**Rental Housing Prices Prediction in Nairobi Area Using Regression**

**Agunda Priscilah Naliaka**

**123854**

**ICS 4 C**

**Supervisor Name**

**Dr. Esther G. Khakata**

**Submitted in Partial Fulfilment of the Requirements of the Bachelor of Science in Informatics and Computer Science at Strathmore University**

**School of Computing and Engineering Science**

**Strathmore University**

**Nairobi, Kenya**

**July 2022**

# Declaration and Approval

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the research proposal contains no material previously published or written by another person except where due reference is made in the research proposal itself.

Student Name: Agunda Priscilah Naliaka

Admission Number: 123854

Student Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The Proposal of **Agunda Priscilah Naliaka** has been reviewed and approved by **Dr. Esther G. Khakata**

Supervisor Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Acknowledgement

I acknowledge the Almighty God for his grace during this study. My special gratitude goes to my supervisors, Dr. Esther G.Khakata and Mr. Kevin Omondi Ochieng for their scholarly comment, continuous guidance and critic that gave direction to this study as well as continuous encouragement that contributed to the successful completion of this project.

My sincere thanks go to my colleagues and friends; for their moral support and always encouraging me to move forward and complete this study.

May the almighty God bless you.

# Abstract

In many developing countries, house prices have recently been a major source of economic and social dispute. Broadly speaking, one of the key factors that affect the values of real property is the state of the economy (Nguyen, 2021) . An economic metric called GDP (Gross Domestic Product) gauges the economy's overall health. GDP specifically calculates the total amount of goods and services generated in a nation, including both current consumption goods and capital goods produced as investments in future output. Housing is a major part of both current consumption and private investment. Research shows that Kenya's real estate sector grew by 5.2 percent in the third quarter of 2021. This represented an increase in comparison to the corresponding quarter in 2020. In that quarter, the value added by real estate activities to the country's GDP expanded by 3.7 percent. This number is expected to rise since it forms part of humans’ basic needs, shelter, and is also in line with the government’s ‘big four’ agenda, one of them being affordable housing. This project aims at coming up with a way in which the prices of rental properties in Nairobi County can be predicted using regression analysis. This method was chosen instead of the commonly used price forecasting models such as single predictor models which are prone to over-fitting and low accuracy levels emanating from their inability to handle noisy data. The dataset used is the Apartment Prices in Kenya readily available on Kaggle. It is expected that the results will give rise to a model which almost accurately predicts prices of rental properties in some areas in Nairobi, based on selected features. This research is paramount in decision making procedures of investors as they eye on setting up developments within the constituency. The county government may benefit heavily as they may be able to work on certain areas that may increase revenue in the sector.

Table of Contents

[Declaration and Approval i](#_Toc108596135)

[Acknowledgement ii](#_Toc108596136)

[Abstract iii](#_Toc108596137)

[List of Figures vi](#_Toc108596138)

[List of Abbreviations vii](#_Toc108596139)

[Chapter 1: Introduction 1](#_Toc108596140)

[1.1 Background Information 1](#_Toc108596141)

[1.2 Problem Statement 2](#_Toc108596142)

[1.3 Objectives 3](#_Toc108596143)

[1.3.1 General Objective 3](#_Toc108596144)

[1.3.2 Specific Objectives 3](#_Toc108596145)

[1.3 Research Questions 3](#_Toc108596146)

[1.5 Justification 3](#_Toc108596147)

[1.6 Scope and Limitations 4](#_Toc108596148)

[1.6.1 Scope 4](#_Toc108596149)

[1.6.2 Limitations 4](#_Toc108596150)

[Chapter 2: Literature Review 5](#_Toc108596151)

[2.1 Introduction 5](#_Toc108596152)

[2.2 Theoretical Review 5](#_Toc108596153)

[2.2.1 Machine Learning Algorithms 5](#_Toc108596154)

[2.3 Empirical Studies 6](#_Toc108596155)

[2.4 Conceptual Framework 7](#_Toc108596156)

[Chapter 3: Methodology 8](#_Toc108596157)

[3.1 Introduction 8](#_Toc108596158)

[3.2 Applied Methodology 8](#_Toc108596159)

[9](#_Toc108596160)

[3.2.1 Empathise 9](#_Toc108596161)

[3.2.2 Define 9](#_Toc108596162)

[3.2.4 Prototype 9](#_Toc108596163)

