

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
SANTIAGO - CHILE



“ANÁLISIS SOBRE LA CLASIFICACIÓN DE TWEETS CON INFORMACIÓN IMPRECISA RELATIVA A CHILE”

RENATO ANDRÉS BASSI CASTILLO

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA

Profesor Guía: Marcelo Mendoza Rocha
Profesor Correferente: José Luis Martí Lara

Diciembre - 2023

DEDICATORIA

Dedico esta memoria a mi familia y amigos, como un reflejo de resiliencia, determinación y las innumerables lecciones aprendidas. Este logro es la culminación de las decisiones que tomé, los obstáculos que superé y el crecimiento personal que experimenté.

AGRADECIMIENTOS

Quiero agradecer a todos los periodistas, profesores y estudiantes que participaron del *PLU210012*, con especial énfasis en los periodistas relativos a los grupos de *fact-checking* como *Fast Check CL*, *Decodificador* y *Fact Checking UC*, dado a que ellos fueron los que proporcionaron el *dataset* que hizo posible esta memoria. Así como también, agradecer al *DI* y a la *UTFSM* por proporcionar credenciales para el uso de la *API* académica de *Twitter* y de *CrowdTangle*.

RESUMEN

Resumen— Análisis de *tweets* con contenido impreciso relativos a Chile, a partir del estudio del contenido de los *tweets* y su propagación, obteniendo la caracterización de los *tweets* imprecisos y su relación con fenómenos de sesgo cognitivo, para determinar la influencia de estos en la polarización de las comunidades en *Twitter* Chile.

Palabras Clave— *Tweet*, Impreciso, Sesgo, Polarización

ABSTRACT

Abstract— Analysis of *tweets* with imprecise content related to Chile, based on the study of the content of the *tweets* and their propagation, obtaining the characterization of the imprecise *tweets* and their relationship with cognitive bias phenomena, to determine the influence of these on the polarization of the communities on *Twitter* Chile.

Keywords— *Tweet*, Inaccurate, Bias, Polarization

GLOSARIO

DI: Departamento de Informática.
UTFSM: Universidad Técnica Federico Santa María.
UC: Universidad Católica.
API: Application Programming Interface.
ANOVA: Analysis of Variance.
URL: Uniform Resource Locator.
COVID-19: Coronavirus Disease 2019.
GFOG: Gunning's Fog Index.
NDCRS: New Dale-Chall Readability Score.
ARI: Automated Readability Index.
FKI: Flesch-Kincaid Index.
DCRS: Dale-Chall Readability Score.
CLI: Coleman-Liau Index
LWRI: Linsear-Write Reading Ease Index.
FRE: Flesch Reading Ease Index.
IFSZ: Index of Flesch-Szigriszt.
FHRI: Fernández-Huerta Readability Index.

ÍNDICE DE CONTENIDOS

RESUMEN	IV
ABSTRACT	IV
GLOSARIO	V
ÍNDICE DE FIGURAS	VIII
ÍNDICE DE TABLAS	VIII
INTRODUCCIÓN	1
CAPÍTULO 1: DEFINICIÓN DEL PROBLEMA	2
1.1 Objetivo general	3
1.2 Objetivos específicos	3
1.3 Impacto inicial de solucionar el problema	3
CAPÍTULO 2: MARCO CONCEPTUAL	5
2.1 <i>Tweet, Twitter</i>	5
2.2 Clasificación no binaria	6
2.3 Contenido a estudiar de los <i>tweets</i>	7
2.4 Test ANOVA	8
2.5 <i>Fact-checking</i>	8
2.6 <i>Fact-checkers</i>	9
2.7 Certificaciones	10
2.8 <i>Fact-checkers</i> chilenos	10
2.9 <i>Dataset</i>	12
2.9.1 Tópicos	13
2.10 Algoritmo de recomendación	15
CAPÍTULO 3: METODOLOGÍA	17
3.1 Contenido y propagación	17
3.2 Reclasificación	18
3.3 Algoritmo, sesgo y polarización	19
CAPÍTULO 4: ANÁLISIS	20
4.1 <i>Tweets</i> por tópico	20
4.2 Características lingüísticas	22
4.3 Propagación de los <i>tweets</i>	23
4.4 Reclasificación	24
4.5 Algoritmo de recomendación	26
4.6 Sesgo de confirmación	27

4.7 Polarización política en Chile	28
CAPÍTULO 5: CONCLUSIONES	32
REFERENCIAS BIBLIOGRÁFICAS	36
6 TIEMPOS SCT	40

ÍNDICE DE FIGURAS

1	Árbol del problema.	4
2	Diagrama del algoritmo de recomendación de <i>Twitter</i>	16

ÍNDICE DE TABLAS

1	Unidades de contenido verificadas por las agencias de fact-checking.	13
2	Conversión de etiquetas de veracidad.	14
3	Categorías de veracidad por agencia de <i>fact-checking</i>	15
4	Categorías de veracidad por tema.	15
5	Categorías de veracidad por tema en <i>Twitter</i>	20
6	Categorías de veracidad por tema en <i>Twitter</i> con porcentajes.	21
7	Funciones lingüísticas para unidades de contenido verificadas en <i>Twitter</i> por categoría de veracidad.	22
8	Funciones lingüísticas para unidades de contenido verificadas en <i>Twitter</i> por dimensión de tema.	23
9	Profundidad de los árboles de propagación en <i>Twitter</i> (respuestas).	23
10	Tamaño de los árboles de propagación en <i>Twitter</i> (respuestas).	24
11	Amplitud de los árboles de propagación en <i>Twitter</i> (respuestas).	24
12	Cantidad de diputados obtenida por partido político en las elecciones parlamentarias, desde 2001 a 2021.	31
13	Cantidad de diputados obtenida por bloque político en las elecciones parlamentarias, desde 2001 a 2021.	31

INTRODUCCIÓN

Este trabajo nace de la necesidad de tener un estudio de los tweets clasificados como imprecisos en el ámbito de temas relativos a Chile. Para esto se estudió el trabajo realizado por el PLU210012 [Mendoza *et al.*, 2023], donde se armó un dataset recolectado por distintos grupos de fact-checking relativos al proyecto, los cuales recolectaron tweets verificados de temas como el estallido social, el COVID-19, las elecciones de 2021 y el proceso constituyente de 2020-2022.

Con esto se procede al estudio del dataset, en cuanto a la proporción de tweets dadas su clasificación de veracidad en los temas anteriormente nombrados relativos a Chile, así como también con respecto al contenido lingüístico, como por ejemplo cantidad de verbos y sustantivos. Además, se ve la propagación de los tweets dependiendo de su clasificación de veracidad, a partir de la profundidad, tamaño y amplitud de los arboles de propagación de los tweets.

Seguidamente, se estudian los fenómenos cognitivos que pueden estar relacionados con la caracterización de los tweets imprecisos, como lo son el sesgo de confirmación y el efecto Dunning-Kruger.

Finalmente, se detalla y explica como los tweets imprecisos son propensos a aparecer y proliferar más en un contexto de polarización, donde los grupos burbujas producidos artificialmente por el algoritmo de recomendación de Twitter tienden a acentuar esta situación.

CAPÍTULO 1

DEFINICIÓN DEL PROBLEMA

El crecimiento de la desconfianza de la gente a los medios masivos de comunicación tradicionales ha generado que las personas tiendan a usar a las redes sociales para informarse sobre la actualidad [Institute, 2020].

Dentro de las redes sociales más usadas a nivel mundial, tenemos a *Twitter*, la cual nace en el año 2006 y desde sus inicios hasta la actualidad, se ha ocupado como una red social para publicar, compartir y opinar sobre distintas cosas, como pueden ser noticias [Campos *et al.*, 2017].

El crecimiento de las redes sociales no necesariamente ha significado el fin del periodismo tradicional, ya que se han presentado estas como nuevas plataformas para subir noticias y presentarlas tal cual se aplicaban con los medios tradicionales. Para el caso de *Twitter*, lo más parecido en la mayoría de los casos de las publicaciones sería “el diario”, ya que al ser todo en su mayoría texto, por lo general se publica información como si fuera una noticia de ese medio. También, *Twitter* puede ser usada para publicar el título de la noticia solamente, y en el cuerpo de la publicación *linkear* a una página web de la plataforma de noticias de alguna radio por ejemplo.

Pero no todo es bueno para *Twitter*, ya que según estudios internacionales, ha habido un incremento de la polarización [Garimella y Weber, 2017], sobre todo en periodos de elecciones, donde han proliferado noticias con contenido que no es completamente cierto, o que la mitad de la noticia es cierta, o que directamente el contenido es totalmente falso.

Es por el anterior punto que han surgido grupos que se dedican a revisar la veracidad de las publicaciones compartidas en *Twitter*. Estos grupos, por lo general compuesto por periodista, se los conoce como *fact-checkers*. Dentro de lo local, tenemos a *Fast Check CL*, *Fact Cheking UC* y *Decodificador*.

Estos *fact-checkers* han recolectado muchos datos a lo largo de los años, donde han recolectado publicaciones de distintas redes sociales, para lo cual las publicaciones se han agrupado en distintos tópicos, de los cuales destacan el estallido social, el COVID-19, el proceso constituyente de 2020-2022 y las elecciones de 2021. Estas fueron reclasificadas, dado que cada grupo de *fact-checking* tiene su propia clasificación. Para el caso de *Twitter*, se dejaron tres categorías de veracidad, las cuales eran: real, falso e impreciso. Hasta la fecha no existe un estudio de análisis sobre la clasificación de *tweets* con información imprecisa relativa a Chile, por lo que en esta memoria el autor se encargará de entregar información nueva y relevante sobre los *tweets* catalogados como imprecisos.

1.1. Objetivo general

Analizar los *tweets* clasificados como imprecisos y estudiar su impacto en los fenómenos políticos y sociales de Chile.

1.2. Objetivos específicos

1. Analizar el *dataset* de *tweets* clasificados.
2. Evaluar si la reclasificación no afectó la caracterización de los *tweets* imprecisos.
3. Investigar la evolución política y social de Chile desde el año 2000.
4. Comparar los resultados obtenidos del proyecto con fenómenos relativos al sesgo de confirmación y la polarización.

