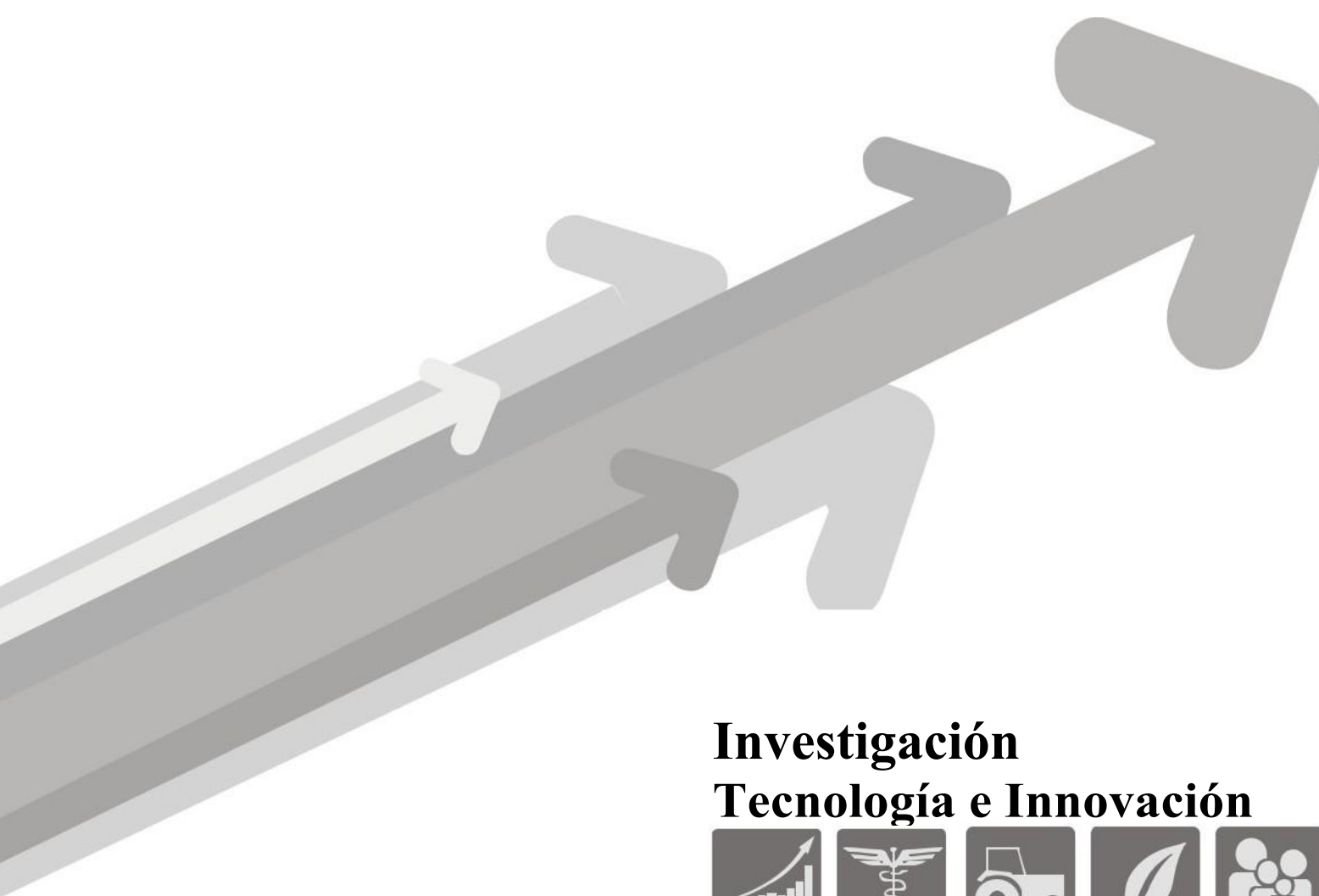


Minería de comentarios en la red social X sobre la percepción de la seguridad en la ciudadanía ecuatoriana

Mining comments from the social network x on the security perceptions of Ecuadorian citizens

Kenia Mabel Navarrete García

Ricardo Ordoñez-Avila



Investigación

Tecnología e Innovación

Minería de comentarios en la red social X para medir la percepción de la seguridad en la ciudadanía ecuatoriana

Mining comments on social network X to measure the security perception among Ecuadorian citizens

Kenia Mabel Navarrete García¹, Ricardo Ordoñez-Avila²

Como citar: Navarrete García, K. M., & Ordoñez-Avila, R. (2025). Minería de comentarios en la red social X sobre la percepción de la seguridad en la ciudadanía ecuatoriana. *Investigación, Tecnología e Innovación*, 17(24), 46-56. <https://doi.org/10.53591/iti.v17i24.2609>

RESUMEN

Contexto: En el actual contexto de creciente atención hacia la seguridad ciudadana en Ecuador, las redes sociales se han consolidado como espacios donde se manifiestan percepciones colectivas sobre la violencia y el crimen. **Objetivo:** Este estudio tiene como propósito analizar publicaciones extraídas de la red social X (antes Twitter), con el fin de identificar patrones discursivos vinculados a la inseguridad, mediante técnicas computacionales aplicadas a datos textuales. **Técnica:** La muestra estuvo compuesta por más de 2.000 publicaciones geolocalizadas en seis cantones: Guayaquil, Quito, Cuenca, Portoviejo, Manta y Durán, recolectadas en seis cortes entre junio y julio de 2025. El preprocesamiento incluyó la depuración de duplicados, normalización léxica y análisis de tokens. Se emplearon técnicas de minería de texto como el análisis de frecuencia de términos, generación de bigramas, clasificación léxica de sentimiento y modelado temático. **Resultados:** El análisis reflejó una alta recurrencia de términos vinculados a hechos delictivos y una predominancia de mensajes con polaridad negativa, particularmente en Guayaquil, Manta y Durán. A través del modelo de clasificación Latent Dirichlet Allocation (LDA) se identificaron cinco tópicos recurrentes que agrupan diferentes expresiones discursivas sobre la inseguridad. Además, se sugiere el empleo de otros modelos de clasificación como Naive Bayes para estimar la polaridad de nuevas publicaciones, mediante una mayor proporción de datos previamente etiquetados. **Conclusión:** Este estudio destaca la importancia del procesamiento de lenguaje natural en combinación con técnicas de clasificación para facilitar el reconocimiento de patrones de interés dentro de la seguridad ciudadana del país, desde la percepción de usuarios de redes sociales.

