	8	4	0.3295	-0.1688	0.7855
8	10	2	0.3989	-0.1125	0.1506
Total	41	12	0.3619	-0.2694	0.1133

Cuadro V
POLARIDAD EN CUENTAS CON ENGAGEMENT BAJO.

Del cómputo de 800 entradas totales referentes a la muestra con un nivel de engagement promedio, se descubre un porcentaje de 43.74 % de entradas neutras, 40 % positivas y 16.25 % negativas (en las tablas solamente se muestran las entradas positivas y negativas por motivos estéticos, sin embargo se calculan fácilmente sabiendo que en total deben ser 200 entradas). La polaridad derivada de las entradas positivas presenta un valor de 39.37 % y las negativas de -24.42 %. En suma, en esta primera escala, se descubre una polaridad media del 12.09 %.

Continuando con la muestra de las cuentas con bajo porcentaje de engagement (100 entradas), tenemos un porcentaje de 47 % entradas neutras, 41 % positivas y 12 % negativas. La polaridad positiva presenta valor medio de 36.19 % y la negativa de -26.94 %. En total, la polaridad de esta segunda muestra es de 11.33 %. Por lo que las cuentas de noticieros y periódicos seleccionados presentan mayor número de entradas positivas que negativas.

En cuanto a la subjetividad, tenemos que para las cuentas con un engagement promedio, la media de subjetividad del total de comentarios positivos es de 57.78 % y de 53.91 % para los negativos. Para la siguiente muestra, tenemos los valores de 49.14 % y 55.54 %, respectivamente. Se percibe una afinidad entre las cuentas con engagement promedio y bajo al resultar un 0.36 y un 0.31 respectivamente en la escala de análisis (0 = muy objetivo; 1 = muy subjetivo). En este orden y, considerando la media de ambas (0.5), se presenta una inclinación a ser comentarios objetivos.

Adicionalmente y, tras comparar en la Figura 2 los resultados relativos a la subjetividad y polaridad, se observa que las entradas subjetivas no son proporcionales al grado de positividad del texto.

Canal	Media de subjetividad en comentarios positivos	Media de subjetividad en comentarios negativos	Media de subjetividad en comentarios de la cuenta
1	0.5596	0.4906	0.3107
2	0.5702	0.5436	0.3618
3	0.5579	0.5788	0.3296
4	0.6235	0.5434	0.4363
Total	0.5778	0.5391	0.3596

Cuadro VI
SUBJETIVIDAD EN CUENTAS CON ENGAGEMENT PROMEDIO.

Canal	Media de subjetividad en comentarios positivos	Media de subjetividad en comentarios negativos	Media de subjetividad en comentarios de la cuenta
5	0.4047	0.7667	0.2814
6	0.6060	0.7988	0.4858
7	0.4531	0.4312	0.2140
8	0.5018	0.2250	0.2447
Total	0.4914	0.5554	0.3065

Cuadro VII
SUBJETIVIDAD EN CUENTAS CON ENGAGEMENT BAJO.

V-D. Resultados del análisis de contenido.

De las dos muestras en conjunto, se recopilan 920 adjetivos (127 de la muestra con bajo engagement y 793 de la muestra con engagement promedio). En la figura 3, se muestra una *nube de palabras* para cada una de las muestras; las palabras con más frecuencia son más grandes, y con menos son más pequeñas.

En la muestra con bajo engagement, se localizaron en orden de densidad: 5 palabras del término *vieja*, 3 de *excelente*, 2 de los términos *preocupados*, *increíble*, *igual*, *buen* y así sucesivamente, hasta reflejar adjetivos únicos (1% o menos). En la muestra con engagement promedio, se compendiaron 68 palabras con el término *bueno*, 53 con *excelente*, 26 con *mejor*, 22 con *gran*, 15 con *misma*, 12 con *peor*, entre otras. En ambos casos, se refleja la reiteración de la palabra «excelente» como uno de los adjetivos más frecuentes en los comentarios de las publicaciones

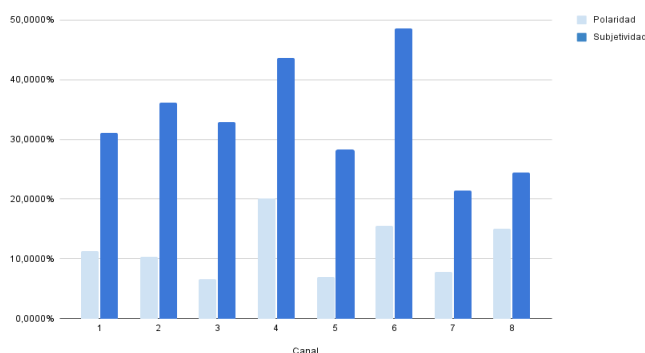


Figura 2. Subjetividad y polaridad de la muestra.



Figura 3. Nube de adjetivos frecuentes en cuentas con bajo engagement (arriba) y engagement promedio (abajo).

Adjetivos	Frecuencia	Porcentaje
Vieja	5	4 %
Excelente	3	2 %
Preocupados	2	2 %
Loca	2	2 %
Increíble	2	2 %
Igual	2	2 %

Cuadro VIII
ADJETIVOS EN CUENTAS CON BAJO ENGAGEMENT

observadas. En los cuadros VIII y IX se muestran algunos de los adjetivos, su frecuencia y el porcentaje en relación al total.

VI. ANÁLISIS, INTERPRETACIÓN Y CONCLUSIONES

Los resultados arrojados tanto por el análisis de sentimiento como por la exploración de adjetivos en las entradas textuales de las publicaciones, presentan un grado prominente de neutralidad (polaridad) en un entorno que se inclina a ser objetivo. En la correlación entre polaridad y subjetividad, no se observan niveles que indiquen que una respuesta positiva y/o negativa vaya proporcionalmente vinculada al modo en que se expresan los usuarios.

Adjetivos	Frecuencia	Porcentaje
Bueno	68	9 %
Excelente	53	7 %
Mejor	26	3 %
Gran	22	3 %
Misma	15	2 %
Peor	12	2 %
Mismo	11	2 %
Nuevo	8	1 %
Primer	7	1 %
Serios	6	1 %

Cuadro IX
ADJETIVOS EN CUENTAS CON ENGAGEMENT PROMEDIO

Por otro lado, los resultados planteados por ciertas investigaciones y encuestas internacionales como ser [15] y [16], por la que se describiría un contexto de emociones positivas expresadas a través de opiniones en relación al tema, queda ratificada por un 40 % y un 41 % de entradas de carácter positivo. Además, las cuentas con bajo engagement como las de engagement promedio, presentan un nivel similar en la escala de subjetividad, donde se observa una media de 0.31 y 0.36 puntos.

Esto también se sustenta con el reiterado uso del adjetivo *excelente* en ambos segmentos, así como otras construcciones relativas al bienestar y a la satisfacción de los usuarios (bueno, mejor, gran, increíble).

VII. LIMITACIONES Y RECOMENDACIONES

Una de las limitaciones principales de la investigación fue la poca disposición de texto en las publicaciones seleccionadas. Esto pudo ser producto de diversos factores; en primer lugar, recordemos que en Instagram el intercambio de información es principalmente visual (imágenes, videos, fotografías, etc.), no como es el caso de Twitter, por ejemplo, en el que se intercambia información sobre todo de forma escrita.

Otro punto importante es el de incorporar un mecanismo para automatizar la revisión ortográfica y así reducir el margen de error. Si bien es cierto se comprobó que al encontrar errores ortográficos en sustantivos, artículos, pronombres o cualquier palabra que no estuviese reconocida por TextBlob como negativa (positiva), no había diferencia en los resultados del AS; si existía cierta diferencia en el valor del resultado cuando el error se encontraba en una de estas palabras. Por tanto, se disminuiría el porcentaje de error al incorporar una técnica o programa que sea capaz de corregir grandes cantidades de texto de manera eficiente.

VIII. TRABAJO FUTURO

En base a los puntos discutidos en la sección anterior, y si se quiere seguir utilizando Instagram como la red social de la investigación (en el caso de twitter sería un poco más accesible por la diversidad de herramientas que existen para realizar AS) se puede solventar el primer problema discutido, ampliando el campo y objeto de estudio. Es decir, se pueden modificar las características de las cuentas seleccionadas de manera que no sean solamente cuentas de periódicos/noticieros hondureños sino, por ejemplo, se admitan periódicos/noticieros internacionales que cuenten con mayor porcentaje de engagement; o bien, en vez de considerar solo el poder ejecutivo, se puede ampliar al gobierno hondureño en general.

Es importante también, el tema de poder tener una referencia o escala por medio de la cual una persona que no sepa cómo interpretar los datos numéricos de TextBlob pueda guiarse para la correcta interpretación de los resultados. Esto podría hacerse, dependiendo de los recursos disponibles para la investigación, incorporando especialistas en el área de las

ciencias sociales y trabajar en conjunto para crear criterios de referencia.

