



**“Implementación de un sistema de recomendación de habitabilidad en la ciudad de Quito orientado a la planificación urbana cognitiva aplicando técnicas de machine learning”**

Rojas Aguirre, Ximena Aracely

Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología

Centro de Posgrados

Maestría en Gestión de Sistemas de Información e Inteligencia de Negocios

Trabajo de titulación, previo a la obtención del título de Magíster en Gestión de Sistemas de Información e Inteligencia de Negocios

Ing. Recalde Cerda, Lorena Katherine, PhD

25 de mayo del 2022



TRABAJO DE TITULACION ESPE\_XIMENA ROJAS AGUIRRE\_MA...

Scanned on: 16:19 May 10, 2022 UTC



Overall Similarity Score



Results Found



Total Words in Text

Identical Words	498
Words with Minor Changes	237
Paraphrased Words	320
Omitted Words	0

LORENA  
KATHERINE  
RECALDE  
CERDA

Firmado  
digitalmente por  
LORENA KATHERINE  
RECALDE CERDA  
Fecha: 2022.05.16  
19:41:49 -05'00'



**Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología**

**Centro de Posgrados**

### **Certificación**

Certifico que el trabajo de titulación: **“Implementación de un sistema de recomendación de habitabilidad en la ciudad de Quito orientado a la planificación urbana cognitiva aplicando técnicas de machine learning”** fue realizado por la señorita **Rojas Aguirre, Ximena Aracely**; el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizado en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

**Sangolquí, 25 de mayo de 2022**

Firma:

.....  
**Ing. Recalde Cerda, Lorena Katherine, PhD**

**Director**

**C.C.: 1720035789**



**Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología**

**Centro de Posgrados**

**Responsabilidad de Autoría**

Yo **Rojas Aguirre, Ximena Aracely**, con cédula de ciudadanía n° 1717599185, declaro que el contenido, ideas y criterios del trabajo de titulación: **Implementación de un sistema de recomendación de habitabilidad en la ciudad de Quito orientado a la planificación urbana cognitiva aplicando técnicas de machine learning** es de mí autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

**Sangolquí, 25 de mayo de 2022**

Firma

.....  
**Ing. Rojas Aguirre, Ximena Aracely**

**C.C.: 1717599185**



**Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología**

**Centro de Posgrados**

**Autorización de Publicación**

Yo **Rojas Aguirre, Ximena Arcely**, con cédula de ciudadanía n° 1717599185, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Implementación de un sistema de recomendación de habitabilidad en la ciudad de Quito orientado a la planificación urbana cognitiva aplicando técnicas de machine learning** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

**Sangolquí, 25 de mayo de 2022**

Firma

.....  
**Ing. Rojas Aguirre, Ximena Arcely**

**C.C.: 1717599185**

### **Dedicatoria**

El resultado de este proyecto va dedicado con amor a mi mamita María Ruano que con su apoyo incondicional ha sabido estar conmigo siempre brindándome toda la fortaleza para continuar, ahora sé que me verá desde el cielo.

A mis padres Ángel y Ximena que siempre han brindado su apoyo incondicional.

## **Agradecimiento**

Agradezco a Dios, por darme cada día la salud y sabiduría para poder desempeñar las actividades diarias con responsabilidad.

A mi ángel, la abuelita María y padres muchas gracias por sus consejos y por hacer de mí una excelente persona.

A la Universidad de las Fuerzas Armadas – ESPE, en especial a los coordinadores: Tatiana Gualotuña y Diego Marcillo.

A mi tutora Lorena Recalde, por haberme guiado correctamente en la realización del presente proyecto de tesis y estar preocupada cada día, cada segundo de los avances.

Al Doctor, Walter Fuertes quien fue siempre mi guía en cada proceso y no sólo educativo, sino como persona.

A varias personas de mi pueblito Archidona, que viven en mi corazón, sin importar en donde estén quiero darles las gracias por formar parte de mi vida, de mis sueños y anhelos, por todo lo que me han brindado y por todas sus bendiciones y buenos deseos para mí.

## Índice de Contenidos

<b>Carátula</b> .....	1
<b>Certificación</b> .....	3
<b>Responsabilidad de autoría</b> .....	4
<b>Autorización de publicación</b> .....	5
<b>Dedicatoria</b> .....	6
<b>Agradecimiento</b> .....	7
<b>Índice de Contenidos</b> .....	8
<b>Índice de Tablas</b> .....	10
<b>Índice de Figuras</b> .....	11
<b>Resumen</b> .....	12
<b>Abstract</b> .....	13
<b>Glosario de Nomenclaturas</b> .....	14
<b>Capítulo I. Introducción</b> .....	15
<b>Antecedentes</b> .....	15
<b>Justificación e Importancia</b> .....	16
<b>Planteamiento del Problema</b> .....	17
<b>Objetivo General</b> .....	17
<b>Objetivos Específicos</b> .....	18
<b>Hipótesis</b> .....	18
<b>Alcance</b> .....	18
<b>Metodología</b> .....	19
<b>Metodología de Inteligencia de Negocios</b> .....	19
<b>Metodología de Minería de Datos</b> .....	19
<b>Capítulo II: Estado del Arte</b> .....	23
<b>Preguntas de Investigación</b> .....	23
<b>Construcción de la Cadena de Búsqueda</b> .....	26
<b>Capítulo III: Marco Teórico</b> .....	31
<b>Fundamentación de la Variable Independiente</b> .....	31
<b>Inteligencia Cognitiva</b> .....	31
<b>Machine Learning</b> .....	32
<b>Técnicas de Aprendizaje:</b> .....	32



<b>Algoritmos de Predicción .....</b>	<b>33</b>
<b>Fundamentación de la Variable Dependiente.....</b>	<b>33</b>
<b>Características de Habitabilidad .....</b>	<b>33</b>
<b>Capítulo IV: Desarrollo .....</b>	<b>37</b>
<b>Aplicación de la Metodología.....</b>	<b>37</b>
<b>Obtención de los Datos .....</b>	<b>38</b>
<b>Arquitectura del Sistema .....</b>	<b>51</b>
<b>Diseño de la Base de Datos .....</b>	<b>53</b>
<b>Diagrama de Clases.....</b>	<b>55</b>
<b>Componentes del Sistema Machine Learning.....</b>	<b>56</b>
<b>Capítulo V: Conclusiones y Recomendaciones.....</b>	<b>58</b>
<b>Conclusiones .....</b>	<b>58</b>
<b>Recomendaciones .....</b>	<b>60</b>
<b>Anexo .....</b>	<b>61</b>
<b>Bibliografía.....</b>	<b>68</b>

## Índice de Tablas

<b>Tabla 1.</b> Desarrollo de objetivos específicos.....	20
<b>Tabla 2.</b> Preguntas de Investigación.....	24
<b>Tabla 3.</b> Estudios por Grupos de Control.....	25
<b>Tabla 4.</b> Construcción de la cadena de búsqueda. ....	27
<b>Tabla 5.</b> Descripción de las variables del SRH .....	40
<b>Tabla 6.</b> Caga de Datos al modelo con el 80% de entrenamiento.....	43
<b>Tabla 7.</b> Caga de Datos al modelo con el 70% de entrenamiento.....	44
<b>Tabla 8.</b> Resultados.....	46
<b>Tabla 9.</b> Resultados valor residual.....	47
<b>Tabla 10.</b> Evaluación MAE 20% y 30%.....	50
<b>Tabla 11.</b> Validación de los modelos .....	51
<b>Tabla 12.</b> Detalle de las tablas utilizadas.....	53

## Índice de Figuras

<b>Figura 1.</b> Población total en la ciudad de Quito .....	15
<b>Figura 2.</b> Variable dependiente e independiente .....	31
<b>Figura 3.</b> Proceso de Machine Learning KDD.....	38
<b>Figura 4.</b> Archivo de entrada de Lotes.....	39
<b>Figura 5.</b> Obtención de los datos.....	39
<b>Figura 6.</b> Proceso ETL de limpieza.....	41
<b>Figura 7.</b> Base de Datos MySQL.....	42
<b>Figura 8.</b> Validación Linear Regression MAE .....	48
<b>Figura 9.</b> Validación Logistic Regression MAE .....	49
<b>Figura 10.</b> Arquitectura del Sistema ML.....	52
<b>Figura 11.</b> Diseño Lógico Funcional sistema ML .....	52
<b>Figura 12.</b> Modelo Entidad – Relación.....	53
<b>Figura 13.</b> Diagrama de Componentes de SRH. ....	54
<b>Figura 14.</b> Diagrama de clases.....	55
<b>Figura 15.</b> Componentes del Sistema ML.....	56
<b>Figura 16.</b> Componentes y herramientas del SRH .....	57

## Resumen

La ciudad de Quito actualmente tiene una población aproximada de 2.735.987 habitantes, según datos del INEC (Instituto Nacional de Estadística y Censos), dónde el crecimiento desmesurado de la población genera el aumento del consumo de recursos, servicios básicos, movilidad y salud, existen necesidades que crecen considerablemente, necesidades como: seguridad, tranquilidad, servicios básicos, áreas verdes, entre otras características que son buscadas en un barrio. (Ecuador, 2019)

La sobrepoblación en los barrios debido al crecimiento elevado por cuestiones migratorias desencadena características a ser analizadas en el presente proyecto.

En la ciudad de Quito la problemática que enfrentamos es la habitabilidad en zonas de riesgo, el ciudadano elige vivir en un lugar sin conocer las características del sector. Ante esta problemática, este trabajo propone la aplicación de algoritmos que permitan la selección de las características de habitabilidad y así dar recomendaciones a los ciudadanos de acuerdo a su perfil, a través de la implementación de un Sistema de Recomendación de Habitabilidad en la Ciudad de Quito orientado a la planificación urbana cognitiva aplicando técnicas de Machine Learning para soporte a la toma de decisiones.

Este proceso de planeación fue asistido por SC<sup>1</sup>, se aplicó técnicas de Machine Learning. También se exploró el impacto de la IC<sup>2</sup>, como medio de apoyo en la toma de decisiones. La sobrepoblación es considerada como uno de los principales impulsores para la urbanización de las ciudades. Los resultados que se obtuvo fue la predicción del precio de la vivienda que permite mejorar el proceso de toma de decisiones en la planificación urbana.

**Palabras Clave:** planificación urbana, aprendizaje automático, ciudad cognitiva, ciudades cognitivas.

---

<sup>1</sup> SC (Sistemas Cognitivos) Los sistemas cognitivos son la evolución de la inteligencia artificial, pues tienen la capacidad de pensar, aprender y entender información en su contexto, como lo hacen los seres humanos.

<sup>2</sup> IC (Inteligencia Colectiva) Es una inteligencia repartida en todas partes, valorizada constantemente, coordinada en tiempo real, que conduce a una movilización efectiva de las competencias.

## Abstract

The city of Quito currently has an approximate population of 2,735,987 inhabitants, according to data from the INEC (National Institute of Statistics and Censuses), where the disproportionate growth of the population generates an increase in the consumption of resources, basic services, mobility and health. In this context, there are needs that grow considerably, such as: security, tranquility, basic services, green areas, among other characteristics that are sought after in a neighborhood. (Ecuador, 2019)

Overcrowding in neighborhoods due to high growth given migration issues triggers characteristics to be analyzed in this project.

In the city of Quito, the problem we face is habitability in high-risk areas. Commonly, the citizens choose to live in places without knowing the characteristics of the sector. Faced with this problem, this work proposes the application of algorithms that allow the selection of habitability characteristics and thus give recommendations to citizens according to their profile, through the implementation of a Habitability Recommendation System in the City of Quito oriented to cognitive urban planning applying machine learning techniques to support decision-making.

This planning process was assisted by CS, applying machine learning techniques. The impact of CI as a means of support in decision-making was also explored. Overpopulation is considered one of the main drivers for urbanization in cities. The results that were obtained were the prediction of the price of the house that allows to improve the decision-making process in urban planning.

**Key words:** urban planning, machine learning, cognitive city, cognitive cities.

## **Glosario de Nomenclaturas**

**API:** Interfaces de Programación de Aplicaciones.

**SC:** Sistemas Cognitivos.

**IC:** Inteligencia Colectiva.

**SLR:** Revisiones sistemáticas de literatura.

## Capítulo I. Introducción

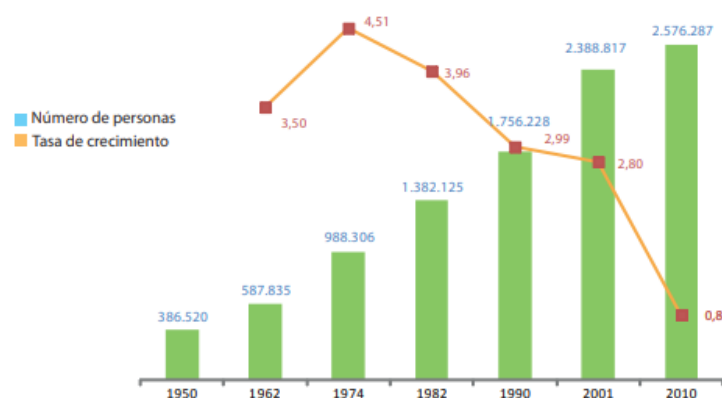
### Antecedentes

Los censos de población en el Ecuador se han levantado con una periodicidad promedio de diez años, constituyéndose en la única fuente de información para niveles geográficos menores y proporcionan información relevante para el análisis y evaluación del crecimiento poblacional, demanda de servicios básicos, condición socio-económica de las personas, entre otras múltiples aplicaciones. Los censos permiten la formulación de políticas, programas y estrategias de desarrollo social por parte del Estado (Villacís, 2011).