Palabras clave: seguridad ciudadana, minería de texto, redes sociales, sentimiento, LDA, Ecuador.

ABSTRACT

Context: In the current context of growing attention to citizen security in Ecuador, social media has become a space where collective perceptions about violence and crime are expressed. **Objective:** The purpose of this study is to analyze posts extracted from social media platform X (formerly Twitter) in order to identify discursive patterns linked to insecurity, using computational techniques applied to textual data. **Technique:** The sample consisted of more than 2,000 geolocated posts in six cantons: Guayaquil, Quito, Cuenca, Portoviejo, Manta, and Durán, collected in six rounds between June and July 2025. Preprocessing included duplicate removal, lexical normalization, and token analysis. Text mining techniques such as term frequency

¹ Maestría en Ciencia de Datos y Máquinas de Aprendizaje, Facultad de Posgrado, Universidad Técnica de Manabí, Portoviejo, Ecuador. Correo electrónico: knavarrete7643@utm.edu.ec

² Departamento de Sistemas Computacionales, Universidad Técnica de Manabí, Portoviejo, Manabí, Ecuador. Correo electrónico: ermenson.ordonez@utm.edu.ec



analysis, bigram generation, lexical sentiment classification, and thematic modeling were used. **Results:** The analysis revealed a high recurrence of terms linked to criminal acts and a predominance of messages with negative polarity, particularly in Guayaquil, Manta, and Durán. Using the Latent Dirichlet Allocation (LDA) classification model, five recurring topics were identified that group together different discursive expressions about insecurity. Additionally, the use of other classification models, such as Naive Bayes, is suggested to estimate the polarity of new posts, utilizing a higher proportion of previously labeled data. **Conclusion:** This study emphasizes the significance of natural language processing in conjunction with classification techniques to identify patterns of interest within the country's citizen security, as perceived by social media users.

Keywords: public safety, text mining, social media, sentiment analysis, LDA, Ecuador.

Fecha de recepción: Agosto 30, 2025.

Fecha de aceptación: Octubre 16, 2025.

INTRODUCCIÓN

Durante la última década, las redes sociales se han convertido en espacios centrales para compartir opiniones, emociones y experiencias cotidianas (Ferrara & Yang, 2015; Lagrange, 2025). Entre estas plataformas, la red social X (antes Twitter) destaca como un escenario de uso frecuente para el debate público, donde los usuarios comentan en tiempo real sobre diversos acontecimientos sociales (Highfield & Miltner, 2023). Esta dinámica ha despertado el interés en los investigadores por explorar el potencial de los comentarios en redes sociales como fuente para el análisis de casos específicos, como el de la percepción ciudadana sobre la seguridad (Chaparro et al., 2021; Valla, 2022). El estudio de mensajes breves, directos y con carga social permite acercarse a la forma en que las personas interpretan su entorno.

En este trabajo se analizaron los comentarios publicados por usuarios en el Ecuador en la red social X. El enfoque de este trabajo fue analizar particularmente las publicaciones vinculadas con la seguridad ciudadana. Para ello, se aplicaron técnicas de minería de texto orientadas a organizar y examinar información no estructurada, con el fin de obtener información útil e interpretable para la toma de decisiones en marcos prospectivos de gestión de la seguridad ciudadana (Alghamdi & Alfalqi, 2015; Avasthi et al., 2022; Buenano-Fernandez et al., 2020; Kobayashi et al., 2018; Kumar et al., 2021).

La selección de X como red social responde a su amplio uso en el país y a la accesibilidad de sus publicaciones mediante herramientas tecnológicas que facilitan su recopilación y procesamiento (Adams et al., 2021; Shakhayev & Seyidova, 2023). Con la aplicación de técnicas de minería de texto se podría identificar patrones asociados a preocupaciones recurrentes, emociones y percepciones de los ciudadanos sobre la seguridad, que se interpretan en diversas formas y polaridades. Además, se puede ampliar las formas de comprender diferentes fenómenos sociales como el de la inseguridad. Sus resultados se convierten en recursos complementarios que consolidan los registros oficiales de entidades de gobierno. Por otro lado, reconocer el sentimiento colectivo presente en la ciudadanía en torno a la seguridad puede ser un insumo de información importante para la toma de decisiones (Adams & Lubbers, 2022; Bai, 2022). En este sentido, el análisis de datos generados en redes sociales puede complementar los enfoques estadísticos tradicionales y contribuir con elementos útiles para la formulación de políticas públicas.

Para guiar este estudio, se plantearon las siguientes preguntas de investigación: ¿Qué patrones de percepción sobre la seguridad ciudadana pueden observarse a partir del análisis de estas publicaciones? y ¿Qué técnicas de minería de texto resultan adecuadas para analizar los comentarios realizados en la red social X? Este trabajo se ha organizado en las secciones: Introducción y Trabajos relacionados, Metodología, Resultados, Discusión, Limitaciones y Conclusiones principales.

Trabajos relacionados

Algunas investigaciones han analizado la relación entre lo que se publica en redes sociales y los hechos que generan preocupación en la ciudadanía (Guo et al., 2021; Van Nguyen et al., 2025). En el contexto ecuatoriano, se ha detectado un aumento en la actividad en línea cuando ocurren situaciones vinculadas a la inseguridad (Figueroa-Campoverde, 2024; Goyanes et al., 2021; Naranjo-Zolotov et al., 2021). Este comportamiento ha



sido abordado por estudios que analizan cómo se comunican estos temas en entornos digitales (Hodorog et al., 2022; McCarthy et al., 2023).

