

## USO DE CATEGORÍA GRAMATICAL EN LA IDENTIFICACIÓN DE SENTIMIENTO

### *Use of Gramatical Category for Sentiment Identification*

*Emmanuel Abarca Jiménez*<sup>1</sup>

*Rocío Ruiz Ramón*<sup>2</sup>

*Edgar Casasola Murillo*<sup>3</sup>

#### RESUMEN

Esta investigación analiza el impacto que tiene el filtrado de n-gramas de palabras, usando su categoría gramatical, sobre la identificación de sentimiento a partir de comentarios de texto provenientes de redes sociales en idioma español. Se investigó el impacto de filtrar n-gramas que contienen adjetivos, adverbios e interjecciones. Se logró determinar que es posible reducir el volumen de datos procesado y, a la vez, lograr una mejora de hasta un 30 % en la precisión al clasificar un corpus anotado de comentarios separando los que expresan sentimiento de los que no.

**Palabras clave:** Procesamiento del lenguaje natural, categoría gramatical, análisis de sentimiento.

#### ABSTRACT

This research analyzes the impact of the filtering of word n-grams, using their grammatical category, on the identification of sentiment based on text comments from social networks in Spanish. The impact of filtering n-grams containing adjectives, adverbs and interjections was investigated. It was determined that it is possible to reduce the volume of processed data and at the same time achieve an improvement of up to 30 % in the accuracy when classifying an annotated corpus of test comments separating those that contain sentiment from those that do not.

**Key Words:** Natural Language Processing, Part of Speech, Sentiment Analysis.

---

<sup>1</sup>Universidad de Costa Rica. Programa de Posgrado en Computación. Costa Rica.

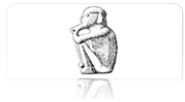
Correo electrónico: [emmanuel.abarca@ucr.ac.cr](mailto:emmanuel.abarca@ucr.ac.cr)

<sup>2</sup>Universidad de Costa Rica. Programa de Posgrado en Lingüística. Costa Rica.

Correo electrónico: [rocio.ruiz@ucr.ac.cr](mailto:rocio.ruiz@ucr.ac.cr)

<sup>3</sup>Universidad de Costa Rica. Escuela de Ciencias de la Computación e Informática y Centro de Investigaciones en Tecnologías de la Información y Comunicación (CITIC). Costa Rica.

Correo electrónico: [edgar.casasola@ucr.ac.cr](mailto:edgar.casasola@ucr.ac.cr)



## 1. Introducción

El análisis de sentimiento es un campo de estudio que se ha desarrollado en los últimos años, debido a su relevancia para diferentes ámbitos tanto académicos como comerciales. Gracias a un análisis de los comentarios publicados en medios electrónicos se pueden estudiar las percepciones de usuarios, con el fin de determinar tendencias y opiniones, particularmente la opinión sobre entidades, tales como instituciones, productos comerciales, productos de entretenimiento o incluso sobre personalidades de la farándula o del mundo de la política. Por este motivo, resulta relevante desarrollar este tipo de estudios en español, debido a que no solo es un campo en desarrollo, sino que también existen pocos análisis realizados específicamente para el español (Casasola y Leoni, 2017). Actualmente, la mayor parte de la investigación en español se concentra en el trabajo llevado a cabo por la Sociedad Española para Procesamiento de Lenguaje Natural (SEPLN) a través del taller TASS mencionado por Martínez-Camara *et al.* (2016), sin embargo, los datos de prueba están compuestos por comentarios de Twitter usando español peninsular solamente. Por lo tanto, aún queda mucho por investigar para contar con estudios de otras variedades del español en América.

El objetivo de la presente investigación es generar datos relevantes para el análisis de sentimiento en español, por medio de pruebas realizadas a partir de categorías gramaticales que se consideraron adecuadas para ayudar a reconocer eficazmente la subjetividad en las frases. La finalidad es definir cuáles categorías son más apropiadas para detectar sentimiento y para esto se eligieron los adjetivos, adverbios e interjecciones. Por otra parte, se tomó en cuenta la puntuación para analizar si esta característica influye de alguna manera en la manifestación de subjetividad. Las pruebas se realizaron utilizando el corpus del Taller de Análisis de Sentimiento (TASS) y se partió de investigaciones previas pertenecientes al campo, que resalta la importancia de dicho tipo de análisis para diferentes objetivos.

