

**CHATBOT PARA LA BÚSQUEDA Y COMPARACIÓN DE PRODUCTOS
TECNOLÓGICOS A TRAVÉS DE LA PLATAFORMA MESSENGER**

**RUBEN DARIO VILLAMIZAR LAGUADO
DANIEL FERNADO RICO LEON**

**UNIVERSIDAD AUTONOMA DE BUCARAMANGA
FACULTAD DE INGENIERIAS
PROGRAMA DE INGENIERIA DE SISTEMAS
PROYECTO DE GRADO
BUCARAMANGA 2019**

**CHATBOT PARA LA BÚSQUEDA Y COMPARACIÓN DE PRODUCTOS
TECNOLÓGICOS A TRAVÉS DE LA PLATAFORMA MESSENGER**

**RUBEN DARIO VILLAMIZAR
LAGUADO DANIEL FERNADO
RICO LEON**

**DIRECTOR
RENE ALEJANDO LOBO QUINTERO**

**UNIVERSIDAD AUTONOMA DE BUCARAMANGA
FACULTAD DE INGENIERIAS
PROGRAMA DE INGENIERIA DE SISTEMAS
PROYECTO DE GRADO
BUCARAMANGA 2019**

TABLA DE CONIDO

1. Introducción	4
2. Planteamiento del problema	4
3. Pregunta(s) de Investigacio	7
4. Objetivos	7
4.1. General.....	7
4.2. Específicos	7
5. Marco teórico	7
5.1. Web scraping.....	7
5.2. Tecnologías para hacer web scraping.....	10
5.3. Chatbot.....	11
5.4. E-commerce (Comercio Electrónico).....	14
5.5. Data analytics.....	15
6. Metodología	18
7. Resultados obtenidos	19
8. Desarrollo	20
8.1. Objetivo 1.	20
8.1.1. Estado del arte.....	20
8.1.1.1. Plataformas.....	20
8.1.1.2. Artículos científicos.....	18
8.2. Objetivo 2.	27
8.3. Objetivo 3.	32
8.4. Objetivo 4.	38
8.5. Objetivo 5.	48
9. Conclusiones	51
10. Trabajos futuros	52
11. Agradecimientos	53
12. Bibliografía	54

CHATBOT PARA LA BÚSQUEDA Y COMPARACIÓN DE PRODUCTOS TECNOLÓGICOS A TRAVÉS DE LA PLATAFORMA MESSENGER

1. Introducción

El comercio online es una tendencia que poco a poco se ha ido convirtiendo en parte de la sociedad de consumo como una herramienta básica e indispensable a la hora de adquirir un bien o servicio o buscar información de los mismos antes de comprarlos o adquirirlos. En Europa y América del Norte es un mercado que ya se ha estabilizado ya que en estos países se ha adoptado este tipo de ventas mucho antes que en otras partes del mundo, incluida LA. Es por eso que en este documento se presenta el desarrollo de un chatbot para la comparación de productos y sus precios y otros criterios de artículos tecnológicos que se venden en tiendas online en Colombia el cual permite al usuario mediante ciertos criterios como lo son el precio, la calidad, la reputación de la tienda online, reputación de la marca. Escoger el producto que está buscando o encontrar uno que se acomode a sus gustos o presupuesto, este sistemas se desarrollará por medio de un chat inteligente el cual mediante una serie corta de preguntas concretas le mostrará al usuario todas la opciones que hay en el mercado online de ese producto, productos que serán extraídos de las diferentes tiendas online con la tecnología web scraping. Este sistema es además una herramienta para la ayuda de toma de decisiones a la hora de comprar productos tecnológicos. Para la selección de los productos se utilizarán ciertas fórmulas matemáticas, como la de ponderar o medir la importancia de cada una de las variables o criterios (precio, calidad del producto, comentarios, reputación de la tienda online) que el usuario le dé a las mismas con el fin de mostrar el producto más indicado.

2. Planteamiento del problema

Comparar y cotizar bienes y servicios antes de adquirirlos es una actividad que los consumidores han ido adoptando y priorizando a lo largo de estos últimos años, debido a muchos factores entre ellos el de que un consumidor puede ahorrarse una importante suma de dinero al conocer todas las ofertas que existen en el mercado. El comercio electrónico es uno de los que más ha tomado ventaja de estas necesidades, ofreciendo comodidad, ahorro de

tiempo, dinero en desplazamientos y buena cantidad de ofertas. Según un estudio de la INE(Instituto Nacional de Estadística) más de 17 millones de personas (españoles), cerca del 49,9% de la población realizaron operaciones de comercio electrónico durante el año 2017. Por otra parte, *“Colombia es el país de Latinoamérica con mejores precios para comprar tecnología esto por los bajos impuestos y la gran oferta que hay. En el estudio realizado por el sitio Linio.com se compraron los precios de 14 dispositivos en 71 países dejando a Colombia en el 15 posición a nivel global”*(«Índice precios tecnología 2016 - Linio», 2016). También en un estudio realizado por la cámara colombiana de comercio electrónico en el año 2016 *“se pudo evidenciar que el 76% de los personas que visitan la web en Colombia han comprado al menos un producto o servicio online desde el 2015, y siendo la electrónica una de las preferencias de los usuarios teniendo un 28% del todas de las compras online y siendo la segunda en la que el gasto promedio es de \$1.681.734. Con respecto a la experiencia de usuario la comodidad de los usuarios es del 50% y facilidad de compra del 48%”* (Pichponreay, Kim, Choi, Lee, & Cho, 2016). Por esto se puede evidenciar que las compras online son una preferencia para el usuario que va en crecimiento.

Debido a la gran cantidad de tiendas online que existen en el mercado y a su vez la cantidad de información que ofrecen por cada uno de los distintos tipos de productos que tienen a la venta en cada tienda online, se hace para el usuario el momento de realizar la búsqueda y comparación entre las diferentes ofertas y especificaciones del producto de su preferencia en cada una de las tiendas sea confuso, ya que el usuario debe estar navegando por cada una de las tiendas y seleccionando los productos que se hacen semejantes al que busca, Esto puede generar sensaciones de remordimiento y arrepentimiento en el consumidor al comprar un producto a un precio mayor del ofrecido en otras tiendas o con características deficientes para lo que lo necesitaba.

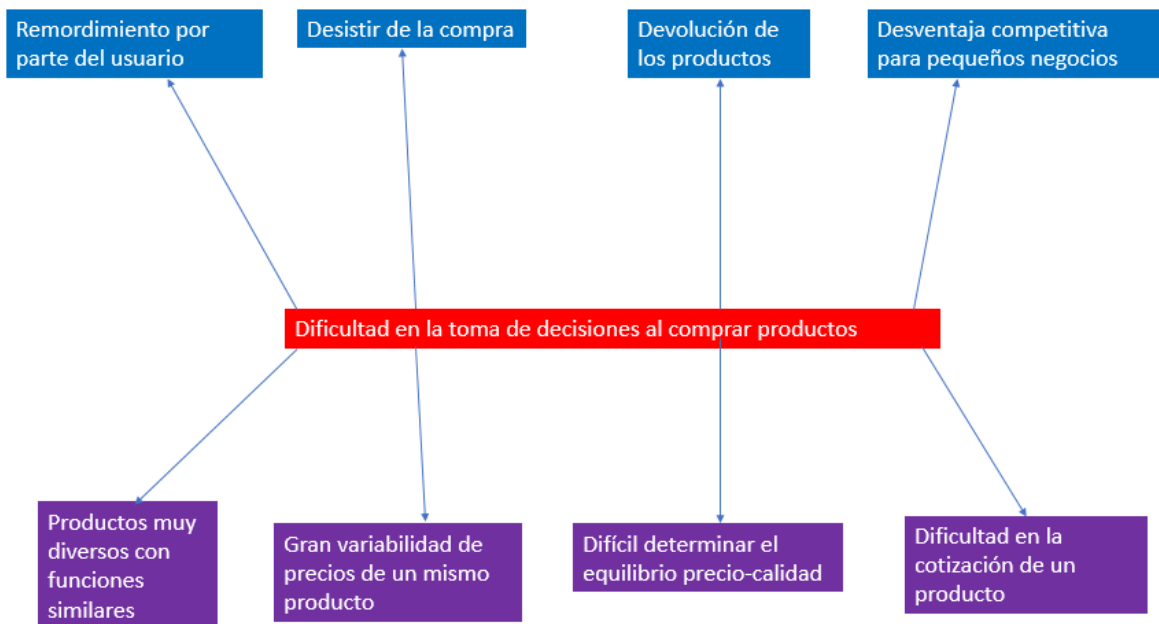


Figura 1: Árbol de problemas, autoría propia.

Para el consumidor es importante el precio y la calidad del producto, por esto desea encontrar variedad de ofertas del producto con las características que lo necesita y cotizar en un solo lugar. Ya que para el usuario la valoración de diferentes alternativas según sus preferencias siempre es importante para comprar y ahorrar su dinero (*la decisión de compra del consumidor*, 2016.).

Este proyecto plantea el desarrollo de un chatbot mediante el cual el usuario pueda buscar los artículos tecnológicos que prefiere mediante diferentes criterios de búsqueda como marca, precio o especificaciones y le muestre los productos con las características que se están ofreciendo en el momento en las diferentes tiendas online, de modo que pueda hacer una comparación de un mismo producto teniendo en cuenta precios y características del artículo. Además de comparar calidad-precio el consumidor podrá usar otras variables como comentarios, prestigio del sitio para tomar la decisión de compra del producto que mejor se adapte a sus necesidades, gustos o presupuesto. Con el fin de ser una herramienta para la ayuda de toma de decisiones en este caso a la hora de comprar un artículo tecnológico.

3. Pregunta(s) de Investigación

¿Cómo mostrar al usuario productos tecnológicos de muchos sitios web en un solo lugar?

¿Cómo facilitar el proceso de búsqueda de productos tecnológicos de forma que sea más natural para el usuario y logre facilitar la toma de decisiones?

4. Objetivos

4.1. General

Desarrollar un chatbot que permita al usuario buscar y comparar productos tecnológicos de diferentes sitios online usando la tecnología de web scraping con el objetivo de ser una herramienta para la ayuda de toma de decisiones; en este caso, en la compra de artículos tecnológicos.

4.2. Específicos

1. Realizar un estado del arte de plataformas y servicios similares de forma que se puedan extraer elementos a favor del proyecto.
2. Diseñar la base de conocimiento y la experiencia de usuario del chatbot basados en las características presentes de las tiendas online.
3. Desarrollar los algoritmos de web scraping para poblar la base de conocimiento.
4. Desarrollar el chatbot utilizando los diseños de interacción y la base de conocimiento.
5. Realizar pruebas funcionales y de usuario del chatbot.

5. Marco teórico

5.1. Web scraping

Como es definido por Haddaway la tecnología del web scraping “*Es la manera como se extraen datos de archivos HTML en Internet. Normalmente, estos datos están con patrones, particularmente listas o tablas. Los programas que interactúan con páginas web y extraen datos utilizan conjuntos de comandos conocidos como interfaces de programación de aplicaciones (API). Estas API pueden 'enseñarse' para extraer datos con patrones de páginas web individuales o de todas las páginas similares en*

todo un sitio web. Alternativamente, las interacciones automatizadas con los sitios web pueden integrarse en las API, de modo que los enlaces dentro de una página puedan 'hacerse clic' y los datos extraídos de las páginas subsiguientes”(Haddaway, 2016). Esto es particularmente útil para extraer datos de varias páginas de resultados de búsqueda. Además, esta interactividad permite a los usuarios automatizar el uso de los sitios de búsqueda de sitios web, extraer datos de múltiples páginas de resultados de búsqueda y sólo requerir que los usuarios ingresen los términos de búsqueda en lugar de tener que navegar y buscar primero en cada sitio web”

Desde que los sitios web no proveen una API bien estructurada o no la ofrecen para nada, la tecnología web scraping es la única opción a tomar cuando las compañías quieren obtener información de sus competidores, precios, novedades, recolección masiva de correos electrónicos, etc. En la figura 1 podemos ver el proceso general de web scraping el cual inicia con unos datos no estructurados los cuales se encuentran en los HTML de las páginas web, luego se observan las diferentes tecnologías que existen para realizar web scraping a estas páginas y por último vemos que los datos que se obtuvieron pasan de ser no estructurados a estructurados en diferentes formatos posibles (XML, JSON, CSV, etc.), para luego darles el uso por el cual se extrajeron de los HTML de esas páginas web(«WEB SCRAPING – Emmitta – Medium», 2017, Figura).



Figura 2: Arquitectura del Web Scrapping, «WEB SCRAPING – Emmitta – Medium». <https://medium.com/@Emmitta/web-scraping-7f87930face4> (31 de mayo de 2019).

