



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE ENGENHARIA

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA

LICENCIATURA EM ENGENHARIA INFORMÁTICA

Trabalho de Licenciatura

Desenvolvimento de uma aplicação Web para avaliação automática de preços de imóveis habitacionais na Cidade de Maputo com base na técnica de *Clustering*

Autor:

Munguambe Júnior, Rui Alberto

Supervisor:

Engº Ruben Manhiça

Maputo, Novembro de 2024



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE ENGENHARIA

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA

LICENCIATURA EM ENGENHARIA INFORMÁTICA

Trabalho de Licenciatura

Desenvolvimento de uma aplicação Web para avaliação automática de preços de Imóveis habitacionais na Cidade de Maputo com base na técnica de *Clustering*

Autor:

Munguambe Júnior, Rui Alberto

Supervisor:

Engº Ruben Manhiça

Maputo, Novembro de 2024



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE ENGENHARIA

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA

TERMO DE ENTREGA DE RELATÓRIO DO TRABALHO DE LICENCIATURA

Declaro que o estudante **Rui Alberto Munguambe Júnior** entregou no dia ____ / ____ / 2024 às ____ cópias do relatório do seu Trabalho de Licenciatura com a referência: _____ intitulado: Desenvolvimento de uma aplicação Web para avaliação automática de preços de Imóveis habitacionais na Cidade de Maputo com base na técnica de *Clustering*.

Maputo, ____ de _____ de 2024

O chefe de Secretaria



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE ENGENHARIA

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA

DECLARAÇÃO DE HONRA

Declaro sob compromisso de honra que o presente trabalho é resultado da minha investigação e que foi concebido para ser submetido apenas para a obtenção do grau de Licenciatura em Engenharia Informática na Faculdade de Engenharia da Universidade Eduardo Mondlane.

Maputo, ____ de _____ de 2024

O Autor

(Rui Alberto Munguambe Júnior)

Dedicatória

Dedico este trabalho aos meus pais

Rui Alberto Munguambe e

Aldevina Lucas Nhantumbo

Agradecimentos

Em primeiro lugar, agradeço aos meus pais Rui Alberto Munguambe e Aldevina Lucas Nhantumbo que enquanto em vida por tudo fizeram para garantir um futuro melhor para mim, o seu apoio e motivação para que eu alcance os meus sonhos e me torna em quem eu sou hoje.

Agradeço aos meus irmãos que me apoiaram e continuam apoiando no meu crescimento, bem como pelo aprendizagem que tenho no meu dia a dia.

Agradeço a todo o corpo docente da Universidade Eduardo Mondlane, principalmente aos docentes do curso de Engenharia informática que partilharam as suas experiências e conhecimentos que servem de uma fonte contínua de nutrientes para o meu aprendizado contínuo fora do meio académico.

Agradeço, em especial, ao Eng.º Rúben Manhiça pelo aprendizado e orientações para o meu sucesso na realização do presente trabalho.

Agradeço aos meus colegas, pelos momentos que tivemos juntos dentro e fora da faculdade, bem como pela amizade que partilhamos.

Epígrafe

"Keep Calm and Keep Walking"

Resumo

O mercado imobiliário é um sector no qual são negociados imóveis, com participação de proprietários, interessados e corretores para a venda e compra. Em Moçambique, os métodos tradicionais de avaliação de imóveis são amplamente utilizados, baseando-se em dados históricos e na experiência dos avaliadores. No entanto, essa abordagem pode gerar resultados imprecisos, especialmente em contextos onde o avaliador tem conhecimento limitado ou os dados comparativos são escassos. Além disso, a subjetividade do processo pode levar a inconsistências nas avaliações.

Com o avanço da tecnologia e a necessidade de avaliações mais rápidas e consistentes, explorou-se o uso da aprendizagem de máquina para modernizar esse processo. Realizou-se uma análise do cenário atual de avaliação imobiliária em Maputo, identificando o método comparativo como o mais utilizado, mas com limitações devido à falta de ferramentas tecnológicas e à variação de preços que não refletem as condições reais do mercado.

Dessa forma, propôs-se o desenvolvimento de uma aplicação web que aplica técnicas de aprendizagem de máquina para estimar o valor de imóveis com base em dados históricos e características relevantes. O desenvolvimento da solução incluiu a preparação de dados, engenharia de atributos e treinamento de modelos, seguido pela implementação de um protótipo que permite efectuar estimativas com base em informações introduzidas pelo utilizador.

Durante o desenvolvimento, ferramentas como Django, Google Colab e MySQL foram utilizadas para o desenvolvimento da aplicação, análise de dados e visualização dos resultados. A solução final oferece uma alternativa prática e objetiva, reduzindo a dependência de avaliadores tradicionais e fornecendo aos usuários uma referência confiável para a tomada de decisões.

Palavras-chave: avaliação automática de imóveis, *clustering*, *machine learning*, eficiência na avaliação de imóveis, aplicação web.

Abstract

The real estate market is a sector in which properties are traded, with the participation of owners, interested parties and brokers for sale and purchase. In Mozambique, traditional property valuation methods are widely used, based on historical data and the experience of appraisers. However, this approach can generate inaccurate results, especially in contexts where the evaluator has limited knowledge or comparative data is scarce. Furthermore, the subjectivity of the process can lead to inconsistencies in assessments.

With the advancement of technology and the increasing need for faster and more consistent assessments, the use of machine learning was explored to modernize this process. An analysis of the current real estate valuation scenario in Maputo was carried out, through which the comparative method was identified as the most used, however, this approach has shown limitations due to the lack of technological tools and price variations that do not reflect real market conditions.

Therefore, it was proposed to develop a web application that applies machine learning techniques to estimate the value of properties based on historical data and relevant characteristics. The development of the solution included data preparation, feature engineering and model training, followed by the implementation of a prototype that allows estimates to be made based on information input by the user.

During development, tools such as Django, Google Colab and MySQL were used to develop the application, analyze data and visualize results. The final solution offers a practical and objective alternative for evaluation, reducing dependence on traditional evaluators and providing users with a reliable reference for decision-making.

Keywords: automatic property valuation, clustering, machine learning, efficiency in property valuation, web application.

Índice

1. Capítulo I - Introdução	1
1.1. Contextualização.....	1
1.2. Motivação	2
1.3. Descrição do problema	2
1.4. Objectivos.....	3
1.4.1. Objectivo Geral	3
1.4.2. Objectivos Específicos	3
1.5. Metodologia	4
1.6. Organização do Trabalho	6
2. Capítulo II – Revisão da Literatura	7
2.1. Mercado Imobiliário	7
2.1.1. Avaliação de Imóveis	7
2.1.2. Desafios	8
2.2. Utilização da Inteligência Artificial (IA) na Avaliação de Imóveis.....	9
2.3. <i>Clustering</i> (Agrupamento)	13
2.4. Métricas de Avaliação de Modelos de <i>Machine Learning</i>	17
2.5. Ferramentas e Tecnologias para o Desenvolvimento de uma Aplicação Web .	17
2.5.1. <i>Front-end</i>	18
2.5.2. <i>Back-end</i>	18
3. Capítulo III – Caso de Estudo.....	19
3.1. Descrição do Cenário Actual.....	19
3.2. Resultados Obtidos no Campo através de Entrevistas	19
3.3. Comparação entre os <i>Stacks</i> Tecnológicos	21
3.4. Proposta de Solução	23

4. Capítulo IV – Construção do Modelo de Avaliação Automático Baseado em <i>Clustering</i>	25
4.1. Colecta de dados	25
4.2. Pré Processamento de dados.....	26
4.3. Transformação dos dados.....	28
4.4. Aplicação das técnicas de <i>Clustering</i>	29
4.5. Elaboração do Módulo de Avaliação de Imóveis.....	31
5. Capítulo V – Desenvolvimento do Protótipo	33
5.1. Levantamento de Requisitos.....	33
5.1.1. Requisitos Funcionais.....	33
5.1.2. Requisitos Não Funcionais	34
5.2. Análise	34
5.2.1. Análise dos Requisitos	34
5.2.2. Casos de Uso.....	35
5.2.3. Modelagem de Dados.....	38
5.3. Design.....	39
5.4. Implementação.....	39
5.5. Testes	40
6. Capítulo VI – Conclusões e Recomendações	44
6.1. Conclusões	44
6.2. Recomendações	45
Fontes Bibliográficas.....	46
Apêndices	A1.1
Apêndice 1: Interfaces do Utilizador.....	A1.1
Apêndice 2: Roteiro de Entrevista	A2.1

Apêndice 3: Casos de Uso.....	A3.1
-------------------------------	------

Lista de Tabelas

Tabela 1 - Aplicação dos Modelos de Avaliação Automática. <i>Fonte:</i> (Zacarias, 2022)..	12
Tabela 2 - Utilização de modelos automáticos de avaliação em alguns Países. <i>Fonte:</i> (Zacarias, 2022).....	13
Tabela 3 - Comparação de Bases de Dados.....	22
Tabela 4 - Comparação de ambientes de desenvolvimento do modelo de avaliação....	22
Tabela 5 - Comparação de <i>frameworks</i> para o desenvolvimento da aplicação proposta	23
Tabela 6 - Descrição dos requisitos funcionais	33
Tabela 7 - Descrição dos requisitos não funcionais	34
Tabela 8 - Actores do sistema.....	36
Tabela 9 - Descrição do caso de uso: Introduzir dados do imóvel.....	36
Tabela 10 - Caso de uso: Avaliar o valor do imóvel	36
Tabela 11 - Caso de uso: Apresentar o resultado da avaliação	37
Tabela 12 - Caso de uso: Armazenar os dados dos imoveis avaliados.....	37
Tabela 13 - Caso de uso: Extrair o relatório da avaliação	37

Lista de Figuras

Figura 1 - Percentagem de valores em falta de atributos com pelo menos um valor em falta	26
Figura 2 - Gráfico com a assimetria calculada por atributo	27
Figura 3 - Visualização do preço em relação ao nº de casas de banho	27
Figura 4 - Resultados obtidos após a transformação logarítmica	28
Figura 5 - Resultados obtidos após aplicar a técnica DBSCAN	30
Figura 6 - Resultados obtidos após aplicar a técnica <i>K-Medoids</i>	30
Figura 7 - Resultados obtidos após aplicar a técnica <i>K-Means</i>	30
Figura 8 - Resumo referente ao valor dos imóveis em cada um dos clusters nos dados de teste	31
Figura 9 - Diagrama de casos de uso	38
Figura 10 - Diagrama da arquitectura em 3 camadas da aplicação	39
Figura 11 - Função de Teste:Carregamento do formulário de avaliação.....	41
Figura 12 - Função de Teste: Avaliação de Imóveis	41
Figura 13 - Função de Teste: Submissão e carregamento de dados na base de dados....	42
Figura 14- Função de Teste: Criação de uma avaliação associada ao imóvel avaliado.....	42
Figura 15 - Página principal da aplicação	A1.1
Figura 16 - Interface para visualização do resultado.....	A1.1
Figura 17 - Introdução de dados do imóvel.....	A3.1
Figura 18 - Avaliar o valor de um imóvel	A3.1
Figura 19 - Apresentar resultados da avaliação.....	A3.2
Figura 20 - Extrair o relatório	A3.2

1. Capítulo I - Introdução

Neste capítulo, são abordados aspectos introdutórios bem como a definição do problema, contextualização do problema, os objectivos do trabalho e a metodologia aplicada.

1.1. Contextualização

O mercado imobiliário é um sector no qual são negociados imóveis para a venda e compra. Normalmente, participam da negociação três agentes principais, nomeadamente, o proprietário do imóvel, os interessados e os correctores dos mesmos. Os proprietários procuram ganhar valor com o imóvel, os interessados em adquirir o imóvel e o corretor que auxilia o processo de transação do bem.

Em Moçambique, assim como em muitos outros países, os métodos tradicionais de avaliação de imóveis são amplamente utilizados. Estes métodos incluem a abordagem baseada em comparação de mercado, que é uma das principais opções para avaliação, e consiste em determinar o valor de um imóvel comparando-o com imóveis semelhantes que foram recentemente vendidos na mesma área ou bairro. Nesta abordagem, o avaliador efectua ajustes no preço do imóvel de acordo com as diferenças que este possui em relação aos objectos de comparação.

De acordo com Naiene (Naiene, 2015), a relação quantitativa de três componentes (imóvel, vendedor e comprador) formadores do mercado é uma relação determinante na formação dos preços, sendo a situação ideal aquela onde a oferta e a procura é equilibrada, ou seja, um mercado de concorrência perfeita. Por outro lado, pode ocorrer o mercado de concorrência imperfeito, onde ocorrem casos de monopólio. Neste caso, o mercado é comandado por um único vendedor. Já no caso do oligopólio (mais comum) ocorre a situação de mercado em que a oferta é controlada por um pequeno número de vendedores, e em que a competição tem por base, não as variações de preços, mas sim a propaganda e as diferenças de qualidade. Isto faz os preços serem “puxados” para cima. No caso de monopsonio (raro) onde há apenas um comprador e oligopsonio (mais comum) alguns compradores, os preços têm uma tendência para baixo (Naiene, 2015).

Portanto, o mercado imobiliário é um sector em constante mudança, impactado por vários fatores, como a localização da propriedade, o tipo de imóvel, as condições do imóvel, as taxas de juros, e outros factores que tornam a determinação de um preço óptimo para venda e compra mais difícil visto que o mercado oferece poucas ferramentas disponíveis para o público interessado em avaliar os preços de imóveis habitacionais.