Chakravarty & Arifuzzaman (2024), junto con Hawkins et al. (2016), exploraron sentimientos expresados en publicaciones de X sobre el sistema de salud en Estados Unidos, utilizando técnicas de procesamiento de lenguaje natural y aprendizaje automático. Herramientas como Tweepy y SpaCy han sido empleadas en estudios similares para recolectar y procesar datos textuales en plataformas sociales (Jani et al., 2025; Khemani et al., 2024; Sewalk et al., 2018). Asimismo, estudios como el de (Jani et al., 2025; Kolyshkina & Simoff, 2021; Usman et al., 2024) detallan el uso combinado de técnicas supervisadas, no supervisadas y análisis de redes en entornos sociales digitales.

Salah et al. (2019) destacan cómo la minería de opiniones puede complementar los mecanismos institucionales de monitoreo y respuesta frente a situaciones relacionadas con la seguridad. Los estudios de Greco & Polli (2021); Hodorog et al. (2022); Pandya y Dave, (2025) muestran que el análisis de publicaciones en redes sociales no solo es factible desde el punto de vista técnico, sino también significativo en contextos donde la percepción de seguridad influye en la calidad de vida.

En la literatura reciente se reporta el uso de algoritmos supervisados como Support Vector Machines (SVM), Random Forest y redes neuronales recurrentes (RNN) para el análisis de sentimientos. En la etapa inicial, varios estudios integran técnicas de web scraping o APIs de plataformas sociales para la recopilación automatizada de datos (Salah et al., 2019; Chakravarty & Arifuzzaman, 2024).

Estudios enfocados en el análisis de sentimientos identifican seis emociones básicas: alegría, tristeza, miedo, enojo, sorpresa y disgusto (Ekman, 1992). Estas categorías han sido retomadas por investigaciones recientes en redes sociales para modelar las respuestas emocionales ante situaciones de inseguridad o violencia (Battle et al., 2025; Khemani et al., 2024). En este estudio se emplea un enfoque léxico de las polaridades: positiva, neutral y negativa; trazando una potencial ampliación del análisis hacia emociones más específica en futuras investigaciones.

Metodología

Este estudio adoptó un enfoque cuantitativo, con un diseño de tipo descriptivo y un alcance exploratorio, desarrollado en cuatro fases principales: recolección de datos, preprocesamiento, análisis textual y modelado. La recolección se llevó a cabo mediante la plataforma Fedica, que permitió obtener publicaciones de dominio público de la red social X, geolocalizadas en seis cantones del Ecuador: Guayaquil, Quito, Cuenca, Portoviejo, Manta y Durán. Dentro de la recolección de datos, se estableció un grupo de control de términos candidatos para la búsqueda de comentarios relacionados con la seguridad. Los términos de búsqueda iniciales se definieron a partir de una revisión exploratoria de publicaciones sobre seguridad en medios digitales y reportes policiales, seleccionando aquellas palabras con mayor frecuencia de aparición y relevancia contextual, como 'inseguridad', 'violencia', 'crimen', 'tiroteo', 'asesinato' y 'balacera'. Con ello, se realizaron seis extracciones entre junio y julio de 2025, considerando en cada corte publicaciones del día y de los siete días anteriores, lo que permitió consolidar un corpus superior a 2.000 registros.

Se utilizó el entorno de desarrollo de Python de Google Colab empleando bibliotecas como Pandas, Numpy, TextBlob, Scikit-learn, Corpora, Matplotlib y Seaborn, para el preprocesamiento, análisis textual, modelado y la generación de visualizaciones. Durante el preprocesamiento, se eliminaron duplicados, registros vacíos y textos ajenos a la temática de seguridad. Se aplicaron procesos de limpieza y normalización textual, que incluyeron la conversión a minúsculas, eliminación de signos de puntuación, palabras vacías y caracteres no alfabéticos, así como una tokenización básica para estructurar el texto en listas de palabras.

En la etapa de análisis textual se realizó el cálculo de frecuencias de términos mediante unigramas y bigramas, facilitando la identificación de las palabras más recurrentes y de combinaciones léxicas representativas del discurso. Dentro del modelado, se implementó el análisis de sentimientos utilizando el enfoque léxico de TextBlob, con el cual se clasificaron las publicaciones según las polaridades: negativa, positiva y neutral. Esta clasificación fue agregada posteriormente por cantón y fecha para su visualización y comparación.



Adicionalmente, se aplicó un modelo de clasificación temática utilizando Latent Dirichlet Allocation (LDA), con el propósito de identificar agrupaciones de tópicos recurrentes en el corpus. El modelo se entrenó con el conjunto de textos preprocesados, y los temas fueron interpretados a partir de las palabras clave más representativas en cada grupo. Las etiquetas temáticas resultantes permitieron observar patrones comunes en las publicaciones entre cantones y caracterizar las principales formas de percepción de la seguridad.

Finalmente, se construyó un modelo de clasificación automática con el algoritmo Naive Bayes, planteando el uso de un subconjunto etiquetado del corpus derivado de los resultados del análisis de sentimientos. El modelo se evaluó utilizando una partición de entrenamiento y prueba en proporción 80/20. Se calcularon métricas de desempeño como: Precisión, para identificar las predicciones positivas; Recall, para recuperar los positivos reales; F1-score, como balance entre la Precisión y Recall. Así mismo, se utilizó la matriz de confusión para identificar patrones de error de resultados falsos positivos y falsos negativos, y posibles sesgos. Esta propuesta permitió observar la capacidad del modelo para reconocer patrones generales de polaridad en nuevas publicaciones.