Una de las fuentes principales fue el texto *Sentiment Analysis and Opinion Mining*, de Liu (2012). Este autor aclara que el análisis de sentimiento es importante, ya que puede ser aplicado en situaciones de la vida real. Sumado a esto, se resalta que antes



del año 2000 no fue un tema de estudio popular en el procesamiento de lenguaje natural ni en la lingüística, debido a que para ese tiempo no era fácil acceder a grandes cantidades de corpus de opiniones de forma digital. No obstante, a partir del año 2000, este campo de investigación ha crecido rápidamente y se ha hecho popular también en el campo de las ciencias administrativas y en muchas otras áreas como la política o las opiniones según el género (Liu, 2012).

En cuanto a los elementos en los que se basa este tipo de investigación, Liu menciona la importancia medular de las llamadas *palabras de opinión*, ya que estas son las que comúnmente se utilizan para referirse de forma positiva o negativa a temas, cosas y personas (2012, p. 12). Algunas de estas palabras pueden ser *mal*, *bien*, *excelente*, *terrible*. Además de las palabras, también hay frases que pueden ser utilizadas para expresar opiniones. Sin embargo, las complejidades que presenta este tipo de investigación surgen cuando se resalta que las palabras de opinión no siempre son suficientes para detectar la polaridad de un enunciado (polaridad se refiere a las emociones positivas o negativas que una persona transmite cuando opina sobre algo). No obstante, para la presente investigación, se tomarán en cuenta solamente los enunciados que incluyan una polaridad explícita. Aunado a esto, los sistemas utilizados para dicho análisis no requieren de un conocimiento total de la semántica de cada frase, ya que solo las características formales son necesarias para poder llevar a cabo el estudio.

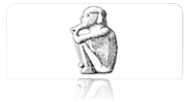
Por otra parte, Phan y Cao (2014) destacan que el análisis de sentimiento es de valor para el estudio de comentarios de usuarios en redes sociales y valoraciones de entidades. De acuerdo con Liu (2012), una entidad se refiere a cualquier producto, servicio, persona, evento, organización o tema sobre el que se emite una opinión. En el caso de Phan y Cao, se analizaron comentarios de usuarios con respecto a servicios de turismo y alimentos. Para la fase de recolección del corpus, en primer lugar, estos investigadores extrajeron de forma automática los comentarios de los tipos de sitios web indicados por medio de procesos de análisis de web. En segundo lugar, realizaron el preprocesamiento de los textos en el cual eliminaron errores y emoticones. En tercer lugar, efectuaron un etiquetado de los comentarios a nivel de frase para marcar las categorías gramaticales correspondientes. Después, ejecutaron un proceso de minería de opinión. Por último, se aplicó un modelo de *skip-grama* para extraer grupos de palabras



para el análisis. Los resultados de esa investigación demostraron que el sistema de búsqueda de restaurantes fue mejorado gracias a la minería de opinión de los comentarios y calificaciones de usuarios, esto fue logrado gracias al modelo de *skip-gramas* y a la categorización de oraciones; las frases más útiles resultaron ser las que tenían mayor descripción.

Aparte de esta investigación, en el trabajo de Wiebe *et al.* (2006) se profundizó en el tema de la subjetividad en el lenguaje natural por medio de una investigación que buscaba identificar oraciones subjetivas y documentos que contienen opiniones. En ese trabajo, se utilizó el lenguaje subjetivo presente en un corpus para identificar elementos contextuales para producir conocimiento y reconocer las oraciones subjetivas y documentos de opinión. Además, se menciona que los sistemas de extracción de información deberían ser capaces de distinguir cuándo se trata de hechos y cuándo no, ya que solo estos primeros deben ser extraídos. Agregan que los sistemas de respuesta a preguntas deberían ser capaces de presentar diferentes respuestas al usuario basándose en especulación u opiniones de diferentes fuentes. Se destaca también que para realizar un buen análisis automático de subjetividad es necesario contar con buenas claves para que el sistema de procesamiento de lenguaje natural sea capaz de desambiguar estas expresiones en el contexto. Para la investigación, utilizaron un corpus de artículos de opinión y comentarios del *Wall Street Journal*. Para la selección de información se enfocaron en tres tipos de pistas de subjetividad. La primera es el grupo de palabras que aparecen solo una vez en el corpus o *palabras únicas* o *hápax legómenon*. El segundo tipo son las colocaciones que son identificadas por medio de n-gramas. Por último, el tercer tipo son los adjetivos y los verbos y parten de la hipótesis de que dos palabras con una distribución similar son potencialmente subjetivas. Los resultados demostraron que las palabras únicas son más subjetivas de lo que se espera y pueden ofrecer pistas valiosas, también su procedimiento fue útil para identificar colocaciones potencialmente subjetivas. Asimismo, su análisis de palabras con distribución similar también fue provechoso, de modo que los tres pasos incrementaron el éxito de la detección de subjetividad.

