



Análisis de comentarios y revisiones de usuario usando redes neuronales recurrentes y procesamiento de lenguajes naturales.

Julio Barberan León¹ Fabricio Rivadeneira Zambrano² Jhonny Larrea Plua³

¹Universidad Laica Eloy Alfaro de Manabí, Manta, Manabí, ²Universidad Laica Eloy Alfaro de Manabí, Manta, Manabí, ³Universidad Laica Eloy Alfaro de Manabí, Manta, Manabí

¹barberanjulio@icloud.com, ¹fabricio.rivadeneira@uleam.edu.ec,

¹Johnny.larrea@live.uleam.edu.ec

Recibido: 10/01/2018

Aceptado: 18/01/2018

RESUMEN

La minería de opinión tiene como principal herramienta el procesamiento de lenguajes naturales y mediante técnicas lingüísticas computacionales realizar grandes series de operaciones sobre masivas cantidades de texto para un objetivo común. Esta investigación se centra en la minería de opinión y la utilización del procesamiento de lenguajes naturales y redes neuronales para identificar y extraer información subjetiva de los entornos que se presenten para su análisis, bajo un grado de polaridad comprobar la connotación que esta técnica puede influir en toma de decisiones económicas que puedan estar relacionada con la información. Se hará uso de la red social Facebook, la red social más grande de la cual tiene un promedio de 10.2 millones de comentarios cada 20 minutos y mediante Amazon una de las compañías mas grandes de comercio electrónico del mundo la cual publica contenido comercial alojado en Facebook poder analizar el contenido sintáctico bajo un nivel connotativo de polaridad en correspondencia con la información relacionada en Amazon estableciendo categorías de productos que se ofrecen en Facebook y sus influencias en las revisiones que tienen los mismos productos ofrecidos en Amazon y determinar el grado de relación sobre las preferencias que tienen los usuarios bajo estas tecnologías.

PALABRAS CLAVES: Inteligencia artificial, Redes neuronales, Machine learning.

1. Introducción

La información que genera la sociedad ha mejorado sustancialmente las capacidades para generar y recolectar datos de diversas fuentes. La cantidad tremenda de datos que ha inundado casi todos los aspectos de la vida cotidiana. Este crecimiento explosivo de datos almacenados o transitorios ha generado una necesidad urgente de nuevas técnicas y herramientas automatizadas que pueden ayudar inteligentemente a transformar la gran cantidad de datos en información y conocimientos útiles.

La minería de datos, también conocida como descubrimiento de conocimiento a partir de datos, es la extracción automatizada o conveniente de patrones que representan el conocimiento almacenado o capturado implícitamente en grandes bases de datos, almacenes de datos, paginas web u otros repositorios de información masivos o flujos de datos. [1]

Bajo este precedente existen distintas técnicas algorítmicas que dirigen la minería de datos en distintos hábitos que puedan presentarse los datos, PLN también conocido como procesamiento de lenguajes naturales, establece un conjunto de técnicas de lenguaje computacional que se utilizan en la comunicación diaria de los seres humanos.[2]



Las tecnologías basadas en PLN son ampliamente utilizadas en motores de búsqueda web, dan acceso a la información encerrada en texto que no es estructurada.

En el primer trimestre de 2010, el sitio web Facebook superó a Google para la mayoría de las visitas en la página, lo que confirma un cambio definitivo en cómo la gente consume su tiempo en Internet, lo que indica que la web se ha convertido ahora en un medio social más que un instrumento para la investigación y la información.

Sin embargo, indudablemente indica que los sitios web de redes sociales están satisfaciendo algunos deseos humanos muy básicos en una escala masiva en formas que los motores de búsqueda nunca fueron diseñados para cumplir.

Con antecedentes irrefutables, la aglomeración masiva de datos en nuevos medios informáticos que no solo son social sino de información donde el porcentaje de uso de estas redes ha ido en aumento exponencial surge como respuesta una técnica de la minería de datos, como la minería de redes sociales la cual introduce conceptos básicos y principales algoritmos adecuados para investigar la cantidad masiva de datos de medios sociales bajo teorías y metodologías de diferentes disciplinas como es la informática.