Resultados

El análisis se realizó sobre un conjunto de más de 2.000 publicaciones extraídas de distintos usuarios de la red social X, con relación a la seguridad ciudadana de seis cantones del Ecuador: Guayaquil, Quito, Cuenca, Portoviejo, Manta y Durán. La recolección se desarrolló en seis cortes durante los meses de junio y julio del año 2025. Se realizaron iteraciones de recolección diaria y semanal para evaluar datos duplicados y consolidar registros de un periodo aproximado de un mes. Se realizó una selección preliminar de términos en publicaciones generales de seguridad para la conformación final de las palabras de búsqueda: inseguridad, violencia, asesinato, balacera, tiroteo, crimen. Una vez concluido el proceso de limpieza y depuración de duplicados o textos incompletos, se obtuvo un corpus homogéneo que permitió desarrollar las distintas etapas del análisis. Como resultado, más de 31942 palabras fueron procesadas en el análisis textual siguiente.

En primer lugar, se examinaron los términos más frecuentes mediante el análisis de N-Grams, de nivel 1 en unigramas y de nivel 2 en bigramas. La Figura 1, presenta las palabras con mayor presencia en el corpus, encabezadas por “crimen”, “violencia” e “inseguridad”, y seguidas por otras palabras como “organizado” y “asesinato”. Estas expresiones reportan la recurrencia temática temporal relacionada con los meses de extracción de los datos, por tanto, no se puede considerar como una posición discursiva sostenida en torno a eventos delictivos y percepción de riesgo general. El análisis de bigramas (Ver Figura 2) evidenció la asociación de términos en frases como “crimen organizado” y otras que aportan elementos sobre el tipo de discurso de la ciudadanía en torno a problemáticas temporales o sostenidas en al menos más de 3 semanas.

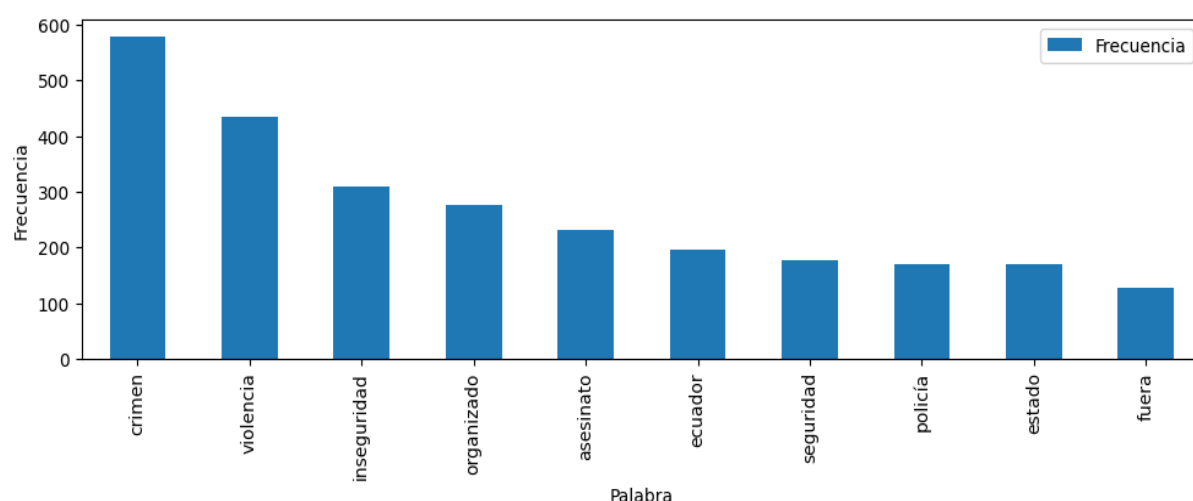
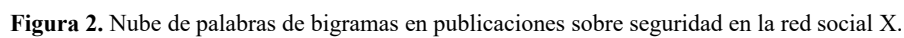


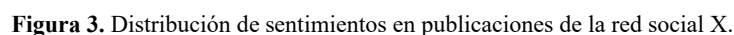
Figura 1. Frecuencia de los 10 principales unigramas en publicaciones sobre seguridad.

Fuente: Autores





Posteriormente, como parte del modelado de datos con técnicas de aprendizaje supervisado, se aplicó el análisis de sentimientos para la clasificación léxica de polaridades en las publicaciones. La Figura 3, muestra los resultados del análisis de sentimientos, destacando una alta proporción de mensajes de carga negativa. La Figura 4, por su parte, presenta la distribución desagregada por cantón, evidenciando variaciones entre territorios. Los reportes siguen el orden de mayor polaridad negativa en las ciudades de Guayaquil, Quito y Manta. Las ciudades de Durán, Cuenca y Portoviejo, presentaron cierta diversidad en todas las polaridades, lo que podría tener relación con una distribución temática más heterogénea.



