Universidad
Tecmilenio®

Universidad Tecmilenio

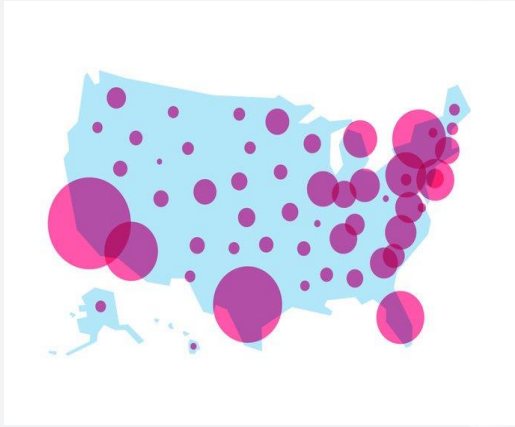
# Visualización y programación en Python

## Tema 7. Geolocalización

# Introducción

Recientemente, un poco antes de la pandemia por el Covid-19 se podían ver estadísticas de movilidad de las personas, contagios, muertes, etc. Además, ubicándolas geográficamente para que tuvieran un sentido más puntual y así poder analizar estrategias, acciones, impactos, incluso ayudando a conocer el comportamiento del virus en cada país.

**Con esto nos damos una idea de la utilidad de la geolocalización**, destacando la información que se puede visualizar más rápidamente y con más sentido en una gráfica que da idea de dónde proviene y utilizando indicadores con formas y colores para dar una idea más clara de la información.



## Geolocalización: propósito, posicionamiento y modos de recolección

La **geolocalización** es una tecnología que utiliza datos obtenidos de la computadora o dispositivo móvil para identificar o describir su ubicación física real, siendo una de las funcionalidades más populares actualmente.

Tabla 1. Propósitos de la geolocalización.

| Propósito | Descripción | Ejemplos |
|---|---|---|
| Georreferenciación o posicionamiento. | Determina la ubicación física de un objeto o una persona con relación a un sistema de coordenadas almacenándola para, posteriormente, acceder a información específica. | • Navegación vehicular.<br>• Vigilancia de prisioneros. |
| Geocodificación. | Busca información sobre objetos o servicios en un mapa. | • Localizar restaurantes de cierto tipo de comida.<br>• Ubicar un banco o negocio cercanos. |
| Geoetiquetado. | Agrega información geográfica a un objeto. | • Ubicación en fotografía.<br>• Ubicación en video.<br>• Ubicación en SMS. |

Fuente: Fox, P. (2021). *Geolocalización*. Recuperado de: https://es.khanacademy.org/computing/ap-computer-science-principles/x2d2f703b37b450a3:online-data-security/x2d2f703b37b450a3:user-data-tracking/a/geolocation

A ubicar un dispositivo y poder analizar la información con la que cuenta, se les llama **sistema de posicionamiento de dispositivos.**

Tabla 2. Sistemas de posicionamiento de dispositivos.

