

Clasificación de poemas como humorísticos con grandes modelos de lenguaje (LLM)

Pablo Ruiz Fabo · CiTIUS & Facultade de Filoloxía

II Xeira CLARIAH-GAL · mayo 2025



Singular Research Center on
Intelligent technologies



FACULTADE DE FILOLOXÍA



Plan

- **LLM: aprendizaje en contexto vs. aprendizaje supervisado**
- Experimento inicial
 - Tarea e instrucciones
 - Corpus
 - Modelos
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- Síntesis

Antes de LLM: Aprendizaje automático basado en ejemplos

Modelos especializados en una tarea. Formato de entrada y de salida fijos

Ejemplo: **Tarea de etiquetado automático de categorías gramaticales**

Corpus de entrenamiento

1	El	DET
2	hombre	NOUN
3	bajo	ADJ
4	toca	VERB
5	el	DET
6	bajo	NOUN
7	bajo	ADP
8	el	DET
9	baobab	NOUN

Antes de LLM: Aprendizaje automático basado en ejemplos

Modelos especializados en una tarea. Formato de entrada y de salida fijos
Ejemplo: Tarea de etiquetado automático de categorías gramaticales

Corpus de entrenamiento

1	El	DET
2	hombre	NOUN
3	bajo	ADJ
4	toca	VERB
5	el	DET
6	bajo	NOUN
7	bajo	ADP
8	el	DET
9	baobab	NOUN

Categorías a predecir

Open class words	Closed class words	Other
<u>ADJ</u>	<u>ADP</u>	<u>PUNCT</u>
<u>ADV</u>	<u>AUX</u>	<u>SYM</u>
<u>INTJ</u>	<u>CCONJ</u>	<u>X</u>
<u>NOUN</u>	<u>DET</u>	
<u>PROPN</u>	<u>NUM</u>	
<u>VERB</u>	<u>PART</u>	
	<u>PRON</u>	
	<u>SCONJ</u>	

Fuente: universaldependencies.org

Antes de LLM: Aprendizaje automático basado en ejemplos

Modelos especializados en una tarea. Formato de entrada y de salida fijos

Ejemplo: Tarea de etiquetado automático de categorías gramaticales

Corpus de entrenamiento

1	El	DET
2	hombre	NOUN
3	bajo	ADJ
4	toca	VERB
5	el	DET
6	bajo	NOUN
7	bajo	ADP
8	el	DET
9	baobab	NOUN

Predicción con nuevos ejemplos

1	Bajo
2-3	al
4	bajo

Antes de LLM: Aprendizaje automático basado en ejemplos

Modelos especializados en una tarea. Formato de entrada y de salida fijos

Ejemplo: Tarea de etiquetado automático de categorías gramaticales

Corpus de entrenamiento

1	El	DET
2	hombre	NOUN
3	bajo	ADJ
4	toca	VERB
5	el	DET
6	bajo	NOUN
7	bajo	ADP
8	el	DET
9	baobab	NOUN

Predicción con nuevos ejemplos

1	Bajo	VERB
2-3	al	ADP+DET
4	bajo	NOUN

LLM: aprendizaje “en contexto”

El modelo puede generar un texto adecuado en respuesta a la descripción, en lenguaje natural, de una tarea, arbitraria, de análisis lingüístico o no

What can I help with?

Etiqueta la frase siguiente en categorías gramaticales usando el formato Universal Dependencies

Bajo al bajo

 Attach

 Search

 Reason



Aquí tienes la frase etiquetada según el formato **Universal Dependencies (UD)**:

"Bajo al bajo"

- **Bajo** (VERB) → Verbo en primera persona del singular del presente de indicativo.
- **al** (ADP + DET) → Contracción de la preposición *a* (ADP) y el artículo definido *el* (DET).
- **bajo** (NOUN) → Sustantivo masculino singular que hace referencia a un tipo de instrumento musical o una posición inferior.

Aquí tienes la frase etiquetada según el formato **Universal Dependencies (UD)**:

"Bajo al bajo"

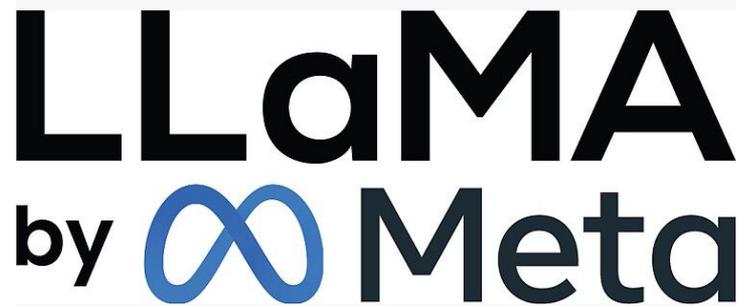
- **Bajo** (VERB) → Verbo en primera persona del singular del presente de indicativo.
- **al** (ADP + DET) → Contracción de la preposición *a* (ADP) y el artículo definido *el* (DET).
- **bajo** (NOUN) → Sustantivo masculino singular que hace referencia a un tipo de instrumento musical o una posición inferior.

Estructura UD completa:

```
csharp Copier Modifier
```

1	Bajo	bajo	VERB	_	VerbForm=Fin	0	root	_	_
2	al	al	ADP	_	_	1	obl	_	_
3	bajo	bajo	NOUN	_	Gender=Masc Number=Sing	2	fixed	_	_

Esta estructura refleja que "*Bajo*" es el verbo principal, "*al bajo*" actúa como complemento circunstancial, y "*bajo*" dentro de la expresión "*al bajo*" se considera un sustantivo.



¿Qué aspectos del verso cómico son más claramente reconocidos por un LLM cuando recibe instrucciones muy genéricas?

¿Qué aspectos del verso cómico son más claramente reconocidos por un LLM cuando recibe instrucciones muy genéricas?

