



**NATURAL LANGUAGE PROCESSING COMO HERRAMIENTA PARA
DETECTAR FAKE NEWS: APLICADO A LAS ELECCIONES DE ALCALDÍA
DE CALI PARA EL PERIODO 2020-2023**

JUAN DAVID PRADA

DIRECTORA DE PROYECTO

LUCIANA MANFREDI

UNIVERSIDAD ICESI

FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONOMICAS

ECONOMÍA

SANTIAGO DE CALI

2019

Tabla de Contenido

Resumen	3
Justificación y viabilidad de la investigación.....	5
Objetivos	5
Definición de términos.....	6
Natural Language Processing.....	6
Fake News	7
Marco Normativo Colombiano	7
Efectividad de las FN	8
Casos relevantes de FN en Colombia.....	8
Casos relevantes de FN en el mundo.....	9
La lucha contra las FN	9
Proceso de clasificación de las FN.....	11
Algoritmos Clasificadores.....	12
Caso.....	16
Métricas	17
Selección de algoritmo.....	19
Resultados.....	20
Conclusión.....	20
Bibliografía	22

Resumen

En el ambiente político y económico, las noticias falsas o *fake news* que son creadas principalmente con el objetivo de desinformar, han encontrado en las redes sociales una plataforma que facilita su “propagación” masiva, logrando exceder la capacidad de validar manualmente las noticias, debido a la velocidad con la que son generadas. El hecho de que a pesar de la reducción de los niveles de analfabetismo y del avance de la tecnología, las personas continúan presentando dificultades en determinar si una noticia es falsa o no, junto a la facilidad de propagación en las redes sociales, hacen de las *fake news* un instrumento de desinformación masiva, que logra crear opiniones sesgadas hacia una persona, empresa o una idea. En la actualidad, el *Natural Language Processing* es un área que está generando grandes avances en la validación automática del grado de veracidad de una noticia, lo cual reduce la facilidad con la que las *fake news* son propagadas y logran su propósito de desinformar.

Palabras Claves: fake news, algoritmos, clasificación, redes sociales

ABSTRACT

In the political and economic environment, the fake news are created mainly with the objective of misinformation, they have found in social networks a platform that facilitates its massive “propagation”, exceeding the ability to manually validate the news, due at the speed with which they are generated. The fact that despite the reduction in illiteracy levels and the advancement of technology, people continue to present difficulties in determining whether a story is false or not, together with the ease of propagation in social networks, make fake news an instrument of mass misinformation, which manages to create biased opinions towards a person, company or an idea. Currently, Natural Language Processing is an area that is generating great advances in the automatic validation of the degree of veracity of a story, which reduces the ease with which fake news is propagated and achieve its purpose of misinformation.

Keywords: fake news, algorithms, classification, social networks

Justificación y viabilidad de la investigación

La mayor parte de los adelantos en la automatización en la detección de noticias falsas se ha llevado a cabo en el idioma inglés, es por este motivo que este trabajo pretende hacer la clasificación de noticias falsas para el idioma español y mostrar el gran reto que se tiene por delante debido a la escasez de herramientas. Precisamente debido a la dificultad de encontrar bases de datos con noticias falsas en español, se decidió por usar la red social Twitter como fuente primaria de información y aplicarlo a sujetos que tienen grandes incentivos a promover noticias falsas como los candidatos a una alcaldía.

Objetivos

Objetivo General

- Crear un algoritmo que permita la clasificación de tweets escritos en español como falsos o no.

Objetivo Especifico

- Comprender la necesidad de mejorar las condiciones necesarias para que los algoritmos efectúen una mejor clasificación de las FN escritas en español.
- Comprender el rol de las redes sociales en la masificación de las FN.
- Aplicar el algoritmo creado a los principales candidatos para el periodo 2020-2023 de la Alcaldía de Cali.

Definición de términos

Para entender los conceptos de Natural Language Processing (en adelante NLP) y Fake News (en adelante FN), se referencian las definiciones propuestas por algunos autores reconocidos en el estudio de la misma.

