



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

Maestría en Ciencia de Datos

Universidad Autónoma de Nuevo León

Tarea 2 - Análisis de sentimiento de los comentarios del primer debate presidencial

Por

Juan Jesús Torres Solano

Profesor: José Alberto Benavides Vázquez

Índice general

ÍNDICE DE FIGURAS	2
1 Introducción	3
1.1 Objetivos	3
2 Análisis de sentimiento	4
3 Resultados	5
4 Conclusiones	8
References	9

Índice de figuras

Figura 3.1	Distribución de polaridad y subjetividad.....	5
Figura 3.2	Distribución de frecuencia de voto y su polaridad	6
Figura 3.3	Distribución de frecuencia de voto y objetividad	7

1. INTRODUCCIÓN

Los comentarios en videos en vivo son parte importante de la respuesta de la audiencia y se entiende que estos comentarios este cargados de palabras emotivas o que representen los sentimientos que las personas intentan comunicar al resto de espectadores, incluso a los mismos creadores del contenido o participantes de la transmisión en vivo, por lo que es de esperar que estos se expresen de forma positiva o negativa, de igual forma estos comentarios pueden ser poco objetivos, claramente objetivos o muy subjetivos[1].

El análisis de sentimientos implica el procesamiento del lenguaje natural de modo que se logre identificar información subjetiva de un texto en particular, en esta tarea buscamos identificar el sentimiento subjetivo expresado por los usuarios que dejaron un comentario en la transmisión en vivo del primer debate presidencial [2][1].

1.1. OBJETIVOS

El objetivo de este trabajo es identificar el sentimiento expresado por los usuarios en los comentarios del debate, teniendo en cuenta los votos que recibieron los comentarios y el sentimiento presente en aquellos mas votados, al igual que la respuesta general de los sentimientos de la audiencia; y por lo tanto se requiere establecer su identificación, distribución y cuantificación:

1. Identificar y cuantificar la polaridad de los comentarios que la audiencia dejo en el video del debate.
2. Identificar y cuantificar el nivel de objetividad y subjetividad de los comentarios en el video.

2. ANÁLISIS DE SENTIMIENTO

Para el análisis de sentimiento se empleo la librería de TexBlod, la cual tiene como objetivo ofrecer un entorno familiar, cómodo e intuitivo empleado Python, esta librería nos permite emplear su modelos entrenados para clasificar el sentimiento subjetivo reflejado en las palabras empleadas en los comentarios, así como la objetividad de los mismos [1].

A partir de esta clasificación obtendremos una asignación de peso para cada comentario, el cual esta expresado como "Polaridad". este valor representa sentimientos positivos o negativos, haciendo referencia al concepto de polaridad [2], la escala esta representada de la siguiente manera, donde $c = \text{comentario}$:

- Se consideran con polaridad muy negativa

- $-1 \leq C < 0$

- Se consideran con polaridad neutral

- $C = 0$

- Se consideran con polaridad muy positiva

- $0 < C \leq 1$

Mientras que la objetividad presenta dos limites donde un comentario de puede interpretar como objetivo, es decir presenta una estructura la cual contiene información sin incluir opiniones ni sentimientos, o subjetivo expresando su punto de vista, opinión personal o sentimientos, se clasifica mediante la siguiente escala, donde $c = \text{comentario}$:

- Se consideran objetivo

- $C \geq 0$

- Se consideran subjetivo

- $C \leq 1$

3. RESULTADOS

A partir de los resultados obtenidos mediante el análisis de sentimientos, podemos observar la distribución de frecuencias de los pesos asignados a la objetividad y polaridad de cada comentario, a partir de esta distribución se observa lo siguiente (ver fig 3.1).

Para una mejor visualización de la distribución las frecuencias esta se representa empleando escala logarítmica.

