

CIENTÍFICA Y TECNOLÓGICA

In Geni 

eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

Análisis de patrones de búsqueda de los usuarios en un sitio web de comercio electrónico usando Web Analytics

Analysis of search patterns of users on an e-commerce website using Web Analytics

Ariosto Vicuña

<https://orcid.org/0000-0002-1459-1759>

avicuna@uteq.edu.ec

Universidad Técnica Estatal de Quevedo. Facultad de Ciencias de la Ingeniería. Quevedo. Ecuador

Andrés Florencia

<https://orcid.org/0000-0002-7920-8095>

andres.florencia2013@uteq.edu.ec

Universidad Técnica Estatal de Quevedo-FCI. Ecuador

Jean C. Bazurto

<https://orcid.org/0000-0002-7786-1840>

jean.bazurto@uteq.edu.ec

Universidad Técnica Estatal de Quevedo-FCI. Ecuador

Orlando Erazo

<http://orcid.org/0000-0001-5642-9920>

oerazo@uteq.edu.ec

Universidad Técnica Estatal de Quevedo. Facultad de Ciencias de la Ingeniería. Quevedo. Ecuador

RESUMEN

La masificación del uso del Internet y la frecuencia con la que las personas acceden a sitios web, particularmente los dedicados al comercio electrónico, han dado origen a una necesidad de conocer la forma en que los usuarios interactúan con ellos. Esta situación conduce a que diariamente se genere una cantidad enorme de información que puede convertirse en

Ingenio

Enero - Diciembre Vol. 2 Núm. 1 (2019)

<https://revistas.uteq.edu.ec/index.php/ingenio>

eISSN: 2697-3642

ingenio@uteq.edu.ec

Recepción: 4 de febrero 2018

Aprobación: 8 septiembre 2018

Pág. 1-11

Esta obra está bajo una Creative Commons Atribución/Reconocimiento-NoComercial-CompartirIgual 4.0 Licencia Pública Internacional — CC BY-NC-SA 4.0

<https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.es>

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

un factor clave para las empresas siempre y cuando sea analizada adecuadamente. Es por ello que, este trabajo se enfoca al análisis de patrones de búsquedas realizadas en sitios web de comercio electrónico basado en Web Analytics y Data Science. Para ello, se ha generado un conjunto de datos al que se le ha aplicado el algoritmo seleccionado para encontrar reglas de asociación. Particularmente, se ha aplicado la herramienta Google Analytics en el estudio. Como resultado se han encontrado patrones de búsquedas que podrían realizar los usuarios de un sitio web de comercio electrónico.

Palabras clave: ciencia de datos; Web Analytics; reglas de asociación; a priori; comercio electrónico.

ABSTRACT

The widespread use of the Internet and the frequency with which people access websites, particularly those ones dedicated to electronic commerce, have given rise to a need to know how users interact with them. This situation leads to a huge amount of information being generated daily that can become a key factor for companies as long as it is analyzed properly. That is why, this work focuses on the analysis of search patterns performed on e-commerce websites based on Web Analytics and Data Science. With this aim, a data set has been generated to which the selected algorithm has been applied to find association rules. In particular, the Google Analytics tool has been applied in the study. As a result, patterns of searches that users of an e-commerce website could perform were found.

Keywords: Data Science; Web Analytics; association rules; a priori; E-commerce.

INTRODUCCIÓN

El creciente uso del Internet y la frecuencia con la que las personas recurren a él han generado una necesidad de conocer la forma en que los usuarios interactúan con los sitios web (Fang, 2007). De hecho, el uso de herramientas para el análisis web es un factor importante para las empresas que desean incrementar sus ventas y ofrecer un mejor servicio en su sitio web. Para ello, las empresas, a través de información recopilada en sus portales web, buscan observar, comprender y controlar mejor varios aspectos del comportamiento de sus usuarios. Sin embargo, es tal la cantidad de información generada a diario que el uso de herramientas para el análisis de datos se convierte en un factor clave para las empresas.

