



Equipo Deutsche Welle - Presentación del Proyecto: Análisis de Sentimiento

Sarah
Armstrong

Carlos
Gómez Guijarro

Martín
Pérez Bernal

Sophia
Lunetto



Resumen de DW

- La emisora internacional de Alemania, fundada en 1953
- Financiada con dinero público, con oficinas principales en Bonn y Berlín

Algunos datos

- Millones de contactos semanales de usuarios a través de TV, radio y en línea (2023)
- DW transmite en 32 idiomas

Misión

- Brindar información imparcial para mentes libres





Contenido

Estructura de la presentación

1. 📌 Introducción
2. 🔍 Motor de búsqueda
3. 🗨️ Análisis de sentimiento
4. ⚖️ Análisis de sesgo
5. 📊 Resultados
6. ✅ Conclusiones
7. 🚀 Perspectivas

🔧 Entorno técnico del proyecto

- Lenguaje: Python
- Librerías principales:
 - nltk, pandas, NumPy
 - openai, transformers, PyABSA
 - MLflow



Objetivos



Meta

Automatizar el análisis de sentimiento para identificar la tonalidad (positiva, neutral o crítica) de una entidad específica mencionada en artículos periodísticos.



Entregables

- ❑ Datos de evaluación etiquetados
- ❑ Motor de búsqueda
- ❑ Pipeline de análisis de sentimiento
- ❑ Evaluación de modelos
- ❑ Seguimiento de rendimiento con la plataforma MLflow



Análisis Exploratorio de Datos y Limpieza



Datos sin etiquetar de DW:

ID del artículo, idioma, título, resumen, texto completo

Conjuntos de datos externos (etiquetados por expertos):

- PerSenT
- NewsMTSC

- Datos de evaluación (tarea de etiquetado)

EDA basico:

- *Filtrar articulos cortos*
- *Artículos unicos*
- *Frecuencia mensual de publicación*

EDA Avanzado:

- modelado de temas
- NER (Reconocimiento de Entidades Nombradas)



Preprocesamiento General y Específico del corpus



Datos sin etiquetar de DW:

ID del artículo, idioma, título, resumen, texto completo

Conjuntos de datos externos (etiquetados por expertos):

- PerSenT
- NewsMTSC

• Datos de evaluación (tarea de etiquetado)

EDA basico:

- *Filtrar articulos cortos*
- *Artículos unicos*
- *Frecuencia mensual de publicacion*

EDA Avanzado:

- modelado de temas
- NER (Reconocimiento de Entidades Nombradas)

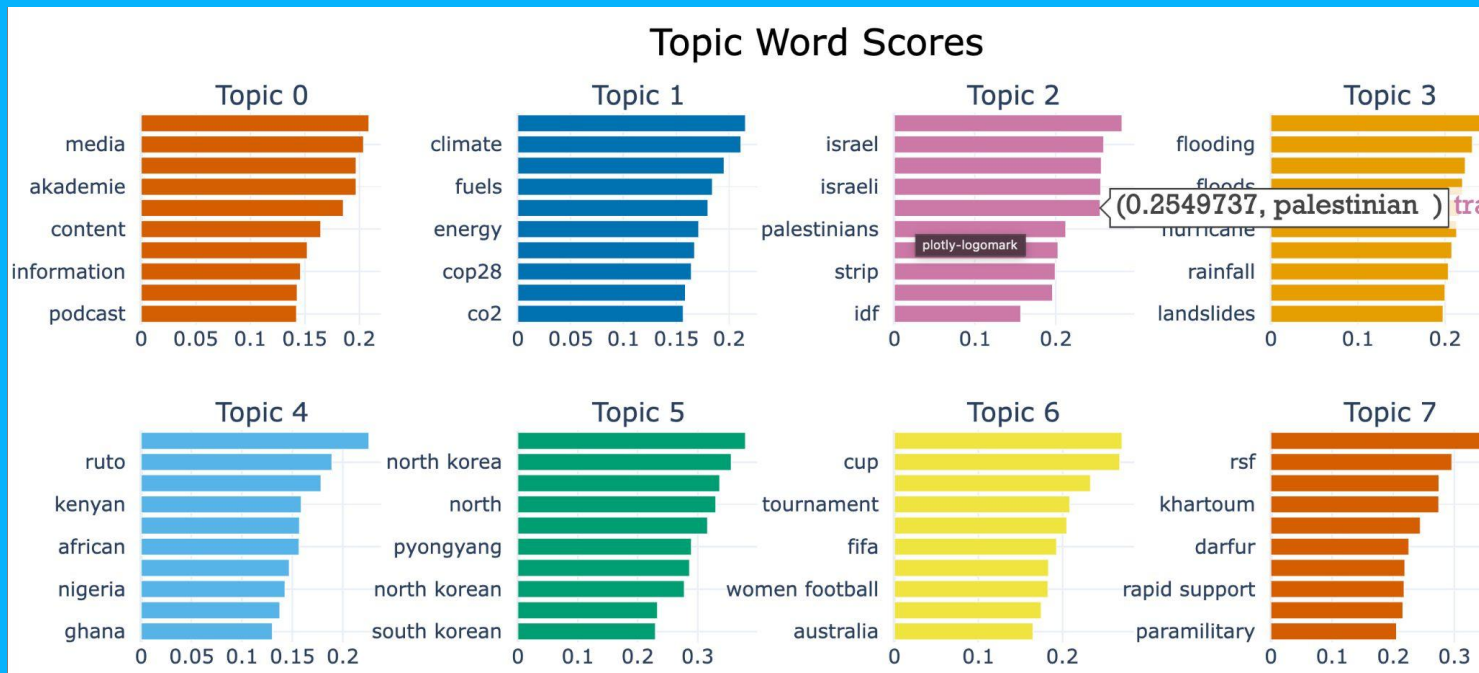
Preprocesamiento general:

- Manejo de valores faltantes
- Manejo de textos muy cortos o muy largos

Preprocesamiento específico del modelo:

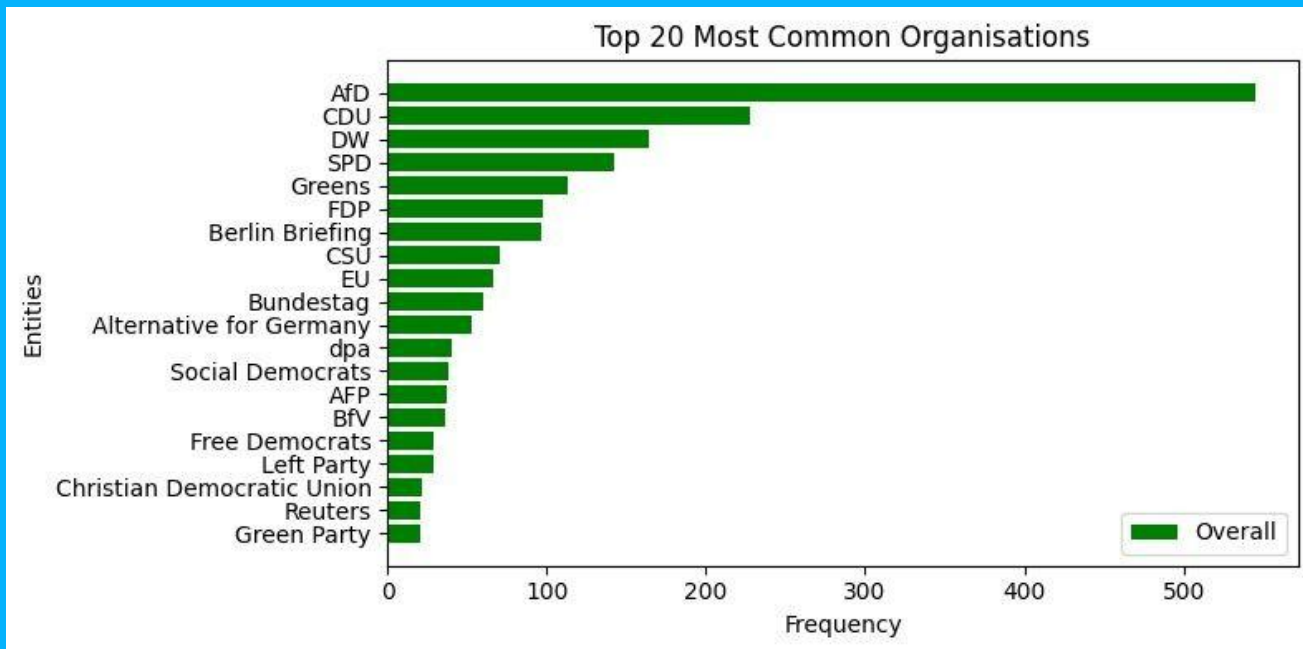
- *NLTK library*
- *Hugging Face*

EDA Avanzado: Modelado de Temas



El modelado de temas hace clusters semanticos

Análisis Exploratorio Avanzado: Reconocimiento de Entidades Nombradas



NER (Reconocimiento de Entidades Nombradas) resalta las entidades principales y está integrado en el motor de búsqueda para obtener perspectivas más profundas



Datos de Evaluación

Crear un conjunto de datos de evaluación (entregable clave)

- Ground Truth
- Mejoramiento del Modelo
- Confianza en el Modelo

Desafíos:

- Complejidad del sentimiento
- Menciones indirectas
- Lenguaje dependiente del contexto

“Merkel es realmente genial, mientras que Steinmeier es un incompetente...”





Proceso Colaborativo de Etiquetado

Selección de un sub
set para evaluación

- **Tema:** “Política alemana” (138 artículos)
- **Entidades clave:** Olaf Scholz (Canciller alemán), AfD (partido de extrema derecha) y el Papa Francisco

Desarrollo
de Guidelines

- Se crearon instrucciones claras para el etiquetado
- Se realizaron reuniones de equipo para asegurar la consistencia y obtener retroalimentación

Primera
ronda

- Cada persona etiquetó aproximadamente 46 artículos
- Cada artículo fue etiquetado dos veces
- Se presentaron desacuerdos en el 37% de los artículos

Segunda ronda

- **Ronda de desempate:** Resolución de discrepancias
- Artículos sin acuerdo fueron eliminados del conjunto de datos

Final de
Evaluación

- Se obtuvo un conjunto de datos de alta calidad (109 artículos) para la evaluación del modelo



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Motor de búsqueda

Ingresar query

I. Encontrar artículos relevantes

Búsqueda por palabra clave

Búsqueda por palabra clave mejorada

Búsqueda Semántica



Entidad nombrada

Coincidencias exactas

Conjunto ampliado semánticamente

Coincidencia basada en similitud semántica



Coincidencia con entidades identificadas por el modelo



Ingresar query

I. Encontrar artículos
relevantes

II. Análisis de
sentimiento

III. Resultados

Motor de búsqueda

Búsqueda por palabra clave

Ingresar query

I. Encontrar artículos
relevantes

query =
"Germany"

Búsqueda por
palabra clave

resultado =
artículos que
contienen:
"Germany"

Pros: Simple, lightweight.

Cons: Rígido, el tiempo requerido escala linealmente con el tamaño del corpus.



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Motor de búsqueda

Búsqueda por palabra clave mejorada

Ingresar query

I. Encontrar artículos relevantes

query =
"Global
warming"

Búsqueda por
palabra
clave
mejorada

resultado =
Articulos que
contienen: "Global
warming" O "Climate
change" O
"Greenhouse gases"

Pros: Captura sinonimos, lightweight.

Cons: Resultados con baja relevancia, el tiempo requerido escala linealmente con el tamaño del corpus.