Natural Language Processing

Para Chowdhury (2003) el NLP es:

Un área de investigación y aplicación que explora cómo se pueden usar las computadoras para comprender y manipular el texto o el habla del lenguaje natural para hacer cosas útiles. Los fundamentos del NLP se encuentran en una serie de disciplinas como: la informática y las ciencias de la información, la lingüística, las matemáticas, la ingeniería eléctrica y electrónica, la inteligencia artificial y robótica, psicología, etc. Las aplicaciones del NLP incluyen una serie de campos de estudios, como la traducción automática, el procesamiento y resumen de textos en lenguaje natural, interfaces de usuario, multilingüe y recuperación de información en varios idiomas (CLIR), reconocimiento de voz, inteligencia artificial y sistemas expertos, etc. (p.51)

De forma similar para Cohen y Hunter (2004): “El NLP es el procesamiento, o tratamiento por computadora, del lenguaje natural, es decir, los lenguajes humanos, en oposición a los lenguajes de programación” (p.147). Bird, Klein y Loper (2009) amplían un poco más la definición y “toma el NLP en un sentido amplio para cubrir cualquier tipo de manipulación informática del lenguaje natural. En un extremo, podría ser tan simple como contar frecuencias de palabras para comparar diferentes estilos de escritura. En

el otro extremo, el NLP implica "comprender" las expresiones humanas completas, al menos hasta el punto de poder darles respuestas útiles" (párr.1). Para este trabajo vamos a definir el NLP como: el conjunto de herramientas informáticas que permiten analizar el lenguaje humano con el fin de darle un sentido a lo que está escrito.

Fake News

Para la definición de FN vamos a usar la planteada por Allcott y Gentzkow (2017) quienes postulan que "definimos "noticias falsas" como artículos de noticias que son intencional y verificablemente falsos, y que pueden confundir a los lectores" (p.213). Rubin, Chen y Conroy (2015) nos indican que para poder detectar las FN, debemos hacer una distinción entre las diferentes clases de FN que existen: a) fabricaciones serias (descubiertas en los medios de comunicación principales o participantes, prensa amarilla o tabloides); b) engaños a gran escala; c) falsificaciones humorísticas (sátira de noticias, parodia, programas de juegos)" (p.83).

Marco Normativo Colombiano

Respecto al marco normativo, el Congreso de la Republica dicta que: 'los concesionarios de los noticieros y los espacios de opinión en televisión, durante la campaña electoral, deberán garantizar el pluralismo, el equilibrio informativo y la **imparcialidad**' (Ley 130,1994, art.27), aunque la Ley 130 no menciona explícitamente a las redes sociales, se entiende que estas deben estar sujetas a los mismos principios.

Efectividad de las FN

Casos relevantes de FN en Colombia

El Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (2019), estima que 6 millones de colombianos utilizan la red social Twitter, la cual se ha convertido en el medio de comunicación preferido por los medios de comunicación y las campañas políticas entre otras. Un claro ejemplo de la fuerza de Twitter en la opinión pública colombiana, se evidenció en el Plebiscito sobre los acuerdos de paz en el 2016 , Rincón (2016) nos señala que dicha campaña contó con imágenes alteradas de los principales actores políticos del país que los mostraba haciendo campaña a el espectro político que no representan, además de un sinnúmero de argumentos como “la «entrega del país a las FARC», el «peligro del socialismo del siglo XXI» y la «ideología de género» pesaron más que la estadística y la lógica común. El poder de las redes se muestra en la entrevista de Ramírez Prado (2016) a Juan Carlos Velez (jefe de la campaña del No), donde afirma que:

Descubrimos el poder viral de las redes sociales. Por ejemplo, en una visita a Apartadó, Antioquia, un concejal me pasó una imagen de Santos y ‘Timochenko’ con un mensaje de por qué se le iba a dar dinero a los guerrilleros si el país estaba en la olla. Yo la publiqué en mi Facebook y al sábado pasado tenía 130.000 compartidos con un alcance de seis millones de personas.

Hicimos una etapa inicial de reactivar toda la estructura del Centro Democrático en las regiones repartiendo volantes en las ciudades. Unos estrategas de Panamá y Brasil nos dijeron que la estrategia era dejar de explicar los acuerdos

para centrar el mensaje en la indignación. En emisoras de estratos medios y altos nos basamos en la no impunidad, la elegibilidad y la reforma tributaria, mientras en las emisoras de estratos bajos nos enfocamos en subsidios. En cuanto al segmento en cada región utilizamos sus respectivos acentos. En la Costa individualizamos el mensaje de que nos íbamos a convertir en Venezuela. Y aquí el No ganó sin pagar un peso. En ocho municipios del Cauca pasamos propaganda por radio la noche del sábado centrada en víctimas. (párr.1)

Este caso es un ejemplo perfecto del poder de las FN en el ámbito nacional y ayuda a comprender el poder que tiene la desinformación en momentos tan críticos como la aprobación de un proceso de paz para todo un país.

Casos relevantes de FN en el mundo

Las FN tienen la capacidad de invadir cualquier medio de comunicación, durante las elecciones a la presidencia de los EEUU en el 2016, cadenas de noticias como la WTOE 5 News, difundieron noticias donde el candidato Donald Trump recibía un supuesto apoyo del Papa Francisco, hecho que generó mucho debate dado a que precisamente el Papa Francisco tiene posiciones contrarias a las de Donald Trump (Caracol Radio, 2016).

La lucha contra las FN

El crecimiento exponencial de las FN, comenzó a llamar la atención de la comunidad académica, lo que conllevó a la realización del primer Fake News Challenge en el 2017, que contó con 50 grupos de participantes, y desarrollaron diferentes algoritmos con el fin de clasificar diferentes noticias. Además, se ha visto el crecimiento de

páginas como *PolitiFact.com*, quienes se enfocan en mirar declaraciones específicas hechas por los políticos y calificarlas por su precisión. PolitiFact está conformado por periodistas que se encargan de realizar las verificaciones (Drobnic Holan, 2018) . En Colombia también existen iniciativas como Colombiacheck.com con el mismo enfoque de PolitiFact.

El problema de estos modelos de verificación de noticias, radica en que la tasa con la que son generadas las FN, sobrepasan la capacidad de los grupos investigativos en determinar el grado de precisión de las noticias. De ahí surge la necesidad de crear algoritmos que den la mano en la mitigación esta problemática.

Además de la dificultad de clasificar las FN , existen maravillosos proyectos duales como Grover (visitar <https://grover.allenai.org/>) , que hace uso de la Inteligencia Artificial para generar una FN que pareciera estar escrita por un escritor de primer nivel. Para generar la FN se parte de un simple encabezado y el programa se encarga de lo demás (esto es conocido como user-generated FN), de forma paralela Grove también permite clasificar con un alto grado de precisión , si la noticia es escrita por medio de Inteligencia Artificial o no, como afirma Zellers et al (2019) “de forma contra-intuitiva, la mejor defensa contra Grover resulta ser Grover, con una precisión del 92%, lo que demuestra la importancia de la liberación pública de generadores potentes.”(p.1)

Cabe resaltar que en las políticas de Twitter (2019), está prohibido promover las FN, y la preocupación que estas generaban para sus usuarios, los llevo a adquirir en junio del 2019 la start-up FABULA, que cuenta con un modelo de Inteligencia Artificial capaz de identificar en pocas horas el 93% de las FN que se propagan por esta red social (Moon, 2019).

Proceso de clasificación de las FN

Para lograr identificar una FN, Atodiresei, Tănăselea y Iftene (2018) nos plantean la arquitectura general para lograr este propósito, especialmente aplicado para la red social Twitter (ver Ilustración 1), esta arquitectura se divide en los siguientes pasos : primero se hace el proceso de web-scraping , que es una técnica utilizada mediante programas de software para extraer información de sitios web mediante el uso de un extractor de Tweets (o crawler), segundo se crea la base de datos ,la cual debe incluir tanto fuentes confiables como de FN, y por ultimo un módulo de proceso NLP .

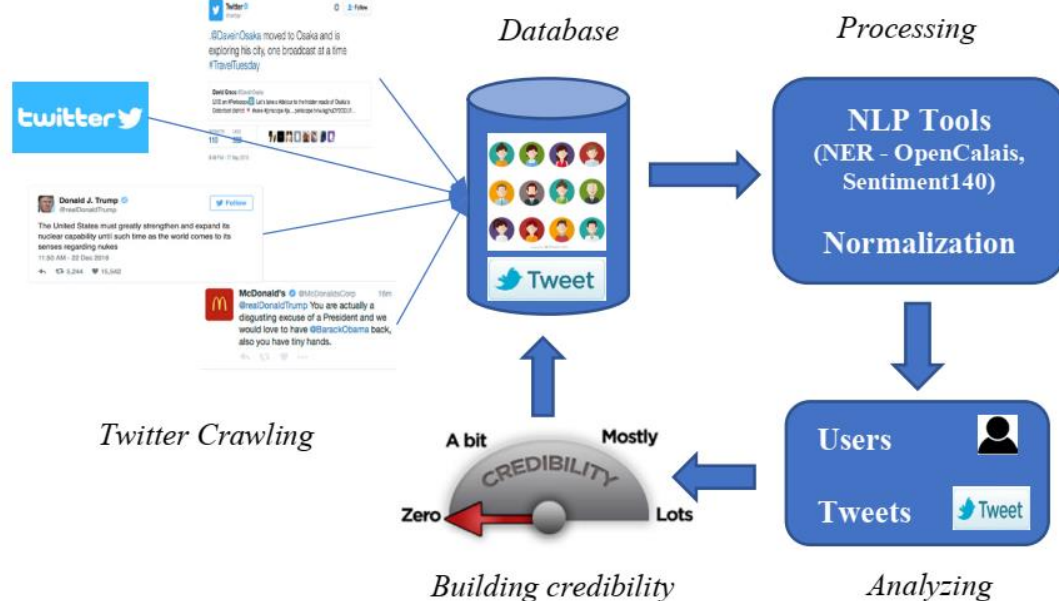


Ilustración 1 Proceso de clasificación de las FN , tomado de Atodiresei, Tănăselea y Iftene (2018)

Respecto al proceso de NLP, una técnica relevante es la de text-mining (minado de texto), según Kumar y Paul (2016), esta técnica es aplicable con el software estadístico R , y para lograr clasificar un texto es necesario seguir los siguientes pasos : el de

tokenización , el cual es proceso de dividir una secuencia de texto, una secuencia de caracteres o una unidad de documento definida, en frases, palabras, símbolos u otros elementos significativos llamados tokens ; de segundo se hace un proceso de limpieza o normalización , donde se remueven signos no útiles , números ,y se coloca todo el texto en minúscula ; tercero se hace el proceso de stemming que consiste en reducir las palabras a su raíz (por ejemplo las palabras : consultar, consultaría, consultado , se entenderá como la palabra consulta) y por último se aplica un algoritmo de clasificación , que permita mirar la probabilidad que el texto sea una FN.

Algoritmos Clasificadores

Después de que el texto está listo para que el software estadístico pueda aplicar los algoritmos, Kumar y Paul (2016) nos plantean aquellos que son útiles para realizar la clasificación de texto como lo son : el Naive Bayes, la Maquina de Soporte Vectorial (en adelante SVM¹) , los Arboles de decisión y por último el de Regresión Logística.

Naive Bayes: Para definir este algoritmo seguimos la definición propuesta por González-Meneses, Pedroza-Méndez, López-Briones, Pérez-Corona y Ramírez-Cruz (2011):

Uno de los métodos supervisados que estiman probabilidades a posteriori es el algoritmo Naive Bayes. Este clasificador asume, para una muestra x , que sus atributos x_1, x_2, \dots, x_n presentan una independencia condicional dado el valor de la clase, por lo que la probabilidad condicional puede expresarse como el

¹ Support Vector Machine en ingles

producto de funciones de probabilidad condicional de cada atributo por separado
(p.16)

En términos simples, un clasificador de Naive Bayes asume que la presencia o ausencia de una característica particular no está relacionada con la presencia o ausencia de cualquier otra característica. Por ejemplo, un animal puede ser considerado como un perro si tiene cola y cuatro patas, el Naive Bayes considera que cada una de estas características contribuye de manera independiente a la probabilidad de que el animal sea un perro, independiente de la presencia o ausencia de las otras características.

SVM: Para definir este algoritmo seguimos la definición propuesta por Betancourt (2005):

Una SVM aprende la superficie decisión de dos clases distintas de los puntos de entrada. Como un clasificador de una sola clase, la descripción dada por los datos de los vectores de soporte es capaz de formar una frontera de decisión alrededor del dominio de los datos de aprendizaje con muy poco o ningún conocimiento de los datos fuera de esta frontera. Los datos son mapeados por medio de un kernel Gaussiano u otro tipo de kernel a un espacio de características en un espacio dimensional más alto, donde se busca la máxima separación entre clases. Esta función de frontera, cuando es traída de regreso al espacio de entrada, puede separar los datos en todas las clases distintas, cada una formando un agrupamiento (p.1)

En términos simples (como podemos apreciar en la Ilustración 2), el SVM intenta trazar un hiper-plano que logre clasificar diferentes datos dado un grado de similitud.

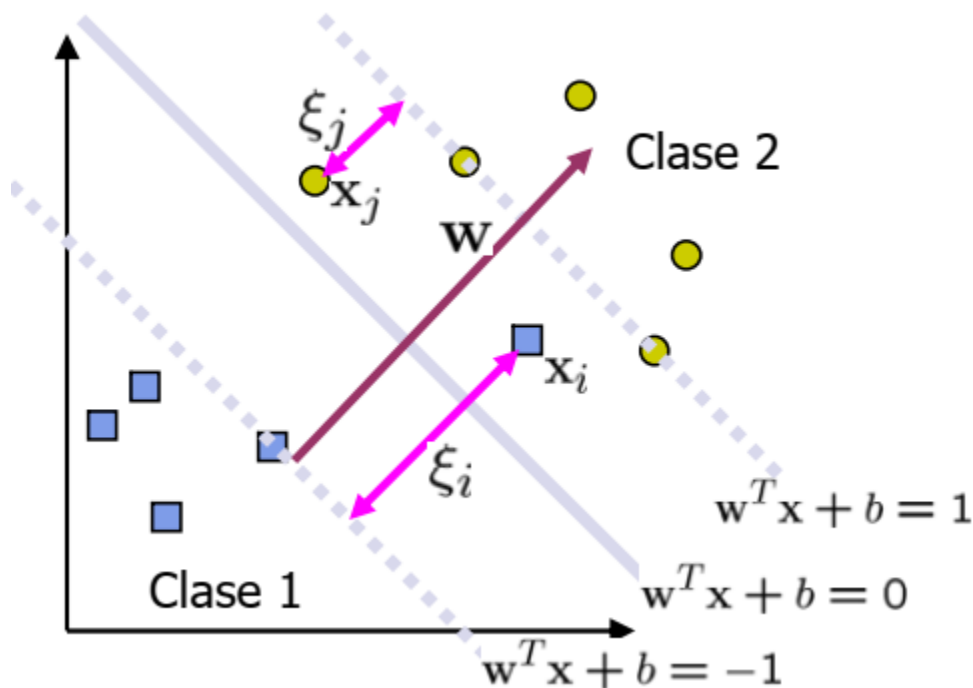


Ilustración 2 Modelo de SVM, tomado de Betancourt (2005)

Arboles de decisión: Para definir este algoritmo seguimos la definición propuesta por Berlanga (2013):

Un árbol de decisión es una forma gráfica y analítica de representar todos los eventos (sucesos) que pueden surgir a partir de una decisión asumida en cierto momento. Nos ayudan a tomar la decisión más “acertada”, desde un punto de vista probabilístico, ante un abanico de posibles decisiones. Estos árboles permiten examinar los resultados y determinar visualmente cómo fluye el modelo. Los resultados visuales ayudan a buscar subgrupos específicos y

relaciones que tal vez no encontraríamos con estadísticos más tradicionales
(p.63)

En términos simples (como podemos apreciar en la Ilustración 3), los arboles de decisión se basan en diagramas de flujo los cuales clasifican o pronostican valores de una variable dependiente a partir de las probabilidades de las variables independientes.

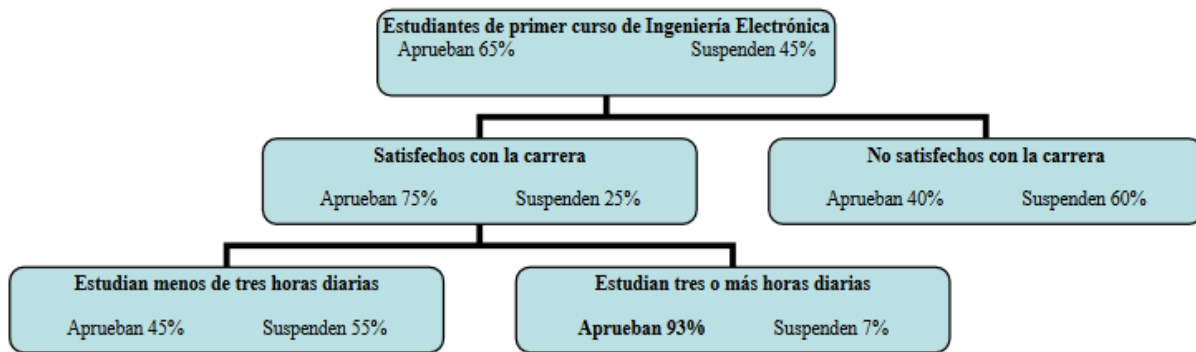


Ilustración 3 Árbol de clasificación del rendimiento académico , tomado de Berlanga (2013)