En la figura 2. Podemos apreciar de una manera más detallada el paso a paso que se hace a la hora de extraer datos. En el primer paso, la aplicación

web scraping (Import.io, Apify) mediante una serie de funciones realizadas en javascript realiza una petición a la página web. Luego, en el 2do paso, se obtiene la respuesta (datos) que arroja la función; en el 3er paso la tecnología web scraping toma esos datos obtenidos y los convierte en datos estructurados(Json, XML, etc.), para ya en el cuarto y último paso tomar esos datos y guardarlos en base de datos con el fin de poderlos leer y manipular.

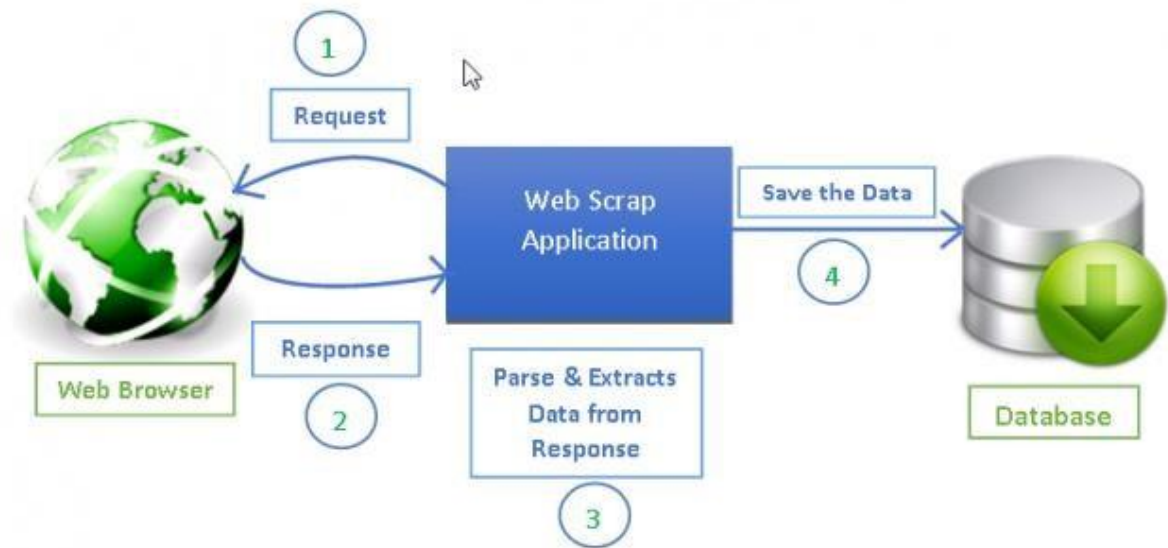


Figura 3: Pasos del Web Scrapping

En la siguiente tabla se muestran las diferentes herramientas existentes para realizar web scraping, se muestra el nombre de la herramienta, su autor, sus características, sus pros y sus contras. Cabe destacar que estas herramientas son diseñadas con el fin de automatizar los procesos de extracción de datos, haciendo este proceso mucho más sencillo.

Tabla 1:

5.2. Tecnologías para hacer web scraping

Nombre	Autor	Características	Pros	Contras
Import.io	-Matthew Painter, (CTO) -Andrew Fogg, (CDO) -David White	-Reportes de comparación -Visualización de datos con gráficos y tablas -Reportes personalizados - Fácil integración con otras herramientas	Almacenar los resultados en formato se necesite.	Alto costo y la versión gratuita solo dura 48 horas
Dexi.io	Henrik Hofmeister	-Debugging, para en caso de fallas de los robots	-Permite realizar la extracción de	-Para aprender a usar las
		Poder detectar el error y corregirlo. Scale, si quieres incrementar la cantidad solo se debe añadir más colaboradores - Extraer información desde cualquier sitio web	datos desde 4 herramientas diferentes - Plan gratuito muy completo	herramientas es necesario entender muy bien los tutoriales - Necesario saber inglés

Parsehub.com	Angelina Fomina Serge Toarca Peter Prelich	-eCommerce, Extrae datos de millones de productos y precios de vendedores online. Obtener comentarios, precios, descripción de estos productos en las tiendas online -Ciencia de los datos. -Se usa como app que se descarga en el ordenador - Rotación ip automática	-Fácil de usar y no se necesita programar nada -extracción de datos de cualquier página web. Incluso de aquellas más complejas. -Gratis -Tiene documentación	Poca potencia por lo cual la tendencia es a usar planes de suscripción
Apify.com	Jan Čurn Jakub Balada	-Conocimiento necesario en JS - Extracción de los datos en formatos csv,json,xml y rss - Diseñado principalmente para proyectos de investigación y monitorización de la competencia	Contiene plan gratuito y los planes para comenzar empiezan desde los 19\$ mensuales. - Tiene documentación	

5.3. Chatbot

Un chatbot es una tecnología capaz de simular una conversación humana a través de una interfaz conversacional. Un bot, en general, es una pieza de software diseñada para automatizar una tarea específica. Los chatbots funcionan bajo la misma premisa, automatizando la respuesta de un usuario por medio del intercambio de mensajes en lenguaje natural. Un chatbot realiza dos tareas principales: entender lo que un usuario escribe en una

conversación y ser capaz de responder de forma acertada. Donde se utilizan algoritmos de Machine Learning combinados con técnicas de Procesamiento de Lenguaje Natural. (Benotti, 2017) Por ejemplo los bancos utilizan chatbots para que los usuarios comenten sus inquietudes y se pueden resolver sin que el usuario se tenga que desplazar hasta la sucursal.

La siguiente tabla muestran los 5 mejores frameworks para construir chatbots

Tabla 2: Frameworks para construir de Chatbots

Nombre	Autor(es)	Características	Pros	Contras
Microsoft Bot	Microsoft azure	Programar en .Net, Visual Studio. Integra inteligencia artificial del bot Luis. Servidor de azure. («Microsoft Bot Framework», 2017.)	Se puede implementar para múltiples plataformas. Diseñado específicamente para el desarrollo de bot con los conectores de Microsoft Bot Framework y los SDK de BotBuilder.	Interfaz complicada para el desarrollo.
Botkit	Comunidad botkit	Integrated NLP from LUIS.ai. Código abierto. («Botkit: Bloques de construcción para construir Bots», 2017)		No se encuentra cantidad suficiente de documentación para el desarrollo.

Dialog Flow	Google	<p>Desarrollado por el aprendizaje automático de Google.</p> <p>Se integra a cualquier plataforma. («Dialogflow», 2017)</p>	<p>Interfaz de desarrollo con fácil comprensión.</p> <p>Aprendizaje automático.</p> <p>En la página de Dialog Flow se encuentra documentación para el desarrollo, complementada con tutoriales y ejemplos.</p>	
Botpress	Brightscout	<p>Editor de flujo. Implementa algoritmos de inteligencia artificial. Se puede cargar sobre cualquier servidor. («Botpress - Un Marco de Desarrollo de Chatbot Maker & Bot», 2018)</p>	<p>Código abierto. Multiplataforma.</p>	
Rasa Stack	Rasa Technologies GmbH	<p>Se conecta con varias APIs.</p> <p>Modelos personalizados. («Rasa Stack: AI conversacional de código abierto», 2018)</p>	<p>Los datos se pueden tener en un servidor privado</p>	<p>No se puede implementar para todas las plataformas de mensajería.</p>

5.4. E-commerce (Comercio Eletrónico)

Esta manera de comercio Online se define como “Es la actividad de comprar y vender bienes y servicios a través de internet y la transferencia de dinero para ejecutar esta actividad” (Manuel Rado Quiroz, 2018). Este término a veces es confundido con e-business el cual es un concepto mucho más complejo, es todo de lo que se hace detrás de la ventas y las estrategias necesarias a hacer para mantener negocios online, es decir el término e-commerce es solo una parte del modelo de e-business. “El comercio electrónico (‘e-commerce’) ha multiplicado por tres las ventas registradas durante los últimos seis años en América Latina. Además, las estimaciones apuntan que el comercio electrónico crecerá un 24,5% en la región durante 2018” (Manuel Rado Quiroz, 2018). Por otra parte, según la International Data Corporation (IDC) “el ‘e-commerce’ ha triplicado sus ventas en la región durante los últimos seis años, pasando a ingresar 40.000 millones de dólares”. Y según pronósticos de la misma “el comercio electrónico, compuesto por la distribución, venta, compra, marketing y suministro de información de productos o servicios a través de Internet, crecerá un 25,4% en 2018.”. Aunque en Europa y Norteamérica es una tendencia que se ha estado estabilizando ya que ellos llevan más tiempo con este modelo de negocio que AL.

En la figura 3 se observa una gráfica en donde se compara el crecimiento del comercio electrónico en los diferentes países de AL de un año respecto al otro. Por ej: En el 2018 el comercio electrónico ha crecido un 25% más aprox vs año 2017 en Colombia («WEB SCRAPING – Emmita – Medium», 2017).

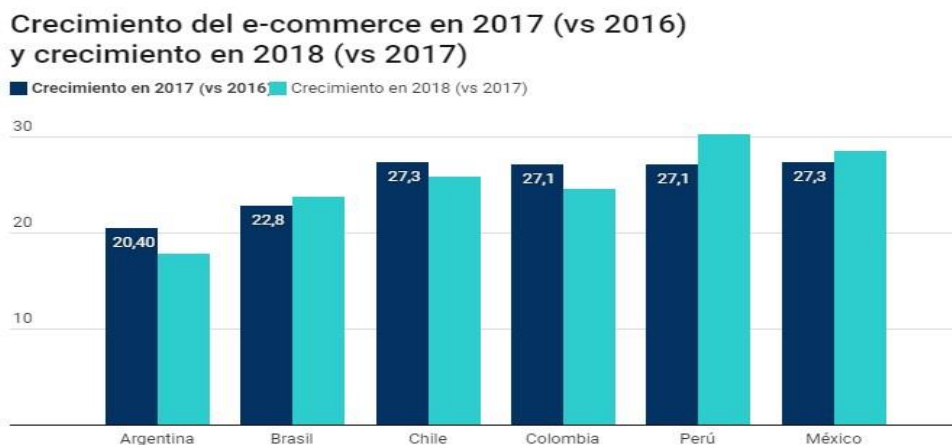


Figura 4: Crecimiento del e-commerce en 2017 (vs 2016) y crecimiento en 2018 (vs 2017)

Además cabe destacar que el 77,6% de las transacciones se realizaron con tarjetas, más en países como Colombia, Perú o Chile el uso de efectivo para realizar las transacciones es más equilibrado vs tarjetas.

Tipos de e-commerce («WEB SCRAPING – Emmita – Medium»,):

- Business to Consumer (B2C): Es el más famoso y consiste en que el consumidor compra un producto o adquiere un servicio por medio de un productor o distribuidor. Ej: Amazon
- Business to Business (B2B): Es básicamente el negociar entre 2 empresas diferentes es cuando una empresa le ofrece a otra un servicio o producto que esta otra necesita para su funcionamiento. Ej: Facebook empresarial
- Consumer to Consumer (C2C): Es cuando un consumidor le vende un producto o suministra un servicio a otro consumidor. Ej: Cuando tu vendes tu celular de segunda en mercadoLibre o e-Bay a otra persona.
- Consumer to Business (C2B): Es lo contrario a B2C, ya que en este modelo el consumidor le comunica que es lo que necesita a una empresa y esta se lo ofrece a un precio determinado. Ej: sectores de ocio y turismo. Ej: Trivago y Agropool

5.5. Data analytics

En los últimos años la importancia de los datos que se general con el uso de la web ha crecido y con eso se ha generado la necesidad de descubrir tendencias las cuales son estudiadas con el análisis de datos que se define como “el proceso de examinar información usando principalmente un computador, con el objetivo de averiguar algo o ayudar a la toma de decisiones” (Cambridge dictionary, 2017). Otra definición algo más técnica dice que es la ciencia de la obtención de datos en bruto, para representarlos después de ciertos procesos de limpieza y categorización por medio de tablas y gráficos, con el fin de sacar conclusiones para la toma de decisiones dependiendo de los requerimientos o necesidades de quien lo solicita ya sea una empresa, una organización o una persona natural. Además la ciencia lo usa para comprobar o descartar teorías o modelos existentes.

“El análisis de datos es usado en varias industrias para permitir que las compañías y las organizaciones tomen mejores decisiones empresariales y también es usado en las ciencias para verificar o reprobando modelos o teorías existentes”. El análisis de datos se distingue de la extracción de datos por su alcance, su propósito y su enfoque sobre el análisis. Los extractores de datos

clasifican inmensos conjuntos de datos usando software sofisticado para identificar patrones no descubiertos y establecer relaciones escondidas. El análisis de datos se centra en la inferencia, el proceso de derivar una conclusión basándose solamente en lo que conoce el investigador. Las compañías de comercio electrónico (Ecommerce) examinan el tráfico en el sitio web o los patrones de navegación para determinar qué clientes son más o menos propensos a comprar un cierto producto o servicio, basándose en compras previas o patrones de visualización. El análisis de datos moderno normalmente usa tableros de información que se basan en flujos de datos en tiempo real. El llamado análisis en tiempo real implica análisis e informes dinámicos basados en los datos que introducidos en un sistema un minuto antes del tiempo actual de uso (Margaret Rouse, 2017). Y este proceso es muy usado en el modelo B2C del e-commerce.