El comercio electrónico es precisamente un área en la cual las empresas requieren analizar grandes volúmenes de datos. El análisis de datos generados en los sitios web permite medir las interacciones del usuario con productos durante toda su experiencia de compra. Ejemplos de esto son las impresiones del

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

producto, los clics en el producto, la consulta de información del producto, la inclusión de un producto en un carrito de la compra, el inicio del proceso de pago, y las transacciones o las devoluciones. La obtención de datos como éstos que permitan efectuar el análisis necesario puede lograrse a través de *Google Analytics*.

Google Analytics es una herramienta para realizar Análisis Web (*Web Analytics*) ofrecida por Google (Mejía, 2013) (Anaytics, 2016). Su aplicación puede ir desde hacer que un sitio web sea más robusto (Gaur, Singh, & Kumar, 2016) hasta, por ejemplo, descubrir campañas maliciosas (Starov, Zhou, Zhang, Miramirkhani, & Nikiforakis, 2018). Para ello, esta herramienta ofrece reportes que proporcionan información sobre el origen de los visitantes, las páginas visitadas, el tiempo de permanencia en cada página, la profundidad del sitio explorado, la fecha de finalización de las visitas y su destino (Ballesteros & Ballesteros, 2007). Una vez obtenidos los datos será necesario elegir una técnica para el análisis correspondiente. Precisamente, una de las utilidades que se le puede dar a Google Analytics se encuentra en el comercio electrónico. Por ello, este trabajo tiene como fin realizar un análisis de un conjunto de transacciones generadas de comercio electrónico en un sitio web (como por ejemplo Amazon). Este análisis apunta a encontrar la relación que existe entre las búsquedas de productos previas a una compra por parte de los usuarios, de tal manera que permita mejorar la forma en que se promocionan los mismos.

Para el efecto, luego de proporcionar el fundamento teórico considerado necesario, se describen las etapas sugeridas a considerar para realizar un análisis del tipo en mención. Luego, estas etapas son aplicadas en un primer estudio con el conjunto de datos creado para el efecto. Los resultados obtenidos son presentados y discutidos posteriormente. Estos resultados constituyen una evidencia del provecho que se podría obtener al usar Google Analytics en la obtención de patrones de las búsquedas que realizan los usuarios de sitios web de comercio electrónico. El comercio electrónico consiste en transacciones de compra y venta de productos o servicios a través de un medio electrónico. La práctica del comercio electrónico existe desde 1965 cuando los consumidores pudieron retirar dinero de los cajeros automáticos y realizar compras en terminales de venta usando tarjetas de crédito (Molla & Licker, 2001). El comercio electrónico es “el pago electrónico de un bien adquirido a través de Internet” (Ballesteros & Ballesteros, 2007). Generalmente los medios de pago de esta tecnología incluyen tarjetas de crédito o débito, cheques virtuales, tarjetas prepago, entre otros (Ballesteros & Ballesteros, 2007). Existen varios tipos de comercio electrónico, entre los cuales se encuentra el *Business-to-Consumer* (B2C).

B2C es el comercio de las empresas con el cliente que se realiza en las tiendas online que distribuyen al consumidor final. B2C es un sitio en la web por el cual un consumidor puede comprar un producto o servicio (Ranganathan & Ganapathy, 2002). Esto incluye sitios de compra online manejados por infomediarios, comerciantes *brick-and-mortar* y tiendas virtuales (Ranganathan & Ganapathy, 2002). Los datos generados por las transacciones realizadas en sitios B2C proporcionan información útil que puede ser utilizada para mejorar el comercio electrónico mediante un análisis de datos. El análisis de datos web es la técnica implementada en sitios web con el fin de medir, recopilar, analizar y generar informes de datos para comprender y optimizar el uso de los mismos (Phippen, Sheppard, & Furnell, 2004) (Gonçalves & Ramasco, 2008). La forma en la que se obtienen estos datos puede lograrse a través de herramientas como *Google Analytics* (Mejía, 2013) (Anaytics, 2016).