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Motor de búsqueda

Búsqueda por entidad nombrada

Ingresar query

I. Encontrar artículos relevantes

query =
"Olaf Scholz"

Selección de la entidad mas relevante desde una lista: "Olaf Scholz"

resultados =
Articulos que contienen:
"Olaf Scholz" como entidad

Pros: Precision alta, extrae entitiades especificas.

Cons: Reconocimiento de entidades requiere bastante computacion. Mejora la precisión del análisis de sentimiento.



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Motor de búsqueda

Búsqueda semántica

Ingresar query

I. Encontrar artículos relevantes

query =
"Pope Francis
tour of Sudan"

Búsqueda semántica

resultado =
Articulos ordenados por
relevancia (conteniendo
frases "suficientemente"
similares)

Pros: Comparacion al nivel del significado, alta flexibilidad/relevancia

Cons: Requiere Embeddings, los cuales son intensivos en términos de recursos.



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Motor de búsqueda

similaridad semántica seguido de búsqueda por entidad nombrada

Ingresar query

Ia. Encontrar artículos relevantes

query =
"German
elections"

Búsqueda Semántica

Ingresar query

Ib. Encontrar artículos relevantes

query =
"Olaf Scholz"

Busqueda por
entidad
nombrada

resultado =
Articulos sobre
"Olaf Scholz"
en el contexto
de
"German elections"



Ingresar query

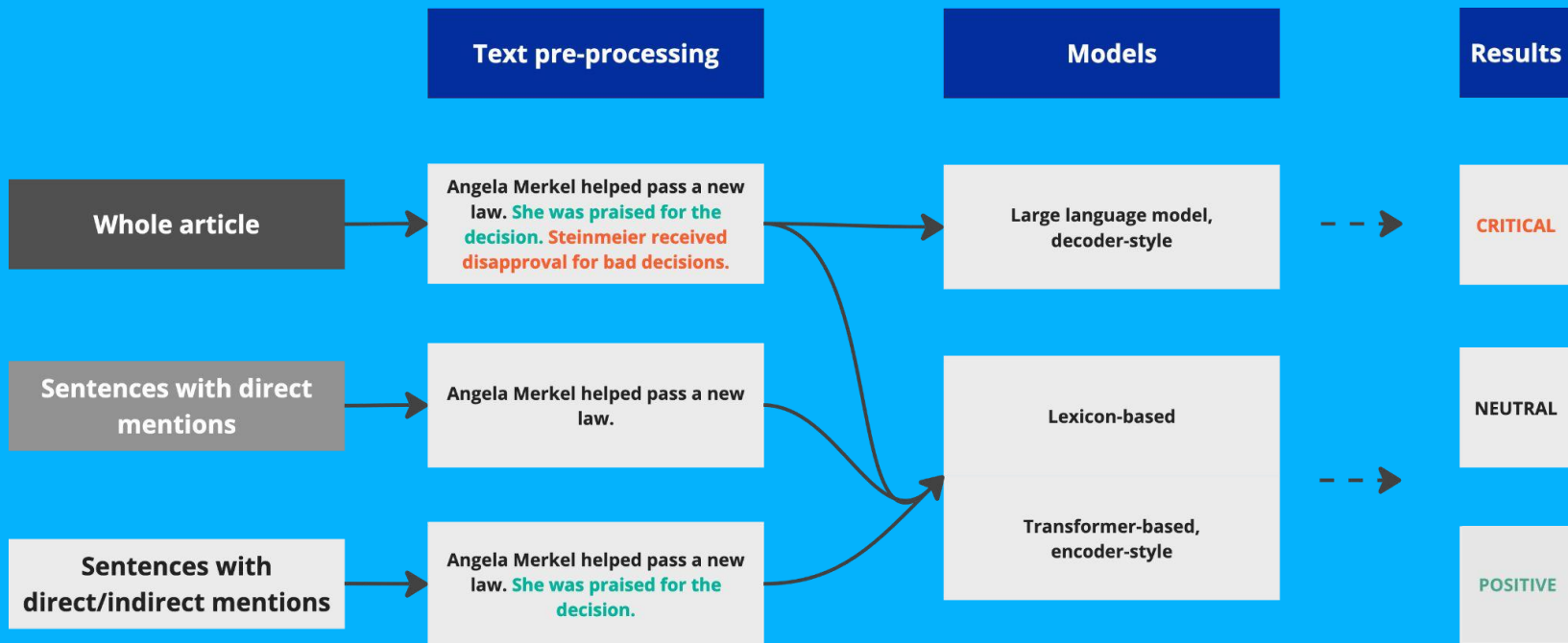
I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Modelos

Preprocesamiento de texto





Ingresar query

I. Encontrar artículos
relevantes

II. Análisis de
sentimiento

III. Resultados

Modelos “Lexicon”

Valence Aware Dictionary for Sentiment Reasoning (Vader)



Léxico de puntuación

desde -4 hasta 4

“horrible”: -2.5

“great”: 3.1

“okay”: 0.9



Heurísticas

intensidad (como “very”)

puntuación (como “great!!!”)

capitalización (como “GREAT”)

Puntaje agregado entre -1 y 1

entre +0.05 y +1 es POSITIVO

entre -0.05 y 0.05 es NEUTRAL

entre -1 y -0.05 es CRITICO



Ingresar query

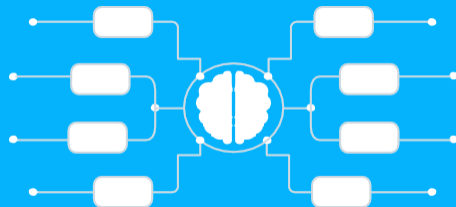
I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Modelos “transformer”

Los transformer son NN especializados con 100M-300M de parametros.



Pros: Bajos prerequisites computacionales.