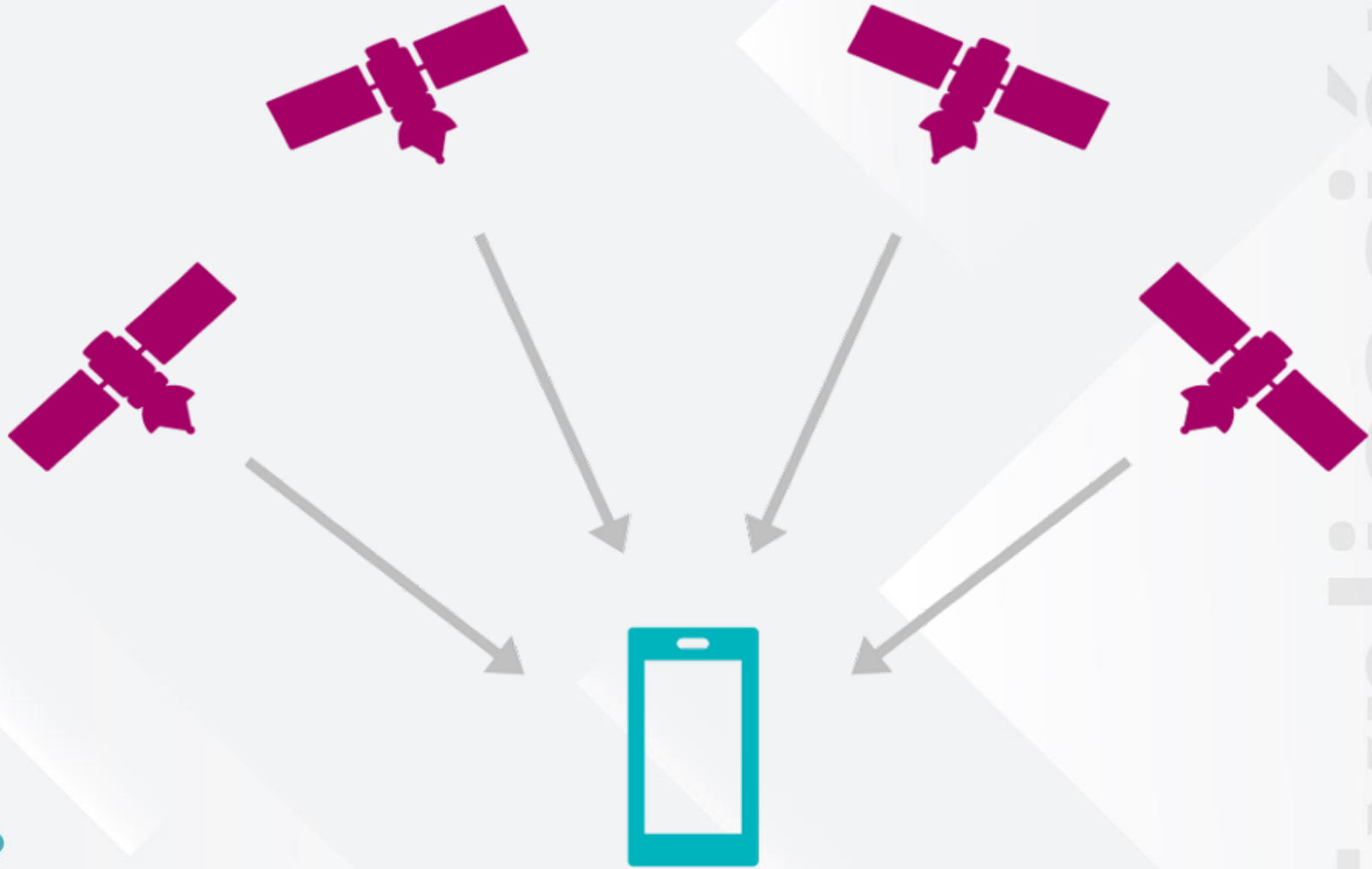
| Sistema de posicionamiento de dispositivos | Descripción | Características |
|---|---|---|
| **GPS (Sistema de Posicionamiento Global)** | Este tipo de posicionamiento nace a partir del gobierno de Estados Unidos, según menciona Fox (2021), en 1970 y ahora controla alrededor de 30 satélites GPS en órbita alrededor de la tierra.<br><br>Consisten en pequeños sensores con antenas que reciben señales de radio de los satélites GPS en la órbita del cielo, en la mayoría de los casos puede recibir señales de al menos cuatro satélites y el receptor puede calcular su posición mediante una técnica llamada trilateración. | El GPS funciona mejor al aire libre y con cielo despejado.<br><br>Un dispositivo inteligente puede normalmente registrar una geolocalización de 4.9 metros a cielo abierto, según Fox (2021).<br><br>También destaca que no funciona tan bien en interiores o entre edificios grandes por la interferencia y, sin embargo, sigue siendo muy utilizado. |
| **Wifi** | Un dispositivo con WiFi escanea puntos de acceso Wi-Fi y mide la fuerza de la señal de cada red.<br><br>También usa la trilateración, determinando la ubicación al buscar en una base de datos de WiFi, basándose en las ubicaciones encontradas, la fuerza de sus señales y la huella digital del dispositivo. | Funciona mejor en zonas urbanas con bastantes redes Wifi, contrario al GPS. |

| | | |
|---|---|---|
| **Trilateración de torre celular** | Estas torres celulares hacen que las redes celulares sean posibles.<br><br>Cada torre tiene tres arreglos de antenas direccionales en forma triangular y puede estimar la distancia entre la torre y el teléfono.<br><br>Entre más torres cercanas más exacta la ubicación, sin embargo, una sola torre puede ser suficiente para tener una ubicación aproximada. | Si un teléfono celular no tiene GPS para informar su ubicación se puede usar trilateración de torre celular. |
| **Basada en IP** | Cada que un dispositivo se conecta a internet tiene una dirección IP, que puede describir un lugar fijo en el mundo, es decir, es posible asignar IP a un área geográfica.<br><br>Se almacenan en bases de datos de geolocalización por IP creadas por empresas que toman como fuente registros regionales de direcciones IP, ubicaciones de usuarios en sitios web, datos de ISPs y estimaciones basadas en rutas de red. | Es el último recurso, pues es la más inexacta de todas las técnicas.<br><br>Dado que si un dispositivo accede desde una VPN (Red privada virtual) puede simular estar en otro lugar diferente al que realmente está. |

Fuente: Fox, P. (2021). *Geolocalización*. Recuperado de: https://es.khanacademy.org/computing/ap-computer-science-principles/x2d2f703b37b450a3:online-data-security/x2d2f703b37b450a3:user-data-tracking/a/geolocation

Explicación

Figura 1. Trilateración para tener la posición con GPS.



Fuente: Fox, P. (2021). *Geolocalización*. Recuperado de: https://es.khanacademy.org/computing/ap-computer-science-principles/x2d2f703b37b450a3:online-data-security/x2d2f703b37b450a3:user-data-tracking/a/geolocation

Explicación

Universidad Tecmilenio

Tabla 3. Modos de generación y recolección de datos de geolocalización.

