International Journal of Innovation Scientific Research and Review
Vol. 06, Issue, 10, pp.7217-7234, October 2024
Available online at http://www.journalijisr.com
SJIF Impact Factor 2023: 6.599

ISSN: 2582-6131

Research Article

RESIDENTIAL RENT PRICE PREDICTION BASED ON MACHINE LEARNING

[1,]* Oyeleke Samuel Oluwafemi, [2]Wang Jing, [3]Olumayowa O. Adedara, [3]Isatou K. Njie, [3]Obafemi Oluwaseun Adewale

[1]MSc Software Engineering (College of Software), Nankai University, Tianjin, China.
[2]Professor, College of Software, Nankai University (Supervisor), Tianjin, China.
[3]MSc., College of Software, Nakai University, Tianjin, China.

Received 01st August 2024; Accepted 02nd September 2024; Published online 30th October 2024

ABSTRACT

The dynamic and competitive real estate sector of Lagos, Nigeria, has become a growing area concern, particularly, the issue regarding fluctuations in the cost of rent. This research conducts comprehensive data analysis about factors that are influencing residential rent prices in Lagos, with the aim to develop a predictive model. The model seeks, not only to enrich comprehension but also foster making informed decisions by stakeholders, renters, landlords, policymakers, and investors. This research utilized varied datasets and advanced machine learning techniques to analyze the complex interplay of socioeconomic, urban development, and market-driven variables influencing prices of house rent in Lagos. Having the understanding of these relationships is of utmost importance for accurate predictive modelling in such a dynamic environment. Through examination of feature exhaustively and rigorous analysis of model performance, five machine learning technologies were tested in order to identify the most accurate predictive model. The study also investigates the implications of rent price prediction models for various stakeholders including landlords, policy makers, investors, and renters, offering insights for informed decision-making. Summarily, this thesis contributes to the field of house rent prediction in Lagos by leveraging advanced methodologies and comprehensive datasets to enhance understanding of the housing market dynamics and empower stakeholders with actionable insights.

Keywords: Machine learning, Statistical metrics, Flask server, predictive modeling, XG Boosting, Preprocessing.

INTRODUCTION

Lagos is the most populous city in Africa and the economic hub of Nigeria. It is also among the fastest-growing cities in the world, with a population expected to reach 30 million by 2030. This rapid growth is putting a strain on the city's housing supply. House rent costs in Lagos have been steadily increasing, particularly in desirable areas of the city, making housing affordability a major concern. Predicting these costs is not so easy as a result of many factors that include size of the property, location, property type, economic growth, unemployment, and inflation. Understanding these factors is crucial for informed decision-making, business planning, and effective government housing policies [1]. Utilizing machine learning algorithms and leveraging previous data, the primary objective of this research is to build a model that will accurately predict house rent prices in Lagos. The model will be developed using Python programming language and will utilize Flask Server Technology for seamless integration with web platform. Using machine learning algorithms to predict residential rent prices is a relatively new and innovative approach. It allows for the analysis of various factors that influence the rental market, such as property features, location, and market trends. Additionally, previous market research and published literature have shown a limited application of machine learning approaches, such as Extreme Gradient Boosting (XGBoosting), in analyzing housing prices and rents in the context of land use and transportation modeling. Therefore, this thesis set to contribute to existing literature, making use of machine learning methods and volunteered geographic information for house rent prediction in Lagos.

***Corresponding Author: Oyeleke Samuel Oluwafemi,**
1MSc Software Engineering (College of Software), Nankai University, Tianjin, China.

Overall, this thesis seeks to contribute to the field of house rent prediction in Lagos, Nigeria, by utilizing machine learning algorithms, leveraging previous data and volunteered geographic information, and providing accurate predictions that can benefit both renters and real estate companies [2].

Background of the Study

The Lagos housing market is significant for a number of reasons. For instance, it contributes substantially to the city's economy. One of the largest contributors to the Lagos GDP is real estate sector, and it employs millions of people. Again, the housing market is essential for providing shelter to the city's growing population. Furthermore, the housing market is a major investment vehicle for both individuals and companies. The Lagos housing market is also significant because it has potential to impact other sectors of the economy. For ex-ample, a well-functioning housing market can help to attract and retain skilled workers, which can boost productivity and economic growth. Additionally, a healthy housing market can help to create a more stable and prosperous society. The cost of house rent in Lagos is influenced by many factors which does not exclude location of the property, the size of the property, the condition of the property, the amenities offered, rapid growth of Lagos, economic growth, investment opportunities in Lagos, and the overall demand for housing.

Statement of Problems

Lagos, Nigeria's economic powerhouse, plays a pivotal role in driving the nation's economic growth and development. As the country's commercial nerve center, Lagos attracts a significant influx of settlers seeking employment opportunities, business prospects, and a higher standard of living. However, the city's burgeoning population and rapid urbanization have exacerbated challenges in the housing sector, leading to an acute shortage of affordable housing options. Accurately predicting residential rent prices in Lagos is essential for sustaining the city's economic vitality and ensuring equitable access to housing. By providing policymakers, real estate developers, investors, and tenants with reliable insights into rent price trends, predictive models can facilitate in-formed decision-making, foster sustainable urban development, and mitigate housing disparities. Despite the wealth of housing data available, existing predictive models often fail to capture the intricate socio-economic dynamics and geographical nuances that characterize the Lagos housing market. As such, there is a pressing need to develop robust machine learning techniques tailored specifically to the unique challenges and opportunities presented by Lagos. Therefore, this project is going to deploy to via web application for easy accessibility, residential predictive model that is based on machine learning to accurately predict rental prices for the State. The prediction model seeks to tackle the enumerated challenges below.

Price Variability

Rental prices in Lagos exhibit significant variability, making it challenging for tenants to assess whether a listed property is reasonably priced. A predictive model can provide transparency and help individuals make informed decisions.

Inconsistent Pricing Factors

The factors influencing house rent in Lagos are diverse and may include location, property size, amenities, neighborhood safety, and other socio-economic variables. Identifying and quantifying these factors will contribute to a more accurate prediction model.

Data Scarcity and Quality

Availability of reliable and comprehensive data on rental properties in Lagos can be a hurdle. The model development process addresses data scarcity is-sues and ensure that the available dataset is representative and reflective of the diverse housing market in the city.

Market Dynamics

The real estate market in Lagos is affected by many external factors like trends of the economic, infrastructure development, and changes in policy. The developed model is robust enough to adapt to changing market dynamics and provide up-to-date predictions.

Easily Accessible User-Friendly Application

The end goal is to lunch user-friendly web application that will allow tenants and landlords to features of the type of property they want to rent and get ac-curate prediction for their quest. When accessibility and simplicity is ensured in the application, it will certainly enhance its usability.

A Successful implementation of such a data driven web application could potentially serve as a template for addressing similar challenges in other dynamic real estate markets.

Architectural Framework

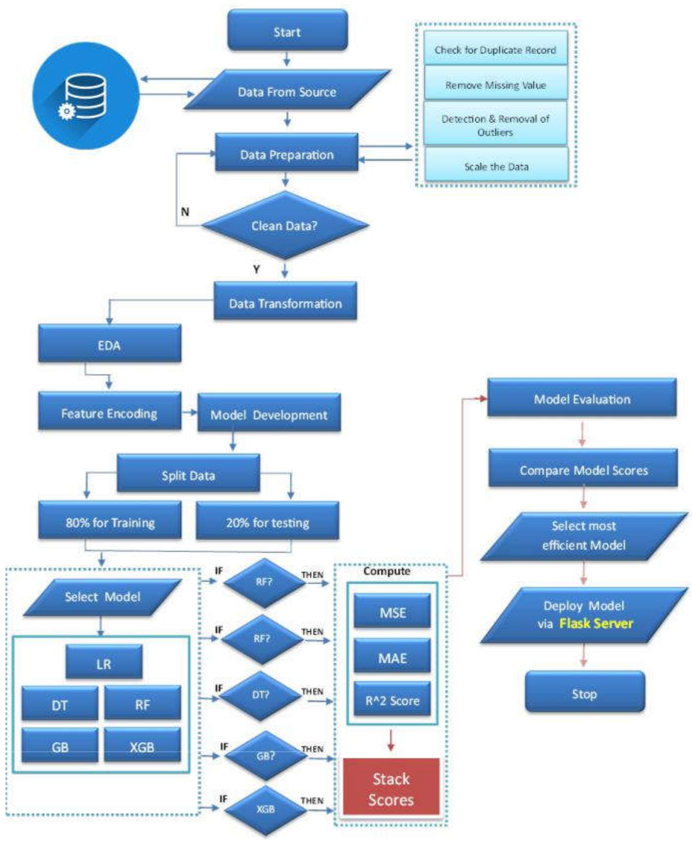


Figure 1.1: Model Development and Deployment Architecture

Criticality of Residential Rent Price Prediction in Lagos

- With its rapid urbanization and population growth, Lagos faces significant challenges in housing provision and affordability.
- Predicting residential rent prices accurately is crucial for tenants seeking affordable housing options and landlords aiming to optimize rental income in a competitive market.
- A well-functioning rental market is crucial for economic efficiency.
- The predictive models can assist in aligning rental prices with the actual value of properties, promoting a more efficient allocation of resources in the real estate sector. Landlords can benefit from the study when they have better understanding of factors that are influencing rental prices of different areas in Lagos. This knowledge allows them to set competitive and reasonable rent rates, attracting more potential tenants.
- Insights into the impact of external factors on rental prices can inform urban planning and development strategies.
- Government authorities and policymakers can use this information to guide infrastructure projects and housing policies.
- This study is significant because it provides a new data-driven model for predicting the cost of house rent in Lagos.
- The model enhances transparency in the rental market which helps both tenants and landlords make more informed decisions.
- It contributes to the broader field of real estate analytics and helps address rental pricing challenges in various urban settings.
- Generally speaking, the importance of this study rests in its potential to foster a more transparent, efficient, and equitable rental market, benefiting both individual stakeholders and the broader community in Lagos, Nigeria.

Motivation

Lagos State, located in southwestern Nigeria, is one of the country's most populous and economically vibrant regions. It serves as Nigeria's commercial and financial hub, attracting a diverse population seeking opportunities in various sectors such as finance, technology, and entertainment. Hence, the impetus driving this research stems from a fundamental desire to address the complexities inherent in the Lagos real estate market, particularly concerning the determination of equitable house rent prices. By developing a sophisticated predictive model harnessing machine learning methodologies, my aim is not only to offer stakeholders a tool for informed decision-making but also to catalyze a transformative shift towards greater market efficiency and transparency. This predictive model is poised to serve as a multifaceted solution, offering insights into prevailing rental trends within specific locales, bridging informational gaps that may exist between landlords and tenants, and providing forecasts of rental prices with a forward-looking perspective. By sidestepping the need for intermediaries and offering direct access to predictive rental data, this approach not only empowers stakeholders but also fosters a more equi-table and informed rental landscape. Ultimately, the envisioned outcome of this endeavor is to facilitate more harmonious and efficient interactions within the Lagos housing market, fostering a climate of transparency and equitable decision-making for all involved parties.