Finalmente, se menciona el trabajo de Hatzivassiloglou y Wiebe (2000), cuyo tema trata la importancia de la subjetividad para la extracción y la recuperación de



información. El estudio se centra en adjetivos y su efecto al utilizar un clasificador de subjetividad simple. Además, se enfoca en el efecto que tienen cierto tipo de adjetivos como los dinámicos, los orientados semánticamente y los graduables. Su fin computacional era encontrar la diferencia entre oraciones que utilizan opiniones y otras formas de subjetividad, y oraciones que son utilizadas para presentar información de hechos. Los resultados de su investigación demostraron que los objetivos semánticos y los graduables fueron de gran utilidad para demostrar la subjetividad de un enunciado.

## 2. Marco conceptual

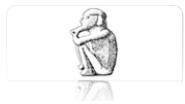
Para los lineamientos conceptuales que sirvieron de base para este trabajo, se partió de lo propuesto por Liu (2012), quien plantea una definición para el *análisis de sentimiento* o *minería de opinión*. Liu propone que el análisis de sentimiento puede definirse como el campo de estudio que analiza las opiniones, sentimientos, evaluaciones, actitudes y emociones de las personas hacia productos, servicios, organizaciones, individuos, y sus atributos. El análisis de sentimiento se centra principalmente en opiniones que expresan o implican sentimientos positivos o negativos. Como se mencionó en el apartado anterior, el análisis de sentimiento tiene una gran variedad de aplicaciones que pueden ir de lo político a lo comercial. Además, presenta muchos retos para la investigación que aún no han sido abordados.

Para Liu (2012), la opinión consiste en dos elementos clave: la *entidad* y un *sentimiento* que se expresa sobre esta. La entidad es entonces el objeto de la opinión y puede ser una persona, una institución, un evento, un servicio, un producto u otro.

Liu (2012) propone los componentes de la estructura de una opinión, los cuales se muestran en la Tabla 1 con su respectiva definición.

**Tabla 1.** Estructura de una opinión

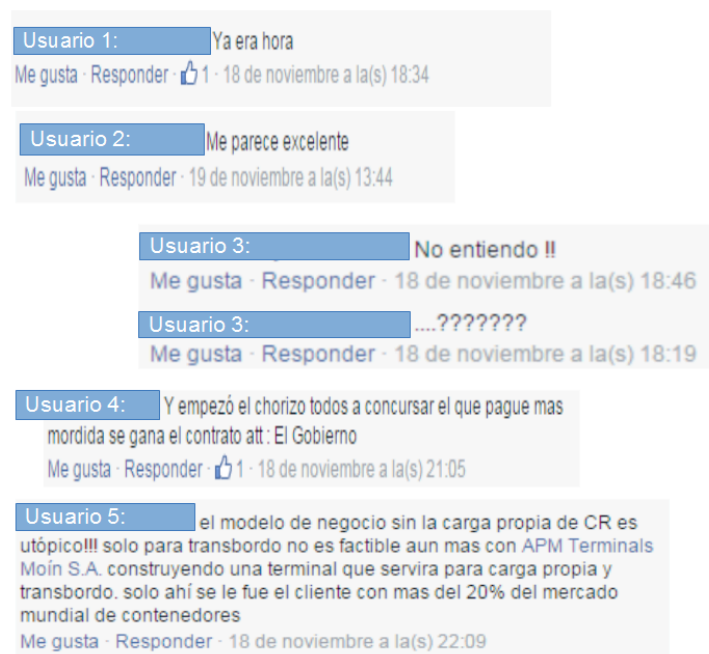
|                |  |
|----------------|--|
| <b>Entidad</b> | Objeto de la opinión.  |
| <b>Aspecto</b> | Propiedad, atributo o característica sobre la que se emite la opinión. |
| <b>Autor</b>   | Individuo que opina.   |



|                   |  |
|-------------------|--|
| <b>Polaridad</b>  | Orientación positiva, negativa o neutral del comentario. También puede usarse una escala de 5 valores: <i>N+</i> , <i>N</i> , <i>NEU</i> , <i>P</i> , <i>P+</i> que van desde muy negativo <i>+N</i> a muy positivo <i>+P</i> pasando por neutral <i>NEU</i> . |
| <b>Día y hora</b> | Indica el momento en que se emite la opinión. Esto permite que el autor pueda cambiar de opinión a lo largo del tiempo.  |

Fuente: Elaboración propia a partir de Liu (2012).

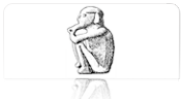
Casasola y Leoni (2017) ejemplifican esta composición de la opinión usando comentarios de Facebook en una noticia relevante a Costa Rica, publicada en la página de Facebook de Amelia Rueda el 18 de noviembre de 2014, según se puede ver en la Figura 1.



**Figura 1.** Ejemplo de tipos de comentario emitidos por usuarios sobre una noticia en Facebook

Fuente: Tomado de Casasola y Leoni (2012).

Es posible estudiar los comentarios aplicando la estructura anterior. Por ejemplo, el comentario “Me parece excelente” va dirigido a la entidad *Mega terminal de*

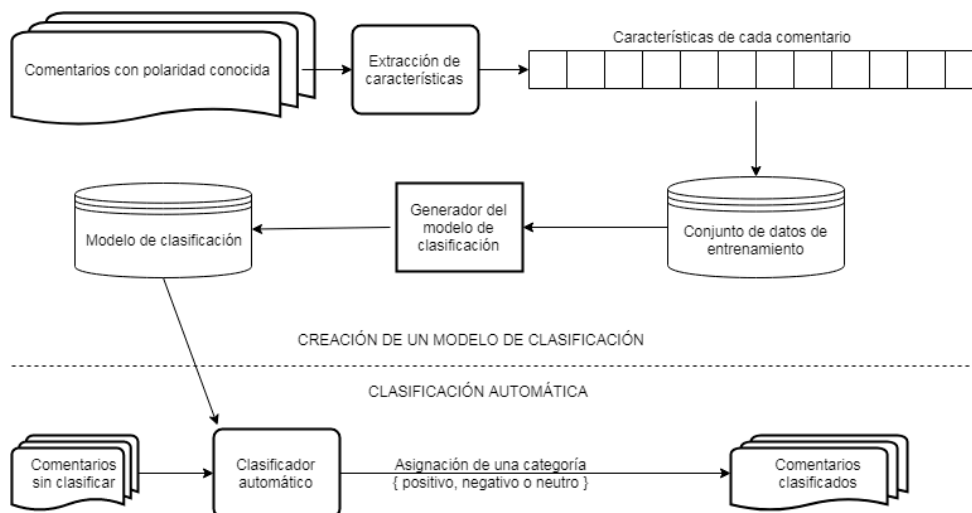


*transbordo de Moín*. El aspecto en ese caso es *general*, ya que se refiere a la entidad como un todo. El usuario 2 es quien emitió la opinión y es considerada muy positiva o P+. La hora y lugar son el 19 de noviembre de 2014 a la 1:44 p. m.

Otro ejemplo es el comentario “Y empezó el chorizo todos a concursar el que pague mas mordida se gana el contrato att: El Gobierno”. En este caso, la entidad es la misma que en el caso anterior, *la Mega Terminal de Moín*, el aspecto es *la corrupción* y es una opinión muy negativa o +N, emitida por el usuario 4 el 18 de noviembre de 2014 a las 9:05 p. m.

Sin embargo, un ejemplo muy relevante para este trabajo es el caso del comentario “No entiendo!!”. Este comentario no corresponde a ninguna opinión. Se sabe que el usuario 3 emitió ese comentario el 18 de noviembre de 2014 a las 6:46 p. m., pero este texto no corresponde a ningún aspecto en particular y debería ser descartado y no ser tomado en cuenta como una opinión. En este caso, este comentario debería indicar que no corresponde a ninguna categoría o polaridad. En algunos casos a este tipo de comentarios se les clasifica usando una etiqueta en inglés con el término *NONE*.

Para realizar la identificación de la polaridad de un comentario de manera automática, se puede usar un clasificador automático, tal como lo describe la Figura 2. El primer paso del proceso que se muestra en la Figura 2 es el mismo que se aplica a comentarios con una clasificación (polaridad) ya conocida como a comentarios que se desean clasificar. Nótese que el procesamiento de texto que se realiza en la parte superior de la figura muestra cómo se pretende obtener un vector de características a partir del comentario. Estas características pueden variar según el modelo de lenguaje que se desee, por ejemplo, puede ser una *bolsa de palabras* que contiene el conjunto de las palabras que conforman el comentario.



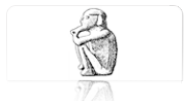
**Figura 2.** Diagrama que muestra el proceso que lleva a cabo un clasificador

Fuente: Elaboración propia.

Posteriormente, se nota en la Figura 2 que el *clasificador* se alimenta con cada lista de características y su polaridad conocida; este paso se lleva a cabo en la etapa de entrenamiento. Cabe aclarar que el grupo completo de datos con los que se *entrena* el clasificador se conoce como datos de entrenamiento. A lo interno, la función del clasificador es determinar una asociación entre los valores de las características correspondientes a una polaridad determinada para la opinión. Por esto, cuando al clasificador entrenado se le presenta un texto sin categoría asignada, este clasificador procederá a asignarle una categoría o polaridad partiendo de los datos con los que ha sido entrenado, tomando en cuenta a cuál polaridad es más probable que pertenezca el comentario.

Para este experimento en particular, se utilizó un clasificador de tipo *Bayes ingenuo*, un subtipo de los clasificadores bayesianos que buscan la probabilidad más alta para cada lista de características. De hecho, según menciona Rish (2001), este tipo de clasificador bayesiano simplifica esta tarea, ya que supone que no hay dependencia entre las características.



**Tabla 2.** Ejemplo de comentarios de texto con categoría asignada

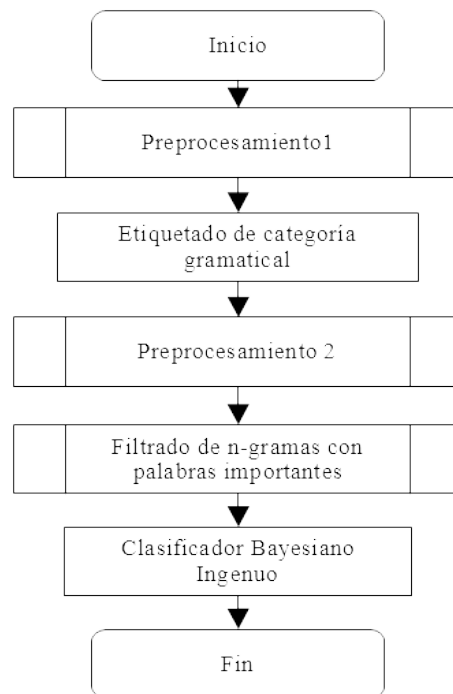
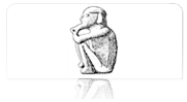
| <b>Categoría</b> | <b>Comentario</b>                |
|------------------|----------------------------------|
| P+               | GRACIAS AMIGOS YA SOMOS 200000;) |
| NONE             | Retazos de una primera dama      |
| N                | En el blog Lágrimas tecnócratas  |
| P+               | Gracias Venezuela                |

Fuente: Elaboración propia.

La Tabla 2 muestra una selección de comentarios de entrenamiento del corpus de TASS, que sirven para ejemplificar lo que hace un clasificador, pues si se ha entrenado un clasificador automático con estos datos y se desea que clasifique el comentario “Gracias a todos”, este no sería capaz de encontrar las palabras “a” ni “todos” dentro de los datos que maneja, pero sí encontraría “gracias”. “Gracias” está solo en comentarios P+ por lo que clasificaría el comentario como P+ al encontrar que, de acuerdo con el entrenamiento dado, “Gracias” tiene una alta probabilidad de aparecer en comentarios P+.

### 3. Metodología

La Figura 3 muestra el diseño del sistema propuesto para identificar sentimiento de manera automática. El corpus utilizado fue etiquetado agregando una categoría para la polaridad del texto. La polaridad puede tener los siguientes valores: N+ (muy negativo), N (negativo), NEU (neutral), P (positivo), P+ (muy positivo) o NONE (sin sentimiento). Cabe aclarar que se buscó distinguir los comentarios con sentimiento (N+, N, NEU, P, P+) de aquellos carentes de sentimiento (NONE) con este sistema.



**Figura 3.** Fases del clasificador desarrollado

Fuente: Elaboración propia.

Como se puede observar en la Figura 3, este sistema consta de varias fases. En la fase de preprocesamiento 1, se eliminan los enlaces a páginas de internet, debido a que sin un estudio más profundo, el análisis del contenido de la página vinculada no tendrá significado. En segundo lugar, para la fase de etiquetado de categoría gramatical, se utilizó Freeling, una herramienta de análisis multilingüe elaborada por Padró y Stanilovsky (2011) que fue utilizada para separar componentes de la oración e identificar su categoría gramatical. Estos componentes son usualmente palabras, pero para esta investigación también se consideró la puntuación y los pronombres clíticos como componentes separados.

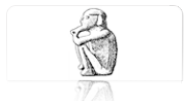
Una importante distinción de Freeling con respecto de otras herramientas similares (como Stanford POS Tagger) es que distingue pronombres clíticos. Por ejemplo, el verbo “cómétela” es analizado como: **come** (verbo segunda persona imperativo) **te** (pronombre segunda persona) **la** (pronombre tercera persona femenino



acusativo). Este elemento resulta importante, ya que aporta más información y, pese a que aumenta la distancia entre palabras, es posible que gracias al uso de trigramas se disminuyan las implicaciones negativas de este hecho. Aunque Freeling provee información detallada, como género y tiempo, se decidió usar solamente la categoría gramatical principal, puesto que se observó que la cantidad de subcategorías correctamente asignadas era menor que para la categoría principal. En la fase de preprocesamiento 2, se agruparon los componentes de tipo puntuación, esto con tal de que fuera posible reconocer emoticones en fases siguientes. Por otra parte, en la fase de filtrado de n-gramas con palabras importantes, se generaron los unigramas, los bigramas y los trigramas correspondientes, pero solamente se conservaron los n-gramas que contuvieran al menos un componente de alguna de las categorías gramaticales que resultan de interés según cada tratamiento a evaluar en el cuasi experimento realizado.

La metodología utilizada para evaluar los resultados en este trabajo fue de tipo cuasi experimental, pues no evaluó datos completamente basados en muestras aleatorias, sino que se utilizó un conjunto de datos de prueba basado en un corpus. Se estudió si la reducción de la cantidad de n-gramas basada en las categorías gramaticales permite obtener precisiones similares o mejores al clasificar los comentarios con la correspondiente disminución en el volumen de datos que se procesa. Una reducción en el volumen de información procesada permite reducir los tiempos de procesamiento y los requerimientos de espacio de almacenamiento de la información.

Se entrenaron los clasificadores usando n-gramas de dos y tres términos. Los bigramas y los trigramas permiten modelar y obtener información referente a las relaciones que existen entre grupos de términos consecutivos en un texto. Los bigramas son todos los grupos de dos términos consecutivos que aparecen en el texto. Y los trigramas, los grupos de tres términos consecutivos; dado que se utilizó Freeling, un término puede ser entonces una palabra o un signo de puntuación. Por ejemplo, en la frase “hoy, como ayer, estoy feliz”, los bigramas posibles, si se toman en cuenta los signos de puntuación, corresponden a: (“hoy” , “,” ) , ( “,” , “como” ) , (“como” , “ayer”), (“ayer” , “,”), (“,” , “estoy”), (“estoy” , “feliz”). Nótese que el primer bigrama se compone de la palabra hoy y el signo correspondiente a la coma. En la etapa de asignación de categoría gramatical, la presencia de los signos de puntuación implica un



problema, ya que formalmente el signo de puntuación no tiene asociada una categoría gramatical, sin embargo, para efectos de esta investigación se consideró evaluar la presencia o la ausencia de signos de puntuación como parte de los n-gramas, pues Freeling les asigna una etiqueta particular.

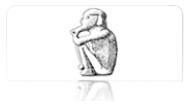
Es así como se representó el texto de cada comentario conservando bigramas y trigramas que tenían combinaciones de términos donde, al menos, uno pertenece a las categorías especificadas en cada combinación o **tratamiento** que fue evaluado. Las categorías utilizadas se codificaron según la categoría asignada por Freeling, tal y como se muestra en la Tabla 3.

Los n-gramas utilizados en cada tratamiento son aquellos que contienen, al menos, uno de sus términos perteneciente a alguna de las categorías indicadas en el tratamiento. Las categorías posibles son adjetivo, adverbio, interjección, sustantivo, preposición y se incluye el signo de puntuación como una etiqueta adicional.

**Tabla 3.** Categorías asignadas a los términos de los n-gramas

| <b>Categoría</b> | <b>Descripción</b> |
|------------------|--------------------|
| A                | adjetivo           |
| R                | adverbio           |
| I                | interjección       |
| F                | puntuación         |
| N                | sustantivo         |
| S                | preposición        |

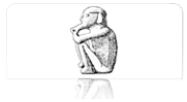
Un tratamiento corresponde entonces con el conjunto de categorías gramaticales evaluadas, cuya aparición en un n-grama hace que el mismo sea conservado, mientras que todos los demás son filtrados y no se toman en cuenta por parte del clasificador. Los siete tratamientos que se evaluaron fueron el *caso base*, que conserva todos los n-gramas de hasta tres términos, y demás tratamientos correspondientes al filtrado de n-gramas basado en combinaciones de categorías, a saber: ARIFN, ARIFS, ARI, AR, ARIF, ARF. A continuación, estos tratamientos se describen en la Tabla 4.