Técnicas o tipos de análisis de datos:

- **Análisis de datos cualitativo:** Los datos cualitativos se presentan de manera verbal (en ocasiones en gráficas). Se basa en la interpretación. Las formas más comunes de obtener esta información es a través de entrevistas abiertas, grupos de discusión y grupos de observación, donde los investigadores generalmente analizan patrones en las observaciones durante toda la fase de recolección de datos.
- **Análisis de datos cuantitativos:** Los datos cuantitativos se presentan en forma numérica. Se basa en resultados tangibles.

	Tipos de datos	Análisis	Ejemplos
Cualitativo	Se centra en las opiniones, actitudes y creencias.	Preguntas y respuestas a preguntas como: ¿Por qué? ¿Cómo?	Paneles en donde se da una discusión y se entrevista a consumidores sobre lo que les agrada o no del lugar.
Cuantitativo	Se centra en los datos duros e información que pueda contabilizarse.	Se obtiene mediante preguntas similares a: ¿Cuántos? ¿Quién? ¿Con qué frecuencia? ¿Dónde?	Encuestas enfocadas a medir las ventas, tendencias, reportes o percepciones.

Figura 5: Tipos de análisis de datos.

En la figura 5 se pueden ver los 5 pasos necesarios para realizar análisis de datos. En el 1er paso se realizan las preguntas correctas a las cuales se les podría dar respuesta con este análisis, se define el “qué” pasar saber el “como” solucionarlo, además de que en este paso con base a las preguntas se define si es necesario o no realizar DA. En el 2do paso se conoce que datos se tienen, en qué cantidad, naturaleza de los datos además de recolectar esos datos por medio de encuestas, bases de datos, etc. Esto con

el objetivo de conocer qué soluciones le podría ofrecer este conjunto de datos. En el 3er paso se debe conocer cuáles son los datos que se necesitan para sacar una conclusión o responder a una pregunta. Ej: Si el problema es “¿cuál de los empleados ha vendido más en los últimos seis meses? El conjunto de datos ideal es una lista de empleados, y la suma de ventas para cada uno, para el último semestre vigente. Al conocer el entorno de datos, se identifica qué datos le hacen falta, y una estrategia para dar con estos y consumirlos. También, se deben “limpiar” las variables para evitar confusiones (empleados repetidos, faltas de ortografía o formatos). Una buena estrategia es crear análisis exploratorios, con datos simulados, para poder justificar la compra u obtención de los datos faltantes, acá se convierten los datos en gráficos para una mayor interpretación. En el 4to paso: con base a los resultados obtenidos se implementan las soluciones, acá se debe evitar un análisis erróneo. “Garbage in, garbage out” (si entra basura, sale basura). Un análisis mal planteado conlleva a malos resultados. Y el último y sexto paso implica en medir qué tanto han sido beneficiosas esas medidas tomadas para a partir de esto crear una “guía” a la hora de seguir haciendo este proceso, proceso que debe ser pensado a largo plazo.



Figura 6: Pasos para hacer análisis de datos.

6. Metodología

Objetivo 1:

- Consultar en las base de datos científicas(Scopues, IEEE) artículos en base a las keywords del proyecto
- Seleccionar una cantidad considerable de artículos, basados en el abstract de cada uno de ellos

- Hacer una selección más minuciosa de los artículos y a estos hacerles un estudio en búsqueda de aportes científicos, matemáticos, tecnológicos y teóricos.

Objetivo 2:

- Realizar una búsqueda de los sitios online que más venden tecnología en Colombia para tomarlos como referente al elaborar la base de datos.
- Realizar una búsqueda de las características en común de los productos tecnológicos que venden los sitios online escogidos.
 - Crear el modelo Entidad-Relación de la plataforma
 - Crear la Base de conocimiento.

Objetivo 3:

- Consultar las diferentes tecnologías que permitan hacer web scrapping
- Realizar web scraping a las tiendas online a partir de la tecnología seleccionada
- Insertar los resultados obtenidos de los web scrapings a la base de conocimiento.

Objetivo 4:

- Seleccionar los filtros con los cuales el usuario realizará su comparación (Precio, calidad, reputación marca, tienda online, etc.)
- consultar los mejores frameworks para realizar el front-end del chat bot.
- Diseñar e implementar la UI del chatbot a partir del framework seleccionado.

Objetivo 5:

- Desarrollar el backend del chatbot. (El backend recibe mensajes, piensa en una respuesta y la devuelve al usuario) □ Conectar front y back mediante webhooks.

Objetivo 6:

- Seleccionar una muestra de estudiantes de la universidad para que usen el chat bot en busca de errores, deficiencias, calidad de la user experience, performance del chatbot, etc.

- Realizar los respectivos ajustes obtenidos de las pruebas realizadas

7. Resultados obtenidos

1. Estado del arte el cual contiene elementos a favor del proyecto.
2. Base de datos relacional para almacenar información de los productos tecnológicos que se venden en las tienda online seleccionadas.
3. Conjunto de algoritmos de extracción de datos de tiendas online.
4. Interfaz gráfica por medio de la cual los usuarios podrán interactuar con nuestro sistema de comparación.
5. Chatbot para la comparación de productos tecnológicos.
6. Diagramas con resultados de las pruebas realizadas con los usuarios.

8. Desarrollo

A continuación se muestra el desarrollo que se ha realizado durante el proyecto para dar cumplimiento con los objetivos que fueron planteados.

8.1. Objetivo 1: Realizar un estado del arte de plataformas y servicios similares de forma que se puedan extraer elementos a favor del proyecto.



8.1.1. Estado del arte


Para tener una mayor referencia de lo que ya está hecho con respecto al tema se ha desarrollado un estado del arte basado en las plataformas que ya existen tomando de estas sus funciones y que aporte dan al proyecto. También, se ha hecho una revisión en bases de datos académicas sobre investigaciones que traten sobre web scrapping, chatbot y e-commerce para poder tener una referencia a cerca de las tecnologías que se pueden implementar en el desarrollo del proyecto.


Tabla 3:

8.1.1.1. Plataformas

Nombre	Que hace	Aporte	Descargas
TRIVAGO	Es una aplicación que permite a los usuarios comparar precios de hoteles trayendo la	Trivago es una plataforma que aporta una característica muy importante de	120 millones de visitas mensuales

	<p>información de las páginas web de los hoteles de todo el mundo, le muestra la comparación haciendo una búsqueda por fecha, ubicación y precio. Es un medio por el cual el usuario puede hacer su búsqueda en tiempos más cortos, esta plataforma muestra las diferentes opciones y envía al usuario a la página web de la oferta, no se encarga de pagos de las reservas o la reputación de cada hotel.(«trivago.com.co - Comparar precios de hoteles en todo el mundo», 2018)</p>	<p>nuestro proyecto y es el uso de filtros para personalizar y mejorar aún más la comparación de precios entre productos, ellos lo hacen con hoteles, nosotros con productos tecnológicos. Además, de tomarlo como guía para la construcción de la interfaz de usuario de nuestro proyecto.</p>	
<p>Pricekart</p> 	<p>Es una app para comparar precios en la India y está disponible en la Play Store. App de compras para compradores inteligentes antes de que tomen la decisión de comprar online. Clientes online pueden encontrar, buscar, y comprar los precios de millones de productos de más de 100 tiendas en la India. Esta app fue creada con inteligencia artificial y machine learning que ayuda a los usuarios a tomar la mejor decisión a la hora de comprar. («Best Price Comparison shopping</p>	<p>Realizar una selección de todas las características (Marcas, rangos de precios, tamaños de pantallas, ram, pesos, dimensiones) de algún tipo de producto tecnológico (TV, portátil, Tablet, celular) para luego mostrarme todos los artículos que hay en base a esa configuración previa a la búsqueda.</p>	<p>5.000+</p>

	online in India - Pricekart.com», 2018)		
<p>Best Price Comparison Shopping</p> 	<p>App hindú para la comparación de precios a la hora de comprar online. Compara precios online en las principales tiendas online. (Flipkart, Amazon, Snapdeal, eBay, PayTM, Shopclues, etc.). Más de 10 millones de productos se enlistan por categorías como lo son: Teléfonos móviles, tablets, computadores portátiles, tablets, cámaras, televisores, dispositivos personales, libros entre otros. En esta app su pueden encontrar cupones de ahorro, promociones, ofertas con buen descuento de cada una de las tiendas online que existen en el mercado. Además cuentan con su propia página web donde vender los productos. («Shopping.com UK - Find, Compare, and Buy Anything in Seconds», 2000-2018)</p>	<p>-Obtener alertas en tiempo real cuando los precios de los productos bajen. -Comparar productos que no tienen que ser los mismos. Por ejemplo: Un galaxy s8 con un huawei 10. -Algunos segmentos de la parte gráfica.</p>	+1'000.000

<p>Mobile Price Comparison App</p> 	<p>91mobiles en una app móvil hindú para la comparación de precios de teléfonos móviles y dispositivos. En esta el usuario podrá encontrar los mejores precios de teléfonos, laptops, tablets, etc. Con base al presupuesto del usuario. Compara precios, especificaciones, y ofrece las diferentes ofertas de sitios de venta online como los son: Flipkart, Amazon, Snapdeal, eBay, PayTM, Shopclues, etc. («Mobile Phones Mobile Prices in India Online Mobile Shopping 91mobiles.com», 2016)</p>	<p>Poder ir añadiendo productos a una lista para luego compararlos y dar la información de cada uno y en qué tiendas vende cada uno de ellos. Cuando un producto está agotado, puedo activar para que se me notifique cuando vuelva a estar disponible.</p>	<p>1.000.000+</p>
<p>Kelkoo</p>	<p>Servicio web de comparación de precios Europeo fundando en Francia en el año 1999. Ellos lo que hacen es buscar las mejores ofertas para el usuario según lo que está y con todo tipo de productos. («Kelkoo Comparador de precios, guías de compra, ofertas de vacaciones online», 2018)</p>		
<p>Idealo</p>	<p>Idealo es una web que compara precios podrás buscar un sinfín de productos y servicios. Desde aerolíneas y</p>	<p>Contiene una gráfica del historial de precio de ese producto, en un rango hasta de un año hasta la fecha</p>	

	<p>artículos de primera necesidad hasta de ocio, seguros y demás. La plataforma de la página funciona buscando entre más de 7.400 tiendas online. Y no son solo precios lo que encontrarás, también podrás obtener los comentarios, opiniones y recomendaciones del resto de los usuarios y filtros especiales para buscar. Además de que funciona en varios países de Europa. («IDEALO – die Nr. 1 im Preisvergleich», 2018)</p>	<p>actual, además de que permite al usuario fijar un precio ideal por el que pagaría ese producto y dejando su correo se le enviará uno al mismo cuando el producto llegue a ese precio ideal fijado por el cliente.</p>	
--	---	--	--

8.1.1.2. Artículos científicos

- II. Julian, L. R., & Natalia, F. (2016). *The use of web scraping in computer parts and assembly price comparison. En CONMEDIA 2015 - International Conference on New Media 2015 (pp. 1-6). IEEE***

Esta aplicación también cuenta con una función de comparación de precios basada en fuentes de datos recuperadas de cinco tiendas de informática para que los usuarios puedan ahorrar los costos de comprar piezas de PC y ensamblar la computadora más fácilmente. Buscando que el usuario pueda comprar las piezas a menor costo y buena calidad.

- III. Slamet, C., Andrian, R., Maylawati, D. S., Suhendar, Darmalaksana, W., & Ramdhani, M. A. (2018). *Web Scraping and Naïve Bayes Classification for Job Search Engine. En IOP Conference Series: Materials Science and Engineering (Vol. 288, p. 012038). IOP Publishing.***

En esta solución quiere simplificar la manera de buscar trabajo en los diferentes sitios online mostrando en un solo lugar las ofertas que el usuario está buscando mediante Web Scrapping y NAive BAYes que se utiliza para la clasificación de los datos. Para ellos es muy importante conocer la navegación de las páginas web a la hora de hacer Web Scrapping y guardar todos los datos que se extraen en bases de datos con historiales de paquetes, presentan un esquema con los diferentes pasos que hace el web scrapping para obtener los datos y guardarlos en una base de datos para después hacer la una clasificación de los datos ellos utilizando el algoritmo de NAive Bayes que fue implementado por ellos por la velocidad y precisión al momento de clasificar.

IV. B V S, U., Gaiind, B., Kundu, A., Holla, A., & Rungta, M. (2017). Classification-Based Adaptive Web Scraper. En 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 125-132). IEEE

El Web Scrapping sufre consecuencias con la constante evolución de las páginas web que hace que sus atributos cambien y esto hace que no funcione correctamente el Web Scrapping, dentro de esta solución proponen una forma diferente de hacerlo para que se pueda adaptar a los cambios desarrollando una técnica de bloques repetitivos, donde cada bloque representa un objeto con sus atributos y un nuevo enfoque basado en clasificación. Para esto primero se desarrolla un clasificador para recopilar datos basados en css-selector y utilizando Python BeautifulSoup, después se hace una limpieza de los datos, en el siguiente paso se hace una conversión de los datos a vectores característicos.

