



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

MEDICIÓN DEL NIVEL DE CONTROVERSIA EN REDES SOCIALES EN CHILE

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

ESTEBAN EDUARDO ROJAS MOYANO

PROFESOR GUÍA:
Andrés Abeliuk Kimelman

MIEMBROS DE LA COMISIÓN:
Francisco Rivera Serrano
Felipe Bravo Márquez

SANTIAGO DE CHILE
2022

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: ESTEBAN EDUARDO ROJAS MOYANO
FECHA: 2022
PROF. GUÍA: ANDRÉS ABELIUK

MEDICIÓN DEL NIVEL DE CONTROVERSIA EN REDES SOCIALES EN CHILE

La motivación de este trabajo se manifiesta a partir del interés de estudiar la polarización que se ha pronunciado a través de los años en Chile, debido a todos los sucesos políticos y sociales que han ocurrido y han generado polémicas en sus respectivos tiempos. Es por ello, se tiene como objetivo final de esta memoria cuantificar la controversia y analizar cómo varía este fenómeno por cada evento en la red social de *Twitter* y cómo los usuarios de esta plataforma interactúan entre sí a partir de estos temas que pueden o no ser controversiales.

El problema a resolver se basa en la implementación de metodologías que permitan dilucidar la polarización entre dos grupos con opiniones distintas en base a métricas, que puedan ser obtenidas a partir de diversas maneras, como grafos o procesamiento de lenguaje natural, para cuantificar la controversia, y, a su vez, construir gráficos donde se observen cómo varían estas medidas con respecto al periodo de tiempo en el que se estudian los eventos escogidos.

La solución que se implementó se puede separar en varios pasos. Inicialmente, se buscan todos los términos necesarios para realizar las consultas en la *Twitter* API y obtener los *tweets* que estén relacionados al tema en cuestión. Todo esto en función de un algoritmo de similitud que permite encontrar coincidencias entre los *tweet* y el conjunto de términos encontrados. Luego, con estos datos, se realizan dos tareas, siendo una de ellas la construcción de grafos de *retweets* que enlacen usuarios entre sí y, eventualmente, calcular dos métricas capaces de cuantificar la polarización que se observan en estas redes de nodos. La otra tarea es calcular una tercera métrica, pero empleando el texto de los *tweets* con procesamiento de lenguaje natural y la distancia euclidiana. Finalmente, se crean gráficos que permiten visibilizar la variabilidad de la mejor métrica calculada con grafos y la calculada con PLN con respecto al tiempo del periodo del evento.

Los resultados expresan la presencia y la medición de la polarización en temas controversiales, como también la diferencia de estos con los tópicos en lo que no existe polarización alguna, además del apreciable cambio de ambas métrica en cada semana de los eventos, lo que lleva a la conclusión de que esta manifestación que se produce en las redes sociales no es fija por cada grupo, o entre estos, y que varía con el tiempo.

A todos quienes me apoyaron a lo largo de este camino.

Agradecimientos

En primera instancia, quiero agradecer a mi familia. A mis padres, Ethel y Wilson, por darme la oportunidad de escoger el camino que quería y siempre estar a mi lado en las decisiones que tomara, además de entregarme todo el cariño y amor que necesitara para seguir adelante. Como también a mis hermanas Daniela, Javiera y Melisa, por también facilitarme su apoyo y afecto en todos estos años y siempre buscar formas de pasar tiempo juntos.

Además, agradecer a todos los amigos que hice en mi paso por la universidad pero, por sobre todo, a Luna, Mati y Rodrigo, quienes hicieron más amenos estos últimos años entre conversaciones, risas y buenos momentos.

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Descripción del problema	2
1.3. Objetivos	2
1.3.1. Objetivo General	2
1.3.2. Objetivos Específicos	2
1.4. Formalización del problema	3
1.5. Estructura de la Memoria	3
2. Marco Teórico y Estado del Arte	4
2.1. Estado del Arte	4
2.1.1. The Political Blogosphere and the 2004 U.S. Election: Divided They Blog	4
2.1.2. Quantifying Controversy on Social Media	4
2.1.3. A Motif-Based Approach for Identifying Controversy	5
2.1.4. A Measure of Polarization on Social Media Networks Based on Community Boundaries	6
2.1.5. Measuring Political Polarization: <i>Twitter</i> shows the two sides of Venezuela	6
2.1.6. Every Colour You Are: Stance Prediction and Turnaround in Controversial Issues	7
2.1.7. Quantifying Content Polarization on <i>Twitter</i>	8
2.2. Marco Teórico	9
2.2.1. Grafos	9
2.2.2. Lematización	9
2.2.3. Tokenización	10
2.2.4. Word2Vec	10
2.2.5. Análisis Sentimental	11
2.2.6. Twitter y tweets	11
2.2.7. <i>Twitter</i> API	12
3. Metodología	13
3.1. Obtención de Datos	13
3.2. Construcción de Grafos de <i>Retweets</i>	16
3.3. Métricas basadas en Grafos	16
3.3.1. Random Walk	16
3.3.2. Betweenness	18

3.4. Métrica basada en NLP	19
3.4.1. Normalized Cut	19
4. Resultados y Análisis	21
4.1. Temas a Estudiar	21
4.2. Grafos de Reweets	22
4.3. Métricas de cada Tema	24
4.4. Random Walk vs Tiempo del Evento	25
4.5. Normalized Cut vs Tiempo del Evento	28
4.6. Normalized Cut de los Sentimientos vs Tiempo del Evento	31
5. Conclusiones	34
Bibliografía	36

Índice de Tablas

3.1.	Columnas extraídas y ejemplos.	14
4.1.	Temas de observación donde los primeros tres se consideran de gran controversia: el cuarto, que puede posiblemente serlo; y los últimos dos, que no lo son. . . .	21
4.2.	Puntajes que cuantifican el nivel de controversia de temas controversiales. . . .	24
4.3.	Puntajes que cuantifican el nivel de controversia de temas controversiales. . . .	24

Índice de Ilustraciones

2.1.	Ejemplo de grafo.	9
2.2.	Explicación de la técnica Word2Vec.	10
3.1.	Búsqueda de usuarios para la creación de grafos.	15
3.2.	Grafos construido a partir de uno o todos los datasets.	15
4.1.	Grafos de <i>retweets</i> de los siguientes temas: a) Torneo primera división en Chile del año 2021, b) Programa televisivo Pasapalabra en sus meses de mayor rating, c) Incendios forestales ocurridos en el verano del 2017.	22
4.2.	Grafos de <i>retweets</i> de los siguientes temas: a) Primeros meses luego del homicidio del comunero mapuche, Camilo Catrillanca b) Proceso legislativo de las tres causales del aborto, c) Campaña del apruebo y rechazo antes del plebiscito del año 2020.	23
4.3.	Polarización de la red a través del tiempo, donde la línea representa el valor de random walk de cada semana dentro del periodo establecido, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	25
4.4.	Polarización de la red a través del tiempo, donde la línea representa el valor de <i>random walk</i> de cada semana dentro del periodo establecido, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	26
4.5.	Polarización de la red a través del tiempo, donde la línea representa el valor de <i>random walk</i> de cada semana dentro del periodo establecido, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	26
4.6.	Polarización de la red a través del tiempo, donde la línea representa el valor de <i>random walk</i> de cada semana dentro del periodo establecido, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	27
4.7.	Polarización del contenido a través del tiempo, donde la línea representa el valor de <i>normalized cut</i> de cada semana dentro del periodo establecido, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	28
4.8.	Polarización del contenido a través del tiempo, donde la línea representa el valor de <i>normalized cut</i> de cada semana dentro del periodo establecido, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	29

4.9.	Polarización del contenido a través del tiempo, donde la línea representa el valor de <i>normalized cut</i> de cada semana dentro del periodo establecido, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	29
4.10.	Polarización del contenido a través del tiempo, donde la línea representa el valor de <i>normalized cut</i> de cada semana dentro del periodo establecido, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	30
4.11.	Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de <i>normalized cut</i> de un sentimiento específico, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	31
4.12.	Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de <i>normalized cut</i> de un sentimiento específico, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	32
4.13.	Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de <i>normalized cut</i> de un sentimiento específico, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	32
4.14.	Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de <i>normalized cut</i> de un sentimiento específico, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.	33

Capítulo 1

Introducción

1.1. Motivación

A través de los años se ha podido observar un incremento en la polarización de las opiniones en general, siendo esto reflejado más claramente en redes sociales, existiendo varias causas que son capaces de explicar este fenómeno. Sin embargo, una razón relevante de esto es la facilidad de estas plataformas para permitir a grupos de personas, que deseen compartir una opinión, mantenerse en un contacto constante y, a su vez de ignorar o bloquear a quienes tienen puntos totalmente contrarios a los suyos, lo que lleva que se formen "burbujas sociales", donde se es incapaz de lograr entablar un dialogo sin que se generen controversias en estos hilos de Internet. Además de que, los actuales algoritmos utilizados en las redes sociales, cada vez más tienden a mostrar a sus usuarios solamente contenido que les interese, incluyendo temas, ideas y opiniones de carácter social y político [1–3].

En un país como Chile, cuya población se dividió en dos por una serie de sucesos, como lo fue el estallido social ocurrido en 2019 y el plebiscito por una nueva constitución, se ha presentado una oleada constante de información y usuarios polarizantes. No obstante, no se ha presentado una solución exacta que permita detectar y cuantificar a los usuarios o publicaciones que cumplan con las condiciones que demuestren que estos sean polarizantes.

En uno de los primeros artículos respecto a esta problemática tenemos a Adamic et al. [4] que estudian los patrones de enlaces y los temas de discusión de los blogs políticos, centrándose en publicaciones de blogs sobre las elecciones presidenciales de EE. UU. de 2004. En este se mide el grado de interacción entre blogs liberales y conservadores, proporcionando evidencia que los blogs conservadores se enlazan entre sí con más frecuencia y en un patrón más denso. Estos hallazgos son confirmados por el estudio más reciente de Conover et al. [5], quienes también estudian la controversia en la comunicación política con respecto a las elecciones parlamentarias del Congreso. Utilizando datos de *Twitter*, Conover identifica una estructura partidista altamente segregada, con conectividad limitada entre usuarios de izquierda y derecha.

También es importante diferenciar cada trabajo de acuerdo a los tipos de resultados que presentan. En el caso de este trabajo, se enfatizó en particular el uso de dos investigaciones. En primer lugar, se tiene el de Garimella et al. [7], el cual es uno de los trabajos respecto a la medición de controversia en redes sociales más relevantes que existen, haciéndose uso de sus

metodologías basadas en redes de usuarios interconectadas a través *retweet*, seguimientos, menciones y contenido. Por último, se tiene el trabajo de Yang et al. [9], siendo este un trabajo que se centra en las elecciones de EE.UU. de 2016 y presenta certeras métricas que se basan netamente en el texto presente en los *tweets*.

1.2. Descripción del problema

Tal como se mencionó anteriormente, en los últimos años las personas han encontrado en las redes sociales un lugar para difundir sus ideas, lo que ha generado un nuevo escenario de conflicto, dado que se da la entrada a hablar públicamente de temas sensibles, como raza, género, política, religión, entre otros, que van surgiendo de acuerdo a las circunstancias, al mismo tiempo que existen problemas relacionados a la información en redes sociales como lo son las noticias falsas, perfiles fantasmas, difusión de odio, *cyberbullying*, etc. Sumado a esto, está además el hecho de que es inevitable que las personas civiles y las personas directamente ligadas a este tipo de movimientos, usen las redes sociales con el fin de debatir o esparcir su ideas; por lo tanto, es de interés cuantificar, de alguna forma, qué temas tienden a tener una discusión que vaya hacia la polarización.

Por consiguiente, realizando una búsqueda a través de distintos papers, se consideró que la red social propicia para este estudio es *Twitter*, puesto que allí las personas son abiertas, escriben acerca de sus pensamientos y afiliaciones, y por ello, será más sencillo segmentar y encontrar patrones de similitud que permitan cumplir la meta. De este modo, este trabajo de título busca analizar las tendencias que hay en *Twitter* y los *tweet* asociados a ellas, permitiéndose aplicar las métricas que parezcan necesarias para la solución de esta problemática.

1.3. Objetivos

1.3.1. Objetivo General

Calcular métricas y obtener gráficos haciendo uso de técnicas de machine learning, que representen el nivel de polarización en *tweet* con respecto al contexto chileno.

1.3.2. Objetivos Específicos

- Construir un arreglo de términos que representen un solo tema que se esté discutiendo.
- Construir grafos que permitan visualizar la polarización de un tema.
- Buscar y calcular métricas que sirvan para categorizar que tópicos son controversiales y cuales no.
- Realizar gráficos de las métricas calculadas a nivel temporal para analizar el cambio de la controversia en respuesta a ciertos eventos.

1.4. Formalización del problema

Obtención de métricas y gráficos que permitan ilustrar la polarización en *tweet* con respecto a un tema en particular, haciendo uso de técnicas de machines learning.

1.5. Estructura de la Memoria

El segundo capítulo de esta memoria se centra en explicar las investigaciones utilizadas para este trabajo y el cómo aportaron a éste. El tercer capítulo presenta el estado del arte y el marco teórico, que exhibe otros trabajos que sentaron la base de esta memoria y la explicación de términos claves que se harán uso a lo largo de este trabajo, respectivamente. Siguiendo con el cuarto capítulo se tiene la metodología, que explica de qué forma se obtuvieron los resultados de las métricas, redes y gráficos, siendo estos enseñados en el quinto capítulo, para su eventual discusión y comparación en el sexto capítulo. Y de esta manera, en el séptimo y último capítulo, se concluye el trabajo realizado a lo largo de todo el proceso.