50

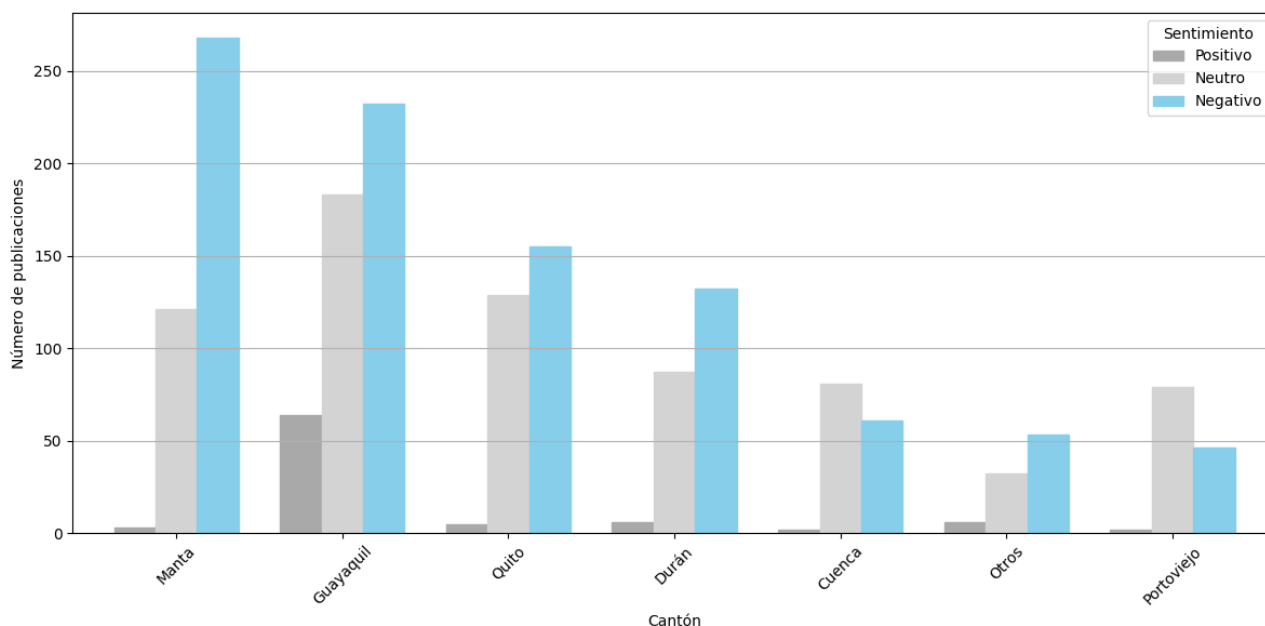


Figura 4. Distribución de sentimientos por cantón (agrupado por categoría).

Fuente: Autores

Adicionalmente, se representó la evolución temporal del sentimiento mediante una escala cualitativa. En ella se incluyeron niveles como “extremadamente negativo” hasta el nivel “positivo”, permitiendo observar ciertas concentraciones de publicaciones con carga crítica en fechas puntuales. Este tipo de visualizaciones puede ofrecer una evidencia temporal de cambios relevantes hacia un tema o situación en análisis sobre la tendencia general de eventos (Ver Figura 5).

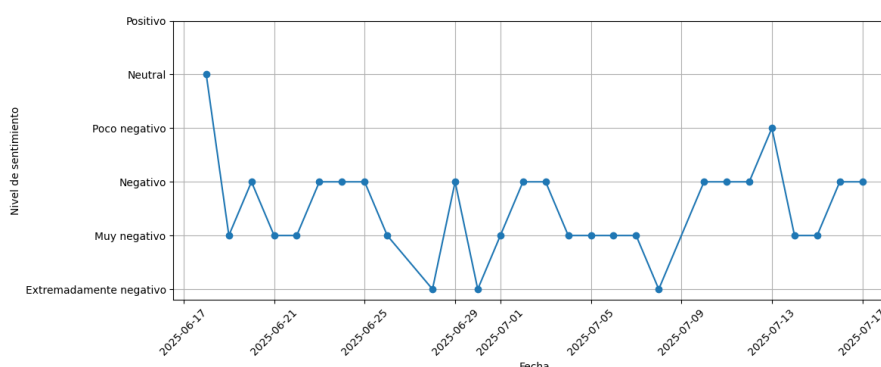


Figura 5. Evolución diaria del sentimiento cualitativo en publicaciones sobre seguridad en la red social X.

Dentro del modelado de datos, Fuente: Autores

Como parte del proceso de modelado de datos, se realizó el análisis temático mediante el modelo de tópicos Latent Dirichlet Allocation (LDA), identificando cinco categorías recurrentes. Estos tópicos fueron rotulados de acuerdo a las palabras más representativas: “Violencia policial y asesinatos”, “Crítica política y conflicto institucional”, “Crimen organizado en Ecuador”, “Delincuencia y percepción ciudadana” e “Inseguridad urbana y violencia”. Esta agrupación permitió identificar patrones discursivos y líneas narrativas comunes entre cantones, aun cuando la distribución de tópicos no fue homogénea en todos los casos. La Figura 6, muestra la representación visual de la proximidad entre temas mediante escalamiento multidimensional, mientras que la Figura 7 presenta la cantidad de publicaciones agrupadas por tópico identificado.

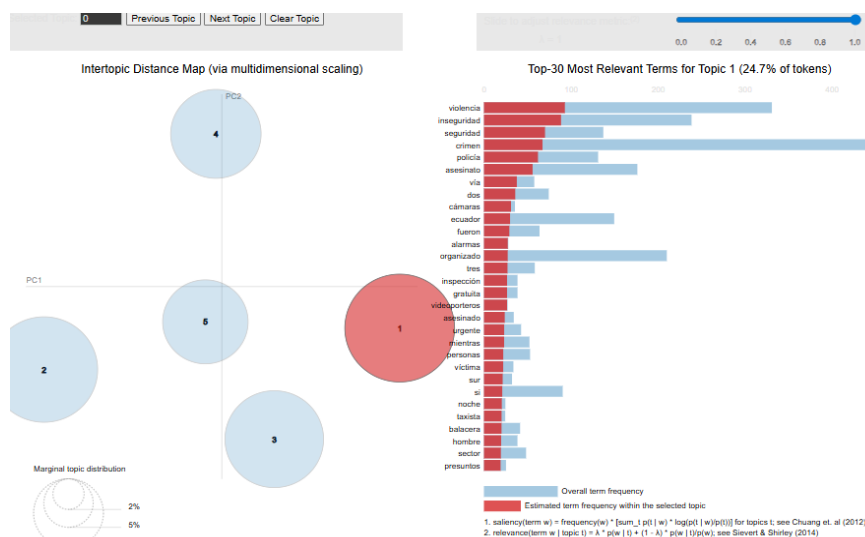


Figura 6. Distribución de tópicos por proximidad (modelo LDA).

