



UNIVERSIDAD DE GUAYAQUIL
FACULTAD DE INGENIERIA INDUSTRIAL
CARRERA DE INGENIERÍA EN TELEINFORMÁTICA

TRABAJO DE TITULACIÓN
PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERIA
EN TELEINFORMÁTICA

ÁREA
TECNOLOGÍA DE LOS ORDENADORES

TEMA
“PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT
UTILIZANDO WEBSCRAPING”

AUTORA
APOLO ARAGUNDI RUTH DENNISE

DIRECTOR DEL TRABAJO
ING. PLAZA VARGAS ANGEL MARCEL, MG

GUAYAQUIL, ABRIL 2021



**ANEXO XI.- FICHA DE REGISTRO DE TRABAJO
DE TITULACIÓN
FACULTAD DE INGENIERÍA INDUSTRIAL
CARRERA INGENIERÍA EN TELEINFORMÁTICA**



REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA			
FICHA DE REGISTRO DE TRABAJO DE TITULACIÓN			
TÍTULO Y SUBTÍTULO:	Prototipo de asistente virtual tipo Chatbot utilizando Webscraping		
AUTOR(ES) (apellidos/nombres):	Apolo Aragundi Ruth Dennise		
REVISOR(ES)/TUTOR(ES) (apellidos/nombres):	In. Castillo León Rosa Elizabeth / Ing. Vargas Plaza Ángel		
INSTITUCIÓN:	Universidad de Guayaquil		
UNIDAD/FACULTAD:	Facultad de Ingeniería Industrial		
MAESTRÍA/ESPECIALIDAD:			
GRADO OBTENIDO:	Ingeniera en Teleinformática		
FECHA DE PUBLICACIÓN:	28 de septiembre del 2021	No. DE PÁGINAS:	125
ÁREAS TEMÁTICAS:	Tecnología de los ordenadores		
PALABRAS CLAVES/ KEYWORDS:	Web Scraping, Facebook Messenger, Python, Asistente Virtual		
<p>Resumen (100-150 palabras): Actualmente nos enfrentamos una crisis de salud global que nos obliga a quedarnos en casa por más tiempo, el presente proyecto tiene como objetivo diseñar un prototipo de asistente virtual para búsqueda de productos de tecnología (categoría laptops) que ayude al usuario a consultar de manera rápida el artículo que desee comprar y sin que salga de casa. Los detalles de los productos son extraídos de páginas web de empresas muy reconocidas en el Ecuador como son: La Ganga, Artefacta, Novicompu, Pycca y Computron y que disponen del servicio de venta online. Para la extracción de datos se usará una técnica denominada web scraping que facilitara la recopilación de información y nos ayudará a que sea más precisa y sin errores.</p> <p>Abstract: We are currently facing a global health crisis that forces us to stay home longer, The objective of this project is to design a prototype of a virtual assistant to search for technology products (laptops category) that helps the user to quickly consult the item they want to buy and without leaving home. The details of the products are extracted from the websites of well-known companies in Ecuador such as: La Ganga, Artefacta, Novicompu, Pycca y Computron and that they have the online sales service. For data extraction a technique called web scraping will be used and it will help us to be more precise and without errors</p>			
ADJUNTO PDF:	SI X	NO	

CONTACTO CON AUTOR/ES:	Teléfono: 0998336510	E-mail: dennise.apoloa@ug.edu.ec
CONTACTO CON LA INSTITUCIÓN:	Nombre: Ing. Ramón Maquilón Nicola, MG	
	Teléfono:593-2658128	
	E-mail: direccionTi@ug.edu.ec	



**ANEXO XII.- DECLARACIÓN DE AUTORÍA Y DE
AUTORIZACIÓN DE LICENCIA GRATUITA INTRANSFERIBLE Y
NO EXCLUSIVA PARA EL USO NO COMERCIAL DE LA OBRA
CON FINES NO ACADÉMICOS
FACULTAD DE INGENIERÍA INDUSTRIAL
CARRERA INGENIERÍA EN TELEINFORMÁTICA**



LICENCIA GRATUITA INTRANSFERIBLE Y NO COMERCIAL DE LA OBRA CON
FINES NO ACADÉMICOS

Yo, **APOLO ARAGUNDI RUTH DENNISE**, con C.C. No. **0941168734**, certifico que los contenidos desarrollados en este trabajo de titulación, cuyo título es **“PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING.”** son de mi absoluta propiedad y responsabilidad, en conformidad al Artículo 114 del CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN*, autorizo la utilización de una licencia gratuita intransferible, para el uso no comercial de la presente obra a favor de la Universidad de Guayaquil.

APOLO ARAGUNDI RUTH DENNISE

C.C.No. 0941168734



**ANEXO VII.- CERTIFICADO PORCENTAJE
DESIMILITUD
FACULTAD DE INGENIERÍA INDUSTRIAL
CARRERA INGENIERÍA EN
TELEINFORMÁTICA**



Habiendo sido nombrado ING. ANGEL MARCEL PLAZA VARGAS, tutor del trabajo de titulación certifico que el presente trabajo de titulación ha sido elaborado por APOLO ARAGUNDI RUTH DENNISE, con mi respectiva supervisión como requerimiento parcial para la obtención del título de INGENIERO EN TELEINFORMÁTICA.

Se informa que el trabajo de titulación: PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING, ha sido orientado durante todo el periodo de ejecución en el programa Antiplagio URKUND quedando el 1 % de coincidencia.

The screenshot shows the Urkund web interface. On the left, document details for 'tesis2.docx' (ID: D112386988) are listed, including the presenter (DENNISE.APOLDA@ug.edu.ec) and the recipient (angel.plazav.ug@analysis.orkund.com). A message states that 1% of the 33 pages consist of text present in 2 sources. On the right, a table titled 'Lista de fuentes' (List of sources) shows two sources with their respective URLs. The main content area displays a snippet of text from a document, mentioning 'Fundamentación legal Para el desarrollo de este proyecto se buscará el apoyo legal de la Ley Orgánica de defensa del consumidor en el Capítulo I, capítulo 2 en donde señala que: "Todo persona natural o jurídica que, como destinatario final adquiera, utilice o disfrute bienes o servicios, o bien reciba oferta para ello." En la presente ley se menciona al consumidor, quien hace referencia al usuario, que es quien hace uso de un servicio público. Es decir, la ley orgánica de defensa del consumidor, además, resguarda a los usuarios que disponga de un servicio procedente de una organización o empresa pública. El consumidor depende de...

<https://secure.orkund.com/view/107085662-677068-371005>



Firmado electrónicamente por:

**ANGEL
MARCEL PLAZA
VARGAS**

ING. ANGEL MARCEL PLAZA
VARGASDOCENTE TUTOR
C.C. 0915953665
FECHA: 10/9/2021



**ANEXO VI. - CERTIFICADO DEL DOCENTE-TUTOR
DEL TRABAJO DE TITULACIÓN**

**FACULTAD DE INGENIERÍA INDUSTRIAL
CARRERA INGENIERÍA EN ELEINFORMÁTICA**



Guayaquil, 13 de septiembre del 2021.

Sr (a).

Ing. Annabelle Lizarzaburu Mora, MG.

Director (a) de Carrera Ingeniería en Teleinformática / Telemática

**FACULTAD DE INGENIERÍA INDUSTRIAL DE LA UNIVERSIDAD
DEGUAYAQUIL**

Ciudad. -

De mis consideraciones:

Envío a Ud. el Informe correspondiente a la tutoría realizada al Trabajo de Titulación

“PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING” del estudiante **APOLO ARAGUNDI RUTH DENNISE**, indicando queha (cumplido con todos los parámetros establecidos en la normativa vigente:

- El trabajo es el resultado de una investigación.
- El estudiante demuestra conocimiento profesional integral.
- El trabajo presenta una propuesta en el área de conocimiento.
- El nivel de argumentación es coherente con el campo de conocimiento.

Adicionalmente, se adjunta el certificado de porcentaje de similitud y la valoración del trabajo de titulación con la respectiva calificación.

Dando por concluida esta tutoría de trabajo de titulación, **CERTIFICO**, para los finespertinentes, que la estudiante está apta para continuar con el proceso de revisión final.

Atentamente,



Firmado electrónicamente por:

**ANGEL
MARCEL
PLAZA
VARGAS**

ING. ANGEL MARCEL PLAZA VARGAS, MSC
2021
CC: 0915953665

13 de septiembre del



ANEXO VIII.- INFORME DEL DOCENTE REVISOR

FACULTAD DE INGENIERÍA INDUSTRIAL CARRERA INGENIERÍA EN TELEINFORMÁTICA



Guayaquil, 23 de septiembre de 2021

Sr (a).

Ing. Annabelle Lizarzaburu Mora, MG.

Director (a) de Carrera Ingeniería en Teleinformática / Telemática

FACULTAD DE INGENIERÍA INDUSTRIAL DE LA UNIVERSIDAD DE GUAYAQUIL
Ciudad. -

De mis consideraciones:

Envío a Ud. el informe correspondiente a la REVISIÓN FINAL del Trabajo de Titulación **"PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING"** del estudiante **APOLO ARAGUNDI RUTH DENNISE**. Las gestiones realizadas me permiten indicar que el trabajo fue revisado considerando todos los parámetros establecidos en las normativas vigentes, en el cumplimiento de los siguientes aspectos:

Cumplimiento de requisitos de forma:

El título tiene un máximo de 8 palabras.

La memoria escrita se ajusta a la estructura establecida.

El documento se ajusta a las normas de escritura científica seleccionadas por la Facultad.

La investigación es pertinente con la línea y sublíneas de investigación de la carrera.

Los soportes teóricos son de máximo 5 años.

La propuesta presentada es pertinente.

Cumplimiento con el Reglamento de Régimen Académico:

El trabajo es el resultado de una investigación.

El estudiante demuestra conocimiento profesional integral.

El trabajo presenta una propuesta en el área de conocimiento.

El nivel de argumentación es coherente con el campo de conocimiento.

Adicionalmente, se indica que fue revisado, el certificado de porcentaje de similitud, la valoración del tutor, así como de las páginas preliminares solicitadas, lo cual indica el que el trabajo de investigación cumple con los requisitos exigidos.

Una vez concluida esta revisión, considero que el estudiante está apto para continuar el proceso de titulación. Particular que comunicamos a usted para los fines pertinentes.

Atentamente,



Firmado electrónicamente por:
**ROSA
ELIZABETH
CASTILLO LEÓN**

ING. ROSA ELIZABETH CASTILLO LEÓN, MG.

C.C: 0922372610

FECHA: 23 DE SEPTIEMBRE DE 2021

Dedicatoria

Primeramente, a Dios, por darme las fuerzas y las ganas durante todo este proceso, porque he sentido su presencia y su manifestación en cada paso. A mi madre Sara Aragundi Jara por ser el motor que necesito para seguir adelante, quien ha estado presente en cada etapa de mi vida, que me ha guiado con amor y con paciencia y me ha formado como una mujer con valores. A Joffre Suarez Acosta, mis hermanos, mi papá, mis tíos y primos que me han apoyado y depositaron en mí la confianza que necesitaba para seguir, así mismo, este proyecto me lo dedico a mí, por el sacrificio que día a día dedicaba y que me ha permitido culminar esta etapa académica con éxito.

Agradecimiento

Mis agradecimientos en primera instancia al ser perfecto y poderoso, Dios, porque el tiempo de él es perfecto y en su voluntad me ha permitido concluir este trabajo de titulación. Le agradezco a mi familia por ser incondicional conmigo en todo momento, especialmente mi mamá que siempre ha depositado su confianza en mí y sus oraciones me han permitido llegar hasta donde estoy ahora. A mis profesores, mi tutor de tesis por guiarme durante todo este proceso. A mi jefa, Narcisa Vásquez, por ser siempre comprensible durante mi formación académica. Gracias a mis amigos y compañeros de universidad por mostrar su apoyo cuando lo he necesitado. Muchas gracias a todos por ser algo fundamental en mi vida.

Índice General del Contenido

N°	Descripción	Pág.
	Introducción	1

Capítulo I

El problema

N°	Descripción	Pág.
1.	El problema	3
1.1.	Planteamiento de Problema	3
1.1.1.	Pregunta principal de la investigación	4
1.2.	Objetivo de la investigación	4
1.2.1.	Objetivo general	4
1.2.2.	Objetivos Específicos	4
1.3.	Justificación	4
1.4.	Alcances	5
1.5.	Limitaciones	5

Capítulo II

Marco teórico

N°	Descripción	Pág.
2.	Marco Teórico	6
2.1.	Antecedentes de estudio	6
2.1.1.	Antecedentes Internacionales	6
2.1.2.	Antecedentes Nacionales	10
2.2.	Fundamentación teórica	13
2.2.1.	Inteligencia Artificial	13
2.2.1.1.	Machine Learning	13
2.2.1.1.1.	Aprendizaje supervisado	14
2.2.1.1.2.	Aprendizaje no supervisado	14
2.2.1.1.3.	Aprendizaje Reforzado	15
2.2.1.2.	Deep Learning	15
2.2.1.3.	Natural Language Processing (NLP)	16

Nº	Descripción	Pág.
2.2.1.3.1.	Logicos	16
2.2.1.3.2.	Probabilistico	16
2.2.1.4.	Agentes conversacionales o Chatbot	17
2.2.2.	ManyChat	19
2.2.2.1.	Tipos de interfaces	19
2.2.2.1.1.	Basic Builder	19
2.2.2.1.2.	Flow Builder	20
2.2.2.2.	Entendimiento del Lenguaje Natural	20
2.2.2.3.	Metodos de Entrada	21
2.2.2.4.	Metodos de Salida	21
2.2.3	Google Drive	21
2.2.3.1	Botsheets	21
2.2.4.	Faebook Messenger	22
2.2.4.1.	Facebook Messenger Bots	23
2.2.5.	Web Scraping	23
2.2.5.1.	Técnicas utilizadas para web scraping	24
2.2.5.1.1.	Recolección manual	24
2.2.5.1.2.	Expresiones regulares	25
2.2.5.1.3.	Protocolo HTTP	25
2.2.5.1.4.	Parsers de HTML	25
2.2.5.1.5.	Reconociemiento de web semántica	25
2.2.5.1.6.	Aplicaciones para web scraping	25
2.2.5.2.	Herramientas para realizar Web scraping	25
2.2.5.2.1.	Usuarios no programadores	25
2.2.5.2.2.	Usuarios programadores	26
2.2.5.3.	Impedimentos al utilizar Web scraping	27
2.2.5.3.1.	Limitaciones al usar Web scraping	27
2.2.5.3.2.	Limitaciones legales al usar Web scraping	28
2.2.5.3.3.	Técnicas para protegerse del Web scraping	29
2.2.6.	Lenguaje de programación	29
2.2.6.1.	Python	29
2.2.7.	Comercio electronico	30

N°	Descripción	Pág.
2.2.7.1.	Tipo de venta E-commerce	31
2.2.7.1.1.	Business-to-Consumer (B2C)	31
2.2.7.1.2.	Business-to-Business (B2B)	32
2.2.7.1.3.	Consumer-to-Business (C2B)	32
2.2.7.1.4.	Consumer-to-Consumer (C2C)	33
2.2.7.1.5.	Business-to-Administration (B2A)	33
2.2.7.2.	Tipo de plataforma según E-commerce	34
2.2.7.2.1.	Sitio Web	34
2.2.7.2.2.	Mobile e-commerce	34
2.2.7.2.3.	Social e-commerce	35
2.2.7.3.	Tiendas online con productos tecnológicos en Ecuador	36
2.2.7.3.1.	Artefacta	36
2.2.7.3.2.	La Ganga	37
2.2.7.3.3.	Novicompu	37
2.2.7.3.4.	Computron	38
2.2.7.3.5.	Pycca	38
2.3.	Fundamentación legal	39

Capítulo II

La Propuesta

N°	Descripción	Pág.
3.1.	Metodología del proyecto	42
3.1.1.	Diseño de la investigación	42
3.1.1.1.	Enfoque mixto	43
3.1.1.2.	Diseño cuasi experimental	44
3.1.2.	Tipo de investigación	44
3.1.2.1.	Investigación aplicada	44
3.1.2.2.	Investigación exploratoria	45
3.1.2.3.	Investigación descriptiva	45
3.1.3.	Diseño de la investigación y método de muestreo	45
3.1.3.1.	Diseño de la investigación	45
3.1.3.1.1	Diseño probabilísticos	45

N°	Descripción	Pág.
3.1.3.1.2	Diseño no probabilísticos	46
3.1.3.2	Método de muestreo	46
3.1.4.	Técnica: Encuesta	46
3.1.4.1.	Población y muestra	46
3.1.4.1.1	Población	46
3.1.4.1.2	Muestra	47
3.1.4.2	Análisis de la encuesta	47
3.1.5.	Método y técnica	54
3.1.5.1.	Método de PLN y extracción de datos	54
3.1.5.1.1.	Método de reconocimiento de nombres de entidades no anidados	54
3.1.5.1.2.	Método de extraer datos	54
3.1.5.2.	Técnica de PLN y Web scraping	55
3.1.5.2.1.	Técnica utilizada para el PLN	55
3.1.5.2.2.	Técnica para extraer datos	55
3.2.	Descripción del proyecto	55
3.2.1.	Factibilidad técnica	56
3.2.2.	Factibilidad legal	57
3.2.3.	Factibilidad económica	57
3.2.4.	Factibilidad Operacional	58
3.3.	Esquema general de la extracción de datos y asistente virtual	58
3.4.	Recursos para el desarrollo del proyecto	59
3.4.1.	Recursos para la extracción de datos	59
3.4.1.1.	Scrapy	60
3.4.1.2.	Páginas web	60
3.4.1.3.	Lenguaje de programación: Python	61
3.4.1.4.	Formato CSV	61
3.4.2.	Recursos para la creación del asistente virtual	62
3.4.2.1.	Página de Facebook	62
3.4.2.2.	Botsheets	63
3.4.2.3.	Manychat	63
3.5.	Procedimiento: Aplicación de la metodología escogida	64
3.5.1.	Identificación de Sitios web	64

N°	Descripción	Pág.
3.5.2.	Web scraping	65
3.5.3.	Limpieza de datos	67
3.5.4.	Cargar los datos	67
3.5.5.	Crear y conectar el bot con la base de datos y Facebook Messenger	67
3.5.5.1.	Conectar el bot con la página de Facebook Messenger	67
3.5.5.2.	Conectar el bot con la base de datos	68
3.5.5.3.	Creación del Asistente Virtual	70
3.6.	Pruebas del funcionamiento del prototipo	75
3.7.	Conclusiones	80
3.8.	Recomendaciones	81
	Glosario de palabras	82
	Anexos	83
	Bibliografía	103

Índice de tablas

N°	Descripción	Pág.
1.	Trabajos internacionales aplicando técnica de web scraping y asistente virtual.	8
2.	Trabajos nacionales aplicando técnica de web scraping y asistente virtual.	12
3.	Características de herramientas para la creación de chatbot.	18
4.	Características de plataformas de mensajería instantánea.	23
5.	Características de Python	30
6.	Estudiantes de la carrera de Ingeniería en Telemática	47
7.	Frecuencia de la pregunta ¿La herramienta le proporcionó la información de manera rápida?	47
8.	Frecuencia de la pregunta ¿Considera usted que el flujo de búsqueda es claro y entendible?	48
9.	Frecuencia de la pregunta ¿Tuvo algún inconveniente al momento de realizar la búsqueda del producto?	49
10.	Frecuencia de la pregunta ¿La herramienta le ayudó a reducir el tiempo de cotización del producto?	50
11.	Frecuencia de la pregunta ¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?	51
12.	Frecuencia de la pregunta ¿Considerarías usar un asistente virtual para sus búsquedas futuras?	52
13.	Frecuencia. Coloque del 1 al 5 que tanto le agrado la herramienta	53
14.	Hardware usado para el desarrollo del proyecto	56
15.	Hardware usados en la extracción de datos	57
16.	Software usados para el asistente virtual	57
17.	Costo total del requerimiento del software	57
18.	Comparativa de herramientas y servicios de raspado	59
19.	Páginas web para extraer datos	60

Índice de Figuras

Nº	Descripción	Pág.
1.	Como funciona Machine Learning	14
2.	Deep Learning	15
3.	Usos de Natural Language Processing	17
4.	Icono ManyChat	19
5.	Interfaz ManyChat	20
6.	Botsheets	22
7.	Evolución del logo Messenger	23
8.	Que es Web scraping	24
9.	Como extraer datos de un sitio web	27
10.	e-commerce	31
11.	Comercio electrónico B2C	31
12.	Business-to-Business	32
13.	Comportamiento de compra de negocios	33
14.	Consumer-to-consumer	33
15.	Plataformas e-commerce	34
16.	Mobile e-commerce	35
17.	Comercio electrónico	36
18.	Artefacta	37
19.	La Ganga	37
20.	Novicumpu	38
21.	Computron	38
22.	Pycca	39
23.	Diseño de la investigación	42
24.	Enfoque Mixto	43
25.	Diseño cuasiexperimental	44
26.	Gráficos estadísticos de la pregunta ¿La herramienta le proporciono la información de manera rápida?	48
27.	Gráficos estadísticos de la pregunta ¿Considera usted que el flujo de búsqueda es claro y entendible?	49
28.	Gráficos estadísticos de la pregunta ¿Tuvo usted algún inconveniente al momento de realizar la búsqueda del producto?	50

N°	Descripción	Pág.
29.	Gráficos estadísticos de la pregunta ¿La herramienta le ayudo a reducir el tiempo de cotización del producto?	51
30.	Gráficos estadísticos de la pregunta ¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?	52
31.	Gráficos estadísticos de la pregunta ¿Considerarías usar un asistente virtual para sus búsquedas futuras?	53
32.	Gráficos estadísticos Coloque del 1 al 5 que tanto le agrado la herramienta.	54
33.	Reconocimiento de nombres no anidados	54
34.	Pasos web scraping	55
35.	Esquema de la extracción de datos y asistente virtual	59
36.	Formato csv	62
37.	Datos estadísticos redes sociales	63
38.	Datos seleccionados de extracción	64
39.	Estructura HTML del producto	65
40.	Definir campos python	66
41.	Llenar ítems de extracción	66
42.	Limpiar texto	67
43.	Datos CSV	67
44.	Página de Facebook	68
45.	Herramienta ManyChat	68
46.	Conexión con Botsheets	69
47.	Conexión con hoja de datos	69
48.	Base de datos	70
49.	Crear campo de usuario	70
50.	Creación de flujo	71
51.	Mensaje de presentación	71
52.	Flujo acción	72
53.	Flujo acción configurado	72
54.	Flujo condición	73
55.	Siguiente consulta	73
56.	Mensaje final	74

57.	Flujo final	74
58.	Flujo para búsquedas que no se encuentran	75
59.	Palabras claves	75
60.	Saludo de inicio del chatbot	76
61.	Consultas del chatbot	77
62.	Segunda consulta	77
63.	Tiempo de consulta del chatbot	78
64.	Dato no encontrado	78
65.	Tercera consulta	79
66.	Fin del chat	80

Índice de Anexos

Nº	Descripción	Pág.
1.	Ley orgánica de defensa del consumidor	84
2.	Constitución de la república del Ecuador	85
3.	Ley Orgánica de Protección de Datos Personales	87
4.	Manual: Crear una cuenta de ManyChat	89
5.	Manual: Pasos para instalar Python	92
6.	Manual: Pasos para instalar Scrapy	96
7.	Manual: Pasos para usa la herramienta de asistente virtual	98
8.	La encuesta	101



ANEXO XIII.- RESUMEN DEL TRABAJO DE TITULACIÓN (ESPAÑOL)



FACULTAD DE INGENIERÍA INDUSTRIAL CARRERA INGENIERÍA EN TELEINFORMÁTICA

“PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING.”

Autor: Apolo Aragundi Ruth Dennise

Tutor: Ing. Comp. Plaza Vargas ángel Marcel, MSc.

Resumen

Actualmente nos enfrentamos una crisis de salud global que nos obliga a quedarnos en casa por más tiempo, el presente proyecto tiene como objetivo diseñar un prototipo de asistente virtual para búsqueda de productos de tecnología (categoría laptops) que ayude al usuario a consultar de manera rápida el artículo que desee comprar y sin que salga de casa. Los detalles de los productos son extraídos de páginas web de empresas muy reconocidas en el Ecuador como son: La Ganga, Artefacta, Novicompu, Pycca y Computron y que disponen del servicio de venta online. Para la extracción de datos se usará una técnica denominada web scraping que facilitara la recopilación de información y nos ayudará a que sea más precisa y sin errores.

Palabras Claves: Web Scraping, Facebook Messenger, Python, Asistente Virtual.



**ANEXO XIV.- RESUMEN DEL TRABAJO DE
TITULACIÓN (INGLÉS)**



**FACULTAD DE INGENIERÍA INDUSTRIAL
CARRERA INGENIERÍA EN TELEINFORMÁTICA**

**“PROTOTYPE OF A CHATBOT-LIKE VIRTUAL ASSISTANT USING
WEBSCRAPING”**

Author: Apolo Aragundi Ruth Dennise

Advisor: SE Plaza Vargas ángel Marcel, MSc.

Abstract

We are currently facing a global health crisis that forces us to stay home longer, The objective of this project is to design a prototype of a virtual assistant to search for technology products (laptops category) that helps the user to quickly consult the item they want to buy and without leaving home. The details of the products are extracted from the websites of well-known companies in Ecuador such as: La Ganga, Artefacta, Novicompu, Pycca y Computron and that they have the online sales service. For data extraction a technique called web scraping will be used and it will help us to be more precise and without errors.

Keywords: Web Scraping, Facebook Messenger, Python, Virtual Assistant.

Introducción

El proceso para realizar una compra en los consumidores ha ido cambiando constantemente con el tiempo cuando pretenden conseguir algún producto o artículo por lo que se ven en la necesidad de salir y cotizar en varios establecimientos hasta encontrar el lugar ideal que tenga el producto de preferencia y que se ajuste a su presupuesto para efectuar su compra.

En la actualidad se vive una emergencia sanitaria sin retorno que obliga a las personas a realizar nuevas actividades que no tenían acostumbrados por cumplir el aislamiento. Una de ellas el hacer compras frecuentes pagando en línea, lo que para muchos quizás sea algo cotidiano para otros significa experiencias nuevas. El comercio habitual ha optado por acogerse a nuevas estrategias para distribuir sus productos, como el de efectuar pedidos hasta el domicilio e implementar e-commerce o también conocido como comercio electrónico para así conservar el rendimiento de sus negocios. (Rodríguez, Ortiz, Quiroz & Parrales, 2020)

Por lo que las ventas online se han convertido en una estrategia muy importante en el área comercial y es por ello que muchas tiendas o empresas han visto en el marketing digital una oportunidad para lograr mantenerse en el mercado y no dejar de generar ingresos mientras sigan las restricciones.

El proceso de intercambio de información o la comunicación que se da a través de las redes sociales se ha transformado en un apoyo directo de las empresas en cualquier lugar del mundo ya que permite una conexión inmediata con los consumidores y el desarrollo de publicidad es suma importancia para cualquier emprendimiento. Las situaciones de catástrofes inesperadas en una empresa pueden ser unas tras otras, pero siempre hay que contar con destrezas y prácticas necesarias para asumirlas para adaptarse a los requerimientos y necesidades de los consumidores. Es por ello que durante la emergencia sanitaria las empresas decidieron optar por el marketing digital es decir a publicidad en redes sociales y páginas web ya que les permite tener una comunicación con los consumidores sin necesidad que sea de manera presencial sino a través de imágenes, post de promociones, ofertas, audios, llamadas con la ayuda de los equipos de telecomunicaciones y así realizar los pedidos y para ellos cada empresa han podido adoptar sus estrategias de envío ya sean estos con costo adicional o gratis. (Suárez, 2020)

De tal manera que, debido a tanto confinamiento muchas empresas han tenido que adaptarse a nuevas estrategias de marketing para buscar otro tipo de comercialización y que los usuarios puedan cotizar y buscar los productos en los diferentes sitios web o redes sociales en donde las empresas publican sus artículos y mejores ofertas sin que el consumidor tenga que salir de su domicilio.

Es por ello que en este documento se presenta la implementación de un prototipo de asistente virtual que permita optimizar el tiempo de búsqueda de un producto. Estos productos serán extraídos con la técnica web scraping para que los datos de consulta que genere el consumidor sean exactos y sin errores.

Capítulo I

El problema

1.1 Planteamiento del problema

El efecto de covid-19 es el causante del confinamiento y el cierre físico de locales lo que provocó un gran crecimiento en el área de e-commerce. El consumo del comercio electrónico se convirtió en algo esencial en sectores como la alimentación, salud y limpieza. (Revista Espacios, 2020)

En un estudio realizado se comprobó que el 13% de personas encuestadas solo realizaba una compra online una vez por semana, pero luego de la emergencia sanitaria y debido a las restricciones este porcentaje se duplico. Pese a todos estos acontecimientos se puede asegurar que el comercio electrónico se conservará hasta después de la emergencia sanitaria ya que un 40% de encuestados aseguro que comprarán en tiendas físicas, pero con poca frecuencia, mientras que el 32,4% se mantendrán con las compras en línea. Es por ello que muchas empresas ya sean grandes o pequeñas han optado por seguir con este tipo de comercio online teniendo a su disposición paginas o sitios web en donde exponen los diferentes artículos y productos que tiene a la venta por lo que el usuario podrá navegar por cada una de las páginas que dispongan los artículos de interés y hacer las respectivas comparaciones para escoger el artículo que mejor se ajuste a sus necesidades y a su bolsillo. Pero esto puede ocasionar fatiga por el tiempo que se toma en buscar en cada tienda y realizar la comparación de cada uno de los productos. (CECE, 2020)

Para un usuario es necesario el poder cotizar y comparar productos antes de adquirirlos y así tener la mejor opción en cuanto a calidad y precio referente al dispositivo que sea de su necesidad ya que en muchas ocasiones un mismo dispositivo en una tienda no va a tener el mismo costo en otra tienda por motivo de que cada empresa maneja estrategias diferentes que les permite mejorar sus ingresos.

