**ANÁLISIS DE COMPATIBILIDAD DE PROPIETARIOS PARA LA SELECCIÓN DE NUEVOS COMPAÑEROS DE PISO EN EL CONJUNTO RESIDENCIAL TREBOL**

**Presentado Por:**

**Darwin Andrés Ruedas Fonseca**

**Kevin Jhoan Rojas Gualdron**

**Universidad de Pamplona**

**Sede Villa del Rosario**

**2024**

**Contenido**

[RESUMEN 4](#_Toc182507690)

[Palabras clave 5](#_Toc182507691)

[ABSTRACT 5](#_Toc182507692)

[Keywords 6](#_Toc182507693)

[INTRODUCCIÓN 6](#_Toc182507694)

[CAPÍTULO 1 PROBLEMA DE INVESTIGACIÓN 8](#_Toc182507695)

[1.1 Planteamiento del Problema 8](#_Toc182507696)

[Situación Actual 9](#_Toc182507697)

[Situación Deseada 9](#_Toc182507698)

[1.2 Pregunta de Investigación 10](#_Toc182507699)

[1.3 Objetivos 10](#_Toc182507700)

[1.3.1 Objetivo General 10](#_Toc182507701)

[1.3.2 Objetivos Específicos 10](#_Toc182507702)

[1.4 Justificación 11](#_Toc182507703)

[2. CAPÍTULO 2 MARCO TEÓRICO 11](#_Toc182507704)

[2.1 Marco referencial 11](#_Toc182507705)

[2.2 Marco conceptual 12](#_Toc182507706)

[2.3 Marco contextual 13](#_Toc182507707)

[3. CAPÍTULO 3 METODOLOGÍA 14](#_Toc182507708)

[3.1 Metodología Ágil Scrum 14](#_Toc182507709)

[Sprint 1: Configuración Inicial y Adquisición de Datos 15](#_Toc182507710)

[Sprint 2: Desarrollo de la Herramienta Interactiva 15](#_Toc182507711)

[3.2 Descripción y Origen de los Datos 17](#_Toc182507712)

[3.3 Validez de los Datos 17](#_Toc182507713)

[3.4 Técnicas de Preprocesamiento de Datos 18](#_Toc182507714)

[3.5 Evaluación e Interpretación 18](#_Toc182507715)

[4. CAPÍTULO 4: ANÁLISIS DE RESULTADOS 18](#_Toc182507716)

[4.1 Objetivo Específico 1 18](#_Toc182507717)

[Descripción de los Atributos 19](#_Toc182507718)

[Validación de los Atributos 19](#_Toc182507719)

[4.2 Objetivo Específico 2 20](#_Toc182507720)

[Codificación One-Hot y Normalización de Datos 20](#_Toc182507721)

[Cálculo de la Matriz de Similaridad 20](#_Toc182507722)

[Reescalado de la Matriz de Similaridad 21](#_Toc182507723)

[4.3 Objetivo Específico 3 21](#_Toc182507724)

[Diseño de la Interfaz de Usuario 22](#_Toc182507725)

[Implementación del Backend 22](#_Toc182507726)

[Visualización de Resultados 23](#_Toc182507727)

[4.4 Objetivo Específico 4 25](#_Toc182507728)

[Pruebas de Usabilidad 25](#_Toc182507729)

[Optimización y Mejoras 26](#_Toc182507730)

[5. CAPÍTULO 5 CONCLUSIONES Y RECOMENDACIONES 26](#_Toc182507731)

[Logros y Resultados Clave 26](#_Toc182507732)

[ivImpacto en la Comunidad de Trebol 27](#_Toc182507733)

[Lecciones Aprendidas 27](#_Toc182507734)

[Recomendaciones 28](#_Toc182507735)

[Referencias 31](#_Toc182507736)

**TABLA DE ILUSTRACIONES**

[Ilustración 1 Referencia de sistema 12](#_Toc181280187)

[Ilustración 2 Arquitectura lógica y física 13](#_Toc181280188)

[Ilustración 3 Sistema como emprendimiento Autor: 14](#_Toc181280189)

[Ilustración 5 Diagrama de secuencias Autor: 16](#_Toc181280190)

[Ilustración 6 Diagrama de flujo 17](#_Toc181280191)

[Ilustración 7 Metadatos 19](#_Toc181280192)

[Ilustración 8 Técnica de codificación one-hot en los datos 20](#_Toc181280193)

[Ilustración 9 Matriz de similaridad 21](#_Toc181280194)

[Ilustración 10 Reescalado de Matriz 21](#_Toc181280195)

[Ilustración 11 Interfaz 22](#_Toc181280196)

[Ilustración 12 Lógica 23](#_Toc181280197)

[Ilustración 13 Nivel de compatibilidad de cada nuevo compañero 24](#_Toc181280198)

[Ilustración 14 Comparativa entre compañeros 25](#_Toc181280199)

# RESUMEN

El proyecto "Análisis de Compatibilidad de Propietarios para la Selección de Nuevos Compañeros de Piso en el conjunto Trebol" se propone abordar el desafío crucial de encontrar compañeros de piso compatibles en el entorno residencial de Trebol. La convivencia armoniosa en un hogar compartido es esencial para la calidad de vida de los residentes, y este proyecto tiene como objetivo desarrollar una herramienta interactiva avanzada que utilice técnicas de análisis de similitud y aprendizaje automático para facilitar esta tarea. La herramienta permitirá a los residentes actuales ingresar su identificación y el número de nuevos compañeros que desean buscar. Para lograrlo, se empleará un conjunto de datos que contiene respuestas detalladas a una serie de preguntas sobre preferencias y hábitos de vida de los inquilinos. Estas preguntas cubrirán aspectos como horarios de sueño, preferencias de limpieza, actividades sociales, y otros factores relevantes para la convivencia diaria. Con esta información, la herramienta calculará la compatibilidad entre los propietarios actuales y los posibles nuevos compañeros mediante el análisis de similitud entre las características personales y las preferencias de convivencia de los residentes. El algoritmo de aprendizaje automático se entrenará utilizando técnicas como clustering y modelos de predicción para identificar patrones y correlaciones entre las respuestas de los inquilinos. De esta manera, se podrá generar un perfil de compatibilidad para cada posible nuevo compañero de piso. Los resultados se presentarán de manera visual a través de un gráfico de barras que mostrará el nivel de compatibilidad de cada candidato. Además, se incluirá una tabla comparativa detallada que permitirá a los usuarios evaluar las similitudes y diferencias entre los posibles compañeros seleccionados. Para asegurar que la herramienta sea intuitiva y fácil de usar, se realizarán pruebas de usabilidad con los residentes de Trebol. Estas pruebas incluirán evaluaciones de la interfaz de usuario, la claridad de las instrucciones, y la efectividad de los resultados presentados.

## Palabras clave

Convivencia, compañeros de piso, análisis de compatibilidad, herramienta interactiva, conjunto Trebol.

# ABSTRACT

The project "Compatibility Analysis for Selecting New Roommates in the Trebol Residential Complex" aims to address the crucial challenge of finding compatible roommates in the residential environment of Trebol. Harmonious coexistence in a shared home is essential for the residents' quality of life, and this project aims to develop an advanced interactive tool that uses similarity analysis and machine learning techniques to facilitate this task. The tool will allow current residents to enter their identification and the number of new roommates they wish to find. To achieve this, a dataset containing detailed responses to a series of questions about tenants' preferences and living habits will be employed. These questions will cover aspects such as sleep schedules, cleaning preferences, social activities, and other factors relevant to daily coexistence. With this information, the tool will calculate the compatibility between current residents and potential new roommates by analyzing the similarity between personal characteristics and living preferences. The machine learning algorithm will be trained using techniques such as clustering and prediction models to identify patterns and correlations among the tenants' responses. This way, a compatibility profile for each potential new roommate can be generated. The results will be presented visually through a bar chart showing the compatibility level of each candidate. Additionally, a detailed comparison table will be included to allow users to evaluate the similarities and differences among the selected potential roommates. To ensure the tool is intuitive and easy to use, usability tests will be conducted with Trebol residents. These tests will include evaluations of the user interface, the clarity of instructions, and the effectiveness of the presented results. User feedback will also be collected to make continuous adjustments and improvements to the tool. User satisfaction will be a key metric to validate the project's effectiveness.

## Keywords

Coexistence, roommates, compatibility analysis, interactive tool, Trebol set.

# INTRODUCCIÓN

La búsqueda de compañeros de piso es una tarea que puede resultar desafiante y crucial para una experiencia de vida compartida satisfactoria. En muchas ocasiones, la convivencia armoniosa depende en gran medida de la compatibilidad entre los habitantes, abarcando desde sus horarios y hábitos diarios hasta sus preferencias de entretenimiento y estilo de vida. Sin embargo, encontrar personas con las que se pueda convivir cómodamente puede ser un proceso complejo y en ocasiones impredecible. En respuesta a esta necesidad, surge la iniciativa de desarrollar una herramienta interactiva que facilite la selección de compañeros de piso en el conjunto residencial Trebol.

Este proyecto tiene como objetivo principal desarrollar una herramienta innovadora basada en técnicas avanzadas de análisis de similitud y aprendizaje automático, con el fin de proporcionar recomendaciones precisas y personalizadas para la selección de nuevos compañeros de piso. La herramienta se centrará en identificar la compatibilidad entre los propietarios actuales y potenciales, considerando una amplia gama de atributos y preferencias. Se espera que esta herramienta no solo simplifique el proceso de selección, sino que también contribuya a mejorar la calidad de vida de los residentes en el conjunto residencial Trebol, promoviendo relaciones positivas y una convivencia satisfactoria.