V. Argal, A., Gupta, S., Modi, A., Pandey, P., Shim, S., & Choo, C. (2018). Intelligent travel chatbot for predictive recommendation in echo platform. En 2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018 (Vol. 2018-Janua, pp. 176-183). IEEE.

Dentro de esta solución se propone un chatbot basado en lenguaje natural que mediante la voz se puedan hacer consulta acerca del viaje, para facilitar al usuario la búsqueda de información sin tener que llamar a la agencia, también se busca que el chatbot pueda conocer al cliente para cuando el desee consultar nuevamente se tenga una referencia de sus preferencias. El sistema se divide en tres partes un dispositivo de eco para reconocimiento de la voz, que después es interpretado por Alexa Skill y consulta a una base de conocimiento que fue alimentada con Web Scrapping, también implementa

algoritmos de Machine Learning para la inteligencia del chatbot y para el filtrado colaborativo implementaron la máquina Restricted Boltzmann.

- VI. Bhawiyuga, A., Fauzi, M. A., Pramukantoro, E. S., & Yahya, W. (2018). Design of E-commerce chat robot for automatically answering customer question. En Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology, SIET 2017 (Vol. 2018-Janua, pp. 159-162).**

Dentro de esta solución tecnológica se crea un chatbot para el uso de comercio electrónico que permite a los clientes y las tiendas una comunicación constante de las preguntas que surgen o inconvenientes para los clientes, también se presenta como una solución a los costos que genera tener una persona dedicada a la atención al cliente, y permite que el cliente pueda consultar en cualquier hora del día sin ninguna restricción. Se aplica la técnica de Machine learning para seleccionar las respuestas a las preguntas realizadas por el usuario. Para su desarrollo lo dividen en dos partes. Desarrollan un agente de comunicación por medio de una solicitud HTTP servidor de Telegram y seleccionan mediante un agente inteligente la respuesta a la pregunta mediante de un algoritmo de coincidencia utilizando la distancia de Leventein para medir la similitud del texto.

- VII. Pichponreay, L., Kim, J. H., Choi, C. H., Lee, K. H., & Cho, W. S. (2016). Smart answering Chatbot based on OCR and Overgenerating Transformations and Ranking. En International Conference on Ubiquitous and Future Networks, ICUFN (Vol. 2016-Augus, pp. 10021005). IEEE.**

Es un chatbot basado en reconocimiento óptico de caracteres que se basa en escaneo e identificación de texto impreso o escrito para volverlo en datos numéricos, evitando la necesidad de volver a escribir algo impreso, para eso utilizan la biblioteca de Apache DFBOX OCR para la extracción de texto del documento. Para la generación de preguntas sobre los artículos que se dan se utiliza un proceso que consiste en tres pasos: 1. *“Extraer y derivar oraciones declarativas de un texto fuente; 2. Traduce oraciones declarativas en preguntas usando reglas declarativas de propósito general; 3. Calcule estadísticamente el resultado de las etapas 1 y 2 de sobregeneración para la aceptabilidad”*. Para su respuesta se utiliza inteligencia artificial y un servidor en Apache Tomcat.

VIII. Li, W. (2010). Study on the Typical Application of E-commerce. En 2010 International Conference on E-Business and E-Government (pp. 46-48). IEEE.

En este proyecto se quiere construir un sistema que permita a los usuarios tener sugerencias sobre libros, que sea de interfaz de fácil manejo para el usuario y que permita al usuario encontrar el libro que desea en menor tiempo. Para la construcción del sitio online utilizaron las siguientes tecnologías: La estructura del entorno NET, y el uso de Visual Studio 2005 como herramientas de desarrollo, el uso de C # como lenguaje de programación y utiliza Microsoft SQL Server 2000 como base de datos. Como producto obtuvieron un sistema que está compuesto por una tienda online que le permite al usuario adquirir sus libros de manera rápida, fácil y que le permita resolver sus dudas con respecto a la aplicación, también una plataforma de administrador que tener control de la base de datos y demás funcionalidades.

Tabla 4: **Documentos revisados**

Keyword	Documentos encontrados	Documentos seleccionados	Documentos Analizados	Rango de Fechas de la búsqueda
Web scrapping	247	29	3	2012- 2018
Chatbot	391	20	3	2010-2018
Ecommerce	6158	10	1	2015-2018

8.2. Objetivo 2: Diseñar la base de conocimiento y la experiencia de usuario del chatbot basados en las características presentes de las tiendas online.

1. se han seleccionado las tiendas basándonos en el reconocimiento que tienen estas marcas y las estadísticas de ventas que han tenido en los

últimos años en productos vendidos online, también se tuvo en cuenta su estructura y el tipo de categorías que ofrecen.

Tienda seleccionadas:

- Falabella
- Éxito
- Jumbo

2. De cada una de las tiendas se han seleccionado cuatro categorías de productos tecnológicos teniendo en cuenta que son los más vendidos en Colombia.

Categorías seleccionadas:

- Portátiles
- Televisores
- Celulares
- Video juegos

3. Para realizar el web scraping se debe definir que detalles cuales son los más importantes y se quieren traer de cada producto por eso se tomaron:

- Nombre de producto
- Marca
- Precio
- Especificaciones

4. Se creó una base de datos Mysql donde se guardarán las tiendas y los productos que se ofrecen en cada una, del producto se guardará nombre, marca, precio, especificaciones, foto y el url. La información que se extraerá con los web scraping.

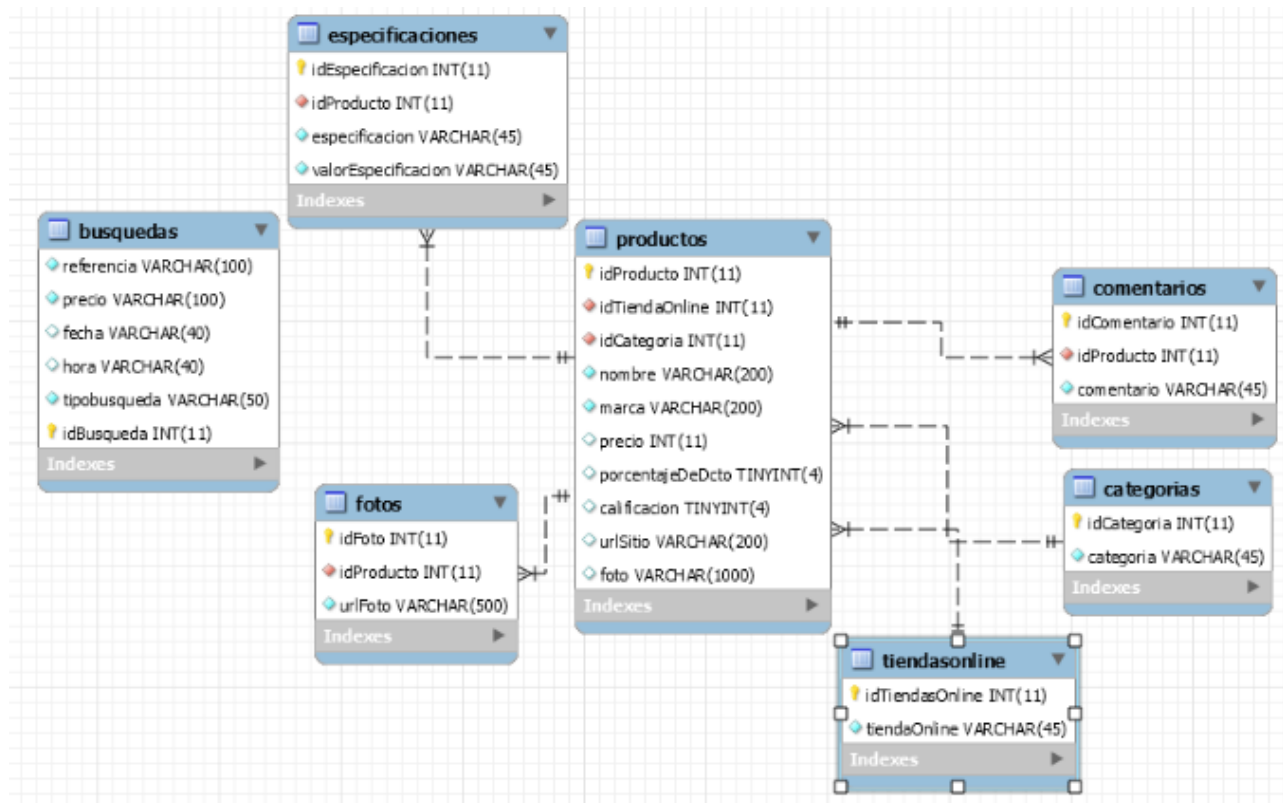


Figura 7: Modelo relacional base de datos, autoría propia.

- Método que toma los resultados de las arañadas y llena la base de datos, en este método tenemos una lista de Urls dadas por Apify las cuales contienen los resultados de cada una de las arañadas que se realizaron. Se recorre cada una de esas listas las cuales contienen una lista de productos extraídos, estos se recorren y se van añadiendo a la base de datos por medio de un insert.

```

$app->post('/apify',function(Request $request, Response $response){
    //lo primero es desocupar la base de datos para volver a llenarla DELETE FROM "productos";
    $query=$this->db->prepare("DELETE FROM productos");
    $query->execute();
    $queryAutoincrement=$this->db->prepare("ALTER TABLE productos AUTO_INCREMENT = 1");
    $queryAutoincrement->execute();
    $listaDeURLs=[];
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/1Tuoupxvhe608x7a/lastExec/results?token=NA6gF9c7Ml8Pul92WvuzEfnR";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/Au8Zd1kash80mxyDS/lastExec/results?token=Kv27Hl66R4K6Crc0Ch8V6dHMA";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/bahlp0EowFfk38e2c/lastExec/results?token=7H8zq5ahF9va0a391ge7GblZ";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/BSns954yv2kov7kkn/lastExec/results?token=8wNHFbRqxf9v64DmEP8W7kC3";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/n1GidEf9MR4KL290E/lastExec/results?token=XaAyTamdAentEGHMCAyX0GNgG";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/PLt5zqk860K8o3cz/lastExec/results?token=5AhaZCPLF6Gc75K8UB8D26u8";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/xmF38b7CYacEwC/lastExec/results?token=oxPC9KwAKw8nd2AvvWah8e9";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/bahlp0EowFfk38e2c/lastExec/results?token=7H8zq5ahF9va0a391ge7GblZ";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/1D9FkwJ1q5FEd2/lastExec/results?token=U3eTQZFG22X43Segp5bFRovr";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/CF8BofTP5VYgG0T/lastExec/results?token=boHm6k7X0eR36gZd13pSryW20";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/qGH4fVZ8Ew0c54K/lastExec/results?token=YkN538ZMeSp2EYwGaeWOT4e";
    $listaDeURLs[]="https://api.apify.com/v1/Nr8VC8j21zZd8Gw6/crawlers/Nr28v5YETC1W8vPv/lastExec/results?token=9enc8edYasvnaAmvCas7uqJh";
    foreach ($listaDeURLs as $url) {
        $productos=[];
        $productos = json_decode(file_get_contents($url));
        foreach ($productos as $producto) {
            $miproducto=$producto->pageFunctionResult;

            if($miproducto->precio==""){
                $miproducto->precio=rand(600000,3000000);
            }
            if($miproducto->dcto==""){
                $miproducto->dcto=rand(10,50);
            }
            $query=$this->db->prepare("INSERT INTO productos(idTiendaOnline,Idcategoria,nombre,marca,precio,porcentajeDeDcto,calificacion,urIsitio,Foto) VALUES ('$miproducto->
            $query->execute();
        }
    }
    return $response->withJson(["status":200]);
});

```

Figura 8: código del método para tomar las arañadas, autoría propia.