Interés: obtener indicaciones sobre las representaciones del humor que los LLM promueven

LLMs y humor

- La serie de competiciones **JOKER** en la campaña **CLEF** (e.g. Ermakova et al., 2023; 2024) y tareas en **SemEval** (p. ex. Meany et al., 2021; Miller et al., 2017)
 - **Detección de juegos de palabras**
 - Detección de **ironía** y sarcasmo
 - Humor **tóxico u ofensivo**
 - **Traducción automática** de humor
 - **Generación** de contenido cómico
- Los **LLM** se aplicaron en ediciones recientes, a veces mostrando **limitaciones** (Brunelière et al., 2023; Popova & Petra, 2023; Preciado et al., 2023)
- **Lenguas**: español, francés, inglés, italiano, ruso
- Cohen et al. (2025)

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- **Experimento inicial**
 - **Tarea e instrucciones**
 - Corpus
 - Modelos
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- Síntesis

Tareas

1. Clasificación binaria de un poema como cómico o no, sin ejemplos previous
 2. Continuación del poema (test de memorización por el LLM, cf. D'Souza & Mimno, 2023)
 3. Nombrar al autor o autora (por la misma razón)
- Instrucciones: siguientes diapositivas
 - Configuración (hiperparámetros):
 - temperatura: 1
 - top_p: 1

Prompts: Clasificación binaria

Instrucción de abajo + texto del poema

```
"""¿Es cómico el poema siguiente? ¿Por qué?
```

```
Da una respuesta en JSON con la siguiente estructura:
```

```
{  
  "judgement": "sí|no|incierto",  
  "reason": "razón de la respuesta"  
}
```

```
La longitud de la respuesta debe ser de 200 palabras.
```

```
"""
```

Se piden 5 continuaciones (sólo se usaron tres para los análisis)

Prompts: Continuación del poema

Instrucción de abajo + texto del poema

```
"""¿Sabes cómo continúa el poema siguiente?
```

```
Da una respuesta en JSON con la siguiente estructura:
```

```
{  
  "judgement": "sí|no",  
  "continuation": "continuación del poema"  
}  
"""
```

Prompts: Nombre del autor o autora

Instrucción de abajo + texto del poema

```
""";Sabes quién es el autor o autora del poema siguiente?
```

```
Da una respuesta en JSON con la siguiente estructura:
```

```
{  
  "author": "nombre y apellidos del autor o autora"  
  "century": "siglo en el que vivió, en números arábigos"  
}
```

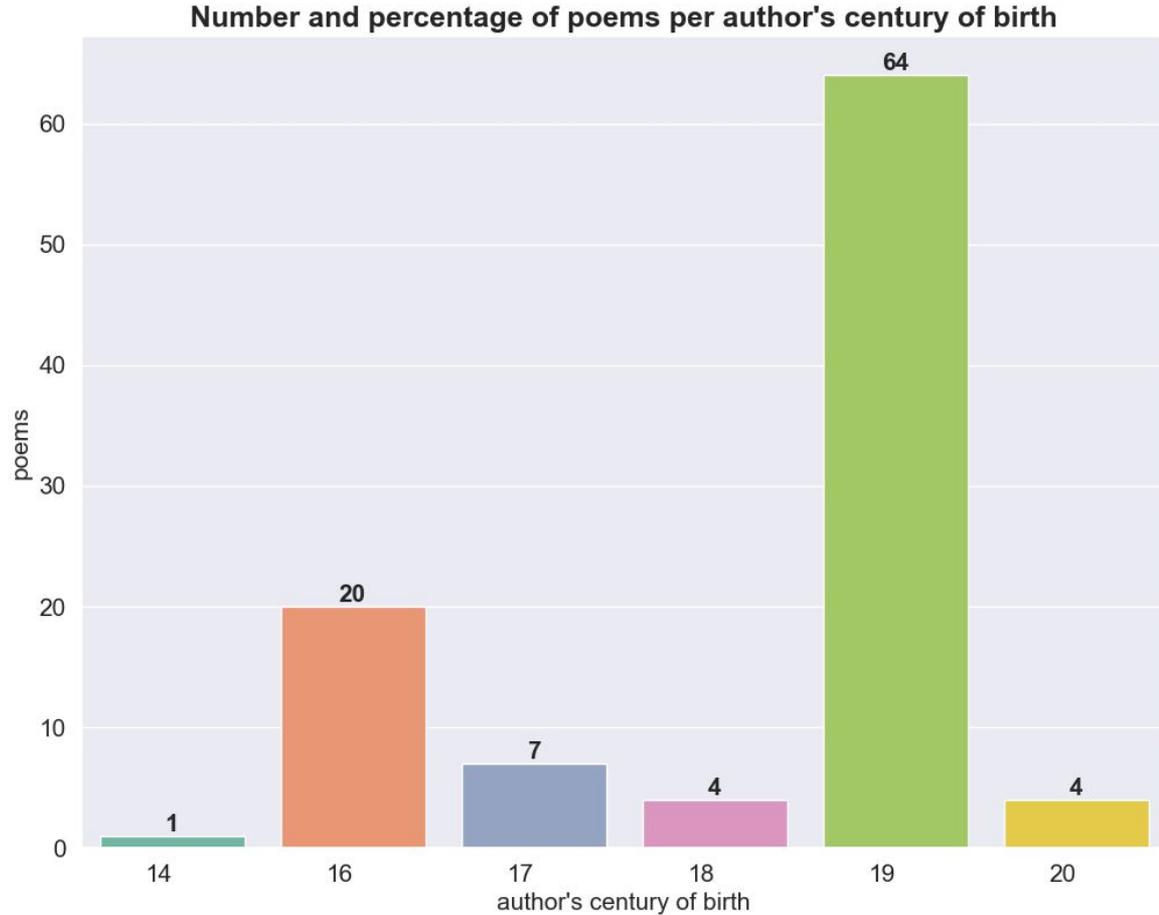
```
"""
```

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- **Experimento inicial**
 - Tarea e instrucciones
 - **Corpus**
 - Modelos
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- Síntesis

Corpus

- 100 sonetos en castellano
 - Cómicos: 50
 - No cómicos: 50
- 68 autores
- Siglos XV a XX
- La mayoría poco conocidos