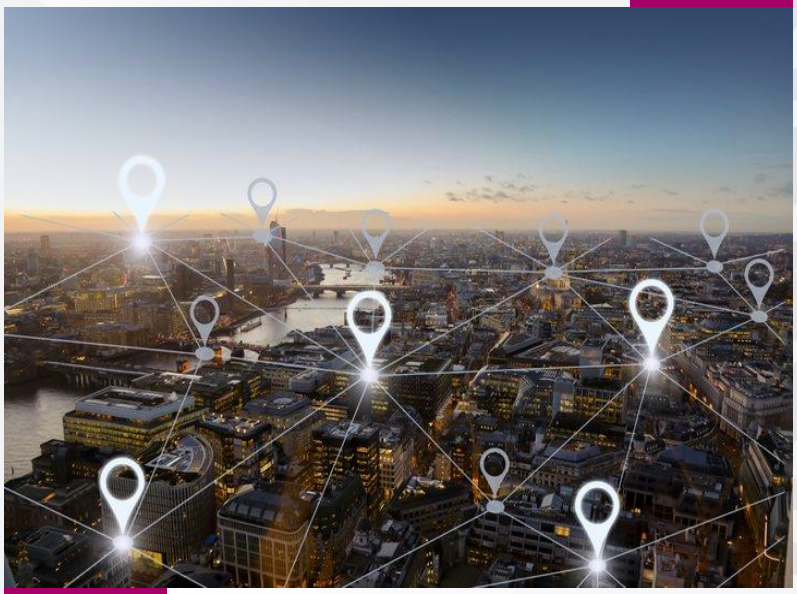
| Modo | Método de recolección | Tecnologías |
|---|---|---|
| **Activo: Basado en el dispositivo del usuario.** | • Utiliza el firmware y software del dispositivo inalámbrico del usuario.<br>• La ubicación se determina con GPS y/o triangulación mediante torres celulares.<br>• Modelo de petición-respuesta. | • GPS.<br>• GPS asistido (A-GPS).<br>• Wifi.<br>• 3G/G4.<br>• Aplicaciones móviles: iPhone, dispositivos Android, etc. |
| **Pasivo: búsqueda de datos en servidores.** | • Involucra proveedores de servicios de geolocalización como: Quova, NatGeo y Bering Media.<br>• Basado en la dirección IP.<br>• La correlación con las bases de datos de IP´s o SSID (de registros de compras, por ejemplo) y análisis de rutas de red (DNS). | • Ubicación IP.<br>• 3G/G4.<br>• Wifi. |

Fuente: EvaluandoSoftware. (2022). *Qué es la geolocalización y cómo funciona*. Recuperado de: https://www.evaluandosoftware.com/la-geolocalizacion-funciona/

Explicación

## **Usos** de la geolocalización

- Salud pública y emergencias
- Juegos
- Seguimiento de adultos mayores.
- Monitoreo de bienestar mental.
- Gestión de la cuarentena.
- Aplicaciones de chat basadas en la ubicación
- Gestión de densidad de multitudes.



En el lado empresarial también se está aprovechando esta funcionalidad en diferentes ámbitos, por ejemplo:

Seguimiento de activos y gestión de flotas:

Esto beneficia al sector logístico, por ejemplo, localizando vehículos y cargas para estimar su llegada a cierto lugar. Algunas de las aplicaciones son: Rhino Fleet Tracking, One Step GPS Fleet Tracking, ClearPathGPS.

Marketing (incentivos basados en proximidad):

Esto se trata de dar cupones a las personas que estén cercanas al negocio, que incluso se han trasladado al mercado digital y ya también propone nuevas tiendas que ofrecen lo mismo en menor costo tales como: Tada, Honey, Amazon Coupons, etc.

Explicación

Universidad Tecmilenio

## Ejemplos de geolocalización Python

Con la librería **geopy de Python** se puede hacer uso de varias funcionalidades para usar la localización. En este ejemplo con las coordenadas muestra la colonia, ciudad, estado y país de la ubicación:

Tabla 4. Usando la librería geopy para obtener información de una coordenada.

```
#Accederemos a información de una ubicación
#Paso 0: Cargar la librería para geolocalización
from geopy.geocoders import Nominatim
ap_nom = "mi localizador" # Necesario para identificar el acceso a Geopy
geolocator = Nominatim(user_agent=ap_nom)
#Obtener coordenadas en https://mylocation.org/
coordenadas= '25.66, -100.200' # Se colocan latitud y longitud
try:
    lugar = geolocator.reverse(coordenadas)
    print (lugar)
except: print("Revisa las coordenadas, deben tener un formato similar a este
: '3.0000,50.0000' ")
```

Fuente: GeoPy. (s.f.). GeoPy's documentation. https://geopy.readthedocs.io/en/stable/#nominatim
HRCOONECT. (s.f.). People Analytics: la geolocalización como herramienta para proteger a nuestros trabajadores https://www.hrconnect.cl/desarrollo/people-analytics-la-geolocalizacion-como-herramienta-para-proteger-a-nuestros-trabajadores/

Explicación

Figura 1. Mostrando los datos de la coordenada dada en el código.

```
Presa las Villitas, Acapulco, Guadalupe, Nuevo León, 67202, México
```

Al practicar este ejemplo, en el paso de coordenadas, entra al sitio https://mylocation.org/ para obtener las coordenadas de tu ubicación actual para que posteriormente se te desplieguen los datos de ese lugar.

Explicación

En este otro ejemplo, se ven los datos de ubicación que tenemos almacenados para cada tweet.m esto permite procesar las coordenadas de geolocalización y obtener aquellas ubicaciones con más publicaciones en los datos dados.