El desarrollo de un chatbot que se plantea podrá mejorar la búsqueda de dispositivos tecnológicos que ofrezcan las diferentes tiendas online con el que el usuario podrá comparar precios y características del artículo que requiera y así podrá minimizar el tiempo de toma de decisiones y adquirir el producto acorde a sus necesidades y presupuesto.

Actualmente se busca poder facilitar el poder obtener gran cantidad de datos de manera automatizada y minimizar los errores que se puede ocasionar al tratar de obtener información

de manera manual. Desarrollar con rapidez técnicas, métodos o procesos nos ayuda a reducir carga de trabajos.

1.1.1 Pregunta principal de la investigación

¿Qué funcionalidades puede tener un asistente virtual para optimizar el tiempo de búsqueda y comparación de artículos tecnológicos?

1.2 Objetivo de la investigación

1.2.1 Objetivo general

Implementar un prototipo de chatbot basado en web scraping para la comparación de productos tecnológicos (categoría laptops).

1.2.2 Objetivos Específicos

- Investigar los sitios web de las empresas que venden productos de tecnología.
- Determinar que framework es el adecuado usar para la extracción de datos.
- Determinar que lenguaje de programación se adecua a la técnica que se requiere implementar
- Desarrollar metodología para recuperación de datos.
- Implementar prototipo de chatbot en la plataforma seleccionada.
- Evaluar funcionalidad de chatbot.

1.3 Justificación

Los procesos automatizados y precisos es una oportunidad que no se debe desaprovechar al momento de manejar datos ya que estos son muy importantes en cualquier ámbito y el raspado web nos ayuda a obtener información de manera más rápida y sin errores humanos.

Al implementar estos datos en un chatbot podremos dar una buena atención a los usuarios que los ayude a optimizar el tiempo de búsqueda de artículos ya que es una herramienta inteligente capaz de dar respuestas a las solicitudes de los consumidores proporcionando complacencia y efectividad al momento de querer adquirir un dispositivo lo que será de gran comodidad para los usuarios ya que esta se implementara en Facebook Messenger, plataforma gratuita, muy utilizada mundialmente y que se adapta a cualquier tipo de

dispositivo facilitando la manipulación a cualquier tipo de usuario que disponga de una cuenta en esta red social.

1.4 Alcances

Al obtener las páginas web con el contenido necesario se alcanzará la recolección de datos que son importante para el uso de una herramienta que permita realizar una búsqueda automatizada de artículos tecnológicos con sus respectivas características minimizando el tiempo de búsqueda que puede tomarle a un usuario si lo hace de manera manual y así poder satisfacer las necesidades de los consumidores.

Se presentará un prototipo de asistente virtual tipo chatbot que se implementara en una plataforma seleccionada que tendrá como objetivo dar un servicio de búsqueda automatizada para los usuarios que requieran adquirir un producto o dispositivo tecnológico y reducir el tiempo de cotización y comparación con datos exactos de sitios web de ventas de tecnología.

1.5 Limitaciones

Se debe tener en cuenta la complejidad de cada sitio web de donde se va a extraer la información ya que existen limitaciones al utilizar este tipo de herramientas ya que hay páginas web que pueden ser más completas que otras lo que puede provocar dificultar al momento de extracción de datos. También es recomendable tener en cuenta los términos y condiciones de cualquier sitio web antes de aplicar dicha técnica ya que algunas pueden detectar con facilidad el proceso.

CAPITULO II

2 MARCO TEORICO

2.1 Antecedentes de estudio

El enfoque de este capítulo se centra en la síntesis de antecedentes que son importantes para la realización de un profundo estudio de esta investigación sobre el uso de un asistente virtual o chatbot en algunos ámbitos y la importancia de obtener información a partir de una herramienta que permita extraer datos de manera precisa y sin errores, esto con el beneficio de ahondar nuestro modelo de estudio.

2.1.1. Antecedentes internacionales

Según (Chesñear & Estevez, 2018) en su artículo expreso el interés que debe tener el comercio electrónico en aplicar una herramienta como la inteligencia artificial ya que puede ser capaz de captar de mejor manera las necesidades de los consumidores, asimismo, detalla uno de los múltiples beneficios que ha desarrollado durante la última década, el aprendizaje automatizado que no es más que técnicas computacionales que permite a la maquina adquirir conocimiento a partir de grandes cantidades de datos y para ello se utiliza la minería de datos que nos facilita el proceso de obtención de grandes cantidades de información con las cuales se puede hacer clasificaciones de dichos datos. Así es como la inteligencia artificial favorece al progreso del comercio electrónico porque pueden lograr mejorar la experiencia de los consumidores al instante de pretender adquirir un producto o servicio de manera online de una forma simple e inmediata. Es importante reconocer el tipo de chatbot que sea de mejor utilidad para una entidad dependiendo de las necesidades que requiera cubrir y mejorar la experiencia. El uso de técnicas como la inteligencia artificial nos ayuda a tener soluciones eficientes por ello este tipo de herramientas pueden seguir en evolución e integrar otras capacidades como reconocer emociones a través de su tono de voz. El comercio electrónico es uno de los más beneficiarios en este ámbito ya que lo más probable es que nacerán nuevos modelos de negocio en un futuro.

Se encontró una tesis de la universidad tecnológica de Perú, realizada por (Huaman & Quispe, 2019) el proyecto se basa en la creación de un chatbot que está enfocado en buscar productos de abarrotes utilizando una plataforma como Messenger ya que es una aplicación en la que la mayoría de los usuarios está familiarizado y es más sencillo de utilizar para mucho de ellos, además de esto los datos son obtenidos de páginas web de empresas seleccionadas y dicha información será obtenida a través de una técnica eficaz que permite

tener resultados precisos de la información que necesita el asistente virtual para la búsqueda de productos, como lo es Web Scraping que es muy utilizado para cuando se necesita recopilar grandes cantidades de datos. Para la visualización de la búsqueda de productos dicha información estará clasificada y cada uno se mostrará con costo, link, nombre de la entidad, imagen y descripción dependiendo de la búsqueda que requiera el consumidor. La finalidad de esto es que los usuarios puedan reducir ese tiempo de búsqueda puede tomar al tratar de cotizar un producto que sea de necesidad para el usuario y tenga detalles actualizados y reales de la descripción en cada búsqueda de manera eficaz y dando una mayor satisfacción al cliente.

(Cortés & Edward, 2019) propone en su artículo la implementación de un prototipo de chatbot que será para Facebook Messenger y tendrá como objetivo principal detectar síntomas de trastornos psiquiátricos, ya que explica que la inteligencia artificial ha sido importante para muchos campos, incluyendo el campo de la medicina teniendo grandes beneficios en esta incluyendo la salud mental que es una dificultad que se ha vuelto muy frecuente actualmente en todas partes, como la depresión. A partir de estudios se pudo conocer que los chatbots pueden ser un instrumento muy importante si son destinados como terapeuta virtual lo que conlleva a que este sea un gran soporte y apoyo para expertos en la detección de salud mental. Para ello los estudios que se exponen en este artículo tiene como propósito llevar a cabo la creación y diseño de un chatbot que permita dar a conocer si un usuario está presentando síntomas del trastorno depresivo ya sea este de alcance mayor o suicidio a partir de preguntas detalladas que fueron obtenidas de la Entrevista Neuropsiquiátrica Internacional (MINI). Para la creación de chatbot a quien se lo nombro como Juliana se utilizó técnicas como Natural Language Processing, Natural Language Generation y Natural Language Understanding y servicio de Watson Assistant, así es como en conclusión de ciertos estudios realizados y probado por diferentes personas se pudo deducir que Juliana es una buena opción como ayuda a la detección de problemas mentales sin embargo se da a conocer que se puede mejorar dicho prototipo.

Se encontró un repositorio de la Universidad César Vallejo, realizada por (Ariste & Ramírez, 2020) en donde el autor da a conocer los problemas de aprendizaje que tienen los estudiantes actualmente. Por ello propone la implementación del uso de IA a través de un chat Bot que permita la enseñanza de la fotosíntesis con la finalidad de que los alumnos puedan adquirir dichos conocimientos de manera diferente y dinámica para así incentivarlos a aprender de una manera sencilla y fácil. El proyecto está dirigido para los estudiantes con

un rango de edad de 13 a 15 años por motivo de un bajo rendimiento. Para que se lleve a cabo se lo implemento en la plataforma telegram por su sencillo manejo y fácil acceso y la información respectiva se tomó haciendo uso de las técnicas de recolección de datos o web scraping para facilitar el proceso de adquisición de información. Se determino la factibilidad del proyecto mediante pruebas para los alumnos, antes y después de usar el asistente virtual de aprendizaje en donde se comprobó la mejoría a nivel académico de los estudiantes después de interactuar con el chat Bot.

Se encontró otro artículo científico elaborado por (Aranda & Hernández, 2020) en donde se realizó un estudio que describe algunas de las técnicas que más se usa para actualmente para hacer web scraping y algunos ejemplos de sitios web en donde se han extraído información en el sector turístico ya que es el tema relevante en este artículo. Indican que actualmente el internet ha ido cambiando constantemente y que cada vez son más las personas que adquieren un dispositivo tecnológico, lo que ocasiona que cada una de ellas genere una cantidad excesiva de datos en diferentes formatos como audio, imágenes, videos y texto que si son manejados de la mejor manera pueden ayudar a resolver algunas inconvenientes. La extracción de datos es uno de los mayores inconvenientes en el sector de turismo porque dichos datos pueden ser limitados y de alto costo, es por ello que para resolver estos problemas las herramientas de extracción de datos es una de las mejores opciones. En su estudio dan a conocer algunas de las herramientas existentes para realizar web scraping como las herramientas de programación, instaladores, aplicaciones web y extensiones para navegadores.

Tabla 1. Trabajos internacionales aplicando técnica de web scraping y asistente virtual.

Referencia	Tema	Metodología	Técnica	Herramientas
(Chesñear & Estevez, 2018)	El comercio electrónico en la era de los bots.	Método explicativo.	Cuantitativa.	Encuestas, documentos, estadísticas.

(Huaman & Quispe, 2019)	Modelo de búsqueda de productos alimenticios en supermercados online categoría abarrotes utilizando chatbot y extracción de datos con Web Scraping.	Método de PLN y extracción de datos	Interfaces de lenguaje natural y web scraping	Python, Dialogflow, Firebase y Facebook Messenger.
(Cortés & Edward, 2019)	Chatbots para apoyar la detección de síntomas de trastornos psiquiátricos.	Método de Natural Language Processing (NLP), Natural Language Understanding (NLU) y Natural Language Generation (NLG)	Evaluación al usuario.	Watson Assistant, Facebook Messenger.
(Ariste & Ramírez, 2020)	Chatbot para el aprendizaje de la fotosíntesis utilizando la técnica web scraping.	Crystal, Método de desarrollo dinámico del sistema (DSDM), Feature Driven Development (FDD), Scrum y Extreme	Encuestas de recolección de datos.	BotFather, telegram, PHP, MySQL

(Aranda & Hernández, 2020)	Programming (XP)			Tablas comparativas, textos, documentos.
	Herramientas Web	Método explicativo.	Cuantitativa.	
	Scraping para la Adquisición de Datos Turísticos.			

Información tomada de los repositorios y artículos investigados. Elaborado por Ruth Apolo

2.1.2. Antecedentes nacionales

Se encontró una tesis de Universidad Politécnica Salesiana sede Quito realizada por (Montufar & Mullo, 2021) en donde plantea desarrollar una aplicación web que permita realizar la extracción de información en el área de procesos para la matriculación de los vehículos en los sitios web oficiales de instituciones de Quito, Guayaquil y Cuenca haciendo uso de las técnicas de Web Scraping para ahorra el tiempo que tomaría obtener la información de manera manual y a su vez para evitar errores humanos involuntarios. Dichos datos serán almacenados y usados para entrada a un chat Bot que se creó con la finalidad de que este pueda ser de ayuda para las personas que deseen acceder a los procesos de matrícula de una manera rápida y sencilla a la información requerida por los usuarios. Por ello a lo largo de todo el proceso del desarrollo del proyecto, realizaron pruebas de funcionalidad para verificar que todo lo planteado cumple con lo requerido y que los usuarios queden satisfechos.

El proyecto realizado por (Barbosa & Banchón, 2017) está enfocado en implementar un asistente virtual que proporcione ayuda para los usuarios en el área de soporte técnico de equipos informáticos, hardware, con la finalidad de acelerar las respuestas a dichos inconvenientes en el área de mesa de ayuda de empresas ecuatorianas con el objetivo de abastecerse con la cantidad de peticiones que generan los usuarios. Este asistente virtual es capaz de tener un dialogo virtual simulando una conversación como un humano para generar un ambiente de confianza con los clientes. Cuando el usuario interactúe con el Bot y requiera de una consulta o ayuda para resolver su inconveniente este en cuestión de segundos realizara una búsqueda que se encontrara registrada en su base de datos y dará la respuesta requerida por el usuario haciéndole saber qué pasos debe seguir para resolver el problema ya sean de primer o segundo nivel de su equipo y si necesita de un soporte especializado este se encargara de enviar información del técnico encargado de la empresa para que se encargue asistir al problema del equipo.

Este proyecto realizará con metodología Scrum con lenguaje de programación C# que tendrá una base de datos en MySQL y con una interfaz en PHP.

(Pérez, 2019) Realizo una tesis dirigida a la empresa Mundirepuesto en la que implementa un chat Bot integrado en la plataforma Facebook Messenger para que sus clientes tengan una mejor atención y asesoría técnica de primer nivel ya que esta empresa se dedica a vender y distribuir repuestos automotrices de ciertas líneas de vehículos. El objetivo de este es brindar un mejor servicio para sus clientes y que pueda imitar las acciones del personal que se encarga de las ventas de la empresa para conseguir un aumento en sus comercializaciones alcanzar más clientes. En él se podrán consultar repuestos disponibles del inventario de la compañía y también responder interrogaciones habituales de los usuarios.

Para el desarrollo del chat Bot se escogió Dialogflow como instrumento principal ya que después de ciertos estudios e investigaciones fue la mejor opción, con mejores características para efectuar el trabajo, así como también se trabajó con Proceso de Software Personal que ayudo a mejorar la gestión del desarrollo del sistema y su vez mejorar la productividad del mismo.

Según (Males & Villacrés, 2021) en su proyecto realizado, utilizan la técnica Web Scraping con la finalidad de recolectar información para el análisis de datos climatológicos que son presentados en la central meteorológica Yanarumi que se encuentra en la provincia de Chimborazo, la información se genera cada 30 minutos en la estación y estos se almacenan en una base de datos MySQL. Hay que tener en cuenta que esta información es muy importante ya que les permite monitorias el cambio climático, por lo que estos deberían tener el mínimo porcentaje de error en la colectación y poder hacer el análisis respectivo a través de un sitio web donde se pueda con la visualización de datos actualizados en cualquier sitio a través de internet.

Este proyecto se llevara a cabo utilizando un herramienta de software estadístico R y los datos podrán ser obtenidos mediante una representación gráfica en tiempo real ya que estos estarán actualizados cada 30 minutos en un sitio web que establece una conexión con la base de datos MySQL que posee la estación meteorológica y esto podrá favorecer directamente a un aproximado de 200 familias que podrán tener la oportunidad para realizados nuevos planes o proyectos contribuyan a la conservación del ecosistema.

Tabla 2. Trabajos nacionales aplicando técnica de web scraping y asistente virtual.

Referencia	Tema	Metodología	Técnica	Herramientas
(Montufar & Mullo, 2021)	Creación de una aplicación web para la extracción de datos de páginas web acerca de procesos de la matriculación vehicular utilizando la técnica de web scraping.	Metodología ágil Extreme Programming (XP)	Web scraping.	Octoperf Python Servidor LINUX DEBIAN Servidor Apache
(Barbosa & Banchón, 2017)	Desarrollo de un servicio web chatbots basado en mesa de ayuda para las empresas ecuatorianas.	Metodología Ágil Scrum.	Procesamiento del Lenguaje Natural (PLN)	Interfaz PHP C# (si Sharp) MySQL Bot Framework QNA
(Pérez, 2019)	Implementación de un agente conversacional para negocio de repuestos automotrices integrado a plataformas de mensajería instantánea.	metodología (PSP) Proceso de Software Personal	El procesamiento del lenguaje natural (NLP)	DialogFlow Facebook Messenger API
(Males & Villacrés, 2021)	Desarrollo del sistema web de análisis de información	Metodología SCRUM	Web Scraping	Software estadístico R MySQL Arduino

climatológica para
la central
meteorológica
Yanarumi
utilizando el
software estadístico
R.

Visual Studio
Code

Información tomada de los repositorios investigados. Elaborado por Ruth Apolo

2.2 Fundamentación teórica

2.2.1 Inteligencia Artificial

Con el pasar de los años sea buscado la forma de crear sistemas que permitan lograr tener las capacidades de razonamiento de un ser humano, de este modo se podría ejecutar procesos de una manera más sencilla. El inventor y ajedrecista Johann Von Kempelen fue quien dio sus inicios con la inteligencia artificial quien creo un androide que simulaba jugar ajedrez de manera autónoma, pero aquello fue una idea farsa ya que este androide realmente era manipulado por una persona detrás del mecanismo, pero, en base a eso nació la idea de elaborar auténticos autómatas. Fue así como surgió la idea de Leonardo Torres Quevedo quien le dio vida a El Ajedrecista, el primer juego de ordenador sin intervención humana creado en 1912. (Alvarado, 2015)

La inteligencia artificial surgió con la finalidad de poder imitar el pensamiento del ser humano. Debido a esto ha surgido una interrogación ¿pueden las maquinas pensar? Aun no se tiene una respuesta clara a esta pregunta, porque se debe considerar en primera instancia que es “pensar” y si solo puede ser una habilidad humana o no. (Serna, 2017)

La inteligencia artificial o también conocida como IA, es un tipo de inteligencia que es efectuado por maquinas cuando imitan o copian las habilidades de un ser humano, es decir, aprender, solucionar problemas, reconocer, examinar, entre otras destrezas y así alcanzar tareas concretas que se necesitan. Además, que gracias a la big data, la IA es capaz de analizar grandes cantidades de datos de manera precisa y rápida de manera automática.

2.2.1.1 Machine learning

Según (González, 2015) Machine learning estudia la edificación de sistemas que logran el aprendizaje a partir de datos.

Machine learning o aprendizaje automático es aquello que permite que las máquinas puedan aprender ya sea a base de ejemplos y experiencias de como alcanzar dicho trabajo o tarea a través de grandes cantidades de datos que le permite descubrir patrones. Machine learning es una rama de la IA que tiene la finalidad desarrollar métodos que permitan que las máquinas adquieran conocimiento a base de datos proveídos. (Andrade & Cañar, 2019)

Machine learning es la habilidad que le da los algoritmos a las máquinas de reconocer patrones en grandes cantidades de datos a base de predicciones, permite a los ordenadores ejecutar trabajos determinados de manera autónoma sin que estos necesiten ser programados, es decir, pueden iniciar tareas nuevas deduciendo o haciendo predicciones a partir de situaciones de ejemplos anteriores.



Ilustración 1. Como funciona Machine Learning. Obtenido de Repositorio. Elaborado por Ruth Apolo

2.2.1.1.1 Aprendizaje supervisado

Tiene a su disposición una base de datos tanto como de entrada y salida y que se los usa para “entrenar” y así este pueda seguir un modelo teniendo conocimiento de los valores de salida para un conjunto de entradas, es decir que realiza predicciones de comportamiento a partir de nuevas entradas. (Pereira, 2019)

2.2.1.1.2 Aprendizaje no supervisado

Tiene a su disposición una base de datos, pero solo de entrada, por lo que este puede lograr reconocer patrones teniendo solo valores de entrada lo que hace posible alcanzar predecir nuevas entradas al sistema. (Pereira, 2019)

2.2.1.1.3 Aprendizaje reforzado

No tiene a su disposición una base de datos previamente, debido a que el algoritmo concede al modelo aprender a través de observaciones del sistema de una manera iterativa ya que quiere decir que aprende a partir de ensayo-error. El modelo desarrolla el aprendizaje siguiendo acciones que resultan deseadas o no, es decir a base de recompensas o castigos, con la finalidad de aumentar las recompensas para que este realice buenas acciones. (Pereira, 2019)

2.2.1.2 Deep learning

Es una técnica que indica a las maquinas a realizar lo que para una persona es algo natural, aprender con el ejemplo o ya sea ejecutando trabajos de clasificación a través de sonidos, textos o imágenes. (Centeno, 2019).

Deep learning es un algoritmo que está compuesto por niveles jerárquicos en donde cada nivel inicia desde lo más simple hasta lo más complejo para así poder alcanzar ciertos conocimientos a base de entrenamientos y en base a los datos que va obteniendo, evolucionando su aprendizaje sin necesidad de la intervención humana.

El aprendizaje profundo es utilizado normalmente para aplicaciones de identificación, como pueden ser el reconocimiento facial, voz o traducción de texto, por ellos se lo designa “profundo” por la cantidad de capas que posee la red neuronal ya que en una red habitual solo contienen de 2 o 3 capas, pero en aprendizaje profundo pueden llegar a poseer cientos de ellas. (Alcaraz, 2018).

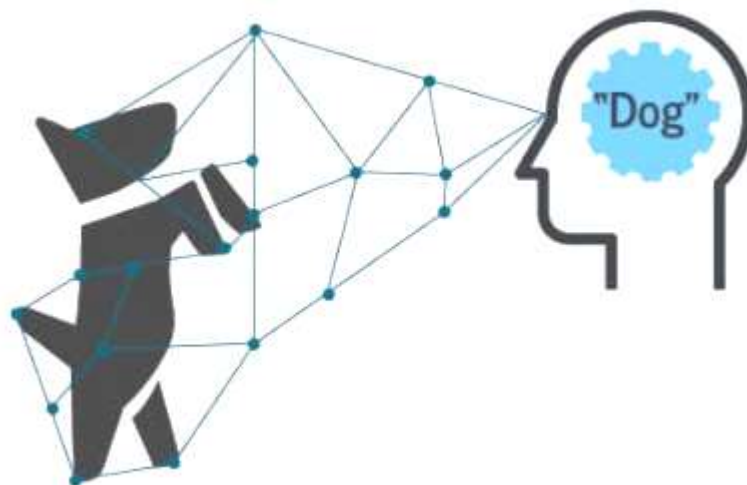


Ilustración 2. Deep Learning. Tomado de Olhar Digital. Elaborado por el Autor

2.2.1.3 Natural Language Processing (NLP).

Es una rama de la IA que busca la interacción entre los ordenadores y el lenguaje humano a través del proceso de datos creando así, aplicaciones que son capaces de analizar una gran cantidad de dichos datos. (Alias & Cassanelli, 2019)

El PLN tiene como finalidad usar el lenguaje natural para comunicarnos con el ordenador, de tal manera que esta pueda entender las oraciones que le suministren. Haciendo uso de los lenguajes naturales, proporciona facilidad al desarrollo de aplicaciones que realicen trabajos de relacionados con el lenguaje. (Cortez & Vega, 2009)

El procesamiento de lenguaje natural se enfoca en crear mecanismos computacionales que puedan perfeccionar las comunicaciones entre el ser humano y los ordenadores, siendo capaz de tener entendimiento de lo que le hablan o escriben para proporcionar una respuesta apropiada y correcta, tratando de que llegue a ser lo más natural posible.

Es claro que pensemos que las maquinas solo pueden entender bytes y dígitos binarios, es por ello que para comprender el lenguaje humano es importante darle al ordenador un modelo lingüístico.

Existen dos tipos de modelos para el procesamiento del lenguaje natural:

2.2.1.3.1 Lógicos.

Este tipo de modelo está enfocado en específicas formas gramaticales, por ello los lingüistas especializados tienen diferentes patrones precisos para que los ordenadores logren identificarlos. Si agregamos información que ya existe en los diferentes diccionarios computacionales, proporcionamos la posibilidad de definir modelos capacitados para buscar información, resumir, traducir etc.

2.2.1.3.2 Probabilístico.

Este tipo de modelo a diferencia del lógico, recolecta grandes cantidades de datos que se utilizarán principalmente para el análisis y localización de los diferentes elementos lingüísticos que pueden aparecer con frecuencia, así como las letras o palabras. Partiendo de este análisis es posible predecir que unidad surgirá en el texto, en el campo de la inteligencia artificial, a esto se lo denomina como aprendizaje automático.

A continuación (Ecorfan, 2020) describe la arquitectura de un PLN:

- *Fonológico*: está encargada de ver como las palabras se relacionan con los sonidos que representan.
- *Morfológico*: está encargada de revisar como las palabras se edifican a partir de morfemas.
- *Sintáctico*: está encargada de tratar de unir cada palabra para formar una oración.
- *Semántico*: está encargada de ver el significado de las palabras y de que manera puede darle sentido a una oración.
- *Pragmático*: está encargada de ver como el sentido de una oración puede ser afectado por las oraciones anteriores.

Aplicaciones del PLN

- Traducción automática.
- Recuperación de información.
- Extracción de información.
- Resolver de problemas.
- Tutores inteligentes.
- Reconocimiento de voz.



Ilustración 3. Usos de Natural Language Processing. Obtenido de Repositorio. Elaborado por Ruth Apolo

2.2.1.4 Agentes conversacionales o Chatbot

El termino Chatbot tiene aproximadamente 60 años desde que Alan Turing propusiera en su Test que las maquinas podrían ser indistinto de un ser humano a partir de una

conversación textual, para ello se usó el procesamiento de lenguaje natural que logro la evolución de los chatbots, en base a dichos procedimientos se desarrolló el primer chatbot en 1966 quien fue nombrado Eliza, quien fue creada por Joseph Weizenbaum a través del lenguaje C y estaba basado en patrones quienes simulaban el comportamiento humano. En 1989 se inspiraron en Eliza para el desarrollo de Chatterbot quien fue denominado TYNMUD. (Ganan, 2021)

Un chatbot es una entidad artificial desarrollada para tener conversaciones reales con el ser humano, dichas conversaciones se logran a través de texto, oral e incluso no verbal. Estos chatbots los podemos hallar en pantallas, mundos virtuales y proyecciones holográficas, así como también juguetes y robots. (Morales, 2011)

Los agentes conversacionales son herramientas que tienen múltiples usos actualmente ya que están basados en inteligencia artificial y procesamiento de lenguaje natural quienes logran procesar la información que reciben y generar una respuesta de carácter racional que forma una conversación con sentido o coherente.

Según (García, 2018) el crecimiento de las herramientas de agentes conversacionales ha avanzado tanto que es posible categorizarlos, siendo así, los tipos de chatbots son:

- *Chatbot de ventas*, este tipo de chatbots están encaminados a la comercialización tanto de servicios como productos de distintas organizaciones o empresas que se ven en la necesidad de implementar estos asistentes virtuales.
- *Chatbot de servicio al cliente*, este tipo de chatbots están encaminados a resolver dudas de los usuarios que disponen de algún tipo de servicio.
- *Chatbot de noticias y contenido*, este tipo de chatbots se han alcanzado a implementar a través de medios de mensajería instantánea, tiene como finalidad primordial enviar contenido masivo a través de diferentes canales de mensajería.

Los avances de este tipo de sistemas han hecho posible el desarrollo de algunas aplicaciones en la era actual, el mundo habitualmente se ve sumergido en este uso debido a sistemas como Siri, Google now, entre otros.

Tabla 3. Tabla comparativa de características de herramientas para la creación de chatbot.

Manychat	Dialogflow	Wit.ai	Watson
----------	------------	--------	--------

Idioma español.	No	Si	Si	Si
Gratuito.	Si, limitado	Si, limitado	Si	Si, limitado
Integración Facebook Messenger.	Si	Si	Si	No directamente
Visualización de flujos conversación.	Si	No	Si	Si

Información tomada de los repositorios investigados. Elaborado por Ruth Apolo

2.2.2 ManyChat

Es una herramienta que brinda un servicio sin costo para desarrollar un Bot para tu página de Facebook sin necesidad que seas un experto en programación. También dispone de una versión de pago, en donde podrás tener a disposición una versión mejorada y con más herramientas para mejorar el servicio de mensajería a los usuarios.

Lo que lo hace interesante es que es una herramienta muy fácil de usar y práctica, no es necesario tener conocimientos técnicos avanzados para su desarrollo, solo es cuestión de ingresar con credenciales de Facebook y empezar con el proceso del desarrollo del Bot.



Ilustración 4. Icono ManyChat. Tomado de ManyChat. Elaborado por el Autor

2.2.2.1 Tipos de interfaces:

Manychat presenta dos tipos de interfaces para la creación del Bot:

2.2.2.1.1 Basic Builder

Basic Builder nos ayuda a diseñar o desarrollar en interfaces de una manera visual y sencilla. Nos permite efectuar todo tipo de configuraciones como si se tratara de un formulario ya que disponen de algunos componentes visuales que facilitan el trabajo en el diseño dejando de lado la codificación lo que nos ayuda a ahorrar tiempo.

Es decir que para el proceso del desarrollo del Bot en manychat esta herramienta nos va a permitir realizar o armar los mensajes en secuencia desde una interfaz de manera que lo podamos visualizar con aspecto de chat.

2.2.2.1.2 Flow Builder

Flow Builder es una interfaz declarativa que nos ayuda a el proceso de desarrollo de flujos individuales. Se lo puede emplear para desarrollar una lógica equivalente a un código sin que se posea conocimientos previos de lenguaje de programación.

La interfaz usa un mapa drag & drop, lo que quiere decir, arrastrar y soltar, dicha técnica permite que el usuario tenga una facilidad para interactuar con el programa para que este pueda conectar mensajes y acciones entre sí y pueda llevar los elementos de un lugar a otro para agilizar el desarrollo permitiendo ver los resultaos en tiempo real.

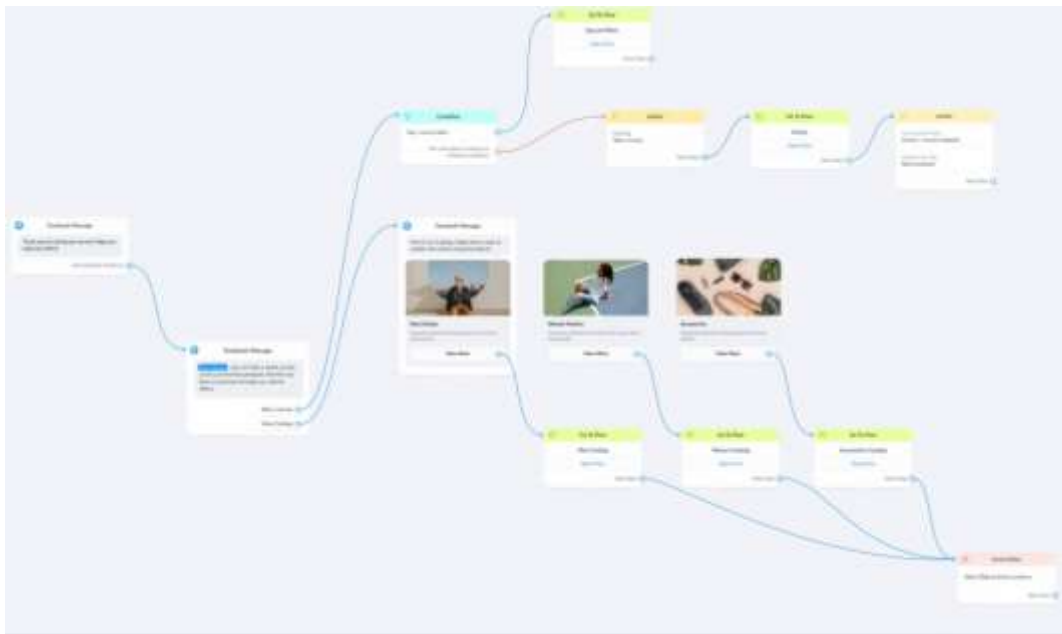


Ilustración 5. Interfaz ManyChat. Tomado de ManyChat. Elaborado por el Autor

2.2.2.2 Entendimiento del Lenguaje Natural

Manychat puede percibir una consulta como un ingreso de datos, esto quiere decir que dicha consulta va a ser un texto en lenguaje natural y este será comparado tomando en consideración el método de aprendizaje, de esta manera la Manychat logrará transformar los datos de texto en información que sea procesable.

Todo este proceso de transformación de lenguaje natural a información que se pueda procesar por el ordenador o computador, lo designamos como entendimiento del lenguaje natural (NLU).

2.2.2.3 Métodos de Entrada

Manychat toma la información que recibe de texto como entradas. Cuando la información que se requiere enviar sea un archivo de audio, se solicitaría el uso de un tercero para transformar dicho archivo de audio en un texto.

2.2.2.4 Métodos de Salida

Manychat ofrece una sección en donde podemos realizar la configuración de la interacción del Bot con el humano. Hay muchas posibilidades de configuración para enviar los mensajes, no solo se puede enviar texto, manychat proporciona herramientas de edición como imágenes, archivos, audio, videos, entre otras, algunas en versión gratuita y otras disponibles en versión pro.

2.2.3 Google Drive

Google drive es un servicio creado por Google utilizada para el almacenamiento de datos en la nube, dispone de algunas aplicaciones de utilidad como editor de texto, hoja de cálculo, presentación y calendario, todos estos servicios en una sola plataforma que al que puedes acceder desde cualquier dispositivo que tenga conexión a internet.

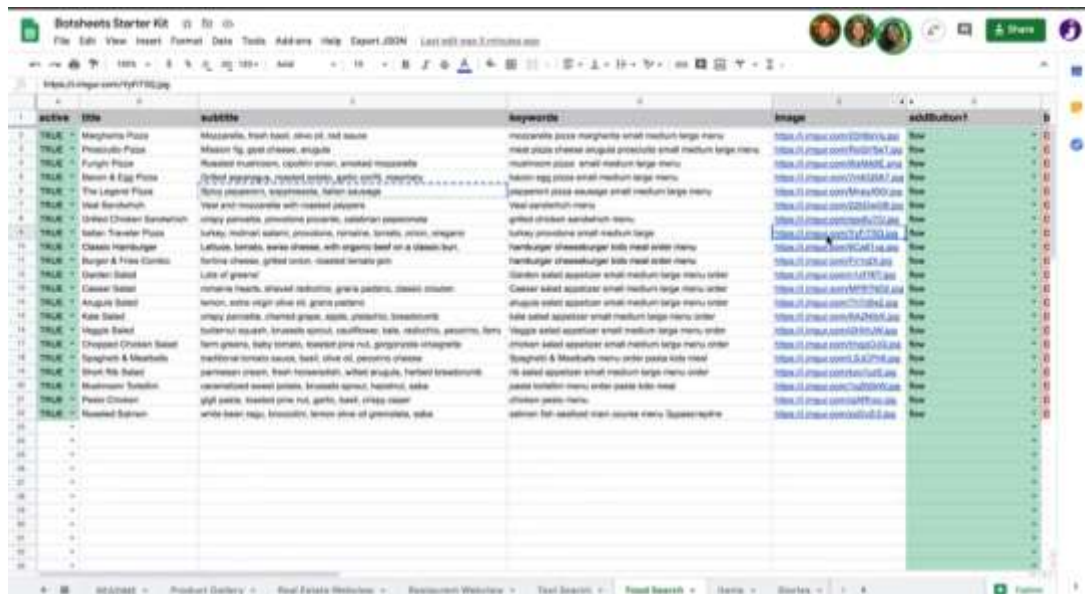
Esta herramienta tiene como ventajas el poder crear y editar documentos en tiempo real desde cualquier dispositivo en diferentes formatos ya que dichos archivos se encuentran alojados en los servidores de Google, solo es cuestión de acceder a la nube desde tu cuenta de correo y disfrutar de dichos servicios ya que su uso es muy práctico.

2.2.3.1 Botsheets

Botsheets es una herramienta de hoja de cálculo, dicha hoja de cálculo dispone de algunas características que te permitirán administrar tu contenido y mejorarán tu experiencia al ejecutar alguna tarea de búsqueda ya que es una plantilla previamente estructurada que permite administrar tu contenido.

Esta aplicación se puede habilitar e instalar en tu bot de Manychat para generar respuestas con contenido dinámico para Facebook Messenger lo que pueden ayudar a mejorar la experiencia de interacción entre el usuario y el Bot ya que con dichos datos el usuario puede obtener información amplia y completa., puede ser usado para editar tablas para la realización de algunos proyectos. Uno de sus ventajas es la conectividad instantánea, podemos tenerlo a nuestro alcance a partir de cualquier equipo que tengamos a disposición ya sea este un ordenador y smartphone ya que se puede crear una copia de la plantilla de la

hoja de cálculo en nuestra cuenta de Google Drive y manipular la información desde cualquier dispositivo inteligente.



active	title	subtitle	keywords	image	substitution
TRUE	Margherita Pizza	Mozzarella, fresh basil, olive oil, red sauce	mozzarella pizza margherita small medium large menu	https://i.imgur.com/1234567.jpg	True
TRUE	Prosciutto Pizza	Mozzarella, fig, goat cheese, arugula	meat pizza cheese arugula prosciutto small medium large menu	https://i.imgur.com/2345678.jpg	True
TRUE	Fajita Pizza	Mozzarella, chicken, pepperoni, onion, tomato sauce	meat pizza small medium large menu	https://i.imgur.com/3456789.jpg	True
TRUE	Beef & Egg Pizza	Grilled pepperoni, tomato sauce, eggs, onion, mozzarella	beef egg pizza small medium large menu	https://i.imgur.com/4567890.jpg	True
TRUE	The Legend Pizza	Strong tomatoes, mushrooms, extra cheese	meat pizza small medium large menu	https://i.imgur.com/5678901.jpg	True
TRUE	Veal Sandwich	Veal and mozzarella with tomato sauce	Veal sandwich menu	https://i.imgur.com/6789012.jpg	True
TRUE	Grilled Chicken Sandwich	Grilled chicken, provolone cheese, tomato sauce	grilled chicken sandwich menu	https://i.imgur.com/7890123.jpg	True
TRUE	Salmon Sandwich	Salmon, tomato, onion, tomato sauce, extra cheese	salmon sandwich small medium large menu	https://i.imgur.com/8901234.jpg	True
TRUE	Classic Hamburger	Lettuce, tomato, onion, tomato sauce, with organic beef on a classic bun	Hamburger cheeseburger with meat extra menu	https://i.imgur.com/9012345.jpg	True
TRUE	Burger & Fries Combo	Tomato, cheese, grilled onion, tomato sauce, onion	Hamburger cheeseburger with meat extra menu	https://i.imgur.com/0123456.jpg	True
TRUE	Chicken Salad	Let's go of great!	Chicken salad appetizer small medium large menu order	https://i.imgur.com/1234567.jpg	True
TRUE	Cheese Salad	Tomato, onion, tomato, tomato sauce, onion	Cheese salad appetizer small medium large menu order	https://i.imgur.com/2345678.jpg	True
TRUE	Angus Beef	Angus, extra virgin olive oil, tomato sauce	Angus beef appetizer small medium large menu order	https://i.imgur.com/3456789.jpg	True
TRUE	Kale Salad	Grilled chicken, tomato, onion, tomato sauce, onion	Kale salad appetizer small medium large menu order	https://i.imgur.com/4567890.jpg	True
TRUE	Veggie Salad	Butter, onion, tomato, tomato sauce, onion	Veggie salad appetizer small medium large menu order	https://i.imgur.com/5678901.jpg	True
TRUE	Chopped Chicken Salad	Tomato, onion, tomato, tomato sauce, onion	Chicken salad appetizer small medium large menu order	https://i.imgur.com/6789012.jpg	True
TRUE	Spaghetti & Meatballs	Mozzarella, tomato sauce, beef, olive oil, tomato sauce	Spaghetti & Meatballs menu order pasta keto meal	https://i.imgur.com/7890123.jpg	True
TRUE	Beef & Egg Salad	Grilled pepperoni, tomato sauce, eggs, onion, tomato sauce	Beef & Egg salad appetizer small medium large menu order	https://i.imgur.com/8901234.jpg	True
TRUE	Meatloaf Sandwich	Meatloaf, tomato, onion, tomato sauce, onion	Meatloaf sandwich menu order pasta keto meal	https://i.imgur.com/9012345.jpg	True
TRUE	Pasta Chicken	Grilled pepperoni, tomato sauce, eggs, onion, tomato sauce	Chicken pasta menu	https://i.imgur.com/0123456.jpg	True
TRUE	Mozzarella Sandwich	White bean, onion, tomato, tomato sauce, onion	Mozzarella sandwich menu Supermarket	https://i.imgur.com/1234567.jpg	True

Ilustración 6. Botsheets. Tomado de Botsheets. Elaborado por el Autor

2.2.4 Facebook Messenger

Facebook Messenger es una plataforma de servicio de mensajería que representa una parte de Facebook y nos ayuda a establecer contactos con otros usuarios de la red social y podemos acceder a él a partir de la web y sus aplicaciones nativas.

Facebook inicio el 4 de noviembre de 2004, el desarrollador del proyecto es conocido como Mark Zuckerberg, junto a Eduardo Saverin, Dustin Moskovits y Chris Hudhes. Inicialmente fue un sitio destinado para alumnos de la universidad de Harvard y su finalidad era crear un espacio en donde los estudiantes logren una comunicación fluida y compartir contenido de manera sencilla a través de la web, a este sitio lo denominaron “thefacebook.com”. Se extendió a otras universidades y aproximadamente en el 2005 se integró a escuelas de EE. UU y otros lugares del mundo, acogiendo el nombre decisivo de Facebook. (Sforzin, 2018).

Facebook Messenger apareció en 2008 con el nombre de Facebook Chat. Empezó como un servicio básico de mensajería de Facebook. En 2010 Facebook actualiza su servicio haciendo lanzamientos en plataformas para smartphones. En 2014 para que los usuarios puedan disfrutar del servicio debían instalar la aplicación ya que esta se desliga de Facebook ofreciendo una interfaz más innovadora y llamativa. Con el paso de los tiempos la interfaz de Messenger ha ido mejorando hasta llegar a lo que es hoy en día, además de pasar por

muchas ediciones para adaptarse a las necesidades de los usuarios que usan diferentes versiones como de Android y iOS.

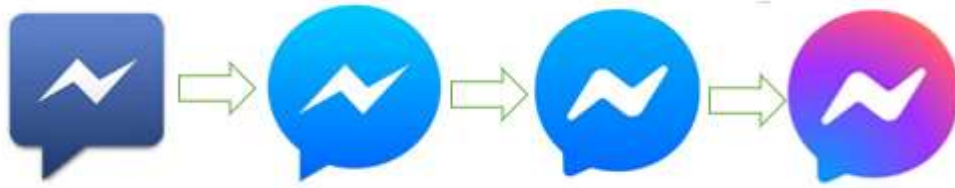


Ilustración 7. Evolución del logo Messenger. Obtenido de Repositorio. Elaborado por Ruth Apolo

2.2.4.1 Facebook Messenger Bots

Facebook suministra recursos útiles e imprescindibles para ayudar a los perfiles comerciales y fidelizar a sus clientes, subir las ventas a partir de internet.

En este caso un Bot para Facebook Messenger es un asistente virtual que va a poder entablar una conversación con algunos de los 1.3 mil millones de usuarios que utilizan la plataforma y así ofrecer un servicio de atención al cliente de manera automatizada las 24 horas del día con mensajes e información importante para el usuario.

Tabla 4. Tabla comparativa de características de plataformas de mensajería instantánea.

	WhatsApp	Telegram	Facebook Messenger	Line	Skype
Multidispositivo	no	Si	si	no	Si
Licencia	Software privado	Servidor: Código cerrado	Freeware	Software privado	Freeware
Dispositivos	Android, iOS y PC	Android, iOS y PC	Android, iOS y PC	Android, iOS y PC	Android, iOS y PC
Soporte Bots	si	Si	si	si	Si

Información tomada de los repositorios investigados. Elaborado por Ruth Apolo.

2.2.5 Web Scraping

Es una técnica que sirve para extraer datos de las páginas web y los almacena con la finalidad de que estos puedan ser analizados o manipulados en otras herramientas.

Existen diferentes métodos para realizar la extracción de información de la página web y para que cada uno logre sus objetivos, realiza un proceso en donde hace uso de bots para extraer la información, dicha información es la recopilación del código HTML y datos almacenados en bases de datos.

Existen dos maneras en el que Web Scraping pueda ejecutarse. Primero con la creación de un bot con la ayuda de un lenguaje de programación y como segunda opción hacer uso de una herramienta con la que ya no será necesario tener conocimientos técnicos en programación y disfrutar de la extracción de otros sitios web.

Esta herramienta es totalmente legal y lícito, aunque existen sitios web que tratan de protegerse de manera preventiva del web scraping ya que en muchas ocasiones se lo ha relacionado como una técnica de robo de información, por ello es muy importante leer las condiciones y términos del uso de una página web a la que nos interese extraer datos.

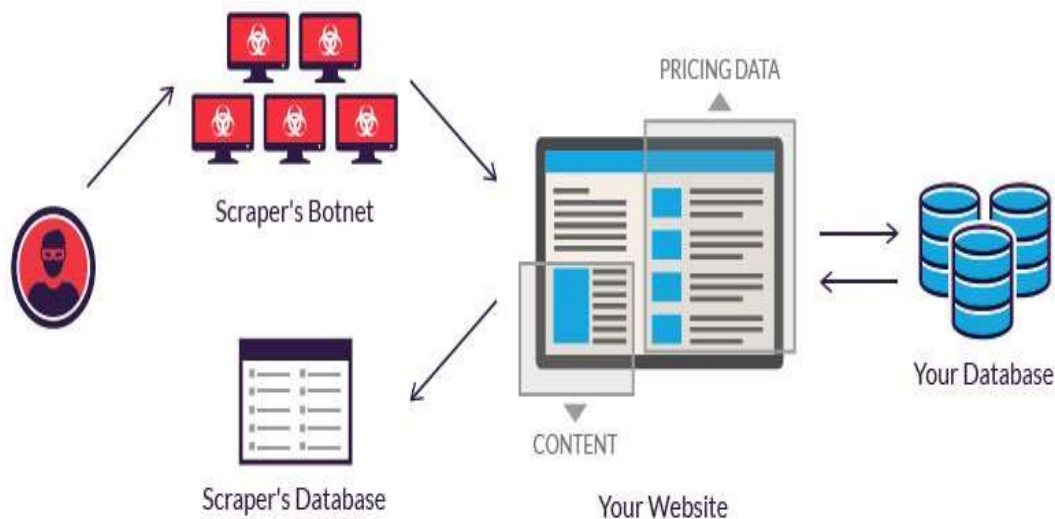


Ilustración 8. Que es Webscraping. Tomado de Antevenio. Elaborado por el Autor

2.2.5.1 Técnicas utilizadas para web scraping

2.2.5.1.1 Recolección manual

Consiste en copiar los datos que sea de nuestro interés de una página web y pegarla. Llevar esto a cabo puede ser cansado y tedioso ya que realizarlo en miles de páginas sería una tarea imposible. A pesar de esto es el único método que nos va a permitir realizar la extracción de datos si una página contiene barreras para prevenir dicha tarea de recopilación de información.

2.2.5.1.2 Expresiones regulares

Es serie de caracteres que establecen un patrón de búsqueda. Por lo general es utilizado para buscar cadenas de caracteres, entonces, usando motores de búsqueda se logra obtener la búsqueda de expresiones habituales de un sitio web, es un método sencillo y potente.

2.2.5.1.3 Protocolo HTTP

Se consigue a partir de peticiones HTTP a un servidor a través de sockets, de esta manera se logra obtener todos los datos que puede contener el sitio web para que estos puedan ser almacenados y clasificados.

2.2.5.1.4 Parsers de HTML

A través del lenguaje de programación se puede lograr procesar documentos HTML, rescatando y transformando la información. Un ejemplo de lenguaje que nos permiten efectuar este tipo de metodologías son Xquery o HTQL.

2.2.5.1.5 Reconocimiento de web semántica

Ciertos sitios web tienen información semántica, es decir, anotaciones o comentarios y puede ser utilizada para extraer datos deseados. Dichas anotaciones pueden contenerse en el mismo sitio, lo que significa que es de utilidad cuando se procesa el DOM (Document Object Model) del documento logrando obtener los esquemas antes del análisis de los documentos.

2.2.5.1.6 Aplicaciones para web scraping

Se puede encontrar una gran cantidad de plataformas disponibles que permiten extraer la información. Para lograrlo tienen que tener la capacidad de conocer la estructura del sitio web o conceder al usuario seleccionar los campos que necesite en un documento.

2.2.5.2 Herramientas para realizar Web scraping

Existen muchas herramientas que tiene como finalidad la extracción de datos o información de la web, cada una de estas herramientas están desarrolladas a partir de un lenguaje de programación como Python, Ruby, PHP... dependiendo del tipo de usuario se puede acceder a cualquiera de ellos.

2.2.5.2.1 Usuarios no programadores

Las herramientas para estos usuarios son sencillas y con interfaces amigables para su uso, de modo que, a través de “clicks” se puede extraer la información sin necesidad de programar, son bastante limitadas con respecto a los que están destinadas a los usuarios programadores.

- *Web HTTrack*: es una herramienta que nos permite descargar sitios web por completo, es decir obtiene el HTML, imágenes, archivos de directorio y otros archivos importantes desde el ordenador, puede ser utilizado en macOS, Android, Windows y Linux.

En el caso de que las descargas hayan sido interrumpidas esta herramienta permite reanudarlas realizando una actualización en el sitio existente.

- *Mozenda*: esta herramienta es un servicio de web scraping, es capaz de realizar múltiples tareas, nos ayuda a recopilar datos e imágenes de cualquier página web y para ello utiliza una interfaz para crear y editar de una manera fácil los agentes que obtiene la información.
- *WebHarvy*: es una plataforma de extracción de datos que puede recopilar información automáticamente de imágenes, textos y URL de una manera sencilla y cómoda ya que es fácil de usar ya que se puede manipular a través de “clicks” y es compatible con todo sitio web.
- *Visual web Ripper*: este programa de web scraping tiene un buen funcionamiento para algunos sitios web que son difíciles de extraer información, además de eso contiene una interfaz muy interactiva y muy sencilla de usar para ayudar a los usuarios que no dispongan de conocimientos técnico de programación. Permite exportar la información en diferentes formatos como (CSV, Excel, Base de datos...).

2.2.5.2.2 Usuarios programadores

Las plataformas que son destinadas para estos usuarios son aquellos que les permite programar la información que necesite extraer de los sitios web.

- *Scraper.rb*: Es un programa que extrae la información navegando y recopilándola para ordenarla y almacenarla en un formato XML haciendo uso de dos librerías, Watir y Nokogiri.

Watir es el que maneja el comportamiento en el navegador, es decir, que logra abrir la página y navegar por ellas haciendo “clicks” sobre las secciones que han sido determinados en el archivo con formato XML.

Por su lado Nokogiri permite conseguir los datos recorriendo la página y reunirla en formato XML.

- *Scrapy*: es una plataforma que logra construir pequeños scripts que navegan por las páginas web recopilando su contenido sin ayuda humana. Es de código libre

escrito en Python usado para diferentes aplicaciones como la minería de datos. Esta herramienta nos permite procesar una cantidad grande de páginas ya que da la posibilidad de descargar varias páginas simultáneamente.

- *Beautiful Soup*: se trata de una biblioteca de código abierto de Python que usa su analizador HTML/XML y transforma el sitio web HTML/XML en un conjunto de elementos, valores y atributos. Esta herramienta puede conseguir información extrayendo cualquier parte del sitio web de una manera procesable.
- *Selenium*: es una herramienta que permite la automatización de navegadores web, pero también se puede usar para realizar scraping de cualquier página web analizando como primera instancia la estructura HTML y por consiguiente llamar funciones para extraer etiquetas explícitas denominadas por algún ID.



Ilustración 9. Como extraer datos de un sitio web. Tomado de Morioh. Elaborado por el Autor

2.2.5.3 Impedimentos al utilizar Web Scraping

Existen ciertas dificultades en el proceso del desarrollo de Web scraping y se dividen en dos grupos.

2.2.5.3.1 Limitaciones al usar web scraping

- La estructura de la página web están en constantes cambios, la información que se extraen de un sitio web se organiza acorde a la estructura de esta, pero, en ciertas ocasiones al volver a visitar la página el diseño ha cambiado debido a que algunos diseñadores actualizan con frecuencia los sitios web con la finalidad de mejorar la interfaz para el usuario o incluso con el objetivo de prevenir el uso de

web scraping. Estos cambios pueden estropear la extracción de información ya sea este un cambio pequeño o grande.

- Páginas webs complejas, no todos los sitios web son sencillos de scrapear, casi el 80% de las páginas web son sencillas y moderadas para hacer uso de web scraping, pero también existe un 20% de ellas que son bastantes complejas lo que provoca dificultad al utilizar esta herramienta para extraer información. Hoy en día los sitios web están incluyendo contenidos dinámicos, como un desplazamiento infinito o un botón para “cargar más” para poder seguir visualizando el contenido del sitio, por ello, en este tipo de situaciones los usuarios deben hacer uso de herramientas que le permitan scrapear de una manera más funcional.
- La página web puede prohibir tu IP, existen sistemas que autenticación que pueden lograr distinguir una máquina de un humano y detectar tu IP ya que usar la técnica de extracción de datos puede formar mucho tráfico y esto conlleva a la sobrecarga de un servidor web. Los sistemas más conocidos son captcha y paywalls y para poder evitarlos se debe configurar la técnica de scrapeo para que este pueda simular una navegación normal de un ser humano.

2.2.5.3.2 Limitaciones legales al usar web scraping

Hoy por hoy no se puede afirmar o negar para dar una respuesta ante la interrogante que existe ¿legal o ilegal?, se deberá hacer un análisis de cada caso para decretar esa legalidad. El problema no radica justamente en el uso de crawlers para analizar y guardar los datos, sino en el uso que le vamos a dar a esa información obtenida.

Tomando en cuenta esto, ciertas paginas optan por proteger su contenido, añadiendo captcha, términos legales, monitoreo de tráfico bloqueando IP, añadiendo usos comerciales autorizados y no autorizados en sus páginas, sean estos lucrativos o no.

De esta manera podemos indicar que no se evidencian leyes que exijan legalidad ante el control de procesos de Web scraping.

Hay casos en donde se ha podido apreciar el abuso en la cantidad de datos de puede ser usada y manipulada por terceros, como el caso de American Airlines, que expuso una queja a FareChase quien ofrecía lograra tener acceso a tarifas de sus vuelos. Dicha acusación radicó en la intrusión de los servidores de la aerolínea, sin un permiso de esta.

2.2.5.3.3 Técnicas para protegerse del Web Scraping

- Bloqueo de direcciones IP.
- Monitoreo de exceso de tráfico.
- Implementar sistema de verificación.
- Implementar servicios de antibots.
- Añadir condiciones legales de tu página web.
- Diseño de página web con estructuras que dificulten el web scraping.

2.2.6 Lenguaje de programación

Es una manera en la que el usuario y ordenador pueden mantener una comunicación a través de una serie de instrucciones o algoritmos que nuestro ordenador comprende, procesa e interpreta en lenguaje máquina.

Este lenguaje es un conjunto de símbolos y reglas sintácticas y semánticas interpretadas como instrucciones lógicas en donde se desarrolla el código fuente de un programa. Existen muchos lenguajes de programación que son utilizados actualmente y que comparten algunas características similares. Normalmente los programas se desarrollan en un entorno de desarrollo integrado (IDE) que es quien dispone de un compilador, editor de texto y depuradora.

2.2.6.1 Python

Python remonta a inicios de los años 90, Guido Van Rossum fue el desarrollador de este lenguaje que tenía como finalidad que este sea sencillo de aprender, escribir y entender. Este desarrollo duro varios años y para el 2000 ya se disponía de un producto muy completo.

Es un lenguaje de programación que tiene licencia de código abierto, es un lenguaje multiparadigma ya que es orientado a objetos donde podemos desarrollar todo tipo de programa gracias a sus amplias posibilidades en trabajar en muchos campos ya sean esta inteligencia artificial, data science, entre otros.

Es perfecto para trabajar con gran cantidad de información o datos, debido a que Python es un lenguaje multiplataforma, ayuda al procesamiento y extracción, es por ese motivo que entidades de Big Data lo prefieren.

Python es un lenguaje de programación usado por muchas compañías en el mundo ya que dispone de múltiples características que ayuda al usuario a desarrollar un código simple, claro y poderoso para dar soluciones a problemas de gran relevancia.

Tabla 5. Características de Python.

- Es multiplataforma, se puede utilizar en varios dispositivos y sistemas operativos.
- Es gratis y libre
- Es multiparadigma, podemos desarrollar un programa en el orientado a objetos, estructurada y funcional.
- Es tipado dinámico, es decir, no hay que especificar a que tipo de dato corresponde, las variables se acoplan a lo que desarrollamos al ejecutar el programa.
- Posee una amplia biblioteca de librerías y funciones.

Información tomada de Pythones. Elaborada por Ruth Apolo.

2.2.7 Comercio electrónico

Es un modelo de negocio que se enfoca en la compra y venta de productos o servicios usando redes computacionales, es decir, a través de internet ya sean por redes sociales medios electrónicos o sitios web.

El uso del internet nos da la oportunidad para desarrollar negocios y con el paso de los años se han desarrollado avances tecnológicos que ayudan a la compra electrónica como son e-business y m-business, que son aquellos que permiten transacciones electrónicas sean estos de un servicio o de un producto.



Ilustración 10. e-commerce. Tomado de Empireecommerce. Elaborado por el Autor

2.2.7.1 Tipo de venta E-commerce

2.2.7.1.1 Business-to-Consumer (B2C)

Las tiendas online son las más comunes en internet, es un tipo de comercio electrónico que ofrecen sus productos o servicios por medio de la web y crean una relación entre la empresa y el consumidor.

Se debe tomar en cuenta que para este tipo de E-commerce retener o ganar clientes depende exclusivamente de sus estrategias y técnicas de marketing y ventas. La compañía es responsable de la compra y distribución de los productos y servicios ofrecidos. Algunos ejemplos de B2C serían Amazon, MediaMarkt, etc.