El crecimiento demográfico en la ciudad de Quito es notable, de acuerdo con las proyecciones poblacionales del Instituto Nacional de Estadística y Censos (INEC), Entre 1974 Y 1982 Quito creció a un elevado ritmo del 4.6% anual (Dupedier, 1995). En los resultados del censo del INEC se evidencia el crecimiento poblacional desde el año 1950 hasta el 2010, información que muestra la Figura 1 (Villacís B, 2011).

### Figura 1.

*Población total en la ciudad de Quito*



Nota: Población total en la ciudad de Quito según la división Política Administrativa vigente en cada año (Villacís B, 2011)

Debido al crecimiento de la población en la ciudad de Quito especialmente en zonas de riesgo, se implementó nuevos métodos y mecanismos para mejorar la Planificación Urbana. Al mencionar riesgo, se refiere a la seguridad de cada ciudadano para su hábitat y seguridad social. Se analizó la información para presentar como resultado final recomendaciones de habitabilidad en la ciudad de Quito mediante las Técnicas de Machine Learning.

El uso de Sistemas de Información Geográfica SIG en países Europeos y en países de América Latina desarrollados como Brasil, Argentina, ha tenido algunas aplicaciones prácticas en la planificación colaborativa, donde se destacan los beneficios de un prototipo que integra SIG y herramientas de toma de decisiones. Este prototipo involucra a diferentes actores como: el ciudadano, el gobierno, instituciones financieras y la industria de la construcción. Además, engloba a la comunidad para resolver complejos problemas de planificación y generar consenso (Pece, 2014).

La Agenda de la Conferencia Hábitat III relacionada con la planificación territorial de la vivienda ha sido escasamente explorada. La Nueva Agenda Urbana, representa una visión compartida para un futuro mejor y sostenible. Si está bien planificada y administrada, la urbanización puede ser una herramienta poderosa para el crecimiento sostenible tanto para los países en desarrollo como para los desarrollados (Clos, 2018).

### **Justificación e Importancia**

El trabajo colaborativo que integre tecnología y personas que trabajan como un todo, da lugar al surgimiento de la Inteligencia Colectiva (IC). En la temática de acceso a vivienda de la conferencia Hábitat III, se priorizan varios elementos y actores básicos que deben ser analizados en profundidad para mejorar los problemas discutidos. Los elementos principales que deben ser analizados como parte de los desafíos propuestos guardan relación con: la tenencia, servicios, acceso, habitabilidad, accesibilidad, ubicación y desarrollo cultural, en el



cual deberían ser partícipes los ciudadanos (comunidades, barrios), gobiernos locales y nacionales (municipios, gobiernos provinciales y gobierno central), instituciones financieras, cámaras de la construcción e instituciones de manejo de información geográfica y estadística públicas y/o privadas. Urbanistas y teóricos de la planificación urbana, coinciden que no hay modelos generalizables para una planificación urbana colaborativa.

El máximo potencial de esta investigación se centra en la Inteligencia Colectiva, y la creación de una base de conocimiento que permitirá la toma de decisiones de habitabilidad.

### **Planteamiento del Problema**

El crecimiento desmesurado de la población genera el aumento del consumo de recursos, servicios básicos, servicios de movilidad y servicios de salud, necesidades que crecen considerablemente.

La ciudad de Quito tiene sectores de riesgo de habitabilidad que son habitados por ciudadanos que desconocían las características del sector.

El ciudadano que busque una vivienda en la Ciudad de Quito podrá obtener las recomendaciones del precio de la vivienda en venta de acuerdo a su búsqueda de características de habitabilidad y posibilidades de pago.

### **Objetivo General**

Implementar un Sistema de Recomendación de habitabilidad en la ciudad de Quito orientado a la Planificación Urbana Cognitiva aplicando técnicas de Machine Learning, para soporte a la toma de decisiones de habitabilidad.

## **Objetivos Específicos**

**OE1:** Realizar un análisis de la literatura mediante una revisión inicial, para determinar las técnicas de sistemas de recomendaciones existentes y seleccionar la que mejor se ajuste a las necesidades de la población de Quito.

**OE2:** Recolectar y analizar los datos sobre las necesidades de los ciudadanos en una Planificación Urbana Colaborativa asistida por Sistemas Cognitivos (SC), para extraer recomendaciones de habitabilidad en la ciudad de Quito.

**OE3:** Evaluar diversas técnicas de Machine Learning mediante la aplicación de métricas para seleccionar el modelo que se ajuste al procesamiento de datos referente a las necesidades de la población de Quito.

**OE4:** Probar el modelo del algoritmo en el Sistema de Recomendación de Habitabilidad para la presentación de las respectivas sugerencias.

## **Hipótesis**

La implementación de un Sistema de Recomendación de Habitabilidad mediante el procesamiento de información permitirá identificar las características de hábitat y generar recomendaciones orientadas a la Planificación Urbana Colaborativa.

## **Alcance**

El tema “Implementación de un Sistema de Recomendación de Habitabilidad en la Ciudad de Quito Orientado a la Planificación Urbana Cognitiva Aplicando Técnicas de Machine Learning” comprende realizar algoritmos sobre la información obtenida y visualizarlos en un aplicativo, a fin de que el ciudadano Quiteño pueda elegir una vivienda adecuada.

## **Metodología**

### **Metodología de Inteligencia de Negocios**

### **Metodología de Minería de Datos**

#### **KDD**

KDD (Knowledge Discovery in Databases – KDD), es el proceso de investigación cuantitativa que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Requiere de una metodología denominada: KDD PROCESS; esta metodología trabajará con los datos consolidados.

Para la ejecución de la investigación se aplicará la metodología Investigación – Acción. Esta metodología hace un acercamiento al objeto de estudio, a partir de un diagnóstico inicial, punto de análisis, juicios de expertos, sobre una problemática, consistente en una práctica reflexiva social en la que interactúan la teoría y la práctica con miras a establecer cambios apropiados en la situación estudiada y en la que no hay distinción entre lo que se investiga, quién investiga y el proceso de investigación (Gómez, 2007).

#### **Metodología del Proyecto**

Para la consecución de los objetivos se utilizó la Metodología de Investigación Analítica, porque se ha buscado resolver un problema de regresión a través del análisis de datos. Y también es Aplicada porque modelos ya existentes se usan en un campo de investigación determinado. De carácter explicativa porque los resultados obtenidos reflejan la realidad en cuanto al comportamiento del mercado de la vivienda en Quito, debido a que se trabajó con información de viviendas y detalles de cada una y el sector de ubicación. En la Tabla 1 se muestra el desarrollo de los objetivos específicos.

**Tabla 1.***Desarrollo de objetivos específicos*

<b>Ciclo</b>	<b>Objetivo</b>	<b>Acciones</b>
1	OE1: Realizar un análisis de la literatura mediante una revisión inicial, para determinar las técnicas de sistemas de recomendaciones existentes y seleccionar la que mejor se ajuste a las necesidades de la población de Quito.	Revisión de literatura sobre Planificación Urbana, ciudad cognitiva, Inteligencia Colaborativa, Algoritmos de predicción mediante SLR.
2	OE2: Recolectar y analizar los datos sobre las necesidades de los ciudadanos en una Planificación Urbana Colaborativa asistida por Sistemas Cognitivos (SC) para extraer recomendaciones de habitabilidad en la ciudad de Quito.	Analizar las fuentes de datos con que cuenta el proyecto, el cual está relacionado con temas referentes a Hábitat III.

Ciclo	Objetivo	Acciones
3	OE3: Evaluar diversas técnicas de Machine Learning mediante la aplicación de métricas para seleccionar el modelo que se ajuste al procesamiento de datos referente a las necesidades de la población de Quito.	<p>Evaluar mediante los siguientes items:</p> <ul style="list-style-type: none"> <li>▪ Técnicas de Modelado y entrenamiento del modelo.</li> <li>▪ Análisis de Alternativas.</li> <li>▪ Estimación de resultados.</li> <li>▪ Relaciones históricas.</li> </ul> <p>Seleccionar la metodología a utilizar: KDD-Process.</p> <p>Tomar 2 técnicas de Machine Learning, y mediante una comparativa sacar la que tenga el mínimo error.</p>
4	OE4: Probar el modelo del algoritmo en el Sistema de Recomendación de Habitabilidad para la presentación de las respectivas sugerencias.	<p>Evaluar y validar el modelo a través de pruebas funcionales.</p> <p>Evaluar los resultados de la planificación urbana que comprende: parques, centros de vigilancia, vías y detalles de la vivienda en diferentes sectores de la ciudad de Quito.</p>

En el desarrollo del Sistema de Recomendaciones de Habitabilidad se aplicó los pasos para la construcción de un sistema de aprendizaje automático, debido a que con la información recolectada y tratada se siguió con el manejo de modelos y testeo de algoritmos, según (González, 2015) puede dividirse en:

1. Recolección de los datos. Proceso que ocupa gran parte del desarrollo del sistema puesto que generalmente se hace necesario recopilar grandes cantidades de datos para poder asegurar que la muestra usada es representativa del conjunto estudiado.

2. Elección de características. Se trata de un paso crítico dado que es necesario extraer aquellas variables que sean útiles para distinguir los patrones de cada categoría (técnica conocida como feature selection), en nuestro caso dichos patrones serán usados para generar las recomendaciones de habitabilidad.

3. Elección del modelo. En este paso se elegirá el modelo que más se ajuste a nuestro problema y que consiga el rendimiento esperado sobre el conjunto de test.

4. Entrenamiento del modelo. En esta fase se construye el clasificador, cuyos parámetros se ajustan a partir del conjunto de datos de entrenamiento. Encontrar los parámetros que se ajusten a nuestro modelo constituye un problema de optimización ya que el objetivo es siempre minimizar una cierta función objetivo, en dependencia de ciertas métricas.

5. Evaluación del modelo. Al hacer uso del conjunto test, se fijará una medida de error y se obtendrá el rendimiento del modelo. Si el resultado no es el esperado, se deberá probar y regresar a cada uno de los puntos anteriores y retomar el proceso de nuevo.

6. Aplicación de las técnicas de Machine Learning. El objetivo es mejorar la predicción para poder generar unos datos más fiables y de mayor calidad.

## Capítulo II: Estado del Arte

En el presente capítulo se presentan los resultados de la revisión bibliográfica sobre el tema planteado, donde intervienen los criterios de inclusión y exclusión. Se revisó las siguientes fases:

- **Definición del Objetivo:** Se relacionan las preguntas de investigación en concordancia con los objetivos específicos que direccionan al objetivo de la investigación planteada.
- **Revisión Inicial:** Se realiza una búsqueda inicial en los distintos repositorios académicos para buscar estudios relacionados con las preguntas de investigación.
- **Validación Cruzada de Estudios:** En esta fase se procede a verificar que los estudios cumplan con los criterios de inclusión y exclusión, con lo cual finalmente se obtiene el listado inicial de documentos académicos con los cuales se va a trabajar en las siguientes fases del estudio.

### Preguntas de Investigación

Se realizó las siguientes preguntas de estudio para cada objetivo, preguntas que han sido analizadas y formuladas de acuerdo a la Conferencia de la nueva Agenda Urbana (Clos, 2018). Se menciona en la Tabla 2.