Fuente: Autores

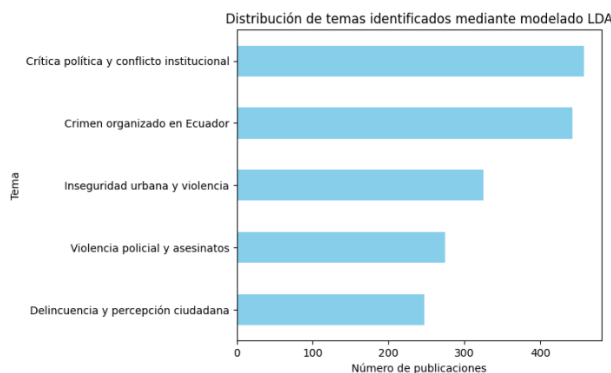


Figura 7. Cantidad de publicaciones por tópico identificado.

Fuente: Autores

Finalmente, con el propósito de ofrecer una referencia adicional basada en los resultados del análisis de sentimientos, se construyó un modelo de clasificación supervisada utilizando el algoritmo Naive Bayes, entrenado con un subconjunto de publicaciones etiquetadas por polaridad. El modelo permitió clasificar nuevos textos según su orientación, con un enfoque exploratorio del estudio. Esta aproximación de resultados, se puede realizar con datos de mayor tiempo y de mayor volumen, para mejorar la explicación del discurso, en términos de polaridades de sentimientos en el tiempo. La Figura 8, presenta la matriz de confusión de este modelo exploratorio para evaluar mejoras en trabajos futuros, tales como el balanceo de clases, y el aprendizaje con una mayor proporción de datos.

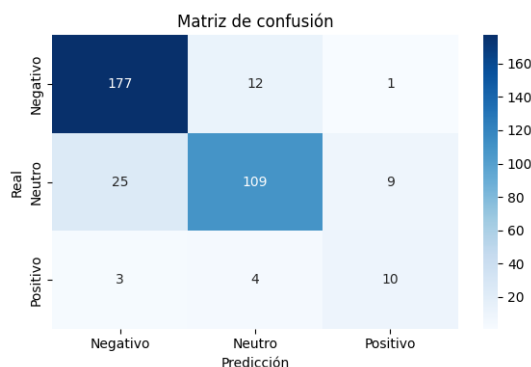


Figura 8. Matriz de confusión del modelo Naive Bayes para clasificación de publicaciones por polaridad.

Fuente: Autores

Discusión

El análisis de publicaciones en la red social X permitió identificar patrones discursivos vinculados a la seguridad ciudadana en distintos contextos locales. Siguiendo lo propuesto por Chaparro et al. (2021), el estudio confirma que los datos generados en plataformas digitales constituyen una fuente alternativa para aproximarse a percepciones sociales sobre fenómenos de interés público. En este caso, la recurrencia de términos asociados a hechos violentos y la orientación mayoritariamente negativa de las publicaciones ofrecen indicios sobre cómo se estructura el discurso digital en torno a la inseguridad.

Se evidenciaron diferencias entre cantones tanto en la frecuencia de publicaciones como en la polaridad obtenida del análisis de sentimientos, reflejando una posible relación entre los contextos territoriales y las narrativas que pueden publicarse en línea. Estos hallazgos son coherentes con lo señalado por Guo et al. (2021), quienes destacan que las condiciones locales pueden influir en la manera en que las personas interactúan con plataformas gubernamentales o expresan sus opiniones sobre temas públicos.

Desde una perspectiva metodológica, el uso de herramientas de minería de texto, análisis de frecuencia y técnicas de clasificación, has sido adecuadas para procesar un volumen considerable de información no estructurada. Tal como sostienen Adams et al. (2021), este tipo de aproximaciones son de utilidad para observar dinámicas sociales en entornos digitales en apoyo a los métodos tradicionales de levantamiento y exploración de incidencias.

El modelado temático mediante LDA permitió categorizar los mensajes en grupos temáticos recurrentes, contribuyendo a una mejor comprensión del contenido predominante. La propuesta de entrenamiento con modelos supervisados como el algoritmo Naive Bayes puede clasificar nuevas publicaciones según la orientación léxica de sentimientos en la percepción ciudadana, ampliando las posibilidades de análisis en trabajos futuros. En consecuencia, las técnicas de clasificación ofrecen valor al momento de explorar discursos amplios y no estructurados, como los que suelen encontrarse en redes sociales (Buenano-Fernandez et al. 2020).

Los resultados obtenidos pueden ser de utilidad para instituciones interesadas en complementar sus diagnósticos sobre percepción ciudadana. El monitoreo de publicaciones digitales podría integrarse como una herramienta adicional en estrategias de análisis situacional, comunicación institucional o intervención comunitaria. Como sugieren Bai (2022) y Goyanes et al. (2021), el aprovechamiento responsable de estos datos puede aportar recursos para el diseño de políticas públicas, particularmente en áreas sensibles como la seguridad.

Además, integrar el monitoreo digital en procesos de diagnóstico contribuiría al diseño de políticas públicas más receptivas y centradas en las preocupaciones de la población. Como sugieren McCarthy et al. (2023), las redes sociales pueden actuar como espacios donde emergen demandas, tensiones o percepciones que los marcos institucionales aún no logran captar plenamente.

Este tipo de aproximaciones no busca sustituir otras formas de levantamiento de información, sino ampliar el espectro de herramientas disponibles para captar el pulso social, especialmente en contextos de alta sensibilidad como el de la seguridad. Reconocer estos patrones emocionales y discursivos puede servir como insumo en estrategias de intervención comunicacional o en el fortalecimiento de canales de diálogo con la comunidad.

