
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Real Estate Web Scraping With EDA(Exploratory Data Analysis)

[1]GopiReddy Prakash Reddy, [2]Jullapally Shashank, [3]J. Koti Sasidhar, [4]G. Rakesh Goud, [5]Mr. GONUGUNTA RAJ KUMAR,
[1,2,3,4] U.G.Scholor, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.
[5]Research Guide, Department of ECE, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

ABSTRACT

Web scraping with exploratory data analysis (EDA) in the real estate domain involves extracting data from various real estate websites and performing an analysis to gain insights into the market trends, property prices, and other relevant information By combining web scraping with EDA in the real estate domain, you can gain a deeper understanding of market dynamics, make informed decisions. This project presents a comprehensive study on "Real Estate Web Scraping with Exploratory Data Analysis (EDA)," aiming to create a robust and data-driven solution for the real estate industry. By conducting a thorough literature survey, we evaluate existing web scraping methods and data analysis techniques, offering insights into their limitations and potential improvements. The project introduces an innovative system that not only overcomes the disadvantages of current solutions but alsobringsseveraladvantagestothereal estatedomain.Theintegrationofweb scraping and EDA promises to provide real estate professionals and enthusiasts with valuable insights and comprehensive datafor more informed decision-making.

1. INTRODUCTION

The real estate industry continuously relies on data to make informed decisions about property investments, pricing, and market trends. To meet these data demands, web scraping and exploratory data analysis (EDA) have become critical tools. This project explores the fusion of web scraping andEDA,presentinganinnovativeapproach to provide comprehensive insights into the real estate market. As part of our investigation,wedelveintoexisting systems to uncover their limitations, and we introduce a system that offers significant advantages. By integrating web scraping techniques to gather up-to-date real estate dataandapplyingEDAtoextract

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

meaningful insights, our proposed system promises to revolutionize the way real estate data is harnessed.

APPLICATIONS

Web scraping can help real estatebusinesses to find and compare properties based on various criteria, such as location, price, size, type, and amenities. This can help them to provide better services to their customers, and to increase their sales and revenue. Web scraping can help real estate businesses to understand the trends and patterns of the real estate market, such asthe demand, supply, price, and profitability of the properties. This can help them to optimize their decision making and their investment strategies, and to identify the opportunities and risks of the market.

2.LITERATURESURVEY

1: Title: "Web Scraping Techniques: A Comprehensive Review" Authors: John M. Smith, Sarah L. Davis

2: Title: "Exploratory Data Analysis in Real Estate: An Overview" Authors: Michael J. White, Emily S. Parker

3: Title: "Challenges in Real Estate Data Collection:AComparativeAnalysis" Authors: David W. Anderson, Mary E. Green

4: Title: "Data-Driven Decision-Making in Real Estate: An In-Depth Review"Authors: Laura M. Johnson, Mark A. Turner

5: Title: "Innovations in Real Estate Data Analysis: Challenges and Opportunities" Authors: Richard C. Lewis, Jennifer R. Hall

3. SYSTEMDESIGN

SYSTEMARCHITECTURE

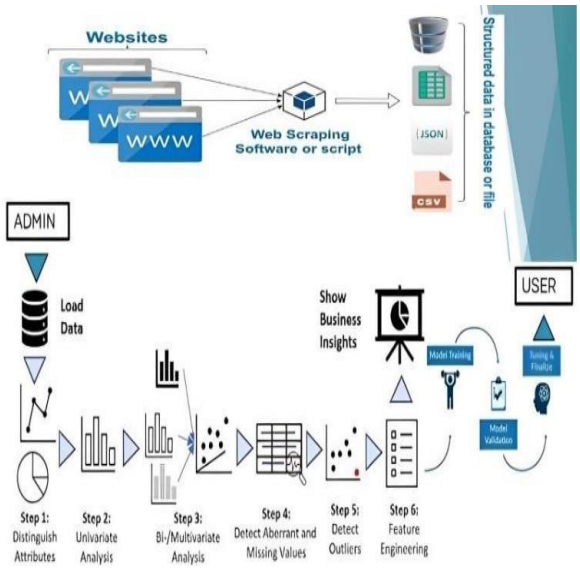


Fig-1SystemArchitecture

3.3 UNIFIED MODELLINGLNGUAGE(UML) DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

object-orientedsoftwareengineering.The standard is managed, and was created by, the Object Management Group.The goal is forUMLtobecomeacommonlanguagefor creatingmodelsofobject orientedcomputer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may alsobe added to; or associated with, UML.The Unified Modeling Language is a standard languageforspecifying,Visualization, Constructinganddocumentingthe artifacts of software system, as well as for business modeling and other non-software systems.