[3.2.5 Test 10](#_Toc108596164)

[3.3 Analysis Diagrams 10](#_Toc108596165)

[3.3.1 Use-Case Diagram 10](#_Toc108596166)

[3.3.2 Sequence Diagram 10](#_Toc108596167)

[3.3.3 System Sequence Diagram 10](#_Toc108596168)

[3.3.4 Entity Relationship Diagram 10](#_Toc108596169)

[3.3.5 Context Diagram. 10](#_Toc108596170)

[3.3.6 Data Flow Diagram 11](#_Toc108596171)

[3.4 Design Diagrams 11](#_Toc108596172)

[3.4.1 System Architecture 11](#_Toc108596173)

[3.5 Deliverables 11](#_Toc108596174)

[3.5.1 Concept Defense 11](#_Toc108596175)

[3.5.2 Proposal Document 11](#_Toc108596176)

[References 13](#_Toc108596177)

# List of Figures

[Figure 2.1 Conceptual Framework 8](#_Toc108596924)

[Figure 3.1 Phases of Design Thinking 9](#_Toc108596925)

# List of Abbreviations

GDP – Gross Domestic Product

SDG – Sustainable Development Goals

UN – United Nations

# Chapter 1: Introduction

## 1.1 Background Information

One of the measures of adequate standard of living is the availability of proper housing which is prominent as a basic human right. The ultimate choice towards owning a home is one of the most important decisions made by a potential owner in respect to the cost involved. A home not only provides a shelter to the family but also provides an opportunity for investors to participate in.

From the onset of the 2007/2008 financial crunch that turned out to be a global crisis and which has often been referred to as the great recession, it is now widely acknowledged by empiricists and practitioners that the role played by house prices in generation of business cycles and financial dynamics is greatly significant (Valadez, 2011) . United Nations (2019) says that in 2015, all United Nations members endorsed the 17 Sustainable Development Goals (SDGs), which were a call to all countries to reduce poverty, safeguard the environment, and ensure peace and prosperity for all by the year 2030.

SDG 11 underlines the significance of creating cities and human settlements that are more inclusive, safe, resilient, and sustainable regarding housing as a form of shelter. This is done by ensuring that everyone has access to decent housing that is both safe and inexpensive while yet offering the basics. Notably, more than 50% of the world's population lives in cities, and that percentage is predicted to reach 60% by 2030, contributing to 60% of the global economy.

In essence, this establishes a global focus on affordability and financing as essential components of providing housing for the urban population. (Bah, 2018) observes that the rate of urbanisation in the African continent is much higher than anywhere else in the world. Further, there is a major deficit of access to proper housing for urban poor and middle income earners.

According to Julia Faria in a report made in January 2022, Kenya's real estate sector grew by 5.2 percent in the third quarter of 2021. This represented an increase in comparison to the corresponding quarter in 2020. In that quarter, the value added by real estate activities to the country's Gross Domestic Product (GDP) expanded by 3.7 percent.

This number is expected to rise since it forms part of humans’ basic needs, shelter, and is also in line with the government’s ‘big four’ agenda, one of them being affordable housing.

## 1.2 Problem Statement

Africa Housing Finance Yearbook, 2019 indicates that the rate of urbanisation in Kenya stood at 4.3 per cent as of 2017 which is higher compared to Sub-Saharan Africa average that stands at 4.1 per cent. Access to decent housing in urban areas is one of the current social amenities that is being placed under excessive demand by the growing urban population. This has highlighted an increase in housing demand, leading to an increase in land and home prices.

A house's price is typically negotiated, and high transaction costs are a feature of the market. To establish a foundation for estimating the price, it can be necessary to do an online search for information on comparable homes or receive the same from real estate agencies. The procedure is time-consuming and demanding, and time may not always be available. If the price estimating process is delayed, another customer may buy the indicated house, resulting in the loss of the option to buy it. Untrustworthy real estate agents take advantage of would-be homeowners by bilking them out of their hard-earned cash by asking outrageous charges while setting the final price. This indicates that home prices are constrained to what sellers are willing to offer as there are no comparable pricing.

There are multiple studies on house price prediction which use either single predictor models or multiple predictor models. Single predictor models involve the use of a single algorithm in the prediction process while multiple predictor models rely on multiple algorithms such as the use of multiple trees in a random forest algorithm. The use of multiple algorithms is the basis upon which ensemble learning is defined.

