UNIVERSIDAD
DE LA REPUBLICA
URUGUAY

Te Ahorré Un Click: Caracterización y detección automática de clickbait en español

Gabriel Mordecki

Tutores:
Dr. Ing. Javier Couto
Dr. Ing. Guillermo Moncecchi

Director académico:
Dr. Ing. Guillermo Moncecchi

Tesis de Maestría presentada a la Facultad de Ingeniería de la Universidad de la República y PEDECIBA, como parte de los requisitos necesarios para la obtención del título de Magíster en Informática.

Montevideo – Uruguay
Diciembre de 2024

A las Berrys

Agradecimientos

Agradezco en primer lugar a Javier Couto, gracias a que me convenció; entre Chartreuses, discusiones sobre Chomsky y Cointreaus; de que podía hacer una maestría y (pese a) que era el momento indicado. También a Guillermo Moncecchi que se puso el equipo al hombro cuando se necesitó.

A la persona detrás de aquella cuenta anónima de Twitter que insinuó, allá por el año 2019, que la cuenta de *@TeAhorre1Click* era un bot sencillo de replicar, y de esa forma inspiró la temática de esta tesis.

A mi querida FING, al Pedeciba y toda la educación pública de Uruguay, de la que estoy muy orgulloso.

A la ANII, a través de su portal Timbó, y a Alexandra, mi mejor amiga kazaja, por permitirme acceder a los trabajos necesarios para crear esta tesis.

A mi familia por la curiosidad (esa que hace que tanto me cueste no caer en la trampa del clickbait) y el gusto por la academia, la lectura y el estudio.

A Maite, que está, como tantas veces a lo largo de estos casi dos años, dormida en mis brazos mientras escribo estas palabras y hace que me cueste teclear bien.

Y finalmente a Vero, por todo.

The cure for boredom is curiosity
There is no cure for curiosity

(Atribuida a) Dorothy Parker

RESUMEN

El clickbait es un fenómeno muy frecuente en las noticias en Internet: es una forma de titulación y presentación de las noticias que busca captar la atención de los lectores para generar más tráfico, relegando la función de informar a un segundo plano. Sin embargo, no existe un consenso sobre qué es exactamente; las definiciones suelen confundirlo con otros fenómenos similares como el sensacionalismo e incluso son contradictorias entre sí. En esta tesis argumento que la teoría de la brecha de información de Loewenstein, que explica que la curiosidad surge cuando una persona se hace consciente de una diferencia entre la información que conoce y la que no, es el concepto clave para distinguir el clickbait y propongo una nueva definición: *El clickbait es un método para generar avances de noticias, especialmente online, que consiste en omitir deliberadamente parte de la información con el objetivo de generar curiosidad mediante una brecha de información y así atraer la atención de los lectores y conseguir que hagan click.*

Presento un nuevo enfoque en la creación de conjuntos de datos sobre detección automática de clickbait. Los trabajos anteriores abordan el problema de la subjetividad de la decisión confiando en criterios externos o mediante el promedio de múltiples anotaciones. En su lugar, propongo criterios específicos que abordan varios de los casos límite y logran reducir la subjetividad de la anotación. En esta tesis presento y describo TA1C (por *Te Ahorré Un Click*): el primer conjunto de datos de detección de clickbait disponible en español. Se trata de 3.500 tweets de 18 medios con buena reputación en español, manualmente anotados siguiendo los criterios establecidos alcanzando un acuerdo entre tres anotadores de Fleiss' κ de 0,825.

Utilizando TA1C implementé diversos métodos para la detección de clickbait, logrando un modelo que obtiene un 0.93 de precisión promedio (*average precision*) y 0.87 de F1-score sobre la clase positiva. Presento tres abordajes, el primero basado en aprendizaje automático clásico, con características elaboradas manualmente adaptando al español las presentes en la literatura y aportando algunas nuevas. Un segundo abordaje con modelos de aprendizaje profundo, haciendo *fine-tuning* de *transformers*. Finalmente, un método

híbrido que toma los vectores generados por los *transformers* y los combina con las características de los modelos clásicos para luego aplicar una Regresión Logística y es la que logra los mejores resultados.

Palabras claves:
Clickbait, Defínción de Clickbait, Detección de Clickbait, Procesamiento de Lenguaje Natural, Clasificación de Texto, Conjunto de Datos, Español, Noticias, Redes Sociales.

ABSTRACT

Clickbait is a widespread phenomenon in online news: it is a way of creating headlines and teasers aimed at capturing readers' attention in order to increase traffic, relegating the function of informing to a secondary role. However, there is no consensus at the moment over what exactly it is. Clickbait definitions usually confuse it with other similar phenomena such as sensationalism and are even contradictory. In this thesis, I argue that Loewenstein's information gap theory, which explains that curiosity arises when a person becomes aware of a gap between the information they know and what they do not, is the key concept to distinguish clickbait. I propose a new definition: *Clickbait is a method for generating teasers, especially online, that deliberately omits part of the information with the goal of generating curiosity by creating an information gap, thereby attracting the readers' attention and making them click.*

I introduce a new approach in the automatic detection of clickbait datasets. Previous works address the subjectivity problem of the decision by relying on external criteria or by averaging multiple annotations. Instead, I propose specific criteria that tackles several of the borderline cases reducing the subjectivity of the annotation. In this thesis, I present and describe TA1C (short for *Te Ahorré Un Click*, *Saved You A Click* in Spanish): the first clickbait detection dataset publicly available in Spanish. It consists of 3,500 tweets from 18 reputable Spanish-language media outlets, manually annotated following the established criteria, achieving a Fleiss' κ agreement of 0.825 among three annotators.

Using TA1C, I implemented many automatic clickbait detection methods, achieving a model that obtains an average precision of 0.93 and an F1-score of 0.87 for the positive class. I present three approaches to it. The first one is based on classic machine learning, with manually crafted features created by adapting to Spanish those present in the literature and introducing some new ones. A second approach with deep learning models, by fine-tuning transformers. Finally, a hybrid method that takes embeddings from the transformers, combines them with the classical features and applies a Logistic Regression to

it, achieving the best results.

Keywords:
Clickbait, Clickbait Definition, Clickbait Detection, Natural Language Processing, Text Classification, Dataset, Spanish, News, Social Media.

Tabla de contenidos

1 Introducción **1**

1.1 Uso y relevancia . . . 2
1.2 Rechazo e impacto . . . 4
1.3 Detección automática y conjuntos de datos . . . 8
1.4 Contribuciones . . . 8
1.5 Organización del documento . . . 9

2 Definición de Clickbait **11**

2.1 Definiciones en la literatura . . . 11
2.1.1 Tipo de contenido . . . 14
2.1.2 Decepción . . . 16
2.2 Curiosidad y brecha de información . . . 18
2.3 Creación de brechas de información en los medios . . . 26
2.4 Propuesta de definición . . . 30
2.5 Criterios de identificación de clickbait . . . 31
2.5.1 Definición operativa . . . 33
2.5.2 Tener en cuenta todo el contexto de publicación . . . 33
2.5.3 Insinuar con detalles desconectados genera clickbait . . . 38
2.5.4 La adjetivación intensa puede provocar clickbait . . . 40
2.5.5 No todo el sensacionalismo es clickbait . . . 41
2.5.6 Las preguntas pueden generar clickbait, pero no todas lo son . . . 42
2.5.7 Las expresiones apelativas pueden generar clickbait . . . 43
2.5.8 Algunos editoriales pueden ser clickbait . . . 45
2.5.9 Los titulares incomprensibles no son clickbait . . . 46
2.5.10 Considerar todo el texto del avance . . . 46

3 Clasificación de clickbait: contexto y antecedentes **48**
3.1 Conjuntos de datos disponibles . . . 48
3.1.1 Métodos de recopilación . . . 48
3.1.2 Binario vs continuo . . . 49
3.1.3 Corpus existentes . . . 51
3.2 Detección automática de clickbait . . . 52
3.2.1 Aprendizaje automático clásico . . . 53
3.2.2 Aprendizaje profundo . . . 57

4 Conjuntos de datos **62**
4.1 Recolección de datos . . . 62
4.1.1 Características . . . 63
4.1.2 Twitter/X . . . 63
4.1.3 Medios . . . 65
4.2 TA1C: Descripción del conjunto de datos etiquetado . . . 66
4.3 TA1C-unlabeled: Descripción del corpus sin etiquetar . . . 68

5 Clasificación de clickbait y resultados **71**
5.1 Diseño experimental . . . 71
5.1.1 Preprocesamiento . . . 72
5.1.2 Métricas . . . 73
5.2 Métodos propuestos . . . 74
5.2.1 Baselines . . . 74
5.2.2 Aprendizaje automático clásico . . . 76
5.2.3 Aprendizaje profundo . . . 88
5.2.4 Método híbrido . . . 91
5.3 Resultados . . . 93

6 Conclusión **96**
6.1 Trabajo futuro . . . 97

Referencias bibliográficas **99**

Apéndices **122**
Apéndice 1 Historia del clickbait . . . 123
1.1 Antecedentes . . . 123
1.2 Origen . . . 125

Apéndice 2 Impacto del clickbait 129
2.1 Rechazo popular 129
2.2 Discusión en medios de prensa 130
2.3 Calidad periodística 132
2.4 Función de la titulación 134
2.5 Desinformación 135
2.6 Noticias falsas 140
2.7 Distracción 142
2.8 Fuera de la prensa 142
Apéndice 3 Cuentas incluidas en TA1C 144
Listado de Tweets 149

Capítulo 1

Introducción

Aunque clickbait es una palabra de uso común, popularizada en la segunda mitad de la década del 2010[1] no hay un consenso sobre su definición exacta. Muchas personas lo reconocen y tienen una intuición de qué es clickbait y qué no (Y. Chen y Rubin, 2018). Sin embargo, las definiciones formales suelen diferir entre sí e incluso ser contradictorias.

Es un fenómeno con el que cualquier consumidor de noticias en Internet está familiarizado, con lo que convive diariamente, las personas que conocen el concepto pueden fácilmente identificar las noticias de la Figura 1.1 como clickbait. Pero no está claro cuál es el límite, la o las características exactas que diferencian al clickbait del resto de los titulares. Esa delimitación es además en todos los casos, en definitiva, subjetiva. Por eso los acuerdos inter-anotador sobre clickbait en la literatura son similares a los que se obtienen en otras tareas típicamente subjetivas del procesamiento de lenguaje natural como la detección de ironía y humor. Los resultados de acuerdo entre anotadores reportados van desde un Fleiss' κ[2] de 0.36 hasta 0.73 para anotación binaria (Broscoţeanu y Ionescu, 2023; Chakraborty et al. 2017; Li et al. 2022; Potthast, Gollub, Komlossy et al. 2018; William y Sari, 2020).

Existen, sí, ciertos consensos sobre qué es el clickbait. Se trata de titulares o avances de noticias (incluyendo, además del titular, texto en redes sociales y/o imágenes que lo acompañan) que difieren de los tradicionales en su objetivo y estilo, dejando en segundo plano el propósito de informar a favor de atraer

[1] Google Trends - Clickbait

[2] Medida estadística de la concordancia entre un número fijo de anotadores que etiquetan un conjunto de datos. Cero indica un acuerdo similar al del puro azar y uno indica un acuerdo perfecto. Véase Fleiss' kappa en Wikipedia

la atención (Bazaco et al. 2019; Dor, 2003; García Orosa et al. 2017; Ifantidou, 2009; Oxford English Dictionary, 2024; Palau-Sampio, 2016), a menudo sacrificando la calidad periodística (García Serrano et al. 2019; Molyneux y Coddington, 2020; Palau-Sampio, 2016). Como su nombre sugiere, el clickbait provoca, manipula, actúa de cebo para atraer lectores a clickear (Bazaco et al. 2019; Blom y Hansen, 2015; J. Zheng et al. 2021). También se asocia habitualmente al clickbait con la decepción que genera que el artículo no cumpla con las expectativas generadas por el titular. En este trabajo argumento que la decepción es una característica habitual pero no necesaria del clickbait. En el segundo caso de la Figura 1.1 se muestra como la noticia entrega exactamente el dato que promete su avance, y eso no hace que no sea clickbait.

A lo largo del documento sostengo que la teoría de la brecha de información de Loewenstein (1994) no es solamente relevante al tema, sino que es la diferencia fundamental entre el clickbait y otros fenómenos similares, tales como el sensacionalismo. Loewenstein explica que la curiosidad surge cuando una persona se hace consciente de una diferencia entre la información que conoce y la que no. La curiosidad es el deseo de llenar esa brecha y terminar con la sensación desagradable que genera esa 'privación cognitiva'. Luego de explorar las diferentes definiciones en la literatura, ejemplos concretos y criterios de identificación para crear un conjunto de datos, concluyo que la esencia del clickbait no es el tipo de contenido, ni el lenguaje empleado, ni características léxico-sintácticas, ni la generación de decepción; sino la retención de parte de la información con el objetivo de provocar curiosidad.

1.1. Uso y relevancia

El clickbait es característico de la era digital y de la difusión de noticias por Internet. Está, de todas formas, íntimamente relacionado con otros fenómenos que preceden largamente la invención de Internet como el sensacionalismo, la prensa tabloide y el amarillismo. Son, en definitiva, respuestas al problema de cómo captar fácilmente más atención. La web tiene algunas características particulares que propician el crecimiento de estas prácticas. Por una lado, la capacidad de medir con mucha precisión y en tiempo real las interacciones. Por otra parte, los usuarios están acostumbrados a consumir noticias desde redes sociales, foros o agregadores (Amoedo-Casais et al. 2023; Nic et al. 2023; Pew Research Center, 2022), donde los publicadores deben competir fuertemente

por la atención debido a la sobreabundancia de información.

Figura 1.1: Dos ejemplos de clickbait. Arriba, un clickbait sobre un tema muy importante como la respuesta de un presidente ante la ejecución de un golpe de estado en su contra; pese a la idea de que el clickbait es sólo sobre *noticias blandas*. Abajo, un ejemplo en el que se entrega exactamente lo prometido, como se ve en un recorte del texto del cuerpo de la noticia; pese a la idea de que el clickbait se acompaña siempre de una noticia que decepciona y no entrega lo prometido.

Por estas razones el clickbait nació en medios nativos digitales. Pero los medios tradicionales se adaptaron a su uso debido a la efectividad que tiene en incrementar el tráfico, acompañando la migración de los lectores desde el papel al formato digital. La competencia de los medios más prestigiosos del mundo pasó a ser en Internet e incluyendo a los nuevos medios digitales (BuzzFeed News, 2014; Kilgo y Sinta, 2016; Palau-Sampio, 2016). El Apéndice 1 profundiza en la historia del clickbait desde sus antecedentes hasta la evolución a lo que es hoy en día.

Actualmente el clickbait tiene una presencia notoria, incluso abrumadora para algunos, en las noticias en la web a nivel mundial. Potthast, Gollub, Komlossy et al. (2018) compilaron más de 38 mil tweets de 27 medios en inglés y casi 45 % de los titulares fueron etiquetados como clickbait. García Orosa et al. (2017) encuentran que entre 28 de los medios más leídos de Europa en el 60 % de los titulares «priman las sensaciones y emociones frente a la información» y «cerca de dos terceras partes de los titulares no tienen sentido en sí mismos ni responde a los criterios de noticiabilidad». Rony et al. (2017) afirman que el 33.5 % de casi 900 mil titulares de «prensa convencional» contienen clickbait. Específicamente en trabajos hechos sobre noticias en español, Bazaco et al. (2019) encuentra que 38.2 % de los tweets del periódico El Mundo y 53 % de los de El País de España utilizan «recursos gancho». Finalmente, en TA1C, el conjunto de datos creado en este trabajo bajo los criterios especificados en el Capítulo 2, encuentro que

el 28.6 % de las noticias compartidas por medios de prensa en español utilizan clickbait. Es importante notar que los números de los trabajos no son siempre comparables entre sí debido a las diferencias en la definición de clickbait. Sin embargo, está claro que que se mida como se mida la presencia es notoria y considerable.

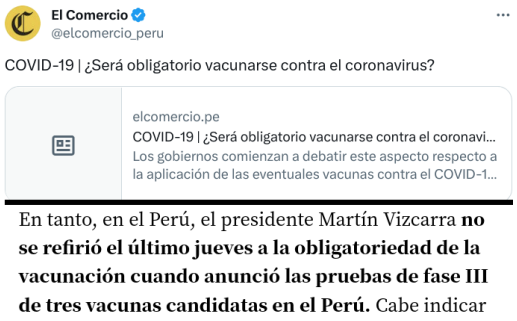

Figura 1.2: Ejemplo de clickbait sobre un tema importante. Aquí, el cuerpo de la noticia no entrega lo prometido, una respuesta sobre la obligatoriedad de la vacuna anti Covid-19 en Perú, sino que se limita a informar que el presidente no se refirió al tema.

El hecho de que medios tradicionales y prestigiosos hayan recurrido al clickbait generó que su asociación inicial a noticias blandas, poco importantes, haya desaparecido en la actualidad. Hay muchos ejemplos de noticias duras presentadas en forma de clickbait como la que muestran Figura 1.2 sobre la vacunación contra el Covid-19 o Figura 1.1 sobre las palabras de un presidente derrocado por un golpe de estado. Este hecho acentúa el impacto generado: el clickbait se coloca como barrera para acceder a información relevante e importante.

1.2. Rechazo e impacto

No es sólo por su presencia masiva que el clickbait vale la pena como objeto de estudio. Los efectos de su uso y proliferación en los medios de comunicación han sido discutidos y analizados partiendo desde diversas perspectivas, desde el rechazo de los lectores hasta su influencia en la calidad periodística y la desinformación.

Pese a que funciona para generar más clicks (no existiría de lo contrario), los lectores dicen no preferirlo (Karadzhov et al. 2017; Mukherjee et al. 2019; Scacco, Joshua and Muddiman, Ashley, 2016). Facebook, aunque sin explicar cómo llevó a cabo la encuesta, reportó en 2014 que el 80 % de sus usuarios prefería ver menos clickbait (El-Arini, Khalid and Tang, Joyce, 2014). Su presencia masiva logra distraernos de lo que más nos importa e interesa (Chakraborty et al. 2016; Mark, Gloria, 2014; Paasonen, 2016). Esta discrepancia entre la

efectividad que tiene el clickbait y el rechazo que produce se explica, desde mi punto de vista, porque el clickbait no genera en el lector un interés genuino, un deseo real de conocer la información prometida por el titular. Por lo contrario, genera una incomodidad, una molestia; y un impulso imposible de controlar y difícil de combatir de terminar con ella. Ese deseo por llenar la brecha de información logra desviar nuestra atención hacia noticias que no nos hubieran despertado interés de otra forma. La curiosidad es distinta que el interés (Shin y Kim, 2019; Tang et al. 2022). Mientras que el interés es un sentimiento «estable y sustentable» basado en las «preferencias personales», la curiosidad es un sentimiento activo y «una emoción que usualmente implica una fuerte necesidad de actuar»[1] (Aslan et al. 2021). Esa curiosidad generada crea expectativas de que la noticia sea más relevante de lo que es en realidad (Scott, 2021), creando la decepción. Muchas veces parece hecho de forma intencional. Por ejemplo, es el caso mostrado en la Figura 1.2 donde el titular promete contestar si la vacuna de Covid va a ser obligatoria, y el texto de la noticia sólo dice que el presidente «no se refirió» al asunto. Como afirma Blom y Hansen (2015) esta práctica se encuentra «más cerca de la manipulación que de la estimulación».

Este tipo de manipulación puede no ser intencional, y ser una consecuencia de pensar únicamente en aumentar el tráfico y los ingresos. Seguramente en ocasiones es a propósito pero se intenta ocultar la intención. Pero hay ejemplos en los que es explícita. En su libro «Memecracia: Los virales que nos gobiernan» Delia Rodríguez (2013), periodista de Huffington Post y del diario El País (España) entre otros, fundadora y directora de la sección Verne, explica: «Conozco bien cómo elegir una foto de Scarlett Johansson para que pinchéis antes de que os deis cuenta, cómo redactar un tuit que os lleve a un lugar donde os tendré entretenidos durante más tiempo del que deberíais o qué declaración de un político os enfurecerá tanto como para comentar en mi noticia. He aprendido a explotar las debilidades e intereses naturales de los cerebros para crear contenidos difíciles de resistir».

A veces esta discrepancia entre titular y artículo puede llevar a casos de desinformación. Como ejemplo, en un caso extremo, la figura Figura 1.3 muestra una noticia donde el titular habla de un voluntario que murió en los ensayos de una vacuna de Covid. Sin embargo, es necesario leer el cuerpo de la noticia para enterarse de que el voluntario estaba en el grupo de control, por lo que

[1] Traducción propia. Este criterio se aplica a todas las citas de textos originales en inglés.

Figura 1.4: Dos ejemplos de clickbait donde se sugiere información que la noticia desmiente, en una práctica desinformativa. A la izquierda, es necesario leer la nota para entender que se sabe que la hidroxicloroquina no es eficiente para combatir el virus de Covid-19. A la derecha, hay que leer toda la noticia para enterarse que a la modelo la mató el novio, y nada tiene que ver con lo que le dijo a Putin.

recibió un placebo.

Las dos funciones tradicionales de los titulares son resumir el artículo, ya sea mediante una síntesis o resaltando el punto principal, y llamar la atención (Dor, 2003; Kuiken et al. 2017b). El clickbait es una de las técnicas que desbalancean la prioridad hacia la segunda función, descuidando la primera. En ese contexto suele también habitualmente asociarse su uso con un descenso en la calidad periodística (Bazaco et al. 2019; Frampton, 2015; García Orosa et al. 2017; García Serrano et al. 2019; Gunnels, 2017; Lazauskas et al. 2018; Molyneux y Coddington, 2020; Mukherjee et al. 2019) y la generación de una pérdida de credibilidad (Hurst, 2016; Pengnate, 2016; Pengnate et al. 2018; Zhang et al. 2020).

Figura 1.3: Ejemplo de un caso extremo de cuando el clickbait se transforma en desinformación (no forma parte de *TA1C*). El título oculta el por qué de la muerte y sólo leyendo la noticia se puede entender que la muerte no tiene ninguna relación con la vacuna, ya que el fallecido recibió placebo en los ensayos.

Los títulos y avances de noticias deben reflejar la información que se trans-

mite. En ocasiones, intentando generar clickbait, se omite información crucial o se genera una ambigüedad. Estas diferencias pueden parecer sutiles pero sus efectos no lo son, porque los titulares terminan implicando información distinta a la que cuenta la noticia (Chesney et al. 2017; Mukherjee et al. 2019; Silverman, 2015). Por un lado, las personas leen habitualmente únicamente los titulares de muchas noticias. Es lo que se comparte en redes sociales y lo que se ve en la página principal de los medios y la información que normalmente tenemos para decidir si leer o no el artículo. Y antes de leer el texto de la noticia no es posible saber si el titular se corresponde. Pero además el titular, por ser lo primero que se lee, genera una primera impresión que sesga nuestra memoria, interpretación y entendimiento y puede persistir incluso luego de leer el artículo completo (Ecker et al. 2014; Hurst, 2016; Silverman, 2015; Wagner-Egger y Gygax, 2018; Wegner et al. 1981). La imagen Figura 1.4 muestra dos ejemplos en los que el titular sugiere algo que la noticia desmiente. En el primer caso el título vincula la hidroxicloroquina y el combate al Covid-19, pero el artículo advierte que «la OMS y Francia frenaron el uso de la hidroxicloroquina para tratar la covid-19 a la luz de un amplio estudio que advierte sobre sus peligrosas contraindicaciones». El segundo caso muestra cómo se puede generar confusión dando exclusivamente información cierta, pero omitiendo parte del contexto crucial para entender la noticia. En este caso, que los dos hechos no tienen ninguna relación entre sí y se informa que fue su novio quién la mató y ya confesó el delito.

Estos efectos son a veces generados a propósito porque la noticia que se sugiere llama más la atención que la real, o porque la propia ambigüedad es la que genera curiosidad. Es imposible juzgar la intención de quien escribe. Sin embargo, el engaño se produce de todas formas. Muchos, incluso, consideran al clickbait como una categoría de noticias falsas (Baptista y Gradim, 2020; Bourgonje et al. 2017; Edson C. Tandoc Jr. y Ling, 2018; Ghanem et al. 2020; Karadzhov et al. 2017; Molina et al. 2021; Mourão y Robertson, 2019; Park et al. 2020; Volkova et al. 2017; Wikipedia, 2024b; Zannettou et al. 2019; X. Zhou et al. 2020), aunque en mi opinión no debería serlo: es una herramienta de difusión e incremento de tráfico que se puede utilizar tanto para noticias reales como falsas.

El Apéndice 2 profundiza sobre el impacto del clickbait mediante un repaso del rechazo popular que genera, su función en la titulación y la pérdida de calidad periodística y la desinformación, su relación con las noticias falsas y

cómo genera distracción en los lectores.

1.3. Detección automática y conjuntos de datos

La tarea de detectar el clickbait de forma automática ha concitado atención en los últimos años, principalmente luego del *Clickbait Challenge* en 2017 (Potthast, Gollub, Hagen y Stein, 2018). La proliferación e impactos del clickbait lo hacen un buen motivo de estudio y su detección es un problema difícil de resolver.

Un desafío extra en la resolución radica en la diversidad de definiciones y criterios existentes sobre qué es clickbait y qué no. Además, son populares distintos métodos de recolectar y etiquetar noticias que generan conjuntos de datos de distinta calidad. Todo esto genera una dificultad en la comparación entre diferentes investigaciones. En este trabajo propongo una definición y criterios específicos que cubren varios de los casos borde reduciendo la subjetividad de la decisión al mínimo posible. Este criterio permite generar un conjunto de datos con un acuerdo entre anotadores Fleiss' κ de 0.825.

Además, pese a la gran cantidad de trabajos y conjuntos de datos disponibles (ver Tabla 3.1), no hay ninguno, hasta donde tengo conocimiento, en español. Este trabajo presenta *TA1C*, el primer conjunto de datos abierto para detección de clickbait en español.

Sobre esos datos, experimenté con diversas técnicas de aprendizaje automático logrando un modelo que obtiene un 0.93 de precisión promedio (*average precision*) y 0.87 de F1-score sobre la clase positiva. Lo hice tomando ideas de los métodos más populares y exitosos del estado del arte. Por un lado, de trabajos específicamente de detección de clickbait de los que compilé, implementé y traduje al español de ser necesario las características usadas. Por otro lado de la clasificación de textos en general, específicamente de los métodos que pueden ser replicados para usar en ese problema.

1.4. Contribuciones

En resumen, las contribuciones del trabajo son:

1. Un estudio del clickbait como concepto y una propuesta de definición.

2. Criterios claros de identificación de clickbait que contextualizan y ejemplifican la definición para maximizar la consistencia en un problema subjetivo.
3. La publicación de TA1C, el primer conjunto de datos anotado manualmente para la detección de clickbait en español. Se encuentra disponible en github.
4. La publicación de un conjunto de datos de más de medio millón de tweets con noticias en español, que puede servir de recurso para muchos problemas de procesamiento de lenguaje natural. Se encuentra disponible en github.
5. Un repaso del estado del arte en clasificación de clickbait incluyendo los conjuntos de datos disponibles, las características (*features*), modelos y métodos más utilizados y sus resultados.
6. Métodos de clasificación que obtienen un 0.87 de F1-score y 0.93 de *Average Precision* en detección de clickbait en español, usando el conjunto de datos TA1C.

1.5. Organización del documento

El Capítulo 2 repasa las diferentes definiciones de clickbait existentes en la literatura y los conceptos de la curiosidad y la brecha de información y su relación con el clickbait. Luego, contiene una propuesta de definición de clickbait y criterios concretos de identificación que permiten la anotación.

El Capítulo 3 repasa el estado del arte de la clasificación de clickbait y los conjuntos de datos existentes.

En el Capítulo 4 se describen los conjuntos de datos construidos y publicados en este trabajo: uno sin etiquetar y otro etiquetado para detección de clickbait, junto con el método de recolección y las decisiones tomadas en su construcción.

El Capítulo 5 contiene la descripción de los métodos utilizados para la clasificación de clickbait y los resultados obtenidos.

El Capítulo 6 presenta conclusiones y propone lineamientos para el trabajo futuro.

Finalmente, el Apéndice 1 y Apéndice 2 profundizan en la historia y el impacto del clickbait respectivamente y el Apéndice 3 lista todas las cuentas

utilizadas para generar el conjunto de datos *TA1C*.

Capítulo 2

Definición de Clickbait

La palabra clickbait se utilizó por primera vez en el año 2006, en un blogpost titulado «Definition of Click Bait»[1], y casi como un sinónimo de *linkbait*, palabra a la que luego sustituyó. El término se popularizó en la mitad de la década del 2010[2] y hoy en día es comúnmente utilizado en los medios de comunicación, la comunidad científica y hasta fue incluido en varios diccionarios. Sin embargo, la definición de clickbait no es un consenso (Bazaco et al. 2019; García Orosa et al. 2017; Kuiken et al. 2017a; Mormol, 2019; Potthast, Gollub, Komlossy et al. 2018).

Este capítulo repasa las definiciones existentes en la literatura con sus similitudes y diferencias, discutiendo su idoneidad. Luego repasa el fenómeno de la curiosidad; clave para entender y, desde mi punto de vista, para definir el clickbait; y presenta una propuesta de definición. Finalmente expone ejemplos que ilustran las dificultades que existen para delimitar el alcance del clickbait y criterios para determinar ciertos casos borde, con el objetivo de disminuir al mínimo posible la inevitable subjetividad que generan algunas de esas decisiones.

2.1. Definiciones en la literatura

No existe un acuerdo en la definición de clickbait. Existen variadas propuestas con características distintas e incluso en algunos casos contradictorias. Como explican Y. Chen y Rubin (2018): «Pese a que la gente entiende intui-

[1] El blogpost fue escrito por Jay Geiger y actualmente no se encuentra disponible, su URL era http://www.jaygeiger.com/index.php/2006/12/01/definition-of-click-bait/

[2] https://trends.google.com/trends/explore?date=all&q=clickbait

tivamente qué es el clickbait, no hay una descripción formal del concepto. Los ejemplos que definitivamente son o no son clickbait no son ambiguos en general, pero hay desacuerdo y falta de certeza en el medio.» En (Potthast, Gollub, Komlossy et al. 2018) etiquetan un dataset de tweets de noticias y obtienen un acuerdo de Fleiss' κ de 0.36 para dos clases, muy similar al 0.35 de (Potthast et al. 2016). De hecho, proponen etiquetar en cuatro clases en lugar de dos para lidiar con los casos borde. En otros trabajos como (Li et al. 2022) se reporta un 72.6 % de acuerdo, mostrando que existe acuerdo entre los anotadores pero confirmando ciertas discrepancias entre ellos, en (Chakraborty et al. 2017) anotan 200 tweets y llegan a un Fleiss' κ de 0.72 y en (William y Sari, 2020) un Fleiss' κ de 0.42 con tres anotadores. El nivel de acuerdo entre anotadores es similar e incluso inferior al encontrado en variados trabajos de detección de características subjetivas del lenguaje como la ironía (Zeng y Li, 2022) y el humor (Chiruzzo et al. 2021). En (D. Molina et al. 2021) se comparan cuatro clasificadores automáticos en las mismas noticias y el acuerdo no llega a la mitad.

Es difícil realizar una lista exhaustiva de características que definan al clickbait por extensión y delimiten claramente las diferencias con el resto de los titulares (Potthast, Gollub, Komlossy et al. 2018). Además, si bien se pueden listar las propiedades más frecuentes, la lista se debería actualizar permanentemente, debido a que los medios buscan en cada momento nuevas formas de generarlo.

Muchas de las definiciones existentes intentan diferenciar al clickbait de los titulares *tradicionales* del periodismo. Normalmente, los titulares buscan informar, además de captar la atención, resumiendo la noticia o resaltando un aspecto de ella (Dor, 2003; Ifantidou, 2009). El clickbait, por el contrario, tiene el objetivo de que la mayor cantidad posible de personas hagan click. Con ese propósito, los titulares clickbait dejan de lado muchos de los principios característicos de la titulación en la prensa tradicional.

Por ejemplo, García Orosa et al. (2017) explica que «el titular cebo tendría como objetivo principal la comercialización o difusión de la información, mientras que el titular periodístico mantendría el rol principal de informar a los usuarios» y define clickbait como «el titular de una noticia que no responde a los criterios periodísticos tradicionales y cuyo objetivo final es mantener al receptor en la página el mayor tiempo posible, no informar.»

La definición de Bazaco et al. (2019) deja claro el concepto de que se pierde

calidad informativa en pos de una ganancia comercial: «un fenómeno comunicativo dinámico que recurre a contenidos pseudo informativos elaborados mediante estrategias de economía de la atención que se aproximan al sensacionalismo y al infoentretenimiento. Sus mensajes se diseñan como gancho, con fines más propios del marketing, a través de recursos tales como titulares llamativos, apelación directa al lector y contenidos superficiales o exagerados.»

Potthast, Gollub, Komlossy et al. (2018), que también destacan lo difícil de definir el clickbait, al que califican de «concepto elusivo», dan 4 definiciones similares entre sí, cada una basada en un punto de vista:

Aparición Clickbait es el resultado de optimizar los mensajes de avance para maximizar los clicks que se hacen.

Intención Mensajes de avance cuyos autores pretenden atraer a tantas personas como sea posible a una página web, sin tener en cuenta el público objetivo del contenido.

Efecto Mensajes de avance que manipulan a sus lectores para que hagan clic en un enlace.

Percepción Mensajes de avance percibidos por (algunos) lectores como cebo para hacer clic en un enlace.

Se puede notar que el concepto al que se refieren los trabajos es el mismo, algo distingue al clickbait del resto de los titulares. Sin embargo, no es fácil explicar exactamente qué, y en este lugar es cuando comienzan a verse las diferencias. Muchos se refieren a una intención, una *manipulación*, una búsqueda explícita de engañar (Bazaco et al. 2019; Y. Chen y Rubin, 2018; García Orosa et al. 2017; Oxford English Dictionary, 2024). Otros, a características del texto en sí, y son muchas las que se describen. Pero las definiciones de la literatura se distinguen principalmente en dos nociones:

1. Si el concepto se refiere únicamente al titular (eventualmente acompañado de una imagen o un pequeño texto) y su forma de presentación, o también al tipo de contenido.
2. Si el clickbait incluye el hecho de decepcionar, prometiendo algo que al final no entrega.

En las siguientes secciones se analizan cada una de estas dos discusiones.

2.1.1. Tipo de contenido

Una discrepancia entre las variadas definiciones disponibles de clickbait es si alude únicamente al estilo literario del titular, o en su caso al mensaje en una red social que puede ir acompañado con un texto secundario y/o una imagen; o también al tipo y los temas del contenido de las noticias.

La literatura se divide entre una noción más restrictiva, que restringe el alcance del clickbait como estrategia a la formulación de titulares (y el contenido que los acompaña), y otra «más general e inclusiva que alberga distintas fórmulas para captar la atención y conseguir incrementar el tráfico a una página web.» (Bazaco et al. 2019)

Por ejemplo, el *Oxford English Dictionary (2024)* define clickbait como «contenido cuyo principal propósito es atraer la atención y alentar a los visitantes a clickear en un link a una página en particular». Indurthi y Oota (2017) como «un término peyorativo que describe el contenido web destinado a generar ingresos por publicidad, especialmente a expensas de la calidad o la precisión, y se apoya en titulares sensacionalistas o imágenes llamativas para atraer clicks y alentar el reenvío del material a través de redes sociales». Mormol (2019) como «un término paraguas para un número de técnicas usadas para atraer la atención y despertar curiosidad.» También el *Merriam-Webster dictionary (2024)*, una definición usada frecuentemente en la literatura (Mukherjee et al. 2022; Munger et al. 2020; Rajput, 2019), lo describe como «algo (como un titular) diseñado para hacer a los lectores querer hacer click en un link especialmente cuando lleva a contenido de dudoso valor o interés».

Sin embargo, hay muchos ejemplos de noticias serias, formales, importantes, que serían sin dudas calificadas como hard news presentadas en sus titulares en forma de clickbait, tales como los de la Figura 1.1 o

> (1) *Elecciones en Estados Unidos: por qué el recuento de votos podría demorarse más de los normal*

Algunos trabajos estudian este fenómeno, como Araujo et al. (2021) donde encuentran que el 31 % de las *noticias duras* en medios nativos digitales españoles son clasificadas como clickbait. El hecho de que muchos medios tradicionales en sus ediciones web utilicen esta estrategia para competir con los nativos digitales hace claro que el clickbait no es exclusivo de las *noticias blandas*.

Es lógico que clickbait y conceptos como *sensacionalismo*, *tabloidización* y *contenido blando* se confundan porque están estrechamente relacionados: son

la respuesta a una pregunta muy habitual en los medios: *¿cómo se genera más tráfico fácilmente?* Es claro al repasar los orígenes del clickbait que siempre hubo una conexión. Los primeros sitios en aplicar el clickbait lo hacían principalmente con noticias sensacionalistas.

Numerosos trabajos recogen la fuerte correlación entre la forma de atracción y el contenido sensacionalista. Palau-Sampio (2016) encuentra que el diario El País en sus secciones creadas para competir en lo digital usa títulos clickbait además de «contenidos anecdóticos y curiosos o de estilos de vida, mientras que las llamadas hard news están ausentes». Kilgo y Sinta (2016) compara las noticias de dos periódicos con las de dos medios nativos digitales y concluye que en los digitales hay más clickbait (catáforas, listículos), personalización y noticias blandas.

Incluso Y. Chen y Rubin (2018) afirman que quienes clasificaron, «en términos de contenido, fueron más proclives a clasificar titulares de soft news como entretenimiento, deporte e historias poco convencionales como clickbait» confirmando que no sólo aparecen comúnmente juntos sino que están asociados en la percepción popular.

A pesar de la fuerte correlación, es importante distinguir al clickbait como, a lo sumo, una parte del sensacionalismo y definitivamente no como el mismo fenómeno. Kilgo y Sinta (2016) señalan que «la forma sensacionalista debe verse como un elemento de empaque, y considerada una herramienta de distribución de información ... [Hay artículos que] no están alineados con las evaluaciones negativas del sensacionalismo pese a que están presentados en una forma sensacionalista.»

En el mismo sentido Molek-Kozakowska (2013) afirma que «es importante no confundir dos aspectos del sensacionalismo: la elección sensacionalista de los tópicos que se cubren (escándalos, crímen, sexo), y una manera sensacionalista de empaquetar la información para hacerla parecer más relevante o interesante.» El clickbait es parte del segundo aspecto, el de la presentación de la noticia, y dentro de él únicamente referido al titular o el avance (texto e imagen que lo acompaña en redes sociales o una página web), pero no al contenido y la temática.

Distinguir al clickbait del sensacionalismo mediante la restricción de su definición a las formas de escritura y presentación de los titulares y/o avances de las noticias permite enriquecer la variedad de los conceptos con los que se cuenta para estudiar el tema. Sensacionalismo, tabloidización y hasta in-

foentretenimiento son términos muy conocidos que carecen de una definición concreta (Bird, 2009; Kilgo y Sinta, 2016; Molek-Kozakowska, 2013; Palau-Sampio, 2016), que describen prácticamente lo mismo y que refieren a estrategias que se utilizan desde hace décadas. Definir clickbait como un tipo de contenido sólo agregaría una palabra más para llamarle a un mismo concepto.

Al pensar en el clickbait como una forma relativamente novedosa, y nacida en Internet, de captar la atención y apelar a las emociones y la curiosidad para empaquetar el avance de una noticia, incitando al usuario a clickear; se adquiere un nuevo concepto y se enriquece el lenguaje.

2.1.2. Decepción

La otra gran corriente de disimilitud en las definiciones de clickbait se refiere al hecho de que la noticia no satisfaga las expectativas generadas en el titular. Muchas definiciones exigen que el artículo no cumpla con lo prometido, generando una decepción al lector; mientras que otras no lo consideran una condición necesaria.

Por ejemplo, Biyani et al. (2016) define clickbait como «artículos con títulos engañosos, que exageran el contenido de la página de destino. Su objetivo es atraer al usuario a hacer click en el título para monetizar en la página, cuyo contenido es usualmente de mala calidad». Según Rony et al. (2017) se refiere a «una forma de contenido web que emplea fórmulas de escritura y técnicas lingüísticas en los titulares para engañar a los lectores a clickear los links, pero no entrega lo esperado». Chakraborty et al. (2016) afirma que «el clickbait usualmente no cumple las expectativas de los lectores y los deja decepcionados». También Jonah Berger (2014), en una columna en el New York Times opina que «el clickbait es malo porque entrega menos de lo que promete. En contenido generalmente no está a la altura de lo que se jacta ser y al final quedamos decepcionados.» Además, Wikipedia (2024a) lo define como «una forma de publicidad engañosa que usa el texto de un link para generar atención y atraer usuarios a seguirlo y leer, ver o escuchar el contenido con una característica definitoria de ser engañoso, típicamente sensacionalista o falaz».

Nuevamente la característica de decepcionar, al igual que con el contenido sensacionalista, está notoriamente correlacionada con el clickbait. Y otra vez la relación viene del hecho de que contestan a la misma pregunta, en este caso: *¿qué avance generará que más gente haga click en este link?* Entre las opciones

con las que cuenta quién escribe el titular (y lo que lo acompaña) está la de, en caso de que la noticia no sea lo suficientemente atractiva, exagerar, resaltar notoriamente un aspecto poco importante, o incluso inventar o sugerir algo que no es real para hacer que su información parezca más atractiva.

