

## Análisis Exploratorio de Datos

### Distribución de Clases

El dataset de entrenamiento contiene 7,613 tweets distribuidos de la siguiente manera:

No Desastre (0): 4,342 tweets (57.03%)

Desastre (1): 3,271 tweets (42.97%)

La distribución está relativamente balanceada, lo cual es favorable para el entrenamiento de los modelos.

### Preprocesamiento de Texto

Se implementó una función de limpieza que incluye:

Conversión a minúsculas

Eliminación de URLs

Cambio de símbolos @ y #

Eliminación de caracteres especiales y números

Filtrado de stopwords

### Análisis de Frecuencias

#### Palabras Más Comunes

En tweets de desastre: Las palabras más frecuentes fueron altamente predictivas, incluyendo términos como "fire" (180), "disaster" (117), "suicide" (110), "police" (107), "killed" (93), "storm" (85) y "crash" (84). Estas palabras tienen una clara relación semántica con eventos de emergencia.

En tweets de no desastre: Dominaron palabras de uso cotidiano como "like" (253), "im" (248), "new" (170), "get" (163) y "love" (90), reflejando conversaciones personales y opiniones.

El análisis de bigramas reveló patrones aún más específicos:

Desastre: "suicide bomber", "oil spill", "burning buildings", "california wildfire"

No Desastre: "cross body", "youtube video", "body bag"

### Resultados de Rendimiento

Modelo	Accuracy	Macro F1-Score	Precision (Desastre)	Recall (Desastre)
Regresión Logística	0.82	0.82	0.85	0.71
Naive Bayes	0.81	0.80	0.86	0.67
Linear SVC	0.79	0.79	0.77	0.74

### Análisis de Resultados

La Regresión Logística emergió como el mejor modelo, mostrando el mejor balance entre precision y recall. Naive Bayes mostró alta precision pero bajo recall para la clase de desastre, indicando que era conservador en sus predicciones positivas.

Linear SVC, sorprendentemente, tuvo el rendimiento más bajo.

## **Análisis de Sentimiento**

Se implementó análisis de sentimiento usando VADER para explorar si la polaridad emocional podría mejorar la clasificación.

- Tweets de desastre tuvieron una puntuación promedio de sentimiento de -0.265
- Tweets de no desastre tuvieron una puntuación promedio de -0.061

Sin embargo, el análisis reveló casos interesantes:

- Tweets de desastre con sentimiento positivo
- Tweets de no desastre muy negativos

## **Impacto en el Modelo**

Cuando se añadió la puntuación de sentimiento como característica adicional, no se observó mejora en el rendimiento:

- Macro F1-Score se redujo ligeramente de 0.82 a 0.81
- La accuracy se mantuvo en 0.82

## **Fortalezas del Modelo Final**

Buen balance entre precision y recall

Capacidad de generalización demostrada en validación

Interpretabilidad clara de las características más importantes

Procesamiento eficiente de texto

## **Limitaciones**

Dependencia de palabras clave específicas podría limitar la generalización a nuevos tipos de desastres

El recall de 0.71 indica que se pierden algunos desastres reales (29% de falsos negativos)

Sensibilidad al preprocesamiento de texto