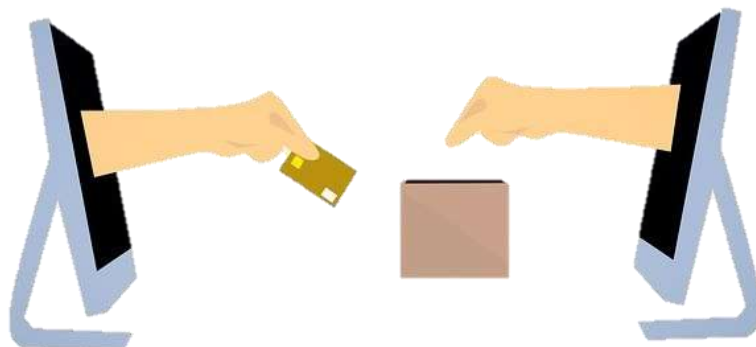


Ilustración 11. Comercio electrónico B2C. Tomado de diariodepontevedra. Elaborado por el Autor

2.2.7.1.2 Business-to-Business (B2B)

Es aquel en donde la transacción comercial se genera entre dos empresas, es decir, que una empresa le vende a otra empresa sin necesidad de la intervención de un consumidor final, por ejemplo, cuando una empresa mayorista le vende al minorista y es este quien se encarga de vender el producto o servicio al consumidor final.

Este tipo de comercios electrónicos se pueden dar en 3 modalidades

- Un minorista que busca mayorista
- Un mayorista que busca minorista
- Intermediario que busca producir una transacción comercial entre un minorista y un mayorista.



Ilustración 12. Business-to-Business. Tomado de esan. Elaborado por el Autor

2.2.7.1.3 Consumer-to-Business (C2B)

Este modelo de negocio se da cuando una persona brinda sus servicios o producto a una entidad u organización, esta básicamente encaminada a contenidos digitales, por ejemplo, cuando una persona de medio o influencer recomienda un producto o servicio poniendo en su contenido un enlace para la venta de este y dicha persona recibe un pago de la organización por compartirlo.



Ilustración 13. Comportamiento de compra de negocios. Tomado de mind42. Elaborado por el Autor

2.2.7.1.4 Consumer-to-Consumer (C2C)

Este modelo de negocio tiene como finalidad la venta de productos que, por lo general, son artículos de segunda mano, alquilar viviendas, vehículos, entre otros, pero en buen estado y se da entre consumidor que son quienes establecen un precio y el usuario final. Un ejemplo es la plataforma OLX.



Ilustración 14. Consumer-to-consumer. Tomado de shopivo. Elaborado por el Autor

2.2.7.1.5 Business-to-Administration (B2A)

Este tipo de modelo contiene transacciones muy concretas entre una organización y una administración pública, por ejemplo, los portales de transparencia como de seguridad social, empleo, docencia o cualquier ente que tenga la necesidad de cubrir la administración pública en el área digital.

2.2.7.2 Tipo de plataforma según E-commerce

2.2.7.2.1 Sitio Web

Es el método más antiguo de comercio electrónico y se basa en la creación de una página web en donde la organización o empresa da a conocer los productos o servicios que se encuentran a la venta, el usuario o consumidor observa, consulta y elige un artículo de su preferencia y lo compra. La entidad es el encargado de hacer llegar el producto enviándoselo al consumidor en el respectivo punto de entrega.



Ilustración 15. Plataformas e-commerce. Tomado de neolo. Elaborado por el Autor

2.2.7.2.2 Mobile e-commerce

Este comienza al recibir publicidad de descuentos de algún producto o servicio a nuestro dispositivo inteligente ya sea este un celular o Tablet. Mobile e-commerce no se trata de hacer una llamada o enviar un mensaje, un correo o WhatsApp, va mucho más allá de todo esto, se trata de acceder a las plataformas móviles de las empresas o al site de estas, pero con la adaptación móvil (m.site) pudiendo tener acceso a estos a cualquier hora y en cualquier lugar.



Ilustración 16. Mobile e-commerce. Tomado de neothek. Elaborado por el Autor

2.2.7.2.3 Social e-commerce

Es una manera de comercio electrónico a través de las redes sociales en donde las compañías u organizaciones han utilizado sus estrategias de venta online más convenientes y seguras para la comercialización de sus productos o servicios.

Tanto las redes sociales como los negocios online están relacionándose entre sí en la vida de los usuarios. Es por ello que las empresas utilizan este medio que ha tenido mucha acogida a lo largo de estos años en el mundo digital para llegar a sus consumidores gracias a la comunicación y exposición de sus productos.

Debido a la relevancia que tiene las redes sociales, e-commerce ha ganado una gran valoración por parte de los consumidores y clientes al introducirse en este campo ya que pueden realizar sus compras sin necesidad de acceder a otro tipo de plataformas, solo es cuestión de utilizar la misma red social que usan normalmente.



Ilustración 17. Comercio electrónico. Tomado de uptilus. Elaborado por el Autor

2.2.7.3 Tiendas online con productos tecnológicos en Ecuador

En Ecuador existen diversas tiendas que ofrecen productos de diversas categorías, entre ellas artículos tecnológicos. A continuación, se listarán las tiendas más conocidas y recomendadas.

2.2.7.3.1 Artefacta

Nació bajo el nombre de ELTESA como una planta de ensamblaje de electrodomésticos, no obstante, la empresa cambió su estrategia comercial centrándose en la venta de electrodomésticos.

Para alcanzar dicho objetivo inauguró dos establecimientos en las dos ciudades principales del Ecuador, Quito y Guayaquil y para ello decidieron cambiar el nombre de la empresa a Artefacta. Con el pasar de los años la compañía ya disponía de productos con proveedores nacionales e importados. En 2002 amplía su línea de artículos en motos y celulares y en 2004 incursiona en ordenadores para lograr variedad en las categorías de sus productos.

Actualmente la empresa cuenta con 152 establecimientos en Ecuador posicionándose como la primera del Market Share de mercado de electrodomésticos del país.



Ilustración 18. Artefacta. Tomado de Artefacta. Elaborado por el Autor

2.2.7.3.2 La Ganga

Inicio 1983, inaugurando su primera tienda en Guayaquil en la “Bahía”. Desde su comienzo se dedicó al comercio de electrodomésticos de las principales marcas internacionales, innovando en ideas que otros establecimientos no proporcionaban a sus clientes, como garantía en sus artículos.

El tiempo fue pasando y la compañía tuvo mucho éxito, logrando abarcar muchas provincias del país. Al día de hoy cuenta con un aproximado de 130 establecimientos en el Ecuador y su objetivo es seguir expandiéndose a lo largo de nuestro país.



Ilustración 19. La Ganga. Tomado de La Ganga. Elaborado por el Autor

2.2.7.3.3 Novicompu

Surgió como un negocio familiar que hoy en día cuenta con establecimientos en Ecuador y Perú, contiene una gran variedad en sus artículos de tecnología. Hoy en día cuenta con 15 locales en 7 ciudades del Ecuador.

Son importadores de las más importantes marcas en ordenadores, celulares inteligentes, tablets a precios accesibles para los usuarios. Además, cuenta con un equipo de profesionales que otorgan servicio de asesoramiento de las marcas que comercializa.



Ilustración 20. Novicompu. Tomado de Novicompu. Elaborado por el Autor

2.2.7.3.4 Computron

La empresa se encarga de la comercialización de productos de tecnología como, ordenadores, portátiles, impresoras, smartphone entre otros artículos y a su vez ofrece servicio técnico y asesoría comercial.

Es una compañía desarrollada por algunos accionistas en el año 2000 y ya en 2005 estaban extendiéndose empezando a inaugurar una sucursal en la ciudad, desde ese momento ha ido progresando continuamente abriendo sucursales en otras provincias del Ecuador. Actualmente la empresa ya cuenta con 23 locales en todo el país, en ciudades como Guayaquil, Quito Riobamba y Machala.



Ilustración 21. Computron. Tomado de Computron. Elaborado por el Autor

2.2.7.3.5 Pycca

Se edificó en 1957 en el centro de la ciudad de Guayaquil, nació como una tienda con sin número de productos surtidos a precios bajos. Lo innovador de esto y captó la atención de

los clientes es que les consintiera tocar y examinar los artículos, así como también la implementación de aire acondicionado.

Con el pasar de los años la empresa ha ido teniendo transformaciones por lo que se enfocó en formar una cadena de almacenes en todo el Ecuador. Actualmente la empresa cuenta con un aproximado de 36 sucursales situados en diferentes ciudades del país ofreciendo productos de diferentes categorías satisfaciendo las necesidades de sus clientes.



Ilustración 22. Pycca. Tomado de Pycca. Elaborado por el Autor

2.3 Fundamentación legal

Para el desarrollo de este proyecto se buscará el apoyo legal de la Ley Orgánica de defensa del consumidor en el Capítulo I, capítulo 2 en donde señala que “Todo persona natural o jurídica que, como destinatario final adquiera, utilice o disfrute bienes o servicios, o bien reciba oferta para ello.” En la presente ley se menciona al consumidor, quien hace referencia al usuario, que es quien hace uso de un servicio público. Es decir, la ley orgánica de defensa del consumidor, además, resguarda a los usuarios que disponga de un servicio procedente de una organización o empresa pública. El consumidor depende de una entidad que le provee un producto o servicio determinado lo que provoca un entorno de superioridad del proveedor frente al consumidor poniendo al usuario en desventajas en ciertas condiciones al no contar con la experiencia del proveedor, es por ello que la ley de defensa del consumidor beneficiara al usuario por ser la parte débil de la relación de consumo. (Pérez, 2012).

Los productos de consumo ecuatoriano y los que se importan al país, tienen que llevar a cabo todas las normas de calidad, tener todos los requisitos para el consumo del usuario, siendo estos evaluados en cuanto características y contenido. Así lo sitúa en su Art.52: “Las personas tienen derecho a disponer de bienes y servicios de óptima calidad y a elegirlos con libertad, así como a una información precisa y no engañosa sobre su contenido y características”. Es por ello que este proyecto se enfoca en la recopilación de datos de

artículos de tecnología de páginas que ofrecen y tienen a su disposición productos de empresas que disponen de artículos con estándares de primera calidad para sus consumidores.

Asimismo, el artículo 53, manifiesta lo siguiente “Las compañías, instituciones y organismos que proporcionen servicios públicos se comprometen a incorporar sistemas de medición de satisfacción de los usuarios y consumidores, y poner en práctica sistemas de atención y reparación. El Estado responderá civilmente por los deterioros y perjuicios que se pueden causar a los usuarios por su negligencia y descuido en la atención de los servicios públicos que estén a su cargo, y por la carencia de servicios que han sido pagados”.

En el artículo 54 también señala que “Las personas o entidades que presten servicios públicos o que produzcan o comercialicen bienes de consumo, serán responsables civil y penalmente por la deficiente prestación del servicio, por la calidad defectuosa del producto, o cuando sus condiciones no estén de acuerdo con la publicidad efectuada o con la descripción que incorpore. Las personas serán responsables por la mala práctica en el ejercicio de su profesión, arte u oficio, en especial aquella que ponga en riesgo la integridad o la vida de las personas.” Logramos observar que el Estado avala al consumidor garantizando que hace responsable a los proveedores civil y penalmente en cuanto a la prestación de bienes y servicios que estén en pésimas condiciones. También notamos el manifiesto de los servicios prestados por profesionales que se comprometen a entregar un servicio que no ponga en peligro su integridad. (Velín, 2013).

Teniendo en cuenta la cantidad de información y datos que puede manipular una compañía con respecto a sus clientes y el peligro que conlleva el mal manejo de estos, la Asamblea Nacional aprobó la Ley Orgánica de Protección de Datos Personales en donde manifiesta en su capítulo I, disposiciones directivas, artículo 2 “La finalidad de la presente Ley es procurar el adecuado tratamiento y flujo de datos personales para garantizar los derechos fundamentales y las libertades individuales, promover el progreso económico y sociales; impulsar la producción nacional y la cooperación internacional: fomentar la competitividad, la innovación y productividad, elevar la eficiencia de los servicios públicos y/o privados; y, mejorar la calidad de vida.”

Según Constitución de la República del Ecuador. Capítulo sexto. Derechos de libertad

Art. 66.- Se reconoce y garantizará a las personas: “El derecho a la protección de datos de carácter personal, que incluye el acceso y la decisión sobre información y datos de este

carácter, así como su correspondiente protección. La recolección, archivo, procesamiento, distribución o difusión de estos datos o información requerirán la autorización del titular o el mandato de la ley.”

CAPÍTULO III

3 DESARROLLO DE LA PROPUESTA

3.1 Metodología del proyecto

Para este capítulo se realizará la propuesta, a través de un conjunto de metodologías que ayuden a solucionar el problema de estudio y evaluar el funcionamiento del modelo y su validación.

3.1.1 Diseño de la investigación

Se debe tomar en cuenta que la evolución al mundo digital está desarrollando cambios en la sociedad gracias al uso de los dispositivos inteligentes que se han convertido en herramientas necesarias para nuestro diario vivir y con las que día a día producimos una gran cantidad de información o datos al mantenernos enlazados a internet, por ello, se desarrolla un método de búsqueda de productos que están almacenados en sitios web, haciendo uso de una técnica para la extracción de información y que tenga como objetivo reducir el tiempo de cotización de artículos para los consumidores que necesiten adquirir algunos de los productos de alguna tienda online.

Para el diseño de la investigación, se establece un enfoque de estudio mixto debido a que se pretende examinar algunos métodos para analizar un mismo problema y abordarlo a partir de distintos puntos de vista ya sean mediciones numéricas y opiniones de los encuestados. Asimismo, se realizará una investigación cuasi-experimental ya que se necesita trabajar sobre algo ya establecido que facilite el proceso de nuestro desarrollo sin necesidad que este sea preciso.

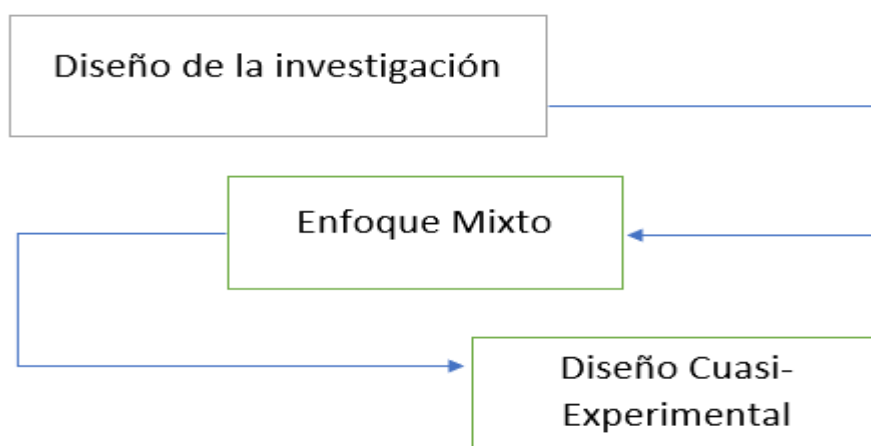


Ilustración 23. Diseño de la investigación. Información tomada de la investigación previa. Elaborada por Ruth Apolo

La investigación es una actividad que busca descubrir nuevos conocimientos, es un proceso que persigue una respuesta a problemas de conocimiento. La investigación establece un enfoque mixto que se explica a continuación.

3.1.1.1 Enfoque mixto

El objetivo del enfoque mixto no es reemplazar la investigación cualitativa y la cuantitativa, sino usar ambas minimizar las debilidades. Recolecta datos cualitativos y cuantitativos necesarios para el análisis e interpretación de los mismos. (Otero, 2018)

Se está comprobando que el uso de estudios mixtos se puede alcanzar una figura más amplia y profunda del problema ya que la investigación se soporta en las fortalezas de cada método de investigación, produce información más rica y variada a partir de las diversas observaciones, ya que se toma en cuenta algunas fuentes y tipos de datos para su análisis. (Cedeño, 2012)

Entre sus principales características se encuentran las siguientes:

- Está compuesto por los métodos de investigación cuantitativos y cualitativos en un solo estudio.
- Analiza conjuntamente la información cuantitativa y cualitativa.
- Su método investigativo basado en información mixta, consigue un mejor entendimiento del fenómeno en el estudio.

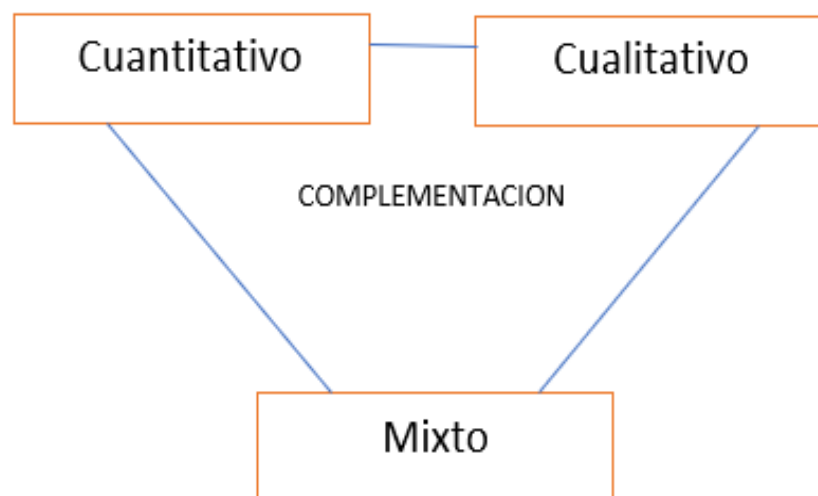


Ilustración 24. Enfoque Mixto. Información tomada de la investigación previa. Elaborada por Ruth Apolo

3.1.1.2 Diseño cuasi experimental

Es un plan de trabajo con que se aspira estudiar el efecto de impacto de los procedimientos en circunstancias en donde los sujetos a los grupos con un sistema aleatorio, si no, se usan en grupo naturales.

Tal como confirma Campbell (1988), “podemos diferenciar los cuasiexperimentos de los auténticos experimentos por la ausencia de asignación aleatoria de las unidades a los tratamientos.” (p. 191)

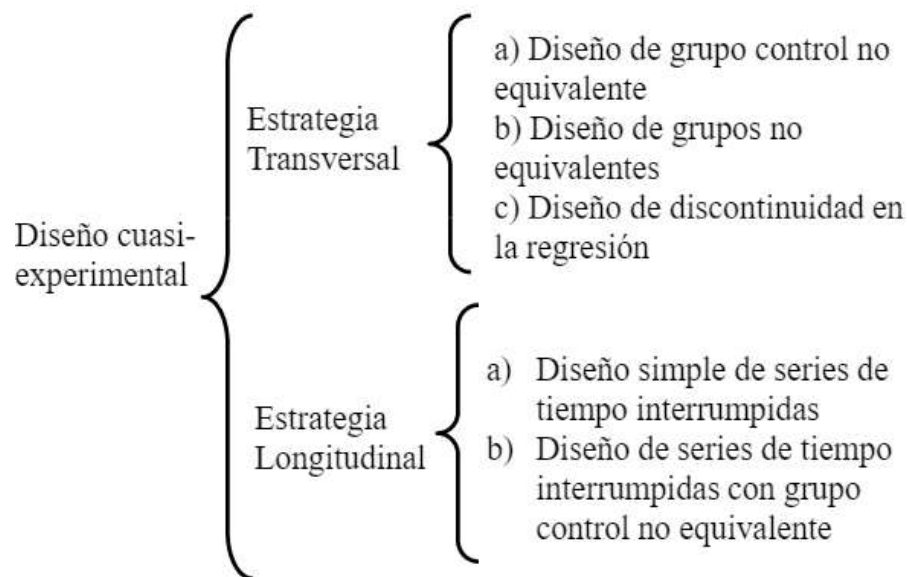


Ilustración 25. Diseño Cuasiexperimental. Obtenido de Repositorio. Elaborada por el Autor

3.1.2 Tipo de investigación

Para efectuar el proceso de este proyecto, es necesario complementar algunas metodologías, por ello, se detalla los tipos de investigación a tratar durante este desarrollo.

Los tipos de investigación que se llevaran a cabo son:

3.1.2.1 Investigación aplicada

A partir de este tipo de investigación se utilizan y aplican todos los conocimientos teóricos conseguidos durante el transcurso de la investigación y a su vez los conocimientos obtenidos durante la implementación de la práctica de dicha investigación. Buscar averiguar de qué manera funcionan las cosas para su uso final. Este tipo de investigación puede lograr aportar hechos nuevos si realizamos suficientemente bien nuestra investigación.

3.1.2.2 Investigación exploratoria

La investigación exploratoria se usa cuando no existen estudios previos de un tema en particular. El termino explorar da significado a incursionar en un territorio desconocido. Este tipo de investigación se desarrolla sobre un tema poco difundido, por ello, pretende la recopilación de información, identificar antecedentes, examinar variables que puedan estudiarse a profundidad en las investigaciones.

La investigación exploratoria nos permite ampliar y conocer los conocimientos sobre el tema de investigación para ser precisos en el problema a indagar.

3.1.2.3 Investigación descriptiva

Este tipo de investigación nos va a ayudar a comprender las situaciones importantes a través de la descripción de procesos identificando la relación que existen entre dos o más variables. Se recopilan los datos basándose en una hipótesis que se exponen de forma cuidadosa para ser analizados para la contribución de los conocimientos.

3.1.3 Diseño de la investigación y método de muestreo

3.1.3.1 Diseño de la investigación

En la investigación científica, el tamaño de la población y la muestra debe estimarse siguiendo los criterios que ofrece la estadística, y por ello es necesario conocer algunas técnicas o métodos de muestreo. El método de muestreo utilizado para estimar el tamaño de una muestra depende del tipo de investigación que desea realizarse y, por tanto, del diseño de investigación que se hayan definido para desarrollar el estudio.

En el caso de la presente investigación se seleccionó entre los siguientes probables diseños de investigación:

3.1.3.1.1 Diseño probabilísticos

Muestreo aleatorio simple: Es la forma más sencilla de seleccionar una muestra. Aquí, cada miembro tiene las mismas posibilidades de formar parte de la muestra. Los objetos de esta muestra son elegidos al azar y cada miembro tiene exactamente la misma probabilidad de ser elegido.

Muestreo estratificado: Es un proceso de división de los encuestados en parámetros distintivos pero predefinidos. En este método, los encuestados no se superponen, sino que representan colectivamente a toda la población.

3.1.3.1.2 Diseño no probabilísticos

Muestreo de conveniencia: El muestreo de conveniencia representa la conveniencia con la que el investigador puede llegar al encuestado. Los investigadores no tienen la autoridad para seleccionar las muestras y se realizan únicamente por motivos de proximidad y no representatividad.

Muestreo deliberado, crítico o por juicio: En este tipo de muestreo el investigador juzga y desarrolla su muestra sobre la naturaleza del estudio y la comprensión de su público objetivo.

Sólo se seleccionan las personas que se ajustan a los criterios de investigación y al objetivo final.

3.1.3.2 Método de muestreo

Cosiste en una selección de individuos para poder conocer algún dato y realizar estadísticas a partir de ellos para estimar información de toda una población.

Para llevar a cabo la recopilación de la información, se comparó los diferentes métodos de diseño de muestreo, tomando en consideración el acceso a las unidades muestrales y la facilidad de realizar una selección que cumpla el criterio de ser aleatorio (en el caso del muestreo aleatorio simple), consideración de todas estas variables se escogió un diseño no probabilísticos que recopile información a través de un muestreo por conveniencia (en este caso acceso a los paralelos y estudiantes a ser encuestados)

3.1.4 Técnica: Encuesta

A través de esta encuesta se pudo validar la funcionalidad, usabilidad y calidad de la herramienta y el nivel a conformidad y aprobación del usuario, es decir, que luego de interactuar e implementar las funciones de este modelo de búsqueda, se podrá obtener la información desde el punto de vista de la persona encuestada, su nivel de satisfacción frente al uso de este tipo de herramientas.

3.1.4.1 Población y muestra

3.1.4.1.1 Población

Para llevar a cabo el presente trabajo de investigación se ha considerado como población objetivo a estudiantes de la Universidad de Guayaquil, Carrera de Telemática que están cursando el noveno semestre y egresados.

Tabla 6. Estudiantes de la carrera de Ingeniería en Telemática

Semestre	Cantidad de estudiantes
Noveno Semestre	45
Egresados	50
Total	85

Información adaptada de la Universidad de Guayaquil. Elaborado por el autor

Como resultado de la suma de ambos cursos de noveno semestre, tenemos una cantidad de 85 estudiantes de ambas materias.

3.1.4.1.2 Muestra

De acuerdo con nuestra población establecida, la muestra es de 85 estudiantes, 45 alumnos de noveno semestre y 50 alumnos egresados, que son quienes realizarán la encuesta en la que podremos adquirir información significativa e importante.

3.1.4.2 Análisis de la encuesta

En el anexo 8, se puede visualizar el modelo de la encuesta que se realizó a los estudiantes de noveno semestre y egresados. La información que se obtuvo de 7 preguntas a 85 estudiantes encuestados se analizó con la finalidad de que ayuden al desempeño y cumplimiento de los objetivos trazados en el desarrollo de esta investigación.

Se presentan a continuación los resultados:

1.- ¿La herramienta le proporciona la información de manera rápida?

Tabla 7. Tabla de frecuencia de la pregunta ¿La herramienta le proporciona la información de manera rápida?

	Frecuencia	%Respuestas
Si	77	91,7 %
No	3	3,6 %
Tal vez	5	6 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

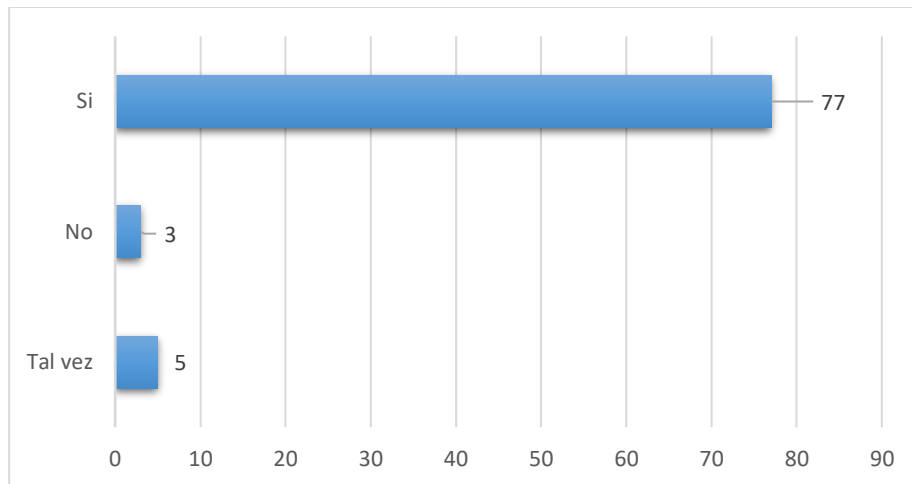


Ilustración 26. Gráficos estadísticos de la pregunta ¿La herramienta le proporciona la información de manera rápida?. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: Según los datos de la pregunta si la herramienta proporciona información de manera rápida, se puede observar que un total de 77 estudiantes respondieron “sí”, un total de 3 estudiantes respondieron “no” y 5 de ellos respondieron “tal vez”, por tanto, se aprecia que más del 50 % de usuarios encuestados manifiestan que la herramienta realiza una entrega rápida de información.

2.- ¿Considera usted que el flujo de búsqueda es claro y entendible?

Tabla 8. Tabla de frecuencia de la pregunta ¿Considera usted que el flujo de búsqueda es claro y entendible?

	Frecuencia	%Respuestas
Si	75	88,2 %
No	1	1,2 %
Tal vez	9	10,6 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

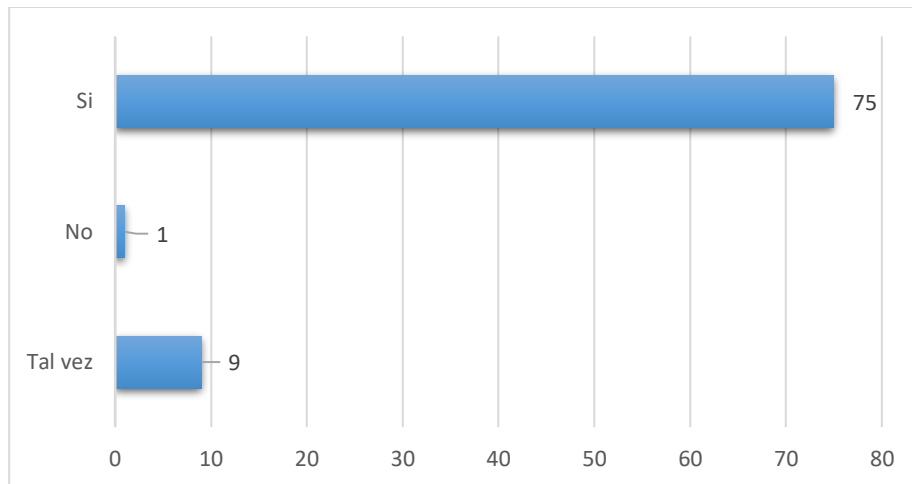


Ilustración 27. Gráficos estadísticos de la pregunta ¿Considera usted que el flujo de búsqueda es claro y entendible?. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: De acuerdo con la encuesta, acerca de la pregunta si el flujo de búsqueda es claro y entendible, se puede visualizar un total de 75 estudiantes que respondieron “si”, un estudiante que respondió “no” y 9 de ellos que respondió “tal vez”. Es decir que más del 50% de usuarios encuestados consideran que el modelo de búsqueda es claro y entendible.

3.- ¿Tuvo algún inconveniente al momento de realizar la búsqueda del producto?

Tabla 9. Tabla de frecuencia de la pregunta ¿Tuvo algún inconveniente al momento de realizar la búsqueda del producto?