Este sector é tendencialmente visto como o impulsionador do desenvolvimento do espaço urbano das cidades, possibilitando maior qualidade de vida para a sociedade, o que o torna importante para o desenvolvimento da economia do País (Augusto, 2019).

Este trabalho está focado nos mecanismos para determinação de preços ótimos para a venda e compra de imóveis habitacionais que sejam utilizados para tornar a negociação mais rápida e satisfatória para as partes interessadas.

1.2. Motivação

A principal motivação para a concepção deste projecto é a dificuldade verificada das pessoas no meio do autor em efectuar de forma clara uma avaliação de imóveis nos quais demonstrou interesse em saber mais sobre.

1.3. Descrição do problema

Em Moçambique, o processo de avaliação de imóveis enfrenta vários desafios dos quais destacam-se os seguintes a falta de dados adequados (confiáveis e atualizados) e a presença de imóveis informais cujo valor pode ser subestimado ou superestimado.

O processo de avaliação tradicional depende de profissionais qualificados, como avaliadores imobiliários, engenheiros e arquitetos, que utilizam sua expertise para determinar o valor dos imóveis. No entanto, assim como em muitos outros países, os métodos de avaliação de imóveis normalmente aplicados por estes são baseados em dados históricos, tornando-se o processo fortemente dependente da experiência e integridade do avaliador, o que na maioria das vezes leva a resultados imprecisos, especialmente em

contextos onde o conhecimento do avaliador ou a disponibilidade de dados comparativos são limitados.

Diante deste cenário, torna-se essencial o desenvolvimento de uma solução que permita às partes interessadas obter estimativas de valor de imóveis de forma confiável, rápida e acessível, sem depender exclusivamente de profissionais.

Este trabalho visa desenvolver uma aplicação web que utilize técnicas de *Clustering* para a avaliação automática de imóveis habitacionais na Cidade de Maputo, abordando as limitações do sistema atual e promovendo uma maior confiança e eficiência no mercado imobiliário.

A questão de pesquisa colocada é:

- Como tornar o processo de avaliação de preços de Imóveis automático e acessível para o público interessado?

1.4. Objectivos

1.4.1. Objectivo Geral

- Desenvolver uma aplicação web para avaliação automática de Imóveis habitacionais na Cidade de Maputo com base na técnica de *Clustering*;

1.4.2. Objectivos Específicos

- Descrever os mecanismos actuais de avaliação de imóveis na Cidade de Maputo e os seus desafios;
- Avaliar técnicas de *Clustering* que podem ser usadas para avaliação automática de imóveis na Cidade de Maputo;
- Identificar um *stack* tecnológico Web para o desenvolvimento de uma aplicação Web confiável, transparente, eficiente e acessível para a avaliação de imóveis habitacionais na Cidade de Maputo;
- Desenvolver um protótipo funcional de um sistema de avaliação de imóveis em Maputo com base no *stack* tecnológico Web e técnica de *Clustering* escolhidos.

1.5. Metodologia

No âmbito desta pesquisa, será adoptada uma abordagem quantitativa que segundo Da Silva & Menezes (2005), considera que tudo pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las, entretanto, requer o uso de recursos e de técnicas estatísticas (percentagem, média, moda, mediana, desvio-padrão, coeficiente de correlação, análise de regressão, etc.).

Primeiro serão recolhidas informações sobre os mecanismos actuais e os seus desafios, serão efetuadas pesquisas bibliográficas e entrevistas a profissionais da área. A entrevista será elaborada de acordo com os passos a seguir, baseados na pesquisa de Manzini (2012):

1. Descrição das informações que se pretende adquirir;
2. Determinação das pessoas certas para entrevistar;
3. Elaboração do Roteiro que inclui a introdução, perguntas principais e uma conclusão;
4. Realização da entrevista;

De seguida, serão avaliadas as técnicas de *Clustering* de acordo com os seguintes procedimentos:

1. Identificação dos critérios de avaliação;
2. Recolha de dados sobre os imóveis na Cidade de Maputo, incluindo o preço;
3. Preparação dos dados através do tratamento de valores em falta, *outliers*, inconsistências e conversão dos mesmos em formatos adequados para a análise;
4. Escolha de técnicas de *Clustering* que podem ser adequados baseado na natureza dos dados e a distribuição dos dados;
5. Definição de métricas para avaliação dos *clusters* gerados pelas técnicas de *Clustering* seleccionadas;
6. Aplicação, optimização e avaliação através de métricas definidas das técnicas de *Clustering* nos dados colecionados;
7. Interpretação dos resultados por meio da análise dos *clusters* gerados pelas diferentes técnicas com vista a verificar se estes estão alinhados com as características conhecidas dos imóveis na Cidade de Maputo;
8. Com base na análise dos resultados, serão efectuados ajustes nos parâmetros dos algoritmos ou na preparação dos dados de modo a melhorar a qualidade dos *clusters*;

9. Por fim, a técnica com os melhores resultados será selecionada para que seja aplicada no desenvolvimento do módulo de avaliação de imóveis na Cidade de Maputo.

Após a escolha da técnica mais adequada, irá se proceder com a escolha do *stack* tecnológico de acordo com os passos a seguir:

1. Identificação dos requisitos da aplicação;
2. Análise das necessidades técnicas como os recursos necessários para a aplicação, manipulação de dados, escalabilidade e desempenho;
3. Pesquisa de tecnologias existentes (*frontend*, *backend*, banco de dados, infraestruturas de hospedagem e ferramentas de desenvolvimento);
4. Avaliação de Tecnologias com base em critérios como:
 - **Confiança e Maturidade** - estabilidade, a comunidade de desenvolvedores e a adoção pela indústria;
 - **Eficiência e Desempenho** - desempenho da tecnologia em relação aos requisitos de escalabilidade e tempo de resposta;
 - **Acessibilidade e Usabilidade** - suporte a boas práticas de acessibilidade e experiência de usuário;

Por fim, será desenvolvido um protótipo funcional de um sistema de avaliação de imóveis na Cidade de Maputo com base no *stack* tecnológico Web e técnica de *Clustering* escolhida aplicando a metodologia de desenvolvimento *Waterfall*, de acordo com os passos identificados por Pontes e Arthaud (2018):

1. **Levantamento de Requisitos** - serão identificados os requisitos de acordo com as necessidades da aplicação;
2. **Análise** - os requisitos serão analisados e transformado em especificações técnicas que servirão como base para o *design* do sistema;
3. **Design** - será definida a arquitetura do sistema com base nas especificações técnicas. A definição da arquitetura irá envolver a elaboração de diagramas e a respectiva documentação;
4. **Implementação** - o protótipo da aplicação será desenvolvido de acordo com a arquitetura definida e os requisitos identificados;

5. **Testes** - serão efectuados testes no protótipo implementado, com o objectivo de garantir que os requisitos são satisfeitos e a aplicação esteja a funcionar corretamente.

1.6. Organização do Trabalho

Para melhor percepção dos capítulos e os tópicos abordados, refere-se a organização do trabalho e a respectiva descrição de cada ponto. O trabalho está organizado da seguinte forma:

- **Capítulo I – Introdução:** neste capítulo, são abordados aspectos introdutórios bem como a definição do problema, contextualização do problema, os objectivos do trabalho e a metodologia aplicada.
- **Capítulo II – Revisão da Literatura:** neste capítulo, apresenta-se tópicos importantes para a realização do presente trabalho.
- **Capítulo III – Caso de Estudo:** neste capítulo, será descrito o cenário actual, os desafios, as tecnologias existentes e a solução proposta.
- **Capítulo IV – Construção do Modelo de Avaliação Automático Baseado em *Clustering*:** neste capítulo, serão apresentados os processos aplicados no desenvolvimento do modelo de avaliação, bem como os resultados obtidos e o modelo final.
- **Capítulo V – Desenvolvimento do Protótipo:** neste capítulo, será apresentada a documentação técnica do protótipo desenvolvido, incluindo os requisitos e a arquitectura.
- **Capítulo VI – Conclusões e Recomendações:** neste capítulo, será avaliado o grau de satisfação com cada um dos objetivos e serão partilhadas recomendações para pesquisas futuras.

2. Capítulo II – Revisão da Literatura

Neste capítulo é possível encontrar conceitos de avaliação de imóveis, aplicações de inteligência artificial e *clustering*. Primeiramente, será apresentado o enquadramento referente ao mercado imobiliário, incluindo o processo de avaliação de imóveis e os seus desafios. De seguida serão descritos conceitos básicos de inteligência artificial e a sua aplicação, onde serão abordados os diferentes métodos de aprendizagem e a forma como diversos países aplicam este conhecimento para a elaboração de modelos de avaliação automática de imóveis. Por fim, será descrito a visão geral de uma aplicação web e as ferramentas e tecnologias necessárias para construir uma.

2.1. Mercado Imobiliário

O mercado imobiliário é o sector onde ocorrem transações de bens imóveis. Quando alguém compra, vende ou aluga algum desses bens, consequentemente está participando de uma negociação no mercado imobiliário.

O mercado imobiliário é um sector no qual são negociados imóveis para a venda e compra, pelo que, de acordo com Sousa (2022), quando alguém compra ou vende algum imóvel está, consequentemente, efectuando uma negociação no mercado imobiliário. Normalmente, participam da negociação três agentes principais, nomeadamente, o proprietário do imóvel, os interessados e os correctores dos mesmos. Os proprietários procuram ganhar valor com o imóvel, os interessados em adquirir o imóvel e o corretor que auxilia o processo de transação do bem.

Segundo Sousa (2022), o mercado imobiliário funciona com ciclos de demanda e oferta, e nesses ciclos, por vezes, a demanda é superior à oferta em uma região específica, quando as pessoas da região procuram por um tipo específico de imóvel para compra ou venda.

2.1.1. Avaliação de Imóveis

Para que seja efectuada uma transação sobre um imóvel no mercado imobiliário, é essencial que seja efectuada a avaliação do mesmo com o objectivo de determinar o valor óptimo para a sua compra ou venda. De acordo Gonzalez (2022), a avaliação de imóveis pode ser classificada em individual ou colectiva, sendo que, as avaliações individuais e colectivas

diferem apenas em termos de escala, e os métodos aplicados são essencialmente os mesmos.

Segundo o autor Gonzalez (2022), o método comparativo de dados de mercado é o mais utilizado para avaliação individual, e consiste em uma inspeção detalhada do imóvel, busca e seleção de propriedades similares, ajustamento dos preços e documentação da avaliação através de um formulário ou relatório, sendo que, em alguns países, como os Estados Unidos, é comum a avaliação baseada apenas em três dados similares.

Na avaliação coletiva é elaborado um modelo representativo do mercado que deverá refletir teoria da avaliação e o comportamento do mercado no ciclo em que se pretende avaliar os imóveis, no entanto, a análise efectuada nesse tipo de avaliação é mais geral e normalmente é elaborado um relatório geral (González, 2022).

2.1.2. Desafios

De acordo com a pesquisa de Żróbek , Sabina, et al (2014), o processo de avaliação de imóveis enfrenta vários desafios dos quais se destacam:

- **Acesso a Dados Confiáveis** - o difícil acesso a dados actualizados e precisos sobre imóveis para comparação pode impactar significativamente no resultado de avaliações efectuadas;
- **Complexidade de Factores** – o processo de avaliação envolve a análise de múltiplos factores, incluindo características físicas do imóvel, mercado local, infraestrutura da região, e tendências econômicas, o que torna o processo de análise manual demorado e complexo, sendo suscetível a erros;
- **Subjetividade na Avaliação** - diferentes profissionais poderão estimar valores distintos para o mesmo imóvel, devido a diferenças na interpretação dos dados obtidos, experiência e ate mesmo a preferências pessoais dos mesmos.

2.2. Utilização da Inteligência Artificial (IA) na Avaliação de Imóveis

Segundo Gomes (2010), a inteligência artificial é uma área multidisciplinar da Ciência da Computação que envolve a psicologia, biologia, lógica matemática, linguística, engenharia, filosofia, etc, com o objectivo de permitir que os computadores possam agir ou pensar de forma inteligente. No entanto, para se adquirir inteligência deve-se considerar dois elementos fundamentais, nomeadamente: o aprendizagem e a experimentação (Sanchez, 2023).

Sanchez (2023) evidencia a importância destes elementos utilizando um bebê como exemplo, sendo que, ao nascer, o mesmo possui todos membros e órgãos em funcionamento sendo capaz de exercer tarefas involuntárias como respirar, entretanto, para efectuar tarefas complexas como falar, este irá precisar de observar, repetir e treinar. As etapas de observação (colecta de dados), repetição e treino, compõem o processo de aprendizagem que é essencial para os sistemas de Inteligência Artificial, entretanto, este processo é sustentado pela análise de dados, uma vez que estes sistemas são baseados em dados.

De acordo com Da Silva e Vanderlinde (2010), somente um sistema capaz de aprender poderá ser chamado de inteligente, sendo que, deste modo a aprendizagem de máquina é um processo de aprendizagem por experiência através da qual, conforme a tarefa é executada, a máquina aprende a melhor maneira de resolver.