Tabla 5. Usando la librería geopy y pandas para los lugares con más tweets.

```
import pandas as pd

df = pd.read_csv('https://raw.githubusercontent.com/BBKdatasciencetaster/DS/
main/data/twitter_data_olio_UK_b.csv')
df.head()

df['place_name'].describe()

places = df.groupby(['place_name']).size().nlargest(10)
places

bchart = places.plot(kind='bar')
bchart.set_title('Tweeted locations')
bchart.set_ylabel('Count')
bchart

df.groupby(['place_type']).size()

df.loc[df['place_type'] == 'admin'].sample(10)
```

Fuente: Pandas. (s.f.) pandas.core.groupby.SeriesGroupBy.nlargest. https://pandas.pydata.org/docs/reference/api/pandas.core.groupby.SeriesGroupBy.nlargest.html

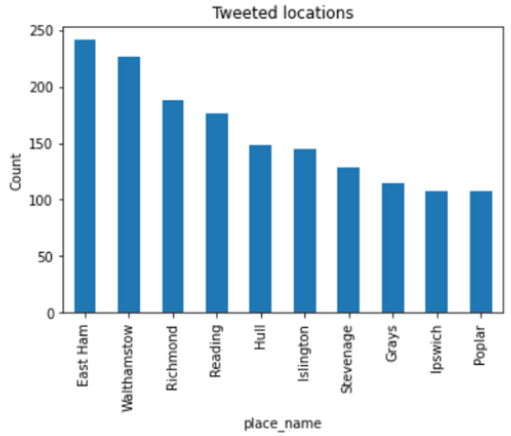
Explicación

Universidad Tecmilenio

Lo que ofrece de salida no solo los datos sino una gráfica que permite ver estos diez lugares más específicamente y sus cifras:

Figura 2. Mostrando la información de los lugares con más tweets.

| | date | _id | place_name | place_type | lat | lng | text |
|---|---|---|---|---|---|---|---|
| 963 | Tue Sep 29 10:20:24 +0000 2020 | 1310887144135766016 | South East | admin | 51.600193 | -1.263759 | "White thick toaste bread 800g" in #Didcot #Un... |
| 3380 | Thu Oct 01 08:00:47 +0000 2020 | 1311576781267308544 | West Midlands | admin | 52.366900 | -1.701310 | "Watercress" in #Solihull #UnitedKingdom https... |
| 465 | Tue Sep 29 16:26:28 +0000 2020 | 1310979266604142592 | London | admin | 51.671963 | -0.016852 | "Bacon roll " in #Enfield #UnitedKingdom https... |
| 3375 | Thu Oct 01 08:02:55 +0000 2020 | 1311577320776499205 | West Midlands | admin | 52.366900 | -1.701310 | "Ice ring doughnuts " in #Solihull #UnitedKing... |
| 4196 | Wed Sep 30 15:41:30 +0000 2020 | 1311330337717354497 | Bailiwick of Jersey | admin | 49.201000 | -2.143830 | "Tortilla chips" in #Jersey #UnitedKingdom htt... |
| 1022 | Tue Sep 29 10:00:28 +0000 2020 | 1310882125281996800 | South East | admin | 51.600193 | -1.263759 | "Fair trade bananas " in #Didcot #UnitedKingdo... |
| 485 | Tue Sep 29 16:18:57 +0000 2020 | 1310977737773406210 | London | admin | 51.671963 | -0.016852 | "Spinach pastry" in #Enfield #UnitedKingdom ht... |
| 2594 | Mon Sep 28 07:53:03 +0000 2020 | 1310487673686163457 | East | admin | 52.604352 | 1.172783 | "Cheese straws" in #Norwich #UnitedKingdom htt... |
| 3395 | Thu Oct 01 07:56:08 +0000 2020 | 1311575613405360128 | West Midlands | admin | 52.366900 | -1.701310 | "Romaine hearts lettuce " in #Solihull #United... |
| 3397 | Thu Oct 01 07:55:16 +0000 2020 | 1311575395536441345 | West Midlands | admin | 52.366900 | -1.701310 | "Butterhead salad " in #Solihull #UnitedKingdo... |



Universidad Tecmilenio

**Video**

**Objetivo:** Realizar un video corto de máximo 1 minuto donde expliques qué es la geolocalización, cómo hacer usos de ella con el lenguaje de programación Python junto con un ejemplo.

**Instrucciones:**

- Para realizar esta actividad es importante que consultes el contenido del tema.
- Con base en lo aprendido elabora un video corto que incluya los conceptos más relevantes y significativos sobre el uso y aplicación organizacional de la geolocalización
- Debe ser un video didáctico, creativo e interesante.
- Utiliza música, elementos gráficos y tipografías para llamar la atención.
- Máximo 1 minuto.

Actividad

Universidad Tecmilenio

En este notebook de Google colab puedes ver todo el código desarrollado en el tema:

https://colab.research.google.com/drive/1u-W4Cg2ddaDoq7_Ghp8zuncOO3B-EzxC?usp=sharing

Explicación

Es evidente que contar con la **geolocalización** es una ventaja para una gran variedad de objetivos. Específicamente en la **analítica de personas o people analytics**, también se pueden tener diferentes usos:

- Seguimiento de asistencia.
- Análisis de trabajo remoto.
- Análisis de ventas y servicio de campo.
- Programación de turnos.
- Análisis de compromiso de los empleados.
- Análisis de contratación.
- Análisis de desplazamiento.