Pero también existe la posibilidad de apelar a las emociones, a la intriga, a un estilo de escritura que llame la atención o a cualquier estrategia que incite al lector a querer entrar a la noticia sin engañarlo. Por ejemplo, cuando un titular en lugar de presentar un dato lo anuncia en forma de pregunta, provoca curiosidad e induce a satisfacerla leyendo el artículo. Muchas veces no puede decirse que sea decepcionante: el artículo da exactamente lo que prometía, la respuesta a la pregunta. Sin embargo, el titular será sin dudas clasificado como clickbait por cualquiera por la forma en la que se presenta. Es el caso del ejemplo presentado en la Figura 1.2, y también de

(2) *El bebé de Zaira Nara ya tiene nombre y seguro no es el que te imaginás*

En ocasiones incluso la respuesta en el artículo es acompañada de una larga explicación de buena calidad periodística como en

(3) *África erradica la polio: cómo consiguió este histórico hito*

pero la presentación es sin dudas un clickbait.

Bazaco et al. (2019) coloca la exageración y la información incompleta, variables que pueden llevar a la decepción del lector, como dos de las seis principales categorías de fórmulas utilizadas en el clickbait. En otras palabras, lo coloca como una característica frecuente pero no indispensable. En esa misma línea García Orosa et al. (2017) genera ocho categorías de análisis para definir si un titular es clickbait y dos de ellas son «el titular se identifica con el texto de la noticia» y «el titular hace referencia a la idea principal de la historia». Las otras seis no se refieren a la concordancia entre título y artículo.

Incluso muchos de los artículos que presentan definiciones de clickbait condicionadas por la decepción intentan una clasificación de titulares como clickbait sin tomar en cuenta el texto del artículo, lo cual sería imposible de determinar siguiendo las definiciones que los propios artículos proporcionan (Chakraborty et al. 2016; Rony et al. 2017).

El límite es fino porque, continuando con el caso anterior, es posible que si el titular hubiera presentado ese dato en particular, si el título fuera *Zaira*

Nara le pondrá Viggo Silvestre a su hijo muchos más lectores habrían elegido pasarlo de largo. Esa es la razón que lleva a los medios a elegir este tipo de titulares sobre los tradicionales. La curiosidad hace que el contenido parezca más relevante y, por lo tanto, el lector termina consumiendo una noticia que no habría clasificado, de tener un titular tradicional, entre sus elegidas para leer (Scott, 2021). Como explican Munger et al. (2020) eso genera una sensación «como el arrepentimiento, si el consumidor de clickbait parara a pensar sobre su decisión de clickear probablemente no lo haría.» No obstante, la decepción se debe en estos casos al hecho de haber invertido tiempo en leer esa noticia por sobre otras y no sobre una discrepancia entre lo prometido y lo entregado.

Al igual que en la discusión sobre el tipo de contenido, el clickbait y la decepción son fenómenos se dan frecuentemente juntos pero esa correlación no quiere decir que sean lo mismo. Puede considerarse que el hecho de prometer algo que no se cumple es una condición suficiente para el clickbait, pero no debe considerarse una condición necesaria.

Nuevamente, también, la asociación tiene raíces en la propia historia de la creación del clickbait. En este caso, sin embargo, la diferenciación es anticuada y eso se hace cada vez más notorio. Por un lado, debido a que Facebook y luego otros sitios lograron controlar este tipo de clickbait decepcionante mediante la penalización a los sitios en los que los lectores cerraban rápidamente la página que abrían (Munger et al. 2020). Simultáneamente, la apropiación de la técnica del clickbait por parte de los medios tradicionales aumentó la cantidad de noticias de buena calidad con titulares que cumplen varias de las otras características.

2.2. Curiosidad y brecha de información

Hasta ahora este capítulo trató sobre características que no deberían considerarse parte de la definición de clickbait. Pero entonces, si no se trata del tipo de contenido ni de la decepción qué genera, ¿qué distingue al clickbait de el resto de los titulares? ¿Cómo se puede definir? Para responder estas preguntas, hace falta entender el fenómeno de la curiosidad y sus efectos desde un punto de vista psicológico. Particularmente, sobre el concepto de la brecha de información y su relación con el clickbait.

Al momento de escribir esta tesis no es mucha la literatura específica sobre la razón psicológica que hace que el clickbait dé resultado, con (Li et al. 2022;

Scott, 2021; Venneti y Alam, 2018) y la mención al clickbait en la conclusión de (Golman et al. 2022) como algunos ejemplos. Sin embargo, sí se puede entender este fenómeno aplicando las teorías psicológicas sobre la búsqueda de información, la tentación, la atención y principalmente la curiosidad. Si bien la gran mayoría de los trabajos no lo mencionan o tratan directamente, se pueden aplicar al clickbait los conocimientos en estas áreas.

La curiosidad atrajo el interés de la psicología desde hace más de un siglo, principalmente con el estudio de su función como motivadora del aprendizaje. Sin embargo, sigue siendo un fenómeno poco entendido (Kidd y Hayden, 2015). La investigación específicamente sobre la curiosidad se intensificó en los años 50 y 60 cuando se realizaron variados estudios en animales y Berlyne (1954, 1966) distinguió entre la curiosidad perceptual, la que nos lleva a buscar nuevos estímulos y disminuye con la exposición continua, y la curiosidad epistémica, que nos hace buscar la adquisición de nuevo conocimiento. Además, distinguió la curiosidad específica, el deseo por una información particular; de la diversiva, la referida al deseo por cualquier estimulación perceptual o cognitiva que sucede por ejemplo como consecuencia del aburrimiento.

Luego, en la década de los 70, las investigaciones se centraron en medir la curiosidad, aunque con poco éxito (Loewenstein, 1994). Fue en 1994 cuando Loewenstein (1994) generó una nueva interpretación de la curiosidad que se convirtió en el estado del arte, que fue luego respaldada con diferentes estudios en animales y humanos (Jepma et al. 2012; Jordan Litman y Russon, 2005; Kang et al. 2009; Kruger y Evans, 2009; Lau et al. 2020), y que se ajusta perfectamente como explicación del clickbait. Loewenstein describió la curiosidad como una privación cognitiva inducida que surge de la percepción de una brecha entre el conocimiento y la comprensión. La teoría se basa en el concepto de la brecha de conocimiento, la diferencia entre lo que uno sabe y lo que quiere saber de un tema específico. Cuando una persona es consciente y presta atención a una brecha de conocimiento, surge la curiosidad: el deseo de llenar esa brecha mediante la obtención de la información faltante.

Uno de los experimentos que apoyan esta teoría es el realizado por Kang et al. (2009) donde encuentran que la curiosidad con respecto a la respuesta de preguntas triviales es una función en forma de U invertida de la seguridad de la persona en la respuesta, como se puede ver en la Figura 2.1. La curiosidad es poca ante la confianza absoluta de conocer la respuesta, pero también ante el desconocimiento total del tema de la pregunta y es máxima en un punto

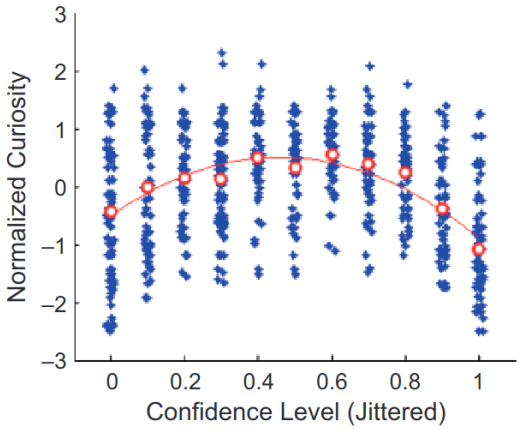

Figura 2.1: Muestra de la U invertida que forma la curiosidad de las personas en función de la seguridad que tienen en conocer una información. La curiosidad es más alta cuando la confianza en la respuesta a una pregunta es media, y más baja cuando las personas están seguras de saber o no saberla contestar. Imagen tomada de (Kang et al. 2009)

medio: cuando es más amplia la brecha de información.

En Golman y Loewenstein (2015) se explora aún más esta teoría poniendo énfasis en el hecho de que la atención de las personas es limitada. Si bien uno podría hacerse infinitas preguntas, sólo es consciente de una cantidad finita de ellas en un momento dado, las que tienen la atención en ese momento y que definen como *activadas*. La curiosidad se genera cuando existe una brecha de información en una pregunta *activada*, a la que se le está prestando atención. La activación de una pregunta depende de tres factores: importancia, la utilidad de conocer la respuesta; notoriedad, si surge en el contexto, por ejemplo si se escucha a alguien hablar del tema; y sorpresa, que es lo que sucede cuando la adquisición de un nuevo conocimiento inesperado genera que se modifiquen las probabilidad asignadas a priori de las respuestas a preguntas relacionadas.

En el caso del clickbait, se producen entonces dos efectos al mismo tiempo: la activación del tema de la noticia y la creación de la brecha de curiosidad. Por un lado, el tema de la noticia se activa, al igual que sucedería con cualquier titular, porque cobra notoriedad que deriva en atención. Con la titulación tradicional, el lector normalmente decide si leer la noticia según la importancia que le asigne al tema del que se trata. El sensacionalismo apela a la sorpresa y por eso recurre a la exageración, los eventos inusuales y el excesivo destaque de aspectos secundarios. El clickbait, por otro lado, recurre a dar la suficiente información, pero también ocultar lo necesario para generar una brecha de

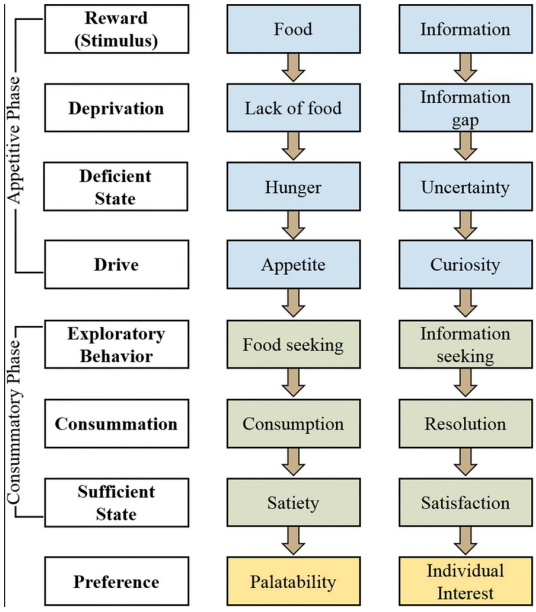

Figura 2.2: Comparación entre el apetito y la curiosidad, tomada de (Shin y Kim, 2019). En esta comparación, la brecha de información toma el lugar de *privación*, al igual que la falta de comida para el hambre.

información que genere que el lector sienta la curiosidad, y por lo tanto el deseo, de leer la noticia para saciar la necesidad de llenar la brecha.

Una de las analogías más utilizadas para entender la curiosidad es su comparación con el apetito. Ya los filósofos Jeremy Bentham y Edmund Burke se referían al «apetito de la curiosidad» (Loewenstein, 1994) y también Loewenstein (1994) utilizó la comparación. Shin y Kim (2019) explica que «el hambre (incertidumbre) debido a falta de comida (información) estimula el apetito (curiosidad) lo que conduce a un comportamiento de búsqueda de alimentos (búsqueda de información) y saciedad (satisfacción) cuando se consume comida (se llena la brecha de información).» y extiende la comparación en la imagen de la Figura 2.2. Así como el apetito es el deseo de comida para reducir la incomodidad psicológica del hambre, la curiosidad es el deseo de información para reducir la incomodidad psicológica de la falta de certeza.

Profundizando la metáfora, podría pensarse en el sensacionalismo como la comida chatarra del mundo de la información: son noticias de fácil consumo, seductoras a la vista y que calman el apetito sin dejar mucho valor (nutricional, o informacional). De hecho el término *junk food news* fue acuñado en el año 1983 (Wikipedia, 2024c). Y en este contexto el clickbait se puede comparar con la publicidad de comida que apela a seducir al consumidor con imágenes tenta-

doras para generar apetito. Nuevamente se puede ver que el sensacionalismo y el clickbait están relacionados (la mayor parte de este tipo de publicidad de comida es de comida chatarra como la mayor parte del clickbait es con contenido sensacionalista), pero perfectamente pueden existir este tipo de publicidades para comida de alto valor nutricional, así como se puede generar intencionalmente una brecha de información con el titular de una noticia típicamente dura.

Li et al. (2022) realizan estudios de ERP mediante electroencefalogramas a personas que leen noticias que son o no clickbait. Concluyen que el clickbait (en particular las catáforas, véase Sección 2.3) aumenta el potencial del click de los usuarios y, lo que es más importante, de «una fuerte atención involuntaria que atrae muchos recursos de la atención» desde un punto de vista cognitivo. Los usuarios, estimulados por la información faltante le dan involuntariamente su atención al texto que genera esa brecha.

Hay un aspecto que permite ahondar aún más en la analogía de la comida: la posibilidad de que ambas puedan surgir sin una necesidad o un motivo externo (en el caso de la información nunca hay una necesidad real como es el caso del hambre, nadie muere directamente por tener poca información). Como dicen Shin y Kim (2019) «aunque la incertidumbre y el hambre son antecedentes claves de la curiosidad y el apetito, respectivamente, la curiosidad y el apetito aún pueden surgir en su ausencia». Así como ver u oler comida que nos gusta puede generar apetito pese a ya haber comido suficiente, las brechas de información generan curiosidad sobre información que no nos es relevante. «Así como la comida es más atractiva cuando alguien tiene hambre, la información es más valiosa cuando es óptimamente incierta».

Desde el punto de vista neurobiológico, la investigación ha demostrado que el cerebro reacciona ante la información como ante una recompensa neuronal (Blanchard et al. 2015a; Kang et al. 2009; Kobayashi y Hsu, 2019a). Las recompensas modelan nuestro comportamiento generando placer ante ciertas acciones u objetos. La presión evolutiva generó recompensas motivando las acciones que permiten la sobrevivencia, como las homeostáticas (comida, bebida, etc) y sexuales, las recompensas primarias. Además, existen otros estímulos que generan recompensas: desde objetos materiales como el dinero o artículos de lujo, hasta subjetividades estéticas, como el arte, y actividades sociales (Schultz, 2015). Estos estímulos son placenteros por la reacción cerebral mediante señales en las neuronas dopaminérgicas.

Kang et al. (2009) experimentan con personas que reciben preguntas triviales mientras se someten a una resonancia magnética funcional. Luego de leer las preguntas, los sujetos evaluaron su nivel de curiosidad y confianza con respecto a la respuesta, que recibirían luego de un tiempo. Las imágenes de la resonancia muestran que la curiosidad activa los circuitos de recompensa y su lectura es «consistente con la idea de que la curiosidad está asociada con un valor intrínseco de aprender».

Se puede ver además que la curiosidad es un sentimiento de anticipación. La respuesta cerebral en una situación en la que se satisface una curiosidad específica no comienza cuando se recibe la información, sino cuando recibe señales que le indican que la va a recibir (Baranes et al. 2015; Blanchard et al. 2015b; Kang et al. 2009; Oudeyer et al. 2016; Shin y Kim, 2019). Bromberg-Martin y Hikosaka (2009) ofrecen un premio (agua) de magnitud variable a monos y estudian el deseo que generan en saber con anticipación qué tanta agua recibirán. «Las mismas neuronas de dopamina que indican la cantidad esperada de agua también indican la expectativa de información».

Es decir, en el caso del clickbait, se puede pensar que el cerebro comienza a reaccionar cuando, luego de haber leído el titular y haberse generado la curiosidad, se percata de que está a apenas un click de lograr cerrar la brecha de información creada. Antes de leer la noticia, incluso antes de hacer click, se genera curiosidad, que desencadena la actividad dopaminérgica. «La curiosidad se confunde frecuentemente con una experiencia positiva porque la anticipación de resolver la curiosidad, como la anticipación de la recompensa, conduce a un efecto positivo» (Shin y Kim, 2019).

Incluso Ming Hsu, uno de los autores de Kobayashi y Hsu (2019b) comenta el trabajo en un artículo y afirma que «si bien la investigación no aborda directamente el consumo excesivo de información digital, el hecho de que la información involucra el sistema de recompensa del cerebro es una condición necesaria para el ciclo de adicción» (University of California - Berkeley Haas School of Business, 2019). Y agrega «la forma en que nuestros cerebros responden a la anticipación de una recompensa placentera es una razón importante por la cual las personas son susceptibles al clickbait. Al igual que la comida chatarra, esta podría ser una situación en la que los mecanismos de adaptación anteriores se explotan ahora que tenemos un acceso sin precedentes a nuevas curiosidades».

Otro aspecto a tener en cuenta es que la búsqueda de información está

motivada tanto por razones instrumentales (esa información puede servir para obtener algo) como no-instrumentales (Kobayashi y Hsu, 2019b). La utilidad percibida de la información es una función subjetiva que «puede incorporar diversas influencias, incluyendo riesgo, demora, esfuerzo e interacción social» (Schultz, 2015). Por lo tanto, «si se percibe que el costo de resolver la curiosidad es alto, las personas están menos dispuestas a buscar la información» (Shin y Kim, 2019). Esto explica que pueda surgir la curiosidad ante una noticia de un tema que no tiene valor instrumental, que no nos interesa o genera valor (véase Sección 2.7 del Apéndice 2). Conocer la facilidad e inmediatez que implica el apenas hacer un click para conocer la respuesta y saber que el artículo seguramente no requiera prácticamente un esfuerzo cognitivo en su lectura, aumenta la sensación de utilidad. Es, además, una nueva razón para explicar la asociación entre el clickbait y el sensacionalismo. Incluso, la decisión sobre si obtener la información capaz de saciar la curiosidad o no puede variar si esa información llega de inmediato o con un retraso (Kruger y Evans, 2009): la inmediatez puede jugar un rol relevante.

No es solamente que la noticia puede no generar valor instrumental sino que puede además tampoco gustarnos. Hay una diferencia psicológica y neurobiológica entre los conceptos de querer y gustar. El querer algo implica la prominencia de incentivos que motivan la consumición de algo que no necesariamente tiene que gustar (Berridge et al. 2009). De hecho, Kruger y Evans (2009) muestra que la curiosidad puede hacer que las personas no sólo busquen información inútil sino también información previsiblemente dañina para satisfacer su curiosidad.

El deseo de satisfacer la curiosidad es difícil de controlar, explica Kruger y Evans (2009): «así como la satisfacción inmediata del hambre, el deseo o la excitación pueden hacer que uno coma, beba o tenga relaciones sexuales a pesar del deseo genuino (y a veces desesperado) de mantenerse delgado, sobrio o abstinente, también la curiosidad puede hacer que uno busque una información con la que uno sinceramente cree que estaría mejor sin tener conocimiento.»

La curiosidad incita también al consumo complaciente, la elección del consumidor de permitirse seleccionar y disfrutar el placer de una opción que se considera darse un gusto en comparación con las opciones alternativas (Cavanaugh, 2014) asociado a actos innecesarios e irracionales (Suzuki et al. 2018). En Wiggin et al. (2018) uno de los estudios expone a un grupo de personas a imágenes borrosas y los reta a adivinar qué hay en la original para incitar

la curiosidad (se puede ver la brecha de información activada), mientras que al grupo de control no. Luego, se les hace elegir a todos una película para ver y el grupo *curioso* es más proclive a elegir una complaciente frente a una virtuosa. Acompañado de otros estudios similares, se provee evidencia de que «un sentimiento de curiosidad sirve como un ímpetu para caer en la tentación de la indulgencia creando un deseo por premios», incluso cuando no están relacionados con lo que generó la curiosidad.

Ese deseo es un «factor visceral», que explica el por qué podemos actuar en contra de nuestro propio interés incluso siendo conscientes de ello (Loewenstein, 1996). Una brecha de información puede generar «un deseo irrazonablemente fuerte de llenar el vacío» (Loewenstein, 1994). Tan irracionales e impulsivas pueden ser las decisiones tomadas en un estado de curiosidad que «las personas están dispuestas a apostar a someterse a riesgos físicos sólo para satisfacer su curiosidad por conocimiento trivial sin valor instrumental aparente» (Lau et al. 2020). Un experimento le hizo ver a algunas personas un truco de magia o una pregunta trivial y a otras imágenes de comidas apetitosas (nuevamente aparece la comparación entre la curiosidad y el apetito) y les permitieron apostar: de ganar podían ver la resolución del truco, la respuesta a la pregunta o comer uno de los platos mostrados, de perder recibirían un shock eléctrico. Los resultados mostraron que, en igual medida para apetito y curiosidad, varias personas estaban dispuestas a correr el riesgo de recibir un castigo físico con tal de poder saciar su incomodidad inmediata producida por la curiosidad (o el apetito). El experimento incluye también un estudio de resonancia magnética funcional que confirma la activación de las áreas de recompensa del cerebro y son consistentes con que «la importancia de los incentivos juega un papel en los comportamientos impulsivos causados por la curiosidad de manera similar a como lo hace con los incentivos extrínsecos (por ejemplo, la comida).»

Por otra parte existe una diferencia entre la curiosidad epistémica y la curiosidad por carencia de información, la que surge de enfrentarse con una brecha de información (Litman, 2008, 2010). En Zedelius et al. (2022) muestran que esta curiosidad por carencia está relacionada con emociones negativas y correlaciona con la falta de capacidad de distinguir información real de inventada, la creencia y ganas de compartir noticias falsas, la receptividad a la pseudociencia y la exageración. En este contexto, menciona el clickbait como uno de los fenómenos que nos exponen a este tipo de curiosidad.

Nuevamente, vale la pena aclarar que la mayoría de estos trabajos no tienen

directamente al clickbait como objeto de estudio. Sin embargo, a mi entender, describen las razones por las que el clickbait funciona, el por qué se genera esa sensación de curiosidad, qué pasa psicológica y neurocientíficamente cuando nos invade la curiosidad y la intensidad de los sentimientos de necesidad de satisfacerla.

2.3. Creación de brechas de información en los medios

Entender las razones y los efectos de la curiosidad que genera ver un titular donde falta parte de la información permite valorar su importancia. Los medios aprovechan el sentimiento de curiosidad y las fuertes ganas de llenar esa brecha de información para hacer que la gente se sienta casi que obligada a cliquear en sus noticias, y así ganar la batalla por la atención y los clicks.

Muchas veces la brecha de información se genera con la forma de escribir los titulares y/o avances. En otras palabras, entregando la misma información que entrega un titular tradicional, se logra insinuar algo más, atraer, u otras estrategias para crear esa brecha.

Las preguntas se usan en general para insinuar información pero no entregarla. En el primer ejemplo de la Figura 1.1 se informa que Evo Morales escribió una carta donde presentó su «renuncia obligada», pero se pregunta qué dice. Esta sugerencia de que hay más contenido, pero sin contar qué es, genera la curiosidad por saber más. Un titular tradicional como *Evo Morales publicó una carta donde presentó su renuncia obligada* da la misma información pero no crea esta necesidad.

Las expresiones apelativas son otra de las estrategias de redacción que se utilizan con este fin. En el ejemplo de

> (4) *WhatsApp dejará de funcionar en varios modelos de iPhone: ¿te va a tocar cambiar de móvil?*

aquí la pregunta apelativa no agrega ni quita nada de información, pero es la generadora de la curiosidad.

Otra estrategia, correctamente, muy asociada al clickbait es la de utilizar catáforas[1]. Es un método muy efectivo para generar brechas de información,

[1] «Relación de identidad que se establece entre un elemento gramatical y una palabra o grupo de palabras que se nombran después en el discurso.» Fuente: RAE

Figura 2.3: Ejemplo de clickbait formado por una deixis, en este caso modal. También llamado *forward-reference* en inglés, es un recurso muy efectivo para generar brechas de información, y por lo tanto muy frecuente en el clickbait, estudiado en profundidad en Blom y Hansen (2015) y Li et al. (2022)

estudiado en profundidad en (Blom y Hansen, 2015; Li et al. 2022), y frecuente entre los titulares clickbait. En el titular

> (5) *Esto es lo que tienes que ahorrar al mes para poder pagar la universidad de tus hijos*

esto refiere a algo que todavía no se nombró. Además el titular tiene una expresión apelativa (*tus hijos*). En

> (6) *#Moquegua: así se desarrolla el segundo día de protestas contra el proyecto minero #Quellaveco*

que se ve en la Figura 2.3, es la deixis modal *así* la que insinúa algo desconocido. El titular *#Moquegua: segundo día de protestas contra el proyecto minero #Quellaveco* (que de hecho es el titular que se ve en la imagen del tweet, véase Figura 2.3) contiene la misma información, casi las mismas palabras, pero no es clickbait. Al igual que en los ejemplos anteriores, la catáfora aquí no agrega ni quita información, pero es decisiva en la clasificación.

Similarmente, en

> (7) 🚨 *Esto es lo que se sabe de la detonación de un artefacto que dejó seis menores heridos en Buenaventura* 👇

el título de la noticia (a diferencia del texto del tweet) es *Seis menores resultaron heridos por explosión en Buenaventura* que aporta la misma información

sin generar curiosidad y está implícito que la nota aporta los detalles que se conocen. Este tipo de uso de deixis para generar catáforas está explicado en detalle también por Palau-Sampio (2016).

Los *listículos*, o artículos-lista, que han sido estudiados como un fenómeno en sí mismo (Sadri, 2019; Vijgen, 2014) son sistemáticamente utilizados para generar clickbait, e incluidos en las distintas tipologías. Se trata de titulares que prometen enumeraciones y generan curiosidad por saber qué o quién las integra. Muchas veces el avance incluye la cantidad de contenidos como en

> (8) *Estas son 5 plataformas de entretenimiento para diversificar sus horas de diversión*

y

> (9) *Final de la NBA: 5 claves del duelo entre los Golden State Warriors y los Cleveland Cavaliers por el trono del baloncesto mundial*

pero a veces es implícito como en

> (10) *Claves para encontrar las mejores ofertas en el Black Friday 2019*

La adjetivación exagerada, grandilocuente o sin conexión es otro recurso de redacción muy común en el clickbait. En

> (11) *Los elegidos de Ariel Schettini: virginidad, desborde y tragicomedia*

no hay nada que nos indique por qué *los elegidos* se describen de esa forma, pero genera curiosidad saber qué será que tiene esas tres cualidades. Es en estos aspectos donde el clickbait se parece y confunde más con el sensacionalismo. Pero, nuevamente, es importante notar que lo que convierte la noticia en clickbait no son los adjetivos en sí sino su uso con el objetivo de generar una brecha de información. En

> (12) *Insólito: Kyllian Mbappé terminó en muletas por el patadón de un rival que vio la roja y se retiró del fútbol*

es esa calificación de *insólito* la que llama la atención y genera más curiosidad, o en

(13) *La vertiginosa vida de Christian Nodal tras revelar su romance con Belinda*

es el *vertiginosa* lo que hace querer saber más, para entender el por qué del epíteto.

Finalmente, existe una estrategia de redacción que Bazaco et al. (2019) define como *ambigüedad deliberada* que es muy utilizada para generar clickbait. Al usar este método, se pueden crear casos donde el límite entre lo que es clickbait y lo que no es muy subjetivo y también se puede caer en titulares engañosos (fenómeno discutido en la Sección 2.5 del Apéndice 2).

El caso de

(14) *Asesinan de un tiro en la cabeza a un repartidor de pizzas en La Plata*

describe cómo la clave en ciertos clickbaits es la redacción. El texto del tweet es un titular totalmente normal. Sin embargo, el título de la noticia es *Entregó una pizza y lo asesinaron de un tiro en la cabeza.* La noticia, el hecho sobre el que se informa, es claramente la misma, pero la presentación es sutilmente muy distinta. El título sugiere que ambos hechos están relacionados, que existe una relación causal, aunque el desarrollo de la noticia no sugiere ninguna relación más allá de la temporal. La causalidad entre un hecho tan común como entregar una pizza y algo tan trágico como un asesinato genera la curiosidad: ¿por qué un hecho llevó al otro? Este tipo de casos se relaciona también con lo discutido en la Sección 2.5 sobre la relación del clickbait y la desinformación, porque al leer la noticia completa se encuentra que no existe tal relación.

Estas ambigüedades se generan frecuentemente con el recurso de la nominalización: titulares que se componen únicamente por un sintagma nominal o que no contienen verbo conjugado (Palau-Sampio, 2016). Los siguientes tweets

(15) *Steve Irwin: las últimas palabras del Cazador de Cocodrilos antes de morir hace 15 años*

(16) *La sorprendente confesión de Juana Viale sobre su vínculo con los futbolistas*

(17) *Las dudas en el lanzamiento del equipo de fútbol de David Beckham en Miami*

son ejemplos de nominalización. La falta de una acción o explicación nos hace preguntarnos *¿qué pasa con ellos?, ¿por qué son noticia?*

No siempre este tipo de nominalizaciones genera una brecha informativa. En

> (18) *Ronaldinho y Gimnasia: el deseo del brasileño de volver al fútbol y ser dirigido por Diego Maradona*

la propia existencia del deseo de Ronaldinho es la noticia, la redacción no genera curiosidad sino que no es más que un recurso de titulación. El límite que define cuándo la nominalización oculta o sugiere información es difuso y por lo tanto existe un componente subjetivo en la decisión de la clasificación. En el intento de generar curiosidad, los medios pueden incluso generar titulares que, por sí mismos, sugieren algo distinto a lo que dice realmente la nota. Esta práctica fue abordada en la Sección 1.2 y más en profundidad en la Sección 2.5.

2.4. Propuesta de definición

Varias definiciones y descripciones del clickbait en la literatura nombran la omisión parcial de información como parte importante del concepto. Por ejemplo, la definición que dio Facebook cuando anunció que intentaría reducir la cantidad de clickbait en su plataforma (Peysakhovich, Alex and Hendrix, Kristin, 2016), «titulares que intencionalmente dejan afuera información crucial, o engañan a la gente, forzándola a hacer click para obtener una respuesta», apunta claramente en esa línea (aunque agrega el engaño). Biyani et al. (2016) y luego García Serrano et al. (2019) consideran la «omisión de detalles para crear suspenso» como uno de los tipos de clickbait. También Potthast, Gollub, Komlossy et al. (2018) afirman que «el clickbait retiene información crucial» y Bazaco et al. (2019) que considera «información incompleta» como la primera variable caracterizadora.

Scott (2021) estudia el clickbait desde el punto de vista de la teoría de la relevancia lingüística y encuentra evidencia de sobreutilización de ciertas características que favorecen la brecha de información. Explica que se puede entender «en términos de una promesa de una relevancia futura y el poder del clickbait en hacernos clickear puede entenderse como una explotación de una tendencia natural a buscar relevancia».

Sin embargo, a diferencia de estos trabajos y por lo expuesto a lo largo de este capítulo, considero que el ocultar parte de la información no es una característica más del clickbait sino el elemento crucial que lo distingue de otras estrategias para llamar la atención. De hecho Peter Koechley (2012), cofundador de Upworthy (uno de los medios más relacionados con el clickbait, véase el Apéndice 1) describió su técnica para escribir buenos titulares como crear unos que «seducen a la gente a hacer click contándoles suficiente para despertar su curiosidad, pero no lo suficiente para satisfacerla. Los titulares en redes sociales necesitan crear una brecha de información. Demasiado vago, a nadie le interesa. Demasiado específico, y nadie necesita hacer click.»

El clickbait es un fenómeno propio de la digitalización, y por lo tanto no puede tener la misma definición que el sensacionalismo que nació más de un siglo antes. Pese a que está relacionado íntimamente y es muy frecuente encontrarlo con estrategias sensacionalistas como la exageración, la voz imperativa y la temática de noticias blandas, el clickbait tiene algo que lo distingue. Es importante marcar la diferencia para, en lugar de sumar un sinónimo más para lo mismo, poder distinguir claramente un método distinto.

La gran diferencia entre el clickbait y otros tipos de formas de titular que buscan atención, como el sensacionalismo o la prensa amarilla, es que se llama la atención mediante la curiosidad que se logra en el lector al generar una brecha de información a través de la omisión de un fragmento de la información.

Tomando en cuenta lo argumentado, propongo una definición de clickbait:

> **El clickbait es un método para generar avances de noticias, especialmente online, que consiste en omitir deliberadamente parte de la información con el objetivo de generar curiosidad mediante una brecha de información y así atraer la atención de los lectores y conseguir que hagan click**.

2.5. Criterios de identificación de clickbait

Incluso contando con una definición de consenso y trabajando sobre un mismo conjunto de datos, cuando diferentes personas, incluso quienes cuentan con más conocimiento en el área, etiquetan textos en problemas complejos y subjetivos, el acuerdo entre ellas dista mucho de ser perfecto. Por esta razón

es muy importante la construcción de criterios claros para generar consensos y aumentar el grado de consistencia al máximo posible. La línea que divide el clickbait continúa siendo de todas formas difusa y subjetiva, lo que se evidencia en algunos casos concretos.

Una de las formas de reducir la magnitud de este problema es generar una descripción exhaustiva de los casos que caracterizan el objeto de estudio, y utilizarlo como definición operacional. En muchos trabajos se especifican los criterios de etiquetado o análisis que generan una tipología que marca cuándo una noticia se trata de un clickbait. Estas distintas tipologías difieren entre sí dependiendo principalmente de la definición que se haya tomado en cuenta al construirlas (Araujo et al. 2021; Bazaco et al. 2019; Biyani et al. 2016; Palau-Sampio, 2016; Zhang et al. 2020).

Por la diferencia en definiciones y el intento de cubrir la totalidad de los casos, las tipologías de las que tengo conocimiento incluyen casos que no se ajustan a la definición propuesta. En la mayoría de los casos se debe a características que suelen comunes o estar relacionadas con el clickbait pero no tienen, según mi definición propuesta, relación de causalidad. Por ejemplo, muchos incluyen las preguntas (Araujo et al. 2021; Biyani et al. 2016; García Serrano et al. 2019; Mihelich, 2019; Palau-Sampio, 2016; Zhang et al. 2020), hipérboles y exageración (Araujo et al. 2021; Bazaco et al. 2019; Biyani et al. 2016; García Serrano et al. 2019), lenguaje coloquial y palabras genéricas (Palau-Sampio, 2016), variables de contenido (Bazaco et al. 2019; García Orosa et al. 2017), expresiones de inmediatez (tales como *esto* y *ahora*) (Mihelich, 2019) o características de redacción y formato (tales como signos de puntuación, juegos de palabras, uso de mayúsculas, lenguaje vulgar, etc) (Araujo et al. 2021; Bazaco et al. 2019; Biyani et al. 2016; García Serrano et al. 2019; Palau-Sampio, 2016; Zhang et al. 2020).

Estos criterios apuntan a basar cada decisión en la definición propuesta y únicamente en ella: son los titulares que generan curiosidad con una brecha de información. De todas formas también con este enfoque se generan subjetividades y existen muchos casos límite, grises y decisiones a tomar, algunas ya mencionadas en las tipologías y algunas otras descritas en esta sección junto con ejemplos reales para ilustrarlos.

2.5.1. Definición operativa

La definición de clickbait propuesta habla de una omisión deliberada, lo que a la hora de clasificar textos como clickbait genera el problema de inferir la intención de quien escribe. Es, como mínimo, muy difícil de probar la intencionalidad sin tener la posibilidad de preguntarle a quien escribió el titular. Por esta razón propongo una definición operativa que se usa en el trabajo para la clasificación, atenuando el problema de la subjetividad:

> **El clickbait es un método para generar avances de noticias que omiten parte de la información que no es obvia por contexto, generando curiosidad mediante una brecha de información, atrayendo la atención de los lectores y aumentando las posibilidades que hagan click.**

2.5.2. Tener en cuenta todo el contexto de publicación

Con la definición operativa desaparece el problema de inferir la intención de quien escribe, pero surge uno nuevo: ¿qué información es considerada obvia por contexto? Es habitual que el grado de subjetividad en la decisión se origine en la información de contexto que tienen los distintos receptores. Ese contexto es lo que define si la información que no se explicita falta porque es obvia o para generar curiosidad.

El criterio definido es tomar como obvio cualquier información que el lector promedio de esa noticia seguramente conozca, considerando factores como cuándo y dónde se publicó la noticia y con qué público objetivo. En

> (19) *La masacre en el aeropuerto de Kabul deja al menos 170 muertos*

es crucial tomar en cuenta la fecha de publicación: el tweet es del 27 de agosto del 2021, un día después de los atentados en el aeropuerto[1] por lo que la información que está transmitiendo el titular es la actualización en la cantidad de muertes producidas y no el acontecimiento en sí, por lo que no es clickbait.

Se debe tener siempre en cuenta el contexto del país en el que fue publicada la noticia. Hay casos en los que el contexto es obvio: en

> (20) *Vicentín: una decisión con un rédito político incierto*

[1] Atentados del Aeropuerto Internacional Hamid Karzai de 2021

la *decisión* fue un tema muy discutido en Argentina y no era necesario en ese momento aclarar cuál fue la decisión, no se explicita por obvia y no por generar clickbait. También es el caso de

> (21) *Se retiró Mascherano: el extraordinario viaje del caudillo al que sólo le quedó una materia pendiente en la Selección*

cualquier argentino sabe que la materia pendiente es haber ganado un título oficial con la selección.

Es común que los medios, para economizar espacio, eviten nombrar eventos notoriamente ya conocidos en sus titulares. En el titular

> (22) *Estados Unidos: un escándalo familiar de Donald Trump empaña el comienzo de la convención republicana*

uno se pregunta en qué consistió el escándalo. Un titular muy similar, si se lee fuera de contexto, como

> (23) *Unanimidad en Diputados: suspenden al diputado salteño Juan Emilio Ameri tras el escándalo*

puede generar también curiosidad sobre cuál fue *el escándalo*, pero se trató de un hecho de mucha notoriedad y, por lo tanto, en Argentina y en ese día era difícil que alguien no supiera de qué se trataba. En este caso la noticia no es clickbait: se debe considerar el contexto *normal* de la noticia en su momento y lugar para tomar la decisión.

Hay veces en las que la información de contexto es obvia para cualquier persona, pero puede ser de todas formas desafiante para un sistema automático. En

> (24) *Científicos logran leer, sin abrirla, una carta de hace 300 años*

nada de la redacción indica que se esté ocultando información, podría ser una noticia cualquiera, excepto por el hecho de que cualquiera sabe que leer una carta sin abrirla es difícil y surge inmediatamente la pregunta *¿cómo hicieron?*. De hecho otro medio cubrió la misma noticia con un titular no-clickbait que explica, brevemente, el cómo:

> (25) *Sin abrir, revelan contenido de carta de hace 300 años con un rayos X dental*

Un caso similar es el de

(26) *Identifican una hormona que regula el envejecimiento*

alguien podría preguntarse *¿qué hormona?, ¿cómo se llama?* Pero lo relevante de la noticia en este caso es el descubrimiento en sí y no tanto el nombre. Es una diferencia sutil, que requiere conocimiento del mundo para entenderlo e incluso compartiendo el conocimiento se puede diferir. En

(27) *Hallaron el método más increíble para adelgazar: comer tierra*

sucede algo similar, la curiosidad surge de lo inesperado de la respuesta. Lo que falta no es una respuesta sino una explicación básica de cómo puede ser posible (además de que se lo califica de «increíble»).

Un ejemplo claro de la dependencia en el contexto sucede con las noticias sobre deportistas que contraen Covid-19. En

(28) *#Turf — Quince jockeys dieron positivo de coronavirus en California y cancelaron las carreras del fin de semana próximo*

queda claro que lo importante no es de qué jockeys se trata, sino el hecho de que hubo positivos, lo que llevó a suspender la actividad. En

(29) *La golfista mexicana María Fassi dio positivo por Covid-19*

lo importante es quién fue la afectada, aunque no hay ocultamiento de la información y por lo tanto no hay clickbait. Sin embargo, esta diferencia entre si lo relevante es que hubo positivos o quién exactamente los tuvo no es tan clara cuando la información no está presente.

(30) *Un golfista dio positivo de coronavirus en pleno torneo del PGA Tour*

es seguramente una noticia en sí misma, principalmente teniendo en cuenta que es de junio 2020 y fue, como dice la nota, «el primer caso de contagio en el circuito que había vuelto a disputarse desde hace una semana». Pero esa conclusión está basada en conocimiento contextual del mundo: es necesario saber que la epidemia había comenzado sólo unos meses antes y que no había casos anteriores en el circuito de golf profesional. Algo similar pasa con

(31) *Juventud de las Piedras confirmó 17 casos positivos de COVID-19*

con el conocimiento de que Juventud de las Piedras es un club de fútbol cuyos planteles suelen tener cerca de 25 o 30 futbolistas, queda claro que lo importante es que casi todos fueron positivos y no los nombres en particular. En otros casos la conclusión no es tan clara o requiere de un conocimiento mucho más fino. En

(32) *Un jugador del Real Madrid dio positivo de coronavirus en la previa del duelo con el Inter*

es necesario saber si los partidos se suspendían en ese momento por un caso de covid, lo que haría del titular un hecho completo, o si sólo se lo perdía el jugador afectado, lo que convierte la ausencia de su nombre en una importante omisión. Este conocimiento de contexto puede depender también de los intereses de los lectores:

(33) *Posponen partido de la NBA por positivo al Covid-19*

¿Es importante saber que se suspendió *un* partido de NBA, o *qué* partido exactamente, o *qué jugador* exactamente fue el afectado? Seguramente dependerá del conocimiento e interés que se tenga por la NBA. Similarmente, en

(34) *Un atleta argentino que ya tiene su lugar en los Juegos de Tokio dio positivo de coronavirus*

si se tiene conocimiento sobre los atletas argentinos que irían a los Juegos Olímpicos la ausencia del nombre del afectado generará curiosidad, pero sin conocimiento alguno podría considerarse que el hecho de que alguno de ellos esté en duda es una información completa en sí misma.

Muchas de estas decisiones finas suelen surgir de situaciones en las que un sujeto omitido realiza una acción y la pregunta es si lo importante de la información es la acción en sí o quién la realizó. En

(35) *Un congresista demócrata demanda a Trump por el asalto al Congreso*

no importa mucho quién fue sino que alguien demandó a Trump. Pero ya no es tan claro en

(36) *En Día del Trabajo, Donald Trump critica al líder del Sindicato de Trabajadores más grande Estados Unidos*

podría considerarse como un clickbait al omitir quién es el líder o cuál sindicato, podría considerarse información obvia para algunas personas, o quizás lo importante es la crítica a quien ocupa esa posición sin importar tanto quién es (pese a que no sea obvio). A la hora de etiquetar esta noticia, consideré la tercera opción como la correcta en este caso. Algunos ejemplos similares son

(37) *Elecciones en Bolivia: tres candidatos piden inhabilitar al partido de Evo Morales*

donde no es claro si la acción es lo único importante o también quién lo pidió (nótese que son también candidatos),

(38) *El campeonato ecuatoriano de fútbol se quedó sin su patrocinador principal*

donde es difícil decir si el patrocinador es tan principal que es obvio de quién se trata o si se oculta la marca para generar curiosidad y

(39) *Vandalizan otra escultura en Berlín; esta vez grafitean nombre de un rapero y equipo de fútbol*

donde la redacción sugiere conocimiento previo (*otra*, *esta vez*) por lo que me inclino a pensar que lo importante es que se continúa vandalizando esculturas y no qué rapero y equipo de fútbol en particular fueron los elegidos.

Hay ejemplos en los que es el conocimiento de una persona el que define el contexto.

(40) *Leila Sucari: una repentina sensualidad*

podría parecer el típico artículo prometiendo una imagen de redes sociales. Sin embargo, se trata de una reseña sobre el nuevo trabajo de una escritora argentina. Más difícil es el caso de

(41) *Esmeralda Mitre se hizo un nuevo hisopado y dio negativo: Esto es un complot*

donde hay que conocer el contexto previo (hubo al menos otro hisopado) para entender si el supuesto *complot* surge de forma *obvia* de los acontecimientos o se intenta generar curiosidad al no explicar por qué.

Es el contexto también el que nos hace conocer en ciertos casos que algunos aspectos que se destacan son muy específicos e inconexos. Y son esas características las que generan una brecha de información: conocemos un detalle

muy particular pero no la historia completa, insinúa información pero no la cuenta. Se puede ver en

(42) *Joe Biden, el aliado de la luz marcado por la tragedia familiar*

Muchos de los casos de uso de esta técnica pertenecen a un tipo particular de clickbaits que son las crónicas presentadas sugiriendo aspectos específicos.

2.5.3. Insinuar con detalles desconectados genera clickbait

Un recurso muy usado es titular un artículo de crónica o relato de una historia destacando algunos detalles específicos que suelen ser llamativos, extraños, a veces exagerados y hasta estrafalarios, presentados sin contexto alguno. Las crónicas no son necesariamente sobre noticias blandas, tampoco los detalles pueden siempre asociarse a prácticas sensacionalistas. Pero allí presentados, sin nada que los acompañe, son el ingrediente perfecto para la activación de una brecha de información.

Frecuentemente se destacan varios detalles no relacionados como en

(43) *La fundación Slim, el empresario Sigman desde Madrid y un reto a Ginés González García: cómo se gestó el anuncio de la vacuna de Oxford*

y a veces es una única particularidad como en los casos de

(44) *El estudiante que resolvió dos grandes problemas matemáticos por impuntualidad*

(45) *25 años de la PlayStation: el enojo que cambió la historia*

Es interesante notar que algunas definiciones de clickbait o caracterizaciones en tipologías hablan de «omitir detalles» o «retener información crucial» (Biyani et al. 2016; García Serrano et al. 2019; Potthast, Gollub, Komlossy et al. 2018). Sin embargo, este tipo de clickbait genera la brecha de información no ocultando detalles sino ofreciéndolos sin la información necesaria para entenderlos y conectarlos con el resto de la información disponible. Además, en la gran mayoría de estos casos la lectura total de la nota aclara el contexto y por lo tanto es un tipo de clickbait que da exactamente lo que promete y no

se ajusta a las definiciones que requieren la decepción como fue discutido en la Subsección 2.1.2.

Algunos de los casos son titulares que combinan este recurso con el de un listículo o artículo lista como en

> (46) *De la inquietante gambeta a Vladimir Putin a la foto sin revancha con David Beckham: cinco anécdotas inéditas de Diego Maradona*

Se destacan algunos de los puntos más llamativos de la lista para generar aún más curiosidad.

El límite entre cuánto contexto dar o cuán extravagante es un detalle es, nuevamente, difuso y genera, nuevamente, subjetividades. Por ejemplo en

> (47) *La falta envido que desató un crimen en un bar de pueblo*

el cuerpo de la noticia concuerda casi que exactamente con el título, lo único que no explica es que fue una discusión sobre las reglas del juego de truco, y específicamente las de la falta envido, la que generó la discusión que terminó en asesinato y no el juego en sí. Quizás es la nominalización (recurso analizado en la Sección 2.3) la que inclina la balanza hacia el clickbait en este caso. También se puede volver al ejemplo de *Entregó una pizza y lo asesinaron de un tiro en la cabeza* discutido en la Sección 2.3, genera curiosidad el por qué se destaca la entrega de la pizza y qué tiene que ver lo que sucedió después, pero no deja de ser un hecho real y no es claro qué se oculta. La redacción es similar a la de

> (48) *Caminaba con su mujer y sus hijas, quedó en medio de un cruce entre bandas y lo mataron a balazos*

que no genera brecha de información, la clave está como fue explicado anteriormente en la sugerencia de causalidad que tiene el primer título entre un hecho (la entrega de la pizza) y el otro (el asesinato).

Ese límite también depende del contexto y el conocimiento previo del lector, por lo que está muy relacionado con la sección anterior (Subsección 2.5.2). En

> (49) *La agricultura consume alrededor del 70 % del agua disponible en el planeta y la forma en la que la obtiene no siempre es legítima*

la obtención de agua para riego de forma ilegítima puede considerarse algo conocido o también llamativo, dependiendo de quién lo lea.

En pos de hacer que los detalles sean atractivos, a veces se recurre a la adjetivación exagerada. Desde casos como

> (50) *Bacterias vampiro, caníbales zombis... y otros microbios de miedo. Sorprendentes características de los microorganismos*

hasta el más común

> (51) *La fuga rocambolesca del ex presidente de Renault deja atónitos a Japón y a Francia*

donde la adjetivación, precisamente, rocambolesca genera la curiosidad de cómo habrá sido esa fuga.

2.5.4. La adjetivación intensa puede provocar clickbait

Adjetivar y calificar excesivamente una noticia o un detalle de ella funciona muchas veces para activar la brecha de información. Es uno de los puntos de más consenso dentro de las tipologías de clickbait existentes. Pero hay una línea fina y, nuevamente, subjetiva entre lo que es excesivo y lo que no lo es. Un titular como

> (52) *Cómo es un huracán espacial, el espectacular fenómeno detectado por primera vez en la Tierra*

no explica de ninguna forma por qué es *espectacular*, por lo que es clickbait.

Habitualmente se exageran características de una foto, video o declaración como en los casos de

> (53) *La foto de un zorro y un cóndor juntos se hizo viral y conmueve al mundo entero*

> (54) *La sorprendente confesión de Juana Viale sobre su vínculo con los futbolistas*

El recurso es el mismo: el adjetivo es exagerado, lo que llama la atención, y no se explica por qué, lo que genera curiosidad.

No siempre adjetivar es generar clickbait, e incluso la misma palabra puede generarlo o no dependiendo del contexto, retomando la discusión de la Subsección 2.5.2. Las noticias

(55) *Las bolsas mundiales registran fuertes alzas impulsadas por los nuevos planes de estímulos contra el coronavirus*

(56) *La fuerte acusación de Thomas Müller al Atlético Madrid del Cholo Simeone*

tienen el mismo adjetivo *fuerte*, pero en la primera se entiende que *fuertes* alzas es un crecimiento grande, mientras que la acusación *fuerte* le da sensación de gravedad pero sin explicar absolutamente el por qué, generando curiosidad.

La exageración es muy típica del sensacionalismo y este recurso es quizás la principal razón de su confusión con el clickbait. Sin embargo, si bien hay algunas noticias que claramente entran en ambas categorías como

(57) *Ballena en Puerto Madryn: el impresionante video de una ballena jugando con una mujer que estaba remando en el sur de Argentina. No te pierdas en este video el entrañable momento de este insólito encuentro*

existen noticias con todas las otras características del sensacionalismo e incluso la adjetivación llamativa como

(58) *La cruenta batalla por el Cártel de Sinaloa: el Mayo, los hijos del Chapo y el grupo de Caro Quintero se disputan su control*

donde los calificativos, pese a ser atractivos y/o provocativos se entienden en el contexto dado. Es muy conocido por qué las batallas entre cárteles son *cruentas* y no se necesita de más información para entenderlo, por lo que sí quizás se apele al morbo u otros sentimientos como lo hace el sensacionalismo, pero no a la curiosidad por una brecha de información, por lo que no es clickbait.

2.5.5. No todo el sensacionalismo es clickbait

Como se discutió en la Subsección 2.1.1, el clickbait se confunde habitualmente con el sensacionalismo, pero no son lo mismo. Por esa razón, algunas noticias que seguramente fueran consideradas como clickbait en otros trabajos, como

(59) *Leticia Brédice publicó fuertes mensajes contra su novio: lo acusó de ladrón y mala gente*

no deberían serlo. No hay ninguna brecha de información.

En la misma línea, si es la noticia en sí la que genera curiosidad, y no la forma en la que está presentada, tampoco se considera clickbait. Es el caso de

(60) *Video: cocodrilo se comió a dos tiburones en una playa de Australia*

donde el titular explica perfectamente lo que se encuentra en la noticia, sin omitir ninguna información importante.

2.5.6. Las preguntas pueden generar clickbait, pero no todas lo son

Formular titulares como preguntas es una de las formas más notorias y utilizadas de generar clickbait. Es común asociar al clickbait todas las noticias que se plantean en forma de interrogación, pero la presencia de una pregunta no implica necesariamente que se trate de clickbait.

Uno de los casos más evidentes es cuando la pregunta es una cita, como en

(61) *Las comunidades lamentan el pacto de mínimos: ¿Para qué hace falta un Ministerio de Educación si no coordina?*

También es habitual que la pregunta sea contestada en el mismo avance, como en

(62) *El Salvador — ¿Cuál es el perfil de estudios de los diputados? 82 % del actual pleno legislativo reporta una carrera de formación profesional. Y 25, de 84 diputados, reportan una carrera relacionada con el ramo jurídico.*

por lo que la brecha de información se cierra inmediatamente y no es necesario cliquear para satisfacer la curiosidad. Estos casos no son considerados como clickbait, y se profundiza el análisis en la Subsección 2.5.10.

Otra forma que suele repetirse es una pregunta que introduce un tema pero el resto del texto deja claro que la interrogante no se puede resolver aún con la información disponible. En el caso de

(63) *¿Deja a Messi por CR7? Juventus quiere a Luis Suárez*

Aquí, *Juventus quiere* implica que la transacción no es confirmada y por lo tanto no se sabía todavía si Suárez pasaría a jugar con Cristiano Ronaldo (CR7). Esto no es considerado clickbait. Similar es el caso de

> (64) *¿Fútbol o rugby? Fútbol gaélico, la combinación de dos deportes en la que la Argentina se destaca*

Es importante diferenciarlo del caso en el que la respuesta sí se puede obtener pero se oculta del titular, como en

> (65) *¿Cuál es la forma correcta de acariciar a tu gato? La ciencia tiene la respuesta*

donde sí se trata de clickbait.

Un caso similar al anterior es cuando la pregunta no se puede responder, pero la certeza no viene del texto en sí sino del conocimiento previo del lector. En

> (66) *¿Te acordás qué hacías, dónde estabas cuando Milton Wynants ganó la medalla de plata en Sídney 2000?*

es claro que quien escribe la nota no puede contestar la pregunta: la respuesta depende de cada lector. Este titular, pese a ser únicamente una pregunta y además apelativa, no es clickbait. Es distinto cuando el artículo no contesta la pregunta pero nada indica a priori, antes de hacer click, que así será. Por ejemplo en la nota a la que conduce

> (67) *Graciela Alfano confesó que tiene nuevo novio; ¿quién es?*

se dice que «prefirió guardarse para sí el nombre del caballero en cuestión». Aquí sí estamos ante clickbait y uno de los tipos que más suelen molestar por el efecto de decepción que se profundizó en la Subsección 2.1.2.

2.5.7. Las expresiones apelativas pueden generar clickbait

En varios de los ejemplos anteriores se puede ver otra característica repetida en el conjunto de datos: expresiones apelativas[1]. Utilizando este recurso quienes

[1] Dicho de una expresión lingüística, de un texto, etc.: Que pretenden influir en el receptor. (RAE)

escriben los titulares pretenden llamar la atención del lector mediante el pedido de que realice alguna acción o sencillamente se sienta aludido.

En las diferentes tipologías suele incluirse «expresiones apelativas» (Araujo et al. 2021; Bazaco et al. 2019) o similares como «apelación directa al interlocutor» (Palau-Sampio, 2016) sugiriendo o afirmando que todos los avances de noticias que las contienen son clickbait.

Por ejemplo, *WhatsApp dejará de funcionar en varios modelos de iPhone* sería un título completamente normal, sin embargo

> (68) *WhatsApp dejará de funcionar en varios modelos de iPhone: ¿te va a tocar cambiar de móvil?*

logra crear una brecha de información. Es esa pregunta directa al lector la que genera curiosidad. Otro caso similar es

> (69) *Al menos el 40 % de los establecimientos están cerrados por coronavirus ¿Quieres saber qué locales están abiertos en tu barrio? Pincha aquí* 👇

Ambos son clickbait.

Existen casos, por otro lado, donde las expresiones apelativas no generan brecha de información, el hecho de que muchos de los titulares apelativos sean clickbait no quiere decir que exista una relación causal. De hecho, hay medios como El Tiempo[1] que utilizan el modo imperativo en la gran mayoría de sus tweets, y eso no quiere decir que sean todos clickbait. Por ejemplo

> (70) *Al ritmo de El baile de los que sobran. De esta forma cientos de chilenos siguen celebrando el triunfo del voto a favor de cambiar la Constitución. Aquí le contamos más sobre esta histórica jornada en el país austral.* ➡️

> (71) 🚨 *Más de 15.000 peruanos marcharon para rechazar el Gobierno interino. Este Congreso no me representa, fue una de las consignas de quienes protestaron. Le contamos* 👉

sólo prometen ampliar la información, pero no ocultan nada para generar curiosidad. Logran sí, en todo caso, llamar la atención con la apelación directa

[1] https://www.eltiempo.com/ con su cuenta @ELTIEMPO

al lector y además los emojis. De forma similar a lo discutido en la Subsección 2.1.1 son estrategias que comúnmente se utilizan junto al clickbait porque tienen el mismo objetivo de llamar la atención y aumentar el tráfico, pero no necesariamente una implica la otra.

Hay también casos donde se apela al lector a expresar su opinión, con el objetivo de generar interacciones, pero sin ocultar información como en los siguientes casos que tampoco se consideran clickbait

(72) *Memo Ochoa regresa para el juego contra Tigres 🚨 ¿Les agrada la idea, americanistas? 🤔📥*

(73) *Instituto Geofísico reporta sismo de 3.5 grados en Simón Bolívar, Guayas #ATENCIÓN — Instituto Geofísico reporta sismo en Guayas este 28 de noviembre. ¿Sintió el temblor?*

.

2.5.8. Algunos editoriales pueden ser clickbait

Los medios de prensa combinan habitualmente notas de información con algunas de opinión. Estas editoriales suelen tener un estilo distinto de escritura y también de presentación. Algunos medios incluso los distinguen explícitamente cuando los comparten como en

(74) *[OPINIÓN] La ambigüedad y el mal entendido en la política: ecos de la carta de Cristina Kirchner — Por Federico González*

(75) *Columna — "El partido ha dejado sus huellas en la corrosión de la democracia, cuya consecuencia es Jair Bolsonaro. El PSDB no es sólo otro partido que ha contribuido a la más reciente escalada autoritaria de Brasil" ✍ @brumelianebrum*

En algunos trabajos deciden dejar las editoriales fuera del análisis. Por ejemplo García Serrano et al. (2019) lo hace argumentando que son «de carácter subjetivo y personal, difícilmente encasillables en una clasificación de veracidad informativa o de concreción frente al clickbait». Sin embargo, esta idea está basada en una definición de clickbait basada en contenido.

En ocasiones las editoriales también son redactadas de forma de ocultar una parte de la opinión para generar curiosidad mediante una brecha de información, como se ve en

(76) *Columna de Sebastián Rodríguez, psicólogo: Mi yerno se fue en la pandemia... ¿qué hago con mi hija?*

(77) *¿Por qué el Gobierno vuelve a atacar el proceso de paz?*

En otras palabras, también hay editoriales clickbait, aunque no sea el caso más común.

2.5.9. Los titulares incomprensibles no son clickbait

Un caso particular de la ausencia de información de contexto se da cuando el titular resultante, en lugar de sugerir una noticia, termina siendo directamente incomprensible o muy difícil de entender sin leer el artículo.

Puede darse por titulares *artísticos*, usualmente titulando crónicas o críticas culturales como sucede en el caso de

(78) *La primavera atrapada en un viejo sobretodo*

(79) *Sapiens, cautivos de nuestro ADN*

También suelen originarse en editoriales como

(80) *Apóstrofes*

(81) *Nicole ahora es experta*

Es claro que la omisión de contexto no se hace para generar curiosidad. Para que exista brecha de información tiene que faltar parte de la información, no toda. Por lo que estos titulares no son considerados clickbait.

2.5.10. Considerar todo el texto del avance

La definición propuesta habla de *avances* y no de titulares, porque el clickbait es muy común en redes sociales donde los títulos vienen usualmente acompañados de más texto y/o contenido multimedia. Además, dependiendo incluso del dispositivo y según si hay imágenes o no y otros factores, a veces el título de la noticia se ve completamente en la pantalla y otras veces no.

Es común que los medios publiquen la bajada de la nota acompañando el titular, incluso muchas veces se comparte la noticia más de una vez con distintos textos. Algunas veces esos avances cierran la brecha de información que el titular genera, *arruinando* el clickbait. Por ejemplo

(82) *¿Cuál es la distancia ideal que hay que correr por semana?*

es claramente un clickbait por sí mismo, pero no cuando se presenta acompañada del texto

(83) *Si el objetivo es mejorar la salud, no hay que correr más de 9 kilómetros por semana*

En un caso similar, no es lo mismo que el texto que acompaña al título *¿En qué están los dirigentes sudamericanos acusados por el caso FIFAGate?* sea

(84) *¿Cuál es la situación de todos los dirigentes sudamericanos involucrados en el FIFAgate de 2015? ¿Cuántos siguen presos y cuántos recuperaron su libertad parcial o total?*

que

(85) *"Tres de los acusados fallecieron, uno está preso, dos esperan sentencia y Eugenio Figueredo goza de libertad provisional y dijo haber devuelto lo que tenía que devolver"*

En estos casos en los que se puede ver la respuesta, no es necesario cliquear para leer la nota para resolver la incertidumbre y no se debe considerar, por lo tanto, como clickbait.

En este capítulo propuse una definición de clickbait, luego de repasar las definiciones existentes junto con sus contradicciones y problemas. Además, propuse una definición operativa y criterios de identificación más precisos que permiten disminuir la subjetividad en las decisiones sobre la clasificación de un texto como clickbait o no. Pese a las diferencias existentes en la definición, hay muchos trabajos en la literatura sobre detección de clickbait. En el siguiente capítulo se repasa el estado del arte con especial atención en las características y modelos usados para la detección y los conjuntos de datos utilizados, especialmente sus características de construcción.

Capítulo 3

Clasificación de clickbait: contexto y antecedentes

Existen diversos y numerosos trabajos sobre clasificación de clickbait, muchos de ellos surgidos a raíz del Clickbait Challenge en 2017 (Potthast, Gollub, Hagen y Stein, 2018). Este capítulo los repasa, enfocándose primero en los datos utilizados y luego en los modelos, características y métodos utilizados y los resultados obtenidos.

3.1. Conjuntos de datos disponibles

Todo problema de clasificación en texto natural necesita de un conjunto de datos, como mínimo para validar la calidad de los resultados. Desde el año 2016 a la fecha distintos trabajos han publicado conjuntos de datos. Esta sección recopila los conjuntos disponibles, además de exponer y discutir sus distintas características.

3.1.1. Métodos de recopilación

La clasificación de clickbait destaca por la heterogeneidad en sus conjuntos de datos disponibles, y en los usados en los trabajos de detección. La forma de recopilar los textos a usar incide sustancialmente en la calidad de los datos generados y por lo tanto en la posible interpretación de los resultados obtenidos. Hay ocasiones en las que se busca un criterio de representatividad y generalidad en los textos y hay veces en las que se obtienen las noticias de forma de facilitar el etiquetado, sacrificando la calidad final.

Potthast, Gollub, Komlossy et al. (2018) identifican tres categorías en las que clasificar los procesos de recopilación de datos y construcción de datasets de clickbait:

Basados en reputación Se recopilan noticias de medios con buena reputación y se clasifican todas como no-clickbait. Por otro lado, se seleccionan ciertos medios y se toman todas sus publicaciones como clickbait.

Basados en intermediarios Se utiliza algún foro o cuenta de redes sociales que se dedique a marcar o arruinar clickbait para generar los ejemplos positivos y cuentas de noticias generales para los ejemplos negativos.

Basados en importancia Se toman las noticias más relevantes según un determinado criterio y no se asume la categoría sino que se etiqueta manualmente.

En los dos primeros casos la etiqueta se obtiene implícitamente en la construcción de los datos, lo que facilita el proceso y probablemente eso es lo que los hace dos métodos muy populares. Luego, en algunas ocasiones se repasan manualmente para corregir posibles errores. Sin embargo, como explica Agrawal (2016) ese etiquetado está sesgado por el hecho de que los textos ya están previamente asignados a una categoría.

Además, estos dos métodos comparten un problema: los ejemplos que son clickbait no difieren únicamente en esta característica de los demás. En el caso del método basado en reputación se trata de diferentes medios que usualmente tienen distintas formas de escribir y presentar las noticias, temas que cubrir, etc. En el caso de los basados en intermediario también difieren en los criterios con los que los distintos intermediarios eligen qué noticias y de qué medios publicar. Estas diferencias hacen que los modelos de aprendizaje automático puedan diferenciar las dos clases por características que no tienen que ver con el clickbait, obteniendo muy buenos resultados que no reflejan en realidad la capacidad de detectar clickbait (este análisis se profundiza en la Subsección 3.2.1).

3.1.2. Binario vs continuo

Otra característica que difiere entre varios de los conjuntos de datos existentes es si se considera la etiqueta como un binario *Clickbait* vs *No clickbait*

o si se permiten más categorías intermedias o incluso un valor continuo que matice la decisión.

El clickbait tiene un componente subjetivo. Los titulares deben ser concisos, no pueden tener toda la información y es normal evitar hechos obvios y conocidos por todos. La línea que divide lo obvio de lo que no, lo que es conocido de lo que se oculta para generar una brecha de información es fina, la percepción no es la misma para todos y depende fuertemente del contexto (véase la Sección 2.5).

Con el objetivo de considerar estos niveles subjetivos en algunos trabajos se etiqueta en más de dos categorías. El Clickbait Challenge (Potthast, Gollub, Hagen y Stein, 2018) decidió considerar cuatro categorías: *not*, *slightly*, *considerably* y *heavily* clickbait. Incluso hay quienes modelan el problema, utilizando este conjunto de datos, como uno de regresión (Cao et al. 2017; Indurthi et al. 2020; Wiegmann et al. 2018). Indurthi et al. (2020) argumenta que permite «predecir la intensidad del clickbait con fineza» y Cao et al. (2017) explican que su modelo comete la mayoría de errores en anotaciones que se encuentran entre 0.27 y 0.66, los casos más ambiguos que «son desafiantes incluso para los humanos». C. Wu et al. (2020) etiquetan un conjunto de datos con valores del 0 al 5 argumentando que «un grado bajo de clickbait puede animar la atmósfera en ciertas condiciones, mientras que un alto grado de clickbait puede generar indignación en el lector. Por lo tanto, es necesario e importante llevar a cabo una detección de clickbait a un nivel de granularidad fino, en lugar de una clasificación binaria.»

Pese a que los argumentos para afirmar que el clickbait es una tarea continua son válidos, puede trazarse un paralelismo con otras tareas subjetivas del procesamiento de lenguaje natural como el análisis de sentimiento. Si bien existen muchos conjuntos de datos no-binarios, existen también los binarios y son de mucha utilidad. En primer lugar porque muchas de las decisiones que se toman con los resultados, tales como *mostrar o no este tweet, y ahorrar al lector de la curiosidad que le genera*, son intrínsecamente binarias. Además, porque a partir de datos binarios se pueden, y es lo más común, entrenar modelos que entregan una probabilidad o puntaje continuo que permite lidiar con el resto de las decisiones.

En este trabajo consideré el etiquetado como binario porque, además, para captar las subjetividades es necesario que el etiquetado sea ejecutado por un colectivo amplio, y eso quedó fuera del alcance del proyecto.

3.1.3. Corpus existentes

Existen múltiples conjuntos de datos sobre clickbait de diversas características. La cantidad de noticias, la fuente de los datos, la escala de anotación, el idioma y principalmente el método de recopilación y etiquetado (véase Subsección 3.1.1) son muy importantes para poder evaluar debidamente los resultados de los trabajos que los utilizan. En la Tabla 3.1 se recopilan varios de los conjuntos de datos más utilizados, especialmente los disponibles, en una actualización de la tabla presentada por Potthast, Gollub, Hagen y Stein (2018).

Paper	Idioma	Platforma	Método	Tamaño	Anotación	Año	Disponible
Potthast et al. (2016)	Inglés	Twitter	Importancia	2.992	Binario	2016	Sí
Agrawal (2016)	Inglés	Reddit, Facebook, Twitter	Intermediarios	2.388	Binario	2016	Sí
Chakraborty et al. (2016)	Inglés	Sitios web	Reputación	32.000	Binario	2016	Sí
Karadzhov et al. (2017)[1]	Búlgaro	Sitios web	Importancia	2.815	Binario	2017	Sí
Potthast, Gollub, Komlossy et al. (2018)	Inglés	Twitter	Importancia	38.517	4 niveles	2018	Sí
Wongsap et al. (2018)[2]	Tailandés	Sitios web	Reputación	5.000	Binario	2018	Sí
H.-T. Zheng et al. (2018)	Chino	Sitios web, WeChat, BBS	Importancia	14.922	Binario	2018	Sí
Mathur (2019)	Inglés	Sitios web	Reputación	12.000	Binario	2019	Sí
Kaur et al. (2020)	Inglés	Reddit, Facebook	Intermediarios	1.800	Binario	2020	Sí
William y Sari (2020)	Indonesio	Sitios Web	Importancia	15.000	Binario	2020	Sí
Genç y Surer (2023)	Turco	Twitter	Intermediarios	48.060	Binario	2021	Sí
Marreddy et al. (2021)	Télugu	Sitios web	Reputación	112.657	Binario	2021	Sí
Broscoţeanu y Ionescu (2023)	Rumano	Sitios web	Importancia	8.313	Binario	2023	Sí
TA1C	**Español**	**Twitter**	**Importancia**	**3.500**	**Binario**	**2024**	**Sí**
Biyani et al. (2016)	Inglés	Yahoo	Reputación	4.073	Binario	2016	No
Rony et al. (2017)	Inglés	Facebook	Reputación	1.67 M	Binario	2017	No
Chakraborty et al. (2017)	Inglés	Twitter	Reputación	288.000	Binario	2017	No
Kongyoung et al. (2019)	Tailandés	Sitios web	Reputación	132.938	Binario	2018	No
Indurthi (2020)	Español	Sitios web	Reputación	60.000	Binario	2020	No
Li et al. (2022)	Chino	Sitios web	Importancia	3.300	4 niveles	2021	No
Al-Sarem et al. (2021)	Árabe	Twitter	Importancia[3]	12.321	Binario	2021	No[4]

Tabla 3.1: Conjuntos de datos de detección de clickbait

Como es común en el área del procesamiento de lenguaje natural la mayoría de los trabajos se realizan sobre textos en inglés. Sin embargo, existen publicaciones donde detectan o analizan el clickbait en otros idiomas, como el chino (Li et al. 2022; Lu y Pan, 2021; Zhang et al. 2020; H.-T. Zheng et al. 2017), holandés (Kuiken et al. 2017a), danés (Blom y Hansen, 2015), búlgaro (Karadzhov et al. 2017), árabe (Al-Sarem et al. 2021), turco (Geçkil et al. 2020; Genç y Surer, 2023), tailandés (Kaothanthong et al. 2021; Kongyoung et al. 2019; Wongsap et al. 2018), indonesio (Fakhruzzaman et al. 2021; William y Sari, 2020), telugu (Marreddy et al. 2021), húngaro (Vincze y Szabó, 2020), filipino (Dimpas et al. 2017) y el rumano (Broscoţeanu y Ionescu, 2023).

No hay, hasta donde tengo conocimiento, datasets públicos y disponibles

[1] El conjunto de datos fue construido como uno de noticias falsas, donde su definición está mezclada con la de clickbait.

[2] Es un subconjunto del publicado luego en (Kongyoung et al. 2019) que no está disponible públicamente

[3] Aplicando «técnicas de aprendizaje de pseudo-etiquetado»

[4] La URL fue reportada pero no funciona en este momento

de clickbait en español. Varios trabajos como (Araujo et al. 2021; Bazaco et al. 2019; García Orosa et al. 2017; Palau-Sampio, 2016) estudian el clickbait en medios en español aunque de forma cualitativa, sin crear un conjunto de datos cuantitativamente significativo.

Indurthi (2020) sí genera un conjunto de 60 mil noticias en español utilizando el método basado en reputación. Las noticias de clickbait se toman de Buzzfeed (en español)[1] y los ejemplos negativos de los diarios El Mundo y El País. Sin embargo, como se explica en la Subsección 3.1.1 este método de recopilación genera que las diferencias entre los textos de una y otra categoría difieran en más atributos que ser clickbait o no. En este caso, los distintos estilos periodísticos, de redacción, el tipo de noticias que se elige cubrir y hasta el componente regional son claramente diferenciadores entre los dos conjuntos. Por ejemplo, los textos que hablen sobre regiones de España tienen más probabilidades de ser de El Mundo que de Buzzfeed independientemente de si alguno de ellos es clickbait. Además, los medios elegidos para la clase negativa sí recurren al clickbait (Bazaco et al. 2019) por lo que las dos categorías se pueden confundir.

3.2. Detección automática de clickbait

Las diferentes definiciones del concepto de clickbait que conviven y la ausencia de un consenso (véase la Sección 2.1) dificultan la comparación de los distintos trabajos y la capacidad de reproducir sus resultados en distintos conjuntos de datos. Los resultados obtenidos dependen en gran medida de la calidad y características de los datos, por lo que no es posible comparar la efectividad de los métodos aplicados. Este panorama tiene la notable excepción de los artículos que utilizan el dataset del Clickbait Challenge que se ha posicionado como lo más similar a un conjunto estándar de detección de clickbait.

Los problemas resueltos en el campo del procesamiento de lenguaje natural como la detección de spam, la clasificación de temáticas, el etiquetado gramatical (*POS-tagging*) o el reconocimiento de entidades con nombre (*NER*) pueden decidirse utilizando información léxica y morfosintáctica. Sin embargo, otros problemas como la generación de lenguaje (C. Dong et al. 2022), la detección

[1] https://www.buzzfeed.com/mx

de la ironía (Zeng y Li, 2022), del humor (Chiruzzo et al. 2021), de expresiones de odio (Jahan y Oussalah, 2023) o lenguaje abusivo (Alrashidi et al. 2022) están en una etapa más temprana de la resolución pese a los grandes avances ocurridos en los últimos años. Estos problemas presentan un mayor grado de complejidad, lo que implica la necesidad de abordarlos con otros métodos y tener en cuenta el contexto y la pragmática del lenguaje, exigiendo una fineza de interpretación que requiere conocimiento del mundo. Por lo expuesto en el Capítulo 2 la detección de clickbait es un problema que califica en este último grupo.

Un repaso de la literatura permite ver que, como en la gran mayoría de los problemas del área, las primeras soluciones se plantearon mayormente con métodos de aprendizaje automático clásico y características léxicas y sintácticas y luego pasó a dominar el aprendizaje profundo. También que la abrumadora mayoría de los trabajos disponibles se enfocan en el idioma inglés.

3.2.1. Aprendizaje automático clásico

Las publicaciones que utilizan métodos de aprendizaje clásico difieren, además de en los conjuntos de datos, en los modelos y principalmente en las características utilizadas.

En la Subsección 4.1.1 hay un repaso de las características empleadas con su detalle y referencias a algunos de los trabajos que utilizan cada una de ellas. Se puede ver que dominan las características de estilo de escritura (signos de puntuación, longitud, mayúsculas, categorías gramaticales, cantidad de palabras, formalidad, etc). Muchas intentan encontrar algunos rasgos distintivos del clickbait como los listículos, signos de interrogación, intensificación y listas de expresiones comunes (*"te sorprenderá"*) o léxicas (hipérboles, palabras provocativas, etc). Además, algunas apuntan más directamente a detectar las brechas de información, como la búsqueda de demostrativos, pronombres personales o artículos determinados. Otras están específicamente diseñadas para otras definiciones de clickbait, como la presencia de hashtags, el sentimiento del titular, detección de blasfemias o juegos de palabras y la similaridad (o diferencia) entre título y contenido.

Algunas de las características presentes en la literatura se encuentran en artículos que no son sobre detección automática de clickbait, sino de análisis y descripción del fenómeno. Por eso, algunas de ellas, como la detección de voca-

tivos[1] y juegos de palabras (Araujo et al. 2021; Palau-Sampio, 2016), atracción emocional (Mormol, 2019), surrealismo o ambigüedad (García Serrano et al. 2019) y neologismos (Araujo et al. 2021) son difíciles de implementar y podrían merecer un trabajo en sí mismas. Otras características igualmente complejas como la detección de insinuación o hipérboles se pueden intentar resolver de forma sencilla con una lista de palabras como en (Zhang et al. 2020) pero eso no deja de ser apenas una aproximación sobre la que incluso no se reportan resultados obtenidos.

Blom y Hansen (2015), si bien no realizan detección de clickbait, hacen una descripción profunda y detallada de una de sus características más comunes: la referencia anticipada mediante deixis discursivas y catáforas, tema analizado en la Sección 2.3. Encuentran, además, que se usan más en medios comerciales, de noticias blandas y tabloides (que suelen tener también más clickbait) que en los que no.

Potthast et al. (2016) presentaron el primer trabajo de detección automática de clickbait. Construye un dataset de casi tres mil tweets etiquetados a mano. Realizan la detección usando más de 200 características que van desde léxicas (n-gramas, listas de palabras) hasta información sobre los tweets (si tiene un adjunto, hora de envío) y utilizan una Regresión Logística, Naive Bayes y Random Forests, que obtienen el mejor resultado con 0.76 de precision y recall. Este dataset fue luego ampliado en (Potthast, Gollub, Komlossy et al. 2018) para crear el conjunto de datos del Clickbait Challenge.

Otro trabajo inicial en el tema fue (Chakraborty et al. 2016) donde utilizan árboles de decisión, Random Forests y SVM, que obtiene los mejores resultados con un 93 % de F1-score. El conjunto de datos usado fue obtenido basado en reputación (véase Subsección 3.1.1) donde las noticias de clickbait se toman de medios como Buzzfeed, Upworthy y Viralstories y las no-clickbait de Wikinews. La Figura 3.1 presenta un cuadro, tomado del trabajo, que muestra los 20 sujetos gramaticales más comunes para cada categoría y se ve cómo la mayor diferencia entre una y otra categoría son los temas tratados y no la presentación de la noticia, mucho menos si existe una brecha de información.

Las diferencias temáticas y también de estilo entre diferentes medios que conforman este tipo de corpus hacen que los métodos de aprendizaje automático puedan distinguir entre los dos conjuntos por atributos que nada tienen que ver con el clickbait. No es correcta, además, la suposición de que ciertos me-

[1]Dicho de una expresión nominal: Que se usa en función apelativa. Fuente: RAE

WNG	Score
how to	0.11
as	0.11
sex	0.11
dies	0.11
5	0.11
that	0.13
an	0.13
to	0.14
...	0.15
in	0.15
here	0.15
things	0.18
heres	0.20
is	0.28
at	0.40
a	0.57
!	0.82
how	0.83
these	1.10
?	1.38
.	1.50
this	3.16

Clickbait	I, you, dog, everyone, girls, guys, he, here, it, kids, men, mom, one, parent, photos, reasons, she, something, that, they
Non-clickbait	bomb, court, crash, earthquake, explosion, fire, government, group, house, U.S., China, India, Iran, Israel, Korea, leader, Obama, police, president, senate

Figura 3.1: A la izquierda, los principales N-gramas (*Word N-Grams*) para distinguir clases en el conjunto de datos del Clickbait Challenge presentado por (Wiegmann et al. 2018). Se puede ver como la mayoría refieren a estilos de escritura y no a temáticas, como sucede con conjuntos de datos construidos por el método basado en reputación. A la derecha, los 20 sujetos gramaticales más comunes para cada categoría (clickbait o no) en (Chakraborty et al. 2016). Las mayores diferencias son los temas sobre los que tratan las noticias de cada categoría, en lugar de la existencia de una brecha de información.

dios publican únicamente clickbait: Potthast, Gollub, Komlossy et al. (2018) reportan el porcentaje de clickbait por medio y sólo uno de ellos supera el 50 % y no es Buzzfeed, uno de los tres sitios *clickbait* en (Chakraborty et al. 2016). Pero incluso aunque publicaran únicamente clickbait, esas métricas no reflejan la capacidad de los modelos de detectarlo sino, en realidad, de detectar si un titular pertenece a un conjunto de medios u otro, un problema distinto y más sencillo de resolver. La Figura 3.1 muestra un cuadro presentado en Wiegmann et al. (2018) donde se ven los n-gramas que mejor distinguen clickbait en datos anotados por humanos, como los del Clickbait Challenge, donde mayoritariamente no se trata de palabras que refieran a temas sino a formas de escribir.

Los trabajos con datos de este tipo reportan excelentes resultados, algunos incluso con niveles de error en accuracy y F1-score menores al 3 % o incluso al 1 % (Anand et al. 2017; Brogly y Rubin, 2018; Bronakowski et al. 2023; Kongyoung et al. 2019; Manjesh et al. 2017; Marreddy et al. 2021; Pujahari y Sisodia, 2021; Rony et al. 2017; Singh Sisodia, 2019; Venneti y Alam, 2018; Wongsap et al. 2018), pero todos estos resultados deben tomarse con la misma precaución.

Varios artículos utilizan el conjunto de datos del Clickbait Challenge (Potthast, Gollub, Komlossy et al. 2018), por lo que sí se pueden comparar sus resultados. Cao et al. (2017) obtiene 0.61 de F1-score con Random Forests y agregando a las características más habituales, surgidas del artículo original, algunas apuntando a la relación entre el título y el contenido. Khater et al. (2018) reporta 0.54 F1-score (en la clase positiva) con un SVM sobre características léxicas y de formalidad y legibilidad del texto. Grigorev (2017) aplica un ensamblado de modelos SVM, cada uno de ellos aplicado a cada texto (título, tweet, artículo, etc). Wiegmann et al. (2018) se enfoca en la selección de características para un modelo de regresión, con un modelo Ridge, obteniendo buenos resultados.

Algunos trabajos no obtienen buenos resultados, incluso siendo superados por el baseline del desafío, pero pueden destacarse algunas propiedades: Coste et al. (2020) trabaja con características que no dependen del lenguaje como signos de puntuación y distintos conteos de largo y cantidad de palabras, caracteres, etiquetas gramaticales (*POS tags*) y más. El enfoque es interesante pero se prueba únicamente con los datos del Clickbait Challenge donde no superan el baseline. Papadopoulou et al. (2017) prueban una clasificación en dos niveles con 65 regresiones logísticas sobre distintas características de texto e imágenes y luego una sobre estos resultados, que llega apenas al .43 de F1-score. Plantean el uso de modelos de detección de objetos en la imagen que acompaña la noticia para agregarlo como feature. Liao et al. (2019) intenta resolver el problema con un enfoque federado.

Biyani et al. (2016) crean un dataset basado en ocho tipos de clickbait que, en algunos casos, difiere de la definición de este trabajo. Utilizan Gradient Boosting Decision Trees (GBDT) sobre características tanto del título como del artículo y consiguen un .60 de F1-score en la categoría positiva. El concepto de los ocho tipos influenció otros trabajos como el de Al-Sarem et al. (2021) donde crean un conjunto de datos de once mil noticias de twitter en árabe. Modelan con características léxicas, datos de los tweets y similaridad título-contenido y el mejor resultado lo logran con un SVM.

H.-T. Zheng et al. (2017) agregan características del «comportamiento del usuario», como la información de qué usuarios hicieron click y el desempeño de los links, además de las léxicas y sintácticas aunque logran una mejora muy menor al punto porcentual. Trabajan con un conjunto de datos de más de 43.000 noticias en chino etiquetadas manualmente.

3.2.2. Aprendizaje profundo

La progresión de trabajos en detección de clickbait utilizando aprendizaje profundo siguió naturalmente las tendencias en la clasificación de texto y procesamiento de lenguaje natural. Primero con fuerte importancia de los vectores densos de palabras con modelos convolucionales o recurrentes, luego incluyendo mecanismos de atención y finalmente con transformers y aprendizaje semi-supervisado.

La primera aplicación de la que tengo conocimiento del aprendizaje profundo a este problema específico fue la de Agrawal (2016), donde se aplica una red neuronal convolucional (*CNN*) con tamaños de 3, 4 y 5 palabras sobre vectores de word2vec (Mikolov et al. 2013). Además, crea un conjunto de datos basado en intermediarios que, si bien fue re-etiquetado a mano, tiene el mismo inconveniente que los basados en reputación: los temas y estilos de las noticias de una clase y otra difieren en otras características además de ser clickbait o no, que los clasificadores pueden utilizar para distinguir entre ellas aumentando las métricas sin necesidad de aprender realmente el problema de fondo. Fue, sin embargo, muy influyente. William y Sari (2020), por ejemplo, presentan un dataset de 15.000 noticias etiquetadas en indonesio y utilizan como clasificador una réplica de lo hecho en este trabajo.

Existen otros artículos basados en CNNs, entre los que se destaca el de H.-T. Zheng et al. (2018), donde además de word2vec agregan el tipo de artículo (noticia, blog, o cuentas de BBS o WeChat). También construyen un dataset, en este caso de noticias en chino, aunque lamentablemente no se especifica el método de etiquetado por lo que es difícil interpretar las métricas reportadas.

En cuanto a redes neuronales recurrentes (*RNN*), Anand et al. (2017) utilizan redes neuronales convencionales, LSTMs y GRUs, todas bidireccionales sobre vectores de word2vec en el dataset de (Chakraborty et al. 2016). Obtienen resultados de más de 0.98 de F1-score, superando al trabajo original, aunque nuevamente es importante recalcar que los datos fueron obtenidos por reputación.

Muchas otras publicaciones siguieron el camino de las RNNs. En Omidvar et al. (2019) aplican los tres mismos modelos pero al dataset del Clickbait Challenge obteniendo un buen resultado de F1-score de 0.67. Kaur et al. (2020) usan los vectores de GloVe (Pennington et al. 2014) con una combinación de CNNs y RNNs y crean otro conjunto de datos basado en intermediario. Pandey

y Kaur (2018) crean un nuevo dataset por intermediario y utilizan una LSTM con sus hiperparámetros elegidos con un algoritmo genético. Además de los vectores de palabras utilizan características léxicas y sintácticas. Naeem et al. (2020) usan LSTMs sumando etiquetas gramaticales a los vectores de palabras y lo aplican sobre un conjunto de datos por intermediario tomado de Reddit. También Dimpas et al. (2017) usan una LSTM con word2vec en un conjunto de datos creado especialmente que mezcla inglés con filipino. Sin embargo, no reportan los criterios de etiquetado y, en particular sabiendo que se descargan noticias por intermediario, hace imposible comparar sus resultados con otros clasificadores.

Otro trabajo influyente es el de Rony et al. (2017), donde crean un corpus de más de 1.6 millones de noticias tomadas de Facebook, etiquetado por reputación. Luego, crean vectores densos de palabras utilizando word2vec (Skip-Gram) y clasifican con un softmax sobre el promedio de los vectores de cada noticia. Obtienen resultados de más de 98 % de ejemplos bien clasificados, aunque, nuevamente, hay que interpretar los resultados teniendo en cuenta la forma de creación de los datos. Similarmente, Indurthi y Oota (2017) aplica una regresión logística a vectores de Glove sumado algunas características léxicas y sintácticas al conjunto de datos del Clickbait Challenge, obteniendo .65 de F1-Score.

Luego de la publicación de Vaswani et al. (2017) y algunos otros artículos posteriores como Devlin et al. (2019) y Yang et al. (2019) las arquitecturas basadas en Transformers pasaron a ser el estado del arte en muchos problemas de procesamiento de lenguaje natural. La detección de clickbait no fue una excepción. Y. Zhou (2017) presenta un modelo basado en redes recurrentes GRU bidireccionales con una capa de self-attention, lo aplica al Clickbait Challenge y logra .68 de F1-Score en la clase positiva y .86 de accuracy, ganando la competencia. También Wei y Nguyen (2022) generan un modelo basado en *self-attention* junto con *joint-attention* entre los textos de título y contenido con el objetivo de que aprenda si son coherentes. Lo aplican al Clickbait Challenge obteniendo muy buenos resultados con un .89 de accuracy. C. Wu et al. (2020) combinan los valores de un transformer sobre el texto del artículo y otro sobre el título, un transformer a nivel de carácter sobre el título con el objetivo de encontrar patrones de estilo, y un valor resultante de aplicar co-atención a los vectores de título y artículo para encontrar similaridad entre ellos. Luego, una red densa clasifica usando esos cuatro puntajes. Logran muy buenos resultados

con .89 de accuracy y .73 de F1-score en el Clickbait Challenge (probando en el conjunto de validación, sin hacer una entrega oficial). Finalmente, J. Zheng et al. (2021) aplican un modelo de soft-attention sobre una LSTM unido a una comparación de las distancias de los vectores generados por transformers para título y contenido del artículo (para medir similaridad del contenido) por una red neuronal feed-forward. Lo aplican al Clickbait Challenge, con buenos resultados de .88 en accuracy. También lo aplican a un conjunto propio de datos en chino de cinco mil noticias etiquetado manualmente, aunque no se explica con qué criterios ni cómo se seleccionaron las noticias.

Por su parte, muchos trabajos se basan en modelos de Transformers preentrenados, aprovechando las ventajas del aprendizaje por transferencia, que permite reducir el tiempo de entrenamiento y la cantidad de datos requeridos para una tarea. Indurthi et al. (2020) prueban muchos de los modelos más utilizados como Elmo (Peters et al. 2018), Universal sentence encoder (Cer et al. 2018), BERT (Devlin et al. 2019), RoBERTa (Zhuang et al. 2021) y GPT-2 (Radford et al. 2019) utilizando los datos del Clickbait Challenge para realizar un *fine-tuning* y aplicar distintos modelos de regresión; vale la pena recordar que la principal métrica del desafío es error cuadrático (*MSE*), una métrica de regresión. Obtienen excelentes resultados, con la combinación de RoBERTa con Gradient Boosted Regression llegando a .74 de F1-score en la clase positiva y colocándose al tope del challenge también en MSE. Por su parte, Rajapaksha et al. (2021) también entrenan sobre algunos modelos preentrenados y exploran varias formas de utilizar los vectores generados. Nuevamente RoBERTa genera los mejores resultados, aunque deciden mezclar los conjuntos de entrenamiento y validación del clickbait challenge y usarlos para entrenar y probar con el conjunto de datos de Chakraborty et al. (2016). Otro artículo que hace uso del transfer learning es el de Fakhruzzaman et al. (2021) donde usan un BERT multilenguaje para clasificar el conjunto de datos en indonesio de William y Sari (2020) obteniendo muy buenos resultados pese a que, como descartan los datos que no fueron etiquetados de igual forma por todos los anotadores, se corre riesgo de estar trabajando con un subconjunto más fácil que el original. También Marreddy et al. (2021) crean un conjunto de datos en Telugu de más de cien mil noticias tomadas por intermediario aunque reetiquetadas manualmente. Luego usan características léxicas y POS tags con muchos modelos y también una gran cantidad de embeddings entre los que Elmo (Peters et al. 2018) junto con un LightGBM (Ke et al. 2017) genera los

mejores resultados.

Otros artículos se basaron en este tipo de modelos pero agregando alguna característica particular en el intento de mejorar los resultados. M. Dong et al. (2019) le ponen especial énfasis a la diferencia entre título y contenido, comparando vectores calculados sobre los dos textos y también entre distintos bloques de estos mismos vectores para capturar «distancias locales». También Oliva et al. (2022) hacen énfasis en esa distancia. Por otro lado Setlur (2018) ataca el problema de la escasez de datos generando un modelo semi supervisado, que usando sólo un subconjunto de las etiquetas prácticamente logra los mismos resultados que con todas ellas: .68 de F1-score en el clickbait challenge con el 100 % y .67 y .66 con el 50 % y 30 % de las etiquetas respectivamente. Thomas (2017) usa el clickbait challenge pero se enfoca en características que no dependen del idioma: tokeniza únicamente por los espacios (no sirve para algunos idiomas como el chino, pero sí para la mayoría) y aplica una LSTM a vectores de palabras inicializados aleatoriamente. La ventaja es que el modelo podría ser fácilmente extensible a muchos idiomas, aunque no obtiene buenos resultados en el inglés. Glenski et al. (2017) agrega a las características de texto habituales un vector de características de las imágenes de las noticias generado con ResNet (He et al. 2016), además de utilizar una red de detección de objetos para agregar más características. Las pruebas en los datos del clickbait challenge, sin embargo, no dan buenos resultados. Kumar et al. (2018) agregan redes siamesas para encontrar diferencias entre texto e imágenes de título/*post* y artículo/imágen de contenido distintos vectores obtenidos de redes preentrenadas, sin obtener tampoco buenos resultados. M. Zhou et al. (2022) entrenan una red convolucional, y también usando atención, sobre un Grafo de Conocimiento creado a partir de las entidades (sustantivos) de las noticias junto con *atributos* como adjetivos, adverbios y palabras auxiliares. Lo prueban sobre un conjunto de datos creado específicamente con más de 8 mil noticias en chino, aunque no se brindan muchos detalles sobre el proceso de etiquetado más allá de que lo hacen 7 «expertos». Alcanzan un 0.69 de F1-score, ligeramente mejor que una LSTM, aunque por ser datos distintos los resultados no son comparables. Los datos tienen la particularidad de tener etiquetas del 0 al 5 en lugar de binarias y el modelo la ventaja de brindar explicabilidad en las decisiones. Por su parte Jain et al. (2021) se centran en el uso de las imágenes, usan datos tomados de Instagram junto al subconjunto de las noticias del Clickbait Challenge que tienen imagen, combinan un vector de la imagen generado con

ResNet con uno del texto y aplican modelos clásicos entre los que XGBoost (T. Chen y Guestrin, 2016) genera los mejores resultados.

Algunos artículos, al igual que con el aprendizaje clásico, trabajan con conjuntos de datos que no representan (por lo menos no sólamente) clickbait contra no clickbait, sino otras características como el estilo de los medios que generaron las noticias. Especialmente con técnicas de aprendizaje profundo que tienden a sobreajustar a los datos, se generan resultados cercanos al 100 % de correctitud, pero que no son un reflejo de un modelo perfecto de detección. Es el caso de Karadzhov et al. (2017) donde crean un conjunto de datos que tiene un 98 % de correlación en clickbait con noticias falsas y varios artículos que utilizan los datos de en el dataset de Chakraborty et al. (2016), Kaur et al. (2020) donde se agregan datos también recolectados por reputación, Anand et al. (2017), López-Sánchez et al. (2018) y Chawda et al. (2019); muchos de ellos ya mencionados previamente.

En este capítulo repasé los trabajos y conjuntos de datos disponibles en detección automática de clickbait. Debido a la falta de un conjunto de datos en español y de uno que se ajuste a las definiciones y criterios explicados en el Capítulo 2, construí uno que cumpliera con ambas condiciones. Este conjunto de datos, además de uno no-etiquetado, se describe en el siguiente capítulo.

Capítulo 4

Conjuntos de datos

Una de las contribuciones de este trabajo es la publicación de dos conjuntos de datos de noticias en español. Uno de ellos, llamado TA1C (por *Te Ahorré Un Click)* es el primer dataset público de clickbait en español. Contiene 3500 noticias, cada una de ellas etiquetada como clickbait o no siguiendo la definición y los criterios de identificación introducidos en el Capítulo 2. El segundo conjunto tiene más de medio millón de noticias sin etiquetar y puede usarse para ampliar TA1C, para etiquetar otros problemas de texto sobre noticias en español o para aplicar un *finetuning* de un modelo de lenguaje y adaptarlo a noticias en español.

Existen muchas distintas formas de recolectar noticias y criterios para categorizarlas como clickbait por lo que la construcción de estos conjuntos implicó la toma de numerosas decisiones. Este capítulo detalla la construcción de los conjuntos de datos, el etiquetado y los criterios empleados.

4.1. Recolección de datos

Los conjuntos de datos fueron concebidos para representar todo el lenguaje español y sus medios más influyentes. Este proceso requiere la elección de qué datos incluir, su recolección y selección así como la elección de criterios para la categorización, temas a tratar en las siguientes secciones.

El primer paso de la creación de un corpus es la definición de qué datos reunir. En este trabajo, algunas características vienen dadas por la propia definición del problema, como el idioma y la necesidad de recopilar noticias de prensa en internet.

4.1.1. Características

La selección de textos intenta cumplir con cinco características:

Noticias online El clickbait es un fenómeno, aunque presente en otras disciplinas Sección 2.8, principalmente de las noticias escritas y de Internet, por lo que los datos son noticias de prensa online.

Representativo del español El trabajo es en español y por lo tanto también los textos del conjunto de datos. Además, es importante que los medios elegidos sean un conjunto representativo, sin sesgos hacia ninguna variante, temática ni zona específica, para representar todo el idioma español.

Basado en importancia Los medios se eligen por su importancia y representatividad y no por tener más o menos clickbait. Luego, las noticias se clasifican individualmente, evitando los sesgos y problemas explicados en el Capítulo 3.

Medios masivos y prestigiosos Se priorizan los medios con más llegada, ya sea en ventas como en redes sociales. Se evitan los medios nacidos específicamente para el clickbait y estrategias similares tales como Buzzfeed, debido a que sus estilos son distintos de los medios que incluyen noticias duras. Sí se incluyen las secciones de noticias blandas de los periódicos seleccionados.

Medios tradicionales y nativos digitales Más allá del punto anterior, la presencia en el corpus de medios nativos digitales junto con los tradicionales genera mayor heterogeneidad y representatividad.

4.1.2. Twitter/X

Ambos conjuntos de datos se componen de artículos tomados de Twitter[1], con toda la información que integra el tweet sumado a la que hay en el artículo linkeado. Obtener los datos de esta red social tiene algunas ventajas: la limitación rigurosa de las formas en las que se presenta una noticia, lo que aporta homogeneidad a los datos; la presencia histórica de los posteos y, hasta hace un tiempo atrás, el acceso a una API que permite recopilarlos convenientemente.

La mayoría de los medios tienen una presencia masiva en la plataforma y suelen colocar allí todas sus noticias por lo que no hay riesgo de generar sesgo

[1] Ahora llamada X

hacia cierto tema o tipo de noticia. Además, es cada vez más común consumir noticias a través de las redes sociales en lugar de hacerlo directamente desde los sitios web de cada medio (Amoedo-Casais et al. 2023; Nic et al. 2023; Pew Research Center, 2022), y es allí donde más se genera la competencia por la atención. Por estas mismas razones muchos de los datasets más utilizados del tema en otros idiomas son también tomadas de Twitter (véase Tabla 3.1).

Descarté todos los retweets, formando parte de los conjuntos de datos únicamente los tweets directamente publicados por las cuentas oficiales. Además descarté todos los tweets que no tienen un link hacia una página del medio, como los que contienen sólo texto, imágenes y/o video o dirigen hacia la página principal u otros contenidos. Al igual que en Potthast, Gollub, Komlossy et al. (2018) tomé en cuenta un único link por tweet y descarté aquellos en los que la extracción de contenido no fue exitosa. Algunas de las razones de extracciones fallidas incluyen páginas especiales (para eventos específicos, fotogalerías, secciones, etc), bloqueos al acceso al texto debido a artículos exclusivos para suscriptores y noticias en las que el la noticia linkeada fue actualizada luego de la publicación del tweet, como es el caso de

(86) *Uruguay-Brasil, por las eliminatorias sudamericanas: horario, TV y formaciones*

donde el tweet promete datos de la previa del partido, pero el link conduce a una nota titulada *Canto a la efectividad: Brasil atacó dos veces y derrotó a Uruguay en el Centenario* escrita luego de terminado el evento.

Consideré a la hora de etiquetar todos los textos presentes tanto el titular del link a la noticia como el texto del tweet, pero no el contenido multimedia que está en el tweet. Por un lado, con el objetivo de simplificar el problema. Además, dependiendo incluso del dispositivo y según si hay imágenes o no y otros factores, a veces el título de la noticia y el contenido multimedia se ve completamente en Twitter y otras veces no por lo que fue necesario unificar el criterio.

No se consideran las imágenes, ni las del artículo (que a veces se ven en el avance) ni las embebidas en el tweet para etiquetar la noticia, incluso si contienen texto. Esta decisión genera algunos problemas en tweets como

(87) *#ATENCIÓN — Messi y un comunicado que pone en alerta al FC Barcelona*

o el que muestra la Figura 4.1 donde los textos de los tweets podrían no ser clickbait pero las imágenes tienen escritos los textos «¿Messi se va del Barcelona?» y «¿Sin saludos y abrazos?» respectivamente, donde ambos son clickbait claramente y no forman parte de la decisión de etiquetado, la utilización de un OCR[1] quedó fuera del alcance del proyecto. De todas formas los casos son mínimos y cuando la imagen tiene un clickbait el texto que lo acompaña lo suele tener también.

4.1.3. Medios

Figura 4.1: Ejemplo de tweet en el que si se tiene en cuenta exclusivamente el texto no se trata de clickbait, pero si se mira la imagen sí.

Los medios elegidos para formar parte de los conjuntos de datos deben contar con algunas características para adecuarse a lo descrito anteriormente: estar escritos en español; disponer de presencia en la web con un formato que permita la descarga, al menos, del texto y tener al menos una cuenta en Twitter donde publiquen sus noticias.

Además, deben tener algunas cualidades como que se trate de medios nacionales o internacionales (no regionales o locales), generales (no restringidos a un tema) y populares, en particular en su versión digital.

La empresa Comscore elabora un ranking de los medios más leídos en español en la web que, si bien no es público, se puede acceder en algunas notas de prensa[1] y los ocho medios cumplen con todas características deseadas y se encuentran incluidos en ambos conjuntos de datos.

De la misma forma, los conjuntos de datos también incluyen a los 10 primeros medios del ranking de *4imn*[2] algunos de los cuales ya están comprendidos en el ranking de Comscore (Infobae, Clarín, La Nación y El Universal). Excluí a Las Últimas Noticias que ocupa el tercer lugar pero, al momento de descar-

[1] Reconocimiento óptico de caracteres del inglés *Optical Character Recognition*.

[1] En marzo, Infobae fue el sitio de habla hispana más visto del mundo, Comscore reveló que Infobae es la web de noticias en español más consumida

[2] Top 100 Newspapers in Latin America by Web Ranking

gar las noticias, no era posible hacerlo debido a que se subían escaneadas del papel y O Estado de São Paulo que ocupa el séptimo lugar pero es en portugués. Para ocupar sus lugares, agregué los diarios El Comercio de Ecuador, El Comercio de Perú, El Tiempo de Colombia y Prensa Libre de Guatemala.

Luego, con el objetivo de aumentar la representatividad geográfica, los corpus incluyen un medio de cada uno de los siguientes países hispanoparlantes:

Venezuela El Universal, primero en el ránking *4imn* entre medios venezolanos.

Chile La Tercera, segundo medio chileno en el ránking *4imn* y tercero en lectura papel más digital[1] en ambos casos luego de medios que sólo ofrecen sus noticias web escaneadas del papel, lo que imposibilita su participación en este trabajo.

República Dominicana Diario Libre, número 11 en el ránking *4imn*, primero dominicano.

El Salvador La Prensa Gráfica, número 13 en el ránking *4imn*, primero salvadoreño.

Uruguay El País, número 33 en el ránking *4imn*, primero uruguayo.

Finalmente, incluí a BBC en español, por ser un medio reconocido internacionalmente y consumido desde todos los países de habla hispana.

En todos los casos consideré la cuenta principal de Twitter de cada medio, pero también cuentas secundarias, normalmente referidas a secciones temáticas o regionales, que se listan en el Apéndice 3. Es común que en algunas secciones, principalmente las referidas a *noticias blandas*, se utilice más el clickbait como muestra (Palau-Sampio, 2016).

4.2. TA1C: Descripción del conjunto de datos etiquetado

El conjunto de datos TA1C está disponible en https://github.com/gmordecki/TA1C. Se compone de 3500 tweets, una muestra tomada aleatoriamente con la garantía de que cada uno de los 18 medios contribuye con al

[1] Valida Chile

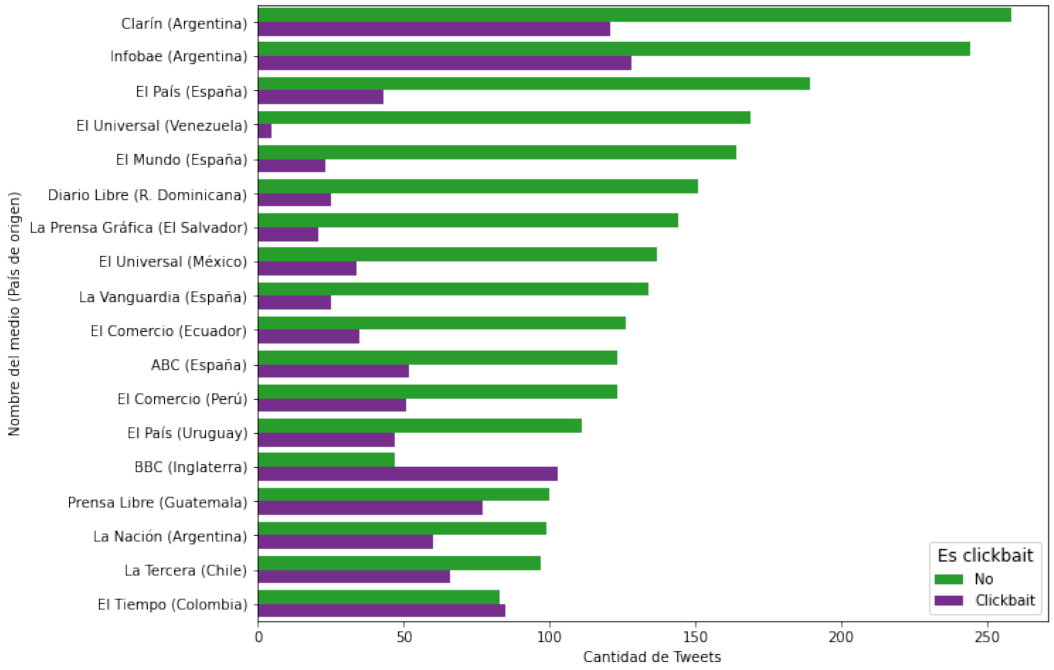

Figura 4.2: Cantidad de tweets en el conjunto de datos TA1C según el medio de origen y según la etiqueta. Se puede notar como la proporción de noticias clickbait varía según el medio.

menos 150 noticias. Esos 3500 están divididos aleatoriamente, aunque manteniendo el balance entre ejemplos positivos y negativos, en un conjunto de entrenamiento de 2100 y dos conjuntos de validación y test de 700 cada uno.

Fue etiquetado por tres anotadores independientes: el autor de esta tesis además de un investigador en procesamiento de lenguaje natural y una licenciada en comunicación. Las anotaciones se hicieron siguiendo estrictamente la definición operativa de clickbait y los criterios explicitados en la Sección 2.5. Primero etiquetamos una muestra de 100 ejemplos y luego discutimos los ejemplos que se habían anotado distintos para asegurar que fueran diferencias de subjetividad y no de criterio. De esta forma, pudimos generar un conjunto de datos con un alto acuerdo entre anotadores Fleiss' κ de 0,825.

El etiquetado es binario (véase discusión en la Subsección 3.1.2) y la etiqueta final se decide por mayoría. Sin embargo, se puede acceder a la etiqueta de cada anotador.

En la Figura 4.2 se aprecia la distribución de noticias entre cada uno de los medios y también la proporción de clickbait encontrada en cada uno. El total de clickbait en el conjunto de datos es de 29 %.

Publiqué dos versiones que contienen los mismos tweets y etiquetas y di-

fieren únicamente en la cantidad de información que los acompaña. La versión mínima contiene:

- ID del tweet
- Fecha del tweet
- Nombre del medio
- País de origen del medio
- Texto del adelanto (Combinación de título y texto del tweet descrita en la Subsección 5.1.1)
- Etiqueta de clickbait

La versión completa añade:

- Cuenta del tweet
- URL del tweet
- URL de imágen del tweet (si existe)
- Texto del tweet
- URL del artículo vinculado
- Título del artículo extraído
- Subtítulo del artículo extraído
- Texto limpio extraído del artículo (sólo el cuerpo de la noticia)
- Imágenes extraídas del artículo con sus descripciones (en una lista JSON)
- URL incorporadas extraídas del artículo (generalmente enlaces a redes sociales)
- Etiqueta de cada anotador

4.3. TA1C-unlabeled: Descripción del corpus sin etiquetar

El conjunto de datos *TA1C-unlabeled* reúne todos los tweets, junto con la noticia que lo acompaña, descargados en el proceso de recolección de datos sin procesamiento ni etiqueta. Estos datos pueden usarse para clickbait o para cualquier tarea de procesamiento de lenguaje natural sobre noticias y/o redes sociales de medios de comunicación en idioma español.

Tiene 568.128 datos descargados entre octubre de 2020 y octubre de 2021, por lo que la mayoría de las noticias son de ese rango de fechas. Sin embargo,

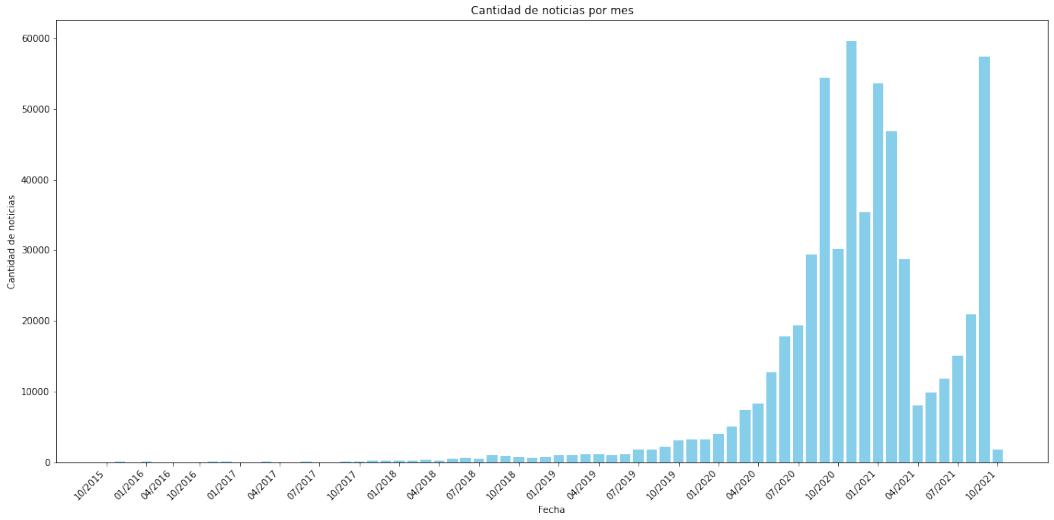

Figura 4.3: Cantidad de tweets en el conjunto de datos TA1C-unlabeled por fecha de publicación. Se puede notar que la mayoría de las noticias pertenecen al año transcurrido entre octubre de 2020 y octubre de 2021.

dependiendo de la frecuencia de publicación de cada cuenta, se recuperaron algunos tweets anteriores, incluso desde octubre 2015. En la Figura 4.3 se puede ver la cantidad de tweets (y noticias) descargadas en cada mes.

La cantidad de noticias por cada medio depende totalmente de su frecuencia de publicación, no hay ningún tipo de elección de noticias ni procesamiento destinado a generar balance. En la Figura 4.4 se describe la cantidad de noticias por medio de prensa, siendo la totalidad de medios los 18 descritos anteriormente en la Subsección 4.1.3.

Cada fila del conjunto de datos contiene

- ID del tweet
- Fecha del tweet
- Texto del tweet
- Nombre del medio
- País de origen del medio
- URL del artículo enlazado
- Titular del artículo
- Subtítulo del artículo
- Texto limpio extraído del artículo (sólo el cuerpo de la noticia)

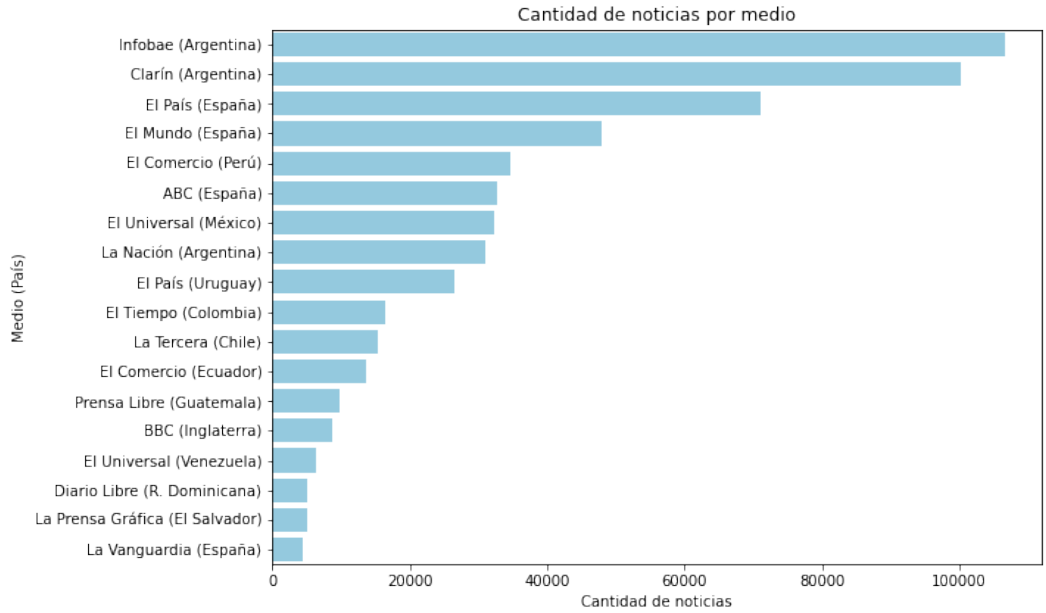

Figura 4.4: Cantidad de tweets en el conjunto de datos TA1C-unlabeled según el medio de origen. La cantidad de noticias varía dependiendo de su frecuencia de publicación en Twitter en el momento de descargados los datos.

En este capítulo repasé los conjuntos de datos que generé y publiqué en este trabajo, junto con las definiciones tomadas en su construcción. En el siguiente capítulo describo los métodos usados y los resultados obtenidos en clasificación de clickbait en español usando el conjunto TA1C.

Capítulo 5

Clasificación de clickbait y resultados

Este capítulo detalla los métodos utilizados y los resultados obtenidos en la clasificación automática de clickbait en noticias en español, usando el corpus especificado en el capítulo anterior. Primero define el diseño experimental y el preprocesamiento de los datos, luego especifica los baselines y los métodos aplicados, tanto de aprendizaje tradicional como de aprendizaje profundo. Finalmente, repasa los resultados junto con un análisis de los errores cometidos por el modelo elegido.

5.1. Diseño experimental

Gran parte de las decisiones que típicamente se toman en el diseño experimental fueron tomadas con la construcción del corpus y descritas en el Capítulo 4. Por ejemplo, la partición del conjunto de datos en tres conjuntos de train (2100 tweets), uno de validación (700) y otro de test (700).

Todos los modelos entrenan únicamente con la partición de entrenamiento y la de validación se reserva para elección de hiperparámetros y modelado (por ejemplo, para seleccionar el mejor modelo de *zero-shot* entre los varios disponibles). Finalmente, reporto la ejecución por única vez del modelo de mejor rendimiento en el conjunto de test.

Por otro lado, descarté en este trabajo la utilización de algunos datos como la imagen que acompaña la noticia, el texto del artículo y las fechas y estadísticas de los tweets por simplicidad y para asegurar la extensibilidad de

los resultados a otras plataformas.

5.1.1. Preprocesamiento

Todas las soluciones, resultados y análisis presentados aquí fueron posteriores a un preprocesamiento común de los datos, excepto cuando sea explicitado.

Tanto para el título de la noticia como para el texto del tweet, el preprocesamiento consiste en:

- Sustitución de emojis por la palabra *EMOJI*
- Sustitución de hashtags *#Columna*, *#Editorial* y *#Opinión* por las palabras *Columna*, *Editorial* y *Opinión*, respectivamente
- Sustitución de hashtags restantes por la palabra *HASHTAG*
- Eliminación del texto "*RT—por—vía @cuenta*"
- Sustitución de cuentas (*@cuenta*) por la palabra *NombrePropio*
- Eliminación de caracteres extraños y espacios múltiples
- Eliminación de URLs

Luego, el texto utilizado surge como una combinación de ambos. Twitter da la posibilidad de mostrar una previsualización de un link o no, así como de acompañar la publicación con una imagen. Muchas veces los periódicos utilizan estas herramientas para no revelar la información clave y, por ejemplo, no mostrar previsualización si el título estropea la brecha de información que genera el posteo. Por lo tanto, la selección de qué textos utilizar no es trivial y un criterio único puede generar inconsistencias con lo que lee el usuario. Sin embargo, ante la imposibilidad técnica de saber cómo lo ven los usuarios (entre otras cosas porque depende del dispositivo que tengan) y por consistencia, la combinación se hace siempre con los siguientes criterios:

- Si el título está incluido en el texto del tweet o el texto el tweet es más largo que el título y la distancia de edición (o distancia de Levenshtein[1]) es menor que 5 o la similaridad de Levenshtein normalizada, calculada por la biblioteca textdistance[2] es mayor que 0.9, se usa el texto del tweet.
- Si el texto del tweet está incluido en el título o el título es más largo que el texto el tweet y la distancia de edición es menor que 5 o la similaridad de Levenshtein normalizada es mayor que 0.9, se usa el título.

[1] Levenshtein distance
[2] textdistance

- En caso contrario, se usa el título seguido por el texto del tweet.

5.1.2. Métricas

El problema a resolver es de clasificación binaria y relativamente balanceada, con una relación de aproximadamente 29 % a 71 % entre la cantidad de ejemplos positivos y negativos respectivamente.

Considero que la métrica más representativa de la calidad de los modelos para este problema particular es la precisión promedio (*average precision* en inglés), definida en la Ecuación 5.1, donde R_n y P_n se refieren a la precisión y el recall en el n-ésimo umbral respectivamente.

$$\text{AP} = \sum_n (R_n - R_{n-1}) P_n \tag{5.1}$$

La precisión promedio resume la curva de *precision-recall* mediante un promedio de las precisiones obtenidas en los distintos umbrales de clasificación, ponderadas por el incremento en el *recall* que se obtiene en cada uno. Esta implementación es similar pero no exactamente igual a calcular el área bajo la curva de *precision-recall* (también conocida como *AUPRC*, por sus siglas en inglés) interpolando linealmente, que puede generar resultados «demasiado optimistas» (Scikit Learn Documentation, 2024). Es representativa de la calidad de los modelos porque resume la posible compensación entre *precision* y *recall*, permitiendo entender qué tan bien funciona la detección en diferentes contextos de uso. Para un sistema como el creado en Chakraborty et al. (2016), una extensión de navegador para bloquear lo que se detecta como clickbait y ocultarlo al usuario, se precisa alto *recall* aunque venga de la mano de bloquear algunas noticias más de las necesarias. O quizás podría ser una configuración que elija el usuario. Para obtener clickbaits ya sea para estudiarlos, publicarlos en algún foro específico de clickbaits, o para responderlos (manual o automáticamente) se necesita muy buena precisión, aunque se sacrifique *recall*. La precisión promedio muestra cómo se comportaría el modelo en estos distintos casos.

Además, con el fin de poder comparar los resultados con otros trabajos que no suelen reportar la precisión promedio, reporto *precision* (definida en la Ecuación 5.2), *recall* (Ecuación 5.3, sensibilidad en español), *F1-score* (Ecuación 5.4), otra forma de combinar *precision* y *recall*, y *accuracy* (Ecuación 5.5,

precisión en español, pero le llamo *accuracy* para evitar la confusión con presicion)

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5.2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5.3}$$

$$\text{F1-score} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.4}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.5}$$

Donde, en todos los casos:

- **TP** (True Positives): Verdaderos Positivos, el número de textos positivos correctamente clasificadas como positivos.
- **FP** (False Positives): Falsos Positivos, el número de textos negativos incorrectamente clasificadas como positivos.
- **TN** (True Negatives): Verdaderos Negativos, el número de textos negativos correctamente clasificadas como negativos.
- **FN** (False Negatives): Falsos Negativos, el número de textos positivos incorrectamente clasificadas como negativos.

5.2. Métodos propuestos

Del repaso de la literatura sobre detección de clickbait surge que al igual que en muchos problemas del procesamiento de lenguaje natural conviven soluciones de aprendizaje automático clásico con otras que utilizan aprendizaje profundo en alguna de sus variantes, como se profundiza en el Capítulo 3. Basado en esta observación y con el objetivo de cubrir todo el espectro de alternativas propuestas previamente, implementé modelos de ambos tipos, además de combinaciones entre ellos.

5.2.1. Baselines

Presento cinco modelos de línea base para proporcionar puntos de referencia en la evaluación de las soluciones. Sus resultados se encuentran en la Tabla 5.1.

Zero-shot Utilización de varios de los modelos más utilizados en clasificación de texto zero-shot para el idioma español o multilenguaje. La tarea es, dado el texto, responder a la hipótesis *Este ejemplo es clickbait*. Debido a que para estos modelos en particular podría influir negativamente, presento los resultados con y sin pre-procesamiento de texto. El rendimiento de los modelos es muy malo, lo que puede explicarse por la poca *familiaridad* de los modelos con el concepto de clickbait, seguramente no muy presente en los textos de entrenamiento, y por la dificultad inherente del problema.

Tf-idf + Modelo lineal Uno de los modelados más utilizados en problemas de clasificación de texto. TF-idf con n-gramas de hasta 3 tokens para generar las características y luego un modelo lineal, tanto Regresión Logística como SVM.

Tf-idf + Gradient Boosting Trees Otro modelo muy utilizado como línea base. La misma forma de generar las características pero aplicando luego árboles de decisión con gradient boosting utilizando la librería XGBoost (T. Chen y Guestrin, 2016) con búsqueda de hiperparámetros usando hyperopt (Bergstra et al. 2013). Al igual que en Zero-shot, hice pruebas con y sin preprocesamiento, y se puede ver que el proceso aporta buenos resultados.

Fasttext Clasificación con la librería fastText(Joulin et al. 2017) en su versión de clasificación de texto supervisado, con los hiperparámetros elegidos por el método de *autotuneValidation*.

Large Language Models Clasificación con los modelos de lenguaje ofrecidos por OpenAI[1] y Cohere AI[2]. En el caso de OpenAI a través de GPT-4 (OpenAI et al. 2024) usando el prompt de sistema *You are a helpful assistant. You are designed to understand text in English and Spanish and answer in English. Your answer is just one word and one number: yes or no, followed by the probability that the answer is yes* y el prompt de usuario *Is the headline "{}" clickbait? Answer* con el texto del avance. Si bien GPT-4 está disponible sólo para inglés, OpenAI asegura que es «lo suficientemente robusto para generar buenos resultados para una variedad de idiomas» (OpenAI, 2024).

En el caso de Cohere, por un lado usando la generación de texto con

[1] https://openai.com/
[2] https://cohere.com/

un prompt muy similar: *Your answer is just one word and one number: yes or no, followed by the confidence that you have in your answer, expressed in a decimal between 0 and 1. For example "Yes, {decimal}" or "No, {decimal}". Is the headline "{}" clickbait? Answer*. Por otro lado, utilizando su opción de *few-shot learning*. Envié los 3 primeros ejemplos de cada etiqueta del conjunto de entrenamiento como ejemplos, usando el modelo *embed-multilingual-v3.0*. En todos los casos, ejecuté pruebas con y sin el pre-procesamiento de los textos y utilicé una temperatura de 0.2.

Los malos resultados de ambos modelos pueden reflejar la complejidad inherente del problema así como la dificultad para conciliar los distintos criterios de clickbait que, con seguridad, aparecen en los textos de entrenamiento de los modelos y su probable diferencia con el criterio tomado en la construcción del conjunto de datos en este trabajo.

Modelo	Preprocesamiento	Average Precision	Precision	Recall	F1-score	Accuracy
xlm-roberta-large-xnli[1]	Sí	0.26	0.26	0.16	0.20	0.62
xlm-roberta-large-xnli	No	0.26	0.26	0.15	0.19	0.63
Zero-shot SELECTRA[2]	Sí	0.30	0.34	0.07	0.11	0.69
Zero-shot SELECTRA	No	0.30	0.34	0.08	0.13	0.69
mDeBERTa-v3-base-xnli-multilingual-nli-2mil7[3]	Sí	0.28	0.30	0.06	0.11	0.68
mDeBERTa-v3-base-xnli-multilingual-nli-2mil7	No	0.28	0.33	0.07	0.12	0.69
XLM-RoBERTa-large-XNLI-ANLI[4]	Sí	0.32	0.56	0.01	0.05	0.71
XLM-RoBERTa-large-XNLI-ANLI	No	0.31	0.47	0.04	0.08	0.71
Tf-idf + Logistic Regression	Sí	0.73	**1.00**	0.11	0.20	0.74
Tf-idf + SVM	Sí	0.73	0.81	**0.55**	**0.65**	**0.83**
Tf-idf + XGboost	Sí	0.74	0.85	0.48	0.61	0.82
Tf-idf + XGboost	No	0.68	0.77	0.44	0.56	0.80
fastText	-	**0.76**	0.77	**0.55**	0.64	0.82
LLM - GPT4 - ZeroShot	Sí	0.49	0.71	0.25	0.37	0.75
LLM - GPT4 - ZeroShot	No	0.49	0.72	0.21	0.32	0.75
LLM - Cohere AI - Text Generation	Sí	0.35	0.61	0.09	0.16	0.72
LLM - Cohere AI - Text Generation	No	0.35	0.50	0.05	0.10	0.71
LLM - Cohere AI - Few Shot Classification	Sí	0.25	0.43	0.47	0.45	0.66
LLM - Cohere AI - Few Shot Classification	No	0.25	0.42	0.47	0.44	0.66

Tabla 5.1: Resultados de los modelos baseline aplicados al conjunto de datos TA1C

5.2.2. Aprendizaje automático clásico

Los métodos de aprendizaje automático clásico dependen en gran medida de la extracción de características. En esta sección describo brevemente las características utilizadas, una combinación de algunas surgidas de la revisión literaria y otras de la exploración de los datos, luego la selección de cuáles

[1] https://huggingface.co/joeddav/xlm-roberta-large-xnli
[2] https://huggingface.co/Recognai/zeroshot_selectra_medium
[3] https://huggingface.co/MoritzLaurer/mDeBERTa-v3-base-xnli-multilingual-nli-2mil7
[4] https://huggingface.co/vicgalle/xlm-roberta-large-xnli-anli

utilizar en el modelo y finalmente los distintos modelos experimentados y sus resultados.