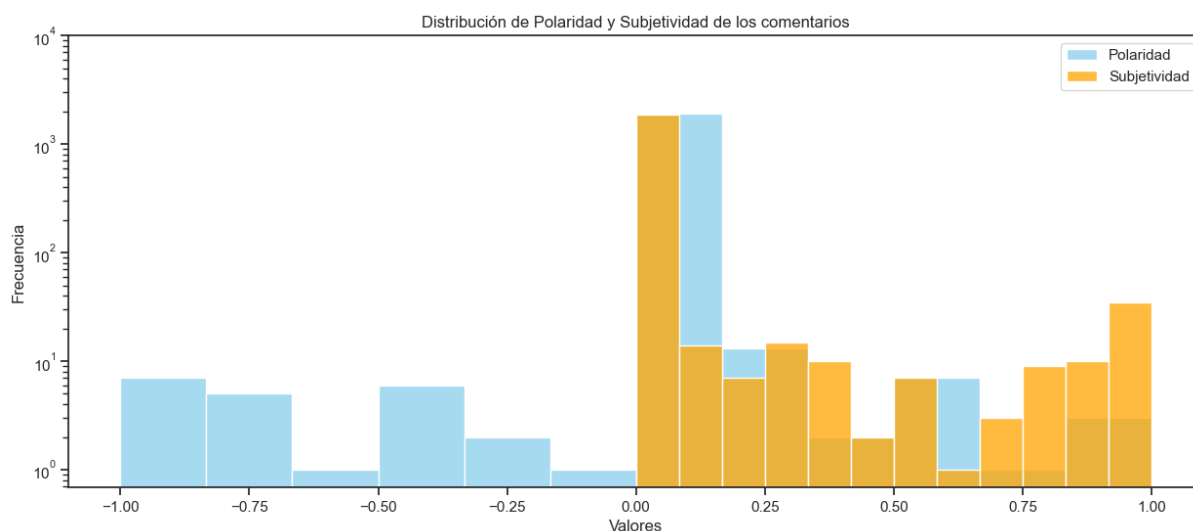


Figura 3.1: Distribución de la polaridad y subjetividad de los comentarios del primer debata presidencial.

Podemos observar que la distribución de polaridad se concentra en la porción central, es decir se identifica predominantemente comentarios con mayor tendencia de polaridad neutral, sin embargo se observa un numero mayor de comentario con tendencia negativa que aquellos con tendencia positiva, a partir de este comportamiento podemos concluir que los comentarios de los usuarios son mayormente neutrales (ver fig 3.1).

La subjetividad por otro lado muestra coherencia en la distribución con respecto a la polaridad, se identifica un numero alto de comentarios objetivos mientras que la distribución de los comentarios subjetivos presenta una menor distribución, sin embargo, en conjunto es notorio un aumento de la subjetividad dentro del rango próximo al valor de 1 (ver fig 3.1).

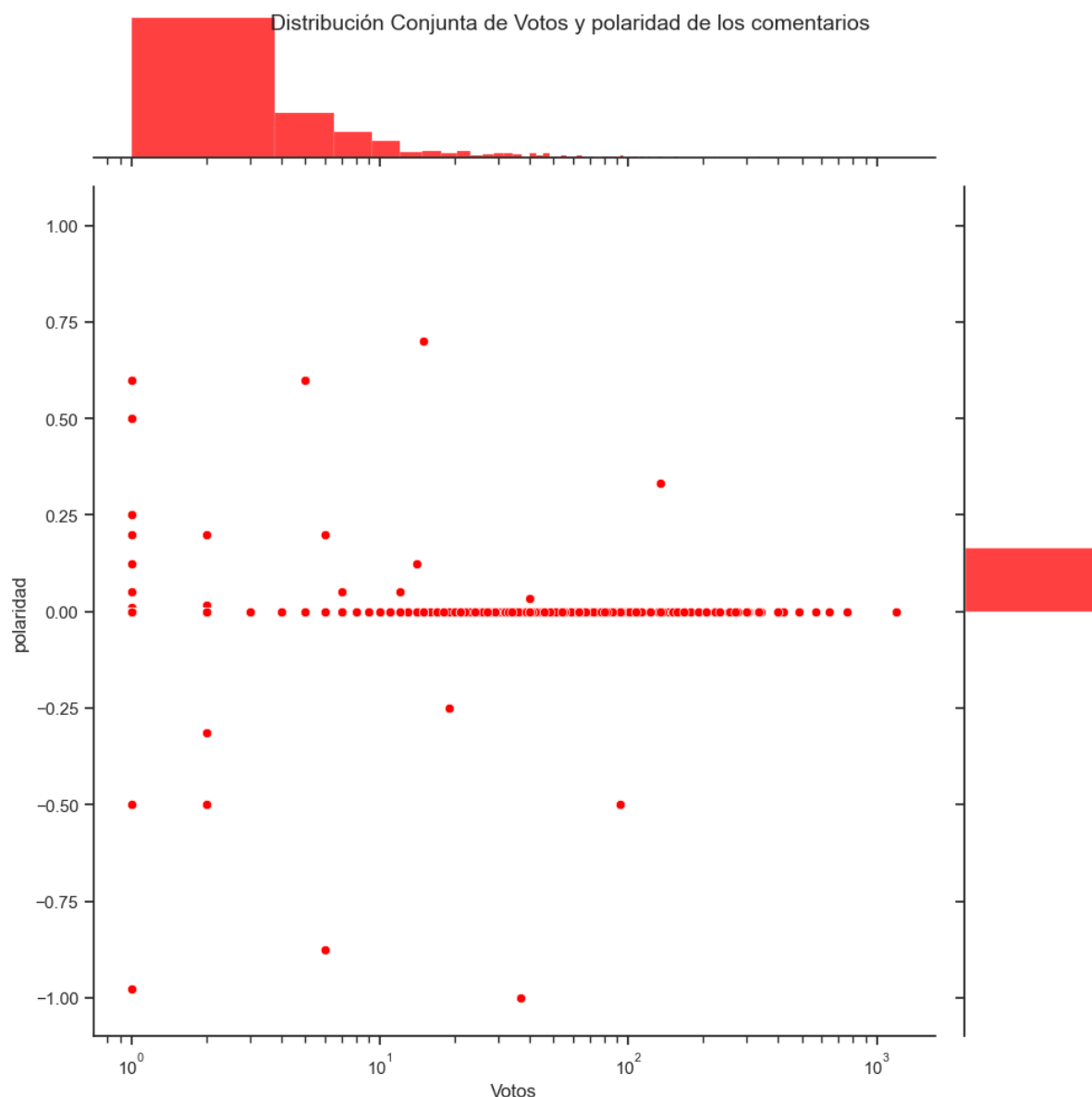


Figura 3.2: Distribución de votos y la polaridad asociada con el interés y aprobación del publico.

Al revisar la distribución de los comentarios mas votados y su correspondiente valor de polaridad podemos notar una polaridad dispersa en comentarios con un menor numero de votos, pudiendo significar esto la poca aprobación de comentarios demasiado negativos o positivos.

En cuanto a la distribución conjunta de frecuencias de votos y objetividad, podemos destacar la los comentarios con menor numero de votos presentan subjetividad en el texto, sin embargo la distribución de comentarios con algún grado de subjetividad presentan una distribución dentro del rango de 0 a 300 votos, encontrándose lejos del top 20 de comentarios mas votados.