IX. AGRADECIMIENTOS

A mis padres, por siempre inculcar en mi el deseo de aprender. A Gabriel, por ser mi constante apoyo. Al Dr. Fredy Vides, por ser maestro y líder.

X. DISPONIBILIDAD DE DATOS

El código utilizado para realizar el análisis de sentimiento en TextBlob, así como los datos de la investigación en formato .txt están disponibles en [14].

REFERENCIAS

- [1] ¿Cuáles son las redes sociales más usadas en Honduras en 2021? (s.f). Recuperado de <https://ilifebelt.com/cuales-son-las-redes-sociales-mas-usadas-en-honduras-en-2021/2021/02/#:~:text=La%20siguiente%20red%20social%20elegida,imperecedero%20Twitter%2C%20con%20290.000%20perfiles>.
- [2] Román Mastrángelo, Eric; Aguaded, Ignacio (2018). Herramienta integral de análisis de sentimientos en redes sociales. Tesis de pregrado. Universidad Nacional del Centro de la Prov. de Bs.As.
- [3] Garcés Chaparro, Tomás Ignacio (2019). Análisis de sentimientos en redes sociales orientado a la percepción de la calidad de servicios de internet, redes móviles, tv cable y electricidad.
- [4] Vizcaino-Verdu, Arantxa; Aguaded, Ignacio (2020). *Análisis de sentimiento en Instagram: polaridad y subjetividad de cuentas infantiles*. Zer. <https://doi.org/10.1387/zer.21454>
- [5] H. Blasco Fontecilla (2021). *El impacto de las redes sociales en las personas y en la sociedad: redes sociales, redil social, ¿o telaraña?*. Tarbiya, Revista De Investigación E Innovación Educativa, (49). <https://doi.org/10.15366/tarbiya2021.49.007>
- [6] M. Lamba, M. Madhusudhan (2021). *Text Mining for Information Professionals: An Uncharted Territory*. Springer.
- [7] S. De, S. Dey, S. Bhattacharyya, S. Bhatia, (2021). *Advanced Data Mining Tools and Methods for Social Computing*. ELSEVIER.
- [8] Y. Shi (2022). *Advances in Big Data Analytics: Theory, Algorithms and Practices*. Springer.
- [9] B. Gallardo-Paúl, S. Enguix Oliver (2016). *Pseudopolítica: el discurso político en las redes sociales*. Guada Impresores S.L.
- [10] R. Colle (2011). *El análisis de contenido de las comunicaciones 2. Técnicas de análisis*. Sociedad Latina de Comunicación Social. Creative Commons. <http://www.revistalatinacs.org/067/cuadernos/artesanos.html>
- [11] *Making Natural Language Processing easy with TextBlob* Sociedad Latina de Comunicación Social. Creative Commons. (2021, Octubre 9) Recuperado de <https://www.analyticsvidhya.com/blog/2021/10/making-natural-language-processing-easy-with-textblob/>
- [12] Alemán Viteri, S. (2021, Octubre 10). Análisis de sentimientos para Twitter con Vader y TextBlob. ODIGOS, Vol. 2 (Núm. 3), pp. 12-13.
- [13] De la Vega, R. (s.f.). Análisis de sentimientos en Python con TextBlob. Recuperado de <https://pharos.sh/analisis-de-sentimientos-en-python-con-textblob/>
- [14] A. Meza. SA-con-TextBlob. Recuperado de <https://github.com/ammr98/SA-con-TextBlob>
- [15] Hondureños califican con un 6 sobre 10 al Gobierno de Castro, según un sondeo. (05 julio, 2022). Recuperado de https://www.swissinfo.ch/spa/honduras-encuesta_hondure%C3%B1os-califican-con-un-6-sobre-10-al-gobierno-de-castro--seg%C3%BAAn-un-sondeo/47729764
- [16] DATOWORLD [@Datoworld]. (20 junio, 2022). Aprobación de los presidentes LATAM https://twitter.com/Datoworld/status/1538876653585281030?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1538876653585281030%7Ctwgr%5E213ec3d8fe7f4957ef22748fa50e07408fb27fb8%7Ctwcon%5Esl_1&ref_url=https%3A%2F%2Ftnh.gob.hn%2Fgobierno%2Fdatoworld-presidenta-castro-entre-los-mandatarios-mejor-evaluados-de-america-latina%2F