Limitaciones

El estudio se basó en publicaciones extraídas de la red social X, lo cual implicó una dependencia directa de los datos disponibles públicamente en dicha plataforma. Si bien se aplicaron filtros geográficos para delimitar los cantones de interés, no fue posible verificar con total certeza la ubicación real de quienes emitieron los mensajes, ya que la información de procedencia de la publicación puede estar incompleta, imprecisa o ser auto declarada por el usuario. Adicionalmente, el análisis textual se apoyó en un diccionario de polaridad y en modelos de emociones que operan sobre vocabularios predefinidos. Esto pudo haber limitado la capacidad del sistema para interpretar expresiones locales, ironías o construcciones lingüísticas propias del contexto analizado o de la naturaleza de los datos.

El alcance exploratorio de este estudio no consideró un tramo de tiempo mayor para una mejor aproximación a la realidad o identificación de patrones en las publicaciones sostenidas. Sin embargo, la ruta de trabajo técnico ofrece una referencia de trabajo mensual para siguientes estudios, y la consideración de mayor tiempo y número de publicaciones para predicciones.

CONCLUSIONES

El análisis de publicaciones en la red social X permitió explorar las formas léxicas y estilos de expresiones de la ciudadanía frente a situaciones relacionadas con la seguridad en distintos cantones del Ecuador. A través de técnicas de procesamiento de texto y modelos de clasificación, fue posible identificar términos frecuentes, emociones predominantes y variaciones en la polaridad del discurso según el contexto geográfico. La frecuencia de palabras, el análisis de sentimientos, el modelado temático, y otras técnicas de clasificación, son herramientas útiles de la minería de texto, para la exploración de comentarios de los ciudadanos en ámbitos relacionados con la seguridad y otros de importancia social.

La observación de tendencias emocionales, así como la organización de grandes volúmenes de información no estructurada, aporta elementos que pueden complementar otros enfoques tradicionales de medición, como las encuestas o los registros institucionales. Esta aproximación puede ser de utilidad para quienes buscan comprender las formas en que ciertos temas circulan y se configuran en entornos digitales.

El flujo de trabajo guiado desde el grupo de control de selección de términos de búsqueda, la combinación del análisis léxico y el modelado temático, proporcionan un enfoque metodológico de minería de texto que puede ser replicado en trabajos futuros orientados a examinar percepciones sociales en contextos locales sobre opinión pública y de seguridad en países con características similares.

En futuras investigaciones, se podrían incorporar modelos preentrenados para el análisis de una mayor cantidad de texto y mejorar la comprensión semántica de las publicaciones. En consecuencia, se podría extender el estudio hacia otras plataformas sociales, con el fin de contrastar narrativas o identificar patrones comunes desde diferentes entornos digitales. Estos enfoques podrían contribuir al desarrollo de herramientas más robustas sobre la naturaleza de datos en un contexto local, y su ampliación centrada en la percepción social de la seguridad.

REFERENCIAS BIBLIOGRÁFICAS

- Adams, jimi, & Lubbers, M. J. (2022). Social Network Data Collection: Principles and Modalities. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.4216936>
- Adams, jimi, Santos, T., & Williams, V. N. (2021). Strategies for Collecting Social Network Data. *The Oxford*



- Handbook of Social Networks*, 117–136. <https://doi.org/10.1093/OXFORDHB/9780190251765.013.10>
- Alghamdi, R., & Alfalqi, K. (2015). A Survey of Topic Modeling in Text Mining. *International Journal of Advanced Computer Science and Applications*, 6(1). <https://doi.org/10.14569/IJACSA.2015.060121>
- Avasthi, S., Chauhan, R., & Acharjya, D. P. (2022). Topic modeling techniques for text mining over large-scale scientific and biomedical text corpus. *International Journal of Ambient Computing and Intelligence*, 13(1). <https://doi.org/10.4018/IJACI.293137>
- Bai, J. (2022). Design and Implementation of Data Analysis System of Social Network. *International Journal of Frontiers in Sociology*, 4(2). <https://doi.org/10.25236/IJFS.2022.040215>
- Battle, H., Álvarez-Mon, M. Á., Lara-Abelenda, F. J., Perez-Araluce, R., & Pinto Da Costa, M. (2025). Attitudes towards mental health professionals in social media: infodemiology study. *The British Journal of Psychiatry*, 1–6. <https://doi.org/10.1192/BJP.2024.261>
- Buenano-Fernandez, D., Gonzalez, M., Gil, D., & Lujan-Mora, S. (2020). Text Mining of Open-Ended Questions in Self-Assessment of University Teachers: An LDA Topic Modeling Approach. *IEEE Access*, 8, 35318–35330. <https://doi.org/10.1109/ACCESS.2020.2974983>
- Chakravarty, U. K., & Arifuzzaman, S. (2024). Sentiment analysis of tweets on social security and medicare. *Social Network Analysis and Mining*, 14(1). <https://doi.org/10.1007/S13278-024-01248-3>
- Chaparro, L. F., Pulido, C., Rudas, J., Victorino, J., Reyes, A. M., Estrada, C., Narvaez, L. A., & Gómez, F. (2021). Quantifying Perception of Security through Social Media and Its Relationship with Crime. *IEEE Access*, 9, 139201–139213. <https://doi.org/10.1109/ACCESS.2021.3114675>
- Ferrara, E., & Yang, Z. (2015). Measuring Emotional Contagion in Social Media. *PLOS ONE*, 10(11), e0142390. <https://doi.org/10.1371/JOURNAL.PONE.0142390>
- Figueroa-Campoverde, D. S. (2024). *Análisis de sentimientos sobre la percepción de seguridad para la ciudad de Cuenca durante el año 2023* [Pontificia Universidad Católica del Ecuador]. <https://repositorio.puce.edu.ec/server/api/core/bitstreams/496f879a-0dab-4ac0-bc9c-569852a7962a/content>
- Goyanes, M., López-López, P. C., & Demeter, M. (2021). Social Media in Ecuador: Impact on Journalism Practice and Citizens' Understanding of Public Politics. *Journalism Practice*, 15(3), 366–382. <https://doi.org/10.1080/17512786.2020.1724180;PAGE:STRING:ARTICLE/CHAPTER>
- Greco, F., & Polli, A. (2021). Security Perception and People Well-Being. *Social Indicators Research*, 153(2), 741–758. <https://doi.org/10.1007/S11205-020-02341-8>
- Guo, J., Liu, N., Wu, Y., & Zhang, C. (2021). Why do citizens participate on government social media accounts during crises? A civic voluntarism perspective. *Information & Management*, 58(1), 103286. <https://doi.org/10.1016/J.IM.2020.103286>
- Hawkins, J. B., Brownstein, J. S., Tuli, G., Runels, T., Broecker, K., Nsoesie, E. O., McIver, D. J., Rozenblum, R., Wright, A., Bourgeois, F. T., & Greaves, F. (2016). Measuring patient-perceived quality of care in US hospitals using Twitter. *BMJ Quality & Safety*, 25(6), 404–413. <https://doi.org/10.1136/BMJQS-2015-004309>
- Highfield, T., & Miltner, K. M. (2023). Platformed solidarity: Examining the performative politics of Twitter hashflags. *Convergence*, 29(6), 1641–1667. https://doi.org/10.1177/13548565231199981/ASSET/FF235F2C-9F65-4354-A4D6-4AFC1B9EBC31/ASSETS/IMAGES/LARGE/10.1177_13548565231199981-FIG11.JPG
- Hodorog, A., Petri, I., & Rezgui, Y. (2022). Machine learning and Natural Language Processing of social media data for event detection in smart cities. *Sustainable Cities and Society*, 85. <https://doi.org/10.1016/J.SCS.2022.104026>
- Jani, J. A., Cowan, D., Ouonkap, L., Adesina, D., Ma, T., Chen, S., Aldakhil, S., & Hoang, K. B. (2025).