**Tabla 2.***Preguntas de Investigación.*

<b>Objetivos Específicos</b>	<b>Preguntas de Investigación</b>
<b>OE1:</b> Realizar un análisis de la literatura mediante una revisión inicial, para determinar las técnicas de sistemas de recomendaciones existentes y seleccionar la que mejor se ajuste a las necesidades de la población de Quito.	<b>OE1-RQ 1.1:</b> ¿Cómo se emplean los Sistemas Cognitivos (SC), en la Planificación Urbana Cognitiva?
<b>OE2:</b> Recolectar y analizar los datos sobre las necesidades de los ciudadanos en una Planificación Urbana Colaborativa asistida por Sistemas Cognitivos (SC), para extraer recomendaciones de habitabilidad en la ciudad de Quito.	<p><b>OE2-RQ 2.1:</b> ¿Cómo la planificación urbana colaborativa, y sistemas cognitivos (SC), podrían mejorar la toma de decisiones para una planificación urbana?</p> <p><b>OE2-RQ 2.2:</b> ¿Cuál es la técnica para la generación de sistemas de recomendaciones de hábitat que mejor se ajuste a la realidad de la ciudad en estudio y por qué?</p> <p><b>OE2-RQ 2.3:</b> ¿Cuáles son las fuentes de datos con las que se contará?</p> <p><b>OE2-RQ 2.4:</b> ¿Cuáles son los análisis que se tienen acerca de Planificación Urbana en la ciudad de Quito?</p>
<b>OE3:</b> Evaluar diversas técnicas de Machine Learning mediante la aplicación de métricas para seleccionar el modelo que se ajuste al procesamiento de datos referente a las necesidades de la población de Quito.	<p><b>OE3-RQ 3.1:</b> ¿Cuáles son las necesidades más relevantes en cuanto a habitabilidad de los quiteños en la actualidad?</p> <p><b>OE3-RQ 3.2:</b> ¿Qué técnica de Machine Learning se acopla mejor a las necesidades de cada barrio en Quito?</p> <p><b>OE3-RQ 3.3:</b> ¿Existen valores atípicos que puedan afectar el resultado de la ejecución de los algoritmos de técnicas de Machine</p>



Objetivos Específicos	Preguntas de Investigación
	<p>Learning aplicadas en Sistemas de Recomendación?</p> <p><b>OE3-RQ 3.4:</b> ¿Será posible validar la efectividad del modelo generado?</p> <p><b>OE3-RQ 3.5:</b> ¿Qué herramientas de minería de datos y/o visualización adoptará para alcanzar un mejor resultado?</p> <p><b>OE3-RQ 3.6:</b> ¿Cuáles serán las características del modelo de algoritmo apropiado?</p>
<p><b>OE4:</b> Probar el modelo del algoritmo en el Sistema de Recomendación de Habitabilidad para la presentación de las respectivas sugerencias.</p>	<p><b>OE5-RQ 4.1:</b> ¿Qué tipo de reportes de diagnóstico de requerimientos se obtendrá mediante el análisis del modelo?</p>

**Integración del Grupo de Control:** El grupo de control está conformado por los estudios que cumplen con los criterios de inclusión y exclusión para lo cual se procedió a realizar un análisis inicial del título de los estudios, introducción, conclusiones y palabras claves. Los estudios seleccionados para el grupo de control son los siguientes, detallados en la Tabla 3:

**Tabla 3.**

*Estudios por Grupos de Control*

Grupo Control	Título	Palabras Claves
EC1	Emotional Cognition in Urban Planning and Design	Complexity, cognition, urban planning, design.
EC2	A Spatial Analysis Approach for Evaluating the Service Capability of	Greenway, service capability, accessibility, regular grid, spatial

<b>Grupo Control</b>	<b>Título</b>	<b>Palabras Claves</b>
	Urban Greenways-A Case Study in Wuhan	distribution.
EC3	Overview of Three-dimensional Geological Modeling Technology	Three-dimensional, Geological Modeling, Visualization, Mathematical modeling;
EC4	Multi-Scale Urban Analysis Using Remote Sensing and GIS	GIS, Remote Sensing, Land Use Change, Urban Growth, Multi-Scale Urban, Analysis.
EC5	On-line learning of predictive kernel models for urban water demand in a smart city	Smart cities, urban water demand, kernel regression, on-line learning
EC6	Applying the fuzzy analytical hierarchy process in cognitive cities	Cognitive City, Decision Support System, E-Governance, Fuzzy Analytical Hierarchy Process, Granular Computing, Smart City.
EC7	Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges	Machine Learning, Smart Cities, Big Data

### **Construcción de la Cadena de Búsqueda**

Para la construcción de la cadena de búsqueda se usaron las palabras que más se repiten en cada contexto definido a partir de los estudios del grupo de control, para el presente estudio se definieron los siguientes contextos: Urban Planning, design, model, Machine Learning, cognitive city, cognitive cities, Smart city. Detallada en la Tabla 4.

**Tabla 4.***Construcción de la cadena de búsqueda.*

Contexto	Palabra Clave	EC1	EC2	EC3	EC4	EC5	EC6	EC7	Número de Repeticiones
Planificación Urbana	Urban Planning	x	x	x	x	-	x	x	6
	design	x	x	x	-	x	x	x	6
	cognition	x	-	-	-	-	x	x	3
	Model	x	x	x	x	x	x	x	7
	Complexity	x	-	x	-	x	-	-	3
Análisis espacial	Greenway	-	x	-	-	x	-	-	2
	service capability	-	x	-	-	-	-	-	1
	accessibility	x	-	-	x	x	x	-	4
	spatial distribution.	x	x	-	x	-	-	-	3
Modelos predictivos	Smart cities	x	-	x	x	x	x	x	6
	kernel regression	x	-	-	-	x	-	-	2
	on-line learning	x	x	-	-	x	-	-	3
	Machine Learning	-	x	x	-	x	x	x	5

La cadena de búsqueda está formada por la unión de las palabras claves que más se repitieron en cada contexto, los conectores usados son OR para las palabras que están dentro del mismo contexto y el conector AND para las palabras que están en contextos distintos, de esta manera se estableció la siguiente cadena de búsqueda:

**(((urban planning) AND Machine Learning) AND cognitive city) OR cognitive cities) NOT Three-dimensional Geological)**

Se realizó el resumen de artículos seleccionados y publicados en: IEEEExplore, y ACM Digital Library:

**Blockchain and IoT-based Cognitive Edge Framework for Sharing Economy Services in a Smart City** (Rahman, 2019)

Proponen una infraestructura basada en Blockchain para respaldar los servicios de contrato inteligente espacio-temporales orientados a la seguridad y la privacidad para una economía de Internet sostenible en Internet de las cosas (IoT) habilitada en mega ciudades inteligentes. La infraestructura aprovecha los nodos de niebla cognitiva para alojar y procesar la carga y la carga multimedia 'multimedia geo etiquetadas' descargadas desde una aplicación móvil y los nodos de IoT, utiliza Inteligencia Artificial para procesar y extraer información de eventos significativos, produce analítica digital semántica y guarda resultados en Blockchain y utiliza repositorios en la nube descentralizados para facilitar el intercambio de servicios económicos.

**Fuzzy Reasoning in Cognitive Cities: An Exploratory Work on Fuzzy Analogical Reasoning Using Fuzzy Cognitive Maps** (D'Onofrio, 2018)

Los autores analizan el razonamiento difuso global que se desarrolla. El marco conceptual para los sistemas de diálogo urbano permite a tales sistemas emular el razonamiento analógico humano mediante el uso de la computación cognitiva. También

realizaron dos experimentos para confirmar la estabilidad de la base teórica del marco, y ocho expertos lo evaluaron. Sobre la base de los resultados experimentales, los autores han ampliado el marco con un componente complementario centrado en la gestión del conocimiento. Por lo tanto, se presenta una memoria similar a una red que se puede utilizar para recuperar la información necesaria para el proceso de razonamiento analógico difuso.

### **Applying the fuzzy analytical hierarchy process in cognitive cities** (Kaltenrieder, 2014)

Este documento presenta una aplicación móvil, como la primera parte de un marco interactivo. El marco mejora la interacción entre las ciudades y sus ciudadanos, introducir el Proceso de Jerarquía Analítica Fuzzy (FAHP) como un método potencial de adquisición de información para mejorar los esfuerzos existentes de gestión ciudadana para ciudades cognitivas. La gestión ciudadana se mejora con la visualización avanzada mediante el uso de mapas cognitivos difusos.

### **Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges** (Mohammadi, 2018)

Se aplica el aprendizaje automático tradicional, las técnicas asumen un modelo de entrenamiento fijo y un contexto estático. Se analiza que son sistemas críticos para el aprovisionamiento óptimo y eficiente, los datos de capacitación evolucionan a lo largo del tiempo, estos desafíos incluyen: reciclaje de datos, muestreo eficiente y diseño de modelos escalables.

### **Conclusión:**

Al realizar la revisión de literatura se identificó una problemática y es el desarrollo de las ciudades y su ritmo acelerado, esto genera grandes cantidades de datos que no son analizados, la información es volátil y no procesada por los individuos u organismos (distritos), para aplicar inteligencia cognitiva, se tiene que realizar una semi-supervisión, esta puede

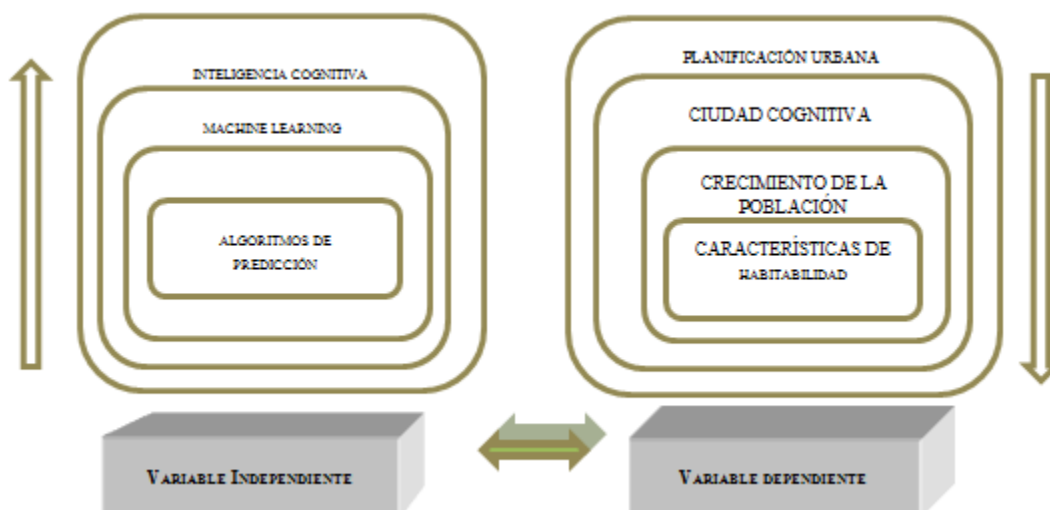
manejar el lado cognitivo de servicios de ciudad inteligente y mejorar su rendimiento, con casos de uso que abarcan los diferentes dominios de ciudades inteligentes. Se identifica que existen algoritmos de aprendizaje automático que pueden ser utilizados para aprender de los grandes datos recopilados, algoritmos que no han sido aplicados en Ecuador. En conclusión, la literatura propone la gestión del conocimiento basado en el uso de mapas cognitivos difusos que pueden almacenarse en bases de datos de gráficos y trabajará conjuntamente con las técnicas de Machine Learning.

### Capítulo III: Marco Teórico

La fundamentación teórica tiene el propósito de generar una congruencia de la teoría con la hipótesis, para esto se realizó un análisis de la teoría a partir de las variables de la hipótesis, con el fin de investigar jerárquicamente cada categoría hasta llegar a la categoría que comprende y explicar las variables dependientes e independientes del tema de estudio, para lo cual se planteó la jerarquía de estudio que se muestra en la Figura 2.

**Figura 2.**

*Variable dependiente e independiente*



La fundamentación teórica busca la congruencia con la hipótesis, a medida que se analiza y desarrolla el sistema de recomendaciones analizaré las variables dependiente e independiente.

#### Fundamentación de la Variable Independiente

**Inteligencia Cognitiva:** Es aquella que tiene que ver con diversas capacidades del ser humano como son la memoria, la atención o el lenguaje. (Drenth, 2015).

Las posibilidades de vivir mejor, sin que nadie se quede atrás, depende grandemente de la capacidad de determinar y priorizar los problemas que afectan al distrito y a su población y, a partir de las potencialidades de sus recursos humanos, territoriales e institucionales, formular en consenso las soluciones que viabilizarán mejores estándares de vida.

**Machine Learning:** Machine Learning o Aprendizaje Automático es aprender de los datos, es descubrir la estructura y los patrones que subyacen en ellos. El objetivo principal del aprendizaje automático es la extracción de la información contenida en un conjunto de datos con el fin de adquirir conocimiento que permita tomar decisiones sobre nuevos conjuntos de datos (González, 2015). Machine Learning tiene actualmente una multitud de aplicaciones, tales como: motores de búsqueda, diagnósticos médicos, análisis de los mercados financieros, detección del fraude, clasificación de secuencias de ADN, aplicación de la robótica, en fin Machine Learning establece técnicas que hacen posible que los dispositivos aprendan por si solos. Por estas razones y con la data existente de las viviendas que se trabajó, es fácil que ahora el aplicativo lance un resultado predictivo del precio de la vivienda según su sector, con un margen mínimo de error, porque aprendió, tras las diferentes cargas y ejecuciones.

El aporte que realiza este aplicativo es la automatización de construcción de modelos analíticos, debido a que tiene algoritmos que interactivamente aprenden de los datos.

### **Técnicas de Aprendizaje:**

Entre las técnicas que se revisará en primera instancia será: Aprendizaje Supervisado donde los algoritmos de esta clase utilizan datos etiquetados para predecir lo que han aprendido, los datos y las clases de datos tales como: categorías y etiquetas que se utilizarán en la capacitación son conocidos de antemano. El sistema aprende e interpreta los nuevos datos entrantes con lo que han aprendido.



La fase de aprendizaje supervisado que consume más tiempo es la preparación de los datos. El sistema entrenado con datos de entrenamiento mal preparados hará malas predicciones. Se clasifican en dos grupos:

- Método de clasificación;
- Método de regresión;

Según el desarrollo del trabajo de tesis, se realizó el tratamiento de la data, se buscó el algoritmo adecuado "Supervisado" y se seleccionó las características dominantes, y se preparó la data para un entrenamiento del modelo.

**Algoritmos de Predicción:** La minería de datos ("Data Mining" en inglés) es el proceso de analizar y resumir los datos desde distintas perspectivas. Trabaja para encontrar patrones de datos y relaciones en grandes conjuntos de información con algoritmos, que son conjuntos de reglas para resolver un problema mediante una serie de pasos concretos.

## **Fundamentación de la Variable Dependiente**

### **Características de Habitabilidad:**

**Indicadores de Asociatividad.-** La población de Quito lleva consigo una larga tradición de asociatividad entre cooperativas de vivienda, ligas deportivas barriales, grupos juveniles y otros, que forman un tejido social vivo; sin embargo se aduce que las sociedades contemporáneas han caminado hacia la falta de integración y cohesión social, lo cual tiende al debilitamiento de dicho tejido. Factores políticos y brechas sociales que existen en determinados sectores o actores abonan en esta situación.

**Indicadores de Convivencia.** - Los problemas de convivencia ciudadana dan cuenta de una alta percepción de inseguridad; los hechos de muertes violentas; la presencia de la violencia intrafamiliar, de género, maltrato infantil y violencia sexual, acoso y violencia escolar,

incrementan esta percepción. La mayor connotación de delitos son los asaltos comunes, robo de domicilios, autos y auto partes; los espacios públicos están abandonados o deteriorados, y existen asentamientos irregulares (Recalde et al., 2019).

### **Ciudad Cognitiva:**

Las ciencias cognitivas estudian los procesos de la mente humana: cómo se aprende, cómo recuerda y cómo procesa la información. De la misma manera se puede aplicar un sistema cognitivo a una ciudad que sea capaz de entender, razonar, aprender e interactuar en lenguaje natural a través de todo un universo de datos, donde ayuda a empresas y organizaciones a aumentar experiencia para vencer nuevos y cambiantes retos, una ciudad cognitiva es la que aprende cada día a través de sus actores (Thagard, 2008).

### **Planificación Urbana:**

La Planificación Urbana es una profesión que articula y media el complejo sistema de relaciones e intereses que se establecen entre las diversas disciplinas y múltiples factores que influyen en el diseño y desarrollo de las ciudades. Se centra en la comprensión del fenómeno urbano, con la finalidad de optimizar su uso y lograr un desarrollo sustentable, a través de la gestión pública y privada. Por tanto, involucra la distribución espacial de todas las actividades humanas y la toma de decisiones que inciden en el espacio urbano en sus distintas escalas (López, 2020).

### **Sistema de Recomendación:**

Los Sistemas de Recomendaciones son algoritmos destinados a sugerir a los usuarios elementos relevantes, como películas para ver, texto para leer, productos para comprar o cualquier otra cosa depende de la industria (Pérez, 2015).

Los Sistemas de Recomendaciones son realmente críticos en algunas industrias ya que pueden generar una gran cantidad de ingresos cuando son eficientes o también pueden ser una forma de diferenciarse significativamente de la competencia (Pérez, 2015).

### **Beneficios de la implementación de un Sistema de Recomendación:**

Como usuarios de internet, todos interactúan con los Sistemas de Recomendación de productos casi todos los días, durante las búsquedas en Google, cuando se usa servicios de transmisión de películas o música, compra en línea, se navega en los medios sociales y se usa cosas como aplicaciones de citas (Pérez, 2015).

Los Sistemas de Recomendación de productos son una de las aplicaciones más exitosas y extendidas de Machine Learning en la empresa. Cuando se configuran correctamente, pueden aumentar significativamente las ventas, los ingresos, las tasas de clics, las conversiones y otras métricas importantes. Esto se debe a que la personalización de las recomendaciones de productos o contenidos a las preferencias de un usuario en particular crea un efecto positivo en la experiencia del usuario. Y esto, a su vez, se traduce en métricas que son más difíciles de medir: satisfacción del cliente, lealtad, afinidad con la marca (Aggarwal, 2016).

### **Tipos de Sistema de Recomendación:**

Los modelos de filtrado colaborativo utilizan el poder colaborativo de las calificaciones proporcionadas por múltiples usuarios para hacer recomendaciones. El principal desafío en el diseño de métodos de filtrado colaborativo es que las matrices de calificaciones subyacentes son escasas.

Hay dos tipos de métodos que se utilizan comúnmente en el filtrado colaborativo, que se conocen como:

**Métodos Basados en Memoria:** Los métodos basados en memoria también se conocen como algoritmos de filtrado colaborativo vecinos. Estos fueron algunos de los primeros algoritmos de filtrado colaborativo, en los que las clasificaciones de las combinaciones de elementos de usuario se predicen sobre la base de sus vecindarios (Hu, 2014).

**Métodos Basados en Modelos:** En los métodos basados en modelos, el aprendizaje automático y los métodos de extracción de datos se utilizan en el contexto de los modelos predictivos. En los casos en que el modelo está parametrizado, los parámetros de este modelo se aprenden en el contexto de un marco de optimización. Algunos ejemplos de esos métodos basados en modelos son los árboles de decisión, los modelos basados en normas, los métodos bayesianos y los modelos de factores latentes. Estos métodos, como los modelos de factores latentes, tienen un nivel de cobertura incluso para matrices de calificaciones escasas (Hu, 2014).

## Capítulo IV: Desarrollo

### Aplicación de la Metodología

El proceso de investigación cuantitativa que intenta descubrir patrones en grandes volúmenes de conjuntos de datos requiere de la metodología KDD PROCESS, ésta metodología trabajo con los datos consolidados, se realizó una selección y pre procesamiento con los datos preparados, luego se realizó una minería de datos para encontrar el modelo adecuado, y así finalmente poder evaluar el algoritmo.

El objetivo de este capítulo es el desarrollo completo del sistema Machine Learning orientado a la WEB según los objetivos planteados en el proyecto, con tecnologías de código abierto como son Python, Pentaho.

Se usó Python como lenguaje de desarrollo para el prototipo. Python tiene una sintaxis consistente y es fácil de leer y entender su código. Existen muchas librerías para Python que facilitan la manipulación de datos y el procesamiento matemático que es requerido al momento de implementar algoritmos de aprendizaje de máquina (González, 2018).

Jupyter Notebook es una aplicación cliente-servidor. Permite crear y compartir documentos en formato JSON que siguen un esquema versionado y una lista ordenada de celdas de entrada y de salida. Estas celdas albergan, entre otras cosas, código, texto, fórmulas matemáticas y ecuaciones, o también contenido multimedia. El programa se ejecuta desde la aplicación web cliente que funciona en cualquier navegador estándar. El nombre del proyecto “Jupyter” proviene de unir los 3 lenguajes de programación de núcleo Julia, Python y R (Díaz García, 2018).

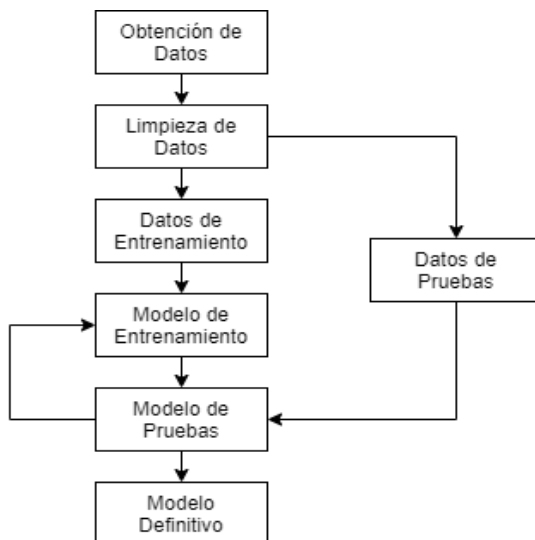
ETL (Extracción, Transformación y Carga) hace referencia a las herramientas o piezas de software que son responsables de la extracción de los datos desde diferentes fuentes, su

limpieza, personalización e inserción en un almacén de datos (Data Warehouse). En el desarrollo de este proyecto se utilizó la herramienta Pentaho,

En la Figura 3 se visualiza el proceso KDD para realizar el modelo de Machine Learning.

### Figura 3.

*Proceso de Machine Learning KDD. Tomado de (Moine, 2011).*



En relación con la metodología KDD PROCESS se detallan los siguientes pasos:

#### Obtención de los Datos

Se dispone de información en la cual se visualiza en un aplicativo en la WEB, El método de extracción de datos se realizó con un archivo con formato CSV (Comma Separated Values), esta información se puede extraer por medio de herramientas informáticas como “Pentaho”, que nos permite obtener datos depurados en modelo de datos relacional.

Las fuentes de datos involucradas en el proyecto son las datas de lotes, otorgadas y obtenidas por El Instituto Nacional de Estadísticas y Censos, estas detallan información de lotes, que contienen datos como áreas de construcción, zonas y parroquia. Esta información fue de gran ayuda para extraer los sectores de la ciudad de Quito e implementarlo en el

sistema de recomendación, para poder escoger según el sector los barrios a visualizarse en el aplicativo. La figura 4 representa la información inicial de los lotes con que se trabajó.

**Figura 4.**

*Archivo de entrada de Lotes*



La otra data que se visualiza en la figura 5, se obtuvo por recolección de las diferentes inmobiliarias que publicaron sus precios en la web, en la cual se recabó la información histórica de las viviendas a ser vendidas.

**Figura 5.**

*Obtención de los datos.*



La Figura 5 representa la información inicial de las viviendas en venta, con que se trabajó en el proyecto habitacional.

En esta etapa se realiza el análisis de las viviendas en venta, con información extraída desde enero 2020. Cabe destacar que, si existen datos similares sobre la vivienda con diferentes precios, no es que el aplicativo o modelo en ejecución está mal, sino que el usuario por tratar de vender publica en diferentes plataformas inmobiliarias.

En la tabla 5 se detallan los campos que se obtiene del registro de Excel de las inmobiliarias, para ser procesados:

**Tabla 5.**

*Descripción de las variables del SRH*

<b>Campos:</b>	<b>Detalle del campo</b>
Casa	Se detalla si es casa o departamento, se aplica variables dummy. Donde 0= departamento y 1=casa.
Sector	Sector donde está ubicada la vivienda.
Detalle sector	Detalle donde está ubicada la vivienda, sea en el Centro, Norte, Sur y Valles.
Años de antigüedad	Años de construcción que tiene la vivienda. Cuando es nueva 0=Nueva sin estrenar.
Plantas	El número de pisos que tiene la vivienda.
número de habitaciones	Descripción del número de habitaciones.
número de baños	Descripción del número de baños.



<b>Campos:</b>	<b>Detalle del campo</b>
total_metros	Descripción del total de metros de construcción.
servicios	Detalles de los servicios que ofrece la vivienda, puede ser: Guardianía, Zona residencial.
Garaje	Detalle si cuenta con garaje, se aplica variables dummy. Donde 0= No tiene y 1=Si tiene.
Precio	Detalles de los precios de las viviendas.
Fecha	Fecha en que fue extraída la información.

### Limpieza de los Datos

A continuación, se redacta como se realizó la limpieza de los datos con el siguiente proceso de Extracción, Transformación y Carga (ETL) realizado en Pentaho.

Para la extracción de los datos de las tablas origen, se realizó el análisis de los siguientes campos: Tipo de Vivienda, Servicios Básicos, Cuenta con Garage, Número de habitaciones, Número de Baños, Número de Plantas, Años de antigüedad, Área de la Vivienda, Precio Vivienda. En la figura 6 se visualiza el flujo del proceso de extracción de los datos.

### Figura 6.

*Proceso ETL de limpieza*

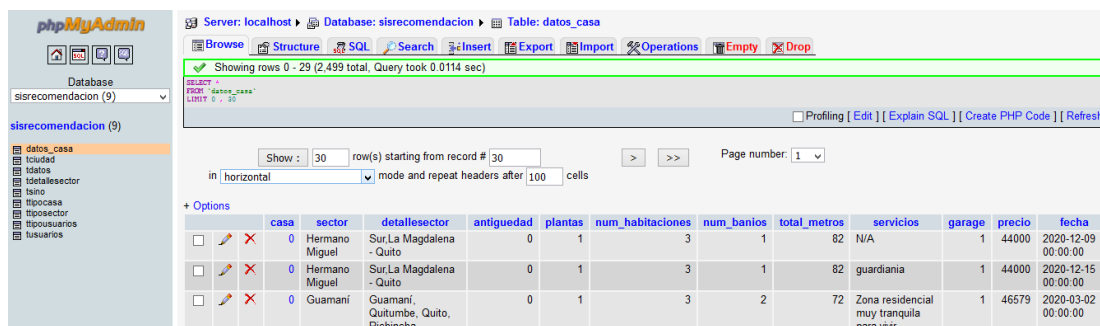


*Nota.* La figura representa el flujo final del ETL de la limpieza de la información. Derivado del Software Pentaho.

En la figura 7 se visualiza la base de datos que se obtiene después de realizar el proceso de extracción en Pentaho, la información es ingresada a la base de datos de MySQL, información que va a ser procesada y visualizada en el aplicativo SRH.

**Figura 7.**

### Base de Datos MySQL



casa	sector	detallesector	antiguedad	plantas	num_habitaciones	num_banios	total_metros	servicios	garage	precio	fecha
0	Hermano Miguel	Sur,La Magdalena - Quito	0	1	3	1	82	N/A	1	44000	2020-12-09 00:00:00
0	Hermano Miguel	Sur,La Magdalena - Quito	0	1	3	1	82	guardiania	1	44000	2020-12-15 00:00:00
0	Guamani	Guamani, Quitumbe, Quito, Pichincha	0	1	3	2	72	Zona residencial muy tranquila para vivir	1	46579	2020-03-02 00:00:00

### Selección del Algoritmo y Pre-Procesamiento

Debido a los documentos que se analizaron y se hacen referencia en este estudio, permitió determinar cuál modelo tiene menos error y cual se adapta mejor al sistema de recomendación de habitabilidad. Se escoge lineal regresión y logistic regresión los cuales se utilizó incorporándolo en el aplicativo WEB. Para determinar cuál modelo tiene el mínimo error en las diferentes ejecuciones se incrementó el número de datos.