I) USECASEDIAGRAM

A use case diagram in the UnifiedModeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and anydependenciesbetweenthose use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors inthe system can be depicted

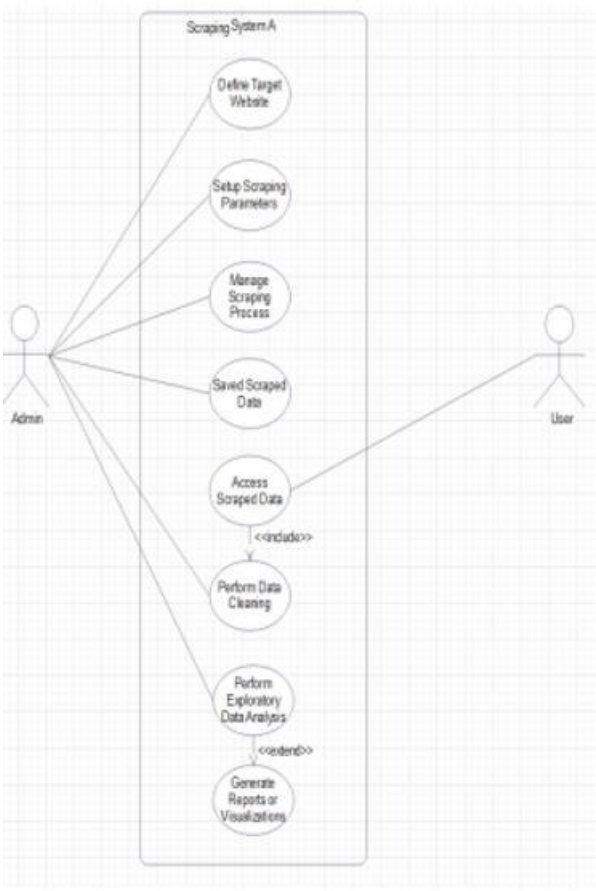


Fig-2Usecase Diagram

II) CLASSDIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

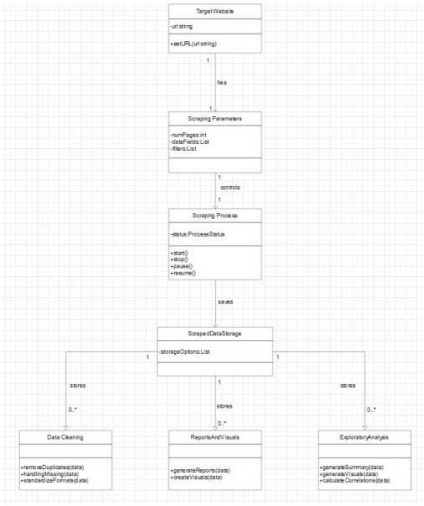


Fig-3ClassDiagram

iii) SEQUENCEDIAGRAM

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

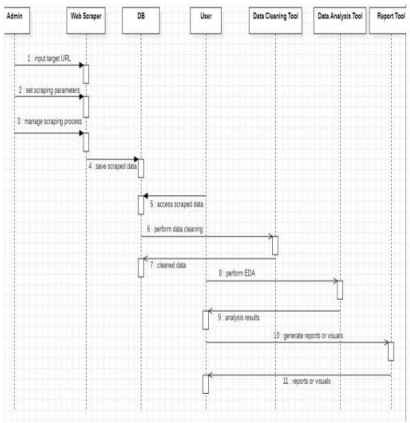


Fig-4Sequence Diagram

iv) ACTIVITYDIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

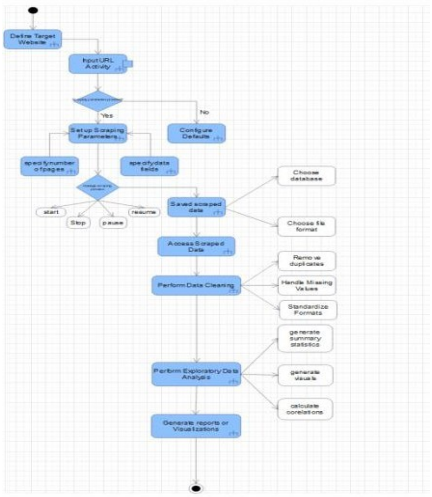


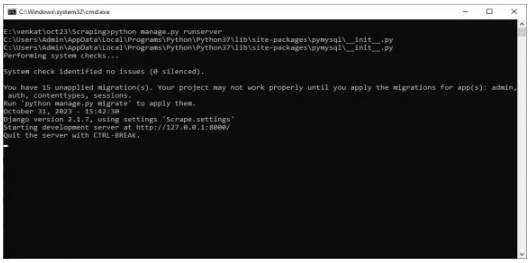
Fig-5Activity Diagram

4. OUTPUTSCREENS

Screen1:

To run project copy content from DB.txtfile and then paste in MYSQL console to create database and then double click on ‘run.bat’ file to start python server and get below page

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025



In above screen python server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and press enter key to get below page

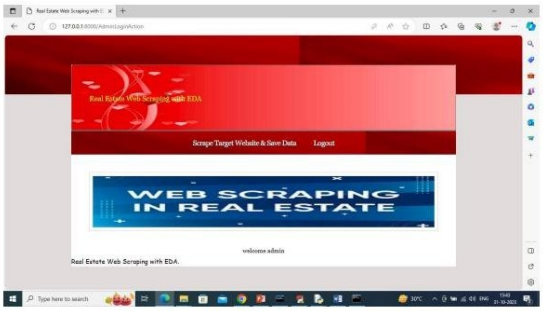
Screen2:

In above screen click on 'Admin Login'link to get below page

Screen3:

In above screen admin is login and after login will get below page

Screen4:



In above screen click on 'Scrape Target Website & Save Data' link to scrape the website and to save the scraped data after that we will get below page

Screen5:

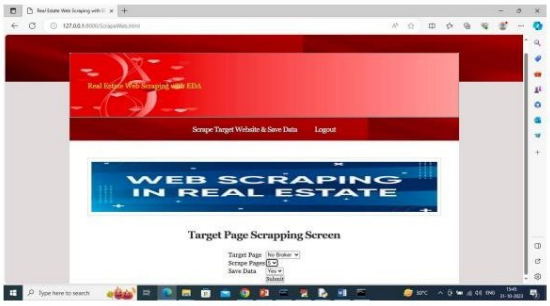
In above screen I selected website as Magic Brick and then selected number of pages with save option and then press button toget below output

Screen6:

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

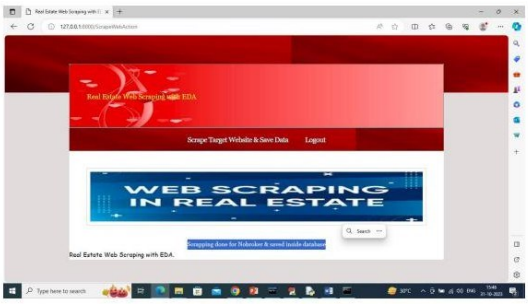
In above screen in blue colour text we can see scraped data issaved in database and similarly we can scrape for other website

Screen7:



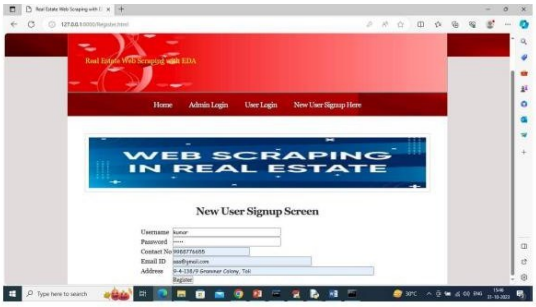
In above screen I selected website as NoBroker and then selected number of pages with save option and then pressbutton to get below output

Screen8:



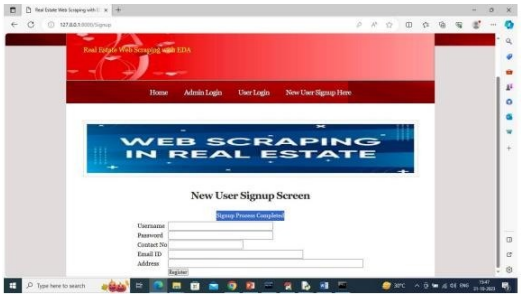
In above screen No Broker data is also saved in database and now logout and signup as new user so that user can access the scraped data.

screen9:



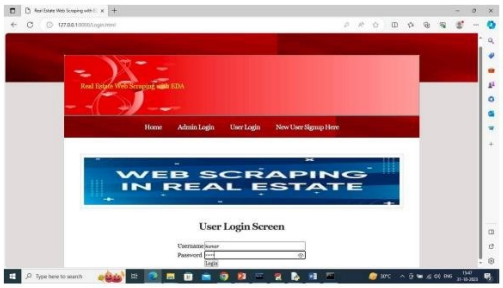
In above screen user is registering and after signup will get below page

