



3.2 Minería de datos en Twitter para sentiment análisis.

- 3.2.1 Creación de cuenta para desarrollador, Application Programming Interface (API).
- 3.2.2 Paquetería para Twitter Mining.
- 3.2.3 Caso práctico: El estado de ánimo de los tuiteros respecto al “NAFTA”, “Trump”.

Minería de Datos en Twitter:

La minería de datos en Twitter implica el uso de técnicas computacionales para extraer información útil y significativa de los datos generados en la plataforma. Esto puede incluir la extracción de texto de tweets, identificación de temas, detección de tendencias y, crucialmente, análisis de sentimientos.

El análisis de sentimientos se refiere a la tarea de determinar la actitud emocional expresada en un texto, ya sea positiva, negativa o neutral. En el contexto de Twitter, esto puede aplicarse para comprender cómo se sienten las personas sobre un tema específico, un producto, un evento o incluso una marca.



Herramientas y técnicas:

Para llevar a cabo la minería de datos en Twitter para análisis de sentimientos, se emplean una variedad de herramientas y técnicas. Estas pueden incluir algoritmos de aprendizaje automático, procesamiento del lenguaje natural (NLP), y técnicas específicas de análisis de sentimientos como el análisis léxico basado en diccionarios o modelos de aprendizaje profundo.

Las API proporcionadas por Twitter permiten a los investigadores y desarrolladores acceder a grandes cantidades de datos en tiempo real, facilitando así el proceso de recopilación de datos para el análisis.

Aplicaciones y beneficios:

Las aplicaciones de la minería de datos en Twitter para análisis de sentimientos son diversas y van desde la investigación académica hasta la toma de decisiones empresariales. Por ejemplo, las empresas pueden utilizar esta técnica para evaluar la percepción del consumidor sobre sus productos o servicios, identificar problemas potenciales y tomar medidas correctivas de manera proactiva. Los gobiernos pueden monitorear el sentimiento público sobre políticas específicas o eventos importantes, lo que les permite adaptar sus estrategias de comunicación y políticas en consecuencia.

Limitaciones y desafíos:

A pesar de sus numerosos beneficios, la minería de datos en Twitter para análisis de sentimientos enfrenta varios desafíos y limitaciones. Por un lado, la naturaleza ruidosa y a menudo ambigua de los tweets puede dificultar la precisión del análisis de sentimientos, especialmente en contextos donde el sarcasmo, la ironía o el lenguaje coloquial son comunes. Además, la representatividad de los datos puede ser un problema, ya que los usuarios de Twitter no son necesariamente una muestra representativa de la población en general.



Pasos para crear una cuenta de desarrollador y obtener acceso a la API de Twitter:

1. Registrarse en el Portal de Desarrolladores de Twitter:

- Accede a <https://developer.twitter.com/en/portal/dashboard>.
- Haz clic en "Iniciar sesión" y selecciona "Crear una cuenta".
- Introduce tu nombre, correo electrónico, contraseña y acepta los términos y condiciones.
- Verifica tu cuenta de correo electrónico haciendo clic en el enlace que te enviará Twitter.

2. Crear una aplicación:

- En el panel de control del desarrollador, haz clic en "Crear aplicación".
- Introduce el nombre de tu aplicación, una descripción y selecciona la categoría que mejor la represente.
- Acepta los términos y condiciones y haz clic en "Crear aplicación".

3. Obtener las claves de la API:

- En la pestaña "Detalles de la aplicación", busca la sección "Claves y tokens de acceso".
- Haz clic en "Crear token de acceso".
- Revisa los permisos que se le otorgan a tu aplicación y haz clic en "Crear".
- Guarda las claves de la API (API Key y API Secret) en un lugar seguro.

4. Solicitar acceso elevado (opcional):

- Si necesitas acceder a datos más sensibles de Twitter, como tweets históricos o información de usuarios, puedes solicitar acceso elevado.
- Para ello, ve a la pestaña "Permisos" de tu aplicación y haz clic en "Solicitar acceso elevado".
- Completa el formulario de solicitud y espera la aprobación de Twitter.



Universidad
de la Ciudad de
Aguascalientes

Importante:

A partir del 9 de febrero de 2024, la API de Twitter ya no es gratuita. Se requiere un plan de pago para acceder a ella.

Información sobre los planes de precios y las opciones de pago en <https://developer.twitter.com/en/products/twitter-api>.