- Se crearon las diferentes consultas a la base de datos para poder obtener los datos necesarios que se deben mostrar a las diferentes peticiones que el usuario haga por medio de dialogflow. Este es el ejemplo del end-point que contiene el resultado de la búsqueda que el usuario hace dentro del chatbot de un producto en específico, este como todos los métodos recibe unos parámetros y según estos devuelve una respuesta al chatbot.

```

function option2($parametros,$contexto,$fecha,$hora){
    $producto=$parametros["producto"];
    $query=$contexto->db->prepare("INSERT INTO busquedas(referencia,precio,fecha,hora,tipobusqueda) VALUES ('$producto','0','$fecha','$hora','Referencia exacta')");
    $query->execute();
    $respuesta = new stdClass;
    $respuesta->fulfillmentMessages=[];
    $consulta= $contexto->db->prepare("SELECT * FROM 'productos' WHERE nombre LIKE '%{$producto}%' LIMIT 10");
    $consulta->execute();
    $resultados = $consulta->fetchAll();
}

```

```

foreach ($resultados as $resultado) {
    $tienda="";
    switch ($resultado["idTiendaOnline"]) {
        case 1:
            $tienda="Jumbo";
            break;
        case 2:
            $tienda="Exito";
            break;
        case 3:
            $tienda="Falabella";
            break;
    }
    $card=new stdClass;
    $atributosCard=new stdClass;
    $atributosCard->subtitle=$tienda;
    $atributosCard->title=$resultado["precio"];
    $atributosCard->imageUri=$resultado["foto"];
    $buttons=[];
    $atributosButton=new stdClass;
    $atributosButton->text="Comprar";
    $atributosButton->postback=$resultado["urlSitio"];
    $buttons[]=$atributosButton;
    $atributosCard->buttons=$buttons;
    $card->card=$atributosCard;
    $card->platform="FACEBOOK";
    $respuesta->fulfillmentMessages[]=$card;
    return $respuesta;
}

```

Figura 9: código para realizar consultas a la base de datos, autoría propia.

7. Creamos un servidor web en AWS para alojar nuestra aplicación web dinámica.

En este servidor en blanco se configuro Apache y PhpMyAdmin para que pudiera sobre este alojar nuestro código back-end y nuestra base de datos relacional MySQL. Por otra parte, debido a los requerimientos de dialogflow nuestra instancia debía tener certificado SSL para poder hacer la integración entre nuestro back-end y el chatbot. Para poder tener un certificado SSL se necesita tener un dominio y se creó por medio de Name Cheap el cual tuvo un costo de \$10 dólares, dominio el cual se llama www.botdecompras.com. (AWS | Elastic compute cloud (EC2) de capacidad modificable en la nube s. 2019)

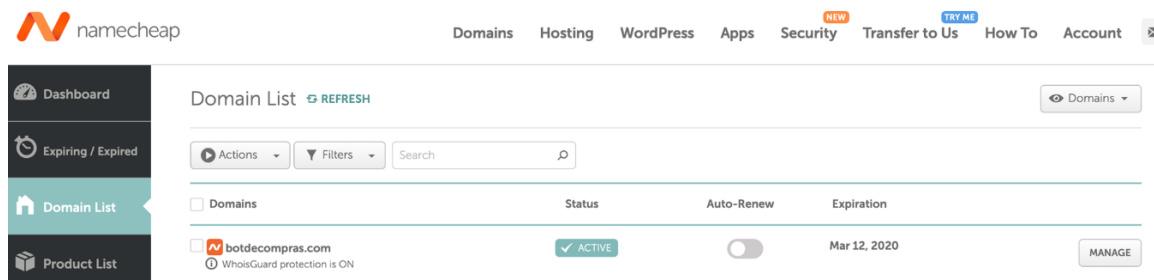


Figura 10: Dominio creado en NameCheap, autoría propia.

A continuación podrán ver link de los pasos necesarios para obtener lo ya mencionado.

- Tutorial: Instalación de un servidor web LAMP con la Amazon Linux AMI
https://docs.aws.amazon.com/es_es/AWSEC2/latest/UserGuide/install-LAMP.html
- Tutorial: Conexión a la instancia de Linux desde Windows mediante PuTTY
https://docs.aws.amazon.com/es_es/AWSEC2/latest/UserGuide/putty.html
- Tutorial de como vincular la instancia de AWS con nombre de dominio Namecheap.
<http://paoloibarra.com/2017/03/11/Route53-To-NameCheap-Domain/>

8.3. Objetivo 3: Desarrollar los algoritmos de web scraping para poblar la base de conocimiento.

Como resultado de este objetivo tenemos que:

1. Se trabajara en la plataforma apify.com para la construcción de los algoritmos web scraping del proyecto, ya que realizando la consulta y comparando entre cada una de las tecnologías que en el momento nos permiten hacer web scraping se pudo evidenciar que es la más completa, dinámica, los planes de plago no tienen tanto costo a diferencia de las demás y su documentación es bastante amplia permitiendo que sea fácil el aprendizaje para realizar nuestro proyecto.
2. Se realizaron los algoritmos de web scraping de algunas de las categorías de las diferentes tiendas seleccionadas de las cuales se seleccionaron los siguientes detalles:
 - Nombre
 - Marca

- Precio
- Especificaciones

A continuación se explican los pasos para construir el algoritmo de Web Scraping para obtener los datos tomando como ejemplo los portátiles de la página web del éxito:

The screenshot shows a web scraping configuration interface. At the top left is a blue 'Save' button. Below it are two input fields: 'Custom ID' with the value 'ExitoTV' and 'Internal ID' with the value 'xmKFq38b7cYAcENFC'. A 'Comments' field contains the text 'Optional notes'. The 'Start URLs' section has a table with one row: 'Optional label' and 'https://www.exito.com/Tecnologia-TV_y_Video-Televisores/_/N-2csn'. Below this is a '+ Add' button. The 'Pseudo-URLs' section has a table with two rows: 'paginas' with 'https://www.exito.com/Tecnologia-TV_y_Video-Televisores/_/N-2csn?No=[\d+]' and 'detalles' with 'https://www.exito.com/products/[\d+]/[{\w|-}]*'. Below this is another '+ Add' button. At the bottom, the 'Clickable elements' field contains the selector 'a:not([rel=nofollow])'.

Figura 11: Algoritmo de web Scraping para televisores de la tienda éxito, autoría propia.

La etiqueta Start Urls contiene la URL desde donde apify (arañador) va a empezar a hacer la extracción de los datos

Las etiquetas Pseudo-URLs contienen Urls también pero su función es navegar entre las paginas (1,2,3....n) y para navegar dentro de los componentes que existen dentro de cada página, estas etiquetas contienen expresiones regulares que admiten cierto tipo de caracteres dependiendo de la estructura URL de la página a la que estamos arañando.

Ejemplo:

|
[https://www.exito.com/Tecnologia-Computadores-_impresoras_y_tablets-Portatiles/_/N-2b2n?No=\[\d+\]&Nrpp=20](https://www.exito.com/Tecnologia-Computadores-_impresoras_y_tablets-Portatiles/_/N-2b2n?No=[\d+]&Nrpp=20)

Contiene como se señala una expresión regular que admite números enteros positivos. Ahora, la navegación entre paginas contiene esa estructura, simplemente cambia el numero en el que va la expresión regular, de ese modo entraremos a todas las páginas que haya dentro de esa URL.

Ahora, dentro de cada página necesitamos entrar en los productos que hay en cada página por eso necesitamos otra Psuedo-URL y es la siguiente:

`https://www.exito.com/products/[(\d+|\w)*]/[(\w|-)*]`

En esta vemos expresiones regulares para poder entrar dentro de cada producto que hay en cada página.

Ejemplo:

`https://www.exito.com/products/0003087421927226/zenbook-pro-intel-ci9-8950hks`

Como podemos ver los productos que hay dentro de cada página tienen una estructura similar, lo que cambia es lo que se encuentra señalado por eso en vez de poner valores fijos, fijamos expresiones regulares que admitan todos esos productos sin importar que caracteres contengan y así nos aseguramos de entrar en cada producto, por medio de las expresiones regulares.

Para la construcción del algoritmo se usa lenguaje JS que nos permite la extracción de información de cada producto y para eso además se usa JQuery Slector (\$), para poder obtener de acuerdo a la clase del ítem extraerle la información, este proceso de obtener la clase del ítem a extraer se tiene que hacer viendo la estructura HTML de la página a de donde se está haciendo el web scraping y una vez obtenida la clase se procede a poner en el código

```

4     var especificaciones=[];
5     var comentarios=[];
6
7     $("#tabSpecifications tr").each(function(rowIndex) {
8     $(this).each(function(cellIndex) {
9         var result={
10            llave:$(this).find("th").text().trim(),
11            valor:$(this).find("td").text().trim()
12        }
13        especificaciones.push(result);
14    });
15    });
16    if (context.request.label === "detalles") {
17        var precio=$( ".product .price.offer" ).text().trim()||$( ".product .price" ).text().trim();
18        precio=precio.replace(/[,]/g, '');
19        var result = {
20            //jumbo=1 exito=2 faLabella=3
21            //Portatiles =1 Videojuegos=2 Televisores=3 Celualres=4
22            idTiendaOnline:2,
23            idCategoria:3,
24            foto:$( ".product .image img" ).attr('data-src'),
25            nombre:$( ".name" ).text().trim(),
26            marca:$( ".brand" ).text().trim(),
27            precio:precio,
28            dcto:$( ".discount-rate" ).text().trim(),
29            especificaciones:especificaciones
30        }
31        return result;
32    }

```

Figura 12: Código Js del algoritmo de Web Scraping, autoría propia.

Como se puede ver en el screenshot vemos que la clase que contiene el nombre del producto se llama "name" ya con esto, procedemos a escribir el código para extraer ese nombre del portátil

Y ahora en el screenshot del arañador podemos ver como creamos una variable que contiene el texto que haya dentro de la clase "name" es decir el nombre del producto, así también se hace con las demás variables que queramos obtener en la extracción de los datos (precios, fotos, marca, etc.). Lo demás es con el método return devolver los resultados que obtuvimos del scraping. Información que es devuelta en formato JSON, ya que así lo definimos en la arañada, pero puede ser devuelta o transformada en diferentes formatos.

El web scraping nos devuelve una tabla de resultados con los datos que logro capturar de la web con estructura de datos Json.

URL	comentarios/0/fecha	especificaciones/0/llave	especificaciones/0/valor	especificaciones/1/llave
https://www.falabella.com.co/Falabella-co/product/3645267/MacBook-Air-11.6-Intel-Core-i5-4GB-128GB-M3VM2LL-A/3645267	hoy	Procesador	Intel Core i5	Modelo tarjeta de video
https://www.falabella.com.co/Falabella-co/product/3633682/Portatil-15.6-Intel-Core-i5-4GBx2816GB-INTEL-OPTANE-1TB-A315-53-55T1/3633682	hoy	Procesador	Intel Core i5	Modelo tarjeta de video
https://www.falabella.com.co/Falabella-co/product/3411326/Notebook-15.6-AMD-A9-4-GB-1-TB-X555BP-X02477/3411326	hoy	Procesador	AMD A9	Modelo tarjeta de video

especificaciones/2/llave	especificaciones/2/valor	especificaciones/3/llave	especificaciones/3/valor	especificaciones/4/llave
Memoria total (RAM + Intel Optane)	4GB	Tamaño de la pantalla	11.6	Disco duro
Memoria total (RAM + Intel Optane)	4GB+16GB INTEL OPTANE	Tamaño de la pantalla	15.6"	Disco duro
Tamaño de la pantalla	15.6"	Disco duro	1TB	Núcleos de

Figura 13: Tabla de resultados del algoritmo Web scraping, autoría propia.

Se creó un web scraping por cada categoría y de cada tienda, para obtener todos los datos necesarios de cada producto.

Automatización de los algoritmos web scraping. Lo que quiere decir que los web scraping con esta configuración se ejecutaran automáticamente todos los días a partir de las 7 pm hora colombiana con un intervalo de una hora entre cada uno

+ Create new Filter

Name ▲	Enabled	Cron expression	Last run	Next run	Crawlers	Actors	Tasks	
ExitoCelulares		0 4 * * *	2019-04-13 23:00:00		ExitoCelulares	RubenExitoCelulares		🔍
ExitoComputadores		0 5 * * *	2019-04-14 00:00:00		ExitoPortatiles	RubenExitoComputadores		🔍
ExitoTV		0 6 * * *	2019-04-14 01:00:00		ExitoTV	RubenExitoTV		🔍
ExitoVideoJuegos		0 7 * * *	2019-04-14 02:00:00		ExitoVideoJuegos	RubenExitoVideoJuegos		🔍
FalabellaCelulares		@daily	2019-04-13 19:00:17		FalabellaCelulares	ReneFalabellaCelulares		🔍
FalabellaComputadores		0 3 * * *	2019-04-13 22:00:00		FalabellaPortatiles	ReneFalabellaComputadores		🔍
FalabellaTV		0 1 * * *	2019-04-13 20:00:00		FalabellaTV	ReneFalabellaTV		🔍
FalabellaVideoJuegos		0 2 * * *	2019-04-13 21:00:00		FalabellaVideoJuegos	ReneFalabellaVideoJuegos		🔍
JumboCelulares		0 9 * * *	2019-04-14 04:00:00		jumboCelulares	DanielJumboCelulares		🔍
JumboComputadores		0 11 * * *	2019-04-14 06:00:00		JumboPortatiles	DanielJumboComputadores		🔍
JumboTV		0 8 * * *	2019-04-14 03:00:00		jumboTelevisores	DanielJumboTV		🔍
JumboVideoJuegos		0 10 * * *	2019-04-14 05:00:00		JumboVideoJuegos	DanielJumboVideoJuegos		🔍

Figura 14: Automatización de los algoritmos Web Scraping, autoría propia.

Automatización de la llamada al End-point que recoge todos los resultados de todas las arañas y llena la base de datos del sistema.