Capítulo 2

Marco Teórico y Estado del Arte

2.1. Estado del Arte

Como se ha dicho anteriormente, el estudio de la polarización en redes sociales es un desafío que presenta una amplia literatura donde es posible consultar metodologías que parezcan aptas para la obtención de resultados concluyentes.

2.1.1. The Political Blogosphere and the 2004 U.S. Election: Divided They Blog

Este debe ser el primer trabajo que trata la controversia en el internet, siendo realizado por Adamic et al. en el año 2005, donde se estudian blogs conservadores y liberales y cómo interactúan entre sí, durante el periodo de las elecciones presidenciales del año 2004. Para esto, se recolectan un amplio conjunto de URL de blogs políticos mediante varios directorios en línea, donde se lograron obtener un total de 1492 blogs, siendo 759 liberales y 735 conservadores. Pero para realizar un estudio más minucioso solo se tomaron en cuenta los 20 blogs más relevantes de cada partido, considerando el número de citas que realizaban a otros blogs y su clasificación por relevancia. Es así como se recolectan 12470 posts de blogs con inclinación hacia la izquierda y 10414, hacia la derecha.

En consecuencia a lo anterior, se analizan las citas de cada post y se revisan sus interacciones con otros blogs, permitiéndose ver que los blogs liberales se citan unos con otros 1511 veces, mientras que los conservadores, 2110 veces. Además, se presenta el hecho de que las citas cruzadas solo representan un 15% del total de enlaces, con liberales citando a conservadores 247 veces, y conservadores citando a liberales 312 veces. Finalmente, se concluye que los blogs con tendencia a la derecha exhiben una red mucho más robusta, con una mayor interconexión, que los de izquierda. No obstante, no es posible observar patrones de uniformidad en ambos partidos. Más bien, se ven ambas comunidades actuando como leves cámaras de eco al discutir frecuentemente diferentes páginas web con respecto al otro.

2.1.2. Quantifying Controversy on Social Media

Considerado uno de los trabajos más extensos con respecto a esta problemática, realizado por 2015 por Garimella et al. [7], debido a que aplica una variedad de formas de medir el nivel

de controversia, presentando una metodología que se basa principalmente en la construcción de grafos a través de todas las posibles formas de relacionarse que tienen a su disposición los usuarios, como las menciones, *retweets*, *likes*, seguidores, entre otros. Para esta tarea exploran distintos cálculos que dependen en su mayoría de la distancia que presentan los usuarios entre sí en las redes creadas, como *random walk*, *edge betweenness*, modularidad, momento de dipolo, etc. Así también, aplican algunas técnicas de procesamiento de lenguaje natural, como *bag of words* y análisis sentimental, al contenido de los *tweets*, de manera que también obtener métricas basadas en texto.

La principal influencia de este trabajo en la presente memoria fue la construcción de los grafos de *retweets*, que se consideraron más prometedores para su uso en posteriores métricas. Además de esto, se tiene la aplicación de las métricas de *random walk* y *edge betweenness*, que fueron las que presentaron los mejores resultados con una relevante robustez, respecto a las medidas con metodologías similares, por lo que pareció pertinente utilizarlas para este trabajo.

2.1.3. A Motif-Based Approach for Identifying Controversy

En este trabajo realizado el año 2017 por Coletto et al. [17] detectan la controversia haciendo uso de motivos de redes, o más bien, patrones locales de redes, como la direccionalidad en seguimientos y *retweets*. Para ello, utilizan un dataset compuesto de un gran rango de dominios, como noticias, políticos, celebridades, chismes y entretenimiento, donde se les extrajo a cada página los 200 *tweets* más recientes y se evaluó cada uno manualmente si es que se consideraban controversiales o no. Se consideraron solamente las páginas que contienen principalmente *tweets* de carácter controversial, llegando a obtener más de 190 mil *tweets* en total.

Dada una red social, el trabajo se enfoca en modelar las interacciones entre usuarios y la dinámicas que ocurre entre estos a través de distintas redes. En primer lugar, se tiene un grafo de seguidores, donde los nodos representan a usuarios y las aristas si es que un individuo sigue a otro. Luego se tiene un árbol que une a los elementos de un set de contenidos, es decir, *tweet*, a partir de un arco que representa que si es que un tweet es la respuesta a otro. Finalmente, se tiene otro grafo donde cada nodo es un usuario y cada arista indica que un usuario a respondido a algún contenido de otro. La hipótesis que se presenta es que todas las estructuras construidas se pueden caracterizar por motivos simples de interacciones de usuarios locales, o también características del *baseline* que permitan identificar la existencia de controversia, como estructura de la red, propagación de contenido y características temporales.

De acuerdo a la estructura de la red, se decide por utilizar dos tipos de métricas: el tamaño de la red y el grado de los nodos. Para el primer caso, se obtiene que la cantidad de nodos y aristas no guarda una correlación con si un tema es controversial o no, pero si ocurre con respecto al nivel de grado promedio de la red. Para la propagación de contenido se buscan los subárboles de mayor tamaño para cada tema con el fin de observar con claridad cuál es el nivel de profundidad de las interacciones de usuarios. Se obtiene que efectivamente las conversaciones acerca de temas controversiales presentan una mayor propagación. Para las características temporales se toma el tiempo de demora de los usuarios en responder a otro y el porcentaje de respuestas dentro de una hora. Finalmente, para la hipótesis, se estudian los patrones locales como posible identificador de controversia, donde para ello se hace uso

del grafo de seguidores y de respuestas, siguiendo la idea de que los usuarios que existe una mayor probabilidad de que un usuario le responda a otro que no sigue, dado si el tema de conversación es controversial. Además, se evalúa el nivel de complejidad de ambas redes combinadas, es decir, de acuerdo a las siete posibles combinaciones de interacciones entre usuarios, como por ejemplo, si un usuario responde a otro, pero no lo sigue. Teniendo ya todas las características pertinentes, se aplican distintos tipos de modelos de clasificación de manera de obtener la característica más relevante, para lo cual se escogió AdaBoost. Se concluye que dentro de las características que presentan más utilidad al momento de identificar y cuantificar si un tema es controversial o no, se tiene el tiempo de respuesta de usuarios y el nivel de grado máximo relativo como las mejores opciones.

2.1.4. A Measure of Polarization on Social Media Networks Based on Community Boundaries

Este trabajo, realizado en 2013 por Guerra et al. [18], tiene como enfoque la medición de la controversia a través de la modularidad. Para ello, utilizan datasets ya existentes y públicamente disponibles de distintas redes sociales, con usuarios que se unen por distintas mecánicas, como *retweets*, seguidores y respuestas, y de diversos contextos; como redes de amistades de una universidad, seguidores de equipos de deportes, blogs de política y control de armas.

Para ello, se calcula la modularidad para cada par de comunidades, siendo esta capaz de cuantificar el grado a la que los vértices se agrupan en grupos comunitarios y, cuanto mayor sea su valor, más modular es la red. El enlace entre un valor alto de modularidad y de polarización tiende a la implicancia de que la ausencia de interacciones entre los distintos nodos se explica como una oposición entre los usuarios. El fin de la modularidad es comparar la conectividad interna y externa de dos grupos. Es por ello que, para comprender este fenómeno, este trabajo estudia a partir de la construcción de grafos que representen a dos comunidades para cada tema. Para cumplir esto, se aplica una segunda métrica que permita cuantificar la polarización de manera de comparar con la modularidad, donde se logra observar cierta correlación entre ambas métricas; sin embargo, se concluye que la modularidad por sí sola difícilmente puede considerarse una métrica infalible para cuantificar la polarización, dado que esta mide simultáneamente la homofilia, la tendencia de las personas por la atracción a sus homónimos, y el antagonismo entre grupos, lo que no permite evidenciar fácilmente la oposición entre ambos.

2.1.5. Measuring Political Polarization: *Twitter* shows the two sides of Venezuela

Este trabajo, realizado el año 2015 por Morales et al. [19], se basa principalmente en el caso de la muerte del presidente venezolano Hugo Chávez, obteniéndose más de 16.383.490 mensajes escritos por 3.173.090 usuarios entre el 4 de febrero de 2013 y el 4 de mayo de 2013; siendo un período de tiempo que toma en cuenta un mes antes de su muerte, el anuncio de esta y los eventos posteriores a este punto, incluyendo las próximas elecciones presidenciales.

Se propone una medida de polarización a partir de la idea de qué diferenciación, que existe dentro de un conjunto de individuos, se produce cuando hay dos grupos con un tamaño similar entre sí y con opiniones opuestas. Por lo tanto, se plantea un método que cuantifica

ambos efectos para una distribución obtenida de cierto modelo, basado en el momento de dipolo eléctrico. Esta métrica es aplicada a 56 grafos de *retweets* construidos a partir de cada uno de los días dentro del periodo de observación, donde cada enlace entre nodos representa un *retweet*, acción que puede considerarse como un apoyo hacia el otro usuario.

Observando los resultados, se puede apreciar como la polarización es alta en general antes de la muerte de Hugo Chavez; sin embargo, luego de este evento el nivel de controversia crece en gran medida y no vuelve bajar cómo a los días anteriores, lo que se debe a como ese suceso gatilló que usuarios de otros países se unieran a la discusión, que permitió que aumentara la polarización.

Finalmente, luego de evaluar el modelo y validar los datos al comparar con la distribución geográfica de los usuarios polarizados, se logra concluir que las áreas con la mayor concentración de usuarios que estaban alineados con el oficialismo, corresponden a las partes de la ciudad con las concentraciones más grandes de los vecindarios más pobres. Consiguiente con esto, los usuarios de la oposición se concentran en las regiones urbanas más desarrolladas. Todo esto permite inferir que existe una correlación entre la polarización de la población venezolana y los factores socioeconómicos y su geolocalización.

2.1.6. Every Colour You Are: Stance Prediction and Turnaround in Controversial Issues

Graells-Garrido et al. [20], del año 2020, realiza un estudio de la controversia del caso del proceso legislativo del aborto tanto en Chile como en Argentina. Para este propósito, se realizó una búsqueda exhaustiva desde el año 2015, que fue el año en que la expresidente de Chile, Michelle Bachelet, propuso el proyecto de las tres causales de aborto, siendo esta aprobada en 2017; hasta el año 2018, donde se sostuvo el debate legislativo del aborto libre en Argentina durante el gobierno de Macri.

Los *tweets* que se analizaron en este trabajo fueron extraídos a través de la *Twitter Streaming API*, donde los parámetros de consulta fueron palabras clave relacionadas con el aborto utilizando el operador OR, además de incluir términos como frases, hashtags y usuarios; lo que llevó a conseguir sobre 31.4 millones de *tweets* de 1.8 millones de usuarios. De esta manera, se procedió a procesar el dataset y filtrar de acuerdo a ciertos aspectos, como el género y el país de origen, por lo que se descartaron todos los perfiles que no cumplían con tener estas características de manera clara. A su vez también se filtraron todos los usuarios que no pertenecieran al componente conectado más grande (LLC), eliminando así un 16 % de los nodos de la red; por lo que, finalmente, el dataset quedó con 6 millones de *tweets* de 663 mil usuarios.

Luego de tener los datos, se procede a etiquetar cada usuario con su localización, edad, género y postura con respecto al aborto, utilizando un conjunto de términos que pertenecían a ambos grupos. Para el caso de los que no se pudieron etiquetar con este método, se aplica un modelo que permite predecir estas características con un alto umbral de confianza. Entonces, se entrenan distintos modelos para lograr predecir la postura de cada usuario con el objetivo de evaluar que características son más importantes que las demás. Finalmente, se infiere la relación entre aspectos de los perfiles de los usuarios y los cambios de ideologías, es decir, si es que un individuo, que pertenece a un grupo, cambió su mentalidad hacia el bando contrario.

Dentro de los resultados, se observó que los *emojis* expresan de manera significativa más como manifestaciones físicas, que como símbolo de adherencia o soporte a su grupo. Mientras que otros términos como “mujeres” y “derecho a decidir” suelen aplicarse más a los grupos a favor del aborto, y “vida/muerte” y “derecho a nacer” a quienes estén en contra.

Por último, se logra concluir que el cambio de postura realmente ocurre, que fue posible de medir en base a características disponibles públicamente, como *emojis* y términos dentro de la biografía; además de atributos predichos, como localización, edad y género, permitiendo dar una base sólida para analizar los cambios de postura de manera cuantitativa

2.1.7. Quantifying Content Polarization on *Twitter*

Uno de los trabajos más relevantes en cuanto al tema con respecto a análisis de controversia haciendo uso exclusivamente del contenido de los *tweets*, siendo creado en 2018 por Yan et al. [9]. Este paper se enfoca en las elecciones de EEUU de 2016, para lo cual se extrajeron los datos requeridos de la siguiente manera: se buscaron los seguidores de Trump y Clinton en dos puntos temporales, entre el 30 de agosto de 2016 y 15 de noviembre de 2016. Después, se identifican los usuarios que siguen exclusivamente a cada candidato, sin seguir al contrincante. Es así entonces como, luego de descartar los usuarios con menor actividad, se obtuvieron datos de un poco más de 87 mil seguidores de Trump y 74 mil, de Clinton. Por último, se extraen todos los posts dentro del periodo de tiempo establecido para cada usuario.

