

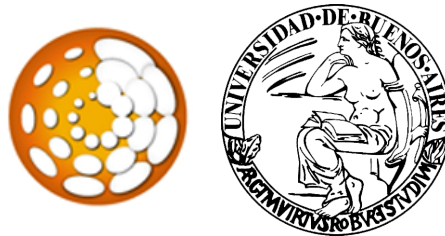
Aplicación de modelos de lenguaje para la identificación de emociones en Twitter durante las elecciones presidenciales de 2022 en Colombia

Tesis presentada para optar por el título de

**Magister en Explotación de Datos y
Descubrimiento del Conocimiento**

por

Juan Jose Iguaran Fernandez



Universidad de Buenos Aires

**Facultad de Ciencias Exactas y Naturales
Departamento de Ciencias de la Computación**

[Insert Month and Year]

AGRADECIMIENTOS

Agradezco

Resumen

El estudio de Twitter como medio para el análisis de fenómenos sociales mediante técnicas de procesamiento del lenguaje natural (NLP) es de gran interés. Dentro de estas técnicas, la detección de emociones en el texto es especialmente relevante, ya que permite identificar la respuesta subjetiva de las personas ante los eventos sociales. Los modelos de lenguaje basados en redes neuronales, como los Transformers y específicamente BERT, han reemplazado las técnicas tradicionales de NLP debido a su capacidad para capturar el contexto y las relaciones entre las palabras en el texto. Sin embargo, en español, hay pocos estudios de detección de emociones en redes sociales y ninguno que utilice modelos de lenguaje basados en Transformers en un contexto político. Este trabajo tiene como objetivo emplear modelos basados en BERT para detectar emociones en Twitter durante las elecciones presidenciales de Colombia en 2022. Se utilizará una base de datos de tweets descargados a través de la API, etiquetados manualmente con emociones y se realizará un fine tuning de varios modelos de lenguaje preentrenados para seleccionar el mejor. Este modelo se utilizará para clasificar toda la base de datos y analizar la respuesta emocional de los tweets asociados a sectores políticos y su variación temporal.

Palabras Clave: [BERT, Colombia, Elecciones, Detección de emociones]

Índice general

Índice de figuras	4
Índice de cuadros	6
1. Introduccion	1
2. Trabajos Previos	4
2.1. Aprendizaje supervisado en la detección de emociones	4
2.2. Redes neuronales para el análisis de texto	6
2.3. Sentimientos y emociones en redes sociales	10
3. Metodología	13
3.1. Datos	13
3.1.1. Obtención y análisis exploratorio de datos	13

3.1.2. Etiquetado	16
3.2. Modelos	24
3.2.1. Modelos pre entrenados	24
3.2.2. Entrenamiento de los modelos	24
3.2.3. Evaluación de los modelos	26
4. Resultados	29
4.1. Distribución de emociones por sector	30
4.2. Emociones a lo largo del tiempo	32
5. Conclusiones	38
Bibliografía	41
A. bla	48
A.1. Anexos	48

Índice de figuras

2.1. Arquitectura de los Transformers	8
2.2. Arquitectura de BERT	9
3.1. Porcentaje de Hashtags clasificados según sector político	16
3.2. Porcentaje de Tweets clasificados según sector político	17
3.3. Porcentaje de tweets clasificados según sector político a lo largo del tiempo	18
3.4. Interfaz de etiquetado	23
3.5. Porcentaje de coorelación entre etiquetadores de las emociones asignadas a los tweets'	28
4.1. Matriz de confusión entre etiquetas asignadas y predicciones del algoritmo para el dataset etiquetado	31
4.2. Porcentaje de tweets clasificado según emociones por el algoritmo	33

4.3. Porcentaje de tweets clasificados según emoción por el algoritmo, separado por sector político del hashtag	34
4.4. Porcentaje diario del total de tweets clasificados en cada emoción a lo largo del tiempo	35
4.5. Porcentaje diario del total de tweets clasificados como Alegría, para cada sector político	36
4.6. Porcentaje diario del total de tweets clasificados como Miedo, para cada sector político	36
4.7. Porcentaje diario del total de tweets clasificados como Tristeza, para cada sector político	37
4.8. Porcentaje diario del total de tweets clasificados como Asco, para cada sector político	37

Índice de cuadros

3.1. Ejemplos de tweets con respecto a orientacion politica	15
3.2. Descripción de las emociones usadas	19
3.3. Indice de Fleiss Kappa en cada emocion	24
3.4. Descripción de los modelos usados	25
4.1. Metricas de los modelos entrenados	30
4.2. Ejemplos de tweets clasificados	32
A.1. Hashtags	48

Capítulo 1

Introduccion

Twitter ha sido analizado como un medio particularmente interesante para el estudio indirecto de fenómenos sociales a través del uso de técnicas del procesamiento del lenguaje natural (NLP) pues es capaz de captar a una gran cantidad de usuarios, que discuten sobre una gran numero de tópicos, usando en una cantidad limitada de palabras con estilo propio. Esto ha sido estudiado en el pasado por ejemplo, al analizar como los textos provenientes de este medio, coinciden con lo que arrojan otras aproximaciones de las realidades sociales tales como las encuestas de opinión, discursos políticos o eventos de relevancia popular.

Una tarea en particular dentro del NLP, la detección de emociones, resulta de particular interés al estudiar la respuesta individual a fenómenos sociales, pues dentro del texto existe tanto información objetiva, que corresponde a hechos verificables, como información subjetiva, que corresponde a los procesos internos que los individuos experimentan y son plasmados en el texto, tal como las opiniones. Las emociones son parte de esta información subjetiva, y en ese contexto, la detección de las emociones

presentes en el texto, se define como un sub campo del análisis de sentimiento, que busca determinar la polaridad y el grado de las distintas dimensiones de la subjetividad presentes en el texto.

El estudio del análisis de sentimientos en general y de emociones en particular, se ha hecho usualmente a través de técnicas tradicionales de NLP tales como el empleo de modelos de aprendizaje supervisado usando de features construidos a partir del texto. Durante los últimos años, estas técnicas están siendo remplazadas por modelos de lenguajes usando redes neuronales, en particular arquitecturas como los Transformers, debido a su capacidad de tener en cuenta el contexto dentro del cual se encuentran las palabras en el texto, es decir, su relación con otras palabras. Una aplicación particular de los Transformers BERT, es particularmente útil pues consiste en un modelo pre entrenado que es capaz de desempeñar la tarea de NLP particular, para la cual se haga un entrenamiento final.

En español existen pocos estudios de detección de emociones en redes sociales, y no se conoce de ninguno que use modelos de lenguaje basado en Transformers para este fin en un contexto político. EL presente trabajo tiene por objetivo el empleo de modelos de lenguaje, específicamente BERT que es un una red neuronal pre entrenada con la Wikipedia basada en Transformers para detectar emociones presentes en twitter durante las elecciones presidenciales en Colombia en 2022.

Para poder obtener los datos de entrenamiento, se hará uso de una base de datos obtenida mediante la descarga tweets, a través de la API, utilizando hashtags relacionados al tema de las elecciones, entre el 22 de mayo y el 22 de junio del 2022, periodo que contiene la primera y segunda vuelta presidencial de las elecciones en Colombia. Estos hashtags, y por consiguiente los tweets relacionados a estos, serán

asociados por el autor a los sectores políticos izquierda, derecha o neutro, basándose en el contenido observado en los tweets y el apoyo o rechazo que estos muestren. A partir de ahí, se procederá a tomar una muestra de 1200 tweets para ser etiquetados manualmente por el autor y los directores, en alguna de las emociones disponibles. Con los tweets etiquetados, se procederá a realizar el fine tuning de varios modelos de lenguaje pre entrenados y evaluar su desempeño para escoger el mejor. Este modelo, será utilizado para clasificar toda la base de datos y a partir de esta clasificación, realizar agrupaciones que permitan determinar la respuesta emocional de los tweets asociados a cierto sector político, así como la variación temporal de esta respuesta.

Capítulo 2

Trabajos Previos

2.1. Aprendizaje supervisado en la detección de emociones

La detección automática de las respuestas internas que las personas reflejan en el texto, ante cierto fenómenos, ha sido estudiada con mucho interés en los últimos años. Es así como en Pang et al. (2002) se establece la importancia de desarrollar, ante textos que reflejen opiniones subjetivas, sistemas capaces de identificar si dicha opinión es negativa o positiva. Para este propósito, se emplea el dominio de las reviews online de películas, que facilitan la tarea, al contar con una clasificación negativa o positiva por parte del usuario, construyendo algoritmos de aprendizaje supervisado que utiliza como features principalmente unigramas, que es un elemento individual del texto y como variable objetivo la clasificación asignada. Por otro lado, Turney (2002) pretende establecer a través de aprendizaje no supervisado, si determinada

review sobre diversos temas online, presenta un sentimiento negativo o positivo. Para ello, emplea el concepto de orientación semántica presente en Hatzivassiloglou and McKeown (1997) para determinar si una frase tiene orientación negativa o positiva y así poder determinar si la review en su conjunto es positiva o negativa. En general, la determinación de la orientación negativa o positiva dentro del texto, es conocida como análisis de sentimiento

Así como en el análisis de sentimiento, contar con un corpus de términos que cuenten con una clasificación previa de su orientación, es clave para poder llevar a cabo automáticamente la tarea, para los estados emocionales es igualmente importante, por lo que Strapparava et al. (2004) realiza una anotación manual de estados emocionales basados en las categorías de Ortony et al. (1987) sobre algunos términos encontrados en WordNet Miller (1995), que es una base de datos de términos en inglés agrupados por grupos de sinónimos y con relación semántica entre grupos. A partir de las anotaciones manuales de las emociones de ciertos términos presentes en la base, se establece la categoría emocional de nuevos términos gracias a los sinónimos y las relaciones semánticas, construyendo así una base de datos de estados emocionales llamada WordNet-Affect. Dentro de esta misma línea, Wiebe et al. (2005) elabora una anotación manual de los estados emocionales presentes en las oraciones de un gran volumen de noticias, en donde se tiene en cuenta el contexto.

En el contexto de la clasificación automática de las emociones en el texto, Alm et al. (2005) utilizan los cuentos infantiles para desarrollar un modelo de aprendizaje supervisado capaz de detectar emociones en las frases. Para ello se elabora una anotación manual de estas frases que constituyen el set de datos y luego, se generan un set de features para dichas frases que pasaran a entrenar un clasificador lineal.

2.2. Redes neuronales para el análisis de texto

Las técnicas tradicionales de análisis de emociones en el texto, están principalmente basadas en la relación que existe entre los términos que constituyen el texto y determinado estado emocional. Este tipo de análisis, si bien ha sido de gran utilidad, no dispone de la capacidad para tener en cuenta el contexto, es decir la relación que existe entre las palabras que constituyen el texto, además del orden en que estas se presentan. Debido a esto, recientemente el campo de estudio ha ido migrando hacia el uso de los algoritmos capaces de captar esta relación contextual, específicamente las redes neuronales.

Haciendo un recuento del estado del arte de la detección de emociones en el texto, Acheampong et al. (2021) hablan de las distintas arquitecturas de redes neuronales usadas recientemente para esta tarea, entre ellas las RNN (redes neuronales recurrentes). Estas redes tienen un tipo de arquitectura cuyo uso es especial para datos secuenciales, tales como las tareas de NLP. En estas, para un ejemplo nuevo, es posible utilizar el resultado del procesamiento de un dato anterior, para el del siguiente dato (como en una secuencia de caracteres por ejemplo), es decir la misma topología de pesos recibirá para cada ejemplo nuevo, el resultado anterior además de la nueva entrada. Sin embargo, debido a su arquitectura, la optimización de los pesos que constituyen la topología de las RNN, se vuelve compleja para el cálculo del gradiente del error, pues se deberá pasar tantas veces por los pesos de la red como pasos en el tiempo haya (como número de caracteres en una palabra) haciendo que en secuencias particularmente largas, este cálculo crezca en demasía.

