DESDE 2013
https://repository.uaeh.edu.mx/revistas/index.php/icbi/issue/archive
Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI
ISSN: 2007-6363

Publicación Semestral Pädi Vol. 10 No. Especial 3 (2022) 94–102

Clasificación de opiniones provenientes de la web considerando el uso de información de tipo personal
Opinions's clasification from the web considering the using of personal information

C. A. Cázares-Pérez [a], A. Franco-Árcega [a,*], R. M. Ortega-Mendoza [b], H. Castillejos-Fernández [a], J. Suárez-Cansino [a], V. López-Morales [a]

[a] *Área Académica de Computación y Electrónica, Universidad Autónoma del Estado de Hidalgo, 42184, Pachuca, Hidalgo, México.*
[b] *Investigación y Posgrado, Universidad Politécnica de Tulancingo, 43629, Tulancingo de Bravo, Hidalgo, México.*

Resumen

Los engaños han estado en la sociedad desde hace tiempo y han llegado a converger más, debido a las capacidades tecnológicas que actualmente tienen los dispositivos y el uso generalizado de la web que hacemos. Esta situación ha ocasionado muchos efectos negativos, como la *desinformación* entre usuarios, afectando la perspectiva o pensamiento ante actividades cotidianas. Recientemente, se han realizado diversas investigaciones se han propuesto para detectar engaños en publicaciones en línea. Este trabajo propone contemplar un elemento importante para apoyar la detección de engaños: explorar los pronombres personales en las publicaciones (u opiniones), ya que su uso tiende a mostrar honestidad entre las personas. La metodología propuesta separa las opiniones en oraciones que usan pronombres personales, para analizar su valor en la tarea. Se utilizan dos conjuntos de opiniones para evaluar la propuesta: *Op Spam* y *Amazon*. Los resultados son alentadores, dado que muestran que los pronombres personales son relevantes para identificar esta tarea.

Palabras Clave: Minería de datos, Detección de engaños, Procesamiento de lenguaje natural

Abstract

Deception has been in society since time ago, converging more nowadays due to the current technological capabilities of the devices and the widespread usage of the web. This situation has caused many adverse effects, such as misinformation between users, affecting society's perspective and thinking about daily activities. Recently, different researches have been proposed to detect the deception in online opinions. This work proposes an important element to face the automatic deception detection task: exploring personal pronouns in opinions (or reviews) because their use is related to honesty among the persons. The proposed methodology divides the opinions into sentences using personal pronouns to analyze their value for the task. Two benchmark datasets are used to evaluate the proposal and the deception detection degree: *Op Spam* and *Amazon*. The results are encouraged; they show that information stated in sentences with personal pronouns is relevant for the task.

Keywords: Data mining, Detection deception, Natural language processing

1. Introducción

Anteriormente, los instrumentos de comunicación eran diversos, y tenían sólo como objetivo transmitir información, tales cómo los teléfonos locales, el correo, telegramas, vídeo conferencias, etc. En la actualidad, todas esas vías de comunicación se pueden encontrar juntas, por ejemplo en los teléfonos inteligentes. Estos dispositivos se han vuelto indispensables y comunes, debido a que ofrecen la capacidad de procesamiento de las computadoras que están en el hogar, además de la capacidad de movilidad y la accesibilidad para adquirir alguno. Por estas razones, han aparecido muchas herramientas de comunicación que puedan ser usadas en estos dispositivos, como las *Redes Sociales*, las cuales son de acceso gratuito y permiten generar círculos sociales, agrupando personas con características en común, cómo parentesco, trabajo, pasatiempos, gustos, etc. Debido a las herramientas de comunicación que están apareciendo,

*Autor para correspondencia: afranco@uaeh.edu.mx
Correo electrónico: carlos.cazares.89@gmail.com(Carlos Andrés Cázares-Pérez), afranco@uaeh.edu.mx(Anilú Franco-Árcega), rosa.ortega@upt.edu.mx(Rosa María Ortega-Mendoza), heydy_castillejos@uaeh.edu.mx(Heydy Castillejos-Fernández), jsuarez@uaeh.edu.mx(Joel Suárez-Cansino), virgilio@uaeh.edu.mx(Virgilio López-Morales)

Historial del manuscrito: recibido el 14/04/2022, última versión-revisada recibida el 13/07/2022, aceptado el 02/08/2022, publicado el 31/08/2022. **DOI:** https://doi.org/10.29057/icbi.v10iEspecial3.8998

cualquier persona puede tener acceso a la información y conocimiento, tales como noticias, artículos, recomendaciones, etc. Sin embargo, a causa de que hay muchas facilidades en la consulta de información, surgen ciertos enigmas: ¿La información que se divulga es fidedigna? y ¿Quiénes son los encargados de determinarlo? Al principio se pensaría que la información que se encuentra es 100 % confiable, sin embargo en distintos medios se ha encontrado que lo que se comparte es engañoso (Ahmed et al., 2018). Un ejemplo muy sonado fue el esparcimiento de *Noticias Falsas* o *Fake News* durante las elecciones presidenciales de Estados Unidos de Norte América del año 2016, donde las principales vías de comunicación fueron las redes sociales de Facebook y Twitter (Krishnamurthy et al., 2018), pero este tipo de situaciones no sólo logra afectar en el aspecto política, si no también puede llegar a perjudicar la integridad de una persona, como el caso de un suicidio generado por divulgar información falsa (Fuller et al., 2011). El engaño y la mentira a menudo son confundidos como sinónimos, no obstante estos están relacionados. La mentira se puede definir como un *estado falso* (Holland and Quinn, 1995), mientras que el engaño *es el acto que trata de fomentar a alguien en creer algo que es falso* (Zuckerman et al., 1981; Rill García et al., 2019). Un engaño ocurre cuando el que comunica tiende a manipular la información que contienen sus mensajes, con el fin de transmitir un mensaje apartado de la verdad (Buller and Burgoon, 1996). Para resolver el caso de identificación de engaños, se han realizado varios estudios, algunos desde el enfoque psicosocial, donde se ha tratado de determinar si un engaño está presente o no en algún texto (Fuller et al., 2011; G., 2007), y otros desde un enfoque computacional, donde se estudia el contenido de publicaciones en redes sociales (Altunbey Ozbay and Alatas, 2020; Cabrejas et al., 2019; Mbaziira and Jones, 2016), algunos más considerando un enfoque lingüístico, donde se extrae el contenido y se analiza con ciertos patrones que pueden tener los mensajes engañosos (Song et al., 2005; Sánchez Junquera et al., 2017; Rill García et al., 2019). Este trabajo tiene el objetivo de proponer una metodología que permita llevar a cabo la detección de engaños. Esta metodología permite integrar conocimiento de dos áreas importantes. Por un lado, la *Minería de Datos (MD)* la cual permite analizar grandes cantidades de datos y que ha sido usada para detectar actividades delictivas cómo: violaciones de tráfico, crimen sexual, robo, fraude, daños a inmuebles, narcóticos, crímenes de violencia y ciber delincuencia (Fuller et al., 2011; Wanumen Silva, 2010; Ruiz, 2006). Por otro lado, el *Procesamiento de Lenguaje Natural (PLN)*, permite realizar un análisis de textos para identificar características léxicas y sintácticas. tal como dice Gelbukh (Gelbukh, 2010), el *PLN* es una disciplina que estudia la habilidad que tiene la máquina para procesar la información comunicada, no solo en letras, sino también en los sonidos del lenguaje. Adicional a lo anterior, estudios recientes demuestran que el uso de pronombres personales en tareas relacionadas al *PLN* obtiene buenos resultados, tal es el caso de la tarea de Identificación del perfil de autores y en detección de depresión (Ortega Mendoza et al., 2007, 2018, 2022). En estos trabajos se ha observado que tomar en cuenta estos elementos del lenguaje permite obtener buenos resultados de clasificación, dado que los pronombres personales en los textos suelen referir a la personalidad de quien escribe.

2. Metodología

Se propone una metodología que permite la identificación de publicaciones engañosas en medios sociales mediante una perspectiva de clasificación de textos, donde se tienen dos clases: positiva y negativa, representando un mensaje con engaño y sin engaño, respectivamente. El proceso general que se lleva a cabo se muestra en la Figura 1. Se inicia con la lectura de alguna colección de documentos. Estos mensajes reciben un preprocesamiento, para crear una representación estructurada que alimentará algoritmos de aprendizaje automático. Este procesamiento consiste en la separación de los términos que componen a los mensajes, para proceder a la construcción del vocabulario de datos. Una vez que se haya realizado esta identificación, cada término en cada mensaje será representado por una estrategia de pesado de términos. Dado que en este tipo de problemáticas existe siempre una alta dimensionalidad, por la gran cantidad de elementos que se pueden encontrar en los mensajes, el siguiente paso es realizar una selección de términos. Por esta razón, es necesario llevar a cabo esta selección. Posterior a este paso se aplicarán diversos algoritmos de clasificación supervisada, los cuales permitirán crear modelos de clasificación y por lo tanto, identificar el tipo de mensaje que se introduzca en el modelo.

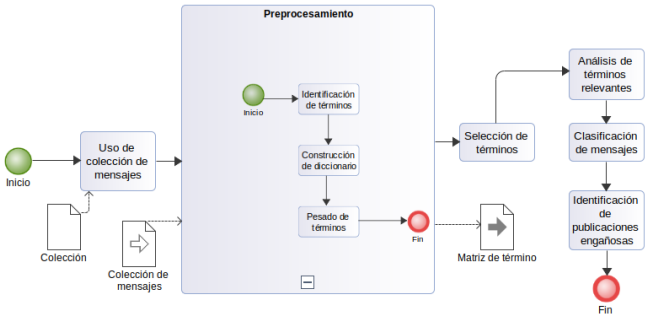

Figura 1: Proceso general para la clasificación.

3. Construcción del vocabulario

Para realizar la creación del vocabulario de términos que permitirán representar a los mensajes a procesar, se ha planteado un preprocesamiento que incluye cinco fases: (i) tokenización y transformación de palabras a minúsculas, (ii) retiro de símbolos y puntuaciones, (iii) retiro de palabras repetidas, (iv) retiro de stopwords y (iv) aplicación del proceso de stemming. Considerando estos procesos y para poder analizar la identificación de publicaciones engañosas, se propone la creación de diferentes vocabularios de datos. Estos fueron creados considerando las siguientes especificaciones:

- Vocabulario *ALL*: Se usan todos los términos encontrados en la identificación de palabras usadas en las opiniones.
- Vocabulario *STW*: Se retiran términos considerados como palabras vacías (conocidos como stopwords) las cuales no tienen un significado importante en el contexto de la opinión.

- Vocabulario *STE*: Se aplica el proceso de stemming, en donde se considera la reducción de términos que compartes raíz gramatical.
- Vocabulario *STESTW*: Se retiran los stopwords y se aplica el proceso de stemming.

4. Uso de pronombres personales

El uso de pronombres personales, puede ser una caraterística que pueda definir los engaños escritos, ya que se omiten por que el autor no quiere relacionarse (Zuckerman et al., 1981), sin embargo se han encontrado el uso de pronombres para hacer más creíbles (Fuller et al., 2014), ya que el uso de pronombres personales tiende hacia la honestidad (Newman et al., 2003). Para determinar si el uso de pronombres personales ayuda en la tarea de detección de publicaciones engañosas, se propone identificar la aparición de este tipo de términos en las publicaciones a procesar. Para este trabajo, se tomarán en cuentan tanto los pronombres personales en su forma singular como en su forma plural, aquellos considerados para esta revisión son:

- Singulares: *"I"*, *"me"*, *"mine"*, *"my"*, *"myself"*, *"im"* y *"I'm"*
- Plurales: *"we"*, *"us"*, *"our"*, *"ours"*, *"ourselves"*

Se propone que la identificación de estos términos se realice de manera independiente. En cada opinión del conjunto de datos, se realiza la búsqueda de la existencia de alguno de los pronombres y si se encuentra, la opinión es separada por oraciones. Este proceso implica que la opinión resultará en dos fragmentos de texto. El primero será aquel que contenga las oraciones en donde los pronombres están presentes, y el segundo, aquel que sólo integre oraciones sin pronombres. Un ejemplo de esta separación sería de la siguiente forma:

Resultado:

- Opinión con pronombres: *In January, I was looking for a spot for quiet leisure and relaxation. My choice turned out to be a 100 % hit. I had a great rest!*
- Opinión sin pronombres: *A wonderful neighbourhood. Amazing hotel conditions, perfect cuisine and very nice staff.*

El proceso de separación propuesto se puede visualizar en el diagrama de la Figura 2, donde se puede observar la forma en cómo se obtienen los dos conjuntos de datos diferentes a partir de uno solo, los cuales serán procesados de manera independiente.

5. Conjunto de datos: *Op Spam*y *Amazon*

Se utilizaron dos conjuntos de opiniones para observar el comportamiento de la propuesta realizada en este trabajo, los cuales son mostrados en la Tabla 1. El primer conjunto está constituido por opiniones hacia hoteles y se denomina *Op Spam* (Ott et al., 2013), el cual está compuesto de 800 opiniones falsas (clase F) y 800 opiniones verdaderas (clase V), por lo que se aprecia un balance entre las dos clases. La construcción de este conjunto de datos obtuvo las opiniones falsas a través de la plataforma *Amazon Mechanical Turk*, diseñada por *Amazon*. Para obtener las opiniones verdaderas se empleó un proceso más sencillo, ya que se obtuvieron a través de distintas fuentes, entre las que destacan *Expedia*, *Hotels.com*, *Orbitz*, *Priceline*, *TripAdvisor* y *Yelp*, plataformas cuya función es evaluar por medio de opiniones de distintos usuarios todo lo relacionado al tema turístico. El segundo conjunto de datos utilizado es *Amazon*, el cual fue obtenido de un repositorio público de *Github*[2]. Está construido por opiniones de usuarios hacia los productos que ofrece *Amazon*. El conjunto de datos está compuesto de 21,000 opiniones de productos, de manera que se desconoce la cantidad de opiniones verdaderas y falsas que hay, por ello, se toman aleatoriamente 1,000 opiniones verdaderas y 1,000 opiniones falsas, haciendo un conjunto balanceado.

Tabla 1: Conjuntos de datos empleados en la experimentación.

Conjunto	Opiniones totales	Opiniones Verdaderas	Opiniones Falsas
Op-Spam	1600	800	800
Amazon	2000	1000	1000

6. Configuración

Los conjuntos de datos son procesados por la metodología mencionada. Se construyen los cuatro vocabularios correspondientes (*ALL*, *STW*, *STE*, *STESTW*), para que posteriormente se aplique el pesado de términos, usando el método *TF*.

$$TF = \frac{n_{ij}}{|d_i|}, \tag{1}$$

donde $|d_i|$ es la suma de todos los pesos de los términos en el documento i y n_{ij} es el número de veces que aparece la palabra j en el documento i.

Al obtener las matrices de cada uno de los vocabularios, se les aplicarán distintos métodos de selección de atributos, con la finalidad de reducir la dimensionalidad de los conjuntos, en cuestión de términos. Los selectores que se aplicaron fueron *Ganancia de Información (Info Gain)*, *Proporción de Ganancia de Información (Gain Ratio)* y *Selección de Subconjuntos de Características basada en Correlación (CfsSubsetEval)*. Para realizar la clasificación de las opiniones se aplicaron cuatro clasificadores, empleando la herramienta Weka: *Naive Bayes (NB)*, Máquinas de vectores de soporte usando el algoritmo *Lib Linear (LL)*, *Random Forest (RF)* y *Naive Bayes Multinomial Updateable (NBMU)* utilizando una validación cruzada de 10 pliegues. La métrica reportada y considerada en estos resultados es *F-Measure* o *F-1 score*, la cual involucra las medidas de *Precision* y *Recall*. La primera parte de la sección de experimentación ayudó a detectar cuál es la mejor configuración

[2]https://github.com/aayush210789/Deception-Detection-on-Amazon-reviews-dataset

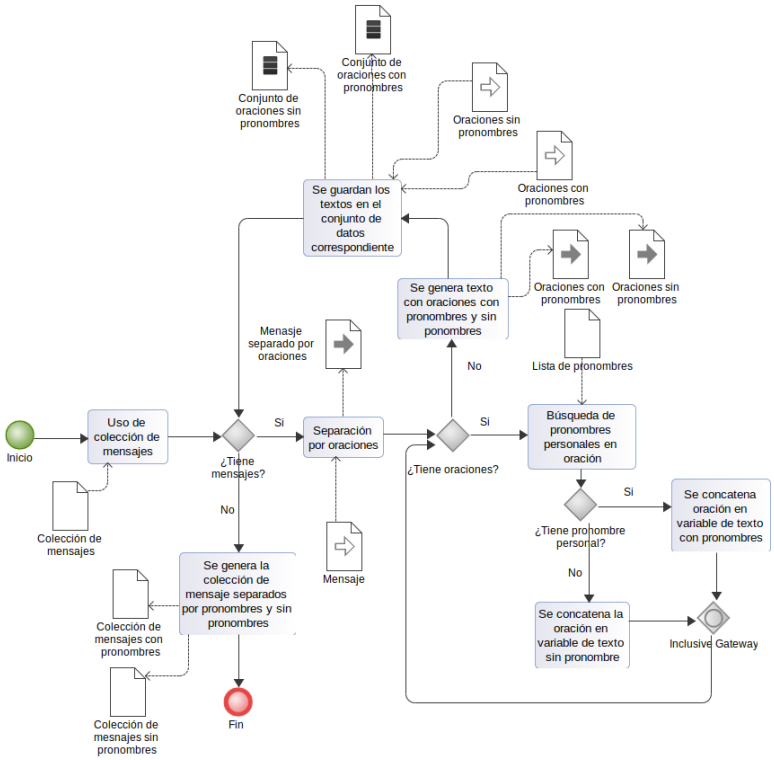

Figura 2: Proceso de búsqueda de pronombres personales por oraciones.