La metodología utilizada incluye dos pasos importantes. En primer lugar, se aplica lematización, tokenización y Word2Vec *embedding* para transformar las palabras de los *tweets* en vectores. Entonces se valida la representación de estos *tweets* con *stance detection*, donde se aplican una variedad de modelos de clasificación, como Naive Bayes, SVM, C4.5 Tree, y redes neurales de una capa, donde el *support vector machine* presenta los mejores resultados. El fin de este método es observar la dificultad de los modelos al predecir dependiendo de que tan diferente son los vectores entre sí, lo que se traduce en la disimilitud de las palabras utilizadas. Luego, para cuantificar la polarización de contenido, se calculan las distancias promedio entre todos los *tweet* que pertenecen a un mismo o distintos grupos. Entonces, se evalúa la correlación entre las distancias obtenidas y la métrica AUC de los modelos de clasificación para examinar la efectividad de la cuantificación.

Como se pueden ver en los resultados, existe en efecto una correlación entre ambas características, por lo que realmente el modelo de *stance detection* tiene mejores resultados mientras más similares son las palabras empleadas por los usuarios. Además, se analiza la polarización del contenido utilizando una métrica que se basa en las distancias promedio de grupo y entre grupos, denominada *normalized cut*, que permite ver la polarización para cada semana dentro del periodo de tiempo estudiado.

La conclusión final de esta investigación es que los Word2Vec *embedding* tienen un buen desempeño al momento de cuantificar la controversia del texto encontrado en los *tweets*. Además que eventos relevantes durante el periodo observado gatillan la polarización y como tópicos negativos tienden más a la controversia que los neutrales.

2.2. Marco Teórico

A continuación se describen ciertos términos y técnicas que son relevantes para el trabajo actual, siendo de las más sobresalientes, el Word2Vec y el análisis sentimental.

2.2.1. Grafos

Los grafos o redes sirven como modelos matemáticos para analizar con éxito muchos problemas concretos del mundo real. Ciertos problemas de física, química, ciencias de la comunicación, informática, genética, psicología, sociología y lingüística pueden formularse como problemas de teoría de grafos [16].

Sea un grafo G un par ordenado $G = G(V, E)$, donde:

- V es un conjunto de vértices o nodos, y
- E es un conjunto de aristas que relacionan estos nodos
- El orden de un grafo G se denomina a su número de vértices
- El grado de un vértice o nodo, $v \in V$, es igual al número de aristas que lo tienen como extremo

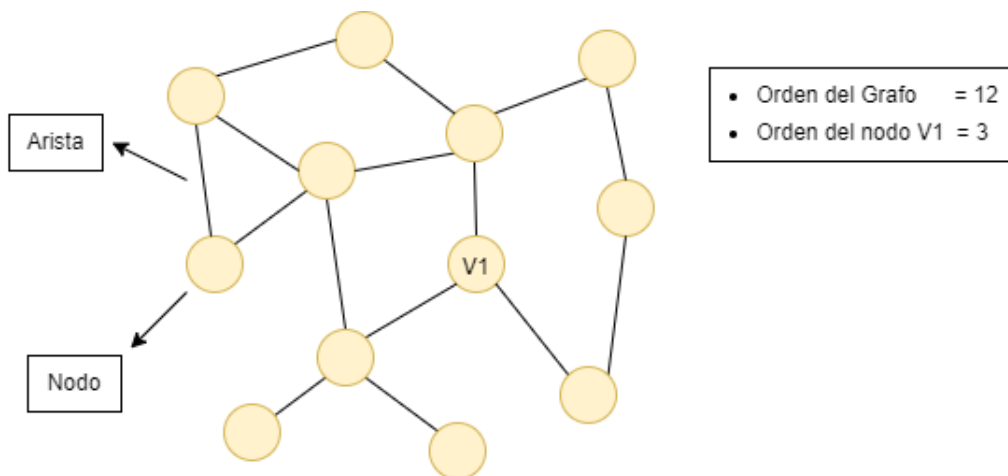


Figura 2.1: Ejemplo de grafo.

2.2.2. Lematización

La lematización es el proceso de determinar la forma de diccionario de una palabra (por ejemplo, nadar) dada una de sus variantes conjugadas (por ejemplo, nadando, nadaría, nadó). Los lematizadores basados en datos enfrentan dos desafíos principales: primero, generalizar más allá de los datos de entrenamiento para lematizar palabras invisibles; y segundo, para eliminar la ambigüedad de formas de palabras ambiguas de su contexto de oración [13].

2.2.3. Tokenización

Antes de definir este concepto, es importante explicar en primer lugar la diferencia entre un lexema y un token. Un lexema es una secuencia de caracteres en el programa fuente que coincide con el patrón de un token y que el analizador léxico identifica como una instancia de ese token. Mientras que un token es un par que consta de un nombre de token y un valor de atributo opcional. El nombre del token es un símbolo abstracto que representa un tipo de unidad léxica, por ejemplo, una palabra clave particular o una secuencia de caracteres de entrada que denotan un identificador (sea *number* un token, una muestra de lexema puede ser 3.14159 o 1). Los nombres de los tokens son los símbolos de entrada que procesa el analizador (o *parser*). Otra forma de pensar en la relación entre los dos es que un token es una estructura de programación utilizada por el analizador que tiene una propiedad llamada lexema que contiene el carácter/símbolos de la entrada [12].

Luego, la tokenización es el proceso de demarcar y posiblemente clasificar secciones de una cadena de caracteres de entrada. Los tokens resultantes, luego, se pasan a alguna otra forma de procesamiento. El proceso puede considerarse una subtarea del *parsing*.

2.2.4. Word2Vec

Word2vec es un grupo de modelos relacionados que se utilizan para producir incrustaciones de palabras (*word embeddings*). Estos modelos son redes neuronales superficiales de dos capas que están entrenadas para reconstruir contextos lingüísticos de palabras. Word2vec toma como entrada un gran corpus de texto y produce un espacio vectorial, normalmente de varios cientos de dimensiones, en el que a cada palabra única del corpus se le asigna un vector correspondiente en el espacio, tal como se observa en la Figura 2.2. Los vectores de palabras se colocan en el espacio vectorial de manera que las palabras que comparten contextos comunes en el corpus se ubican cerca unas de otras en el espacio [14].

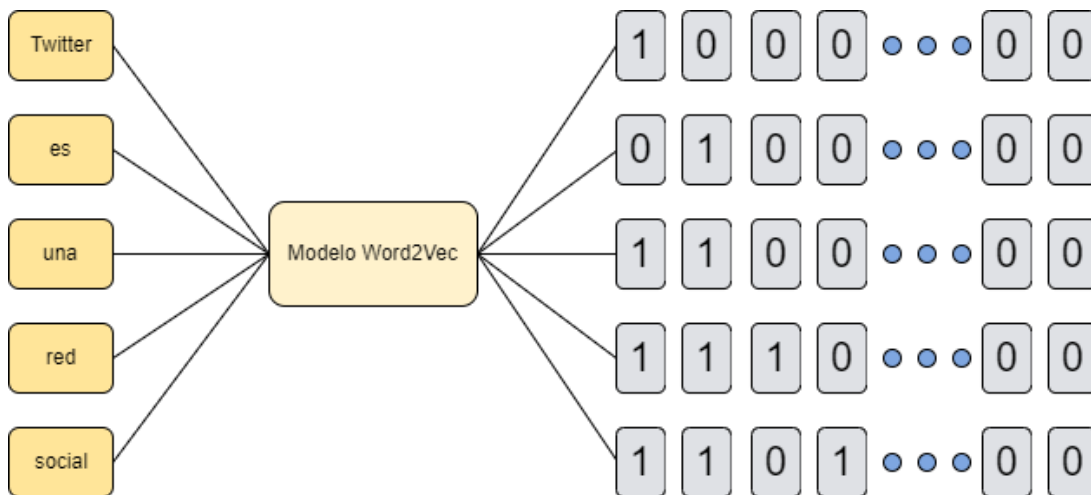


Figura 2.2: Explicación de la técnica Word2Vec.

Word2vec puede utilizar dos tipos de arquitectura para producir una representación distribuida de palabras: *continuous bag-of-words* (CBOW) o *skip-gram* continuo. En la arquitectura de CBOW, el modelo predice la palabra actual desde una ventana de palabras de contexto

circundantes. El orden de las palabras del contexto no influye en la predicción. En la arquitectura de *skip-gram*, el modelo usa la palabra actual para predecir la ventana circundante de palabras de contexto. Además, en ésta pesan más las palabras de contexto cercanas que las palabras de contexto más distantes.[14] [15]

Siguiendo el trabajo de Yan et al. [9], es que se prefiere utilizar CBOW por sobre *skip-gram*, dado que este primero se desempeña de mejor manera para documentos cortos, lo cual es bien aplicable en el caso de este trabajado, considerando que el tamaño máximo de un tweet es de 280 caracteres.

2.2.5. Análisis Sentimental

La polaridad de sentimiento es una característica particular de un texto. Suele ser dicotomizada en dos, como positivo y negativo, pero también puede ser pensada como un rango. En realidad, un documento puede contener una variedad de opiniones, que incluso pueden tener una polaridad mixta entre sí, que es diferente de no tener ningún tipo de polaridad, que es donde recaería el concepto de objetividad. También, es importante recalcar la diferencia entre la polaridad de sentimiento y su nivel de intensidad. A modo de ejemplo, una persona puede sentirse muy bien con un producto o servicio no tan bueno, o mal con uno que puede ser bastante bueno, pero no excelente [11].

El análisis sentimental tiene distintas tareas a considerar, siendo la primera, la detección del sentimiento u opinión en sí, lo que puede verse como una clasificación de texto entre lo que es objetivo y lo que no. Aunque, generalmente este tipo de detección es basada en la examinación de los adjetivos y algunos sustantivos en las oraciones de un texto.

La segunda tarea es la clasificación de polaridad. Dado un extracto de texto de opinión, el objetivo final es clasificar la opinión de dentro dos polaridades de sentimiento totalmente opuestas, o ubicarlas entre medio de ambas, lo que se denomina como neutralidad. Siguiendo esta idea, el problema del análisis sentimental puede resolverse en base a distintas cantidades de clases, ya sea de manera binaria tomando solo los aspectos positivos y negativos del texto, trinaria si se agrega la neutralidad, o varias clases si se toma en cuenta la intensidad del sentimiento. Por ejemplo, demostrar el nivel de positividad o negatividad a través de valores del 1 al 4 y -1 al -4, respectivamente, siendo 0 si se trata de una opinión objetiva y neutral. Siendo este método el que se aplicó a los *tweet* de este trabajo.

2.2.6. Twitter y tweets

Twitter es una red social ampliamente utilizada en todo el mundo, basada en el microblogueo, es decir, en la publicación de textos de no más de 240 caracteres (anteriormente de máximo 140 caracteres). En esta plataforma los usuarios se conectan mediante el seguimiento de otros usuarios, además de otras interacciones que se incluyen como: el uso del “me gusta” o “like” en un tweet; el “retweet”, que permite compartir un tweet en la línea temporal de la cuenta del usuario, dando la posibilidad de agregar una respuesta, empezando así un hilo de conversación entre dos individuos; y por último, las menciones, en las que un usuario puede etiquetar la cuenta de otro, agregando un texto de manera que éste pueda leerlo y así, creando un nuevo hilo de conversación.

2.2.7. *Twitter* API

Siendo una API o *application programming interface* (interfaz de programación de aplicaciones en español) un conjunto de definiciones y protocolos que se usa para diseñar e integrar el software de las aplicaciones. En este trabajo se hace uso en particular de la API desarrollada por *Twitter*, la cual es el método utilizado para la obtención de los datos necesarios para su eventual procesamiento y análisis.

La *Twitter* API tiene una variedad de planes que permiten extraer datos dependiendo del propósito y origen del proyecto en el que se utilizarán. Para el propósito de esta memoria, la mejor opción era el plan *Academic Research Access*, el que permite el acceso a full-archive search y otros endpoints, que da el acceso para buscar *tweet* desde marzo del 2006 hasta la actualidad, con un tope mensual de 10 millones de *tweet*. Para este fin, la plataforma de desarrollador de *Twitter* entrega una “bearer token key”, un código capaz de acceder a todos los puntos de salida entregados.

Capítulo 3

Metodología

Este capítulo tiene como objetivo describir detalladamente varios procesos realizados a lo largo de este proyecto. En primer lugar, la extracción de datos a través de la API de *Twitter*, luego la obtención de todos los términos, usuarios y hashtags relevantes, después la explicación de como se calcularon las métricas asociadas, así como también la construcción de los grafos de nodos.

3.1. Obtención de Datos

Con respecto a la descarga de los datos, se consideró pertinente en un principio hacer uso de términos “semilla”, es decir, palabras o frases que engloben de manera general el tema a analizar y través de éste, buscar otros términos, usuarios y hashtags que puedan aportar en la expansión de la búsqueda de los bandos que se contraponen entre sí. La principal razón de este enfoque es debido a la limitancia que implica usar un solo concepto, dado que unos pocos términos pueden dar información acerca del tópico, pero no específica la postura de partidarios opuestos en el debate, dado que estos harán uso de palabras cargadas de significado para su bando, con la intención de dar a entender de que lado están ellos de la discusión.

Considerando que esta memoria tiene como objeto de estudio la población chilena, y sin importar que tan específicos puedan ser ciertos tópicos para este país, siempre pueden haber posts con los mismos términos. Se decidió utilizar un diccionario con todas las ciudades, comunas y regiones en Chile con el fin de filtrar aquellos *tweet* de usuarios que no pertenecen al país. Conforme a lo anterior, se logró el objetivo propuesto.