Aim and Objectives

Aim and objectives of this research is as follows;

Aim

The sole aim of this study is to develop and deploy via web application, residential rent price predictive model based on machine learning technologies.

Objectives

(a) Engaging different predictive algorithms in order to have options of get-ting best model for the rental prediction.
(b) Evaluate models on several statistical metrics to be able to ascertain the weaknesses and strength of each selected algorithms.
(c) No to base judgment of the assessment of the models on only one parameter but to combine and compare result of each algorithms on these metrics before model selection.
(d) To select the best performing model for the prediction
(e) While training the models, to put into consideration, factors influencing rent prices in Lagos, including population density, proximity to amenities, transportation infrastructure, and economic indicators.

Scope and Limitations

In recent years, the demand for housing in Lagos, Nigeria, has surged significantly due to rapid urbanization, population growth, and economic development. As a result, leveraging data-driven techniques and machine learning algorithms, this research envisioned to carry out an in-depth investigation into the factors influencing residential rent prices and the development of predictive models to forecast house rents in Lagos.

Scope

(a) This study focuses on the city of Lagos, Nigeria, considering the unique characteristics and dynamics of its real estate market. The findings and predictive model may not be directly applicable to other cities or regions with distinct housing market dynamics.
(b) The scope is limited to rental housing, and the predictive model aims to estimate rental prices. Factors influencing property purchase prices are not included in this study.
(c) It is also worthy of note that this study primarily considers historical data and aims to develop a predictive model based on past trends. While the model may account for some dynamic factors, its ability to predict rental prices is inherently tied to historical patterns.
(d) Furthermore, the study focuses on a selection of machine learning algorithms suitable for regression tasks. The scope does not include an exhaustive exploration of all possible algorithms, because the choice of algorithms is based on their suitability for the specific task.
(e) The development of the user-friendly application is within the scope, but detailed design aspects or extensive user interface testing may not be covered comprehensively. The emphasis is on functionality and usability rather than extensive design considerations.
(f) The study considers factors such as trends of economy, infrastructure development, and changes in policy to a certain extent. However, it may not encompass the entire spectrum of external influences affecting the housing market.

Limitations

(a) Limited availability of comprehensive and up-to-date data on rental properties in Lagos may pose some challenges. The predictive model's accuracy is contingent on the quality and credibility of the available data.
(b) Again, the housing market is subject to constant changes influenced by economic conditions, government policies, and other external factors. The predictive model may have limitations in adapting quickly to abrupt market shifts.
(c) The findings and predictive model developed for Lagos may not generalize seamlessly to cities and regions in other places that have different economic contexts and real estate.
(d) The study's user-friendly application development relies on technology, and limitations such as platform compatibility, internet access, or techno-logical literacy may affect its reach and impact.
(e) Moreso, this study may not delve deeply into legal and regulatory aspects that influence the housing market. Compliance with local laws and regulations is assumed but not exhaustively explored.

Research Structure

This project is structured in seven comprehensive chapters, each serving a distinct yet interconnected purpose.

Chapter 1 give introduction to the research topic, provides background and context, states the research problem, objectives, and significance, and outlines the structural flow of the research work.

Chapter 2 gives comprehensive review of the existing studies that are related to residential rent price prediction, machine learning algorithms, and relevant studies conducted in similar contexts. This chapter lays foundation for the re-search by synthesizing previous findings and identifying gaps in the literature.

Chapter 3 details the procedures involved in collecting and preprocessing the data used for rent price prediction. It elaborates on data sources, data cleaning techniques, feature engineering methods, and any transformations applied to the dataset to ensure its suitability for machine learning analysis.

Chapter 4 covers the process of training machine learning models for rent price prediction. It discusses the selection of algorithms, algorithm performance evaluation techniques, model selection criteria, and testing methodologies employed to assess the effectiveness of the developed models.

Chapter 5 focuses on the deployment of the trained machine learning models through Flask server. It describes the setup of the deployment environment, integration of the models into real-world applications, technologies used, soft-ware environment, and the development of user interfaces for accessing and utilizing the rent price prediction system.

Chapter 6 presents a detailed analysis of the results obtained from applying the trained machine learning models to the dataset. It includes insights derived from the analysis, discussions on the implications of the findings, and recommendations for stakeholders. This chapter synthesizes the key findings of the study, discusses their significance, and provides guidance for future research and practical applications.

Finally, Chapter 7: draws the curtains of the thesis, offering a comprehensive synthesis of the findings, reiterates the significance of the research, and concludes with reflections on the study's contributions to the field. It also outlines limitations and suggests avenues for further research.

Furthermore, recommendations were offered for future research endeavors, culminating in a nuanced and reflective conclusion that encapsulates the essence of my inquiry.

LITERATURE REVIEW

This chapter gives a comprehensive review of previous studies on the residential rent price prediction, with a focus on the context of Lagos, Nigeria. The review encompasses relevant technologies used, software environment, various methods, models, and factors influencing house rent predictions, gaining insights from academic research and practical applications in the real estate industry.

Residential Rent Price Prediction Overview

Initially, in a typical machine learning scenario, the aim is to identify a model that accurately captures the interaction that exists between data features and quality levels. While this concept may seem abstract, breaking down the process into logical steps reveals a systematic approach to completing a machine learning project. Subsequently, our project focuses on a supervised regression machine learning problem. To initiate such a task, it is essential to have both features and targets within the dataset. Features comprise descriptive data information, while targets represent the information the model is trained to pre-dict. This process, termed supervised learning, guides the model in discerning patterns between features and targets, enabling it making predictions in line with data features. Moreover, supervised machine learning uses data that is set aside for training to train the model, followed by validation using test data. Upon successful validation, the trained model is ready for deployment with production data.

Factors Influencing Cost of House Rent

Several factors influence house rent costs, including location, property characteristics, neighborhood amenities, economic conditions, and demographic trends. Location is a significant determinant, with rents varying significantly across neighborhoods and cities according to [3], [4] and [5]. Property characteristics, such as size, age, and amenities, also play crucial roles in knowing rent costs. Additionally, neighborhood amenities like schools, parks, and public transportation, can affect rental demand and prices. Economic factors, such as employment expansion, inflation rates, and interest rates, play crucial roles in ascertaining how affordable and dynamics is the rental market. Addition-ally, demographic shifts, including population growth, household formation, and migration patterns, significantly affect both rental demand and supply dynamics.

Previous Studies Reviewed

In contemporary times, the quest for suitable housing gained significant attention across all age groups and family statuses, especially for individuals relocating to bustling business hubs like Lagos, Nigeria, for residence or employment [6]. The primary concern during the house-hunting process revolves around whether the rental cost aligns with one's budget and whether the cost-effectiveness of a residence is reasonable. Consequently, numerous residential housing search applications and websites prioritize displaying house rental costs as crucial outputs of their search tools. Moreover, the integration of in-creasing numbers of smartphone users with mapping applications and Geo-graphic Information System (GIS) software facilitates the real-time visualization of rental cost distribution across different areas of a city, thereby enhancing users' spatial understanding [7]. Geospatial technology has emerged as a popular research domain in spatial science and various application areas, with researchers recognizing residential rent price prediction as a significant issue in both academic and industrial contexts due to its multitude of related factors and complex spatial techniques [8]. The problem statement elucidates the rationale behind choosing Rresidential Rent Price Prediction Based on Machine Learning (RRPPBML) as the research topic. The aim is to furnish individuals with improved recommendations for finding satisfactory residential accommodations [9]. Numerous studies affirm the paramount importance of house rent prediction in the real estate sector, offering invaluable insights for landlords and tenants alike [10]. Accurate predictions empower landlords to establish competitive rental rates, optimize rental income, and make informed property investment decisions. Similarly, tenants benefit from reliable rent predictions, aiding them in making well-informed choices regarding affordability, location preferences, and budget allocation. Consequently, this area of research has attracted considerable attention from scholars and industry professionals, resulting in the development of various methodologies and frameworks for house rent prediction. A review of existing literature reveals several studies utilizing real house rent price data and implementing machine learning models, employing a diverse array of techniques [11]. Notably, regression-based models have emerged as successful tools for house rent prediction, outperforming other methodologies such as graph-based methods. Among regression-based models, decision trees exhibit superior predictive accuracy compared to alternative techniques [7].

Consequently, decision tree models have gained popularity in recent years, prompting the adoption of various methodologies to assess decision-making processes, with a particular focus on analyzing elasticity and demand shifts, especially within the context of Lagos City, Nigeria [6].

Significance of Residential Rent Price Prediction

The significance of residential rent price prediction cannot be overstated in the realm of real estate. This predictive tool holds immense value for landlords, tenants, and property investors alike, guiding them towards well-informed decision-making processes. Through the aid of the residential rent price predictive model, stakeholders have the privilege of fine-tuning rental rates, promote profitability, and mitigate financial risks that could be linked to property in-vestments.

Methods for House Rent Prediction

In the world of house rent prediction, varieties of methodologies have been investigated, ranging from traditional methodologies to contemporary techniques. This include statistical regression models, machine learning technologies, and hybrid techniques. Statistical regression models which include linear regression and generalized linear models, have gotten recognition as a result of their simplicity and being easy to interpret. On the other hand, machine learning technologies such as gradient boosting, random forest, decision trees etc. are known for their ability to effectively handle complex data relation-ships, as emphasized by [12]. Moreover, hybrid methods combined strengths of both statistical and machine learning techniques to enhance predictive accuracy. [12].

Traditional Approach for Rent Price Estimation

The fundamental approach to traditional rent price estimation entails sampling diverse properties within a particular area and collecting data on rental rates along with relevant attributes such as the number of bedrooms and bathrooms, square footage, and amenities like a pool or gym. Following this, techniques such as simple linear regression are employed to predict rental prices based on one or more chosen features. A study conducted in 2017 proposed an iterative approach that incorporates geographical proximity into rent price estimation [13]. Initially, rent prices are estimated using traditional methods, and for each property, formula residuals are computed as the disparities between the estimated and actual rent prices. Subsequently, a multivariate spatial autoregressive model is employed to select features and estimate parameters, leveraging spatial correlation between properties to enhance rent price estimations significantly. The study asserted that this novel method outperforms classical approaches in the area of prediction computational and accuracy efficiency. However, within the literature, conventional methods are often criticized for yielding biased and inconsistent parameter estimates and reduced prediction accuracy. These shortcomings are attributed to violations of under-lying assumptions in parameter statistical models, such as linearity between the dependent variable and features, and independence of errors. Additionally, many pertinent features influencing rent prices are often overlooked due to constraints in manual feature selection. In the contemporary landscape, there's been a discernible shift in research focus and effort from traditional methods towards machine learning techniques in rent price prediction. As evidenced in my review, studies aimed at enhancing traditional methods, like the 2017 spatial model proposal, primarily focus on developing innovative strategies to integrate new knowledge or features into parameter models [13] and [14]. However, there are limited endeavors to wholly replace traditional methods with a new paradigm. The emergence of big data and the capacity of modern computational resources to handle extensive data processing have accelerated the evolution in the fundamental methodology of rent price estimation. Nonetheless, it's imperative to acknowledge that such a transformation necessitates retraining domain researchers and practitioners to stay abreast of new statistical and computational techniques. Furthermore, methodological justifications and comparisons among diverse rent price estimation strategies must be addressed comprehensively.