1.3. Impacto inicial de solucionar el problema

Al resolver el problema, el primer impacto inmediato sería que se podrá saber, la influencia de los *tweets* catalogados como imprecisos en la polarización de las comunidades en *Twitter* Chile.

A continuación, se mostrará el árbol del problema, para poder entender mejor las causas y efectos de lo que se está trabajando.



Figura 1: Árbol del problema.
Fuente: Elaboración propia.

CAPÍTULO 2

MARCO CONCEPTUAL

A modo de apoyar el trabajo, se definirán los siguientes conceptos relativos al presente trabajo, como pilares para la explicación de la construcción del análisis.

Primeramente, se describirá el objeto a estudiar, que en este caso es el *tweet*, explicando la composición de este y el por qué se está estudiando. Luego se explicará que es el *fact-checking* en redes sociales, así como también a los conceptos que estén relacionados a este, como los *fact-checkers*. De estos últimos se hablará de las organizaciones internacionales dedicadas a hacer *fact-checking*, mostrando y explicando su metodología, así como también se hará con las organizaciones chilenas.

En resumen, se seguirá el siguiente orden para explicar el estado del arte:

- *Tweet, Twitter*
- Clasificación no binaria
- Contenido a estudiar de los *tweets*
- Test ANOVA
- *Fact-checking*
- *Fact-checkers*
- Certificaciones
- *Fact-checkers* chilenos
- *Dataset*
- Algoritmo de recomendación

2.1. *Tweet, Twitter*

Como se estudiarán y analizarán los *tweets*, hay que entender como está compuesto y de donde se origina. Por el lado del origen, este viene de la red social *Twitter*, en donde los usuarios pueden compartir mensajes públicos breves con sus seguidores. Desde su lanzamiento en 2006, se ha convertido en una plataforma popular para compartir noticias, opiniones y actualizaciones personales. Los usuarios para comunicarse pueden publicar *tweets*, los cuales pueden ser visibles públicamente para cualquier persona, o bien, los usuarios pueden

optar por enviar mensajes directos a otros usuarios. Estos *tweets* pueden contener hasta 280 caracteres, incluidos texto, *hashtags*, enlaces, fotos y videos [Twitter FAQ, 2023].

Además de compartir mensajes, los usuarios de *Twitter* también pueden interactuar con los *tweets* de otros al darles me gusta, retuitearlos y responderlos. Retuitear permite a los usuarios compartir los *tweets* de otros usuarios con sus propios seguidores, mientras que responder permite a los usuarios comentar un *tweet* directamente. Por otro lado, tenemos a los *hashtags*, que son palabras o frases precedidas por el símbolo #, que permiten a los usuarios categorizar sus *tweets* y hacerlos más fáciles de descubrir por otros.

2.2. Clasificación no binaria

No todas las declaraciones u opiniones pueden clasificarse binariamente como reales o falsas. Si bien algunas declaraciones son sencillas y pueden clasificarse fácilmente como reales o falsas, muchas otras son más complejas y pueden ser subjetivas o estar abiertas a interpretación. Por ejemplo, las declaraciones que se basan en información incompleta o incierta, como "podría llover mañana", no se pueden clasificar como completamente reales o falsas.

En muchos casos, las declaraciones y opiniones pueden ser una mezcla de elementos reales y falsos o pueden depender del contexto, lo que dificulta clasificarlas como totalmente reales o falsas. Es importante considerar el contexto, la evidencia y el matiz de una declaración antes de intentar clasificarla como real o falsa.

Esto mismo también ocurre con los *tweets* que contienen o citan noticias, ya que lo ideal es que todas las noticias se basen en hechos verificables y se presenten de manera objetiva y precisa, pero, no todas las noticias pueden clasificarse binariamente como reales o falsas, especialmente cuando se trata de eventos complejos o en curso.

Los reportajes de noticias a menudo implican la presentación de información de múltiples fuentes, algunas de las cuales pueden ser contradictorias o estar basadas en información incompleta o incierta. En tales casos, puede que no sea posible clasificar definitivamente un informe de noticias como real o falso.

Además, los reportajes de noticias a menudo implican la interpretación y el análisis de eventos, lo que puede introducir subjetividad y diferentes perspectivas. En tales casos, puede ser más apropiado evaluar la calidad y la credibilidad del informe, en lugar de tratar de clasificar el informe como totalmente real o falso.

Es importante ser crítico y perspicaz al consumir noticias y buscar múltiples fuentes para obtener una perspectiva completa. Las habilidades de alfabetización mediática, como la capacidad de evaluar fuentes e identificar sesgos, también pueden ayudar a las personas a emitir juicios informados sobre la precisión y credibilidad de los informes de noticias.

En fin, una clasificación binaria de la veracidad de los *tweets* puede no ser el enfoque más

apropiado para estudiar noticias falsas porque:

- Las noticias falsas suelen ser más complejas, donde además pueden adoptar muchas formas, incluidas la sátira, el *clickbait*, la propaganda y las falsedades absolutas. Puede que no siempre sea fácil clasificar un *tweet* como completamente real o falso, especialmente si contiene elementos de verdad o si tiene la intención de inducir a error de una manera más sutil.
- El contexto es importante, dado que el significado y la veracidad de un *tweet* pueden depender del contexto en el que se comparte, como la audiencia prevista, el momento y la fuente. Una clasificación binaria puede no tener en cuenta estos importantes factores contextuales, que pueden influir en cómo se interpreta y recibe un *tweet*.
- La clasificación binaria puede simplificar demasiado la naturaleza compleja de las noticias falsas y cómo se comparten en las redes sociales. Es posible que no tenga en cuenta los matices de cómo las personas consumen y comparten información en las plataformas de redes sociales, lo que puede involucrar una variedad de factores como el atractivo emocional, el sesgo de confirmación y la influencia social.

En cambio, estudiar las noticias falsas en *Twitter* puede requerir un enfoque más matizado y contextual que tenga en cuenta una variedad de factores, como la fuente de la información, el encuadre y la audiencia prevista. Es posible que los investigadores deban considerar el uso de métodos más sofisticados para analizar *tweets*, como el procesamiento del lenguaje natural o el análisis de sentimientos, para capturar toda la gama de información y perspectivas que se comparten.

2.3. Contenido a estudiar de los tweets

Dado que se quiere diferenciar los *tweets* imprecisos de los reales y falsos, es importante definir qué elementos del contenido de los *tweets* pueden ayudar a caracterizar aquellos *tweets* calificados como imprecisos.

Por un lado, tenemos a Horne [Horne y Adali, 2017] hablando de que las noticias falsas suelen concentrar el grueso de su contenido en el título, ya que las personas tienden a solo leer el título de las noticias. También menciona que las noticias falsas suelen usar menos sustantivos comunes y más sustantivos propios.

Inspirados por estos trabajos, el autor en conjunto con un equipo de investigación buscara reglas entre las distintas categorías de veracidad, a partir de los distintos tipos de elementos gramaticales, como los sustantivos, verbos, cantidad de palabra y el largo del contenido en cuanto a cantidad de caracteres. Para lograr esto, se hará uso de la siguiente métrica.

2.4. Test ANOVA

ANOVA significa "*Analysis of Variance*", que es un método estadístico utilizado para comparar las medias de dos o más grupos. Es un método común utilizado en el análisis de datos para determinar si existe una diferencia significativa entre las medias de varios grupos en función de su variación. Los datos se dividen en grupos y la variación dentro de cada grupo se compara con la variación entre los grupos. Si la variación entre grupos es significativamente mayor que la variación dentro de los grupos, entonces se sugiere que las medias de los grupos son diferentes y que las diferencias no se deben al azar.

Se puede utilizar en una variedad de entornos [Stanley, 2008], como estudios médicos, investigación educativa y control de calidad. Es una herramienta poderosa para identificar diferencias entre grupos y puede ayudar a los investigadores a determinar si una intervención experimental ha tenido un efecto significativo.

Hay varios tipos diferentes de pruebas ANOVA, que incluyen ANOVA unidireccional, ANOVA bidireccional y ANOVA de medidas repetidas. El tipo específico de ANOVA utilizado dependerá de la naturaleza de los datos y de la pregunta de investigación que se esté abordando.

2.5. *Fact-checking*

El *fact-checking* es el proceso de verificar la precisión y la veracidad de las afirmaciones realizadas en noticias, publicaciones en redes sociales y otras formas de medios. El objetivo de este es garantizar que la información que se presenta al público sea precisa y esté respaldada por pruebas, y combatir la difusión de información falsa, engañosa o imprecisa.

En base al artículo de Lucas Graves [Graves, 2016] y a la metodología utilizada por distintos *fact-checkers* que se detallarán más adelante, el proceso de *fact-checking* generalmente implica:

1. Una revisión del reclamo, en donde los *fact-checkers* identificarán el reclamo o la declaración específica que se está haciendo y evaluarán su exactitud.
2. Luego, se inicia una investigación del reclamo, en donde los *fact-checkers* investigarán el reclamo revisando fuentes relevantes, como datos gubernamentales o investigaciones académicas, y consultando con expertos o fuentes primarias.
3. A continuación, se inicia un análisis de la evidencia, para lo cual los *fact-checkers* evaluarán la evidencia en apoyo del reclamo, y evaluarán su credibilidad y relevancia.
4. Finalmente, se llega a una conclusión, con base en la evidencia y el análisis, donde los *fact-checkers* llegarán a una conclusión sobre la exactitud de la afirmación y presentarán sus hallazgos de una manera clara y accesible.

Como describe Amazeen [Amazeen, 2015], no cabe duda de que es una herramienta importante para promover la transparencia y la precisión en los medios, y para ayudar a las personas a tomar decisiones informadas sobre la información que consumen. Es particularmente importante en la era digital, donde la rápida difusión de información en las redes sociales ha facilitado la difusión generalizada de información falsa o engañosa.

2.6. *Fact-checkers*

Un *fact-checker* es una persona u organización que verifica la precisión y veracidad de la información en noticias, publicaciones en redes sociales y otras formas de medios. Los *fact-checkers* generalmente revisan las afirmaciones hechas en un contenido en particular y determinan si están respaldadas por evidencia o si son engañosas o falsas. Estos utilizan una variedad de técnicas para determinar la precisión de la información, incluida la investigación del reclamo, la consulta de expertos o fuentes primarias y el análisis de datos. También pueden usar herramientas como búsquedas inversas de imágenes o detectores de plagio para identificar instancias de información errónea o desinformación. Además, trabajan en una variedad de contextos, ya sea en noticieros, organizaciones sin fines de lucro y plataformas de redes sociales. Su objetivo es promover la precisión y la transparencia en los medios y combatir la difusión de información falsa o engañosa.