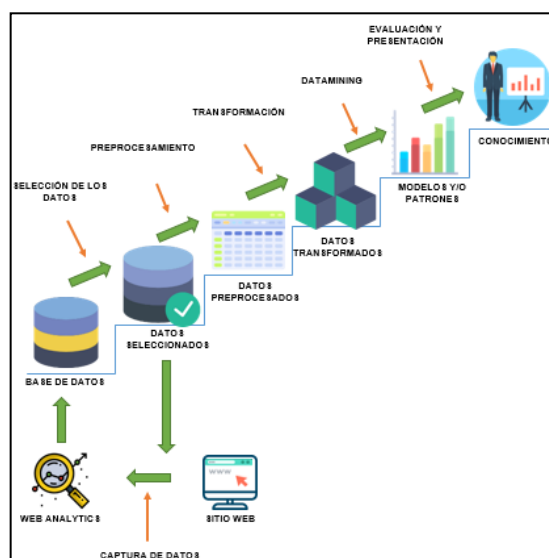


Figura 1. Etapas propuestas para minería de datos

Google Analytics permite recopilar información relevante de un sitio web de forma sencilla. Para ello basta insertar un segmento de código al sitio web para obtener datos de usuarios como las páginas visitadas, cantidad de clics, tiempo de permanencia en el sitio, origen de los visitantes, entre otros datos. Google Analytics agrupa todos estos datos en dos grupos: métricas y dimensiones. Las métricas son mediciones cuantitativas (por ejemplo clics/sesión), mientras que las dimensiones son mediciones cualitativas (por ejemplo, navegador) (Mejía, 2013). Con ellas se pueden generar datos transaccionales de los sitios web, los cuales permiten descubrir conocimiento o patrones realizando tareas de análisis de datos como las reglas de asociación.

Las reglas de asociación ayudan en el descubrimiento de relaciones entre atributos de un conjunto de datos que superan unos determinados umbrales. Ellas son una de las técnicas más conocidas y utilizadas en el análisis de datos por su aplicación en diferentes áreas temáticas (Phippen, Sheppard, & Furnell, 2004) como la salud y la educación. Sin embargo, su aplicación más típica se encuentra en los sistemas de comercio electrónico para informar sobre las preferencias de compra de los clientes (Morales, Soto, & Martínez, 2005). Un ejemplo de esto es el estudio de los patrones de productos vendidos de un supermercado (Gonçalves & Ramasco, 2008).

Un algoritmo muy usado para encontrar reglas de asociación es *a priori*, mismo que será aplicado en este estudio a un conjunto de datos de transacciones. Para lograrlo, se necesitan una serie de pasos, tales como una metodología adaptable a cualquier caso de estudio similar. En la siguiente sección se propone una serie de pasos aplicados en este trabajo.

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

En la aplicación de cualquier técnica de minería de datos es necesario una guía o pasos a seguir que permita cumplir los objetivos de negocio propuestos. A continuación se describen cada uno de los pasos o etapas aplicadas (Figura 1), procurando que los mismos sean adaptables a trabajos similares.

1.- *Captura de datos*: Este paso consiste en obtener datos de un sitio web haciendo uso de herramientas de web analytics, como Google Analytics. Estas herramientas permiten obtener todo el tráfico de un sitio.

2.- *Almacenamiento de datos*: Luego de la captura de los datos será necesario el almacenamiento de los mismos. Para ello se hará uso de gestores de bases de datos, como PostgreSQL. Los datos generados por las herramientas de análisis web se obtienen en un formato de tabla (por ejemplo, de MS Excel) y pueden ser importados a una tabla del gestor de base de datos.

3.- *Selección de los datos*: En este paso se empieza el proceso de la minería de datos o la extracción de conocimiento. Aquí se seleccionarán todos aquellos datos que ayudarán al cumplimiento del objetivo propuesto. Cabe recalcar que los pasos del 1 al 3 pueden ser parte de un proceso iterativo debido a que se pueden necesitar datos adicionales o seleccionar los mejores atributos.

4.- *Pre-procesamiento*: Este punto trata sobre la limpieza y el pre-procesamiento de datos. El pre-procesamiento consta de operaciones básicas tales como la eliminación de "ruido" (valores faltantes, datos duplicados, etc.).

5.- *Transformación*: Luego de pre-procesar los datos será necesario transformar las variables que ayudarán a la aplicación de las técnicas de minería de datos, debido a que muchos atributos son cadenas de caracteres y afectan a ciertos algoritmos de minería. Entre las técnicas de transformación que podrían usarse se tienen: balanceo de clases, reducción de datos y selección de atributos. Todo este proceso se puede traducir en la reducción de la dimensionalidad del conjunto de datos.