La integración de los datos de geolocalización con el análisis de personas puede dar una imagen más completa del comportamiento y el rendimiento de la fuerza laboral.

# Referencias


EvaluandoSoftware. (2022). *Qué es la geolocalización y cómo funciona*. Recuperado de https://www.evaluandosoftware.com/la-geolocalizacion-funciona/

Fox, P. (2021). *Geolocalización*. Recuperado de: https://es.khanacademy.org/computing/ap-computer-science-principles/x2d2f703b37b450a3:online-data-security/x2d2f703b37b450a3:user-data-tracking/a/geolocation

GeoPy. (s.f.). *GeoPy's documentation*. Recuperado de https://geopy.readthedocs.io/en/stable/#nominatim

Pandas. (s.f.) *pandas.core.groupby.SeriesGroupBy.nlargest*. Recuperado de https://pandas.pydata.org/docs/reference/api/pandas.core.groupby.SeriesGroupBy.nlargest.html


Referencias

Universidad Tecmilenio

# Visualización y programación en Python

## Tema 8.
## Análisis de sentimientos

# Introducción

El origen del análisis de sentimientos se remonta a 1950 según, en donde se analizaban documentos en papel para ver si **las opiniones de las personas eran positivas, neutrales o negativas.**

Posteriormente surge, con la digitalización; ésta área que quiere saber cómo obtener valor de toda esta información escrita para que al analizarla se puedan **tomar mejores decisiones** y llevar a cabo las acciones en beneficio de la organización, las personas y la sociedad.

Universidad Tecmilenio

## ¿Qué es el análisis de sentimientos?

También conocido como **minería de opiniones**, es un campo del NLP que utiliza técnicas computacionales para extraer y comprender las actitudes, opiniones y emociones de las personas a partir de datos de texto basado en modelos que se entrenan con datos para identificar qué emoción se puede transmitir en un texto.

Su objetivo es determinar el sentimiento general de un texto, ya sea positivo, negativo o neutral. Esto generalmente se hace analizando las palabras y frases utilizadas en el texto e identificando patrones que indican un sentimiento particular.

Explicación

Universidad Tecmilenio

Figura 1. Análisis de sentimientos.

Hoy en día, el análisis de sentimientos se usa para **extraer información subjetiva del contenido de Internet.**

**Incluidos** textos, tweets, blogs, redes sociales, artículos de noticias, reseñas y comentarios.

Explicación

Tabla 1. Beneficios que brinda el análisis de sentimientos.

| Tipo de análisis | Beneficios |
|---|---|
| **Análisis de sentimientos en redes sociales** | Ayuda a analizar la intención de la audiencia y las opiniones expresadas en varias plataformas sociales.<br>Puede obtener un análisis de mercado granular de los gustos y disgustos de los clientes sobre productos, marcas, contenido publicitario y más a través de técnicas como la escucha social de TikTok y la escucha social de Instagram.<br>Puede brindarle información muy necesaria sobre productos, su grupo demográfico objetivo, los temas comunes en los comentarios y su comparación en diferentes plataformas sociales, y más. |
| **Perspectivas de experiencia de marca** | Puede darle una idea del sentimiento oculto del mercado sobre una marca y lo que los clientes esperan de ella.<br>Permite comprender los problemas de brechas de mercado, retención de clientes, crear una base de clientes leales y aumentar las conversiones de ventas. |
| **Perspectivas de un paciente** | El análisis de la voz de un paciente por ejemplo, se utiliza para medir la eficiencia de la prestación de atención médica, descubrir cualquier brecha en el servicio ambulatorio y hospitalario, suavizar las operaciones en las farmacias, etc.<br>Las encuestas de pacientes y cuidadores ayudan a las organizaciones de atención médica a conocer las necesidades cambiantes de los pacientes, mejorar la atención crítica y la atención de hospicio, así como brindar atención médica en áreas que carecen de servicios primarios. |

Explicación

# Explicación

| | |
|---|---|
| **Mejorar el servicio al cliente** | Historiales de chatbots de minería de emociones, transcripciones de llamadas de servicio al cliente, correos electrónicos de quejas de clientes, comentarios de clientes de devoluciones y reembolsos y encuestas de clientes, todos pueden aprovecharse para mejorar la satisfacción del cliente y ayudarlo a desarrollar una experiencia brillante para el cliente. |
| **Información multilingüe** | Utilizar para analizar los datos de los clientes en distintas geografías, etnias y culturas para obtener los conocimientos más intrínsecos sobre la popularidad del producto, las marcas competitivas, las tasas de éxito de las campañas publicitarias y más.<br>Útil en el comercio electrónico y la hospitalidad de una manera especial porque ambas industrias a menudo tienen clientes de diferentes orígenes, especialmente cuando las audiencias se encuentran en ciudades cosmopolitas. |
| **Análisis de tendencias de noticias** | Extrae tendencias emergentes de sitios web de noticias, videos, artículos y revistas, plataformas en línea que incluyen blogs, redes sociales como Twitter y Facebook, y otras fuentes para anticipar el comportamiento del mercado. Esto incluye noticias sobre temas de actualidad, como nuevos escenarios políticos, comercio internacional y movimientos de acciones de empresas.<br>Industrias como la banca, los seguros, los bienes raíces, la automotriz, la cosmética, etc. utilizan el análisis de sentimientos basado en mercado libre para comprender y analizar dichas noticias con el fin de especular, planificar y estar preparados para cualquier situación. Esto incluye la planificación de la cadena de suministro, la gestión de las relaciones públicas, la modificación de los lanzamientos de nuevos productos y otras funciones operativas y de marketing. |
| **Perspectivas de sentimientos en tiempo real** | Se utiliza para analizar la participación y satisfacción de la audiencia durante espectáculos en vivo, eventos corporativos en vivo, seminarios, eventos promocionales como exhibiciones comerciales y de autos, transmisiones de radio en vivo y otros. |

