

Caso de Uso: “Análisis de Sentimiento de Opiniones sobre ChatGPT”

Contexto: ChatGPT es un modelo avanzado de procesamiento de lenguaje natural desarrollado por OpenAI, que interactúa con usuarios a través de conversaciones escritas, ofreciendo respuestas en una variedad de temas. A medida que ChatGPT se utiliza en diversas aplicaciones, desde asistencia virtual hasta generación de contenido, es crucial para OpenAI comprender cómo los usuarios perciben la calidad de sus interacciones.

Dado que los usuarios pueden tener experiencias variadas, es fundamental capturar sus opiniones y sentimientos respecto a la precisión, utilidad, empatía y relevancia de las respuestas de ChatGPT. Estas opiniones proporcionan un valioso feedback directo que puede ser utilizado para mejorar el modelo, asegurando que se alinee con las expectativas de los usuarios y continúe ofreciendo valor en sus interacciones.

El análisis de sentimiento aplicado a estos comentarios permite a OpenAI identificar tendencias positivas o negativas, evaluar el impacto de actualizaciones recientes y priorizar áreas de mejora en el modelo. Al utilizar técnicas de NLP, el objetivo es extraer insights accionables que guíen el desarrollo futuro de ChatGPT, mejorando la satisfacción del usuario y manteniendo la competitividad del producto en el mercado.

Objetivo: El objetivo del caso de uso es analizar las opiniones y comentarios de los usuarios sobre su experiencia con ChatGPT mediante técnicas de procesamiento de lenguaje natural (NLP) y análisis de sentimiento. Este análisis tiene como finalidad identificar patrones de satisfacción o insatisfacción, capturar la percepción general de los usuarios, y proporcionar insights que permitan mejorar la precisión, relevancia y experiencia general del modelo ChatGPT. A través de este proceso, OpenAI busca optimizar continuamente el rendimiento de ChatGPT, alineándolo con las expectativas y necesidades de los usuarios.

Requisitos funcionales:

1. **Recopilación de Opiniones:** El sistema debe ser capaz de recopilar y almacenar automáticamente los comentarios y opiniones que los usuarios expresan sobre ChatGPT durante sus interacciones. Esto incluye la captura de texto de forma continua y la organización de estos datos en un formato que permita un análisis posterior eficiente.
2. **Preprocesamiento de Texto:** Para preparar los datos para el análisis de sentimiento, el sistema debe realizar un preprocesamiento de las opiniones. Esto incluye la tokenización del texto (división del texto en palabras o frases), la eliminación de palabras vacías (como "y", "el", "de"), la normalización del

texto (conversión de todo el texto a minúsculas) y la lematización (reducción de las palabras a su forma base o raíz).

3. **Análisis de Sentimiento:** El sistema debe aplicar un modelo de análisis de sentimiento que clasifique las opiniones de los usuarios en categorías como "Positiva", "Negativa", o "Neutral". Este análisis debe poder identificar no solo la polaridad del sentimiento, sino también su intensidad, permitiendo una evaluación más precisa de las opiniones.
4. **Asignación de Puntuación de Sentimiento:** Cada opinión debe recibir una puntuación de sentimiento que refleje la intensidad del sentimiento expresado, ya sea positivo o negativo. Esta puntuación ayudará a cuantificar el sentimiento y a realizar comparaciones más detalladas entre diferentes conjuntos de opiniones.
5. **Agregación de Resultados:** El sistema debe ser capaz de agregar los resultados del análisis de sentimiento para proporcionar una visión general de las opiniones sobre ChatGPT. Esto incluye la capacidad de mostrar la distribución porcentual de opiniones positivas, negativas y neutrales, así como identificar los temas más recurrentes mencionados por los usuarios.
6. **Visualización de Resultados:** Para facilitar la interpretación de los resultados, el sistema debe proporcionar herramientas de visualización, como gráficos de barras, gráficos de pastel o gráficos de tendencias a lo largo del tiempo. Estas visualizaciones deben ayudar a identificar patrones y tendencias en las opiniones de los usuarios de manera clara y comprensible.

Los datos se encuentran en el portal de la asignatura comprimidos en un fichero *.zip.

Requisitos para la realización del trabajo:

- Grupos de máximo 3 personas
- Se puede usar cualquier plataforma con python
- Se debe de presentar un documento de máximo 5 diapositivas (sin incluir portadas). Si se presenta más de 5 hojas el trabajo no será evaluado.
- El trabajo debe de contener una solución de viable. No intensificar en la parte técnica.
- El plagio será sancionado con el suspenso automático de la tarea.