Por contra parte la utilización de la informática como medio de aprendizaje surge con la necesidad de poder comprender que patrones y o preferencias son los que impulsa a las personas a realizar ciertas acciones bajo un cierto contexto la Inteligencia Artificial tiene por objeto el estudio del comportamiento inteligente en las máquinas.

Esta investigación tomó 2 enfoques la utilización del procesamiento de lenguajes naturales en virtud de una aplicación de la inteligencia artificial como son redes neuronales y bajo la utilización de software libre con el objetivo de poder determinar mediante un algoritmo la connotación de opiniones en la red social Facebook y su influencia en las preferencias de compras que tienen cada usuario en Amazon.

2. Minería Web

La minería Web se la utilizo para la recuperación de información mediante la extracción de un conjunto de documentos XML que son generados para cada usuarios de la red. Previamente la minería web se basa en la clasificación del conjunto de documentos por lo general de acuerdo con sus puntuaciones de relevancia.[1] En la siguiente figura se muestra el proceso de captura y transformación de este tipo de información.

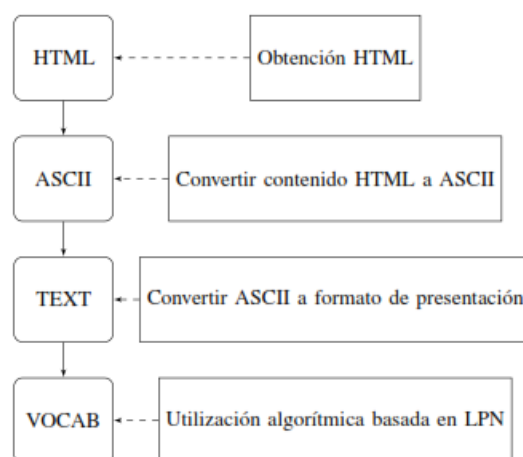


Figura 1: Recuperación contenido HTML



3. Recursos Léxicos

Un recurso léxico, es una colección de palabras o junto a las frases con la información asociada, como las definiciones de parte de la palabra y del sentido.[3] Los recursos léxicos son secundarios a los textos, y usualmente son creados y enriquecidos con la ayuda de textos del mismo modo, una concordancia o acordado como búsqueda mediante algún patrón de características que pueda brindar información sobre el uso de la palabra y que podría ayudar en la preparación de un diccionario para su análisis posterior o entrenamiento. [4]

3.1. Categorización y Etiqueta

Es la técnica dentro del computo morfológico lingüístico que permite dentro de un conjunto previamente almacenado mostrar que etiqueta dentro del lenguaje equivalente a la determinación variable de una oración como un articulo, sea la correspondencia a un sustantivo o un adjetivo [5], cualquier etiqueta con su categorización morfológica respectiva, la siguiente tabla representa la etiqueta para cada variante morfológica dentro del significado de la palabra.

Tabla 1: Categorización léxica y morfológica

Categoría	Etiqueta	Variante morfológica
Categoría léxica	ADJ	Adjetivo
Categoría léxica	ADV	Adverbio
Categoría funcional	CNJ	Conjunción
Categoría léxica	N	Sustantivo
Categoría funcional	DET	Determinador
Categoría funcional	FW	Palabra externa
Categoría funcional	NUM	Número
Categoría léxica	V	Verbo
Categoría léxica	VD	Verbo pasado
Categoría léxica	MOD	Verbo modal
Categoría léxica	P	Pronombre
Categoría léxica	PRO	Preposición

Mediante estas técnicas se podrá encontrar partes significativas que permitan aportar que grados de polaridad en el texto se encuentran y poderlas llevar al campo de la inteligencia artificial.

4. Aprendizaje Profundo

Para la previa utilización de la técnica de red neuronal se utiliza un flujo de procesos de aprendizajes de tipos no supervisados de los cuales aprender bajo un determinado número de entradas x y salidas y dando resultado un conjunto par entrada salida $D = (x_i, y_i)$ N donde D es llamado conjunto de entrenamiento y N es el número de entrenamientos [6].