Explicación

# Explicación

| | |
|---|---|
| **Comentarios de clientes** | Análisis de datos recopilados de plataformas de revisión de clientes como Reddit, OpenTable, Trustpilot, GoogleMyBusiness, Amazon y otras, las empresas pueden obtener información crítica sobre los gustos y disgustos de los productos, los puntos débiles de los clientes y las deficiencias de productos y servicios. |
| **Gestión de reputación en línea** | Al ser automatizado e impulsado por la inteligenica artificial (IA) ayuda a las empresas a administrar la reputación de su marca al permitirles tomar decisiones oportunas sobre cómo responder a las menciones negativas de la marca y, por lo tanto, evitar riesgos . Por ejemplo, al configurar alertas en la plataforma de análisis de sentimientos, puede recibir una notificación inmediata de una mención negativa en las plataformas que está rastreando. Por lo tanto, pueden abordar una situación antes de que se convierta en un problema mayor. Esto también lo ayuda a atraer mejores candidatos y clientes, aumentar su base de clientes y catapultar las ventas. |
| **Información de marketing basada en datos** | No solo de métricas cualitativas como: número de ¨me gusta¨, seguidores, compartidos, etc., sino también de métricas cuantitativas obtenidas del análisis de comentarios como: porcentaje de sentimiento positivo y negativo sobre varios aspectos de un negocio, estrategias de marketing, etc., tomando esto en cuenta se tiene una mayor probabilidad de éxito. |
| **Desarrollo de contenido publicitario relacionado** | Saber qué tipo de contenido resuena más con el público objetivo es crucial para desarrollar las estrategias y campañas publicitarias adecuadas. |

Explicación

# Explicación

| | |
|---|---|
| **Desarrollo de productos** | Puede brindar información inteligente sobre cómo puede mejorarse el producto para que llegue a más audiencias. Esto también es muy importante al lanzar un nuevo producto, abrir una tienda en una nueva ubicación, cambiar los modelos comerciales, etc. |
| **Análisis de competencia** | Ayudan a monitorear y rastrear las tendencias cambiantes en la industria para que pueda relacionarse con otras audiencias para la longevidad del negocio. |
| **Perspectivas de la experiencia de empleado** | Aumentar el compromiso y la satisfacción de los empleados para mejorar la productividad de la fuerza laboral.<br>Las organizaciones han comenzado a darse cuenta de que la fatiga y la soledad de los empleados, que experimenta el 41 % de los empleados solo en los EE. UU., pueden tener graves repercusiones en los objetivos comerciales.<br>Basado en el mercado libre para medir la experiencia de los empleados y poder usar los datos para crear mejores entornos de trabajo, fomentar una cultura laboral más humanizada y mejorar los programas de capacitación y reconocimiento de los empleados. |

Fuente: Repustate. (s.f.). *Benefits of Sentiment Analysis for Impactful Growth*. Recuperado de: https://www.repustate.com/blog/sentiment-analysis-benefits/

Explicación

Universidad Tecmilenio

Existen varias librerías en Python que permiten trabajar con análisis de sentimientos y con forme va pasando el tiempo se van creando cada vez más, tales como:

Tabla 2. Principales librerías de análisis de sentimientos.

| Librería | Características |
|---|---|
| **Pattern** | • Maneja NLP, minería de datos, análisis de red, aprendizaje automático y visualización.<br>• Incluye la búsqueda de superlativos y comparativos. También puede llevar a cabo la detección de hechos y opiniones.<br>• Devuelve la polaridad y la subjetividad de un texto dado.<br>• Maneja un rango de polaridad de muy positivo a muy negativo. |
| **VADER** | • Preconstruida de un analizador de sentimientos de fuente abierta basada en reglas/léxico.<br>• Diseñada específicamente para los sentimientos expresados en las redes sociales y utiliza una combinación de un léxico de sentimientos y una lista de características que generalmente se etiquetan según su orientación semántica como positivas o negativas.<br>• Calcula el sentimiento del texto y devuelve la probabilidad de que una oración de entrada determinada sea positiva, negativa o neutra.<br>• No requiere datos de entrenamiento. |
| **BERT** | • Es un modelo superior de aprendizaje automático utilizado para tareas de NLP, incluido el análisis de sentimientos.<br>• Desarrollado por Google.<br>• Se entrenó con WIkipedia en inglés y BooksCorpus.<br>• Tiene una mejor capacidad para comprender el lenguaje y aprender la variabilidad en los patrones de datos. |