El *fact-checking* se ha vuelto cada vez más importante en la era digital, ya que las plataformas de redes sociales han facilitado que la información errónea y la desinformación se propaguen rápida y ampliamente. Muchas organizaciones de noticias y plataformas de redes sociales ahora cuentan con equipos dedicados a revisar el contenido y evitar la difusión de información falsa.

Hay varias organizaciones de verificación de hechos que son bien conocidas y respetadas por su trabajo en la verificación de la exactitud de las afirmaciones hechas en los medios. Algunos de los verificadores de hechos más destacados incluyen:

- **FactCheck.org:** es un proyecto del *Annenberg Public Policy Center*, donde se tiene una organización no partidista que verifica las afirmaciones hechas por políticos, grupos de interés y otras fuentes [FactCheck, 2023].
- **PolitiFact:** es un proyecto del *Instituto Poynter*, en donde se creó un sitio web de verificación de hechos que califica la precisión de las afirmaciones hechas por políticos y otras figuras públicas en una escala de "Truth-O-Meter" [PolitiFact, 2023].
- **The Washington Post Fact Checker:** es una columna publicada por *The Washington Post* que evalúa la precisión de las declaraciones políticas hechas por figuras públicas [Post, 2023].
- **Snopes:** es un sitio web de verificación de hechos que investiga leyendas urbanas, rumores y otras afirmaciones hechas en la cultura popular [Snopes, 2023].

- **Reuters Fact Check:** es un equipo de periodistas que verifica las afirmaciones virales de las redes sociales, la información errónea y las teorías de conspiración [Reuters, 2023].
- **The Associated Press Fact Check:** es un equipo de periodistas que evalúa la exactitud de las afirmaciones hechas por políticos, figuras públicas y otras fuentes [NEWS, 2023].

Estas organizaciones, junto con otras, han desempeñado un papel importante en la promoción de la transparencia y la precisión en los medios y en ayudar a las personas a tomar decisiones informadas sobre la información que consumen, así como también han ayudado a que se realicen diferentes estudios para el análisis de sus metodologías y resultados [Pavleska *et al.*, 2023].

2.7. Certificaciones

No se requiere una sola certificación o acreditación formal para convertirse en *fact-checker*. Sin embargo, existen organizaciones y programas que brindan capacitación y certificación en *fact-checking* y habilidades relacionadas.

Por ejemplo, la *International Fact-Checking Network (IFCN)* [IFCN, 2023] ofrece un programa de certificación para *fact-checkers* que incluye una serie de cursos en línea y un examen de *fact-checking*. El programa cubre temas como técnicas de verificación, análisis de datos y alfabetización mediática, y está diseñado para ayudar a los *fact-checkers* a desarrollar las habilidades y el conocimiento necesarios para evaluar con precisión la veracidad de las afirmaciones.

Además del programa de certificación de la *IFCN*, también hay programas de periodismo y medios que ofrecen cursos de *fact-checking* y habilidades relacionadas. Por ejemplo, el *Poynter Institute* ofrece un curso de *fact-checking* [Poynter, 2023], y el *American Press Institute* ofrece un programa para periodistas [American Press Institute, 2023].

Si bien no existe un enfoque único para convertirse en un *fact-checker*, una combinación de capacitación, experiencia y habilidades de pensamiento crítico es esencial para una verificación de hechos efectiva. Muchos *fact-checkers* tienen experiencia en periodismo, comunicaciones o campos relacionados, y tienen experiencia en investigar y evaluar afirmaciones hechas en los medios.

2.8. Fact-checkers chilenos

El *dataset* con el que se trabajara cuenta con contenido verificado suministrado por tres agencias de *fact-checking*.

La primera es *Fast Checking CL* [Fast Check CL, 2023], la cual cuenta con una certificación dada por la *International Fact-Checking Network (IFCN)*. Ellos desarrollaron una metodología de 10 pasos de para la calificación de tweets, basada en medios como *Politifact*.

A continuación, se mostrará las distintas clasificaciones de veracidad que ocupan para etiquetar un tweet:

- Real
- Impreciso
- Engañoso
- Falso

Seguido de esta se cuenta con Decodificador [Decodificador Chile, 2023], un medio digital independiente, la cual cuenta con una metodología de chequeo tomada y adaptada de medios como *Polifact*, en donde se siguen los siguientes pasos:

1. La mesa de trabajo seleccionará y aprobará la información a chequear, ajustándose a los filtros del medio.
2. Se confrontarán los datos de la información inicial con los recopilados a través del trabajo de verificación.
3. De ser necesario, se consultará la voz de alguna persona o entidad experta para aportar a las conclusiones de la información.
4. Se redactará la publicación y se harán explícitas las conclusiones de algunos datos, si la información lo amerita.
5. Se presenta el informe con una calificación propuesta por el redactor, para ser aprobada o para recibir sugerencias de la mesa de trabajo.

A continuación, se mostrará las distintas clasificaciones de veracidad que ocupan para etiquetar un tweet:

- Verificado
- Impreciso
- Cuestionable
- Engañoso

- Inchequeable
- Falso

Finalmente tenemos a *Fact Cheking UC* [Fact Checking UC, 2023], una iniciativa de verificación implementada por la Facultad de Comunicaciones de la Pontificia Universidad Católica de Chile desde 2013. En este caso ellos ocupan una metodología que se sustenta en los siguientes tres ejes, la relevancia de las afirmaciones seleccionadas, la validez de las fuentes escogidas y la transparencia del proceso.

A continuación, se mostrará las distintas clasificaciones de veracidad que ocupan para etiquetar un *tweet*:

- Creíble
- Creíble, pero
- Sería creíble, pero
- Se puso creativo
- No es creíble
- Ciencia ficción

2.9. *Dataset*

El *dataset* compuesto de contenido verificado de *Twitter* por las tres agencias de *fact-checking* mencionadas en el anterior punto con información relativa a Chile, en donde los *tweets* están clasificados en distintas temáticas, lo cual se representa en el *dataset* con una columna, donde podremos ver los siguientes temas:

- Estallido social.
- COVID-19.
- Elecciones en Chile de 2021.
- Nueva Constitución.
- Otros.

2.9.1. Tópicos

Primeramente, se definirá y hablará sobre los tópicos anteriormente mencionados, tanto como para entender el contexto del *dataset*, como para un posterior análisis y discusión de estos temas en sí. El único tópico del cual no se definirá y escribirá ahora es “Otros”, debido a que no es posible caracterizar a este por contener a los tweets que no pudieron clasificarse en los otros temas.

- **Estallido social:** Chile experimentó un estallido social en octubre de 2019, donde esto comenzó con protestas por el aumento de las tarifas del metro en la ciudad de Santiago. Las protestas fueron provocadas por el anuncio de aumentar el precio del metro con cuestiones ampliadas para incluir elementos políticos y sociales [Jiménez-Yañez, 2020].
- **COVID-19:** La enfermedad, conocida popularmente como COVID-19, es causada por el virus SARS-CoV-2. La enfermedad fue descubierta en diciembre de 2019 en la ciudad de Wuhan, provincia de Hubei, China. Este virus generó una pandemia que ha afectado la salud, la economía y el estilo de vida internacionales en todo el mundo [Marco Ciotti y Bernardini, 2020].
- **Elecciones en Chile de 2021:** Fueron las elecciones que se realizaron el 21 de noviembre de 2021, en donde se buscaba elegir un nuevo presidente, en conjunto con la elección parlamentario y la de consejeros regionales [BCN, 2021].
- **Nueva Constitución:** Una de las consecuencias del estallido social anteriormente nombrado, fue que representantes políticos de la gran mayoría de partidos llegaron a un acuerdo, para llevar a cabo un proceso en el cual se buscaba reemplazar la actual constitución chilena [UC, 2023].

Cabe destacar que la información sacada está dentro del periodo de octubre de 2019 a octubre de 2021.

Tópico	Fast Check	Decodificador	Fact Checking UC	Total
Estallido Social	102	16	12	130
COVID-19	214	53	86	353
Elecciones 2021	67	12	0	79
Constitución	57	36	39	132
Otros	216	52	38	306
Total	656	169	175	1000

Tabla 1: Unidades de contenido verificadas por las agencias de fact-checking.
Fuente: [Mendoza et al., 2023].

Se cuenta con un total de 1000 publicaciones con contenido verificado proveniente de diferentes fuentes. Para el trabajo en cuestión se usarán aquellas publicaciones relativas a *Twitter*, las cuales serían en total 341.

Cabe notar que, para poder trabajar más fácilmente, se realizó una homogeneización de las categorías de veracidad, donde se siguieron los lineamientos propuestos por Amazeen y Lim. A partir de las definiciones de las calificaciones entregadas por los tres medios de *fact-checking* con los que se trabaja, el equipo de investigación de PLU210012 reagrupó el *dataset* con las siguientes categorías: Real, Falso, Impreciso e Inchequeable, en donde los contenidos clasificados con la última categoría fueron borrados.

Etiqueta Original	Etiqueta Nueva
ciencia ficción	False rumor
creíble	True rumor
creíble, pero cuestionable	Impreciso
engañoso	Inchequeable
fake	Impreciso
impreciso	False rumor
no es creíble	Impreciso
real	False rumor
se puso creativo	True rumor
sería creíble, pero verificado	Impreciso
falso	True rumor
inchequeable	False rumor
	Inchequeable

Tabla 2: Conversión de etiquetas de veracidad.

Fuente: Elaboración propia.

Cabe destacar que los contenidos que fueron agrupados en la categoría de impreciso eran aquellos que contenían medias verdades y contenido desactualizado. Estos imprecisos son en total 249, lo cual representa un 25 % del *dataset*.

Veracidad	Fast Check	Decodificador	Fact Checking UC	Total
Real	284	38	53	375
Falso	250	86	37	373
Impreciso	122	42	85	249
Inchequeable	0	3	0	3
Total	656	169	175	1000

Tabla 3: Categorías de veracidad por agencia de *fact-checking*.
Fuente: [Mendoza *et al.*, 2023].