Machine Learning Techniques in Residential Rent Price Prediction

The integration of cutting-edge technologies in real estate prediction models has become a subject of significant research in recent time. Studies by [5] and [15] clearly elucidated the efficacy of employing machine learning algorithms for accurately predicting residential rent prices. Machine learning provides an attractive alternative to traditional hedonic regression analysis. There are several reasons why machine learning is a promising technique in this area. First of all, machine learning algorithms are highly flexible and can easily model non-linear patterns in data. In contrast, traditional regression methods require the researcher to specify in advance, the functional form of the model. By being able to handle non-linearity, machine learning methods can potentially uncover more complex relationships between the dependable variable and undependable variable. Secondly, machine learning techniques have built-in methods for handling large numbers of predictor variables. The third advantage of ma-chine learning is that these techniques often provide automatic ways to deal with missing data. In our dataset, there are quite a few missing records for the year when the house was built and the year when it was renovated. Rather than deleting these records or doing imputation, which might bias the results, some machine learning methods can naturally accommodate missing data in the algorithms. The last but not the least advantage of machine learning is that it can enjoy the benefits of parallel computing. Modern machine learning technologies can easily harness the power of parallelism when building models, and this significantly reduces the computational time when we com- pare with traditional regression, especially when the sample size is very large. These four reasons suggest that machine learning provides a good opportunity to model the complexity of house rent price and yield more accurate prediction results. With the recent rapid development in computational solutions and machine learning tools, such prediction could be conducted in a much more efficient and effective way. However, it should be noted that these algorithms usually involve careful tuning of parameters and a lot of attention should be paid to avoid over fitting. Next, a brief overview of the methodology behind the rent price prediction will be provided, followed by a description of the machine learning algorithms used in this research work. Moreover, research by [12] emphasized the significance of feature engineering and data preprocessing techniques in enhancing the performance of predictive models. It is possible for researchers to improve the predictive capability of their models and prevent the impact of irrelevant information and noise as they identify and transform related features from the unprocessed dataset. Studies by [3] and [4] clearly emphasised the importance of using visualization tools like Tableau public. These tools allowed stakeholders to explore spatial trends and patterns in real estate markets through interactive charts and intuitive dashboards. Visualizing complex data in a comprehensible way, researchers can communicate insights effectively and promote data-driven decision-making processes. Additionally, the inclusion of web technologies such as Flask, HTML, JavaScript, and jQuery for building interactive web applications has also been explored in the studies by [16]. These web-based platforms make researchers to share predictive model results and engage stakeholders in a dynamic and interactive manner. Providing user-friendly interfaces and real-time updates, web plat-forms enhance the accessibility and usability of predictive models, ultimately empowering users to take good decisions in the real estate domain. In general, previous studies reviewed so far highlighted house rent prediction as a vital component of real estate analysis and decision-making, with significant implications for landlords, tenants, and property investors. Leveraging diverse methodologies and frameworks, it becomes easy for researchers to develop ro-bust predictive models that brings market transparency, optimize rental pricing strategies, and facilitate efficient allocation of resources in the real estate sector [17].

Research Gaps

Residential rent price prediction is an important area for research in Nigeria, most especially, Lagos State, given its status as a major economic and commercial hub of the country. Several studies have addressed this topic, aimed at enriching our understanding of rent dynamics and provide valuable insights for stakeholders in the real estate sector. However, there are still a lot of gaps that warrant further investigation.

For instance;

(a) Studies on Geospatial Analysis of Rent Patterns conducted by Olatunji 2018 [7] mainly focused on geospatial examination of rental trends in Lagos State, Nigeria. The study utilized Geographic Information System (GIS) technology to analyze spatial variations in rent prices across different neighborhoods. Though, the research shed light on the geographical factors influencing rent prices but there is obvious limited scope of anal-ysis. Considering other important factors such as housing quality, amenities, or socio-economic indicators that could also influence rental trends is of great importance. Exploring a broader range of variables could pro-vide a more comprehensive understanding of the rental market dynamics in Lagos State.

(b) Furthermore, research conducted on Machine Learning Approaches to Rent Prediction by Adeleke 2020 [6] explored the application of ma- chine learning algorithms for rent price prediction in Nigeria. Leveraging datasets containing various property attributes, the study developed predictive models to forecast rent prices accurately. Beautifully enough, the research highlighted the potential of machine learning techniques in improving rent prediction accurately, but did not synchronise the model with web application that could make it easily accessible to the stakeholders.

(c) Studies on Impact of Socio-Economic Factors on Rent Prices by Okonkwo 2017 [10] extensively investigated the impact of socio-economic factors on rent prices in Lagos State. Through statistical analysis, the study identified key determinants such as income levels, employment rates, and population density, providing valuable insights for landlords and property investors. However, there exists need to include building of models that can give accurate prediction on residential rent price leveraging on emerging technologies.

(d) Access to data and it's quality is another major research gaps in residential rent price prediction in Nigeria. Limited availability of comprehensive datasets that is up-to-date hinders the development of accurate predictive models. Therefore, most previous predictive models are limited to just the current year of the research. There is need for future studies to focus on improving data collection methods and enhancing data quality to facilitate more robust rent prediction models.

(e) In addition, Many existing studies focus primarily on economic and demographic factors influencing rent prices, overlooking the role of socio-cultural factors. Obviously, future research need to explore further on the integration of socio-cultural variables, such as cultural preferences, lifestyle trends, and community dynamics, to be able to place in the hands of stakeholders and policy makers, a comprehensive understanding of rent in Lagos, Nigeria.

Summarily, while several studies have contributed greatly to our understanding of residential rent price prediction in Nigeria, there remain significant research gaps that warrant further exploration. Most of the relevant studies carried out have been focusing on either fundamental analysis for long-term real estate in-vestments or improved hedonic models by incorporating spatial analysis, neither of which is directly relevant to those determinants in rent price prediction nor completely addressing the drawbacks of the traditional methods mentioned previously.

Moreover, despite the significance of Lagos State to the Nation's economy, no research has been able to deployed predictive model on residential rent price prediction on web application for Lagos State. This is an obvious gap that this research aimed at addressing. Having a well trained model that can accurately predict is one step, another higher expected step is to make the predictive model seamlessly accessible to stakeholders.

Addressing these gaps through interdisciplinary research efforts envisioned to birth a more accurate and comprehensive predictive model, deployed to a web application using emerging technologies is the motive behind this present re-search work. This ultimately will benefit stakeholders in the real estate sector and contributing to sustainable urban development in Nigeria.

METHODOLOGY

This chapter outlines the measures that were employed in evaluating the accuracy of the predictive model. The data collection process describes the variables used in the analysis as well as presenting the methodology employed for the residential prediction. It provides a detailed overview of the data sources, data preprocessing steps, and the predictive models utilized in this study. Also, a comprehensive discussion is provided concerning the performance metrics used to access accuracy and precision of the predictive model. By explaining the roles and limitations of each metric, insight is provided into how a combination of these measures offers a comprehensive evaluation of the predictive model's performance. Moving forward, model selection procedure is elaborated upon, with an emphasis on the adoption of "cross-validation." This method examines how the outcome of a statistical analysis generalize to in-dependent dataset. In a five-fold cross-validation process, various machine learning methodologies which include linear regression, random forest, decision tree, extreme gradient boosting, and gradient boosting, were used to evaluate the dataset. Subsequently, Predictions generated by these models are compared with the true values using RMSE as the performance metric. Following model evaluation, the best performing model is fitted to the entire training data to generate final predictive results. In the subsequent phase, feature selection is conducted to identify and select a subset of the most relevant features. This process aids in constructing new variables based on the chosen features, a method known as feature engineering. Notably, a new variable called "Eco-Class" is introduced, representing the economic per income level of people in specific towns and communities in Lagos. This variable provides insights into the economic context of different areas within Lagos, leading to more accurate predictions and informed decision-making in the real estate market. During the data preprocessing phase, missing data was addressed by removing variables with a large proportion of missing values and filling missing values in numeric columns with the median of the respective column. Descriptive statistics are then computed to summarize the central tendency, dispersion, and distribution shape of the dataset, enhancing understanding of each attribute. In summary, this study offers a detailed exploration of performance metrics, model selection procedures, feature engineering techniques, and data preprocessing methodologies employed in residential rent price prediction.

Source of Data

The required data are the rental listings in the city of Lagos, Nigeria. The first challenge is deciding which website to scrape the data from. A private real estate was chosen over a crowd-sourced one due to availability and consistency [18]. But as expected, the data is messy. A lot of the fields in the dataset contain missing or inconsistent values. Some of the critical fields that need to be cleaned up include the number of bedrooms, bathrooms, names of towns and community and the rental price. So, the next step was preprocessing the data. Missing values were dealt with in different ways. For example, for the "number of bathrooms" column, if the value was missing, it was replaced with the most common number of bathrooms in the dataset. If a standard value was not available, the specific row was dropped from the dataset. However, due to the presence of vast varieties in the data size from different types of houses, apartments, or other types of rental options, the method of replacing missing bathroom values using the mean or the most common value gave similar results. Another approach to handle missing values is a simple solution - dropping the rows that contain missing values. By viewing the data, it is clear that the missing values for columns such as "Laundry in building" are missing because Boolean fields in CSV file are filled with 1's and 0's. The flow data source and preprocessing is pictorially shown in figure 3.1 below.

