

---

## PROYECTO 3

---

**202209622 1 – Pablo Isai Matusalen Cutzal Mazariegos**

### Resumen

This essay describes the development of a sentiment analysis tool in social networks for the company Tecnologías Chapinas, S.A. By means of a predefined dictionary of words associated with positive, negative and neutral feelings, a classification of the messages published on social networks is implemented. The tool uses advanced XML processing techniques, object-oriented programming (OOP), and MVT and MVC design patterns, in Django and Flask, respectively. The results of the analysis are stored in an XML file for consultation and reporting. The application provides a comprehensive approach to monitoring and evaluating public perception of companies and their services, allowing information to be managed efficiently. As a result, the company obtains useful information about customer satisfaction and its areas of opportunity, promoting better service and adaptation to market needs.

### Palabras clave

Sentiment Analysis, XML, API, Social Media, Flask

*Traducir al idioma inglés, el resumen redactado en la columna de la izquierda.*

*La traducción debe ser revisada con un profesional en ingeniería con amplios conocimientos del idioma inglés, en caso que en forma personal no se posean.*

*Evitar la utilización del traductor de google u otra similar.*

*El abstract y las keywords deben abarcar solamente esta columna.*

### Keywords

*Traducción al idioma inglés de las palabras clave.*

### Abstract

## Introducción

En la actualidad, la gestión de la reputación en redes sociales se ha vuelto fundamental para empresas que buscan conocer la percepción del público sobre sus servicios. Tecnologías Chapinas, S.A. ha decidido implementar una herramienta capaz de analizar los mensajes publicados en redes sociales y evaluar el sentimiento de los usuarios hacia la empresa y sus servicios. Esta herramienta, desarrollada con el lenguaje de programación Python y los frameworks Django y Flask, procesa archivos XML para clasificar los mensajes en positivos, negativos y neutros, con base en un diccionario de palabras clave. La solución propuesta es de gran relevancia, ya que permite a la empresa monitorear de forma automatizada la satisfacción del cliente y ajustar su estrategia de servicio.

Debe contener un máximo de 150 palabras.

## Desarrollo del tema

La herramienta de análisis de sentimientos para redes sociales desarrollada para Tecnologías Chapinas, S.A. se compone de varios elementos que trabajan en conjunto para cumplir con el objetivo de identificar y analizar la percepción del público sobre la empresa y sus servicios. Los principales componentes son el procesamiento de sentimientos, la persistencia de datos en XML, el diseño y consumo de un API HTTP y la visualización de resultados a través de una interfaz web. Cada uno de estos elementos incorpora principios de programación orientada a objetos (POO), junto con patrones de diseño de software, para asegurar una solución eficiente y modular.

### Procesamiento de Sentimientos

La clasificación de sentimientos se lleva a cabo utilizando un diccionario de palabras positivas y negativas en formato XML. Este diccionario sirve como una referencia para identificar la naturaleza del sentimiento expresado en los mensajes de redes sociales. Por ejemplo, palabras como "bueno", "excelente" o "satisfecho" clasifican un mensaje como positivo, mientras que palabras como "malo", "pésimo" o "decepcionado" lo clasifican como negativo. Si un mensaje contiene tanto palabras positivas como negativas, o ninguna de las dos, se clasifica como "neutro". Este enfoque permite una clasificación simple pero efectiva, que puede ajustarse y personalizarse fácilmente agregando o eliminando palabras en el diccionario XML.

La herramienta utiliza expresiones regulares para extraer contenido específico de los mensajes. Se busca capturar información clave como la fecha, el nombre del usuario, la red social de origen y el contenido del mensaje. Las expresiones regulares permiten normalizar estos datos, eliminando posibles errores causados por variaciones en la estructura del mensaje. Este procesamiento inicial es crucial para que los datos sean consistentes y se integren correctamente en el sistema de análisis.

### Persistencia de Datos en XML

Los mensajes analizados y sus clasificaciones se almacenan en un archivo XML, que actúa como una base de datos estructurada para la herramienta. La elección de XML como formato de almacenamiento permite organizar los datos en una estructura jerárquica, lo cual facilita la realización de consultas y la generación de reportes. La estructura XML incluye nodos para cada mensaje, con información sobre el sentimiento, la empresa mencionada y el servicio relacionado. También incluye un resumen general que muestra el total de mensajes clasificados por sentimiento y la frecuencia de menciones para cada empresa y servicio.

Además, el sistema actualiza el archivo XML cada vez que se procesan nuevos mensajes, asegurando

que la información esté siempre disponible para futuras consultas. Esta persistencia de datos garantiza que el historial de sentimientos y menciones de la empresa esté siempre a la mano para su análisis, lo que permite monitorear tendencias a lo largo del tiempo.