	Frecuencia	%Respuestas
Si	12	14,1 %
No	69	81,2 %
Tal vez	7	8,2 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

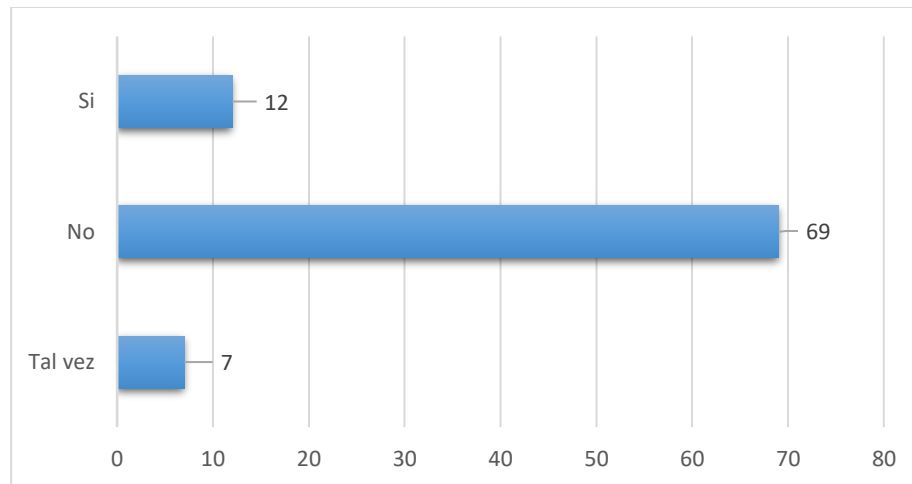


Ilustración 28. Gráficos estadísticos de la pregunta ¿Tuvo usted algún inconveniente al momento de realizar la búsqueda del producto?. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: Como se observa en los datos de la pregunta de la encuesta, si tuvo inconvenientes al momento de realizar la búsqueda del producto, tenemos un total de 12 estudiantes que responden “sí”, un total de 69 estudiantes que responden “no” y 7 de ellos que responden “tal vez”. Por lo tanto, se muestra que hay más del 50 % de usuarios que manifiesta que no tuvieron inconvenientes al momento de realizar la búsqueda del producto.

4.- ¿La herramienta le ayudo a reducir el tiempo de cotización del producto?

Tabla 10. Tabla de frecuencia de la pregunta ¿La herramienta le ayudo a reducir el tiempo de cotización del producto?

	Frecuencia	%Respuestas
Si	78	91,8 %
No	5	5,9 %
Tal vez	4	4,7 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

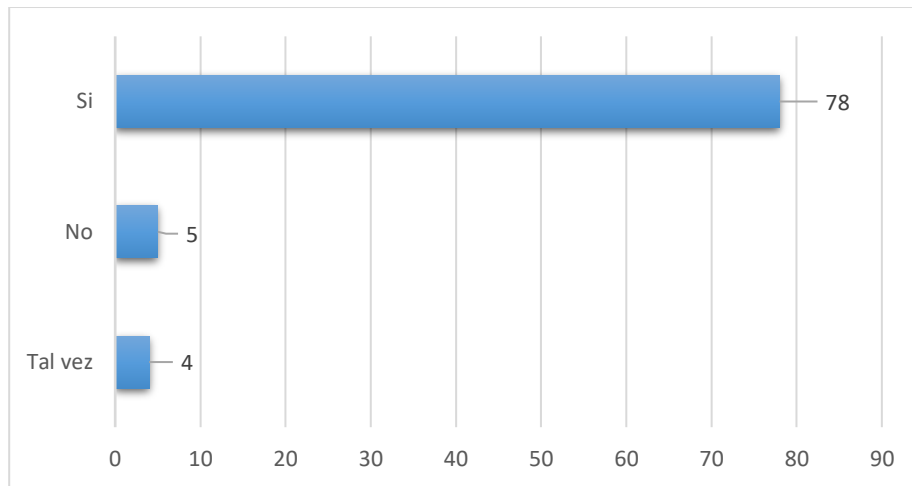


Ilustración 29. Gráficos estadísticos de la pregunta ¿La herramienta le ayudo a reducir el tiempo de cotización del producto?. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: De acuerdo a la encuesta acerca de que, si la herramienta ayuda a reducir el tiempo de cotización del producto, encontramos que hay un total de 78 estudiantes que respondieron “sí”, un total de 5 estudiantes que respondieron “no” y 4 estudiantes que respondieron “tal vez”. Por tanto, se llega a la determinación de más del 50% de usuarios consideran que la herramienta ayuda a reducir el tiempo de cotización del producto?

5.- ¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?

Considerando un rango del 1 al 5 donde 1 es No recomendable y 5 muy recomendable

Tabla 11. Tabla de frecuencia de la pregunta ¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?

	Frecuencia	%Respuestas
1	0	0 %
2	1	1,2 %
3	9	10,6 %
4	35	41,2 %
5	40	47,1 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

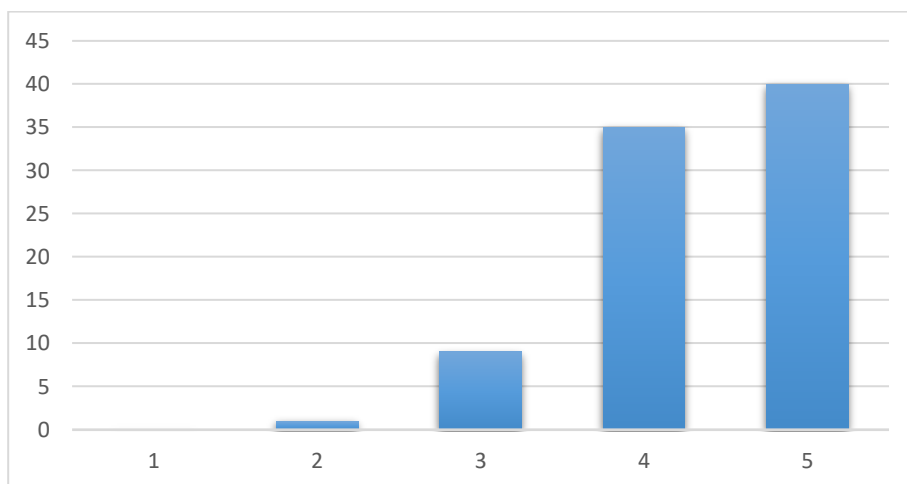


Ilustración 30. Gráficos estadísticos de la pregunta *¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?*. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: De acuerdo con los datos de los encuestados, los estudiantes consideran que si es recomendable el uso de una herramienta de asistente virtual

6.- ¿Considerarías usar un asistente virtual para sus búsquedas futuras?

Tabla 12. Tabla de frecuencia de la pregunta *¿Considerarías usar un asistente virtual para sus búsquedas futuras?*

	Frecuencia	%Respuestas
Si	73	85,9 %
No	1	1,2 %
Tal vez	12	14,1 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

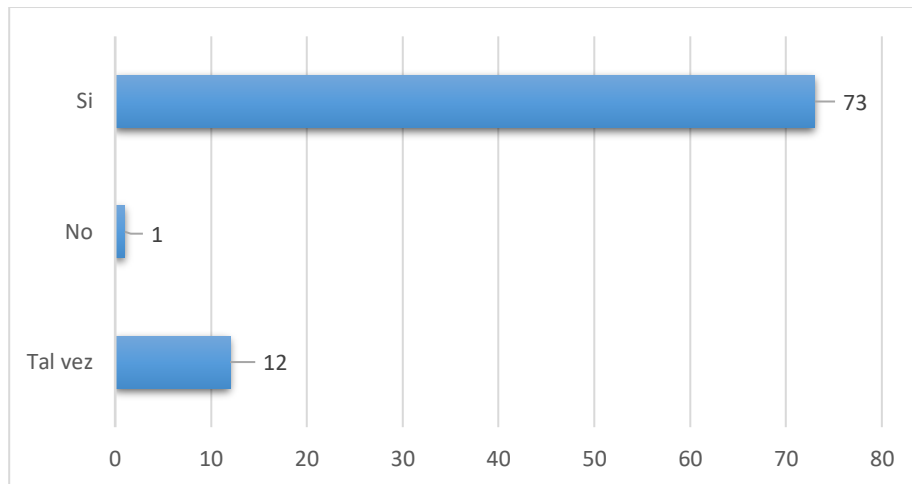


Ilustración 31. Gráficos estadísticos de la pregunta ¿Considerarías usar un asistente virtual para sus búsquedas futuras?. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: Según se visualiza en la pregunta si se consideraría usar un asistente virtual para búsquedas futuras, se aprecia que 73 estudiantes respondieron “sí”, 1 respondió “no” y 12 de los estudiantes respondió “tal vez”. Por lo tanto, se considera que más del 50 % de usuarios consideraría usar un asistente virtual para búsquedas futuras.

7.- Coloque del 1 al 5 que tanto le agrado la herramienta

Considerando un rango del 1 al 5 donde 1 es No recomendable y 5 muy recomendable

Tabla 13. Tabla de frecuencia. Coloque del 1 al 5 que tanto le agrado la herramienta

	Frecuencia	%Respuestas
1	1	1.2 %
2	1	1,2 %
3	4	4,7 %
4	39	45,9 %
5	40	47,1 %

Información tomada de la encuesta realizada en forms.office.com. Elaborada por Ruth Apolo

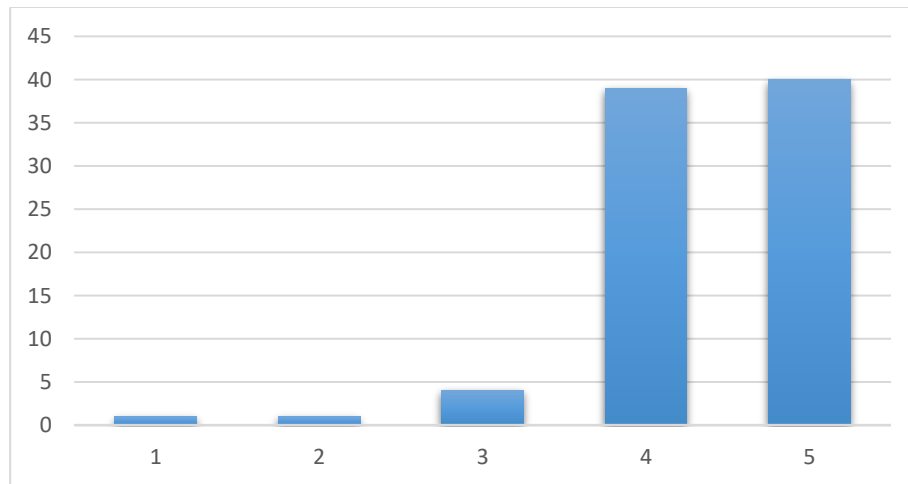


Ilustración 32. Gráficos estadísticos Coloque del 1 al 5 que tanto le agrado la herramienta. Información tomada de la entrevista realizada en forms.office.com. Elaborada por Ruth Apolo

Respuesta: De acuerdo con los datos de los usuarios encuestados, estos consideran que la herramienta de asistente virtual si es de su agrado.

3.1.5 Método y técnica

3.1.5.1 Método de PLN y extracción de datos

3.1.5.1.1 Método de reconocimiento de nombres de entidades no anidados

Comienza al obtener una serie de frases de una oración procesada. Estas frases son candidatas a ser los nombres de personas, lugares, organizaciones, etc. Luego de ello, se procede a clasificar la secuencia de frases candidatas.

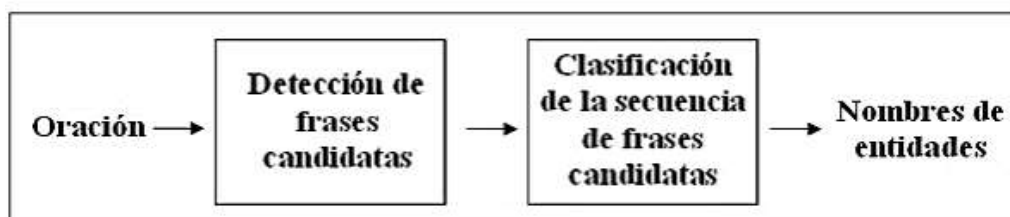


Ilustración 33. Reconocimiento de nombres no anidados. Obtenido de Repositorio. Elaborado por el Autor

3.1.5.1.2 Método de Extraer datos

Para la realizar la técnica de extracción de datos se aplicarán los siguientes pasos que se muestran en la siguiente ilustración 24.



Ilustración 34. Pasos webscraping. Obtenido de Repositorio. Elaborado por Ruth Apolo

3.1.5.2 Técnica de PLN y Webscraping

3.1.5.2.1 Técnica utilizada para el PLN

- *Interfaces en lenguaje natural:* es una aplicación que facilita la interacción con las computadoras haciendo uso del lenguaje natural, en donde el usuario enviar órdenes a la máquina y esta debe ser capaz de entenderlos y ejecutarlos.

3.1.5.2.2 Técnica para extraer datos

- *Webscraping:* una vez analizado la estructura de la página web en donde se desea obtener la información y teniendo en cuenta que campos son los que se necesita extraer, lograremos utilizar la aplicación conveniente que nos permita recolectar dicha información.

3.2 Descripción del proyecto

La propuesta de este proyecto, radica en usar un asistente virtual como herramienta de búsqueda de artículos de tecnología, la base de datos que contiene detalles de cada uno de los artículos, estarán conectados al asistente virtual y se realizará la respectiva configuración para que este pueda realizar la búsqueda que el usuario requiera y ejecute al momento de interactuar con la herramienta. Los datos serán obtenidos a partir del uso de una técnica de extracción de información, denominada web scraping, lo que nos asegura que la extracción de dicha información será real, rápida y sin errores, facilitando así la recaudación de datos que serán recopilados de algunas páginas web seleccionadas de venta online y ofrecerle al

consumidor un servicio de búsqueda precisa y que se ajuste a los requerimientos que este necesite.

3.2.1 Factibilidad técnica

Para el proceso de desarrollo de la base de datos, se tomará en cuenta algunos factores técnicos que ayudaran a la extracción de información.

El framework de Scrapy, es una herramienta de rastreo de sitios web de código libre, contiene un conjunto de implementaciones poderosas para realizar “scraping” o recopilar datos de la web de forma rápida y sencilla. La funcionalidad clave del desarrollo de Scrapy son los “spiders” o “arañas” que están programadas para que extraigan información de una web en concreto y este pueda ir desglosándose de página en página para su respectivo almacenamiento en un archivo csv. Posteriormente la información recaudada será trasladada a una plantilla botsheets, que es una hoja de cálculo usada como base de datos previamente definida y estructurada para conectarla con el asistente virtual y generar búsquedas de datos almacenados. Scrapy es escrito en Python por lo que es la herramienta seleccionada para escribir nuestro código de extracción de datos, además que contiene puntos clave en el procesamiento de texto y apertura de recursos web que son bases técnicas del web scraping. Una de las ventajas de usar Python para desarrollar nuestro código es que es un software libre y gratuito, por ello es recomendable a la hora de crear programas de extracción de datos. Finalmente se usará Manychat para la creación de nuestro asistente virtual ya que es una plataforma vinculada a Facebook con la que se puede interactuar en una interfaz amigable con el usuario, una de sus ventajas es que no necesitas ser un experto programador para crear tu bot, puedes lograrlo de una manera sencilla y eficiente.

A continuación, se detalla el hardware y software usados para el desarrollo del proyecto:

Tabla 14. Hardware usado para el desarrollo del proyecto

Hardware	Características
HP Laptop 15 da0xxx	Windows 10 Home Single Language Procesador Inter® Core TM i5-7200U CPU @ 2.50 GHz 2.70 GHz RAM 4,00 GB Almacenamiento 1000 GB

Información tomada de la presente investigación. Elaborada por Ruth Apolo

Tabla 15. Softwares usados en la extracción de datos

Software	Version
Python	3
Scrapy	2
Formato CSV	-

Información tomada de la presente investigación. Elaborada por Ruth Apolo

Tabla 16. Softwares usados para el asistente virtual

Software	Version
Manychat	Pro
Botsheets	28
Google Drive	93

Información tomada de la presente investigación. Elaborada por Ruth Apolo

3.2.2 Factibilidad legal

Para el desarrollo de este proyecto, los datos obtenidos por web scraping, es información que procede de fuentes de acceso público y que serán usados para realizar consultas con la ayuda una herramienta de búsqueda, no se busca vulnerar ninguna ley de privacidad de datos o derechos del consumidor.

Posteriormente, se detallan los artículos tomados como referencia, que son respaldo en la constitución de la república del Ecuador, como:

- Ley orgánica de defensa del consumidor
- Ley orgánica de protección de datos personales
- Constitución de la república del Ecuador

3.2.3 Factibilidad Económica

Para el proceso de creación de este proyecto, se utilizó herramientas de software libre y un software de pago a un costo muy accesible. El enfoque que tiene este proyecto es poder ayudar a los consumidores a menorar ese tiempo de cotización y búsqueda que le toma al tratar de conseguir un producto que se ajuste a sus necesidades y que se ajuste a su economía sin necesidad de salir de casa. También se ha hecho necesario el uso de recursos básicos para la ejecución de este proyecto durante una fase de tiempo de un mes.

Tabla 17. Costo total del requerimiento del software

Rubros	Valor	Meses	Subtotal
--------	-------	-------	----------

Recursos de software	\$10	1	\$10
Conectividad a internet	\$30	1	\$30
Servicio básico (energía eléctrica)	\$20	1	\$20
Total			\$60

Información tomada de la presente investigación. Elaborada por Ruth Apolo

La factibilidad económica para los requerimientos de software como, recursos de software, conectividad a internet y servicio básico como la energía eléctrica, nos muestra un total de \$60 que serán invertidos en el presente proyecto y será cubierta totalmente por el estudiante.

3.2.4 Factibilidad Operacional

La elaboración de este proyecto está enfocada en el web scraping y el e-commerce, lo que beneficia a que los consumidores puedan realizar sus consultas de varios productos de diferentes sitios en una sola herramienta para que puedan realizar la compra del artículo sin tener que salir de su hogar y pasar varias horas en diferentes tiendas tratando de conseguir algo que se ajuste a sus requerimientos y a su bolsillo. Este tipo de herramientas también beneficia a pequeñas y grandes empresas que necesiten ofrecer un mejor servicio para sus clientes, agilizando su proceso de comprar.

Actualmente, se presentan acontecimientos que no podemos controlar y que ha perjudicado algunos emprendimientos por lo que dar soluciones a estos problemas beneficia tanto a la empresa de poder vender sus productos y al consumidor de poder adquirirlos de manera fácil y sin riesgo alguno.

La herramienta Manychat es muy recomendable ya que tiene una interfaz amigable que facilita su uso ya que no es necesario tener conocimientos de programación para operarlo, el usuario puede interactuar con la aplicación desde cualquier computador con acceso a internet ya que solo necesita acceder a su cuenta para poder manipular los flujos de conversación creados.

3.3 Esquema general de la extracción de datos y asistente virtual

En el siguiente esquema se detalla el proceso que se va a efectuar para la técnica de web scraping y extraer la información de algunas páginas web seleccionadas de empresas conocidas en el Ecuador. El proceso está constituido por cinco fases, en la ilustración 31 se puede observar desde la primera que corresponde a los sitios web seleccionados para la recopilación de información, hasta la última que corresponde a conectar el bot con la base

de datos y la plataforma Facebook Messenger, asimismo se puede observar los softwares utilizados para este desarrollo.



Ilustración 35. Esquema de la extracción de datos y asistente virtual. Información tomada de la presente investigación. Elaborado por Ruth Apolo

3.4 Recursos para el desarrollo del proyecto

3.4.1 Recursos para la extracción de datos

Para el proceso de extraer los datos fue importante seleccionar y reconocer que herramientas serían las más óptimas para el desarrollo de web scraping, estas fueron detalladas en anterior capítulo II, a través de un análisis se seleccionó las aplicaciones más convenientes y así alcanzar los objetivos ya planificados y lograr la ejecución del proyecto de forma eficiente.

Tabla 18. Tabla comparativa de herramientas y servicios de raspado

	Scrapy	Requests	Beautiful Soup	Selenium
Objetivo	Solución completa de web scraping.	Simplifica la realización de solicitudes HTTP.	Analizador de datos.	Navegador web programable para renderizar JavaScript.
Caso de uso ideal	Desarrollo de proyectos	Tareas de raspado web	Tareas de raspado web	Raspado web a pequeña escala

	de web scraping a gran escala.	simples no recurrentes.	simples no recurrentes.	de sitios web pesados con JavaScript.
Soportes de almacenamiento de datos integrados	JSON, JSON lines, XML, CSV.	Necesitas desarrollar la tuya.	Necesitas desarrollar la tuya.	Personalizable.
Precio	Gratis.	Gratis.	Gratis.	Gratis.

3.4.1.1 Scrapy

Se selecciono esta herramienta ya que es una plataforma que se acopla a nuestras necesidades porque nos va permitir crear scripts que serán los encargados de recorrer los sitios web seleccionados para la creación de este proyecto, extrayendo los contenidos de los campos requeridos. Para alcanzar estos objetivos, este framework será escrito en Python ya que es una plataforma colaborativa de esta herramienta.

3.4.1.2 Páginas web

Para el alcance y desarrollo del presente proyecto es necesario seleccionar cuáles serán los sitios web a lo que podemos extraer información, ya que para realizar este tipo de técnicas se debe tomar en cuenta las políticas de cada sitio. Para ello se escogieron 5 páginas web de empresas reconocidas en el Ecuador que se dedican a la venta de artículos de categoría tecnológica.

A continuación, se presenta las páginas web de donde se extraerá los datos

Tabla 19. Páginas web para extraer datos

Sitios Web	Categoría de productos en venta
Artefacta	-Línea blanca
	-Climatización
	-Tecnología
	-Muebles
	-Motos
	-Hogar

La Ganga	-Línea blanca -Climatización -Tecnología -Muebles -Motos -Hogar
Pycca	-Tecnología -Muebles -Electro hogar -Hogar -Decoración
Computron	-Tecnología
Novicompu	-Tecnología

Información tomada de la presente investigación. Elaborada por Ruth Apolo

3.4.1.3 Lenguaje de programación: Python

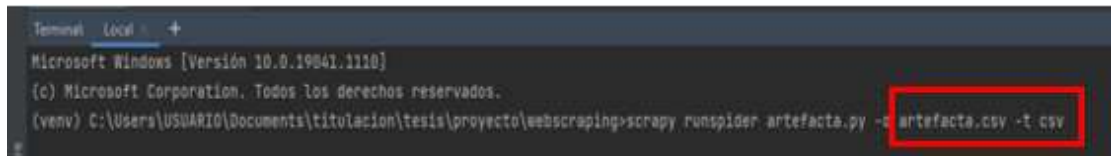
Para la extracción de datos se seleccionó la herramienta de Python ya que es un lenguaje fácil de usar que se puede utilizar en cualquier sistema operativo, es legible y entendible para cualquier usuario programador que desee desarrollar código sobre este tipo de estructuras y, asimismo, contiene librerías para llevar a cabo técnicas de web scraping. Además, Python es una herramienta especializada en el procesamiento de texto y aperturas de recursos web, que son bases técnicas para la extracción de información y que cuenta con bibliotecas y proyectos de código abierto, un extenso ecosistema de programación.

Esta herramienta cumple con todas las características necesarias para el desarrollo y ejecución del presente proyecto, ya que Scrapy se escribe en este lenguaje, que es un framework fundamental para el proceso de este tipo de técnicas, webscraping.

3.4.1.4 Formato CSV

Este tipo de archivos, con extensión .csv, es que utilizaremos para cargar toda la información recaudada de la técnica webscraping y posteriormente almacenarlas en una base de datos que nos permitirá alcanzar el desarrollo de consultas de artículos. Este tipo de archivos presentan los datos que se almacenaron separados por comas, pero puede editarse en Excel

ya que este identifica automáticamente sus distintos separadores en forma de tabla. Al momento de ejecutar el código se debe detallar el formato en el que queremos que se guarden los datos, en este caso, csv como se muestra a continuación en la ilustración 32.



```

Terminal Local +
Microsoft Windows [Versión 10.0.19041.1110]
(c) Microsoft Corporation. Todos los derechos reservados.
(venv) C:\Users\USUARIO\Documents\titulacion\tesis\proyecto\webscraping>scrapy runspider artefacta.py -o artefacta.csv -t csv
  
```

Ilustración 36. Formato cvs. Información tomada de la presente investigación. Elaborado por Ruth Apolo

3.4.2 Recursos para la creación del asistente virtual

3.4.2.1 Pagina de Facebook

Para poder ejecutar las consultas de artículos es necesario disponer de un entorno que brinde una atención personalizada a los consumidores, Facebook es una plataforma muy utilizada en el mundo, millones de personas disponen de una cuenta de esta plataforma y es por ello que a través de él es más fácil llegar a muchas personas, ya que una cuenta puede ser configurada para que toda clase de público pueda tener acceso a tu cuenta o página.

Facebook es una plataforma que ha logrado ayudar en gran manera al comercio electrónico ya que dispone de datos, herramientas publicitarias y recursos para acelerar el proceso de compra, dispone que muchas estrategias para elevar la demanda del e-commerce, entre ellas la adaptación de una bot que permita mejorar la experiencia con el consumidor brindando un mejor servicio para el usuario, lo que nos favorece la creación de nuestro proyecto y nos ayuda a tener mejor resultados siento esta red social la más usada por los usuarios.

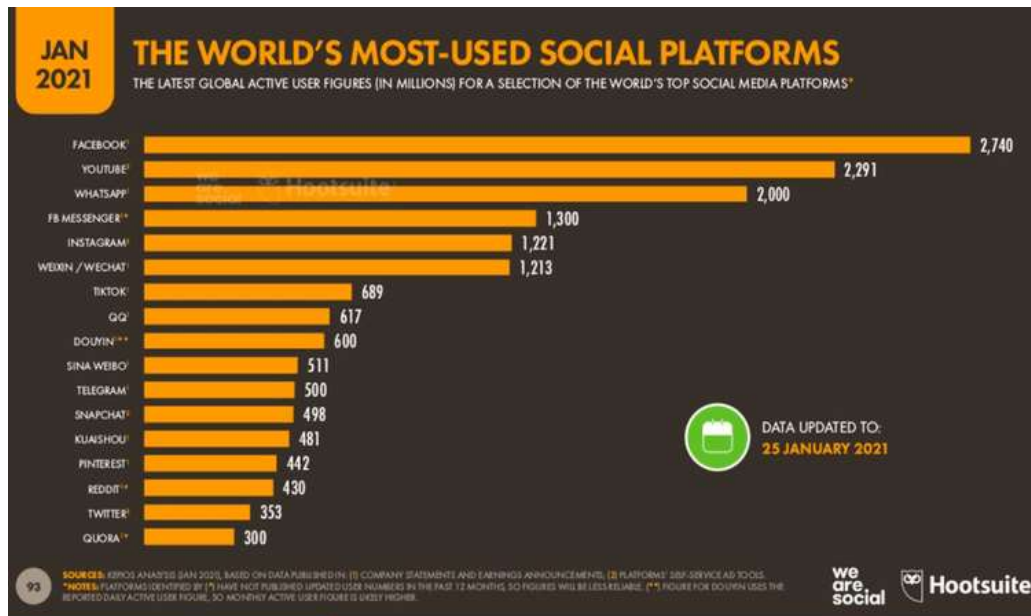


Ilustración 37. Datos estadísticos redes sociales. Tomado de marketing4ecommerce. Elaborado el Autor

3.4.2.2 Botsheets

Esta herramienta será usada para almacenar toda la información recaudada de la técnica de web scraping debido a que la utilizaremos como nuestra base de datos. Esta es una plantilla de hoja de cálculo con características especialmente diseñadas para la consulta y búsqueda de los datos almacenados, lo que nos favorece al momento de conectarla con nuestro asistente virtual y poder tener una mejor interacción entre el usuario y el bot, logrando que podamos ofrecer un servicio con buenos resultados.

3.4.2.3 Manychat

Para la creación de nuestro bot y la interacción con el usuario, emplearemos esta herramienta que nos ofrece un entorno amigable para la ejecución de nuestro proyecto, ya que no necesitaremos tener conocimientos de programación para crear flujos de chat que sean acorde a los objetivos de las consultas de artículos en la que nos hemos enfocado, además, permite instalar plantillas que ayudan a mejorar la experiencia del consumidor y facilitan el desarrollo del presente trabajo.

Es una plataforma con la que podemos interactuar desde cualquier dispositivo inteligente que tenga acceso a internet ya que, con solo enlazarse a nuestra cuenta personal de manychat, podremos realizar las respectivas modificaciones en el caso de que sea necesario. Contiene un conjunto de herramientas con las que podemos crear y desarrollar chat que mejoren nuestro servicio, como usar la opción de base de datos, que es una de las características más importantes para la ejecución de nuestro proyecto y poder llevar a cabo nuestros objetivos.

3.5 Procedimiento: Aplicación de la metodología escogida

Para el desarrollo de este trabajo se efectuará cinco fases que esta compuestas por actividades en donde se detallan los procesos a seguir para cumplir con los objetivos propuesto de este trabajo.

Se explica a continuación cada fase:

3.5.1 Identificación de Sitios web

Es importante identificar los sitios web de donde vamos a extraer los datos que necesitaremos para alcanzar los objetivos propuestos, los paginas seleccionadas tienen que ir acorde a la herramienta que usaremos para realizar web scraping ya que se debe ejecutar de acuerdo a los parámetros de estas.

Se debe tener en cuenta que la aplicación de este tipo de técnicas no lleve a infringir con las políticas y leyes de cada página y que la estructura HTML del sitio no sea compleja para que se nos facilite la ejecución de extracción de información. Para ello se escogió 5 páginas web de empresas reconocidas del Ecuador (Artefacta, la Ganga, Computron, Novicompu, Pycca), que se dedican a la venta de productos de diferentes categorías, entre ellas, tecnología (categoría laptops) que es en la que nos enfocaremos para este proceso.