Sendo assim, a aprendizagem de máquina é um subcampo da inteligência artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitem um programa de computador aprender a partir de uma experiência E com respeito a uma classe de tarefas T e com medida de desempenho P , se seu desempenho na tarefa T melhorar com a experiência E (Corcovia & Alves, 2019). Segundo Corcovi e Alves (2019), a aprendizagem de máquina está dividido em dois principais métodos de aprendizagem, nomeadamente:

- **Aprendizagem supervisionada** - neste método, é fornecida uma classe à qual cada amostra no treinamento pertence; e

- **Aprendizagem não supervisionada** - ao contrário do método supervisionado, no método não supervisionado, o rótulo da classe de cada amostra dos dados de treinamento não é conhecido. Adicionalmente, o número ou conjunto de classes a ser treinado pode não ser conhecido a priori.

De acordo com Sanchez (2023), as principais áreas de estudo e aplicação da Inteligência Artificial são:

- Previsão de resultados - neste caso, a IA é utilizada para prever resultados em diferentes contextos, como a previsão de vendas, previsão de movimentos do mercado financeiro, previsão de doenças com base em sintomas, entre outros;
- Optimização de processos - diferentes sectores aplicam a IA para otimizar processos os seus processos com vista a reduzir custos e melhorar a eficiência dos seus processos;
- Visão computacional - permite aos computadores identificar objectos no ambiente em que se encontram através do processamento de imagens e vídeos, tornando possível o projectos como veículos autónomos;
- Classificação - com a IA é possível classificar dados colectados em diferentes categorias. Esta componente é aplicada em sistemas de recomendação, identificação de spams, entre outros;
- determinação de padrões - a IA pode ser utilizada na identificação de padrões em dados cuja distribuição dificulta identificação dos mesmos através da observação e percepção humana. Pode ser aplicada na detecção de fraudes, diagnóstico médico, entre outros;
- Reconhecimento e síntese de voz - com a IA pode-se reconhecer e sintetizar voz humana, permitindo desenvolver funcionalidades como a interação com sistemas através de comandos de voz ou simplesmente a tradução de automática de voz;
- Tomada de decisão - a IA pode oferecer informações úteis baseadas em dados para o auxílio na tomada de decisões.

Para a avaliação de imóveis, a IA pode ser aplicada para o desenvolvimento de modelos que sejam capazes de efectuar a avaliação automaticamente. Zacarias (2022) descreve modelos automatizados de avaliação (AVM) que são programas informáticos baseados em

algoritmos de IA utilizados para estimar o valor de um imóvel através da análise de fatores como a localização, as condições de mercado ou as características do imóvel, a partir de dados colectados.

De acordo com Zacarias (2022), uma das principais diferenças entre as avaliações tradicionais e as avaliações com AVM é que os peritos realizam inspeções da própria propriedade e os AVM dependem dos dados que são colectados para maior precisão nas avaliações, no entanto, apesar desta dependência, estes modelos possuem várias vantagens, das quais:

- Aumento da transparência do setor;
- Baixo custo;
- Consistência;
- Velocidade;
- Objetividade;
- Ferramenta útil para peritos avaliadores através da comparação de valores.

Diz ainda o autor, que o objectivo da utilização da IA para avaliação de imóveis varia de acordo com o papel do *stakeholder* em relação ao imóvel como indicado na tabela 1.

<i>Stakeholder</i>	Objectivos
Especialista em Avaliação	▪ Complemento do método de avaliação tradicional (permite a comparação dos valores obtidos nos dois métodos); ▪ Revisão de avaliações.
Comprador	▪ Percepção do valor a pagar.
Vendedor	▪ Percepção do valor a receber.
Credor	▪ Definir o valor do crédito à habitação ou da hipoteca.
Investidor	▪ Percepção do valor a investir.

Seguradora	▪ Avaliação de risco e assim definir o valor do seguro do imóvel.
Governo	▪ Determinar o valor do imposto que tem a receber pelo imóvel; ▪ Análise Financeira de despesa pública.

Tabela 1 - Aplicação dos Modelos de Avaliação Automática. Fonte: (Zacarias, 2022)

Em Moçambique não existe uma plataforma pública que permita aos diferentes *Stakeholders* avaliar imóveis de acordo com as suas necessidades em tempo útil independentemente do lugar em que se encontram. Entretanto, em alguns Países existem plataformas que permitem aos usuários efectuar a avaliação de imóveis antes de colocar uma proposta do valor de venda ou compra como as plataformas *housfye* e *imovenda*.

Adicionalmente, de acordo com a pesquisa de Zacarias (2022), alguns Países aplicam modelos baseados em IA para o desempenho de diferentes tarefas, sendo que estas são regidas por determinados regulamentos conforme apresentado na tabela 2.

País	Aplicação
Itália	▪ Aplicados segundo os regulamentos do Banco da Itália usados, principalmente, na gestão de garantias hipotecárias;
EUA	▪ Usados segundo os padrões da <i>International Association of Assessing Officers</i>, da Prática de Avaliação Profissional e os regulamentos de cada Estado; ▪ Geralmente, são utilizados para a tributação da propriedade, segurança hipotecária e seguros;
Reino Unido	▪ São utilizados, maioritariamente, por bancos; ▪ Alguns serviços construíram estes sistemas para que os avaliadores possam verificar as suas avaliações;
Noruega	▪ Existe legislação, onde os Imóveis são registados e avaliados; ▪ No setor público são apenas utilizados por autoridades fiscais;

	▪ Existem empresas privadas que construíram para análise e previsão de preços;
--	--

Tabela 2 - Utilização de modelos automáticos de avaliação em alguns Países. *Fonte: (Zacarias, 2022)*

Como observado na tabela 2, os modelos desenvolvidos apresentam-se como ferramentas fundamentais para determinadas actividades em diversos Países e poderá ter um impacto positivo se for aplicado no contexto Moçambicano.

2.3. **Clustering (Agrupamento)**

Segundo Rai e Singh (2010), *Clustering* é um método de aprendizagem não supervisionada que consiste em combinar dados sem classes em grupos (*clusters*) diferentes constituídos por dados semelhantes. Diz ainda o autor, que os *clusters* formados deverão satisfazer 2 principais critérios, que são:

- Cada *cluster* é homogêneo, ou seja, as amostras pertencentes ao mesmo grupo dever ser semelhantes entre si;
- Cada *cluster* deve ser diferente dos outros, ou seja, as amostras pertencentes a diferentes grupos devem ser diferentes.

Pese embora os critérios descritos sejam importantes para determinar a escolha do algoritmo de *Clustering* a ser implementado, cada algoritmo possui propriedades que, de acordo com as necessidades, poderão determinar o algoritmo mais adequado para o que se pretende desenvolver. Rai e Singh (2010) descrevem as seguintes propriedades:

O *Clustering* possui diversas abordagens cuja aplicação depende do objectivo das análises que se pretende efectuar, sendo que, a interpretação dos resultados obtidos poderá variar com a abordagem escolhida. Dentre as variadas abordagens, Popa e Emmanuel (2014) destacam as seguintes:

- **Clustering Particional ou Optimização** - o objetivo da abordagem de optimização é de formar *clusters* óptimos sobre os dados, dividindo iterativamente as amostras em *k clusters* mutuamente exclusivos. Nesta abordagem, todos os exemplos são utilizados para o cálculo da distância entre os agrupamentos, sendo os algoritmos *k-means* e *k-medoids* os representantes desta abordagem (Metz, 2006);

- **Clustering Baseado em Densidade** - esta abordagem assume que os *clusters* são regiões de alta densidade separadas por regiões com baixa densidade no espaço de exemplos. A ideia dessa abordagem é que cada exemplo do cluster deve manter uma vizinhança com um número mínimo de vizinhos dentro de uma esfera com raio R . Os exemplos que possuem uma vizinhança com densidade mínima, e estão numa distância menor que R , pertencem ao mesmo cluster. O DBSCAN é um exemplo de algoritmos que aplicam esta abordagem (Metz, 2006);
- **Clustering Hierárquico** - esta abordagem se difere das outras mencionadas pelo facto dos resultados obtidos não serem constituídos apenas de uma partição do conjunto de dados inicial, mas sim de uma hierarquia que descreve um particionamento diferente à cada nível analisado. Sendo assim, os variados *clusters* que compõem o conjunto de dados, poderão ser compostos por *sub-clusters*, que por sua vez, poderão também ser formados por *sub-clusters* (Metz, 2006).

Conforme a descrição anterior, a escolha da abordagem de *Clustering* a ser aplicada irá depender do objectivo da análise, no entanto, para o processo de *Clustering* poderão ser realizadas as mesmas etapas sobre os dados colectados. De acordo com Metz (2006), as principais etapas realizadas no processo de *Clustering* são:

1) Pré-processamento dos dados

Esta etapa compreende a preparação dos dados através do tratamento de valores em falta, *outliers*, inconsistências e a transformação dos mesmos em formatos adequados para a análise pretendida.

De acordo com Nascimento, et al (2024), a transformação de dados é uma actividade muito importante desta etapa, visto que maior parte dos algoritmos requerem dados numéricos ao invés de categóricos. O autor destaca as seguintes técnicas para transformar dados categóricos em numéricos:

- **One-hot encoding** – consiste em criar uma nova coluna para cada valor único (categoria) da variável categórica, onde, cada coluna nova indica a presença (com valor 1) ou ausência (com valor 0) daquela categoria na amostra;
- **Dummy encoding** - semelhante ao *one-hot encoding*, excepto que, esta técnica cria uma coluna a menos com o objectivo de evitar a multicolinearidade devido a

existência de colunas altamente correlacionadas, o que pode causar problemas em alguns modelos de aprendizagem de máquina, como a regressão linear.

2) Seleção da medida de similaridade

Nesta etapa, é selecionada a métrica para avaliação da similaridade entre as amostras que compõe os *clusters* gerados pelas técnicas de *Clustering* aplicadas, com vista a avaliar a qualidade e a adequação dos *clusters* obtidos em relação aos dados e ao problema em questão;

3) Execução do algoritmo de *Clustering*

O algoritmo é aplicado sobre os dados, otimizado e avaliado através da métrica selecionada;

4) Avaliação dos resultados

Esta fase compreende a avaliação dos resultados com vista a determinar se os mesmos satisfazem as necessidades do problema que se pretende resolver;

5) Interpretação dos *clusters* identificados

Nesta fase, os resultados são interpretados por meio da análise dos *clusters* gerados pelas diferentes técnicas com vista a verificar se estes estão alinhados com as características conhecidas.

Existem diferentes algoritmos de *Clustering*, cada um com suas próprias características e pressupostos sobre a estrutura dos dados. Neste trabalho serão utilizados 3 algoritmos de *Clustering* descritos por Rai e Singh (2010) sobre os dados colectados, nomeadamente:

▪ *K-means*

Neste algoritmo, cada *cluster* é representado por um centróide que é a média das amostras pertencentes ao mesmo, sendo que, o número de *clusters*, k , que se pretende criar representa a quantidade de centróides que serão criados. Para formação de *clusters*, o algoritmo procede da seguinte forma:

- 1) São selecionadas as posições iniciais dos centróides.

- 2) Para cada amostra no conjunto de dados, o algoritmo calcula a distância das mesmas e cada um dos centróides e de seguida, cada amostra é alocada para o *cluster* representado pelo centróide mais próximo.
- 3) Após a primeira iteração (passo 2), a posição dos centróides é recalculada como a média das amostras do *cluster* que o mesmo representa.
- 4) Os passos 2 e 3 são repetidos até que a posição dos centróides não altere significativamente.

Segundo o autor, este algoritmo funciona apenas com atributos numéricos e pode ser afectado negativamente por *outliers*.

- ***K-medoids***

Este algoritmo é semelhante ao *K-means*, no entanto, ao invés de centróides, os *clusters* são representados por k amostras no conjunto de dados. Diferente do *k-means*, este algoritmo não é limitado pelo tipo de atributo e é a existência de *outliers* tem menor impacto na análise visto que a escolha do representante de cada *cluster* é determinada pela maior parte das amostras.

- **DBSCAN (Density Based Spatial Clustering of Application with Noise)**

Para este algoritmo, é essencial a distribuição das amostras, visto que a sua proximidade é utilizada para medir a densidade e conectividade no conjunto de dados.

Para cada amostra, o algoritmo calcula o número de amostras mais próximas (vizinhos) de acordo com um raio definido e verifica se o mesmo é igual ao mínimo de amostras vizinhas definido. caso seja igual ou maior, o ponto é identificado como sendo um objecto central e todas as amostras acessíveis a partir do mesmo podem ser conectados, se tornando um *clusters*. Todos os pontos que não estejam conectados a objectos centrais são considerados *outliers*.

Este algoritmo não é limitado pelo tipo de dados.

2.4. Métricas de Avaliação de Modelos de *Machine Learning*

Pereira, M. A. P. (2022). Previsão de volumes pulmonares não mobilizáveis com base em parâmetros espirométricos (Doctoral dissertation).

Existem várias métricas aplicáveis para avaliar modelos de *Machine Learning*, no entanto, para avaliação de modelos com a finalidade de estimar valores, o autor Pereira (2022) destaca as seguintes métricas:

- **Erro Médio Absoluto (MAE)** - representa o módulo da diferença média entre o valor real e o previsto;
- **Erro Quadrático Médio (MSE)** - assim como o erro médio absoluto, esta métrica calcula a diferença média entre o valor real e o valor previsto, entretanto, a diferença média é elevada ao quadrado, o que permite penalizar os erros maiores em relação ao valor real;
- **Coeficiente de Determinação r^2** - o valor desta métrica indica se o modelo se ajusta bem aos dados e explica mais variância. É muito fácil de interpretar porque os seus valores variam entre 0 e 1, sendo que 1 representa que o modelo explica perfeitamente a variância dos dados na amostra.