La regresión es un enfoque de aprendizaje supervisado. Se puede utilizar para modelar variables continuas y realizar predicciones. Ejemplos de aplicación del algoritmo de regresión lineal son los siguientes: predicción del precio de bienes inmuebles, predicción de ventas, predicción de los puntajes de los exámenes de los estudiantes, predicción de movimientos en el precio de las acciones en bolsa. En Regresión se tiene los conjuntos de datos etiquetados y el valor de la variable de salida está determinado por los valores de la variable (o variables) de entrada, por lo que es el enfoque de aprendizaje supervisado (Ray, 2019).

En este proyecto de tesis se evaluó los algoritmos de regresión lineal y regresión logística en la predicción del precio de la vivienda. A partir de variables como: Tipo de Vivienda, Servicios Básicos, Cuenta con Garage, Número de habitaciones, Número de Baños, Número de Plantas, Años de antigüedad, Área de la Vivienda, Precio Vivienda y sector de ubicación. Estos datos que son de importancia debido a que describen las características de la propiedad y por lo tanto impactan en el precio de la vivienda. Además, son datos que manejan las empresas inmobiliarias en Ecuador.

En la Tabla 6 se detalla la carga de los datos para entrenar el modelo de regresión lineal y regresión logística con el 80% de entrenamiento de los datos.

**Tabla 6.**

*Carga de Datos al modelo con el 80% de entrenamiento.*

<b>Nombre Perfil de Carga</b>	<b>Fecha Inicio</b>	<b>Fecha Fin</b>	<b>Hora Inicio</b>	<b>Hora Fin</b>	<b>Resultados</b>
Testing_xr01	04/04/2020	04/04/2020	11:30	12:40	Linear Regression Logistic regression
Testing_xr02	14/12/2020	14/12/2020	09:00	17:40	Linear Regression Logistic regression
Testing_xr03	15/12/2020	15/12/2020	08:30	18:30	Linear Regression Logistic regression
Testing_xr04	16/12/2020	16/12/2020	08:30	19:00	Linear Regression Logistic regression

*Nota.* La tabla representa diferentes sets de datos cargados y ejecutados con el 20% testing y 80% de entrenamiento.

Además, de entrenar el modelo con el 80% de la data, se entrenó el modelo con el 70% de los datos para tener un mejor rendimiento. En la Tabla 7 se detalla la carga de los datos.

**Tabla 7.**

*Carga de Datos al modelo con el 70% de entrenamiento.*

<b>Nombre Perfil de Carga</b>	<b>Fecha Inicio</b>	<b>Fecha Fin</b>	<b>Hora Inicio</b>	<b>Hora Fin</b>	<b>Resultados</b>
Testing_xr01_nueva_val	22/12/2020	22/12/2020	07:00	10:40	Linear Regression Logistic regression
Testing_xr02_nueva_val	22/12/2020	22/12/2020	08:00	10:30	Linear Regression Logistic regression
Testing_xr03_nueva_val	22/12/2020	22/12/2020	08:30	18:30	Linear Regression Logistic regression
Testing_xr04_nueva_val	22/12/2020	23/12/2020	08:30	14:00	Linear Regression Logistic regression

*Nota.* La tabla representa diferentes sets de datos cargados y ejecutados con el 30% testing y 70% de entrenamiento.

### **Modelo de Regresión Lineal:**

El modelo más usado para hacer predicciones en sistemas de recomendación de habitabilidad es el modelo de regresión lineal múltiple que se encuentra en los cursos y libros de texto de introducción a la estadística. Este modelo se utiliza para ajustar una relación entre una variable de resultado numérica (Shmueli, 2019).

### **Modelo de Regresión Logística:**

La regresión logística no solo se usa para la clasificación, sino también para la regresión y por eso fue uno de los modelos usados para la tesis. El resultado es binario, ya que da la probabilidad de que ocurra un evento o no (en términos de 0 y 1) en función de los valores de las variables de entrada. Por ejemplo, predecir si un tumor es maligno o benigno o si un correo electrónico se clasifica como spam o no son las instancias que pueden considerarse como

resultado binomial de la regresión logística. También puede haber un resultado multinomial de regresión logística, predicción del tipo de cocina preferido: chino, italiano, mexicano, etc. Puede haber un resultado ordinal, así como: calificación de producto de 1 a 5, etc. Por lo tanto, la regresión logística se ocupa de la predicción de la variable objetivo que es categórica. Mientras que la regresión lineal se ocupa de la predicción de valores de la variable continua e, g. predicción del precio de la propiedad inmobiliaria en un lapso de 3 años (Ray, 2019).

En el proyecto que se desarrolló se analizó los datos de entrada que son las características de la vivienda, se ejecutó algoritmos de ML, y se considera el error absoluto mínimo para su evaluación y se obtiene un resultado final del precio.

### **Ejecución y Validación de los Modelos**

Una de las principales características de la Minería de datos, es tener un modelo con error mínimo y un mejor rendimiento. Para la validación del modelo se usó el MAE y RMSE.

MAE (Mean Absolute Error), es el promedio sobre la muestra de prueba de las diferencias absolutas entre la predicción y la observación real (González Arroyave, 2019).

RMSE (Root Mean Squared Error), es la métrica popular en la evaluación de la precisión de predecir ratings (González Arroyave, 2019).

Después de haber corrido las diferentes pruebas se visualiza que a medida que se ejecuta el algoritmo con mayor número de registros el valor del MAE mejora con 18118.6840 el rendimiento del modelo, mientras que el valor de RMSE en la cuarta ejecución se reduce a 66740.5551. En la Tabla 8 se visualiza los resultados de la ejecución y validación de los modelos logistic regression y linear regression.

**Tabla 8.***Resultados*

<b>Nombre Perfil de Carga</b>	<b>Número de Registros</b>	<b>Técnicas de machine learning</b>	<b>Valor de MAE</b>	<b>Valor de RMSE</b>
Testing_xr01 vs Testing_xr01_nueva_val	98	Linear Regression	102111.8563	131746.3614
		Logistic regression	56108.6000	131746.3614
Testing_xr02 vs Testing_xr02_nueva_val	579	Linear Regression	52960.5272	97553.2157
		Logistic regression	34834.4655	103792.9103
Testing_xr03 vs Testing_xr03_nueva_val (1248)	1248	Linear Regression	48198.5400	89360.6074
		Logistic regression	28190.1920	91495.9472
Testing_xr04 vs Testing_xr04_nueva_val (2499)	2499	Linear Regression	42164.4116	109148.7547
		Logistic regression	18118,6840	66740.5551

Se visualiza en la tabla 9 que en la segunda ejecución con 579 registros se obtuvo el valor residual con valores significativos, mientras que con 2499 registros los valores del residuo disminuyen, el término residual es el resultado del valor de la vivienda real menos el valor predicho.

**Tabla 9.***Resultados valor residual.*

Nombre Perfil de Carga	Número de Registros	Residual					
		TOTAL_VALUE	predicted	residual			
Testing_xr02	579	192	35900	35900	0		
		419	225000	180000	45000		
		350	55000	55000	0		
		410	225000	180000	45000		
		480	132500	132500	0		
		...	...	...	...		
		197	96000	91000	5000		
		247	92800	92800	0		
		407	225000	180000	45000		
		74	72628	28000	44628		
		334	170000	170000	0		
		Testing_xr04	2499	860	114000	114000	0
				710	60000	60000	0
799	114000			114000	0		
432	72628			65628	7000		
236	92800			96900	-4100		
...	...			...	...		
228	92800			100000	-7200		
1195	178000			178000	0		
894	160000			160000	0		
1024	79000			79000	0		
808	79800			74000	5800		
[250 rows x 3 columns]							

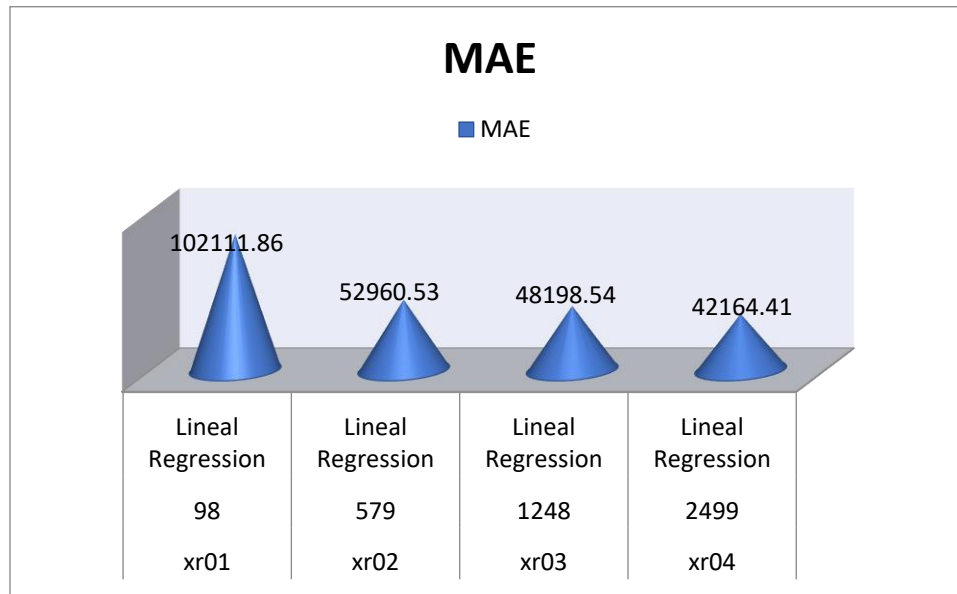
**Validación del Modelo**

Una vez que se obtuvo los resultados en la simulación, se procede a analizar el resultado del valor del MAE, con la Figura 8 y Figura 9:

## Linear Regression:

**Figura 8.**

*Validación Linear Regression MAE*



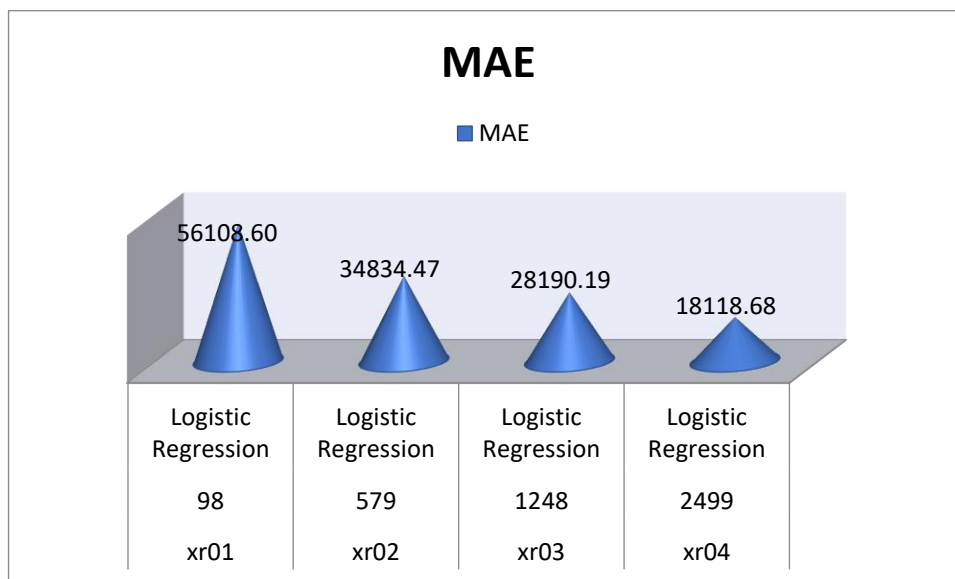
El algoritmo mejor evaluado con Linear Regression con el valor 42164,41. En estas figuras se muestra que, redondear el valor previsto puede disminuir los valores de MAE para todos los algoritmos de recomendación. Los valores de MAE se reducen significativamente después de redondear los resultados MAE.



## Logistic regression:

**Figura 9.**

*Validación Logistic Regression MAE*



El algoritmo Logistic Regression con el valor 18118,68 es el que mejor rendimiento presenta con el 20% de pruebas y el 80% de entrenamiento entre todo el conjunto de carga de datos al modelo.

Se optimiza los resultados, mientras más datos tiene el modelo se entrena y lanza mejor resultados, al minimizar el error del MAE.

A detalle se observa en la Tabla 10, se separa el 20% para pruebas y el 80% para entrenamiento. Además, se probó con 30% pruebas y 70% de entrenamiento, con el fin de analizar si los resultados tienen un cambio significativo.

**Tabla 10.***Evaluación MAE 20% y 30%*

<b>Nombre Perfil de Carga</b>	<b>Técnicas de machine learning</b>	<b>Valor de MAE 20%</b>	<b>Valor de MAE 30%</b>
Testing_xr01 vs Testing_xr01_nueva_val	Linear Regression	102111.8563	117015.3927
	Logistic regression	56108.6000	95053.1000
Testing_xr02 vs Testing_xr02_nueva_val	Linear Regression	52960.5272	52960.5272
	Logistic regression	34834.4655	34834.4655
Testing_xr03 vs Testing_xr03_nueva_val (1248)	Linear Regression	48198.5400	47369.1364
	Logistic regression	28190.1920	28090.7627
Testing_xr04 vs Testing_xr04_nueva_val (2499)	Linear Regression	42164.4116	43651.1624
	Logistic regression	18118.6840	18377.0947

El modelo de habitabilidad cumple con los objetivos que son, evaluar diversas técnicas de Machine Learning mediante la aplicación de métricas para seleccionar el modelo que se ajuste al procesamiento de datos referente a las necesidades de la población de Quito. Y probar el modelo del algoritmo en el Sistema de Recomendación de Habitabilidad para la presentación de las respectivas sugerencias de precios.