Otro desafío a tomar en cuenta era la necesidad de evitar escoger manualmente los términos específicos de cada postura, por lo que se decidió por realizar el cálculo de una métrica, o también denominado algoritmo de *clustering*, que permita encontrar coincidencias entre los diversos *tweet* de manera cuantitativa, la cual se define a continuación, de acuerdo al trabajo de Garimella [7]:

$$sim(h_s, h_t) = \frac{1}{1 + \log(df(h_t))} (\alpha \cos(W_s, W_t) + (1 - \alpha) \cos(H_s, H_t)) \quad (3.1)$$

Donde h_s es el hashtag semilla; h_t es el hashtag candidato; W_x y H_x son el set de palabras

y hashtags que coexisten con el hashtag h_x , respectivamente; \cos es la similaridad de coseno entre dos vectores; df es la frecuencia de documento, es decir, la cantidad de *tweet* en las que se encuentra el término dividido el total de *tweet*; y α es un parámetro que equilibra la importancia de palabras comparada a la de hashtags en un tweet.

Conforme a ello, es importante mencionar que, considerando que la cantidad de *tweet* no es de tal envergadura comparada a la utilizada en otros trabajos de la misma índole, y además de que siendo la población de estudio ubicada netamente en Chile y, por lo tanto, el uso de hashtags no es tan frecuente como en todo el mundo, se optó por no aplicar la ecuación con un set de hashtags solamente, sino que por un set de tanto hashtags como palabras relevantes, siendo escogidas de acuerdo a frecuencia absoluta con respecto al resto de *tweet*. Entonces, considerando los cambios, el puntaje de similitud quedaría de la siguiente forma:

$$sim(T_s, T_t) = \frac{1}{1 + \log(df(T_t))} (\alpha \cos(W_s, W_t) + (1 - \alpha) \cos(H_s, H_t)) \quad (3.2)$$

Siendo T_s y T_t las palabras o hashtags semilla y palabras o hashtags candidatos, respectivamente.

Posteriormente, de manera de tener un panorama claro de tanto el tema como de los usuarios, se decide aprovechar lo más posible la API procediendo a extraer la mayor cantidad de datos con respecto a cada tweet, obteniéndose las siguientes columnas, en la Tabla 3.1:

Tabla 3.1: Columnas extraídas y ejemplos.

Columnas	Descripción	Ejemplos
author_id	ID del autor del tweet	100146769
created_at	Fecha del tweet	2019-10-11 T17:37:16Z
geo	Geolocalización del usuario o del tweet	Santiago de Chile
id	ID del tweet	1182532797363294210
like_count	Nº de Likes	10
quote_count	Nº de Menciones	10
reply_count	Nº de Comentarios	10
retweet_count	Nº de <i>Retweet</i>	10
source	Tipo de Dispositivo en el que se tuiteó	Twitter for iPhone
ref_id	ID del autor del tweet original	1182477877591318528
tweet	Texto que fue posteoado por el usuario	Sr #Presidente y su #rechazó a los #terroristas y #delincuentes que destruyen y quemar propiedad privada sin respeto al #derecho ajeno en la #ZonadelaAraucania https://t.co/EzFO6JyhLv

Luego de realizar una segunda búsqueda de *tweets* con los términos y hashtags obtenidos, gracias al algoritmo explicado anteriormente, se comienzan nuevas búsquedas con el fin de crear conexiones entre los usuarios. Para este caso, se hace uso de dos columnas dentro de los datasets extraídos: *author_id* y *ref_id*. La metodología se explica en el diagrama de bloques que se puede observar en la Figura 3.1:

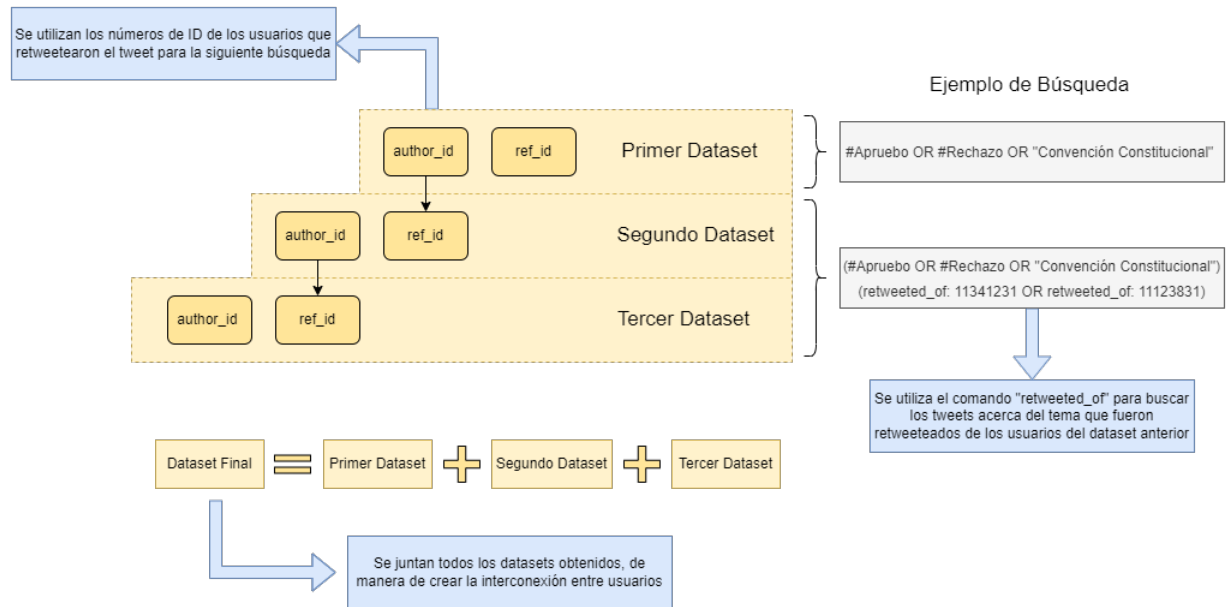


Figura 3.1: Búsqueda de usuarios para la creación de grafos.

Tal como se explica en la Figura 3.1, se realiza una búsqueda consecutiva de los usuarios a través del comando “retweeted_of”, esto debido a que, al extraer datos con respecto a un tema, no implica que necesariamente los usuarios que *retweetean* van a estar interconectados entre sí. Por lo tanto, se ideó este plan que permite obtener los números de ID requeridos para la construcción de los grafos y se produzcan ramificaciones, en vez de que un solo nodo esté conectado a todos los demás, tal como se explica en la Figura 3.2:



Figura 3.2: Grafos construido a partir de uno o todos los datasets.

Y así, por como se aprecia en la Figura 3.2, es posible crear los grafos, dado que se logran unir usuarios con otros, permitiendo la creación de los grafos requeridos con los que sea posible calcular las métricas propuestas con anterioridad.

3.2. Construcción de Grafos de *Retweets*

Ya teniendo a disposición todos los ID necesarios, se empieza esta tarea calculando el *Large Connected Component*, es decir, la red de nodos interconectados más grande de todas, dado que fácilmente pueden formarse redes tipo estrella en donde varios nodos convergen a solo uno de estos, generando así varios grafos en vez de uno. Para este fin, se utilizó la librería *networkx* de Python, que presenta una variedad de herramientas para la construcción de grafos. En consecuencia, se procede a buscar las particiones existentes dentro de la red, para lo que se hace uso de la librería *METIS*, que se especializa en darle un grupo a cada nodo perteneciente a un grafo, de manera que se aplica la cantidad de particiones de acuerdo a las que se le indique a la librería. Finalmente, se crean dos archivos, uno que tiene los nodos con los números de ID para identificarse y el número que dice a qué partición pertenecen. Y otro que tiene las aristas que se crearon entre los nodos. Ambos archivos se suben al software *Gephi*, que permite crear visualizaciones de grafos en base a varios tipos de algoritmos, basándose solamente en la estructura del grafo, siendo el caso de uso de esta memoria, el algoritmo *ForceAtlas2*.

3.3. Métricas basadas en Grafos

A continuación, se explican las métricas, y su respectivo cálculo, con las que se decidió cuantificar el nivel de controversia en todos los temas planteados:

3.3.1. Random Walk

Sea un grafo y un nodo en particular, se selecciona un vecino de él de manera aleatoria y se mueve a este, luego se selecciona un vecino de este nuevo nodo y se mueve a este, entonces se repite esta acción consecutivamente. La secuencia aleatoria de puntos que fueron seleccionados de esta forma es lo que se denomina un *random walk* [10].

Esta métrica utiliza la noción de que en una discusión de carácter controversial, existen usuarios autoritarios en ambos grupos, que se evidencian por un alto grado en el grafo. Así, es posible capturar que tan probable un usuario aleatorio de cualquier grupo sea expuesto al contenido de un usuario autoritario que pertenece al grupo contrario [7].

Sea entonces $G(V,E)$ el grafo construido y las dos particiones X e Y , ($X \cup Y = V$, $X \cap Y = \emptyset$). En primer lugar, se distingue el nodo k de más alto grado de cada partición. El alto grado permite ver el nivel de autoridad de un nodo, dado que implica que el tweet del usuario fue apoyado una cierta cantidad de veces a través de un *retweet*. Posteriormente, se selecciona una partición de manera aleatoria con una probabilidad de 0.5 para cada una y se toma un nodo donde se considerará el comienzo del *random walk* en esa partición. La caminata termina cuando logra encontrar un nodo de alto grado, ya sea en cualquier lado.

Se define entonces la métrica basada en *random walk* de la siguiente manera: se consideran dos *random walks*, una que termina en la partición X y otro en Y . La métrica entonces sería la diferencia de las probabilidades de dos eventos: (i) ambos *random walks* empiezan desde la partición en la que terminaron y (ii) ambos *random walks* empiezan en una partición contraria a la que terminaron. La métrica se define matemáticamente de la siguiente forma:

$$RWC = P_{XX}P_{YY} - P_{YX}P_{XY}, \quad (3.3)$$

donde P_{AB} , $A, B \in X, Y$ es la probabilidad condicional,

$$P_{AB} = Pr[\text{start in partition A} \mid \text{end in partition B}] \quad (3.4)$$

Por consiguiente, RWC es cercano a uno cuando la probabilidad de cruzar grupos es baja, cercana a cero, cuando la probabilidad de cruzar grupos es comparable con la de quedarse en el mismo lado.

Siguiendo el curso de otros trabajos con fines similares a esta memoria, se encontró el gran costo computacional que implica el cálculo de esta métrica en particular, dado que uno de los métodos más directos para realizar un cómputo del *Random Walk* es través del muestreo de Monte Carlo, el cual se considera lento y poco eficiente. Es por ello que se decide aplicar una variante en el cálculo, tal como se explica en el trabajo de Garimella et al. [7], donde se mencionan los problemas con el *random walker*, y como éste puede quedar fácilmente atascado en un nodo, debido a la estructura de estrella de algunos a causa de su nivel de autoridad con respecto al esparcimiento de información que generan estos (ej: cuentas de noticieros).

Con respecto a esta variante de *Random Walk*, se comienza asumiendo la existencia de dos lados en la discusión, que se definen como dos sets de nodos X e Y. Para este caso, tal como se explicó en el capítulo anterior, en vez de que múltiples *random walkers* detengan su movimiento cuando llegan a un nodo con un alto grado, este algoritmo decide que reinicien su camino, con el fin de que no queden atascados rápidamente.

Se consideran entonces dos instancias de un *random walk* con restart, en función de si comienzan (y reinician) desde X (start = X) o Y (start = Y). Observando ambos casos, cuando ocurre que start = X, el método tiene un vector de reinicio distribuido uniformemente sobre X y cero para los vértices en Y (la situación es simétrica para inicio = Y). Además, el *random walk* se ejecuta en un grafo que fue modificado anteriormente con todas las aristas que salen de los vértices de alto grado eliminados. Esto transforma los nodos de alto grado en nodos colgantes, o nodos de solo un grado, lo que obliga a que el *random walk* se reinicie una vez que alcanza uno de estos nodos.

Con el fin de definir las siguientes ecuaciones, sean P1 y P2 las distribuciones estacionarias del *random walk* obtenido para start = X y start = Y, respectivamente. Consideramos la probabilidad condicional $Pr[\text{start} = A \mid \text{end} = B^+]$ que el *random walk* había comenzado en el lado $A \in X, Y$, dado que en algún paso en estado estacionario se encuentra en uno de los nodos de alto grado del lado $B \in X, Y$ (denotado como B^+). Por lo tanto, se consideran las siguientes cuatro probabilidades:

$$P_{X, X^+} = Pr[\text{start} = X \mid \text{end} = X^+] = \frac{\frac{|X|}{|V|} \sum_{v \in X^+} P_1(v)}{\frac{|X|}{|V|} \sum_{v \in X^+} P_1(v) + \frac{|Y|}{|V|} \sum_{v \in X^+} P_2(v)}, \quad (3.5)$$

$$P_{X,Y^+} = Pr[start = X|end = Y^+] = \frac{\frac{|X|}{|V|} \sum_{v \in Y^+} P_1(v)}{\frac{|X|}{|V|} \sum_{v \in Y^+} P_1(v) + \frac{|Y|}{|V|} \sum_{v \in Y^+} P_2(v)}, \quad (3.6)$$

$$P_{Y,Y^+} = Pr[start = Y|end = Y^+] = \frac{\frac{|Y|}{|V|} \sum_{v \in Y^+} P_1(v)}{\frac{|X|}{|V|} \sum_{v \in Y^+} P_1(v) + \frac{|Y|}{|V|} \sum_{v \in Y^+} P_2(v)}, \quad (3.7)$$

$$P_{Y,X^+} = Pr[start = X|end = X^+] = \frac{\frac{|Y|}{|V|} \sum_{v \in X^+} P_1(v)}{\frac{|X|}{|V|} \sum_{v \in X^+} P_1(v) + \frac{|Y|}{|V|} \sum_{v \in X^+} P_2(v)}, \quad (3.8)$$

La variante del puntaje RWC ahora se puede definir finalmente como:

$$RWC = P_{X X^+} P_{Y Y^+} - P_{X Y^+} P_{Y X^+}, \quad (3.9)$$

que, al igual que la versión original, captura intuitivamente la diferencia en la probabilidad de permanecer en del mismo lado y cruzando el corte entre particiones.