Proxy: None

Custom data:

Finish webhook URL: Test

Finish webhook data:

Cookies persistence: Per single crawling process only

Save

Figura 15: Automatización del End-point, autoría propia.

Ya con esto se puede observar que nuestra base de datos contiene los productos que se están extrayendo con los web scraping de las tiendas online.

	idProducto	idTiendaOnline	idCategoria	nombre	marca	precio	porcentajeDeDato	calificacion	urlSitio	foto
Editar Copiar Borrar	1	3	4	Celular Huawei Mate 20 Lite 64GB	Huawei	2543940	19	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...
Editar Copiar Borrar	2	3	4	Celular Samsung A50 64GB	Samsung	1314457	31	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...
Editar Copiar Borrar	3	3	4	Celular Samsung Galaxy S9	Samsung	1202145	41	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...
Editar Copiar Borrar	4	3	4	Celular Samsung Galaxy AT 2018 64 GB	Samsung	1974116	29	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...
Editar Copiar Borrar	5	3	4	Celular Motorola One 32GB	Motorola	1656401	46	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...
Editar Copiar Borrar	6	3	4	Celular Honor View 10 128GB	Honor	2070062	41	1	https://www.falabella.com.co/falabella-co/producto/...	https://falabella.scene7.com/is/image/FalabellaCO/...

Figura 16: Base de datos con los resultados de los algoritmos Web Scraping, autoría propia.

8.4. Objetivo 4: Desarrollar el chatbot utilizando los diseños de interacción y la base de conocimiento.


Después de consultar y comparar entre las diferentes plataformas que existen para el desarrollo conversacional del chatbot se ha seleccionado la plataforma dialogflow una plataforma que nos ofrece aprendizaje automático, despliegue del chatbot en diferentes plataformas de mensajería, una fácil interfaz de desarrollo y buena documentación que nos permite aprender rápido debido a nuestros pocos conocimientos acerca del desarrollo de chatbot.

A continuación se enumeran los pasos que se siguieron para el desarrollo del chatbot:

1. se crea un agente que es un módulo para la comprensión de lenguaje natural, y nos permite traducir las solicitudes de texto en datos procesables cuando la expresión que utiliza el usuario coincide con una intención que existe dentro del agente. Dentro del agente se configura el lenguaje que está permitido para la conversación. Para nuestro agente se llamara ChatBotProductosTeconologicos y aceptara lenguaje español. (Dialogflow 2019)

ChatBotProductosTeconologicos SAVE

General Languages ML Settings Export and Import Speech Share



DESCRIPTION

Describe your agent

DEFAULT TIME ZONE

(GMT-5:00) America/New_York

Date and time requests are resolved using this timezone.

GOOGLE PROJECT

Project ID	chatbotproductosteconologicos (Google Cloud Actions on Google)
Service Account	dialogflow-aaqbdm@chatbotproductosteconologicos.iam.gserviceaccount.com

Figura 17: Configuración del agente en dialogflow, autoría propia.

2. Para el entrenamiento de conversación del chatbot se crean intents que dentro del agente se utilizan para definir las conversaciones, dentro de estos se colocan las posibles palabras que el usuario va a utilizar y posibles repuestas que devolverá. (Dialogflow 2019)

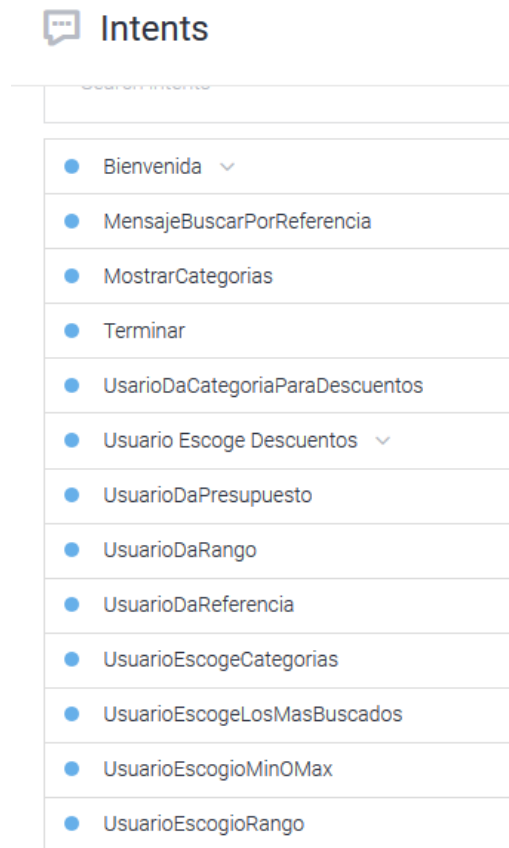


Figura 18: Intents creados para el funcionamiento del chatbot, autoría propia.

3. A su vez se crean las entidades que nos permiten seleccionar una información específica del lenguaje natural que reciben las intents, se utilizan para tomar los datos importantes que el usuario requiere en una solicitud. (Dialogflow 2019)

Entities

Search entities

@ Categoria

@ filtroPrecio

Figura 19: Entities creadas para el funcionamiento del chatbot, autoría propia.

4. Para que nuestro dialogflow interactúe con el back-end es necesario agregar un fulfillment que nos permite tener respuestas dinámicas basadas en el lenguaje natural que utiliza el usuario y hacer consultas a nuestra base de datos. Se debe agregar la dirección url de donde se encuentra alojado nuestro back-end y habilitar el webhook.

Fulfillment

Webhook

ENABLED

Your web service will receive a POST request from Dialogflow in the form of the response to a user query matched by intents with webhook enabled. Be sure that your web service meets all the [webhook requirements](#) specific to the API version enabled in this agent.

URL*	<u>https://www.botdecompras.com/chatbotproductostecnologicos/public/simple</u>	
BASIC AUTH	<input type="text" value="Enter username"/>	<input type="text" value="Enter password"/>
HEADERS	<input type="text" value="Enter key"/>	<input type="text" value="Enter value"/>
	<input type="text" value="Enter key"/>	<input type="text" value="Enter value"/>
	+ Add header	
SMALL TALK	Disable webhook for Smalltalk <input type="checkbox"/>	

Figura 20: Configuración del webhook para la conexión con el back-end, autoría propia.

Cuando es necesario que un intents envíe un parámetro al back-end o que devuelva respuesta se activa el webhook en la intents de esa manera logramos que nuestro chatbot interactúe con el back-end.

La plataforma de dialogflow permite la integración de chatbot para las plataformas de mensajería más destacadas del momento, entre las que se encuentra Facebook messenger que en la actualidad según el periódico Dinero en Colombia es la plataforma que más visitan los colombianos la cual tiene un aproxima de más de 20 millones de usuarios por esto para nuestro proyecto hemos decidido integrarla debido a que estaremos a la vista de una gran cantidad de usuarios que visitan Facebook (Cuáles son las aplicaciones más usadas por los colombianos. 2019.).

Para que nuestro chatbot se pueda mostrar en el messenger de Facebook se debe crear una App en la consola de desarrollo de Facebook y llegar los requisitos que nos exige para poder integrar webhook para messenger.(Documentos para desarrolladores de Facebook | API, SDK y guías de Facebook 2019)



Figura 21: Configuración para crear una App en Facebook, autoría propia.

Uno de los numerales que exige Facebook es crear nuestra página para que los usuarios puedan encontrarnos y sea nuestra identificación.



Figura 22: Pagina publica creada en Facebook

Para establecer la conexión entre dialogflow y la App creada en Facebook se debe:

1. Obtener su token de acceso a la página de Facebook e insértelo en el campo a continuación.
2. Crear tu propio token de verificación.
3. Hacer clic en 'START' a continuación.
4. Usar la URL de devolución de llamada y el token de verificación para crear un evento en la configuración de Facebook Messenger Webhook.

Callback URL	https://bots.dialogflow.com/facebook/7e2ebfe4-4cf7-48c2-85f4-a7c4db11e0c2/webhook
Verify Token	botdecompras
Page Access Token	EAAGFgVMWu3QBAGZBjn1vx0vGEIvyO6ZCTJSKtVS9AH7woiAtJ8nG1bNsHknnRgk3SdrqeZA
	<div style="border: 1px solid #ccc; padding: 5px 20px; display: inline-block;">STOP</div>

Figura 23: Conexión entre dialogflow y Messenger, autoría propia.

Para que nuestra aplicación sea pública y que los usuarios de Facebook puedan interactuar con el chatbot se debe enviar una solicitud a Facebook para que entre a revisión y nos den el respectivo permiso, para esto se deben crear las políticas de privacidad de la App, llenar formularios de datos personales del desarrollador y enviar un video del paso a paso que se sigue para utilizar el chatbot, también se

debe firmar contrato con Facebook aceptando las políticas que maneja para este tipo de desarrollos.



Figura 24: Solicitud para ser una App publica en Facebook, autoría propia.

Nuestro chatbot funciona de la siguiente manera:

Para hacer lo más semejante posible la conversación con el chatbot a una conversación natural se ha utilizado lenguaje cotidiano y de fácil entender para los usuarios.

Lo primero al iniciar nuestra conversación es saludar:

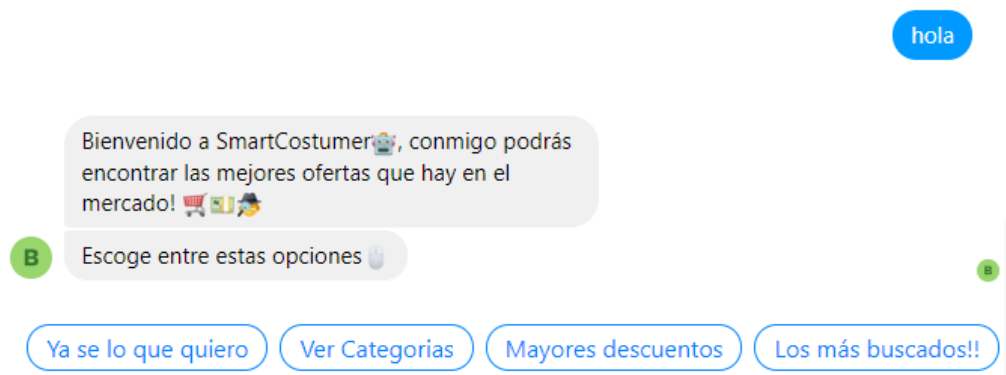


Figura 25: Inicio de conversación con el chatbot, autoría propia.

Como se observa el responderá con una bienvenida y cuatro posibles escenarios por los que el usuario puede optar.

Escenarios:

- Ya sé lo que quiero: este escenario permite al usuario hacer una búsqueda rápida del artículo el cual él ya conoce con una referencia exacta.

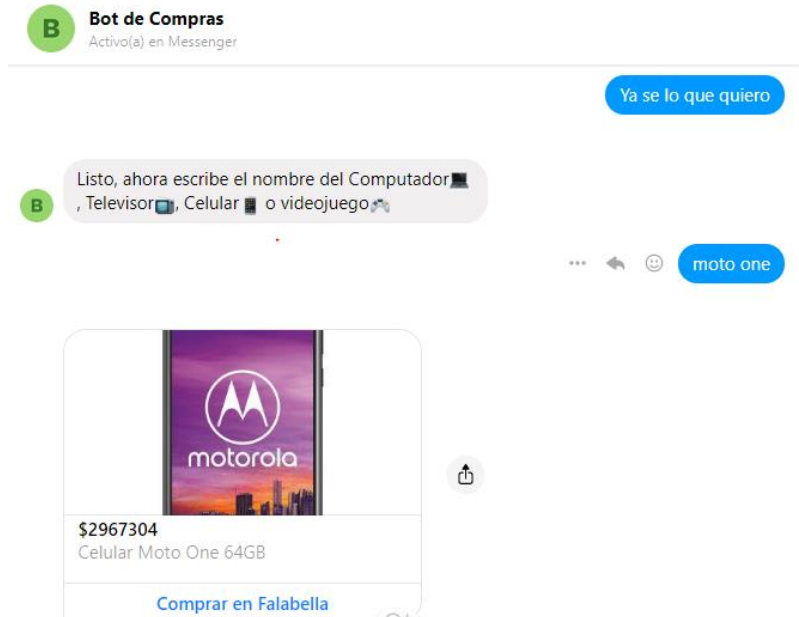
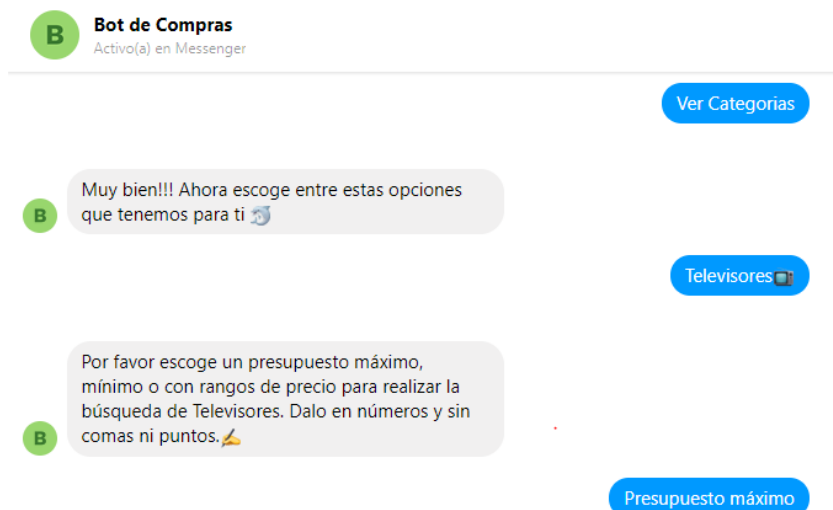


Figura 26: Búsqueda por el escenario “Ya sé lo que quiero”, autoría propia.

