

# Uso de minería de datos a través de Twitter para análisis del servicio de transporte en el Distrito Metropolitano de Quito

Pablo Pico<sup>1</sup> 

<sup>1</sup>Instituto Superior Tecnológico Quito

[pablo.pico@itq.edu.ec](mailto:pablo.pico@itq.edu.ec)

Corresponsal: [pablo.pico@itq.edu.ec](mailto:pablo.pico@itq.edu.ec); Telf.: +593 993689332

**Resumen:** Las redes sociales, como Twitter, han transformado la manera de producir y difundir información sobre eventos, noticias y opiniones para gran parte de la población mundial. Los datos generados por estas plataformas pueden tener diversas aplicaciones. En este estudio, se exploró una herramienta de análisis que permita evaluar la opinión de los usuarios de transporte público acerca de la calidad del servicio en Quito, mediante los datos obtenidos de Twitter al usar palabras clave que filtran los tweets relacionados con el tráfico. Se realizó una revisión de la literatura científica y gris. Se concluyó que esta herramienta puede ser muy útil para las agencias de tránsito urbanas, al facilitar la toma de decisiones sobre planes de movilidad.

**Palabras claves:** Twitter, minería de datos, procesamiento de datos, análisis de datos, transporte público.

**Abstract:** Social networks, such as Twitter, have transformed the way of producing and disseminating information about events, news and opinions for a large part of the world's population. The data generated by these platforms can have various applications. In this study, a tool for analyzing the opinion of public transport users about the quality of service in Quito was explored, using the data obtained from Twitter by using keywords that filter tweets related to traffic. A review of the scientific and gray literature was carried out. It was concluded that this tool can be very useful for urban transit agencies, by facilitating decision-making on mobility plans.

**Keywords:** Twitter, data mining, data processing, data analytics, public transportation.



Check for updates

## Artículo de revisión

**Cita:** Pico. Uso de minería de datos a través de Twitter para análisis del servicio de transporte en el Distrito Metropolitano de Quito. Revista DOXA ITQ, 1(2), 001

[https://itq.edu.ec/wp-content/uploads/2024/01/2023-09-17\\_doxa\\_1-2-1-.pdf](https://itq.edu.ec/wp-content/uploads/2024/01/2023-09-17_doxa_1-2-1-.pdf)

**Recibido:** 22/10/2022

**Aceptado:** 16/01/2023

**Publicado:** 17/09/2023

Santiago del Castillo G., MSc.  
Editor en jefe, Revista DOXA ITQ  
Quito, Ecuador

**Nota del editor:** La Revista DOXA ITQ mantiene una posición neutral con respecto a cualquier reclamo legal que pueda surgir del contenido publicado. La responsabilidad de la información recae enteramente en los autores.

## 1. Introducción y metodología

Una gran cantidad de estudios se han enfocado en el uso de las redes sociales para la investigación relacionada con el servicio de transporte (Bevilacqua & Neira-Villena, 2021). Estos estudios se pueden clasificar en las siguientes categorías: estimación de la demanda de viajes, evaluación del servicio de movilidad, monitoreo del nivel de tráfico y modelado sobre desastres naturales (Motos Cascales, 2019). A la fecha, pocos estudios han utilizado información de las redes sociales para el análisis del servicio de transporte público (Tantaleán Olano, 2021).

Dentro de los estudios que se han realizado, el enfoque ha sido principalmente el análisis de sentimientos para evaluar la opinión de los usuarios sobre el sistema de transporte (Collins et al., 2013). En el trabajo de “Planning and social media”, Schweitzer, (2014) utilizó tweets para evaluar las opiniones de los usuarios sobre el transporte público, descubriéndose que los usuarios de Twitter expresan sentimientos más negativos sobre el transporte público en relación con otros servicios públicos como policía, médico o bomberos. También, se encontró que las agencias de tránsito, cooperativas o gremios de transporte que responden directamente a las preguntas y críticas de sus usuarios, demuestran sentimientos más positivos (Amaya Molinar et al., 2017). Otros estudios analizaron los datos de Twitter para evaluar la satisfacción de los pasajeros de tránsito recopilando los tweets que contenían palabras clave de nombres de trenes en la ciudad de Chicago (Luong & Houston, 2015).