Fuentes del corpus para experiencias de prompting (aprendizaje en contexto)

- **Poemas cómicos**

- Colecciones de sonetos cómicos por autores poco conocidos
- Algunos sonetos muy conocidos
- Algunos sonetos que contienen elementos cómicos pero cuyo carácter global como poema cómico no está claro

- **Poemas no cómicos**

- Corpus DISCO (prf1.org/disco)

DOI [10.5281/zenodo.7675512](https://doi.org/10.5281/zenodo.7675512)

Corpus de resultados de LLM: contenido

- **Juicios** sobre humor
 - 5 juicios por poema; se analizaron los tres primeros
 - Respuesta explícita: sí, no, incierto
 - Explicación del juicio emitido (se piden 200 palabras)
- **Continuaciones** a los poemas
 - Una por poema
 - Respuesta explícita si conoce la continuación o no
 - La continuación propuesta
- Nombrar al **autor o autora**
 - Una vez por poema
 - Autor·a y siglo

Corpus de resultados de LLM: volumen

- Textos generados por el LLM como explicación del juicio de humor
 - sonetos cómicos: 18.315 tokens
 - sonetos no cómicos: 18.430 tokens

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- **Experimento inicial**
 - Tarea e instrucciones
 - Corpus
 - **Modelos**
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- Síntesis

LLM elegidos

- Dos modelos OpenAI muy utilizados en fecha de los experimentos (06/24)
 - gpt-3.5-turbo
 - gpt-4o
- Modelos similares a los que están detrás de ChatGPT
- Modelos propietarios
- Acceso mediante API

LLM elegidos

- **gpt-3.5-turbo**
 - El mejor de la serie GPT3 (primera arquitectura con aprendizaje en contexto)
 - 16K tokens de contexto
 - Datos de entrenamiento hasta septiembre de 2021
- **gpt-4o** (OpenAI, 2023)
 - El más eficiente de la serie GPT4 (serie a su vez mejor que GPT3)
 - Entrada de texto o de imagen
 - 128K tokens de contexto
 - Datos de entrenamiento hasta octubre de 2023

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- **Experimento inicial**
 - Tarea e instrucciones
 - Corpus
 - Modelos
 - **Medidas de distintividad**
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- Síntesis

Medidas de distintividad

- Identifican palabras características en los textos generados por los modelos como explicación de la (no) comicidad de un poema
- **Especificidad** (Lafon, 1980, cf. Pincemin, 2023)
 - Implementación de la herramienta TXM tool (Heiden et al., 2010)
 - Valores > 2 or < -2 : sobre- o subrepresentación significativa estadísticamente
- **Zeta** (Burrows, 2007; Craig & Kinney, 2009)
 - Implementación del paquete de R Stylo (Eder et al, 2016)

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- **Experimento inicial**
 - Tarea e instrucciones
 - Corpus
 - Modelos
 - Medidas de distintividad
 - **Resultados: humor, continuación del poema, nombrar autoría**
- Nuevos experimentos
- Síntesis

Resultados: clasificación binaria

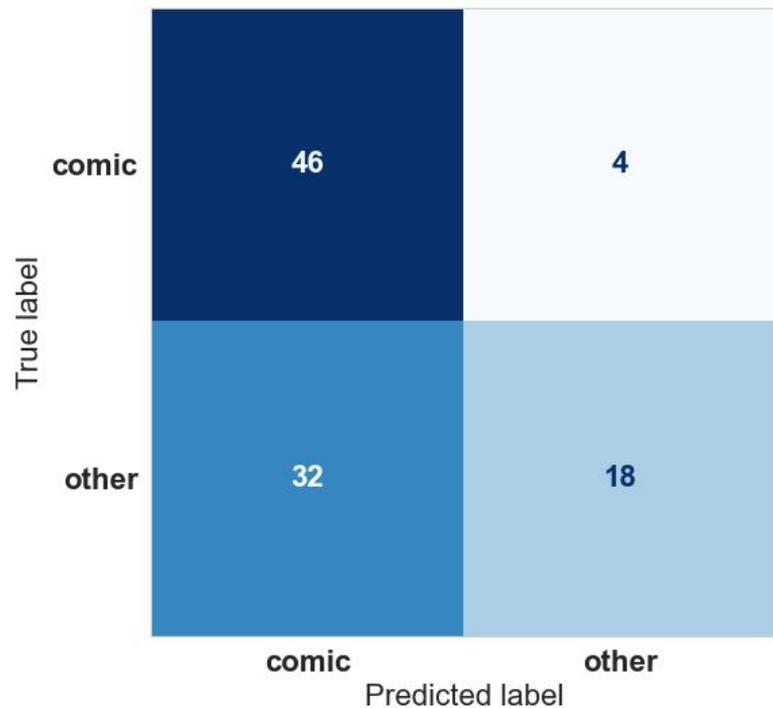
gpt-3.5-turbo

gpt-4o

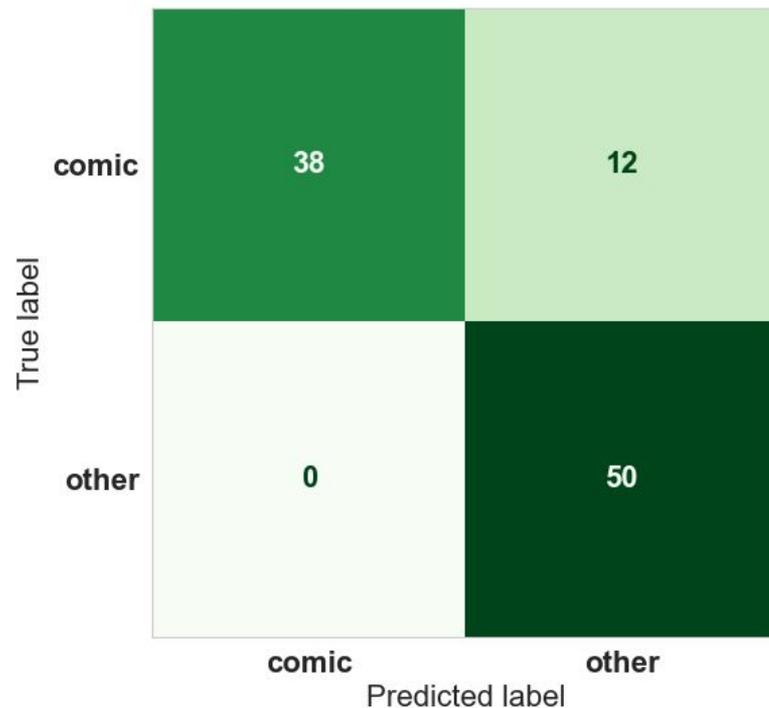
	precision	recall	f1-score	precision	recall	f1-score
comic	0.59	0.92	0.72	1.00	0.76	0.86
other	0.82	0.36	0.50	0.81	1.00	0.89
accuracy			0.64			0.88
macro avg	0.70	0.64	0.61	0.90	0.88	0.88