Missing the message to brain tumor patients: a 2023 twitter analysis among patients, informal caregivers, and healthcare professionals in glioblastoma multiforme. *Journal of Neuro-Oncology*. <https://doi.org/10.1007/S11060-025-04948-8>

- Khemani, B., Malave, S., Patil, S., Shilotri, N., Varma, S., Vishwakarma, V., & Sharma, P. (2024). Sentimatrix: sentiment analysis using GNN in healthcare. *International Journal of Information Technology*. <https://doi.org/10.1007/S41870-024-02142-Z>
- Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & Den Hartog, D. N. (2018). Text Mining in Organizational Research. *Organizational Research Methods*, 21(3), 733–765. <https://doi.org/10.1177/1094428117722619>
- Kolyshkina, I., & Simoff, S. (2021). The CRISP-ML Approach to Handling Causality and Interpretability Issues in Machine Learning. *Proceedings - 2021 IEEE International Conference on Big Data, Big Data 2021*, 2306–2312. <https://doi.org/10.1109/BIGDATA52589.2021.9671754>
- Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *Int. J. Inf. Manag. Data Insights*, 1(1). <https://doi.org/10.1016/J.JJIMEI.2021.100008>
- Lagrange, B. (2025). Emotions on Social Media as Catalysts for Change: Epistemic and Motivational Potentialities for Gender Equality. *Media and Communication*, 13, 8591. <https://doi.org/10.17645/MAC.8591>
- McCarthy, S., Rowan, W., Mahony, C., & Vergne, A. (2023). The dark side of digitalization and social media platform governance: a citizen engagement study. *Internet Res.*, 33(6), 2172–2204. <https://doi.org/10.1108/INTR-03-2022-0142>
- Naranjo-Zolotov, M., Turel, O., Oliveira, T., & Lascano, J. E. (2021). Drivers of online social media addiction in the context of public unrest: A sense of virtual community perspective. *Computers in Human Behavior*, 121, 106784. <https://doi.org/10.1016/J.CHB.2021.106784>
- Pandya, A., & Dave, B. (2025). INFLUENCE OF SOCIAL MEDIA ON STRESS AND QUALITY OF LIFE IN GENERATION Z-AN OBSERVATIONAL STUDY. *Indian Journal of Physical Therapy*, 6(1), 40–42. <https://doi.org/10.63299/IJOPT.060112>
- Salah, Z., Al-Ghuwairi, A. R. F., Baarah, A., & Aloqaily, A. (2019). A systematic review on opinion mining and sentiment analysis in social media. *International Journal of Business Information Systems*, 31(4), 530–554. <https://doi.org/10.1504/IJBIS.2019.101585>
- Sewalk, K. C., Tuli, G., Hswen, Y., Brownstein, J. S., & Hawkins, J. B. (2018). Using Twitter to Examine Web-Based Patient Experience Sentiments in the United States: Longitudinal Study. *Journal of Medical Internet Research*, 20(10). <https://doi.org/10.2196/10043>
- Shakhayev, S. S. S., & Seyidova, I. S. I. (2023). BIG DATA PROCESSING WITH PYTHON IN SOCIAL NETWORKS. *ETM - Equipment, Technologies, Materials*, 16(04), 76–82. <https://doi.org/10.36962/ETM16042023-76>
- Usman, M., Mujahid, M., Rustam, F., Flores, E. S., Mazón, J. L. V., de la Torre Díez, I., & Ashraf, I. (2024). Analyzing patients satisfaction level for medical services using twitter data. *PeerJ Computer Science*, 10. <https://doi.org/10.7717/PEERJ-CS.1697>
- Valla, L. G. (2022). Citizens' Perceptions of Security Issues: New and Old Actors in the National Security Framework. *Journal of Human Security*. <http://www.librelloph.com/journalofhumansecurity/article/view/johs-18.1.18>
- Van Nguyen, P., Vrontis, D., Nguyen, L. D. P., Nguyen, T. T. U., & Salloum, C. (2025). Unraveling the Role of Citizens' Concerns and Cognitive Appraisals in E-Government Adoption: The Impact of Social Media and Trust. *Strategic Change*.