3.3.2. Betweenness

Dentro del análisis de grafos, es de carácter relevante tener conocimiento de qué aristas o nodos tienen una mayor importancia dentro de una red. Es aquí donde recae el uso de la medida *betweenness*, que se define como la cantidad de los caminos más cortos que pasan a través de una arista en particular, y la cual permite cuantificar esto mediante la siguiente ecuación [8]:

$$b(e) = \sum_{\{k,l\} \in V} \frac{\sigma_{k,l}(e)}{\sigma_{k,l}}, \quad (3.10)$$

donde $\sigma_{k,l}$ es la cantidad de los caminos más cortos entre vertices k y l , y $\sigma_{k,l}(e)$ es el número de los caminos más cortos que conectan k y l que pasen a través de la arista e

La aplicación de esta métrica es que se entiende que al existir dos particiones que estén bien separadas entre sí en un grafo. Serán pocas las aristas que se formen entre ambas y, por lo tanto, el valor del *betweenness* será mayor, a diferencia de una red con un gran cantidad de conexión entre particiones, de lo cual se intuye que el *betweenness* será parecido tanto en el corte del grafo como en el resto de los nodos, tal como lo explica Garimella et al. [7].

De manera de cuantificar estas diferencias es que se separan las distribuciones de los *betweenness* entre las aristas, de las que se forman en el corte y en el resto del grafo, y se calcula la función de densidad de probabilidad haciendo uso de estimación de densidad de kernel, con el fin de obtener la divergencia KL d_{KL} y realizar un muestreo de 10000 puntos por cada

distribución, se define entonces la métrica final utilizada en este trabajo:

$$BCC = 1 - e^{-d_{KL}}, \quad (3.11)$$

que asume valores cercanos a cero cuando la divergencia es pequeña y cercanos a 1 cuando la divergencia es grande.

3.4. Métrica basada en NLP

Antes de introducir la métrica que se explica a continuación, se debe tomar en cuenta que a cada tweet fue procesado mediante una lematización y se le extrajeron las *stop words*, con el fin de simplificar el trabajo de ambos procesos.

3.4.1. Normalized Cut

Además de los procesos aplicados anteriormente, para el caso de esta métrica se utilizó la librería Word2Vec para transformar todas las palabras y en vectores con 100 elementos con el fin de realizar los cálculos de las distancias euclidianas. Posteriormente, se promedian entre sí todas las palabras pertenecientes a un tweet con el objetivo de tener una representación adecuada de cada uno y calcular distancias euclidianas entre estos. Se considera esta métrica por sobre otras, como la similitud de coseno, por funcionar mejor con documentos de corta extensión, tal como se explica en Yan et al.[9].

Para cuantificar el contenido inherente en los *tweet*, se consideran dos métricas de *clustering*. La distancia promedio de grupo, que se define como todos las distancias entre par de *tweet* que pertenecen a un mismo grupo o partición. Mientras tanto, la distancia promedio entre grupos, definida como el promedio entre par de *tweet* que pertenecen a distintas particiones. Sean entonces los grupos identificados como A y B, la distancia entre grupos se puede definir como:

$$D_{AB} = \frac{\sum_{p \in A, q \in B} \|\vec{V}_p - \vec{V}_q\|}{N}, \quad (3.12)$$

donde V_p y V_q son los vectores que representan a los *tweet* p y q, respectivamente, y $\|\vec{V}_p - \vec{V}_q\|$, denota la distancia euclidiana entre vectores. Luego, la distancia de grupo se define como:

$$D_X = \frac{\sum_{p \in A, q \in B, p \neq q} \|\vec{V}_p - \vec{V}_q\|}{N}, \quad (3.13)$$

Las distancias a nivel de grupo fueron diseñadas con el objetivo de cuantificar las relaciones y diferencias entre el texto de los *tweet* de cada grupo. Esta definición viene de la intuición que cada grupo o partición es capaz de representarse a través de burbujas de información que no necesariamente tienden a interactuar entre sí, en el caso dado que el tema en cuestión sea de carácter controversial y, por lo tanto, de acuerdo si los valores de estas distancias son bajos o altos, implica que existe una gran o pequeña semejanza, respectivamente, entre los *tweet*. De esta manera, la distancia de grupo permite explicar la diversidad de lenguaje utilizado dentro de un mismo grupo. En cambio, la distancia entre grupos dice que tan diferentes son los grupos en vocabulario. Así, se consideran ambas métricas como aspectos ortogonales en la polarización de contenido, de forma que se utilizan ambas medidas para construir una métrica denominada *Normalized Cut* (NC), que se define a continuación:

$$NC = \frac{\sum_{i=1}^n d_{AB}(i)}{\sum_{i=1}^{n_A} d_{AB}(i) + \sum_{i=1}^n d_A(i)} + \frac{\sum_{i=1}^n d_{AB}(i)}{\sum_{i=1}^{n_B} d_{AB}(i) + \sum_{i=1}^n d_B(i)} \quad (3.14)$$

donde d_{AB} representa la distancia entre dos pares de *tweet* que pertenecen a grupos distintos, siendo n el indicador del número de pares; d_X representa un único par perteneciente a un mismo grupo, ya sea A o B, con n_X indicando el número de pares. De acuerdo a la definición plantada, mientras más alto es, más grande es la distancia entre grupos es comparada a la de ambas distancia de grupo, por lo que se puede concluir que la información es más diversa, y siendo al revés en el caso contrario.

Capítulo 4

Resultados y Análisis

En este capítulo se presentan los resultados obtenidos en este trabajo, que incluye las visualizaciones de los grafos de *retweets*, los puntajes de nivel de controversia explicados anteriormente y los gráficos del mejor de ambos puntajes con respecto al tiempo de duración del evento, como también del *normalized cut* con respecto al tiempo igualmente y, finalmente, gráficos de esta misma métrica, pero las curvas en este caso se encuentran separadas de acuerdo a los sentimientos positivos, negativos y neutrales de los *tweet*.

4.1. Temas a Estudiar

Para el propósito de estas memoria, se decidió por escoger seis temas en particular, los cuales se describen a continuación en la Tabla 4.1 junto a las características de sus datos, respectivamente:

Tabla 4.1: Temas de observación donde los primeros tres se consideran de gran controversia: el cuarto, que puede posiblemente serlo; y los últimos dos, que no lo son.

Descripción del Evento	Fechas	Nº de Tweets	Nodos	Aristas
Proceso de campaña del apruebo y rechazo por el plebiscito durante la pandemia	01/05/2020 - 31/10/2020	332891	89523	316841
Homicidio del comunero mapuche Camilo Catrillanca a manos del comando jungla	13/11/2018 - 30/05/2019	103674	9809	166949
Legalización del aborto en tres causales en 2017	01/07/2017 - 31/10/2020	88889	5271	25097
Torneo Primera División en Chile del año 2021	01/08/2021 - 31/12/2022	60101	19002	37752
Incendios forestales ocurridos en el verano del 2017 en la zona centro y sur de Chile	01/12/2016 - 30/04/2017	124035	45482	99453
Programa de televisión Pasapalabra en sus meses de mejor rating	01/08/2018 - 30/11/2018	74347	40577	72721

4.2. Grafos de Reweets

En vista de los datos obtenidos es que se construyen a partir de estos las siguientes visualizaciones que se pueden apreciar en las Figuras 4.1 y 4.2 realizadas en el software Gephi con el algoritmo ForceAtlas2:

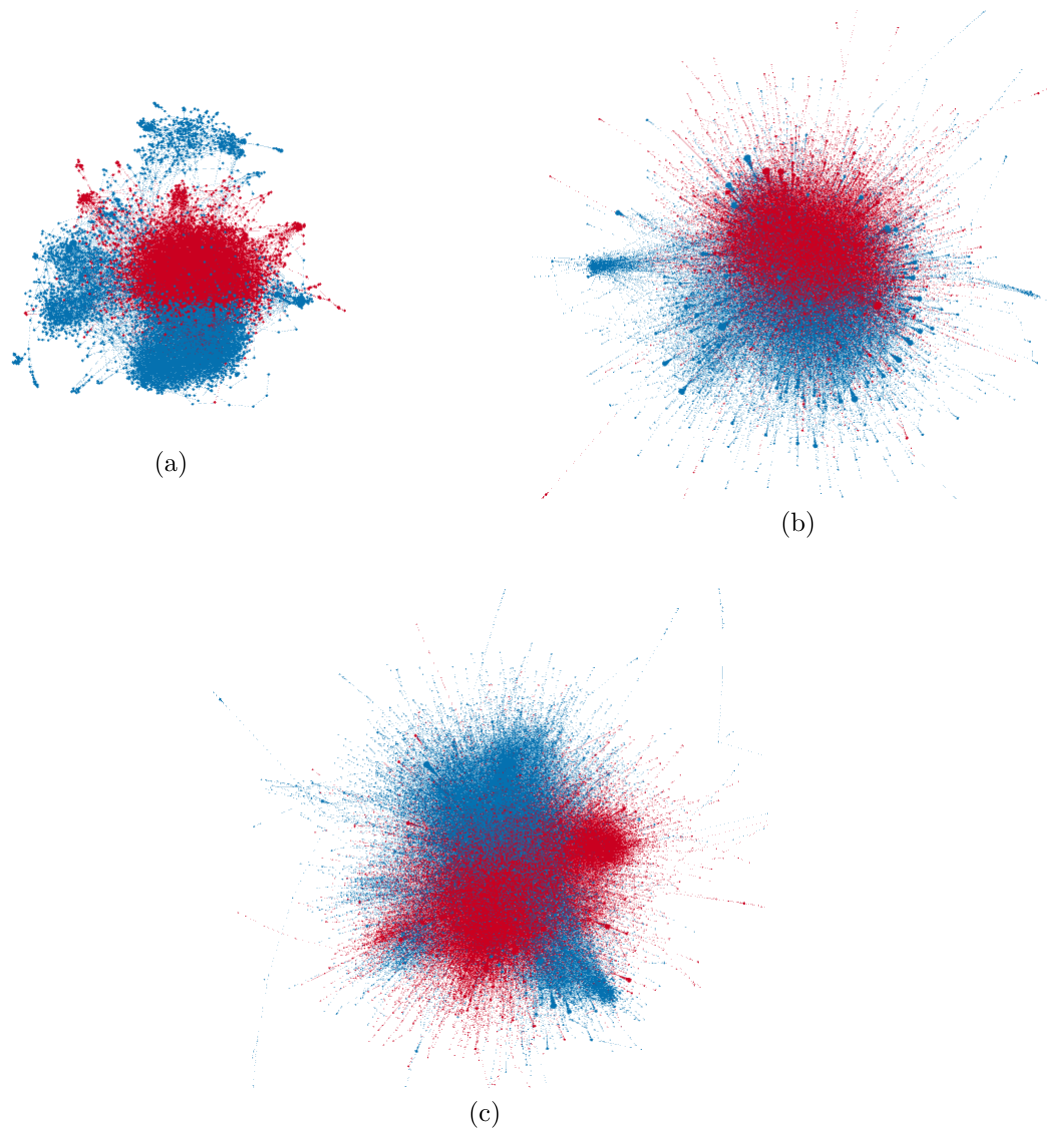


Figura 4.1: Grafos de *retweets* de los siguientes temas: a) Torneo primera división en Chile del año 2021, b) Programa televisivo Pasapalabra en sus meses de mayor rating, c) Incendios forestales ocurridos en el verano del 2017.