6.- *Minería de datos*: Luego de obtener un conjunto de datos "limpio" y de haber definidos los objetivos del análisis a realizar, se llega al punto de la elección y aplicación de las técnicas o algoritmos que permitan obtener patrones o encontrar modelos de predicción. Este punto puede ser iterativo con el paso 5 (Transformación).

7.- *Evaluación e interpretación*: Una vez aplicadas las técnicas del paso 6 se procede a interpretar los resultados obtenidos. Este paso depende de la experiencia de la persona que aplique la minería a los datos. Por ello, puede resultar conveniente emplear varias técnicas y comparar los resultados para determinar cuál resultó ser la mejor.

MATERIALES Y MÉTODOS

Las etapas para minería de datos antes propuestas fueron aplicadas a un conjunto de datos con las métricas y dimensiones ofrecidas por Web Analytics. Los datos necesarios fueron almacenados en una base de datos para posteriormente aplicar técnicas de minería de datos con el fin de encontrar patrones de búsqueda. La base de datos utilizada en el presente trabajo estuvo compuesta de 5 millones de registros. Los datos

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

empleados fueron generados a partir de datos simulados tomando como referencia a Amazon y Google Analytics.

Haciendo uso de la herramienta de análisis web, Google Analytics, se capturaron los datos necesarios para este análisis. Para obtener estos datos se usó también la herramienta de análisis estadístico R Project, donde se estableció una conexión con Google Analytics. Una vez obtenidos los datos se almacenaron como se describe en el siguiente paso.

Luego de la captura de los datos se procedió a la creación de una base de datos que permita el registro de los mismos. Se utilizó el gestor de base de datos PostgreSQL. La herramienta R Project también permitió la conexión con este gestor de base de datos, de tal manera que se realizaron los registros desde R, y de igual manera la lectura para su posterior análisis.

Luego, desde la consola de R Project, se realizó la conexión para extraer los datos relevantes que describen las búsquedas que realizan los usuarios en el sitio web. Para ello se utilizaron funciones como *dbDriver("PostgreSQL")* que permite crear la instancia del driver de conexión con PostgreSQL, *dbConnect()* que permite conectarse a una base de datos, *dbGetQuery()* que permite la ejecución de consultas a una base de datos de PostgreSQL (ver, por ejemplo (Zumel & Mount, 2014), para más detalles). De esta manera se obtuvo el conjunto de datos principal al cual se le aplicaron técnicas de limpieza de datos, explicadas en los siguientes pasos. Las dimensiones y métricas tomadas inicialmente se muestran en la *Tabla 1*. Finalmente los atributos elegidos para el análisis se muestran en la *Tabla II*.

Tabla I: Métricas y dimensiones iniciales

DIMENSIONES	MÉTRICAS
keyword	newUsers
searchKeyword	sessionDuration
searchCategory	pageViews
searchKeywordRefinement	searchUniques
country	

Tabla II: Métricas y dimensiones finales

DIMENSIONES	MÉTRICAS
searchKeywordRefinement	searchUniques
searchKeyword	

Con el fin de obtener mejores resultados al aplicar las técnicas de minería de datos, fue necesario aplicar técnicas de limpieza de los mismos, tales como: eliminación o imputación de valores faltantes, eliminación de datos duplicados, entre otras.



Según Phippen (Phippen, Sheppard, & Furnell, 2004) existen muchos algoritmos que trabajan de forma más eficiente con datos numéricos. Para ello se recomienda transformar ciertos atributos de tipo cadena de texto a numéricos, tales como discretización y normalización.