Cons: Tienen mejor precisión cuando se hace un preprocesamiento adecuado (depende del modelo).

Output: Devuelve un label en forma de Critical, Neutral, o Positive.



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Modelos “transformer”

Bidirectional Encoder Representations from Transformers (BERT).

NLPtown

NewsSentiment

CardiffNLP

PyABSA

Tipo de Modelo

BERT-based transformer

BERT-based transformer

RoBERTa-based transformer

BERT-based transformer

Uso principal

Análisis de sentimiento general (diferentes idiomas)

Análisis dirigido a entidades en artículos periodísticos

Análisis de sentimiento en social media y tweets

Análisis dirigido a entidades

Training Data

Amazon product reviews, movie reviews, etc.

News datasets

Twitter data

ABSA datasets como SemEval, restaurant reviews

Entity-based

✗

✓

✗

✓



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Large Language Models (LLMs)

NNs masivas como GPT y Claude con 1B-14B de parametros.



Pros: Generalizable. Entiende sentimientos de artículos enteros. No hay prerequisites locales.

Cons: Inferencia más costosa (\$), Requieren conocimientos técnicos adicionales (API).

Output: Puntuación de polaridad y, en algunos casos, explicación sobre cómo se llegó a la predicción (útil para prompt engineering)





Ingresar Query

I. Encontrar artículos relevantes

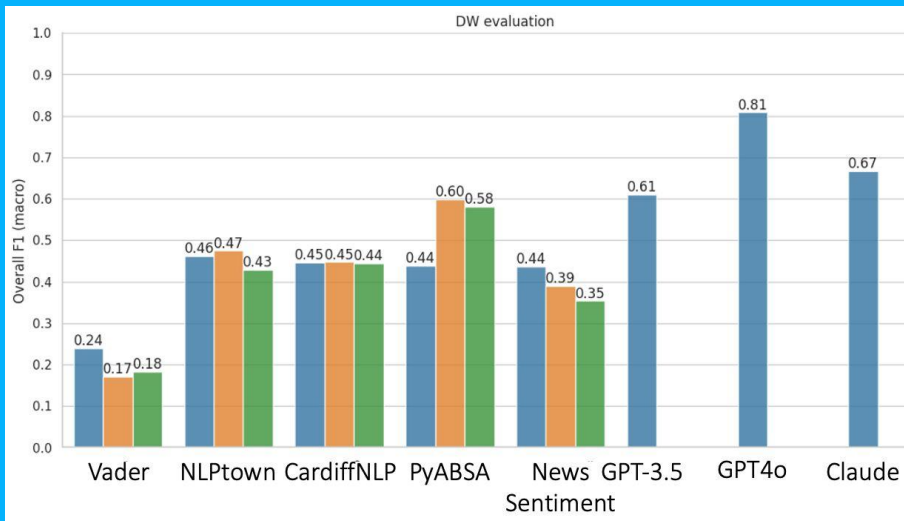
II. Análisis de sentimiento

III. Resultados

Evaluación de los Modelos

F1 score

Promedio de precision and recall, igualdad de importancia en cada clase



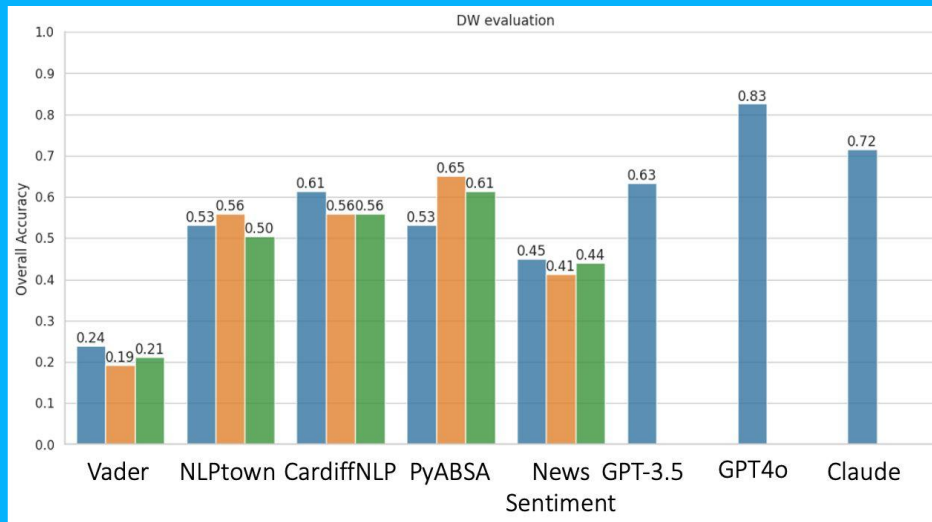
Lexicon-based

Transformer-based

LLM

Accuracy

Proporción de predicciones correctas



whole article



direct mentions



direct/indirect mentions

DW test dataset (109 artículos)