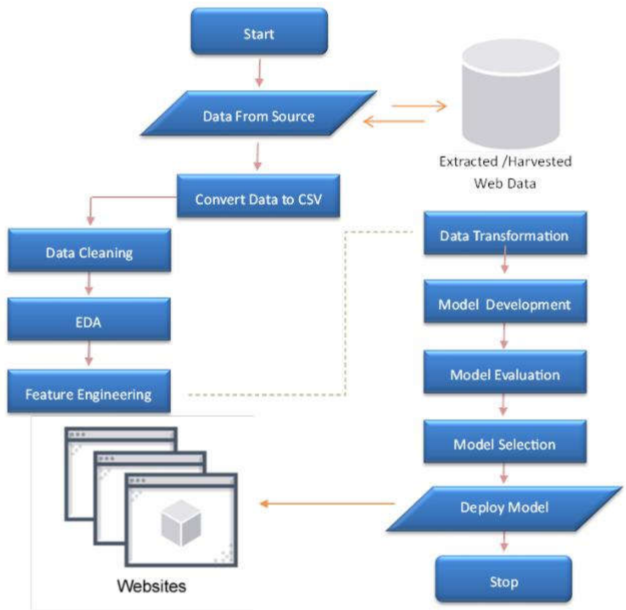


Figure 3.1: Data Analysis Flow

Data Collection

The predictive model's accuracy and reliability hinge on the richness and relevance of data sources and variables employed. Hence, the dataset used for this research work originates from PrivateProperty.com.ng Ltd; a property listing company in Nigeria (https://www.privateproperty.com.ng). The dataset collected contains over 90000 properties (before cleaning). The dataset is de-tailed with columns including date of listing, number of bedrooms, bathrooms, toilets, location (address) and rent cost. The Real estate websites serve as goldmines for data, providing valuable insights into property listings, market trends, and rent fluctuations [4], [3] and [15]. This section explores the integration of data from real estate websites, emphasizing the significance of web scraping techniques.

Web Scraping Techniques

To harness the full potential of real estate websites, web scraping techniques are employed. Web scraping involves extracting data directly from web pages, facilitating the retrieval of real-time information. Python libraries such as Beautiful Soup and Scrapy are commonly used to automate this process. In this research work, specifically, I used Beautiful Soup which enable the collection of up-to-date and accurate data. The source code screen shot is shown in figure ?? below.

Scraped Data Stored as CSV Format

The meticulously extracted data from Private Property estate is saved as a CSV file format. This is done in order to preserve the crucial information obtained from the estate's listings. Saving the data in this structured format makes it easily accessible for further analysis and manipulation. This systematic organization of data underline a commitment to accuracy and efficiency in managing the estate's information.

Ethical Consideration

While web scraping provides valuable data, ethical considerations regarding data ownership and usage are paramount. This study adheres to ethical standards, ensuring that data extraction respects privacy policies and terms of use of the targeted websites. Figure 3.2 shed more light.

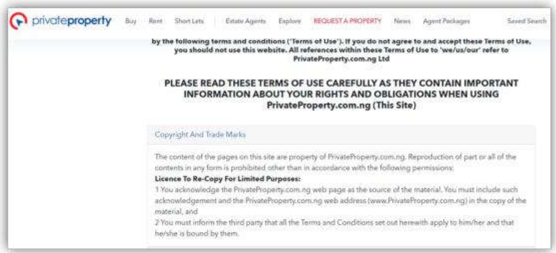


Figure 3.2: Private Property Ethical Policy

The combination of web scraping and a comprehensive set of variables enhances the predictive model's accuracy and applicability to dynamic real estate markets.

Overview of the Harvested Data

The dataset utilized for this research work spans from 2014 to 2024. This comprehensive time frame covers a decade of real estate data, enabling me to cover long-term trends and instabilities that exist in rental prices. The dataset contains information on various house attributes and their corresponding rent prices in Lagos. Key features include the number of bedrooms, bathrooms, square footage, location (neighborhood or district), amenities, and other relevant factors. The target variable is the rent price, represented as a continuous numerical value.

Variables for Prediction

The variables considered in the prediction model encompass a spectrum of factors influencing house rents. These include but are not limited to;

(a) Location: The geographical aspect is a pivotal determinant of rental prices. Proximity to amenities, public services, and employment hubs significantly impacts housing costs.

(b) Property Features: Details about property itself, such as number of bed-rooms, bathrooms, square footage, and other amenities are all important when it comes to predicting rent.
(c) Indicators of the Economy: Variables like local employment rates, inflation, and economic growth contribute to understanding the financial dynamics of a region, influencing housing demand and prices.
(d) Amenities and Neighborhood Characteristics: Proximity to schools, parks, public transportation, and safety indices are essential variables capturing the quality of life and attractiveness of a location.

Data Preprocessing

Data preprocessing techniques are implemented to cleanse and refine the dataset in preparation for analysis. This involves addressing missing values, outliers, and categorical variables, as well as transforming categorical variables into numerical representations. One-hot encoding is employed specifically for neighborhood categories to facilitate numerical processing.

Cleaning and Formatting

The rental listings data underwent cleaning and formatting to tackle issues of missing values, outliers, and inconsistencies. Missing values were entered through mean interpolation. Outliers were identified and treated by statistical methods. Categorical variables were encoded, and numerical variables were scaled and normalized to ensure uniformity across the dataset.

Treating Missing Values

The data required extensive cleaning and transformation as some of the necessary information were not explicitly available. One of the initial challenges encountered in the dataset was the presence of missing values across several features. Missing values are capable of affecting the performance of predictive models significantly, if not appropriately addressed. The following steps were taken to handle missing values;

(a) Identification: Missing values were identified across all features using descriptive statistics and visualization techniques.
(b) Imputation: For numerical features, missing values were imputed using appropriate techniques such as mean, median, or mode imputation. Categorical features were imputed with the most frequent category or a separate 'missing' category.
(c) Validation: After imputation, the dataset was validated to ensure that missing values were effectively handled and no significant data loss occurred.

Addressing Outliers

Data points that deviate significantly from the rest of the data can negatively affect statistical analyses and model predictions. In this house rent prediction task, outliers refers to high or low rent prices or extreme values for house at-tributes. The following steps were taken to handle outliers;

(a) Identification: Outliers were identified using visual inspection, statistical method IQR and domain knowledge.
(b) Remove the Outliers: Outliers were treated using appropriate techniques of transformation and removal (if deemed erroneous or irrelevant). Ex-ample is in figure 3.3 below.

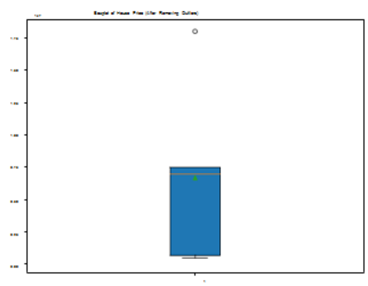


Figure 3.3: Box plot of House Price after removing Outliers

The removal of outliers from the 'House Price' feature is resulted in dataset that is more robust, representative, and suitable for further analysis. How-ever, validate of the impact of outlier removal is inevitable in this re-search.
(c) Validation: Having treated the outliers, dataset was validated to ensure the distribution of data remained consistent and the integrity of the dataset is preserved.

Handling Inconsistencies and Anomalies

Inconsistencies and anomalies in the dataset, such as typographical errors, in-correct data types, or erroneous entries, are capable of compromising the validity of analyses and predictions, hence I did not overlook them.

The following steps were taken to address inconsistencies and anomalies;

(a) Data Type Conversion: This was done to ensure variables have correct data type so as to facilitate analysis and modeling.
(b) Standardization: Standardize categorical variables (e.g., neighborhood names) was done to ensure consistency and eliminate duplicate entries.
(c) Error Correction: Correct typos, misspellings, and formatting errors in the dataset using manual inspection and automated methods.
(d) Multiple and inconsistent values for Date Added column. The different cases include multiple dates (Added date and Updated date), Up-dated/Added Today instead of actual date etc.
(e) The House Price column contains special characters such as currency symbol, commas, slashes, words like month, year.
(f) Address also contains special characters and HTML codes.
(g) The data combines diverse property cases including properties like shops, office space, halls etc. It also lists a few properties for sale. The focus of this research work is for Rents only.
(h) Removal of irrelevant records that does not align to the goal of this project (House Rent prediction in Lagos). Records for Shops and business areas rental, Event Hall rentals, warehouse and land for sale were removed.
(i) Date Added column was transformed to be valid date type.
(j) House Price was formatted appropriately as integer and currency symbols removed.
(k) All special characters attached to numeric columns such as number of bedrooms, bathrooms and toilets were removed.
(l) Some rents were listed as 'per day', 'per month' and others 'per year'. This was totally standardized for all to be 'per year' which is the popular duration of rent payment in Lagos.
(m) As Town and Community of property do not explicitly exist in the dataset, except for details address which contains this information. A comprehensive list of Lagos communities and towns was gotten and used appropriately to match and extract the town and community which each listed property belonged to.

A thorough work of data mapping was done in this aspect to ensure that no property was wrongly tagged with an associated town and community. In addition to an advantage of vast knowledge of the domain being studied, relevant resources were accessed also in the course of the data cleaning.

Exploratory Data Analysis

This refers to process of analyzing and visualizing data for better understanding of its underlying patterns, relationships, and distributions. This phase of data analysis is to give visual insights into the dataset and identify potential trends of data that are out of order before applying more advanced statistical techniques.

In this research, the data exploration stage expounded on specific columns and their association such as;

(a) Houses by bedrooms: Total number of houses for each category of house type (i.e. by number of bedrooms) and percentage that each house type represents.

(b) Average cost of House Price by House type: This analysis aims to pro-vide a detailed understanding of how housing prices vary across different types of houses, shedding light on market trends and preferences. In order offer insights into any variations or outliers present within the dataset, bar charts visualizations such was employed to visually represent the distribution of house prices across different types.

(c) Number of Houses by Town: Number of houses listed by towns and the percentage that each represents in the entire dataset.

(d) Number of house listed by year : The number of houses listed for rent can vary significantly from year to year due to various factors such as economic conditions, population growth, housing market trends, and government policies.

(e) Average cost of house by year by house type: In other words, the aver-age cost of houses for each combination of year and house type.

(f) Number of Houses by Town:

The analysis begins by categorizing houses based on their location within various towns across Lagos State. This segmentation allows for a detailed examination of housing availability and distribution patterns within each town.

The figures 3.4 below shed more light the Exploratory Data Anaysis (EDA) carried out on the dataset for this research.