2.5. Ferramentas e Tecnologias para o Desenvolvimento de uma Aplicação Web

De acordo com Santiago et al (2020), o desenvolvimento de uma aplicação Web pode ser descrito como a construção, criação e manutenção de aplicações na Internet ou em uma intranet. Diz ainda o autor, que este desenvolvimento envolve um *front-end (client-side)*, responsável pelo código que interage com o utilizador final (cliente) através de um navegador, e um *back-end (server-side)*, que é o código executado no servidor Web, que encapsula a lógica de negócios da aplicação e a interação com a base de dados.

Cada componente da aplicação Web envolve o uso de diferentes ferramentas, cuja escolha depende da complexidade e/ou necessidades específicas do projecto, sendo que, em alguns casos, a escolha poderá se basear em preferências do desenvolvedor.

2.5.1. *Front-end*

O *front-end* de uma aplicação é a parte responsável pela interação directa com os utilizadores finais da aplicação, sendo a mesma responsável pela interface da aplicação e a experiência do utilizador. Segundo o Santiago, Veras e Aragão, o *front-end* concentra-se *no layout, design e interatividade da aplicação Web*, sendo assim, essencial a utilização de linguagens como HTML (linguagem de marcação), CSS (linguagem de estilo) e JavaScript (linguagem de script/programação) são utilizadas. O autor descreve estas linguagens do seguinte modo:

- **HTML (*HyperText Markup Language*)** - é a linguagem que fornece a estrutura para o conteúdo de um site e define palavras, títulos, parágrafos, imagens etc. O HTML consiste em um conjunto de tags pré-definidas, representando funções diferentes que transformam o conteúdo em um formato legível na tela;
- **CSS (*Style Cascading Sheets*)** - é uma “folha de estilos” que descreve como os elementos HTML aparecerão em uma página da Web. Usa-se o CSS para controlar a apresentação, o estilo e a formatação dos elementos HTML em um site, configurando cores, bordas, imagens de fundo, etc;
- **JavaScript** - é a linguagem responsável pelo controle do comportamento de uma página Web. O uso do JavaScript torna os sites interativos, manipulando as especificidades dos elementos HTML e CSS. Com JavaScript, um usuário pode clicar em um botão e disparar uma ação, fazer o site rolar para a parte inferior ou exibir fotos aleatórias, por exemplo.

2.5.2. *Back-end*

De acordo com Santiago et al (2020), o *back-end* consiste no servidor Web que hospeda o site, uma aplicação para executá-lo e uma base de dados para armazenar informações. Os programas escritos nesta parte garantem que o servidor, a aplicação Web e a base de dados funcionem correctamente em conjunto de acordo com as necessidades de negócio dos *stakeholders*, fornecendo soluções de programação eficientes e seguras. De acordo com autor, as linguagens utilizadas no *back-end* incluem Python, PHP e Java.

3. Capítulo III – Caso de Estudo

Este capítulo apresenta o a síntese da informação obtida por meio de entrevistas e as acções tomadas com base na mesma. A informação descrita inclui o cenário atual da avaliação de imóveis em Maputo, destacando o processo tradicional de avaliação e os desafios enfrentados pelos profissionais do sector, e a proposta de solução com ênfase no desenvolvimento de uma aplicação web com recurso a diferentes tecnologias como Google Colab para o treinamento de modelos, Django para o desenvolvimento da aplicação web, e MySQL como base de dados.

3.1. Descrição do Cenário Actual

Normalmente, para efectuar a avaliação de um imóvel em Maputo o interessado deve, primeiramente, procurar um avaliador que será responsável por todo o processo de avaliação. Dando seguimento ao processo, o avaliador realiza uma inspeção detalhada do imóvel através da qual efectua a colecta de dados relevantes para a avaliação. De seguida, este procede com uma análise comparativa do imóvel em análise com outros imóveis vendidos ou alugados que apresentam dados semelhantes aos que foram obtidos durante a inspeção. Nesta análise, o avaliador utiliza o preço dos imóveis semelhantes como base e o ajusta de acordo com as características únicas do imóvel, estimando assim o valor mais adequado para o mesmo.

No entanto, considerando o processo acima descrito é muito importante escolher um avaliador confiável para garantir uma avaliação precisa.

3.2. Resultados Obtidos no Campo através de Entrevista

Com vista obter mais informações para enriquecer a solução proposta, efectuou-se uma entrevista (roteiro colocado no apêndice 2) a um profissional da área com aproximadamente 20 anos de experiência no mercado imobiliário, especializado em corretagem de imóveis e gestão de base de dados de imóveis. A partir da entrevista, pode-se constatar o seguinte:

1) Método aplicado na avaliação de imóveis:

Para realizar uma avaliação, o entrevistado utiliza o método comparativo de dados de mercado para avaliar imóveis, através do qual, o mesmo efectua uma comparação entre os dados colectados sobre o imóvel que se pretende avaliar e os dados de

outros que possuem características semelhantes em sua base de dados. Com base neste histórico, este efectua ajustes baseando-se características únicas do imóvel em avaliação.

Diz ainda o entrevistado, que o número de amostras utilizadas para comparação não é fixo, uma vez que o mesmo depende da experiência e confiança do avaliador.

2) Variáveis observadas no processo:

- *Design* do imóvel – se é um imóvel *flat* (Apartamento), uma vivenda (casa isolada) ou uma geminada (casa que compartilham uma parede com outra unidade);
- Finalidade – o propósito da transação, venda ou aluguer do imóvel;
- Localização do imóvel;
- Tipologia - a configuração e ao número de quartos do imóvel, como T1 (um quarto), T2 (dois quartos), T3 (três quartos), etc.
- Móveis;
- Guarda-fatos;
- Ar-condicionado;
- Garagem;
- Cozinha Americana;
- Número de casas de banho;
- Suite;
- Vista;
- Piscina; e
- Obras.

Para além dos factores indicados acima, o entrevistado mencionou ainda a sazonalidade do local, como um factor que pode impactar significativamente no resultado da avaliação, como por exemplo, o final do ano.

3) Desafios e Necessidades:

O entrevistado destaca a sua dificuldade em entender, dominar e aplicar novas tecnologias para ganhar acesso ao mercado digital, razão pela qual, o mesmo apresenta a necessidade de uma plataforma digital segura para registro e

disponibilização de imóveis na internet, com vista a facilitar a primeira abordagem de clientes.

Tendo em conta as constatações acima, pode-se concluir que para o desenvolvimento de uma aplicação web focada na avaliação de imóveis em Maputo deve-se tomar em conta os seguintes pontos:

- O método aplicado para avaliação deve ser de fácil perceção;
- A aplicação deve adaptar-se às necessidades dos avaliadores e do mercado, proporcionando uma ferramenta complementar ao trabalho manual de avaliação.
- Fornecer uma experiência de usuário otimizada para diferentes dispositivos.

3.3. Comparação entre os *Stacks* Tecnológicos

Para implementação da solução proposta serão utilizadas algumas ferramentas, cuja escolha será baseada nos seguintes critérios:

- Facilidade de Uso;
- Escalabilidade;
- Desempenho;
- Caso de uso;

Para a escolha do *stack* tecnológico adequado, serão consideradas tecnologias com as quais o desenvolvedor está familiar, nomeadamente:

a) Base de Dados

- **PostgreSQL** - um sistema de gestão de base de dados relacional avançado de código aberto;
- **MySQL** - um dos sistemas de gestão de base de dados relacional mais populares do mundo e ensinado aos alunos da Universidade;
- **MongoDB** - uma base de dados NoSQL orientado a documentos que armazena dados em formato BSON.

b) Ambiente de Desenvolvimento do Modelo de Avaliação

- **Jupyter Notebook** – uma aplicação web que permite a criação e compartilhamento de documentos que contenham código, visualizações e texto narrativo;
- **Google Colab** - um serviço do Jupyter notebook hospedado pelo Google que permite a execução de código Python a partir do navegador.

c) Frameworks para Desenvolvimento da Aplicação

- **Django** - *framework* web de alto nível em Python que promove o desenvolvimento rápido e limpo.
- **Flask** - microframework web em Python que é leve e fácil de usar.
- **Node.js com o Express.js** – neste caso, o Node.js é um runtime de JavaScript, enquanto Express.js é o *framework* web minimalista para Node.js.

Com os critérios e possíveis tecnologias identificadas, apresenta-se a seguir as tabelas comparativas do para escolha das tecnologias adequadas para cada componente da aplicação.

a) Base de Dados

Critério	PostgreSQL	MySQL	MongoDB
Desempenho	Alto	Alto	Variável (excelente para leituras rápidas)
Escalabilidade	Boa	Boa	Excelente
Facilidade de Uso	Moderada	Alta	Moderada
Casos de Uso	Aplicações empresariais, transações financeiras	Aplicações web, transações financeiras	Aplicações com dados não estruturados, Big Data

Tabela 3 - Comparação de Bases de Dados

b) Ambiente de Desenvolvimento do Modelo de Avaliação

Critério	Google Colab	Jupyter Notebook
Desempenho	Alto (com suporte a GPUs)	Alto (local)
Escalabilidade	Alta (escalável com Google Cloud)	Moderada (local)
Facilidade de Uso	Alta	Alta
Casos de Uso	Desenvolvimento e treinamento de modelos colaborativos e a partir de diferentes dispositivos	Desenvolvimento de modelos locais

Tabela 4 - Comparação de ambientes de desenvolvimento do modelo de avaliação

c) *Frameworks* para Desenvolvimento da Aplicação

Critério	Django	Flask
Desempenho	Alto	Alto
Escalabilidade	Boa	Boa
Facilidade de Uso	Alta	Alta
Casos de Uso	Aplicações completas, rápido desenvolvimento	APIs simples, micro-serviços

Tabela 5 - Comparação de *frameworks* para o desenvolvimento da aplicação proposta

Da comparação efectuada nas tabelas acima, foram seleccionadas as seguintes tecnologias:

- **Google Colab** – a ferramenta foi escolhida por permitir o desenvolvimento colaborativo e treinamento de modelos com suporte a GPUs, o que poderá ajudar com a necessidade crescente de poder de processamento;
- **Django** – escolhido devido a facilidade de integração com ferramentas de *machine learning* em python e domínio da ferramenta pelo desenvolvedor;
- **MySQL** – será utilizada a base de dados MySQL devido a familiaridade do desenvolvedor e a sua ampla adoção em aplicações web que poderá ser útil para futuras colaborações.

3.4. Proposta de Solução

De modo a solucionar o problema descrito no Capítulo I propõe-se a criação de uma aplicação web que disponibilize aos utilizadores um formulário para o preenchimento dos dados referentes ao imóvel que pretendem avaliar, e após a submissão a aplicação apresenta os resultados da avaliação.

A implementação da solução irá compreender as seguintes actividades:

- **Preparação dos dados** – esta actividade deverá incluir a colecta e limpeza dos dados;
- **Engenharia e seleção de atributos** - avaliar de atributos e a extração de novos caso seja necessário/possível;

- **Aplicação de um algoritmo de *clustering*** – aplicar o algoritmo escolhido para criar *clusters* das amostras coletadas baseado nos atributos das casas.
- **Cálculo da média dos valores das casas em cada cluster** - para cada cluster será efectuado o cálculo da média que irá servir de estimativa para o valor de novas casas alocadas à cada *cluster*.
- **Elaboração do módulo de avaliação de imóveis** – será elaborado o módulo de avaliação que será posteriormente integrado na aplicação para utilização pelo utilizador final;
- **Desenvolvimento da aplicação Web** - após a elaboração do módulo de avaliação, o próximo passo será o desenvolvimento da aplicação web. Esta aplicação irá permitir aos utilizadores finais introduzir os dados do imóvel e receber a estimativa do seu valor de mercado, baseada nos *clusters* e nas análises realizadas.

Para realização destas actividades, serão combinados uma metodologia robusta de preparação e análise de dados com a implementação de um modelo utilizando a técnica de aprendizagem de máquina *clustering* que irá permitir oferecer uma solução inovadora e eficiente para os intervenientes do mercado imobiliário em Maputo, promovendo transações mais transparentes e justas.

4. Capítulo IV – Construção do Modelo de Avaliação Automático Baseado em *Clustering*

Neste capítulo serão apresentados os procedimentos aplicados para o desenvolvimento do módulo de avaliação de imóveis com recurso ao Google Colab Notebook, desde a colecta, pré-processamento e transformação dos dados até a aplicação de algoritmos de *clustering* e a apresentação dos resultados.

4.1. Colecta de dados

Para realização do projecto, solicitou-se um conjunto de dados referentes a imóveis junto da Empresa Multicasa. O conjunto de dados possui informações sobre diferentes características de imóveis incluindo a sua tipologia, *design* e outros aspectos observados para avaliação de imóveis.