There needs to be a faster and more accurate way to predict housing prices without having to depend on the current methods of price estimation. This project proposes the use of stacking models, that is the use of Linear Regression and Regression Enhanced Random Forest to come up with a model for predicting rental housing property. Furthermore, the proposed project will attempt to identify the common significant characteristics that are present in the majority of urban setups. These frequently occurring factors will result in a model that can be used to forecast property values in different urban regions.

## **1.3 Objectives**

### 1.3.1 General Objective

The main objective of the study is to develop a model that predicts the prices of rental housing properties in Nairobi area

### 1.3.2 Specific Objectives

1. To identify significant characteristics that influence the price of urban residential houses
2. To study the challenges associated with current methods of rental value estimation
3. To develop a model for predicting urban residential housing prices
4. To test and evaluate the developed model

## Research Questions

1. Which characteristics influence urban residential housing prices?
2. What are the challenges associated with current methods of rental value estimation?
3. Which is the best ensemble learning model for predicting urban residential housing prices?
4. What is the performance of the developed model?

## 1.5 Justification

The project is motivated by the need to develop an efficient model that can be utilized by the housing sector market players in decision making.

Further, the model will be used to provide a price estimate that will be used by potential homeowners hence arresting their continued exploitation by unscrupulous property brokers. This will partly help address the affordability issue in respect to housing prices by providing a price estimate for reference purposes

The identification of significant characteristics will help potential homeowners and investors determine the best combination of characteristics that would fetch better prices for a future sale. This will provide beforehand characteristics that are preferred upon by potential customers thus influencing the type of house to be constructed.

.

## Scope and Limitations

### 1.6.1 Scope

The project seeks to focus on urban residential housing sector where there is a huge demand for decent and affordable housing. Urban areas have a high population growth rate and population density as well as high costs of property. The study will be limited to the urban housing sector with a focus on Nairobi County in Kenya.

### Limitations

There is a limited number of up-to-date datasets on rental property in Nairobi. Most datasets cover other regions, mostly outside Kenya

The project mainly relied on secondary data obtained from Kaggle as such the researcher placed high reliability strictly on this data.

Though useful, the project results may not be used to make generalisations about other property financiers in Kenya; thus, the variables identified are tentative suggestions of the variables that determine house prices across real estate industry in Kenya.

# Chapter 2: Literature Review

## 2.1 Introduction

This chapter covers literature on residential real estate pricing and other relevant topics. It lays a foundation for the study area by offering a review of studies undertaken in relation to housing prices prediction while focusing on the gaps that exist in these studies. This assessment will assist in creating a framework for identifying key factors responsible for the constantly varying home prices. The factors that have been highlighted will be crucial in creating a conceptual framework.

## 2.2 Theoretical Review

As interest has grown over the recent years, academicians have become more interested in real estate development. This has made it necessary to value property according to its nature and the conditions under which the specific property would trade on the open market. According to (Jier-Haur Chen, 2017), studies undertaken in the early stages of forecasting housing prices tended to follow the law of demand and supply with some adjustment latency.

### 2.2.1 Machine Learning Algorithms

In 1959, Arthur Samuel a computer scientist defined machine learning as the ability of computers to learn without being explicitly programmed. Tom M. Mitchell (1997) defined machine learning as the following: “A computer program being able to learn from experience E with respect to some tasks T and a performance measure P, if its performance at tasks in T, as measured by P, improves with experience E”. This in essence means providing computer systems with an ability to learn automatically hence continue improving through experience without any human intervention.

Machine learning has several categories of algorithms which include supervised, unsupervised and reinforcement learning. Supervised learning involves application of what has been learnt in the past to new data. This in essence leverages on labelled data to help predict occurrence of a future event. Supervised learning algorithms usually model existing relationships between the input variables and the target variable to help predict target output for unseen data. Such algorithms are further divided into two categories namely classification and regression-based algorithms. Regression algorithms generate predictive models using labelled output label that is continuous in nature.