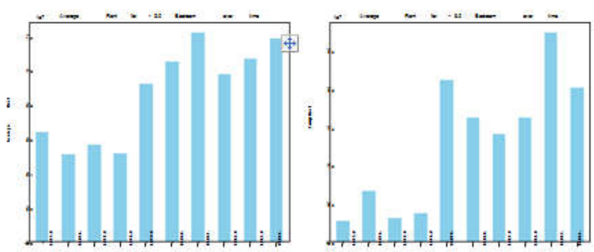
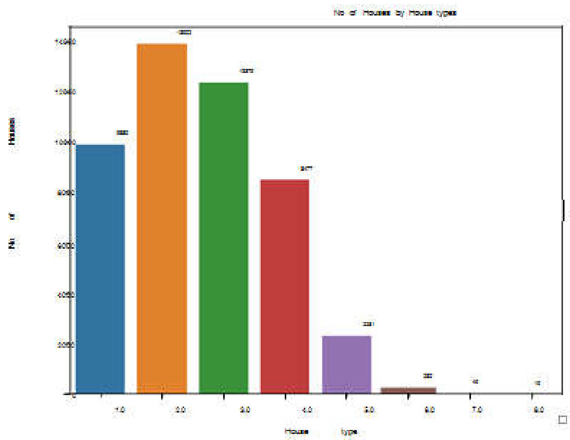


Figure 3.4: Instances of Exploratory Data Analysis



Feature Engineering

Feature engineering involved creating new variables or features from the existing data to enhance predictive performance. For instance, the research raw data was having initial feature of Title, Description, Date Added, and Price only as shown in the table 3.1 below.

Table 3.1: Data before Feature Engineering

Title	Description	Date Added	House Price
Premium 3 Bed-room Apartment +bq+pool+gym	3 BEDROOM FLAT & APARTMENT For Rent	Updated 20 Aug 2023, Added 23 May 2023	7,000,000
Grade A Commercial Office Space In Victoria Island	OFFICE FOR RENT	Updated Today	150,000/sqm
Contemporary 5 Bed- room Detached Duplex	5 BEDROOM DUPLEX FOR RENT	BEDROOM	DUPLEX
Luxury 3 Bedroom Flat With Bq	3 BEDROOM BLOCK OF FLATS FOR RENT	Updated Today	8,000,000/year
Tastefully Furnished Office In Vi	OFFICE FOR RENT	Added Today	75,000/month

However, during the feature engineering exercises, new features such as Bed-room, Bathroom, Toilet, Address, Year Added, Payment Duration, Community, Town, and Ecco Class were added as seen in the table 3.2.

Table 3.2: Dataset with new features added

Date Added	House Price	Bedroom	Bathroom	Toilet	Year Added	Community	Town	Eco_class
18 Jan 2024	60000000.0	2.0	2.0	3.0	2024.0	Ologolo	Lekki	Upper
18 Jan 2024	60000000.0	2.0	2.0	3.0	2024.0	Ologolo	Lekki	Upper
10 Jan 2024	1000000.0	2.0	3.0	3.0	2024.0	Bariga	Gbagada	Middle
10 Jan 2024	1000000.0	2.0	2.0	3.0	2024.0	Bariga	Gbagada	Middle

Furthermore, Some extra works was done on the features to help improve the model performance. To populate new features that were added, data was extracted from Title in the existing data to populate Bedrooms, Birthrooms, Toilets and Address, while the Year Added feature was popullated from the Date Added. So also, Duration feature was populated from the House Price feature which has amount lumped together with the duration. Eco class feature was also added which refers to the economic per income level of people in specific towns and communities in Lagos.

Areas such as upper Lekki, Ikoyi and Victoria Island are known to be towns of the Upper Class while areas in and around Ikeja, Surulere, Ajah, Maryland and Ogba and for the Lower Upper. The middle class dwell in majority of the remaining areas with the exception of locations like Mushin, Ajegunle, Oshodi, Makoko, Bariga and Ketu which is known to be settlings for the poor.

MODEL TRAINING

In these days of intelligent computing and artificial intelligence technologies, being able to design, train model is of paramount importance. and development holds pivotal importance in designing a robust and efficient predictive systems. Model training requires exposing a machine learning algorithm to data, enabling it to discern underlying patterns and relationships within the dataset to autonomously make predictions or decisions. As data becomes increasingly ubiquitous across various domains and industries, the importance of proficient model training and development cannot be over emphasised. This chapter extensively discussed the fundamental principles of model training and development, elucidating the essential techniques, and best practices employed in this all important stage of machine learning workflows.

Machine Learning

At this age when machine learning is at the fore front of modern technologies and innovation, having the technical knowhow to extract valuable information, and to generate precise predictions, with streamline decision-making processes has become crucial. Machine learning technologies is a strong tool that brings a paradigm shift in problem-solving techniques. Computer is made to learn from given data and take informed actions independently. Instead of explicitly programming a computer to perform a task, machine learning algorithms learn from examples or experiences provided in the form of dataset.

Categories of Machine Learning

Basically, there exist three categories of machine learning, guided learning, unguided learning and reward-based learning.

Guided Learning

Guided learning is all about training model with annotated data and each in-put is associated with corresponding output. The algorithm learns patterns to make predictions or classifications through the labeled dataset. regression or classification trees, logistic regression, ordinary least squares regression, sup-port vector machines, random forest, gradient boosting and neural networks are all examples of supervised learning.

Unsupervised Learning

Unlike the guided learning, unguided learning does not depend on annotated data. alternatively, it concentrates on bringing out composition and structures from unlabeled datasets. This it does by bringing together, similar data points in clusters or reducing the Parameter Space of the data. Hierarchical clustering, and autoencoders are all examples of unsupervised learning category of machine learning.

Reinforcement or Adaptive Learning

Adaptive learning is all about an operative that is learning how to make decisions by interacting with a setting. The agent gets feedback as rewards or penalties due to the actions it takes. progressively, the agent targets to learn a policy that will maximize cumulative rewards. Reinforcement learning applications are in gaming, robotics, autonomous vehicles, and recommendation systems. Popular examples of adaptive learning algorithms are Q-learning, Deep Q-Networks (DQN), actor-critic, and policy gradients methods.

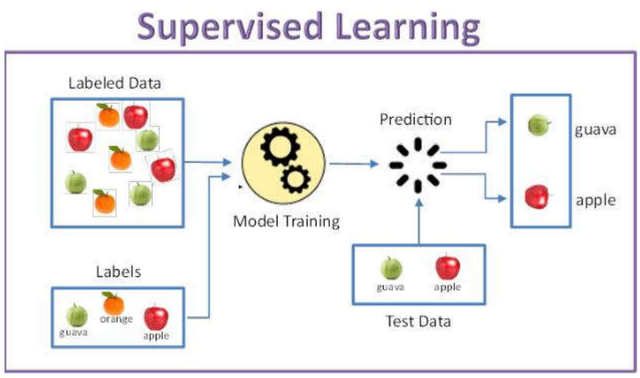
Supervised Learning

The most suitable category of machine learning for this research work of residential rent price prediction project is supervised learning. This is because, this research has to draw knowledge from historical data to learn pattern necessary for prediction. Among the machine learning categories, supervised learning algorithms are the ones that can learn from historical data where both the in-put features (like size, location, number of bedrooms, amenities, etc.) and the corresponding output labels (rent prices) are available.

How does it work?

In the guided learning, models are trained with annotated data and learns about different types of data. After training, model is tested with data that is prepared for testing (which is a subset of the training dataset), and then predicts the output.

The diagram below explains further, how guided learning technologies function.



Suppose we are given different types of fruits as our set of data such as guava, apple and orange. The procedure is to train the model on every type of fruit in the dataset thus;

- if the given fruit shares same shape, size, colour texture and properties as it pertain to an orange, then it will be labelled as an orange.
- or if the selected fruit has same shape, size, colour texture and properties as it pertain to a guava, then it will be labelled as a guava.
- or the fruit has same shape, size, colour texture and properties as it pertain to an apple, then it will be labelled as an apple.
- repeats the steps above until the model is fully familiar with the distinct shape, size, colour texture and properties as it pertain to each category of fruit in the dataset

After completing the training, the subsequent step is to test the model with the test set, and the task of the model is to identify the fruit.

Now that the machine model is already trained with all types of fruits, and then finds a new fruit, it classifies the fruit on the bases of shape, size, colour texture and properties as it pertain to the fruit at hand, then predicts the output.

Steps in Supervised or Guided Learning

(a) Determine the type of training dataset to use
(b) Gather the annotated training dataset.
(c) Divide the dataset into training, testing dataset, and validation of the dataset.
(d) Determine what the input features are in the training dataset, which should have enough knowledge so that the model can accurately predict the out-put.
(e) Determine the algorithm that is most suitable for the model, whether the classification tree, simple regression, random forest, etc.
(f) Execute the algorithm on the training dataset. Sometimes validation sets is needed as the control parameters, which are the subset of training datasets.
(g) Evaluate the accuracy of the model by providing the test set. If the model predicts the correct output, which means the model is accurate.

Benefits of Guided Learning

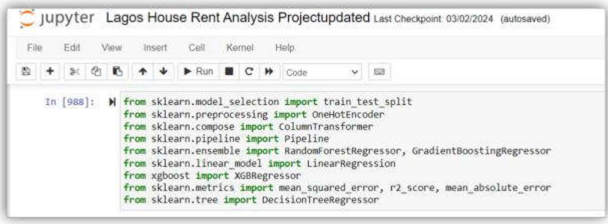
(a) Through the assistance of guided learning, the model can predict the out-put base on prior experiences.
(b) In guided learning, it is possible to have exact idea about the classes of objects.
(c) Guided learning assists us to tackle different real-world problems which include residential rent price prediction, fraud detection, etc.

Machine Learning and Models

Machine learning and models go hand in hand. It is a broad field that en-compasses algorithms and techniques that give rooms for computers to study patterns and infer predictions or decisions implicitly programmed to do so. Models, in the context of machine learning, are mathematical representations or structures that capture these patterns in the data. Machine learning models offer a data-driven approach to rent price prediction, leveraging historical rental data along with various property features to generate accurate forecasts. These models analyze large datasets containing information like location of the property, size, amenities, neighborhood demographics, market trends. Identifying patterns and relationships within the data, machine learning models can make predictions that are tailored to individual properties and reflect the cur-rent market conditions.

Arrays of Algorithms

In the pursuit of accurately prediction, this research work utilizes diverse array of supervised machine learning techniques ranging from linear regression, decision trees, random forest, gradient boosting, to XGBoost in order to develop accurate and reliable models for the rental market prediction, especially, Lagos State, Nigeria house rent.



Regression Analysis

The project on residential rent price prediction is a regression task with the main objective of forecasting a continuous target variable depending on one or more independent variables. This necessitates the utilization of regression analysis to establish a mathematical relationship between house rent costs and diverse predictor variables. This relationship is explained with the expression;

$Y = \beta 0 + \beta 1X1 + \beta 2X2 + . . . + \beta nXn + \varepsilon.$

where;

Y represents predicted house rent cost, X1, X2, . . . , Xn is the independent variables (the predictors), β1, β2, . . . , βn whereas the coefficients is the impact that each predictor has on the rent prices, and ε represents the error term that captures the unexplained variability in the rent prices.

The focus of regression analysis is to ascertain values of coefficients β0, β1, . . . , βn that reduce the estimation of squared differences that is among the actual and forecasted values of the outcome variable. This is done through the use of gradient descent. Gradient descent is an optimization algorithm used to minimize the cost function of a machine learning model by iterative adjusting the model parameters.