Luego de obtener un conjunto de datos limpio y de haber establecido los objetivos del análisis a realizar, se procedió a escoger la técnica o algoritmo que más se ajustaba a la necesidad del análisis, el cual fue el algoritmo *a priori* de la librería *arules* de R Project. Este algoritmo necesita dos columnas principales: un identificador por transacción, al cual se le estableció la métrica *newUsers*, y la transacción que corresponde a los ítems, al cual se le estableció la dimensión *searchKeywordRefinement*. Otros parámetros requeridos son: (1) el soporte mínimo (*support*) el cual fue establecido como 0.01 indicando así que una *regla* será válida si por lo menos aparece en un 1% de las transacciones, y (2) con un mínimo de confianza (*confidence*) de 0.5, el cual indica la probabilidad mínima de *Y* en una transacción. Se eligió este algoritmo porque permite encontrar asociaciones entre elementos en grandes conjuntos de datos.

Después de obtener las reglas de asociación se procede a su respectiva interpretación, esperando a que cumpla con las expectativas del análisis. Estos resultados serán evaluados más a fondo en la sección siguiente.

RESULTADOS

Luego de realizar todo el proceso de análisis de los datos para la obtención de patrones de búsquedas se obtuvieron los siguientes resultados resumidos en la *Tabla III*:

En la *Figura 2* se puede apreciar de forma gráfica las reglas de asociación encontradas en forma de grafo, generadas mediante la función *plot* y el método *“graph”*:

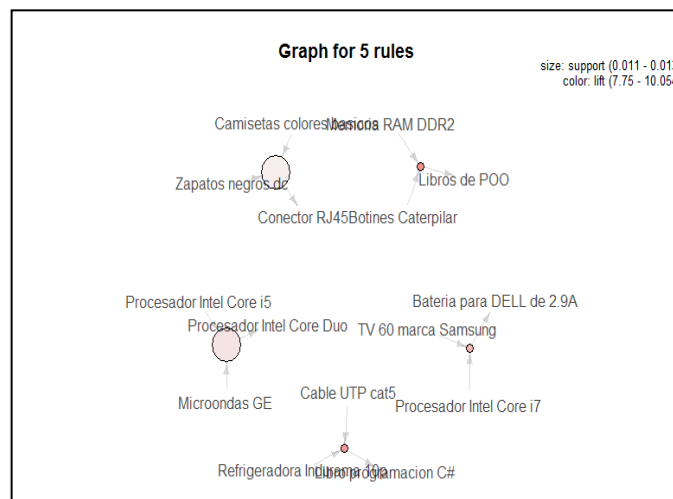


Figura 2: Reglas de asociación en forma de grafo



Tabla III: Reglas obtenidas. Primeras cinco reglas de asociación obtenidas con un $\text{minSupport}=0.01$ y $\text{minConfidence}=0.5$ ordenadas por el nivel de confianza.

ANTECEDENTE (LHS)	CONSECUENTE (RHS)	SUPP	CONF
{camisetas colores básicos, zapatos negros dc}	{Conector RJ45}	0,0134	1
{microondas ge, procesador intel core i5}	{Procesador Intel Core Duo}	0,0134	1
{cable utp cat5, refrigeradora indurama 10p}	{Libro programación C#}	0,0107	1
{botines caterpillar, memoria ram ddr2}	{Libro de POO}	0,0107	1
{procesador intel core i7, tv 60 marca samsung}	{Batería para DELL de 2.9A}	0,0107	1

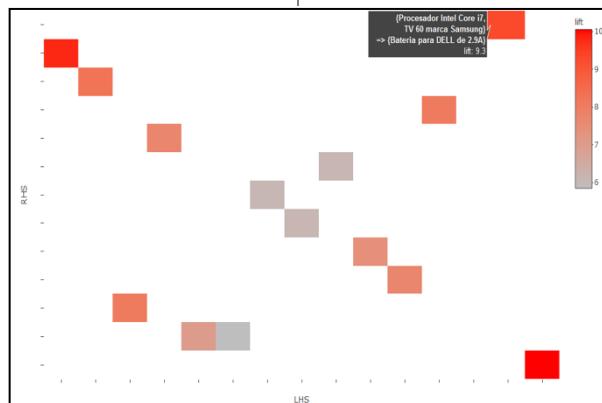


Figura 3: Visualización de 15 reglas de asociación de forma dinámica

La *Figura 3* permite la visualización de forma dinámica de las reglas con su antecedente, consecuente, confianza, soporte e importancia. No obstante, para la visualización ideal de la información contenida en esta imagen es necesario disponer de la herramienta computacional que lo permita.