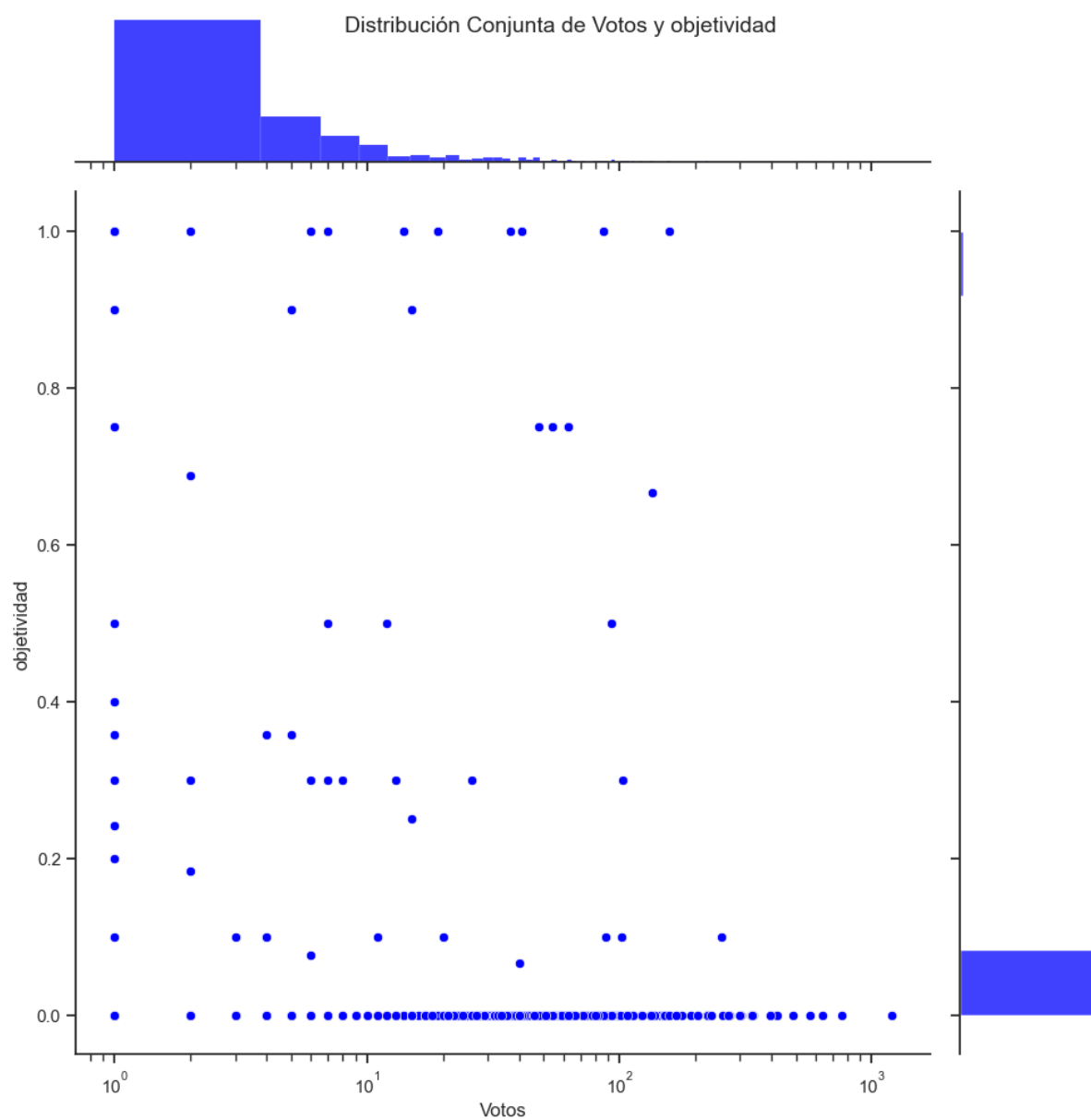


Figura 3.3: Distribución de la frecuencia de los votos y la objetividad de los comentarios

4. CONCLUSIONES

Se puede concluir que los comentarios objetivos y neutrales son aquellos mas votados por el publico que observo y participo en el chat del debate.

De igual forma se observa un apoyo menos concentrado en aquellos comentarios que presentan algún sesgo de preferencia o repudio sobre algún candidato, tema, proceso, etc...

Estas observaciones

BIBLIOGRAFÍA

- [1] S. Loria, Tutorial: Quickstart, TextBlob.
URL `https://textblob.readthedocs.io/en/latest/quickstart.html#quickstart`
- [2] M. A. R. Quiroga, D. V. Ayala, D. Pinto, M. Tovar, B. Beltrán, Análisis de sentimientos basado en aspectos: un modelo para identificar la polaridad de críticas de usuario, *Research in Computing Science* 115 (2016) 171–180.