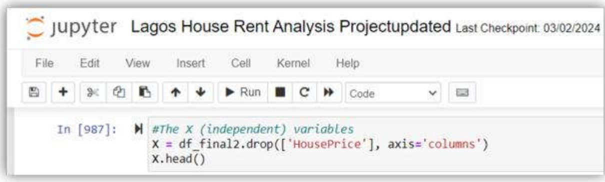


Figure 4.1: The X (independent) variables

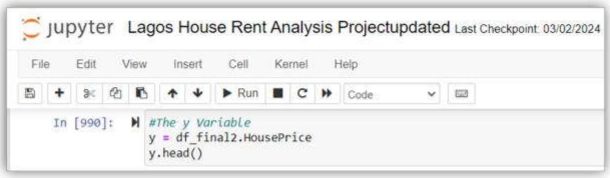


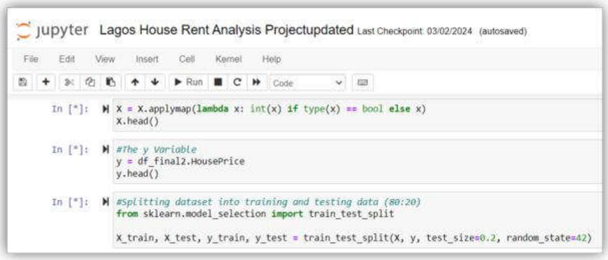
Figure 4.2: Dependent variable Y

Lately, regression analysis has gained notable prominence across diverse do-mains, encompassing the modeling and prediction of house rent prices. This statistical approach offers valuable insights into the underscoring factors that influence rental rates, thereby promoting data-driven decision-making in the real estate industry. Harnessing historical data on housing attributes like number of bedrooms, square footage, location, and more, regression models can discern patterns and trends, enabling accurate predictions of rent prices for new or unseen properties. Thus, regression analysis plays a pivotal role in in-forming rental market dynamics and aiding stakeholders in making informed decisions regarding property valuation and investment strategies. Regression analysis serves as a cornerstone statistical method, enabling the exploration of the intricate relationship between house rent costs and a myriad of predictor variables [19]. This foundational technique aims to quantify the interdependence between a dependent variable (in other words, rent prices) and one or more independent variables (predictors), encompassing various housing features [15].

Analyzing historical data pertaining to these attributes alongside corresponding rent prices, regression models discern patterns and trends, thereby enabling accurate predictions of rent prices for new or unseen properties [19] and [15]. Thus, regression technique plays crucial role in promoting efficiency and effectiveness of decision-making processes within the real estate domain.

Splitting Data to Training and Testing Dataset

Splitting predictive data to training and testing datasets is a an important phase in developing robust and reliable machine learning models for rent prediction. In this research work, the dataset is splitted into 80% of the dataset for Training while 20% of the same dataset is used for testing.



4.7.1 Essence of splitting the dataset

(a) For model performance evaluation Splitting the data helps to evaluate the level of adaptability of our model to new and unseen data. The dataset prepared for training is used to train the model, while the dataset prepared for testing serves as an independent dataset to evaluate the model's performance. This allows us to estimate how well the model will perform on future data as well.

(b) Avoiding Overfitting: Overfitting occurs when a model memorizes the dataset prepared for training rather than capturing the underlying com-positions. Evaluating the model on the dataset prepared for testing can help to detect whether the model has overfitting the training data or not. If the model performs well on the dataset prepared for training but per-formed poorly on the dataset prepared for testing, then, it means there is overfitting.

(c) Parameter Tuning and Model Selection: The splitting techniques facilitate model selection and parameter tuning. Multiple models will be trained and evaluated and the best among them will be selected.

(d) Leakage Prevention: Leakages occur when information from the testing set inadvertently affects the training process, leading to inaccurate performance estimates. Keeping the testing set away from the training set, this can prevent data leakage and obtain unbiased estimates of model performance

Model Selection

In the context of predicting residential rent prices in Lagos State, a variety of algorithms are considered suitable. This section provides a comprehensive overview of these algorithms.

Linear Regression

One algorithm that is commonly used for prediction is linear regression, especially when there's a linear relationship that exists among the features (such as size, number of bedrooms, location, etc.) and the rent price. Linear regression assumes a simple relationship that exists among the predictors and the response variable. If the relationship is non-linear, more complex regression models or other machine learning algorithms may be more suitable. The performance of linear regression algorithm on this research dataset is as shown in the figure ?? below.

Table 4.1: Performance Metrics of Linear Regression Model

Algorithm	MAE	MSE	R^2
Linear Regression	32.38	657.75	0.75

R2 Score of 0.75 shows that the coefficient of determining the measure of the extent of the explained variance in the dependent variable from the predictor variable in this model is 75%. Certainly, this is good but not satisfactory. Hence, the need to test another model for possible better performance.

Decision Tree

Decision Trees are a fundamental tool in the realm of machine learning, widely recognized for their simplicity and interpretability. These models partition the feature space into distinct regions, enabling straightforward decision-making processes. In the Lagos House Rent Prediction Project, Decision Trees serve as a foundational approach for understanding feature importance and initial model exploration. However, due to their inherent limitations in handling complex relationships and tendency to overfit, in this analysis, the decision tree model outperforms linear regression with lower MAE and MSE. Additionally, the R2 Score of 0.86 indicates that the decision tree model explains 86% of the variance in the data, suggesting better predictive performance. The table 4.2 below explains further.

Table 4.2: Performance Metrics of Decision Tree Model

Algorithm	MAE	MSE	R^2
Decision Tree	12.75	356.95	0.86

Random Forest

The random forest model has a higher MAE and MSE compared to the decision tree model, indicating slightly poorer predictive performance. The lower R2 Score is 0.52 suggests that the model explains only 52% of the variance in the data as shown in the figure 4.3. Certainly, this is the poorest performance so far.

Table 4.3: Performance Metrics of Random Forest Model

Algorithm	MAE	MSE	R^2
Random Forest	13.55	1260.75	0.52

Gradient Boosting

Gradient Boosting algorithms operate by sequentially training weak learners to correct the errors of their predecessors. This iterative process results in highly accurate predictive models capable of capturing complex relationships within the data. By leveraging Gradient Boosting in the Lagos House Rent Prediction Project, I was able to harness the predictive power of ensemble learning to achieve superior rent prediction performance of R2 Score is 0.72 as shown in the figure 4.4. The model performs slightly worse than the linear regression and decision tree model, with higher MAE and MSE.

However, the R2 Score is 0.72 indicates that the model still explains a significant portion of the variance in the data.

Table 4.4: Performance Metrics of Gradient Boosting Model

Algorithm	MAE	MSE	R^2
Gradient Boosting	17.36	751.25	0.72

XGBoost

XGBoost is an extension of Gradient Boosting methods, a well known machine learning technique used for regression and classification tasks. It is well known as a result of its scalability and efficiency. By incorporating regularization techniques and advanced optimization strategies, XGBoost further enhances the predictive accuracy and computational efficiency of Gradient Boosting algorithms. In this project, XGBoost serves as a state-of-the-art solution for rent prediction, offering unparalleled performance and versatility of R2 Score 0.90 indicating superior predictive performance as shown in the figure 4.5 below.

Table 4.5: Performance Metrics of Extreme Gradient Boosting (XGBoost) Model

Algorithm	MAE	MSE	R^2
Extreme Gradient Boosting (XGBoost)	12.67	260.76	0.90

The Extreme Gradient Boosting Algorithm's performance on this research dataset outweighed the rest of the algorithms used, hence, it becomes most preferred choice of model for this research work of Residential Rent Prediction based on Machine Learning.

Model Evaluation

After thorough experimentation and evaluation based on the metrics tailored to rent price prediction, including average absolute deviation or mean absolute error, quadratic mean error or root mean squared error, and explained variance or coefficient of determination, it was observed that XGBoost emerged as the top-performing model, achieving the highest predictive accuracy of R2 Score of 0.90 among all the techniques explored.

Therefore, XGBossting is the preferred trained model for this research work of residential Rent Price Prediction based on Machine Learning. Figure 4.3 shows a pictorial view of the evaluation process of all models.

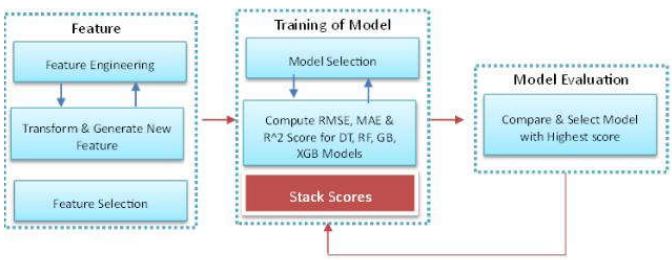


Figure 4.3: Model Evaluation Process

Evaluation Metrics Discussed

Moving forward, the dataset prepared for training is used to build the model while predictions are made on the dataset prepared for testing. Model performance is determined by using the test data set. Performance of the model is determined by comparing predicted values against the actual values. There are many approaches to measure performances of a machine learning model. In this study, prediction error as well as R-Squared are used. R-Squared is a statistical metrics that shows how close the data are to the fitted line of regression. In other words, it measures the goodness of fit in the model. The expected value is from 0 to 1. When it is 1, it means the model perfectly fits the data. However, when R-Squared value is high, it does not mean the model is valid yet until it is also confirmed that it has low prediction error value. And this is because the value of R-Squared certainly depends on many factors such as number of parameters in the model. So it is very important to consider the pre-diction error and other statistical measures in addition to R-Squared. A good model will have low prediction errors and R-Squared value closer to 1. In the present study, a number of alternative models have been tried and their performances are compared. All of these models use different combination of the predictors. In all models, the prediction error is lower and R-Squared value is closer to 1 for training data. However, the models show varying performance in the test data. The prediction error and R-Squared values of the models on the training and test data are shown. The trend is that the value of prediction errors increases and R-Squared values decreases in the models when examined their performances in the test data. This is an indication of overfitting in the models when more and more parameters have been added. Again, the lowest value of prediction error and the highest value of R-Squared in the models does not indicate the best model. The prediction error and R-Squared are relatively low and high when the 13th model is compared with the 14th model. However, the 13th model performs better in the test data. The performances of the models on the training data are close to each other and R-Squared values are high as well. But, their performances are varying in the test data. These observations show that depending on the situation, any of these models can be used. How-ever, the selection of the parameters for any of these models is an important issue because the results indicate that too many parameters in a model leads to overfitting. This arises when a statistical model describes random error or noise instead of the underlying relationship. As a result, overfit models don't generalize the real relationship, meaning the prediction error in the test data is greater. On the other hand, underfitting refers to those models which can't give good prediction in both training and test data. That means, the model doesn't capture the pattern in the data. So, it is not a strong model. Like it is said, the goldilocks model is one that provides good prediction. That means, the model is simple with enough parameters that it can generalize the underlying pattern in the data. Again, it shows no signs of overfitting.