Si bien las RNN permiten preservar de alguna manera la relación contextual del texto, al tener una estructura secuencial, se impide la paralelización de su cómputo,

pues se necesita las salidas de ejemplos anteriores para llevar a cabo el siguiente paso en el tiempo. Además, debido a este mismo funcionamiento secuencial, si existe un gran número de pasos en el tiempo, es poco probable que la red tenga en cuenta información presente al inicio. Para solucionar estos inconvenientes, Vaswani et al. (2017) desarrollan los Transformers, que son un tipo de arquitectura en la que múltiples datos de entrada son ingresados a la red de manera simultánea, como todas las palabras en una frase por ejemplo. Esto se logra mediante la representación de dichas palabras en la forma de un vector y a través de una arquitectura que se define como atención. A partir de ahí, se realiza una transformación de los datos, de manera que la representación de cada dato, en este caso cada palabra, tenga en cuenta que tan importante es la misma para todas las demás palabras de la oración. Esta transformación, aplicada tanto en los datos de entrada como en los de salida, es luego usada para entrenar los pesos de la red. De esta manera la red es capaz de procesar datos en simultáneo así como tener en cuenta toda la información disponible. Un esquema de esta arquitectura se puede apreciar en la imagen 2.1 Vaswani et al. (2017).

Gracias a la capacidad de paralelización y de conservar la relación entre palabras distantes del texto que presentan los Transformers, es posible entrenar un modelo de lenguaje con grandes volúmenes de información para de esta manera, contar con una red con un gran poder predictivo. Es así como Devlin et al. (2018) desarrolla un modelo basado en Transformers capaz de aprender el contexto del lenguaje de una manera general y luego utilizar lo aprendido para distintas tareas de NLP denominado BERT (Bidirectional Transformers for Language Understanding). Para ello se utilizó una arquitectura en donde las capas de entrada son las representaciones vectoriales de las palabras (Embeddings) y a partir de ahí, contiene múltiples capas de Transformers. Para su entrenamiento, se suministran dos oraciones consecutivas

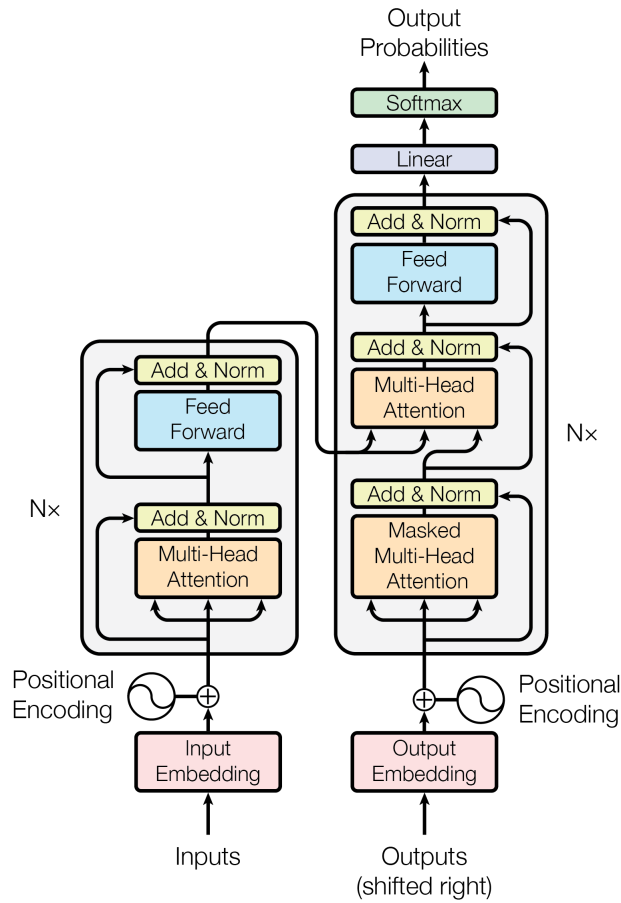


Figura 2.1: Arquitectura de los Transformers

con palabras faltantes y se asignan dos tareas simultaneas: saber que palabra podría ser la faltante, además de identificar el orden de las oraciones. Esto le permite al modelo aprender del contexto del lenguaje a nivel de palabras usando las que vienen antes y después como fuente de información, así como del contexto de las oraciones al identificar su orden. Para ello se utilizó toda la wikipedia. Este modelo pre entrenado es luego capaz de utilizar esta representación contextual del lenguaje para realizar distintas tareas de NLP tales como análisis de sentimiento, al agregar una capa final a la red que produzca un posible resultado, una función softmax por ejemplo que

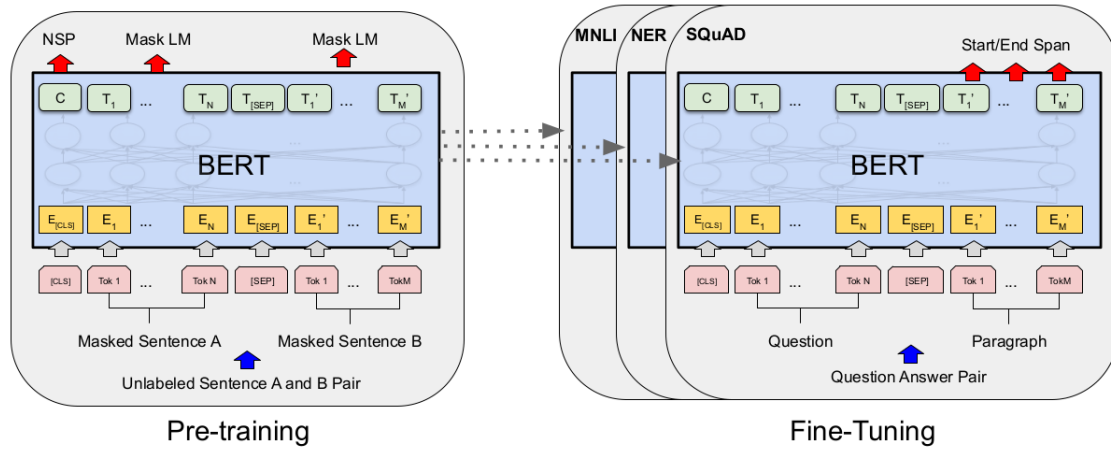


Figura 2.2: Arquitectura de BERT

permite asignar la probabilidad de pertenencia a algunas de las clases y comparar este resultado con su respectivo valor esperado. El esquema de esta arquitectura se puede apreciar en 2.2 Devlin et al. (2018).

En cuanto a la aplicación de modelos tipo BERT en español, en Canete et al. (2020), los autores resaltan que no existía hasta la fecha un modelo de este tipo entrenado específicamente para este lenguaje, además del BERT para múltiples lenguajes. Por ello, se proponen entrenar dicho modelo usando además de la Wikipedia en español, texto de publicaciones de las naciones unidas, gobiernos y charlas TED. El resultado fue un modelo que supera al BERT en múltiples idiomas para el español en casi todas las tareas evaluadas.

2.3. Sentimientos y emociones en redes sociales

El campo de estudio de análisis de sentimiento del texto en Internet, ha venido ganando importancia tal como lo describe Pang et al. (2008) , tanto para usuarios individuales como para la industria de la publicidad, el mercado financiero y la academia, por lo que se hace un recuento de las distintas técnicas y aplicaciones que son consideradas relevantes por los autores hasta la fecha.

Un espacio particular en donde los usuarios individuales pueden generar grandes volúmenes de información sobre distintos contenidos, y por consiguiente, una fuente rica de datos para el análisis de sentimiento son los blogs. Por esta razón, Aman and Szpakowicz (2007) utiliza texto proveniente de ellos para realizar detección de emociones presentes en las oraciones de los mismos. Para este propósito recurren primero a una anotación manual de las mismas y luego a la construcción de features para entrenar distintos modelos supervisados.

Twitter, un sitio de blogs en particular cuyo formato es el microblogging, es decir, publicaciones pequeñas que se denominan tweets, ha cobrado una gran relevancia en el estudio del texto debido a su gran popularidad entre los usuarios de Internet de todo tipo, desde marcas hasta usuarios individuales y políticos, para abarcar cualquier tema. En ese contexto, Pak and Paroubek (2010) extrae tweets de diferentes usuarios sobre distintos temas para realizar un análisis de sentimientos sobre estos. Para ello se procede a la identificación de tweets que contengan emoticones felices y tristes, los cuales etiqueta como reflejando un sentimiento positivo o negativo respectivamente, así como tweets provenientes de cuentas de medios de noticias los cuales etiqueta como neutrales. luego se procede a la construcción de features usando n-gramas, a partir de las palabras presentes en el tweets con los que se entrenan varios

clasificadores.

En Roberts et al. (2012) se identifica la necesidad de contar con un corpus etiquetado que sirva de base para la tarea de identificación de emociones en twitter. Para ello, se seleccionan 14 temas que para los autores tienen un fuerte contenido emocional y las palabras clave asociados a estos para ser usados como hashtags en la extracción. A partir de ahí, manualmente se etiquetaron los tweets con su respectiva emoción. Esto sirvió de base para entrenar un modelo de aprendizaje supervisado y así verificar el poder predictivo de los modelos partiendo de datos etiquetados.

La identificación del sentimiento presente en Twitter puede servir como fuente de aproximación a la realidad social en la que los usuarios se desenvuelven. Por esto, O'Connor et al. (2010) se plantea la pregunta si existe una correlación entre el sentimiento encontrado en twitter y las encuestas de opinión. Para ello, se toma una muestra de mil millones de tweets entre 2008 y 2009 y se toman aquellos tweets que contengan palabras claves asociadas a los temas que se están investigando. EL sentimiento de estos tweets se determina a partir de la proporción de palabras con asociación negativa o positiva presentes en el tema que se esta analizando en un día en particular. Los resultados dan una correlación alta entre el sentimiento encontrado a través del texto y las encuestas. Esta relación entre la realidad social y el sentimiento presente en los tweets es también abordada por Bollen et al. (2011), donde se procede a realizar una medida del estado emocional de una muestra de tweets entre agosto y diciembre del 2008, en donde el mismo se mide a través de la similaridad presente entre las palabras de los tweets y ciertos términos claves asociados a estados emocionales. Esto permite encontrar que determinados eventos relevantes generan un impacto emocional significativo y durante un periodo de tiempo considerable en los usuarios.

Al ser el escenario político un caso particular de los fenómenos sociales, cuyo impacto se traduce en el estado emocional de las personas, el sentimiento presente en tweets de contenido político puede dar un indicio de la percepción pública de dicho escenario. Con esto en mente, Tumasjan et al. (2010) se planteó el uso de twitter como plataforma de medición de la percepción pública respecto a la política durante las elecciones parlamentarias en Alemania en el 2009. Una de sus preguntas de investigación estuvo relacionada con los sentimientos que se reflejan en los tweets que mencionan a los políticos que hacen parte de las campañas, y para esto, usando el texto proveniente de twitter, se utilizó un software capaz de identificar palabras claves asociadas a estados emocionales y cognitivos en el texto. El resultado fue un perfil emocional para cada político que en líneas generales, parece estar de acuerdo con su discurso político.

En Mohammad et al. (2015) se utilizan las elecciones presidenciales de 2012 en Estados Unidos para elaborar un corpus emocional a partir de tweets relacionados a este tema y etiquetados manualmente con 19 emociones, que luego son agrupadas en 8. Estas tweets sirven para entrenar un modelo de aprendizaje supervisado, siendo de esta manera el trabajo más cercano y una referencia importante para el presente trabajo.

Capítulo 3

Metodología

3.1. Datos

3.1.1. Obtención y análisis exploratorio de datos

El dataset obtenido esta inicialmente constituido por 585001 tweets, recolectados entre el 22 de mayo y el 22 de junio del 2022, periodo dentro del cual ocurrieron la primera y la segunda vuelta de las elecciones presidenciales en Colombia , el 29 de mayo y 19 de junio respectivamente. Para la extracción de los datos, se utilizaron 173 hashtags con contenido político que tuvieron lugar durante este periodo. Este dataset fue filtrado para remover aquellos tweets que tuvieran menos de 5 palabras, aquellos que tuvieran una proporción de menciones o hashtags mayor al 20 % del total del texto y aquellos que tuvieran links o que provinieran de usuarios con un numero atípico de posteos. Esto redujo la base a 193348 tweets. Los hashtags utilizados

fueron etiquetados en uno de tres sectores políticos: Izquierda, Derecha y Neutro, dependiendo del contenido asociado a dichos hashtags y su tendencia política. En la tabla A.1 se muestran los hashtags, así como su etiqueta política y cuantos tweets hubo para cada hashtags. En adelante durante el presente trabajo, cuando se hable de sector político, se estará haciendo referencia a esta clasificación.