**Tabla 4.** Descripción de los tratamientos evaluados

| Tratamiento | Descripción   |
|-------------|---|
| Base        | Se usan todos los bigramas y los trigramas.   |
| ARIFN       | Se usan solo los n-gramas que contengan al menos un término con categoría: Adjetivo (A), Adverbio (R), Interjección (I), Puntuación (F) o Sustantivo (N). |
| ARIFS       | Similar al anterior para: Adjetivo (A), Adverbio (R), Interjección (I), Puntuación (F) o Preposición (S).   |
| ARI         | Conserva n-gramas con al menos un: Adjetivo (A), Adverbio (R) o Interjección (I).   |
| AR          | N-gramas con al menos un: Adjetivo (A) o Adverbio (R).  |
| ARIF        | Similar conservando: Adjetivo (A), Adverbio (R), Interjección (I) o Puntuación (F).   |
| ARF         | Conserva N-gramas con: Adjetivo (A), Adverbio (R) o Puntuación (F).   |

Con los n-gramas que pasan la etapa de filtrado se entrena un clasificador automático específico por tratamiento.

Para evaluar la funcionalidad de este sistema, se utilizó como conjunto de datos de prueba el corpus de comentarios con polaridad etiquetada del TASS, proveído por la SEPLN, cuya creación se describe el trabajo de Martínez-Cámara *et al.* (2016). El corpus contiene 68 000 comentarios extraídos de Twitter, escritos en español sobre 150 celebridades y personas descritas por el autor como importantes en el mundo de la política, la economía, la comunicación y la cultura. Ese corpus fue compilado entre noviembre de 2011 y marzo de 2012, según describe Martínez-Cámara *et al.* (2016). Cabe aclarar que pese a que los autores de estos comentarios son de distintas nacionalidades de habla española, los evaluadores del corpus fueron peninsulares, lo que



puede producir casos de evaluaciones que resulten confusas para un hablante no peninsular.

Se buscó distinguir los comentarios con sentimiento (N+, N, NEU, P, P+) de aquellos carentes de sentimiento (NONE) de forma automática. La calidad de los resultados se determina entonces para cada tratamiento observando el impacto que tiene sobre la *precisión* del clasificador entrenado con los n-gramas específicos de cada tratamiento. La precisión se estima como la proporción de comentarios, cuya categoría fue correctamente asignada entre el total de comentarios evaluados. Por tanto, la calidad de cada tratamiento se evalúa basada en la comparación de la precisión que obtiene cada clasificador.

#### 4. Resultados

La Figura 4 muestra la precisión obtenida como resultado de la evaluación para cada tratamiento. Según se detalla en la Tabla 4, el caso *base* es aquel donde los resultados obtenidos corresponden a realizar la prueba sin utilización de filtrado por categoría gramatical, esto resultó en un 45.23 por ciento de precisión, por lo que las comparaciones de los demás tratamientos se llevan a cabo tratando de observar la mejora con respecto a este rendimiento.



**Figura 4.** Precisión obtenida para cada tratamiento

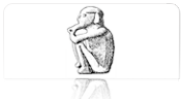


Para el tratamiento ARIFN (adjetivo, adverbio, interjección, puntuación y sustantivo) se obtuvo 49.30 por ciento de precisión. Mientras que con ARIFS (adjetivo, adverbio, interjección, puntuación y preposición) se obtuvo 52.52 por ciento de precisión. Con ARI (adjetivo, adverbio, interjección), 56.35 por ciento. Y en el caso del tratamiento AR (adjetivo, adverbio) se logró obtener 56.45 por ciento, mientras que con ARIF (adjetivo, adverbio, interjección, puntuación) se logró obtener un 59.48 por ciento. No obstante, el mejor rendimiento fue obtenido con ARF, que fueron los n-gramas donde alguno de sus componentes pertenecía a la categoría gramatical adjetivo, adverbio o el término era un signo de puntuación; con estos se obtuvo 59.57 por ciento de precisión.