Regresión Logística (Logit): Para definir este algoritmo seguimos la definición propuesta por Alderete (2006):

La regresión logística, al igual que otras técnicas estadísticas multivariadas, da la posibilidad de evaluar la influencia de cada una de las variables independientes sobre la variable dependiente o de respuesta y controlar el efecto del resto. Tendremos, por tanto, una variable dependiente, llamémosla Y, que puede ser dicotómica y una o más variables independientes, llamémoslas X, que pueden ser de cualquier naturaleza, cualitativas o cuantitativas. Si la variable Y es dicotómica, podrá tomar el valor "0" si el hecho no ocurre y "1" si el hecho ocurre. Este proceso es denominado binomial ya que sólo tiene dos

posibles resultados, siendo la probabilidad de cada uno de ellos constante en una serie de repeticiones. (p.53)

En términos simples la regresión logística nos pronostica la probabilidad de que un evento tenga “éxito” o no, con base a las variables dependientes.

Caso

Como planteamos en la justificación del trabajo, que, a pesar de los avances en la detección de las FN, la mayoría se han realizado respecto al idioma inglés, y para el español no existen suficientes herramientas, ni bases de datos que faciliten la labor. Surge la motivación para hacer uso del NLP, y aplicarlo a los tweets de los principales candidatos a la alcaldía de Cali: Alejandro Eder, Jorge Iván Ospina, Roberto Ortiz, para mirar el grado de veracidad de los tweets que publican, se recogieron los tweets de los candidatos desde el 1 de enero de 2019 hasta el 1 de julio de 2019².

Siguiendo la arquitectura planteada por Atodiresei, Tănăselea y Iftene (2018), se utilizó el programa WebScrapper (disponible en <https://webscraper.io>), para capturar 911 tweets (de 18 de junio de 2018 al 17 de abril de 2019) del usuario de Twitter @ActualidadPanamericana, un portal dedicado a generar noticias falsas de tipo sátira (ante las escasas de bases de datos en español que contuviesen FN, se convirtió en el punto de referencia de qué es una FN), estos tweets tendrán la variable de respuesta categórica de tipo binaria: “falsa”, para las noticias verdaderas se usaron de igual forma 678 tweets de @ELTIEMPO, un portal con muchos años de trayectoria y el cual

² Los datos al igual que el código en R están disponibles en <https://mega.nz/#F!rgU13KRB!F-UYiSCHqv-zSOGsqXribg>

se caracteriza por la veracidad de la información que transmite , estos tweets tendrán la variable de respuesta categórica tipo binaria : “verdadera” y por último los tweets de los 3 candidatos. Una vez creada la base de datos, se procede a aplicar las técnicas de limpieza (tokenización, normalización y stemming). El paso final es la de entrenar el modelo con la aplicación de los algoritmos: Random Forest, Naive Bayes, Neural Network y Arboles de Decisión. Una vez el software termina de entrenar los modelos, se proceden a aplicar en los tweets de los aspirantes a la alcaldía.

Se puede esperar un alto porcentaje de FN debido a la “guerra sucia” que actualmente se presenta entre los candidatos, con diferentes series de acusaciones entre ellos, como lo expone la Revista Semana (2019):

Estas elecciones han sido una de las más polarizadas en la historia de la ciudad con escenarios como la huelga de hambre del candidato Ospina (luego del llamado a juicio) que señaló a medios de comunicación y a su contendor de guerra sucia; o una orden de rectificación de un juez contra Ortiz por las acusaciones que relacionaban a Ospina con el Cartel de la Toga, y las pancartas que salieron en la ciudad con señalamientos e insultos a los candidatos. (párr.1)

Métricas

Como plantean González-Meneses, Pedroza-Méndez, López-Briones, Pérez-Corona y Ramírez-Cruz (2011), es necesario tener unas métricas para mirar cual algoritmo se ajusta mejor a los datos y logra hacer una mejor clasificación, estas son:

1. **TP (True positives)**: ejemplos positivos clasificados correctamente. Por ejemplo, supongamos que tenemos un clasificador binario de imágenes de gatitos que nos indica si la imagen es de un gatito o no. El TP es el número de imágenes de gatitos clasificadas positivamente.
2. **FP (False positives)**: ejemplos negativos clasificados incorrectamente como positivos. En el ejemplo de los gatitos, sería el número de imágenes sin gatitos que son clasificadas positivamente como imagen de gatito. Se puede considerar este valor como aquellos fallos cometidos debido a clasificaciones muy optimistas (muy sesgadas hacia el caso positivo).
3. **TN (True negatives)**: ejemplos negativos clasificados correctamente. De nuevo, en el ejemplo de los gatitos, sería el número de imágenes sin gatitos que son clasificadas negativamente (imagen sin gatito).
4. **FN (False negatives)**: ejemplos positivos clasificados incorrectamente como negativos. En nuestro ejemplo, es el número de imágenes de gatitos que son clasificadas como negativamente (imagen sin gatito). Podemos pensar en este indicador como el número de ejemplos correctos que escapan a nuestro "radar".
5. **Precisión**: son los verdaderos positivos divididos por los positivos pronosticados (los verdaderos positivos más los falsos positivos).
6. **Sensibilidad (o Recall)**: es el número de predicciones positivas correctas dividido por el número total de positivos
7. **Puntaje F1**: es una puntuación única del modelo, y se especifica como la multiplicación por dos de la precisión y sensibilidad, dividido a la precisión más la sensibilidad.

Para este trabajo se utilizará el Puntaje F1, para determinar cuál es el mejor modelo que se ajusta a los datos.

Selección de algoritmo

Ya que elegimos la variable Puntaje F1 como aquella que nos permite elegir el algoritmo que mejor se ajusta a los datos, según la Tabla 1, el algoritmo que se selecciona para clasificar los tweets de los candidatos es la Regresión Logística con un puntaje F1 de 74.37%.

	Arboles de Decisión		SVM		Naive Bayes		Logit	
	Positivos Reales	Negativos Reales	Positivos Reales	Negativos Reales	Positivos Reales	Negativos Reales	Positiv os Reales	Negativos Reales
Positivos Predichos	43	18.5	43	24.1	44.5	22.1	53.1	32.8
Negativos Predichos	14.4	24.2	9.7	19.2	13.4	20	3.8	10.3
	Arboles de Decisión	SVM	Naive Bayes	Logit				
Precisión	69.92%	64.08%	66.82%	61.82%				
Sensibilidad	74.91%	81.59%	76.86%	93.32%				
F1	72.33%	71.79%	71.49%	74.37%				

Tabla 1 Resumen resultados de los modelos , Elaboración Propia

Resultados

Luego de escoger la Regresión Logística como el algoritmo para clasificar los Tweets, el software nos arroja los resultados que se pueden apreciar en la Tabla 2.

Alejandro Eder		Roberto Ortiz		Jorge Iván Ospina	
Tweets Verdaderos	Tweets Falsos	Tweets Verdaderos	Tweets Falsos	Tweets Verdaderos	Tweets Falsos
92	265	55	323	55	424

Tabla 2 Clasificación de tweets según la Regresión Logística, Elaboración Propia

Según el modelo, el candidato que tiene mayor porcentaje de tweets verdaderos es Alejandro Eder, mientras que el de mayor porcentaje de tweets falsos es Jorge Iván Ospina seguido de Roberto Ortiz, lo cual puede mostrar una correlación con la “guerra sucia” entre los dos candidatos.

	Alejandro Eder	Roberto Ortiz	Jorge Iván Ospina
% de verdaderos	25.77%	14.55%	11.48%
% de falsos	74.23%	85.45%	88.52%
Total de tweets	357	378	479

Tabla 3 Porcentajes de tweets verdaderos y falsos, Elaboración Propia

Conclusión

Aunque se consiguieron buenos niveles de precisión en los modelos, falta mucho desarrollo que permita tener bases de datos de buena calidad en español para generar mayores avances, esto queda en evidencia cuando los desarrollos para el idioma inglés pueden alcanzar niveles de detección de FN superiores al 90%. Todo esto se debe hacer con el fin de brindar una herramienta a la población en general que le permita

tomar decisiones más acertadas y no caer en mano de pequeños grupos que solo persiguen intereses particulares.