Ingresar Query

I. Encontrar artículos relevantes

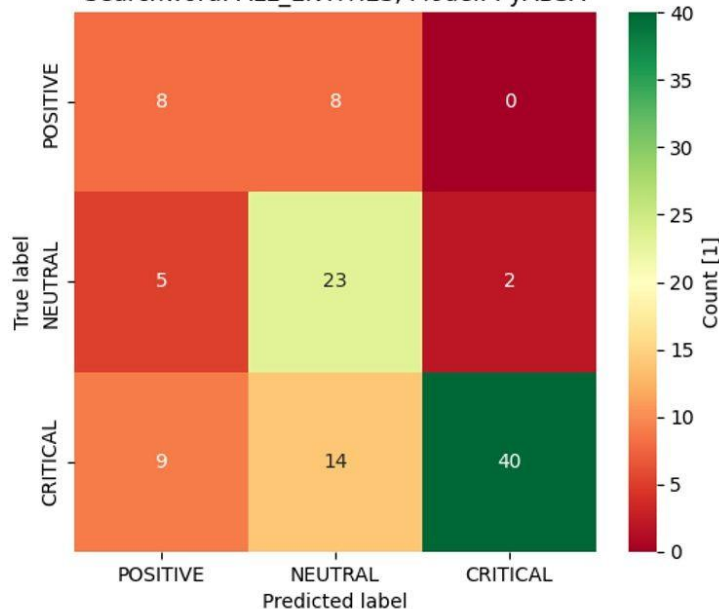
II. Análisis de sentimiento

III. Resultados

Evaluación de los Modelos

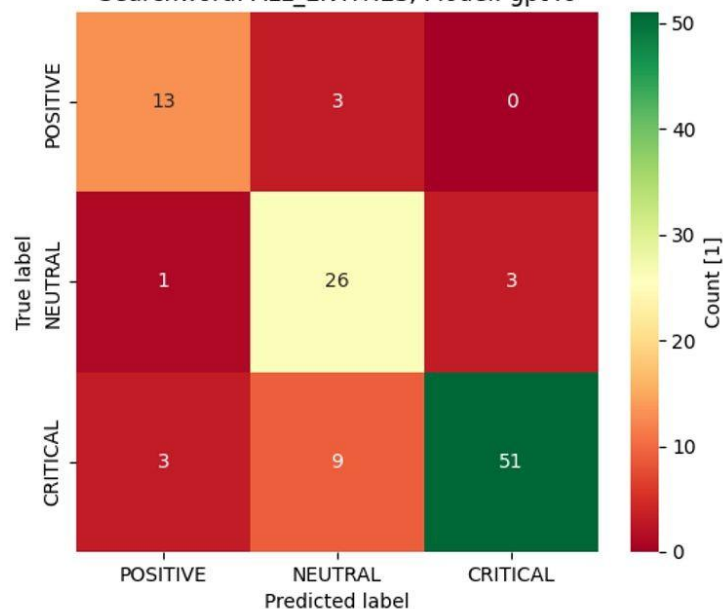
Filtro palabra clave f+ PyABSA

Searchword: ALL_ENTITIES, Model: PyABSA



GPT-4o

Searchword: ALL_ENTITIES, Model: gpt4o





Ingresar query

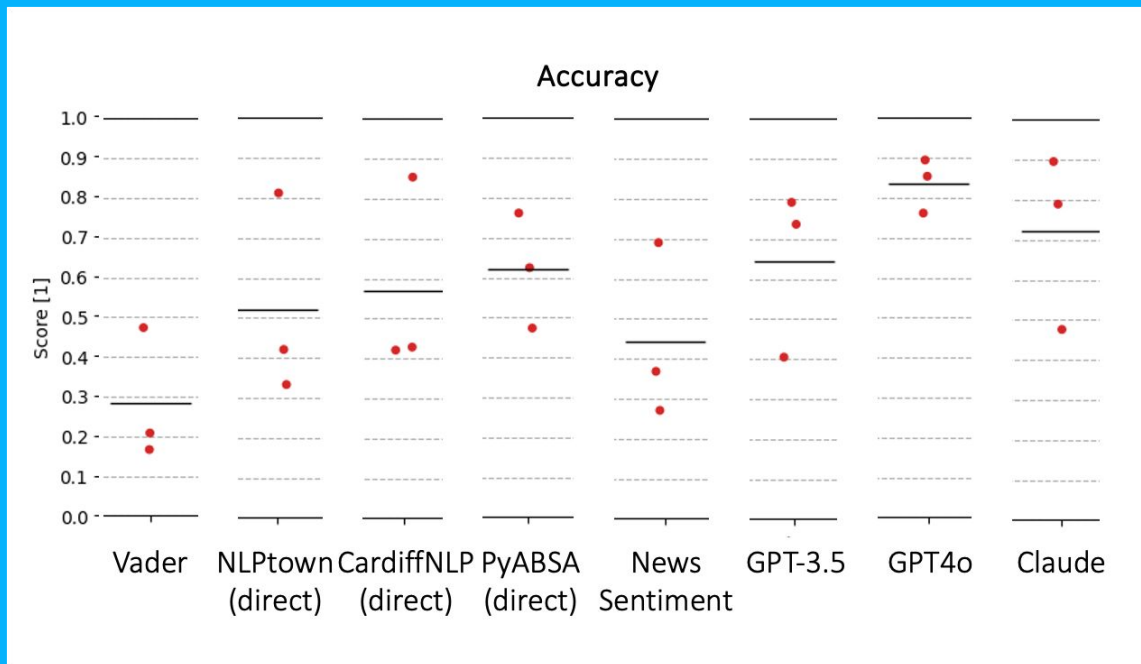
I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Evaluación de los Modelos

Consistencia entre Entidades



Entidades:

AfD (89.6% **CRITICAL**)

Olaf Scholz (50.0% **NEUTRAL**)

Pope Francis (68.4% **POSITIVE**)

Hallazgo:

La mayoría de los modelos tienen dificultades para mantener un rendimiento constante entre distintas entidades.