Tópico	Real	Falso	Impreciso	Inchequeable
Estallido social	56	52	22	0
COVID-19	114	136	101	2
Elecciones 2021	30	30	19	0
Constitución	38	58	36	0
Otros	137	97	71	1

Tabla 4: Categorías de veracidad por tema.
Fuente: [Mendoza *et al.*, 2023].

En el caso de *Twitter*, el equipo de investigación recuperó el texto completo del *tweet* e información estructural de este como son los *retweets*, respuestas y comentarios. Para lo que respecta a este trabajo, se le dará prioridad a estudiar el contenido del texto completo por sobre la otra información.

2.10. Algoritmo de recomendación

El 31 de marzo de 2023 se publicó el código fuente de forma abierta del algoritmo de recomendación de *Twitter* [Twitter GITHUB, 2023]. Es importante estudiarlo, debido a que del análisis de este se pueden encontrar relaciones entre el comportamiento de los *tweets* por verificación, así como también de otros fenómenos sociales o cognitivos que se puedan manifestar en la red social de Elon Musk.

Primeramente, se mostrará el modelo general a forma de diagrama, para mostrar las piezas del rompecabezas que conforman este algoritmo:

Del mismo *blog* de donde se sacó la figura, se resume el *pipeline* del algoritmo en tres principales fases:

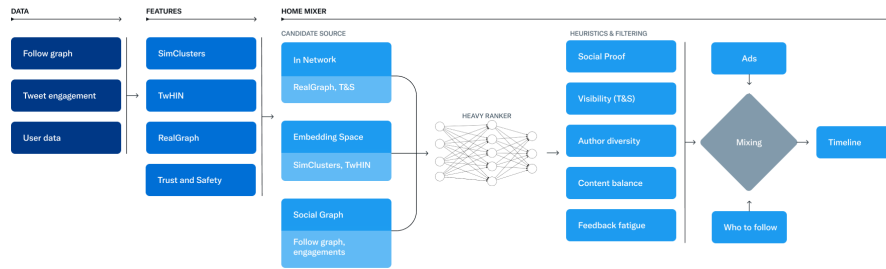


Figura 2: Diagrama del algoritmo de recomendación de *Twitter*.
Fuente: [Twitter Blog, 2023].

1. Se obtiene los mejores *Tweets* de diferentes fuentes.
2. Luego, se clasifican los *tweets* con un modelo de aprendizaje automático
3. Finalmente, se aplican heurísticas y filtros

CAPÍTULO 3

METODOLOGÍA

3.1. Contenido y propagación

Primeramente, se debe estudiar el dataset de *tweets*, contabilizando la cantidad de *tweets* disponibles para el estudio. Dada la antigüedad de algunas publicaciones, algunos *tweets* no son accesibles con la API Académica de *Twitter*, por lo que será importante notar la cantidad de *tweets* a los cuales si se les pudo sacar información con esta API. También, hay que notar la cantidad de *tweets* totales, contabilizando los *retweets* y las respuestas a la publicación original. Finalmente, se debe contabilizar la cantidad de *tweets* raíces (entiéndase *tweet* raíz, como el *tweet* de la publicación original) por tópico y por el etiquetado, donde este último se debe entender como la clasificación de veracidad que se la da a un *tweet*.

Con el test ANOVA se busca sacar distintas medidas relativas a los árboles de respuesta y los textos originales de las aplicaciones a estudiar. Para los árboles de respuesta se estudiará la profundidad, el tamaño y la amplitud de estos. Por otro lado, para los textos de los *tweets*, se buscará estudiar distintos elementos lingüísticos, para la caracterización de los *tweets* imprecisos, de tal forma que se pueda diferenciar a estos de los otros *tweets* raíces catalogados como reales o falsos. Los elementos lingüísticos para estudiar son los siguientes:

- Cantidad de caracteres
- Cantidad de palabras
- Valence
- Arousal
- Dominance
- Entropía
- Emoticones
- Verbos
- Determinadores
- Sustantivos
- Sustantivos propios
- Ad-positions

- Personas
- Lugares
- Organizaciones
- Misceláneo

Luego, para acompañar el estudio de caracterización de los *tweets* catalogados como imprecisos, se estudiarán las métricas de legibilidad computadas en [Mendoza *et al.*, 2023], las cuales son las siguientes:

- GFOG
- NDCRS
- ARI
- FKI
- DCRS
- CLI
- LWRI
- FRE
- IFSZ
- FHRI

Cabe destacar que las métricas anteriores sirven para caracterizar a las personas que publican los *tweets* raíces, lo cual será muy rico en datos para analizar a lo largo del trabajo. Estos, al igual que los elementos lingüísticos, serán estudiados por categoría de clasificación.

3.2. Reclasificación

Por otro lado, se hará una reclasificación de aquellos *tweets* raíces catalogados como imprecisos, para estudiar si existen *tweets* que puedan estar cerca de ser reales o falsos, con el objetivo de ver si dependiendo de la cantidad encontrada, determinar si existe alguna influencia negativa al momento de caracterizar los *tweets* imprecisos, en el sentido de que se pudiera estar solapando los *tweets* imprecisos con las otras categorías y por lo tanto se debería hacer una separación de los *tweets* de esa categoría para poder hacer una correcta caracterización.

Para esto, se seguirán los siguientes pasos:

1. Se filtrará el contenido de *Twitter*.
2. Se seleccionarán los *tweets* raíces catalogados como imprecisos.
3. Se comentará cada *tweet* seleccionado, para ver si es un buen o un mal ejemplo de impreciso.
4. Aquellos *tweets* seleccionados como mal ejemplo se estudiarán, para determinar si se debiese realizar una reclasificación de estos.

3.3. Algoritmo, sesgo y polarización

Una vez terminado el análisis relativo al *dataset*, se procederá a analizar el algoritmo de recomendación de *Twitter*, el cual se centra en recomendar *tweets* “Para ti” en la bandeja de entrada. Para esto, se estudiará el contenido de [Twitter Blog, 2023], donde se buscará aquellas secciones del algoritmo que más relación tengan con la evidencia encontrada en el *dataset*.

Luego, se analizarán fenómenos cognitivos que podrían estar relacionado con los resultados encontrados, tales como el sesgo de confirmación y los fenómenos relacionados a este. Para esto se definirá este fenómeno y los efectos relativos de este, con el objetivo de que en la discusión y conclusión se intente relacionar con la evidencia encontrada.

Finalmente, se procederá a relacionar la polarización política en Chile con los anteriores puntos, para lo cual se estudiarán los diferentes partidos políticos existentes desde el año 2000 en adelante, considerando a solo aquellos que obtuvieron un escaño en la cámara de diputados a partir de las elecciones realizadas desde la fecha mencionada.

CAPÍTULO 4

ANÁLISIS

Los siguientes datos son sacados del *dataset* de pluralismo el cual se filtró en base a la columna de “*Link del contenido*”, en donde se buscaron aquellos que contaran con alguna dirección *url* de *Twitter*. Cabe recalcar también, que los datos fueron recopilados entre octubre de 2019 y octubre de 2021.

Cabe destacar que, si bien se contaba con 341 publicaciones, solo se pudieron acceder a 307 de estos con la *API Académica de Twitter*. De todas formas, se pudieron extraer 397.253 *tweets* en total, los cuales 94.469 corresponden a respuestas y 302.784 a *retweets*, donde cabe notar, que hay un porcentaje mucho mayor de *retweets* (76.22 %) versus las respuestas con respecto a la totalidad de *tweets*.

4.1. Tweets por tópico

Siguiendo lo descrito en la metodología, se contabilizó la cantidad de *tweets* por tópico y por clasificación de veracidad.

Tópico	Real	Falso	Impreciso
Estallido social	20	15	7
COVID-19	43	37	14
Elecciones 2021	7	19	1
Constitución	20	23	29
Otros	39	19	14
Total	129	113	65

Tabla 5: Categorías de veracidad por tema en *Twitter*.

Fuente: [Mendoza *et al.*, 2023].

Con respecto a la tabla, primeramente, hay que ver la cantidad de *tweets* por tópico y porque unos tienen más que otros. Para este análisis se dejaron de lado los *tweets* clasificados en “otros”, debido a que cuenta con *tweets* de diferentes tipos de temas, por lo que se es más difícil concluir y definir algo tan heterogéneo.

Por un lado, se tiene al *COVID-19*, que fue un tema que dio de hablar por mucho tiempo en *Twitter*, dado a la longitud de su duración, desde antes de su llegada a Chile el día 1 de marzo de 2020 [Arrese, 2020], manteniéndose en boca de todos en la red social, cubriendo su duración los años 2020, 2021, 2022 y 2023 [Gob.cl, 2023]. Es por esto por lo que es el tema con más publicaciones del *dataset*.

Seguido de este, se tiene a los *tweets* con temática relativa al proceso constituyente, el cual está relacionada con el tercer tópico con más publicaciones al ser una consecuencia de este, el cual es el estallido social, que a pesar de no tener la misma viralidad, termino teniendo una mayor cantidad de *tweets* debido a que el proceso en sí duró más tiempo [UC, 2023], mientras que el fenómeno del estallido social tuvo una duración de prácticamente de dos meses [Pauta, 2022].

Finalmente tenemos a las elecciones de 2021, cuya duración es la menor con respecto a los otros tópicos, debido a que el proceso en sí duro solo un mes para las elecciones parlamentarias y de consejeros regionales, mientras que la presidencial duro dos meses [BCN, 2021].

Tópico	Real	Falso	Impreciso
Estallido social	42.61 %	35.71 %	16.67 %
COVID-19	45.74 %	39.36 %	14.89 %
Elecciones 2021	25.93 %	70.37 %	3.70 %
Constitución	27.78 %	31.94 %	40.28 %
Otros	54.17 %	26.39 %	19.44 %

Tabla 6: Categorías de veracidad por tema en *Twitter* con porcentajes.

Fuente: [Mendoza *et al.*, 2023].

Con respecto a esta tabla, se busca detallar los porcentajes de cada tópico versus la categoría de veracidad, con el objetivo de encontrar tendencias.