O conjunto de dados possui 287 amostras de imóveis em Maputo com os seguintes atributos:

- **Design do imóvel** – se é um imóvel *flat* (Apartamento), uma vivenda (casa isolada) ou uma geminada (casa que compartilham uma parede com outra unidade);
- **Finalidade** – o propósito da transação, venda ou aluguer do imóvel;
- **Bairro** – o bairro em que o imóvel se localiza;
- **Tipologia** – referente a configuração e ao número de quartos do imóvel, como, T0, T1 (um quarto), T2 (dois quartos), T3 (três quartos), etc.
- **Móveis** – indica se o imóvel possui móveis;
- **Guarda-fatos** – indica se o imóvel possui guarda-fatos;
- **Ar-condicionado** – indica se o imóvel possui ar-condicionado;
- **Garagem** – indica se o imóvel possui garagem;
- **Cozinha Americana** – indica se o imóvel possui cozinha americana;
- **Número de casas de banho** – indica o nº de casas de banho que o imóvel possui;
- **Suite** – indica se o imóvel possui uma suite;
- **Vista** – indica se o imóvel possui uma boa vista;
- **Piscina** – indica se o imóvel possui uma piscina;
- **Obras** – indica se o imóvel possui obras em andamento; e

- **Preço** – o valor do imóvel.

4.2. Pré Processamento de dados

Para pré-processar os dados, primeiro identificou-se os valores em falta, calculando a percentagem de valores em falta por atributo, conforme o recorte apresentado na figura 1.

Missing Ratio (%)	
Garagem	0.696864
Suite	0.696864
Moveis	0.348432
Ar-Condicionado	0.348432
Cozinha Americana	0.348432
Vista	0.348432
Piscina	0.348432
Obras	0.348432
Preço (MZN)	0.348432

Figura 1 - Percentagem de valores em falta de atributos com pelo menos um valor em falta

Uma vez que o nº de valores em falta em cada atributo é inferior a 5%, as amostras referentes a estes valores foram eliminadas permanecendo apenas 283 amostras.

De seguida, foram computados os valores de assimetria (*skewness*) dos atributos numéricos (preço e o nº de casas de Banho). Ambos atributos apresentaram valores uma assimetria grande, com valores acima de 1, pelo que foi aplicada a transformação logarítmica. Este processo visa reduzir o impacto de *outliers* no modelo criado.

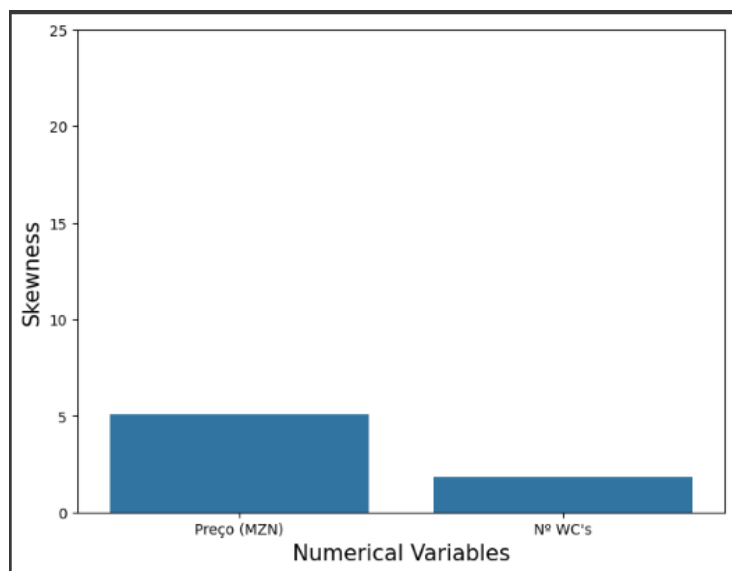


Figura 2 - Gráfico com a assimetria calculada por atributo

Destacar que estes atributos não possuem uma correlação forte entre eles, uma vez que o nº de WC's não tem uma influência directa no valor das amostras no conjunto de dados obtidos, conforme apresentado na figura 3.

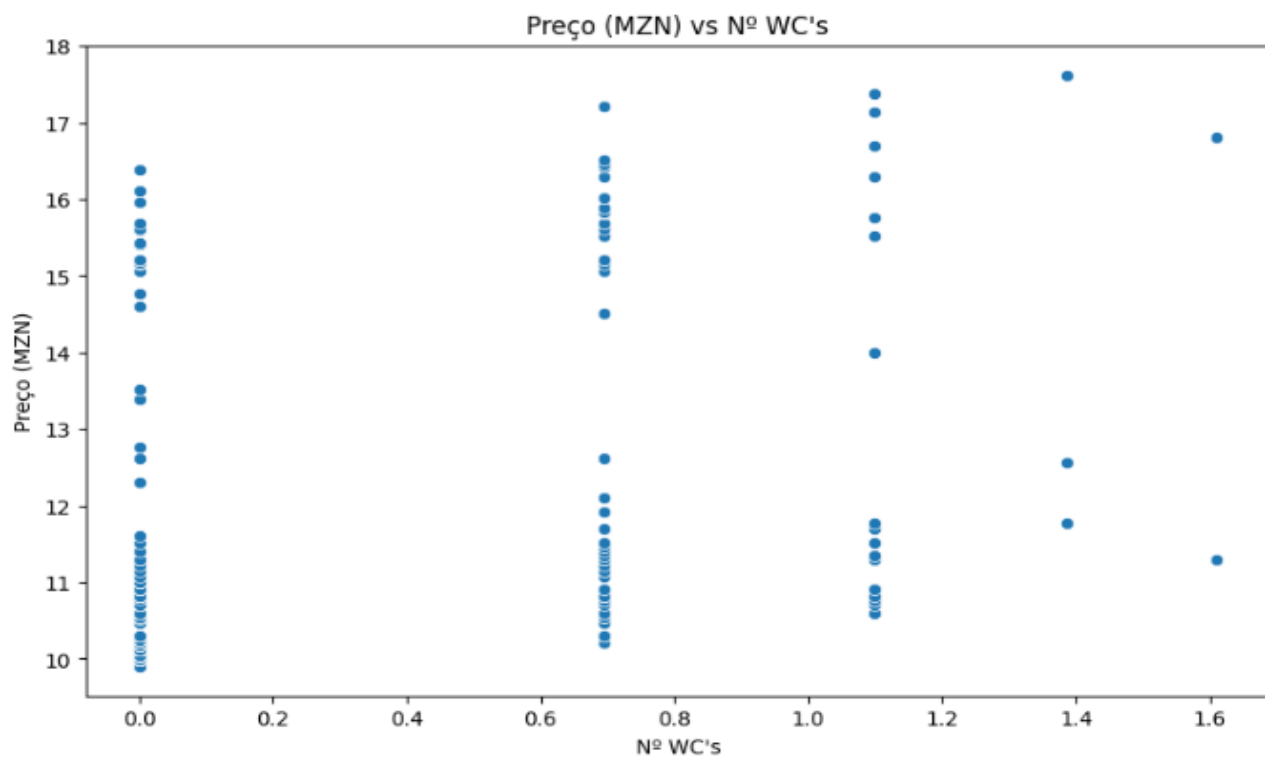


Figura 3 - Visualização do preço em relação ao nº de casas de banho

Após aplicar a transformação logarítmica para reduzir a assimetria, obteve-se dados com valor de assimetria relativamente menor, conforme a figura 4.

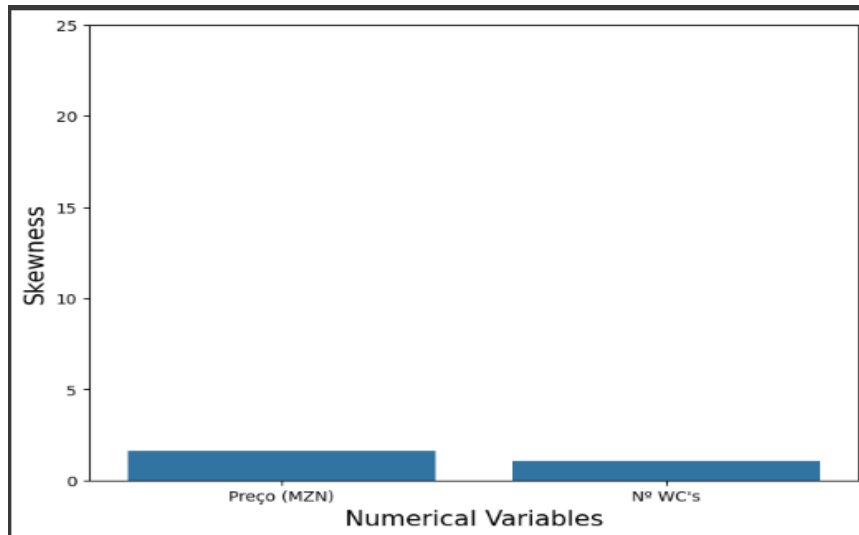


Figura 4 - Resultados obtidos após a transformação logarítmica

Com a assimetria reduzida, procedeu-se com a transformação dos dados, descrita na próxima etapa.

4.3. Transformação dos dados

Nesta etapa, os atributos categóricos foram transformados em numéricos considerando o seguinte raciocínio:

- Os atributos com apenas dois valores únicos foram directamente mapeados para valores binários;
- Os atributos com mais de dois atributos foram transformados através da técnica *one-hot encoding* que consiste em converter as categorias do atributo em novas colunas binárias cuja designação é resultante da combinação do nome da coluna original e a categoria a que se refere.

Como resultado deste processo, o nº de atributos aumentou de 15 para 56. No entanto, somente 37 atributos foram seleccionados, baseado no grau de correlação que os mesmos

possuem com o valor do imóvel. Através da análise de correlação foi possível determinar todos atributos com o grau de correlação entre -0.05 e 0.05 e remove-los da tabela.

De seguida, os dados foram separados em *inputs* e *targets*, os *inputs* foram normalizados de modo a uniformizar os valores de todos atributos, possuindo deste modo, valores em um intervalo semelhante. Este processo ajuda a melhorar a performance dos modelos de *machine learning*.

4.4. Aplicação das técnicas de *Clustering*

Com os dados prontos para treinar e testar os modelos, procedeu-se com a aplicação das técnicas de *clustering* apresentadas anteriormente (K-means, K-medoids e DBSCAN). Para aplicar as técnicas, procedeu-se da seguinte forma:

- Os dados foram separados em dados para o treino e teste do modelo;
- Segmentou-se os dados de acordo com a finalidade de cada amostra (venda ou aluguer);
- Para cada segmento, aplicou-se a técnica de *clustering* e de seguida, os rótulos de cada *cluster* foram atribuídos ao conjunto de dados antes da segmentação.
- Após atribuir rótulos para cada amostra, efectuou-se o cálculo da média para cada *cluster* tomando em conta a finalidade da amostra. As médias foram armazenadas em um dicionário, tornando-se deste modo, o ponto essencial para estimação de novos valores;
- O modelo de *clustering* foi aplicado nos dados de teste e as médias foram utilizadas para estimar o valor. Os resultados obtidos foram avaliados com base nas seguintes métricas:
 - r^2 – o seu valor indica se o modelo se ajusta bem aos dados e explica mais variância. Quanto maior o valor, melhor;
 - **Erro Médio Absoluto (MAE) e Erro Quadrático Médio (MSE)** - representam o quão próximas as previsões do modelo estão dos valores reais. Quanto mais baixo for o valor, melhor o desempenho do modelo.

Os resultados obtidos através da aplicação de cada técnica serão apresentados nas figuras (recortes) abaixo. Para interpretar os resultados,

```
Cluster -1.0:
  Mean Absolute Error: 0.7063775345885156
  Mean Squared Error: 0.961424346291022
  R-squared: 0.7732445248504157
Cluster 0.0:
  Mean Absolute Error: 0.2937550383134457
  Mean Squared Error: 0.13789668943573272
  R-squared: -2.220446049250313e-16
Cluster 1.0:
  Mean Absolute Error: 0.39313746446640324
  Mean Squared Error: 0.47082368997100404
  R-squared: -2.220446049250313e-16
```

Figura 5 - Resultados obtidos após aplicar a técnica DBSCAN

```
Cluster 0.0:
  Mean Absolute Error: 0.3918689041491231
  Mean Squared Error: 0.28209123893800675
  R-squared: 0.8896669733819789
Cluster 1.0:
  Mean Absolute Error: 0.29542024872602646
  Mean Squared Error: 0.12613583521699273
  R-squared: -0.44510457020261374
Cluster 2.0:
  Mean Absolute Error: 0.4979147643983349
  Mean Squared Error: 0.3935641729796407
  R-squared: 0.9264827826025795
```

Figura 6 - Resultados obtidos após aplicar a técnica K-Medoids

```
Cluster 1.0:
  Mean Absolute Error: 0.5230785748219933
  Mean Squared Error: 0.9433430306188394
  R-squared: 0.8208016838127008
Cluster 2.0:
  Mean Absolute Error: 0.4613957498766053
  Mean Squared Error: 0.38002629742956184
  R-squared: 0.9052093121063274
Cluster 0.0:
  Mean Absolute Error: 0.3527343410093787
  Mean Squared Error: 0.2510139215284576
  R-squared: 0.9244759033174706
```

Figura 7 - Resultados obtidos após aplicar a técnica K-Means

De acordo com os resultados apresentados nas figuras acima, escolheu-se a técnica *K-Means* para implementar o modelo. Encontre abaixo a interpretação dos resultados obtidos através desta técnica:

- O *cluster 1* apresenta um desempenho aceitável, mas possui erros mais elevados em comparação com os outros *clusters*. O valor de r^2 indica que o modelo explica aproximadamente de 82% da variância dos dados, no entanto, o resultado é modesto

se comparado aos *clusters* restantes. Este é o *cluster* com o pior desempenho se comparado com os outros;

- O *cluster* 2 é melhor se comparado ao *cluster* 1 uma vez que a sua explicabilidade é melhor que o mesmo (90%), com uma diferença de aproximadamente igual a 8%. Adicionalmente, este modelo possui menores erros;
- Por fim, temos o *cluster* 0 que tem o melhor desempenho entre todos independentemente da métrica observada. A sua explicabilidade atinge um valor aproximado de 92%;

Observando a variação dos preços em cada *cluster* (figura 8), pode-se ver que no *cluster* 0 do segmento de dados com a finalidade de aluguer, o preço máximo em alguma das amostras é muito elevado se comparado com o valor mínimo e a média. Com isto podemos concluir que as amostras que apresentam um preço próximo ao máximo dentro do *cluster* possuem algumas características desconhecidas que poderiam gerar um novo *cluster* mais adequado para eles, se tornando *outliers* dentro do *cluster*. Se obtivermos uma quantidade maior de dados, com mais atributos específicos dos imóveis, será possível criar *clusters* com melhor performance.