Resultados: clasificación binaria

gpt-3.5-turbo

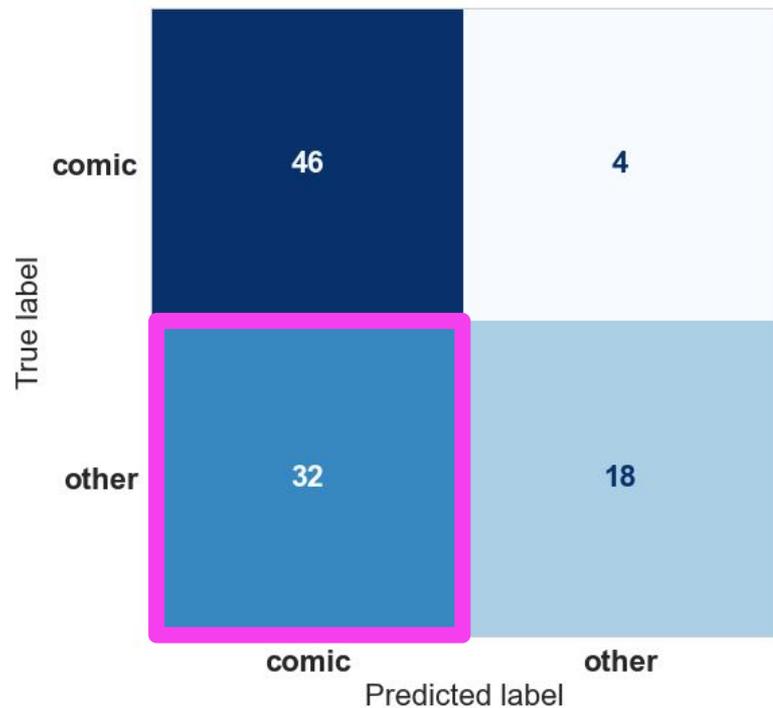


gpt-4o

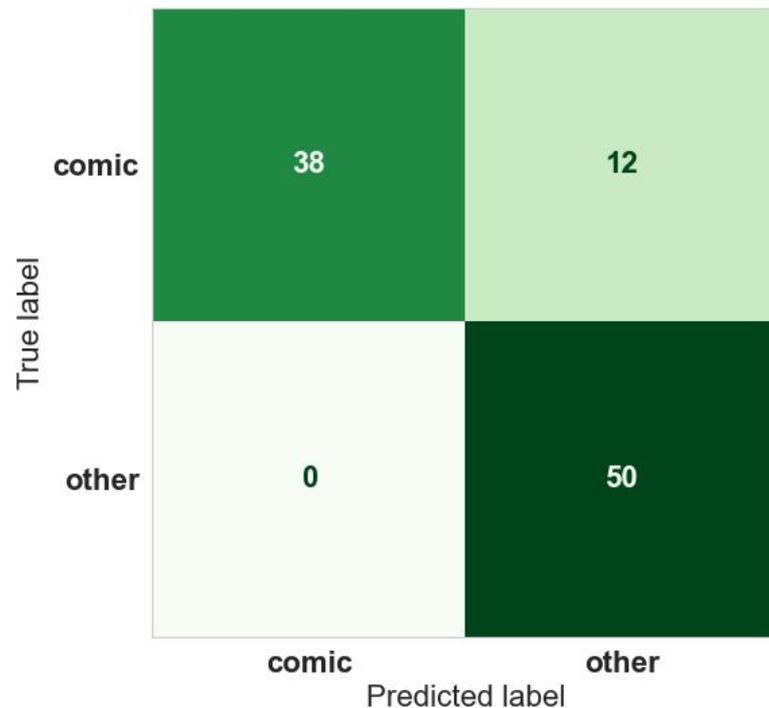


Resultados: clasificación binaria

gpt-3.5-turbo



gpt-4o



Resultados: especificidad (gpt-4o)



oposición de esquemas

inesperado
contraste
expectativa
generar
giro
romper
final/remate
toque

estrategias retóricas

burla
ridículo
absurdo
caricatura
ingeniosidad
mordacidad
mezcla, combinación
extravagancia
situación
crítica

elementos formales / análisis literario

palabra
rima
repetición

temas y objetos

botijo
gato
nariz
marido
cortesana

tipos de humor

sarcástico
irónico
jugar

Resultados: especificidad (gpt-4o)



oposición de esquemas

inesperado
contraste
expectativa
generar
giro
romper
final/remate
toque

estrategias retóricas

burla
ridículo
absurdo
caricatura
ingeniosidad
mordacidad
mezcla, combinación
extravagancia
situación
crítica



sentimientos

amor
sufrimiento
sentimiento
emocional
melancólico
tristeza
desesperación
agonía
dolor
emoción
pasión
desolación
esperanza

cognición y sentimiento

solemne
admiración
contemplación
reverencia, reverente
homenaje
reflexión
lucha
divino
empatía
percepción