## 5. Discusión

Los resultados base obtuvieron un 45.23 por ciento de precisión y, al filtrar los n-gramas que no incluían al menos una palabra de categoría gramatical adjetivo, adverbio o puntuación, se obtuvo una precisión de 59.57 por ciento, lo que corresponde a un aumento de 31.70 por ciento con respecto al resultado base. De acuerdo con la cantidad de n-gramas utilizados (es decir, los n-gramas resultantes luego de filtrar), los resultados base generaban 179 835 n-gramas, mientras que filtrando los n-gramas que no contenían algún ARF, este número se redujo a 57 632 n-gramas, por lo que bajó a 32.04 por ciento el total del volumen de datos necesario para entrenar el clasificador.

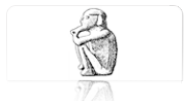
Al construir clasificadores es importante disminuir el volumen de datos, ya que tiene implicaciones en cuanto a la cantidad de tiempo necesario para su entrenamiento. Reducir la cantidad de datos a un 32 por ciento de la cantidad original que se habría tenido que procesar es un logro importante; más si la calidad de la evaluación que se obtiene con el clasificador aumenta de 45 por ciento a casi 60 por ciento. En otras palabras, con un volumen de datos mucho menor se logra obtener una mayor precisión al identificar la emoción presente en los comentarios preclasificados en el conjunto de datos anotado. Esta investigación permite además identificar que efectivamente para distinguir comentarios con sentimiento, identificar los adjetivos, los adverbios y la presencia de signos de puntuación en el texto son buenos indicadores para utilizar en tal tarea.



## 6. Trabajo futuro

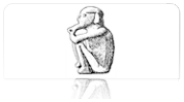
Como se mencionó, el análisis de sentimiento es un campo en desarrollo en el cual aún queda mucho por hacer. Por este motivo, parte de los proyectos a futuro podría ser analizar en la fase de preprocesamiento 1 la página web vinculada para conseguir más contexto, pues esto permitiría identificar más entidades o aspectos de los que se podría estar opinando. Por otra parte, para la generación de n-gramas, se podría hacer uso de *skip-gramas*, de esta forma sería posible analizar las relaciones entre palabras más distantes que no necesariamente aparecen de forma contigua en los comentarios. Además, en la fase de filtrado, sería interesante utilizar información más detallada, como el modo o el tiempo, y esto también conllevaría una mejora necesaria en la precisión de los etiquetadores de categoría gramatical. Finalmente, se podría filtrar por patrones, por ejemplo, conservar trigramas de la forma <sustantivo, verbo, sustantivo>. De realizar todo lo propuesto, se esperaría conseguir un rendimiento superior en la detección de subjetividad en las frases.





## Referencias bibliográficas

- Casasola, E. (2015). [casasola] (2015, Noviembre 26) Tecnologías del lenguaje en Costa Rica [Archivo de video]. Recuperado de <https://www.youtube.com/watch?v=y66DlvEVomE&feature=youtu.be&t=34m13s>
- Casasola, E. y Leoni de León, A. (2017). Transferencia de la polaridad semántica de frases idiomáticas a comentarios de opinión. *Revista Káñina*, 40(3), 65-76.
- Hatzivassiloglou, V. y Wiebe, J. M. (2000, Julio). Effects of adjective orientation and gradability on sentence subjectivity. En Proceedings of the 18th conference on Computational linguistics. New Buswick, NJ. Association for Computational Linguistics. I, 299-305. Recuperado de <http://www.aclweb.org/anthology/C00-1044>
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. San Rafael, California: Morgan & Claypool Publishers.
- Martínez-Cámara, E., García-Cumbreras, M., Villena-Román, J. y García-Morera, J. (2016). TASS 2015 - The Evolution of the Spanish Opinion Mining Systems. *Procesamiento del Lenguaje Natural*, LVI, 33-40. Recuperado de <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/issue/view/218>
- Padró, L. y Stanilovsky, E. (2011). Analizadores Multilingües. *FreeLing. Linguamática*, III(2), 13-20.
- Phan, Dang-Hung y Cao, Tuan-Dung. (2014). Applying skip-gram word estimation and SVM-based classification for opinion mining Vietnamese food places text reviews. En Proceedings of the Fifth Symposium on Information and Communication Technology (pp. 232-239). Recuperado de <http://dl.acm.org/citation.cfm?doid=2676585.2676606>



Rish, I. (2001). An empirical study of the naive Bayes classifier. Recuperado de <http://www.research.ibm.com/people/r/rish/papers/RC22230.pdf>

Wiebe, J., Wilson, T., Bruce, R., Bell, M. y Martin, M. (2006). Learning subjective language. *Learning*, 30(3), 277-308. Recuperado de <http://www.mitpressjournals.org/doi/pdf/10.1162/0891201041850885>



Esta obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)