Cabe resaltar que la asignación de un sector político en particular a un hashtag, fue llevado a cabo a partir de la tendencia política observada en los tweets, lo que no excluye sin embargo la presencia de tweets cuya tendencia política sea contraria. En la tabla 3.1 se pueden encontrar algunos ejemplos de tweets que muestran la relación del hashtag con la orientación política. El hashtag `#EstallidoSocialEs` por ejemplo, fue catalogado como de derecha por presentar en general tweets como el tweet 1 . Sin embargo, hay tweets que lo utilizan, en donde se aprecia un apoyo al candidato de izquierda como el tweet 2 . Del mismo modo el hashtag `#YaEsSuficiente` es utilizado en general como apoyo a la izquierda, como en el tweet 3. Sin embargo, también hay casos en donde se usa como apoyo un candidato de derecha como en el ejemplo 4.

La distribución de los hashtags a través de los sectores políticos se puede evidenciar en el gráfico 3.1 donde se muestra que el sector neutro tiene mas de un 40 % del total de los hashtags, mientras que la izquierda y la derecha, tienen una cantidad semejante, de aproximadamente un 28 % cada uno.

De manera similar, el gráfico 3.2 muestra la distribución de tweets a lo largo de los sectores políticos. Allí se aprecia que el sector neutro tiene la mayoría de los tweets, con mas del 46 %, luego se encuentra la izquierda con un 33 % y finalmente la derecha con cerca de un 29 %,

Al analizar la cantidad de tweets por sector político a lo largo del tiempo, se

Cuadro 3.1: Ejemplos de tweets con respecto a orientacion politica

Numero de Ejemplo	Tweet
1	@lvelez @lafm #EstallidoSocialEs el arma de terror de Petro para obligar a votar por él.
2	Colombia va por el cambio, a redoblar esfuerzos estos 4 días para derrotar a la corrupción. #PetroYFranciaSonElCambio #EstallidoSocialEs
3	#YaEsSuficiente Que los medios proclives al gbno, traten de darle aire boca a boca a un moribundo electoral, FICO. Ante su estancamiento en las encuestas y la distancia que le ha tomado Petro, pretenden en 1 acto de desesperacion, el insuflarlo de votantes de los cuales carece.
4	#YaEsSuficiente de mentir sobre @ingrodolfohdez , vayan a bucaramanga, vean lo que hizo y ahí si hablen..
5	#LoPeorDeEstasElecciones es la división de la gente de este país esta gente y cosas ya que pasan
6	@Zuletalleras Son 4 años O es la. Derecha va realizar un golpe de estado? No cree a sus comentarios son ofensivos e incendiarios.... #PetroEsPresidente

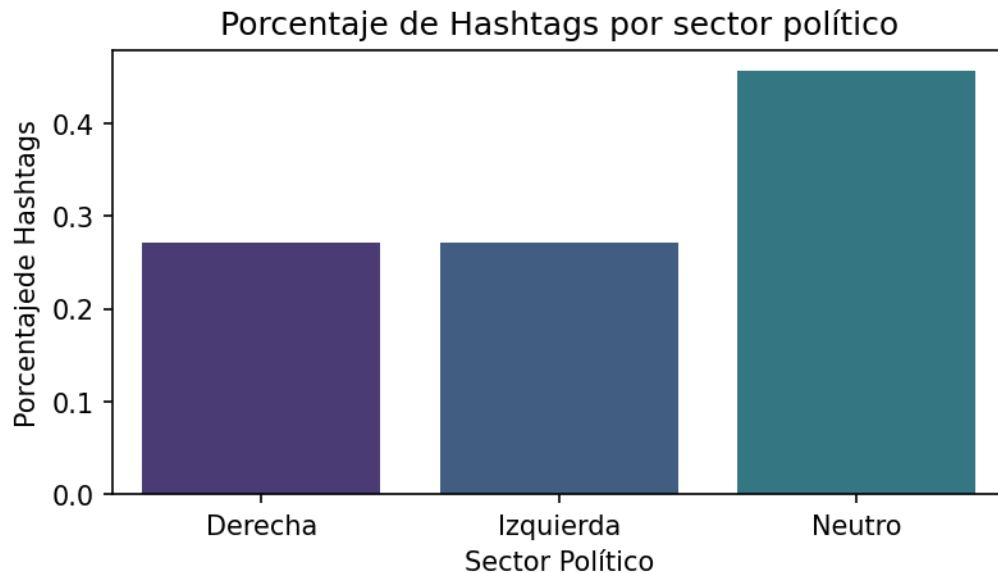


Figura 3.1: Porcentaje de Hashtags clasificados según sector político

obtienen los resultados observados en el gráfico 3.3, en donde se muestra para cada sector, que porcentaje del total de los tweets tuvo cada día. Se puede evidenciar que hubo algunas fechas particularmente importantes: el 24 de mayo fue el día de un debate y sobresale el sector neutro, el 29 de mayo fue el día de la primera vuelta y sobresalen los tres, el 9 de junio fue el día en el que salieron a la luz los llamados Petro videos, que fueron unos videos filtrados en donde se ve al equipo de campaña de Petro discutiendo estrategias políticas y sobresale la derecha y las fechas cercanas al 19 de junio que fue la segunda vuelta, sobresalen los tres.

3.1.2. Etiquetado

Para proporcionar un dataset que sirva de base al eventual entrenamiento del modelo , se procedió al etiquetado manual de 1200 tweets escogidos mediante una

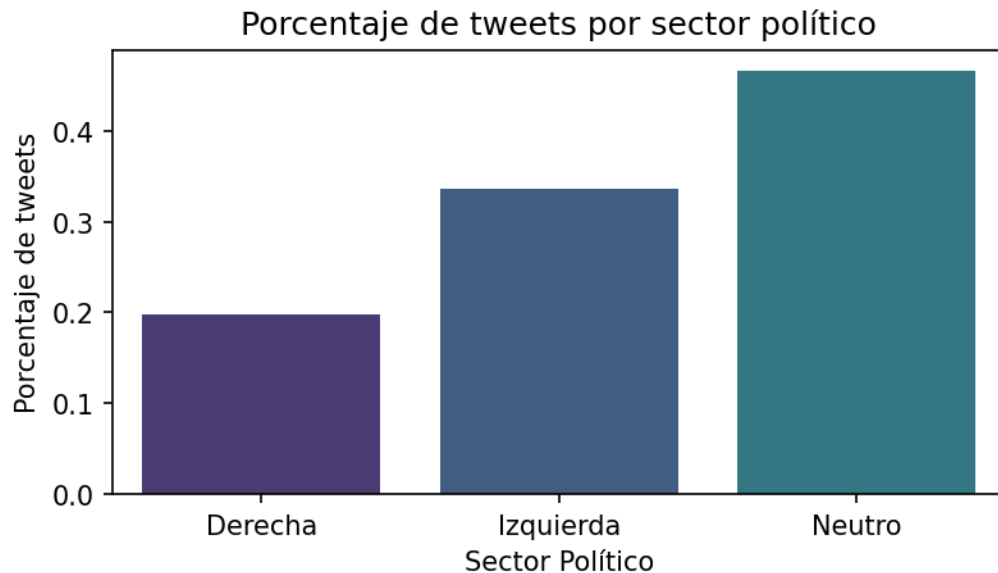


Figura 3.2: Porcentaje de Tweets clasificados según sector político

muestra aleatoria estratificada proporcional a la cantidad de tweets en los hashtags. Las etiquetas a elegir fueron las emociones presentes en el tweet.

Una emoción es un patrón de reacción complejo, que involucra elementos experienciales, conductuales y fisiológicos, mediante el cual un individuo intenta lidiar con un asunto o evento personalmente significativo. La cualidad específica de la emoción (por ejemplo, miedo, ira) está determinada por el significado específico que el individuo asigna al evento. Por ejemplo, si el evento implica amenaza, es probable que se genere miedo. Están comprendidas de tres componentes distintos: una experiencia subjetiva, una respuesta fisiológica y una respuesta conductual o expresiva. Para el contexto del presente trabajo, la experiencia subjetiva y la respuesta fisiológica, no son componentes accesibles. Es así como la atención durante el etiquetado recayó en la respuesta expresiva, concretamente, como el autor del tweet, expresa a través del texto, la respuesta emocional generada en respuesta hacia la entidad en particular a

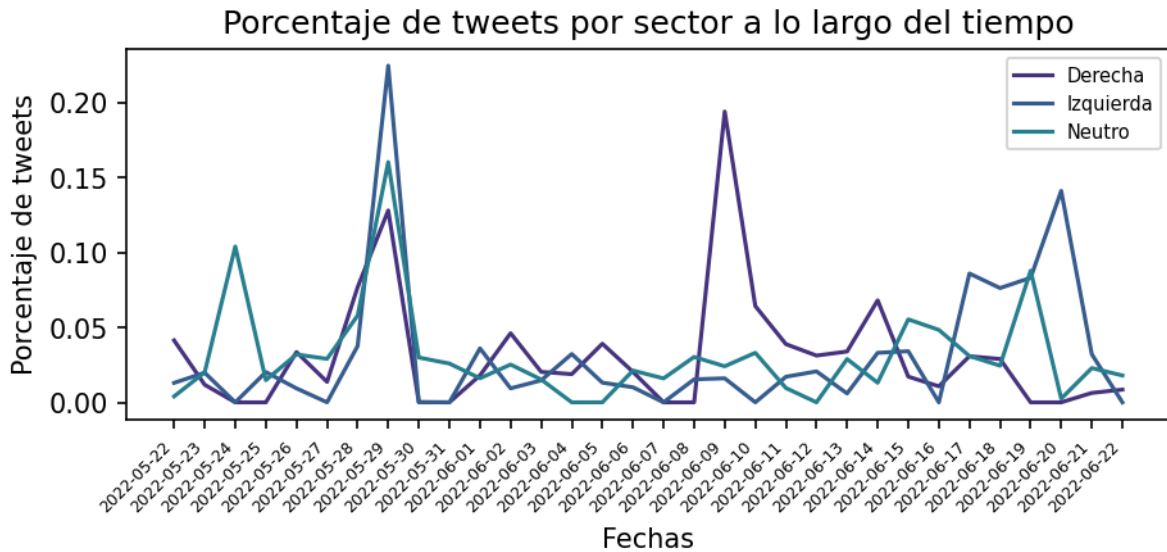


Figura 3.3: Porcentaje de tweets clasificados según sector político a lo largo del tiempo

la cual se dirige el tweet.

LA tarea de etiquetado concretamente consistió en etiquetar un tweet a la vez dentro de una plataforma de etiquetado, en donde se tuvo la posibilidad, mediante un esquema de selección múltiple, de asignar una o varias emociones al tweet, basado en los lineamientos presentes en el manual de etiquetado ¹. Esta tarea fue llevado a cabo por el autor y los directores, quedando así 3 etiquetas independiente para cada uno de los tweets.

Para el esquema de etiquetado, se uso como referencia lo elaborado por Mohammad et al. (2015), en donde el etiquetador responde varias preguntas, y al ser preguntado por la emoción presente en el tweet, puede escoger 19 emociones distintas. En el presente trabajo a través de una prueba iterativa para elegir las etiquetas mas

¹https://docs.google.com/document/d/1hoUYKMaYHSeGeOQ2FqRVyahTin6T09Mu8wt1_HY6200/edit?usp=sharing

adecuadas, se llegó a escoger 14 emociones posibles además de la categoría otra: Alegría, Agrado, Confianza, Admiración, Miedo, Incertidumbre, Sorpresa, Asombro, Tristeza, Decepción, Asco, Desagrado, Ira, Odio, Otra. En la tabla 3.2 se presenta una descripción de las emociones usadas.