Las *Figuras 4* y *5* muestran de forma proporcional la búsqueda refinada de productos (*searchKeywordRefinement*) en base a la búsqueda por palabra clave (*searchKeyword*).

DISCUSIÓN

De las reglas encontradas se tomaron las cinco primeras ordenadas por el nivel de confianza (*confidence*), en donde se puede interpretar que las personas que buscan inicialmente {Camisetas colores básicos, zapatos negros dc} proceden a buscar {Conector RJ45} con un nivel de confianza de 100% y soporte de 1%. De igual manera en la segunda regla donde los usuarios que buscan {Microondas GE, Procesador Intel Core i5} tienden a buscar {Procesador Intel Core Duo}.

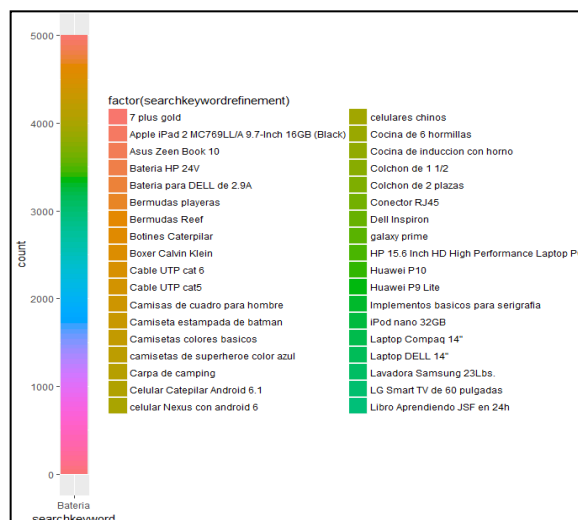


Figura 4: Gráfico proporcional de búsquedas refinadas de productos posterior a la búsqueda de palabras claves.

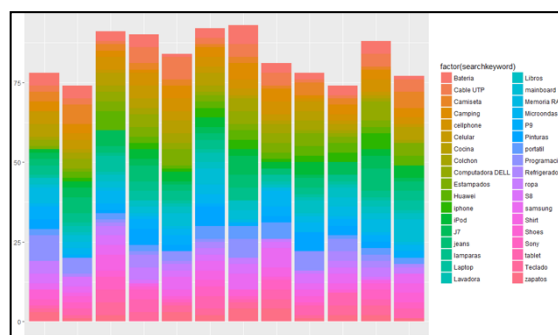


Figura 5: Gráfico proporcional de búsqueda por palabras claves de acuerdo al país

El nivel de soporte de las reglas muestra un porcentaje aproximado del 1%, lo cual no representa una cantidad alta porcentualmente. Debido a la gran cantidad de información que se utilizó para el análisis y la forma en la que se generaron los datos, la frecuencia con la que una regla se repite dentro del conjunto de datos puede no ser alta. Sin embargo en relación con el número de registros utilizados (5`000.000) el porcentaje de soporte de las reglas se considera aceptable para el análisis. Para una mejor interpretación de los resultados es eficiente el uso de herramientas gráficas que permitan describir un problema o la solución del mismo.

El uso de los gráficos avanzados para la descripción de los resultados facilita al analista la interpretación de los datos. Cada uno de los gráficos muestra los mismos resultados en forma diferente para obtener diferentes

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

puntos de vista. Por ejemplo las *Figura 2* y *Figura 3* muestran las reglas encontradas en las cuales una se puede interpretar a simple vista ya que muestra un grafo estático de las reglas (*Figura 2*) y la otra se debe interpretar utilizando una herramienta que permita ver imágenes dinámicas que ocultan información dentro de ellas y es necesario desplazarse sobre sus elementos para obtener conocimiento.

A simple vista, los resultados obtenidos pueden resultar un poco incoherentes. Esto puede discutirse considerando factores como la forma en la que se encuentran agregados los datos en la base de datos: estos fueron insertados de forma aleatoria solamente con fines experimentales. Otro punto a tomar en cuenta es el nivel de confianza y soporte que se aplica al momento de encontrar las reglas; no existen valores exactos a utilizar debido a que esto dependerá de la calidad y cantidad de datos a analizar. En el transcurso del desarrollo de este trabajo se realizaron pruebas con diferentes niveles de confianza, soporte y cantidades de datos, de los cuales la mayoría de reglas encontradas fueron nulas, incoherentes y hasta redundantes.