Ilustración 38. Datos seleccionados de extracción. Tomado de Artefacta. Elaborado por Ruth Apolo

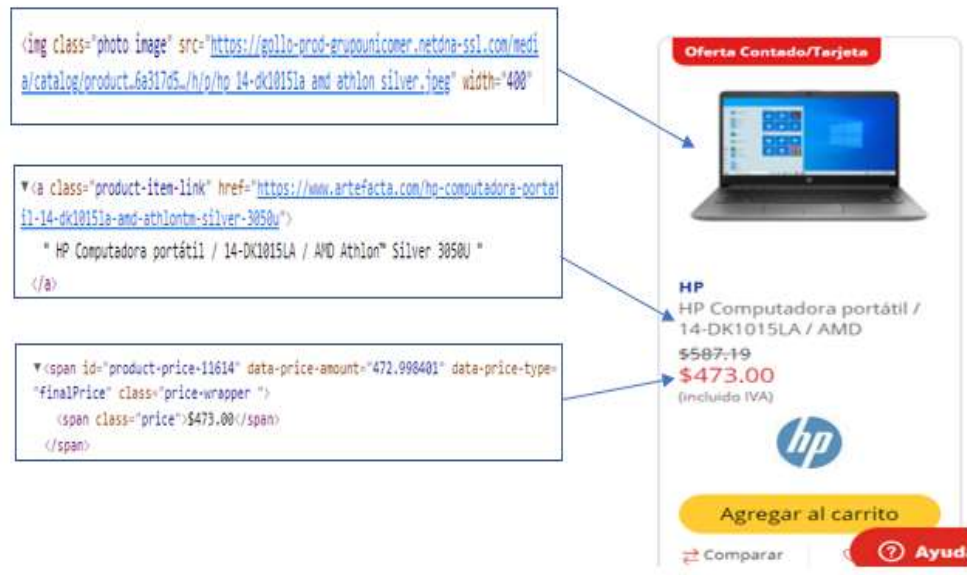
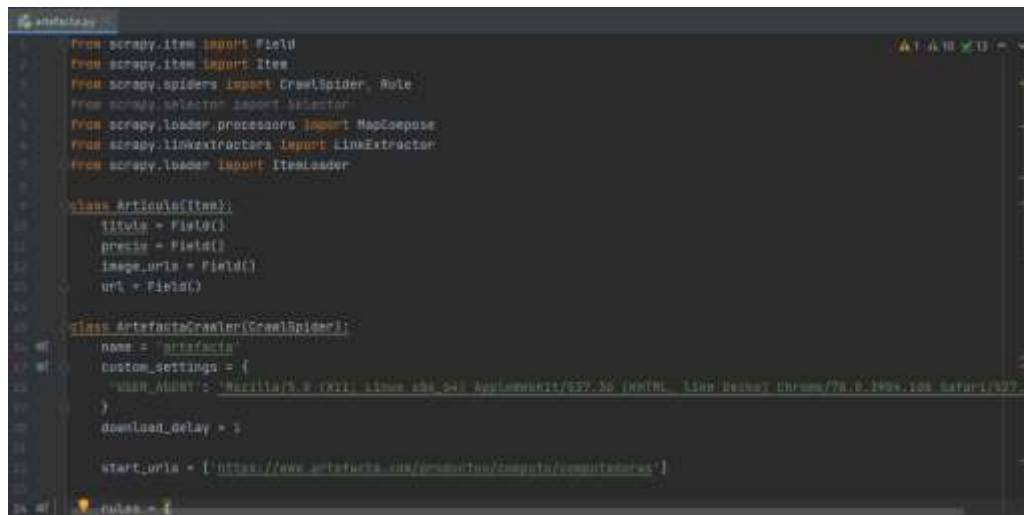


Ilustración 39. Estructura HTML del producto. Tomado de Artefacta. Elaborado por el Autor

3.5.2 Web scraping

Una vez tengamos los recursos para desarrollar la técnica webscraping, en este caso tener instalado tanto el framework de scrapy y la herramienta Python podemos proceder a escribir nuestro código de extracción.

Lo primero que se debe realizar es definir los ítems que vamos a extracción en la página y definir cada uno de los campos. Luego definimos la clase que hará la extracción de los datos, la definimos con un nombre y esta clase puede ser “CrawlSpider” o “Spider” dependiendo de cuál sea la estructura de la página que se necesita extraer. Adicionalmente definimos la url semilla que es la página en donde se encuentra los productos que queremos extraer.



```

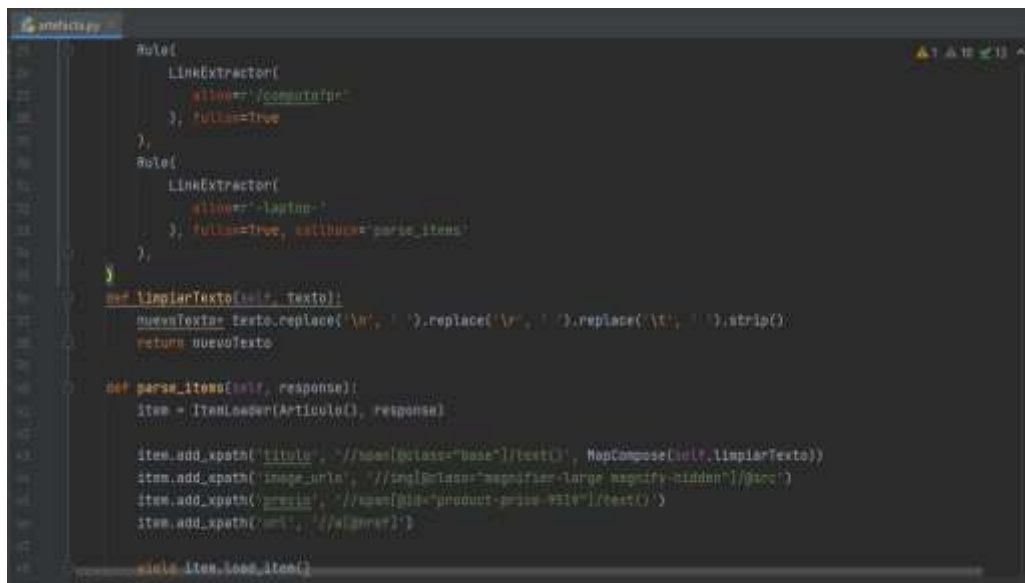
1 from scrapy.item import Field
2 from scrapy.item import Item
3 from scrapy.spiders import CrawlSpider, Rule
4 from scrapy.selector import Selector
5 from scrapy.loader.processors import MapCompose
6 from scrapy.linkextractors import LinkExtractor
7 from scrapy.loader import ItemLoader
8
9 class Artículo(Item):
10     titulo = Field()
11     precio = Field()
12     image_urls = Field()
13     url = Field()
14
15 class ArtifactsCrawler(CrawlSpider):
16     name = 'artifacts'
17     custom_settings = {
18         'USER_AGENT': 'Mozilla/5.0 (X11; Linux i686_386) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/78.0.3904.108 Safari/537.36'
19     }
20     download_delay = 1
21
22     start_urls = ['https://www.artifacts.cha/production/images/computadoras']
23
24     rules = (

```

Ilustración 40. Definir campos python. Información adaptada de Python. Elaborado por Ruth Apolo

Siguiendo con el procedimiento, lo siguiente sería determinar las reglas que se utilizarán para la extracción de datos, en donde se va a detallar los patrones o palabras claves de las url de las páginas para que scrapy pueda dirigirse a la página correcta.

Para terminar con el desarrollo de la codificación se empiezan a llenar cada uno de los campos de los ítems que queremos extraer de la página, por ello, se necesita conocer bien la estructura HTML de la página.



```

25     Rule(
26         LinkExtractor(
27             allow=('/computador')
28         ), follow=True
29     ),
30     Rule(
31         LinkExtractor(
32             allow=('/laptop')
33         ), follow=True, callback='parse_items'
34     ),
35 ]
36
37 def limpiarTexto(self, texto):
38     nuevoTexto = texto.replace('\n', ' ').replace('\r', ' ').replace('\t', ' ').strip()
39     return nuevoTexto
40
41 def parse_items(self, response):
42     itme = ItemLoader(Artículo(), response)
43
44     itme.add_xpath('titulo', '//span[@class="name"]/text()', MapCompose(self.limpiarTexto))
45     itme.add_xpath('image_urls', '//img[@class="magnifier-large magnify-hidden"]/@src')
46     itme.add_xpath('precio', '//span[@id="product-price-9119"]/text()')
47     itme.add_xpath('url', '//a/@href')
48
49     return itme.load_item()
50
51 def crawl(self):

```

Ilustración 41. Llenar ítems de extracción. Información adaptada de Python. Elaborado por Ruth Apolo

Es necesario que para esto tengamos previamente creada nuestra cuenta de Facebook que vamos a conectar a nuestra cuenta de manychat y así los usuarios puedan lograr interactuar con el asistente virtual que les ayudara con la consulta requerida.



Ilustración 44. Página de Facebook. Tomado de Facebook. Elaborado por el Autor

Una vez tengamos lista nuestra página procedemos a conectarla con la cuenta de Manychat para poder seguir con el proceso de creación del bot.

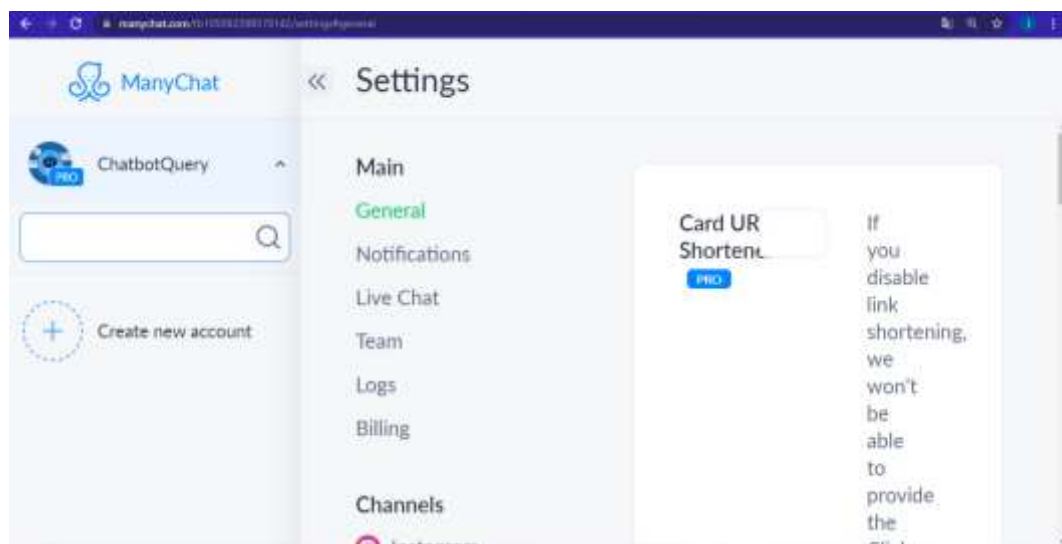


Ilustración 45. Herramienta Manychat. Tomado de Manychat. Elaborado por el Autor

3.5.5.2 Conectar el bot con la base de datos

Manychat dispone de muchas herramientas y apps que le permite brindar un mejor servicio a sus usuarios a la hora que implementar un chatbot, entre ellas se encuentra una denominada botsheets que es la que usaremos como base de datos de toda la información que recopilamos con el uso de la herramienta de webscraping.

Para conectar nuestra base de datos con Manychat, es necesario conectarnos botsheets con nuestra cuenta de Google ya que al momento de obtener la plantilla que nos aporta botsheets podemos crear una copia y guardarla en Google Drive para su modificación respectiva.



Ilustración 46. Conexión con Botsheets. Tomado de Manychat. Elaborado por el Autor

Es necesario que para que la información se muestre en al usuario el enlace de la hoja de cálculo debe estar dentro de los parámetros de la herramienta instalada en Manychat, es decir, dentro de botsheets y a su vez debe estar compartida para todo el público.

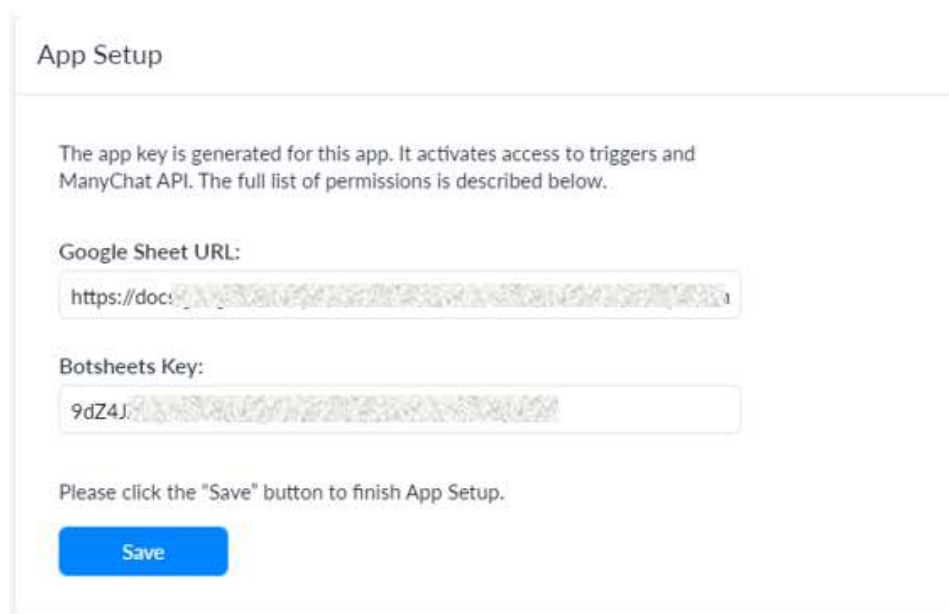


Ilustración 47. Conexión con hoja de datos. Tomado de Manychat. Elaborado por el Autor

Una vez tengamos todo listo procedemos a modificar la plantilla de hoja de cálculo de botsheets. Este integra diferentes hojas con diferentes platillas que podemos modificar según sea nuestra necesidad. Escogimos la plantilla que más se adecua a nuestras necesidades y procedimos a ingresar toda nuestra información (precio, titulo, url de la imagen, url de la página) en nuestra base de datos.

	product_id	price	title	keywords	subtitle	image	website	addbutton
1	1	\$179.11	Laptop HP Aquil Indigo 15 4730	laptop laptops hp hp	PVECA	https://www.pveca.com/bo	https://www.pveca.com/bo	
2	2	\$1,014.0	Laptop HP N0014P000 14"	laptop laptops hp hp	PVECA	https://www.pveca.com/bo	https://www.pveca.com/bo	
3	3	\$1,218.8	Laptop Dell 1390 13.3" Core5	laptop laptops dell dell	PVECA	https://www.pveca.com/bo	https://www.pveca.com/bo	
4	4	299.00 \$	Laptop ENV 14.2 inch, 12gb,	laptop laptops ENV env	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
5	5	895.00 \$	Laptop Dell AMD Ryzen 5, 8gb	laptop laptops dell dell	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
6	6	345.00 \$	Laptop ENV 14.2 inch Intel, 21	laptop laptops ENV	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
7	7	895.00 \$	Laptop HP Core i5 11th, 12gb,	laptop laptops hp hp	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
8	8	670.00 \$	Laptop Acer Gateway Ryzen 5,	laptop laptops acer acer	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
9	9	895.00 \$	Laptop Lenovo Flex 5, Ryzen 7	laptop laptops lenovo lenovo	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
10	10	670.00 \$	Laptop Acer Gateway Core i5,	laptop laptops acer acer	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
11	11	920.00 \$	Laptop Dell Ryzen 5, 8gb, 254g	laptop laptops dell dell	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
12	12	895.00 \$	Laptop HP AMD Ryzen 7, 8gb,	laptop laptops hp hp	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
13	13	525.00 \$	Laptop HP AMD Ryzen 3 325G	laptop laptops hp hp	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
14	14	795.00 \$	Laptop Acer Core i5 10ma, 8gb	laptop laptops acer acer	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
15	15	630.00 \$	Laptop Lenovo Core i3 9ma, 5	laptop laptops lenovo lenovo	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
16	16	895.00 \$	Laptop Dell Core i3 11th, 12gb	laptop laptops dell dell	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
17	17	895.00 \$	Laptop Acer Core i3 10ma, 8gb	laptop laptops acer acer	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	
18	18	635.00 \$	Laptop HP Core i5, 12gb Ram,	laptop laptops hp hp	NOVCOMPU	https://www.novocompu.com/bo	https://www.novocompu.com/bo	

Ilustración 48. Base de datos. Tomado de Google Drive. Elaborado por el Autor

3.5.5.3 Creación del Asistente Virtual

Para este paso lo primero que vamos a crear es un campo de usuario, lo que esto nos va a permitir es que cada consulta que realice el usuario se almacene en este campo, al que me asignamos un nombre “entrada del usuario” como se puede apreciar en la ilustración 44

Create New User Field

Custom attributes let you save info about your subscribers. Store user emails, phones, appointments, behavior or anything else you wish. Later you can segment your audience based on this data.

Name Type

Description (Optional)

Folder

Ilustración 49. Crear campo de usuario. Información adaptada de Manychat. Elaborado por Ruth Apolo

El siguiente paso a seguir es empezar a crear nuestro flujo, y para ello crearemos una carpeta en donde se van a guardar los flujos de chat que crearemos para nuestro asistente virtual de consulta. La carpeta la hemos denominado como “botsheets” y el nombre del flujo “Galería dinámica”.



Ilustración 50. Creación de flujo. Información adaptada de Manychat. Elaborado por Ruth Apolo

Empezamos realizando los flujos de chat que tendrá el bot con el usuario y lo primero que realizamos es desarrollo del mensaje principal, en donde comienza con un saludo y para una mejor experiencia imprime el primer nombre de la persona que esta interactuando con el bot y un retraso de 3 segundos para que simule la escritura del bot, de igual manera se detalla las marcas del producto de tecnología que disponemos (categoría laptops) para realicen la consulta y esa consulta se almacene en el campo de usuario que habíamos creado previamente como “entrada de usuario”.

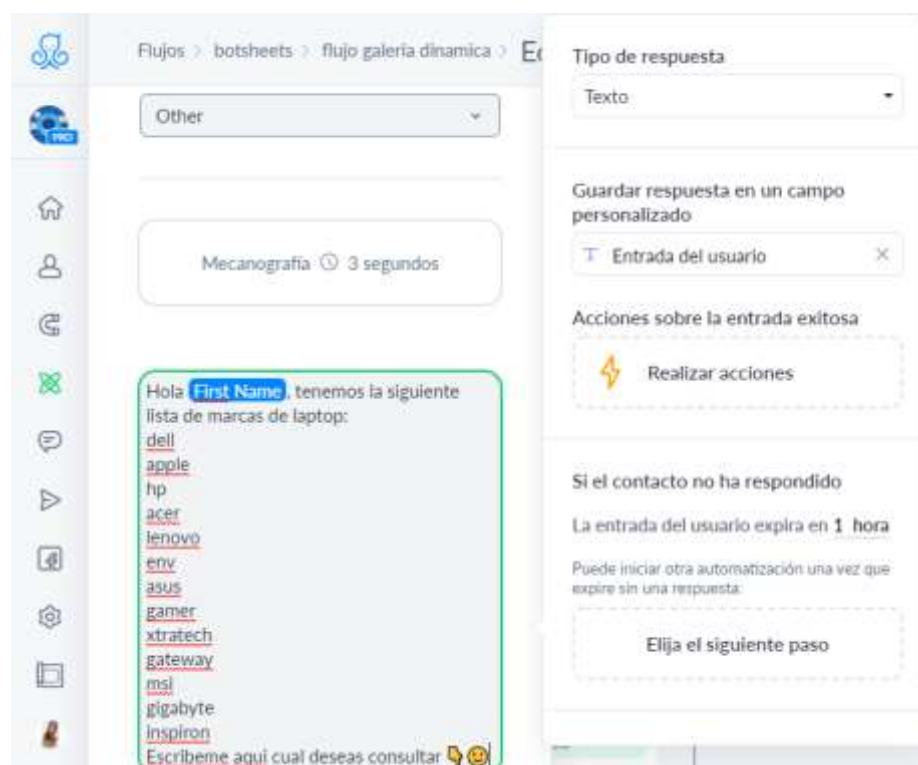


Ilustración 51. Mensaje presentación. Información adaptada de Manychat. Elaborado por Ruth Apolo

Para el segundo flujo lo haremos en uno de “acción” que nos va a permitir realizar la integración del bot con nuestra base de datos en donde se van realizar la consulta respectiva de los usuarios, para ello, seleccionamos botsheets que es nuestra hoja de cálculo en donde tenemos almacenados nuestros datos y realizamos la configuración correspondiente para que este realice la búsqueda en la plantilla de galería con palabras claves, allí nos mostrara detalles de cómo queremos que se realice la búsqueda.

Editar acción de Botsheets

Esta acción utiliza los siguientes campos del sistema: Id. De usuario, Id. De página, Nombre

Nombre de la hoja de trabajo de Google *

¿Imágenes de galería horizontales o cuadradas? *

Número máximo de Gallery Cards *

Si no hay respuesta, ¿a qué flujo recurrir? *

Flujo para enviar después de la galería [Opcional]

Ilustración 52. Flujo acción. Información adaptada de Manychat. Elaborado por Ruth Apolo

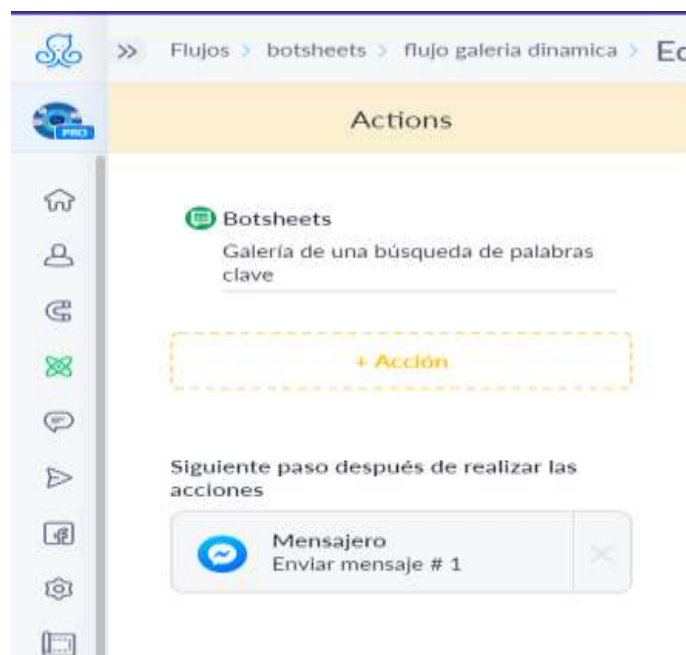


Ilustración 53. Flujo acción configurado. Información adaptada de Manychat. Elaborado por Ruth Apolo

Para el tercer flujo realizamos un mensaje de condición para saber si el usuario desea seguir con sus consultas o ya no necesita seguir haciendo su búsqueda, para este paso

implementamos dos botones que le permite al consumidor seleccionar su respuesta y que el bot pueda seguir ejecutándose, para este flujo también implementamos un retraso de tres segundos que simule la escritura del bot y mejore la experiencia la interacción con el usuario.



Ilustración 54. Flujo condición. Información adaptada de Manychat. Elaborado por Ruth Apolo

Finalmente se crea dos flujos para esta condición, en el caso de que el usuario desee seguir con la búsqueda se presentara un mensaje que tendrá nuevamente la entrada de campo de usuario “entrada de usuario” para que vuelva almacenar la nueva consulta y vuelva a escribir la marca de otro producto que necesite consultar y un mensaje final en el caso de que ya no necesite seguir con la búsqueda de los artículos.

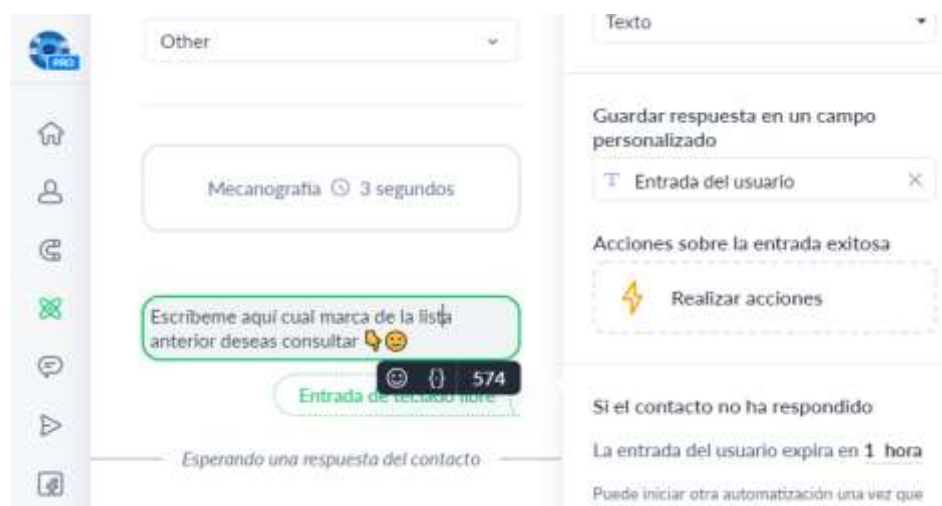


Ilustración 55. Siguiete consulta. Información adaptada de Manychat. Elaborado por Ruth Apolo



Ilustración 56. Mensaje final. Información adaptada de Manychat. Elaborado por Ruth Apolo

A continuación, se podrá visualizar como queda el flujo de búsqueda con el que el usuario podrá interactuar con el bot.

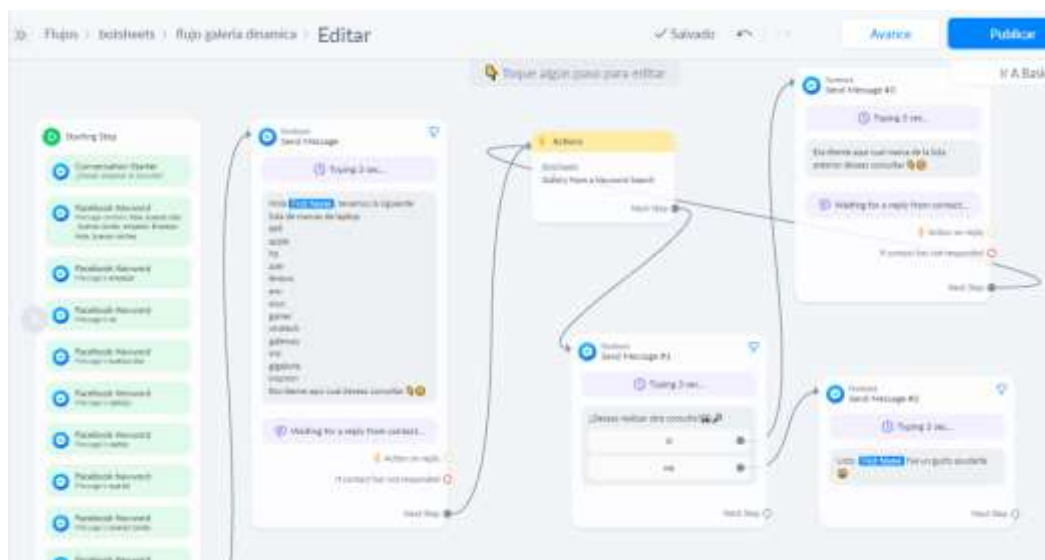


Ilustración 57. Flujo final. Información adaptada de Manychat. Elaborado por Ruth Apolo

En la misma carpeta “botsheets” se creó otro flujo adicional que está conectado con “galería dinámica” para cuando el usuario no escriba bien el nombre de un producto y no esté disponible en la base de datos, este flujo lo denominamos “para búsquedas que no entiende el bot”.



Ilustración 58. Flujo para búsquedas que no se encuentran. Información adaptada de Manychat. Elaborado por Ruth Apolo

Adicionalmente se creó una lista de palabras claves con la que se puede iniciar una conversación, o este caso, una búsqueda, son palabras con las que posiblemente una persona quiera empezar a interactuar con el bot.

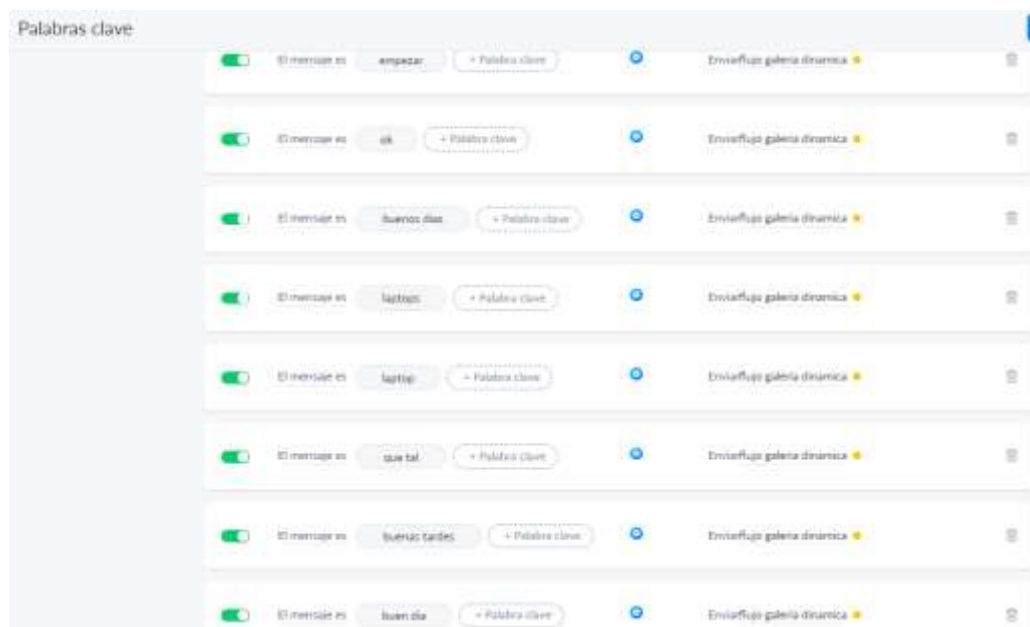


Ilustración 59. Palabras claves. Información adaptada de Manychat. Elaborado por Ruth Apolo

3.6 Pruebas del funcionamiento del prototipo

Para realizar el proceso de prueba del prototipo de nuestro asistente virtual utilizaremos una cuenta de usuario de Facebook para que interactúe con nuestro bot, en donde podemos asegurar que los resultados reflejos se acoplan a los objetivos propuesto ya que la información de búsqueda es rápida.