Se realizó un nuevo análisis con registro de viviendas en venta de los últimos 6 meses, entre el mes de septiembre 2020 y el mes de marzo 2021. Donde con 180 registros, en la ejecución de los modelos: Logistic regression y Linear Regression, se obtuvo los siguientes valores, visualizados en la siguiente Tabla 11.

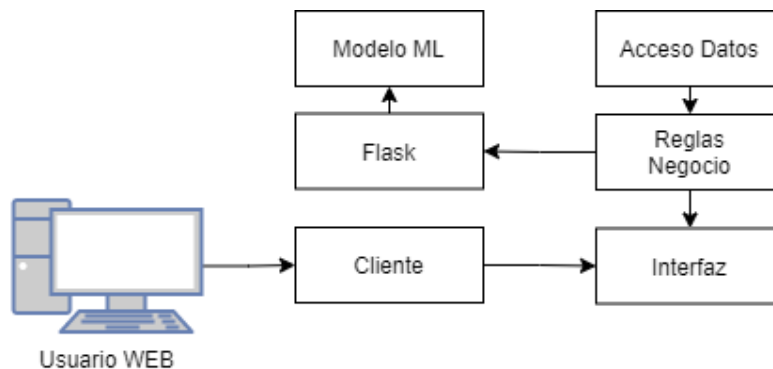
**Tabla 11.***Validación de los modelos.*

<b>Modelo</b>	<b>Detalle de errores:</b>
<b>Logistic regression</b>	Mean Error (ME): 0.0000
	Root Mean Squared Error (RMSE): 104058.7373
	Mean Absolute Error (MAE): 54377.4327
	Mean Percentage Error (MPE): -18.7972
	Mean Absolute Percentage Error (MAPE): 48.1704
<b>Linear Regression</b>	Mean Error (ME): 18611.3887
	Root Mean Squared Error (RMSE): 202994.5452
	Mean Absolute Error (MAE): 115983.3283
	Mean Percentage Error (MPE): -25.0742
	Mean Absolute Percentage Error (MAPE): 74.9851

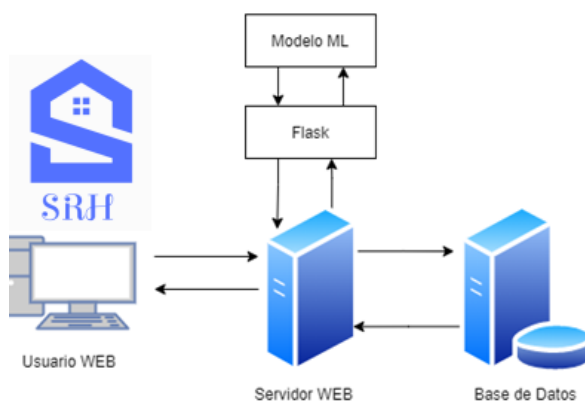
Se realizó un breve análisis y se procedió a embeber el algoritmo Logistic Regression en la aplicación Web, debido a que tiene mejor rendimiento, el error absoluto medio (MAE) es de 54377.4327.

### **Arquitectura del Sistema**

Este sistema ML en la WEB trabaja con una arquitectura cliente-servidor, en donde el Flask framework de Python será el que nos proporcione toda la información de datos por medio de una interfaz de usuario, el cliente consume todos los servicios necesarios para su funcionamiento, envía y recibe información por parte del servidor, este modelo se presenta por medio de las siguientes capas del sistema, las capas que la conforman su funcionalidad y la integración con el sistema ML cómo se visualiza en la figura 10.

**Figura 10.***Arquitectura del Sistema ML.*

Con la arquitectura cliente-servidor que se ha definido, en segundo punto se generó el diseño lógico del funcionamiento de la aplicación ML en la WEB desarrollada y construida bajo un servidor de API REST y servidor de datos, todo integrado en un servidor WEB de alojamiento. A continuación, se presenta en la figura 11, que describe el proceso a ejecutarse del sistema de recomendación de habitabilidad:

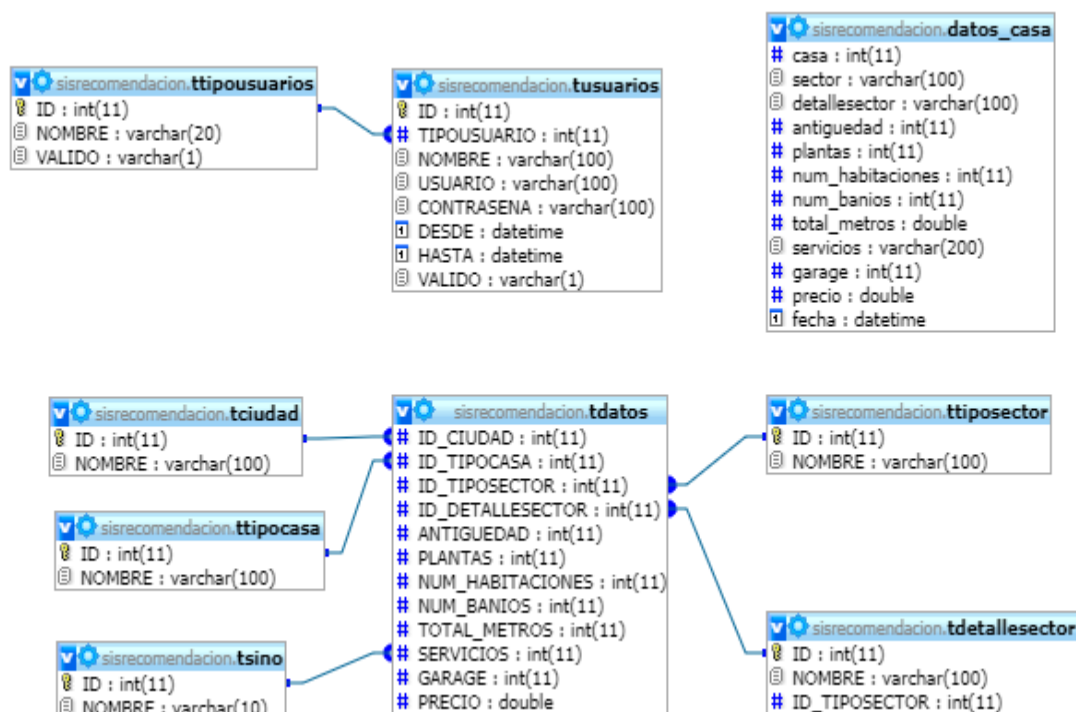
**Figura 11.***Diseño Lógico Funcional sistema ML.*

## Diseño de la Base de Datos

En la figura 12 se visualiza el modelo Entidad – Relación que se utilizó en el sistema de recomendación de habitabilidad. Al analizar el prototipo de sistema SRH que se tendrá, se añade las clases adecuadas, cada una con sus atributos.

**Figura 12.**

*Modelo Entidad – Relación.*



En la tabla 12 se describe las tablas que se crearon en la Base de Datos MYSQL.

**Tabla 12.**

*Detalle de las tablas utilizadas.*

Objeto	Fuente
DATOS_CASA	Datos Obtenidos por CSV
TCIUDAD	En base a la tabla datos_casa llenado por Pentaho
TTIPOCASA	En base a la tabla datos_casa llenado por Pentaho
TTIPOSECTOR	En base a la tabla datos_casa llenado por Pentaho

Objeto	Fuente
TDETALLESECTOR	Datos Obtenidos por CSV
TSINO	En base a la tabla datos_casa llenado por Pentaho
TTIPOUSUARIOS	Administrador
TUSUARIOS	Administrador

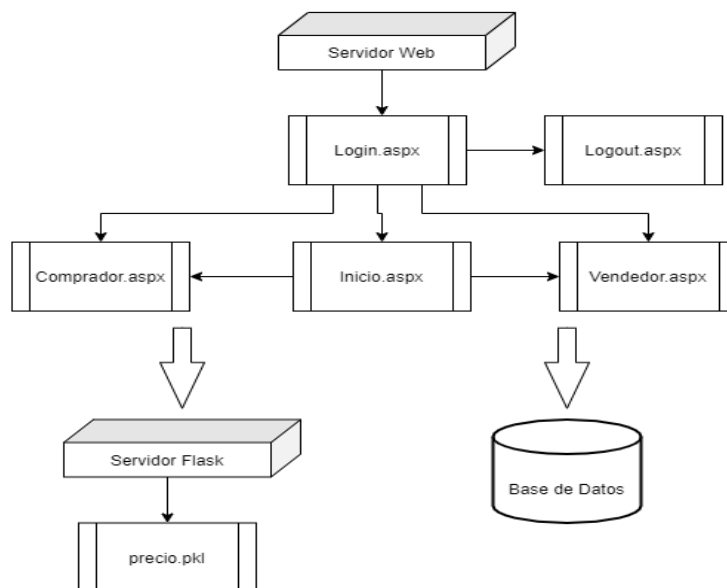
### Diseño del Prototipo

Se realizó el diagrama de componentes del SRH detallada en la figura 13, donde conjuntamente con el servidor Flask se guarda el modelo previamente generado en Python.

Flask es un microframework creado para facilitar el desarrollo de aplicaciones web en Python. Se utilizó para construir servicios web como APIs REST o aplicaciones de contenido estático (Programo Ergo, 2019).

### Figura 13.

*Diagrama de Componentes de SRH.*

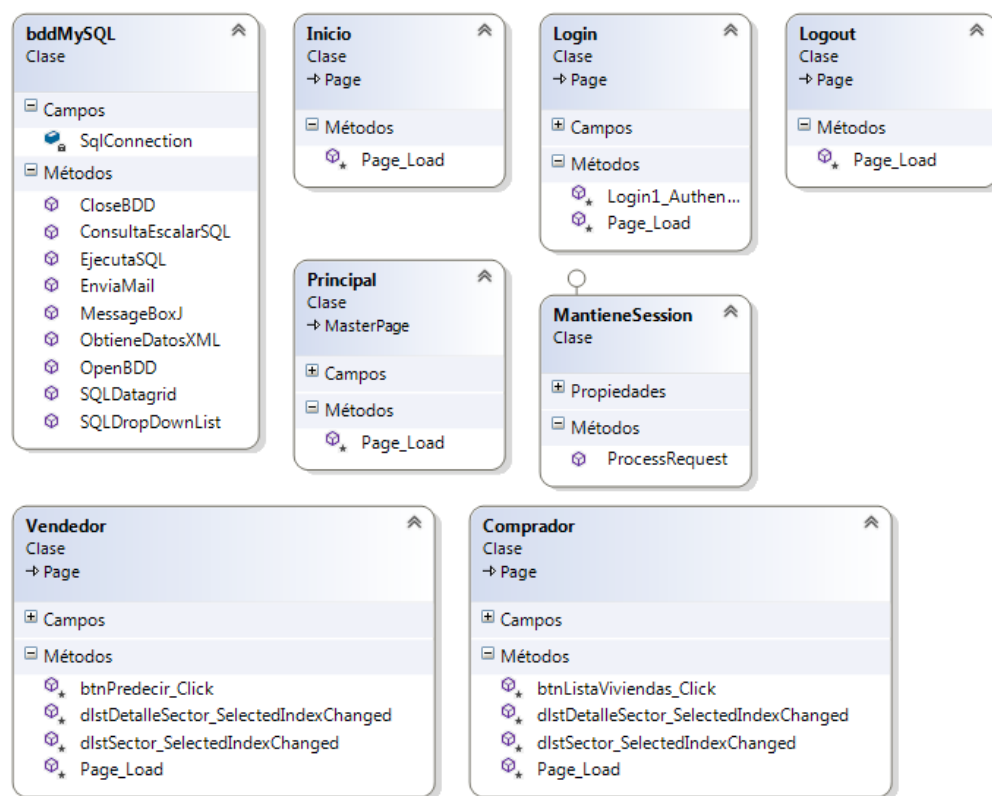


## Diagrama de Clases

En la figura 14 se observa como interactúa cada componente, el consumo del servidor web y sus ventanas de entrada, comprador, inicio, vendedor y salida. Todo se lee desde la base de datos y fusionándolo con el modelo de predicción de Machine Learning, generará un valor predicho del precio de la vivienda, el cual será visualizado en la ventana vendedor. Se utilizó las siguientes clases, que se visualizan en el aplicativo web.

**Figura 14.**

*Diagrama de clases.*



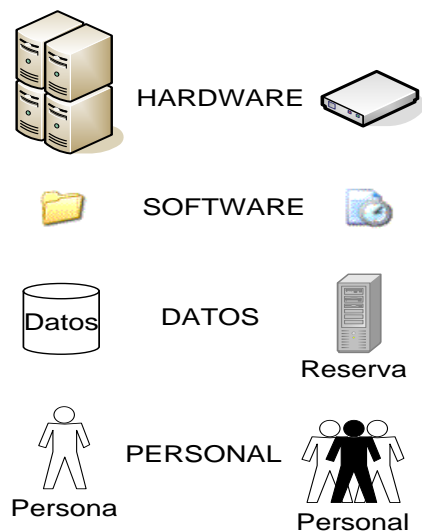
En la figura 14 se observa cómo interactúan las clases bddMySQL, Inicio, Login, Logout, Principal, Mantiene Session, Vendedor y Comprador.