Cuadro 3.2: Descripción de las emociones usadas

Emoción	Descripción
Admiración	La admiración es una emoción social que se siente al observar a personas de competencia, talento o habilidad que superan los estándares. La admiración facilita el aprendizaje social en grupos. La admiración motiva la superación personal a través del aprendizaje de los modelos a seguir.
Agrado	Sensación moderada de felicidad o placer que siente una persona por algo que le gusta.
Confianza	La confianza es la voluntad de una parte (el fideicomitente) de volverse vulnerable a otra parte (el fideicomisario) bajo la presunción de que el fideicomisario actuará de manera que beneficie al fideicomitente. Además, el fideicomitente no tiene control sobre las acciones del fiduciario. Los académicos distinguen entre la confianza generalizada (también conocida como confianza social), que es la extensión de la confianza a un círculo relativamente grande de personas desconocidas, y la confianza particularizada, que depende de una situación específica o una relación específica.
Alegría	Es una emoción positiva que suele ir acompañada de bienestar y alegría. Se genera como resultado de un evento positivo. Se diferencia de la confianza en el grado de entusiasmo y gratitud manifestado por el autor.

Continued on next page

Cuadro 3.2: Descripción de las emociones usadas

Emoción	Descripción
Incertidumbre	La incertidumbre es la falta de seguridad, de confianza o de certeza sobre algo. Aparece en situaciones en las que no tenemos control total, en las que nos faltan respuestas e información, y nos puede generar inquietud, inseguridad, estrés, ansiedad e incluso miedo
Miedo	El miedo surge con la amenaza de daño, ya sea físico, emocional o psicológico. lógica, real o imaginaria. En el texto suele aparecer como una puesta de manifiesto de una amenaza hacia el autor del tweet o hacia lo que éste se refiere, en una situación donde se encuentra relativamente indefenso o en desequilibrio de poder. Puede ser también percibida como incertidumbre respecto al futuro.
Asombro	La condición de estar asombrado; un estado de asombro abrumador, como por sorpresa o miedo repentino, horror o admiración.
Sorpresa	Se define como una reacción provocada por algo inesperado, extraño o novedoso para la persona. En el texto está principalmente asociada a resultados inesperados o descubrimientos singulares respecto al enunciado del tweet.

Continued on next page

Cuadro 3.2: Descripción de las emociones usadas

Emoción	Descripción
Decepción	La decepción es el sentimiento de insatisfacción que sigue al fracaso de las expectativas o esperanzas de manifestarse. Similar al arrepentimiento, se diferencia en que una persona que se arrepiente se enfoca principalmente en las elecciones personales que contribuyeron a un mal resultado, mientras que una persona que se siente decepcionada se enfoca en el resultado mismo. Es una fuente de estrés psicológico.
Tristeza	Esta emoción es una especie de dolor emocional o estado afectivo causado por la decadencia espiritual y que muchas veces se expresa en llanto, abatimiento, falta de apetito, cansancio, etc. Una persona puede sentirse triste cuando no se cumplen sus expectativas o cuando las circunstancias de la vida son más dolorosas que alegres. La emoción opuesta es la alegría.
Desagrado	Una actitud o un sentimiento de disgusto o aversión.
Asco	Contiene una serie de estados con intensidades variables que van desde una leve aprensión hasta una intensa repulsión. Todos los estados de asco se desencadenan por la sensación de que algo es aversivo, repulsivo y/o tóxico.
Odio	El odio es una intensa respuesta emocional negativa hacia ciertas personas, cosas o ideas, generalmente relacionadas con la oposición o repulsión hacia algo. El odio a menudo se asocia con intensos sentimientos de ira, desprecio y disgusto. El odio a veces se ve como lo opuesto al amor.

Continued on next page

Cuadro 3.2: Descripción de las emociones usadas

Emoción	Descripción
Ira	Esta emoción surge cuando se nos impide lograr un objetivo y/o se nos trata injustamente. En su forma más extrema, la ira puede ser una de las emociones más peligrosas debido a su potencial conexión con la violencia. En el texto, el autor está en tono retador o en líneas generales haciendo un reclamo por un derecho vulnerado o exigiendo algún tipo de justicia.

Este etiquetado se llevo a cabo usando la plataforma web Label Studio. La interfaz de etiquetado se puede observar en la figura 3.4.

A partir de los resultados obtenidos, se obtuvo el gráfico 3.5, en donde se calcula la correlacion que tuvieron los etiquetadores en las distintas etiquetas.

Basándose en estos resultados, se construyeron 4 etiquetas agrupadoras, a partir de la correlacion que presentaron. Estas etiquetas fueron: Alegría que contiene las etiquetas Alegría, Agrado, Confianza y Admiración; Miedo que contiene las etiquetas Miedo e Incertidumbre; Tristeza que contiene las etiquetas Tristeza y Decepción; y Asco que contiene las etiquetas Asco, Desagrado, Ira y Odio. A cada tweet se le asigno una o varias de estas etiquetas finales, si al menos dos de los etiquetadores coincidían en dichas etiquetas. Estas fueron las etiquetas con las que se entreno el modelo.

El acuerdo observado entre los etiquetadores para cada una de estas 4 emociones, fue medido utilizando el indice de Fleiss Kappa. Los resultados se encuentran en el cuadro 3.3.

En mi casa esta prohibido hablar de política, por múltiples diferencias, pero por aquí, QUE VIVA PETRO!!!
#PetroPresidenteEnPrimera #PetroEsElCambio

¿Es relevante el presente tweet para el analisis de emociones durante la campaña presidencial?

☒ Si^[1] ☐ No^[2]

¿Existe contenido emocional en este tweet?

☒ Si^[3] ☐ No^[4]

¿Esta el texto escrito de manera ironica?

☐ Si^[5] ☒ No^[6]

¿Que emocion describe mejor el texto?

☐ Admiración^[7]
☐ Asombro^[8]
☐ Ira^[9]
☐ Decepción^[10]
☐ Asco^[11]
☐ Desagrado^[12]
☐ Miedo^[13]
☐ Odio^[14]
☐ Incertidumbre^[15]
☒ Alegría^[16]

☐ Agrado^[17]
☐ Tristeza^[18]
☐ Sorpresa^[19]
☐ Confianza^[20]
☐ Otra^[x]

¿Si en la pregunta anterior eligio Otra, con que emocion describiria mejor el tweet?

Figura 3.4: Interfaz de etiquetado

Se puede apreciar como alegría y asco tuvieron puntajes relativamente altos en comparación con tristeza y miedo. Esto se explica en parte por la cercanía que presentan estas dos ultimas a ascos, como se puede apreciar en el tweet numero 5 de la tabla 3.1. En el, dos de los etiquetadores coincidieron en asignar tristeza mientras el tercero asigno asco. Este mismo fenómeno ocurren con miedo como se aprecia en el tweet numero 6 de la tabla. Allí también fue el caso en donde uno de los etiquetadores asigno asco mientras los otros dos asignaron miedo.

Cuadro 3.3: Indice de Fleiss Kappa en cada emocion

	alegria	miedo	tristeza	asco
Cantidad de Tweets	464	98	103	580
Indice de Fleiss Kapp	0.69	0.47	0.4	0.62

3.2. Modelos

3.2.1. Modelos pre entrenados

Los modelos de lenguaje pre entrenados se encuentran disponibles para su uso en el sitio Hugging Face ². Desde allí, basándose en la relevancia del modelo para la tarea del presente trabajo, se escogieron 6 modelos diferentes para ser probados y escoger aquel que tuviera el mejor desempeño.

Los modelos escogidos fueron robertuito Pérez et al. (2022), bertin la Rosa et al. (2022), BETO Cañete et al. (2020), electricidad Romero (2020), roberta Fandiño et al. (2022a) y twitter-xlm Barbieri et al. (2022). En el cuadro 3.4 se encuentra una pequeña descripción de estos así como su localización dentro del hub de Hugging Face.

3.2.2. Entrenamiento de los modelos

Debido a que el fine tuning de estos modelos requiere un poder de computo importante, se decidió utilizar el servicio de Google Colab ³, en donde se pueden

²<https://huggingface.co/>

³<https://colab.research.google.com/>

Cuadro 3.4: Descripción de los modelos usados

Modelo	Localizacion	Descripcion
robertuito	pysentimiento/robertuito-base-uncased	Modelo de lenguaje preentrenado para contenido generado por usuarios en español, entrenado siguiendo las pautas de RoBERTa en 500 millones de tweets.
bertin	bertin-project/bertin-roberta-base-spanish	Modelo basados en RoBERTa entrenados desde cero en la parte española de mC4 usando Flax
electricidad	mrm8488/electricidad-base-discriminator	Electricidad-base-discriminator es un modelo base tipo Electra entrenado en un gran corpus español (también conocido como corpus de BETO)
beto	dccuchile/bert-base-spanish-wwm-cased	BETO es un modelo BERT formado en un gran corpus español. BETO tiene un tamaño similar a un BERT-Base y fue entrenado con la técnica de enmascaramiento de palabras completas.
roberta	PlanTL-GOB-ES/roberta-base-bne	El roberta-base-bne se basa en el modelo base de RoBERTa y ha sido preentrenado utilizando un corpus en español de 570 GB de texto limpio
twitter-xlm	cardiffnlp/twitter-xlm-roberta-base	Este es un modelo basado en XLM-roBERTa multilingüe entrenado en 198 millones de tweets y ajustado para el análisis de sentimientos.

desarrollar notebooks utilizando GPU de manera gratuita. Allí, se utilizó la librería Transformers de Huggingface para poder acceder a los modelos.

Para poder entrenar los modelos con esta librería, existen dos objetos importantes, el Tokenizer y el trainer. El primero, realiza una transformación de los datos de entrada en vectores tokenizados que puedan ser interpretados por el modelo. El segundo, recibe como entrada estos datos tokenizados, el modelo a utilizar, los hiperparametros y la métrica a utilizar para evaluar el modelo durante el entrenamiento. Tanto para la definición del tokenizer como la del modelo, se debe especificar la ruta del modelo en particular que se desea implementar.

Los modelos se entrenaron usando los hiperparametros por defecto disponibles en Hugging Face. Estos son principalmente Adamw como algoritmo de optimización, con un learning rate de 5e-05, un decay de 0.0, 3 como el número de epochs en train y un batch size por núcleo de 8. Estos pueden consultarse en el sitio web ⁴

3.2.3. Evaluación de los modelos

Los modelos fueron entrenados y evaluados utilizando la métrica Micro F1 score definida de la siguiente manera:

$$\text{Micro F1} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})}$$

En donde TP es el número de verdaderos positivos, FP el número de falsos positivos y FN el número de falsos negativos a lo largo de todas las clases. Se eligió esta

⁴https://huggingface.co/docs/transformers/v4.30.0/en/main_classes/trainer#transformers.TrainingArguments

métrica debido a que el modelo en cuestión permite la clasificación múltiple y con esta métrica se puede evaluar de manera simultanea el desempeño de las distintas clases.

Para entrenar y evaluar los modelos, los datos etiquetados se dividieron en train y test. Se opto por no realizar una tercera partición debido a que no se realiza una hiperoptimizacion de los modelos, sumado al tamaño del dataset etiquetado. Una vez particionados los datos, se procedió a entrenar cada uno de los 6 modelos, 10 veces y a guardar el resultado de su desempeño, con el objetivo de obtener una lectura mas robusta, dada la naturaleza aleatoria de las redes neuronales. El resultado del entrenamiento de cada modelo fue registrado en el sitio Wandb ⁵, cuyo propósito es servir de herramienta de acompañamiento al entrenar modelos.