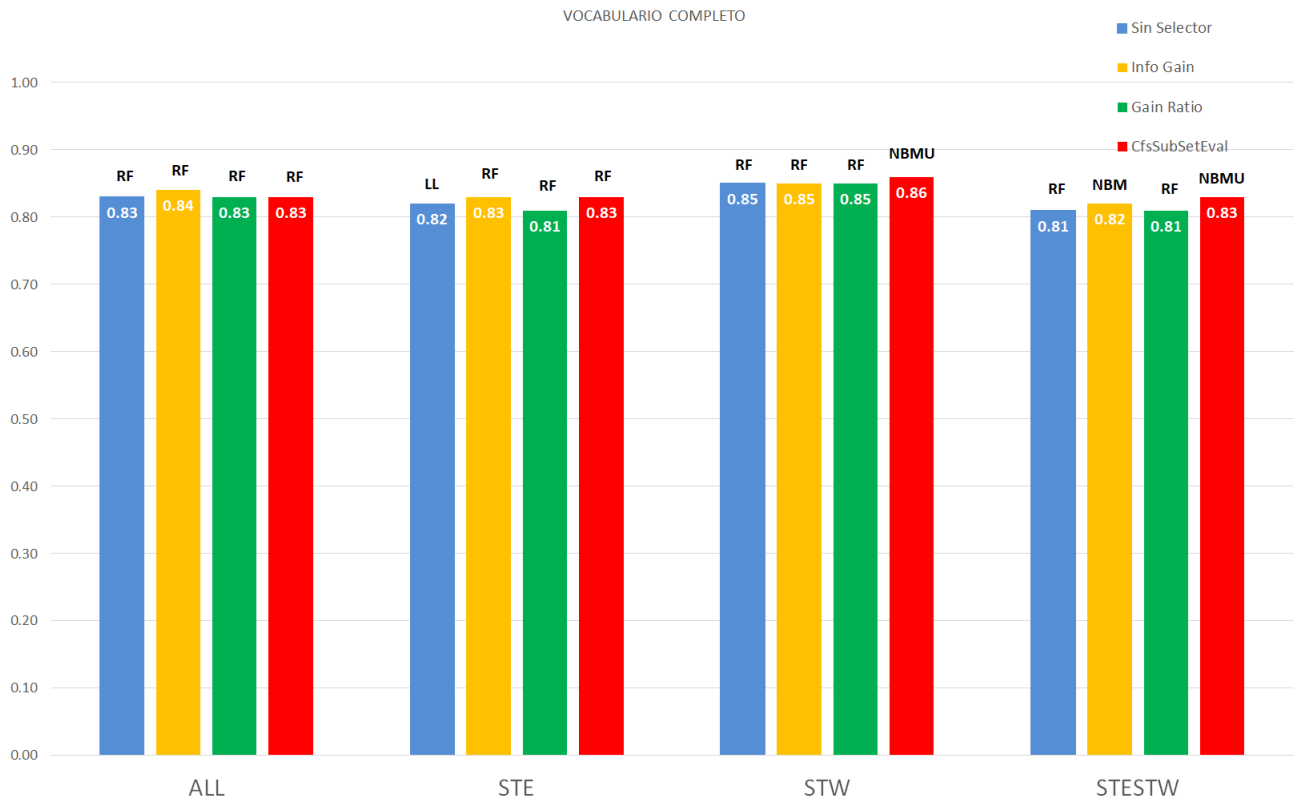

Figura 3: Resultados de cada estrategia con el vocabulario completo para *Op Spam*

selector-clasificador empleada en esta tarea. Por otro lado, se pudo observar que la cantidad de términos que se obtuvieron al generar las matrices de datos fueron variando con respecto a la creación del vocabulario, la aplicación o no de los selectores de atributos y la separación de las opiniones considerando el uso de pronombres personales.

7. Experimentación

Esta sección muestra la experimentación llevada a cabo para observar el comportamiento de la metodología propuesta. Primero se presenta utilizando el vocabulario completo obtenido de los conjuntos de opiniones y posteriormente, se obtienen los resultados considerando la estrategia de separación de opiniones.

7.1. Vocabulario completo

El conjunto *Op Spam* fue procesado y de él se obtuvo el vocabulario completo y posteriormente se le aplicaron los algoritmos de reducción de términos. La Tabla 2 muestra estos resultados, donde se puede observar que existe una importante reducción en el número de términos, ya que se tiene una reducción de hasta 153 veces, con respecto al vocabulario completo y el algoritmo de *Ganancia de Información*. Con *Proporción de Ganancia de Información* se obtiene una reducción de hasta 34 veces más pequeño al conjunto de términos original y con *CfsSubSetEval* se obtiene 127 veces en reducción.

Tabla 2: Reducción de vocabulario completo para *Op Spam*. Sin Selector(SNS), Info Gain (IG), Gain Ratio (GR) y CfsSubsetEval (CSE).

VOCABULARIO	SNS	IG	GR	CSE
ALL	10085	81	313	72
STE	7268	78	267	72
STW	9955	65	288	78
STESTW	7167	69	54	81

Al aplicar los cuatro clasificadores con los cuatro vocabularios con o sin considerar los selectores, se obtienen los resultados de la Figura 3. Al observar los promedios se puede apreciar que el clasificador que tiene mayor presencia en casi todas las configuraciones es *Random Forest*, mostrando que puede ser el más conveniente para usarse en las siguientes etapas.

Para poder decidir cuál fue el mejor resultado entre los vocabularios, se optó por tomar el mejor promedio de cada uno. Esta comparación se muestra en la Tabla 3, donde se eligió la aplicación del selector de atributos *Proporción de Ganancia de Información* usando el clasificador *Random Forest*. Se observó que el mejor resultado en dicha representación fue con el vocabulario *STESTW* arrojando un promedio del 86 % en la clasificación, con una reducción considerable en el tamaño del conjunto de términos. En las tablas de resultados se mostrarán tres resultados diferentes, la métrica de eficiencia para la clase Verdaderos (V), la clase Falsa (F) y la clasificación general de la tarea (G).

Tabla 3: Resultados con mayor promedio de cada vocabulario de la configuración inicial para *Op Spam* con la métrica F-Measure, usando el selector gain ratio y el clasificador random forest.

Vocabulario	Términos	V	F	G
ALL	313	0.85	0.85	0.85
STE	267	0.85	0.86	0.85
STW	288	0.85	0.86	0.85
STESTW	**254**	**0.86**	**0.86**	**0.86**

Con el vocabulario *STESTW* como el más representativo se obtuvieron los términos más representativos, los cuales se muestran a través de una nube de palabras en la Figura 4. Se puede notar que el término mas sobresaliente es “hotel”, esto es por que originalmente es el tema principal de este corpus, además que se muestran otros términos relacionados con la zona o áreas donde se involucran las opiniones cómo “chicago”, “bathroom”, “floor”, etc.. Finalmente, el término “stay”, es un verbo que podría relacionarse con el motivo de estancia dentro de la zona o área.

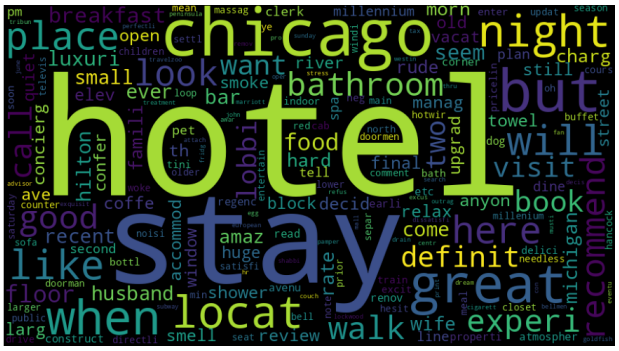

Figura 4: Palabras más representativas con el vocabulario *STESTW* en la configuración inicial de *Op Spam*

7.1.1. Conjunto de datos Amazon

Con el clasificador y selector de atributos ya elegidos, a continuación se implementa en el conjunto de datos *Amazon* esta configuración. La Tabla 4 muestra estos resultados, donde se observa que con el vocabulario *STW* se obtiene el mejor resultado.

Tabla 4: Clasificación con Random Forest y selector Gain Ratio de *Amazon* con la métrica F-Measure.

Vocabulario	V	F	G
ALL	.64	.62	.63
STE	.63	.62	.62
STW	**.69**	**.60**	**.64**
STESTW	.67	.55	.61

Con respecto a los términos más representativos en el vocabulario *STW* se puede observar en la Figura 5 que el que tiene mayor relevancia dentro del conjunto es el término “product”, teniendo mucho énfasis en el dominio al cual pertenece el conjunto de datos, seguido después de “really”, el cual es un adverbio que pretende enfatizar algo verdadero, así que dentro de las opiniones se puede intuir que se trató de calificar el producto de una forma positiva o negativa.

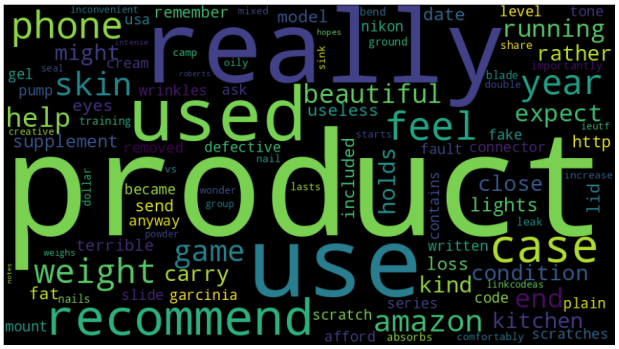

Figura 5: Palabras más representativas con el vocabulario *STW* en la configuración inicial en *Amazon*

7.2. *Uso de pronombres personales del singular*

Tomando los resultados con los vocabularios y la configuración seleccionada, en esta parte de la experimentación se evalúa el comportamiento al considerar el uso de pronombres personales del singular, tomando el proceso de separación mencionado en la Sección 2. Al aplicar el proceso se obtienen dos conjuntos datos, uno que contiene sentencias con pronombres del singular y otro que no los contiene. Al aplicar el proceso, en el conjunto *Op Spam* se obtiene el conjunto de datos que contiene pronombres personales con un total de 637 opiniones verdaderas y 735 opiniones falsas. Como se ve en la Tabla 5, entre las distintas representaciones el que tiene el mejor desempeño en la clasificación es el vocabulario *STW*. Con los otros vocabularios, se observa que al considerar los pronombres personales, se tienen resultados comparables de clasificación. Por su parte, el conjunto que no tiene pronombres personales se conformó de 789 opiniones verdaderas y 780 opiniones falsas. En la Tabla 5 se observa que los resultados con los cuatro vocabularios están equilibrados, llegando a tener aproximadamente el 80 % de resultado en la métrica, sin embargo el que mejor resultado tiene entre los vocabularios es *STESTW*.

Al aplicar el proceso de separación en el conjunto *Amazon*, se obtuvieron 827 opiniones verdaderas y 847 opiniones falsas. En la Tabla 5) se logra observar que hubo una disminución en la clasificación, a diferencia del primer experimento. El mejor resultado fue *STE*, y en los que se retiran los stopwords hubo una disminución menor, ya que como lo menciona Sánchez (Sánchez Junquera et al., 2017), al eliminar los stopwords se puede afectar la clasificación, ya que los pronombres personales se encuentran dentro de esta categoría. Tomando el conjunto sin pronombres personales constituido por 822 opiniones verdaderas y 778 opiniones falsas en la tabla se observa, que a pesar que los términos se redujeron, la clase F de los cuatro vocabularios, tiene un promedio de clasificación bajo con respecto a la media, a diferencia de la clase V que inclusive logra superar al conjunto con pronombres personales. Algo que hay que recalcar es la diferencia en la Clase F entre los conjuntos con y sin pronombres personales, demostrando una diferencia con el uso de pronombres en las opiniones falsas. En la Tabla 6, se observan las estadísticas de los dos conjuntos, tomando en cuenta el número de palabras y oraciones. Con respecto al tamaño, se aprecia una disminución en comparación al tomar las opiniones completas. Al pasar por el clasificador *Random Forest* y el selector *Gain Ratio* los conjuntos con pronombres personales en *Op Spam* se tuvo una reducción de hasta 71 veces, mientras que en *Amazon* se contó con una reducción de hasta 171 veces de los datos originales. Con respecto al conjunto sin pronombres personales, el conjunto *Op Spam* tuvo una reducción de hasta 41 veces y 263 veces de reducción en el conjunto de *Amazon*.

7.3. *Uso de pronombres personales del plural*

A demás de los pronombres personales del singular, en este experimento se agregan los pronombres personales del plural, siendo: *"we"*, *"us"*, *"our"*, *"ours"*, *"ourselves"*, los elegidos. Al separar las opiniones por oraciones, se obtiene en *Op Spam* el conjunto con pronombres personales con un total de 768 opiniones verdaderas y de 769 opiniones falsas. En la Tabla 7, se muestra que la configuración seleccionada mejoró en las representaciones *ALL*, *STE* y *STESTW* entre el 76 % y 77 % a diferencia de los resultados de la Tabla 5.

Comparando los cuatro vocabularios, el que tuvo el mejor resultado fue *STESTW*, mostrando un promedio equilibrado entre las dos clases y el promedio general. Los términos más relevantes en esta representación y que vuelven aparecer fueron "hotel" y "chicago", sin embargo aparecen otros nuevos tales como "night", indicando que los hechos de esas opiniones ocurrieron en ese tiempo, a demás del uso del pronombre personal "us", como puede observarse en la Figura 6.

El conjunto que no tiene pronombres personales se forma de 754 opiniones verdaderas y 732 opiniones falsas, con los resultados obtenidos que se muestra en la Tabla 7, al comparar los cuatro vocabularios, el que tuvo mejores resultados fue *STW*, siendo superior a las representaciones restantes. Al buscar qué términos hay en el vocabulario *STW* (Figura 7) se pueden identificar con mayor relevancia las palabras "hotel", "room" y "staff", teniendo relación con el personal y cuartos de los hoteles, inclusive "great", que es un adjetivo que expresa algo que puede amplificar una expresión o hablar de una manera positiva, que en el caso puede estar ligado con lo segundo, debido a que se encuentran términos como "nice" o "good", que son adjetivos que demuestran una polaridad positiva.

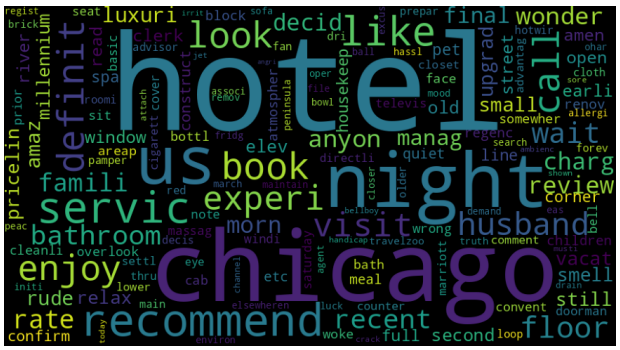

Figura 6: Palabras más representativas del vocabulario *STESTW* con pronombres personales del singular y plural de *Op Spam*

Se puede resaltar que al observar los términos más relevantes dentro de este corpus, a través de las nubes de palabras, se puede ver que las opiniones daban una cierto énfasis en la experiencia en su estancia dentro de los hoteles, con mayor énfasis en la recomendación, que probablemente sería al servicio, al personal y a las posibles zonas que podrían apuntar en los hoteles dentro del estado de Chicago. Conforme al conjunto *Amazon* al aplicar el proceso de separación y añadiendo los pronombres

Tabla 5: Resultados de los conjuntos de datos *Op Spam* y *Amazon*, con y sin pronombres personales del singular, aplicando el clasificador Random Forest y el selector Proporción de Ganancia Información con la métrica F-Measure.