Mejor desempeño:

GPT-4o



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Evaluación de los Modelos

Seguimiento de Experimentos en MLflow

Sentiment analysis ⓘ Provide Feedback ⓘ Add Description Share

Runs Evaluation Experimental Traces Experimental

metrics.rmse < 1 and params.model = "tree" ⓘ Time created ▾ State: Active ▾ + New run

Datasets ▾ Sort: Created ▾ Columns ▾ Group by ▾

☐	Run Name	Created ▾	Dataset	Duration	Source	Models
☐	debonair-toad-350	6 seconds ago	PerSenT_resolved (4d453c52) Testing ...	7.2s	ipykern...	pyfunc
☐	dapper-colt-190	2 minutes ago	DW (24713a74) Testing	7.1s	ipykern...	pyfunc
☐	lyrical-hog-704	4 minutes ago	NewsMTSC (5aa90aba) Testing	13.2s	ipykern...	pyfunc
☐	languid-finch-944	8 minutes ago	DW_resolved (d9840abd) Testing	8.2s	ipykern...	pyfunc
☐	handsome-cow-341	9 minutes ago	DW_resolved (9fc66717) Testing	8.8s	ipykern...	pyfunc
☐	bouncy-squirrel-713	12 minutes ago	DW_resolved (c6fdeb57) Testing	14.6s	ipykern...	pyfunc
☐	charming-fly-298	13 minutes ago	DW_resolved (658f890e) Testing	9.5s	ipykern...	pyfunc





Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Análisis de sesgo

Tema: German politics

Entidades: Olaf Scholz - CDU - AfD

Búsqueda: Combinación de búsqueda semántica y por entidad nombrada

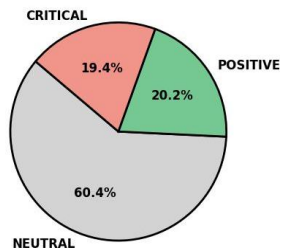
Modelos: PyABSA - GPT-4o

Olaf Scholz (413 artículos)

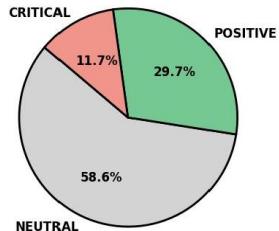
PyABSA

GPT-4o

Sentiment Distribution (%)



Sentiment Distribution (%)

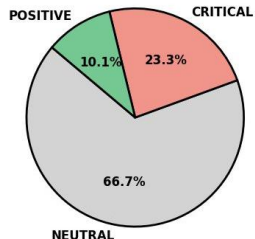


CDU (131 artículos)

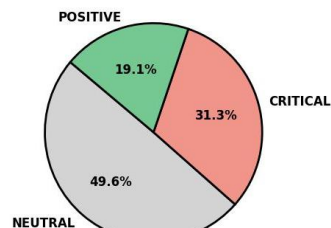
PyABSA

GPT-4o

Sentiment Distribution (%)



Sentiment Distribution (%)

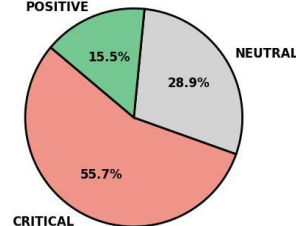


AfD (98 artículos)

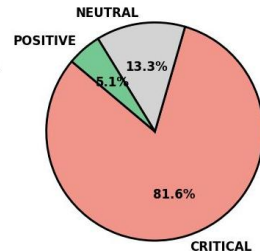
PyABSA

GPT-4o

Sentiment Distribution (%)



Sentiment Distribution (%)





Ingresar query

I. Encontrar artículos
relevantes

II. Análisis de
sentimiento

III. Resultados

Análisis

Tema: German politics

Conclusiones clave:

- Olaf Scholz y CDU están cubiertos predominantemente con un tono neutral

La distribución de tonalidades positiva o crítica varía según el modelo

- AfD recibe una cobertura más crítica, lo que destaca la naturaleza polémica de sus políticas y la percepción pública



Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Análisis

Tema: Israel-Palestine conflict

Entidades: Jews - Arabs - Israel - Netanyahu - Palestinian Authority - Hamas

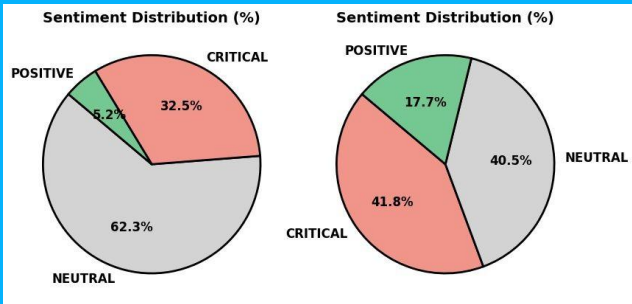
Búsqueda: Combinación de búsqueda semántica y por entidad nombrada

Modelos: PyABSA - GPT-4o

Jews (79 artículos)

PyABSA

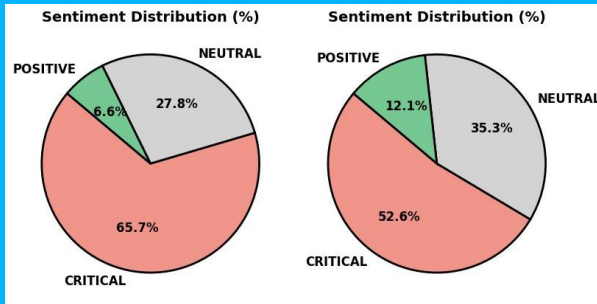
GPT-4o



Israel (372 artículos)

PyABSA

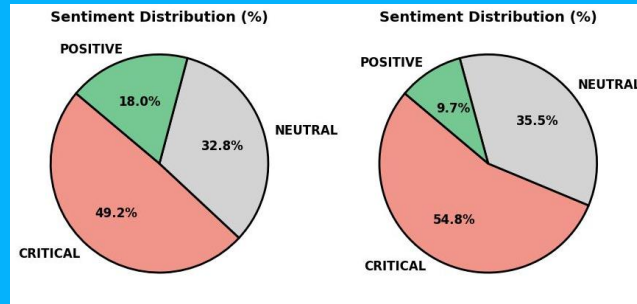
GPT-4o



Netanyahu (62 artículos)