Finalidade -> Aluguel					
	Cluster	Preço Médio	Preço Mínimo	Preço Máximo	Quantidade
0	0.0	61468.26	20000.00	1300000.00	16
1	1.0	33026.31	20000.00	55000.00	20
2	2.0	62646.89	27000.00	130000.00	11
=====					
Finalidade -> Venda					
	Cluster	Preço Médio	Preço Mínimo	Preço Máximo	Quantidade
0	0.0	5376802.02	4900000.00	5900000.00	2
1	1.0	5916079.78	3500000.00	10000000.00	2
2	2.0	9507798.66	3800000.00	30000000.00	6
=====					

Figura 8 - Resumo referente ao valor dos imóveis em cada um dos clusters nos dados de teste

4.5. Elaboração do Módulo de Avaliação de Imóveis

Após escolher o modelo de mais adequado e calcular a média, o próximo passo tomado é a elaboração do módulo que será integrado na aplicação web para estimar valores de imóveis cujas características são introduzidas pelos utilizadores.

Para criar o módulo, primeiro criou-se um ficheiro que possui o modelo de normalização de dados, o modelo de *clustering* e o valor estimado para cada *cluster* utilizando a biblioteca “*Pickle*”. De seguida, criou-se um novo ficheiro em python, no qual foi criado um módulo com dois métodos, um para o pré-processamento dos dados introduzidos pelo utilizador e o outro para estimar o preço do imóvel aplicando o modelo de *clustering* e as médias armazenadas anteriormente.

Destacar que cada método utiliza as informações armazenadas através da biblioteca “*Pickle*”, sendo que, para a sua utilização, as informações devem ser carregadas utilizando a mesma biblioteca.

5. Capítulo V – Desenvolvimento do Protótipo

Neste capítulo será desenhado e desenvolvido o protótipo da aplicação proposta com base nas informações obtidas no caso de estudo e na revisão de literatura.

5.1. Levantamento de Requisitos

Nesta fase, serão identificados os requisitos necessários para o desenvolvimento da aplicação. Os requisitos serão classificados em duas categorias: funcionais e não funcionais. Nas tabelas dos requisitos, serão indicados quais são prioritários e não.

5.1.1. Requisitos Funcionais

Os requisitos funcionais indicam o que a aplicação deverá fazer, isto é, as suas funcionalidades. A aplicação proposta deverá possuir as funções descritas na tabela a seguir:

Identificador	Requisito	Descrição	Prioridade
RF01	Introduzir dados do imóvel	Permitir ao utilizador preencher o formulário com os dados do imóvel que pretende avaliar	Prioritário
RF02	Avaliar o valor do imóvel	Avaliar o imóvel de acordo com os dados introduzidos pelo utilizador	Prioritário
RF03	Apresentar os resultados da avaliação	Visualizar os resultados da avaliação na tela do utilizador	Prioritário
RF04	Armazenar os dados de imóveis avaliados pela aplicação	Armazenar os dados dos imóveis avaliados pela aplicação	Não prioritário
RF05	Extrair o relatório da avaliação	Permitir exportar o resultado da avaliação em formato PDF	Não prioritário

Tabela 6 - Descrição dos requisitos funcionais

5.1.2. Requisitos Não Funcionais

Os requisitos não funcionais diferem dos requisitos funcionais pelo facto de não estarem directamente ligados a funcionalidades específicas da aplicação. Os requisitos identificados estão descritos na tabela abaixo:

Identificador	Requisito	Descrição	Prioridade
RNF01	Multiplataforma	A aplicação deve ser acessível através de diferentes dispositivos (smartphones, tablets, laptops)	Prioritário
RNF02	Eficiência	Carregar informações rapidamente e apresentar resultados de maneira eficiente;	Prioritário
RNF03	Facilidade na utilização	A interface de usuário deve ser amigável e fácil de usar	Prioritário
RNF04	Escalabilidade	A aplicação deve permitir acomodar um número crescente de usuários e dados, conforme a demanda aumenta	Não prioritário

Tabela 7 - Descrição dos requisitos não funcionais

5.2. Análise

Nesta fase, será realizada uma análise detalhada dos requisitos levantados para definir as especificações técnicas e funcionais da aplicação. A análise incluirá a análise dos requisitos e a modelagem de dados.

5.2.1. Análise dos Requisitos

A partir da análise dos requisitos funcionais, foi possível transforma-los em diferentes especificações técnicas e funcionais, conforme apresentado nas alíneas a seguir:

a) Especificações Funcionais

- RF01 (Introduzir dados do imóvel) - formulário de entrada de dados acessível via interface web;
- RF02 (Avaliar o valor do imóvel) - implementação do módulo de avaliação que utiliza o modelo de *clustering* para estimar o valor do imóvel;
- RF03 (Apresentar os resultados da avaliação) - interface de usuário para exibir os resultados da avaliação de forma clara e detalhada;
- RF04 (Armazenar os dados de imóveis avaliados pela aplicação) - armazenamento de dados de avaliações em uma base de dados;
- RF05 (Extrair o relatório da avaliação) - função para exportar os resultados da avaliação em formato PDF, acessível via interface web.

b) Especificações Técnicas

- *Backend*:
 - i. Framework: Django
 - ii. Linguagem: *Python*
 - iii. Base de Dados: MySQL
- *Frontend*:
 - i. HTML, CSS e JavaScript
 - ii. *Bootstrap* para design responsivo
- Modelo de Avaliação:
 - i. Ferramenta para construção: Google Colab
 - ii. Bibliotecas: Pandas, Sklearn, Matplotlib/Seaborn, Pickle
 - iii. Algoritmo de *Clustering*: *K-means*

5.2.2. Casos de Uso

Os casos de uso são uma técnica utilizada para definir claramente o comportamento do sistema em resposta às ações dos usuários, permitindo uma melhor comunicação entre os desenvolvedores e os outros *stakeholders*.

Para melhor descrever os casos de uso são identificados os actores do sistema, que são as entidades que interagem com o sistema. A aplicação proposta possui dois actores, indicados na tabela abaixo:

Actor	Descrição
Utilizador	Representa a todos que pretendem avaliar o seu imóvel através da aplicação

Tabela 8 - Actores do sistema

As tabelas abaixo descrevem os casos de uso da aplicação proposta:

▪ **Caso de uso – Introduzir os dados do imóvel**

Nome	Introduzir os dados do imóvel
Requisitos relacionados	RF01
Descrição	Permite ao utilizador inserir os dados do imóvel que pretende avaliar
Actor	Utilizador
Actor secundário	Aplicação
Prioridade	Prioritário
Fluxo Principal	
Actor	Descrição
Utilizador	O utilizador acede o formulário para introdução dos dados
Utilizador	O utilizador preenche os campos do formulário
Utilizador	O utilizador submete o formulário
Aplicação	A aplicação valida os dados

Tabela 9 - Descrição do caso de uso: Introduzir dados do imóvel

▪ **Caso de uso – Avaliar o valor do imóvel**

Nome	Avaliar o valor do imóvel
Requisitos relacionados	RF02
Descrição	Avaliar o imóvel com base nos dados introduzidos pelo utilizador
Actor	Utilizador
Actor secundário	Aplicação
Prioridade	Prioritário
Fluxo Principal	
Actor	Descrição
Utilizador	O utilizador solicita a avaliação do imóvel
Aplicação	A aplicação processa os dados e aplica o modelo de avaliação
Aplicação	A aplicação retorna o valor estimado do imóvel

Tabela 10 - Caso de uso: Avaliar o valor do imóvel

▪ **Caso de uso – Apresentar o resultado da avaliação**

Nome	Apresentar o resultado da avaliação
Requisitos relacionados	RF03
Descrição	Visualizar o resultado da avaliação do imóvel
Actor	Utilizador
Actor secundário	Aplicação
Prioridade	Prioritário
Fluxo Principal	
Actor	Descrição
Aplicação	A aplicação apresenta o resultado da avaliação na <i>interface</i> do utilizador
Utilizador	O utilizador pode rever os resultados e, caso seja necessário, efectuar uma nova avaliação

Tabela 11 - Caso de uso: Apresentar o resultado da avaliação

▪ **Caso de uso – Armazenar os dados dos imoveis avaliados**

Nome	Armazenar os dados dos imoveis avaliados
Requisitos relacionados	RF04
Descrição	Armazenar os dados dos imoveis avaliados através da aplicação
Actor	Aplicação
Prioridade	Prioritário
Fluxo Principal	
Actor	Descrição
Aplicação	A aplicação armazena os dados do imóvel e o valor estimado na base de dados depois da avaliação

Tabela 12 - Caso de uso: Armazenar os dados dos imoveis avaliados

▪ **Caso de uso – Extrair o relatório da avaliação**

Nome	Extrair o relatório da avaliação
Requisitos relacionados	RF05
Descrição	Exportar o resultado da avaliação em um ficheiro do formato PDF
Actor	Utilizador
Actor secundário	Aplicação
Prioridade	Não prioritário
Fluxo Principal	
Actor	Descrição
Utilizador	O utilizador solicita a exportação do relatório
Aplicação	A aplicação gera o relatório no formato PDF
Utilizador	O utilizador inicia o <i>download</i> do relatório no seu dispositivo

Tabela 13 - Caso de uso: Extrair o relatório da avaliação

Para além da descrição textual, os casos de uso podem ser apresentados através de outros modelos, conforme ilustra o diagrama abaixo:

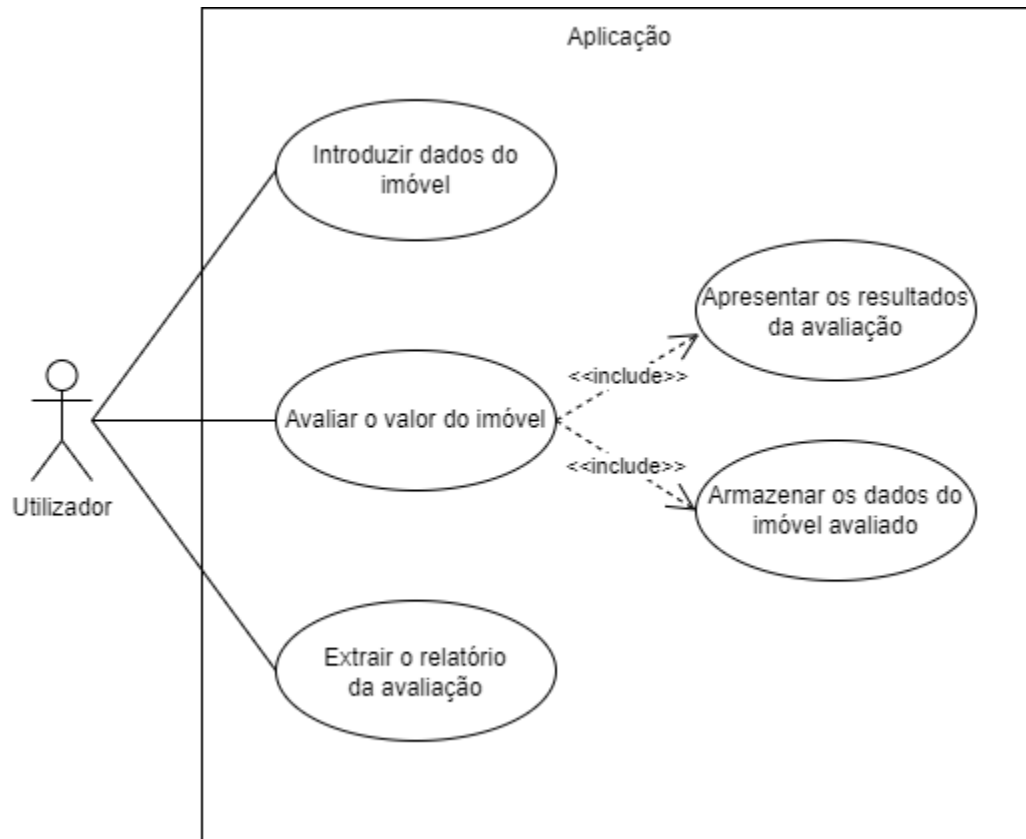


Figura 9 - Diagrama de casos de uso

5.2.3. Modelagem de Dados

A modelagem de dados é uma actividade importante para organizar e estruturar os dados que serão armazenados e processados pela aplicação. A aplicação irá possuir as seguintes entidades:

- **Imóvel** – esta entidade possui os atributos do imóvel, nomeadamente: bairro, tipologia, design, finalidade, e outras características adicionais do imóvel;
- **Avaliação** – é a entidade responsável pelas informações referentes as avaliações realizadas pela aplicação e inclui a data, o valor estimado e o *cluster* ao qual o imóvel foi alocado.