⁵<https://wandb.ai/>

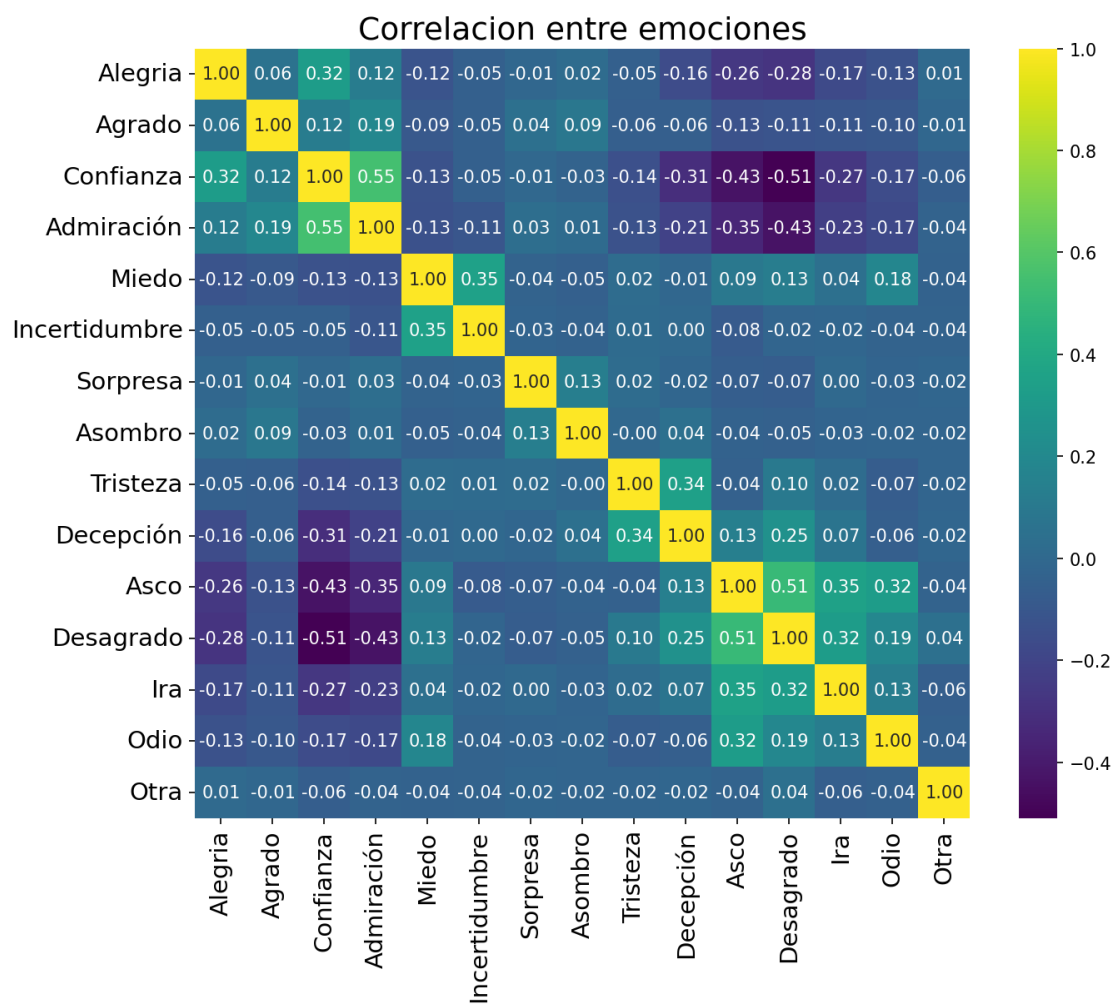


Figura 3.5: Porcentaje de coorelación entre etiquetadores de las emociones asignadas a los tweets'

Capítulo 4

Resultados

En el cuadro 4.1 se puede observar los resultados de la métrica micro F1 para los distintos modelos. Se presentan la media y la desviación estándar de las métricas para cada modelo pues se entrenaron 10 veces cada uno cambiando su semilla. Se observa que el de mejor comportamiento fue el robertuito-base-uncased con un micro f1 de 83.9 %, siendo así el modelo escogido. Esto se debe a que este modelo tiene un dominio de entrenamiento que se encuentra mas cercano al objetivo del siguiente trabajo, pues fueron tweets en español. Algunos de los demás modelos, como beto, estaba entrenado para español pero no para twitter y twitter-xlm para twitter para múltiples idiomas pero no específicamente para español. Así mismo, la tabla muestra el resultado de la métrica F1 para cada una de las emociones, en donde se puede apreciar que todos los modelos tienen buen desempeño para alegría y asco, sin embargo solo robertutito tuvo aciertos en miedo y fue el de mejor desempeño para tristeza.

En el gráfico 4.1 se presenta una matriz de confusión con los resultados de la clasificación sobre el conjunto de test, en donde se muestra el porcentaje de coinci-

Cuadro 4.1: Metricas de los modelos entrenados

Model	Media Micro F1	Alegria	Miedo	Tristeza	Asco
bertin	67.4 +- 23.8	81.5	0.0	0.0	85.2
beto	79.1 +- 1.0	86.9	0.0	36.4	87.3
electricidad	76.6 +- 1.3	82.5	0.0	0.0	84.6
roberta	79.9 +- 0.4	87.3	0.0	12.5	87.4
robertuito	83.9 +- 0.6	84.6	29.6	53.3	88.0
twitter-xlm	79.2 +- 1.2	86.4	0.0	0.0	91.6

dencia entre las etiquetas y las predicciones. Se muestra como muchos de los tweets etiquetados con miedo fueron predichos como asco. Esto se explica en parta por que estos tweets presentaban así mismo esta emoción, pero el modelo solo fue capaz de identificar el asco, como en el tweet numero 1 de la tabla 4.2

Este fue el caso también para tristeza, como se aprecia en el tweet numero 2 de la tabla 4.2. Para Alegría se observa que también algunos fueron predichos como asco. Lo que ocurrió en muchos casos es que en un mismo tweet, se expresaban dos emociones distintas, una en cada frase, como se aprecia en el tweet numero 3 de la tabla.

4.1. Distribución de emociones por sector

Una vez desplegado el modelo y obtenido las etiquetas emocionales para todos los tweets del dataset, se obtienen los resultados obtenidos en el gráfico 4.2, en donde se aprecia el porcentaje de tweets que recibió cada una de las distintas etiquetas. Allí puede apreciarse que la emoción mas preponderante fue el asco, con mas del 50 % de los tweets. Luego fue la alegría, con alrededor del 40 %. Finalmente, el miedo y la

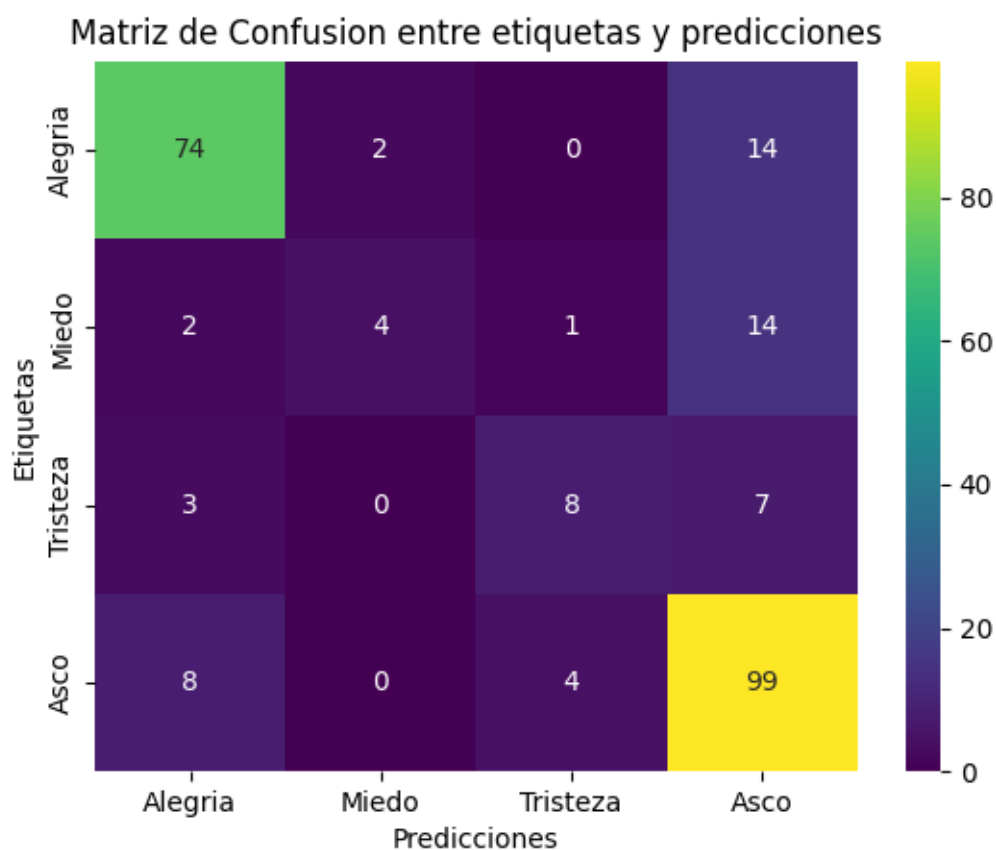


Figura 4.1: Matriz de confusión entre etiquetas asignadas y predicciones del algoritmo para el dataset etiquetado

tristeza estuvieron mucho menos presentes con menos del 3 % cada una.

Al analizar la presencia de cada emoción en los distintos sectores políticos, se obtuvieron los resultados que se presentan en el gráfico 4.3. En dicho gráfico se observa que el sector neutral fue el que mostró una mayor cantidad de tweets etiquetados con miedo, representando aproximadamente el 3 % de sus tweets. En cambio, tanto la izquierda como la derecha tuvieron una presencia menor, con alrededor del 1.2 %

Cuadro 4.2: Ejemplos de tweets clasificados

Numero de Ejemplo	Tweet
1	#AColombiaLePreocupa es la relación de @AndresPastrana_ con la Unión Temporal Disproel, que es la que hace el conteo corrupto desde hace más de 20 años. ¿Pasos de animal grande sobre el control de las elecciones?
2	#TodoVale por eso es que estamos anclados al subdesarrollo, por anteponer todo por los principios. Un mentiroso y sucio Petro no le importa hacer lo que sea con tal de ganar. Los fines no justifican los medios. Petro acabó la posibilidad de un proyecto alternativo por su ego
3	Juan Roberto Vargas es una persona decente. Néstor Morales es una porquería de persona. #DebateFinal

y 0.8% respectivamente. En relación a alegría, se revela que la izquierda fue el sector que mostró una mayor presencia de dicha emoción en sus tweets, alcanzando el 55%. En segundo lugar se encuentra el sector neutral, con un 33%, seguido por la derecha con un 29%. En cuanto al asco, se destaca la participación predominante de la derecha, representando un 69% del total de sus tweets. El sector neutral se posiciona en segundo lugar con un 54% de los tweets, dejando a la izquierda con el menor porcentaje de los tres, un 40%. Finalmente, para la tristeza, el sector neutro fue en donde esta emoción tuvo mayor presencia con un 5% de su total, lo cual fue bastante mayor que la izquierda y la derecha, con un .08 y .04% respectivamente.

4.2. Emociones a lo largo del tiempo

Para determinar en cada emoción, que porcentaje de esta etiqueta tuvo un día en particular, se dividió el numero tweets con esta etiqueta en dicho día sobre el total

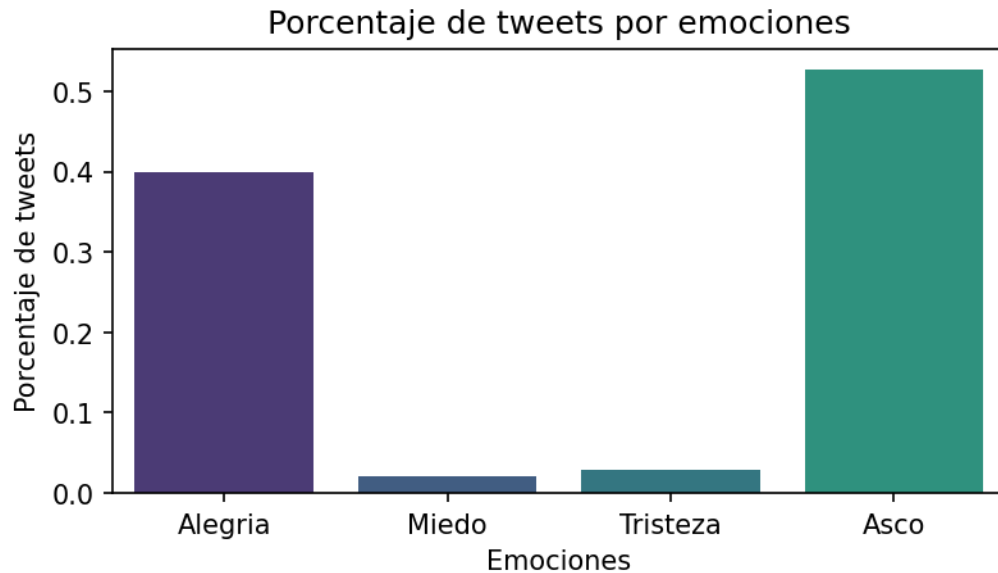


Figura 4.2: Porcentaje de tweets clasificado según emociones por el algoritmo

de tweets con esta etiqueta. De esta manera, se aprecia en el gráfico 4.4, que el 29 de mayo fue particularmente activo pues contó con casi 40 % de los tweets con miedo, y un 26 % de los tweets con alegría y tristeza. Este fue el día de la primera vuelta presidencial. Del mismo modo, se aprecia como los días 9 y 10 de junio tuvieron un repunte de asco y tristeza respectivamente, con un 9 y 11 %. Estos días estuvieron marcados con el evento de los llamados Petro videos. Luego hay un repunte de asco y tristeza al rededor del 16 de junio, fecha en donde se hablo del debate final al cual Rodolfo Hernandez se negó a participar, y finalmente de alegría, tristeza y miedo para el 19 de junio que fue la segunda vuelta.