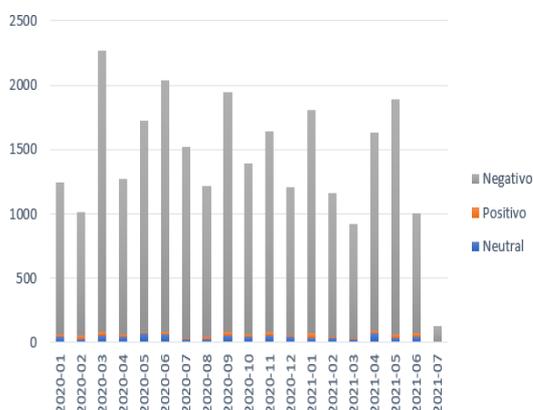
En este trabajo, se pretendió establecer la percepción de la población hacia el sistema de transporte público en el Distrito Metropolitano de Quito, utilizando las opiniones de los usuarios expresadas vía plataforma Twitter. Para su realización, se utilizó revisión de la literatura científica y gris que permita contrastar la información

con el estudio realizado. El procedimiento que se siguió fue el siguiente: en primer lugar, se realizó la escucha de los posts y el procesamiento de los tweets relacionados con información de transporte público, para posteriormente preparar los datos para un análisis semántico. Para esto, se usó un diccionario con palabras claves extraídas de los principales motores de búsqueda. Luego, se utilizó un algoritmo de modelado para extraer tweets relevantes que contengan las opiniones de los usuarios de transporte público sobre la calidad del servicio. Por último, se llevó a cabo un análisis de sentimiento para evaluar los comentarios de los pasajeros sobre el servicio.

## 2. Discusión de resultados

Los resultados que se obtuvieron por las publicaciones relacionadas con la gestión del transporte público permitieron establecer que la percepción de la población hacia el servicio en general es negativa. La extracción de datos se realizó para el periodo comprendido entre enero de 2020 hasta julio de 2021. Mensualmente se extrajeron 300.000 posts, verificando que su origen se encuentre limitado dentro del Distrito Metropolitano de Quito, por medio de un filtrado de las coordenadas que la plataforma registra cuando el usuario realiza las publicaciones. De estas publicaciones, en promedio, solo 1422 presentaron contenido relacionado con el servicio de transporte. De este grupo de datos, el 95.5% fueron clasificados como negativos.

**Figura 1**  
Análisis de sentimiento de las publicaciones extraídas de Twitter



Nota. En la figura se observan los valores de posts mensuales para el transporte público en el Distrito Metropolitano de Quito para el periodo comprendido entre enero de 2020 y julio 2021. Realización propia.

Uno de los hallazgos más importantes en el análisis de las palabras clave a través de los tweets es que generalmente los usuarios expresan impresiones o sentimientos negativos que reflejan la insatisfacción que tienen con el servicio. Las palabras más frecuentes fueron: “exigimos”, “presidente”, “municipal”, “gobernador”, “garanticen”, “segura”, “habitantes”, “aplicación”, “ley”, “investigación”, “atropellamiento”, “vida”, “jacinto león guzmán”, “justiciaparajacinto”, “nomasmuertesviales”. Con los resultados obtenidos en el análisis de los tweets, se podría informar a las empresas públicas de este sector sobre cómo los usuarios los califican y así se podría identificar todas las falencias que actualmente tienen.

Una vez determinadas las palabras clave que más aparecen en las publicaciones de los usuarios, se podrían proponer las mejoras a realizar. Dichas empresas podrían responder de forma adecuada, desde sus distintas plataformas, a cada uno de estos comentarios negativos para transformarlos en experiencias positivas para el usuario. Para las cooperativas o compañías de transporte público, Twitter se convertiría en la mejor herramienta para cambiar la perspectiva

de los clientes. Entre las publicaciones que tuvieron una clasificación positiva, se encuentran: “buenas noticias”, “metrovuelvealaobra”, “generará plazas de empleo”, “suman mil colaboradores”, “aportamos”, “reactivación productiva”, “mientras continuamos construcción”, “obra grande país”, “disciplinaparavolver”. Por último, se observó que en algunos casos de clasificación errónea, se debió generalmente a que se usaron palabras de jerga local.