Ensemble learning is a supervised learning method which takes a weighted average or a majority vote of each base model estimator that has been built on its own using a supervised learning method (Dipanjan Sarkar, 2021). Through this, the resultant ensemble can generalize much better as compared to an individual model, hence providing a better prediction. There are three major categories of ensemble models which include:

a) Bagging: Bagging stands for bootstrap aggregating where predictions of base models are combined based on training done using training samples that are randomly generated. The use of randomly generated samples helps reduce over-fitting and model variance while improving on accuracy. Random forests fall under this category.

b) Boosting: In this method, the ensemble model is constructed incrementally by training each base model sequentially by learning using instances that were previously misclassified. The training of these weak learners is done over multiple iterations with weight modifications during each retraining phase with higher weights being assigned to misclassified instances. These multiple weak base learners are then combined to form a more powerful ensemble model. XGBoost, CatBoost and LightBoost are the commonly used boosting algorithms.

c) Stacking: This involves constructing base model(s) on the training data before building a final ensemble based on the output derived from the predictions of the base models. The output of the base model(s) is used as the input for the final model development.

Unsupervised learning involves the use of unlabelled data to develop a model that defines clear boundaries in a given dataset. There are no defined output labels these set of algorithms use techniques on the input data to detect any existing patterns. Similar patterns are summarized in groups thus deriving meaningful insight that describes the data. there are two major categories of unsupervised learning which include clustering such as Kmeans and association-based algorithms such Apriori.

Clustering algorithms usually group input data points into different classes with similar characteristics. Association algorithms extract rules and patterns mined from data which explain relationships existing between different variables while acknowledging common item sets and patterns in occurrence. Reinforcement learning refers to learning through interactions with the environment through actions leading to either penalties or rewards.

## 2.3 Empirical Studies

(Kimani, 2021)endeavours to analyse the spatial factors that affect rental houses prices in Nyeri town constituency, Kenya and determine their relationship to their current respective prices. They employed the use of spatial hedonic model. Borrowing from this model, the major characteristics as highlighted in the various studies that inform the price of a house are categorized into three major groups which include housing, locational and neighbourhood attributes. Its gaps are that they focus on one factor affecting housing prices rather than diverse factors that have been noted to influence rental house prices, some having higher effect than others.

(Yang & Cao, 2018) highlight that the use of ensemble learning in house price prediction shows improved results. They used the output of the base models and part of original results to make a final prediction using one predictor model. However, the base algorithms did not further clean the dataset for possible information that may have been captured by the other models; location characteristics are not included in the study.

(Oxenstierna, 2017) used ensemble learning technique with clustering whereby the base models’ output was passed directly into the final predictor. This shows that the initial algorithms did not learn from other corresponding model predictions. The original dataset was only screened for important information at the initial model building phase, including characteristics such as House size, year built, Number of rooms, etc. Therefore, loss of critical information may have occurred affecting model performance. Despite Cluster aggregation showing huge potential for further improvement as opposed to models that do not use any form of clustering, neighbourhood characteristics such as distance to roads was not factored.

Even though most of the traditional methods of rent forecasting are widely accepted and used around the globe they all come with their shortcomings. Most of the methods output higher rates of errors for example HPM will output a MAPE value of 38.23 percent. The Hedonic Model, a traditional method for forecasting house prices, has been criticised due to nonlinearity, multicollinearity and heteroskedasticity problems, which were argued to affect estimation accuracy

## 2.4 Conceptual Framework

A conceptual framework illustrates what one expects to find through their research. It defines the relevant [variables](https://www.scribbr.com/methodology/types-of-variables/) for the study and maps out how they might relate to each other.

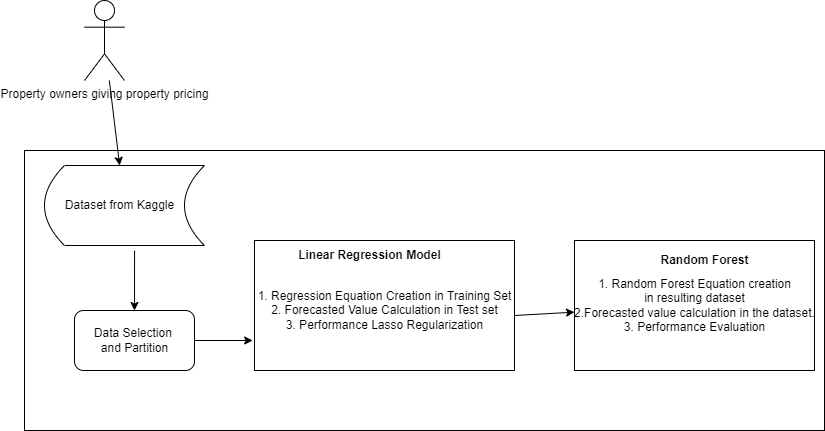


Figure 2.. Conceptual Framework

In this case, data on prices of rental properties as well as the characteristics of the said properties is collected and recorded in a dataset which is available on Kaggle. Using this data, the researcher is able to select the relevant features relevant to the study, train the data by taking it through a series of models, test it and come up with a final model which will predict prices of rental properties based on the selected features.

# Chapter 3: Methodology

## 3.1 Introduction

This chapter describes the various milestones to be achieved in attainment of the stated objectives. This entails providing a detailed description of the methodology to be used as well as data collection methods and analysis performed in undertaking the research. The chapter provides a brief overview of the study area. A discussion of the research design adopted for the research is provided which includes a highlight of data sources and type.

## 3.2 Applied Methodology

The desired methodology for the project is design thinking methodology. Design thinking is an innovative problem-solving process rooted in a set of skills. At a high level, the steps involved in the design thinking process are simple: first, fully understand the problem; second, explore a wide range of possible solutions; third, iterate extensively through prototyping and testing; and finally, implement through the customary deployment mechanisms.

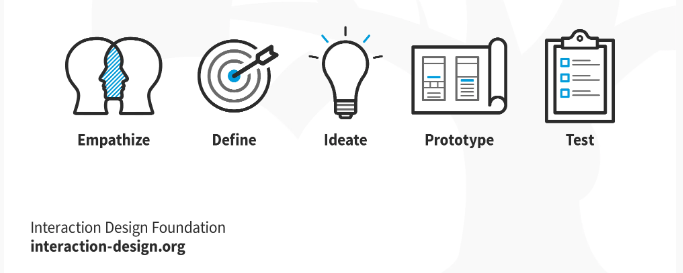


Figure 3.. Phases of Design Thinking

### 3.2.1 Empathise

During this stage the researcher is required to give the customer an understanding of the final product. Requirements gathering processes such as observing, engaging through conversation, and interviewing are used – but with a deeper goal of empathising with the user, to understand the problem and the related issue more thoroughly. In this case, data collection is done by use of a dataset from Kaggle which details Apartment Prices in Kenya. The dataset shows the various apartment listings in various regions in Kenya, Nairobi being one of them; as well as showing certain characteristics of the apartments such as location, number of bedrooms, among others. The goal of this phase is to gather requirements by better understanding the experiences of the users

### 3.2.2 Define

In this phase the information gathered in the empathise stage is used to document the requirements in a clear manner. The define phase should conclude with a statement about the requirements that clearly sets out the scope and parameters of the problems.

**3.2.3 Ideate**

In the ideate phase the researcher brainstorms and generates different possibilities to what could be solutions to the problem at hand. Prototypes can be used as an innovation technique during this phase. This can be achieved through data cleaning to come up with a relevant dataset which can then be put through the testing phase.

### 3.2.4 Prototype

A prototype is an example that serves as a basis for future models. Prototyping gives designers an opportunity to research new alternatives and test the existing design to confirm a product’s functionality prior to production. Prototypes allow teams to recognize flaws in their design thinking progress while having the freedom to iterate their product. In this case, the test set is taken as the prototype.

### 3.2.5 Test

In this phase the researcher gets feedback from the prototypes, this phase allows the iteration to empathy phase. Feedback from the testing phase will help refine prototypes, and ultimately indicate whether the defined problems are addressed appropriately.

## 3.3 Analysis Diagrams

## 3.3.1 Use-Case Diagram

The use case diagram is used to show possible interactions a user can have with the different elements in a system. This model will aid in identifying actors and the functions they will perform on the system in a specific use-case.

### 3.3.2 Sequence Diagram

A sequence diagram is an interaction diagram that shows how objects interact in a sequential order. The sequence diagram will be modelled to display different system-wide methods for a particular class.The interactions between various system objects will be easier to understand with the aid of this graphic.

### 3.3.3 System Sequence Diagram

A system sequence diagram is a dynamic representation of a use case that demonstrates how processes interact over the course of a given period of time. A system sequence diagram illustrates the use-case by displaying the processes, messages, and message timing. This diagram will be created to show the relationships between the system's actors and the system itself.

### 3.3.4 Entity Relationship Diagram

An ERD is a model that depicts the logical relationships and interactions that exist between system entities. This will be drawn to provide a general overview of the system, its entities, their attributes, and the relationships between the system entities and the system itself.

### 3.3.5 Context Diagram.

A context diagram is a high-level view of an information system that shows the boundaries and scope of the system. Process 0 in the context diagram represents the entire information system but does not show its internal workings. This will be drawn to demonstrate the interaction of the system and its entities, as well as how data flows between them.

### 3.3.6 Data Flow Diagram

#### 3.3.6.1 DFD level 0

A level 0 DFD enlarges the system and displays key internal processes, data flows, and data stores. This will be drawn to show how data will flow between processes and how data will be stored.

#### 3.3.6.2 DFD level 1

A level 1 DFD enlarges the system even further to show the most basic processes, data flows, and data stores. This will be drawn to show how data will flow in the most basic processes, such as splitting a dataset in half.

## 3.4 Design Diagrams

### 3.4.1 System Architecture

System architecture translates the logical design of an information system into a physical

structure that includes hardware, software, network support, processing methods, and

security. This will give a brief overview of the proposed system, therefore will be used to

envision all components and their interaction with the system

## 3.5 Deliverables

### 3.5.1 Concept Defense

The concept note is a summary of the general idea of the product to be developed. The

Slides will be prepared, and the idea presented before a panel for approval.

### 3.5.2 Proposal Document

A System Proposal is a document which is presented to get project approved. This document will be needed because it defines the objectives of the solution as well as the steps to be

followed during development.

### 3.5.3 Analysis and Design Diagrams Document

This is the document that contains all diagrams of the system design and architecture of the

approach chosen. The diagrams will be drawn to visualise the whole system and its

processes.

## 3.6 Tools and techniques

In this section provides a list and explanation of the tools and techniques that will be used during this project and why they are important.

### 3.6.1 Python

The proposed solution will be built in Python, a high-level programming language best suited for machine learning projects. Python and itself is backed up with a huge amount of libraries with the ability to work with powerful frameworks such as tensorflow.

### 3.6.2 GitHub

This online collaboration tool will be used to store the code and act as a backup in case of any system failures in the future. GitHub git is a version control system this means that one can track different changes made to their source code. GitHub is designed to manage projects efficiently.

### 3.6.3 Kaggle

Kaggle is a community hosted online for data scientists and machine learning engineers. Kaggle provides dataset to be used in model training and testing.

# Chapter 4: System Analysis and Design

## 4.1 Introduction

This chapter covered the analysis and design of the Rental Prices Prediction model. The specific details of the model as well as how their components to relate to each other are discussed. The chapter has the system diagrams and models that explain and give a view of how the system works. It also has designs of the system and their specified strategy.

## 4.2 System Requirements

This phase involved determining the requirements met based on the objectives identified in this study. Functional and non-functional requirements of the rental property price prediction model were identified. The needs of different stakeholders using the model were an important input in determining the requirements.

### 4.2.1 Functional Requirements

These are specific functions the rental property prediction system performs. Functional requirements identified include:

1. The System allows users to input a rental property type and the area it is located.
2. The system predicts the value of the selected property based on a regression function for that area and property type.
3. The system is hosted as a web service hence it allows users to log in into the system through different web browsers.
4. The system is able to use multiple predictor variables in determining the appropriate function.

### 4.2.3 Non-Functional Requirements

Non-functional requirements are used to define the additional functions that strengthen the prediction system’s reliability. The non-functional requirements of this model includes:

1. The system is user friendly and easily navigable
2. The system allows for multiple users to do prediction without downgrading on the performance
3. The system is accessible from multiple devices and can scale to fit the screen size
4. The application is secure to protect the integrity of the data uploaded by users. Other users should not see individual users’ personal information and data.
5. The system is available when needed by users and shall not have downtimes
6. The application is responsive and efficient.

## 4.3 System Analysis Diagrams

### 4.31 Use Case Diagram

The use case in Figure 4.1 shows how actors of the system will interact with the model. The actor is any user of the system who interacts with the model to get the price prediction by the system.