5.2.2.1. Características

La mayor parte de las características usadas en este trabajo resultan de la revisión de la literatura detallada en la Subsección 3.2.1. Algunas de ellas refieren a propiedades que distinguen al clickbait con otra definición, por lo que las descarté para utilizar con este conjunto de datos y definición del problema. Otras no refieren directamente a la creación de una brecha de información, pero pueden aportar información por reconocer eventos correlacionados al clickbait. A su vez, hay características que son pensadas específicamente para otros idiomas, generalmente el inglés, por lo que algunas de ellas requieren una traducción y otras no pueden traducirse a datos en español.

Como regla general las características son pensadas para tener alta precisión, sin importar si su *recall* es bajo. Existen distintos tipos de clickbait e idealmente cada característica podría identificar uno de ellos, para así recuperar todos los casos con muchas de ellas.

A continuación se detalla la construcción de las características utilizadas.

5.2.2.2. Características categóricas

Etiqueta gramatical (*POS tag*) de la primera palabra Característica categórica, convertida con *one-hot encoding* a varias binarias que guarda información de la etiqueta gramatical de la primera palabra. Información obtenida usando Freeling (Lloberes et al. 2010; Padró y Stanilovsky, 2012).

5.2.2.3. Características binarias

En la Figura 5.1 se muestra un análisis de algunas de las características binarias donde se ve la cantidad de ejemplos que son clickbait en función de si la característica binaria es verdadera o falsa. Además, se puede ver el porcentaje de clickbait dado que la característica es verdadera o falsa, donde grandes diferencias muestran el valor predictivo de cada característica.

Hipérboles Una forma de llamar la atención muy asociada al sensacionalismo pero también al clickbait (Chakraborty et al. 2016; Mormol, 2019; Zhang

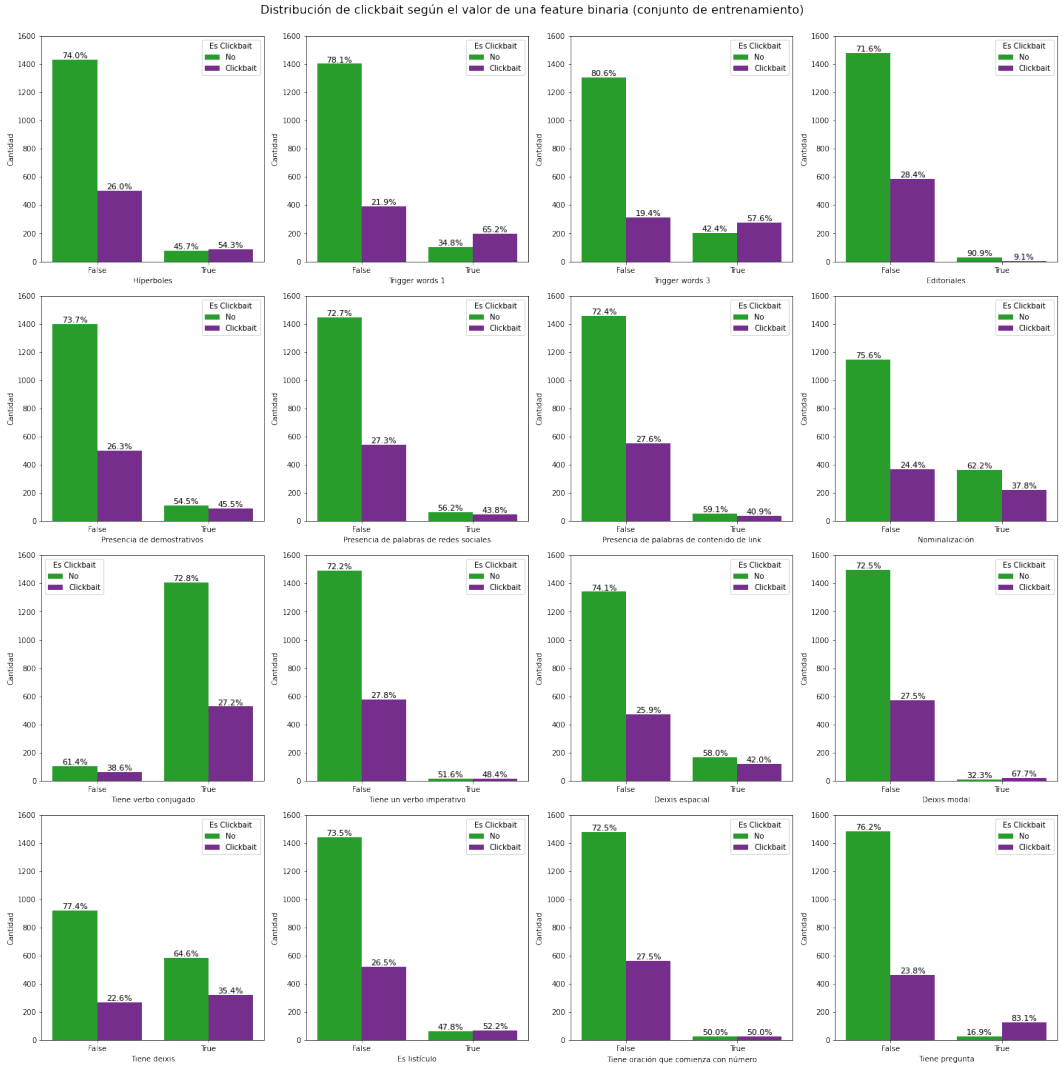

Figura 5.1: Proporción de clickbait en función de si una característica binaria es positiva o negativa, para algunas características seleccionadas. Se puede ver su potencial predictivo dado por el aumento en la proporción de clickbaits cuando la característica es verdadera (o falsa, en el caso de *Tiene verbo conjugado* y *Editoriales*).

et al. 2020). Creado mediante la construcción de una lista de adjetivos y adverbios del conjunto de entrenamiento con frecuencia mayor a 4 apariciones y curada a mano para sólo quedar con los que pueden considerarse hipérboles. La lista completa es: *peligroso, viral, inédito, crítico, insólito, impactante, increíble, único, peor, fuerte, espectacular, fatal, malo, curioso, mucho, polémico, impresionante, emocionante, récord, falso, sorprendente, gran, trágico, así.*

Trigger words 1 Listado de palabras con las que suelen comenzar los titulares clickbait. Luego de ordenar todas las palabras iniciales por porcentaje de clickbait (de estar presentes, qué porcentaje de los artículos son clickbait), seleccioné manualmente, las que se utilizan para generar curiosidad. Por ejemplo, el token *cristiano_ronaldo* aparece en la lista pero quedó descartado por referirse al contenido. La lista completa es: *cómo, qué, quién, conocer, por, este, así, si, video, aquí, haber, acá, día, poder, sin, NUMBER, le, ser.*

Trigger words 3 Listado de palabras que suelen formar parte de las primeras tres palabras de los titulares clickbait. Luego de ordenar por porcentaje de clickbait (de estar presentes, qué porcentaje de los artículos son clickbait), seleccioné manualmente, las que se utilizan para generar curiosidad, al igual que en Trigger words 1. La lista completa es: *qué, quién, cuál, mejor, cuánto, cómo, conocer, video, así, si, explicar, este, te, responder, saber, tomar, ser, historia, pasar, revelar, por, acá, mostrar, le, aquí, poder, contar, estar, para.*

Editoriales Persigue el objetivo de distinguir los tweets que refieren a columnas de opinión o editoriales. Como se detalla en la Subsección 2.5.8 suelen contener menos clickbait que el resto de las noticias. Se compone por la presencia de alguna de las siguientes condiciones:

- Contiene los lemas *editorial, columna, opinión.*
- Contiene bigramas *opina* o *escribe* + nombre propio (recordando que los nombres de cuenta de Twitter, representados por las arrobas, se convierten en nombres propios en el preprocesamiento).
- Contiene *al director*, por el caso de las cartas al director.

Presencia de demostrativos Una de las manifestaciones del clickbait descritas en Blom y Hansen (2015) que traduje al español como presencia de *este, ese, aquel, estos, esos, aquellos, esta, esa, aquella, estas, esas,*

aquellas, esto, eso, aquello.

Presencia de palabras de redes sociales Intenta captar las notas que prometen una publicación en una red social. Se define como si tiene alguna de las siguientes expresiones *video, vídeo, imagen, foto, compartir, seguidor, instagram, facebook, influencer, twitter, postear, fotografía, tiktoker, tiktok, instagramer, memes, red social, tik tok.*

Presencia de palabras de contenido de link Intenta captar las notas que prometen un video, habitualmente a jugadas deportivas, trailers de películas, gráficos, etc. Se define como si tiene alguna de las siguientes expresiones *video, vídeo, imagen, foto, gol, jugada, asistencia, pase, trailer, tráiler, gráfico, gráfica.*

Clickbait de cita La idea es captar si se trata de un clickbait que genera una brecha de información creada por la ausencia de una declaración. Se construye en base a dos categorías de verbos:

Los fuertes Alcanza con su presencia para que la característica sea positiva. Son: *revelar, confesar.*

Los débiles Sólo cuentan si están en tercera persona del pasado y tienen como sujeto a un nombre propio. Esto descarta algunos casos que sería bueno incluir, como *El ministro de Justicia admite* o *La Fiscalía desveló los audios*, pero descarta expresiones como *Estudio relata que restricción a parrillero no reduce el crimen*, que son falsos positivos y los superan ampliamente en cantidad, por lo que no cumplen con la idea de que las características tengan alta precisión. Son: *admitir, contar, denunciar, expresar, indicar, manifestar, pronunciar, reaccionar, reconocer, reflexionar, relatar, testificar, testimoniar, acusar, anunciar, desvelar, comunicar, mensaje.*

Nominalización Intenta detectar titulares que se componen únicamente por un sintagma nominal o que no contienen verbo conjugado. Basado en lo propuesto por Palau-Sampio (2016) y explicado en la Sección 2.3. Son cuatro características en total, implementadas como la búsqueda de textos cuya raíz del árbol de dependencias, creado por Freeling (Lloberes et al. 2010) es un sustantivo y otra si es un verbo sin conjugar. Existe, para cada variante, una característica que captura si alguna oración cumple esta condición y otra si todas las oraciones lo hacen.

Tiene verbo conjugado Otra forma de captar la nominalización, aquí se

espera que la no presencia de verbo conjugado indique nominalización y correlacione con el clickbait.

Preposición + verbo no conjugado Inspirada en la característica «formas no personales del verbo, especialmente las introducidas con la preposición para, que están ligadas a los listados» de Palau-Sampio (2016). La defino como la presencia de una preposición seguida de un verbo no conjugado.

Oraciones subordinadas en el grupo nominal Inspirada en la característica «dentro del sintagma nominal, los adjuntos se manifiestan como oraciones subordinadas» de Palau-Sampio (2016). Para simplificar, la defino como la presencia de un grupo nominal que tiene una oración subordinada.

Encadenado de sintagmas preposicionales Tomado de Palau-Sampio (2016). Definida como un sintagma preposicional seguido de otro o un sintagma preposicional hijo de otro en el árbol de dependencias.

Tiene un verbo imperativo Tomado de Araujo et al. (2021) y Palau-Sampio (2016). Intenta captar las situaciones en las que se le habla directamente al lector, explicadas en la Subsección 2.5.7.

Deixis espacial Uno de los tipos de fórmulas deícticas utilizadas para generar brechas de información (Araujo et al. 2021; Blom y Hansen, 2015; Li et al. 2022; Palau-Sampio, 2016). Definido como la presencia de demostrativos o de los adverbios de lugar *aquí, allí, ahí, allá, acá, arriba, abajo, cerca, lejos, adelante, delante, detrás, encima, debajo, enfrente, atrás, alrededor*.

Deixis temporal Uno de los tipos de fórmulas deícticas utilizadas para generar brechas de información (Araujo et al. 2021; Blom y Hansen, 2015; Li et al. 2022; Palau-Sampio, 2016). Definido como la presencia de los adverbios temporales *antes, después, luego, pronto, tarde, temprano, todavía, aún, ya, ayer, hoy, mañana, anteayer, siempre, nunca, jamás, próximamente, prontamente, anoche, enseguida, ahora, mientras, anteriormente*.

Deixis modal Uno de los tipos de fórmulas deícticas utilizadas para generar brechas de información (Araujo et al. 2021; Blom y Hansen, 2015; Li et al. 2022; Palau-Sampio, 2016). Definido como la presencia del adverbio demostrativo *así*.

Deixis personal Uno de los tipos de fórmulas deícticas utilizadas para ge-

nerar brechas de información (Araujo et al. 2021; Blom y Hansen, 2015; Li et al. 2022; Palau-Sampio, 2016). Definido como la presencia de un pronombre personal o de un posesivo, tanto determinantes, pronombres como adjetivos.

Tiene deixis Presencia de cualquiera de los cuatro tipos de deixis mencionados anteriormente.

Pronombres personales en tercera persona Un tipo específico de deixis personal destacado por Biyani et al. (2016) y Blom y Hansen (2015). Definida como la presencia de *ellos, las, él, ella, les, consigo, le, ello, lo, sí, la, usted, ellas, ustedes, se, los.*

Es listículo Característica diseñada para encontrar los artículos lista (ver la Sección 2.3, definida como la presencia de un número con todas las siguientes condiciones:

- **Está entre 3 y 30** Números muy grandes o muy chicos suelen ser falsos positivos, no existen listículos de una sóla respuesta ni de un millón. Los límites son arbitrarios, elegidos en base a la exploración del conjunto de entrenamiento.
- **Está al principio del titular o luego de un signo de puntuación** Los listículos suelen comenzar con un número, colocarlo al inicio de una oración o muy usualmente luego de unos dos puntos. El límite de *incio de oración* y *luego de dos puntos* son 5 tokens, otro límite arbitrario y elegido en base a la exploración del conjunto de entrenamiento.
- **No es una medida** La siguiente palabra no es una unidad de medida.
- **No está luego de un nombre propio** Con el fin de evitar los falsos positivos generados por, por ejemplo, modelos como *Iphone 11.*

Tiene oración que comienza con número Otra forma de captar los listículos. Es común encontrar la característica «Empieza con número» en muchos trabajos (Biyani et al. 2016; Brogly y Rubin, 2018; Potthast et al. 2016; Wiegmann et al. 2018).

Tiene un número Otra característica relacionada con los listículos.

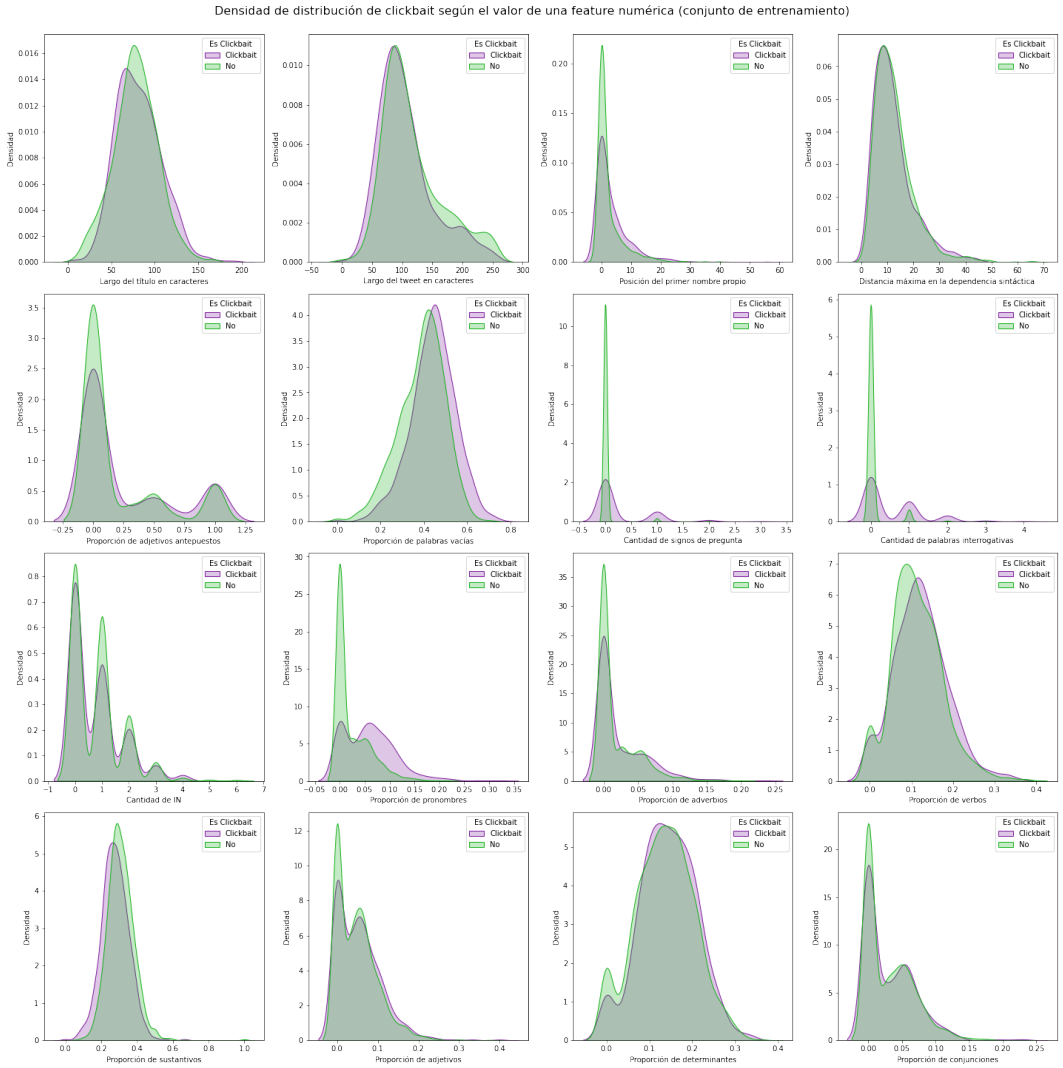

Figura 5.2: Densidad de distribución de clickbait dado el valor de cada característica numérica, según si el texto es clickbait o no. Se puede ver el potencial predictivo de algunas de las características dado por el cambio en las distribuciones.

5.2.2.4. Características numéricas

En la Figura 5.2 se muestra un análisis de algunas de las características numéricas donde se ve la densidad de distribución de la presencia de clickbait según el valor de la característica, distinguiendo si el texto es clickbait o no.

Largo del título en caracteres

Largo del tweet en caracteres

Posición del primer nombre propio Tomada de Brogly y Rubin (2018).

Distancia máxima en la dependencia sintáctica Tomada de Chakraborty et al. (2016). La máxima distancia entre una palabra y su padre en el árbol de dependencias sintácticas. La intuición detrás es que el clickbait tiene oraciones de estructura más compleja y en particular de dependencias largas por el uso de las deixis (*forward-reference*).

Proporción de adjetivos antepuestos Utilizados comúnmente en el clickbait (Araujo et al. 2021; Palau-Sampio, 2016) para generar brecha de información mediante la referencia futura (*forward-reference*). Lo defino como la relación entre cantidad de adjetivos antepuestos y adjetivos totales.

Proporción de palabras vacías La cantidad de palabras vacías (*stopwords*) entre la cantidad de palabras del texto. (Cao et al. 2017; Potthast et al. 2016; Wiegmann et al. 2018; H.-T. Zheng et al. 2017).

Cantidad de signos de pregunta

Cantidad de palabras interrogativas Conteo de las palabras *qué, cuál, cuándo, cuánto, cómo, dónde, quién*, las llamadas *wh words* en inglés. Otra forma de captar la presencia de una brecha de información mediante una interrogante (Biyani et al. 2016; D. Molina et al. 2021). Se implementa antes de la lematización porque no es lo mismo, por ejemplo, la palabra *que* a la presencia de *qué*.

Proporción de etiquetas gramaticales (*POS tags*) La cantidad de ocurrencias de algunas etiquetas gramaticales (adjetivo, adposición, adverbio, conjunción, determinante, interjección, sustantivo, pronombre y verbo) entre la cantidad de palabras del texto. Muchos trabajos incluyen la cantidad, en lugar de la proporción, como característica de al menos alguna de las etiquetas (Brogly y Rubin, 2018; Cao et al. 2017; Chakraborty et al. 2016; Coste et al. 2020; Scott, 2021; H.-T. Zheng et al. 2017).

Cantidad de etiquetas gramaticales IN Cantidad de preposiciones o conjunción subordinantes. Tomada de Cao et al. (2017).

5.2.2.5. Características descartadas del modelo

Implementé pero descarté del modelo algunas de las características surgidas de la literatura por no ajustarse a la definición de clickbait utilizada en este trabajo, por tener alta correlación con otra o por concluir que no aportan valor predictivo al modelo luego de un proceso de selección. A continuación, algunas de las más destacables:

Tf-idf Características léxicas usadas comúnmente en la mayoría de los problemas de clasificación de procesamiento de lenguaje natural y en particular en la detección de clickbait. Sin embargo, su aporte en este problema se restringe a conjuntos de datos creados de formas en las que los temas de las noticias en las diferentes clases difieren, y no es lo que quiero captar aquí, como se explica en el Capítulo 2 y Capítulo 3.

Cantidad de palabras Una característica usada comúnmente en la literatura (Biyani et al. 2016; Brogly y Rubin, 2018; Cao et al. 2017; Chakraborty et al. 2016; Coste et al. 2020; Kuiken et al. 2017a; H.-T. Zheng et al. 2017). Algunos trabajos incluyen también, o en su lugar, el largo en caracteres (Wiegmann et al. 2018).

***Este* + sustantivo** Un tipo específico de deixis muy utilizado en el clickbait.

Cantidad de etiquetas gramaticales NN + IN Bigramas compuestos por un sustantivo y luego una conjunción subordinante o una preposición. Tomada de Cao et al. (2017).

Cantidad de etiquetas gramaticales NNP + VBZ Bigramas compuestos por un nombre propio y luego un verbo en presente y tercera persona del singular. Tomada de Cao et al. (2017).

Cantidad de etiquetas gramaticales IN + NNP Bigramas compuestos por una conjunción subordinante o una preposición y luego un nombre propio. Tomada de Cao et al. (2017).

Cantidad de hashtags

Legibilidad Medidas de legibilidad y formalidad del lenguaje. Si bien refieren a formas de escribir y no a la generación de brechas de información, podría encontrarse correlación y obtuvo buenos resultados como característica en trabajos como (Biyani et al. 2016; Coste et al. 2020; Kuiken

et al. 2017a). Implementada con la librería *textstat*[1] con los cuatro tipos disponibles que aplican al idioma español: Fernandez Huerta, Szigriszt Pazos, Gutiérrez de Polini y Crawford. No aportaron poder predictivo en este caso.

Promedio de largo de palabra Característica muy habitual en modelos de predicción de clickbait, basada en la asunción de que el clickbait usa palabras más cortas (Brogly y Rubin, 2018; Chakraborty et al. 2016; Coste et al. 2020; Kuiken et al. 2017a; Potthast et al. 2016; Wiegmann et al. 2018). Está más relacionada a detectar una forma distinta de escribir que a la generación de una brecha de información.

Promedio de largo de bigramas y trigramas Similar al promedio de largo de palabra pero considerando dos o tres palabras contiguas. De las características con mejores resultados en Brogly y Rubin (2018), aunque no en este trabajo.

Largo máximo de palabra Tomado de Brogly y Rubin (2018).

Tiene comillas

Tiene paréntesis

Tiene palabras genéricas Contiene las palabras *cosa* o *hacer*. Sugerida en Palau-Sampio (2016). Podría expandirse la lista en futuras iteraciones buscando que obtenga mejores resultados.

El texto del tweet incluye el título La generé con la idea de captar los tweets que, pese a tener un título clickbait, contestan la interrogante en el texto que acompaña la noticia, pero no obtuvo los resultados esperados.

Cantidad de emojis

Cantidad de números

Cantidad de signos de exclamación

Tiene dos puntos

Tiene puntos suspensivos Tomada de Palau-Sampio (2016).

Cantidad de expresiones positivas y negativas La cantidad de apariciones de expresiones en el lexicon positivo (o negativo) de Urizar y Roncal (2013). Característica presente en varios trabajos como (Biyani et al. 2016; Kuiken et al. 2017a). Muchos otros trabajos presentan la característica de la polaridad del texto (Cao et al. 2017; Chakraborty et al. 2016, 2017; Potthast et al. 2016; Wiegmann et al. 2018; H.-T. Zheng et

[1] https://github.com/textstat/textstat

al. 2017).

Nombre Propios juntos Una de las características con mejores resultados en Cao et al. (2017).

Cantidad de nombres propios Una de las características con mejores resultados en Cao et al. (2017).

Cantidad de oraciones Por la forma de construcción del texto explicada en la Subsección 5.1.1, depende de si el texto se trata únicamente del título o también había un texto distinto en el tweet, y por lo tanto se correlaciona mucho con la característica *El texto del tweet incluye el título.*

Tiene superlativos Otra forma de llamar la atención asociada al sensacionalismo. También, junto con las hipérboles, intenta captar algunos de los ejemplos de intensificación o «adjetivación intensificadora» (Araujo et al. 2021; Palau-Sampio, 2016). No diferencia entre positivos o negativos como en D. Molina et al. (2021). Definida como superlativos o las palabras *máximo* o *mínimo*.

Alguna oración comienza con un adverbio Tomado de Biyani et al. (2016) como otra forma de detectar referencia anticipada.

Presencia de artículos definidos Una de las manifestaciones del clickbait descritas en Blom y Hansen (2015), traducida al español como presencia de *el, la, los, las*.

5.2.2.6. Modelos y resultados

A continuación describo los modelos utilizados:

LogisticRegression Implementación de scikit-learn[1] (Pedregosa et al. 2011). Mejores hiperparámetros: penalty=l2, dual=False, C=10, cw=balanced.

SVM Implementación de scikit-learn[2] (Pedregosa et al. 2011). Mejores hiperparámetros: penalty=elasticnet, C=10, cw=balanced.

RandomForest Implementación de scikit-learn[3] (Pedregosa et al. 2011). Mejores hiperparámetros: criterion=gini, maxDepth=8, n_estimators=200, min_samples_leaf=2.

XGboost Implementación de la librería XGBoost (T. Chen y Guestrin, 2016) con búsqueda de hiperparámetros

[1] Scikit Learn - LogisticRegression
[2] Scikit Learn - SVC
[3] Scikit Learn - RandomForestClassifier

usando hyperopt (Bergstra et al. 2013). Mejores hiperparámetros: n_estimators=460, learning_rate=0.046663522390317604, max_depth=10, gamma=4.44883294386973, reg_alpha=1, min_child_weight=0, colsample_bytree=0.7507576286875736.

Realicé todos los experimentos usando las características antes mencionadas y el conjunto de entrenamiento. Utilicé el conjunto de validación para encontrar los mejores hiperparámetros para cada modelo. En la Tabla 5.2 se muestran los resultados de la mejor combinación de hiperparámetros para cada modelo, evaluados en el conjunto de test. Se puede ver cómo no hay grandes diferencias en los resultados de las dos métricas principales (Average Precision y F1-Score) según el modelo utilizado, aunque sí cambia el balance entre precisión y recall.

Modelo	Average Precision	Precision	Recall	F1-score	Accuracy
SVM	**0.80**	0.63	**0.73**	0.68	0.80
LogisticRegression	**0.80**	0.63	**0.73**	0.68	0.80
RandomForest	0.79	0.73	0.68	**0.70**	0.83
XGboost	**0.80**	**0.78**	0.63	**0.70**	**0.84**

Tabla 5.2: Resultados de los modelos de aprendizaje automático clásicos aplicados al conjunto de datos TA1C. Resultados reportados sobre el conjunto de test, luego de hacer búsqueda de hiperparámetros en el conjunto de validación.

5.2.3. Aprendizaje profundo

En la Subsección 3.2.2, donde se repasa el estado del arte de los métodos de aprendizaje profundo para la detección de clickbait, puede verse que la mayoría de los artículos se basan en vectores densos de palabras y redes recurrentes o convolucionales. Sin embargo, en los últimos años las arquitecturas basadas en el mecanismo de atención consiguieron el mayor interés y, por lo general, los mejores resultados tanto en detección de clickbait como en el procesamiento de lenguaje natural en general. Además, los métodos de aprendizaje profundo anteriores requieren de una gran cantidad de ejemplos para entrenar, una posibilidad que por el método de etiquetado manual que efectué para generar el conjunto de datos TA1C en este trabajo queda fuera del alcance posible. Por esta razones enfoqué la detección mediante aprendizaje profundo en la utilización de *Transformers*, utilizando la biblioteca huggingface[1].

[1] https://huggingface.co

En todos los casos, exceptuando cuando se indica explícitamente, el entrenamiento se trata de instanciar cada modelo de lenguaje, partiendo de los pesos de sus capas obtenidos previamente y permitiendo el aprovechamiento del aprendizaje por transferencia. Luego, se entrena específicamente para la detección de clickbait usando el conjunto de entrenamiento de TA1C evaluando su desempeño en el conjunto de validación para elegir los hiperparámetros, la cantidad de *epochs* y el modelo que obtengan mejores resultados.

A continuación describo los modelos y arquitecturas utilizadas:

MarIA Descrito en Gutiérrez-Fandiño et al. (2022). Modelos de lenguaje entrenados con datos descargados por la Biblioteca Nacional de España de sitios *.es* pertenecientes a los años 2009 a 2019. Utilicé la versión de RoBERTa-base[1], basado en la arquitectura de RoBERTa (Zhuang et al. 2021). Quizás por estar entrenado en textos muy distintos a los de TA1C, el modelo no logró resolver el problema, clasificando todas las noticias como no-clickbait en todos los entrenamientos. Mejores hiperparámetros: batch size 48, dropout 0.3, learning rate 2e-5.

Narrativa - Detección de noticias falsas en español Descrito en Sofía Sánchez González (2021). Modelo entrenado sobre la base de RoBERTa-large de MarIA (Gutiérrez-Fandiño et al. 2022) con el conjunto de datos de Posadas-Durán et al. (2019), un conjunto de noticias en español mexicano de sitios establecidos y confiables, sitios especializados en validar noticias falsas y sitios que regularmente publican noticias falsas. La tarea de detectar noticias falsas tiene varios puntos de contacto con la de detectar clickbait (véase Sección 2.6) por lo que la transferencia del aprendizaje es más directa que en el caso anterior. El modelo de Narrativa, sin embargo, no logra los mejores resultados y además tiene la desventaja de ser mucho más grande y por lo tanto más lento y computacionalmente costoso que el resto de los modelos utilizados. Mejores hiperparámetros: batch size 16, dropout 0.15, learning rate 2e-5.

Beto Descrito en Cañete et al. (2020). Modelo de lenguaje con la arquitectura de BERT-base (Devlin et al. 2019) entrenado en textos en español tomados de Wikipedia, organizaciones gubernamentales, charlas TED, subtítulos, noticias y más fuentes. Utilicé la versión con mayúsculas y minúsculas llamada dccuchile/bert-base-spanish-wwm-cased[2] con el ob-

[1]Disponible en https://huggingface.co/PlanTL-GOB-ES/roberta-base-bne
[2]Disponible en https://huggingface.co/dccuchile/bert-base-spanish-wwm-cased

jetivo de retener la información que las mayúsculas pueden aportar. Mejores hiperparámetros: batch size 32, dropout 0.2, learning rate 6e-5.

Beto FLM (*Fine-tuned Language Model*) Entrenamiento en dos pasos. Primero realicé un finetuning del modelo de lenguaje de Beto (Cañete et al. 2020) utilizando el dataset de TA1C-unlabeled descrito en la Sección 4.3. El entrenamiento usó 500.000 noticias y 80.000 para validación para ajustar el modelo de lenguaje por 10 *epochs*. Los textos fueron preprocesados igual que como se describe en la Subsección 5.1.1. Los hiperparámetros de entrenamiento siguieron lo usado en el entrenamiento original de Beto. Luego, entrené con TA1C sobre este nuevo modelo de lenguaje. Los resultados usando este modelo de lenguaje ajustado a los textos y el lenguaje específico de este conjunto de datos superan consistentemente el entrenamiento con el modelo original. Mejores hiperparámetros: batch size 48, dropout 0.25, learning rate 8e-5.

RoBERTuito Descrito en Pérez et al. (2022). Modelo de lenguaje con la arquitectura de RoBERTa (Zhuang et al. 2021) entrenado en cerca de 500 millones de tweets de usuarios que escriben en español (no necesariamente todos los textos son en español) que consigue muy buenos resultados en variadas tareas en idioma español. Utilicé la versión robertuito-base-uncased[1]. Mejores hiperparámetros: batch size 32, dropout 0.3, learning rate 6e-5.

CohereAI FineTuning Servicio para realizar fine-tuning sobre un large language model[2]. El entrenamiento se realiza en el servidor del servicio por lo que no es posible conocer datos de cómo evolucionaron las métricas como se muestra en la Figura 5.3 para los otros modelos.

Modelo	Average Precision	Precision	Recall	F1-score	Accuracy
MarIA	0.40	0.00	0.00	0.00	0.70
Narrativa	0.88	0.83	0.78	0.80	0.88
Beto	0.90	0.80	**0.89**	0.84	0.90
Beto FLM	0.92	0.85	0.88	**0.87**	**0.92**
RoBERTuito	**0.93**	0.89	0.84	0.86	**0.92**
LLM - Cohere AI - FineTuning	0.89	**0.93**	0.74	0.83	0.91

Tabla 5.3: Resultados de los modelos de aprendizaje profundo aplicados al conjunto de datos TA1C. Resultados reportados sobre el conjunto de test, luego de hacer búsqueda de hiperparámetros en el conjunto de validación.

[1]Disponible en https://huggingface.co/pysentimiento/robertuito-base-uncased
[2]Fine-tuning with the Python SDK

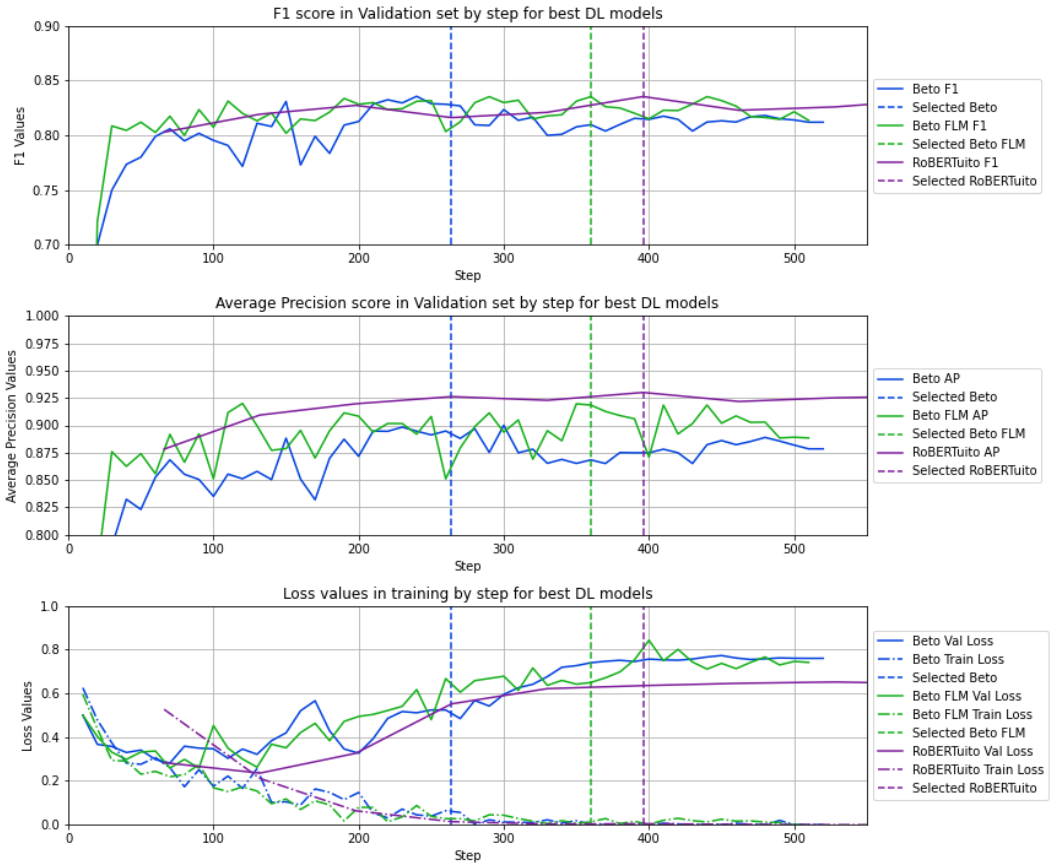

Figura 5.3: F1-score, precisión promedio (*average precision*) y *train loss* de los distintos modelos de aprendizaje profundo según los steps de entrenamiento. Las líneas punteadas horizontales marcan el modelo seleccionado, teniendo en cuenta el balance entre obtener los mejores resultados y evitar el sobreajuste.

En la Tabla 5.3 se reportan los resultados de las pruebas que realicé, todas ellas entrenando con el conjunto de entrenamiento. Todos los entrenamientos fueron de 10 *epochs*, eligiendo luego el modelo en el checkpoint con mejores resultados al evaluarlos en el conjunto de validación y se reportan los resultados de ese mejor modelo sobre el conjunto de test. Los resultados en la evolución del entrenamiento de los mejores tres modelos se pueden ver en la Figura 5.3. En todos los casos los modelos comienzan a sobreajustar y perder calidad predictiva antes de la *epoch* 10, por lo que no fue necesario seguir entrenando.

5.2.4. Método híbrido

Estos modelos combinan las características creadas para el modelado clásico con los vectores aprendidos por los modelos de aprendizaje profundo. La idea

principal es tomar un vector de los modelos descritos en la Subsección 5.2.3, adjuntar las características descritas en la Subsubsección 5.2.2.1 y entrenar un modelo sobre esa combinación.

Modelo	Average Precision	Precision	Recall	F1-score	Accuracy
Beto_FLM - AVG - 1 - XGBoost	0.83	0.87	0.83	0.85	0.92
Beto_FLM - AVG - 4 - XGBoost	0.89	0.88	0.81	0.84	0.91
Beto_FLM - AVG - 9 - XGBoost	0.90	0.89	0.79	0.84	0.91
Beto_FLM - CLS - 1 - XGBoost	0.82	0.87	0.84	0.86	0.92
Beto_FLM - CLS - 4 - XGBoost	0.82	0.86	0.87	**0.87**	0.92
Beto_FLM - CLS - 9 - XGBoost	0.84	0.87	0.87	**0.87**	0.92
RoBERTuito - AVG - 1 - XGBoost	0.90	0.89	0.78	0.83	0.91
RoBERTuito - AVG - 4 - XGBoost	0.89	0.89	0.73	0.80	0.89
RoBERTuito - AVG - 9 - XGBoost	0.89	**0.91**	0.71	0.80	0.89
RoBERTuito - CLS - 1 - XGBoost	0.83	0.89	0.79	0.84	0.91
RoBERTuito - CLS - 4 - XGBoost	0.81	0.89	0.79	0.84	0.91
RoBERTuito - CLS - 9 - XGBoost	0.81	0.89	0.79	0.84	0.91
Beto_FLM - AVG - 1 - LR	**0.93**	0.87	0.86	0.86	0.92
Beto_FLM - AVG - 4 - LR	**0.93**	0.88	0.83	0.85	0.92
Beto_FLM - AVG - 9 - LR	**0.93**	0.87	0.83	0.85	0.91
Beto_FLM - CLS - 1 - LR	0.92	0.87	0.86	0.86	0.92
Beto_FLM - CLS - 4 - LR	**0.93**	0.87	0.87	**0.87**	0.92
Beto_FLM - CLS - 9 - LR	**0.93**	0.87	**0.88**	**0.87**	**0.93**
RoBERTuito - AVG - 1 - LR	0.91	0.90	0.79	0.84	0.91
RoBERTuito - AVG - 4 - LR	0.91	0.90	0.78	0.84	0.91
RoBERTuito - AVG - 9 - LR	0.91	**0.91**	0.76	0.83	0.91
RoBERTuito - CLS - 1 - LR	**0.93**	0.89	0.83	0.86	0.92
RoBERTuito - CLS - 4 - LR	**0.93**	0.89	0.83	0.86	0.92
RoBERTuito - CLS - 9 - LR	**0.93**	0.89	0.83	0.86	0.92

Tabla 5.4: Resultados de los modelos híbridos aplicados al conjunto de datos TA1C. Resultados reportados sobre el conjunto de test, luego de hacer búsqueda de hiperparámetros en el conjunto de validación.

El nombre reportado para cada modelo híbrido se compone de 4 partes que explican su estructura:

Modelo de Aprendizaje Profundo El nombre del modelo de aprendizaje profundo del que se toman los vectores. Usé los dos que obtuvieron mejores resultados: *RoBERTuito* o *Beto FLM* (véase Subsección 5.2.3).

El tipo de vectores tomados del modelo Puede tomar dos valores que representan qué vector se obtuvo del modelo. *CLS* cuando se trata de tomar el vector correspondiente al token especial [CLS], que representa a la secuencia de texto completa. O *AVG* cuando es el promedio de los vectores de cada token, incluyendo el [CLS].

Cantidad de capas Ambos modelos cuentan con 12 capas de atención. Este parámetro se refiere a cuántas capas se promedian para generar los vectores. 1 quiere decir tomar únicamente la capa de salida, 4 un promedio de las últimas 4 y 9 un promedio de las últimas 9.

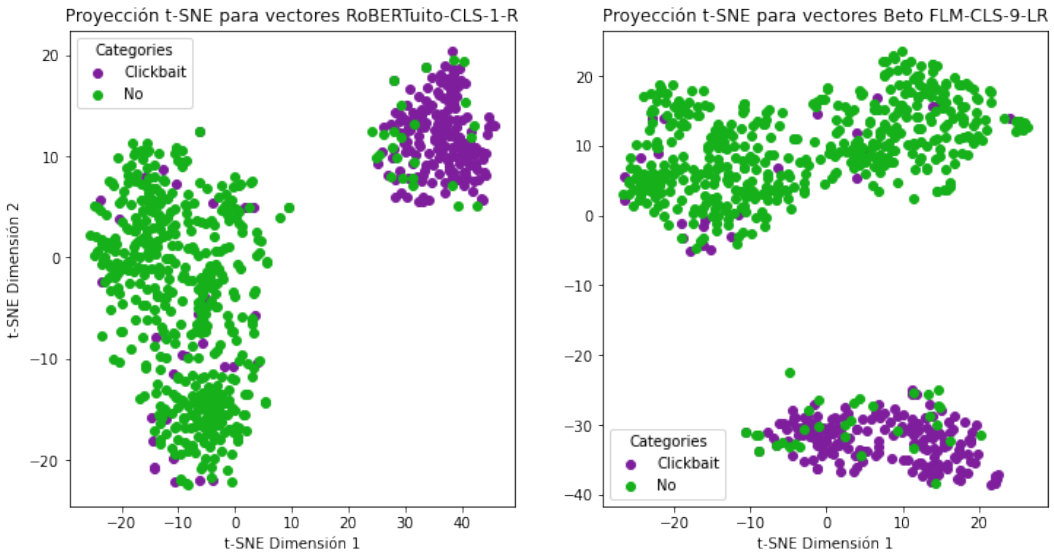

Figura 5.4: Proyección con el algoritmo t-SNE a dos dimensiones para los vectores de las dos combinaciones con mejores resultados. Se puede notar como la mayoría de los textos clickbait o no se logran separar en dos clusters distintos.

Modelo final El modelo que aprende de la concatenación de los vectores obtenidos del modelo de aprendizaje profundo con las características del aprendizaje clásico. Puede ser *xgboost* o *lr* (por Logistic Regression).

En la tabla Tabla 5.4 se reportan los resultados de las pruebas que realicé, todas ellas entrenando con el conjunto de entrenamiento. Los hiperparámetros, incluidos el tipo y cantidad de capas de los vectores a tomar fueron elegidos luego de evaluarlos en el conjunto de validación y se reportan los resultados de ese mejor modelo sobre el conjunto de test.

La Figura 5.4 muestra una proyección a dos dimensiones usando t-SNE (van der Maaten y Hinton, 2008) de los vectores del conjunto de test para los dos modelos con mejores resultados. Estos vectores no incluyen las características del modelado clásico que son luego concatenadas.

5.3. Resultados

A modo de resumen, se reportan los mejores resultados de cada método, con el objetivo de facilitar la comparación, en la Tabla 5.5.

En la Figura 5.5 y Figura 5.6 se ven las curvas de Precision-Recall y de ROC de los modelos seleccionados.

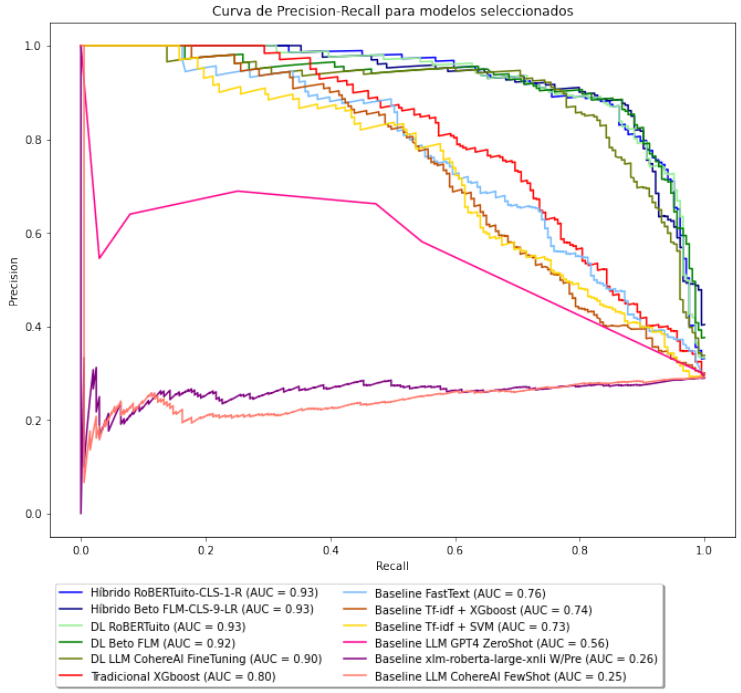

Figura 5.5: Curva de Precision-Recall para una selección de los mejores modelos de cada método explorado.

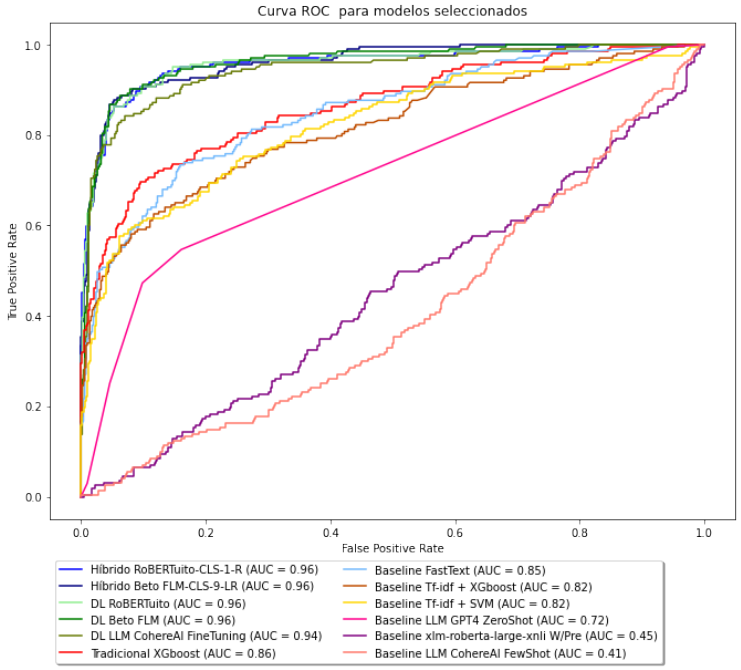

Figura 5.6: Curva de ROC para una selección de los mejores modelos de cada método explorado.

Modelo	Average Precision	Precision	Recall	F1-score	Accuracy
LLM - Cohere AI - Few Shot Classification	0.25	0.43	0.47	0.45	0.66
xlm-roberta-large-xnli	0.26	0.26	0.16	0.20	0.62
LLM - GPT4 - ZeroShot	0.49	0.71	0.25	0.37	0.75
Tf-idf + SVM	0.73	0.81	0.55	0.65	0.83
Tf-idf + XGboost	0.74	0.85	0.48	0.61	0.82
fastText	0.76	0.77	0.55	0.64	0.82
XGboost	0.80	0.78	0.63	0.70	0.84
LLM - Cohere AI - FineTuning	0.89	**0.93**	0.74	0.83	0.91
Beto FLM	0.92	0.85	**0.88**	**0.87**	0.92
RoBERTuito	**0.93**	0.89	0.84	0.86	0.92
RoBERTuito - CLS - 1 - LR	**0.93**	0.89	0.83	0.86	0.92
Beto_FLM - CLS - 9 - LR	**0.93**	0.87	**0.88**	**0.87**	**0.93**

Tabla 5.5: Resultados de una selección de los mejores modelos de cada método explorado.

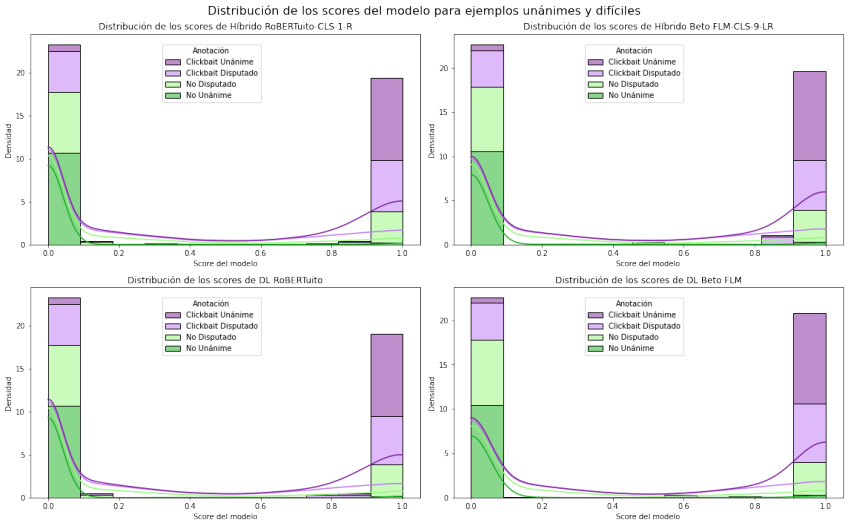

Figura 5.7: Distribución de *score* de los cuatro modelos con mejores resultados según el valor (Clickbait o no) y tipo (unánime o disputado) de etiqueta. Se puede ver como los puntajes suelen ser más extremos cuando las etiquetas son unánimes y cómo la mayor cantidad de errores son en ejemplos con etiquetado disputado.

Finalmente, en la Figura 5.7 se analizan los puntajes (*scores*) de los cuatro modelos con mejores resultados, dos híbridos y dos de aprendizaje profundo, en función de la unanimidad, o no, del etiquetado. Al haber tres etiquetas por tweet se puede asumir que los unánimes deberían ser menos subjetivos que los de votación disputada (2 a 1). Se puede ver que efectivamente los modelos tienden a equivocarse más en los casos disputados y a asignar puntajes más extremos en los ejemplos unánimes.

Capítulo 6

Conclusión

A lo largo de este trabajo mostré cómo no existe un consenso sobre qué es el clickbait, pese a que es un fenómeno que atrae atención de la academia y el público general. Estas contradicciones en las definiciones no sólo hacen que los conjuntos de datos y modelos de detección sean muchas veces incomparables entre sí, sino que dificultan la interpretación y análisis del concepto en general.

Mediante el contraste de las diferentes definiciones junto con el estudio de la curiosidad y la teoría de la brecha de información de Loewenstein (1994), propuse la idea de que esta forma de generar curiosidad mediante la omisión parcial de información es la diferencia fundamental entre el clickbait y otros fenómenos similares como el sensacionalismo. El contenido de noticias blandas, el escándalo, la falta de neutralidad, el amarillismo y otro tipo de tácticas suelen confundirse con el clickbait porque son todas respuestas al mismo objetivo: aumentar el tráfico (o los lectores, si es en papel). Pero no deben considerarse lo mismo, el clickbait tiene sus particularidades y su lenguaje propio. Es un concepto distinto que debe diferenciarse.

Con el objetivo de aportar a reducir la ambigüedad, propuse una definición de clickbait basada en la brecha de información, acompañada de una serie de criterios específicos sobre qué es clickbait y qué no, que logran disminuir la subjetividad. Espero que este aporte pueda utilizarse para unificar el concepto más allá del idioma con el que se trabaje.

Utilizando estos criterios publiqué *TA1C*: el primer conjunto de datos sobre detección de clickbait en español, con 3.500 noticias etiquetadas manualmente por tres anotadores independientes y llegando a un acuerdo inter-anotador muy superior a los reportados anteriormente para el problema.

Además, publiqué *TA1C-unlabeled*: un conjunto de datos sin etiquetar de noticias en español con más de medio millón de textos, que puede utilizarse en diversos problemas de procesamiento de lenguaje natural. Por ejemplo, para hacer fine-tuning de un modelo de lenguaje y adaptarlo a la forma de escribir de varios de los medios más importantes del idioma español.

Finalmente, presenté variados métodos de clasificación de clickbait utilizando los datos de TA1C. Líneas base fuertes que incluyen la utilización de *Large Language Models* y modelos de tres tipos. Por un lado, implementando la mayoría de las características presentes en la literatura para clasificación de clickbait con métodos clásicos, adaptando al español lo que fue necesario y aportando algunas características nuevas. Por otro lado, con modelos de aprendizaje profundo, haciendo *fine-tuning* de *transformers*. Y por último, generando modelos híbridos, que obtienen los mejores resultados de 0.93 de precisión promedio (*average precision)* y 0.87 de F1-score sobre la clase positiva, aplicando una Regresión Logística a una concatenación de las características de los modelos clásicos con un vector denso generado mediante el *finetuning* de un modelo de lenguaje basado en transformers.

6.1. Trabajo futuro

Pese a que los criterios de identificación reducen la subjetividad y la anotación generó un acuerdo inter-anotador Fleiss' κ de 0,825, el clickbait tiene un componente intrínseco de subjetividad. Por esta razón, puede mejorarse el método de anotación del conjunto de datos, ya sea agregando anotadores y/o transformando la anotación en continua o con grados de clickbait. Esto permitiría generar regresiones en lugar de clasificadores, mejoraría la capacidad de elegir entre precisión y recall (que se puede hacer de todas formas actualmente usando el puntaje de salida de los modelos), o aplicar técnicas de perspectivismo como las descritas por Cabitza et al. (2023).

En cuanto a la clasificación, decidí que, por tratarse de un trabajo en el marco del procesamiento de lenguaje natural, las imágenes no fueran consideradas. Sin embargo, ejemplos como el de la Figura 4.1 muestran que las imágenes pueden ser necesarias para clasificar clickbait en algunos casos. También el uso del texto del artículo, tanto para etiquetar como para utilizar en la clasificación, podría aportar más claridad en algunos casos. Por ejemplo, cuando el clickbait se genera por una exageración en la adjetivación, o cuando la omisión

de información es tal que, al leer el artículo, queda claro que no hay noticia alguna como es el caso de

> (88) *¿Qué dijo Marcelo Gallardo al ser consultado sobre la selección uruguaya? Tras el triunfo de River Plate ante Platense, al Muñeco le preguntaron sobre el fuerte rumor que se instaló sobre su arribo a la Celeste*

donde es verdad todo lo que se dice, pero la respuesta de Gallado fue «No voy a hacer ningún comentario». Estos casos se profundizan en la Sección 2.5.

Finalmente, detectar clickbait es útil para poder combatir este fenómeno, pero no es suficiente. Queda claro a lo largo del trabajo que el clickbait no se reduce a noticias *blandas* y temas que podrían no interesar, sino que noticias muy serias e importantes se escriben usando clickbait en los medios más prestigiosos del mundo. Por esta razón, simplemente ocultar el clickbait puede generar un efecto no deseado de perderse una noticia relevante para el lector. La solución está en la respuesta automática de clickbait, que intenta detectar la interrogante, la información omitida que generó curiosidad, y extraerla del texto del cuerpo de la noticia. Algunos trabajos recientes comenzaron con esta idea (Hagen et al. 2022; Maharani et al. 2023; Thirumala y Ferracane, 2023) e incluso hay un desafío para la resolución con un conjunto de datos por Fröbe et al. (2023), aunque al menos en mi conocimiento, todavía no se avanzó en el idioma español.

Referencias bibliográficas

Agrawal, A. (2016). Clickbait detection using deep learning. *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, 268-272. https://doi.org/10.1109/NGCT.2016.7877426

Alrashidi, B., Jamal, A., Khan, I., y Alkhathlan, A. (2022). A review on abusive content automatic detection: approaches, challenges and opportunities. *PeerJ Computer Science*, *8*, e1142. https://doi.org/10.7717/peerj-cs.1142

Al-Sarem, M., Saeed, F., Al-Mekhlafi, Z. G., Mohammed, B. A., Hadwan, M., Al-Hadhrami, T., Alshammari, M. T., Alreshidi, A., y Alshammari, T. S. (2021). An Improved Multiple Features and Machine Learning-Based Approach for Detecting Clickbait News on Social Networks. *Applied Sciences*, *11*(20), 9487. https://doi.org/10.3390/app11209487

Amoedo-Casais, A., Moreno-Moreno, E., Negredo-Bruna, S., Kaufmann-Argueta, J., y Vara-Miguel, A. (2023). *Digital News Report España 2023* (inf. téc.). Universidad de Navarra. Servicio de Publicaciones de la Universidad de Navarra. https://doi.org/10.15581/019.2023

Anand, A., Chakraborty, T., y Park, N. (2017). We Used Neural Networks to Detect Clickbaits: You Won't Believe What Happened Next! En J. M. Jose, C. Hauff, I. S. Altıngovde, D. Song, D. Albakour, S. Watt y J. Tait (Eds.), *Advances in Information Retrieval* (pp. 541-547). Springer International Publishing.

Andrew, B. C. (2007). Media-generated Shortcuts: Do Newspaper Headlines Present Another Roadblock for Low-information Rationality? *Harvard International Journal of Press/Politics*, *12*(2), 24-43. https://doi.org/10.1177/1081180X07299795

Araujo, A. B., Serrano-Puche, J., y Jaso, M. F. N. (2021). Uso del clickbait en los medios nativos digitales españoles. Un análisis de El Confidencial,

El Español, Eldiario.es y Ok Diario. *Dígitos. Revista de Comunicación Digital*, *1*(7), 185. https://doi.org/10.7203/rd.v1i7.184

Aslan, S., Fastrich, G., Donnellan, E., Jones, D. J. W., y Murayama, K. (2021). People's naïve belief about curiosity and interest: A qualitative study. *PLOS ONE*, *16*(9), 1-20. https://doi.org/10.1371/journal.pone.0256632

Baptista, J. P., y Gradim, A. (2020). Understanding Fake News Consumption: A Review. *Social Sciences*, *9*(10). https://doi.org/10.3390/socsci9100185

Baranes, A., Oudeyer, P.-Y., y Gottlieb, J. (2015). Eye movements reveal epistemic curiosity in human observers. *Vision Research*, *117*, 81-90. https://doi.org/https://doi.org/10.1016/j.visres.2015.10.009

Bazaco, Á., Redondo, M., y Sánchez-García, P. (2019). El clickbait, como estrategia del periodismo viral: concepto y metodología. *Revista Latina de Comunicación Social*, (74), 94-115. https://doi.org/10.4185/RLCS-2019-1323

Bergstra, J., Yamins, D., y Cox, D. (2013, junio). Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. En S. Dasgupta y D. McAllester (Eds.), *Proceedings of the 30th International Conference on Machine Learning* (pp. 115-123, Vol. 28). PMLR. https://proceedings.mlr.press/v28/bergstra13.html

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology. General Section*, *45*(3), 180-191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x

Berlyne, D. E. (1966). Curiosity and Exploration: Animals spend much of their time seeking stimuli whose significance raises problems for psychology. *Science*, *153*(3731), 25-33. https://doi.org/10.1126/science.153.3731.25

Berridge, K. C., Robinson, T. E., y Aldridge, J. W. (2009). Dissecting components of reward: 'liking', 'wanting', and learning [Neurosciences]. *Current Opinion in Pharmacology*, *9*(1), 65-73. https://doi.org/https://doi.org/10.1016/j.coph.2008.12.014

Bird, S. E. (2009). Tabloidization: What is it, and does it really matter? [ISBN 9780203878453]. En *The Changing Faces of Journalism*. Routledge.

Biyani, P., Tsioutsiouliklis, K., y Blackmer, J. (2016). "8 Amazing Secrets for Getting More Clicks": Detecting Clickbaits in News Streams Using

Article Informality. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, 94-100.

Blanchard, T. C., Hayden, B. Y., y Bromberg-Martin, E. S. (2015a). Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity. *Neuron*, *85*(3), 602-614. https://doi.org/10.1016/j.neuron.2014.12.050

Blanchard, T. C., Hayden, B. Y., y Bromberg-Martin, E. S. (2015b). Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity. *Neuron*, *85*(3), 602-614. https://doi.org/10.1016/j.neuron.2014.12.050

Blom, J. N., y Hansen, K. R. (2015). Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, *76*, 87-100. https://doi.org/https://doi.org/10.1016/j.pragma.2014.11.010

Bourgonje, P., Moreno Schneider, J., y Rehm, G. (2017, septiembre). From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles. En O. Popescu y C. Strapparava (Eds.), *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism* (pp. 84-89). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-4215

Brogly, C., y Rubin, V. L. (2018). Detecting Clickbait: Here's How to Do It / Comment détecter les pièges à clic. *Canadian Journal of Information and Library Science*, *42*(3), 154-175. muse.jhu.edu/article/743050.

Bromberg-Martin, E. S., y Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron*, *63*(1), 119-126. https://doi.org/10.1016/j.neuron.2009.06.009

Bronakowski, M., Al-khassaweneh, M., y Al Bataineh, A. (2023). Automatic Detection of Clickbait Headlines Using Semantic Analysis and Machine Learning Techniques. *Applied Sciences*, *13*(4). https://doi.org/10.3390/app13042456

Broscoţeanu, D.-M., y Ionescu, R. T. (2023). A Novel Contrastive Learning Method for Clickbait Detection on RoCliCo: A Romanian Clickbait Corpus of News Articles. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

BuzzFeed News. (2014). Exclusive: New York Times Internal Report Painted Dire Digital Picture [https://www.buzzfeednews.com/article/

mylestanzer/exclusive-times-internal-report-painted-dire-digital-picture, Last accessed on 2024-02-15].

Cabitza, F., Campagner, A., y Basile, V. (2023). Toward a Perspectivist Turn in Ground Truthing for Predictive Computing. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(6), 6860-6868. https://doi.org/10.1609/aaai.v37i6.25840

Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., y Pérez, J. (2020). Spanish Pre-Trained BERT Model and Evaluation Data. *PML4DC at ICLR 2020*.

Cao, X., Le, T., Jason y Zhang. (2017). Machine Learning Based Detection of Clickbait Posts in Social Media.

Cavanaugh, L. A. (2014). Because I (Don't) Deserve It: How Relationship Reminders and Deservingness Influence Consumer Indulgence. *Journal of Marketing Research*, *51*(2), 218-232. https://doi.org/10.1509/jmr.12.0133

Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., St. John, R., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., Strope, B., y Kurzweil, R. (2018, noviembre). Universal Sentence Encoder for English. En E. Blanco y W. Lu (Eds.), *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 169-174). Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-2029

Chakraborty, A., Paranjape, B., Kakarla, S., y Ganguly, N. (2016). Stop Clickbait: Detecting and preventing clickbaits in online news media. *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 9-16. https://doi.org/10.1109/ASONAM.2016.7752207

Chakraborty, A., Sarkar, R., Mrigen, A., y Ganguly, N. (2017). Tabloids in the Era of Social Media? Understanding the Production and Consumption of Clickbaits in Twitter. *Proc. ACM Hum.-Comput. Interact.*, *1*(CSCW). https://doi.org/10.1145/3134665

Chawda, S., Patil, A., Singh, A., y Save, A. (2019). A Novel Approach for Clickbait Detection. *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 1318-1321. https://doi.org/10.1109/ICOEI.2019.8862781

Chen, T., y Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. https://doi.org/10.1145/2939672.2939785

Chen, Y., Conroy, N. J., y Rubin, V. L. (2015). Misleading Online Content: Recognizing Clickbait as "False News". *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*, 15-19. https://doi.org/10.1145/2823465.2823467

Chen, Y., y Rubin, V. (2018). Perceptions of Clickbait: A Q - Methodology Approach. *Proceedings of the Annual Conference of CAIS / Actes du congrès annuel de l'ACSI*. https://doi.org/10.29173/cais1046

Chesney, S., Liakata, M., Poesio, M., y Purver, M. (2017). Incongruent Headlines: Yet Another Way to Mislead Your Readers. *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*, 56-61. https://doi.org/10.18653/v1/W17-4210

Chiruzzo, L., Castro, S., Góngora, S., Rosa, A., Meaney, J. A., y Mihalcea, R. (2021). Overview of HAHA at IberLEF 2021: Detecting, Rating and Analyzing Humor in Spanish. *Procesamiento del Lenguaje Natural*, *67*(0), 257-268. http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6394

Coste, C. I., Bufnea, D., y Niculescu, V. (2020). A New Language Independent Strategy for Clickbait Detection. *2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1-6. https://doi.org/10.23919/SoftCOM50211.2020.9238342

D. Molina, M., Sundar, S. S., Rony, M. M. U., Hassan, N., Le, T., y Lee, D. (2021). Does Clickbait Actually Attract More Clicks? Three Clickbait Studies You Must Read. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445753

Devlin, J., Chang, M.-W., Lee, K., y Toutanova, K. (2019, junio). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. En J. Burstein, C. Doran y T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171-4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

Dimpas, P. K., Po, R. V., y Sabellano, M. J. (2017). Filipino and english clickbait detection using a long short term memory recurrent neural network. *2017 International Conference on Asian Language Processing (IALP)*, 276-280. https://doi.org/10.1109/IALP.2017.8300597

Dong, C., Li, Y., Gong, H., Chen, M., Li, J., Shen, Y., y Yang, M. (2022). A Survey of Natural Language Generation. *ACM Comput. Surv.*, *55*(8). https://doi.org/10.1145/3554727

Dong, M., Yao, L., Wang, X., Benatallah, B., y Huang, C. (2019). Similarity-Aware Deep Attentive Model for Clickbait Detection. En Q. Yang, Z.-H. Zhou, Z. Gong, M.-L. Zhang y S.-J. Huang (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 56-69). Springer International Publishing.

Dor, D. (2003). On newspaper headlines as relevance optimizers. *Journal of Pragmatics*, *35*(5), 695-721. https://doi.org/https://doi.org/10.1016/S0378-2166(02)00134-0

Ecker, U. K. H., Lewandowsky, S., Chang, E. P., y Pillai, R. (2014). The effects of subtle misinformation in news headlines. *Journal of Experimental Psychology: Applied*, *20*(4), 323-335. https://doi.org/10.1037/xap0000028

Edson C. Tandoc Jr., Z. W. L., y Ling, R. (2018). Defining "Fake News". *Digital Journalism*, *6*(2), 137-153. https://doi.org/10.1080/21670811.2017.1360143

El-Arini, Khalid and Tang, Joyce. (2014). News Feed FYI: Click-baiting [https://about.fb.com/news/2014/08/news-feed-fyi-click-baiting/, Last accessed on 2024-02-24].

Fakhruzzaman, M. N., Jannah, S. Z., Ningrum, R. A., y Fahmiyah, I. (2021). Clickbait Headline Detection in Indonesian News Sites using Multilingual Bidirectional Encoder Representations from Transformers (M-BERT).

Foundation, 5. (2023). Disrupted Childhood: The cost of persuasive design [https://5rightsfoundation.com/in-action/disrupted-childhood-the-cost-of-persuasive-design-2023.html, Last accessed on 2024-03-16].

Frampton, B. (2015). Clickbait: The changing face of online journalism [https://www.bbc.com/news/uk-wales-34213693, Last accessed on 2024-02-24].

Fröbe, M., Stein, B., Gollub, T., Hagen, M., y Potthast, M. (2023, julio). SemEval-2023 Task 5: Clickbait Spoiling. En A. K. Ojha, A. S. Doğruöz,

G. Da San Martino, H. Tayyar Madabushi, R. Kumar y E. Sartori (Eds.), *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)* (pp. 2275-2286). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.semeval-1.312

García Orosa, B., Gallur Santorun, S., y López García, X. (2017). El uso del clickbait en cibermedios de los 28 países de la Unión Europea. *Revista Latina de Comunicación Social*, (72), 1261-1277. https://doi.org/10.4185/RLCS-2017-1218

García Serrano, J., Romero-Rodríguez, L. M., y Hernando Gómez, Á. (2019). Análisis del çlickbaiting.[en] los titulares de la prensa española contemporánea / Estudio de caso: Diario .[El] País.[en] Facebook. *Estudios sobre el Mensaje Periodístico*, *25*(1), 197-212. https://doi.org/10.5209/ESMP.63724

Geçkil, A., Müngen, A. A., Gündoğan, E., y Kaya, M. (2020). A clickbait detection method on news sites. *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 932-937.

Genç, Ş., y Surer, E. (2023). ClickbaitTR: Dataset for clickbait detection from Turkish news sites and social media with a comparative analysis via machine learning algorithms. *Journal of Information Science*, *49*(2), 480-499. https://doi.org/10.1177/01655515211007746

Ghanem, B., Rosso, P., y Rangel, F. (2020). An Emotional Analysis of False Information in Social Media and News Articles. *ACM Trans. Internet Technol.*, *20*(2). https://doi.org/10.1145/3381750

Glenski, M., Ayton, E., Arendt, D., y Volkova, S. (2017). Fishing for Clickbaits in Social Images and Texts with Linguistically-Infused Neural Network Models.

Golman, R., y Loewenstein, G. (2015). An Information-Gap Framework for Capturing Preferences About Uncertainty. *Conference on Theoretical Aspects of Rationality and Knowledge*, *15*, 141-152. https://www.cmu.edu/dietrich/sds/docs/golman/papers.html

Golman, R., Loewenstein, G., Molnar, A., y Saccardo, S. (2022). The Demand for, and Avoidance of, Information. *Management Science*, *68*(9), 6454-6476. https://doi.org/10.1287/mnsc.2021.4244

Gómez-Mompart, J.-L., Gutiérrez-Lozano, J.-F., y Palau-Sampio, D. (2015). Los periodistas españoles y la pérdida de la calidad de la información:

El juicio profesional. *Comunicar*, *23*(45), 143-150. https://doi.org/10.3916/c45-2015-15

Grigorev, A. (2017). Identifying Clickbait Posts on Social Media with an Ensemble of Linear Models.

Gunnels, K. (2017). The Changing Media Landscape in the United States: A Survey of Journalists at Mainstream National Outlets. https://doi.org/10.13140/RG.2.2.19151.53920

Gutiérrez-Fandiño, A., Armengol-Estapé, J., Pàmies, M., Llop-Palao, J., Silveira-Ocampo, J., Carrino, C. P., Armentano-Oller, C., Rodriguez-Penagos, C., Gonzalez-Agirre, A., y Villegas, M. (2022). MarIA: Spanish Language Models. *Procesamiento del Lenguaje Natural*, *68*(0), 39-60.

Ha, Y., Kim, J., Won, D., Cha, M., y Joo, J. (2018). Characterizing Clickbaits on Instagram. *Proceedings of the International AAAI Conference on Web and Social Media*, *12*(1). https://doi.org/10.1609/icwsm.v12i1.15019

Hagen, M., Fröbe, M., Jurk, A., y Potthast, M. (2022, mayo). Clickbait Spoiling via Question Answering and Passage Retrieval. En S. Muresan, P. Nakov y A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 7025-7036). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.484

He, K., Zhang, X., Ren, S., y Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. https://doi.org/10.1109/CVPR.2016.90

Hurst, N. (2016). *To clickbait or not to clickbait? : an examination of clickbait headline effects on source credibility* [Available at https://mospace.umsystem.edu/xmlui/handle/10355/56080].

Ifantidou, E. (2009). Newspaper headlines and relevance: Ad hoc concepts in ad hoc contexts. *Journal of Pragmatics*, *41*(4), 699-720. https://doi.org/https://doi.org/10.1016/j.pragma.2008.10.016

Indurthi, V., y Oota, S. R. (2017). Clickbait detection using word embeddings. https://doi.org/10.48550/ARXIV.1710.02861

Indurthi, V., Syed, B., Gupta, M., y Varma, V. (2020). Predicting Clickbait Strength in Online Social Media. *Proceedings of the 28th International Conference on Computational Linguistics*, 4835-4846. https://doi.org/10.18653/v1/2020.coling-main.425

Indurthi, V. (2020, abril). *Computational Approaches for Identifying Sensational Soft News* [Tesis de maestría, International Institute of Information Technology, Hyderabad]. https://web2py.iiit.ac.in/research_centres/publications/view_publication/mastersthesis/831

Isikman, E., y Cavanaugh, L. A. (2016). The Effects of Curiosity-Evoking Events on Activity Enjoyment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2771144

Jahan, M. S., y Oussalah, M. (2023). A systematic review of hate speech automatic detection using natural language processing. *Neurocomput.*, *546*(100). https://doi.org/10.1016/j.neucom.2023.126232

Jain, M., Mowar, P., Goel, R., y Vishwakarma, D. K. (2021). Clickbait in Social Media: Detection and Analysis of the Bait. *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, 1-6. https://doi.org/10.1109/CISS50987.2021.9400293

Jepma, M., Verdonschot, R., van Steenbergen, H., Rombouts, S., y Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience*, *6*. https://doi.org/10.3389/fnbeh.2012.00005

John Cassidy. (2015). The Godfather of Clickbait [https://www.newyorker.com/news/john-cassidy/the-godfather-of-clickbait, Last accessed on 2024-03-15].

Jonah Berger. (2014). Why Clickbait Fails [https://jonahberger.com/why-clickbait-fails/, Last accessed on 2024-02-15].

Jordan Litman, T. H., y Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition and Emotion*, *19*(4), 559-582. https://doi.org/10.1080/02699930441000427

Joulin, A., Grave, E., Bojanowski, P., y Mikolov, T. (2017). Bag of Tricks for Efficient Text Classification. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 427-431. https://aclanthology.org/E17-2068

Jung, A.-K., Stieglitz, S., Kissmer, T., Mirbabaie, M., y Kroll, T. (2022). Click me...! The influence of clickbait on user engagement in social media and the role of digital nudging (J. Jankowski, Ed.). *PLOS ONE*, *17*(6), e0266743. https://doi.org/10.1371/journal.pone.0266743

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-y., y Camerer, C. F. (2009). The Wick in the Candle of Learning:

Epistemic Curiosity Activates Reward Circuitry and Enhances Memory [PMID: 19619181]. *Psychological Science*, *20*(8), 963-973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Kaothanthong, N., Kongyoung, S., y Theeramunkong, T. (2021). Headline2Vec: A CNN-based Feature for Thai Clickbait Headlines Classification. *INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)*, *5*(1), 20-31. https://ph02.tci-thaijo.org/index.php/isjet/article/view/240815

Karadzhov, G., Gencheva, P., Nakov, P., y Koychev, I. (2017). We Built a Fake News / Click Bait Filter: What Happened Next Will Blow Your Mind! *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*, 334-343. https://doi.org/10.26615/978-954-452-049-6_045

Kaur, S., Kumar, P., y Kumaraguru, P. (2020). Detecting clickbaits using two-phase hybrid CNN-LSTM biterm model. *Expert Systems with Applications*, *151*, 113350. https://doi.org/https://doi.org/10.1016/j.eswa.2020.113350

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., y Liu, T.-Y. (2017). LightGBM: a highly efficient gradient boosting decision tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 3149-3157.

Khater, S. R., Al-sahlee, O. H., Daoud, D. M., y El-Seoud, M. S. A. (2018). Clickbait Detection. *Proceedings of the 7th International Conference on Software and Information Engineering*, 111-115. https://doi.org/10.1145/3220267.3220287

Kidd, C., y Hayden, B. Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron*, *88*(3), 449-460. https://doi.org/https://doi.org/10.1016/j.neuron.2015.09.010

Kilgo, D. K., y Sinta, V. (2016). Six things you didn't know about headline writing: Sensationalistic form in viral news content from traditional and digitally native news organizations.

Kobayashi, K., y Hsu, M. (2019a). Common neural code for reward and information value. *Proceedings of the National Academy of Sciences*, *116*(26), 13061-13066. https://doi.org/10.1073/pnas.1820145116

Kobayashi, K., y Hsu, M. (2019b). Common neural code for reward and information value. *Proceedings of the National Academy of Sciences*, *116*(26), 13061-13066. https://doi.org/10.1073/pnas.1820145116

Koechley, P. (2012). Why The Title Matters More Than The Talk — by Upworthy [[Online; accessed 26-February-2024]]. https://blog.upworthy.com/why-the-title-matters-more-than-the-talk-867d08b75c3b

Kongyoung, S., Rugchatjaroen, A., y Kaothanthong, N. (2019). Automatic Feature Extraction and Classification Model for Detecting Thai Clickbait Headlines Using Convolutional Neural Network. En T. Endrjukaite, A. Dudko, H. Jaakkola, B. Thalheim, Y. Kiyoki y N. Yoshida (Eds.), *Information Modelling and Knowledge Bases XXX* (pp. 184-194, Vol. 312). IOS Press. https://doi.org/10.3233/978-1-61499-933-1-184

Kruger, J., y Evans, M. (2009). The paradox of Alypius and the pursuit of unwanted information. *Journal of Experimental Social Psychology*, *45*(6), 1173-1179. https://doi.org/https://doi.org/10.1016/j.jesp.2009.06.009

Kuiken, J., Schuth, A., Spitters, M., y Marx, M. (2017a). Effective Headlines of Newspaper Articles in a Digital Environment. *Digital Journalism*, *5*(10), 1300-1314. https://doi.org/10.1080/21670811.2017.1279978

Kuiken, J., Schuth, A., Spitters, M., y Marx, M. (2017b). Effective Headlines of Newspaper Articles in a Digital Environment. *Digital Journalism*, *5*(10), 1300-1314. https://doi.org/10.1080/21670811.2017.1279978

Kumar, V., Khattar, D., Gairola, S., Kumar Lal, Y., y Varma, V. (2018). Identifying Clickbait: A Multi-Strategy Approach Using Neural Networks. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 1225-1228. https://doi.org/10.1145/3209978.3210144

Kümpel, A. S., y Springer, N. (2016). Commenting quality. Effects of user comments on perceptions of journalistic quality. *SCM Studies in Communication and Media*, *5*(3), 353-366.

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., y Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, *4*(5), 531-543. https://doi.org/10.1038/s41562-020-0848-3

Lazauskas, D., Jacka, J., y Kažemėkaitė, I. (2018). Rage, giggles and fishing for clicks: A qualitative study on how clickbaiting affects perceived online

news content quality. https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-75384

Li, X., Zhou, J., Xiang, H., y Cao, J. (2022). Attention Grabbing through Forward Reference: An ERP Study on Clickbait and Top News Stories. *International Journal of Human–Computer Interaction*, *0*(0), 1-16. https://doi.org/10.1080/10447318.2022.2158262

Liao, F., Zhuo, H. H., Huang, X., y Zhang, Y. (2019). Federated Hierarchical Hybrid Networks for Clickbait Detection.

Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, *44*(7), 1585-1595. https://doi.org/https://doi.org/10.1016/j.paid.2008.01.014

Litman, J. A. (2010). Relationships between measures of I- and D-type curiosity, ambiguity tolerance, and need for closure: An initial test of the wanting-liking model of information-seeking. *Personality and Individual Differences*, *48*(4), 397-402. https://doi.org/https://doi.org/10.1016/j.paid.2009.11.005

Lloberes, M., Castellón, I., y Padró, L. (2010). Spanish FreeLing Dependency Grammar. *Proceedings of 7th Language Resources and Evaluation Conference (LREC'10)*.

Lockwood, G. (2016). Academic clickbait: articles with positively-framed titles, interesting phrasing, and no wordplay get more attention online. https://doi.org/10.15200/winn.146723.36330

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75-98. https://doi.org/10.1037/0033-2909.116.1.75

Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272-292. https://doi.org/https://doi.org/10.1006/obhd.1996.0028

López-Sánchez, D., Herrero, J. R., Arrieta, A. G., y Corchado, J. M. (2018). Hybridizing metric learning and case-based reasoning for adaptable clickbait detection. *Applied Intelligence*, *48*(9), 2967-2982. https://doi.org/10.1007/s10489-017-1109-7

Lu, Y., y Pan, J. (2021). Capturing Clicks: How the Chinese Government Uses Clickbait to Compete for Visibility. *Political Communication*, *38*(1-2), 23-54. https://doi.org/10.1080/10584609.2020.1765914

Maharani, N. P. I., Purwarianti, A., y Aji, A. F. (2023). Low-Resource Clickbait Spoiling for Indonesian via Question Answering.

Manjesh, S., Kanakagiri, T., Vaishak, P., Chettiar, V., y Shobha, G. (2017). Clickbait Pattern Detection and Classification of News Headlines Using Natural Language Processing. *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)*, 1-5. https://doi.org/10.1109/CSITSS.2017.8447715

Mark, Gloria. (2014). Click Bait Is a Distracting Affront to Our Focus [https://www.nytimes.com/roomfordebate/2014/11/24/you-wont-believe-what-these-people-say-about-click-bait/click-bait-is-a-distracting-affront-to-our-focus, Last accessed on 2024-02-24].

Marreddy, M., Oota, S. R., Vakada, L. S., Chinni, V. C., y Mamidi, R. (2021). Clickbait Detection in Telugu: Overcoming NLP Challenges in Resource-Poor Languages using Benchmarked Techniques. *2021 International Joint Conference on Neural Networks (IJCNN)*, 1-8. https://doi.org/10.1109/IJCNN52387.2021.9534382

Mathur, S. (2019). clickbait-detector [Last accessed on 2023-09-20]. https://doi.org/10.17605/OSF.IO/T3UJ9

Menon, S., y Soman, D. (2002). Managing the Power of Curiosity for Effective Web Advertising Strategies. *Journal of Advertising*, *31*(3), 1-14. http://www.jstor.org/stable/4189222

Merriam-Webster dictionary. (2024). Clickbait. In Merriam-Webster.com dictionary [https://www.merriam-webster.com/dictionary/clickbait, Last accessed on 2024-02-26].

Mihelich, J. F. (2019). *Curiosity, Clickbait and Their Impact on Political Media Selection* [Available at http://hdl.handle.net/10339/93913].