- Ver categorías: este escenario permite al usuario hacer una búsqueda más guiada de su producto llevándolo por las categorías que se ofrecen, utilizando un filtro de precio hasta llegar al productos que cumplen con lo que quiere el usuario.



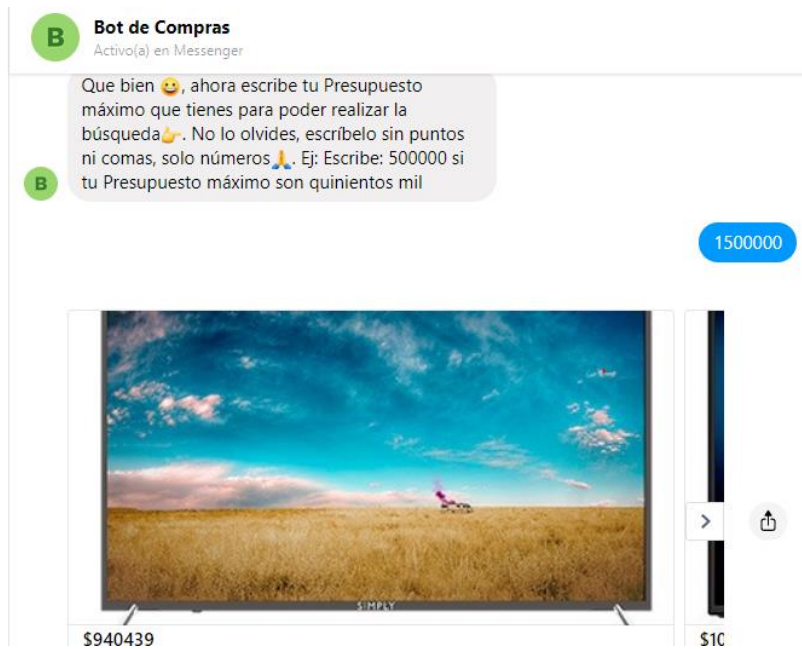


Figura 27: Búsqueda por el escenario de categorías, autoría propia.

- **Mayores descuentos:** cuando el usuario decide esta opción el chatbot lo dirige a los productos de la categoría que el selecciono con los mayores descuentos ofrecidos por las tiendas en el momento.

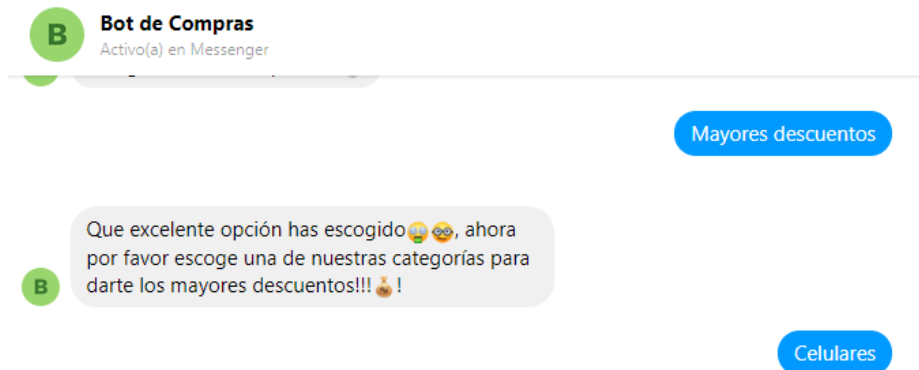




Figura 28: Búsqueda por el escenario de descuentos, autoría propia.

- Los más buscados: cuanto entramos a este escenario encontramos las tendencias de búsqueda de los usuarios que han visitado el chatbot.

Para nosotros tener una recolección de datos de las diferentes búsquedas que realizan los usuarios que interactúan con el chatbot se creó una tabla que contiene atributos de la búsqueda los cuales nos servirán para poder realizar análisis de datos y encontrar tendencias y presupuestos en los usuarios, así como las fechas y horas en que son más buscados los productos. Dentro de esta tabla no se guardan datos personales ni que identifiquen al usuario de alguna manera.

Busquedas ChatBot

#Busqueda	Referencia	Tipo Busqueda	Valor	Fecha	Hora
177	Huawei p20	Referencia exacta	0	30/05/2019	08:47am
176	Portatiles	Maximo	1200000	30/05/2019	12:26am
175	Televisores	Maximo	1500000	30/05/2019	12:07am
174	moto one	Referencia exacta	0	29/05/2019	11:59pm
173	Laptop hp	Referencia exacta	0	26/05/2019	05:36pm
172	VideoJuegos 🎮	Mínimo	20000000	26/05/2019	02:33pm
171	VideoJuegos 🎮	Rango	100-200	26/05/2019	02:31pm
170	p20 lite	Referencia exacta	0	26/05/2019	02:28pm
169	p20 lite	Referencia exacta	0	26/05/2019	02:24pm

Figura 29: Tabla con datos guardados de las búsquedas realizadas en el chatbot, autoría propia.

8.5. **Objetivo 5:** Realizar pruebas funcionales y de usuario del chatbot.

Para realizar pruebas de funcionamiento y recibir retroalimentación de los usuarios se pidió a los usuarios su opinión por medio de una encuesta la cual estaba estructurada de la siguiente manera.

Encuesta acerca de SmartCostumer

De 1 a 5 responda las siguientes preguntas siendo 1 el peor valor y 5 el mejor valor:

- ¿Es claro el propósito del chat bot?

1	2	3	4	5
---	---	---	---	---
- ¿La conversación con el chat bot fue fluida?

1	2	3	4	5
---	---	---	---	---
- La información presentada de los productos fue suficiente

1	2	3	4	5
---	---	---	---	---
- Usaría este chat bot para sus búsquedas futuras

1	2	3	4	5
---	---	---	---	---

Pregunta abierta:
¿Qué le mejoraría al chat bot? (Filtros, Tiendas, tipos de productos)

Figura 30: Preguntas para la encuesta de funcionamiento, autoría propia.

Las encuestas fueron tomadas durante el transcurso de nuestra presentación en IngeniaTe-C y presentación a algunos cursos de estudiantes de ingeniería de sistemas.

El valor de cada pregunta se da como 1 el peor valor y 5 el mejor valor de calificación.

- Pregunta 1:

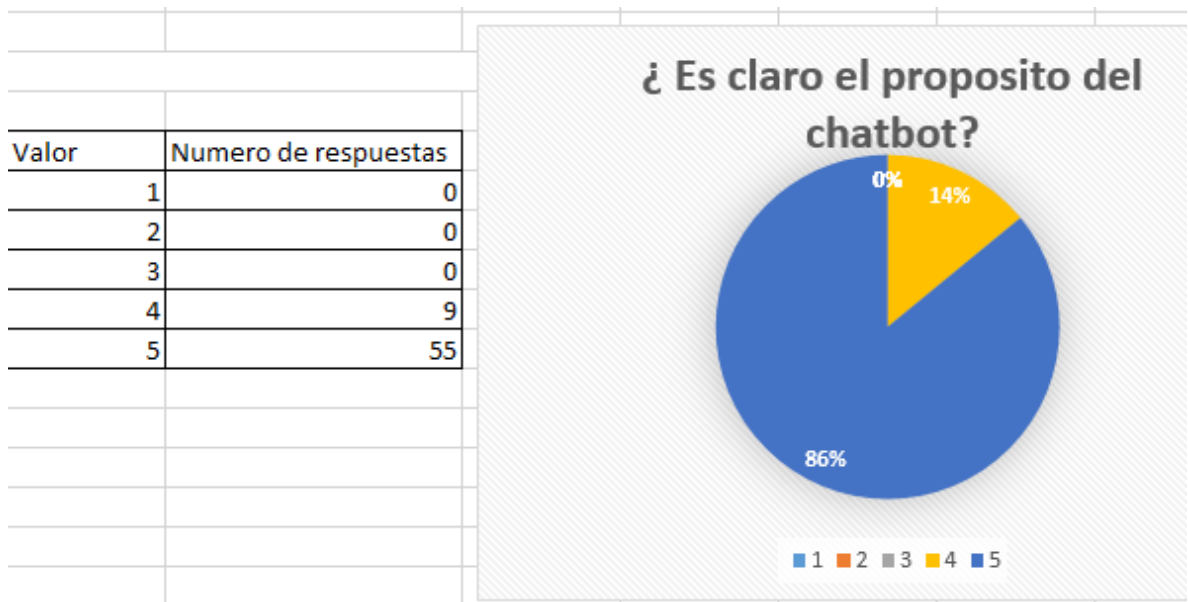


Figura 31: Grafico resultados pregunta uno de la encuesta de funcionamiento del chatbot, autoría propia

La mayoría de las personas comprendió para que se puede utilizar el chatbot.

- Pregunta 2:

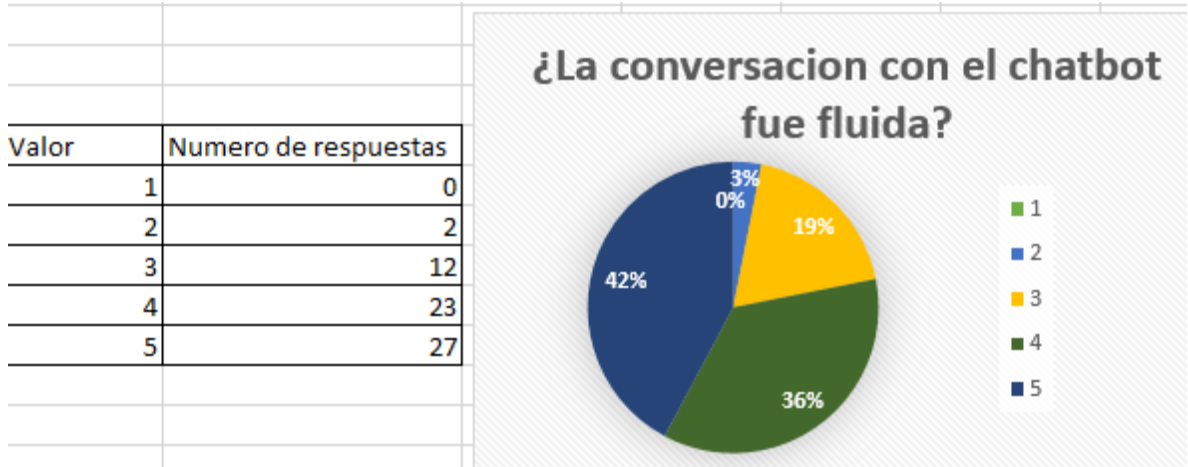


Figura 32: Grafico resultados pregunta dos de la encuesta de funcionamiento del chatbot, autoría propia

Algunas personas nos dieron críticas con respecto a el lenguaje que fuera más natural y que el chatbot diera un poco más de instrucciones.