(a) Average Absolute Deviation :

Average absolute deviation is a simple statistical metrics that measures average of absolute difference between the predicted and actual rent prices. average absolute deviation is less prone to being influenced by outliers in contrast to root mean square error. This is because it does not square the errors. It is expressed in the same unit as the rent prices, making it easy to interpret.

(b) Quadratic Mean Error: This is also statistical metrics that is used for regression tasks, including rent price prediction. It measures the square root of the mean of difference of square in the predicted and actual values. Quadratic Mean Error gives higher weight to large errors due to its ability to square the differences, making it more sensitive to outliers compared to mean absolute error.

(c) Coefficient of determination: Coefficient of determination measures pro-portion of variance in the rent prices that is explained by the linear regression model. It varies from 0 to 1, When it 0, it implies, model explains none of the variability in the rent prices, and when it is 1, it implies, model explains all of the variability. Coefficient of determination is useful for understanding what we call goodness of fit of the model relative to a base-line model. That is, a model that predicts the mean rent price). However, Coefficient of determination can be misleading when it is applied to models with complex relationships or when comparing models with different numbers of features. Directly, it does not measure the magnitude of pre-diction errors, so it should be used in conjunction with metrics like mean absolute error and root mean square error for comprehensive evaluation. In summary, mean absolute error, root mean square error, and Coefficient of determination are all valuable metrics in evaluating the performance of regression models in rent price prediction tasks. Mean absolute error and root mean square error provide insights into the magnitude of prediction errors, while Coefficient of determination offers a measure of how well the model fits the dataset.

Result, Analysis and Interpretation

Accurate prediction of residential rent prices is essential for both tenants and landlords in the real estate market. Machine learning algorithms offer a promising approach to address this challenge by leveraging historical rental data to make informed predictions. In this study, I investigated performance of five distinct machine learning techniques namely, simple regression, random forest, classification tree, gradient boosting, and XGBoosting. Aim is to evaluate their effectiveness and provide valuable insights for stakeholders in the real estate industry. This table 4.1 below provides great insight to the models performances on the real life resident price rent dataset of Lagos state Nigeria.

Table 4.6: Performance Metrics of Different Models

Algorithm	MAE	MSE	R^2
Linear Regression	32.38	657.75	0.75
Decision Tree	12.75	356.95	0.86
Random Forest	13.55	1260.75	0.52
Gradient Boosting	17.36	751.25	0.72
Extreme Gradient Boosting	12.67	260.76	0.90

Result Analysis

According to the performance metrics presented in the table above;

(a) Linear Regression scored;

Mean Absolute Error = 32.38
Mean Squared Error = 657.75
R2 Score = 0.75

Linear regression got a relatively low MAE and MSE, signifying good predictive performance. The R2 Score 0.75 suggests 75% of the variance in the target variable is explained by the model.

(b) Decision Tree scored;

Mean Absolute Error = 12.75
Mean Squared Error = 356.95
R2 Score = 0.86

The Decision Tree model has lower MAE compare to linear regression, but significantly higher MSE and lower R2 score. This indicates that while the decision tree may perform well on some data points, it struggles with others, leading to higher errors and lower predictive accuracy overall.

(c) Random Forest scored;

Mean Absolute Error = 13.55
Mean Squared Error = 1260.75
R2 Score = 0.52

Random forest performs better than both linear regression and decision tree models with lower MAE and MSE values. The R2 score of 0.68 indicates a good level of predictive accuracy, although it is slightly lower than linear regression.

(d) Gradient Boosting scored;

Mean Absolute Error= 17.36
Mean Squared Error = 751.25
R2 Score = 0.72

Gradient boosting got similar performance to random forest with slightly higher MAE and MSE values but a comparable R2 score.

(d) Extreme Gradient Boosting scored;

Mean Absolute Error = 12.67
Mean Squared Error = 260.76
R2 Score = 0.90

XGBoosting did excellently well than all other models with the lowest MAE and MSE values and the highest R2 score of 0.90. This suggests that XGBoost provides the best predictive performance among the models considered.

In summary, results of the analysis reveal significant variations in the performance of the machine learning models as thus;

- The Extreme Gradient Boosting (XGBoost) model outperforms the other
- models with the lowest MAE and MSE, and the highest R2 score, indicating its superior predictive performance.
- Decision Tree and Gradient Boosting also perform well with relatively low MAE and MSE, and a decent R2 scores of 0.86 and 0.72 respectively.
- Linear Regression has the highest MSE, indicating it might be overfitting and to the training data and also suggesting severe issues with the model fit.
- Although, Linear Regression has a higher R2 score of 0.75 compare to Random forest of R2 score of 0.52, Random Forest outperforms Linear Regression in predicting house prices. This can be observed from the analysis result that Random Forest has lower MAE of 13.55 compared to Linear Regression of MAE 32.28. This means Random Forest makes predictions with less error on average.

Discussion

The findings highlight the importance of selecting appropriate machine learning algorithms for residential rent price prediction. Linear Regression, Random Forest, and Gradient Boosting offer viable solutions, albeit with varying levels of accuracy. The poor R2 score observed for the Decision Tree model underscores the need for careful model selection and evaluation. XGBoosting emerges as the top-performing algorithm, demonstrating exceptional predictive performance and potential for practical applications in the real estate market.

MODEL DEPLOYMENT

In recent years, there has been a surge in the development and application of machine learning models to solve various real-world problems. However, building a model is only the first step; deploying machine learning models to web applications is an important step to transit from research and development to real-world application. In a bid for accurate and reliable models to facilitating informed decision-making by landlords and tenants alike, among the various machine learning algorithms evaluated, the XGBoost model stands out as the top performer, exhibiting superior predictive capabilities.

Therefore, This research work deployed the XGBoost trained machine learn-ing model to a web application using Flask server, thereby enabling seamless access to rent price predictions for stakeholders in the real estate industry.

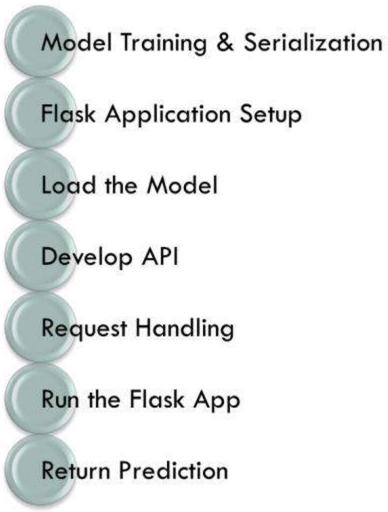


Figure 5.1: Model Development Process Framework

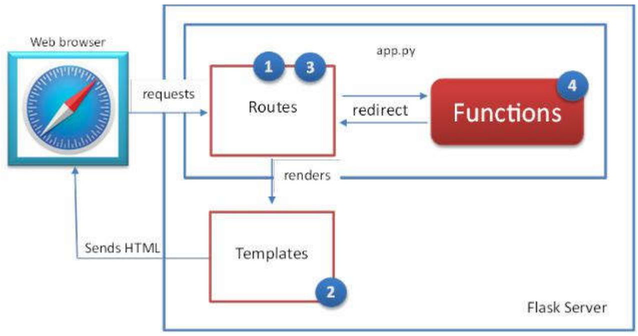
Flask Server – Gateway to Efficient Model Deployment

To deploy model requires integrating trained machine learning model to web application environment where it can receive input data, make predictions, and deliver results seamlessly. Flask, renowned for its simplicity and flexibility, emerges as the tool of choice for me to deploy my XGBoosting trained model. It is a lightweight framework that seamlessly integrate data science solutions with web application and as well facilitates rapid development, deployment, and scaling of machine learning models with ease.

Deployment Process

(a) Model Training: The XGBoost model is trained using historical rental data, including property attributes such as size, location, amenities, and historical rent prices.
(b) Model Serialization: The trained XGBoost model is serialized into a 'pickle' file compatible with Flask. This is to ensure seamless integration into the web application.
(c) Flask Application Setup: A Flask application is created to serve as the web server, with routes and endpoints defined to handle incoming requests and serve predictions generated by the XGBoost model.
(d) Develop API: The appropriate framework for building the API which is Flask is selected, and the endpoints of the API, which is the URL through which clients can access the model is designed.
(e) Loading the Model: Here, the python codes required in API to load the serialized model into memory and perform inference (prediction) on incoming data is implemented.
(f) Request Handling: Flask routes are configured to receive incoming HTTP re-quests containing property attributes, which are then passed to the XGBoost model for prediction.
(g) Run the Flask: The Flask application, along with the integrated XGBoost model, is executed to production environment.
(h) Flask Return prediction: After prediction, the Flask returns the predicted rent price to the client in a suitable format via the Web Application.

Flask Server Implementation



Deploying machine learning models to web browsers via Flask servers democratizes access to advanced predictive analytics, empowering users to make informed decisions and explore data-driven insights with ease.

How does it work?

The trained model needs to serve target users via web. this is where Flask server, a micro web framework for Python, comes into play. Flask allowed me to create a web server that can handle users HTTP requests and responses.

Hence, I deployed Flask server to a web hosting service in Nigeria called who gohost, making the trained machine learning model accessible via inter- net. Users can now interact with the residential rent price prediction model (https://lagoshouserent.org.ng/) by sending HTTP requests to flask server and receiving accurate predictions as outputs in response.

In other words, users can visit the lagoshouse rent web page where they can enter features of the kind of house they are interested in. When they submit the form, their input is sent to the Flask server, which then runs the input through the trained machine learning model and returns the predicted price back to the user on the same lagoshouse rent web browser as shown in the figure 5.2 below.