Las etapas de implementación se aplica de acuerdo a un previo conjunto de entrenamiento de datos que la algorítmica permitirá aprender de sus características para poder mediante las redes neuronales aplicarlos a los datos de evaluación y por la tanto un nivel aceptable de predicción el siguiente gráfico muestra el proceso.

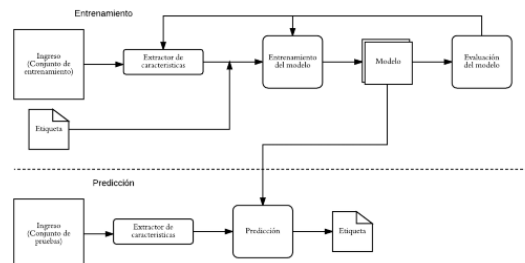


Figura 2: Flujo de procesos de aprendizaje

Partiendo de un conjunto de datos por defecto y estableciendo las características bajo un extractor y un molde de entrenamiento que permita brindar dentro de este proceso una salida o etiqueta que pueda precisar estos conjunto.

5. RED NEURONAL RECURRENTE

La tecnología neural trata de reproducir el proceso de solución de problemas del cerebro, así como los humanos aplican el conocimiento ganado con la experiencia a nuevos problemas o situaciones, una red neuronal toma como ejemplos problemas resueltos para construir un sistema que toma decisiones y realiza clasificaciones [7]. Los problemas adecuados para la solución son aquellos que no tienen solución computacional precisa o que requieren algoritmos muy extensos.

Una red neuronal recurrente tiene como principal objetivo es hacer uso de la información lineal y emplearla de manera iterativa.

En una red neuronal tradicional supone que todas las entradas y salidas son independientes entre sí.

Mediante las salidas que se presenten en el flujo de aprendizaje estas alimentaran la algorítmica neuronal mediante un conjunto de etiquetas que durante el proceso evolutivo permitirán mejorar con mayor precisión el resultado de predicción para cada par entrada salida que se presente.

6. Metodología

Se empleo un modelo empírico del proceso que se ha llevado para lograr los objetivos de esta investigación.

6.1. Recursos

Como lenguaje de programación se utilizo Python con interprete la versión 2.7.10, el siguiente cuadro muestra los recursos de software que se emplearon.



Tabla 2: Bibliotecas empleadas

Librerías	textbfVersión	Descripción
BeautifulSoup	4.4.	Extracción de datos archivos html y XML.
re	7.2.	Utilización de expresiones regulares.
urllib2	20.6.	Acceso a componentes html
json	18.2.	Manejo de datos en formato json
datetime	8.1.	Utilización para formato de fecha
csv	13.1.	Administración de archivos csv
pandas	20.3	Permite realización de operaciones bajo formato csv
numpy	1.13.	Carga de operaciones matemáticas
request	2.18.	Peticiones de carga url mediante páginas web y XML
itertools	9.7.	Serie de bloques de construcción de iteradores
tensorflow	1.2.	Biblioteca de código abierto para aprendizaje automático

6.2. Propuesta

Se emplearon 3 procesos principales para la realización de la propuesta, mediante el siguiente gráfico de procesos se ilustra este procedimiento:

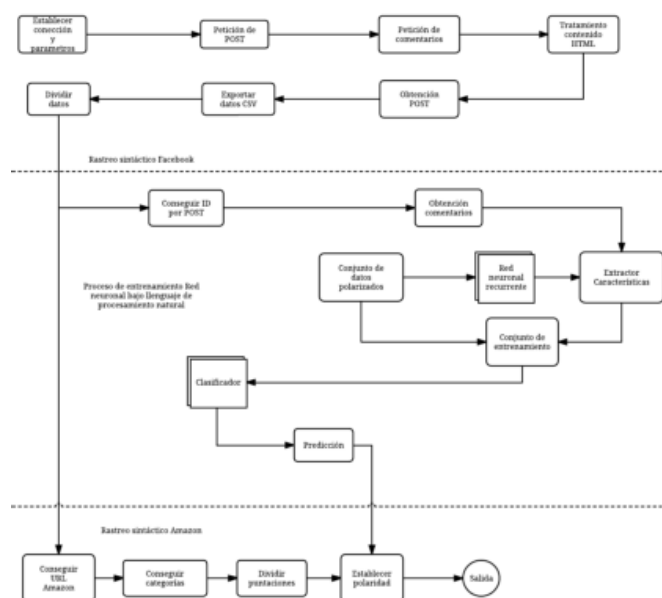


Figura 3: Flujo de procesos

La primera etapa del proceso se realiza mediante un rastreo sintáctico en la página oficial de Amazon en Facebook de la cual se extrajeron 121440 publicaciones desde el año 2012 al presente, Amazon publica dos tipos de publicaciones informativos respecto a los servicios de la empresa, y ofertas de productos de la cual aparece una breve descripción del producto en el encabezado de la publicación y junto una url que dirige en la página de compras de Amazon, para cada publicación se tiene un promedio de 77 comentarios por publicación.

La segunda etapa se basa en el tratamiento de los comentarios previamente categorizados en productos



y la categoría que Amazon da dentro su página de compras y para cada uno de los departamentos que conforman la entidad como se puede apreciar en el siguiente cuadro. Cabe destacar que dentro de las categorías obtenidas dos no pertenecen a algún departamento en específico dentro de Amazon ya que estos se promocionan a nivel aleatorio dentro de sus departamentos.

Tabla 3: Categorización productos Amazon

Categoría	Departamento
lawn and garden	Home, Garden and Tools
electronics	Electronics, Computers
automotive	Automotive and Industrial
wireless	Electronics, Computers
sports	Sports and Outdoors
music	Amazon Music
kitchen	Home, Garden and Tools
jewelry	Handmade
instant video	Amazon Video
services	Home Services
home and garden	Home, Garden and Tools
furniture	Handmade
digital	Electronics, Computers
deal	-
boost	Movies, Music and Games
beauty	Handmade
books	Kindle E-readers and Books
promotion	

Se utilizara un procedimiento de redes neuronal recurrente utilizando técnicas de procesamiento natural de lenguaje como modelo de extracción de características con el objetivo de poder clasificar el texto en 3 tipos de polaridad positiva, negativa, neutral.

La tercera etapa empieza mediante un rastreo sintáctico en la página de compras de Amazon mediante los datos que se extrajeron de Facebook cómo su url con buscar todas las revisiones que hayan hecho los clientes, y poder hacer el mismo proceso que se ha llevado a cabo en la etapa número dos con el objetivo de poder analizar la polaridad que existe para probar algún tipo de indicio que afecte a las compras que se realizan en Amazon.

1. Rastreo sintáctico: Las dos plataformas Web comparten cierta correspondiente información acerca de los productos que lanza por parte de la página de compras y la red social, los campos que comparten y hacen único cada dato obtenido mediante este proceso. El rastreo sintáctico de Facebook este compuesto por 2 procesos, el proceso de captura de publicaciones, comentarios y el tratamiento del contenido HTML que este se emplea para conseguir los campos que producirán el posterior rastreo sintáctico a la plataforma de Amazon como se puede ver en la siguiente ilustración:

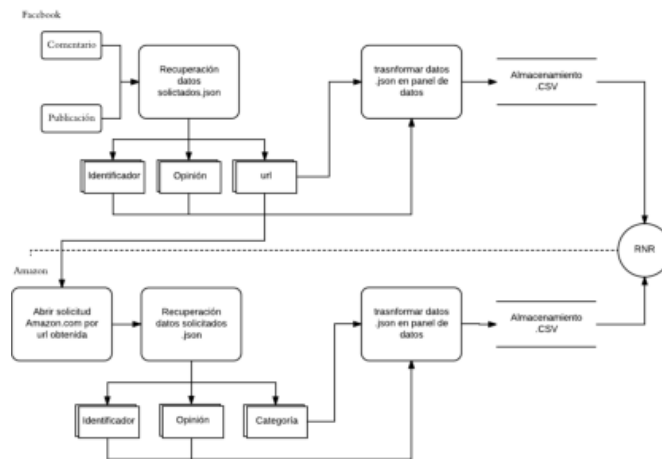


Figura 4: Flujo de procesos rastreo sintáctico

La obtención de datos pasa paralelamente, en la fase de extracción enviando los parámetros como son su identificador, el contenido de opinión y su base url necesarios para la obtención de cada contenido sintáctico a evaluar y extracción de contenido de Amazon compras, la siguiente fase pasa la transformación de estos datos a un tipo de datos manejable para el computador y accesible para su manipulación y almacenamiento.