Developer Platform Products Docs Use Cases Community Support Developer Portal Sign in

Find the right access for you

Free	Basic	Pro	Enterprise
For write-only use cases and testing the X API	For hobbyists or prototypes	For startups scaling their business	For businesses and scaled commercial projects
<ul style="list-style-type: none">Rate limited access to v2 post posting and media upload endpoints1,500 Posts per month - posting limit at the app level1 app IDLogin with XFree	<ul style="list-style-type: none">Rate limited access to suite of v2 endpoints3,000 Posts per month - posting limit at the user level50,000 Posts per month - posting limit at the app level10,000 Posts per month - read-limit rate cap2 app IDsLogin with X\$100 per month	<ul style="list-style-type: none">Rate-limited access to suite of v2 endpoints, including search and filtered stream1,000,000 Posts per month - GET at the app level300,000 Posts per month - posting limit at the app level3 app IDsLogin with X\$5,000 per month	<ul style="list-style-type: none">Commercial-level access that meets your and your customer's specific needsManaged services by a dedicated account teamComplete streams: replay, engagement metrics, backfill, and more featuresMonthly subscription tiers
Get started	Subscribe now	Subscribe now	Apply now



Universidad
de la Ciudad de
Aguascalientes

Developer Portal

Docs Community Updates Support

Developer agreement & policy

Describe all of your use cases of Twitter's data and API:

We need this information for data protection. [Learn more](#)

- ☐ You understand that you may not resell anything you receive via the Twitter APIs
- ☐ You understand your Developer account may be terminated if you violate the Developer Agreement or any of the Incorporated Developer Terms
- ☐ You accept the Terms & Conditions
- By clicking on the box, and by otherwise accessing or using any Licensed Material, you indicate that you have read and agree to this [Developer Agreement](#) and the [Twitter Developer Policy](#)

Back Submit

Developer Portal

Dashboard
Projects & Apps
Products
Account

Docs Community Updates Support

Dashboard

Projects

Default project-1763722089930891264

MONTHLY TWEET CAP USAGE

0 Tweets pulled of 1,500 0% Resets on April 2 at 00:00 UTC

PROJECT APP

1763722089930891264Oswald0D



Sample Apps

Check out our library of sample Apps on [Glitch](#) and clone one to get started.

[View sample Apps](#)



Helpful docs

[Docs home](#)

[Make your first request](#)



Elección de la paquetería:

- rtweet es la opción más recomendable para iniciarse en Twitter Mining en R por su facilidad de uso y amplia gama de funcionalidades.
- twitteR puede ser viable si ya está familiarizado con ella, pero requiere más configuración.
- streamR es ideal para la extracción de datos en tiempo real.
- tidytext y quantdata son excelentes complementos para el análisis avanzado de texto de los tweets obtenidos con rtweet o twitteR.

rtweet:

Es la paquetería más popular para Twitter Mining en R.

Ofrece una interfaz fácil de usar para:

- Autenticación: Conectarse e interactuar con la API de Twitter.
- Búsqueda de tweets: Buscar tweets por palabras clave, hashtags, usuarios y otros criterios.
- Recuperación de datos: Obtener información del usuario, detalles del tweet, datos de metadatos como la ubicación y el idioma.
- Filtrado y limpieza: Limpiar y preprocesar los datos de Twitter para el análisis.
- Es compatible con R 3.0.0 y superior.

<https://cran.r-project.org/package=rtweet>



twitterR:

- Otra opción popular, aunque menos actualizada que rtweet.
- Ofrece funcionalidades similares a rtweet para la búsqueda de tweets, recuperación de datos y filtrado.
- Puede requerir trabajo adicional para configurar la autenticación con la API de Twitter.
- Es compatible con R 2.15.1 y superior.

<https://www.rdocumentation.org/packages/twitterR/versions/1.1.9>

streamR:

- Especializada en la extracción de datos de Twitter en tiempo real (streaming).
- Permite monitorizar hashtags, menciones y palabras clave.
- Requiere un token de acceso especial de la API de Twitter para el streaming.
- Es compatible con R 3.0.0 y superior.

<https://cran.r-project.org/package=streamR>



tidytext:

- Aunque no es específica para Twitter Mining, es una poderosa herramienta para la limpieza y análisis de texto en R.
- Se integra bien con rtweet para el análisis de contenido de tweets.
- Ofrece funciones para tokenización, eliminación de stop words, stemming y lematización, entre otras.
- Es compatible con R 3.1.0 y superior.