### 3. Aportes principales en tecnología e innovación

En análisis planteado en este documento, muestra de forma práctica el uso de minería de datos en las redes sociales y cómo se puede enfocar en un servicio en específico. Este tipo de análisis no se limita a solo el sentir de la ciudadanía, también puede ser usado para monitoreos de accidentes de tránsito, robos, incendios, etc. Resulta muy importante para la mejora de otros servicios al usar las palabras claves encontradas como debilidades o falencias que se deban corregir, en el caso de sentimientos negativos; o como fortalezas que se deban potenciar, en el caso de sentimientos positivos o neutros.

La minería de datos, según Álvarez Gonzaga, (2021) tiene muchas aplicaciones; como el uso en temas políticos, monitoreando el sentimiento de la población hacia un posible candidato, o la popularidad que esté presente en grupos sociales específicos para poder potenciar o redirigir las campañas. Se puede usar para el análisis de mercado, encontrando el posicionamiento de la marca o producto entre la población. También puede usarse para detectar eventos de ciberseguridad por medio de monitoreo de actividad del equipo, de los patrones de interacción usuario-máquina, o registros de funcionamiento de equipos electrónicos (Guerrero Fernández, 2022).

### 4. Conclusiones

Las mejoras que se puedan presentar a este trabajo requerirían un procesamiento previo adicional para identificar las palabras de jerga local, refinamiento de los modelos, y uso de información multimedia como fotos, emojis o audios para obtener un conjunto de datos más uniforme.

Deberían existir iniciativas que permitan usar los datos de redes sociales en busca de encontrar la percepción de la población para los servicios que brinden las organizaciones. Esto ayudaría a mejorar la cultura sobre el uso de datos y permitiría establecer un proceso de mejora continua que ayude a encontrar y determinar posibles falencias dentro de las instituciones.

En el caso específico del servicio de transporte público, la percepción de la población ha sido negativa y presenta un conjunto de palabras claves que las autoridades podrían usar para establecer un reglamento o método de evaluación que esté de acuerdo con la percepción e intereses reales de la población en general.

Este documento se puede utilizar como una recomendación para mejorar el rendimiento del transporte público en las ciudades. Ampliar la investigación, permitiría que los resultados sean más precisos, enriqueciendo el diccionario de palabras claves, como ya se ha mencionado anteriormente, extrayendo información de redes sociales, censos, denuncias y cualquier otra fuente que permita aumentar la calidad y cantidad de los datos.

---

**Contribución del autor:** El autor ha contribuido en todos los apartados de la investigación.

**Financiamiento:** El autor financió totalmente el estudio.

**Conflictos de intereses:** El autor declara no tener ningún conflicto de intereses.

## Referencias

- Alvarez Gonzaga, B. R. (2021). *Análisis comparativo de técnicas de minería de datos aplicadas a business intelligence*.
- Amaya Molinar, C. M., Magaña Espinoza, P., & Ochoa Llamas, I. (2017). Evaluación de destinos turísticos mediante la tecnología de la ciencia de datos. *Estudios y Perspectivas En Turismo*, 26(2), 286–305.
- Bevilacqua, S., & Neira-Villena, J. E. (2021). Percepción del riesgo en viajes compartidos. Efectos de «Ubereconomía» en el transporte de taxis. *Revista EAN*, 90, 213–232.
- Collins, C., Hasan, S., & Ukkusuri, S. v. (2013). A novel transit rider satisfaction metric: Rider sentiments measured from online social media data. *Journal of Public Transportation*, 16(2), 2.
- Guerrero Fernández, M. (2022). *Minería de datos para phishing*.
- Luong, T. T. B., & Houston, D. (2015). Public opinions of light rail service in Los Angeles, an analysis using Twitter data. *IConference 2015 Proceedings*.
- Motos Cascales, G. (2019). *Análisis de los indicadores de movilidad urbana sostenible*.
- Schweitzer, L. (2014). Planning and social media: a case study of public transit and stigma on Twitter.
- Tantaleán Olano, G. M. (2021). *Políticas públicas en movilidad urbana sostenible para prevenir contagios por COVID-19, Trujillo 2021*.