Primeramente, se tiene tanto al estallido social como al *COVID-19*, donde en ambos primaron los *tweets* raíces catalogados como reales (42.61 % y 45.74 % respectivamente). En ambos también se tiene un porcentaje parecido tanto en falsos como imprecisos.

Por otro lado, con respecto a los *tweets* raíces relativos a las elecciones de 2021 y al proceso constituyente, se encontró que en ambos tópicos existen números muy bajos de *tweets* reales (25.93 % y 27.78 % respectivamente). Pero en este caso, estos tópicos difieren mucho en cuanto a la cantidad de falsos e imprecisos, ya que, en el caso de las elecciones, hay más de dos tercios de *tweets* raíces falsos, mientras que contrariamente, en los *tweets* relativos a la constitución hay más imprecisos.

Cabe destacar que por lo definido en el marco teórico y en este capítulo sobre los tópicos del *dataset*, es entendible que se encuentren bajos niveles de *tweets* raíces reales en las elecciones de 2021 como en el proceso constituyente, debido a que estos dos temas son de carácter principalmente político y de elecciones, mientras que el *COVID-19* es relativo a una emergencia sanitaria mundial y el estallido social responde a un fenómeno valga la redundancia social y de protesta más que en un fenómeno político eleccionario en sí. Y es por esto por lo que se puede entender que exista esta tendencia, ya que en el contenido político en general suele tener una gran afluencia de noticias con contenido impreciso, engañoso o falso [Lazer *et al.*, 2018].

4.2. Características lingüísticas

Característica	Rango	Real	Falso	Impreciso
Caracteres*	$\{0, \infty\}$	4171	4086	↑ 4780
Palabras*	$\{0, \infty\}$	687	674	↑ 785
Valence	$[0, 10]$	6.02	5.86	5.94
Arousal	$[0, 10]$	5.28	5.28	5.26
Dominance	$[0, 10]$	5.16	5.09	5.16
Entropía	$] - \infty, 0[$	-4.28	-4.26	-4.23
Emoticones	$[0, 100]$	0	0	0.04
Verbos*	$[0, 100]$	↓ 8.2	9.0	8.8
Determinadores	$[0, 100]$	14.1	14.4	14.2
Sustantivos	$[0, 100]$	20.7	20.7	20.6
Sustantivos propios*	$[0, 100]$	↑ 13.5	12.2	12.0
Ad-positions*	$[0, 100]$	↑ 18.9	17.6	17.8
Personas	$[0, 100]$	1.4	1.4	1.3
Lugares*	$[0, 100]$	↑ 1.8	1.4	1.5
Organizaciones	$[0, 100]$	1.5	1.2	1.4
Misceláneo	$[0, 100]$	2.4	2.6	2.4

Tabla 7: Funciones lingüísticas para unidades de contenido verificadas en Twitter por categoría de veracidad.

Fuente: [Mendoza *et al.*, 2023].

En esta tabla podemos notar que los *tweets* de publicaciones clasificadas como imprecisos, en general, tienen un largo de caracteres y de palabras mayor que los clasificados como reales o falsos. También, los *tweets* imprecisos también suelen usar más verbos que los *tweets* reales, pero, por otro lado, los *tweets* imprecisos usan menos sustantivos propios que los *tweets* reales.

Métrica	Rango	Real	Falso	Impreciso
GFOG*	{0, 20}	16.12	16.11	↓ 15.05
NDCRS	{0, 20}	11.32	11.12	11.42
ARI*	{5, 22}	17.12	↓ 16.12	↓ 15.08
FKI*	{0, 18}	12.57	12.08	↓ 10.96
DCRS	{0, 20}	13.08	13.08	12.87
CLI*	{0, 20}	18.64	↓ 17.58	↓ 17.46
LWRI*	{0, 100}	8.87	8.54	↓ 7.02
FRE*	{1, 100}	↓ 36.12	39.65	40.09
IFSZ*	{1, 100}	↓ 64.11	68.17	67.28
FHRI*	{1, 100}	↓ 66.31	70.12	70.14

Tabla 8: Funciones lingüísticas para unidades de contenido verificadas en Twitter por dimensión de tema.

Fuente: [Mendoza *et al.*, 2023].

Podemos notar que en GFOG, ARI, FKI, DCRS, CLI, LWRI, los tweets raíces catalogados como imprecisos obtuvieron un puntaje menor frente a las otras dos categorías, donde estas métricas comparten en que a menor puntaje, menor nivel de comprensión lectora es necesario tener para entender el contenido del tweet raíz. Por otro lado, en IFSZ y FHRI, los tweets raíces catalogados como imprecisos y falsos obtuvieron un mayor puntaje, donde estas métricas a diferencia de las otras comparten en que, a mayor puntaje, menor nivel de comprensión lectora es necesario tener para entender el contenido del tweet raíz.

4.3. Propagación de los tweets

Profundidad	Media	Std	Min	Max
Real	↓ 6.92	8.54	1	78
Falso	9.13	11.54	1	88
Impreciso	9.92	10.10	1	60

Tabla 9: Profundidad de los árboles de propagación en Twitter (respuestas).

Fuente: [Mendoza *et al.*, 2023].

Tamaño	Media	Std	Min	Max
Real	485	1701	3	14237
Falso	457	669	2	3570
Impreciso	↑ 679	1101	4	5321

Tabla 10: Tamaño de los árboles de propagación en *Twitter* (respuestas).Fuente: [Mendoza *et al.*, 2023].

Amplitud	Media	Std	Min	Max
Real	332	1235	1	10813
Falso	299	469	1	2833
Impreciso	↑ 380	543	1	1883

Tabla 11: Amplitud de los árboles de propagación en *Twitter* (respuestas).Fuente: [Mendoza *et al.*, 2023].

De la tabla 9, 10 y 11 se desprende que los *tweets* raíces catalogados como imprecisos tienen conversaciones más largas, dado que los hilos de *tweets* son más largos y a la vez existe una mayor cantidad de *tweets* de por sí, lo cual en promedio debe significar que exista una mayor participación de personas a estas conversaciones.

En [Mendoza *et al.*, 2023] también se graficó el tiempo promedio que toman los *tweets* por categoría de veracidad en llegar a un número de profundidad, tamaño y ancho de las cascadas de *tweets*. De estas se concluye que el contenido impreciso toma menos tiempo en propagarse que las otras categorías, debido a que para alcanzar una profundidad promedio de un hilo de 20 *tweets*, los *tweets* raíces catalogados como reales y falsos tomaban 15 horas y 5 respectivamente, los imprecisos necesitaban solo 2 horas para alcanzar esa profundidad.

4.4. Reclasificación

Si bien se pudieron encontrar reglas en el test ANOVA, hubo puntos donde no se pudieron determinar reglas para alguna de las métricas. Es por esto por lo que surgió la duda de que tal vez al momento de homologar las clasificaciones entre *fact-checkers* y reducirlas a real, impreciso y falso, se pudieron clasificar en impreciso a *tweets* que tengan un parecido en su estructura y comportamiento con los reales o falsos.

Es por lo anterior, que se llamó para un proceso de reclasificación a distintas personas del grupo de trabajo del PLU210012, que tenían como objetivo estudiar los distintos *tweets* que fueron clasificados como imprecisos, para poder indicar algún patrón dentro de estos que podría dar cuenta de alguna tendencia.

Del original total de 65 tweets raíces imprecisos, se ocuparon 60, debido a que 5 de estos habían sido borrados a la fecha en la que se volvió a revisar. Durante el estudio de estos tweets, se encontró con que podía existir una imprecisión cuantitativa o cualitativa, dependiendo si el dato que estaba haciendo ruido era una cifra, texto o imagen. Nótese que lo siguiente tiene que ver con información que estuviera en el tweet o bien que estuviera en una fuente que fuera citada por el tweet.

En base a los tweets y a los links de verificación de los *fact-checkers* (*fastcheckCL*, *decodificador* y *factcheckingUC*) que se encuentran en el *GITHUB* de [Bassi, 2023], se colocaron comentarios del autor de esta memoria al lado de cada uno de estos, a la vez de determinar si era buen o mal ejemplo de impreciso un tweet. Luego se sintetizaron los comentarios realizados por el autor, determinando que se encontraron tres tipos de imprecisos:

- **Poco impreciso:** Si un tweet no dependía de una cifra que afectara su veracidad, pero una parte pequeña si y esta cifra estuviera desactualizada o difiere muy poco de la real, era clasificado aquí. Por otro lado, también eran clasificados aquí aquellos tweets donde mencionaban noticias que en gran parte de su cuerpo eran reales, pero una pequeña parte no había sido confirmada o verificada la información, pero de todas formas no afectaban al espíritu de la noticia.
- **Impreciso:** Eran clasificados como imprecisos aquellos tweets, donde cercano a la mitad de la información de estos, dependieran de una cifra que estuviera desactualizada o no confirmada su veracidad. También, había tweets donde contenían noticias que se basaban en hechos que no habían sido confirmados o verificado su información, o bien directamente eran falsos.
- **Muy impreciso:** Aquí eran clasificados aquellos tweets, donde gran parte de la información de estos, dependiera de una cifra que estuviera desactualizada o difiriera mucho de la real. Asimismo, eran clasificados aquí a aquellos tweets donde una pequeña parte de la información citaba un hecho real, pero una gran parte declaraba cosas en base a fuentes de la cual no se había confirmado o verificado su información, o que bien directamente tenían información falsa.
- **Información engañosa:** Los tweets clasificados en información engañosa eran aquellos que declaraban cosas en base a información real, pero que no tuvieran relación alguna con lo que se estaba afirmando, o bien omitiendo información clave en gran parte de la noticia (*framing*), para llevar al lector a una conclusión errónea.

Como se puede ver, los tweets que eran clasificados como muy poco imprecisos, tienen una estructura gramatical en su redacción muy parecida a los tweets reales y probablemente con un comportamiento muy parecido en cuanto a los árboles de *retweets* y respuestas. Así mismo, los tweets que eran clasificados como muy imprecisos se parecían mucho a los tweets falsos y es posible que se comporten casi de la misma forma en su propagación.