A relação entre estas entidades é de um-para-um, sendo que, um imóvel deverá possuir apenas uma avaliação.

5.3. Design

A aplicação será desenvolvida utilizando uma arquitectura em três camadas, normalmente utilizada para desagregar o código de acordo com a sua utilidade na aplicação, tornando a sua revisão, manutenção e escalabilidade mais fácil. A figura abaixo apresenta a visão geral da arquitectura que será implementada.



Figura 10 - Diagrama da arquitectura em 3 camadas da aplicação

Na figura 9, a camada de aplicação será responsável pela interação com o utilizador final e irá disponibilizar para ele um formulário para entrada de dados do imóvel e permiti-lo visualizar o resultado da avaliação. Por sua vez, a camada de negócio é o local em que será implementada a lógica principal da aplicação, aplicando os métodos do módulo de avaliação para avaliar o imóvel de acordo com os dados recebidos da camada de apresentação e retornando o resultado da mesma. Adicionalmente, esta camada será responsável por enviar os dados do imóvel e avaliação para a camada de dados, onde serão armazenados.

5.4. Implementação

O desenvolvimento da solução proposta compreendeu a criação do módulo de avaliação baseado no modelo de avaliação de imoveis construído com a aplicação de *Clustering* e a implementação da aplicação web em Django.

Após a criação do módulo conforme indicado no ponto 4.5 do Capítulo IV, procedeu-se com o desenvolvimento da aplicação web utilizando o *framework* Django para permitir que os usuários introduzam os dados do seu imóvel e recebam a estimativa do seu valor.

Para o desenvolvimento da aplicação web efectuou-se o seguinte:

- A Modelagem dos dados – as entidades e os respectivos atributos foram definidos;
- Desenvolvimento de *views* e *templates* – para receber dados do utilizador e visualizar os resultados, criou-se as *views* de introdução de dados e visualização de resultados juntamente com *templates* em HTML para interface do utilizador;
- A lógica para utilizar o módulo de avaliação, bem como para armazenar dados e exportar relatórios foi implementada nas *views* do *framework*.

Com os procedimentos descritos acima, foi desenvolvida uma aplicação web que permite aos utilizadores introduzir os dados dos seus imóveis, realizar avaliações, e visualizar os resultados de forma eficiente. Esta abordagem inovadora melhora significativamente a transparência no mercado imobiliário de Maputo.

5.5. Testes

Para garantir a qualidade da aplicação, efectuou-se testes unitários e de integração. Os testes unitários são utilizados para verificar o funcionamento dos diferentes componentes isolados do código fonte, sendo que, neste caso foram utilizados para garantir o bom funcionamento do módulo de avaliação de imóveis individualmente. Os testes de integração utilizados para verificar o funcionamento conjunto de diferentes componentes da aplicação foi aplicado para garantir a boa interação entre as *views*, a base de dados e o módulo de avaliação.

As alíneas abaixo descrevem os testes efectuados e o resultado esperado:

a) Carregamento do formulário de avaliação

Este teste verifica se a *view* do formulário de avaliação retorna o template correto e o status HTTP esperado (200 OK).

```
def test_evaluation_form_get(self):
    url = reverse('evaluations:evaluation_form')
    response = self.client.get(url)
    self.assertEqual(response.status_code, 200)
    self.assertTemplateUsed(response, 'evaluator/evaluation_form.html')
```

Figura 11 - Função de Teste: Carregamento do formulário de avaliação

b) Avaliação de Imóveis

Este teste verifica se a função responsável por retornar a avaliação retorna o valor esperado.

```
class EvaluationModelTest(unittest.TestCase):

    def setUp(self):
        path = (...) # caminho para localização do modelo
        self.eval_model = Evaluation_model(path)

    def test_get_prediction(self):
        input_data = pd.DataFrame([{...
        }]) # Dados de entrada simulados para a previsão

        self.eval_model.model = MagicMock()

        prediction = float(self.eval_model.get_prediction(input_data)[0])
        self.assertEqual(prediction, 10425506.86)

if __name__ == '__main__':
    unittest.main()
```

Figura 12 - Função de Teste: Avaliação de Imóveis

c) Submissão e carregamento de dados na base de dados

Este teste verifica se, ao submeter um formulário com dados válidos, a avaliação é armazenada na base de dados.

```
def test_evaluation_form_post_valid(self):
    url = reverse('evaluations:evaluation_form')
    data = {
        'venda_ou_aluguel': 'V',
        'design': 'some_design',
        'bairro': 'some_bairro',
        'tipologia': 't1',
    }
    response = self.client.post(url, data)
    self.assertEqual(response.status_code, 200)
    self.assertTrue(Imovel.objects.exists())
```

Figura 13 - Função de Teste: Submissão e carregamento de dados na base de dados

d) Criação de uma avaliação associada ao imóvel avaliado

Este teste simula a avaliação do imóvel usando o modelo de previsão e a criação de uma avaliação associada a um imóvel.

```
def test_evaluation_result_post(self):
    path = (...) # caminho para localização do modelo
    eval_model = Evaluation_model(path)

    input_data = { ...
    } # Dados de entrada simulados para a previsão

    input_df = pd.DataFrame([input_data])
    predicted_price = float(eval_model.get_prediction(input_df)[0])

    Avaliacao.objects.create(
        cod_imovel=self.imovel,
        preco=predicted_price,
    )

    response = self.client.post(reverse('evaluations:result'), self.imovel_data)

    self.assertEqual(response.status_code, 200)

    avaliacao_criada = Avaliacao.objects.filter(cod_imovel=self.imovel).exists()
    self.assertTrue(avaliacao_criada, "Avaliacao não foi criada corretamente")
```

Figura 14- Função de Teste: Criação de uma avaliação associada ao imóvel avaliado

Os testes realizados na aplicação de avaliação imobiliária permitiram validar as principais funcionalidades da mesma, garantindo que todos os componentes essenciais funcionam corretamente. Através destes testes, foi possível confirmar que a aplicação está apta para uso real, com destaque para os seguintes pontos:

- **Validação da Interação com o Usuário** - Os testes de formulário (GET e POST) mostraram que a aplicação reage correctamente face a inserção de dados pelo usuário. A página de formulário de avaliação foi carregada corretamente, e os dados submetidos foram salvos com sucesso no base de dados.
- **Avaliação de Imóveis** - Através do teste de avaliação de preços, foi confirmado que o modelo efectua avaliações de forma consistente e correta. Os dados de entrada foram processados pelo modelo de aprendizagem de máquina e o preço avaliado foi armazenado correctamente, associando-o ao respectivo imóvel. A integração entre a aplicação Django e o modelo foi validada com sucesso.
- **Desempenho da Aplicação** – Pese embora os testes não tenham sido especificamente voltados para o desempenho, estes não revelaram problemas de latência ou falhas inesperadas, indicando que a aplicação pode operar normalmente com as funcionalidades básicas sem quedas de desempenho evidentes.

A abrangência dos testes foi ampla, abordando as principais funcionalidades da aplicação, desde a interação com o usuário (formulários), até a integração com o modelo de avaliação e o armazenamento de dados na base.

6. Capítulo VI – Conclusões e Recomendações

Neste capítulo, será avaliado o grau de satisfação em relação aos objetivos estabelecidos ao longo deste trabalho. Através desta análise, será possível refletir sobre os resultados alcançados e identificar pontos de melhoria. Para além das conclusões, serão partilhadas recomendações para futuras pesquisas e desenvolvimentos, visando a continuação e aprimoramento das soluções apresentadas, bem como a exploração de novos horizontes dentro do tema abordado.

6.1. Conclusões

Em conclusão, os objetivos deste trabalho foram alcançados com sucesso, resultando no desenvolvimento do protótipo de uma aplicação web para avaliação de imóveis conforme a proposta inicial. A aplicação proposta oferece um suporte significativo a indivíduos sem conhecimento do mercado imobiliário, permitindo-lhes realizar avaliações precisas sobre o valor real dos imóveis que pretendem vender ou adquirir. Adicionalmente, a aplicação beneficia as imobiliárias e seus clientes ao proporcionar uma ferramenta eficiente para o apoio na determinação do valor ótimo dos imóveis negociados.

No âmbito deste trabalho, realizou-se uma pesquisa e análise dos métodos actualmente utilizados na avaliação de imóveis em Maputo, identificando o seu processo e desafios, obtendo-se deste modo, uma base sólida para entender melhor as necessidades para resolução do problema identificado.

Com os resultados da pesquisa e análise do processo, avaliou-se diferentes técnicas de clustering, dentre as quais, a técnica selecionada (k-means) demonstrou-se eficaz no agrupamento dos diferentes imóveis com base em características semelhantes, tornando-se uma peça fundamental para o desenvolvimento da aplicação, uma vez que este modelo é o cérebro da mesma.

Adicionalmente, com os resultados da pesquisa no campo, identificou-se o stack tecnológico mais adequado e procedeu-se com a sua implementação no desenvolvimento de um protótipo funcional que aplica o modelo de clustering para avaliar imóveis, garantindo assim que a mesma seja eficiente e acessível para os utilizadores finais.

Entretanto, este trabalho, para além de alcançar os objectivos propostos, estabelece uma base para futuras pesquisas e desenvolvimentos no sector de imobiliário, no que diz respeito a avaliação de imóveis. Adicionalmente, a solução proposta reduz a dependência dos métodos tradicionais e promove maior transparência no mercado imobiliário, facilitando a negociação e a valorização justa dos imóveis.

O protótipo desenvolvido tem potencial para ser expandido e aprimorado através da combinação do Clustering e outras técnicas de Machine Learning, podendo ser adaptado para diferentes contextos e regiões, contribuindo para a modernização e eficiência do mercado imobiliário.

6.2. Recomendações

Para futuras pesquisas, é sugerido aplicar diferentes técnicas de *Machine Learning* combinadas a *Clustering*, de modo a elaborar modelos específicos de avaliação para cada *cluster*, aumentando a precisão do modelo.

É também sugerida a criação de uma API que permite a diferentes aplicações utilizar o modelo de avaliação para diferentes finalidades, como a avaliação de imóveis em massa ou a disponibilização de estimativas em sites de venda de imóveis.

Fontes Bibliográficas

- Adekunle, K. A., & Eke, P. O. (2014). A review of real estate valuation and optimal pricing techniques. *Asian Economic and Financial Review*, 1878-1893.
- Augusto, L. P. (2019). *A IMPORTÂNCIA DO REPORTE DE INFORMAÇÃO FINANCEIRA E NÃO FINANCEIRA PARA A CENTRAL DE BALANÇOS DO BANCO DE MOÇAMBIQUE: O Caso da Visabeira Global, SGPS, S.A.*
- Azevêdo, L. (2018). *Métodos estatísticos em Aprendizado de Máquinas para problemas de Classificação.*
- Büyükkaraciğan, D. N. (2021). Modern methods approach in real estate valuation. *Ankara: Iksad Publishing House*, 130.
- Corcovia, L. O., & Alves, R. S. (2019). Aprendizagem de Máquina e Mineração de Dados: avaliação de métodos de aprendizagem. *Revista Interface Tecnológica*, 90-101.
- Da Silva, E., & Menezes, E. (2005). *Metodologia da pesquisa e elaboração de dissertação.* Florianópolis.
- Da Silva, M. B., & Vanderlinde, M. (2010). *Inteligência artificial, aprendizado de máquina.*
- Gomes, D. d. (2010). Inteligência Artificial: conceitos e aplicações. *Revista Olhar Científico*, 234-246.
- González, M. (2022). *Aplicação de técnicas de descobrimento de conhecimento em bases de dados e de inteligência artificial em avaliação de imóveis.*
- Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data preprocessing for supervised leaning. *International journal of computer science*, 111-117.
- Manzini, E. (2012). *Uso da entrevista em dissertações e teses produzidas em um programa de pós-graduação em educação.*
- Metz, J. (2006). *Interpretação de clusters gerados por algoritmos de clustering hierárquico.* Universidade de São Paulo.
- Naiene, E. (2015). *Avaliação de imóveis com recurso a modelos financeiros e não-financeiros: estudo de um caso.*

- Nascimento, L. B., Gonçalves, A. C., Balbino, M. S., & Nobre, C. N. (2024). Codificação dos Atributos e sua Relação com a Interpretabilidade dos Modelos de Aprendizado de Máquina- Uma Análise em Base de Dados da Saúde Mental. In Simpósio Brasileiro de Computação Aplicada à Saúde. *Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)* (pp. 483-494). SBC.
- Pontes, T., & Arthaud, D. (2018). Metodologias ágeis para o desenvolvimento de softwares. *Ciência e Sustentabilidade*, 173-213.
- Popat, S. K., & Emmanuel, M. (2014). Review and comparative study of clustering techniques. *International journal of computer science and information technologies*, 805-812.
- Prime Yield. (2023). *Flash Mercado Imobiliário*.
- Rai, P., & Singh, S. (2010). A survey of clustering techniques. *International Journal of Computer Applications*, 1-5.
- Sanchez, W. (2023). *Aplicações de inteligência artificial*. São Paulo: Editora Senac.
- Santiago, C. P., Veras, N. d., Aragão, A. P., Carvalho, D. A., & Amaral, L. A. (2020). *Desenvolvimento de sistemas Web orientado a reuso com Python, Django e Bootstrap*. Sociedade Brasileira de Computação.
- Sousa, I. (2022). *Os benefícios e desafios da economia compartilhada para o mercado imobiliário*. Rio de Janeiro.
- Tavares, F. O., Moreira, A. C., & Pereira, E. T. (2012). *Assimetria de informação no mercado imobiliário: uma revisão da literatura*.
- Zacarias, J. M. (2022). *Modelos automatizados de avaliação de imóveis: Aplicação à Área Metropolitana de Lisboa*.
- Żróbek, S., Kucharska-Stasiak, E., Trojanek, M., Adamiczka, J., Budzyński, T., Cellmer, R., . . . Sajnog, N. (2014). *Current problems of valuation and real estate management by value*. Croatian Information Technology Society. Croatian Information Technology Society.