Vocabulario	Pronombres Personales del Singular						Sin Pronombres Personales del Singular					
	Op Spam			*Amazon*			*Op Spam*			*Amazon*		
	V	F	G	V	F	G	V	F	G	V	F	G
ALL	.69	.76	.72	.61	.57	.59	.79	.79	.79	**.68**	**.37**	**.53**
STE	.69	.76	.73	**.62**	**.56**	**.59**	.79	.79	.79	.68	.35	.52
STW	**.72**	**.77**	**.74**	.68	.47	.57	.79	.79	.79	.68	.36	.53
STESTW	.69	.76	.72	.49	.69	.59	**.79**	**.80**	**.80**	.69	.31	.50

Tabla 6: Estadísticas con pronombres personales del singular con los conjuntos *Op Spam (O.S.)* y *Amazon (A)*

	Máx. de palabras		Mín. de palabras		Prom. de palabras		Máx. de oraciones		Mín. de oraciones		Prom. de oraciones	
	O.S.	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*
Opiniones completas	784	1612	25	5	147.27	62.01	58	97	1	1	9.53	4.1
Oraciones con primera persona singulares	519	903	4	2	77.02	44.84	28	40	1	1	4.16	2.41
Oraciones sin primera persona singulares	468	1183	1	1	85	30.62	30	57	1	1	6.11	2.62

Tabla 7: Resultados de los conjuntos de datos *Op Spam* y *Amazon*, con y sin pronombres personales del singular + plural, aplicando el clasificador Random Forest y el selector Proporción de Ganancia Información con la métrica F-Measure

Vocabulario	Pronombres Personales del Singular + Plural						Sin Pronombres Personales del Singular + Plural					
	Op Spam			*Amazon*			*Op Spam*			*Amazon*		
	V	F	G	V	F	G	V	F	G	V	F	G
ALL	.76	.77	.76	.55	.59	.57	.74	.75	.74	**.68**	**.37**	**.53**
STE	.77	.77	.77	**.57**	**.57**	**.57**	.74	.75	.74	.69	.31	.50
STW	.76	.77	.76	.65	.49	.57	**.75**	**.75**	**.75**	.69	.29	.50
STESTW	**.77**	**.78**	**.77**	.48	.68	.58	.75	.74	.75	.69	.29	.50

personales del plural, se obtiene un conjunto con pronombres de 851 opiniones verdaderas y 869 opiniones falsas. Con los resultados obtenidos (Tabla 7) al usar los pronombres personales del plural, la clasificación es comparable, pero no como la obtenida al considerar solo los singulares, aún así el vocabulario que demuestra un gran balance en esta fase es *STE*, teniendo en ambas clases y el promedio general un 57 %. Con el vocabulario *STE* (Figura 8), debido que se anexan los pronombres personales del plural, la carga de datos aumenta y por la aplicación del proceso de stemming se pueden apreciar stopwords. Sin embargo, el sentimentalismo negativo que se muestra en el término "hate" tiende a mostrarse, denotando que podría existir algún descontento en estas opiniones conforme al producto o servicio que ofrecen.

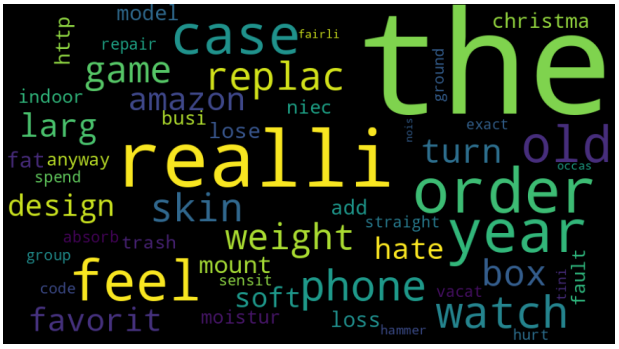

Figura 8: Palabras más representativas con el vocabulario *STE* con pronombres personales singulares y plurales en *Amazon*.

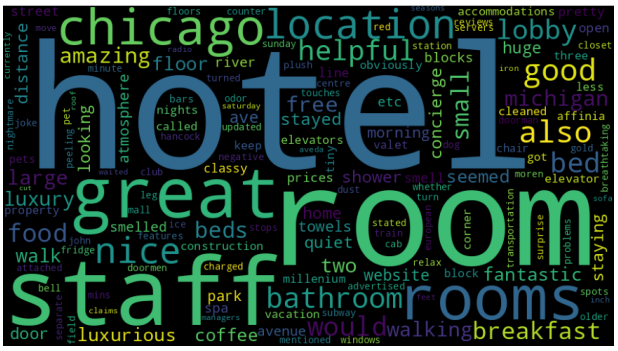

Figura 7: Palabras más representativas del vocabulario *STESTW* sin pronombres personales del singular y plural de *Op Spam*.

Figura 9: Palabras más representativas con el vocabulario *ALL* sin pronombres personales singulares y plurales en *Amazon*.

Con los resultados obtenidos en la Tabla 7, que se asemejan a los de la Tabla 5, se puede apreciar que al descartar los pronombres y querer clasificar las opiniones, las opiniones falsas son afectadas, ya que no se identifican correctamente, pero las opiniones verdaderas resultan mejor en su detección. Lo anterior demuestra una vez más que los pronombres personales son un factor que puede ser usado para la detección de publicaciones engañosas. Tomando el mejor resultado en los cuatro vocabularios, el que tiene un mejor promedio es *ALL*, debido a que, a pesar que la detección de opiniones verdaderas, es menor por un punto, en comparación de los tres restantes con un promedio del 68 %. Tomando los términos del vocabulario *ALL*, en la Figura 9 se aprecia un aumento de ellos teniendo igual los más relevantes como los mostrados en la Figura 8. El término "product" muestra que se trata de evitar mezclarse en las oraciones que contengan un pronombre personal, pero se presentan términos que podrían afectar al producto tales como "stress", "dirt" y "fake".

Con respecto a las estadísticas de los conjuntos, se puede observar en la Tabla 8 que se genera un incremento en el conjunto de oraciones con pronombres personales, debido a que se consideran los pronombres plurales, sin embargo se demuestra la reducción de la carga de datos conforme al conjunto original. Al aplicar el algoritmo de selección y clasificación se tiene, con el conjunto con pronombres en *Op Spam*, una reducción de hasta 54 veces, mientras que con *Amazon* es de 174 veces. Con respecto al conjunto que no tiene pronombres personales se tuvo con *Op Spam* una reducción de hasta 174 veces, entretanto *Amazon* presentó una reducción de 336 veces.