<https://cran.r-project.org/web/packages/tidytext/vignettes/tidytext.html>

quanteda:

- Similar a tidytext, ofrece funcionalidades específicas para el análisis de texto en R.
- Permite el análisis de sentimientos, la detección de temas y la construcción de modelos de tópicos para grandes volúmenes de texto.
- Es compatible con R 3.0.0 y superior.

<https://tutorials.quanteda.io/introduction/install/>



Caso práctico

Hernández Martínez, Robert. (2021). Análisis de Sentimientos en Twitter (actualización 2022). 10.13140/RG.2.2.27465.88169/1.

Las redes sociales han madurado al grado que los gobiernos están formulando políticas sobre cómo deben ser colectados, compartidos y usados estos datos. La minería de redes sociales es tema de interés para emprendedores, tecnólogos y hackers. Basta revisar el impacto de Twitter en temas políticos. Al trabajar en Analytics o Ciencia de Datos es evidente que los datos se generan todo el tiempo y cada vez con mayor rapidez. Los analistas están acostumbrados a trabajar con datos tabulares, generalmente numéricos; sin embargo, actualmente gran parte de los datos es no estructurado y con alto contenido de texto.

Caso práctico

Rodriguez Díaz, Karines & Guerra, Yamile. (2020). Análisis de sentimientos en Twitter aplicado al #impeachment de Donald Trump. Revista Mediterránea de Comunicación. 11. 10.14198/MEDCOM2020.11.2.23.

El presente artículo expone conclusiones sobre los sentimientos emitidos en Twitter con respecto al impeachment de Donald Trump. Asumimos un enfoque mixto, de tipo descriptivo, que permitió la obtención de resultados sobre aspectos emotivos asociados al discurso de usuarios de Estados Unidos en la esfera pública. La perspectiva investigativa utilizada reside en una estrategia transdisciplinar para entender los sentimientos relevantes de los tweets en el contexto de su producción. En el caso de estudio se aplican metodologías de minería de datos y sus tipologías: texto, grafo y multimedia, así como el análisis del discurso multimodal para la extracción y evaluación de valores afectivos respecto al juicio. Proponemos tener en cuenta todos los soportes, códigos y prácticas de producción discursiva en el microblogging: los textos escritos e hipertextos, enlaces sociales, reacciones, contenidos multimediales y las identidades de los usuarios en una relación sistémica de significados. Las publicaciones de las comunidades virtuales en las redes sociales sobre procesos políticos son de vital estudio en situaciones electorales y acontecimientos mediáticos de alto grado de polarización



Universidad
de la Ciudad de
Aguascalientes

Identificación de usuarios ansiosos para detectar condiciones de salud mental en la población tuitera de México

Identification of Anxious Users to Detect Mental Health Conditions in the Twitter Population of Mexico

Ana Miriam Romo Anaya, Víctor Silva Cuevas e Irving Gibrán Cabrera Zamora*

Basándonos en múltiples estudios que muestran que las palabras que usa una persona en sus escritos transmiten información sobre su personalidad, motivaciones, estado de ánimo y emocional e incluso su nivel económico, podemos entonces pretender que la información subyacente de los usuarios de Twitter permite generar indicadores del estado de ánimo e indicadores asociados con la prevalencia de condiciones de salud mental. Como primera aproximación, este trabajo se enfoca en identificar a quienes más expresan sentimientos negativos en sus textos acompañados de palabras relacionadas con ansiedad y que los hacen diferenciables de los demás. Para esto, usamos toda la muestra de tuits recolectados por el INEGI desde el 2017 al 2020 que se utilizó para la construcción del Índice del Estado de Ánimo de los Tuiteros en México junto con el recurso léxico LIWC

Based on multiple studies that show that the words a person uses in their writings convey information about their personality, motivations, moods, emotional states, and even economic level, we can claim that the underlying information of Twitter users allows us to generate mood indicators as well as indicators associated with the prevalence of mental health conditions. As a first approximation, this work focuses on identifying those who express more negative feelings in their texts accompanied by anxiety-related words and who differentiate them from others. For this we use the entire sample of tweets collected by INEGI from 2017 to 2020 that was used for the construction of the Mood Index of Twitter users in Mexico along with the lexical resource LIWC (Linguistic Inquiry Word Count) designed to analyze psychological implications of word use. This resource



Gráfica 9

Top ten de palabras relacionadas con ansiedad en septiembre del 2017 para entidades seleccionadas

Guerrero



Chiapas



Puebla



Tlaxcala