- Pregunta 3:

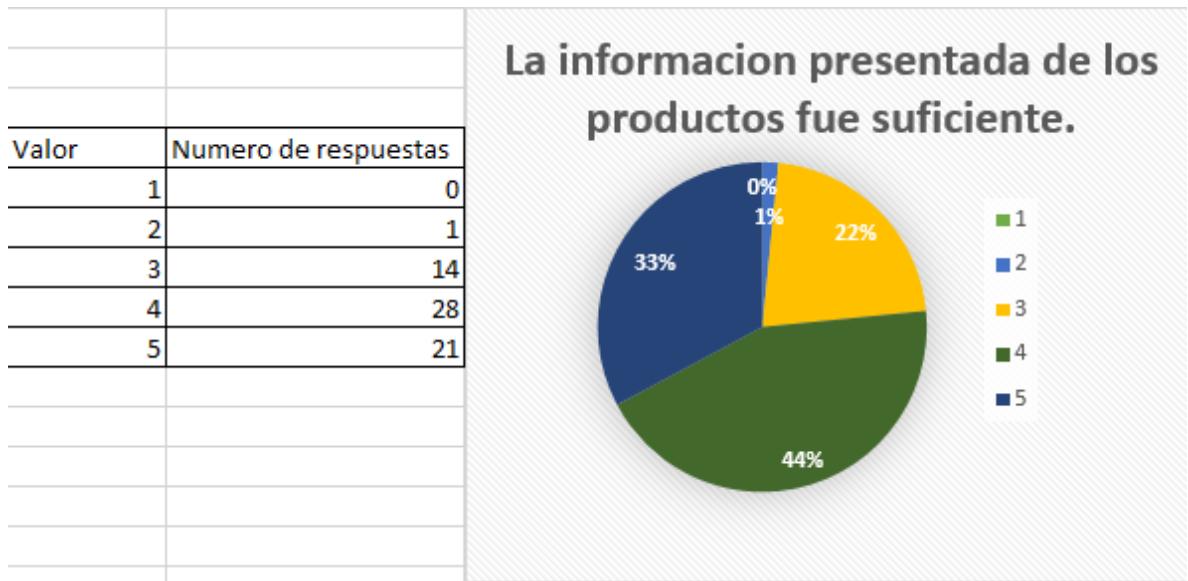
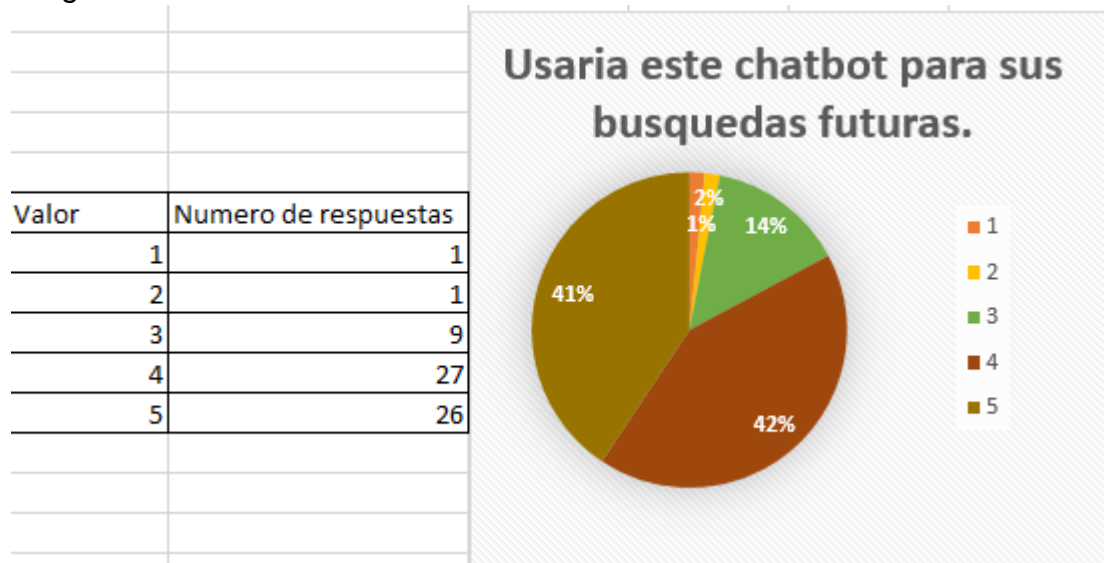


Figura 33: Grafico resultados pregunta tres de la encuesta de funcionamiento del chatbot, autoría propia

La calificación fue buena pero algunas personas quieren que se muestre más acerca de las especificaciones de los productos para conocer más acerca de ellos antes de dirigirse a la tienda.

- Pregunta 4:



- Pregunta 5:
¿Qué le mejoraría al chatbot?

En esta pregunta las personas que interactuaron con el chatbot nos dieron algunas ideas de mejoras como:

- Agregar más tiendas
- Agregar más productos de tecnología o igual de otros tipos de productos.
- Mostrar más información acerca del producto y de sus especificaciones dentro del chatbot antes de dirigirse a la tienda.
- Respecto a la conversación que sea más fluida y natural para tener una mejor comprensión.
- Que se deberían agregar más opciones de filtros para buscar los productos.
- Para los usuarios es importante saber cuándo se finaliza la conversación.
- Los usuarios al momento de hacer varias búsquedas o utilizar diferentes filtros quieren una manera más fácil de buscar sin que el chatbot se tenga que reiniciar con el saludo.

Como se puede evidenciar en los resultados de las gráficas el chatbot fue agradable por la mayoría de las personas que probaron su funcionamiento, la mayoría de críticas se recibieron fueron a que se debería mostrar en poco más de información acerca de los productos dentro del chatbot y que se deberían agregar más productos para tener más variedad.

9. Conclusiones

Construir el chatbot fue una experiencia muy enriquecedora para nosotros como ingenieros ya que pudimos poner en práctica muchas de las cosas que durante nuestra carrera aprendimos. (Bases de datos, programación, etc.) y no solo eso, sino que afrontamos retos nuevos ya que tecnologías como Web Scraping y DialogFlow fueron completamente nuevas para nosotros lo cual implicó no solo una etapa de desarrollo sino de conocimiento de las herramientas y adaptación a las mismas.

Por otra parte, uno de nuestros mayores aciertos fue escoger web Scraping como medio para obtener información de los productos ya que siendo miles los productos que se muestran por internet esa herramienta nos ayudó a obtenerlos de una manera más sencilla.

Adicionalmente, la construcción del chatbot haciendo uso de la herramienta creada por Google para esto (DialogFlow) fue algo bastante interesante ya que a veces se piensa que crear un chatbot es algo súper tedioso, o que se requieren demasiados

conocimientos en programación para ello y concluimos que por medio de esta herramienta muy bien estructurada en el flujo o la forma en la que trabaja una conversación se pudo realizar, por supuesto, no fue fácil pero una vez entendida su forma de trabajar o estructura básica se hizo todo el desarrollo del chatbot sin mayores percances que uno supone se afrontaría al crear un Bot.

Ahora, dimos un paso fuera del cascaron por decirlo así, ya que como desarrolladores de software siempre habíamos trabajado en un servidor local de nuestro pc, y aunque la mayoría de los proyectos de desarrollo de software empiezan de esta manera, hay algo que es alojar nuestro software, nuestro código, nuestro back-end, base de datos en internet de modo que no sea solo para nosotros sino universal, para cualquier persona. Y fue ahí, donde DialogFlow debido a sus características de funcionamiento nos "forzó" a alojar nuestro código y base de datos en internet ya que él debía poder acceder a nuestro código para poder devolver respuestas al usuario que usa nuestro Bot, entonces hubo la necesidad de usar un prestador de servicios web en la nube y el que usamos fue Amazon Web Services un servidor en blanco de Linux que se configuró con las características que necesitábamos para nuestro proyecto. Este proceso de alojar nuestros servicios web en un servidor que no sea el local es algo que como desarrolladores de software estamos destinados, obligados a hacer. Sólo que tuvimos esa ventaja ya que, aunque esto fue algo académico el proceso que se tuvo que hacer nos hizo acercarnos más al mundo real, al mundo de la producción de software; de modo tal que cuando nos enfrentemos a todos estos retos ya tenemos esta experiencia la cual nos podrá ahorrar muchos procesos que afrontamos cuando empezamos el mundo laboral.

10. Trabajo futuro

Finalmente, nuestro producto como tal, nuestra idea, tiene mucho por mejorar y añadirle, eso lo tenemos claro, pero es algo que el tiempo, el mercado, las personas y las nuevas tecnologías nos dirán si tienen un espacio en el mercado o el mundo de la comparación de precios de productos o servicios antes de adquirirlos.

Nuestro primer objetivo es que sea usado, y ya luego si funciona quisiéramos no solo seguir trabajando de la forma en que lo hicimos sino con maneras mucho óptimas como el tema de los datos, que ya no tengamos que nosotros ir a extraerlos sino que los negocios, las empresas vean un medio en nosotros para promover, vender sus productos aparte de sus respectivas páginas web y publicidad en internet y nos den la información de sus productos, para nosotros mostrarlos en nuestro Bot. De esa forma, no gastaríamos tiempo ni recursos obteniendo esa información, sino que los usaríamos en otros procesos mucho más productivos

como el de enriquecer el Bot o generando ideas acerca de este tipo de escenarios como el de cotizar productos antes de adquirirlos y aplicarlas a nuestro proyecto.

11. Agradecimientos.

Nuestro agradecimiento especial a Rene Alejandro Lobo Quintero por compartirnos sus conocimientos y experiencia lo cual nos sirvió muchísimo para poder crear el Bot; somos dos estudiantes exponiendo esta tesis, pero en sí fue entre tres personas que creamos este Bot. Y también a las personas quienes nos ayudaron con las pruebas de funcionamiento por su tiempo y aportes a nuestro proyecto.

12. Bibliografía

- Argal, A., Gupta, S., Modi, A., Pandey, P., Shim, S., & Choo, C. (2018). Intelligent travel chatbot for predictive recommendation in echo platform. En 2018 IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC 2018 (Vol. 2018-Janua, pp. 176-183). IEEE. <https://doi.org/10.1109/CCWC.2018.8301732>
- «AWS | Elastic compute cloud (EC2) de capacidad modificable en la nube». <https://aws.amazon.com/es/ec2/> (31 de mayo de 2019).
- B V S, U., Gaind, B., Kundu, A., Holla, A., & Rungta, M. (2017). ClassificationBased Adaptive Web Scraper. En 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 125-132). IEEE. <https://doi.org/10.1109/ICMLA.2017.0-168>
- Best Price Comparison shopping online in India - Pricemark.com. (s. f.). Recuperado 25 de octubre de 2018, a partir de <https://www.pricemark.com/>
- Bhawayuga, A., Fauzi, M. A., Pramukantoro, E. S., & Yahya, W. (2018). Design of E-commerce chat robot for automatically answering customer question. En Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology, SIET 2017 (Vol. 2018-Janua, pp. 159-162). IEEE. <https://doi.org/10.1109/SIET.2017.8304128>
- Botkit: Bloques de construcción para construir Bots. (2018). Recuperado 26 de octubre de 2018, a partir de <https://botkit.ai/>
- Botpress - Un Marco de Desarrollo de Chatbot Maker & Bot. (2017). Recuperado 26 de octubre de 2018, a partir de <https://botpress.io/>
- «Cuáles son las aplicaciones más usadas por los colombianos». <https://www.dinero.com/tecnologia/articulo/cuales-son-las-aplicaciones-mas-usadas-por-los-colombianos/269930> (30 de mayo de 2019).
- Dialogflow. (2017). Recuperado 26 de octubre de 2018, a partir de <https://dialogflow.com/>
- «Documentos para desarrolladores de Facebook | API, SDK y guías de Facebook». <https://developers.facebook.com/docs/> (31 de mayo de 2019).
- Haddaway, N. R. (2016). The use of web-scraping software in searching for grey literature. Grey Journal, 11(February), 186-190.

Indice precios tecnología 2016 - Linio. (2017). Recuperado 18 de septiembre de 2018, a partir de <https://www.linio.com.mx/sp/indice-precios-tecnologia-2016>

Julian, L. R., & Natalia, F. (2016). The use of web scraping in computer parts and assembly price comparison. En CONMEDIA 2015 - International Conference on New Media 2015 (pp. 1-6). IEEE. <https://doi.org/10.1109/CONMEDIA.2015.7449152>

«Kelkoo | Comparador de precios, guías de compra, ofertas de vacaciones online». <http://www.kelkoo.es/> (25 de octubre de 2018). *la decisión de compra del consumidor.*

Li, W. (2010). Study on the Typical Application of E-commerce. En 2010 International Conference on E-Business and E-Government (pp. 46-48). IEEE. <https://doi.org/10.1109/ICEE.2010.19>

Microsoft Bot Framework. (2017). Recuperado 26 de octubre de 2018, a partir de <https://dev.botframework.com/>

Mobile Phones | Mobile Prices in India | Online Mobile Shopping | 91mobiles.com. (s. f.). Recuperado 25 de octubre de 2018, a partir de <https://www.91mobiles.com/>

Pichponreay, L., Kim, J. H., Choi, C. H., Lee, K. H., & Cho, W. S. (2016). Smart answering Chatbot based on OCR and Overgenerating Transformations and Ranking. En International Conference on Ubiquitous and Future Networks, ICUFN (Vol. 2016-Augus, pp. 1002-1005). IEEE. <https://doi.org/10.1109/ICUFN.2016.7536948>

Rasa Stack: AI conversacional de código abierto. (2016). Recuperado 26 de octubre de 2018, a partir de <https://rasa.com/products/rasa-stack>

Slamet, C., Andrian, R., Maylawati, D. S., Suhendar, Darmalaksana, W., & Ramdhani, M. A. (2018). Web Scraping and Naïve Bayes Classification for Job Search Engine. En IOP Conference Series: Materials Science and Engineering (Vol. 288, p. 012038). IOP Publishing. <https://doi.org/10.1088/1757-899X/288/1/012038>

«Shopping.com UK - Find, Compare, and Buy Anything in Seconds». <http://uk.shopping.com/?sb=1&nrcd=1&rpu=1> (25 de octubre de 2018).

WEB SCRAPING – Emmita – Medium. (s. f.). Recuperado 25 de septiembre de 2018, a partir de <https://medium.com/@Emmitta/web-scraping-7f87930face4>