## Componentes del Sistema Machine Learning

Dentro de la arquitectura del sistema SLR se tiene los componentes que se muestran en la figura 15, componentes que se utilizó para la implementación del sistema.

**Figura 15.**

*Componentes del Sistema ML*



## Requerimientos no Funcionales

Las herramientas y plataformas tecnológicas básicas para la implementación de la aplicación y ofrecer al cliente un sistema ML para la WEB son:

### Requerimientos de Software:

Sistema Operativo: Windows Server 2012

Lenguajes de Programación:

Python

ASPX

C#



Base de Datos: MySql

Servidor API REST: Flask framework de Python

Entre los componentes que se utilizó fue una base de datos con información de los clientes o vendedores, la cual se cargó en MySQL, y con el lenguaje de: Python y el intérprete de comandos: Jupyter, se procesó y trabajó la información para que sea visualizada de manera clara en el Aplicativo SRH, Figura 16.

**Figura 16.**

*Componentes y herramientas del SRH*



## Capítulo V: Conclusiones y Recomendaciones

### Conclusiones

Luego de haber ejecutado el presente proyecto de tesis se puede concluir lo siguiente:

El tipo de investigación desarrollada es analítica porque se ha buscado resolver un problema a través de modelos de regresión de análisis de datos. Además, es aplicada porque modelos ya existentes se usaron en un campo de investigación determinado. Es de carácter explicativa porque los resultados obtenidos reflejan la realidad en cuanto al comportamiento del mercado de la vivienda en Quito, debido a que se trabaja con información de viviendas y detalles de cada una y el sector de ubicación.

El modelo escogido: Se evaluó dos técnicas de Machine Learning mediante la aplicación de métricas para seleccionar el modelo que se ajuste al procesamiento de datos referente a las necesidades de la población de Quito, se trabajó con Logistic Regression y Linear Regression en el lenguaje de: Python y el intérprete de comandos: Jupyter.

Se obtuvo resultados favorables con Linear Regression por los mínimos errores presentados en las diferentes ejecuciones, con registros de un año y también seis meses.

El modelo Machine Learning está embebido en el Sistema de Recomendación de Habitabilidad para la presentación de predicción del valor de una vivienda, de tal forma que el usuario visualice la lista de viviendas en venta por sector, que se ajuste a su perfil o necesidades.

El sistema de recomendación de habitabilidad es fácil de usar y adaptable a nuevos cambios. El ciudadano que busque una vivienda en la ciudad de Quito, podrá ingresar al aplicativo y obtener las recomendaciones de viviendas en venta de acuerdo a su búsqueda de características de habitabilidad.

El sistema se encarga de emparejar el perfil del usuario con el perfil de vivienda gracias a la implementación de algoritmos Machine Learning.

Al seguir los pasos de la metodología KDD, con el uso de la herramienta Jupyter, el algoritmo de Linear Regression se integró con el Flask framework de Python, se proporcionó toda la información de los datos por medio de una interfaz de usuario; esto permitió realizar la predicción de la vivienda elegida por el usuario, con sus características de habitabilidad en la ciudad de Quito.

## **Recomendaciones**

Con los resultados obtenidos en el testeo y evaluación del modelo Machine Learning se recomienda reentrenar el modelo de algoritmo generado, una vez que empiece a estar operativo el aplicativo web, debido a que se interactuará mensualmente con las ventas realizadas.

Buscar una alternativa para que tanto el sistema informático como los algoritmos puedan ser extensibles, colocar un botón donde los usuarios alimenten al sistema con más data y tanto las variables dependiente e independiente reflejen resultados asertivos.

## Anexo

Para realizar la ejecución del SISTEMA DE RECOMENDACIÓN DE HABITABILIDAD EN LA CIUDAD DE QUITO ORIENTADO A LA PLANIFICACIÓN URBANA COGNITIVA APLICANDO TÉCNICAS DE MACHINE LEARNING

Se realizó en .NET VISUAL STUDIO, con 5 pestañas que se detallan a continuación:

**Login:** Se realiza el logeo, ingreso al sistema, debidamente validado.

**Inicio:** Detalla una descripción de lo que realiza el sistema de habitabilidad.

**Propiedades-Inmuebles:** Esta ventana puede visualizar los datos que ingresó el vendedor de la vivienda.

**Predecir Precio:** En esta ventana puede ingresar los datos de la vivienda y también puede visualizar los precios generados por el algoritmo machine learning.

**Logout:** Es necesario para salir del sistema de recomendación de habitabilidad.

Detalle de cada pestaña:

Login

Pantallas:

*Nota.* La figura representa el logeo, ingreso al sistema, debidamente validado por el Administrador.

PANTALLA:

*Nota.* La figura representa el logeo, ingreso al sistema, validación de clave que conste en la tabla de usuarios con privilegios.

PANTALLA:

**SISTEMA DE RECOMENDACIÓN HABITACIONAL**

Login Inicio **Propiedades-Inmuebles** Predecir-Precio Logout

### Comprador

**Recomendar Viviendas**

Ciudad:  Sector:  Detalle Sector:

Tipo de Vivienda:  Servicios Básicos:  Garage:

Precio Mínimo:  Precio Máximo:

Derechos Reservados Ximena Rojas © - 2021

Ingreso de datos del aplicativo:

**SISTEMA DE RECOMENDACIÓN HABITACIONAL**

Login Inicio **Propiedades-Inmuebles** Predecir-Precio Logout

### Comprador

**Recomendar Viviendas**

Ciudad:  Sector:  Detalle Sector:

Tipo de Vivienda:  Servicios Básicos:  Garage:

Precio Mínimo:  Precio Máximo:

Derechos Reservados Ximena Rojas © - 2021

*Nota.* La figura representa la ventana del Comprador, al momento está la ciudad de Quito, este campo a futuro desarrollo puede tener flexibilidad a colocar más ciudades.

## Botón listar viviendas:

### Comprador

**Recomendar Viviendas**

Ciudad:  Sector:  Detalle Sector:

Tipo de Vivienda:  Servicios Básicos:  Garage:

Precio Mínimo:  Precio Máximo:

CIUDAD	TIPO	ZONA	SECTOR	ANTIGUEDAD	PLANTAS	NUM_HABITACIONES	NUM_BANIOS	TOTAL_METROS	SERVICIOS	GARAGE	PRECIO
Quito	Casa	Norte	Batán	0	1	2	2	483	Si	1	72628
Quito	Casa	Norte	Batán	0	2	4	2	100	Si	1	72628
Quito	Casa	Norte	Batán	0	3	6	5	87	Si	1	72628
Quito	Casa	Norte	Batán	0	1	2	2	483	Si	1	72628
Quito	Casa	Norte	Batán	0	2	4	2	100	Si	1	72628
Quito	Casa	Norte	Batán	0	3	6	5	87	Si	1	72628
Quito	Casa	Norte	Batán	0	1	2	2	483	Si	1	72628
Quito	Casa	Norte	Batán	0	2	4	2	100	Si	1	72628
Quito	Casa	Norte	Batán	0	3	6	5	87	Si	1	72628
Quito	Casa	Norte	Batán	0	1	2	2	483	Si	1	72628
Quito	Casa	Norte	Batán	0	2	4	2	100	Si	1	72628
Quito	Casa	Norte	Batán	0	3	6	5	87	Si	1	72628
Quito	Casa	Norte	Batán	0	1	2	2	483	Si	1	72628
Quito	Casa	Norte	Batán	0	2	4	2	100	Si	1	72628
Quito	Casa	Norte	Batán	0	3	6	5	87	Si	1	72628
Quito	Casa	Norte	Batán	0	1	2	1	100	Si	1	79000
Quito	Casa	Norte	Batán	0	1	2	1	100	Si	1	79000
Quito	Casa	Norte	Batán	0	1	2	1	100	Si	1	79000

## Vendedor:

### SISTEMA DE RECOMENDACIÓN HABITACIONAL

[Login](#) [Inicio](#) [Propiedades-Inmuebles](#) [Predecir Precio](#) [Logout](#)

### Vendedor

**Recomendar precio venta**

Ciudad:  Sector:  Detalle Sector:

Tipo de Vivienda:  Servicios Básicos:  Garage:

Número de habitaciones:  Número de Baños:  Número de Plantas:

Años de antigüedad:  Área de la Vivienda (m<sup>2</sup>):



Norte:

 SISTEMA DE RECOMENDACIÓN HABITACIONAL

[Login](#) [Inicio](#) [Propiedades-Inmuebles](#) **Predecir-Precio** [Logout](#)

### Vendedor

**Recomendar precio venta**

Ciudad	<input type="text" value="Quito"/>	Sector	<input type="text" value="Norte"/>	Detalle Sector	<input type="text" value="Batán"/>
Tipo de Vivienda	<input type="text" value="Casa"/>	Servicios Básicos	<input type="text" value="Si"/>	Garage	<input type="text" value="1"/>
Número de habitaciones	<input type="text" value="3"/>	Número de Baños	<input type="text" value="2"/>	Número de Plantas	<input type="text" value="1"/>
Años de antigüedad	<input type="text" value="15"/>	Área de la Vivienda (m <sup>2</sup> )	<input type="text" value="96"/>	<input type="button" value="Predecir Precio"/>	<input type="text" value="102500.93"/>

Derechos Reservados Ximena Rojas © - 2021

Centro:

 SISTEMA DE RECOMENDACIÓN HABITACIONAL

[Login](#) [Inicio](#) [Propiedades-Inmuebles](#) **Predecir-Precio** [Logout](#)

### Vendedor

**Recomendar precio venta**

Ciudad	<input type="text" value="Quito"/>	Sector	<input type="text" value="Centro"/>	Detalle Sector	<input type="text" value="Santa Lucía"/>
Tipo de Vivienda	<input type="text" value="Casa"/>	Servicios Básicos	<input type="text" value="Si"/>	Garage	<input type="text" value="1"/>
Número de habitaciones	<input type="text" value="3"/>	Número de Baños	<input type="text" value="2"/>	Número de Plantas	<input type="text" value="1"/>
Años de antigüedad	<input type="text" value="15"/>	Área de la Vivienda (m <sup>2</sup> )	<input type="text" value="80"/>	<input type="button" value="Predecir Precio"/>	<input type="text" value="125611.07"/>

Derechos Reservados Ximena Rojas © - 2021

Sur:

 SISTEMA DE RECOMENDACIÓN HABITACIONAL

Login Inicio Propiedades- Inmuebles **Predecir Precio** Logout

**Vendedor**

**Recomendar precio venta**

Ciudad	<input type="text" value="Quito"/>	Sector	<input type="text" value="Sur"/>	Detalle Sector	<input type="text" value="Mena Sur"/>
Tipo de Vivienda	<input type="text" value="Casa"/>	Servicios Básicos	<input type="text" value="Si"/>	Garage	<input type="text" value="1"/>
Número de habitaciones	<input type="text" value="3"/>	Número de Baños	<input type="text" value="2"/>	Número de Plantas	<input type="text" value="1"/>
Años de antigüedad	<input type="text" value="15"/>	Área de la Vivienda (m <sup>2</sup> )	<input type="text" value="96"/>	<input type="button" value="Predecir Precio"/>	<input type="text" value="116664.12"/>

Derechos Reservados Ximena Rojas © - 2021

Valles:

 SISTEMA DE RECOMENDACIÓN HABITACIONAL

Login Inicio Propiedades- Inmuebles **Predecir Precio** Logout

**Vendedor**

**Recomendar precio venta**

Ciudad	<input type="text" value="Quito"/>	Sector	<input type="text" value="Los Valles"/>	Detalle Sector	<input type="text" value="Rumiñahui"/>
Tipo de Vivienda	<input type="text" value="Casa"/>	Servicios Básicos	<input type="text" value="Si"/>	Garage	<input type="text" value="1"/>
Número de habitaciones	<input type="text" value="2"/>	Número de Baños	<input type="text" value="2"/>	Número de Plantas	<input type="text" value="2"/>
Años de antigüedad	<input type="text" value="0"/>	Área de la Vivienda (m <sup>2</sup> )	<input type="text" value="150"/>	<input type="button" value="Predecir Precio"/>	<input type="text" value="118272.32"/>

Derechos Reservados Ximena Rojas © - 2021

*Nota.* La figura representa la ventana del Vendedor, en el campo sector, se extrae y separa por sectores: Norte, Centro, Sur y Los Valles. Toda esta data se calculó de más de 65000 datos de los lotes proporcionados por el equipo de trabajo.

En la figura representada la ventana del Vendedor, el campo: Detalle Sector: Luego de escoger el sector, el aplicativo, está validado para que automáticamente seleccione barrios de la ciudad de Quito. Y evitar así la generación de errores cuando existe tipeo de forma manual.

Los siguientes campos: Tipo de Vivienda, Servicios Básicos, Cuenta con Garage, Número de habitaciones, Número de Baños, Número de Plantas, Años de antigüedad, Área de la Vivienda, Precio Vivienda, estos son campos que ayudaran al comprador a ingresar estos datos para que pueda venderla.

El botón de gran ayuda y el plus a adicional del sistema es el botón Predecir Precio, debido a que utiliza algoritmos machine learning.

Logout:

SISTEMA DE RECOMENDACIÓN HABITACIONAL

Login Inicio Propiedades-Immuebles Predecir-Precio Logout

Iniciar sesión

Usuario

Contraseña

Login

Derechos Reservados Ximena Rojas © - 2021

La figura representa la ventana de Salida, es necesario para salir del sistema de recomendación de habitabilidad.