Con estas definiciones, se realizó una reclasificación, en donde se encontró que, de los 60 tweets estudiados, 18 (30.00 %) de estos eran poco imprecisos, 30 (50.00 %) eran imprecisos, 1 (1.67 %) era muy impreciso y 11 (18.33 %) eran información engañosa. Con esto, queda confirmado que el comportamiento encontrado al estudiar las anteriores tablas de los imprecisos era adecuado, ya que lo que se buscaba era ver si existía un porcentaje importante de tweets imprecisos que tuvieran un parecido con tweets falsos y que eso deviniera en el parecido de algunas de las cifras encontradas en las métricas.

4.5. Algoritmo de recomendación

Primeramente, se describirá la primera fase, la cual prácticamente es la entrada de datos inicial del algoritmo. En la fase descrita como *DATA* del diagrama, se tienen los siguientes elementos:

- **Follow Graph:** El grafo de seguimiento entre usuarios.
- **Tweet engagement:** Las métricas de interacción entre *tweets*.
- **User data:** Datos relativos al usuario, como por ejemplo bloqueos o "los me gusta."^a publicaciones.

Se describieron de forma breve, dado que no son tan complejos e imperantes para el estudio a diferencia de la siguiente fase.

Luego, viene la fase de aplicación de cálculos, en donde se destaca el *RealGraph* [Kamath *et al.*, 2014], que es un modelo que calcula la probabilidad de interés en lo publicado por otro usuario. Cabe destacar que esto se aplica incluso si nunca se ha interactuado con ese usuario.

En esta fase también se encuentra el *TweepCred*, que es un algoritmo basado en *PageRank* [Griffiths *et al.*, 2007], cuyo objetivo es encontrar los elementos más influyentes de un grafo, que para el caso de *Twitter* sería el grafo de usuarios.

Los dos puntos anteriores están relacionados en el sentido de que ambos eligen *tweets* candidatos dentro de la red del usuario. Por otro lado, para la elección de *tweets* fuera de la red, se tiene el *Social Graph* y los *Embedding Spaces*. El primero, cuya implementación en *Twitter* se llama *GraphJet* [Sharma *et al.*, 2016], es un método de búsqueda de contenido afín a gustos del usuario, lo que en el caso de *Twitter* serían los *tweets* que han interactuado usuarios que el usuario sigue y *tweets* con los que han interactuado usuarios con características similares al usuario. Con respecto a los *Embedding Spaces*, son más complejos que el primero, aunque tienen el mismo objetivo de encontrar *tweets* y usuarios afines al usuario. La técnica *embedding*, tiene como objetivo transformar en números cadenas de texto,

lo que en este caso sería el contenido de un *tweet*, lo cual sirve para determinar el grado de similaridad entre dos *tweets*, o bien, dos conjuntos de *tweets*. Cabe destacar que hay varios *Embedding Spaces* utilizados en *Twitter*, donde uno de los más importantes para el algoritmo de recomendación es *SimClusters* [Satuluri *et al.*, 2020], el cual crea comunidades artificiales dentro de *Twitter*, las cuales no son visibles, donde están agrupadas en torno a usuarios influyentes. En el *blog* se menciona que se contabilizan en total 145000 comunidades generadas, las cuales se actualizan cada 3 semanas y que tantos los usuarios como los *tweets* pueden pertenecer a más de una comunidad. En resumen y textualmente traducido del *blog*: **Cuanto más les guste un tweet a los usuarios de una comunidad, más se asociará ese tweet con esa comunidad.**

4.6. Sesgo de confirmación

El sesgo de confirmación es un error cognitivo o una preferencia por buscar e interpretar información de tal manera que confirme teorías o ideas ya existentes [Klayman, 1995]. Es común que los humanos graviten inconscientemente hacia la información y enfatizan la evidencia de acuerdo con sus creencias ya mantenidas, descartando o trivializando la evidencia de oposición.

Este fenómeno se nota en muchos lugares de la vida cotidiana, incluidas las creencias/cogniciones personales, las perspectivas políticas e incluso a nivel de investigación científica. Cuando las personas encuentran datos ambiguos u oscuros, favorecen aquellos que respaldan sus prejuicios, mientras descartan o evitan lo contrario. Esta perspectiva también puede afectar negativamente al razonamiento objetivo y al pensamiento crítico.

El sesgo de confirmación está muy extendido debido a varias razones. La exposición selectiva es una de esas cosas en las que las personas tienden a buscar la información que se adapta a sus creencias, mientras descartan la información que no encaja. Otra es la atención selectiva, en la que las personas tienden a priorizar y dar más peso a la información que confirma sus creencias previas. Además, también se aplica la lectura selectiva: la gente verá material poco claro e indeciso de tal manera que ya cree [Sunstein, 2018].

Las consecuencias del sesgo de confirmación pueden ser sustanciales. Esta dinámica puede dar lugar a cámaras de eco, donde las personas se rodean de otras que comparten puntos de vista similares pero que nunca cuestionan sus propias creencias [Vicario *et al.*, 2016]. Lo que potencialmente puede sofocar la exposición a diferentes puntos de vista y obstaculizar la adaptabilidad. En áreas como la ciencia, el sesgo de confirmación puede sesgar los resultados de la investigación, especialmente cuando los investigadores científicos analizan sólo los patrones de datos deseados y no tienen en cuenta otras posibilidades.

Relacionado a lo anterior, está el efecto *Dunning-Kreuger*, que es un sesgo cognitivo en el que las personas con menos habilidades sobrestiman sus propias habilidades [Dunning, 2011]. Los no cualificados son incapaces de reconocer su propia estupidez, y esto es lo que se llama

sesgo metacognitivo. Básicamente es un fenómeno en el que personas sin experiencia ni habilidades en un determinado campo se consideran altamente calificadas o educadas en esa área.

David Dunning y Justin Kruger hicieron la primera descripción de este fenómeno en 1999 después de realizar una serie de estudios [Kruger y Dunning, 1999]. Los investigadores establecieron que las personas con poca capacidad en diversos aspectos tienden a cometer errores y no reconocen estos errores por parte de la autorreflexión. Por tanto, pueden creerse mejores que los profesionales o las personas altamente calificadas. Sin embargo, las personas con mayores habilidades podrían subestimarse a sí mismas o considerarse inadecuadas. Esto es así porque piensan que la facilidad que les resultan estas cosas demuestra que sus habilidades no son únicas, sino que las comparten otros.

4.7. Polarización política en Chile

Cuando se estudia la historia política de Chile de los últimos 20 años, se van a encontrar sorpresas, en el sentido de la gran cantidad de cambios producidos al sistema electoral, así como también a la evolución misma de los partidos políticos y como estos como fuerzas/alianzas han cambiado.

Considerando solamente los partidos políticos que ganaron un escaño mínimo en la cámara de diputados en el momento de las elecciones y no alguna formación de partido en los años intermedios de las elecciones, desde el año 2001 al 2021 han existido en la cámara de diputados chilena un total de 26 partidos políticos de diferentes ideologías.

Antes de comenzar con el análisis, se mostrará la clasificación de partidos políticos en base a rangos ideológicos, con base a lo descrito y estudiado de la descripción de cada partido en la Biblioteca del Congreso Nacional [BCN, 2023].

Primero, se mostrarán aquellos partidos, cuya posición ideológica sea puramente de izquierda:

- Partido Comunista de Chile (PCCh)
- Partido Humanista (PH)
- Izquierda Ciudadana (IC)
- Partido Igualdad (PI)
- Poder Ciudadano (PODER)
- Comunes

Los partidos cuya ideología está dentro del rango de izquierda a centro izquierda son los siguientes:

- Partido Ecologista Verde de Chile (PEV)
- Convergencia Social (CS)
- Revolución Democrática (RD)
- Federación Regionalista Verde Social (FREVS)
- Partido Progresista (PRO)
- Partido Socialista de Chile (PS)

Luego, están los partidos que tienen una ideología de centro izquierda a centro o de centro a centro derecha:

- Partido Liberal de Chile (PL)
- Partido Radical de Chile (PR)
- Partido por la Democracia (PPD)
- Ciudadanos (CIU)
- Partido de Acción Regionalista de Chile (PAR)
- Partido Demócrata Cristiano (PDC)
- Centro Unido (CU)

Seguido, están los partidos que tienen una ideología de centro derecha a derecha:

- Partido de la Gente (PDG)
- Partido Regionalista Independiente Demócrata (PRI)
- Evolución Política (Evópoli)

Finalmente, los partidos cuya posición ideológica son puramente de derecha son los siguientes:

- Renovación Nacional (RN)

- Unión Demócrata Independiente (UDI)
- Partido Social Cristiano (PSC)
- Partido Republicano de Chile (PLR)

Luego, se contabilizó la cantidad de diputados sacados por los partidos anteriormente nombrados en las elecciones parlamentarias de Chile de 2001, 2005, 2009, 2013, 2017 y 2021, donde desde 2001 a 2013 se elegían 120 diputados, mientras que desde 2017 a 2021 se escogían 155. Ténganse en cuenta que en la tabla 12 no se muestran los candidatos independientes.

Es importante destacar de la tabla 13, la cual fue construida a partir de lo anteriormente descrito, la caída que ha tenido en cantidad de diputados las fuerzas políticas que son puramente de centro, lo que vislumbra una creciente polarización en la sociedad chilena. Por otro lado, si se compara la cantidad de partidos políticos, que tenían al menos un diputado electo en el año 2001 versus el año 2021, se ve un aumento de 15 partidos, lo cual es un aumento considerable.

Finalmente, se debe entender que la enunciación de una polarización aparece primeramente en las elecciones de 2017, justamente con el nuevo sistema electoral, promulgado el 25 de abril de 2015 [BCN, 2015], que reemplazaba el antiguo sistema binominal por el sistema proporcional inclusivo. Por último, es interesante ver que en las elecciones post estallido social y las consecuencias de este, se vislumbra un aumento aún mayor de esta polarización.