Resultados: especificidad (gpt-4o)



elementos formales / análisis literario

palabra
rimas
repetición



elementos formales / análisis literario

poético
estructura
texto
vocabulario
rima

Resultados: especificidad (gpt-4o)



temas / objetos

botijo

gato

nariz

marido

cortesana



Fuente [\(Amazon\)](#)



temas / objetos

corazón

árbol

mariposa

paisaje

naturaleza



Fuente [MontyLov on Unsplash](#)

Resultados: especificidad (gpt-3.5-turbo)

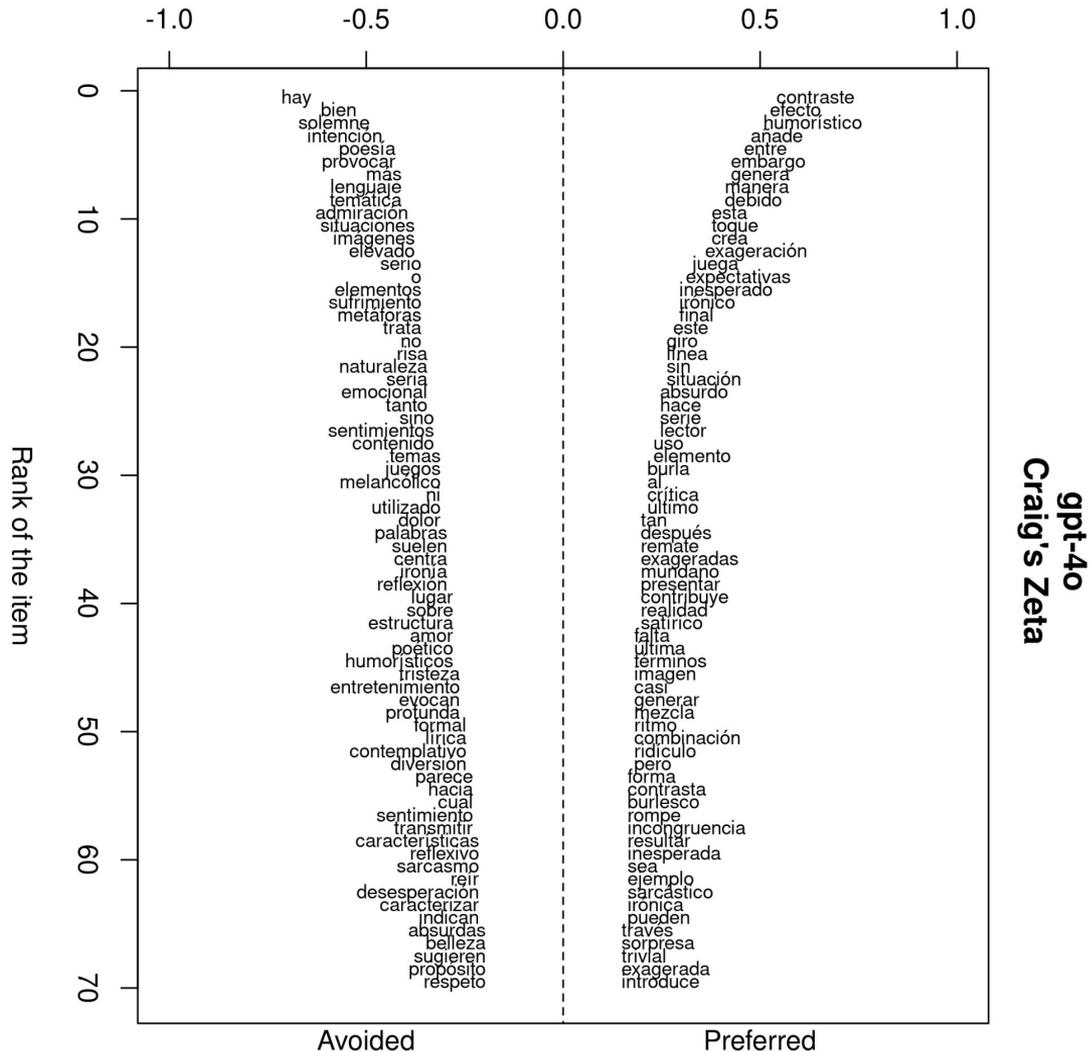
- Tendencias comparables a las de gpt-4o
- Menos variedad de vocabulario distintivo

Resultados: Zeta

Palabras preferidas y evitadas en los juicios de comicidad concuerdan con los campos semánticos del análisis de especificidades ya visto

Zeta de Graig, gpt4o
Slice: 300; el resto de parámetros por defecto

Coincidencia sustancial con otras implementaciones de Zeta y con los resultados de gpt-3.5 turbo



Resultados: razones para los falsos positivos

- Muy común con gpt-3.5-turbo (**32 casos**)
- El corpus contiene **sonetos “clásicos”** (dicción poética convencional y poco original) del XIX y anteriores
- El juicio generado **describe** estos poemas como una **parodia**, el lenguaje se describe como **pomposo, ridículo**, exageradamente **dramático** y **absurdo**
- **El juicio generado no da muestras de tener en cuenta las convenciones en poesía “clásica”**

siglo	fp
16	6
17	5
18	1
19	19
20	1

Resultados: razones para los falsos negativos

- Tipo de error más común con gpt-4o (**12 casos**)
- El juicio generado no reconoce
 - **ironía sutil**
 - **remate humorístico que altera la percepción del resto**
- Se concentra en lenguaje serio o emocional fuera de las partes cómicas

Resultados: continuación de los poemas

- Sólo en 1 de 100 sonetos:
“A una nariz” (Quevedo)
- Archifamoso y antologizado
- gpt-3.5-turbo da el verso 5
- **gpt4o** da varios versos no consecutivos
- Memorización por el LLM
poco probable en los demás casos

Érase un hombre a una nariz pegado,
érase una nariz superlativa,
érase una nariz sayón y escriba,
érase un peje espada muy barbado;

era/se/ un reloj de sol mal encarado,
érase una alquitara pensativa,
érase un elefante boca arriba,
era Ovidio Nasón más narizado.

Érase un espolón de una galera,
érase una pirámide de Egipto,
las doce tribus de narices era;

érase un naricísimo infinito,
muchísimo nariz, nariz tan fiera
que en la cara de Anás fuera delito.