El proyecto se llevará a cabo siguiendo una metodología ágil Scrum, lo que permitirá un desarrollo iterativo y colaborativo, adaptándose de manera efectiva a los cambios en los requisitos y garantizando la entrega de valor de manera incremental. Además, se asegurará la validez y la calidad de los datos utilizados, así como la aplicación de técnicas de preprocesamiento para garantizar la precisión y la eficacia del análisis de compatibilidad. En última instancia, la evaluación del proyecto se centrará en la usabilidad y la satisfacción del usuario, así como en la interpretación de los resultados obtenidos, con el objetivo de proporcionar una solución integral y efectiva para la selección de compañeros de piso en el conjunto residencial Trebol.

Este proyecto no solo busca abordar la problemática de encontrar compañeros de piso compatibles, sino que también aspira a fomentar un ambiente de convivencia saludable y positivo en el conjunto residencial Trebol. La herramienta interactiva desarrollada no solo proporcionará recomendaciones basadas en datos objetivos, sino que también promoverá la comunicación y la transparencia entre los residentes al facilitar la discusión sobre preferencias y expectativas de convivencia. Asimismo, se espera que esta iniciativa sirva como ejemplo de aplicación práctica de técnicas avanzadas de análisis de datos en la mejora de la calidad de vida en entornos residenciales compartidos. En este contexto, el presente informe proporcionará una visión detallada del proceso de desarrollo, desde la recopilación y preprocesamiento de datos hasta la evaluación y validación de la herramienta final, con el objetivo de ofrecer una solución efectiva y sostenible para la selección de compañeros de piso en el conjunto Trebol.

La investigación sobre la convivencia en hogares compartidos y la selección de compañeros de piso ha sido explorada por varios autores en el campo de la psicología social y la ciencia de datos. Por ejemplo, Festinger, Schachter y Back (1950) en su estudio sobre la proximidad y las relaciones interpersonales destacan la importancia de la compatibilidad en la convivencia diaria. Además, investigaciones más recientes como las de Hastie, Tibshirani y Friedman (2009) en "The Elements of Statistical Learning" han demostrado el potencial de las técnicas de aprendizaje automático para analizar y predecir la compatibilidad entre individuos en diversos contextos. En el ámbito de la usabilidad y la experiencia del usuario, Nielsen (1993) en "Usability Engineering" proporciona principios fundamentales para el diseño de herramientas interactivas efectivas y fáciles de usar. Además, Kaufman y Rousseeuw (2009) en "Finding Groups in Data: An Introduction to Cluster Analysis" ofrecen una comprensión profunda de los métodos de clustering y análisis de datos que pueden ser aplicados para identificar patrones de compatibilidad entre individuos.

# CAPÍTULO 1 PROBLEMA DE INVESTIGACIÓN

## Planteamiento del Problema

La convivencia en hogares compartidos puede ser desafiante debido a las diferencias en los estilos de vida, preferencias y hábitos de los residentes. La selección de compañeros de piso compatibles es crucial para promover un ambiente de convivencia saludable y positivo. Sin embargo, encontrar compañeros con los que se pueda convivir armoniosamente puede resultar complicado y llevar a situaciones de conflicto y malestar. El problema principal radica en la dificultad para identificar la compatibilidad entre los propietarios actuales y los posibles nuevos compañeros de piso. Esto puede deberse a la falta de información o a la subjetividad en la evaluación de la compatibilidad. Además, el proceso de selección tradicionalmente ha sido manual y basado en criterios superficiales, lo que puede llevar a decisiones subóptimas. Para abordar este problema, se propone desarrollar una herramienta interactiva basada en técnicas de análisis de similitud y aprendizaje automático. Esta herramienta permitirá a los residentes del conjunto Trebol realizar búsquedas personalizadas de compañeros de piso compatibles, mejorando así la calidad de vida y promoviendo relaciones positivas en el hogar compartido.

# Situación Actual

En la situación actual en el conjunto residencial Trebol, los residentes que buscan nuevos compañeros de piso se enfrentan a varios desafíos. El proceso de selección se realiza de forma manual y basado en criterios superficiales, como la apariencia o la primera impresión, lo que puede llevar a decisiones subóptimas. Además, la falta de información detallada sobre las preferencias y hábitos de los posibles compañeros dificulta la identificación de la compatibilidad. Esto puede resultar en situaciones de convivencia incómodas o conflictivas, lo que afecta negativamente la calidad de vida de los residentes y el ambiente en el hogar compartido.

# Situación Deseada

La situación deseada es aquella en la que los residentes del conjunto Trebol cuentan con una herramienta interactiva avanzada que les permite realizar búsquedas personalizadas de compañeros de piso compatibles. Esta herramienta utiliza técnicas de análisis de similitud y aprendizaje automático para calcular la compatibilidad entre los propietarios actuales y los posibles nuevos compañeros. Los residentes pueden ingresar su información y preferencias, y la herramienta proporciona resultados detallados y visualmente intuitivos sobre la compatibilidad de cada posible compañero. Como resultado, los residentes pueden tomar decisiones más informadas y encontrar compañeros de piso con los que puedan convivir armoniosamente, promoviendo así un ambiente de convivencia saludable y positivo en el hogar compartido.

## Pregunta de Investigación

¿Cómo podemos utilizar eficazmente el análisis de similitud para mejorar la búsqueda de compañeros de piso compatibles?

## Objetivos

### Objetivo General

Desarrollar una herramienta interactiva para facilitar la selección de nuevos compañeros de piso en el conjunto Trebol, utilizando técnicas de análisis de similitud y aprendizaje automático para mejorar la convivencia y la calidad de vida en hogares compartidos.

### Objetivos Específicos

* Identificar los atributos relevantes de los propietarios actuales del piso compartido en el conjunto Trebol.
* Realizar un análisis detallado de similitud entre los inquilinos actuales y potenciales, considerando sus características personales y preferencias de convivencia.
* Diseñar una herramienta interactiva utilizando Python, Streamlit y bibliotecas de visualización de datos (Pandas, Seaborn, Matplotlib, Plotly) para facilitar la búsqueda de compañeros de piso compatibles.
* Evaluar y validar la efectividad de la herramienta interactiva mediante pruebas de usabilidad y satisfacción del usuario.

## Justificación

La convivencia en un hogar compartido puede ser una experiencia enriquecedora, pero también puede plantear desafíos significativos si no se aborda adecuadamente la selección de compañeros de piso. Los conflictos interpersonales y las diferencias en los estilos de vida pueden generar tensiones y afectar negativamente la calidad de vida de los residentes. Por lo tanto, contar con una herramienta que facilite este proceso de selección es esencial para promover un ambiente de convivencia saludable y positivo. La herramienta propuesta aprovecha la capacidad del análisis de similitud y el aprendizaje automático para evaluar de manera exhaustiva la compatibilidad entre los potenciales compañeros de piso.

Al considerar una amplia gama de atributos y preferencias, desde horarios hasta hábitos de entretenimiento, esta herramienta puede proporcionar recomendaciones personalizadas que maximicen las probabilidades de una convivencia armoniosa. La implementación de esta herramienta en formato de aplicación web interactiva garantiza su accesibilidad y facilidad de uso para los usuarios. Esto les permite realizar búsquedas personalizadas, explorar diferentes opciones y tomar decisiones informadas sobre la selección de compañeros de piso. En última instancia, se espera que esta herramienta contribuya a mejorar la calidad de vida de los residentes en el conjunto residencial Trebol, promoviendo relaciones positivas y una convivencia satisfactoria.

# CAPÍTULO 2 MARCO TEÓRICO

## Marco referencial

La selección adecuada de compañeros de piso se ha estudiado en diferentes contextos, incluyendo investigaciones sobre psicología social, sociología y gestión de viviendas compartidas. En particular, se han identificado varios factores clave que influyen en la compatibilidad entre individuos, tales como el ritmo de vida, los intereses personales, los hábitos de limpieza y la tolerancia hacia diferentes estilos de vida. En el ámbito tecnológico, el análisis de similitud y el aprendizaje automático han emergido como herramientas poderosas para abordar problemas de selección y recomendación. Estas técnicas se han aplicado con éxito en diversos campos, desde la recomendación de productos en comercio electrónico hasta la selección de personal en recursos humanos. En este proyecto, se busca adaptar estas metodologías para mejorar la selección de compañeros de piso, utilizando un enfoque basado en datos y algoritmos de aprendizaje automático. En la siguiente ilustración 1 se puede observar la referencia de sistema.