CONCLUSIONES

Tomando en consideración la masificación del uso del Internet, y en particular del comercio electrónico, en este artículo se ha descrito un estudio enfocado a la utilidad de una herramienta en particular: Google Analytics. Esta utilidad ha sido verificada mediante la identificación de patrones de búsquedas que podrían realizar los usuarios de un sitio web de comercio electrónico.

Adicionalmente, el estudio permite emitir otras conclusiones. En primer lugar, la herramienta de análisis de datos proporciona librerías que facilitan la conexión con la base de datos. Los métodos utilizados para obtener los datos son funcionales y sencillos, y no necesitan de muchos parámetros para realizar una transacción. Además, estos métodos mostraron ciento por ciento de efectividad en sus tareas.

Segundo, el algoritmo utilizado para el análisis de los patrones de búsqueda mostró gran velocidad y capacidad de procesamiento. Con su aplicación se logró encontrar patrones de búsqueda que superaron el nivel de soporte y confianza establecidos como parámetros. Esto hace que el resultado del análisis sea más eficiente y productivo mejorando la toma de decisiones en base al conocimiento obtenido.

Tercero, los resultados del análisis efectuado se vieron limitados por la calidad de los datos. Al simular datos transaccionales de una empresa estos muestran cierto grado de incertidumbre en cuanto a su veracidad, por tanto el comportamiento y los patrones de búsqueda pueden mostrar resultados de palabras que no tienen relación entre sí. Es por esto que se recomienda usar datos reales para el análisis. Se espera superar esta limitación a futuro realizando un nuevo estudio con datos reales luego de que se ha verificado la viabilidad de aplicar las etapas propuestas y la utilidad de las herramientas empleadas.

REFERENCIAS

Anaytics, G. (2016). Google. Recuperado el 31 de Agosto de 2017, de <https://support.google.com/analytics/answer/1033861?hl=es>

Ballesteros, D., & Ballesteros, P. (2007). El comercio electrónico y la logística en el contexto latinoamericano. *Scientia et Technica*, 1(1), 269-274.

CIENTÍFICA Y TECNOLÓGICA

In Geni



eISSN: 2697-3642

REVISTA DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

<http://revistas.uteq.edu.ec/index.php/ingenio/index>

- Fang, W. (2007). Using Google Analytics for improving library website content and design: A case study. *Library Philosophy and Practice*, 1-17.
- Gaur, L., Singh, G., & Kumar, S. (2016). Google Analytics: A Tool to make websites more Robust. *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies* (p. 45). ACM.
- Gonçalves, B., & Ramasco, J. J. (2008). Human dynamics revealed through Web analytics. *Physical Review E*, 78(2).
- Mejía, J. (2013). *Curso de formación de Google Analytics*.
- Molla, A., & Licker, P. S. (2001). E-commerce systems success: An attempt to extend and respecify the Delone and MacLean model of IS success. *J. Electron. Commerce Res.*, 2(4), 131-141.
- Morales, C. R., Soto, S. V., & Martínez, C. H. (2005). Estado actual de la aplicación de la minería de datos a los sistemas de enseñanza basada en web. *Actas del III Taller Nacional de Minería de Datos y Aprendizaje, TAMIDA2005*, (págs. 49-56).
- Phippen, A., Sheppard, L., & Furnell, S. (2004). A practical evaluation of Web analytics. *Internet Research*, 14(4), 4-6.
- Ranganathan, C., & Ganapathy, S. (2002). Key dimensions of business-to-consumer web sites. *Information & Management*, 39(6), 457-465.
- Starov, O., Zhou, Y., Zhang, X., Miramirkhani, N., & Nikiforakis, N. (2018). Betrayed by your dashboard: Discovering malicious campaigns via web analytics. *Proceedings of the 2018 World Wide Web Conference. ACM*, (págs. 227-236).
- Zumel, N., & Mount, J. (2014). *Practical Data Science with R*. New York: MANNING.