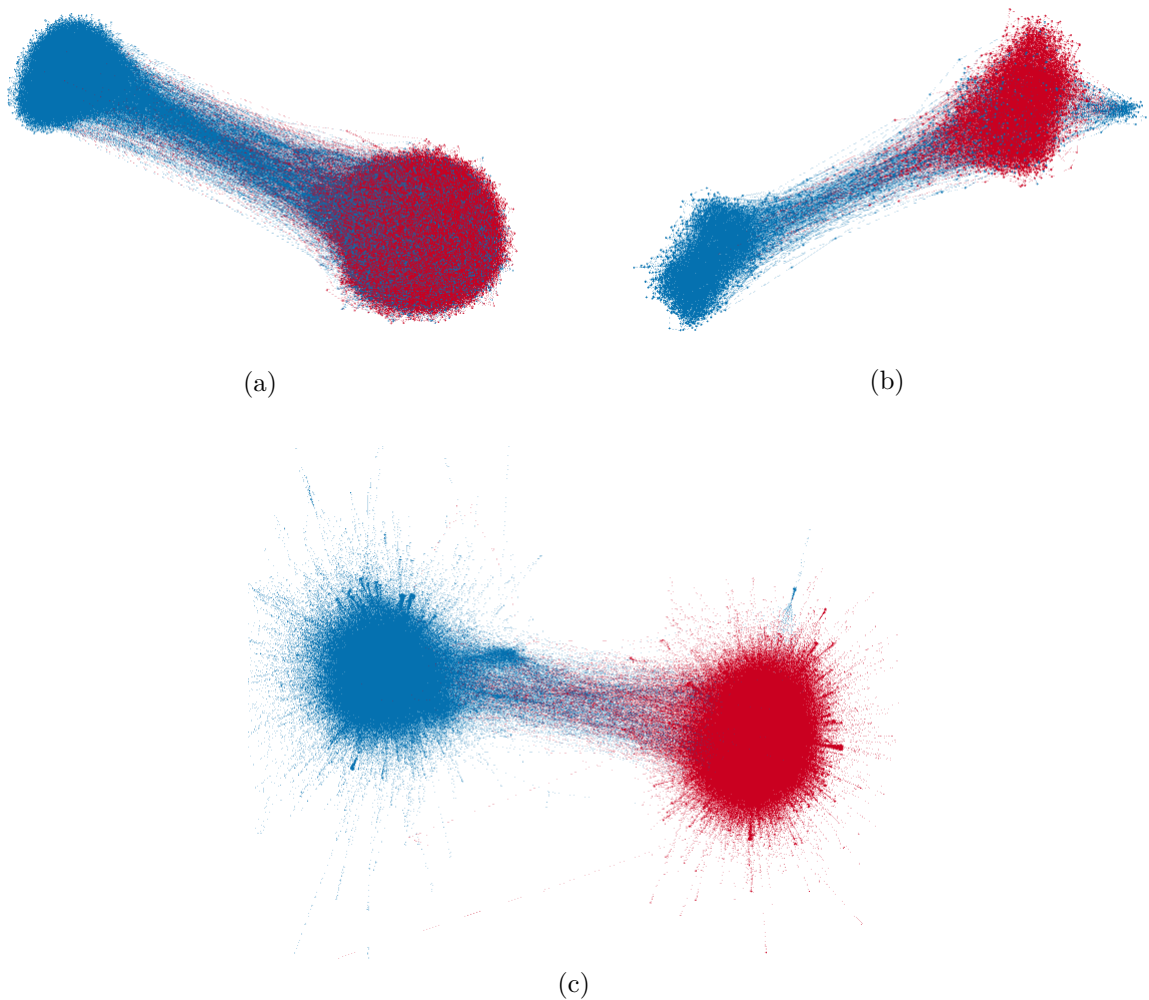


Figura 4.2: Grafos de *retweets* de los siguientes temas: a) Primeros meses luego del homicidio del comunero mapuche, Camilo Catrillanca b) Proceso legislativo de las tres causales del aborto, c) Campaña del apruebo y rechazo antes del plebiscito del año 2020.

Tal como es posible observar en las redes de *retweets* que se observan en las Figuras 4.1 y 4.2, es con fácil inspección notar la clara diferencia entre los temas no controversiales y los que no lo son. Sin embargo, existen ciertos aspectos a considerar con respecto a cada uno de los temas. Por ejemplo, para el tema del torneo de primera división existen cierto nivel de polarización, dado que es posible notar las particiones, a diferencia de los otros dos temas, donde existe una homogeneidad entre ambos. Con respecto a los temas controversiales, existe un caso parecido para el homicidio de Camilo Catrillanca, ya que se puede ver que una cantidad no mínima de usuarios pertenecientes a una partición interactúan con la otra, hecho que no ocurre para los temas de la legalización de las tres causales y la campaña del apruebo y rechazo.

4.3. Métricas de cada Tema

En esta sección se muestran los cálculos obtenidos de las métricas escogidas, *betweenness* y *random walk*, aplicadas a cada uno de los temas pertinentes, a través de las Tablas 4.2 y 4.3.

Tabla 4.2: Puntajes que cuantifican el nivel de controversia de temas controversiales.

Temas Controversiales	Betweenness	Random Walk
Campaña Apruebo/Rechazo	0.92	0.48
Catrillanca	0.30	0.39
Aborto	0.88	0.62

Tabla 4.3: Puntajes que cuantifican el nivel de controversia de temas controversiales.

Temas Controversiales	Betweenness	Random Walk
Campaña Apruebo/Rechazo	0.92	0.48
Catrillanca	0.30	0.39
Aborto	0.88	0.62

De acuerdo a las Tablas 4.2 y 4.3, se puede ver el nivel de variación entre cada tema y métrica. En primera instancia, es evidenciable como se distingue el nivel de controversia entre los dos tipos de tópicos al observar el valor tan bajo de los temas no polarizantes y viceversa. No obstante, existe una clara diferencia entre ambas métricas, lo que es fácil de notar al ver el tema del homicidio de Camilo Catrillanca, donde el puntaje de *betweenness* es más bajo que incluso que los temas no controversiales, dado que al observar el grafo y el número de nodos, comparándose con el de aristas, se debe a la gran cantidad de interconexiones que tienen los usuarios entre sí, por lo que siendo una medida que se basa en encontrar el camino más corto de un nodo a otro, es altamente factible que en la mayoría de los caminos existentes, se deba pasar por menos de dos vértices. Es por esta razón que se decidió considerar el *random walk* como la única métrica a utilizar para un análisis temporal, debido a su clara robustez, esto debido a que no requiere de una cantidad tan grande de nodos y aristas para que funcione adecuadamente.

4.4. Random Walk vs Tiempo del Evento

En esta sección se procede a mostrar los gráficos de la métrica de *random walk* calculada a través de distintos grafos para cada semana perteneciente al periodo del evento.

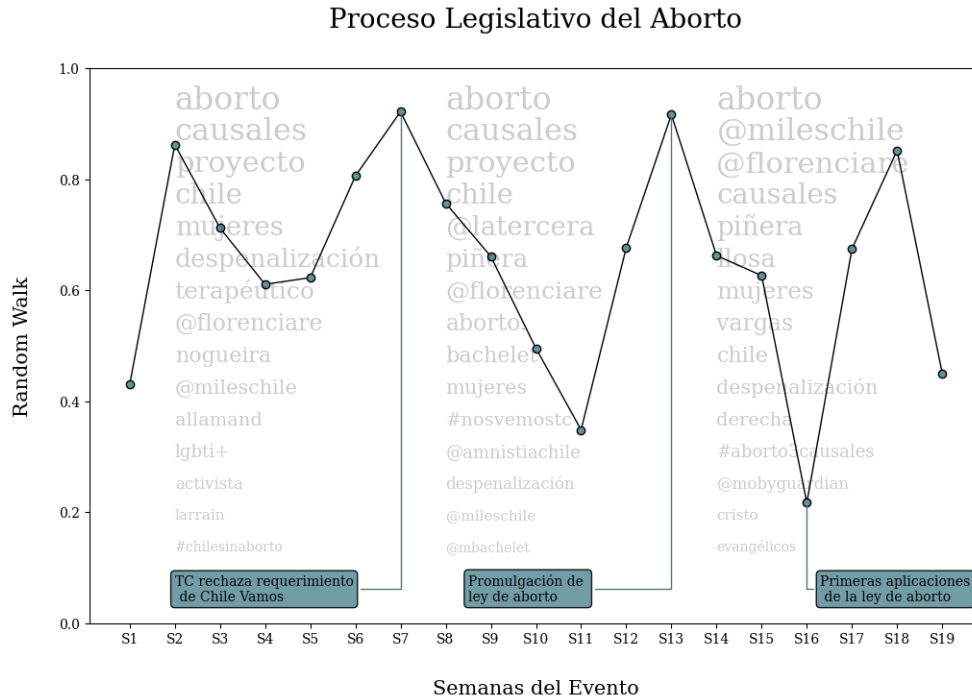


Figura 4.3: Polarización de la red a través del tiempo, donde la línea representa el valor de random walk de cada semana dentro del periodo establecido, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Caso Catrillanca

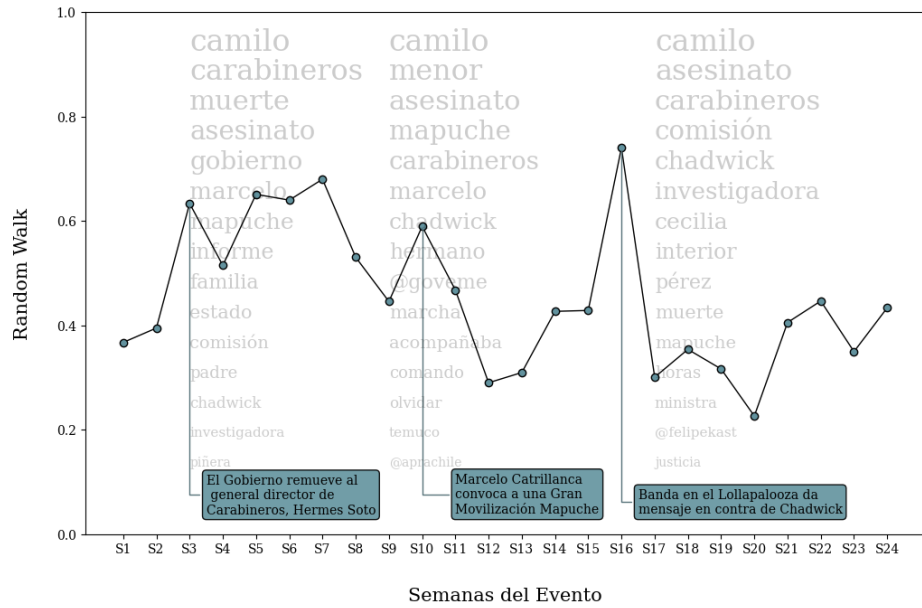


Figura 4.4: Polarización de la red a través del tiempo, donde la línea representa el valor de *random walk* de cada semana dentro del periodo establecido, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Campaña Apruebo vs Rechazo

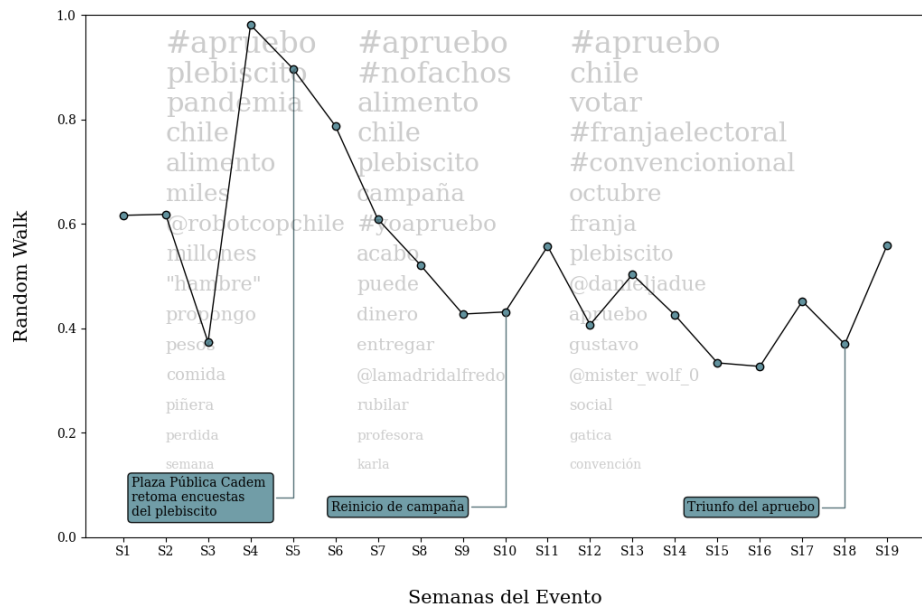


Figura 4.5: Polarización de la red a través del tiempo, donde la línea representa el valor de *random walk* de cada semana dentro del periodo establecido, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Torneo Primera División 2021

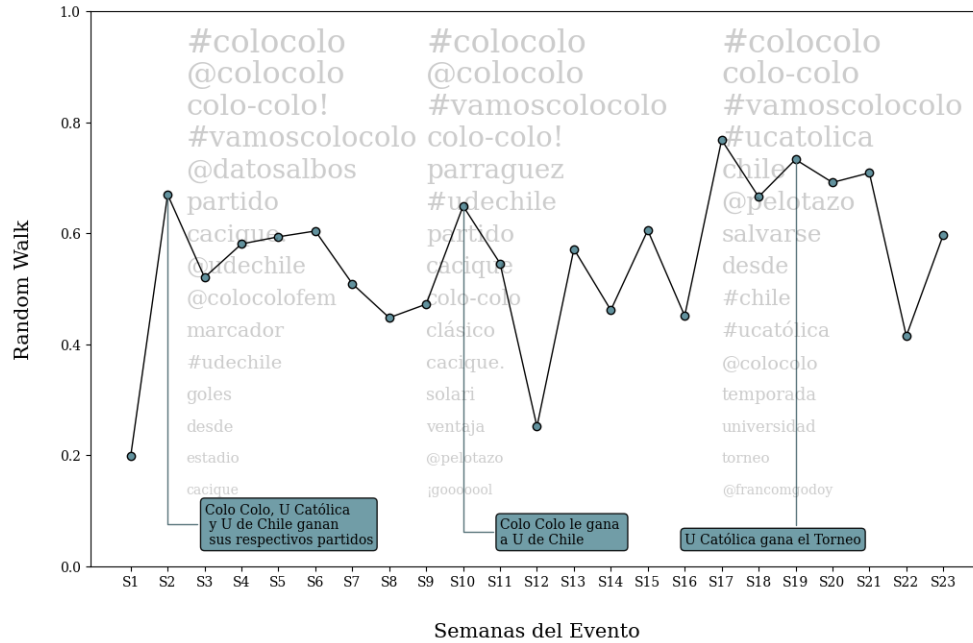


Figura 4.6: Polarización de la red a través del tiempo, donde la línea representa el valor de *random walk* de cada semana dentro del periodo establecido, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Dentro de lo que se puede observar en los gráficos presentes en las Figuras 4.3 a 4.6, es la gran variabilidad de esta métrica alrededor de cada semana para cada uno de los temas, aspecto que se destaca bastante en la Figura 4.3, al notarse los puntos de mayor controversia, como el evento en que el Tribunal Constitucional rechaza el requerimiento de Chile Vamos y evita el frenado del proyecto y cuando, consecutivamente, se promulga la ley en el Diario Oficial. A diferencia de los puntos de menor controversia, como en la primeras aplicaciones de aborto en el transcurso de esas semanas, que se considera un valor bastante bajo, debido a que este caso se trata acerca de la interrupción del embarazo de una menor de edad producto de una violación, hecho que cae dentro de la hipótesis de que la controversia es raro que exista en temas de carácter trágico, como se pudo observar en los puntajes obtenidos en el tema de los incendios forestales del 2017 y las investigaciones de Garimella et al. [7]. Este caso puede ser aplicable también al del homicidio de Camilo Catrillanca y el torneo de primera división que, a pesar de no considerarse finalmente un tema controversial, igualmente tiene momentos en que ambos grupos no discuten entre sí, lo que, al revisar por simple inspección de los *tweet* de cada grupo, se debe a como cada usuario participante realiza *tweet* apoyando a su equipo, más que denostar a quienes celebran los logros del otro. Sin embargo, es importante notar la gran diferencia que tiene la Figura 4.5 con las demás, al ser un tema que empieza con una gran controversia en cuanto a su discusión en *Twitter*. No obstante, a medida que van pasando las semanas, existe una tendencia en la disminución de la controversia con unos leves aumentos, esto debido a que, al observar los datos, existía una mayor interacción entre los grupos a medida que se llega al triunfo del apruebo en las últimas semanas.