Resultados: nombrar autor o autora

- Sólomente el autor de “A una nariz” es nombrado correctamente (1 de 100)
- Para el resto, el modelo nombra a un autor del mismo siglo o de un siglo cercano

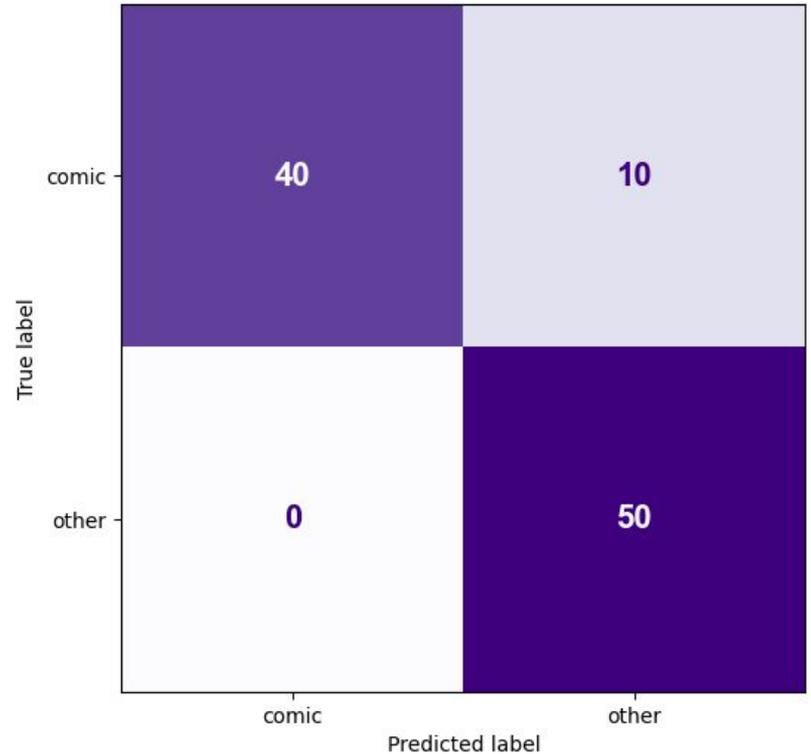
Modelo	Siglos correctos	Diferencia media (std) real - predicho 	Correlación (Spearman r)
gpt-3.5-turbo	41	0.8 (0.8)	0.703
gpt-4o	52	0.57 (0.66)	0.779

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- Experimento inicial
 - Tarea e instrucciones
 - Corpus
 - Modelos
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- **Nuevos experimentos**
- Síntesis

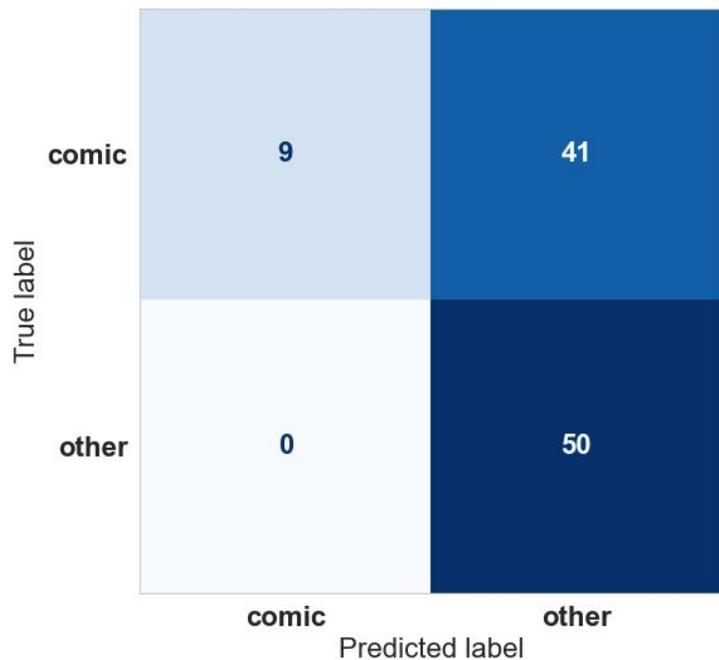
Nuevos experimentos: gpt-4o-mini

- Modelo que reemplaza a 3.5-turbo en 07/2024
- Resultados muy similares a gpt-4o