Se inicia la conversación con el bot como se visualiza en la ilustración 55 y este rápidamente te sugiere una lista detallada de marcas de laptops que tiene en su base de datos para que puedas realizar la consulta.



Ilustración 60. Saludo de inicio del chatbot. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

En la ilustración 61 podemos observar que al escribir la marca que deseamos consultar nuestro asistente virtual envía una galería de productos de la marca que el usuario digito y dichos productos están enlazados con las diferentes casas comerciales en las que decidamos realizar nuestra compra.



Ilustración 61. Consultas del chatbot. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

Para seguir con el proceso de prueba se le pidió seguir con la consulta de productos y para ello nos pidió volver a escribir otra marcar de laptop que necesitemos consultar para seguir con la cotización del producto, cabe mencionar que las respuestas se ejecutan de una manera muy rápida y precisa, a lo que podemos acotar que los objetivos de este proyecto se están cumpliendo claramente con esta prueba.

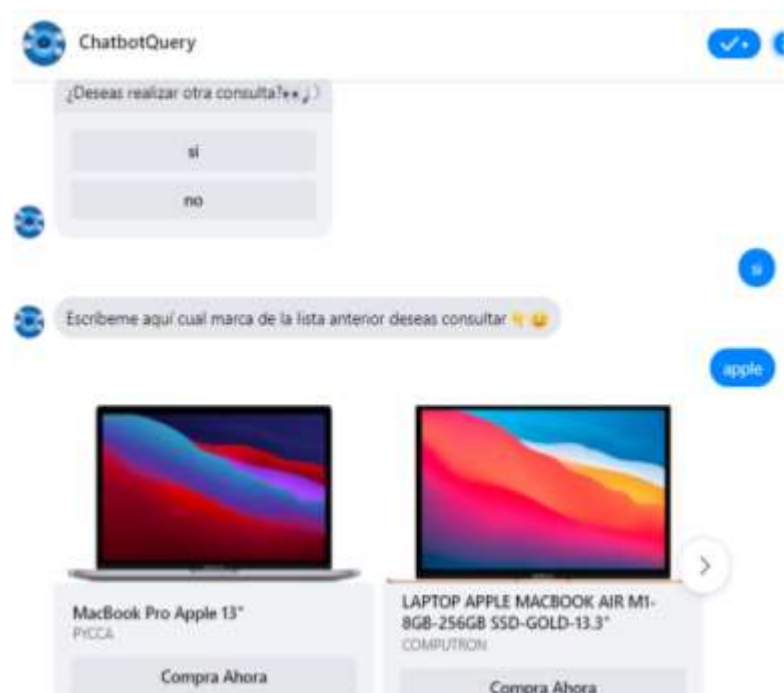


Ilustración 62. Segunda consulta. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo



Ilustración 63. Tiempo de consulta del chatbot. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

Se realizó otra consulta con el objetivo de experimentar el flujo cuando ingresamos un dato que no se encuentra en la lista de marcas que nos había dado al principio del chat, por lo que decidimos poner a prueba el asistente virtual ante una palabra que no registra en su base de datos.

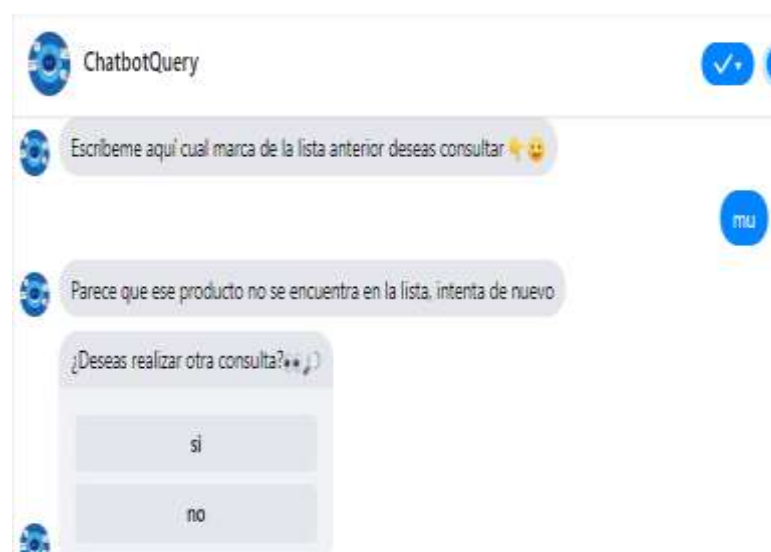


Ilustración 64. Dato no encontrado. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

Como pudimos observar en la ilustración 57, el asistente virtual al no encontrar en su registro la palabra digitada este responde que el producto no se encuentra y que intente probando nuevamente.

Decidimos realizar una última consulta para certificar que el asistente virtual puede buscar un artículo cuantas veces necesite el consumidor o el usuario para realizar su compra y este va a seguir ejecutándose e interactuando de una manera rápida y eficaz cuantas veces sea necesario hasta que se decida ya no seguir con la consulta.

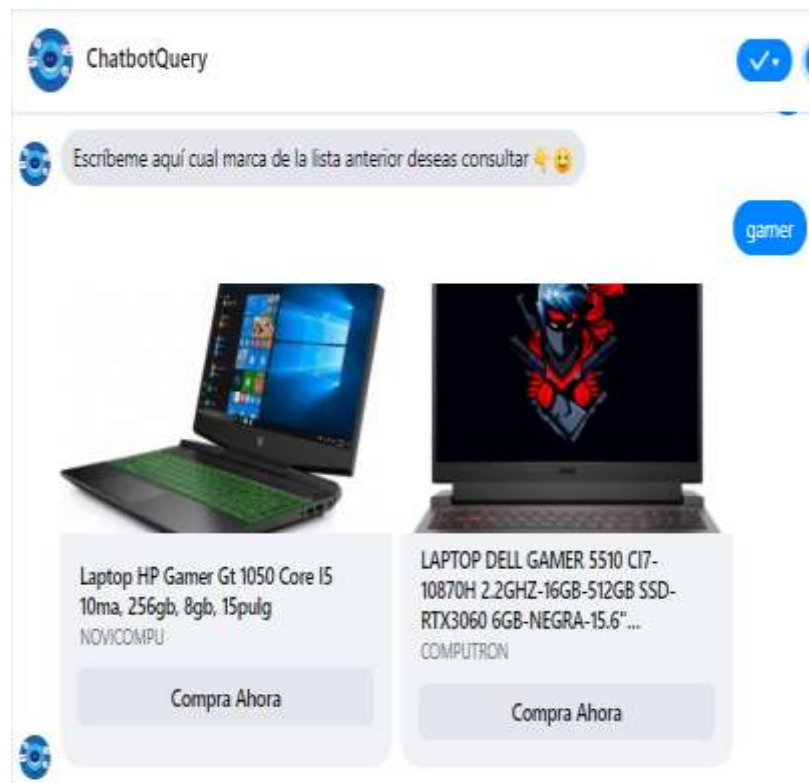


Ilustración 65. Tercera consulta. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

Finalmente, cuando ya el usuario decida que no necesita realizar más consultas y desee comprar el producto puede seleccionar el botón “comprar ahora” que lo llevara directamente a la página en donde encontrara el producto para proceso de compra desde el sitio de la casa comercial en donde se encuentra. Una vez terminada la consulta el flujo de chat termina.

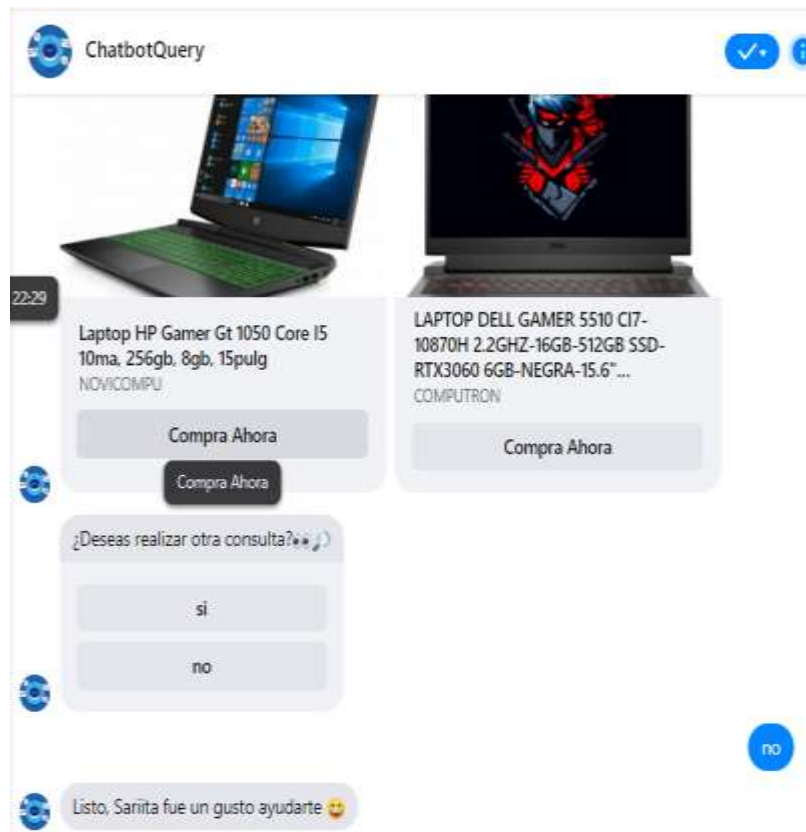


Ilustración 66. Fin del chat. Información adaptada de Facebook Messenger. Elaborado por Ruth Apolo

3.7 Conclusiones

- Se investigaron las casas comerciales que disponen del servicio online para la venta de sus productos ya que se han visto en la necesidad de acoplarse a vender en este tipo de medios por muchas circunstancias que acontecen en la actualidad con el fin de no bajar sus ventas y se pudo determinar que no presentan errores en sus páginas.
- Se investigaron tecnologías como la técnica de extracción de datos que permitió desarrollar nuevos conocimientos y pudimos poner en práctica muchos de los conocimientos aprendidos durante el proceso de estudio de nuestra carrera, este tipo de técnicas ayudo a afrontar nuevos retos ya que son nuevas experiencias la cual implico el desarrollo de los conocimientos y la adopción de estas.
- Se implemento un prototipo de asistente virtual con el que logramos determinar que un bot puede llegar a ofrecer un servicio eficiente rápido y personalizados a los consumidores, logrando ofrecer información en cualquier instante. Se puede llegar a los clientes y brindar un mejor servicio con un chatbot sin necesidad de que seas un experto en programación. Para este modelo de asistente virtual se usó una herramienta fácil de usar con una interfaz muy agradable para el usuario,

adicional a esto para ejecutar este prototipo se implementó una técnica de web scraping que nos ayudó a extraer la información de datos de páginas web que conectamos a nuestro bot con la ayuda de una plantilla de hoja de cálculo que usamos como nuestra base de datos que conectamos a un sistema de mensajería de Facebook Messenger para las entradas de consultas de las búsquedas de los usuarios. El prototipo demostró ser una opción rápida de cotización de productos, sencillo de usar, alcanzando los resultados propuestos y deseados para este trabajo.

- Se comprobó que el asistente virtual consta de un buen nivel de satisfacción por parte de los usuarios ya que estadísticamente se pudo evidenciar a través de las encuestas.

3.8 Recomendaciones

- En base a los resultados que se obtuvo del servicio de consulta de artículos se recomienda que para el desarrollo de futuros trabajos se implemente este tipo de herramientas en otras plataformas que también son muy usadas por los consumidores, como WhatsApp o Instagram.
- Se recomienda la implementación de reconocimientos de voz para el asistente virtual para poder mejorar el servicio al consumidor.
- Se recomienda buscar páginas web con una estructura HTML que no sea compleja y que se facilite su proceso de extracción.
- Se recomienda el uso de la herramienta Manychat ya que por un costo muy bajo pueden tener a tu alcance muchas opciones de creación para tu bot que ayudan a mejorar la experiencia del consumidor, sin necesidad de que seas un programador, ya que esta herramienta no tiene limitaciones para ningún usuario que desee mejorar sus servicios con sus clientes.

Glosario de palabras

CSV: Es un archivo de texto que almacena los datos en forma de columnas, separadas por coma y las filas se distinguen por saltos de línea.

Framework: Es una estructura previa que se puede aprovechar para simplificar la elaboración de un trabajo

Feature Driven Developmen: Es una metodología ágil para el desarrollo de sistemas

Natural Language Understanding: Es un servicio de IBM Cloud que permite analizar textos para extraer metadatos del contenido.

Scrum: Es un proceso en el que se aplican de manera regular un conjunto de buenas prácticas para trabajar colaborativamente.

API: Conjunto de definiciones y protocolos que se utiliza para desarrollar e integrar el software de las aplicaciones.

Extreme Programming: Se basa en realimentación continua entre el cliente y el equipo de desarrollo.

Watson Assitant: Es un servicio en IBM Cloud para ayudarle a crear y desplegar asistentes virtuales.

Xquery: Es un lenguaje de consulta diseñado para colecciones de datos XML.

XML: Es un lenguaje de marcado que define un conjunto de reglas para la codificación de documentos.

ANEXOS

Anexo 1

LEY ORGÁNICA DE DEFENSA DEL CONSUMIDOR

(Ley No. 2000-21)

Fuente: Ediciones Legales

EL CONGRESO NACIONAL

Considerando:

Que la generalidad de ciudadanos ecuatorianos son victimas permanentes de todo tipo de abusos por parte de empresas públicas y privadas de las que son usuarios y consumidores;

Que de la conformidad con lo dispuesto por el numeral 7 del artículo 23 (52) de la Constitución Política de la República, es deber del Estado garantizar el derecho a disponer de bienes y servicios públicos y privados, de óptima calidad; a elegirlos con libertad, así como a recibir información adecuada y veraz sobre su contenido y características;

LEY ORGÁNICA DE DEFENSA DEL CONSUMIDOR

Capítulo I

PRINCIPIOS GENERALES

Art. 2.- Definiciones. – Para efectos de la presente Ley, se entenderá por:

Consumidor. – Toda persona natural o jurídica que, como destinatario final, adquiera, utilice o disfrute bienes o servicios, o bien reciba oferta para ello. Cuando la presente Ley mencione al consumidor, dicha denominación incluirá al usuario.

Anexo 2
CONSTITUCIÓN DE LA REPUBLICA DEL ECUADOR

Título I

ELEMENTOS CONSTITUTIVOS DEL ESTADO

Sección IX

PERSONAS USUARIAS Y CONSUMIDORAS

Art. 52. – Las personas tienen derecho a disponer de bienes y servicios de óptima calidad y elegirlos con libertad, así como a una información precisa y no engañosa sobre su contenido y características.

La ley establecerá los mecanismos de control de calidad y los procedimientos de defensa de las consumidoras y consumidores; y las sanciones por vulneración de estos derechos, la reparación e indemnización por deficiencias, daños o mala calidad de bienes y servicios, y por la interrupción de los servicios públicos que no fuera ocasionada por caso fortuito o fuerza mayor.

Art. 53.- Las empresas, instituciones y organismos que presten servicios públicos deberán incorporar sistemas de medición de satisfacción de las personas usuarias y consumidoras, y poner en práctica sistemas de atención y reparación.

El Estado responderá civilmente por los daños y perjuicios causados a las personas por negligencia y descuido en la atención de los servicios públicos que estén a su cargo, y por la carencia de servicios que hayan sido pagados.

Art. 54.- Las personas o entidades que presten servicios públicos o que produzcan o comercialicen bienes de consumo, serán responsables civil y penalmente por la deficiente prestación del servicio, por la calidad defectuosa del producto, o cuando sus condiciones no estén de acuerdo con la publicidad efectuada o con la descripción que incorpore.

Las personas serán responsables por la mala práctica en el ejercicio de su profesión, arte u oficio, en especial aquella que ponga en riesgo la integridad o la vida de las personas.

Capítulo VI

DERECHOS DE LIBERTAD

Art. – Se reconoce y garantiza a las personas:

19. El derecho a la protección de datos de carácter personal, que incluye el acceso y la decisión sobre información y datos de este carácter, así como su correspondiente protección. La recolección, archivo, procesamiento, distribución o difusión de estos datos o información requerirán la autorización del titular o el mandato de la ley.

Anexo 3



MEMORANDO No. PAN-CLC-2019- 0184

DE: **CÉSAR LITARDO CAICEDO**
Presidente de la Asamblea Nacional

PARA: **JOHN DE MORA MONCAYO**
Prosecretario General Temporal

ASUNTO: Difundir Proyecto

FECHA: Quito D.M., 19 SEP 2019

Según lo dispuesto en el Art. 55 de la Ley Orgánica de la Función Legislativa, envío el **“PROYECTO DE LEY - ORGÁNICA DE PROTECCIÓN DE DATOS PERSONALES”**, remitido por el Presidente Constitucional de la República, Lenin Moreno Garcés, a través del oficio No. T.514-SGJ-19-0740 de 19 de septiembre de 2019, con número de trámite 379637, a fin de que sea difundido a las/los asambleístas y a la ciudadanía a través del portal Web y se remita al Consejo de Administración Legislativa (CAL), para el trámite correspondiente.

Atentamente,


CÉSAR LITARDO CAICEDO

Presidente de la Asamblea Nacional

tr. 379637
jda.

LEY ORGÁNICA DE PROTECCIÓN DE DATOS PERSONALES

CAPÍTULO I DISPOSICIONES DIRECTIVAS

Artículo 1. Objeto: El objeto de la presente Ley es regular el ejercicio del derecho a la protección de datos personales, la autodeterminación informativa y demás derechos digitales en el tratamiento y flujo de datos personales, a través del desarrollo de principios, derechos, obligaciones y mecanismos de tutela.

Artículo 2. Finalidad: La finalidad de la presente Ley es procurar el adecuado tratamiento y flujo de datos personales para garantizar los derechos fundamentales y las libertades individuales, promover el progreso económico y social; impulsar la producción nacional y la cooperación internacional; fomentar la competitividad, la innovación y productividad, elevar la eficiencia de los servicios públicos y/o privados; y, mejorar la calidad de vida.

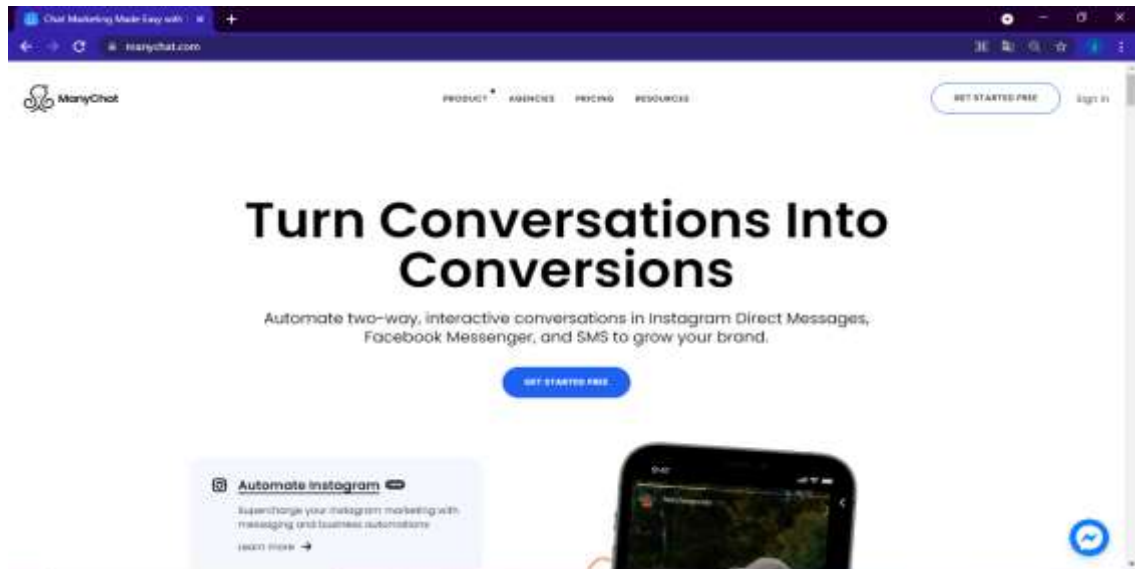
Artículo 3. Ámbito de aplicación material: La presente Ley se aplicará al tratamiento de datos personales contenidos en cualquier tipo de soporte, ya sean totalmente automatizados, parcialmente automatizados o no automatizados y a toda modalidad de uso posterior, por parte de responsables o encargados del tratamiento de datos personales.

Anexo 4

Manual:

Crear una cuenta de ManyChat

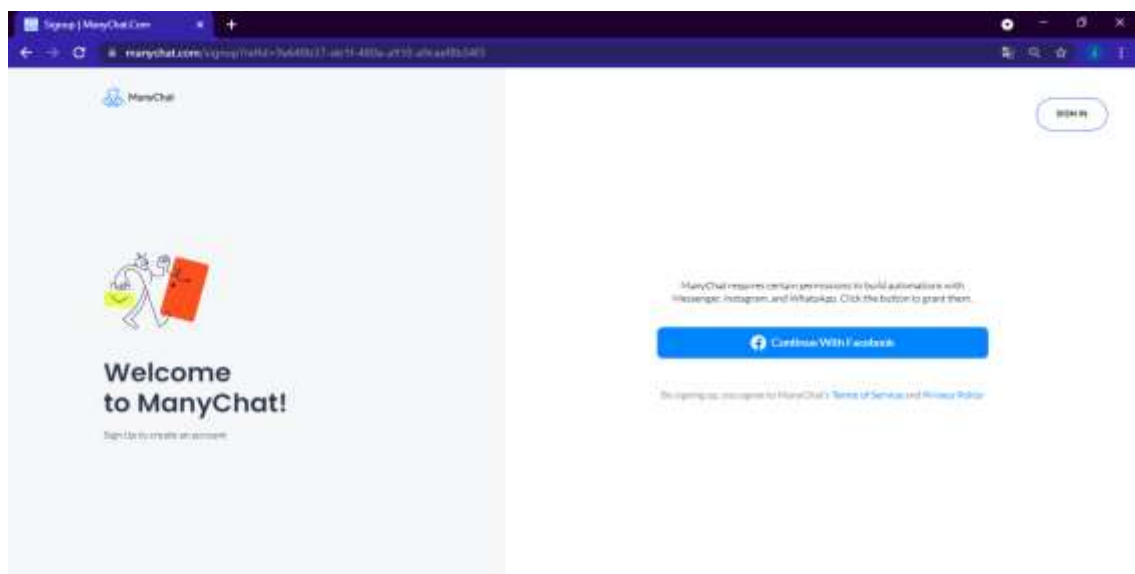
1.- Ingresar al siguiente link: <https://manychat.com/>



Fuente: <https://manychat.com/>

2.- Dirigirse a “GET STARTED FREE”

3.- Hacer clic en “Continue With Facebook”



Fuente: <https://manychat.com/>

4.- Contestar preguntas

Would you share a little bit about yourself?

Sharing info about your business will help us enhance ManyChat experience for you.

What best describes your role? *

How many people work at your company? *

Our website uses cookies to make sure you have the best, most relevant experience. By clicking Accept below, you're agreeing to our cookie use. We

Accept

Fuente: <https://manychat.com/>

5. – Conectar con la página de Facebook de la cuenta

manychat.setupFacebookAccount


ManyChat

☐ Connect Facebook Profile ☐ Describe the Company ☒ Setup Your Account ☐ Choose Marketing Goals ☐ Done

Setup Your Account

We found 1 Facebook pages linked with your profile:

Select a page you want to start working with ManyChat.
If you don't see all your pages, read this.

 ChatbotQuery

You will need to refresh this page after creating a new Facebook Business Page.

Fuente: <https://manychat.com/>

6.- Contestar otro formulario de preguntas

ManyChat

☐ Authorize ☐ Describe yourself ☐ Create account ☒ Set Marketing Goals

Let's talk about your account:

Sharing info about your bot will help us enhance ManyChat experience for you.

Will this account be used for an actual business? *

☐ Yes, this account will be used for an actual business

☐ No, this is a technical/testing/demo account

What type of messages are you going to send from ManyChat? *

☐ SMS

☐ Email

☐ Facebook Messenger

☐ Instagram (coming soon)

☐ WhatsApp (coming soon)

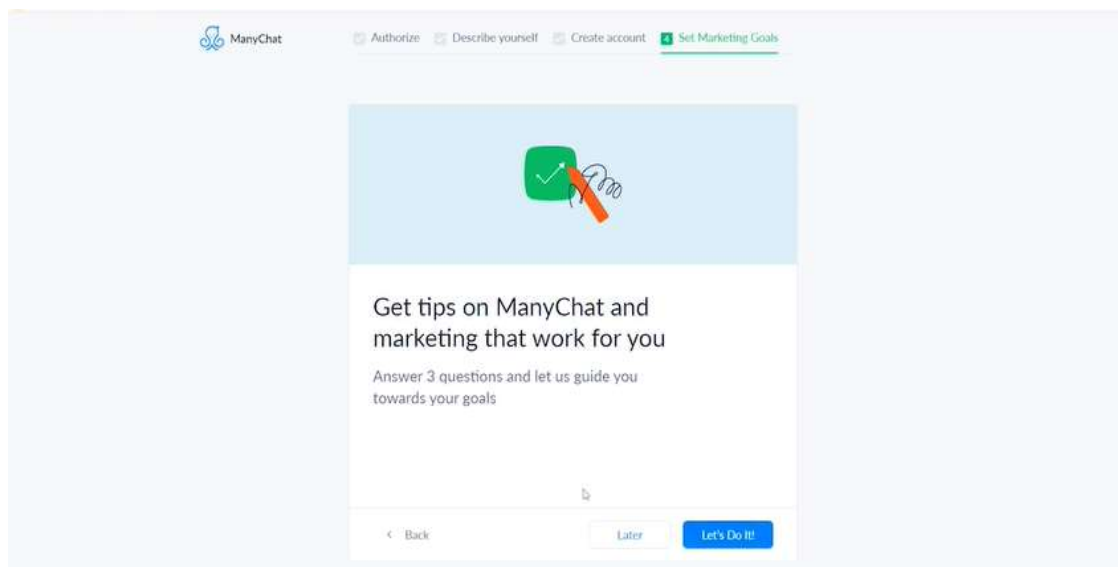
☐ I don't know yet

Our website uses cookies to make sure you have the best, most relevant experience. By clicking Accept below, you're agreeing to our cookie use. We

Accept

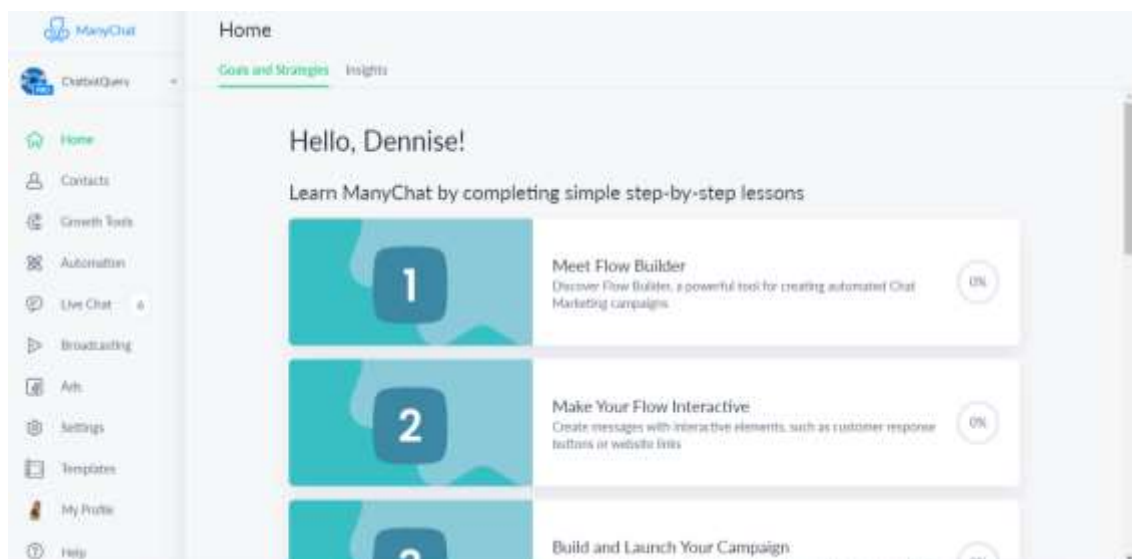
Fuente: <https://manychat.com/>

7. – Hacer clic en “Later”



Fuente: <https://manychat.com/>

8.- La cuenta de ManyChat ya está lista para crear el Bot de la página de Facebook



Fuente: <https://manychat.com/>

Anexo 5

Manual:

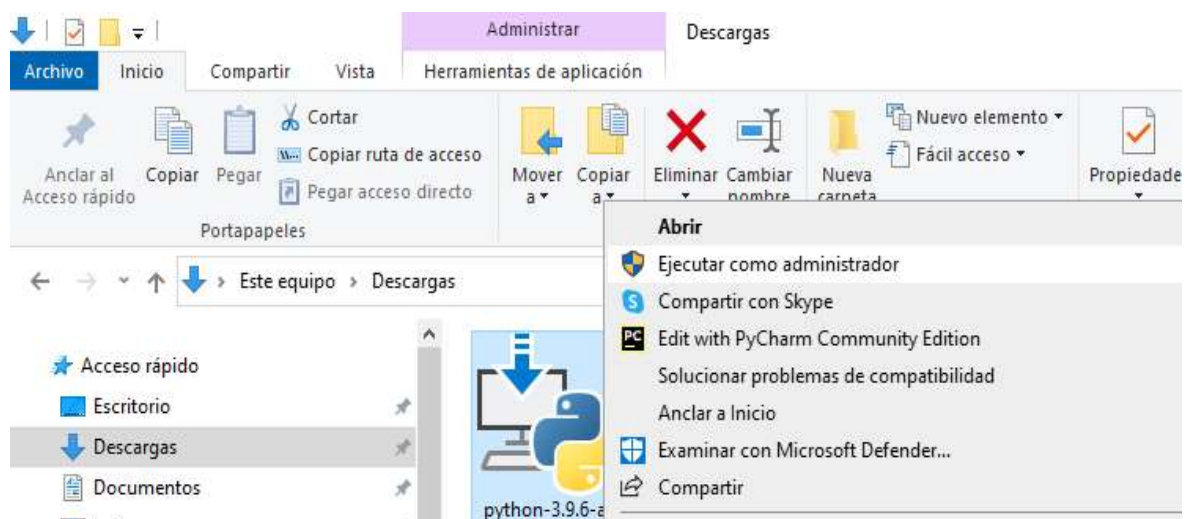
Pasos para instalar Python

- 1.- Ingresar al siguiente link: <https://www.python.org/downloads/>
- 2.- Hacer clic en “Download Python”



Fuente: <https://www.python.org/downloads/>

- 3.- Una vez se descargue el archivo, damos clic derecho y clic en “Ejecutar como administrador”



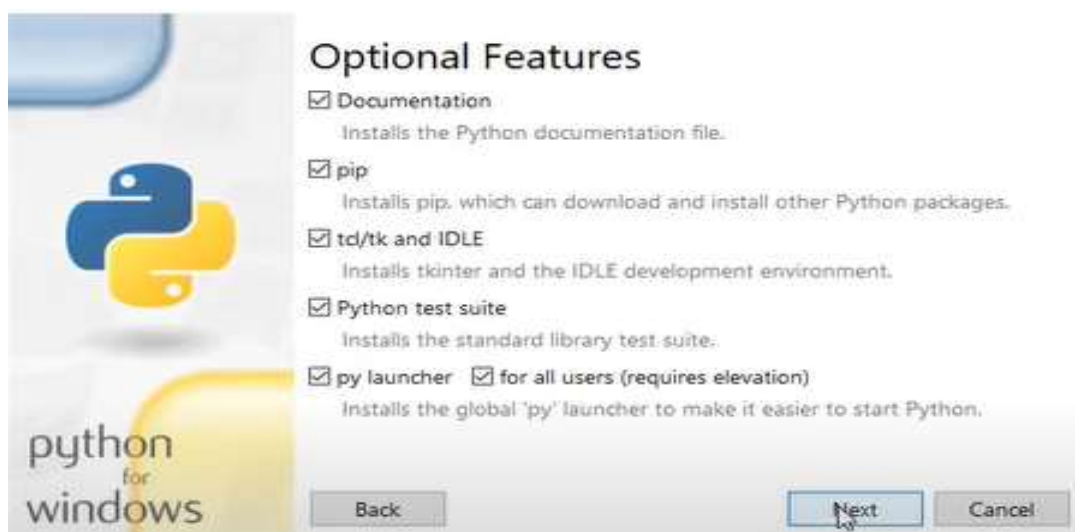
Elaborado por: Ruth Apolo

- 4.- Marcar la casilla “Add Python 3.8 to PATH” y hacer clic en “Customize installation”



Elaborado por: Ruth Apolo

5.- Clic en “Next”



Elaborado por: Ruth Apolo

6.- Poner la ruta de instalación y clic en “Install”



Elaborado por: Ruth Apolo

7.- Una vez termine la instalación damos clic en “close”



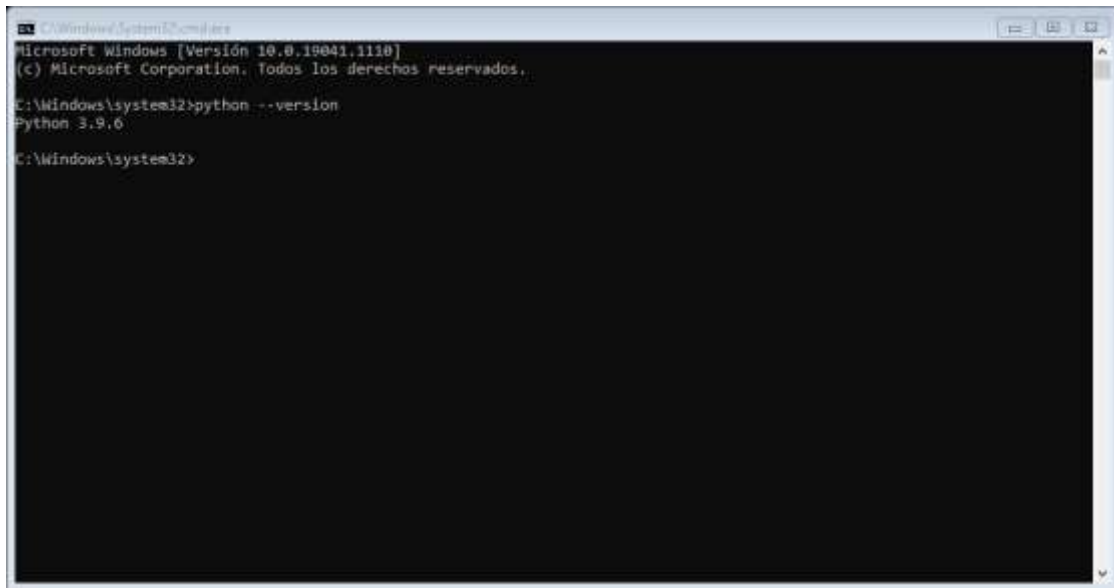
Elaborado por: Ruth Apolo

8.- Para verificar la correcta instalación, abrimos el CMD



Elaborado por: Ruth Apolo

9.- Digitamos “Python --version”



```
C:\Windows\System32\cmd.exe
Microsoft Windows [Versión 10.0.19041.1110]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Windows\system32>python --version
Python 3.9.6

C:\Windows\system32>
```

Elaborado por: Ruth Apolo

Anexo 6

Manual:

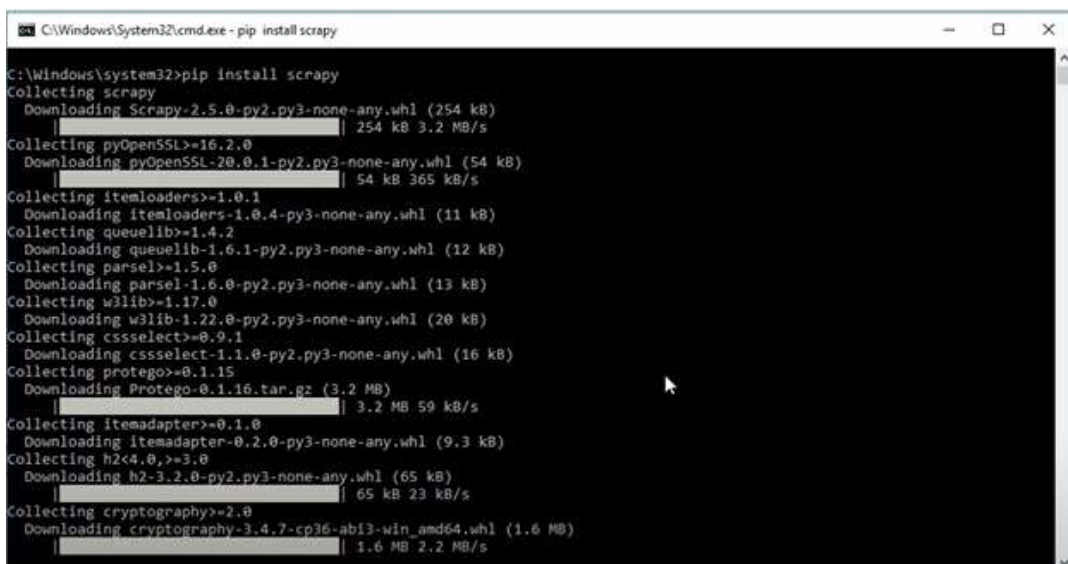
Pasos para instalar Scrapy

1.- Abrimos el CMD



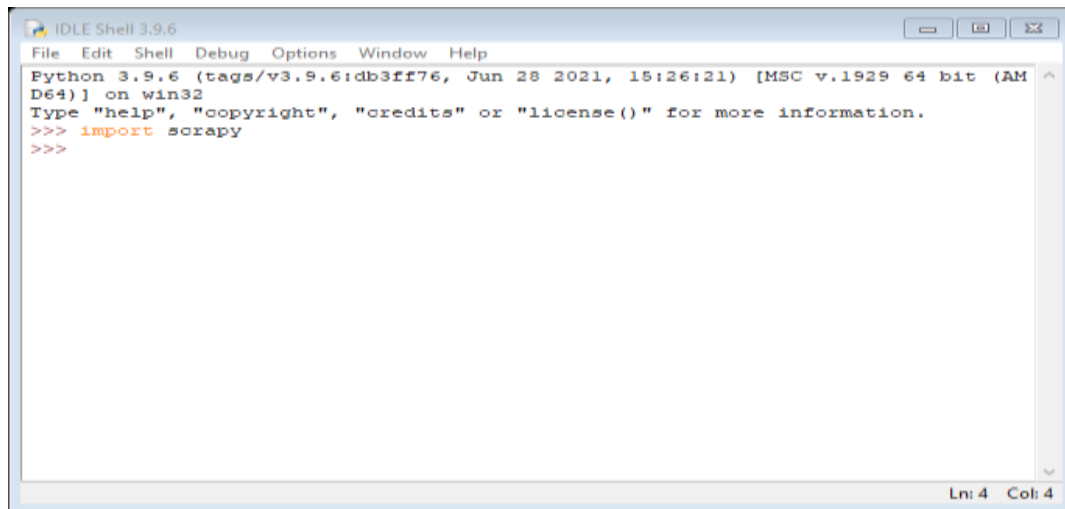
Elaborado por: Ruth Apolo

2.- Digitamos “pip install scrapy”



Elaborado por: Ruth Apolo

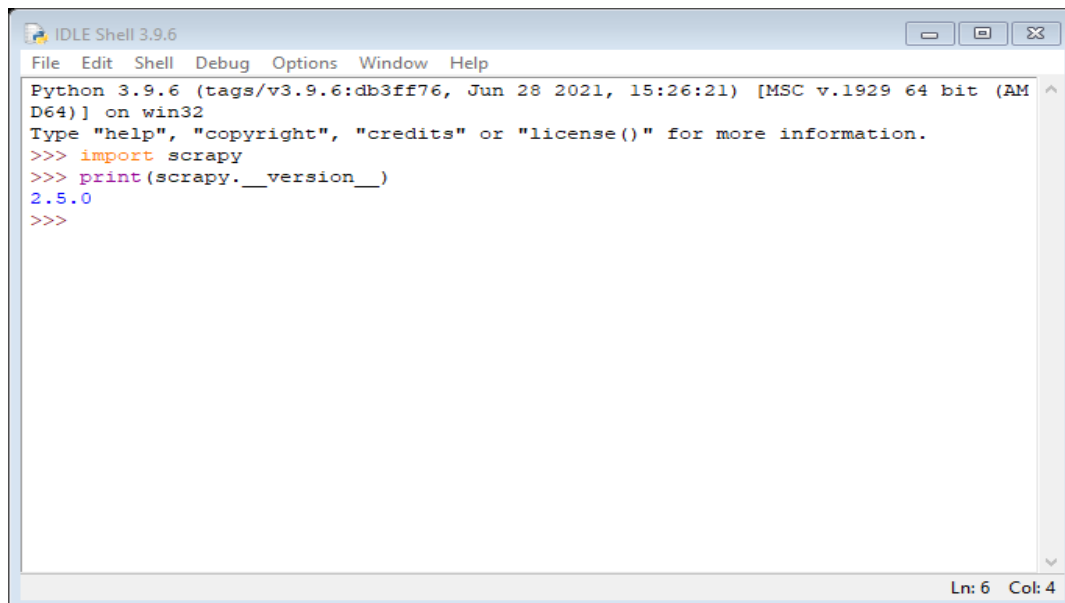
3.- Ahora abrimos el IDLE de Python e importamos scrapy



```
Python 3.9.6 (tags/v3.9.6:db3ff76, Jun 28 2021, 15:26:21) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>> import scrapy
>>>
```

Elaborado por: Ruth Apolo

4.- Para verificar la instalación y la versión digitamos “print (scrapy.__version__)”



```
Python 3.9.6 (tags/v3.9.6:db3ff76, Jun 28 2021, 15:26:21) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>> import scrapy
>>> print(scrapy.__version__)
2.5.0
>>>
```

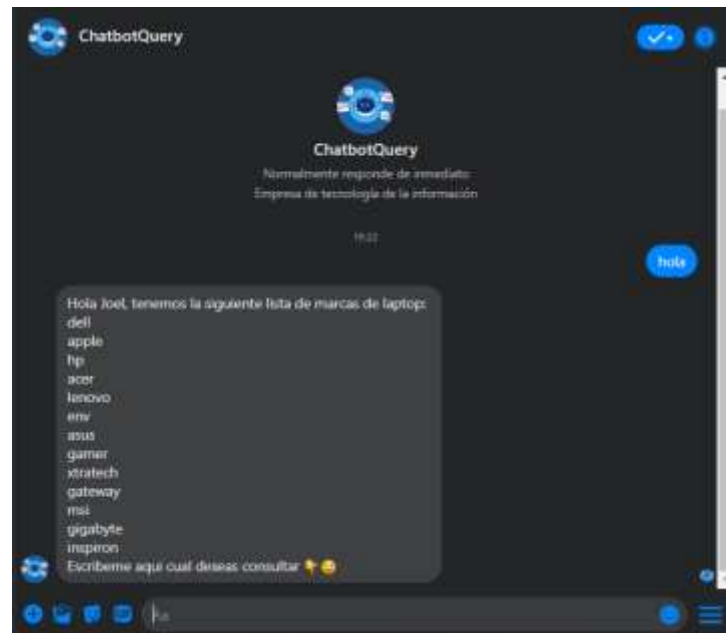
Elaborado por: Ruth Apolo

Anexo 7

Manual:

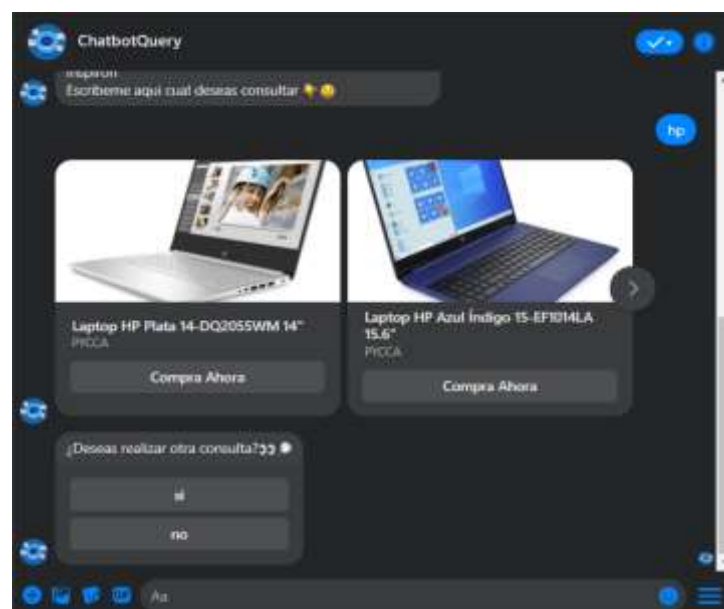
Pasos para usa la herramienta de asistente virtual

1.- Empezamos un saludo inicial con el bot “hola” y nos mostrara una lista de marcas de productos que tiene disponibles para la consulta.



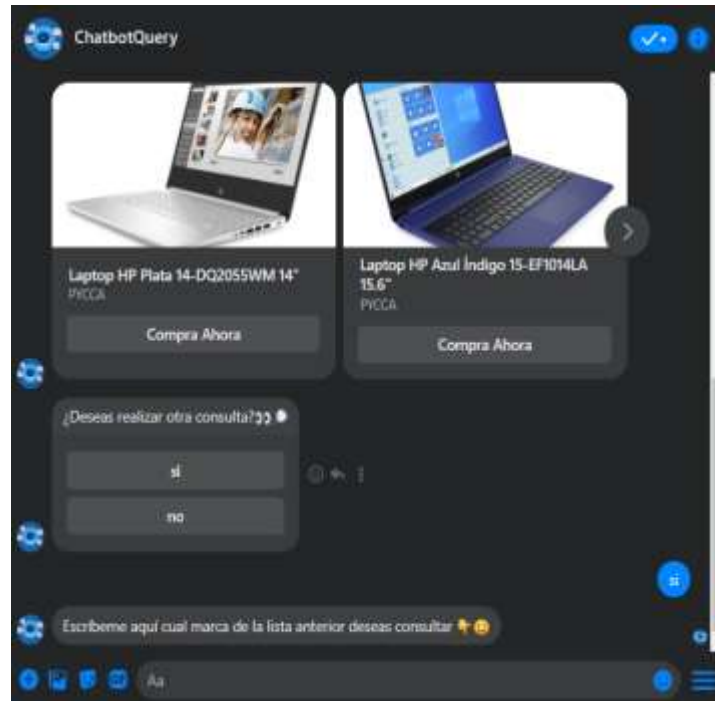
Fuente: <https://www.facebook.com/messages/t/105082398578142>

2.- Escribimos el nombre del productor de deseamos consultar y nos mostrara un carrusel de imágenes de la marca seleccionada.



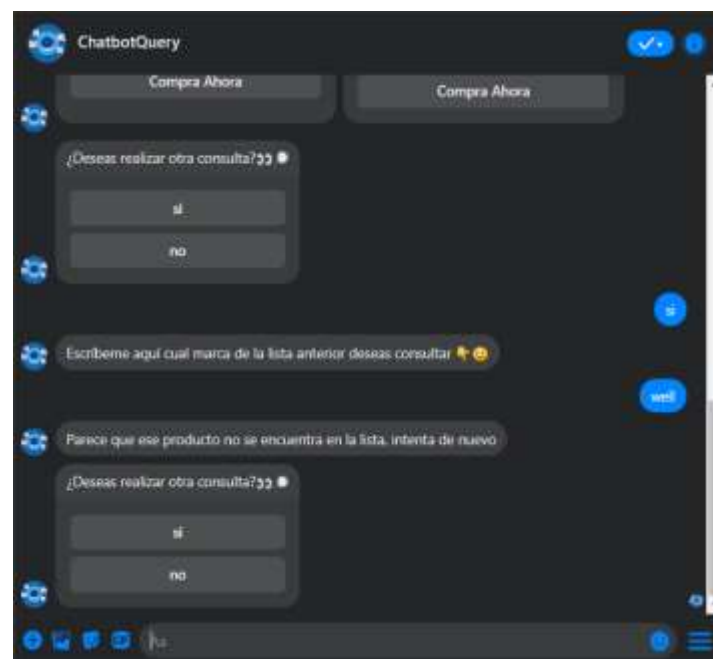
Fuente: <https://www.facebook.com/messages/t/105082398578142>

3.- Si deseamos seguir con la consulta podemos seleccionar “sí” y nos pedirá escribir de nuevo el nombre de la marca que deseamos consultar.



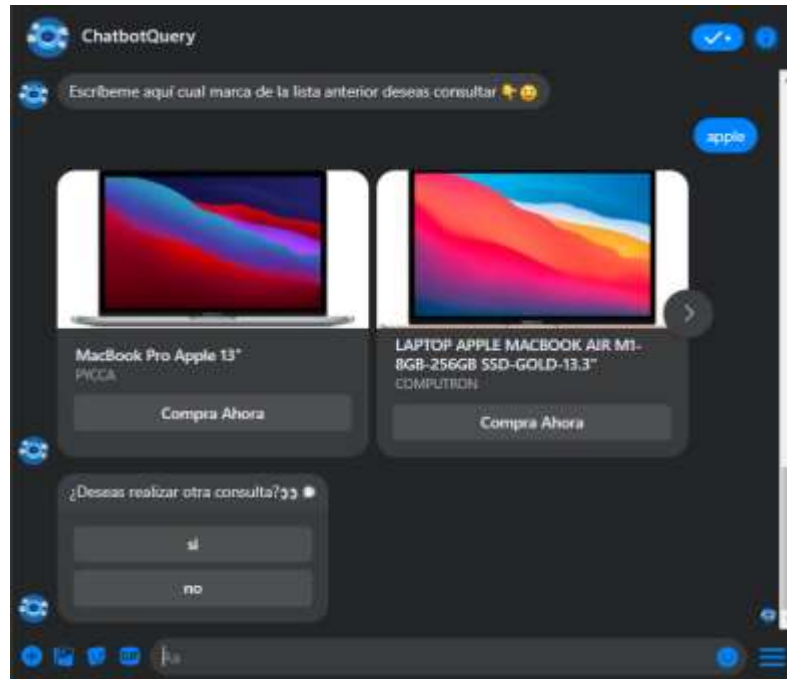
Fuente: <https://www.facebook.com/messages/t/105082398578142>

4.- En el caso de que digitemos mal o digitemos algo que no se localice en la lista nos enviara un mensaje de que no se encuentra en la lista, intenta de nuevo.



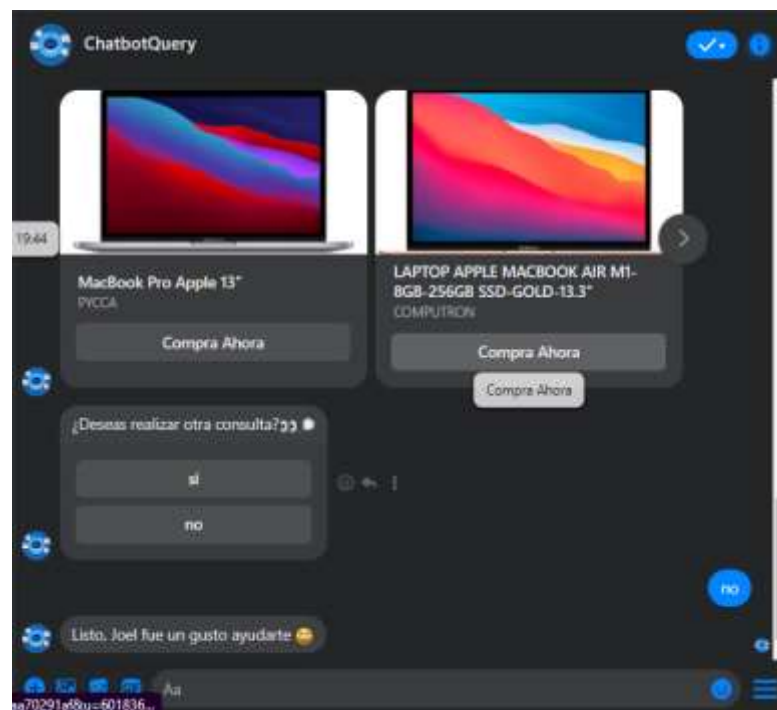
Fuente: <https://www.facebook.com/messages/t/105082398578142>

5.- Si deseamos seguir con la consulta, seleccionamos “si” y escribimos el nombre de la marca del producto deseado.



Fuente: <https://www.facebook.com/messages/t/105082398578142>

6.- Si encontramos el producto deseado, seleccionamos “no” y a su vez seleccionamos “comprar ahora” en el producto que deseamos adquirir y nos enlazara a la pagina web donde se encuentra disponible el producto.



Fuente: <https://www.facebook.com/messages/t/105082398578142>

Anexo 8**La encuesta**

UNIVERSIDAD DE GUAYAQUIL						
CARRERA INGENIERÍA EN TELEINFORMÁTICA						
“PROTOTIPO DE ASISTENTE VIRTUAL TIPO CHATBOT UTILIZANDO WEBSCRAPING”						
1.- ¿La herramienta le proporcionó la información de manera rápida?						
<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Tal vez						
2.- ¿Considera usted que el flujo de búsqueda es claro y entendible?						
<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Tal vez						
3.- ¿Tuvo usted algún inconveniente al momento de realizar la búsqueda del producto?						
<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Tal vez						
4.- ¿La herramienta le ayudó a reducir el tiempo de cotización del product?						
<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Tal vez						
5.- ¿Qué tan recomendable considera usted el uso de este servicio de asistente virtual?						
	1	2	3	4	5	
No recomendable	O	O	O	O	O	Muy recomendable
6.- ¿Considerarías usar un Asistente virtual para sus búsquedas futuras?						
<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Tal vez						

5.- Coloque del 1 al 5 que tanto le agrado la herramienta					
	1	2	3	4	5
Nada					
Mucho	O	O	O	O	O

Bibliografía

- Blanco, M. S. (17 de Septiembre de 2020). *Estrategias digitales utilizadas por 7 empresas durante el covid-19*. Obtenido de <file:///C:/Users/USUARIO/Downloads/10513-Texto%20del%20art%C3%ADculo-39110-1-10-20201219.pdf>
- C, B. (2006). *Metodología de la Investigación*.
- CECE. (2020). Obtenido de Situación de las empresas durante el COVID-19 Ecuador: <https://cece.ec/wp-content/uploads/2020/06/Situacion-de-las-empresas-durante-el-Covid19-en-Ecuador.pdf>
- Chesñevar, E. (2018). *El comercio electronico en la era de los bots*. Obtenido de https://ri.conicet.gov.ar/bitstream/handle/11336/94908/CONICET_Digital_Nro.3de882eb-3f30-4ac2-92db-e9e1a95d66e0_A.pdf?sequence=2&isAllowed=y
- Erl, T., Khattak, W., & Buhler, P. (2016). *Big Data Fundamentals ; Concepts, Drives and Techniques*. New Jersey, United States: Prentice Hall.
- GALIPIENSO, A., ISABEL, M., QUEVEDO, C., ANGEL, M., PARDO, C., OTTO, . . . ANGEL, M. (2003). *Inteligencia artificial: modelos, técnicas y áreas de aplicación*. paraninfo.
- García, C. G. (2018). *En qué consiste el aprendizaje automático (machine learning) y qué está aportando a la*. Obtenido de <http://www.cienciacognitiva.org/files/2017-20.pdf>
- Herrera, A. H. (Junio de 2020). *Herramientas Web Scraping para la Adquisición de Datos Turísticos*. Obtenido de <http://kikame.tecnocientifica.com.mx/index.php/kikame/article/view/26/27>
- Herrero, C. H. (s.f.). *Aplicación de Técnicas de Web*. Obtenido de <https://uvadoc.uva.es/bitstream/handle/10324/5794/TFG-B.512.pdf?sequence=1>
- Loboa, J. E. (2019). *Chatbots para apoyar la detección de síntomas de trastornos psiquiátricos*. Obtenido de <https://repository.usc.edu.co/bitstream/handle/20.500.12421/4542/Chatbots%20para%20apoyar%20la%20detecci%C3%B3n%20de%20s%C3%ADntomas%20de%20trastornos%20psiqui%C3%A1tricos.pdf?sequence=1&isAllowed=y>

- Lopez, J. (12 de Enero de 2018). *Web scraping*. Obtenido de <https://d1wqtxts1xzle7.cloudfront.net/55775125/web-scraping-with-cover-page-v2.pdf?Expires=1631765685&Signature=GSAVDhGVFzOkhv~QmPqswrBCT3NySduVucoCnic~z-A24NC6Y9jZwJ5KklV~epVagL5OBw9aRHcTr~Xws0FurIm~-0-3yQNjWh0m0WtoZq0Dbv~VVJegW8Bra-3j9WhPkkxKpsyImffBVq3>
- Mathivet, V. (2018). *Inteligencia artificial para desarrolladores: conceptos e implementación en C#*. ENI.
- REINA, L. F. (s.f.). *ASISTENTE VIRTUAL TIPO CHATBOT*. Obtenido de https://repository.ucatolica.edu.co/bitstream/10983/17726/1/ASISTENTE%20VIRTUAL%20TIPO%20CHATBOT_final.pdf
- Rivilla, C. (Noviembre de 2020). *El Black Friday y el Cyber Monday en tiempos de COVID-19*. Obtenido de https://marketing.onlinebschool.es/Prensa/Informes/Informe%20OBS_El%20Black%20Friday%20y%20el%20Cyber%20Monday%20en%20tiempos%20de%20COVID-19.pdf
- RODRÍGUEZ, ORTIZ, QUIROZ, & PARRALES. (26 de octubre de 2020). *El e-commerce y las Mipymes en tiempos de Covid-19*. Obtenido de <http://w.revistaespacios.com/a20v41n42/a20v41n42p09.pdf>
- Rouhiainen, L. (2018). *Inteligencia artificial*. alienta.