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., y Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. En C. Burges, L. Bottou, M. Welling, Z. Ghahramani y K. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 26). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

Molek-Kozakowska, K. (2013). Towards a pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication*, *7*(2), 173-197. https://doi.org/10.1177/1750481312471668

Molina, M. D., Sundar, S. S., Le, T., y Lee, D. (2021). "Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *American Behavioral Scientist*, *65*(2), 180-212. https://doi.org/10.1177/0002764219878224

Molyneux, L., y Coddington, M. (2020). Aggregation, Clickbait and Their Effect on Perceptions of Journalistic Credibility and Quality. *Journalism Practice*, *14*(4), 429-446. https://doi.org/10.1080/17512786.2019.1628658

Mormol, P. (2019). 'I Urge You To See This...'. Clickbait as One of the Dominant Features of Contemporary Online Headlines. *Social Communication*, *5*(2), 1-10. https://doi.org/10.2478/sc-2019-0004

Mourão, R. R., y Robertson, C. T. (2019). Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information. *Journalism Studies*, *20*(14), 2077-2095. https://doi.org/10.1080/1461670X.2019.1566871

Mukherjee, P., Dutta, S., y Bruyn, A. D. (2022). Did clickbait crack the code on virality? *Journal of the Academy of Marketing Science*, *50*(3), 482-502. https://doi.org/10.1007/s11747-021-00830-x

Mukherjee, P., Dutta, S., y Mani, D. (2019). The Antecedents and Rebroadcast Consequences of Clickbait. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3421874

Munger, K., Luca, M., Nagler, J., y Tucker, J. (2020). The (Null) Effects of Clickbait Headlines on Polarization, Trust, and Learning. *Public Opinion Quarterly*, *84*(1), 49-73. https://doi.org/10.1093/poq/nfaa008

Naeem, B., Khan, A., Beg, M. O., y Mujtaba, H. (2020). A deep learning framework for clickbait detection on social area network using natural language cues. *Journal of Computational Social Science*, *3*(1), 231-243. https://doi.org/10.1007/s42001-020-00063-y

Nate Silver. (2011). The Economics of Blogging and The Huffington Post [https://archive.nytimes.com/fivethirtyeight.blogs.nytimes.com/2011/02/12/the-economics-of-blogging-and-the-huffington-post/, Last accessed on 2024-03-15].

Nic, N., Fletcher, R., Kirsten, E., Robertson, C. T., y Nielsen, R. K. (2023). Reuters institute digital news report 2023. *Reuters Institute for the Study of Journalism*. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023

Nicholas Carlson. (2014). Upworthy Traffic Gets Crushed [https://www.businessinsider.com/facebook-changed-how-the-news-feed-works--and-huge-website-upworthy-suddenly-shrank-in-half-2014-2, Last accessed on 2024-03-15].

Oliva, C., Palacio-Marín, I., Lago-Fernández, L. F., y Arroyo, D. (2022). Rumor and Clickbait Detection by Combining Information Divergence Measures and Deep Learning Techniques. *Proceedings of the 17th International Conference on Availability, Reliability and Security*. https://doi.org/10.1145/3538969.3543791

Omidvar, A., Jiang, H., y An, A. (2019). Using Neural Network for Identifying Clickbaits in Online News Media. En J. A. Lossio-Ventura, D. Muñante y H. Alatrista-Salas (Eds.), *Information Management and Big Data* (pp. 220-232). Springer International Publishing.

OpenAI. (2024). How do I use the OpenAI API with text in different languages? [https://help.openai.com/en/articles/6742369-how-do-i-use-the-openai-api-in-different-languages, Last accessed on 2024-03-16].

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., ... Zoph, B. (2024). GPT-4 Technical Report.

Oudeyer, P.-Y., Gottlieb, J., y Lopes, M. (2016). Chapter 11 - Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. En B. Studer y S. Knecht (Eds.), *Motivation* (pp. 257-284, Vol. 229). Elsevier. https://doi.org/https://doi.org/10.1016/bs.pbr.2016.05.005

Oxford English Dictionary. (2024). Clickbait definition [https://www.oed.com/search/dictionary/?scope=Entries&q=clickbait, Last accessed on 2024-02-26].

Paasonen, S. (2016). Fickle focus: Distraction, affect and the production of value in social media. *First Monday*. https://doi.org/10.5210/fm.v21i10.6949

Padró, L., y Stanilovsky, E. (2012). FreeLing 3.0: Towards Wider Multilinguality. *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)*.

Palau-Sampio, D. (2016). Reference press metamorphosis in the digital context: clickbait and tabloid strategies in Elpais.com. *Comunicacion y Sociedad*, *29*, 63-79. https://doi.org/10.15581/003.29.35924

Pandey, S., y Kaur, G. (2018). Curious to Click It?-Identifying Clickbait using Deep Learning and Evolutionary Algorithm. *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1481-1487. https://doi.org/10.1109/ICACCI.2018.8554873

Papadopoulou, O., Zampoglou, M., Papadopoulos, S., y Kompatsiaris, I. (2017). A Two-Level Classification Approach for Detecting Clickbait Posts using Text-Based Features. *arXiv preprint arXiv:1710.08528*.

Park, K., Kim, T., Yoon, S., Cha, M., y Jung, K. (2020). BaitWatcher: A Lightweight Web Interface for the Detection of Incongruent News Headlines. En K. Shu, S. Wang, D. Lee y H. Liu (Eds.), *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities* (pp. 229-252). Springer International Publishing. https://doi.org/10.1007/978-3-030-42699-6_12

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., y Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825-2830.

Pengnate, S. ((2016). Measuring Emotional Arousal in Clickbait: Eye-Tracking Approach. *AMCIS 2016 Proceedings*. https://aisel.aisnet.org/amcis2016/HCI/Presentations/3/

Pengnate, S. (, Young, A., y Chen, J. (2018). NEW DETAILS EMERGE! Revealing the Effects of Clickbait Headlines on User Responses. *AMCIS 2018 Proceedings*. https://aisel.aisnet.org/amcis2018/eBusiness/Presentations/6

Pennington, J., Socher, R., y Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *Empirical Methods in Natural Language Processing (EMNLP)*, 1532-1543. http://www.aclweb.org/anthology/D14-1162

Pérez, J. M., Furman, D. A., Alonso Alemany, L., y Luque, F. M. (2022). RoBERTuito: a pre-trained language model for social media text in Spanish. *Proceedings of the Language Resources and Evaluation Conference*, 7235-7243.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., y Zettlemoyer, L. (2018, junio). Deep Contextualized Word Representations. En M. Walker, H. Ji y A. Stent (Eds.), *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)* (pp. 2227-2237). Association for Computational Linguistics. https://doi.org/10.18653/v1/N18-1202

Pew Research Center. (2022). Social Media and News Fact Sheet [https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/, Last accessed on 2023-09-04].

Peysakhovich, Alex and Hendrix, Kristin. (2016). Further Reducing Clickbait in Feed [https://about.fb.com/news/2016/08/news-feed-fyi-further-reducing-clickbait-in-feed/, Last accessed on 2024-02-24].

Portia Roelofs and Max Gallien. (2017). Clickbait and impact: how academia has been hacked [https://blogs.lse.ac.uk/impactofsocialsciences/2017/09/19/clickbait-and-impact-how-academia-has-been-hacked/, Last accessed on 2024-03-15].

Posadas-Durán, J.-P., Gómez-Adorno, H., Sidorov, G., y Escobar, J. J. M. (2019). Detection of fake news in a new corpus for the Spanish language. *Journal of Intelligent & Fuzzy Systems*, *36*(5), 4869-4876.

Potthast, M., Gollub, T., Hagen, M., y Stein, B. (2018). The Clickbait Challenge 2017: Towards a Regression Model for Clickbait Strength. *CoRR*, *abs/1812.10847*. https://arxiv.org/abs/1812.10847

Potthast, M., Gollub, T., Komlossy, K., Schuster, S., Wiegmann, M., Garces Fernandez, E., Hagen, M., y Stein, B. (2018, agosto). Crowdsourcing a Large Corpus of Clickbait on Twitter. En E. Bender, L. Derczynski y P. Isabelle (Eds.), *27th International Conference on Computational Linguistics (COLING 2018)* (pp. 1498-1507). The COLING 2018 Organizing Committee. https://www.aclweb.org/anthology/C18-1127

Potthast, M., Köpsel, S., Stein, B., y Hagen, M. (2016, marzo). Clickbait Detection. En N. Ferro, F. Crestani, M.-F. Moens, J. Mothe, F. Silvestri, G. Di Nunzio, C. Hauff y G. Silvello (Eds.), *Advances in Information Retrieval. 38th European Conference on IR Research (ECIR 2016)* (pp. 810-817, Vol. 9626). Springer. https://doi.org/10.1007/978-3-319-30671-1_72

Pujahari, A., y Sisodia, D. S. (2021). Clickbait detection using multiple categorisation techniques. *Journal of Information Science*, *47*(1), 118-128. https://doi.org/10.1177/0165551519871822

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., y Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. https://github.com/openai/gpt-2

Rajapaksha, P., Farahbakhsh, R., y Crespi, N. (2021). BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits. *IEEE Access*, *9*, 154704-154716. https://doi.org/10.1109/ACCESS.2021.3128742

Rajput, A. E. (2019). Clickbait in Education—Positive or Negative? Machine Learning Answers. En A. Visvizi y M. D. Lytras (Eds.), *Research & Innovation Forum 2019* (pp. 101-108). Springer International Publishing.

Redondo García, M. M. (2011). *El sensacionalismo y su penetración en el sistema mediático español* [Tesis doctoral, Universidad de Valladolid, Facultad de Filosofía y Letras]. https://doi.org/10.35376/10324/891

Roca, M. (2008). Rethinking entertainment. Entrevista a Marty Kaplan. *Tripodos. Blanquerna School of Communication and International Relations-URL*, (23), 135-143. https://raco.cat/index.php/Tripodos/article/view/118924

Rodríguez, D. (2013). *Memecracia Los virales que nos gobiernan*. Gestión 2000.

Rony, M. M. U., Hassan, N., y Yousuf, M. (2017). Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects? *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 232-239. https://doi.org/10.1145/3110025.3110054

Sadri, S. R. (2019). Listicles and the modern news article: comparing the perceived credibility of listicles and traditional articles among millennial media consumers. *Atlantic Journal of Communication*, *27*(2), 83-98. https://doi.org/10.1080/15456870.2019.1574794

Scacco, Joshua and Muddiman, Ashley. (2016). Investigating the influence of "clickbait" news headlines [https://mediaengagement.org/research/clickbait-headlines, Last accessed on 2024-02-24].

Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data [PMID: 26109341]. *Physiological Reviews*, *95*(3), 853-951. https://doi.org/10.1152/physrev.00023.2014

Scikit Learn Documentation. (2024). Average Precision Score [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.average_precision_score.html, Last accessed on 2024-10-08].

Scott, K. (2021). You won't believe what's in this paper! Clickbait, relevance and the curiosity gap. *Journal of Pragmatics*, *175*, 53-66. https://doi.org/https://doi.org/10.1016/j.pragma.2020.12.023

Setlur, A. R. (2018). Semi-Supervised Confidence Network aided Gated Attention based Recurrent Neural Network for Clickbait Detection.

Sevilla, J., y Meyer, R. J. (2020). Leaving Something for the Imagination: The Effect of Visual Concealment on Preferences. *Journal of Marketing*, *84*(4), 109-126. https://doi.org/10.1177/0022242919899393

Shang, L., Zhang, D. (, Wang, M., Lai, S., y Wang, D. (2019). Towards reliable online clickbait video detection: A content-agnostic approach. *Knowledge-Based Systems*, *182*, 104851. https://doi.org/https://doi.org/10.1016/j.knosys.2019.07.022

Shevon Desai and Jo Angela Oehrli. (2022). "Fake News,"Lies and Propaganda: How to Sort Fact from Fiction [https://guides.lib.umich.edu/fakenews, Last accessed on 2024-03-15].

Shin, D. D., y Kim, S.-i. (2019). Homo Curious: Curious or Interested? *Educational Psychology Review*, *31*(4), 853-874. https://doi.org/10.1007/s10648-019-09497-x

Silverman, C. (2015). Lies, Damn Lies and Viral Content. https://doi.org/10.7916/D8Q81RHH

Singh Sisodia, D. (2019). Ensemble Learning Approach for Clickbait Detection Using Article Headline Features. *Informing Science: The International Journal of an Emerging Transdiscipline*, *22*, 031-044. https://doi.org/10.28945/4279

Sofía Sánchez González. (2021). Detect fake news with Narrativa's artificial intelligence [https://www.narrativa.com/detect-fake-news-with-narrativas-artificial-intelligence/, Last accessed on 2024-03-14].

Suzuki, S., Hamamura, T., y Takemura, K. (2018). Emotional fortification: Indulgent consumption and emotion reappraisal and their implications for well-being. *Journal of Consumer Behaviour*, *18*(1), 25-31. https://doi.org/10.1002/cb.1746

Tang, X., Renninger, K. A., Hidi, S. E., Murayama, K., Lavonen, J., y Salmela-Aro, K. (2022). The differences and similarities between curiosity and

interest: Meta-analysis and network analyses. *Learning and Instruction*, *80*, 101628. https://doi.org/https://doi.org/10.1016/j.learninstruc.2022.101628

Thirumala, A., y Ferracane, E. (2023). Clickbait Classification and Spoiling Using Natural Language Processing.

Thomas, P. (2017). Clickbait Identification using Neural Networks.

University of California - Berkeley Haas School of Business. (2019, junio). How information is like snacks, money, and drugs – to your brain: Researchers demonstrate common neural code for information and money; both act on the brain's dopamine-producing reward system [www.sciencedaily.com/releases/2019/06/190619174530.htm, Last accessed on 2023-08-27].

Urizar, X. S., y Roncal, I. S. V. (2013). Elhuyar at TASS 2013. *XXIX Congreso de la Sociedad Española de Procesamiento de lenguaje natural*. http://www.sepln.org/workshops/tass/2013/tasks.php

van der Maaten, L., y Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(86), 2579-2605. http://jmlr.org/papers/v9/vandermaaten08a.html

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., y Polosukhin, I. (2017). Attention is All you Need. En I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan y R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Venneti, L., y Alam, A. (2018). How Curiosity can be modeled for a Clickbait Detector.

Vijgen, B. (2014). The listicle: An exploring research on an interesting shareable new media phenomenon. *Ephemerides*, *59*(1), 103-122. https://www.ceeol.com/search/article-detail?id=204859

Vincze, V., y Szabó, M. K. (2020, diciembre). Automatic Detection of Hungarian Clickbait and Entertaining Fake News. En A. Aker y A. Zubiaga (Eds.), *Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media (RDSM)* (pp. 58-69). Association for Computational Linguistics. https://aclanthology.org/2020.rdsm-1.6

Volkova, S., Shaffer, K., Jang, J. Y., y Hodas, N. (2017). Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 647-653. https://doi.org/10.18653/v1/P17-2102

Wagner-Egger, P., y Gygax, P. (2018). Diana Was Not Involved in the 9/11 Terrorist Attacks!: Or Was She? Newspaper Headlines and the Boomerang Effect. *Swiss Journal of Psychology*, *77*(1), 15-22. https://doi.org/10.1024/1421-0185/a000203

Wegner, D. M. (1984). Innuendo and damage to reputations. En T. C. Kinnear (Ed.), *Advances in Consumer Research, 11* (pp. 691-696). Association for Consumer Research.

Wegner, D. M., Wenzlaff, R., Kerker, R. M., y Beattie, A. E. (1981). Incrimination through innuendo: Can media questions become public answers? *Journal of Personality and Social Psychology*, *40*, 822-832. https://scholar.harvard.edu/dwegner/publications/incrimination-through-innuendo-can-media-questions-become-public-answers

Wei, F., y Nguyen, U. T. (2022). An Attention-Based Neural Network Using Human Semantic Knowledge and Its Application to Clickbait Detection. *IEEE Open Journal of the Computer Society*, *3*, 217-232. https://doi.org/10.1109/OJCS.2022.3213791

Wiegmann, M., Völske, M., Stein, B., Hagen, M., y Potthast, M. (2018). Heuristic Feature Selection for Clickbait Detection.

Wiggin, K. L., Reimann, M., y Jain, S. P. (2018). Curiosity Tempts Indulgence. *Journal of Consumer Research*, *45*(6), 1194-1212. https://doi.org/10.1093/jcr/ucy055

Wikipedia. (2024a). Clickbait — Wikipedia, The Free Encyclopedia [[Online; accessed 25-February-2024]]. https://en.wikipedia.org/w/index.php?title=Clickbait&oldid=1208816291

Wikipedia. (2024b). Fake news — Wikipedia, The Free Encyclopedia [http://en.wikipedia.org/w/index.php?title=Fake%20news&oldid=1206781009, Last accessed on 2024-02-24].

Wikipedia. (2024c). Junk food news — Wikipedia, The Free Encyclopedia [[Online; accessed 26-February-2024]]. https://en.wikipedia.org/w/index.php?title=Junk_food_news&oldid=1208815595

William, A., y Sari, Y. (2020). CLICK-ID: A novel dataset for Indonesian clickbait headlines. *Data in Brief*, *32*, 106231. https://doi.org/https://doi.org/10.1016/j.dib.2020.106231

Wongsap, N., Prapphan, T., Lou, L., Kongyoung, S., Jumun, S., y Kaothanthong, N. (2018). Thai Clickbait Headline News Classification and its Characteristic. *2018 International Conference on Embedded Systems and Intelligent Technology & International Conference on Information and Communication Technology for Embedded Systems (ICESIT-ICICTES)*, 1-6. https://doi.org/10.1109/ICESIT-ICICTES.2018.8442064

Wu, C., Wu, F., Qi, T., y Huang, Y. (2020). Clickbait Detection with Style-Aware Title Modeling and Co-attention. En M. Sun, S. Li, Y. Zhang, Y. Liu, S. He y G. Rao (Eds.), *Chinese Computational Linguistics* (pp. 430-443). Springer International Publishing.

Wu, T. (2016). The Raise of Clickbait [ISBN 9780804170048]. En *The Attention Merchants* (pp. 276-289). Knopf.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., y Le, Q. V. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. En H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox y R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 32). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2019/file/dc6a7e655d7e5840e66733e9ee67cc69-Paper.pdf

Zannettou, S., Chatzis, S., Papadamou, K., y Sirivianos, M. (2018). The Good, the Bad and the Bait: Detecting and Characterizing Clickbait on YouTube. *2018 IEEE Security and Privacy Workshops (SPW)*, 63-69. https://doi.org/10.1109/SPW.2018.00018

Zannettou, S., Sirivianos, M., Blackburn, J., y Kourtellis, N. (2019). The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *J. Data and Information Quality*, *11*(3). https://doi.org/10.1145/3309699

Zedelius, C. M., Gross, M. E., y Schooler, J. W. (2022). Inquisitive but not discerning: Deprivation curiosity is associated with excessive openness to inaccurate information. *Journal of Research in Personality*, *98*, 104227. https://doi.org/https://doi.org/10.1016/j.jrp.2022.104227

Zeng, Q., y Li, A.-R. (2022, octubre). A Survey in Automatic Irony Processing: Linguistic, Cognitive, and Multi-X Perspectives. En N. Calzolari, C.-R. Huang, H. Kim, J. Pustejovsky, L. Wanner, K.-S. Choi, P.-M. Ryu, H.-H. Chen, L. Donatelli, H. Ji, S. Kurohashi, P. Paggio, N. Xue, S. Kim, Y. Hahm, Z. He, T. K. Lee, E. Santus, F. Bond y S.-H. Na (Eds.), *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 824-836). International Committee on Computational Linguistics. https://aclanthology.org/2022.coling-1.69

Zhang, W., Du, W., Bian, Y., Peng, C.-H., y Jiang, Q. (2020). Seeing is not always believing: an exploratory study of clickbait in WeChat. *Internet Research*, *30*(3), 1043-1058. https://doi.org/10.1108/intr-09-2019-0373

Zheng, H.-T., Chen, J.-Y., Yao, X., Sangaiah, A. K., Jiang, Y., y Zhao, C.-Z. (2018). Clickbait Convolutional Neural Network. *Symmetry*, *10*(5). https://doi.org/10.3390/sym10050138

Zheng, H.-T., Yao, X., Jiang, Y., Xia, S.-T., y Xiao, X. (2017). Boost Clickbait Detection Based on User Behavior Analysis. En L. Chen, C. S. Jensen, C. Shahabi, X. Yang y X. Lian (Eds.), *Web and Big Data* (pp. 73-80). Springer International Publishing.

Zheng, J., Yu, K., y Wu, X. (2021). A deep model based on Lure and Similarity for Adaptive Clickbait Detection. *Knowledge-Based Systems*, *214*, 106714. https://doi.org/https://doi.org/10.1016/j.knosys.2020.106714

Zhou, M., Xu, W., Zhang, W., y Jiang, Q. (2022). Leverage knowledge graph and GCN for fine-grained-level clickbait detection. *World Wide Web*, *25*(3), 1243-1258. https://doi.org/10.1007/s11280-022-01032-3

Zhou, X., Jain, A., Phoha, V. V., y Zafarani, R. (2020). Fake News Early Detection: A Theory-driven Model. *Digital Threats: Research and Practice*, *1*(2), 1-25. https://doi.org/10.1145/3377478

Zhou, Y. (2017). Clickbait Detection in Tweets Using Self-attentive Network.

Zhuang, L., Wayne, L., Ya, S., y Jun, Z. (2021, agosto). A Robustly Optimized BERT Pre-training Approach with Post-training. En S. Li, M. Sun, Y. Liu, H. Wu, K. Liu, W. Che, S. He y G. Rao (Eds.), *Proceedings of the 20th Chinese National Conference on Computational Linguistics* (pp. 1218-1227). Chinese Information Processing Society of China. https://aclanthology.org/2021.ccl-1.108

APÉNDICES

Apéndice 1

Historia del clickbait

Para profundizar en el entendimiento del clickbait, es importante conocerlo desde una perspectiva histórica, para entender el por qué de su creación y su evolución a lo largo del tiempo. Este apéndice describe brevemente la historia del clickbait desde sus antecedentes y comienzos hasta su prevalencia en la actualidad.

1.1. Antecedentes

Si bien el clickbait es un fenómeno de la era digital y de la difusión de noticias por Internet, el atraer la atención a una noticia con la utilización de titulares llamativos que generen curiosidad, precede largamente a la invención de Internet.

Los medios de comunicación disputan la atención del público que debe seleccionar qué noticias leer principalmente a través de los titulares. Ya sea en un diario de papel, un sitio de noticias, un kiosco o en redes sociales, los consumidores leen el titular, posiblemente acompañado de un subtítulo, imagen y/o video y deciden si leer o no esa noticia. Los medios, en cualquiera de las plataformas, se benefician de ser elegidos por el lector por lo que es rentable que los titulares sean llamativos, intentando hacer que su información «parezca relevante, urgente o inusual» (Molek-Kozakowska, 2013).

El modelo de negocio de la prensa se ha transformado a lo largo de los años, pero se mantiene invariable que la mayoría de los medios son empresas que buscan generar ingresos y que para eso dependen de atraer la atención a sus noticias. Por esa razón, surge la estrategia de revertir roles: en lugar

de pensar en las noticias más relevantes y luego en cómo poder distribuirlas y convencer al público de que las lea, se generan las noticias en base a su capacidad de atraer. A este fenómeno se lo conoce como sensacionalismo.

Como explica Redondo García (2011) «el gusto por la información escandalosa, el cotilleo, el morbo, lo extraordinario o anómalo ha existido siempre» y se puede rastrear hasta en momentos históricos como el teatro en la antigua Grecia, la *literatura efímera* del siglo XVII o los juglares en la edad media.

Sin embargo, el sensacionalismo como fenómeno en sí mismo surge en la década de 1830. Hasta ese momento, los periódicos apuntaban a las élites sociales y económicas y basaban su modelo de negocio en las suscripciones y en los apoyos de los partidos políticos que formaban su línea editorial. En la década del 30 en Nueva York surge la *penny press*, prensa que toma su nombre del precio de un penique a la que se vendía, principalmente encarnada por el New York Sun (1833) y el New York Morning Herald (1835). Estos periódicos comenzaron a dar lugar a noticias locales, policiales, deportivas, y no únicamente las políticas y económicas que consumían las élites. Además, se escribían de forma más directa y en un lenguaje más simple para que pudiera entender la mayoría de la población. Su modelo de negocio apuntaba a la venta de publicidad y a la masividad. Como consecuencia, describe Redondo García (2011), «esta forma de distribución afectó al diseño y a los contenidos de los diarios, exigiendo que fuesen más impactantes y atractivos al primer golpe de vista.» Los titulares comenzaron a tomar una trascendencia mucho mayor: antes los suscriptores ya habían comprado el diario cuando los leían, ahora generalmente los ojeaban en la tapa o los escuchaban gracias a los gritos de los canillitas y en base a ellos tomaban la decisión de comprar o no.

Además, estas publicaciones encabezadas por el Morning Herald añadieron como temas habituales a su cobertura la crónica social. En particular enfocados en la alta sociedad, se contaban *asuntos íntimos* que empezaron a caracterizar al periodismo moderno y sensacionalista que no diferenciaba los ámbitos público y privado de la vida de las personas sobre las que informaba generando un gran interés en el público (Redondo García, 2011).

Este tipo de sensacionalismo se agudizó en la década de 1890 con Joseph Pulitzer y W. R. Hearst y el nacimiento de la prensa amarilla. Empezaron incluso a exagerar o directamente inventar noticias. También comenzaron a utilizar en los titulares «expresiones llanas y coloquiales», grandilocuencia, definiciones categóricas y «la utilización del cebo que suponía la belleza femenina

o la apelación al misterio imaginado.» (Redondo García, 2011)

1.2. Origen

Se puede ver claramente en el sensacionalismo, la prensa amarilla y posteriormente los tabloides la misma estrategia que llevaría años después al clickbait. Como explica John Cassidy (2015) recordando a Vincent *Vinnie* Musetto; editor de otro de los diarios sensacionalistas más importantes, el New York Post, y creador del famoso titular que se ve en la Figura 1.1; su trabajo era crear titulares que cumplieran un único objetivo: llamar la atención.

Figura 1.1: Uno de los titulares más famosos del *New York Post*, uno de los periódicos sensacionalistas más importantes del mundo. El autor del titular es Vincent Musetto.

Ese mismo objetivo seguía Jonah Peretti, pero a principios del siglo XXI y en la web, un medio aún incipiente en ese momento. Peretti, como cuenta T. Wu (2016), buscaba hacer páginas web «contagiosas» cuyo único objetivo era obtener tantas visitas como fuera posible. De ese modo, generando lo que ahora se conoce como contenido viral, se convirtió en experto en llegar a un público cada vez más grande. En el año 2005 recibió la oferta de Kenneth Lerer, ejecutivo en comunicaciones de AOL, de crear junto con Arianna Huffington; una escritora, comentarista política y ex candidata a gobernadora de California; y Andrew Breitbart, un periodista que trabajaba en el agregador de noticias Drudge Report, un sitio que atrajera la atención del público masivo a contenido con una visión de izquierda estadounidense para beneficiar al partido Demócrata: The Huffington Post.

El *HuffPo* se convertiría en poco tiempo en uno de los sitios de noticias más importantes del mundo (Nate Silver, 2011). Rápidamente sumó a las noticias políticas la información *blanda* y de farándula. Su popularidad se basó en generar contenido hecho por celebridades (*blogs*, columnas, videos), conglomerar noticias de todo tipo que llamaran la atención, tener una fuerte presencia en

redes sociales y «titulares sensacionalistas, emparejados con imágenes provocativas, como una celebridad vestida de bikini ... que parecían tomar el control de la mente causando a la mano a clickear casi involuntariamente».

Figura 1.2: Ejemplos de listículos, artículos lista, una de las especialidades del sitio Buzzfeed.

Otros sitios de noticias llevarían estas tácticas aún más lejos. Peretti dejó Huffington Post para dedicarse de lleno a su nueva creación: Buzzfeed, quizás el sitio más relacionado con el clickbait de la historia de Internet y *el rey de los listículos*[1] (Kilgo y Sinta, 2016), título que se ganó por titulares como los de la Figura 1.2. Por su parte, Upworthy nació en 2012 con una idea similar a la del *HuffPo*, pero enfocado casi exclusivamente en las redes sociales y en particular en Facebook. Popularizaron los titulares largos (muchas veces de dos líneas), que le hablaban en primera persona al lector y principalmente terminando con frases del estilo de «no creerás lo que pasó después». Upworthy llegó a tener más de 80 millones de visitantes únicos en noviembre de 2013 (Nicholas Carlson, 2014).

El éxito de esta nueva manera de atraer tráfico fue tan claro que la competencia se vio forzada a imitar algunas de sus formas. Primero lo hicieron algunos sitios de noticias nativos de Internet como Slate y Salon, pero el fenómeno llegó también a las versiones digitales de varios de los periódicos de papel más importantes. Y no solamente en Estados Unidos ni en el mundo anglosajón, la práctica del clickbait se extendió a lo largo del mundo y particularmente también en el idioma español, tanto es así que en el año 2012 El País de Madrid se asoció con el Huffington Post[2].

Si bien la finalidad e incluso el estilo de los titulares sensacionalistas y la prensa tabloide son muy similares al clickbait actual, la web ha probado ser el ambiente ideal para que estos métodos proliferen y se profundicen. Los medios en línea tienen la posibilidad y la capacidad de medir con mucha facilidad,

[1] Palabra que combina lista con artículo, traducción de la original en inglés *listicle*
[2] 'El Huffington Post', un fenómeno global que desembarca en España

precisión, y en tiempo real, la cantidad de lectores, las interacciones (como *me gustas*, participación mediante comentarios, compartir, etc) y en definitiva cualquier medida de interés.

El modelo de negocio en línea depende casi exclusivamente de la publicidad, excluyendo una tendencia hacia la suscripción, y, por lo tanto, de la cantidad de visitas que reciben las noticias (Kilgo y Sinta, 2016). Además, muchos lectores consumen noticias desde redes sociales, foros o agregadores, donde los artículos de varios medios se presentan juntos, generando una competencia entre ellos por la elección del usuario y, en definitiva, por los ingresos.

En la era digital, donde abunda la información y se hace imposible consumir todo lo que nos llega, «la información no es un recurso que pasivamente espera a que la audiencia lo encuentre y acceda a él, sino que está siempre en competencia por la atención de la gente con otras informaciones» (Bazaco et al. 2019). Como explica Marty Kaplan lo que falta en esta era es el tiempo para procesar tanta información «y todos están intentando captar nuestra atención porque quieren monetizar esa atención una vez que la tienen» (Roca, 2008).

Dada esta nueva realidad, cuando comenzó la era de la información, a los medios tradicionales se les planteó el dilema de cuánto transformarse para imitar a los nativos digitales. En particular, en el comienzo estas tácticas eran vistas como una degradación de la calidad informativa. Como describe Kilgo y Sinta (2016) «la era digital dejó a muchas organizaciones de prensa luchando para mantener sus normas y valores tradicionales».

Sin embargo, la incontrastable efectividad de este tipo de estrategias en Internet, sumado con el imparable cambio de los consumidores desde el diario de papel hacia las pantallas generó que los medios tradicionales no tuvieran otra alternativa que plegarse a ellas. No todos los medios tomaron el camino de adoptar estas tácticas, pero cada momento que pasa están más extendidas que nunca (Kilgo y Sinta, 2016).

Un reporte interno del New York Times del año 2014 (BuzzFeed News, 2014) dejaba en claro que su competencia no era con los otros medios tradicionales sino con los nativos de Internet como BuzzFeed, Vox, Huffington Post, Business Insider, etc. «El NYT está ganando en periodismo ... [Sin embargo,] estamos fallando en un área secundaria crítica: el arte y la ciencia de hacer llegar nuestro periodismo a los lectores ... Mientras tanto, nuestra ventaja periodística se está achicando». El informe sugiere derribar el *muro* imaginario que separa a la editorial del diario de la sección de negocios, una medida que

años atrás se hubiera visto como un sacrilegio a la profesión.

En el idioma español en particular, la transformación y extensión de estas prácticas se dio un poco después que en inglés pero el fenómeno fue muy similar. Por ejemplo, Palau-Sampio (2016) describe cómo el diario El País creó cuatro secciones (*Verne*, *SModa*, *Icon* y *BuenaVida*) específicamente para atraer lectores en la web utilizando estrategias de clickbait y *tabloidización*.

Apéndice 2

Impacto del clickbait

El clickbait se ha convertido en una técnica muy habitual en los medios de comunicación. Como consecuencia, su impacto ha sido objeto de estudio desde diversas perspectivas. Desde la calidad periodística hasta el modelo de negocio, las preferencias de los lectores, su relación con las noticias falsas e incluso las consecuencias de la distracción en los estados de ánimo.

2.1. Rechazo popular

Es común ver en los trabajos que tratan el clickbait la idea de que es una estrategia que la gente rechaza. Por citar tan sólo algunos ejemplos, Karadzhov et al. (2017) afirma que «todo el mundo lo odia», Jung et al. (2022) que «En el discurso científico y social, la estrategia del clickbait es consistentemente presentada como una acción moralmente reprobable» y Mukherjee et al. (2019) que «casi todos los artículos de este tema ... describen al clickbait como una molestia».

Algunas de las afirmaciones tienen respaldo por ejemplo en encuestas o cuestionarios: Scacco, Joshua and Muddiman, Ashley (2016) muestran que los titulares con preguntas generan actitudes más negativas y peor expectativa sobre los artículos que los contienen. Por su parte Facebook, aunque sin explicar cómo llevó a cabo la encuesta, reportó en 2014 que el 80 % de sus usuarios «prefirieron titulares que les ayudaran a decidir si querían leer el artículo completo antes de tener que hacer clic» (El-Arini, Khalid and Tang, Joyce, 2014). Facebook continuó con su lucha contra el clickbait y lo justificó diciendo que escucharon «específicamente de la gente que querían ver menos historias con

clickbait» (Peysakhovich, Alex and Hendrix, Kristin, 2016).

Muchas de las explicaciones con respecto al rechazo se basan en la decepción que generan muchas de las noticias al no entregar lo que prometían. Como fue visto en el Capítulo 2 esto no es una condición necesaria del clickbait, pero sí algo que sucede habitualmente. Como explica Munger et al. (2020) muchos títulos generan una asimetría en el mercado de la información, en la que «los consumidores no pueden evaluar la calidad previamente, consumición y evaluación suceden simultáneamente».

Otra explicación está en el hecho de que, como se vio en la Sección 2.2 se apela a sentimientos muy difíciles de resistir: «desde un punto de vista crítico, esta práctica queda más cerca de la manipulación que de la estimulación» (Blom y Hansen, 2015).

Sin embargo, por una razón u otra, el público consume clickbait y esa es la única razón de su existencia. En un contexto de acceso muy fácil a la información, donde vemos muchas más noticias de las que podríamos llegar a consumir, y la lucha es por llamar la atención, renegar del uso de esta táctica en los medios que hablan de noticias duras, podría significar regalarle la mayor cantidad del tráfico a los medios sensacionalistas.

2.2. Discusión en medios de prensa

La controversia sobre el por qué del rechazo al clickbait se ha tratado tanto en la academia como en los propios medios de prensa. Algunos de ellos, en particular los que más utilizan el clickbait, defienden su uso.

Quizás la más conocida de estas confrontaciones fue la discusión generada en torno a la cuenta de *Twitter @SavedYouAClick*[1], que se dedica a arruinar titulares clickbait revelando las respuestas, y que inspiró a este proyecto. Varios medios de comunicación respondieron a la iniciativa mediante artículos de opinión que generaron un intercambio público.

En un artículo titulado «Dejen de hablar del clickbait»[2] Tim Marchman, editor en jefe de *Deadspin* (sitio inicialmente dedicado al deporte y conocido por su propensión al uso del clickbait) argumenta que «la palabra clickbait, utilizada como un epíteto, presenta una tautología como crítica. Si uno publica algo, quiere también que la gente lo lea.» Agrega que «propone una falsa

[1] https://twitter.com/SavedYouAClick/
[2] Shut Up About "Clickbait"

dicotomía moralista entre unas historias que sirven al interés público y otras presentadas cínicamente con el único motivo de que la gente va a leerlas» y critica la vaguedad del término, afirmando que titulares muy distintos fueron acusados de ser clickbait: «Es un término para todo, lo que significa que es un término para nada».

Otro artículo publicado por Charlie Warzel, uno de los editores del sitio Buzzfeed, titulado «Por favor paren de ahorrarme un click»[1] argumenta en la misma dirección preguntándose «¿cuál es el problema de tener que leer toda la historia para obtener toda la historia?» y afirma que «todos los que trabajan en los medios son por lo menos un poco culpables de intentar que otros presten atención a su trabajo, ¡¿y por qué no?!».

Warzel afirma que el clickbait no es un problema serio, en caso de siquiera ser un problema, debido a que las noticias más vistas en Internet tienen generalmente titulares *francos*. Sin embargo, admite también que el problema del clickbait «no está de ninguna manera resuelto» y que «las presiones de un sistema de medios saturado, gobernado por algoritmos inconsistentes y a veces crueles, ha exacerbado sin dudas el problema y cualquiera que haya visto un muro de noticias de redes sociales tiene derecho a sentirse exasperado».

En una línea similar Nilay Patel, editor en jefe de *The Verge*, criticó[2] el arruinar el clickbait acusando a quienes lo hacen de «no ahorrar un click, sino robar una experiencia». Afirma que las historias detrás de los titulares valen la pena y que esa forma de titular es apenas una más, que agrega «suspenso», y que pensar que no vale la pena clickear en ellas para leerlas sólo por su titular es «una reducción al punto del absurdo».

Por su parte, Jake Beckman, dueño de la cuenta @SavedYouAClick y blanco de la mayoría de estas críticas, ha respondido en entrevistas[3] argumentando que lo que genera el arruinar clickbait es «ahorrar tiempo perdido a los lectores» y apuntar contra «artículos que se posicionan en redes sociales para obtener tráfico fácil». Además habla de que los editores piensen «primero en la experiencia del lector» en lugar de retener información.

En otra entrevista[4], Beckman explica que incluso un estupendo artículo puede estar acompañado de un titular clickbait y que «si una persona con una

[1] Please Stop Saving Me A Click

[2] Are you proud of yourself, @savedyouaclick?

[3] This Guy Hates Clickbaity Headlines. Here's Why, The guy who reads drivel on the Web so you don't have to

[4] Why Vox Media Took on @SavedYouaClick Over Tony Soprano Spoiler

cuenta de Twitter puede trastornar la distribución de un contenido, eso indica un problema mucho mayor».

Toda esta discusión tiene, en mi opinión, un problema de fondo que es la definición de clickbait. Tal como se explica en el Capítulo 2, el concepto no es entendido uniformemente, y varias de las defensas que expresan los editores de los medios aplican a definiciones que se enfocan en el contenido y principalmente en las que exigen la decepción. Esto queda claro en una columna[1], donde el editor en jefe de Buzzfeed, uno de los sitios más asociados con el clickbait, explica cómo en su opinión no debería existir esa asociación y que Buzzfeed de hecho no usa clickbait basado únicamente en el argumento de que los artículos dan exactamente lo que prometen. Es decir, utiliza una la definición basada únicamente en la decepción que se analiza en la Subsección 2.1.2.

El problema de la definición del clickbait atraviesa toda la discusión y forma parte importante también de las discrepancias entre la visiones sobre cómo afecta la calidad periodística.

2.3. Calidad periodística

Es difícil de por sí definir y medir la calidad periodística o informativa. Existen muchos trabajos al respecto, y como recogen Kümpel y Springer (2016) muchos de ellos se resumen en cinco grandes dimensiones: relevancia, exhaustividad, diversidad, imparcialidad y precisión. Las discusiones sobre la pérdida o disminución de la calidad periodística se remontan a mucho tiempo atrás, pero han surgido con intensidad desde la introducción de los medios digitales. Por ejemplo, en Gómez-Mompart et al. (2015) se encuesta a 363 periodistas españoles y el 81 % «reconoce un empeoramiento de los patrones de calidad del producto periodístico», siendo «la adaptación a las nuevas tecnologías» uno de los factores más repetidos como causantes.

El uso del clickbait se asocia comúnmente a un empeoramiento de la calidad periodística (Bazaco et al. 2019; García Orosa et al. 2017; García Serrano et al. 2019; Lazauskas et al. 2018; Molyneux y Coddington, 2020). Gunnels (2017) afirma que el 57 % de los periodistas encuestados concuerda con que «los medios de comunicación valoran demasiado las historias o temas que reciben clicks y generan tráfico en lugar de centrarse en lo más importante». Además,

[1] Why BuzzFeed Doesn't Do Clickbait

agrega que «la necesidad financiera apremiante de obtener clicks ha llevado a que los medios, en ocasiones, bajen la prioridad al periodismo de calidad en profundidad para el contenido que genera un mayor tráfico» con métodos «cuestionables» como el clickbait. Igualmente, la mayoría de estas opiniones se refieren a la elección de contenidos y no específicamente a la presentación o a la creación de una brecha de información.

Bazaco et al. (2019) hacen un estudio profundo del concepto y metodología del clickbait, y concluyen que «es una estrategia que afecta a la calidad informativa» mientras que García Orosa et al. (2017) aseveran que los medios (en este caso europeos) utilizan clickbait «en detrimento de los valores periodísticos tradicionales» y que es una estrategia que ejemplifica cómo «algunas decisiones supuestamente innovadoras» para enfrentar las presiones económicas en el contexto del cambio tecnológico que «se han mostrado como renovadas amenazas para la calidad periodística».

Ambos trabajos advierten en que el clickbait puede ser efectivo para captar audiencia rápidamente pero que puede generar una pérdida de credibilidad a mediano plazo debido al alejamiento de los principios básicos del periodismo. Esta idea concuerda con lo expuesto por Zhang et al. (2020) quienes afirman que «usar clickbait constantemente crea desconfianza entre los seguidores y se convierte en un dilema entre mantener los actuales seguidores y conseguir nuevos». Para respaldarlo, encuentran evidencia de que el tráfico tiene una forma de U invertida con respecto al número de clickbaits publicados por los medios (chinos en el caso de este estudio), ver la Figura 2.1. Es decir: el clickbait al principio atrae gente pero después la aleja.