Para cada emoción y sector, se obtuvo que porcentaje de esta etiqueta tuvo cada sector un día en particular, al dividir el numero de tweets que este sector tuvo con cada etiqueta en dicho día sobre el total de tweets que este sector tuvo con esta etiqueta. De esta manera, se obtiene el gráfico 4.5 para alegría en donde se aprecia

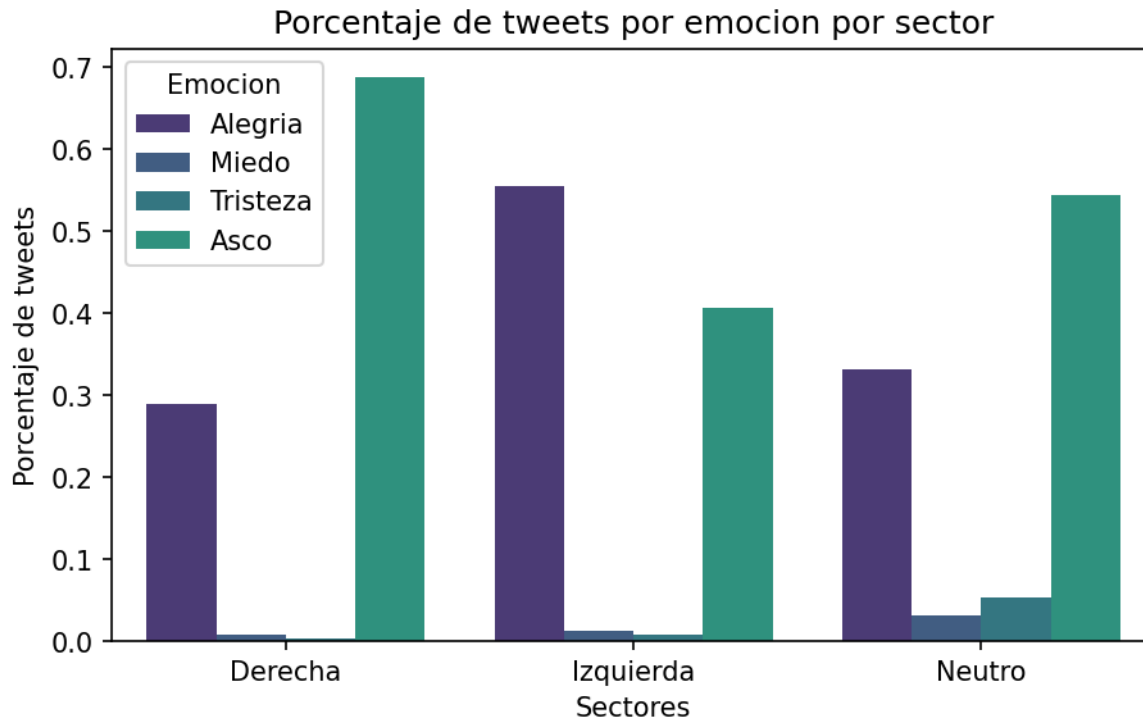


Figura 4.3: Porcentaje de tweets clasificados según emoción por el algoritmo, separado por sector político del hashtag

que el 29 de mayo todos los sectores tuvieron un repunte. Luego los tres sectores van creciendo cercanos al 19 de junio, para decaer eventualmente primero la derecha, luego el sector neutro y finalmente la izquierda.

Para el caso del miedo, se observa un gran pico de los tres sectores el 29 de mayo, día de la primera vuelta. Luego, para el 14 de junio, hubo un pico de miedo en la derecha como consecuencia a los rumores de estallido social, para finalmente haber un pico de miedo del sector neutro y de la izquierda cercano al 19 de junio.

En cuanto a la tristeza, se aprecia que los tres sectores tuvieron un repunte el 29 de mayo, principalmente la izquierda en donde este día se llegó a casi un 40 %.

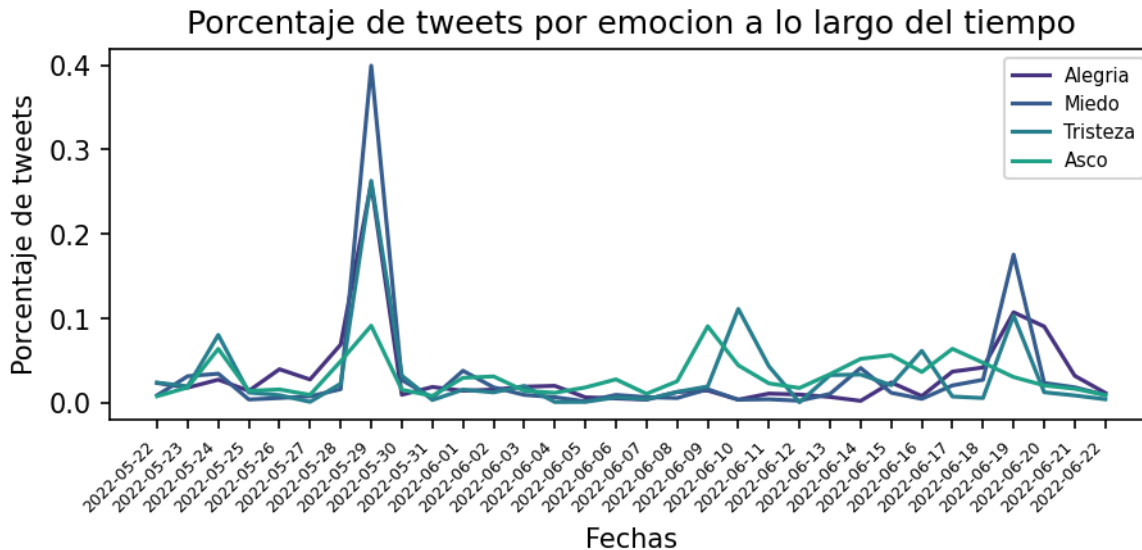


Figura 4.4: Porcentaje diario del total de tweets clasificados en cada emoción a lo largo del tiempo

Luego el 9 de junio hubo un repunte en la derecha, ligado al evento de los petro videos, y luego el 10 un repunte del sector neutro en donde se discutieron temáticas decepcionantes de las elecciones. Para el 14 de junio, la derecha tuvo devuelta un repunte de tristeza ligado, al igual que para el miedo, a la temática del estallido social. Finalmente, los tres sectores tuvieron un incremento de la tristeza para el 19 de junio.

En cuanto al asco, al inicio hay un pico del sector neutro el 24 de mayo debido a las reacciones respecto a un debate. Luego, hubo un repunte de los tres sectores para el 29 de mayo. Se destaca luego el gran pico que tuvo la derecha para el 9 de junio, fecha de los petro videos, con mas del 25 %. Asi mismo, se observa que la izquierda tuvo su pico el 17 de junio con cerca del 20 %. En esta fecha fue en donde se hablo de la negativa de Rodolfo Hernandez a participar en el debate final

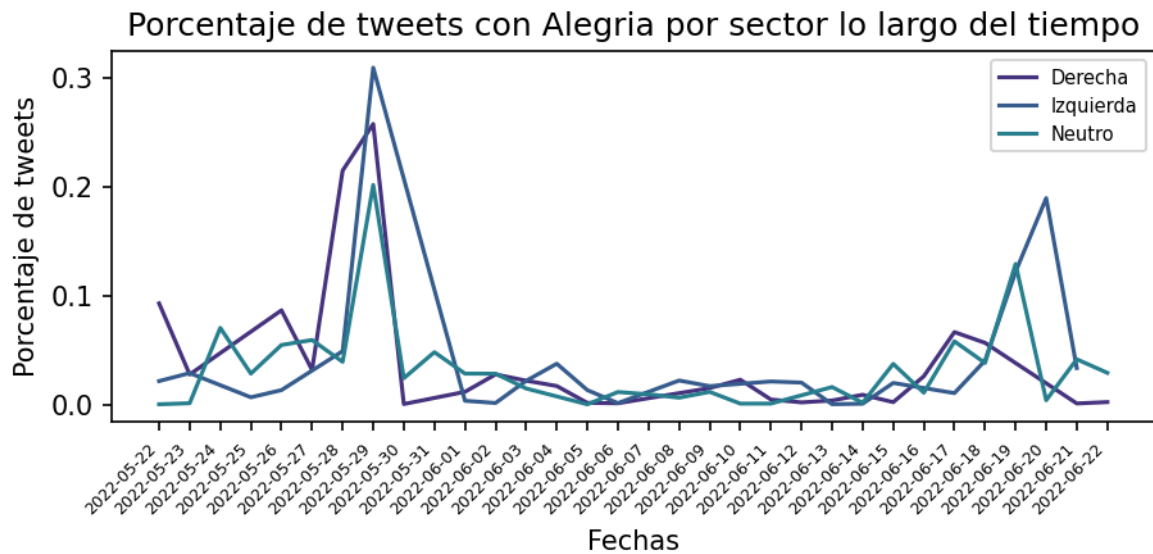


Figura 4.5: Porcentaje diario del total de tweets clasificados como Alegría, para cada sector político

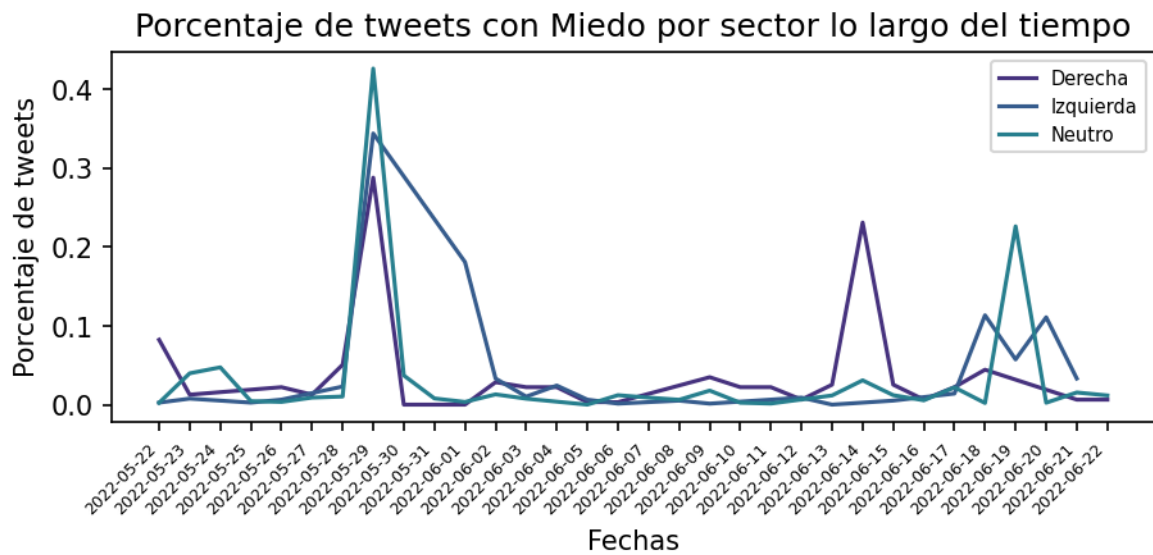


Figura 4.6: Porcentaje diario del total de tweets clasificados como Miedo, para cada sector político