API HTTP

Para integrar la herramienta de análisis con otros sistemas y facilitar el envío de mensajes para su clasificación, se implementa un API HTTP utilizando Flask. Este API permite a otros sistemas o aplicaciones enviar mensajes en formato XML, que luego son procesados y almacenados por la herramienta. El API está diseñado siguiendo principios de REST, lo que facilita la interacción con el sistema a través de métodos estándar como POST (para enviar mensajes) y GET (para consultar datos procesados).

El API también admite consultas específicas, como la recuperación de análisis por fecha o la filtración de resultados para una empresa y servicio determinados. Esto permite a los usuarios acceder a información precisa y personalizada sobre el sentimiento de los clientes hacia la empresa en un momento determinado. Además, el API es seguro y cuenta con validaciones para evitar el procesamiento de datos inválidos o incompletos.

Visualización de Resultados

La interfaz de usuario, desarrollada en Django, permite a los usuarios cargar archivos XML, consultar los datos almacenados y visualizar los resultados del análisis de sentimiento de manera gráfica. Utilizando el patrón Modelo-Vista-Template (MVT), Django ofrece una interfaz organizada y accesible que facilita el uso de la herramienta por parte de los analistas de Tecnologías Chapinas, S.A.

La interfaz permite la selección de fechas específicas o rangos de tiempo para filtrar los mensajes procesados, generando gráficos de barras o pastel que representan la proporción de sentimientos (positivo,

negativo y neutro) y su distribución por empresas y servicios. Esto resulta especialmente útil para identificar cambios en la percepción de los usuarios a lo largo del tiempo. También se pueden generar reportes en PDF, que incluyen un resumen detallado de los mensajes clasificados, junto con los gráficos, lo cual facilita la presentación de la información en reuniones o análisis posteriores.

Escalabilidad y Modularidad

La herramienta está diseñada con principios de programación orientada a objetos (POO) y sigue una arquitectura modular que facilita su mantenimiento y escalabilidad. Cada componente de la herramienta (procesamiento de sentimientos, almacenamiento de datos, API y visualización) está implementado como una unidad independiente. Esto permite que cualquier cambio o mejora en una sección específica no afecte el funcionamiento del resto del sistema.

Por ejemplo, si Tecnologías Chapinas, S.A. decide expandir el análisis de sentimientos para incluir redes sociales adicionales o ajustar el diccionario de sentimientos, puede hacerlo sin modificar la estructura general de la herramienta. Asimismo, la arquitectura modular permite que la herramienta se adapte fácilmente a futuros cambios en los requisitos empresariales, como la adición de nuevas empresas o servicios en el análisis.

ARQUITECTURA

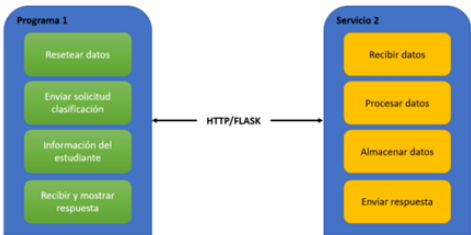


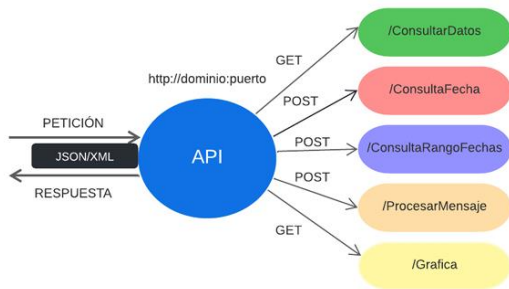
Fig. 1 – Arquitectura general de la aplicación

Figura 1. Título o descripción breve de la figura.

Fuente: elaboración propia, o citar al autor, año y página.

Tabla I.

El título de la tabla debe ser corto y conciso.



Fuente: elaboración propia, o citar al autor, año y página.

Es conveniente describir brevemente el contenido de una tabla, evitando los aspectos obvios.

En el caso de inclusión de fórmulas, éstas deben elaborarse utilizando el editor de ecuaciones disponible en Word, indicando el significado de cada una de las variables o parámetros que se incluyen.

Deben enumerarse entre paréntesis para poder hacer referencia de esta. Por ejemplo, un modelo de crecimiento exponencial

$$y = y_0 e^{kt} \quad (1)$$

donde:

y = cantidad presente en el tiempo t

y<sub>0</sub> = cantidad presente al inicio de la observación

k = tasa específica de crecimiento

t = periodo de tiempo (años, minutos, otros)

## Conclusiones

El desarrollo de esta herramienta proporciona a Tecnologías Chapinas, S.A. una ventaja competitiva al permitirle conocer el sentimiento de los usuarios en tiempo real. La integración de tecnologías como XML, Flask y Django ofrece una solución completa para la gestión de la reputación en redes sociales. Este sistema automatizado facilita el análisis y permite a

la empresa adaptarse rápidamente a las necesidades del mercado, mejorando así su relación con los clientes y optimizando sus servicios.

## Referencias bibliográficas

C. J. Date, (1991). *An Introduction to Database Systems*. Addison-Wesley Publishing Company, Inc.