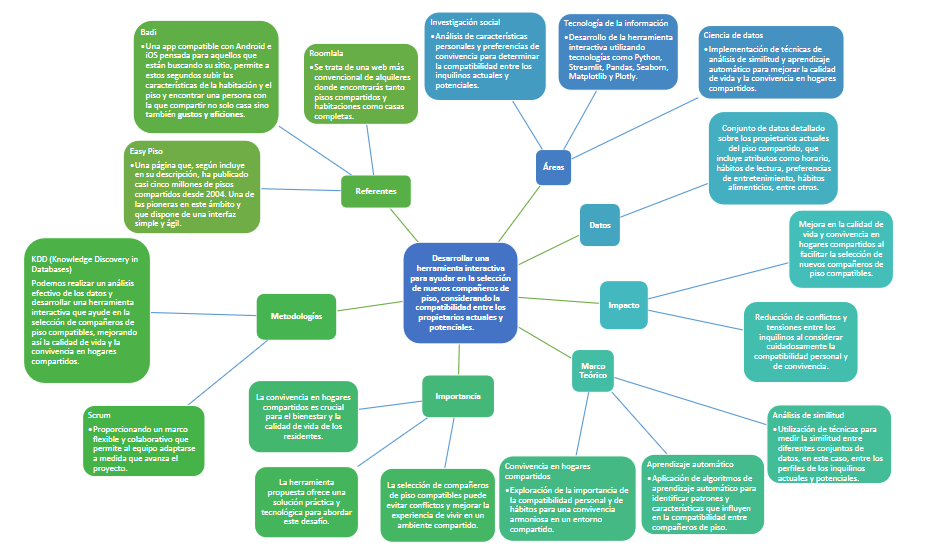


Ilustración 1 Referencia de sistema

## Marco conceptual

**Compatibilidad:** La medida en que dos o más individuos pueden convivir de manera armoniosa, basada en una serie de atributos y preferencias personales. La compatibilidad se evalúa a través de la similitud en factores como horarios, hábitos de limpieza, preferencias de entretenimiento y alimentación, entre otros.   
**Análisis de Similitud:** Técnica utilizada para medir la semejanza entre diferentes entidades (en este caso, inquilinos). Se utiliza para identificar a los compañeros de piso potencialmente más compatibles con los actuales residentes.

**Aprendizaje Automático (Machine Learning):** Subcampo de la inteligencia artificial que se enfoca en el desarrollo de algoritmos que permiten a las computadoras aprender de datos y hacer predicciones o decisiones basadas en esos datos. En este proyecto, se utiliza para analizar y predecir la compatibilidad entre inquilinos.  
**One-Hot Encoding:** Técnica de preprocesamiento de datos que convierte variables categóricas en un formato que puede ser utilizado en modelos de aprendizaje automático. Cada categoría se representa como un vector binario.  
**Matriz de Similaridad:** Matriz que contiene las medidas de similitud entre cada par de inquilinos en el conjunto de datos. Se utiliza para identificar los inquilinos más compatibles.   
**Streamlit:** Framework de Python utilizado para la creación rápida de aplicaciones web interactivas y visualizaciones de datos. Permite desarrollar interfaces de usuario de manera sencilla y eficiente. En la siguiente ilustración 2 se puede observar la arquitectura lógica y física del sistema.

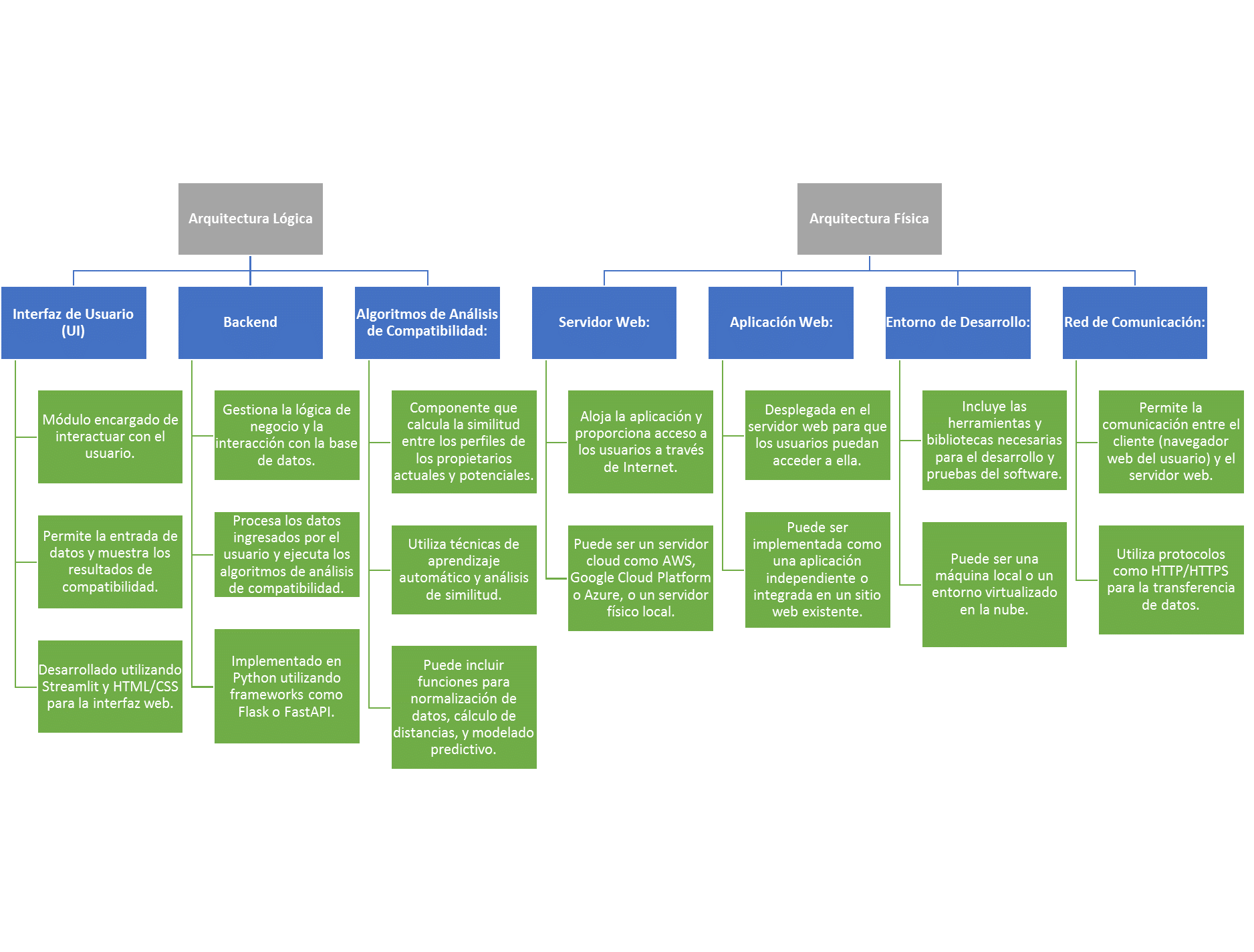


Ilustración 2 Arquitectura lógica y física

## Marco contextual

Una comunidad que busca mejorar la calidad de vida de sus residentes a través de la selección adecuada de compañeros de piso. Trebol se enfrenta a desafíos comunes en comunidades de viviendas compartidas, donde la diversidad de estilos de vida y preferencias puede llevar a conflictos y malestar. La herramienta desarrollada en este proyecto está diseñada para ser una solución práctica y accesible para los residentes de Trebol. Utilizando técnicas avanzadas de análisis de similitud y aprendizaje automático, la herramienta proporciona recomendaciones personalizadas que ayudan a los residentes a encontrar compañeros de piso compatibles. Esto no solo mejora la convivencia, sino que también promueve un ambiente más positivo y cohesionado en la comunidad. La implementación de esta herramienta en un formato de aplicación web interactiva garantiza su accesibilidad y facilidad de uso, permitiendo a los residentes realizar búsquedas personalizadas y tomar decisiones informadas sobre la selección de compañeros de piso. En la siguiente ilustración 3 podemos observar el sistema como emprendimiento.

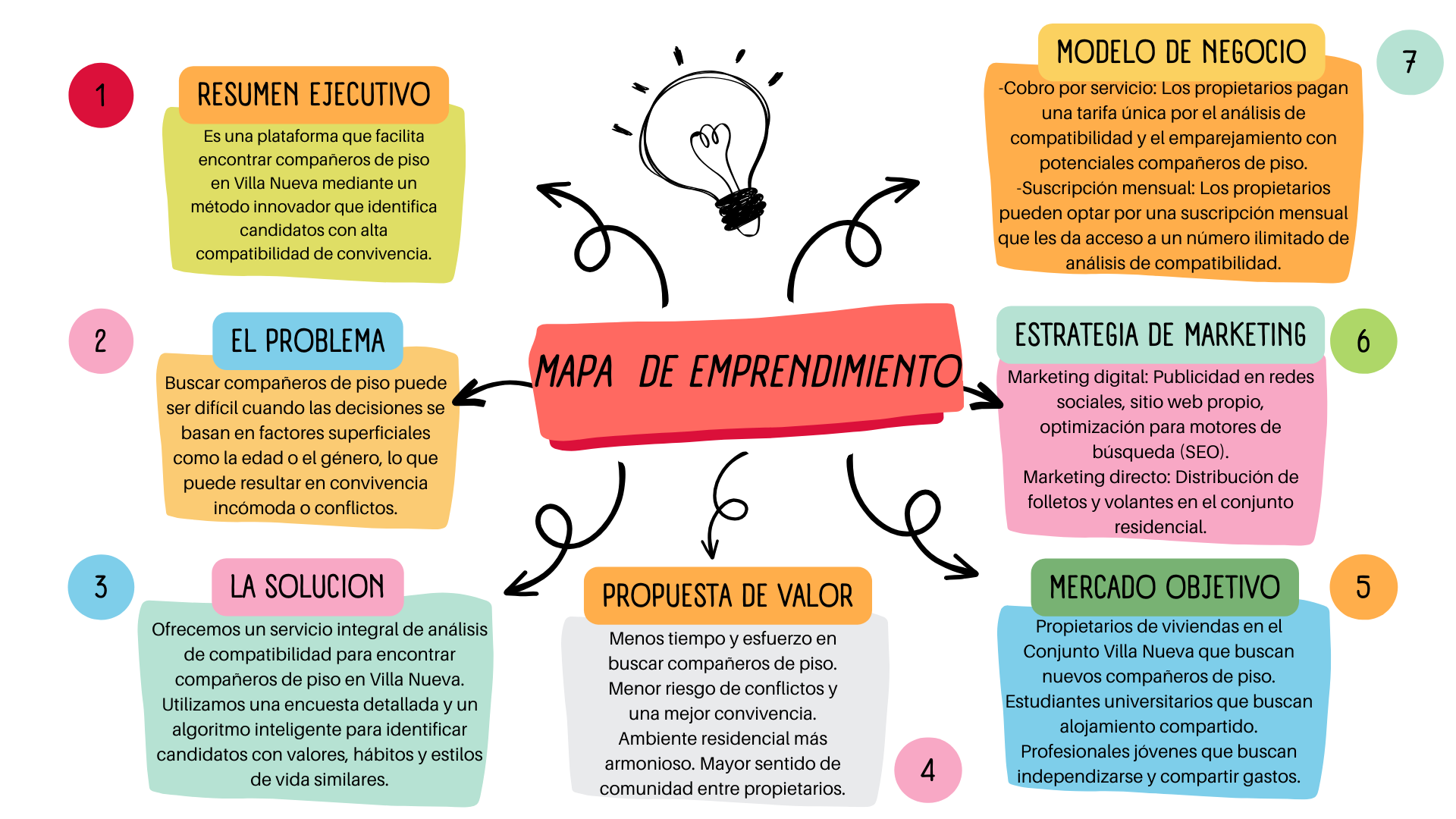


Ilustración 3 Sistema como emprendimiento Autor:

# CAPÍTULO 3 METODOLOGÍA

## Metodología Ágil Scrum

Para el desarrollo de este proyecto, se adoptó la metodología ágil Scrum, que es conocida por su enfoque iterativo e incremental. Se organizó en sprints de trabajo, cada uno con una duración fija y objetivos claros. Durante la planificación de cada sprint, se definieron las tareas a realizar. Al final de cada sprint, se llevará a cabo una revisión para demostrar el trabajo completado y recopilar retroalimentación, lo que permitirá ajustar y mejorar continuamente el producto. Esta metodología ágil facilitara la adaptación a los cambios en los requisitos del proyecto y la entrega iterativa de valor a los usuarios.

### Sprint 1: Configuración Inicial y Adquisición de Datos

**Objetivo del Sprint:** Establecer el entorno de desarrollo, adquirir y preparar los datos necesarios para el análisis de compatibilidad.

**Tareas:**

* Configurar el repositorio de código y el ambiente de desarrollo.
* Familiarizarse con las herramientas y tecnologías necesarias, como Streamlit, Pandas y las bibliotecas de visualización de datos.
* Adquirir y explorar los datos del conjunto residencial Trebol, incluyendo el archivo de encuestas (metadatos.xlsx) y el conjunto de datos (dataset\_inquilinos.csv).
* Realizar la limpieza inicial de los datos para eliminar valores atípicos y datos faltantes.
* Investigar y planificar las técnicas de preprocesamiento de datos necesarias, como la codificación one-hot y la normalización de datos.

**Resultados Esperados:**

* Ambiente de desarrollo configurado y listo para trabajar.
* Datos adquiridos y limpios, listos para su procesamiento.
* Planificación detallada de las técnicas de preprocesamiento de datos a utilizar en los próximos sprints.

### Sprint 2: Desarrollo de la Herramienta Interactiva

**Objetivo del Sprint:** Desarrollar la interfaz de usuario interactiva para la herramienta de selección de compañeros de piso, teniendo en cuenta nuestro diagrama de secuencias observado en la siguiente ilustración 5 “Diagrama de secuencias".

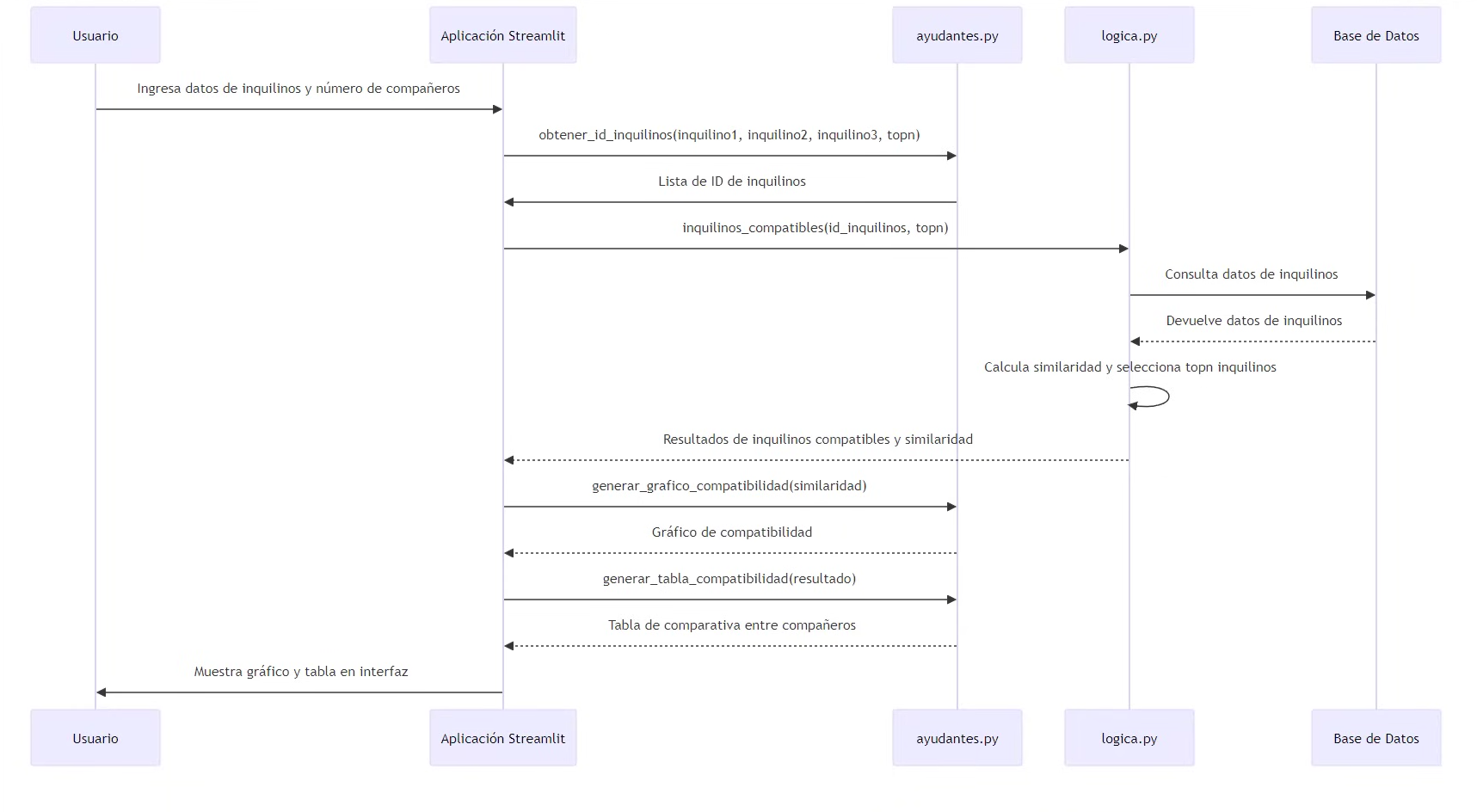


Ilustración 5 Diagrama de secuencias Autor:

**Tareas:**

* Diseñar la interfaz de usuario utilizando Streamlit, con entradas para los identificadores de los propietarios actuales y el número de nuevos compañeros a buscar.
* Implementar la lógica de procesamiento de datos en el backend, incluyendo la función para calcular la compatibilidad entre los inquilinos.
* Integrar las funciones de generación de gráficos y tablas de compatibilidad en la interfaz de usuario.
* Realizar pruebas de integración para garantizar el funcionamiento correcto de la herramienta.
* Recopilar y documentar los comentarios sobre la interfaz de usuario y las funcionalidades implementadas.

**Resultados Esperados:**

* Herramienta interactiva funcional y lista para su implementación.
* Pruebas de integración realizadas y documentadas.
* Retroalimentación del equipo recopilada para futuras mejoras.

Su funcionamiento será de manera intuitiva el usuario ingresara los datos del propietario el cual ya está viviendo en el apartamento y se ingresara cuantos nuevos propietarios serán recibidos después el sistema comparara los resultados y nos dará el análisis de los mejores compañeros de apartamento, podemos observar el proceso en la siguiente ilustración 6 “Diagrama de flujo” .

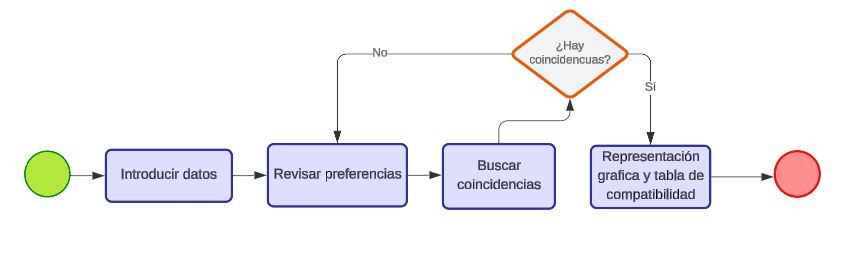


Ilustración 6 Diagrama de flujo

## Descripción y Origen de los Datos

Los datos utilizados en este proyecto provienen de un conjunto de encuestas realizadas a los inquilinos del conjunto residencial Trebol. Estas encuestas, contenidas en el archivo "metadatos.xlsx", consisten en una serie de preguntas diseñadas para recopilar información relevante sobre las preferencias y características de los inquilinos actuales y potenciales. Las respuestas a estas preguntas se recopilaron y se organizaron en el archivo "dataset\_inquilinos.csv", que sirvió como fuente de datos para el análisis de compatibilidad. Es importante destacar que los datos fueron recopilados de manera ética y cumpliendo con las normativas de privacidad y protección de datos.

## Validez de los Datos

Para garantizar la validez de los datos utilizados en este proyecto, se llevaron a cabo varios procesos. En primer lugar, se realizó una revisión exhaustiva de las encuestas y los métodos de recopilación de datos para asegurar la consistencia y la fiabilidad de la información. Además, se llevaron a cabo controles de calidad para identificar y corregir posibles errores o inconsistencias en los datos. Se verificó la integridad de los datos mediante la comparación de los resultados con fuentes externas y la validación cruzada entre diferentes conjuntos de datos. También se consideró la representatividad de la muestra y se tuvieron en cuenta posibles sesgos en la recopilación de datos para evitar conclusiones erróneas.

## Técnicas de Preprocesamiento de Datos

Antes de realizar el análisis de compatibilidad, se aplicaron diversas técnicas de preprocesamiento de datos para garantizar la calidad y la adecuación de los datos. Esto incluyó la limpieza de datos para eliminar valores atípicos, datos faltantes y errores de entrada. Se realizó una normalización de los datos para asegurar que todas las características tuvieran la misma escala y contribución al análisis. Además, se llevó a cabo una codificación one-hot para convertir variables categóricas en variables numéricas que pudieran ser utilizadas en algoritmos de aprendizaje automático.

## Evaluación e Interpretación

La evaluación del proyecto se realizó mediante pruebas exhaustivas de usabilidad y satisfacción del usuario, así como mediante la validación de los resultados obtenidos. Se realizaron pruebas de usuario con un grupo piloto para evaluar la funcionalidad y la facilidad de uso de la herramienta interactiva desarrollada. Se recopiló retroalimentación y se realizaron ajustes según sea necesario para mejorar la experiencia del usuario. Además, se compararon los resultados del análisis de compatibilidad con las expectativas y los criterios establecidos para evaluar su validez y utilidad en la selección de compañeros de piso. La interpretación de los resultados se realizó considerando el contexto del proyecto y las necesidades de los usuarios, con el objetivo de proporcionar recomendaciones precisas y personalizadas para mejorar la convivencia y la calidad de vida en hogares compartidos.

# CAPÍTULO 4: ANÁLISIS DE RESULTADOS

## Objetivo Específico 1

Identificar los atributos relevantes de los propietarios actuales del piso compartido en el conjunto Trebol

Para comenzar el desarrollo de la herramienta, fue esencial identificar los atributos que definen la compatibilidad entre los inquilinos. Estos atributos fueron obtenidos a partir de una encuesta detallada realizada a los propietarios actuales, contenida en el archivo metadatos.xlsx. A continuación, se describen los principales hallazgos:

### Descripción de los Atributos

Los atributos seleccionados incluyen horarios de actividad, biorritmo, nivel educativo, hábitos de lectura, preferencias de entretenimiento (como animación y cine), la posesión de mascotas, habilidades culinarias, práctica de deportes, tipo de dieta, hábito de fumar, frecuencia de visitas, nivel de orden, tipo de música, volumen de la música, actividades ideales para un día perfecto, y la práctica de instrumentos musicales. Como lo podemos observar en la siguiente ilustración 7 “Metadatos”.

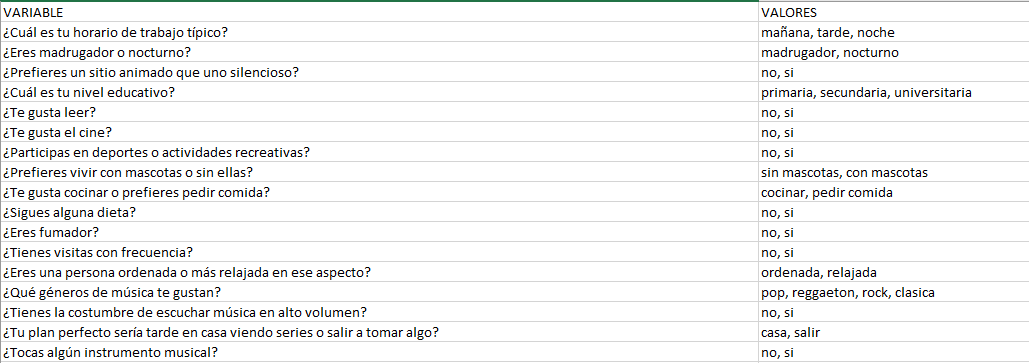


Ilustración 7 Metadatos

### Validación de los Atributos

Estos atributos fueron validados a través de un proceso de revisión y limpieza de datos, asegurando que cada entrada fuera coherente y completa. Se eliminaron valores atípicos y datos faltantes, lo cual permitió que el conjunto de datos fuera representativo y confiable. Este proceso de validación garantizó que los atributos utilizados fueran los más relevantes y precisos para el análisis de compatibilidad.

## Objetivo Específico 2

Realizar un análisis detallado de similitud entre los inquilinos actuales y potenciales

Una vez identificados y validados los atributos relevantes, el siguiente paso fue realizar un análisis detallado de similitud entre los inquilinos actuales y potenciales. Para ello, se implementaron diversas técnicas de procesamiento y análisis de datos.

### Codificación One-Hot y Normalización de Datos

Se utilizó la técnica de codificación one-hot para transformar las variables categóricas en una representación numérica adecuada para el análisis. Este proceso permitió convertir atributos cualitativos en un formato que facilita el cálculo de similitudes. Además, se aplicó la normalización de datos para asegurar que todas las características tuvieran la misma escala, evitando que algunas variables dominaran el análisis debido a diferencias en la magnitud de sus valores. Como podemos observar en la siguiente ilustración “Técnica de codificación one-hot en los datos", la codificación one-hot se utiliza para convertir las variables categóricas en una forma que se pueda utilizar para cálculos numéricos. En este proyecto, se utilizó la biblioteca “OneHotEncoder” de “sklearn” para este propósito. Podemos observar la ilustración 8 técnica de codificación one-hot en los datos.

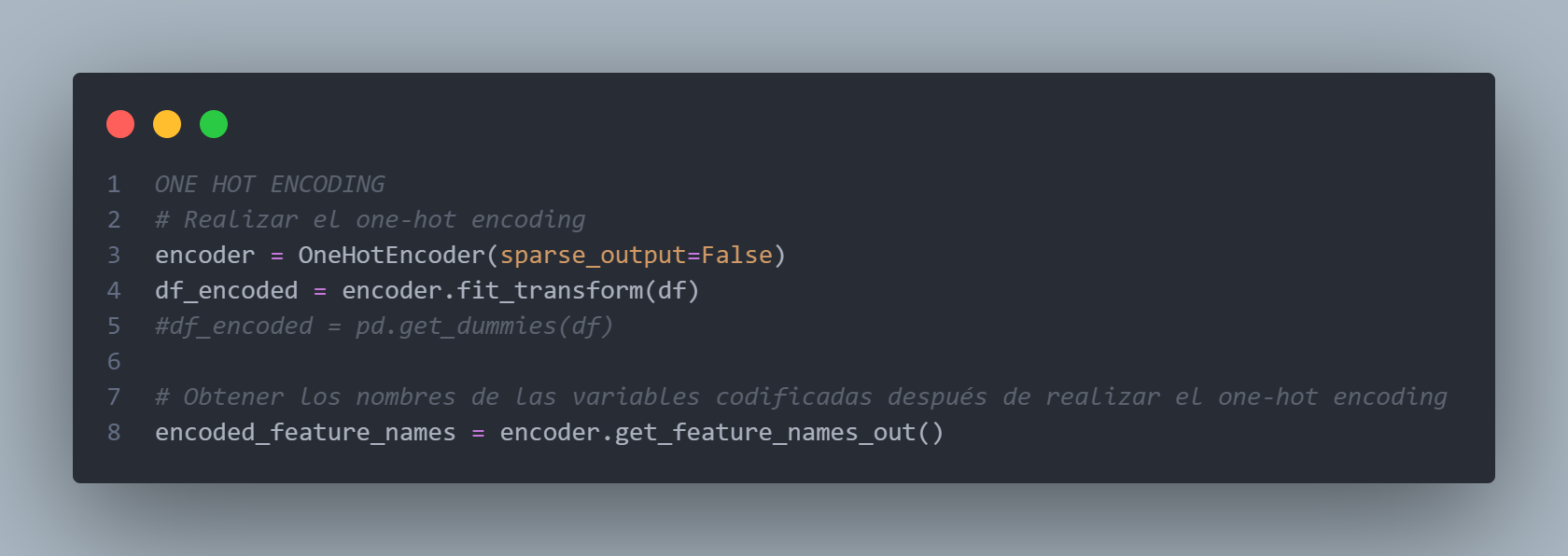


Ilustración 8 Técnica de codificación one-hot en los datos

### Cálculo de la Matriz de Similaridad

Con los datos preprocesados, se calculó una matriz de similaridad utilizando el producto punto entre los vectores de características de los inquilinos. Esta matriz permitió medir la similitud entre cada par de inquilinos, lo que fue crucial para identificar a los compañeros de piso más compatibles. Podemos observar la implementación en el software en la siguiente ilustración “Matriz de Similaridad”, esto da una medida de cuán similares son los inquilinos entre sí. Podemos observar la codificación en el sistema en la siguiente ilustración 9 Matriz de similaridad.