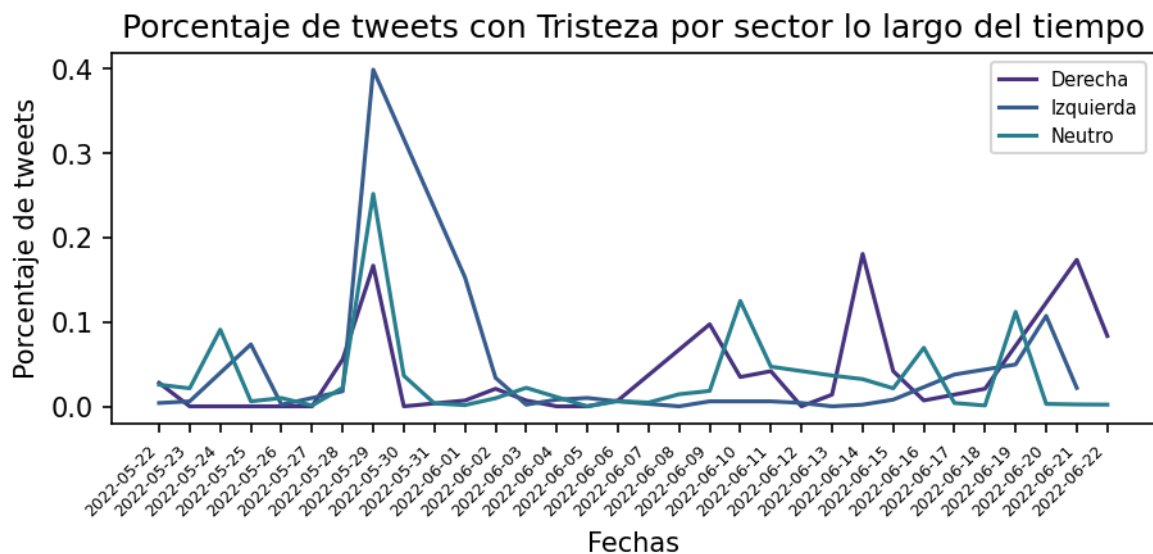


Figura 4.7: Porcentaje diario del total de tweets clasificados como Tristeza, para cada sector político

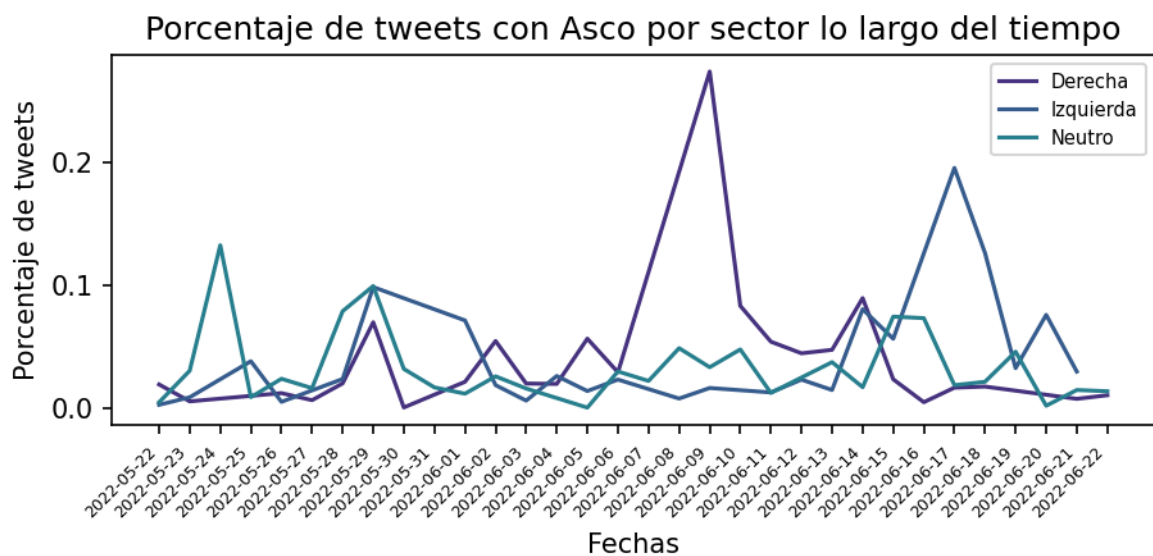


Figura 4.8: Porcentaje diario del total de tweets clasificados como Asco, para cada sector político

Capítulo 5

Conclusiones

El presente trabajo logro a través del fine tuning de un modelo de lenguaje pre entrenado, en este caso robertuito-base-uncased, identificar las emociones presentes en el texto de twitter relacionado con las elecciones presidenciales en Colombia del 2022. El fine tuning fue posible a través de la generación de un recurso propio, que fue un dataset de 1200 tweets etiquetados manualmente por el autor y los directores, usando una interfaz web de etiquetado, guiados por un manual de elaboración interna en donde a cada tweet se le asigno una o varias de las 14 emociones disponibles. Eventualmente, cada una de estas emociones fue asignada al tweet si al menos dos de los etiquetadores coincidieron en ella. Finalmente cada tweet quedo con una de las cuatro etiquetas resultantes de la agrupación de las etiquetas originales, basado en cuan correlacionadas se encontraban. Es necesario mencionar la dificultad presente en la tarea de etiquetado pues al ser la tarea una actividad subjetiva que intenta lograr una clasificación objetiva, se requirió de un proceso iterativo hasta lograr un manual de etiquetado que pudiera aproximar a los etiquetadores a este propósito.

Así mismo, si bien todos los etiquetadores son hispanohablantes nativos, solamente el autor es originario de Colombia, por lo que cierto uso del lenguaje o situaciones contextuales particulares, fueron mas claras para este que para los demás etiquetadores. Las predicciones realizadas por el modelo coinciden con las respuestas emocionales esperadas en ciertos eventos importantes durante las elecciones.

Los resultados del modelo muestran una presencia mucho mayor de alegría y asco que de miedo y tristeza, lo cual corresponde con lo observado durante el proceso de etiquetado siendo estas dos emociones por un lado menos prevalentes y por otro, acompañadas con frecuencia por asco. Estas razones ademas explicarían el desempeño mas pobre en términos de poder predictivo, de estas emociones con respecto a alegría y a asco. Cabe resaltar, que estos fueron los resultados obtenidos con una cantidad de 1200 tweets etiquetados, por lo que agregar mas tweets a la muestra, habría posiblemente ayudado a mejorar el desempeño del algoritmo, principalmente en las emociones que tuvieron poca participación.

Se pudo observar en lineas generales una tendencia mayor a la alegría por parte de la izquierda y al asco por parte de la derecha. Esto en parte se explica en que el candidato victorioso fue un candidato de izquierda, despertando estos sentimientos respectivamente en cada sector político, principalmente en algunas fechas particulares, como el día de las elecciones.

EL presente trabajo disponibiliza el dataset etiquetado así como el modelo utilizado para que pueda ser usado para futuras investigaciones, como por ejemplo la identificación de emociones en español en un contexto diferente al del presente trabajo. De igual manera, muestra a partir del uso de estos recursos, como para el contexto particular de las elecciones presidenciales en Colombia en el 2022, se expresaron emo-

cionalmente los usuarios de twitter que utilizaron los hashtags incluidos.

Bibliografía

- Acheampong, F. A., Nunoo-Mensah, H., and Chen, W. (2021). Transformer models for text-based emotion detection: a review of bert-based approaches. *Artificial Intelligence Review*, 54(8):5789–5829.
- Alm, C. O., Roth, D., and Sproat, R. (2005). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 579–586.
- Aman, S. and Szpakowicz, S. (2007). Identifying expressions of emotion in text. In *International Conference on Text, Speech and Dialogue*, pages 196–205. Springer.
- Barbieri, F., Anke, L. E., and Camacho-Collados, J. (2022). Xlm-t: Multilingual language models in twitter for sentiment analysis and beyond. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 258–266.
- Bollen, J., Mao, H., and Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 450–453.

- Canete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., and Pérez, J. (2020). Spanish pre-trained bert model and evaluation data. *Pml4dc at iclr*, 2020:1–10.
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., and Pérez, J. (2020). Spanish pre-trained bert model and evaluation data. In *PML4DC at ICLR 2020*.
- Cerón-Guzmán, J. A. and León-Guzmán, E. (2016). A sentiment analysis system of spanish tweets and its application in colombia 2014 presidential election. In *2016 IEEE international conferences on big data and cloud computing (BDCloud), social computing and networking (socialcom), sustainable computing and communications (sustaincom)(BDCloud-socialcom-sustaincom)*, pages 250–257. IEEE.
- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Davidov, D., Tsur, O., and Rappoport, A. (2010). Enhanced sentiment learning using twitter hashtags and smileys. In *Coling 2010: Posters*, pages 241–249.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Ekman, P. (1993). Facial expression and emotion. *American psychologist*, 48(4):384.
- Fandiño, A. G., Estapé, J. A., Pàmies, M., Palao, J. L., Ocampo, J. S., Carrino, C. P., Oller, C. A., Penagos, C. R., Agirre, A. G., and Villegas, M. (2022a). Maria: Spanish language models. *Procesamiento del Lenguaje Natural*, 68.

- Fandiño, A. G., Estapé, J. A., Pàmies, M., Palao, J. L., Ocampo, J. S., Carrino, C. P., Oller, C. A., Penagos, C. R., Agirre, A. G., and Villegas, M. (2022b). Maria: Spanish language models. *Procesamiento del Lenguaje Natural*, 68.
- Gil, G. B., Jesús, A. B. d., and López, J. M. M. (2013). Combining machine learning techniques and natural language processing to infer emotions using spanish twitter corpus. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 149–157. Springer.
- Go, A., Bhayani, R., and Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N project report, Stanford*, 1(12):2009.
- Gonzalez, J. A., Hurtado, L.-F., and Pla, F. (2021). Twilbert: Pre-trained deep bidirectional transformers for spanish twitter. *Neurocomputing*, 426:58–69.
- Hatzivassiloglou, V. and McKeown, K. (1997). Predicting the semantic orientation of adjectives. In *35th annual meeting of the association for computational linguistics and 8th conference of the european chapter of the association for computational linguistics*, pages 174–181.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Kouloumpis, E., Wilson, T., and Moore, J. (2011). Twitter sentiment analysis: The good the bad and the omg! In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 538–541.
- la Rosa, J. D., Ponferrada, E. G., Romero, M., Villegas, P., de Prado Salas, P. G., and Grandury, M. (2022). Bertin: Efficient pre-training of a spanish language model using perplexity sampling. *Procesamiento del Lenguaje Natural*, 68(0):13–23.

- Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.
- Mohammad, S. (2012). # emotional tweets. In * *SEM 2012: The First Joint Conference on Lexical and Computational Semantics–Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255.
- Mohammad, S., Bravo-Marquez, F., Salameh, M., and Kiritchenko, S. (2018). Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 1–17.
- Mohammad, S. M., Zhu, X., Kiritchenko, S., and Martin, J. (2015). Sentiment, emotion, purpose, and style in electoral tweets. *Information Processing & Management*, 51(4):480–499.
- O’Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *Fourth international AAAI conference on weblogs and social media*.
- Ortony, A., Clore, G. L., and Foss, M. A. (1987). hatzivassiloglou1997predicting. *Cognitive science*, 11(3):341–364.
- Pak, A. and Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*.
- Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, 2(1–2):1–135.

- Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.
- Pérez, J. M., Furman, D. A., Alonso Alemany, L., and Luque, F. M. (2022). RoBERTuito: a pre-trained language model for social media text in Spanish. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 7235–7243, Marseille, France. European Language Resources Association.
- Picard, R. W. (2000). *Affective computing*. MIT press.
- Plaza-del Arco, F. M., Jiménez Zafra, S. M., Montejo Ráez, A., Molina González, M. D., Ureña López, L. A., and Martín Valdivia, M. T. (2021). Overview of the emoeval task on emotion detection for spanish at iberlef 2021.
- Plaza-del Arco, F. M., Martín-Valdivia, M. T., Ureña-López, L. A., and Mitkov, R. (2020a). Improved emotion recognition in spanish social media through incorporation of lexical knowledge. *Future Generation Computer Systems*, 110:1000–1008.
- Plaza-del Arco, F. M., Strapparava, C., Lopez, L. A. U., and Martín-Valdivia, M. T. (2020b). Emoevent: A multilingual emotion corpus based on different events. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 1492–1498.
- Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., and Harabagiu, S. (2012). Empatweet: Annotating and detecting emotions on twitter. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC’12)*, pages 3806–3813.
- Romero, M. (2020). Spanish electra by manuel romero. <https://huggingface.co/mrm8488/electricidad-base-discriminator/>.