PyABSA

GPT-4o





Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Análisis

Tema: Israel-Palestine conflict

Entidades: Jews - Arabs - Israel - Netanyahu - Palestinian Authority - Hamas

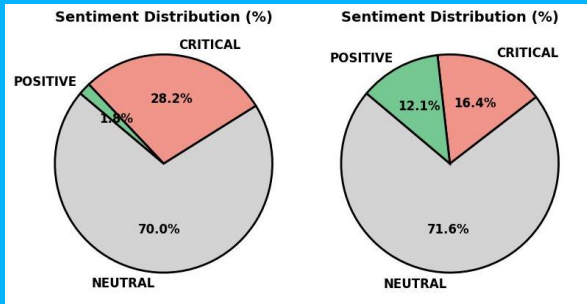
Búsqueda: Combinación de búsqueda semántica y por entidad nombrada

Modelos: PyABSA - GPT-4o

Arabs (116 artículos)

PyABSA

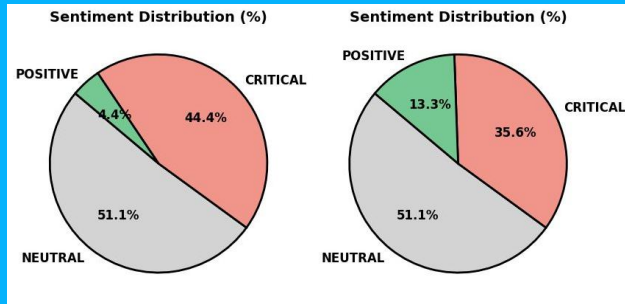
GPT-4o



Palestinian Authority (45 artículos)

PyABSA

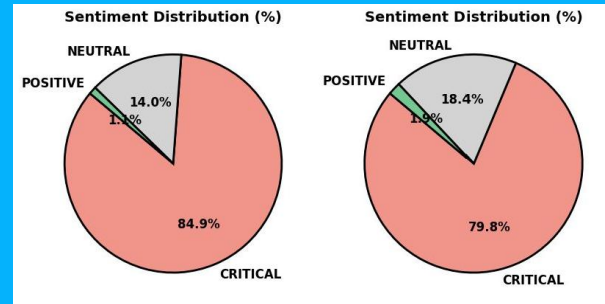
GPT-4o



Hamas (267 artículos)

PyABSA

GPT-4o





Ingresar query

I. Encontrar artículos
relevantes

II. Análisis de
sentimiento

III. Resultados

Análisis

Tema: Israel-Palestine conflict

Conclusiones clave:

- El sentimiento hacia las comunidades judía y árabe depende del modelo utilizado.
- Se necesita más pruebas exhaustivas
- Las entidades con roles de gobierno o conflicto activo (como Netanyahu y Hamas) reciben un sentimiento más **crítico**



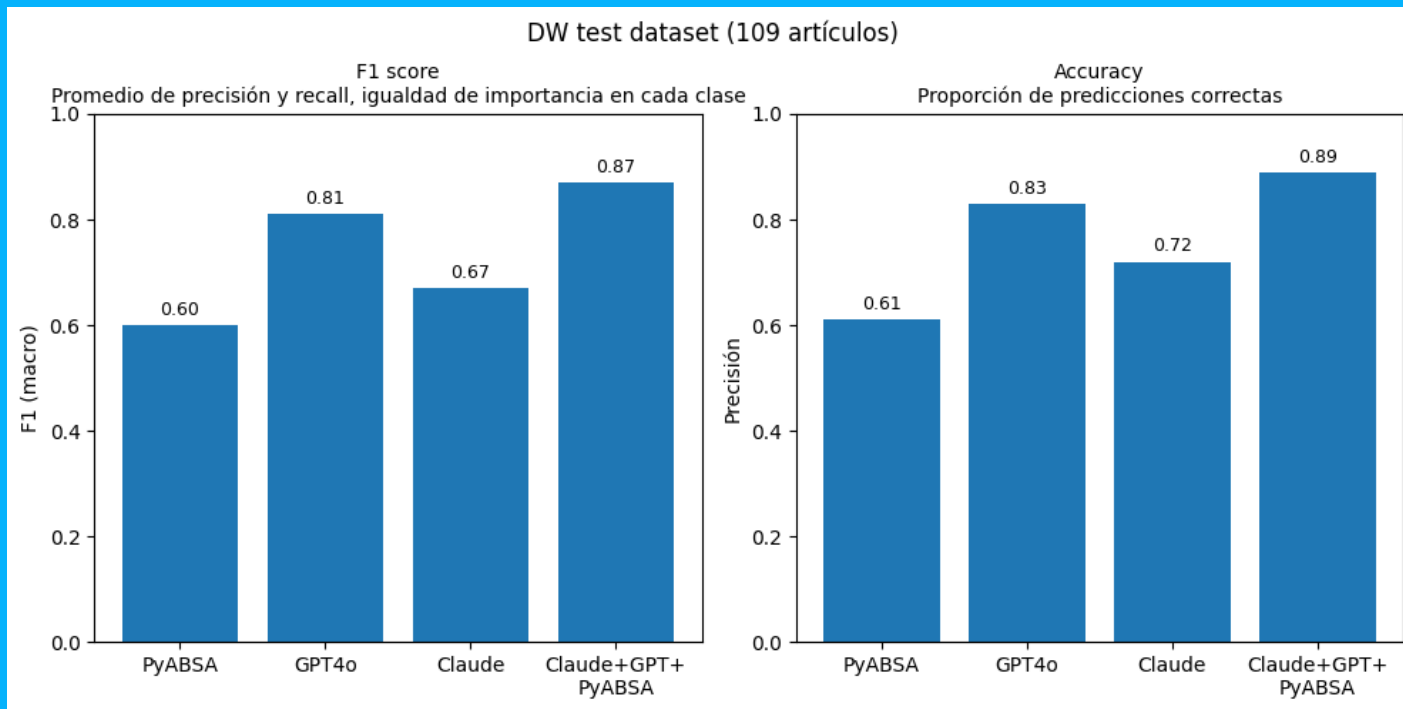
Ingresar Query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Una nueva idea





Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Análisis

Tema: Israel-Palestine conflict

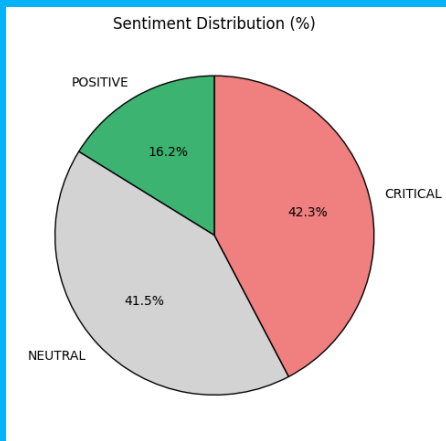
Entidades: Jews - Arabs - Israel - Netanyahu - Palestinian Authority - Hamas

Búsqueda: Combinación de búsqueda semántica y por entidad nombrada

Modelos: claude+gpt+pyabsa

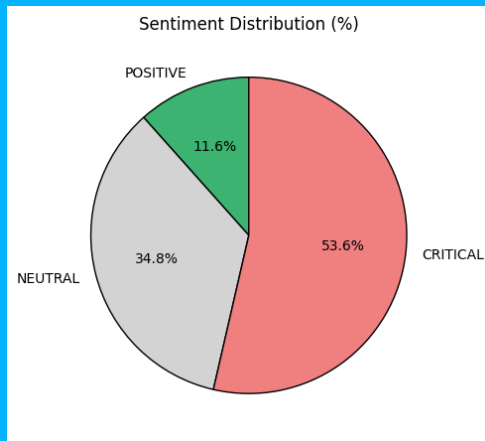
Jews (79 articulos)

claude+gpt+pyabsa



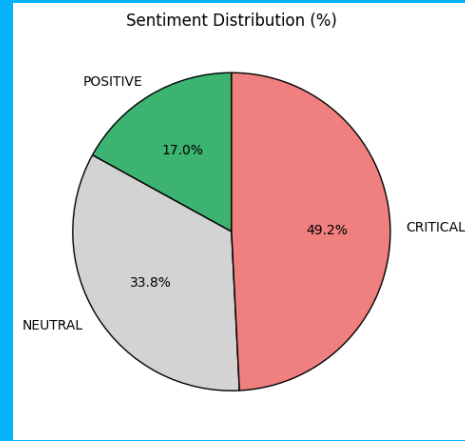
Israel (372 articulos)

claude+gpt+pyabsa



Netanyahu (62 articulos)

claude+gpt+pyabsa





Ingresar query

I. Encontrar artículos relevantes

II. Análisis de sentimiento

III. Resultados

Análisis

Tema: Israel-Palestine conflict

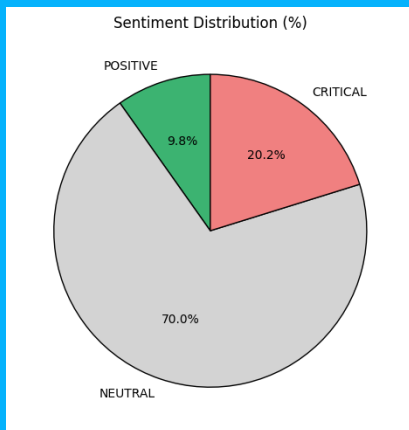
Entidades: Jews - Arabs - Israel - Netanyahu - Palestinian Authority - Hamas

Búsqueda: Combinación de búsqueda semántica y por entidad nombrada

Modelos: claude+gpt+pyabsa

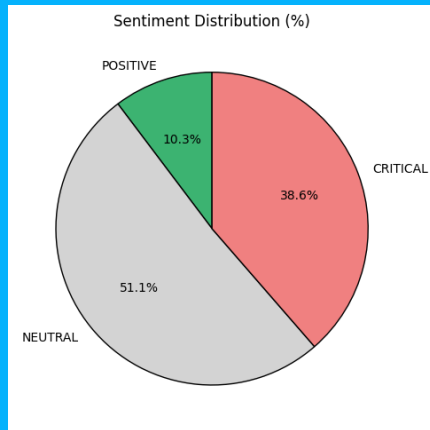
Arabs (116 artículos)

claude+gpt+pyabsa



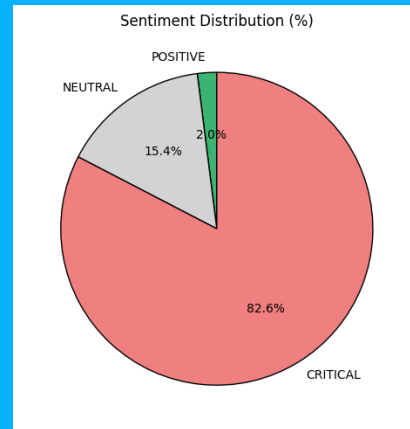
Palestinian Authority (45 artículos)

claude+gpt+pyabsa



Hamas (267 artículos)

claude+gpt+pyabsa





Resumen

- Datos etiquetados de alta calidad son esenciales para evaluar el rendimiento de los modelos
- **Mejor modelo en rendimiento:** GPT-4o
Modelo más eficiente y liviano: PyABSA
Top performer: GPT+Claude+PyABSA
- El sistema de votación de los tres mejores modelos tiene una precisión increíble y esta determino que hay sesgo en algunas entidades



Perspectivas

- Continuar el análisis en temas relevantes y con metadatos
- Mejoras de rendimiento:
 - Ajuste fino de modelos basados en transformadores con corpus de DW
 - Ingeniería de prompts para modelos LLM
 - Estudiar porque el sistema de votación tiene un incremento en precisión
- Facilidad de uso para periodistas:
 - Crear una interfaz de usuario
 - Desarrollar un chat interactivo
- Nuevas funcionalidades:
 - Soporte para múltiples idiomas
 - Mostrar evolución del sentimiento a lo largo del tiempo



Gracias por su atención