La plataforma de Amazon compras ofrece a los usuarios un medio de calificación de compras y experiencia del producto que ofertan con categorías de 1 a 5 estrellas dependiendo del nivel de etiqueta del comentario que publica cada usuario.

- Publicaciones: Identificador de publicación, url
- Comentarios: Identificador de comentario, Identificador de publicación

6.3. Redes neuronal recurrente

Como principal biblioteca se recurrió a tensorflow el cual es una biblioteca de código abierto creado por Google y utilizable para cualquier aspecto que tenga que ver con aprendizaje automático.

1) Conjunto de datos y pre procesamiento:

- Primer proceso es la carga de 2 archivos con contenido polarizado de tipo connotativo positivo y negativo para su entrenamiento posterior, para esto se acudió a un conjunto de datos recolectados por la Universidad de Cornell.
- Segundo proceso es la construcción de un vocabulario el cual es pasado al modelo de tokenización bajo un conteo de la frecuencia de la palabra y bajo un clasificador bayesiano [8]:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (1)$$

Se utiliza la tokenización y (1) como medio para separar previamente las oraciones, para cada palabra que se busca si está se encuentra previamente en la clase entrenada, al cumplir esta afirmación se añade la probabilidad total entre el número de veces que esta repite en el conjunto de datos.

- Tercer proceso es la carga de comentarios y el tratamiento de las cadenas de texto a nivel de características quitando contenido de estos pues rebajarían la precisión del conjunto de entrenamiento al ser inmediatamente dependiente de los datos de entrenamiento estos deberán tener un grado de proximidad al contexto que se se hacen pruebas.



2) Modelo: Las primeras capas añaden bajo el proceso de tokenización palabras sobre los vectores. La capa que prosigue hacer el proceso de convulsión sobre los vectores de palabra utilizando diferentes tamaños de filtro en la siguiente imagen se puede apreciar el modelo de transición expuesto.

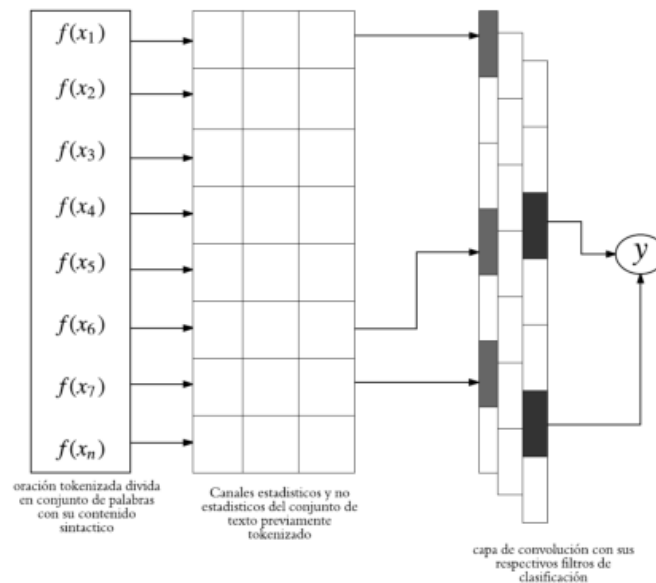


Figura 5: Modelo red neuronal

3) Implementación: Los parámetros que se utilizan para las instancias respectivas de la clase que contiene el modelo:

- Número de clases o etiquetas de salidas que para este caso serán 3 positivo, negativo, neutro
- Tamaño del vocabulario el cual es necesario para definir la capa embebida o incrustada la cual posee forma [tamaño del vocabulario, tamaño de incrustación]
- Tamaño de incrustación
- Tamaño de filtros para los números de palabras que se requiera utilizar dentro de las capas de la red neuronal tanto de incrustación como para la salida

La primera capa que definimos es la capa de incrustación, la cual es introducida dentro de los vectores de división de oraciones.