8. Conclusiones

Con los resultados obtenidos en estos conjuntos de datos, el uso de los pronombres personales ayuda en la identificación de publicaciones engañosas, además que se percibe una reducción considerable de términos. Por medio de la metodología propuesta, se observa que el dominio donde se aplique puede ser beneficiado, debido al enfoque que se da en los pronombres personales, especialmente porque se relacionan con la honestidad, tal y como lo menciona Pennebaker (Newman et al., 2003).

En trabajos relacionados, Sánchez (Sánchez Junquera et al., 2017) menciona que el retiro de stopwords o stemming, podría alterar los resultados de clasificación, sin embargo, en la experimentación mostrada con *Op Spam*, se muestra que los mejores resultados se obtienen cuando el vocabulario de datos se construye retirando los stopwords y aplicando stemming (vocabulario *STESTW*), aún cuando se emplea la separación de oraciones por pronombres del singular y sumando el plural. El considerar el uso de este tipo de pronombres ayuda a identificar el tipo de publicación u opinión que se puede observar en algún sitio en línea. Como se pudo observar en los resultados, cuando se hace la separación de las opiniones en oraciones que contengan o no pronombres personales (ya sea del tipo singular o plural) la clasificación que se obtiene es competitiva a procesar los conjuntos de datos completos.

Por último, al observar las nubes de palabras en cada conjunto de datos, se pudo encontrar que en las opiniones con pronombres personales, las personas tienden a hablar de términos más relacionados a la situación personal en cuestión de lo que

Tabla 8: Estadísticas con pronombres personales del singular y plural con los conjuntos *Op Spam (O.S.)* y *Amazon (A)*.

	Máx. de palabras		Mín. de palabras		Prom. de palabras		Máx. de oraciones		Mín. de oraciones		Prom. de oraciones	
	O.S.	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*	*O.S.*	*A.*
Opiniones completas	784	1612	25	5	147.27	62.01	58	97	1	1	9.53	4.1
Oraciones con primera persona singulares + plurales	642	976	4	2	103.06	45.65	30	46	1	1	5.69	2.47
Oraciones sin primera persona singulares + plurales	389	1121	1	1	52.03	28.87	28	54	1	1	4.4	2.53

está experimentando, mientras que las que no contienen pronombres, tienden a relacionarse a cuestiones interpersonales, es decir todo lo que se involucre en cuestión de espacio e interacciones. Como trabajo futuro, se plantea explorar estas observaciones con el fin de mejorar la detección de engaños.

Referencias

Ahmed, H., Traore, I., and Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy*, 1(1):e9.

Altunbey Ozbay, F. and Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*.

Buller, D. B. and Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, pages 203 –242.

Cabrejas, C., Martí, J. V., Pajares, A., and V., S. (2019). Deception detection in arabic texts using n-grams text mining. In *FIRE (Working Notes)*.

Fuller, M. C., Biros, P. D., and Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications*, 38:8392–8398.

Fuller, M. C., Biros, P. D., and Delen, D. (2014). Automated deception detection of 911 call transcripts. *Security informatics*, 8(7).

G., D. B. (2007). Deception detection expertise. *American Psychology-Law Society*, pages 339 – 351.

Gelbukh, A. (2010). Procesamiento de lenguaje natural y sus aplicaciones. komputer. *Pattern recognition and information forensics*, pages 6 – 11.

Holland, D. and Quinn, N. (1995). *Cultural models in language and thougth*. Cambridge University Press.

Krishnamurthy, G., Majumder, N., Poria, S., and Cambria, E. (2018). A deep learning approach form multimodal deception detection. *arXiv preprint arXiv:1803.00344*.

Mbaziira, A. and Jones, J. (2016). A text-based deception detection model for cybercrime. In *Int. Conf. Technol. Manag*, pages 1–8.

Newman, L. M., Pennebaker, J. W., Berry, D. S., and Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *PSPB*, 29.

Ortega Mendoza, R. M., Hernández Farías, D. I., Montes y Gómez, M., and Villaseñor Pineda, L. (2022). Revealing traces of depression through personal statements analysis in social media. *Artificial Intelligence in Medicine*, 123:102202.

Ortega Mendoza, R. M., López Monroy, A. P., Franco Arcega, A., and Montes y Gómez, M. (2018). Peimex at erisk2018: Emphasizing personal information for depression and anorexia detection. In *CLEF (Working Notes)*.

Ortega Mendoza, R. M., Villaseñor Pineda, L., and Montes y Gómez, M. (2007). Using lexical patterns for extracting hyponyms from the web. In *Mexican International Conference on Artificial Intelligence*, pages 904–911. Springer.

Ott, M., Cardie, C., and Hancock, J. (2013). Negative deceptive opinion spam. *NAACL HLT 2013 - 2013 Conference of the North American chapter of the association for computational linguistics: Human language technologies, proceedings of the main conference*, pages 497–501.

Rill García, R., Villaseñor Pineda, L., Reyes Meza, V., and Escalante, H. (2019). From text to speech: A multimodal cross-domain approach for deception detection. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, pages 164–177.

Ruiz, W. (2006). Técnicas de minería de datos aplicadas en la detección de fraude. pages 1 – 9.

Sánchez Junquera, J., Villaseñor Pineda, L., Escalante, H., Montes, and Gómez, M. (2017). Detección del engaño en notas de opinión a través de técnicas tradicionales de clasificación automática de textos. *Res. Comput. Sci.*, 134:141–150.

Song, F., Liu, S., and Yang, J. (2005). A comparative study on text representation schemes in text categorization. *Pattern analysis and applications*, 8:199–209.

Wanumen Silva, L. F. (2010). Minería de datos para la predicción de fraudes en tarjetas de crédito. pages 44 – 57.

Zuckerman, M., DePaulo, B. M., and Rosenthal, R. (1981). Verbal and non verbal communication of deception. *Advances in expenmental social psychology*, 14.