## Bibliografía

- Felipe Riverón , E., & Suárez Guerra, S. (2014). Recuperado el 12 de 02 de 2019, de <http://148.204.63.111/SABERv3/Repositorios/webVerArchivo/26041>
- Aggarwal, C. C. (2016). *Recommender systems (Vol. 1)*. Cham: Springer International Publishing.
- Arquitectura de BI*. (17 de 03 de 2020). Obtenido de <https://blog.bi-geek.com/arquitectura-comparativa-inmon-y-kimball/>
- BBC Mundo. (21 de 02 de 2017). *BBC News*. Obtenido de <https://www.bbc.com/mundo/noticias-39045803>
- Beltrán Martínez, B. (10 de 01 de 2020). Obtenido de <http://bbeltran.cs.buap.mx/NotasMD.pdf>
- Chimbo, J. (21 de 10 de 2019). *Scribd*. Obtenido de <https://es.scribd.com/document/431235773/Sisalima-Ortega-Fabricio-Roberto>
- Clos, J. (2018). *La Nueva Agenda Urbana*. Obtenido de <http://habitat3.org/the-new-urban-agenda>
- Cmotions*. (10 de 04 de 2017). Obtenido de <https://cmotions.nl/en/knime-vs-rapidminer/>
- Cuenca, E. T.-D. (28 de 08 de 2018). *Diario de Cuenca*. Obtenido de <https://www.eltiempo.com.ec/noticias/ecuador/4/cuantos-venezolanos-entraron-al-ecuador>
- D'Onofrio, S. M. (2018). Fuzzy reasoning in cognitive cities: an exploratory work on fuzzy analogical reasoning using fuzzy cognitive maps. *IEEE International Conference on Fuzzy Systems*.
- Díaz García, E. &. (2018). Manual de uso de Jupyter Notebook para aplicaciones docentes.
- Dupedier, E. (1995). Población y Urbanización Metropolitana. En E. Dupedier, *Población y Urbanización Metropolitana*. Quito: [http://horizon.documentation.ird.fr/exl-doc/pleins\\_textes/divers17-08/010006838.pdf](http://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers17-08/010006838.pdf).
- Ecuador, F. (15 de 01 de 2019). *Población de Quito (Ecuador) 2019 - Número actual de habitantes de Quito*. Obtenido de <http://www.forosecuador.ec/forum/ecuador/educaci%C3%B3n-y-ciencia/166821-poblaci%C3%B3n-de-quito-ecuador-2019-n%C3%BAmero-actual-de-habitantes-de-quito>
- El Universo. (02 de 09 de 2019). *El Universo*. Obtenido de <https://www.eluniverso.com/noticias/2019/09/02/nota/7500762/quito-se-acaba-pico-placa-cambio-habra-hoy-no-circula>
- El Universo*. (16 de 03 de 2020). Obtenido de <https://www.eluniverso.com/noticias/2019/09/02/nota/7500762/quito-se-acaba-pico-placa-cambio-habra-hoy-no-circula>
- Federico, M. (s.f.). *Los Sistemas Cognitivos*. Obtenido de <https://www.javerianacali.edu.co/noticias/los-sistemas-cognitivos-y-su-importancia-en-el-mundo-de-hoy>

- Felipe Riverón, E. M., & Suárez Guerra, S. (2014). Obtenido de <http://148.204.63.111/SABERv3/Repositorios/webVerArchivo/26041>
- Foros Ecuador*. (15 de 01 de 2019). Obtenido de Foros Ecuador.ec: <http://www.forosecuador.ec/forum/ecuador/educaci%C3%B3n-y-ciencia/166821-poblaci%C3%B3n-de-quito-ecuador-2019-n%C3%BAmero-actual-de-habitantes-de-quito>
- García Aquino, C. (11 de 12 de 2019). *Scribd*. Obtenido de <https://es.scribd.com/document/439274742/CAPITULO-2-PARTE1>
- gbadvisors. (27 de 11 de 2017). *gbadvisors*. Obtenido de <https://www.gb-advisors.com/es/cuadrante-de-gartner/>
- Gómez, B. R.-A.-3. (2007).
- Gomez, F. (23 de 03 de 2015). *deusto formacion*. Obtenido de <https://www.deustoformacion.com/blog/disenio-produccion-audiovisual/8-caracteristicas-imagen-digital-que-debes-conocer>
- González Arroyave, D. L. (2019). Estandarización de datos y priorización de algoritmos para un sistema de recomendación (Doctoral dissertation, Universidad EAFIT).
- González, D. (. (2018). Obtenido de González, D. (2019). Estandarización de datos y priorización de algoritmos para un sistema de recomendación (Doctoral dissertation, Tesis de maestría, Universidad EAFIT). Recuperado de <https://repository.eafit.edu.co/bitstream/handle/10784/15326/DiderL>
- González, N. V.-P. (2015).
- Guedez, A. (15 de 01 de 2020). *gbadvisors*. Obtenido de <https://www.gb-advisors.com/es/cuadrante-de-gartner/>
- H, S. (26 de 10 de 2019). *Servidor web con Flask en Python*. Obtenido de <https://www.programoergosum.es/tutoriales/servidor-web-con-flask-en-raspberry-pi/>
- Hernández, L. (12 de 05 de 2017). *Programarfacil*. Obtenido de <https://programarfacil.com/blog/vision-artificial/deteccion-de-movimiento-con-opencv-python/>
- Hu, L. S. (2014). Your neighbors affect your ratings: on geographical neighborhood influence to rating prediction. *In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval* .
- imagendigitallllllll. (01 de 04 de 2018). *imagendigitallllllll*. Obtenido de <https://imagendigitallllllll.wordpress.com/2018/04/01/imagen-digital/>
- INEC. (s.f.). *INEC*. Obtenido de Instituto Nacional de Estadísticas y Censos: <http://www.ecuadorencifras.gob.ec/banco-de-informacion/>
- Infaimon*. (15 de 03 de 2020). Obtenido de <https://blog.infaimon.com/vision-computador-soluciones-permite/>

- Kaltenrieder, P. P.-2. (2014). Applying the fuzzy analytical hierarchy process in cognitive cities. In *Proceedings of the 8th international conference on theory and practice of electronic governance*, 259-269.
- kdnuggets*. (13 de 01 de 2020). Obtenido de <http://datascience.recursos.uoc.edu/es/portal-kdnuggets/>
- KDnuggets*. (15 de 03 de 2020). Obtenido de <https://www.kdnuggets.com/2020/02/gartner-mq-2020-data-science-machine-learning.html>
- KDnuggets*. (14 de 01 de 2020). *KDnuggets*. Obtenido de <http://datascience.recursos.uoc.edu/es/portal-kdnuggets/>
- Landa, J. (16 de 03 de 2020). *Tratamientos de datos*. Obtenido de <http://fcojlanda.me/es/ciencia-de-los-datos/kdd-y-mineria-de-datos-espanol/>
- López, L. A.-r. (2020).
- Ltib*. (04 de 05 de 2010). Obtenido de <https://pytorch.org/docs/stable/torchvision/index.html>
- Martínez, J. (15 de 01 de 2019). *Data Forest*. Obtenido de <https://data-forest.co/2019/01/15/anaconda-que-es/>
- Mohammadi, M. &.-F.-1. (2018).
- Moine, J. M. (2011).
- Molina López, J., & García Herrero, J. (12 de 01 de 2020). *TÉCNICAS DE ANÁLISIS DE DATOS*. Obtenido de [http://matema.ujaen.es/jnavas/web\\_recursos/archivos/weka%20master%20recursos%20naturales/apuntesAD.pdf](http://matema.ujaen.es/jnavas/web_recursos/archivos/weka%20master%20recursos%20naturales/apuntesAD.pdf)
- Municipio del Distrito Metropolitano de Quito. (15 de 02 de 2020). Recuperado el 15 de 01 de 2019, de <http://gobiernoabierto.quito.gob.ec/wp-content/uploads/documentos/pdf/diagnosticomovilidad.pdf>
- Municipio del Distrito Metropolitano de Quito. (21 de 03 de 2020). *DIAGNÓSTICO DE LA MOVILIDAD EN EL DISTRITO METROPOLITANO DE QUITO PARA EL PLAN METROPOLITANO DE DESARROLLO TERRITORIAL (PMOT)*. Obtenido de <http://gobiernoabierto.quito.gob.ec/wp-content/uploads/documentos/pdf/diagnosticomovilidad.pdf>
- Naeem, T. (17 de 03 de 2020). *Astera*. Obtenido de <https://www.astera.com/es/type/blog/data-warehouse-concepts/>
- OpenCV. (18 de 07 de 2020). *OpenCV*. Obtenido de <https://www.filehorse.com/es/descargar-opencv/>
- Pece, A. D. (2014). *SIG participativo*. Obtenido de Revisiones de energía renovable y sostenible: [https://econpapers.repec.org/article/eeerensus/v\\_3a41\\_3ay\\_3a2015\\_3ai\\_3ac\\_3ap\\_3a162-177.htm](https://econpapers.repec.org/article/eeerensus/v_3a41_3ay_3a2015_3ai_3ac_3ap_3a162-177.htm)
- Pereira, T., Arteaga, H., Zambrano, C., Troya, H., & Pérez, A. (29 de 07 de 2016). *El proceso de descubrimiento de conocimiento en bases de datos*. Obtenido de <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=2ahUKewjsyPirp>

aHhAhViuVkkHUbFDTsQFjAEegQIBBAC&url=http%3A%2F%2Fediciones.ucc.edu.co%2Findex.php%2Fucc%2Fcatalog%2Fdownload%2F36%2F40%2F230-1&usg=AOvVaw0SaX6XN6n4ijRC\_Y4KWf41

Pérez, A. A. (2015).

Piatetsky-Shapiro, G., & Mayo, M. (13 de 01 de 2020). *KDnuggets Poll*. Obtenido de <https://www.kdnuggets.com/2016/09/poll-algorithms-used-data-scientists.html>

Prieto Valdés, M. (01 de 05 de 2010). Obtenido de <https://repository.javeriana.edu.co/bitstream/handle/10554/7034/tesis480.pdf?sequence=1&isAllowed=y>

Prieto Valdes, M. A. (2010). Algoritmo para conteo vehicular en tiempo real con base en franjas de interes. 47.

Programo Ergo, S. (26 de 10 de 2019). *Servidor web con Flask en Python*. Obtenido de <https://www.programoergosum.es/tutoriales/servidor-web-con-flask-en-raspberry-pi/>

Rahman, M. A.-b.-1. (2019).

Ray, S. (.-3. (2019). A quick review of machine learning algorithms. In *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 35-39.

Robles Aranda, Y., & R. Sotolongo, A. (12 de 01 de 2020). *Scielo*. Obtenido de [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1807-17752013000200389](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1807-17752013000200389)

Rojas X. (2021).

Romero, J. (11 de 06 de 2019). *JorgeRomero*. Obtenido de <https://jorgeromero.net/metodologias-de-mineria-de-datos/>

Sandra Zapata, Q., Alba Yauri, I., Rodrigo Huamani, A., & Jose Sulla, T. (26 de 07 de 2019). *Prediction for the Car Rental Business with Supervised Techniques*. Obtenido de [http://laccei.org/LACCEI2019-MontegoBay/full\\_papers/FP371.pdf](http://laccei.org/LACCEI2019-MontegoBay/full_papers/FP371.pdf)

Shmueli, G. B. (2019). *Data mining for business analytics: concepts, techniques and applications in Python*. John Wiley & Sons.

Sisalima Ortega, F. (21 de 10 de 2019). *SCRIBD*. Obtenido de <https://es.scribd.com/document/431235773/Sisalima-Ortega-Fabricio-Roberto>

*sistemeduca*. (26 de 03 de 2019). Obtenido de <https://sistemeduca.com/sistemas-kdd/que-es-kdd/>

studylab. (05 de 01 de 2020). *studylab*. Obtenido de <https://studylab.es/doc/5663366/series-temporales>

Tamariz, G. (11 de Julio de 2016). *GKILLCITY - Terremoto 2016*. Obtenido de <http://terremoto.gk.city/index.php/2016/07/11/viviendas-de-manabi-antes-del-terremoto/>

Thagard, P. (. (2008). *La mente. Introducción a las ciencias cognitivas (Vol. 3021)*. Katz Editores.

Turmero, P. (2020). *Monografías*. Obtenido de <https://www.monografias.com/trabajos108/problema-extraccion-automatica-conocimiento/problema-extraccion-automatica-conocimiento.shtml>

Velásquez, D. (2016). Obtenido de [https://www.academia.edu/39826268/Desarrollo\\_e\\_implementaci%C3%B3n\\_de\\_un\\_sistema\\_de\\_visión\\_artificial\\_basado\\_en\\_lenguajes\\_de\\_uso\\_libre\\_para\\_un\\_sistema\\_seleccionador\\_de\\_productos\\_de\\_un\\_centro\\_integrado\\_de\\_manufactura](https://www.academia.edu/39826268/Desarrollo_e_implementaci%C3%B3n_de_un_sistema_de_visión_artificial_basado_en_lenguajes_de_uso_libre_para_un_sistema_seleccionador_de_productos_de_un_centro_integrado_de_manufactura)

Villacís, B. C.-7. (2011). *Estadística demográfica en el Ecuador: diagnóstico y propuesta*. Quito: Instituto Nacional de Estadística y Censos.