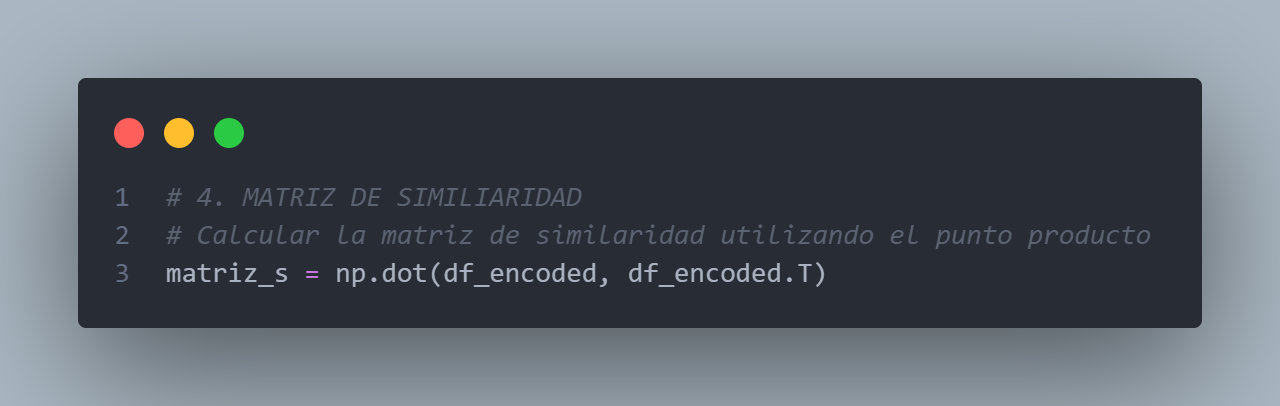


Ilustración 9 Matriz de similaridad

### Reescalado de la Matriz de Similaridad

Para mejorar la interpretabilidad de los resultados, se reescaló la matriz de similaridad a un rango de -100 a 100. Este reescalado facilitó la identificación de los inquilinos más y menos compatibles, proporcionando una medida clara y comprensible de la similitud entre ellos. Podemos observar la implementación en la siguiente ilustración 10 “Reescalado de Matriz” .

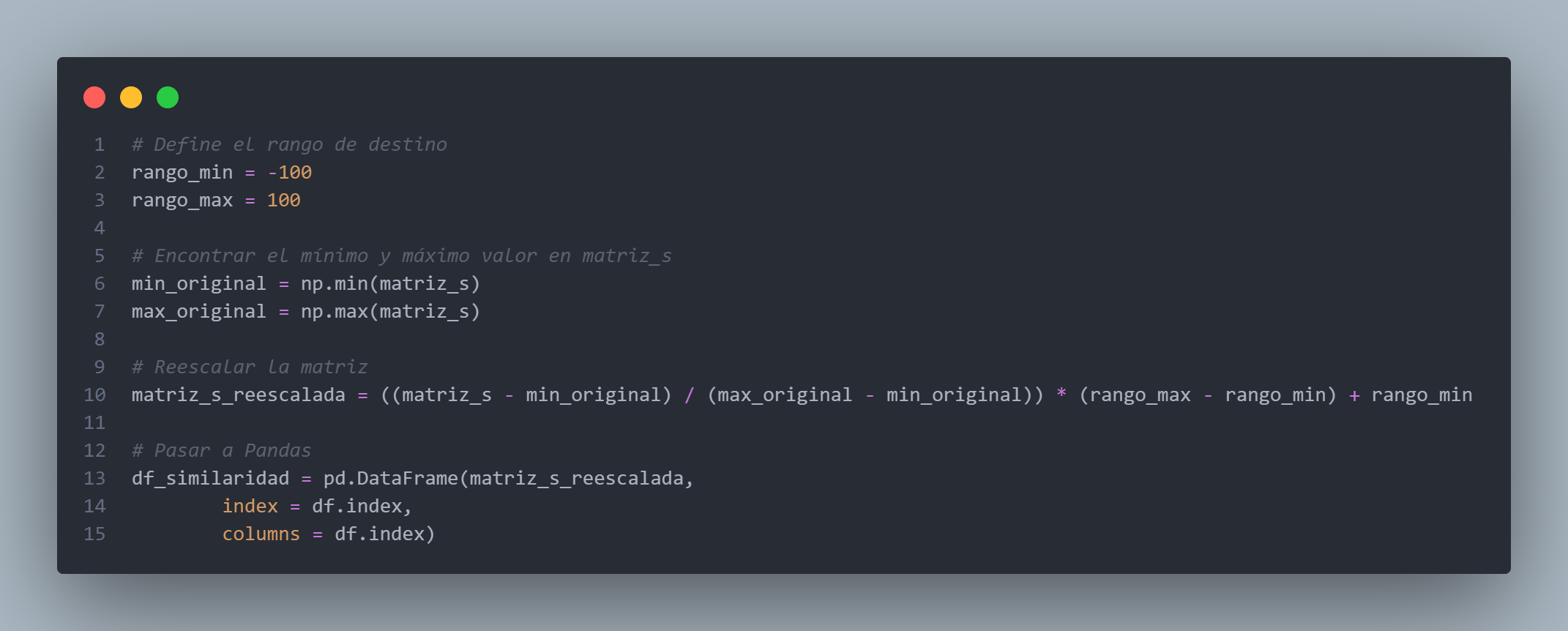


Ilustración 10 Reescalado de Matriz

## Objetivo Específico 3

Desarrollar una herramienta interactiva utilizando Python, Streamlit y bibliotecas de visualización de datos

El desarrollo de la herramienta interactiva fue un componente clave del proyecto, proporcionando una interfaz fácil de usar para los propietarios en la búsqueda de nuevos compañeros de piso.

### Diseño de la Interfaz de Usuario

Utilizando Streamlit, se creó una interfaz de usuario intuitiva que permite a los propietarios ingresar los nombres de los inquilinos actuales y el número de nuevos compañeros que desean buscar. La interfaz incluye campos de entrada y botones de acción claramente etiquetados, facilitando la interacción del usuario con la herramienta. Podemos observar la interfaz gráfica en la siguiente ilustración 11 “Interfaz”.

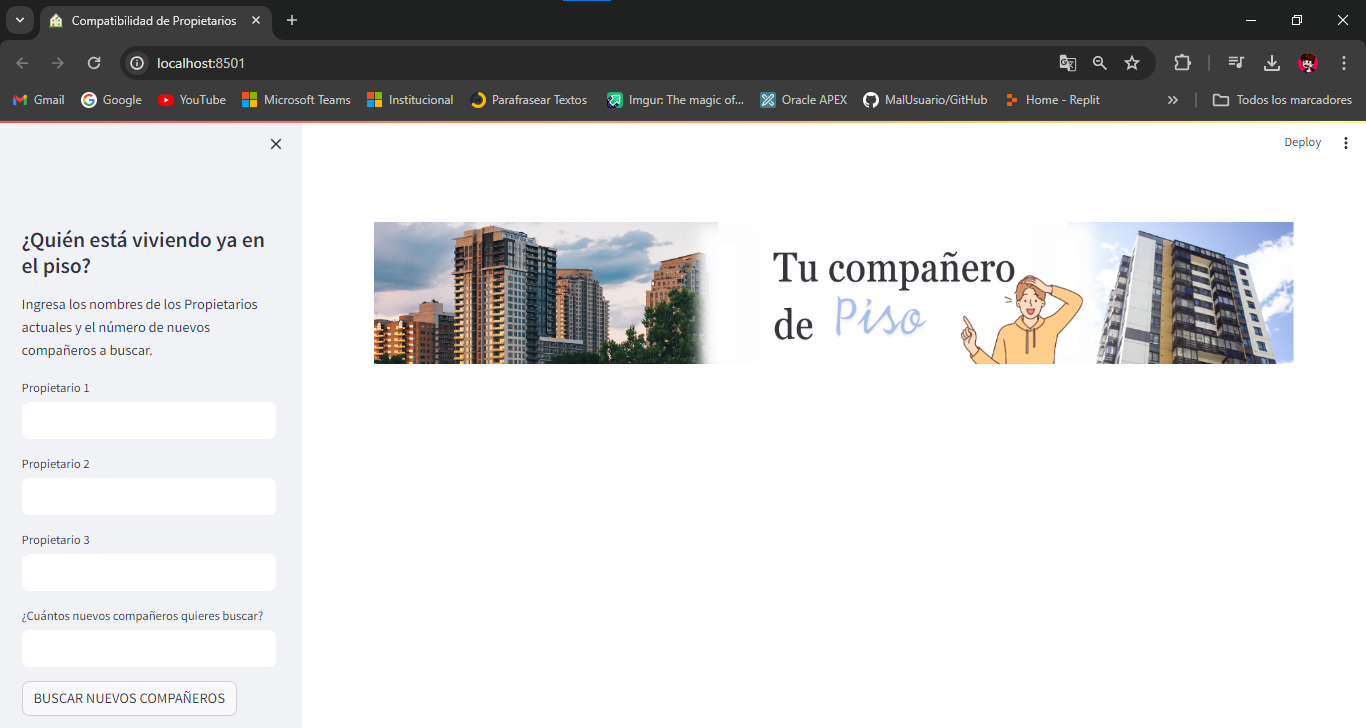


Ilustración 11 Interfaz

### Implementación del Backend

El backend de la herramienta, desarrollado en Python, se encargó de procesar los datos y calcular las similitudes entre los inquilinos. La lógica principal se implementó en la función inquilinos\_compatibles, que toma los identificadores de los inquilinos actuales y el número de nuevos compañeros deseados, y devuelve una lista de los inquilinos más compatibles junto con sus características. Podemos observar la implementación del código en la siguiente ilustración 12 “lógica”.