Explicación

Universidad Tecmilenio

| | |
|---|---|
| **TextBlob** | • Simple y admite operaciones y análisis complejos en datos textuales.<br>• Define un sentimiento por su orientación semántica y la intensidad de cada palabra en una oración, lo que requiere un diccionario predefinido que clasifique las palabras negativas y positivas.<br>• Asigna puntajes individuales a todas las palabras y se calcula un sentimiento final.<br>• Devuelve la polaridad y la subjetividad de una oración, con un rango de polaridad de negativo a positivo.<br>• Las etiquetas semánticas de la biblioteca ayudan con el análisis, incluidos emoticones, signos de exclamación, emojis y más. |
| **spaCy** | • Permite a los desarrolladores crear aplicaciones que pueden procesar y comprender volúmenes masivos de texto.<br>• Se utiliza para construir sistemas de comprensión del lenguaje natural y sistemas de extracción de información.<br>• Permite a los desarrolladores crear aplicaciones que pueden procesar y comprender volúmenes masivos de texto, y se utiliza para construir sistemas de comprensión del lenguaje natural y sistemas de extracción de información. |
| **coreNLP** | • Incorpora herramientas de NLP de Stanford, incluido el análisis de sentimientos.<br>• Admite cinco idiomas en total: inglés, árabe, alemán, chino, francés y español.<br>• Analiza el texto agregando "sentimiento".<br>• Soporte de línea de comando y soporte de entrenamiento de modelos. |
| **scikit-learn** | • Es especialmente útil para los algoritmos clásicos de aprendizaje automático, como los que se usan para la detección de spam y el reconocimiento de imágenes.<br>• Se puede usar para tareas de NLP, incluido el análisis de sentimientos al entrenar un modelo que puede generar si el texto es positivo o negativo.<br>• Proporciona varios vectorizadores para traducir los documentos de entrada en vectores de características y viene con varios clasificadores diferentes ya integrados.<br>• Utilizado por grandes empresas como Spotify.<br>• Basado en SciPy y NumPy. |

Explicación

Universidad Tecmilenio

| | |
|---|---|
| **Polyglot** | • Se basa en Numpy y es increíblemente rápida al tiempo que ofrece una gran variedad de comandos.<br>• Admite el análisis de sentimientos para 136 idiomas.<br>• Es conocido por su eficiencia, velocidad y sencillez.<br>• Se elige para proyectos que involucran idiomas que no son compatibles con spaCy. |
| **PyTorch** | • Creada por el equipo de investigación de inteligencia artificial de Facebook.<br>• Permite realizar muchas aplicaciones diferentes, incluido el análisis de sentimientos, donde puede detectar si una oración es positiva o negativa.<br>• Tiene una ejecución extremadamente rápida y se puede operar en procesadores simplificados o CPU y GPU.<br>• Plataforma y ecosistema en la nube. |
| **Flair** | • NLP simple de código abierto.<br>• Se basa en PyTorch y el equipo de investigación detrás de Flair ha lanzado varios modelos preentrenados para una variedad de tareas.<br>• Uno de los modelos preentrenados es un modelo de análisis de sentimientos entrenado en un conjunto de datos de IMDB (Internet Movie Database) y es fácil de cargar y hacer predicciones.<br>• Puede entrenar un clasificador usando su conjunto de datos.<br>• Admite varios idiomas. |

Fuente: Farland, A. (2022). *10 Best Python Libraries for Sentiment Analysis*. Recuperado de https://www.unite.ai/10-best-python-libraries-for-sentiment-analysis/

Explicación

Universidad Tecmilenio

Tabla 3. Realizando análisis de sentimiento con la librería VADER.

```
Jennifer Aniston is smart, gorgeous, and funny.------------------ {'neg': 0.0, 'neu': 0.294, 'pos': 0.706, 'compound': 0.8625}
Cameron Diaz is smart, gorgeous, and funny!----------------------- {'neg': 0.0, 'neu': 0.288, 'pos': 0.712, 'compound': 0.8718}
Oben Wilson is very smart, handsome, and funny.------------------- {'neg': 0.0, 'neu': 0.348, 'pos': 0.652, 'compound': 0.8545}
Will Smith is VERY SMART, handsome, and FUNNY.-------------------- {'neg': 0.0, 'neu': 0.29, 'pos': 0.71, 'compound': 0.9227}
Adam Sandler is VERY SMART, handsome, and FUNNY!!!---------------- {'neg': 0.0, 'neu': 0.276, 'pos': 0.724, 'compound': 0.9342}
Jim Carrey is VERY SMART, uber handsome, and FRIGGIN FUNNY!!!---- {'neg': 0.0, 'neu': 0.327, 'pos': 0.673, 'compound': 0.9469}
Will Ferrell is not smart, handsome, nor funny.------------------- {'neg': 0.593, 'neu': 0.407, 'pos': 0.0, 'compound': -0.7424}
The book was great.----------------------------------------------- {'neg': 0.0, 'neu': 0.423, 'pos': 0.577, 'compound': 0.6249}
At least it isn't a horrible meal.-------------------------------- {'neg': 0.0, 'neu': 0.678, 'pos': 0.322, 'compound': 0.431}
The book was only kind of good.----------------------------------- {'neg': 0.0, 'neu': 0.697, 'pos': 0.303, 'compound': 0.3832}
The plot was good, but the characters are uncompelling and the dialog is not great. {'neg': 0.327, 'neu': 0.579, 'pos': 0.094, 'compound': -0.7042}
Today SUX!-------------------------------------------------------- {'neg': 0.779, 'neu': 0.221, 'pos': 0.0, 'compound': -0.5461}
Today only kinda sux! But I'll get by, lol------------------------ {'neg': 0.127, 'neu': 0.556, 'pos': 0.317, 'compound': 0.5249}
Make sure you :) or :D today!------------------------------------- {'neg': 0.0, 'neu': 0.294, 'pos': 0.706, 'compound': 0.8633}
Catch utf-8 emoji such as such as 💘 and 💋 and 😁------------------ {'neg': 0.0, 'neu': 0.615, 'pos': 0.385, 'compound': 0.875}
Not bad at all---------------------------------------------------- {'neg': 0.0, 'neu': 0.513, 'pos': 0.487, 'compound': 0.431}
```

