
PROYECTO 3 – HERRAMIENTA DE ANÁLISIS DE CONTENIDO DE REDES SOCIALES

202300596 – Xavi Alexander De León Perdomo

Resumen

El proyecto aborda el desarrollo de una solución tecnológica para el análisis de sentimientos en redes sociales, específicamente para la empresa Tecnologías Chapinas, S.A. La solución comprende una aplicación web desarrollada en Django (frontend) y una API REST implementada con Flask (backend), utilizando programación orientada a objetos en Python y bases de datos XML.

La innovación radica en la capacidad de analizar mensajes de redes sociales para determinar el sentimiento (positivo, negativo o neutro) de los usuarios hacia empresas y sus servicios, mediante un diccionario de palabras clave. El sistema procesa mensajes estructurados que incluyen información sobre lugar, fecha, usuario y red social, generando reportes estadísticos detallados.

El impacto técnico se refleja en la implementación de estructuras de datos abstractas y el uso de tecnologías modernas como Django y Flask. A nivel económico y social, permite a las empresas monitorear la percepción de sus servicios en redes sociales, facilitando la toma de decisiones estratégicas basadas en el feedback de los usuarios.

El proyecto enfatiza la importancia del versionamiento mediante GitHub, la documentación exhaustiva y la adherencia a estándares de desarrollo, asegurando una solución robusta y mantenible.

Palabras clave

análisis de sentimientos, redes sociales, Python, Django, Flask, XML, API REST.

Abstract

The project focuses on developing a technological solution for sentiment analysis in social media, specifically for Tecnologías Chapinas, S.A. The solution comprises a web application developed in Django (frontend) and a REST API implemented with Flask (backend), using object-oriented programming in Python and XML databases.

The innovation lies in the ability to analyze social media messages to determine user sentiment (positive, negative, or neutral) towards companies and their services through a keyword dictionary. The system processes structured messages containing information about location, date, user, and social network, generating detailed statistical reports.

The technical impact is reflected in the implementation of abstract data structures and the use of modern technologies like Django and Flask. At the economic and social level, it allows companies to monitor the perception of their services on social media, facilitating strategic decision-making based on user feedback.

The project emphasizes the importance of versioning through GitHub, comprehensive documentation, and adherence to development standards, ensuring a robust and maintainable solution.

Keywords

sentiment analysis, social media, Python, Django, Flask, XML, REST API.

Introducción

En la era digital actual, donde las redes sociales se han convertido en una plataforma fundamental de comunicación entre empresas y usuarios, la capacidad de analizar y comprender el sentimiento público hacia productos y servicios se ha vuelto crucial para el éxito empresarial. Este proyecto aborda esta necesidad mediante el desarrollo de una solución tecnológica integral para Tecnologías Chapinas, S.A., que permite analizar automáticamente el sentimiento expresado en mensajes de redes sociales.

¿Cómo pueden las empresas procesar eficientemente grandes volúmenes de feedback social para mejorar sus servicios? ¿De qué manera se puede cuantificar objetivamente el sentimiento general hacia una marca o servicio específico? Estas preguntas fundamentales guían el desarrollo de nuestra solución, que implementa un sistema basado en tecnologías modernas como Python, Django y Flask, junto con análisis de texto mediante diccionarios de sentimientos personalizados.

El proyecto se fundamenta en principios de programación orientada a objetos y arquitecturas cliente-servidor modernas, utilizando REST APIs para la comunicación entre componentes. La solución no solo permite el análisis de sentimientos en tiempo real, sino que también proporciona herramientas de visualización y reportería para facilitar la toma de decisiones basada en datos.

A través de este desarrollo, se busca proporcionar a las empresas una herramienta robusta para monitorear y comprender la percepción pública de sus servicios, permitiendo así una mejora continua basada en el feedback directo de los usuarios. La implementación de estándares de desarrollo modernos y prácticas de documentación exhaustiva asegura la escalabilidad y mantenibilidad del sistema a largo plazo.

Desarrollo del tema

La implementación del sistema de análisis de sentimientos se estructura en dos componentes principales: un frontend web desarrollado en Django y un backend API construido con Flask. Esta arquitectura permite una separación clara de responsabilidades y facilita el mantenimiento y escalabilidad del sistema.