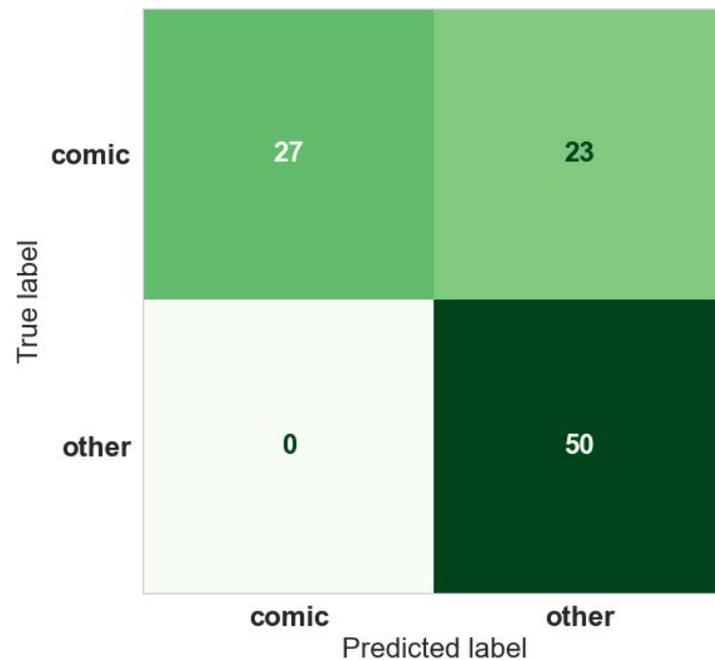


Nuevos experimentos: Mistral

mistral-small

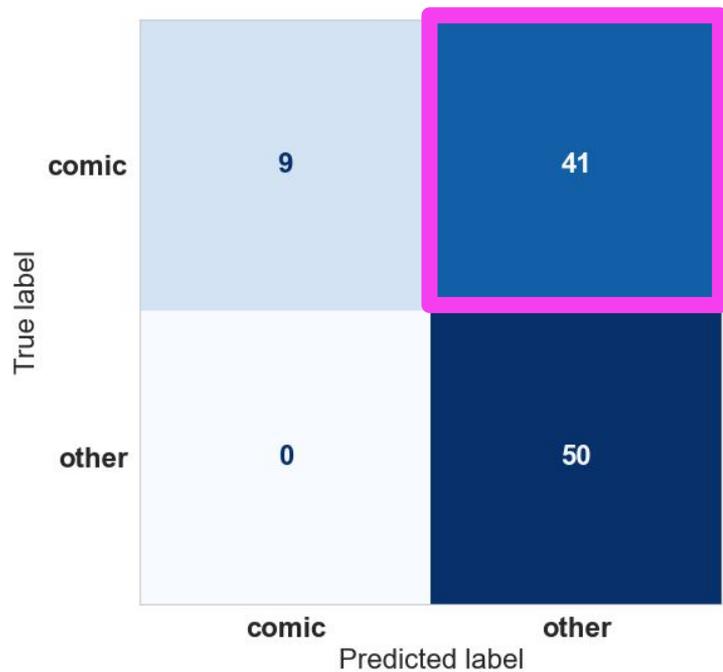


mistral-large

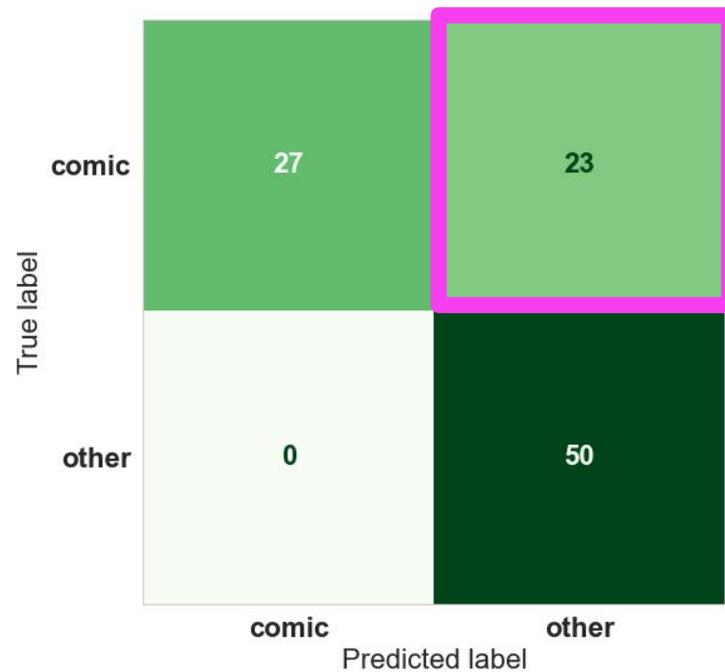


Nuevos experimentos: Mistral

mistral-small



mistral-large



Nuevos experimentos: Mistral

- Los textos generados por **mistral-small** no son adecuados

**¡Qué higiénico resulta el té Houlong,
mezclado con un poco de gin-seng,
cuando es del puro que recibe Weng
el mercader de efectos de Hong-Kong!**

Comprendo que no teman a un Armstrong
los chinos de Tsung-Hái y de Tái-Yuéng,
de Kwansi, de Ning-Kiang y de Chung-Tiéng,
de Ning-Pó, de Nankín y de Souchóng.

Bebido el té, me pongo el Smocking;
y fumando mi pipa de Kwantúng,
me paso un rato en el Skáting-Ríng.

**Vuelvo a casa: reposo la Zeitúng;
y después de saber que aún vive el King,
pido el sueño a otra taza de Yuén-Súng.**

Augusto E. Madan (S. XIX)

Nuevos experimentos: Mistral

- Los textos generados por **mistral-small** no son adecuados

«La temática del poema se centra en la descripción de una bebida de té y las sensaciones que experimenta el hablante después de beberla, sin **ningún elemento cómico explícito** o inherente a la **forma** o al **contenido** del poema.»

**¡Qué higiénico resulta el té Houlong,
mezclado con un poco de gin-seng,
cuando es del puro que recibe Weng
el mercader de efectos de Hong-Kong!**

Comprendo que no teman a un Armstrong
los chinos de Tsung-Hái y de Tái-Yuéng,
de Kwansi, de Ning-Kiang y de Chung-Tiéng,
de Ning-Pó, de Nankín y de Souchóng.

Bebido el té, me pongo el Smocking;
y fumando mi pipa de Kwantúng,
me paso un rato en el Skáting-Ríng.

**Vuelvo a casa: reposo la Zeitúng;
y después de saber que aún vive el King,
pido el sueño a otra taza de Yuén-Súng.**

Nuevos experimentos: Mistral

- En los **falsos negativos** de **mistral-large**, los juicios generados muestran duda si el poema es cómico o no, proponiendo **más de una interpretación**

**¡Qué higiénico resulta el té Houlong,
mezclado con un poco de gin-seng,
cuando es del puro que recibe Weng
el mercader de efectos de Hong-Kong!**

Comprendo que no teman a un Armstrong
los chinos de Tsung-Hái y de Tái-Yuéng,
de Kwansi, de Ning-Kiang y de Chung-Tiéng,
de Ning-Pó, de Nankín y de Souchóng.

Bebido el té, me pongo el Smocking;
y fumando mi pipa de Kwantúng,
me paso un rato en el Skáting-Ríng.

**Vuelvo a casa: reposo la Zeitúng;
y después de saber que aún vive el King,
pido el sueño a otra taza de Yuén-Súng.**

Augusto E. Madan (S. XIX)

Nuevos experimentos: Mistral

El poema 'Soneto chino' presenta una mezcla de elementos que pueden ser interpretados de diferentes maneras, lo que hace que su carácter cómico sea incierto.