Bibliografía

- Alderete, A. M. (2006). Fundamentos del análisis de regresión logística en la investigación psicológica. *Revista evaluar*, 6(1).
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), 211–36.
- Atodiresei, C.-S., Tănăselea, A., & Iftene, A. (2018). Identifying Fake News and Fake Users on Twitter. *Procedia Computer Science*, 126, 451–461.
<https://doi.org/10.1016/j.procs.2018.07.279>
- Berlanga, V., Rubio Hurtado, M. J., & Vilà Baños, R. (2013). Cómo aplicar árboles de decisión en SPSS. *REIRE. Revista d'Innovació i Recerca en Educació*, 2013, vol. 6, num. 1, p. 65-79.
- Betancourt, G. A. (2005). Las máquinas de soporte vectorial (SVMs). *Scientia et technica*, 1(27).
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. O'Reilly Media, Inc.
- Caracol Radio. (2016). El falso apoyo de Francisco a Trump también fue noticia en 2016. Recuperado el 10 de octubre de 2019, de Caracol Radio website:
https://caracol.com.co/radio/2016/12/28/internacional/1482936389_870480.html
- Chowdhury, G. G. (2003). Natural language processing. *Annual review of information science and technology*, 37(1), 51–89.

Cohen, K. B., & Hunter, L. (2004). Natural language processing and systems biology. En *Artificial intelligence methods and tools for systems biology* (pp. 147–173). Springer.

Congreso de Colombia. (23 de marzo de 1994) Artículo 27 [Título I]. Ley de Partidos Políticos. [Ley 130 de 1994]

Drobnic Holan, A. (2018). The Principles of the Truth-O-Meter: PolitiFact's methodology for independent fact-checking. Recuperado el 18 de octubre de 2019, de PolitiFact website: <https://www.politifact.com/truth-o-meter/article/2018/feb/12/principles-truth-o-meter-politifacts-methodology-i/#How%20PolitiFact%20Started>

González-Meneses, Y. N., Pedroza-Méndez, B. E., López-Briones, F., Pérez-Corona, C., & Ramírez-Cruz, J. F. (2011). *Implementación del clasificador Naive Bayes para la Acentuación Automática de Palabras Ambiguas del Español*. CiCos.

Kumar, A., & Paul, A. (2016). *Mastering Text Mining with R*. Packt Publishing Ltd.

Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia. (2019). Colombia es uno de los países con más usuarios en redes sociales en la región. Recuperado el 9 de agosto de 2019, de Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia website: <https://mintic.gov.co/portal/inicio/Sala-de-Prensa/Sabia-Ud-que/2713:Colombia-es-uno-de-los-paises-con-mas-usuarios-en-redes-sociales-en-la-region>

Moon, M. (2019). Twitter acquires AI startup to help it fight fake news. Recuperado de Engadget website: <https://www.engadget.com/2019/06/04/twitter-acquires->

fabula-ai-fake-

news/?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_si

g=AQAAALbUboutRsVVYnULK5NC-

b_EddOhqbYlwYPTp8pr1nU4PjjmY1YeAF_n7r-

LKX1IBUNMLxyVYYSSkteGPZ78q6R4dUTJV16e3iRYy-IHvX92boDYi

Ramírez Prado, J. (2016). El No ha sido la campaña más barata y más efectiva de la historia. Recuperado el 9 de agosto de 2019, de Asuntos Legales website: <https://www.asuntoslegales.com.co/actualidad/el-no-ha-sido-la-campana-mas-barata-y-mas-efectiva-de-la-historia-2427891>

Rincón, W. (2016). Redes sociales en el plebiscito: El fenómeno del yo con yo. Recuperado el 9 de agosto de 2019, de Revista Semana website: <https://www.semana.com/tecnologia/articulo/plebiscito-por-la-paz-la-burbuja-de-las-redes-sociales/497814>

Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: Three types of fakes. *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*, 83. American Society for Information Science.

Semana. (2019). Final de infarto en Cali: ¿Jorge Iván Ospina o Roberto Ortiz? Recuperado el 25 de octubre de 2019, de Revista Semana website: <https://www.semana.com/nacion/articulo/final-de-infarto-en-cali-jorge-ivan-ospina-o-roberto-ortiz/637601>

Twitter. (2019). The Twitter Rules. Recuperado el 9 de agosto de 2019, de Twitter website: <https://help.twitter.com/en/rules-and-policies/twitter-rules>

Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y.

(2019). Defending Against Neural Fake News. *arXiv e-prints*, arXiv:1905.12616.