4.5. Normalized Cut vs Tiempo del Evento

A continuación se enseñan los gráficos obtenidos, en las Figuras 4.7 a 4.10, a partir de la métrica de *normalized cut* calculada para cada semana dentro un tema tomando en cuenta el periodo del evento.

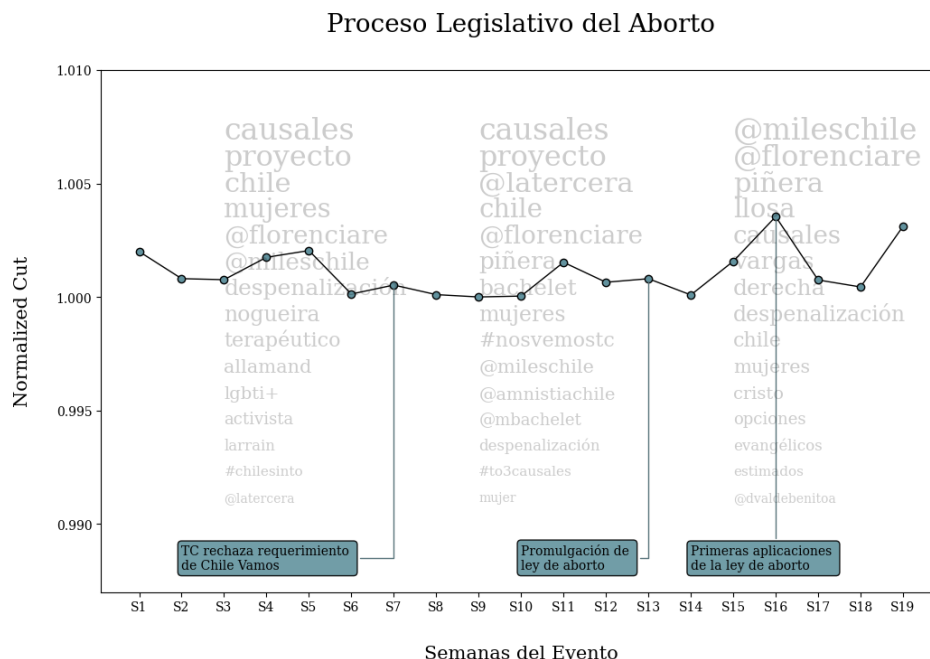


Figura 4.7: Polarización del contenido a través del tiempo, donde la línea representa el valor de *normalized cut* de cada semana dentro del periodo establecido, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Caso Catrillanca

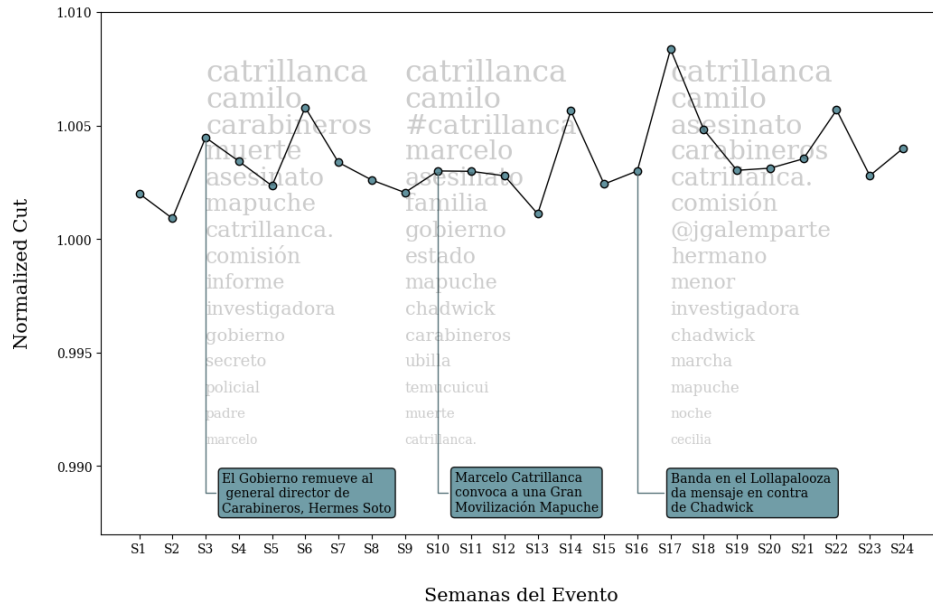


Figura 4.8: Polarización del contenido a través del tiempo, donde la línea representa el valor de *normalized cut* de cada semana dentro del periodo establecido, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Campaña Apruebo vs Rechazo

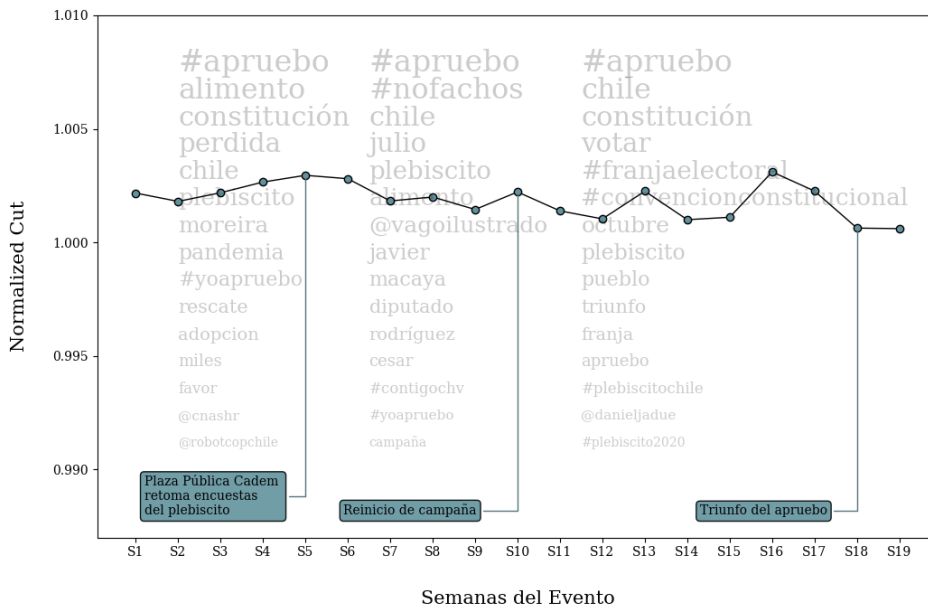


Figura 4.9: Polarización del contenido a través del tiempo, donde la línea representa el valor de *normalized cut* de cada semana dentro del periodo establecido, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Torneo Primera División 2021

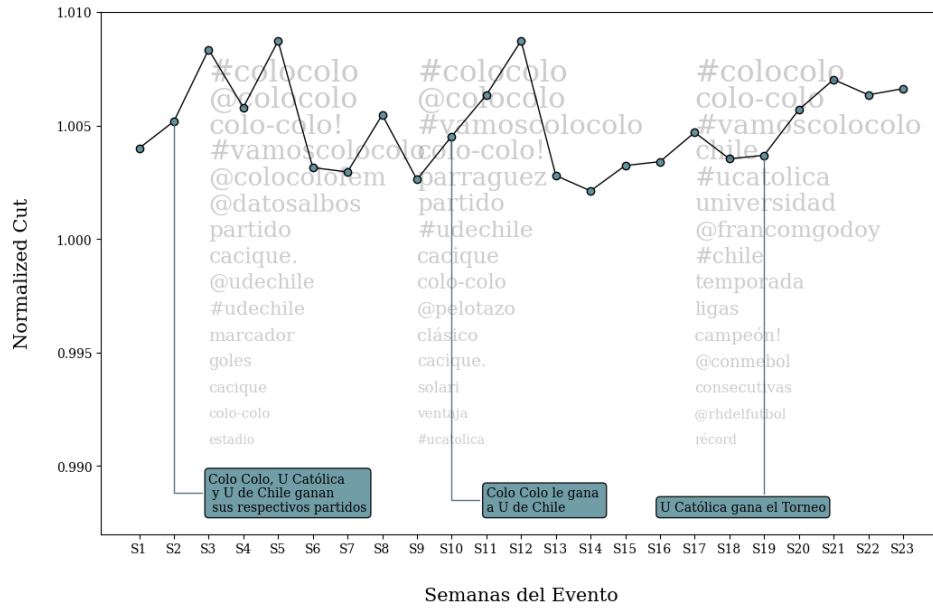


Figura 4.10: Polarización del contenido a través del tiempo, donde la línea representa el valor de *normalized cut* de cada semana dentro del periodo establecido, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Es claro notar la distinción entre la métrica discutida anteriormente y esta, considerando que las cumbres entre ambos gráficos no coinciden con las mismas semanas, como también ocurre que con los gráficos del *normalized cut*, la diferencia entre los máximos y mínimos no es tanta a como los del *random walk*. Esto se debe a que las métricas basadas en procesamiento de lenguaje natural, tal como se explica en Garimella et al. [7], difícilmente pueden utilizarse para cuantificar controversia, sobre todo si se considera que esta métrica se basa en la intuición de que quienes pertenecen a un mismo grupo ideológico expresarán sus opiniones y discutirán de una manera similar entre sí. Sin embargo, se pueden encontrar ciertas coincidencias con los gráficos vistos de *random walk*, como en el caso de la Figura 4.5, donde se pueden observar similitudes con la Figura 4.9, como los puntos máximos de la controversia que coinciden con eventos relevantes que tienden a aumentar ésta; además de un comportamiento de disminución de la polarización a medida que van pasando las semanas.

4.6. Normalized Cut de los Sentimientos vs Tiempo del Evento

Para esta sección se muestran los gráficos obtenidos, en las Figuras 4.11 a 4.14, a partir de la métrica de *normalized cut* calculada, separada de acuerdo al sentimiento de los *tweet*, para cada semana dentro un tema, por el periodo del evento.

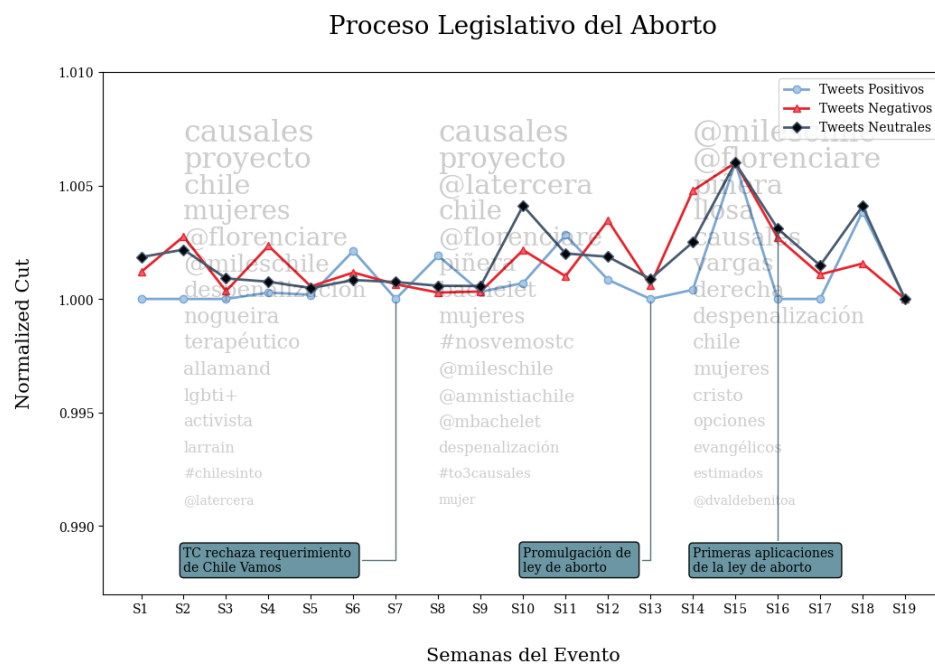


Figura 4.11: Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de *normalized cut* de un sentimiento específico, para la legislación de las tres causales del aborto. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Caso Catrillanca