Fuente: Farland, A. (2022). *10 Best Python Libraries for Sentiment Analysis*. Recuperado de: https://www.unite.ai/10-best-python-libraries-for-sentiment-analysis/

Figura 1. Salida de ejecución de código de la librería VADER.

```
pip install vaderSentiment #se instala la librería
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer #se carga
 la libreria
frases = ["Jennifer Aniston is smart, gorgeous, and funny.",  # positive sente
nce example
            "Cameron Diaz is smart, gorgeous, and funny!",  # punctuation emp
hasis handled correctly (sentiment intensity adjusted)
            "Oben Wilson is very smart, handsome, and funny.", # booster word
s handled correctly (sentiment intensity adjusted)
            "Will Smith is VERY SMART, handsome, and FUNNY.",  # emphasis for
 ALLCAPS handled
            "Adam Sandler is VERY SMART, handsome, and FUNNY!!!", # combinati
on of signals - VADER appropriately adjusts intensity
            "Jim Carrey is VERY SMART, uber handsome, and FRIGGIN FUNNY!!!",
# booster words & punctuation make this close to ceiling for score
            "Will Ferrell is not smart, handsome, nor funny.",  # negation se
ntence example
            "The book was great.",  # positive sentence
            "At least it isn't a horrible meal.",  # negated negative sentenc
e with contraction
            "The book was only kind of good.", # qualified positive sentence
is handled correctly (intensity adjusted)
            "The plot was good, but the characters are uncompelling and the d
ialog is not great.", # mixed negation sentence
            "Today SUX!",  # negative slang with capitalization emphasis
            "Today only kinda sux! But I'll get by, lol", # mixed sentiment e
xample with slang and constrastive conjunction "but"
            "Make sure you :) or :D today!",  # emoticons handled
            "Catch utf-
8 emoji such as such as 🙈 and 🤘 and 😁",  # emojis handled
            "Not bad at all"  # Capitalized negation
            ]

analizador = SentimentIntensityAnalyzer() #se agrega la instancia del analinza
dor
for frase in frases: #en un ciclo se evalua cada frase
    salida = analizador.polarity_scores(frase) #se obtiene la salida de cada f
rase
    print("{:-
<65} {}".format(frase, str(salida))) #se le da formato y se imprime la salida
de cada frase
```

Universidad Tecmilenio

Es importante notar que la **librería VADER no requiere de entrenamiento previo** y por ello puede identificar sin que le digamos qué significa positivo o negativo. Otras las librerías sí lo requieren.

En este notebook de Google colab puedes ver todo el código desarrollado en el tema:

https://colab.research.google.com/drive/1IPIJsCt9lcAOO_79ALP8p9C7j1OwDPac?usp=sharing

Explicación

Universidad Tecmilenio

## Ejercicio práctico

**Objetivo:** Ejecutar un análisis de sentimientos con base al ejemplo proporcionado.

**Instrucciones:**

- Para realizar esta actividad es importante que consultes el contenido del tema.
- Con base en el ejemplo de curso, ejecuta un análisis de sentimientos con los datos de la organización a la que perteneces
- Incluye las bibliotecas de análisis de sentimientos que te resulten mas adecuadas a tu ejemplo
- Utiliza la plataforma python para realizar tu actividad.
- Incluye una reflexión sobre el proceso de ejecución y manejo de la plataforma para llevar a cabo tu actividad.
- Una cuartilla máximo.

Actividad

El **análisis de sentimientos** es uno de los software más poderosos y valiosos que se pueden tener para aportar a la organización. Con la analítica de empleados se puede usar para obtener información sobre:

1. Análisis de compromiso de los empleados.
2. Predicción de la rotación.
3. Sentimiento de los empleados hacia la cultura de la empresa.
4. Sentimiento de los empleados hacia el liderazgo.
5. Identificación de áreas de oportunidad.
6. Sentimiento de los empleados hacia la visión de la empresa.
7. Análisis de diversidad e inclusión.
8. Sentimiento de los empleados hacia los valores de la empresa.

# Referencias


Farland, A. (2022). *10 Best Python Libraries for Sentiment Analysis*. Recuperado de: https://www.unite.ai/10-best-python-libraries-for-sentiment-analysis//

Repustate. (s.f.). *Benefits of Sentiment Analysis for Impactful Growth*. Recuperado de: https://www.repustate.com/blog/sentiment-analysis-benefits/


Referencias

La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECMILENIO), protegida por la Ley Federal de Derecho de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor.

El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECMILENIO.

Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECMILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.