Por convención tensorflow espera un canal bajo 4 dimensiones correspondientes al lote, anchura, altura y canal. El resultado de la incrustación no contiene la dimensión del canal, por lo que se añade manualmente, dejando con una capa de forma [parámetro puesto a nulo, secuencia de tamaño, tamaño de incrustación].

Cada proceso de convulsión produce diferentes punto de flexión o tensores de diferentes tamaño los cuales deben ser iterados bajo el mismo objeto y al mismo tiempo definir y crear una capa para cada uno de estos. Usando un vector de características mediante max-pooling se puede producir predicciones mediante multiplicación matricial y seleccionando la clase con la puntuación más alta. Bajo la pérdida de precisión se puede definir como la medida de error que hace el proceso de red neuronal al tratar de minimizar lo cual se basa en la pérdida de entropía cruzada.

4) Implementación y sesiones: En tensorflow, una sesión es el entorno en el que está ejecutando operaciones de visualización y contiene estado sobre cualquier tipo de datos estructurados como listas, colas y diccionarios. Una sesión funciona bajo solo un gráfico. Al no especificar explícitamente una sesión, se crean variables y operaciones, utilizando la sesión predeterminada actual creada por tensorflow, una



característica importante que posee esta librería es el concepto de resúmenes, permiten hacer un seguimiento y visualizar varias cantidades durante el entrenamiento y la evaluación, además de permitir un seguimiento de las cantidades más complejas.

5) Punto de control: Un componente muy importante dentro de la fase de entrenamiento y seguimiento de la eficiencia del modelo de red neuronal son los puntos de control el cual dentro de la biblioteca empleada como tensorflow permite guardar los parámetros del modelo implementado para restaurarlos más adelante. Los puntos de control se pueden utilizar para continuar con el entrenamiento en un punto posterior, o para escoger el mejor ajuste de parámetros usando retrocesos tempranos.

6) Tiempo de entrenamiento: Como se puede apreciar en la siguiente gráfica se recolecto los puntajes arrojados durante la fase de aprendizaje.



Figura 6: Tiempo de aprendizaje

Bajo la tasa aprendida promedia del 0.71 en conjunto con el tiempo empleado de 6 minutos con 55 segundos obtuvo un porcentaje con un promedio constante del 0.94 bajo el conjunto de datos entrenados

7. RESULTADO

En la presente sección se presentan los datos recopilados tras la propuesta presentada en la sección anterior, las siguientes gráficas muestran estos resultados.



Figura 7: Comentarios extraídos de la plataforma Facebook sitio oficial Amazon

Se denota que un total 251563 comentarios extraídos en 18 categorías obtenidas por un total 10000 publicaciones realizadas en la página oficial de Amazon en Facebook posee un gran volumen de los comentarios entre las promociones realizadas con un aproximado del 84 %, Amazon Books correspondiente



a un aproximado del 6 % y Amazon Kitchen con un aproximado al 6 % los demás porcentajes se reparten en las de más categorías previamente mostradas.



Figura 8: Comentarios extraídos de la plataforma Amazon compras

Dentro de la página de compras Amazon se obtuvieron un total de 24026542 comentarios de los cuales aproximadamente el 88 % corresponden a promociones, un aproximado del 2 % a la línea Amazon Books y un aproximado del 6 % a la línea Amazon Kitchen los demás porcentajes se reparten en las de más categorías previamente mostradas.