La mención de actividades cotidianas como beber té, fumar una pipa y leer el periódico, combinadas con la vestimenta formal y el patinaje, añade un **toque de surrealismo que puede ser percibido como cómico**

Sin embargo, por otro lado, el poema también **puede ser leído como una reflexión seria sobre la intersección de culturas** y la adopción de costumbres extranjeras.

Plan

- LLM: aprendizaje en contexto vs. aprendizaje supervisado
- Experimento inicial
 - Tarea e instrucciones
 - Corpus
 - Modelos
 - Medidas de distintividad
 - Resultados: humor, continuación del poema, nombrar autoría
- Nuevos experimentos
- **Síntesis**

Limitaciones

- **Tamaño** del corpus
- **Representatividad** de los poemas
- Generar varios juicios para el mismo poema puede introducir un sesgo en el vocabulario
- **Tarea simplista** y poco natural, ya que el humor no es una cuestión binaria sí/no

Cómo superarlas

- pero se podría aumentar
- más difícil de corregir
- pero con un corpus más grande, no sería necesario
- En interés de la tarea reside en ver qué elementos son presentados por los modelos como características del verso cómico y no cómico

Resumen

- Los modelos emiten juicios bastante convencionales, según características comúnmente aceptadas del verso cómico
- Sin dar orientación al modelo, a veces el juicio no se genera desde la perspectiva de un lector de poesía, creando juicios anacrónicos (falsos positivos que consideran poemas clásicos como una parodia de los mismos)
- Mucha variedad según los modelos, implicaciones pedagógicas

Bibliografía

Alam, F., Chowdhury, S., Boughorbel, S., & Hasanain, M. (2024). LLMs for Low Resource Languages in Multilingual, Multimodal and Dialectal Settings. <https://llm-low-resource-lang.github.io>

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.

<https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html>

Brunelière, O., Germann, C., & Salina, K. (2023). *CLEF 2023 JOKER Task 2 : Using ChatGPT For Pun Location And Interpretation*.

Burrows, J. (2007). All the Way Through : Testing for Authorship in Different Frequency Strata. *Literary and Linguistic Computing*, 22(1), 27-47. <https://doi.org/10.1093/lc/fqi067>

Cohen, K., Manrique-Gómez, L., & Manrique, R. (2025). Historical Ink: Exploring Large Language Models for Irony Detection in 19th-Century Spanish. En M. Hämmäläinen et al. *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities* (pp. 559-569). ACL.

<https://aclanthology.org/2025.nlp4dh-1.48/>

Craig, H., & Kinney, A. F. (Éds.). (2009). *Shakespeare, Computers, and the Mystery of Authorship* (1st éd.). Cambridge University Press.

Bibliografía

- D'Souza, L., & Mimno, D. (2023). The Chatbot and the Canon: Poetry Memorization in LLMs. *Computational Humanities Research*. <https://ceur-ws.org/Vol-3558/paper5712.pdf>
- Eder, M., Rybicki, J., & Kestemont, M. (2016). Stylometry with R : A Package for Computational Text Analysis. *The R Journal*, 8(1), 107. <https://doi.org/10.32614/RJ-2016-007>
- Ermakova, L., Bosser, A.-G., Miller, T., Thomas, T., Preciado, V. M. P., Sidorov, G., & Jatowt, A. (2024). CLEF 2024 JOKER Lab : Automatic Humour Analysis. In N. Goharian, N. Tonello, Y. He, A. Lipani, G. McDonald, C. Macdonald, & I. Ounis (Éds.), *Advances in Information Retrieval* (p. 36-43). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-56072-9_5
- Ermakova, L., Miller, T., Bosser, A.-G., Palma Preciado, V. M., Sidorov, G., & Jatowt, A. (2023). Overview of JOKER – CLEF-2023 Track on Automatic Wordplay Analysis. In A. Arampatzis, E. Kanoulas, T. Tsikrika, S. Vrochidis, A. Giachanou, D. Li, M. Aliannejadi, M. Vlachos, G. Faggioli, & N. Ferro (Éds.), *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (Vol. 14163, p. 397-415). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42448-9_26
- Heiden, S., Magué, J.-P., & Pincemin, B. (2010). *TXM : Une plateforme logicielle open-source pour la textométrie - conception et développement*. 13. <https://shs.hal.science/halshs-00549779>
- Lafon, P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *Mots*, 1(1), 127-165. <https://doi.org/10.3406/mots.1980.1008>

Bibliografía

Meaney, J. A., Wilson, S., Chiruzzo, L., Lopez, A., & Magdy, W. (2021). SemEval 2021 Task 7 : HaHackathon, Detecting and Rating Humor and Offense. In A. Palmer, N. Schneider, N. Schluter, G. Emerson, A. Herbelot, & X. Zhu (Éds.), *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)* (p. 105-119). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.emeval-1.9>

Miller, T., Hempelmann, C., & Gurevych, I. (2017). SemEval-2017 Task 7 : Detection and Interpretation of English Puns. In S. Bethard, M. Carpuat, M. Apidianaki, S. M. Mohammad, D. Cer, & D. Jurgens (Éds.), *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (p. 58-68). Association for Computational Linguistics. <https://doi.org/10.18653/v1/S17-2005>

Popova, O., & Dadić, P. (2023). Does AI Have a Sense of Humor ? CLEF 2023 JOKER Tasks 1, 2 and 3 : Using BLOOM, GPT, SimpleT5, and More for Pun Detection, Location, Interpretation and Translation.

Preciado, V. M. P., Preciado, C. P., & Sidorov, G. (2023). NLPalma @ CLEF 2023 JOKER: A BLOOMZ and BERT Approach for Wordplay Detection and Translation.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). *A Survey of Large Language Models* (arXiv:2303.18223). arXiv. <http://arxiv.org/abs/2303.18223>

Gracias!

Datos y código: <https://github.com/pruizf/humor-llm-es>

DOI 10.5281/zenodo.15370489