COMPONENTE FRONTEND (DJANGO)

El frontend se diseña como una interfaz intuitiva que permite a los usuarios interactuar con el sistema a través de las siguientes funcionalidades clave:

1. Gestión de Archivos
 - Carga de archivos XML con solicitudes de clasificación
 - Visualización de resultados procesados
 - Interfaz para resetear la base de datos
2. Análisis y Reportería
 - Consulta de datos históricos
 - Generación de reportes por fecha específica
 - Análisis comparativo por rangos de fecha
 - Exportación de informes en formato PDF
 - Pruebas de clasificación de mensajes individuales

COMPONENTE BACKEND (FLASK API)

El backend implementa la lógica de negocio central y se estructura en varios módulos:

1. Procesamiento de Datos
 - Parser XML para lectura de solicitudes
 - Sistema de clasificación de sentimientos
 - Gestión de diccionarios de palabras clave
 - Almacenamiento en base de datos XML
2. Análisis de Sentimientos
 - Identificación de palabras positivas y negativas
 - Cálculo de porcentajes de sentimiento
 - Detección de menciones de empresas y servicios
 - Generación de estadísticas agregadas

IMPLEMENTACIÓN TÉCNICA

La solución utiliza estructuras de datos abstractas propias para optimizar el procesamiento:

- Listas enlazadas para gestión de mensajes
- Árboles para organización de diccionarios

- Tablas hash para búsqueda eficiente de palabras clave

INTEGRACIÓN Y FLUJO DE DATOS

1. Proceso de Análisis
 - Recepción de mensajes XML
 - Tokenización y preprocesamiento de texto
 - Aplicación de reglas de clasificación
 - Generación de resultados estructurados
2. Almacenamiento y Consulta
 - Persistencia en archivos XML
 - Indexación para búsqueda rápida
 - Cache de resultados frecuentes

VISUALIZACIÓN Y REPORTERÍA

El sistema implementa diferentes tipos de visualizaciones:

- Gráficos de tendencias temporales
- Distribución de sentimientos por empresa
- Análisis comparativo de servicios
- Reportes detallados en PDF

CONSIDERACIONES DE DISEÑO

1. Escalabilidad
 - Arquitectura modular
 - Procesamiento asíncrono
 - Optimización de consultas
2. Mantenibilidad
 - Documentación exhaustiva
 - Versionamiento en GitHub
 - Pruebas unitarias
3. Seguridad
 - Validación de entrada
 - Control de acceso
 - Sanitización de datos

IMPACTO Y BENEFICIOS

1. Empresariales
 - Monitoreo en tiempo real de la percepción del cliente
 - Identificación temprana de problemas
 - Mejora continua basada en feedback
2. Técnicos
 - Sistema extensible y adaptable
 - Integración con sistemas existentes
 - Base para análisis más avanzados
3. Operativos
 - Reducción de tiempo de procesamiento

- Automatización de análisis
- Generación automática de reportes

Esta implementación representa una solución integral que combina tecnologías modernas con prácticas de desarrollo robustas, proporcionando una herramienta valiosa para el análisis de sentimientos en redes sociales.

Conclusiones

El desarrollo e implementación de este sistema de análisis de sentimientos demuestra la importancia crítica de las herramientas automatizadas para el procesamiento de feedback social en la era digital actual. A través del proyecto, se han identificado varias conclusiones significativas:

1. Arquitectura y Diseño
 - La separación entre frontend y backend mediante una arquitectura cliente-servidor demuestra ser fundamental para la escalabilidad y mantenibilidad del sistema.
 - El uso de tecnologías modernas como Django y Flask proporciona un equilibrio óptimo entre robustez y flexibilidad.
 - La implementación de estructuras de datos abstractas personalizadas resulta crucial para un procesamiento eficiente de grandes volúmenes de datos.
2. Impacto Empresarial
 - El análisis automatizado de sentimientos permite a las empresas responder proactivamente a las percepciones de los usuarios.
 - La capacidad de generar reportes detallados facilita la toma de decisiones basada en datos.
 - La monitorización continua de la satisfacción del cliente se convierte en una ventaja competitiva significativa.
3. Aspectos Técnicos
 - El uso de XML como formato de almacenamiento proporciona una solución

flexible y estructurada para la persistencia de datos.

- La implementación de expresiones regulares para el procesamiento de texto demuestra ser efectiva para la extracción de información relevante.
- El versionamiento mediante GitHub facilita el desarrollo colaborativo y el mantenimiento del código.

Referencias bibliográficas

Máximo 5 referencias en orden alfabético.

M. Grinberg, (2018). *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media, Inc.

W. S. Vincent, (2022). *Django for Professionals: Production websites with Python & Django*. Django for APIs LLC.

B. Liu, (2020). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.

S. Bird, E. Klein, E. Loper, (2019). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media, Inc.

R. Mitchell, (2018). *Web Scraping with Python: Collecting More Data from the Modern Web*. O'Reilly Media, Inc.

.