7.1. Resultados Análisis connotativo

La implementación de la red neuronal basado en el conjunto de entrenamiento escogido arrojo los siguientes resultados:

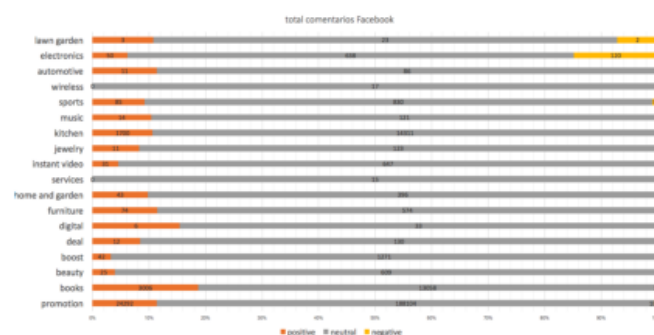


Figura 9: Análisis connotativo Facebook

Dentro de las categorías más representativas de los comentarios extraídos para la plataforma de Facebook revela que tiene una gran influencia la clase neutral seguido de la positiva y por último negativa la forma de evaluación para cada salida fue basarse en el puntaje de 0 a 1 siendo la clase negativa, positiva o neutral con un puntaje mayor al 0.5, los resultados para la línea de promociones bajo la clase neutral con un promedio del 88 %, positiva con un promedio del 11 % y negativa con un promedio del 1 % para la línea de Amazon Books revela un promedio de connotación neutral del 71 %, positiva del 19 %, negativa del 10 %, la línea de Amazon Kitchen revela un promedio de connotación neutral del 90 %, positiva del 10 %

Dentro de las categorías más representativas de los comentarios extraídos para la plataforma de Amazon revela que tiene una gran influencia la clase positiva seguido de la neutral y por último negativa, los resultados para la línea de promociones bajo los resultado de la clase neutral revela un promedio del 9 %, positiva con un promedio del 75 % y negativa con un promedio del 14 % para la línea de Amazon



Figura 10: Análisis connotativo Amazon

Books revela un promedio de connotación neutral del 48 %,positiva del 27 %, negativa del 25 %, la línea de Amazon Kitchen revela un promedio de connotación neutral del 85 %, positiva del 6 % y neutral con el 9 %.

Los resultados muestran que la clase connotativa neutral posee un mayor porcentaje que las demás clases connotativas para la plataforma Facebook mientras que la plataforma Amazon tiene un mayor porcentaje para la clase connotativa positiva seguida de las demás clases connotativas.

7.2. Frecuencias lexicográficas

Realizando un análisis sobre las frecuencias lexicográficas del contenido neutral se podrá evaluar el grado de proximidad a un sentido negativo o positivo, el conjunto de datos a evaluar se procederá a escoger una lista de 50 palabras más usadas dentro de este conjunto de datos como se aprecia en el siguiente gráfico de barras.

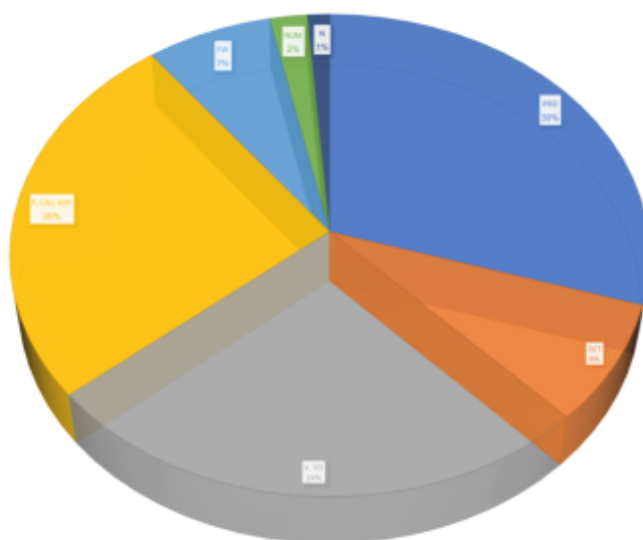


Figura 11: Frecuencias lexicográficas

Dentro del análisis se revela que hay una gran influencia dentro del contenido léxico en cuanto a categoría funcional, el recurso léxico que revela un grado de proximidad a una connotación estaría compuesto por



verbos y adjetivos de los cuales tienen un promedio del 26 % mostrando que el contenido ciertamente no contiene un sentido negativo y no positivo.

Realizando un diagrama de correlación se puede apreciar la relación dentro de los resultados obtenidos para las líneas de Amazon más representativos a nivel de opinión como Amazon Promotions, Books y Kitchen.