Partido	2001	2005	2009	2013	2017	2021
PCCh	0	0	3	6	8	12
PH	0	0	0	0	5	3
IC	0	0	0	1	0	0
PI	0	0	0	0	1	0
PODER	0	0	0	0	1	0
Comunes	0	0	0	0	0	6
PEV	0	0	0	0	1	2
CS	0	0	0	0	0	9
RD	0	0	0	0	10	8
FREVS	0	0	0	0	4	2
PRO	0	0	0	0	1	0
PS	10	15	11	15	19	13
PL	0	0	0	1	2	4
PR	6	7	5	6	8	4
PPD	20	21	18	15	8	7
CIU	0	0	0	0	0	1
PAR	0	1	0	0	0	0
PDC	23	20	19	21	14	8
CU	0	0	0	0	0	1
PDG	0	0	0	0	0	6
PRI	0	0	3	0	0	1
Evópoli	0	0	0	0	6	4
RN	18	19	18	19	36	23
UDI	31	33	37	29	30	25
PSC	0	0	0	0	0	1
PLR	0	0	0	0	0	14

Tabla 12: Cantidad de diputados obtenida por partido político en las elecciones parlamentarias, desde 2001 a 2021.

Fuente: Elaboración propia.

Ideología	2001	2005	2009	2013	2017	2021
Izquierda	0.00 %	0.00 %	2.50 %	5.83 %	9.68 %	13.55 %
Centro Izquierda	8.33 %	12.50 %	9.17 %	12.50 %	22.58 %	21.94 %
Centro	40.83 %	40.83 %	35.00 %	35.83 %	20.65 %	16.13 %
Centro Derecha	0.00 %	0.00 %	2.50 %	0.00 %	3.87 %	7.10 %
Derecha	40.83 %	43.33 %	45.83 %	40.00 %	42.58 %	40.65 %

Tabla 13: Cantidad de diputados obtenida por bloque político en las elecciones parlamentarias, desde 2001 a 2021.

Fuente: Elaboración propia.

CAPÍTULO 5

CONCLUSIONES

Lo primero que se puede notar al estudiar la tabla de tópicos, es que efectivamente la cantidad de publicaciones durante el tiempo no depende exclusivamente de cuanto dure el fenómeno en sí, ya que también depende de la magnitud de este, en el sentido de cuan viral puede llegar a ser. Es por esto por lo que el estallido social tuvo más de la mitad de las publicaciones que tuvo el tópico de proceso constituyente, a pesar de que el primero duró menos tiempo.

Al principio se tenía dudas si se encontrarían diferencias remarcables de los *tweets* clasificados como imprecisos versus los reales o falsos. Por un lado, los imprecisos comparten con los falsos que en su contenido tengan más verbos y menos sustantivos propios que los verdaderos. Pero, por otro lado, los imprecisos se diferencian significativamente de verdaderos y falsos en que usan más caracteres y palabras en su *tweet* raíz.

Durante el estudio de las métricas relativas a la comprensión lectora, se pudo evidenciar que el contenido de los *tweets* raíces imprecisos suelen tener un nivel menor requerido para poder ser entendido por las personas, lo cual si lo sumamos a lo descubierto de que estos mismos *tweets* suelen ser compartidos y discutidos por más personas, a la vez de tener discusiones más largas, se termina conectando indudablemente con el modelo del efecto *Dunning-Kruger*, donde este enfatiza la autoconciencia y las autoevaluaciones objetivas. Esto implica que es probable que los individuos sin habilidades y conocimientos adecuados de un dominio en particular exageren su competencia en dicho dominio porque no poseen habilidades metacognitivas relevantes para una capacidad de autoevaluación efectiva. En base a esto, si una persona que subestima su conocimiento sobre un tema y *tweetea* sobre este, lo más probable es que tienda a ocupar más palabras de las que debería para expresarse y además dará menos datos relevantes sobre la temática, provocando a que la persona tienda a usar menos nombres de personas, lugares u organizaciones en sus argumentos. Esta tendencia de usar más palabras de lo que se debería y el poco uso de sustantivos propios calza con el comportamiento de los *tweets* raíces catalogados como imprecisos.

El desarrollo de grupos burbuja y cámaras de eco en *Twitter* puede verse afectado por el sesgo de confirmación [Flaxman *et al.*, 2016]. Dado lo visto en este trabajo y en otros que se han ido mencionando, uno de las razones de porque se genera esto es por la **exposición selectiva**, donde las personas tienden a buscar algunos detalles que ya creen que cumplirían con sus expectativas; ignorando a los demás. En *Twitter*, los usuarios pueden crear sus listas siguiendo sólo a aquellos con opiniones similares [Twitter Help, 2023]. Esta tendencia a menudo significa que las personas sólo encuentran mensajes que refuerzan sus creencias previas sobre un tema.

También, se tiene el **refuerzo algorítmico**, ya que las redes sociales como *Twitter* dependen de algoritmos para crear *timelines* específicas para el usuario. Estos algoritmos tienen co-

mo objetivo proporcionar a los usuarios contenido que les resultará fascinante y relevante a juzgar por su actividad anterior y su historial de clics. Por ejemplo, si un usuario interactúa repetidamente con información congruente con su opinión, el algoritmo intensifica esto promoviendo más de esas piezas, creando y fortaleciendo así el proceso de una cámara de eco [Samuels, 2012]. Dado que en *Twitter* se hizo público el código fuente de su algoritmo de recomendación, se pudo notar que los *embedding spaces* son los que terminaban ayudando en gran medida a la generación de grupos burbujas y cámaras de eco de comunidades que tuvieran usuarios con características afines, o bien con usuarios que terminaban uniéndose a una comunidad debido a que publicaban *tweets* con un contenido a fin a esa comunidad [Satuluri *et al.*, 2020].

Por otro lado, las métricas de participación en *Twitter*, como **los retweets y los "me gusta"**, pueden aumentar el sesgo de confirmación. Cuando dichos usuarios ven que ese contenido consistente con sus creencias recibe el visto bueno y comenta, lo más probable es que lo transmitan compartiéndolo o *retweeteandolo*, o dándole me gusta, entre otras cosas. Esto es mencionado en la descripción de la entrada de datos con respecto del usuario cuando se le recomiendan *tweets* con el algoritmo.

Hay que considerar que los usuarios también pueden utilizar el **refuerzo social** dentro de estas cámaras de eco publicando contenidos que probablemente sean compartidos por los seguidores. Esto puede desarrollar un círculo vicioso, en el que los usuarios de una burbuja refuerzan las ideas de los demás. Esto se puede sumar a un **desaliento de opiniones diferentes**, lo cual puede terminar. Es posible que estén menos expuestos a ideas alternativas y que descarten o incluso ataquen opiniones opuestas cuando las encuentren. Esto también puede conducir a un ciberespacio polarizado y hostil. Aunque la polarización es evidente dados los resultados mostrados en la tabla que mues la cantidad de diputados obtenida por bloque político en las elecciones parlamentarias (desde 2001 a 2021), si bien la polarización ocurrió ante del surgimiento del fenómeno de las *fake news* y contenido impreciso en las redes sociales [Lindh *et al.*, 2019], está claro que en los últimos años los partidos políticos a nivel mundial han gastado más en redes sociales, ya que son estas las que se han transformado en las principales plataformas de información para mucha gente, debido a que generan más confianza que los medios de comunicación masivos tradicionales [Granese, 2021].

Como consecuencia de todo lo anterior y que además termina alimentando el sesgo de confirmación, es la **formación de identidades** en línea, ya que también puede verse afectada por las cámaras de eco en *Twitter*. El sesgo de confirmación puede profundizarse aún más cuando los usuarios desarrollan un fuerte apego a un grupo u orientación ideológica en su burbuja, creando un refuerzo adicional [Pennycook *et al.*, 2020].

Finalmente, la generación misma de grupos burbujas y cámaras de eco pareciera que se terminara perpetuando por el mismo algoritmo de recomendación de *Twitter* mencionado, ya que este termina generando una polarización, donde dada la evidencia vista en el análisis, es en este tipo de escenarios de polarización en donde se terminan asentando los *tweets* raíces imprecisos. Esto mismo se vio en Turquía [Tufekci, 2017], así como también se vio en Chile con el estallido social, solo que no se evidencio en este en sí dado su aparición efímera pero

contundente en *Twitter*, sino que, durante el proceso constituyente, que de todas formas es una consecuencia directa del estallido. Esto tiene una gran influencia en el discurso público, especialmente en las llamadas cámaras de eco y grupos burbuja. Tales factores pueden reducir la exposición a opiniones variadas, lo que resulta en un diálogo sofocante y una polarización social. Comprender la contribución de los sesgos de confirmación a la formación de estas cámaras de eco ayuda a analizar la dinámica de la comunicación en línea con miras a apoyar comunicaciones más abiertas.

A modo de resumen, se tiene que hay varias razones por las que la información imprecisa puede difundirse más fácilmente que la información verdadera y en algunos casos de la información falsa:

1. **Novedad y sensacionalismo:** La información imprecisa tiene un factor de “novedad” (en comparación con las noticias que son sabidas como verdaderas o falsas) y, por lo tanto, es más fácil para una persona compartirla porque se siente nueva e inusual; esto hace que las historias imprecisas se difundan más rápidamente en comparación con los hechos reales.
2. **Sesgo de confirmación:** Generalmente, las personas prefieren la información verificable que refuerza sus percepciones actuales. Como tal, esto podría resultar en la difusión de información imprecisa a través de ciertos grupos comunitarios o redes sociales.
3. **Algoritmo de recomendación:** El mismo algoritmo que recomienda *tweets* en la bandeja de entada, por el diseño mismo de este termina generando comunidades artificiales.
4. **Emoción e identidad:** La información imprecisa se propaga más rápido, ya que evoca sentimientos o identidades y, por lo tanto, es mucho más probable que se propague al compartirla y ser aceptada como verdad.
5. **Lapsos de atención limitados:** el mundo actual cambia rápidamente. Las personas tienden a no investigar antes de compartir sus datos. Esto puede dar lugar a que hechos con datos erróneos lleguen a las masas antes de que puedan retractarse.
6. **Campañas de desinformación:** Personas u organizaciones hacen circular deliberadamente mentiras con el fin de manipular la opinión pública o provocar disturbios.

En general, la proliferación de información imprecisa ha sido causada (en parte) por puntos ciegos psicológicos, dinámicas impulsadas por la tecnología y esfuerzos conscientes de desinformación. Dicho esto, las personas deben tener cuidado de no compartir información incorrecta de las redes sociales antes de consultarla ellos mismos y las plataformas deben hacer más para evitar que la información errónea se vuelva viral. El efecto *Dunning-Kruger* afecta muchos aspectos como la educación, el desempeño laboral y la toma de decisiones. Destaca lo importante que es ser consciente de uno mismo y tener una evaluación correcta, precisa y válida para la mejora personal y profesional.