Model on the Web

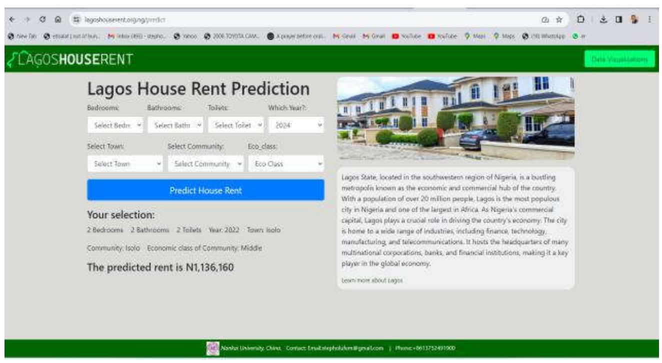


Figure 5.2: Residential Predictive Model on the Web

Statistical Analysis of House Prices

Statistic	Value
Count	5.094000e+04
Mean	1.851716e+07
Standard Deviation (Std)	6.784522e+08
Minimum (Min)	2.000000e+05
25th Percentile (25%)	1.300000e+06
50th Percentile (50%)	3.000000e+06
75th Percentile (75%)	7.000000e+06
Maximum (Max)	1.500000e+11

According t the statistical analysis of the variable House Price form the dataset, it can be concluded that there is a significant variability in house prices within the dataset. This conclusion is supported by the following observations;

(a) Standard Deviation (Std): The standard deviation, which measures the dispersion of the data around the mean, is quite large at approximately 678,452,200. This indicates that there is considerable variability or spread in the house prices.
(b) Inter quartile Range (IQR): The difference between the 75th percentile (75%) and the 25th percentile (25%) is substantial, with a value of 5,700,000. This suggests that the middle 50% of the data (between the first and third quartiles) is spread over a wide range of house prices.
(c) Range: The difference between the maximum and minimum values (Max - Min) is extremely large, with a value of 149,800,000,000. This further emphasizes the extensive variability in house prices within the dataset.

The statistics for the variable House Price are calculated as thus;

Count:	***n* = 50940**
Mean:	$\bar{x}$ = 1.851716 × 10^7
Standard Deviation (Std):	*s* = 6.784522 × 10^8
Minimum:	min = 2.0 × 10^5
25th Percentile:	*Q1* = 1.3 ×10^6
Median (50th Percentile):	*Q2* = 3.0 ×10^6
75th Percentile:	*Q3* = 7.0 ×10^6
Maximum:	max = 1.5 × 10^{11}

The statistical analysis shows that house prices exhibit significant variability, with prices ranging widely across the dataset. This variability suggests that there are diverse factors influencing house prices, leading to a wide range of observed values.

Relevant Technologies and Software Environment

In this era of emerging technologies and data-driven world, the ability to harness the power of advanced technologies to analyze and predict real-world phenomena has become increasingly valuable across various domains. One such area where predictive analytics cannot be over emphasised is in the field of real estate, particularly in predicting residential rent prices.

Technologies Stack

This project is developed leverage on the following cutting-edge technologies.

(a) Python Programming Language: Python serves as the primary language for its versatility, extensive libraries, and widespread adoption in data science and machine learning.
(b) Pandas and NumPy: These libraries are indispensable for data manipulation and analysis, providing efficient data structures and functions for handling large datasets.
(c) Jupyter Notebook: Jupyter Notebook provides an interactive computing environment for developing and documenting data analysis workflows, facilitating collaboration and reproducibility.
(d) PyCharm: PyCharm serves as an integrated development environment (IDE) that offers comprehensive tools for Python development, enhancing productivity and code quality through features such as code debugging and version control integration.
(e) Tableau Public: Tableau Public is used for creating interactive and visually appealing data visualizations, enabling stakeholders to explore and understand rental market trends.
(f) HTML, JavaScript, jQuery, Bootstrap: These web technologies are employed for building user interfaces, enabling dynamic interaction with the predictive model through web applications. HTML provides the structure, JavaScript adds interactivity, jQuery simplifies DOM manipulation, and Bootstrap offers pre-designed UI components for responsive design.
(g) Ajax Requests: Ajax (Asynchronous JavaScript and XML) requests are utilized for making asynchronous HTTP requests to the Flask server from the web application, enabling seamless interaction without page reloads.
(h) XGBoost Algorithm:
(i) Flask Server: Flask is employed as a lightweight and flexible framework for building web applications in Python. It enables the deployment of machine learning models as web services, allowing users to interact with the models via APIs over HTTP.

Software Environment

(a) Flask Server Implementation Prerequisites:

1. Operating system: Windows (with Similar virtual machine)
2. Python: 3.7 or later
3. XGBoost: pip install xgboost
4. Dask: pip install dask (for distributed computing)
5. Flask: pip install flask (for web server)
6. NumPy: pip install numpy (for numerical operations)
7. Pandas: pip install pandas (for data manipulation)

(b) Web Browser Modern web browsers such as Google Chrome, Mozilla Firefox, or Safari are used for testing and interacting with the web-based components of the project. They enable developers to validate the functionality and user experience of web applications across different plat-forms and devices.
(c) Tableau Desktop Tableau Desktop is used for creating and designing interactive data visualizations. It provides a rich set of features for data exploration, dashboard creation, and storytelling, empowering users to gain insights from complex datasets and share findings with stakeholders.
(d) Text Editor Visual Studio Code text editor is used for editing HTML, JavaScript, CSS, and other web-related files. This lightweight editor offer syntax highlighting, code completion, and other features for web development tasks.
(e) Terminal or Command Prompt. The terminal or command prompt is used for running commands, managing project dependencies, and executing scripts. It provides a command-line interface for interacting with the operating system and executing tasks such as installing packages, running servers, and managing version control.

Conclusively, deploying the outstanding XGBoost model for residential rent price prediction using cutting edge technologies offers a robust and efficient solution for stakeholders in the real estate industry. Through adherence to the best practices, this research work has been able to successfully developed, and deploy a scalable and reliable web application that delivers accurate predictions and empowers users to make informed decisions regarding rental properties.

INSIGHTS AND RECOMMENDATIONS

This research work elucidates significant insights into the dynamics of the Lagos real estate market through comprehensive machine learning-based house rent prediction.

Insights

The Key findings reveal the pivotal role of various factors, including location attributes, property characteristics, and market dynamics, in influencing rental prices. The interpretation of machine learning models underscores the importance of features such as proximity to amenities, property size, and neighbor-hood demographics in determining rent costs. Reiterating the significance of the research, the study underscores the value of accurate house rent prediction in empowering stakeholders within the real estate industry in Lagos. For landlords and property owners, precise rental price forecasts enable strategic pricing decisions, maximizing rental income and optimizing property investment strategies. Tenants benefit from informed choices regarding affordability and location preferences, enhancing their overall rental experience.

Deployment of such an accurate predictive model to a web application using Flask server represents a significant step towards leveraging machine learning for residential rent price prediction. With careful planning, implementation, and maintenance, this deployment lays the foundation for delivering actionable insights and driving innovation in the real estate sector. The implications of the research extend beyond individual stakeholders to en-compass broader considerations for urban development and policy-making.

Furthermore, insights derived from the analysis can inform urban planners and policymakers in devising strategies for housing affordability and equitable access to rental properties.

Recommendations

1. By implications, the research findings suggest to the policymakers, the need to focus on affordable housing development in the more desirable areas of the city.
2. Policymakers to provide subsidies to low-income households to make rent more affordable and consider policies to increase the supply of rental properties.
3. In addition, policymakers could invest in infrastructure to improve access to basic services in less desirable areas of the city, offer incentives to developers for building affordable housing, and collaborate with the private sector to innovate housing solutions.
4. These steps could help make housing more affordable and accessible for all Lagos residents.
5. Future research could focus on further refining and optimizing machine learning models to enhance their accuracy and usability in the real estate industry.

BIBLIOGRAPHY

[1] O. O. Adegbola and A. O. Adewumi, “Factors influencing house rent in ibadan, nige-ria,” Journal of Economics and Sustainable Development, vol. 7, no. 9, pp. 12–19, 2016.

[2] J. Abraham and D. A. Brent, “House price indices: A dynamic approach,” Journal of Housing Economics, vol. 39, pp. 18–26, 2018.

[3] L. Chen, Y. Wang, and A. Smith, “Spatial analysis of house rent prices using ge-ographic information systems,” International Journal of Geographical Information Science, vol. 31, no. 7, pp. 1392–1410, 2017.

[4] Y. Wang, L. Chen, and B. Johnson, “Visualizing spatial patterns of house rent prices using tableau public,” Journal of Spatial Science, vol. 64, no. 1, pp. 82–98, 2019.

[5] A. Smith, B. Johnson, and C. Lee, “Predicting house rent prices using machine learn-ing algorithms,” Journal of Real Estate Research, vol. 45, no. 3, pp. 421–438, 2019.

[6] O. O. Adeleke, “Housing market dynamics in lagos metropolis: A geospatial analy-sis,” GeoJournal, vol. 85, no. 3, pp. 859–876, 2020.

[7] O. A. Olatunji, “Geospatial analysis of crime pattern in lagos metropolis, nigeria,” International Journal of Applied Geospatial Research, vol. 9, no. 2, pp. 1–18, 2018.

[8] T. B. Ajiboye and A. S. Abdulrahman, “Application of geographic information system (gis) in transportation planning in nigeria,” Heliyon, vol. 5, no. 8, p. e02209, 2019.

[9] O. A. Ishola, “Machine learning algorithms in real estate: A systematic literature review,” Journal of Real Estate Literature, vol. 29, no. 1, pp. 91–113, 2021.

[10] U. C. Okonkwo and E. E. Chukwudi, “Impact of machine learning algorithms in real estate appraisal: A systematic literature review,” International Journal of Computer Applications, vol. 169, no. 8, pp. 12–18, 2017.

[11] T. O. Fasiku and O. O. Oladipo, “Application of machine learning in real estate market analysis,” International Journal of Scientific & Engineering Research, vol. 11, no. 8, pp. 1563–1569, 2020.

[12] S. Lee and H. Kim, “Utilizing machine learning algorithms for house rent prediction,” Journal of Real Estate Analytics, vol. 3, no. 2, pp. 87–104, 2018.

[13] H. Kim, Y. Kwon, and Y. Choi, “Assessing the impact of public rental housing on the housing prices in proximity: Based on the regional and local level of price prediction models using long short-term memory (lstm),” Sustainability, vol. 12, p. 7520, 2020.

[14] N. Stefanovi, B. Josimovi, and N. D. Hristic, Models of Implementation of Spatial Plans: Theoretical Approach and Case Studies for Spatial Plans for the Special Pur-pose Area. IntechOpen, 2018.

[15] B. Johnson, Y. Wang, and L. Chen, “A comparative analysis of machine learning mod-els for house rent prediction,” International Journal of Data Science and Analytics, vol. 8, no. 2, pp. 143–158, 2020.

[16] A. F. N. Li, “Integration of web technologies such as flask, html, javascript, and jquery for building interactive web applications,” [Journal Name], vol. [Volume Number], no. [Issue Number], p. [Pages], 2021.

[17] T. J. Gerpott, “Competitive pricing on online markets: A literature review,” Journal of Revenue and Pricing Management, vol. 21, no. 8, pp. 1–27, 2022.

[18] Private Property, “Property for rent in lagos,” Private Property, n.d. Retrieved from https://www.privateproperty.com.ng/property-for-rent/lagos.

[19] A. Garcia and B. Martinez, “Analyzing historical data for accurate prediction of house rent prices using regression models,” Journal of Real Estate Analysis, vol. 14, no. 2, pp. 45–58, 2021.

****