Tabla 4: Categorización productos Amazon

Categoría	Amazon (+)	Amazon (-)	Facebook (+)	Facebook (-)
Promotion	0.84	0.16	0.99	0.01
Books	0.75	0.25	0.90	0.10
Kitchen	0.91	0.09	1	0

Los resultados muestran una correlación fuerte dentro de las connotaciones positivas de Amazon compras y la red social Facebook con un coeficiente de correlación mayor del 0.7 mientras que las connotaciones negativas tienden a tener una correlación fuerte de la misma manera que lo hace las connotaciones positivas mediante el siguiente gráfico se ilustra este hecho:

	Amazon (+)	Amazon (-)	Facebook (+)	Facebook (-)
Amazon (+)	1	-1	0,93	-1
Amazon (-)	-1	1	-0,93	0,93
Facebook (+)	0,93	-0,93	1	-1
Facebook (-)	-0,93	0,93	-1	1

Figura 12: Resultados correlación

Bajo estos resultados indica que la red social influye en las preferencias que los consumidores pueden tener a ciertos tipos de productos que se pueden promocionar en esta plataforma, Amazon promociona distintos tipos de publicaciones que se pueden realizar en Facebook ya que esto hace un indicador de los puntajes que Amazon emite a sus productos mediante los comentarios que se publican a través de Amazon ya que estos comentarios deben ser hechos después de la compras de productos etiquetados con un nivel de satisfacción de la compra entre una o cinco estrellas.

8. Conclusiones

Con el desarrollo, implementación y presentación de los resultados obtenidos de este trabajo se resalta la suma importancia del conjunto de técnicas computacionales propuesto para las formas de marketing y el desarrollo de organizaciones y empresas.

Se realizó la extracción de comentarios en la red social y en Amazon utilizando el rastreo Web como medio computacional y posteriormente utilizar técnicas que permitan hacer un análisis connotativo de manera ordenada y metodológicamente siguiendo un proceso el cual arroja un resultado que puede ser



usado e implementado por departamentos de marketing que ofrezcan o tengan vinculado su sector a exposición de productos a cualquier página de compras en línea que se encuentran en Internet.

Como resumen se pueden extraer las siguientes conclusiones:

- Los resultados presentan un promedio de 0.93 % de polaridad positiva neutral entre todo el conjunto de datos los cuales corresponden a las 2 plataformas evaluadas.
- Los resultados presentan un grado de relación fuerte con la influencia que tiene el contenido u opinión que se presenta en la red social Facebook en conjunto con los puntajes y grados de connotación que los consumidores realizan en la página de compra de Amazon, haciendo que los puntajes que Amazon realiza sobre la satisfacción que cada cliente ha tenido durante la compra se refleja del lado de la red social evaluada.
- Se denota que hay un incentivo sobre el efecto que tiene las redes sociales en el entorno del usuario en este caso compras en línea ya que el tiempo promedio de conexión por persona es de aproximadamente 5 horas y 30 minutos al día según un estudio de GlobalWebIndex haciendo inevitable que Amazon aproveche para lanzar publicidad sobre esta plataforma.

Referencias

- [1] Bing Liu. *Web data mining: exploring hyperlinks, contents, and usage data*. Springer Science & Business Media, 2007.
- [2] Yoav Goldberg. "Neural network methods for natural language processing". En: *Synthesis Lectures on Human Language Technologies* 10.1 (2017), págs. 1-309.
- [3] Steven Bird, Ewan Klein y Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. .O'Reilly Media, Inc.", 2009.
- [4] Ryan Mitchell. *Web scraping with Python: collecting data from the modern web*. .O'Reilly Media, Inc.", 2015.
- [5] Ruslan Mitkov. *The Oxford handbook of computational linguistics*. Oxford University Press, 2005.
- [6] Kevin Murphy. *Machine Learning A Probabilistic Perspective*. 1rd. Ed. MIT Press, 2012.
- [7] Pedro Ponce Cruz. *Artificial intelligence with application to engineering*. 1rd. Ed. Alfaomega, 2010.
- [8] Nils J. Nilsson. *Artificial intelligence: a new synthesis*. Elsevier, 1998.