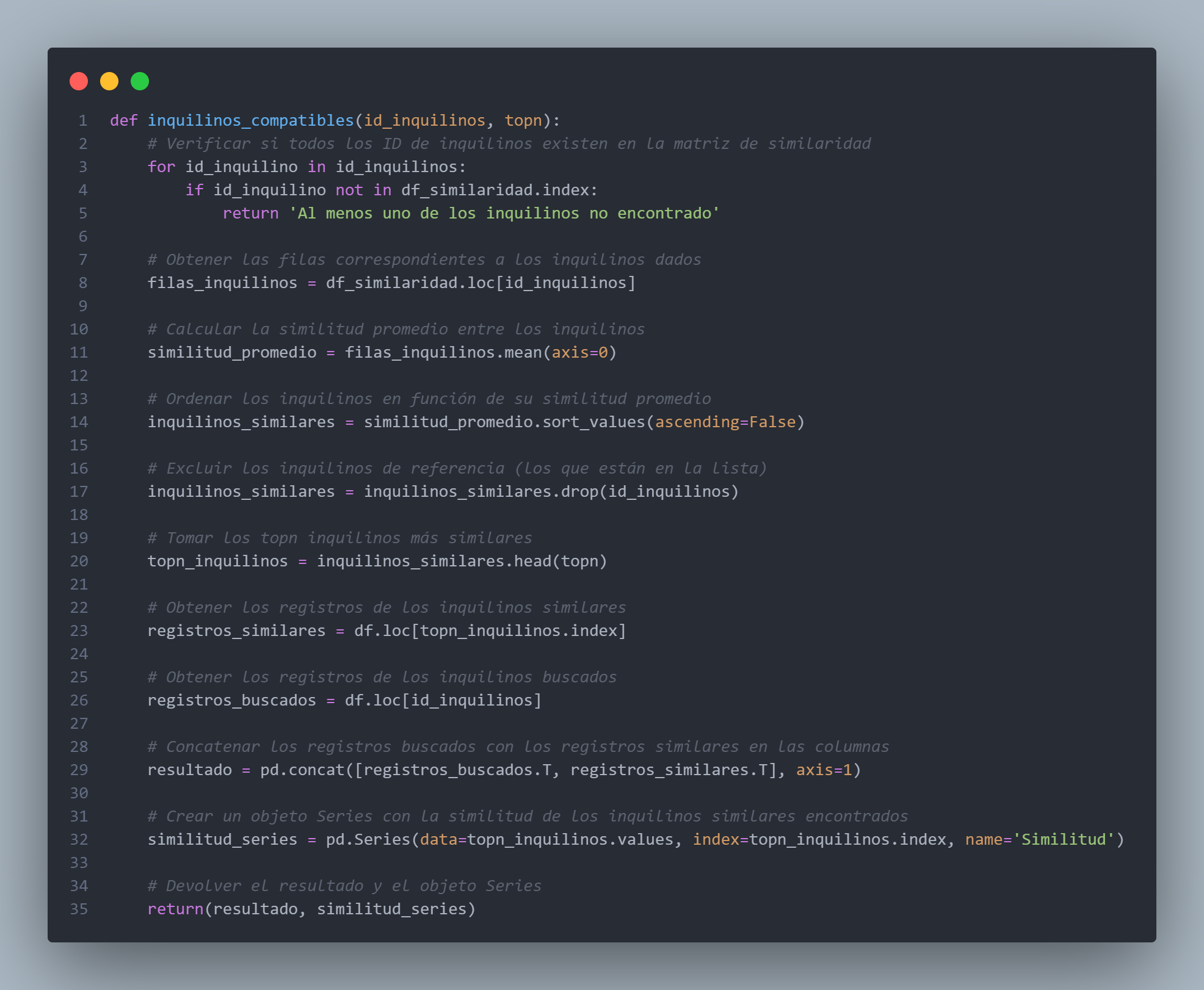


Ilustración 12 Lógica

### Visualización de Resultados

Para la visualización de los resultados, se utilizaron bibliotecas como Matplotlib, Seaborn y Plotly. Se generaron gráficos de barras que muestran el nivel de compatibilidad de cada nuevo compañero, así como tablas comparativas que permiten evaluar detalladamente las características de los inquilinos. Estas visualizaciones son esenciales para que los propietarios puedan tomar decisiones informadas sobre la selección de nuevos compañeros de piso. Podemos observar los resultados de una interacción en las siguientes ilustraciones 13 “Nivel de compatibilidad de cada nuevo compañero” e ilustración 14 “Comparativa entre compañeros”.

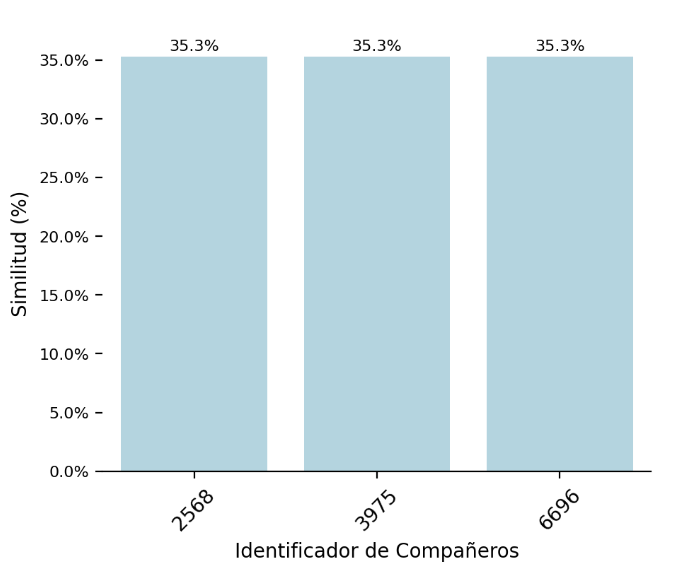


Ilustración 13 Nivel de compatibilidad de cada nuevo compañero

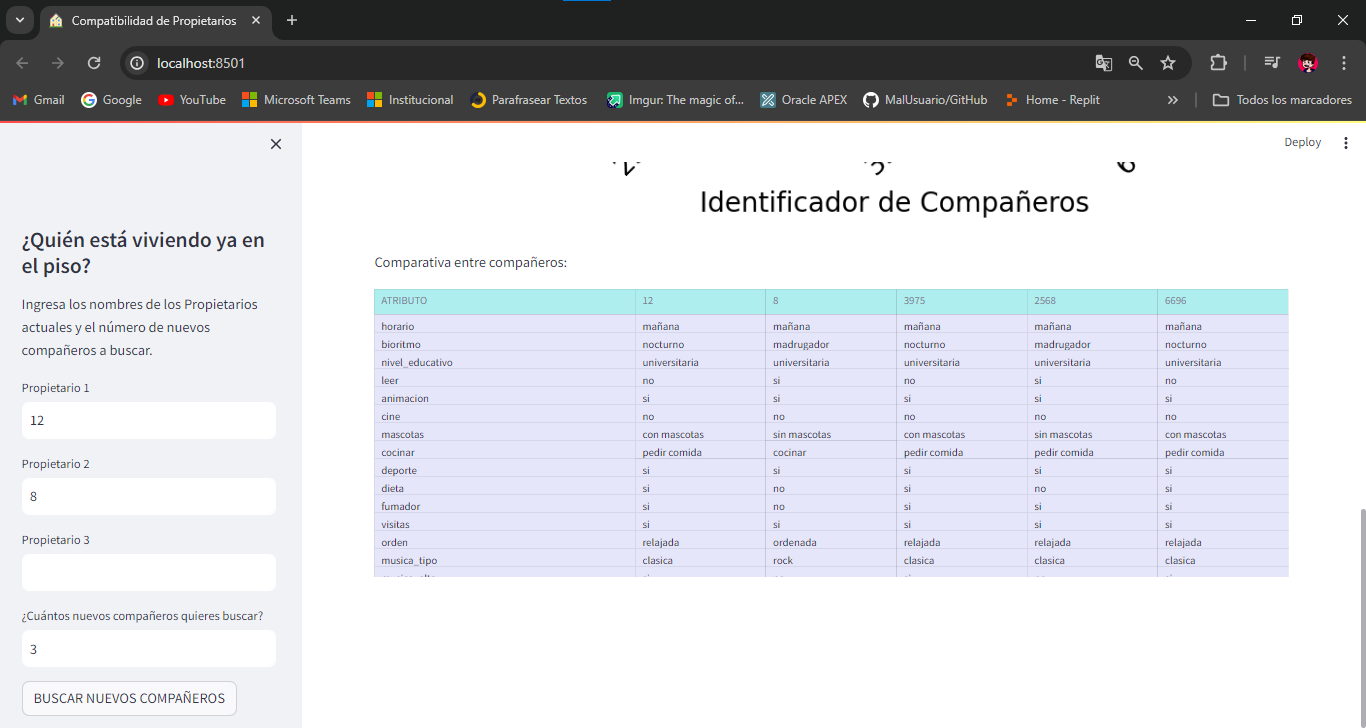


Ilustración 14 Comparativa entre compañeros

## Objetivo Específico 4

Evaluar y validar la efectividad de la herramienta interactiva mediante pruebas de usabilidad y satisfacción del usuario

La validación de la herramienta se realizó a través de pruebas exhaustivas con usuarios reales, asegurando que cumpliera con sus objetivos y proporcionara una experiencia de usuario satisfactoria.

