# Introducción

El análisis de sentimientos mediante un texto nos permite identificar la emoción de un texto (positivo, negativo o neutral). En este proyecto se desarrolló un clasificador de sentimientos usando el algoritmo Naïve Bayes, aplicado a un conjunto de tweets. La solución fue implementada desde cero sin el uso de librerías externas para el modelo, y desplegada a través de una aplicación web.

# Definición del problema y objetivos

## Problema

determinar la emoción expresada en un tweet (positivo, negativo o neutral) en función de su contenido textual.

## Objetivo General

Implementar un modelo de clasificación Naïve Bayes capaz de identificar el sentimiento (positivo, negativo o neutral) en tweets.

## Objetivos específicos

* Realizar el preprocesamiento del dataset de tweets.
* Implementar el modelo Naïve Bayes desde cero.
* Evaluar el desempeño del modelo.
* Desarrollar una plataforma web para consultar el modelo.

# Descripción de Dataset

Se utilizó un dataset (twets.csv) con tweets etiquetados en tres categorías: negativo, neutral, positivo. El dataset fue procesado para eliminar ruido como símbolos, URLs, stopwords, lematizado y vectorizado.

# Descripción del Preporcesamiento Aplicado

Se realizo la limpieza y tratamiento del dataset tweets.csv siguiendo el siguiente pipeline utilizado JupyterNotebook:

1. Limpieza de texto: en este punto se realizo una conversión a minúsculas, eliminación de signos de puntuación, menciones y URLs, almacenados en un dataset modificado.
2. Tokenización: se separaron las palabras del texto limpio del proceso anterior. Almacenados en un dataset modificado.
3. Lematización: se estandarizaron las palabras buscando su forma raíz para un mejor tratamiento. Almacenados en un dataset modificado.
4. Vectorización: se utiliza la libreriza BOW de Python para vectorizar las plabras encontradas en los proceso anteriores, estas palabras son las que dan valor, adicional se eliminan palabras como the, this, etc.

# Se realizo la codificación del algoritmo Naïve Bayes:

El modelo se implementó utlizando la fórmula clásica de Naïve Bayes, fórmula:

X: conjunto de palabras

C: clase

# Red Bayesiana: