



# Análisis visual para la detección de sarcasmo en textos mediante emojis.

Diaz Castro Berly Joel

Orientador:

*Plan de Tesis presentado la Escuela Profesional Ciencia de la Computación como paso previo a la elaboración de la Tesis Profesional.*

UNSA - Universidad Nacional de San Agustín de Arequipa  
Junio de 2025

# 1. Motivación y Contexto

El sarcasmo es una forma de comunicación en la que el significado literal difiere de la intención real, frecuentemente mediante ironía o contradicción. En plataformas como Twitter, Reddit o Facebook, el sarcasmo se expresa a través de:

Incongruencias entre texto y emojis (ej. "Me encanta madrugar 🙄").

Sin embargo, detectar sarcasmo automáticamente es un problema complejo debido a la ambigüedad lingüística, frases neutras pueden volverse sarcásticas con un emoji y falta de herramientas visuales que contrasten texto y emojis para identificar patrones.

Herramientas como StanceVis Prime[Kucher et al., 2020] analizan sentimientos y posturas, pero no están diseñados para visualizar incongruencias entre la polaridad del texto y la de los emojis y mapear la frecuencia y posición de emojis, claves para identificar sarcasmo (según estudios como SarcOji[Grover and Banati, 2022]).

# 2. Problema

La falta de herramientas visuales para analizar el sarcasmo en textos de redes sociales que combinan lenguaje natural y emojis. Los métodos actuales ignoran el rol crítico de los emojis como marcadores de ironía.

# 3. Justificación

El sarcasmo es un fenómeno omnipresente en redes sociales, utilizado para expresar ironía, crítica o humor. Sin embargo, su detección automática sigue siendo un desafío debido a su naturaleza contextual y dependiente de claves lingüísticas y multimodales (como emojis). Este proyecto es válido y necesario porque:

Sistemas como StanceVis Prime o herramientas de sentiment analysis (VADER, TextBlob) no están diseñados para identificar incongruencias entre texto y emojis, clave para el sarcasmo y visualizar la posición y frecuencia de emojis.

Este proyecto cierra esa brecha al proponer un enfoque visual intuitivo para analizar estos patrones.

## 4. Objetivos

### 4.1. Objetivo General

Desarrollar un sistema de visualización analítica que permita identificar y analizar patrones de sarcasmo en textos de redes sociales, mediante la integración de incongruencias texto-emoji y frecuencia, posición y polaridad de emojis.

### 4.2. Objetivos Específicos

- Diseñar un modelo de procesamiento de lenguaje natural (NLP) que:
  - Detecte sarcasmo basado en incongruencias entre texto y emojis (ej., texto positivo + emoji negativo).
  - Incorpore frecuencia y posición de emojis como features clave (según hallazgos de SarcOji).
- Implementar visualizaciones interactivas para:
  - Comparar polaridades de texto vs. emojis (gráficos de barras duales o matrices de incongruencia).
  - Mapear la distribución espacial de emojis (heatmaps o líneas de tiempo que muestren posición en el texto).
  - Resaltar emojis repetidos como indicadores de sarcasmo (nubes de emojis ponderadas por frecuencia).
- Validar el sistema con datasets etiquetados (ej., SarcOji o tweets anotados) para:
  - Medir la precisión en la identificación de sarcasmo frente a métodos automáticos tradicionales.
  - Evaluar la utilidad de las visualizaciones mediante pruebas con usuarios (lingüistas o analistas de redes sociales).

## 5. Conjunto de datos, SarcOji

SarcOji[Grover and Banati, 2022] es una colección curada de publicaciones en redes sociales diseñadas para el análisis y detección automática de sarcasmo con apoyo en el uso de emojis. Este conjunto de datos ha sido específicamente estructurado para facilitar experimentos en procesamiento de lenguaje natural (NLP), y en particular, en tareas de detección de sarcasmo mediante modelos basados en deep learning.

El objetivo principal del dataset SarcOji es explorar cómo los emojis actúan como señales de refuerzo emocional o connotativo en textos breves (como tweets o publicaciones de Facebook) y cómo estos pueden ser aprovechados por modelos para mejorar la detección del sarcasmo, una tarea desafiante por su ambigüedad contextual y semántica.

### 5.1. Preguntas

- ¿Si el sentimiento de un emoji es opuesto al sentimiento del texto, es más probable que el mensaje sea sarcástico?
- ¿La ausencia de emojis en textos emocionalmente neutros puede dificultar la detección de sarcasmo?
- ¿El grado de incongruencia cuantificable (Análisis de sentimiento de Text menos el del Emoji) se correlaciona con la probabilidad de sarcasmo?
- ¿Los Emojis agregan polarización emocional que puede ocultar sarcasmo?

## Referencias

- [Grover and Banati, 2022] Grover, V. and Banati, H. (2022). Understanding the sarcastic nature of emojis with SarcOji. In Wijeratne, S., Lee, J., Saggion, H., and Sheth, A., editors, *Proceedings of the Fifth International Workshop on Emoji Understanding and Applications in Social Media*, pages 29–39, Seattle, Washington, USA. Association for Computational Linguistics.
- [Kucher et al., 2020] Kucher, K., Martins, R. M., Paradis, C., et al. (2020). Stancevis prime: visual analysis of sentiment and stance in social media texts. *Journal of Visualization*, 23(5):1015–1034.