Apêndices

Apêndice 1: Interfaces do Utilizador

Avaliação Automática de Imóveis

Informações Gerais do Imóvel

Finalidade: Venda | Design: Flat | Bairro: Alto-Maé | Tipologia: T1

Comodidades e Outras Características do Imóvel

Móveis: Possui | Guarda-fatos: Possui | Ar-condicionado: Possui | Garagem: Possui

Cozinha Americana: Possui | Nº WC's: 0 | Suite: Possui | Vista: Possui

Piscina: Possui | Obras: Possui

Avaliar ➤

Figura 15 - Página principal da aplicação

Avaliação Automática de Imóveis

Valor Estimado:
10,425,506.86 MZN

Informações Gerais do Imóvel

Finalidade: Venda | Design: Flat | Bairro: Alto-Maé | Tipologia: T2

Comodidades e Outras Características do Imóvel

Móveis: Possui | Guarda-fatos: Possui | Ar-condicionado: Possui | Garagem: Possui

Cozinha Americana: Possui | Nº WC's: 3 | Suite: Possui | Vista: Possui

Piscina: Possui | Obras: Possui

Voltar

Exportar

Figura 16 - Interface para visualização do resultado

Apêndice 2: Roteiro de Entrevista

Roteiro da Entrevista

- **Introdução:**
 - **Apresentação geral do projecto para o entrevistado.**
 - "Gostaria de discutir o processo de avaliação de imóveis em Maputo. Pode compartilhar sua experiência e perspectiva sobre esse tema?"

- **Processo de Avaliação de Imóveis:**
 - "De acordo com o autor Gonzalez (2002), o método comparativo de dados de mercado é o mais utilizado para avaliação individual de imóveis, e consiste em uma inspeção detalhada do imóvel, busca e seleção de propriedades similares, ajustamento dos preços e documentação da avaliação através de um formulário ou relatório, sendo que, em alguns países, como os Estados Unidos, é comum a avaliação baseada apenas em três amostras similares."
 - Você utiliza esse método ou outros no seu processo de avaliação? Poderia descrever os cenários em que os aplica?"
 - "Quais são os principais fatores e características que você analisa ao avaliar um imóvel?"
 - "Você utiliza dados históricos (informações sobre casas vendidas no passado) no processo de avaliação? Se sim, quantas casas no histórico são utilizadas como base de comparação?"

- **Desafios e Necessidades:**
 - "O autor Sousa (2022) menciona desafios como a identificação de demanda referente a necessidade de cada época do ano e a bolha imobiliária caracterizada pelo aumento repentino dos preços dos imóveis devido a mudanças sociais, econômicas ou políticas. Você tem enfrentado desafios semelhantes? Existem outros desafios específicos que tem encontrado?"

- "Quais desafios você antecipa para o futuro no mercado imobiliário?"
- "Poderia descrever mais detalhadamente sua necessidade de acesso ao novo mercado digital?"

▪ **Requisitos da aplicação e desafios Potenciais:**

- "Poderia descrever mais detalhadamente sua necessidade de acesso ao novo mercado digital?"
- "Já utilizou aplicativos para avaliação de imóveis ou outros itens relacionados com o sector imobiliário? Pode descrever a sua experiência com os mesmos?"
- "Considerando a sua experiência digital, quais funcionalidades considera fundamentais para a aplicação que se pretende desenvolver?"
- "Qual informação gostaria de obter em relação ao uso da aplicação?"
- "Existe algum aplicativo, não necessariamente do sector imobiliário, que você acha fácil de utilizar e gostaria que a aplicação proposta tivesse um funcionamento semelhante? "
- "Gostaria que o aplicativo estivesse disponível em diferentes dispositivos, como smartphones, laptops e tablets?"
- "Considera necessário que a aplicação tenha algum tipo de autenticação de usuário?"
- "Quais desafios antecipa ao implementar e utilizar a aplicação proposta para avaliação de imóveis?"
- "Na sua opinião, como será a adaptação dos seus colaboradores e clientes ao uso da aplicação?"
- "Quais barreiras de mercado poderão surgir com a introdução da aplicação no mercado imobiliário"

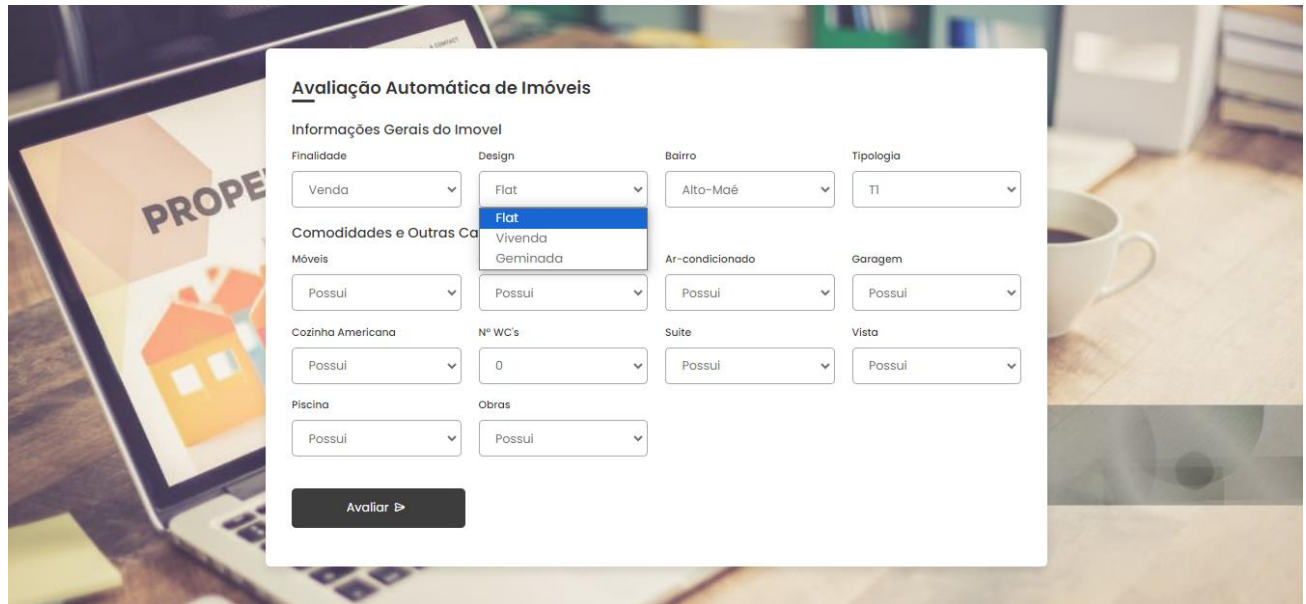
▪ **Experiências Anteriores e Melhores Práticas:**

- "Pode compartilhar exemplos de experiências anteriores ou melhores práticas que você aplicou na avaliação de imóveis?"
- "Quais lições aprendidas ou insights importantes você obteve dessas experiências?"

- **Colaboração e Parcerias:**

- "Você tem colaborado com outras entidades ou setores no processo de avaliação de imóveis? Como essas parcerias contribuem para a eficácia ou transparência do processo de avaliação?"

Apêndice 3: Casos de Uso



Avaliação Automática de Imóveis

Informações Gerais do Imóvel

Finalidade	Design	Bairro	Tipologia
Venda	Flat	Alto-Maé	T1

Comodidades e Outras Características do Imóvel

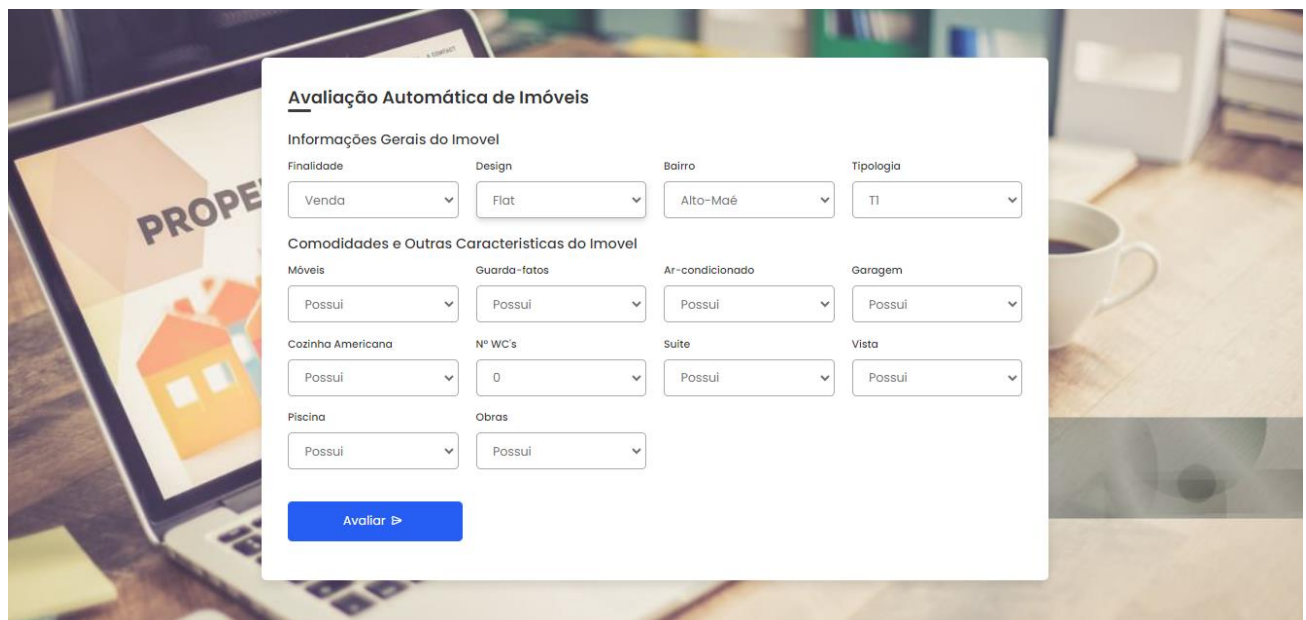
Móveis	Ar-condicionado	Garagem
Possui	Possui	Possui

Cozinha Americana	Nº WC's	Suite	Vista
Possui	0	Possui	Possui

Piscina	Obras
Possui	Possui

Avaliar ➤

Figura 17 - Introdução de dados do imóvel



Avaliação Automática de Imóveis

Informações Gerais do Imóvel

Finalidade	Design	Bairro	Tipologia
Venda	Flat	Alto-Maé	T1

Comodidades e Outras Características do Imóvel

Móveis	Guarda-fotos	Ar-condicionado	Garagem
Possui	Possui	Possui	Possui

Cozinha Americana	Nº WC's	Suite	Vista
Possui	0	Possui	Possui

Piscina	Obras
Possui	Possui

Avaliar ➤

Figura 18 - Avaliar o valor de um imóvel

Avaliação Automática de Imóveis

Valor Estimado:
3,809,726.55 MZN

Informações Gerais do Imóvel

Finalidade Venda	Design Flat	Bairro Alto-Maé	Tipologia T1
---------------------	----------------	--------------------	-----------------

Comodidades e Outras Características do Imóvel

Móveis Possui	Guarda-fatos Possui	Ar-condicionado Possui	Garagem Possui
Cozinha Americana Possui	Nº WC's 0	Suite Possui	Vista Possui
Piscina Possui	Obras Possui		

[Voltar](#)

[Exportar](#)

Figura 19 - Apresentar resultados da avaliação

Avaliação Automática de Imóveis

Valor Estimado:
3,809,726.55 MZN

Informações Gerais do Imóvel

Finalidade Venda	Design Flat	Bairro Alto-Maé	Tipologia T1
---------------------	----------------	--------------------	-----------------

Comodidades e Outras Características do Imóvel

Móveis Possui	Guarda-fatos Possui	Ar-condicionado Possui	Garagem Possui
Cozinha Americana Possui	Nº WC's 0	Suite Possui	Vista Possui
Piscina Possui	Obras Possui		

[Voltar](#)

[Exportar](#)

result (19).pdf
3.3 MB • Done

Figura 20 - Extrair o relatório