La mayoría de estas acusaciones a la pérdida de calidad periodística apuntan hacia el hecho de que los medios que deciden usar clickbait lo hacen como parte de una estrategia para atraer tráfico que normalmente incluye otras medidas. Así opina Ken Smith, presidente del consejo ejecutivo galés de la Unión Nacional de Periodistas[1]: «Inevitablemente, si el criterio para incluir la historia en el sitio web está determinado por la cantidad de clicks, entonces vamos por un camino muy peligroso ... se estará más inclinado a hacer este periodismo de gratificación instantánea en lugar de examinar cuestiones de interés porque toman tiempo de investigar». De hecho, Mukherjee et al. (2019) muestran que el clickbait es más usado en medios con menos personas con antecedentes periodísticos y, como ya fue analizado previamente, en los medios que sólo tienen

[1] Clickbait: The changing face of online journalism

presencia online. También Palau-Sampio (2016) concluye que la utilización de clickbait en el diario El País (España) «pone de manifiesto el efecto perverso que tiene sobre la calidad de un medio el establecimiento de objetivos de audiencia» que derivan del modelo de negocio que prima en los medios digitales, que «contaminan» los contenidos de un medio prestigioso.

Algunas críticas se centran también en la credibilidad que perciben los lectores cuando una noticia es presentada de esta forma, más allá de la calidad del artículo (Pengnate, 2016; Pengnate et al. 2018; Scacco, Joshua and Muddiman, Ashley, 2016). Hurst (2016) estudia los efectos del clickbait en los niveles percibidos de credibilidad de las noticias y muestran que son negativos: los medios son menos más creíbles cuando usan clickbait y eso se acentúa cuando el medio ya partía de una credibilidad baja.

En definitiva, más allá de las notorias diferencias en las definiciones, existe cierto consenso en considerar al clickbait como negativo para la calidad periodística, o por lo menos asociarlo a estrategias más globales que sí atentan contra ella.

Sin embargo, existe una defensa del clickbait que apunta a que puede generar que más gente se sienta predispuesta a informarse debido a la atracción que genera. Chakraborty et al. (2017) compara el clickbait con el efecto de *tabloidización*, y recuerda que los tabloides ayudaron a que mucha población comenzara a leer información sobre política. Hipotetizan con que el clickbait puede tener más llegada en «la sección de la sociedad que muestra poco interés en seguir el periodismo convencional.» En el mismo sentido, Kilgo y Sinta (2016) opinan que, siempre que no se trate del clickbait que promete falsamente y decepciona, «no necesariamente equivale a una disminución en los estándares periodísticos» y ayuda a aumentar la llegada al público.

2.4. Función de la titulación

Otro eje de la discusión en torno al clickbait pasa por la función de los titulares. Tradicionalmente la literatura da dos funciones principales a los titulares de una noticia: resumir el artículo, ya sea mediante una síntesis o resaltando el punto principal, y el de llamar la atención (Dor, 2003; Kuiken et al. 2017a).

Las críticas a los titulares clickbait radican en que se desbalancean hacia la segunda función, descuidando la primera. Esto es algo que sucede no sólo con el clickbait sino que comenzó con los titulares de tabloides. Dor (2003) analiza

los titulares como optimizadores de la relevancia de la historia que cuentan hacia el lector, y opina que el mejor titular «no debería hacer al lector común leer el artículo», sino que debería escribirse con el objetivo de «asegurarse de que el lector ha recibido el mejor negocio al leer el titular en sí.»

Existen muchos manuales de estilo e intentos de generar normas estándar sobre titulación, aunque no hay un consenso al respecto. Los titulares clickbait no cumplen por lo general con esos estándares, y de hecho Palau-Sampio (2016) hace un análisis detallado en la sección 5.2 de cómo los titulares clickbait de El País (España) violan consistentemente el «Libro de Estilo de Titulación» del propio diario.

Sin embargo, también existe la opinión de que esas reglas no deberían existir, o al menos ser tan estrictas. Incluso Ifantidou (2009) encuentra que los lectores no responden positivamente a los titulares que cumplen con los estándares como el valor informativo, ser cortos y concisos y no ambiguos, sino con la relevancia basada en sus gustos e intereses y valorando otros aspectos como la creatividad.

En algunos casos similares, aunque más extremos, el intento de llamar la atención, generar curiosidad y atraer gente, puede llevar a los medios a titular de forma de que si el lector no lee la noticia completa se lleve una impresión errónea.

2.5. Desinformación

Los títulos, o el avance de las noticias incluyendo los textos, imágenes y/o videos incluidos en las redes sociales, deberían reflejar lo que dice la noticia, o por lo menos no decir o sugerir algo contrario. En el ánimo de aumentar el tráfico, aunque no sea de forma deliberada, las estrategias de clickbait suelen encontrarse con este problema que puede generar desinformación.

La desinformación es un fenómeno muy relacionado al de la pérdida de calidad periodística, y muchos de los trabajos discutidos en esta sección se refieren a ambos problemas. La diferencia en este contexto radica en que la referencia es exclusivamente al hecho de leer un titular que implica algo distinto a la noticia. Por ejemplo, Silverman (2015) encuentran que el 13 % de los artículos que estudian incurre en una disonancia entre el texto del titular y el del artículo, Mukherjee et al. (2019) encuentran que los títulos clickbait suelen tener distinto sentimiento que el artículo principal y Chesney et al. (2017)

tratan el tema en profundidad.

Este problema no es exclusivo del clickbait, tan es así que algunos trabajos lo trataban antes de su aparición (Andrew, 2007; Wegner et al. 1981). Sin embargo, el hecho de que se le asigne considerablemente más importancia a la parte de llamador de atención que a la informativa de los titulares, hace que se intensifique. Incluso, la ambigüedad es una de las formas de generar curiosidad y obligar al usuario a leer todo el artículo para encontrar la respuesta por lo que su uso es a veces deliberado (Molyneux y Coddington, 2020).

Las diferencias pueden parecer sutiles pero sus efectos no lo son. El titular (junto con el resto del avance) es el representante de la noticia, lo que se lee siempre, se lea el texto completo o no, y lo que nos permite decidir si seguir leyendo. Y hasta no decidir que sí y realmente leer el artículo, no podemos saber si el titular era preciso o no. Ecker et al. (2014) tratan los efectos de leer un titular con información engañosa (no necesariamente falsa) incluso cuando después sí se lee el artículo completo. Estos titulares «afectan la memoria del lector, su razonamiento inferencial y sus intenciones de comportamiento, así como las impresiones» que se forman. Leer información errónea sesga a los lectores hacia una interpretación específica y «restringen el procesamiento adicional de la información» porque genera dificultades para actualizar su memoria y corregir los conceptos erróneos iniciales.

Silverman (2015) agrega que la «naturaleza clara y compacta» de los titulares genera una impresión inmediata que «enmarca cómo entienden cualquier información posterior, si es que se sigue leyendo. Por lo tanto, es poco probable que el texto detallado del cuerpo mitigue el daño de un título engañoso o falso.» En ese sentido, Hurst (2016) introduce en la discusión el concepto del primado: una teoría basada en la psicología cognitiva que habla de cómo se relacionan los conceptos en el cerebro subiendo la probabilidad de que se nos vengan a la mente ciertos pensamientos o ideas asociadas. Los titulares activan ciertas ideas y por lo tanto facilitan que pensemos en ideas relacionadas, es un proceso que «permanentemente influencia los juicios, percepciones y conductas».

Wegner (1984) describe los procesos cognitivos por los cuales «la información que se transmite a través de preguntas, negaciones o incluso calificaciones (como la posibilidad de que no sea cierta) se entiende primero y luego se reprocesa para agregar la negación» que se puede convertir entonces en «un apéndice superfluo a la información almacenada». Wagner-Egger y Gygax (2018) tam-

bién tratan el mismo problema y muestran cómo este efecto sucede con información relacionada a teorías conspirativas. El clickbait utiliza ampliamente ese tipo de insinuaciones:

(89) *Rosario Castillo habló del rumor de que "usa peluca porque tiene motas"*

omite expresamente el contenido de la declaración para generar curiosidad, pero al hacerlo difunde el rumor sin aclarar que es falso. Una técnica similar es realizar preguntas en lugar de afirmaciones, tan es así que se popularizó la *Ley de los titulares de Betteridge*[1] que dice que «cualquier titular en forma de pregunta puede ser respondido con un no». La ley no se cumple en todas las ocasiones, como muestra

(90) *Daniel Radcliffe y su novia Erin Darke, ¿se casan?*

pero su existencia habla de la cantidad de titulares y/o avances que sugieren la existencia de algo que luego niegan en el desarrollo de la noticia. Por ejemplo,

(91) *Coronavirus — ¿Qué se sabe de la hidroxicloroquina para combatir al virus?*

(que se muestra en la Figura 1.4) sugiere una relación entre la hidroxicloroquina y el combate al virus del covid-19 aunque la nota explica que no da resultados y

(92) *¿Ser desagradable ayuda a ascender en el trabajo?*

sugiere que ser *desagradable* ayuda en el trabajo aunque la noticia explica que «los investigadores hallaron que los individuos con rasgos de personalidad egoístas, engañosos y agresivos no eran más propensos a haber alcanzado más poder que quienes eran generosos, confiables y en términos generales amables».

No solamente con preguntas se puede sugerir información que no es correcta, sino también omitiendo parte del contexto crucial para entender la noticia. Por ejemplo en

(93) *Una modelo rusa que llamó psicópata a Putin apareció muerta dentro de una maleta tras un año desaparecida*

[1] Ley de los titulares de Betteridge

(que también se muestra en la Figura 1.4) ambos datos son correctos. Cualquier lector puede concluir que el gobierno ruso la asesinó, pero en el título se omite la información crucial de que nada tiene que ver una cosa con la otra: fue su novio quién la mató por «un ataque de celos y una disputa por dinero». También en

> (94) *Reino Unido prohibió el sexo entre personas que no vivan en el mismo domicilio*

se sugiere que la prohibición es específicamente a tener sexo, pero en realidad es apenas algo que surge de una medida tomada debido a la pandemia de Covid-19 y prohibió «las reuniones y desplazamientos entre viviendas» de personas que no vivan juntas. La especificidad del titular es diseñada específicamente para generar curiosidad, la información es técnicamente correcta pero claramente engañosa. Por su parte,

> (95) *Los muertos probables por coronavirus se sumaron a los fallecidos certificados*

es un ejemplo de una práctica también muy común: es necesario entrar a la noticia para conocer que el hecho no sucedió en el país del medio de comunicación (Uruguay) sino otro (Chile), y quienes se queden con el titular seguramente asuman que es una noticia local. Finalmente,

> (96) *Tras pedido de vecinos, se retiró una obra del artista @gallinoart en la que mostraba a su mamá con tapaboca. ¿El motivo? Nos lo cuenta en esta nota*

se escribió en medio de una fuerte discusión social por el uso o no de tapabocas y la redacción deja la duda de si ese sería el motivo por el que se censuró la obra, pero en realidad no tiene nada que ver con el contenido sino que querían cobrar la utilización de la pared donde se encontraba el mural.

Hay veces que el engaño es más sutil. En

> (97) *Brasil: Nacional festejó el triunfo de Bragantino al Inter de Diego Aguirre quien pierde apoyo*

el titular sugiere que el festejo de Nacional se debió a que le ganó el equipo que dirigía Diego Aguirre (ex-jugador y técnico muy identificado con el clásico rival de Nacional). Pero la nota explica que «lo festejó» porque la victoria derivó en

la activación de una cláusula del contrato de un jugador que le dió dinero al club. Nuevamente, los hechos que marca la noticia son estrictamente correctos, pero la conclusión que insinúa no. También es común utilizar esta técnica para generar curiosidad sobre algo que no existe en verdad, por ejemplo en

> (98) *¿Qué dijo Marcelo Gallardo al ser consultado sobre la selección uruguaya? Tras el triunfo de River Plate ante Platense, al Muñeco le preguntaron sobre el fuerte rumor que se instaló sobre su arribo a la Celeste*

es verdad todo lo que se dice, pero la respuesta fue «No voy a hacer ningún comentario». Este caso es un claro ejemplo de por qué se relaciona tanto al clickbait con la decepción (véase Subsección 2.1.2). Los lectores esperan una declaración y se encuentran con que no existió ninguna. Pero lo que hace a esta noticia un clickbait no es eso, sino el hecho de que genera curiosidad ocultando parte de la información: la respuesta.

Volviendo a la diferencia entre las dos definiciones propuestas (la definición y la operativa), la gran diferencia entre ellas es que una juzga la intención de quien escribe y la otra no. Si se quisiera realmente saber si la omisión fue a propósito, no siempre es claro y generaría otra zona de subjetividades.

> (99) *Cesan a dos gerentes de ANII: ganaban más que el presidente*[1]

no aclara que se refiere al presidente del propio organismo y no al presidente de la república. Se puede ver en los comentarios del tweet que mucha gente interpretó lo segundo, normalmente cuando se dice «el presidente» sin más contexto se habla del presidente del país, pero ¿fue una omisión intencional? Es difícil saberlo y por eso es tan importante tener una definición operativa y saber que sí se trata de un clickbait según ella.

Chesney et al. (2017) hablan justamente del engaño que se produce cuando un titular no es coherente con la noticia. Sobre titulares que «si bien ocultan información y por lo tanto inducen a un error» afirman que no son clickbait porque «no obligan al lector a hacer click para conocer las conclusiones de la noticia, sino que directamente ofrecen una conclusión engañosa». Lo más importante a tener en cuenta, desde mi punto de vista, es si la omisión genera una historia más atractiva (y no sólo distinta) que la original. En ese caso, es

[1] Ejemplo tomado de un medio que no forma parte de los elegidos para TA1C.

la omisión la que genera curiosidad, aunque no siempre estrictamente mediante una brecha de información, calificando como clickbait. En caso de que la historia engañosa no sea más atractiva, probablemente se trate sencillamente de un error o de manipulación y no debería clasificarse como clickbait. El artículo agrega que este tipo de titulares «carece de las típicas características observables del clickbait» y es por lo tanto «más sutil» y «improbable de ser detectado por estos medios estilométricos». La observación es correcta e importante, pero no puede ser de ninguna manera un argumento para descartar este tipo de noticias: el intento es de definir y caracterizar al clickbait, no sólo a los ejemplos fáciles de clasificar.

Es distinto cuando la omisión es, o por lo menos parece, intencionadamente para que la noticia parezca otra más atrapante o llamativa.

> (100) ●*ÚLTIMA HORA — Muere un voluntario que participaba en los ensayos de la vacuna de Oxford en Brasil*

(que se muesta en la Figura 1.3) podría ser una muy importante información sobre los posibles efectos secundarios de las vacunas contra el Covid-19, si no fuera porque el voluntario había recibido el placebo en los ensayos como indica la nota. De haberse contado este detalle tan importante en el titular, la noticia no tendría siquiera sentido. Estos efectos son a veces generados a propósito porque la noticia que se sugiere llama más la atención que la real. También se pueden dar sin ser buscados, quizás la única intención del titulador es realmente generar curiosidad para aumentar el tráfico y poder recaudar más dinero de publicidad. Sin embargo, el engaño se produce de todas formas. La situación es distinta cuando la desinformación se difunde a propósito.

2.6. Noticias falsas

La difusión deliberada de información falsa es un fenómeno que ha acompañado a la humanidad desde hace siglos. Sin embargo, su propagación en los medios digitales y redes sociales, ha tomado preponderancia en los últimos años, generalmente bajo el nombre de *fake news* y en particular luego de las elecciones presidenciales de Estados Unidos en el 2016.

Muchas de las tipologías de noticias falsas en este contexto incluyen al clickbait como una de las formas de difusión de *fake news* (Baptista y Gradim, 2020; Bourgonje et al. 2017; Ghanem et al. 2020; Karadzhov et al. 2017;

Mourão y Robertson, 2019; Park et al. 2020; Volkova et al. 2017; Zannettou et al. 2019; X. Zhou et al. 2020). También Wikipedia[1] lo incluye. Algunos trabajos explican el por qué de la inclusión. Edson C. Tandoc Jr. y Ling (2018) examinan 37 artículos académicos que definen *fake news* e incluyen al clickbait, ejemplificando con un titular viral en Facebook que en realidad, al clickear, no iba hacia una noticia sino hacia un sitio de marketing, pero que generó miles de comentarios como si fuera real. Y. Chen et al. (2015) también incluyen al clickbait como parte, indicando que «ha estado implicado en la rápida propagación de rumores y desinformación». Shevon Desai and Jo Angela Oehrli (2022) definen que «a veces, estas historias pueden ser propaganda diseñada intencionalmente para engañar al lector, o pueden estar diseñadas como clickbait, escritas para obtener incentivos económicos». Sobre este punto trata una investigación sobre operadores de páginas de noticias falsas y clickbait[2] que operan en redes sociales y tuvieron una importante y negativa ingerencia en la política de algunos países. Por su parte Zannettou et al. (2019) pese a incluirlo como parte de los tipos de noticias falsas, advierte que «es uno de los tipos menos severos, porque si un usuario lee todo el contenido, puede distinguir si el título o la miniatura son engañosos».

La inclusión del clickbait se refiere principalmente al tipo de desinformación tratado en el punto anterior, a las publicidades que se hacen pasar por noticias y al hecho de que muchas de las noticias falsas se difunden utilizando las estrategias del clickbait para generar más tráfico. Sin embargo, nuevamente, que dos estrategias se utilicen habitualmente juntas no quiere decir que sean lo mismo y no puede decirse que una noticia sea falsa por cómo está presentada en su titular. Así lo aclaran también, por ejemplo, Zhang et al. (2020) quienes afirman que «muchos estudios han confundido el concepto de clickbait y el de fake news aunque son completamente distintos en esencia.»

Pese a que muchas de las críticas al clickbait vienen acompañadas de discrepancias en su definición o de confusión con el concepto de noticias falsas, su uso puede llevar, de todos modos, a problemas de desinformación. Es debido a ello, entre otras razones, que existe un rechazo al clickbait. Incluso cuando las noticias son verdaderas, otros motivos como el desvío involuntario de la atención molestan habitualmente al público lector.

[1] Fake News

[2] The tech giants are paying millions of dollars to the operators of clickbait pages, bankrolling the deterioration of information ecosystems around the world

2.7. Distracción

La cantidad de noticias que se consumen y la lucha por capturar el interés del público, que lleva a hablar de la economía de la atención, es en lo que se basa otra de las críticas a las consecuencias del clickbait: la distracción que genera (Chakraborty et al. 2016; Paasonen, 2016).

Una nota del *New York times*[1] discute este problema y, aunque reconoce que el clickbait no es lo único para culpar al respecto, lo posiciona como un facilitador en una relación adictiva que tenemos con la distracción. Afirma que las consecuencias de esta distracción pueden llevar al estrés, mal humor y baja productividad. En este caso, no importa si el artículo cumple o no con lo prometido, el daño está hecho al momento de cambiar el foco de la atención hacia una noticia que de otra forma no hubiéramos consumido. En Isikman y Cavanaugh (2016) muestran como someterse a eventos que evocan curiosidad mientras se está disfrutando de otra actividad tiene efectos negativos sobre el disfrute.

La curiosidad puede fácilmente distraer recursos limitados como nuestro tiempo y atención hacia cosas que no nos interesan. El hecho de que cierta información nos de curiosidad es un efecto de corto plazo, una necesidad de actuar, en lugar de un genuino sentimiento de preferencia (Aslan et al. 2021; Shin y Kim, 2019; Tang et al. 2022).

El clickbait es, por otro lado, apenas una de las tantas estrategias utilizadas para captar la atención y provocar curiosidad. Si bien algunos trabajos lo tratan como uno de los causantes de la sobreexposición a este tipo de eventos, es en cualquier caso uno más de un grupo.

Igualmente, esto muestra como el clickbait puede tener efectos incluso en campos no relacionados con las noticias y la información.

2.8. Fuera de la prensa

Este trabajo se centra en el clickbait en los medios de prensa, el que afecta a las noticias que se difunden tanto en sus sitios web como en redes sociales o agregadores. Sin embargo, es un fenómeno que se ha extendido hacia otras áreas, donde se vieron los buenos resultados que puede generar en el aumento de tráfico y se copió la estrategia.

[1] Click Bait Is a Distracting Affront to Our Focus

Una de esas áreas es la academia. Portia Roelofs and Max Gallien (2017) discuten cómo algunos artículos académicos de baja calidad toman posiciones polémicas para generar discusión. Esa polémica causa muchas citas, mejorando las métricas por las que generalmente se califican los artículos. Asimismo, Lockwood (2016) estudia los artículos publicados en la revista científica *Frontiers in Psychology* y descubre que los que usan algunas características propias del clickbait tienen en promedio mejor desempeño.

El clickbait es usado también para atraer la atención hacia videos, en particular en *YouTube* (Rajput, 2019; Shang et al. 2019; Zannettou et al. 2018). Este tipo de contenidos, a diferencia de las noticias de prensa, son consumidos por niños. Por eso, el reporte Foundation (2023) coloca al clickbait en videos como uno de los tipos de contenido altamente emocional, parte del diseño persuasivo, que impacta en el desarrollo social, mental y físico de los niños. También, habla de cómo los adolescentes tempranos (13 a 15 años) se caracterizan por un «idealismo que tiende al pensamiento polarizado» por lo que son más proclives a responder al contenido cargado de emociones que suele llamar la atención a través del clickbait.

Otros campos donde se pueden encontrar efectos del clickbait son la publicidad, donde también se han aplicado conceptos similares, presentando productos pero ocultando algunas características para apelar a la brecha de información con buenos resultados (Menon y Soman, 2002; Sevilla y Meyer, 2020) y algunas redes sociales como Instagram donde se lo utiliza para mejorar la llegada de la publicidad (Ha et al. 2018).

Apéndice 3

Cuentas incluidas en TA1C

Como se explica en la Subsección 4.1.3, para crear los conjuntos de datos TA1C y TA1C-unlabeled consideré la cuenta principal de *Twitter* de cada medio junto con sus cuentas secundarias, normalmente referidas a secciones temáticas o regionales. A continuación, se listan todas las cuentas consideradas por medio:

El País (Uruguay)

- @elpaisuy
- @Ovaciondigital
- @tvshowuru

Infobae (Argentina)

- @teleshowcom
- @infobae
- @infobaeamerica
- @infobaetecno
- @infobaeeconomia
- @infobaevzla
- @infobaecultura
- @infobaedeportes
- @infobaepolitica
- @infobaemexico
- @infobaecolombia

Clarín (Argentina)

- @clarincom
- @clarinfama
- @Clarin_deportes
- @MundoClarin
- @PoliticaClarin
- @clarinrural
- @ClarinSociedad
- @clarintecno
- @CulturaClarin
- @clarineconomico
- @ClarinBuenaVida
- @PolicialClarin

La Nación (Argentina)

- @LANACION
- @DeportesLN
- @Politica_ln
- @LNcampo
- @LNTecnologia

El País (España)

- @el_pais
- @elpais_economia
- @elpais_america
- @elpais_deportes
- @elpais_tec
- @elpais_inter
- @elpais_espana
- @elpais_cultura
- @elpaismexico
- @BuenaVidaEP
- @smoda
- @verne

BBC (Inglaterra)

- @bbcmundo

- @BBCDeportes

El Comercio (Ecuador)

- @elcomerciocom
- @deportesec

El Mundo (España)

- @elmundoes
- @ElMundoDeportes
- @ElMundoEspana
- @ElMundoData
- @ElMundoInter
- @ElMundoViajes
- @ElMundo_Madrid
- @elmundocultura
- @elmundociencia
- @elmundobcn
- @LunaMetropoli

El Universal (México)

- @El_Universal_Mx
- @Univ_Cultura
- @Univ_Mundo
- @UnivDeportes
- @Univ_Estados
- @Univ_Cartera

El Tiempo (Colombia)

- @ELTIEMPO
- @BogotaET

El Universal (Venezuela)

- @ElUniversal

Prensa Libre (Guatemala)

- @prensa_libre
- @tododeportes_pl
- @Economia_pl

- @Escenario_pl
- @Tecno_pl

El Comercio (Perú)

- @elcomercio_peru
- @Lima_ECpe
- @EconomiaECpe
- @Luces_ECpe
- @Mundo_ECpe
- @PeruECpe
- @Politica_ECpe
- @TVmas_ECpe
- @tecnocienciaEC

La Prensa Gráfica (El Salvador)

- @prensagrafica
- @LPGPolitica
- @LPGDptos
- @LPGJudicial
- @LPGEconomia

Diario Libre (República Dominicana)

- @diariolibre

La Vanguardia (España)

- @LaVanguardia

ABC (España)

- @abc_es
- @abc_tecnologia
- @abc_mundo
- @abc_deportes
- @abc_ciencia
- @abc_cultura
- @abc_familia
- @abc_conocer
- @abc_gente

- @abceconomia

La Tercera (Chile)

- @latercera
- @QuePasaCL
- @CultoLT
- @eldeportivolt
- @revista_paula
- @pulso_tw

Listado de Tweets

Listado de todos los tweets usados como ejemplos en el documento junto con su URL

1. *Elecciones en Estados Unidos: por qué el recuento de votos podría demorarse más de los normal* (https://twitter.com/clarincom/status/1323409639069880321)
2. *El bebé de Zaira Nara ya tiene nombre y seguro no es el que te imaginás* (https://twitter.com/pantallazomvd/status/1225819618465472516)
3. *África erradica la polio: cómo consiguió este histórico hito* (https://twitter.com/bbcmundo/status/1298308806557007872)
4. *WhatsApp dejará de funcionar en varios modelos de iPhone: ¿te va a tocar cambiar de móvil?* (https://twitter.com/abc_tecnologia/status/1372136227705655297)
5. *Esto es lo que tienes que ahorrar al mes para poder pagar la universidad de tus hijos* (https://twitter.com/abc_familia/status/1395408985210097673)
6. *#Moquegua: así se desarrolla el segundo día de protestas contra el proyecto minero #Quellaveco* (https://twitter.com/PeruECpe/status/1192103743266967553)
7. 🚨 *Esto es lo que se sabe de la detonación de un artefacto que dejó seis menores heridos en Buenaventura* 👇 (https://twitter.com/ELTIEMPO/status/1366083060043808769)
8. *Estas son 5 plataformas de entretenimiento para diversificar sus horas de diversión* (https://twitter.com/Tecno_pl/status/1165977886303686658)
9. *Final de la NBA: 5 claves del duelo entre los Golden State Warriors y los Cleveland Cavaliers por el trono del baloncesto mundial* (https://twitter.com/BBCDeportes/status/1001895018142076928)
10. *Claves para encontrar las mejores ofertas en el Black Friday 2019* (https:

//twitter.com/abceconomia/status/1196712216097107968)

11. *Los elegidos de Ariel Schettini: virginidad, desborde y tragicomedia* (https://twitter.com/CulturaClarin/status/1209257462852980736)
12. *Insólito: Kyllian Mbappé terminó en muletas por el patadón de un rival que vio la roja y se retiró del fútbol* (https://twitter.com/Clarin_deportes/status/1286817410193002496)
13. *La vertiginosa vida de Christian Nodal tras revelar su romance con Belinda* (https://twitter.com/InfobaeMexico/status/1298462691158179840)
14. *Asesinan de un tiro en la cabeza a un repartidor de pizzas en La Plata* (https://twitter.com/clarincom/status/1180873338341724162)
15. *Steve Irwin: las últimas palabras del Cazador de Cocodrilos antes de morir hace 15 años* (https://twitter.com/elcomercio_peru/status/1434940106566209538)
16. *La sorprendente confesión de Juana Viale sobre su vínculo con los futbolistas* (https://twitter.com/LANACION/status/1300292910773141505)
17. *Las dudas en el lanzamiento del equipo de fútbol de David Beckham en Miami* (https://twitter.com/BBCDeportes/status/958726707137130496)
18. *Ronaldinho y Gimnasia: el deseo del brasileño de volver al fútbol y ser dirigido por Diego Maradona* (https://twitter.com/clarincom/status/1276518412626341892)
19. *La masacre en el aeropuerto de Kabul deja al menos 170 muertos* (https://twitter.com/DiarioLibre/status/1431368481392447488)
20. *Vicentín: una decisión con un rédito político incierto* (https://twitter.com/PoliticaClarin/status/1270741564659302401)
21. *Se retiró Mascherano: el extraordinario viaje del caudillo al que sólo le quedó una materia pendiente en la Selección* (https://twitter.com/Clarin_deportes/status/1328171071212752896)
22. *Estados Unidos: un escándalo familiar de Donald Trump empaña el comienzo de la convención republicana* (https://twitter.com/clarincom/status/1297557329932619782)
23. *Unanimidad en Diputados: suspenden al diputado salteño Juan Emilio Ameri tras el escándalo* (https://twitter.com/LANACION/status/1309290404584386560)
24. *Científicos logran leer, sin abrirla, una carta de hace 300 años* (https://twitter.com/abc_es/status/1366875974923599874)

25. *Sin abrir, revelan contenido de carta de hace 300 años con un rayos X dental* (https://twitter.com/El_Universal_Mx/status/1366819966868201475)
26. *Identifican una hormona que regula el envejecimiento* (https://twitter.com/clarincom/status/871680794703441920)
27. *Hallaron el método más increíble para adelgazar: comer tierra* (https://twitter.com/ClarinBuenaVida/status/1084481342635536386)
28. *#Turf — Quince jockeys dieron positivo de coronavirus en California y cancelaron las carreras del fin de semana próximo* (https://twitter.com/DeportesLN/status/1283812412009975808)
29. *La golfista mexicana María Fassi dio positivo por Covid-19* (https://twitter.com/UnivDeportes/status/1313609982756044800)
30. *Un golfista dio positivo de coronavirus en pleno torneo del PGA Tour* (https://twitter.com/Clarin_deportes/status/1274124820318486528)
31. *Juventud de las Piedras confirmó 17 casos positivos de COVID-19* (https://twitter.com/Ovaciondigital/status/1335709623135981568)
32. *Un jugador del Real Madrid dio positivo de coronavirus en la previa del duelo con el Inter* (https://twitter.com/Clarin_deportes/status/1323222228905984001)
33. *Posponen partido de la NBA por positivo al Covid-19* (https://twitter.com/ElUniversal/status/1361686783734513668)
34. *Un atleta argentino que ya tiene su lugar en los Juegos de Tokio dio positivo de coronavirus* (https://twitter.com/Clarin_deportes/status/1283463276920213506)
35. *Un congresista demócrata demanda a Trump por el asalto al Congreso* (https://twitter.com/ELTIEMPO/status/1361929872554221570)
36. *En Día del Trabajo, Donald Trump critica al líder del Sindicato de Trabajadores más grande Estados Unidos* (https://twitter.com/Univ_Mundo/status/1036697841635340288)
37. *Elecciones en Bolivia: tres candidatos piden inhabilitar al partido de Evo Morales* (https://twitter.com/clarincom/status/1285330736795594753)
38. *El campeonato ecuatoriano de fútbol se quedó sin su patrocinador principal* (https://twitter.com/elcomerciocom/status/1307071055505371137)
39. *Vandalizan otra escultura en Berlín; esta vez grafitean nombre de un rapero y equipo de fútbol* (https://twitter.com/Univ_Cultura/status/

1320072314814386177)

40. *Leila Sucari: una repentina sensualidad* (https://twitter.com/CulturaClarin/status/1306523832090595330)
41. *Esmeralda Mitre se hizo un nuevo hisopado y dio negativo: Esto es un complot* (https://twitter.com/clarinfama/status/1300970081535614978)
42. *Joe Biden, el aliado de la luz marcado por la tragedia familiar* (https://twitter.com/elpaisuy/status/1323529236154638336)
43. *La fundación Slim, el empresario Sigman desde Madrid y un reto a Ginés González García: cómo se gestó el anuncio de la vacuna de Oxford* (https://twitter.com/infobaepolitica/status/1293711747967143937)
44. *El estudiante que resolvió dos grandes problemas matemáticos por impuntualidad* (https://twitter.com/abc_ciencia/status/1168448479934537728)
45. *25 años de la PlayStation: el enojo que cambió la historia* (https://twitter.com/LNTecnologia/status/1202269723867537408)
46. *De la inquietante gambeta a Vladimir Putin a la foto sin revancha con David Beckham: cinco anécdotas inéditas de Diego Maradona* (https://twitter.com/infobaedeportes/status/1323120277841170432)
47. *La falta envido que desató un crimen en un bar de pueblo* (https://twitter.com/PolicialClarin/status/1156520595120644096)
48. *Caminaba con su mujer y sus hijas, quedó en medio de un cruce entre bandas y lo mataron a balazos* (https://twitter.com/PolicialClarin/status/1175433002358333441)
49. *La agricultura consume alrededor del 70 % del agua disponible en el planeta y la forma en la que la obtiene no siempre es legítima* (https://twitter.com/bbcmundo/status/1306325254613151745)
50. *Bacterias vampiro, caníbales zombis... y otros microbios de miedo. Sorprendentes características de los microorganismos* (https://twitter.com/elcomerciocom/status/1323771690757857290)
51. *La fuga rocambolesca del ex presidente de Renault deja atónitos a Japón y a Francia* (https://twitter.com/MundoClarin/status/1321967438263513094)
52. *Cómo es un huracán espacial, el espectacular fenómeno detectado por primera vez en la Tierra* (https://twitter.com/bbcmundo/status/1367895735593746432)
53. *La foto de un zorro y un cóndor juntos se hizo viral y conmueve al mundo*

entero (https://twitter.com/elpaisuy/status/1269662970830901254)

54. *La sorprendente confesión de Juana Viale sobre su vínculo con los futbolistas* (https://twitter.com/LANACION/status/1300292910773141505)
55. *Las bolsas mundiales registran fuertes alzas impulsadas por los nuevos planes de estímulos contra el coronavirus* (https://twitter.com/infobaeeconomia/status/1272879239411298304)
56. *La fuerte acusación de Thomas Müller al Atlético Madrid del Cholo Simeone* (https://twitter.com/infobaedeportes/status/1319340437715517440)
57. *Ballena en Puerto Madryn: el impresionante video de una ballena jugando con una mujer que estaba remando en el sur de Argentina. No te pierdas en este video el entrañable momento de este insólito encuentro* (https://twitter.com/prensa_libre/status/1433827733738491908)
58. *La cruenta batalla por el Cártel de Sinaloa: el Mayo, los hijos del Chapo y el grupo de Caro Quintero se disputan su control* (https://twitter.com/infobaeamerica/status/1321767749132210182)
59. *Leticia Brédice publicó fuertes mensajes contra su novio: lo acusó de ladrón y mala gente* (https://twitter.com/clarinfama/status/1247899492549984256)
60. *Video: cocodrilo se comió a dos tiburones en una playa de Australia* (https://twitter.com/prensagrafica/status/1362590001737977857)
61. *Las comunidades lamentan el pacto de mínimos: ¿Para qué hace falta un Ministerio de Educación si no coordina?* (https://twitter.com/elmundoes/status/1299523415594934272)
62. *El Salvador — ¿Cuál es el perfil de estudios de los diputados? 82 % del actual pleno legislativo reporta una carrera de formación profesional. Y 25, de 84 diputados, reportan una carrera relacionada con el ramo jurídico.* (https://twitter.com/prensagrafica/status/1364564038710161408)
63. *¿Deja a Messi por CR7? Juventus quiere a Luis Suárez* (https://twitter.com/elpaisuy/status/1299330084273303552)
64. *¿Fútbol o rugby? Fútbol gaélico, la combinación de dos deportes en la que la Argentina se destaca* (https://twitter.com/DeportesLN/status/1309540663080423424)
65. *¿Cuál es la forma correcta de acariciar a tu gato? La ciencia tiene la respuesta* (https://twitter.com/abc_ciencia/status/1155784979126898688)

66. *¿Te acordás qué hacías, dónde estabas cuando Milton Wynants ganó la medalla de plata en Sídney 2000?* (https://twitter.com/Ovaciondigital/status/1307704183928377352)
67. *Graciela Alfano confesó que tiene nuevo novio; ¿quién es?* (https://twitter.com/tvshowuru/status/867080249930125316)
68. *WhatsApp dejará de funcionar en varios modelos de iPhone: ¿te va a tocar cambiar de móvil?* (https://twitter.com/abc_tecnologia/status/1372136227705655297)
69. *Al menos el 40 % de los establecimientos están cerrados por coronavirus ¿Quieres saber qué locales están abiertos en tu barrio? Pincha aquí* 👇 (https://twitter.com/EM_GranMadrid/status/1249975458114023424)
70. *Al ritmo de El baile de los que sobran. De esta forma cientos de chilenos siguen celebrando el triunfo del voto a favor de cambiar la Constitución. Aquí le contamos más sobre esta histórica jornada en el país austral.* ➡️ (https://twitter.com/ELTIEMPO/status/1320555864625991682)
71. 🚨 *Más de 15.000 peruanos marcharon para rechazar el Gobierno interino. Este Congreso no me representa, fue una de las consignas de quienes protestaron. Le contamos* 👉 (https://twitter.com/ELTIEMPO/status/1327416690506944518)
72. *Memo Ochoa regresa para el juego contra Tigres* 🚨 *¿Les agrada la idea, americanistas?* 🤔⬇️ (https://twitter.com/UnivDeportes/status/1320939684927803393)
73. *Instituto Geofísico reporta sismo de 3.5 grados en Simón Bolívar, Guayas #ATENCIÓN — Instituto Geofísico reporta sismo en Guayas este 28 de noviembre. ¿Sintió el temblor?* (https://twitter.com/elcomerciocom/status/1332817818497200133)
74. *[OPINIÓN] La ambigüedad y el mal entendido en la política: ecos de la carta de Cristina Kirchner — Por Federico González* (https://twitter.com/infobaepolitica/status/1323366332604256257)
75. *Columna — "El partido ha dejado sus huellas en la corrosión de la democracia, cuya consecuencia es Jair Bolsonaro. El PSDB no es sólo otro partido que ha contribuido a la más reciente escalada autoritaria de Brasil"* ✍️ *@brumelianebrum* (https://twitter.com/elpais_america/status/1332202622334480384)
76. *Columna de Sebastián Rodríguez, psicólogo: Mi yerno se fue en la pande-*

mia... ¿qué hago con mi hija? (https://twitter.com/QuePasaCL/status/1355206573669269509)

77. *¿Por qué el Gobierno vuelve a atacar el proceso de paz?* (https://twitter.com/ELTIEMPO/status/1440975806411558913)
78. *La primavera atrapada en un viejo sobretodo* (https://twitter.com/clarincom/status/1309614341390577666)
79. *Sapiens, cautivos de nuestro ADN* (https://twitter.com/CulturaClarin/status/1238782629056598016)
80. *Apóstrofes* (https://twitter.com/ElUniversal/status/1322860443191435269)
81. *Nicole ahora es experta* (https://twitter.com/tvshowuru/status/1295417529230295040)
82. *¿Cuál es la distancia ideal que hay que correr por semana?* (https://twitter.com/ClarinBuenaVida/status/682687777851490305)
83. *Si el objetivo es mejorar la salud, no hay que correr más de 9 kilómetros por semana* (https://twitter.com/ClarinBuenaVida/status/683383588738445312)
84. *¿Cuál es la situación de todos los dirigentes sudamericanos involucrados en el FIFAgate de 2015? ¿Cuántos siguen presos y cuántos recuperaron su libertad parcial o total?* (https://twitter.com/Referiuy/status/1245795646143168513)
85. *"Tres de los acusados fallecieron, uno está preso, dos esperan sentencia y Eugenio Figueredo goza de libertad provisional y dijo haber devuelto lo que tenía que devolver"* (https://twitter.com/Referiuy/status/1245323364215947266)
86. *Uruguay-Brasil, por las eliminatorias sudamericanas: horario, TV y formaciones* (https://twitter.com/LANACION/status/1328697104127172608)
87. *#ATENCIÓN — Messi y un comunicado que pone en alerta al FC Barcelona* (https://twitter.com/elcomerciocom/status/1298321907331342337)
88. *¿Qué dijo Marcelo Gallardo al ser consultado sobre la selección uruguaya? Tras el triunfo de River Plate ante Platense, al Muñeco le preguntaron sobre el fuerte rumor que se instaló sobre su arribo a la Celeste* (https://twitter.com/Ovaciondigital/status/1462899197963915278)
89. *Rosario Castillo habló del rumor de que "usa peluca porque tiene motas"*

(https://twitter.com/pantallazomvd/status/1293182367926857734)

90. *Daniel Radcliffe y su novia Erin Darke, ¿se casan?* (https://twitter.com/pantallazomvd/status/841282349623062533)
91. *Coronavirus — ¿Qué se sabe de la hidroxicloroquina para combatir al virus?* (https://twitter.com/tecnoyocienciaEC/status/1265754128527118337)
92. *¿Ser desagradable ayuda a ascender en el trabajo?* (https://twitter.com/ClarinBuenaVida/status/1300879236173934592)
93. *Una modelo rusa que llamó psicópata a Putin apareció muerta dentro de una maleta tras un año desaparecida* (https://twitter.com/infobae/status/1504492321374617600)
94. *Reino Unido prohibió el sexo entre personas que no vivan en el mismo domicilio* (https://twitter.com/elpaisuy/status/1268345989670359041)
95. *Los muertos probables por coronavirus se sumaron a los fallecidos certificados* (https://twitter.com/Subrayado/status/1274440073526853632)
96. *Tras pedido de vecinos, se retiró una obra del artista @gallinoart en la que mostraba a su mamá con tapaboca. ¿El motivo? Nos lo cuenta en esta nota* (https://twitter.com/Canal4_UY/status/1286010198902149120)
97. *Brasil: Nacional festejó el triunfo de Bragantino al Inter de Diego Aguirre quien pierde apoyo* (https://twitter.com/futbolportaluy/status/1469385770767167495)
98. *¿Qué dijo Marcelo Gallardo al ser consultado sobre la selección uruguaya? Tras el triunfo de River Plate ante Platense, al Muñeco le preguntaron sobre el fuerte rumor que se instaló sobre su arribo a la Celeste* (https://twitter.com/Ovaciondigital/status/1462899197963915278)
99. *Cesan a dos gerentes de ANII: ganaban más que el presidente*[1] (https://twitter.com/Subrayado/status/1359820069786836995)
100. 🔴*ÚLTIMA HORA — Muere un voluntario que participaba en los ensayos de la vacuna de Oxford en Brasil* (https://twitter.com/el_pais/status/1318977055833378824?s=03)

[1]Ejemplo tomado de un medio que no forma parte de los elegidos para TA1C.