- Sidorov, G., Haro, S. N. G., and Vázquez, V. A. C. (2016). Construcción de un corpus marcado con emociones para el análisis de sentimientos en twitter en español. *Revista Escritos BUAP*, 1(1).
- Sidorov, G., Miranda-Jiménez, S., Viveros-Jiménez, F., Gelbukh, A., Castro-Sánchez, N., Velásquez, F., Díaz-Rangel, I., Suárez-Guerra, S., Trevino, A., and Gordon, J. (2012). Empirical study of machine learning based approach for opinion mining in tweets. In *Mexican international conference on Artificial intelligence*, pages 1–14. Springer.
- Strapparava, C., Valitutti, A., et al. (2004). Wordnet affect: an affective extension of wordnet. In *Lrec*, volume 4, page 40. Lisbon, Portugal.
- Tumasjan, A., Sprenger, T., Sandner, P., and Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 4, pages 178–185.
- Turney, P. D. (2002). Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Wang, W., Chen, L., Thirunarayan, K., and Sheth, A. P. (2012). Harnessing twitter”big data”for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, pages 587–592. IEEE.

- Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2):165–210.
- Wiebe, J. M. (1994). Tracking point of view in narrative. *arXiv preprint cmp-lg/9407019*.
- Wikipedia contributors (2022). Recurrent neural network — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Recurrent_neural_network&oldid=1109264340. [Online; accessed 26-September-2022].
- Yu, H. and Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 129–136.

Apéndice A

bla

A.1. Anexos

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-05-22	NosUnimosONosJodemos	Derecha	427
2022-05-22	FedericoEsColombia	Derecha	1597
2022-05-22	LoPeorDeEstasElecciones	Neutro	357
2022-05-22	MiVotoEsPetroYFrancia	Izquierda	2131
2022-05-23	MiVotoEsPetroYFrancia	Izquierda	2131
2022-05-23	UnTramposoEs	Neutro	1255
2022-05-23	FedericoEsColombia	Derecha	1597

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-05-23	MeInquieta	Neutro	696
2022-05-24	ElDebateDefinitivo	Neutro	7546
2022-05-24	AColombiaLePreocupa	Neutro	1088
2022-05-24	LeApuestoA	Neutro	326
2022-05-24	UnTramposoEs	Neutro	1255
2022-05-24	GrandezaEs	Neutro	498
2022-05-25	VotoPor	Neutro	1093
2022-05-25	YaEsSuficiente	Izquierda	816
2022-05-25	ElDebateDefinitivo	Neutro	7546
2022-05-25	vanessapregúnteleafico	Izquierda	323
2022-05-25	AColombiaLePreocupa	Neutro	1088
2022-05-25	ConvenceUnoMásPorPetro	Izquierda	174
2022-05-26	FicoElGanador	Derecha	2578
2022-05-26	ElMejorCandidatoEs	Neutro	1681
2022-05-26	ElDebateDecisivo	Neutro	1305
2022-05-26	PetroEsElCambio	Izquierda	457
2022-05-26	MagdalenaVotaPetro	Izquierda	150
2022-05-27	PetroEs	Neutro	877
2022-05-27	EsElMomentoDe	Neutro	1334
2022-05-27	FicoElGanador	Derecha	2578
2022-05-27	ElDomingoVoyA	Neutro	298

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-05-27	ElMejorCandidatoEs	Neutro	1681
2022-05-28	LaVidaEnUnVoto	Izquierda	221
2022-05-28	VoyConFico	Derecha	78
2022-05-28	YoVotoPetroPresidente	Izquierda	3376
2022-05-28	DebateFinal	Neutro	5217
2022-05-28	VoyConFederico	Derecha	4148
2022-05-28	FicoElGanador	Derecha	2578
2022-05-29	EleccionesColombia	Neutro	7088
2022-05-29	CualquieraMenosPetro	Derecha	957
2022-05-29	Elecciones2022	Neutro	5978
2022-05-29	PetroPresidenteEnPrimera	Izquierda	4032
2022-05-29	PetroPresidente2022	Izquierda	1712
2022-05-29	ElCambioEnPrimera	Izquierda	3077
2022-05-29	SegundaVuelta	Neutro	1752
2022-05-29	MiVotoEsSecretosky	Izquierda	734
2022-05-29	VotoPorFicoYLara	Derecha	846
2022-05-29	PetroNuncaSeráPresidente	Derecha	461
2022-05-29	HoyMeIdentiFICO	Derecha	92
2022-05-29	YoVotoPetroPresidente	Izquierda	3376
2022-05-29	ColombiaConPetro	Izquierda	3263
2022-05-29	VoyConFederico	Derecha	4148

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-05-29	ColombiaDecide	Neutro	3906
2022-05-29	HoyGana	Neutro	687
2022-05-29	PetroYFranciaSonElCambio	Izquierda	637
2022-05-29	PetroNUNCAseraPresidente	Derecha	461
2022-05-30	SegundaVuelta	Neutro	1752
2022-05-30	LaPrimeraVueltaFue	Neutro	693
2022-05-30	Elecciones2022	Neutro	5978
2022-05-30	VotoPorFicoYLara	Derecha	846
2022-05-30	HoyGana	Neutro	687
2022-05-31	SegundaVuelta	Neutro	1752
2022-05-31	MiCandidatoEs	Neutro	2123
2022-06-01	FraudeElectoral	Izquierda	2814
2022-06-01	VelezPregunteleAPetro	Derecha	192
2022-06-01	pactosabroso	Izquierda	79
2022-06-01	ElVerdaderoCambioEs	Neutro	787
2022-06-01	RodolfoPresidente2022	Derecha	2720
2022-06-01	CreoEnColombiaPor	Neutro	658
2022-06-01	Antipetrismo	Derecha	176
2022-06-02	MeSumoA	Neutro	307
2022-06-02	ColombiaNecesita	Neutro	2381
2022-06-02	RodolfoPresidente2022	Derecha	2720

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-06-02	UnaPreguntaParaPetroFm	Derecha	850
2022-06-02	FraudeElectoral	Izquierda	2814
2022-06-02	ElDebateFeminista	Izquierda	56
2022-06-03	PetroElCambioSeguro	Izquierda	3145
2022-06-03	RodolfoPresidente2022	Derecha	2720
2022-06-03	ColombiaQuiere	Neutro	600
2022-06-03	ColombiaNecesita	Neutro	2381
2022-06-03	UnDesastre	Neutro	357
2022-06-04	PetroElCambioSeguro	Izquierda	3145
2022-06-04	YoPactoConPetro	Izquierda	761
2022-06-04	RodolfoPresidente2022	Derecha	2720
2022-06-05	LaCorrupcionPetrista	Derecha	1492
2022-06-05	VotarEnBlancoEs	Neutro	1466
2022-06-05	GraciasPero	Neutro	719
2022-06-05	PetroElCambioSeguro	Izquierda	3145
2022-06-06	VotarEnBlancoEs	Neutro	1466
2022-06-06	MeLimpioELCuloConEsaLey	Izquierda	149
2022-06-06	LaCorrupciónPetrista	Derecha	775
2022-06-06	GraciasPero	Neutro	719
2022-06-06	AliasElIngeniero	Izquierda	505
2022-06-07	UnaVerdadEs	Neutro	1504

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-06-07	VotarEnBlancoEs	Neutro	1466
2022-06-08	UnaBurlaEs	Neutro	1467
2022-06-08	YoContratoAPetro	Izquierda	1511
2022-06-08	NoQuiero	Neutro	921
2022-06-08	UnaVerdadEs	Neutro	1504
2022-06-08	ConPetroGanasTú	Izquierda	259
2022-06-09	PetrismoSucio	Derecha	5447
2022-06-09	RumboALaVictoria	Izquierda	271
2022-06-09	PetroVideos	Derecha	2011
2022-06-09	MeDaVergüenza	Neutro	195
2022-06-09	PetroNoMateaRodolfo	Derecha	37
2022-06-09	NoTodoVale	Neutro	257
2022-06-09	PetroMentiroso	Derecha	524
2022-06-09	LosPetroVideosSon	Derecha	3034
2022-06-09	TodoVale	Neutro	796
2022-06-09	YoContratoAPetro	Izquierda	1511
2022-06-09	ColombiaEspera	Neutro	918
2022-06-10	HablaMalDeColombia	Neutro	2231
2022-06-10	PetrismoSucio	Derecha	5447
2022-06-10	PetroVideos	Derecha	2011
2022-06-10	CampañaSuciaEs	Neutro	1149

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-06-10	MeVoyConElIngeniero	Derecha	278
2022-06-11	SoyPetroSoyPaz	Izquierda	2125
2022-06-11	PetroVideos	Derecha	2011
2022-06-11	HablaMalDeColombia	Neutro	2231
2022-06-11	PetrismoSucio	Derecha	5447
2022-06-11	QuéCansancio	Neutro	459
2022-06-12	LaMaldadPetrista	Derecha	2152
2022-06-12	SoyPetroSoyPaz	Izquierda	2125
2022-06-12	ArizabaletaLameCulos	Izquierda	328
2022-06-13	ArizabaletaSicarioMoral	Izquierda	386
2022-06-13	PetroRespondaleAlosNiños	Derecha	332
2022-06-13	VanessaEnSegundaVotoPor	Neutro	199
2022-06-13	UnGolpeBajoEs	Neutro	676
2022-06-13	EsUnaBajeza	Neutro	501
2022-06-13	LaGranEncuesta	Neutro	70
2022-06-13	LaMaldadPetrista	Derecha	2152
2022-06-13	NosFalta	Neutro	1155
2022-06-14	PetrismoAcosaALaPrensa	Derecha	461
2022-06-14	YoBloqueoASemana	Izquierda	3508
2022-06-14	LaCuotaDeBenedetti	Derecha	416
2022-06-14	EstallidoSocialEs	Derecha	1964

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-06-14	NoMeDejaDormir	Neutro	1178
2022-06-15	DebateObligadoEs	Neutro	3491
2022-06-15	YoBloqueoASemana	Izquierda	3508
2022-06-15	LoCorrectoEs	Neutro	1495
2022-06-15	CaracolPetrista	Derecha	409
2022-06-15	presidentepetro	Izquierda	859
2022-06-15	EstallidoSocialEs	Derecha	1964
2022-06-16	IrrespetuosoEs	Neutro	2125
2022-06-16	debate	Neutro	791
2022-06-16	NosHaceFalta	Neutro	985
2022-06-16	RodolfoElPresidentedelPueblo	Derecha	408
2022-06-16	MeDecepciona	Neutro	452
2022-06-17	RHGanador	Derecha	2277
2022-06-17	RodolfoCobarde	Izquierda	4843
2022-06-17	YoVotoPor	Neutro	1748
2022-06-17	EsteDomingo	Neutro	1033
2022-06-17	PetroPresidenteEnsegunda	Izquierda	1359
2022-06-18	ExguerrilleroOIngeniero	Neutro	2760
2022-06-18	RegistraduriaResponda	Izquierda	3107
2022-06-18	MiVotoEsAntiFraude	Izquierda	1592
2022-06-18	RHGanador	Derecha	2277

Continued on next page

Cuadro A.1: Hashtags

date	trending	Sector	Count
2022-06-18	PetroPresidenteEnsegunda	Izquierda	1359
2022-06-19	PetroEsPresidente	Izquierda	15521
2022-06-19	Elecciones2022	Neutro	5978
2022-06-19	EleccionesPresidenciales2022	Neutro	327
2022-06-19	ExguerrilleroOIngeniero	Neutro	2760
2022-06-19	EleccionesColombia2022	Neutro	4484
2022-06-19	RegistraduriaResponda	Izquierda	3107
2022-06-19	VivirSabroso	Izquierda	790
2022-06-20	PetroEsPresidente	Izquierda	15521
2022-06-20	ExguerrilleroOIngeniero	Neutro	2760
2022-06-20	Elecciones2022	Neutro	5978
2022-06-21	PetroEsPresidente	Izquierda	15521
2022-06-21	NuestraPrioridadEs	Neutro	1253
2022-06-21	PetroBastaYa	Derecha	564
2022-06-21	DePetroEspero	Neutro	429
2022-06-21	PetroNecesita	Neutro	377
2022-06-22	EsMuyImportante	Neutro	1255
2022-06-22	PetroBastaYa	Derecha	564
2022-06-22	LosImpuestosDePetro	Neutro	359