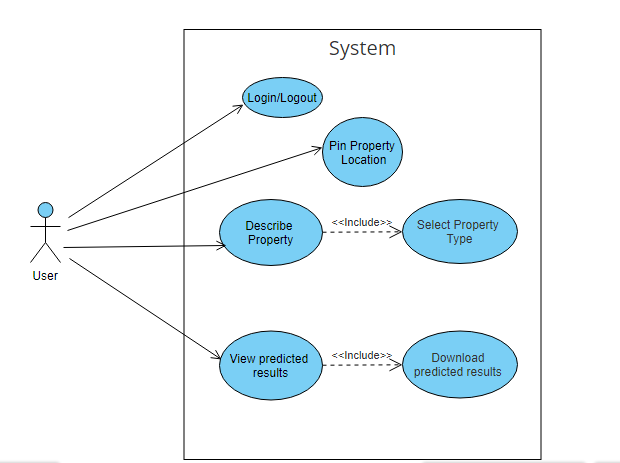


Figure 4. 1 Use Case Diagram

### 4.3.2 Sequence Diagram

A sequence diagram illustrates the sequence of messages between objects in an interaction. It describes how and in what order a group of objects works together. This is shown in Figure 4.2:

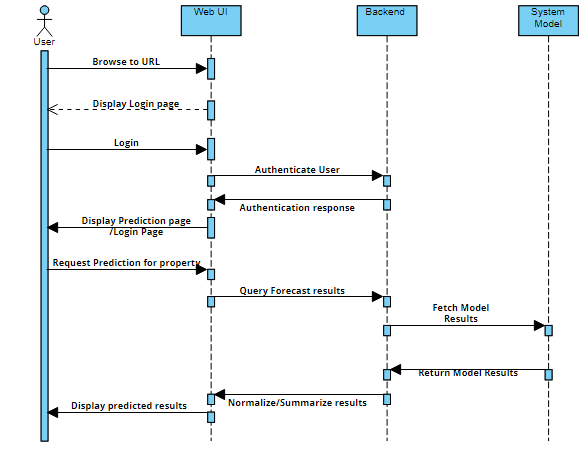


Figure 4. 2 Sequence Diagram

### 4.3.3 System Sequence Diagram

System sequence diagram shows the interaction of the user with the system showing the input and output events. The Figure 4.3 shows the system sequence diagram:

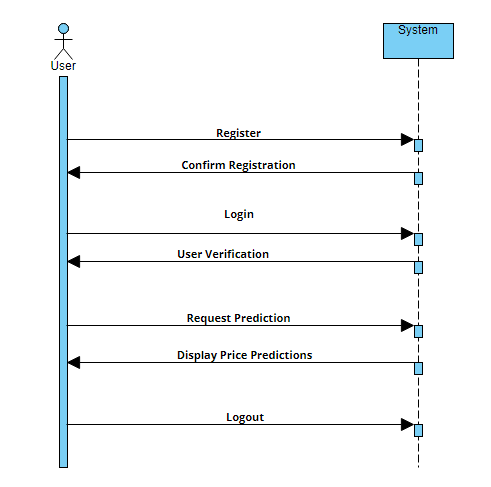


Figure 4. System Sequence Diagram

### 4.3.4 Entity Relationship Diagram

An Entity Relationship Diagram shows the relationship between each entity. It was used to show a model of the final system and attributes. The Figure 4.4 shows the ERD:

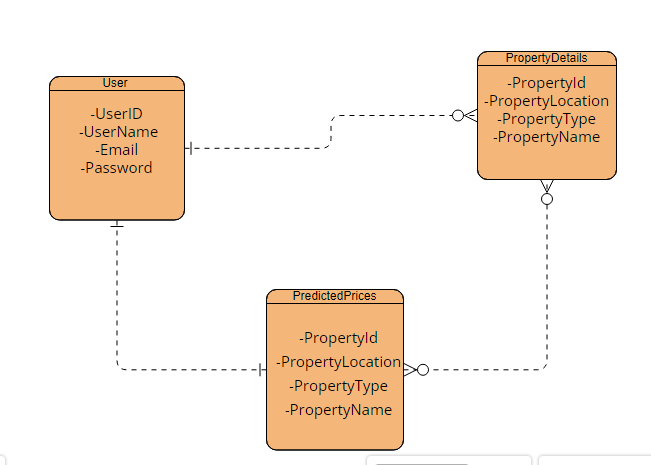


Figure 4. Entity Relationship Diagram

### 4.3.5 Context Diagram, Level 0

A context diagram was used to display the system as a whole. It shows all the external entities and how they interact with the system. The application was put in the middle and the external entities that surround the system without going deep into the system.The Figure 4.5 shows the context diagram:

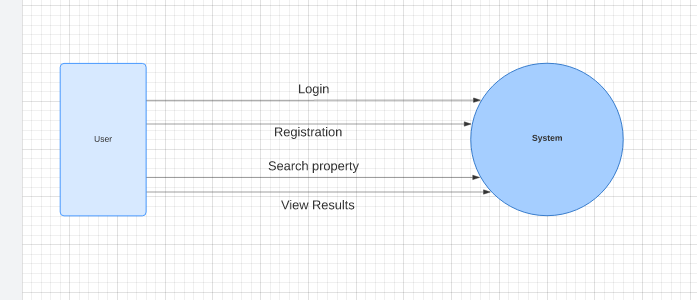


Figure 4. Context Diagram, Level 0

### 4.3.6 Level 1 Diagram

The Level 1 one diagram shows how the entities of the application interact with each other and how they interact with the application.

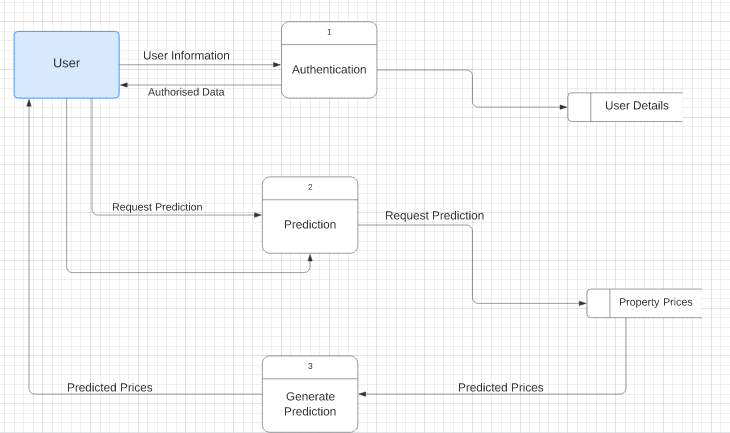


Figure 4. Context Diagram, Level 1

## 4.4 Design Diagrams

### 4.4.1 System Architecture

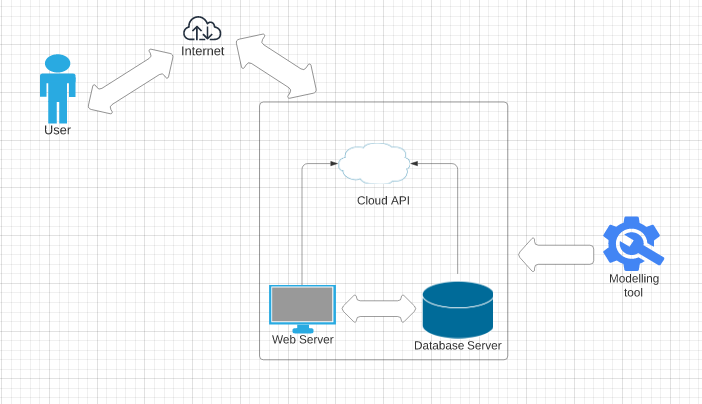


Figure 4. System Architecture

# References

Bah, E. M. (2018). *Housing Market Dynamics in Africa. .* London: PalGrave MacMilllan.

Dipanjan Sarkar, R. B. (2021). *Practical Machine Learning with Python: A Problem-Solver's Guide to Building Real-World Intelligent Systems.* Apress.

Jier-Haur Chen, S.-C. H. (2017). Forecasting spatial dynamics of the housing market using Support Vector Machine. *International Journal of Strategic Property Management*.

Kimani, E. N. (2021). Analysis of Spatial Factors Affecting Rental House Prices: A Case Study of Nyeri Town Constituency, Kenya. *Journal of Geosciences and Geomatics*.

Nguyen, J. (2021). 4 Key Factors that drive the Real Estate Market. *Real Estate Investing*.

Oxenstierna, J. &. (2017). *Predicting house prices using Ensemble Learning with Cluster Aggregations.*

Valadez, R. M. (2011). The housing bubble and the GDP: a correlation. *Journal of Case Research in Business and Economics*.

Yang, B., & Cao, B. (2018). Research on Ensemble Learning-based Housing Price Prediction Model. *Big Geospatial Data and Data Science*.