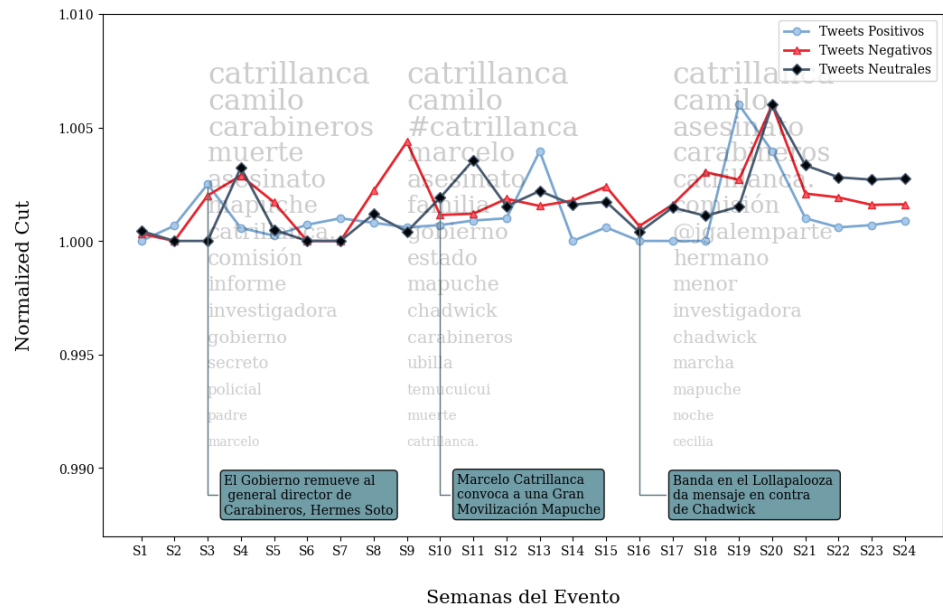


Figura 4.12: Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de *normalized cut* de un sentimiento específico, para el caso del homicidio de Camilo Catrillanca. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Campaña Apruebo vs Rechazo

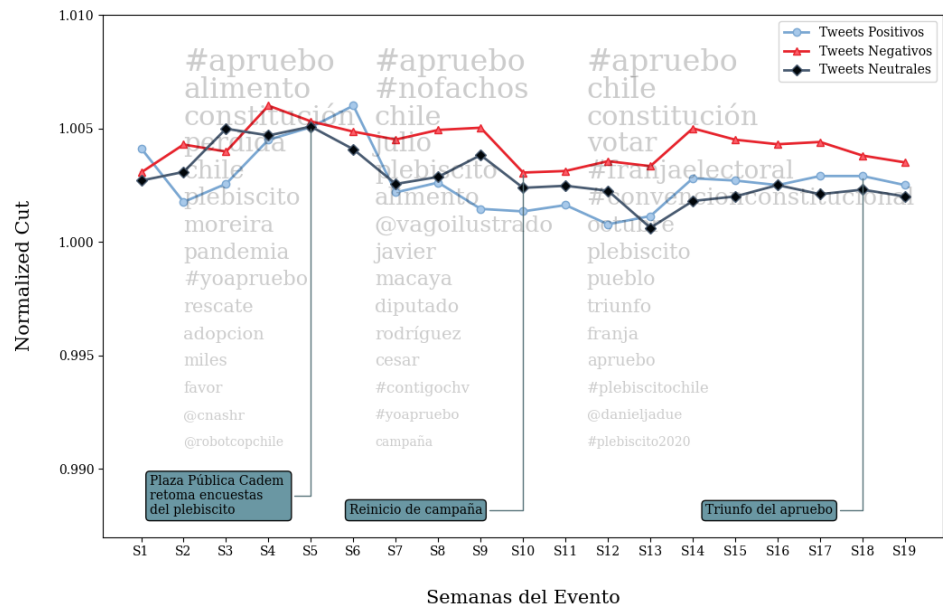


Figura 4.13: Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de *normalized cut* de un sentimiento específico, para la campaña de los últimos meses del apruebo y del rechazo. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Torneo Primera División 2021

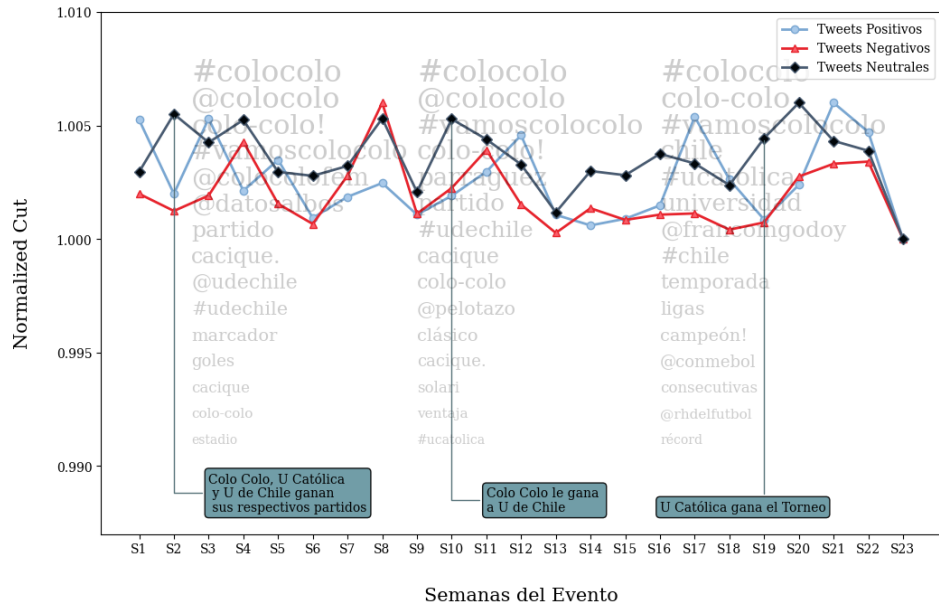


Figura 4.14: Polarización del contenido a través del tiempo por sentimiento, donde cada línea representa el valor de *normalized cut* de un sentimiento específico, para el torneo de primera división del 2021. Se enseñan las palabras, hashtags y usuarios más utilizados dentro de cada cierto rango de semanas.

Realizando una inspección a los gráficos, se puede apreciar una gran variabilidad de estos en todos los temas evaluados. De acuerdo a esto, se debe notar como los *tweets* que son discutidos, o más bien, *retweeteados* por otros usuarios son más de carácter positivo, negativo o neutral. Por ejemplo, en la Figura 4.13 se puede apreciar expresamente como en gran parte del periodo del tema existen interacciones principalmente de manera negativa, es decir, la polarización dentro y fuera de un grupo se manifiesta con sentimientos netamente negativos. No obstante, existen otros temas en que la discusión se exterioriza tanto positiva como negativamente sin importar necesariamente el evento importante que ocurra en aquella semana, como es el caso de las Figuras 4.11 y 4.12, donde se pueden ver como las curvas coinciden en cuanto a su comportamiento.

Capítulo 5

Conclusiones

Para este trabajo se planteó el problema de cómo en las redes sociales se han formado con el tiempo especies de cámaras de eco que han incentivado a que los usuarios en estas plataformas no compartan sus opiniones e ideas con quiénes puedan tener pensamientos distintos o totalmente contrarios a ellos. De manera que, para tener una noción clara que temas puedan ser más controversiales que otros, se propone la medición de este fenómeno, con el fin de analizar y estudiar el nivel con el que tópicos de carácter social, político y religioso se discuten, particularmente en la red social, *Twitter*. Para ello, se obtienen los datos necesarios a través de la API, la cual, nos permite descargar una gran cantidad de *tweet* con toda la información requerida, como el nombre de usuario, su número de ID, el texto del tweet, el usuario que retweeteó este tweet, entre otras características. Por lo que, haciendo uso de estos datos, es que se lograron obtener los resultados anteriormente mostrados, como los grafos de nodos que permitieron visualizar la polarización existente entre dos grupos de un tema en particular, y a través de éstos, es que se pudieron también calcular las métricas de *random walk* y *betweenness* para evidenciar matemáticamente como los usuarios tienden a interactuar más con quienes piensan de una forma similar a ellos que con quienes no lo hacen. Además, con la finalidad de realizar un trabajo más extenso, se calculó también una métrica que utiliza procesamiento de lenguaje natural para medir la diferencia entre el vocabulario que emplean usuarios dentro de un mismo y distintos grupos.

Todos esto conduce a inspeccionar la validez final de los objetivos que fueron planteados al principio del trabajo, donde es posible mencionar que todos fueron logrados e implementados, desde la obtención de los términos necesarios para realizar una búsqueda exhaustiva de tanto usuarios como contenidos de sus respectivos *tweets*, así como también la construcción de grafos de nodos y gráficos que permitan explicar de manera visual como se comportan los grupos en cada tema considerado y la variabilidad de la controversia semana a semana, respectivamente. Sumado a lo anterior, se sostiene que durante el transcurso de este trabajo se logró un análisis importante de esta problemática.

Dentro de los aprendizajes obtenidos en este trabajo, se pueden destacar una gran variedad de estos, que permitieron construir experiencia que pueda ser útil para futuros proyectos. La principal fue la maduración en el manejo de datos, dado que se logró aprender una diversidad de funciones y métodos que permitieron agilizar el tiempo de codificación, además de la optimización en la ejecución de las *pipelines*. Otro importante aspecto fueron los grafos, tanto su construcción a través de librerías especializadas en aquella tarea y su visualización

de manera interactiva.

Un trabajo futuro podría ser realizar no solo grafos a partir de *retweets*, sino que se podrían aplicar en base a seguidores y *likes*, lo que le daría otro valor al análisis de esta memoria, dado que la metodología original solamente utiliza usuarios más activos dentro de la red social considerando que ellos mismos generan su propio contenido al escribir sus propios *tweets* y compartir los de otros. Sin embargo, no se toman en cuenta los usuarios pasivos, es decir, quienes forman parte de un grupo dentro de la discusión de un tema controversial y no profesan su opinión de manera libre, pero solo siguen o dan su apoyo de manera anónima a quienes piensan de manera similar a ellos. Otro alcance que puede tener este trabajo, es el uso de otras plataformas para obtener datos, como *Facebook*, *Instagram*, o blogs que se centren en temas relevantes, puesto que el hecho de utilizar solo *Twitter* resulta limitante en cuanto al estudio si es que se desea llegar a individuos que pertenezcan a una diversidad de contextos, ya sea a nivel etario, socioeconómico o localización. Finalmente, parece de carácter relevante tomar en cuenta que no todos los temas a tratar tendrán únicamente dos grupos contrarios, sino que es innegable la existencia de otros bandos que puedan tener una ideología o pensamientos más moderados que varíen entre ambos extremos.

Bibliografía

- [1] An J., Quercia D., Crowcroft J. 2014. Partisan sharing: Facebook evidence and societal consequences. In Proceedings of the second ACM conference on Online social networks (pp. 13-24).
- [2] Flaxman S., Goel S., Rao J. M. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1), 298-320.
- [3] Grevet C., Terveen L. G., Gilbert E. 2014. Managing political differences in social media. In Proceedings of the 17th ACM conference on Computer supported cooperative work social computing (pp. 1400-1408).
- [4] Adamic L. A., Glance N. 2005. The political blogosphere and the 2004 U.S. election: divided they blog. In Proceedings of the 3rd international workshop on Link discovery (LinkKDD '05). Association for Computing Machinery, New York, NY, USA, 36-43.
- [5] Conover M., Ratkiewicz J., Francisco M., Gonçalves B., Menczer F., Flammini A. 2011. Political polarization on *Twitter*. In Proceedings of the international aaai conference on web and social media (Vol. 5, No. 1, pp. 89-96).
- [6] Mejova Y., Zhang A. X., Diakopoulos N., Castillo C. 2014. Controversy and sentiment in online news. arXiv preprint arXiv:1409.8152.
- [7] Garimella K., Morales G. D. F., Gionis A., Mathioudakis M. 2018. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1(1), 1-27.
- [8] Vukičević D., Škrekovski R., Tepoh A. 2016. Relative edge betweenness centrality. *Ars Mathematica Contemporanea*, 12(2), 261-270.
- [9] Yang M., Wen X., Lin Y. R., Deng L. 2017. Quantifying content polarization on *Twitter*. In 2017 IEEE 3rd international conference on collaboration and internet computing (CIC) (pp. 299-308). IEEE.
- [10] Lovász L. 1996. Random walks on graphs. *Combinatorics, Paul erdos is eighty*, 2(1-46), 4.
- [11] Mejova Y. 2009. Sentiment analysis: An overview. University of Iowa, Computer Science Department.
- [12] Aho A., V., Lam M. S., Sethi R., Ullman J. D. 2007. *Compilers: principles, techniques, tools*. Pearson Education India.
- [13] Bergmanis T., Goldwater S. 2018. Context Sensitive Neural Lemmatization with Lematus. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 1391-1400, New Orleans, Louisiana. Association for Computational Linguistics.

- [14] Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean, J. 2013. Efficient Estimation of Word Representations in Vector Space. ICLR (Workshop Poster).
- [15] Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean, J. 2013. Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems, 26.
- [16] Balakrishnan R., Ranganathan K. 2012. A textbook of graph theory. Springer Science Business Media.
- [17] Coletto M., Garimella K., Gionis A., Lucchese C. 2017. A motif-based approach for identifying controversy. In Eleventh International AAAI Conference on Web and Social Media.
- [18] Guerra P., Meira Jr W., Cardie C., Kleinberg R. 2013. A measure of polarization on social media networks based on community boundaries. In Proceedings of the international AAAI conference on web and social media (Vol. 7, No. 1, pp. 215-224).
- [19] Morales A. J., Borondo J., Losada J. C., Benito R. M. 2015. Measuring political polarization: *Twitter* shows the two sides of Venezuela. Chaos: An Interdisciplinary Journal of Nonlinear Science, 25(3), 033114.
- [20] Graells-Garrido E., Baeza-Yates R., Lalmas M. 2020. Every colour you are: Stance prediction and turnaround in controversial issues. In 12th ACM Conference on Web Science (pp. 174-183).