Reducir el sesgo de confirmación requiere un esfuerzo y una reflexión conscientes. Requiere una curiosidad activa por otros puntos de vista, una apertura a interpretaciones alternativas y críticas de las propias suposiciones. Debatir y dialogar cortésmente con personas que tienen otras opiniones puede abrir puertas a la exposición a diversas ideas, disminuyendo el sesgo de confirmación para cualquiera. Y es fundamental desarrollar una disposición hacia la modestia epistemológica manteniendo la mente abierta a la hora de alterar sus creencias con información nueva. Es por esto, que es importante identificar y gestionar nuestras tendencias hacia esta forma de falacia es fundamental para desarrollar una lente más equilibrada y objetiva con la que tomar decisiones, resolver problemas y percibir la realidad. Ser consciente de este punto ciego permite a las personas intentar contrarrestarlo y fomentar la percepción de la realidad basada en la evidencia.

REFERENCIAS BIBLIOGRÁFICAS

- [Amazeen, 2015] Amazeen, M. (2015). Revisiting the epistemology of fact-checking. *Critical Review*.
- [American Press Institute, 2023] American Press Institute (2023). Training Tools - American Press Institute. Disponible en <https://www.americanpressinstitute.org/training-tools/>.
- [Arrese, 2020] Arrese, S. (2020). Observatorio de enfermedades infecciosas. Disponible en <https://observatorio.medicina.uc.cl/primer-caso-de-covid-19-en-chile/>.
- [Bassi, 2023] Bassi, R. (2023). GITHUB: bassisi/PLU210012. Disponible en <https://github.com/bassisi/PLU210012>.
- [BCN, 2015] BCN (2015). Biblioteca del congreso nacional: Ley 20840. Disponible en <https://www.bcn.cl/leychile/navegar?idNorma=1077039>.
- [BCN, 2021] BCN (2021). Elecciones parlamentarias de 2021. Disponible en https://www.bcn.cl/historiapolitica/elecciones/detalle_eleccion?handle=10221,1%2F87444&periodo=1990-2022.
- [BCN, 2023] BCN (2023). Biblioteca del Congreso Nacional: Historia Política. Disponible en https://www.bcn.cl/historiapolitica/partidos_politicos/index.html.
- [Campos *et al.*, 2017] Campos, E., Miguel, Calvo, D., y Díez, M. (2017). Twitter y la comunicación política. *El profesional de la información*, 26(5):785-794.
- [Decodificador Chile, 2023] Decodificador Chile (2023). Decodificador Chile. Disponible en <https://decodificador.cl/>.
- [Dunning, 2011] Dunning, D. (2011). The dunning-kruger effect: On being ignorant of one's own ignorance. En *Advances in experimental social psychology*, volumen 44, pp. 247-296. Elsevier.
- [Fact Checking UC, 2023] Fact Checking UC (2023). Fact Checking UC. Disponible en <https://factchecking.cl/>.
- [FactCheck, 2023] FactCheck (2023). Factcheck. Disponible en <https://www.factcheck.org/>.
- [Fast Check CL, 2023] Fast Check CL (2023). Fast Check CL. Disponible en <https://www.fastcheck.cl/>.
- [Flaxman *et al.*, 2016] Flaxman, S., Goel, S., y Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(51):298-320.

- [Garimella y Weber, 2017] Garimella, V. R. K. y Weber, I. (2017). A long-term analysis of polarization on twitter. En *Proceedings of the International AAAI Conference on Web and social media*, volumen 11, pp. 528–531.
- [Gob.cl, 2023] Gob.cl (2023). Gob.cl. Disponible en <https://www.gob.cl/noticias/fin-alerta-sanitaria-covid-19-coronavirus-enfermedades-respiratorias-mascarillas-teletrabajo/>.
- [Granese, 2021] Granese, M. (2021). La utilización de redes sociales en campañas políticas: problemas y posibles soluciones. *Punto de Referencia*, 561:1–25.
- [Graves, 2016] Graves, L. (2016). Anatomy of a fact check: Objective practice and the contested epistemology of fact checking. *Communication, Culture Critique*.
- [Griffiths et al., 2007] Griffiths, T. L., Steyvers, M., y Firl, A. (2007). Google and the mind: Predicting fluency with pagerank. *Psychological science*, 18(12):1069–1076.
- [Horne y Adali, 2017] Horne, B. D. y Adali, S. (2017). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news.
- [IFCN, 2023] IFCN (2023). IFCN - Poynter. Disponible en <https://www.poynter.org/ifcn/>.
- [Institute, 2020] Institute, R. (2020). Digital news report: Chile. Disponible en <https://www.digitalnewsreport.org/survey/2020/chile-2020/>.
- [Jiménez-Yañez, 2020] Jiménez-Yañez, C. (2020). # chiledespertó: causas del estallido social en chile. *Revista mexicana de sociología*, 82(4):949–957.
- [Kamath et al., 2014] Kamath, K., Sharma, A., Wang, D., y Yin, Z. (2014). Realgraph: User interaction prediction at twitter. En *user engagement optimization workshop@ KDD*, número ii.
- [Klayman, 1995] Klayman, J. (1995). Varieties of confirmation bias. *Psychology of learning and motivation*, 32:385–418.
- [Kruger y Dunning, 1999] Kruger, J. y Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6):1121.
- [Lazer et al., 2018] Lazer, D., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S., Sunstein, C., Thorson, E., Watts, D., y Zittrain, J. (2018). The science of fake news. *Science*, 359:1094–1096.
- [Lindh et al., 2019] Lindh, J., Fábrega, J., y González, J. (2019). La fragilidad de los consensos. polarización ideológica en el chile post pinochet. *Revista de ciencia política (Santiago)*, 39(1):99–127.

- [Marco Ciotti y Bernardini, 2020] Marco Ciotti, Massimo Ciccozzi, A. T. W.-C. J. C.-B. W. y Bernardini, S. (2020). The covid-19 pandemic. *Critical Reviews in Clinical Laboratory Sciences*, 57(6):365–388. PMID: 32645276.
- [Mendoza et al., 2023] Mendoza, M., Valenzuela, S., Núñez-Mussa, E., Padilla, F., Providel, E., Campos, S., Bassi, R., Riquelme, A., Aldana, V., y López, C. (2023). A study on information disorders on social networks during the chilean social outbreak and covid-19 pandemic. *Applied Sciences*, 13(9).
- [NEWS, 2023] NEWS, A. (2023). Ap news. Disponible en <https://apnews.com/hub/ap-fact-check>.
- [Pauta, 2022] Pauta (2022). Cronología del Estallido Social: Lo que pasó entre el 18-O y el 15-N. Disponible en <https://www.pauta.cl/actualidad/2022/10/18/cronologia-estallido-social-aniversario-hechos-18o-15-noviembre-acuerdo.html>.
- [Pavleska et al., 2023] Pavleska, T., Školkay, A., Zankova, B., Ribeiro, N., y Bechmann, A. (2023). Performance assessment of fact checking organizations - skamba. Disponible en http://skamba.info/wp-content/uploads/2018/02/Performance-assessment-of-fact-checking-organizations_A-critical-overiview.pdf.
- [Pennycook et al., 2020] Pennycook, G., Bear, A., Collins, E. T., y Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management science*, 66(11):4944–4957.
- [PolitiFact, 2023] PolitiFact (2023). Politifact. Disponible en <https://www.politifact.com/>.
- [Post, 2023] Post, T. W. (2023). The washington post. Disponible en <https://www.washingtonpost.com/news/fact-checker/>.
- [Poynter, 2023] Poynter (2023). Training Catalog - Poynter. Disponible en <https://www.poynter.org/shop/>.
- [Reuters, 2023] Reuters (2023). Reuters. Disponible en <https://www.reuters.com/fact-check>.
- [Samuels, 2012] Samuels, M. G. (2012). The filter bubble: What the internet is hiding from you by eli pariser. *InterActions: UCLA Journal of Education and Information Studies*, 8(2).
- [Satuluri et al., 2020] Satuluri, V., Wu, Y., Zheng, X., Qian, Y., Wichers, B., Dai, Q., Tang, G. M., Jiang, J., y Lin, J. (2020). Simclusters: Community-based representations for heterogeneous recommendations at twitter. En *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '20*, p. 3183–3193, New York, NY, USA. Association for Computing Machinery.
- [Sharma et al., 2016] Sharma, A., Jiang, J., Bommanavar, P., Larson, B., y Lin, J. (2016). Graphjet: Real-time content recommendations at twitter. *Proceedings of the VLDB Endowment*, 9(13):1281–1292.

- [Snopes, 2023] Snopes (2023). Snopes. Disponible en <https://www.snopes.com/>.
- [Stanley, 2008] Stanley, L. (2008). Why we should use simpler models if the data allow this: relevance for anova designs in experimental biology. *BMC Physiology*.
- [Sunstein, 2018] Sunstein, C. (2018). *# Republic: Divided democracy in the age of social media*. Princeton university press.
- [Tufekci, 2017] Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.
- [Twitter Blog, 2023] Twitter Blog (2023). Twitter's Recommendation Algorithm. Disponible en https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm.
- [Twitter FAQ, 2023] Twitter FAQ (2023). Preguntas frecuentes para usuarios nuevos. Disponible en <https://help.twitter.com/es/resources/new-user-faq>.
- [Twitter GITHUB, 2023] Twitter GITHUB (2023). Source code for Twitter's Recommendation Algorithm. Disponible en <https://github.com/twitter/the-algorithm>.
- [Twitter Help, 2023] Twitter Help (2023). Información sobre Las Listas de x. Disponible en <https://help.twitter.com/es/using-x/x-lists>.
- [UC, 2023] UC, F. C. (2023). Etapas del proceso constituyente. Disponible en <https://foroconstitucional.uc.cl/cronograma-del-proceso-constituyente/etapas-del-proceso-constituyente/>.
- [Vicario et al., 2016] Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., y Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554-559.

TIEMPOS SCT

Tiempos SCT [horas]	Planificación	Búsqueda Información	Análisis	Desarrollo	Edición	Total
	4	12	9	26	4	55 horas