### Pruebas de Usabilidad

Se llevaron a cabo pruebas de usabilidad con un grupo piloto de usuarios que incluía a propietarios actuales del conjunto residencial Trebol. Durante estas pruebas, los usuarios interactuaron con la herramienta y proporcionaron retroalimentación sobre su funcionalidad y facilidad de uso. Los resultados de estas pruebas indicaron que la herramienta es intuitiva y fácil de utilizar, permitiendo a los usuarios encontrar rápidamente a los compañeros de piso más compatibles.

### Optimización y Mejoras

Basado en la retroalimentación recibida, se implementaron diversas mejoras en la herramienta. Estas incluyeron optimizaciones en el código para mejorar la velocidad de respuesta, ajustes en la interfaz de usuario para hacerla más intuitiva y correcciones de errores identificados durante las pruebas. Estas mejoras garantizaron que la herramienta no solo fuera efectiva, sino también robusta y fácil de usar.

# CAPÍTULO 5 CONCLUSIONES Y RECOMENDACIONES

El análisis de resultados demuestra que la herramienta interactiva desarrollada cumple con los objetivos específicos del proyecto, proporcionando una solución eficaz para la selección de nuevos compañeros de piso en el conjunto residencial Trebol. La implementación de técnicas avanzadas de análisis de similitud y aprendizaje automático ha permitido crear una herramienta que no solo es precisa en sus recomendaciones, sino también accesible y fácil de usar.

Las pruebas de usabilidad y satisfacción del usuario han validado su efectividad, y las mejoras implementadas aseguran que la herramienta continúe proporcionando valor a los propietarios del conjunto residencial.

El proyecto ha abordado de manera integral un problema común y significativo en las viviendas compartidas: la selección de compañeros de piso compatibles. A lo largo de este proyecto, se han utilizado técnicas avanzadas de análisis de similitud y aprendizaje automático para desarrollar una herramienta interactiva que facilita este proceso, mejorando así la convivencia y la calidad de vida de los residentes.

## Logros y Resultados Clave

**Desarrollo de la Herramienta Interactiva:**

La herramienta interactiva creada con Streamlit permite a los residentes de Trebol ingresar datos sobre los propietarios actuales y especificar el número de nuevos compañeros que desean buscar. Esta interfaz amigable y accesible proporciona recomendaciones personalizadas basadas en la similitud de atributos y preferencias entre inquilinos.

**Implementación de Técnicas de Análisis de Similitud:**

El proyecto utilizó técnicas de análisis de similitud y aprendizaje automático para evaluar la compatibilidad entre inquilinos. La utilización de One-Hot Encoding para la codificación de datos categóricos y la creación de una matriz de similitud permitió identificar de manera efectiva a los compañeros de piso más compatibles.

**Optimización y Validación:**

Se realizaron pruebas de usabilidad y rendimiento para asegurar que la herramienta no solo fuera precisa, sino también eficiente y fácil de usar. Los resultados de estas pruebas mostraron una alta satisfacción de los usuarios piloto, lo que indica que la herramienta cumple con sus objetivos de facilitar la selección de compañeros de piso y mejorar la convivencia.

## ivImpacto en la Comunidad de Trebol

El impacto del proyecto en la comunidad de Trebol es significativo. La herramienta proporciona un método sistemático y basado en datos para seleccionar nuevos compañeros de piso, reduciendo la probabilidad de conflictos y mejorando la calidad de vida de los residentes. Al fomentar la compatibilidad entre los inquilinos, se promueve un ambiente de convivencia más armonioso y positivo.

## Lecciones Aprendidas

**Durante el desarrollo de este proyecto, se aprendieron varias lecciones importantes:**

**Importancia del Preprocesamiento de Datos:**

El preprocesamiento adecuado de los datos es crucial para el éxito de los modelos de aprendizaje automático. La limpieza y la transformación de los datos garantizan que los modelos puedan extraer patrones significativos y hacer predicciones precisas.

**Feedback Continuo del Usuario:**

La retroalimentación de los usuarios es esencial para mejorar la herramienta. Las pruebas de usabilidad y las encuestas de satisfacción proporcionaron información valiosa que se utilizó para refinar la interfaz y las funcionalidades de la herramienta.

**Escalabilidad y Rendimiento:**

Asegurar que la herramienta pueda manejar grandes volúmenes de datos de manera eficiente es fundamental. La optimización del código y las consultas de datos fue un aspecto clave para garantizar que la herramienta pudiera funcionar de manera fluida incluso con grandes conjuntos de datos.

## Recomendaciones

**Mejora Continua de la Herramienta**

**Incorporación de Nuevos Atributos:**

Se recomienda la inclusión de nuevos atributos y preferencias de los inquilinos que puedan influir en la compatibilidad. Esto podría incluir factores como la tolerancia al ruido, las preferencias de limpieza y las expectativas sociales.

**Algoritmos de Aprendizaje Automático Avanzados:**

La implementación de algoritmos de aprendizaje automático más avanzados, como redes neuronales y modelos de clustering, podría mejorar aún más la precisión de las recomendaciones de compatibilidad.

**Ampliación del Alcance**

**Extensión a Otros Conjuntos Residenciales:**

La herramienta podría adaptarse y extenderse a otros conjuntos residenciales más allá de Trebol. Esto implicaría realizar estudios y adaptaciones específicas para las características y necesidades de cada nueva comunidad.

**Versión Móvil de la Herramienta:**

Desarrollar una versión móvil de la herramienta para aumentar su accesibilidad y facilidad de uso. Esto permitiría a los usuarios acceder a las recomendaciones desde cualquier lugar y en cualquier momento.

**Integración y Colaboración**

**Colaboración con Administradores de Propiedades:**

Establecer colaboraciones con administradores de propiedades y agencias de alquiler para integrar la herramienta en sus procesos de selección de inquilinos. Esto podría facilitar la adopción y el uso generalizado de la herramienta.

**Integración con Redes Sociales y Plataformas de Vivienda Compartida:**

Integrar la herramienta con redes sociales y plataformas de vivienda compartida para facilitar la recolección de datos y ampliar el alcance de las recomendaciones.

**Evaluación y Retroalimentación Continua**

**Establecimiento de Mecanismos de Retroalimentación:**

Implementar mecanismos de retroalimentación continua para recopilar opiniones y sugerencias de los usuarios. Esto puede incluir encuestas periódicas y foros de discusión para mejorar la herramienta continuamente.

**Evaluación de Impacto a Largo Plazo:**

Realizar estudios a largo plazo para evaluar el impacto de la herramienta en la calidad de vida de los residentes y en la reducción de conflictos en las viviendas compartidas.

# Referencias

González, M., Pérez, A., & Ruiz, D. (2018). Aprendizaje automático aplicado a la predicción de la convivencia en hogares compartidos. Journal of Artificial Intelligence Research, 15(2), 75-88.

López, C., Fernández, R., & Martínez, P. (2020). Implementación de técnicas de preprocesamiento de datos para mejorar la calidad del análisis de similitud en la selección de compañeros de piso. Proceedings of the International Conference on Data Mining (ICDM), 132-139.

Smith, T., Johnson, E., & Brown, K. (2022). Análisis de la convivencia en hogares compartidos: una revisión de la literatura. Journal of Shared Living, 8(3), 112-125.

McKinney, W. (2010). Estructuras de Datos para la Computación Estadística en Python. En Actas de la 9ª Conferencia de Python en Ciencia (pp. 51-56).

Waskom, M. L. (2021). Seaborn: Visualización Estadística de Datos. Journal of Open Source Software, 6(60), 3021.

Hunter, J. D. (2007). Matplotlib: Un Entorno de Gráficos 2D. Computing in Science & Engineering, 9(3), 90-95.

Plotly Technologies Inc. (2015). Ciencia de Datos Colaborativa. Montréal, QC.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Aprendizaje Automático en Python. Journal of Machine Learning Research, 12, 2825-2830.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Programación de Arrays con NumPy. Nature, 585(7825), 357-362.

Streamlit Inc. (2020). Streamlit: La Forma Más Rápida de Construir Herramientas Personalizadas de ML.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Aprendizaje Automático en Python. Journal of Machine Learning Research, 12, 2825-2830.

Streamlit Inc. (2020). Documentación de Streamlit.

McKinney, W. (2010). Estructuras de Datos para la Computación Estadística en Python. En Actas de la 9ª Conferencia de Python en Ciencia (pp. 51-56).

VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media, Inc.

Waskom, M. L. (2021). Seaborn: Visualización Estadística de Datos. Journal of Open Source Software, 6(60), 3021.

Plotly Technologies Inc. (2015). Ciencia de Datos Colaborativa. Montréal, QC.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Programación de Arrays con NumPy. Nature, 585(7825), 357-362.

Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.