Screen10:



In above screen signup task completed and now click on 'User Login' link to get below page

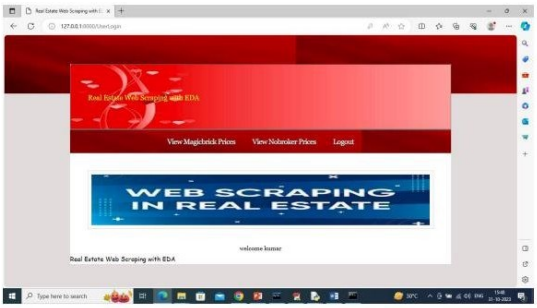
Screen11:



In above screen user login is successful and after login will get below page

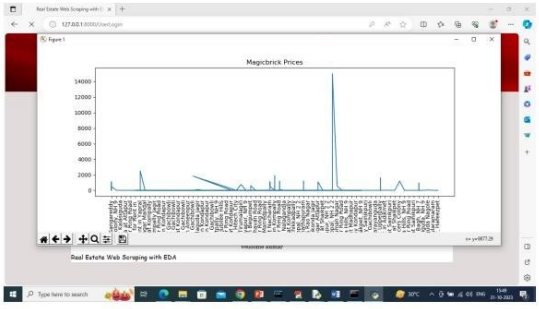
ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Screen12:



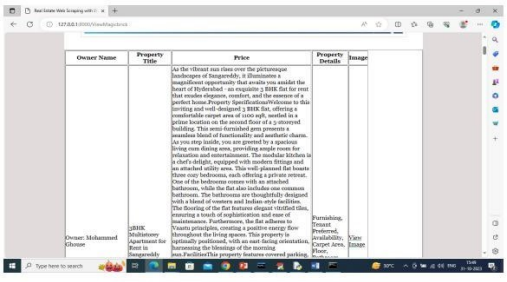
Inabovescreenclickon‘ViewMagicBrick Prices’ link to view all scrape data and get below output

Screen13:



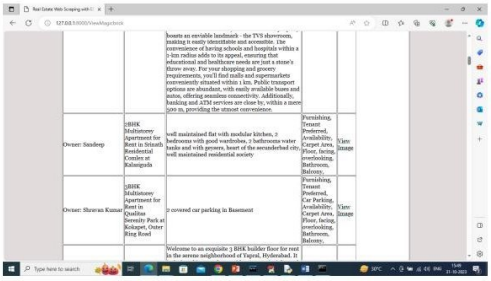
In above graph x-axis represents property names and y-axis represents prices and now close above graph to view Magic Brick property details

Screen14:



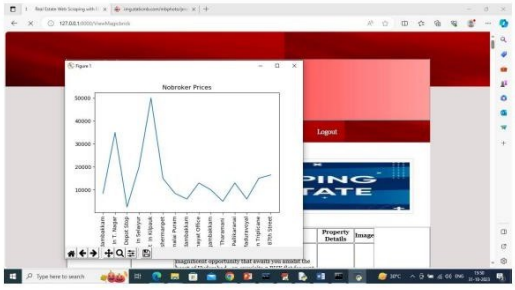
This shows the details along with the property title,owner name,property details with which we can easily view the property prices without visiting number of websites.

Screen15:



Inabovescreenwecanreadpropertydetails from Magic Brick and now click on ‘View No Broker Prices’ link to get below page

Screen16:



In above graph x-axis represents No broker property names and y-axis represents prices and now close above graph to get below page

Screen17:

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Owner Name	Property Title	Price	Property Details	Image
K.V.MOHANKUMAR	1 BHK House for Rent In Adambakkam	8500	2023-09-13, Adambakkam, chennai, 1 BHK House for Rent In Adambakkam, Adambakkam, chennai,	View Image
shankar	1 BHK House for Rent In Kundrathur Bus Depot Stop	7500	2023-10-31, Kundrathur Bus Depot Stop, chennai, 1 BHK House for Rent In Kundrathur Bus Depot Stop, Kundrathur Bus Depot Stop, chennai,	View Image
jaykumar	3 BHK Flat In The Peninsula for Rent In Kilpauk	50000	2023-10-26, Kilpauk, chennai, 3 BHK Flat In The Peninsula for Rent In Kilpauk, Kilpauk, chennai,	View Image
Sooriaraj	1 BHK Flat In Standalone Building for Rent In Raja Annamalai Puram	8500	2023-10-31, Raja Annamalai Puram, chennai, 1 BHK Flat In Standalone Building for Rent In Raja Annamalai Puram, Raja Annamalai Puram, chennai,	View Image
vaishnavi	2 BHK Flat In Devi Flats for Rent In Alapakkam Panchayat Office	13000	2023-09-24, Alapakkam Panchayat Office, chennai, 2 BHK Flat In Devi Flats for Rent In Alapakkam Panchayat Office, Alapakkam Panchayat Office, chennai,	View Image
Uday Kumar	1 RK House for Rent In Tharamani	5000	2023-10-11, Tharamani , chennai, 1 RK House for Rent In Tharamani, Tharamani , chennai,	View Image
vasanthkumar N	1 BHK House for Rent In Maduravoyal	9000	2023-10-17, Maduravoyal, chennai, 1 BHK House for Rent In Maduravoyal, Maduravoyal, chennai,	View Image
suresh babu	2 BHK House for Rent In 87th Street	16500	2023-10-23, 87th Street, chennai, 2 BHK House for Rent In 87th Street, 87th Street, chennai,	View Image
Vimala	2 BHK Flat In Ashirvath Apartment for Rent In Madipakkam	10000	2023-11-14, Madipakkam, chennai, 2 BHK Flat In Ashirvath Apartment for Rent In Madipakkam, Madipakkam, chennai,	View Image
Kulandaivel	2 BHK Flat In Apartment for Rent In Polavakkam	15000	2023-10-27, Polavakkam, chennai, 2 BHK Flat In Apartment for Rent In Polavakkam, Polavakkam, chennai,	View Image

In above screen can see scrape data from NO Broker website. Similarly admin will scrape latest property details and then user will view those details

5. CONCLUSION

The"Real EstateWebScraping with EDA" projectrepresentsasignificantadvancement in harnessing real estate data for more informed decision-making. By addressing the limitations of existing systems, our proposed solution offers a holistic approach to data collection, data cleaning, and exploratory data analysis. This innovative system leverages the strengths of web scraping while mitigating the challenges related to data quality, website structure changes, and legal compliance. Moreover, by integrating EDA, our project opens new avenues for comprehensive insights intoreal estate market trends, risk assessment, and property valuation. As the real estate industry increasingly relies on data-driven decision-making, our project promises to be atransformativetool,enhancingtheway real estate data is gathered, analyzed, and utilized for more strategic and informed choices in the dynamic real estate market

6. FUTUREENHANCEMENT

Using natural language processing and sentiment analysis to extract and analyzethe reviews and ratings of the properties, agents, and developers from varioussources, such as social media, blogs,forums, and websites. This can help to understand the customer feedback, preferences,andsatisfaction,andtoidentify the strengths and weaknesses of the real estatemarket.WebScrapingforReal Estate Data: Complete Guide.Using machine learning and predictive analytics to forecast the trends and patterns of the real estate market, such as the demand, supply, price, and profitability of the properties, based on thehistoricalandcurrentdata.Thiscanhelp to optimize the decision making and the investment strategies of the real estate stakeholders, and to identify the opportunities and risks of the market. 3 Predictions about Future of Web Scraping – ProWebScraper.Using geospatial analysis and visualization to map and compare the properties, locations, and neighborhoods based on various criteria, such as the amenities,facilities,accessibility,safety,

and livability. This can help to provide a comprehensive and interactive view of the real estate market, and to enhance the user experience and engagement of the web scraping platform. Real Estate Market,Web-Scrapping and EDA using Python

7. REFERENCES

[1] J. Ratcliffe, M. Stubbs;, and M. Keeping, Urban Planning and Real Estate Development. Routledge, 2021.

[2] J. S. Chou, D. B. Fleshman, and D. N. Truong, Comparison of machine learning models to provide preliminary forecasts of real estate prices, no. 0123456789. Springer Netherlands, 2022. [3] A. S. Ravikumar, Real Estate Price Prediction Using Machine Learning. 2017.

[4] M.F.Lee,G.S.Chen,S.P.Lin,andW. J. Wang, "A Data Mining Study on House Price in Central Regions of Taiwan Using Education Categorical Data, Environmental Indicators, and House Features Data," Sustain., vol. 14, no. 11, 2022, doi: 10.3390/su14116433.

[5] C. H. Cheng and M. C. Tsai, "An Intelligent Homogeneous Model Based on an Enhanced Weighted Kernel Self-OrganizingMapforForecastingHouse Prices," Land, vol. 11, no. 8, 2022, doi: 10.3390/land11081138.

[6] Y. Ahn and C. K. Uejio, "Modeling Air Conditioning Ownership and Availability," SSRN Electron. J., vol. 46, no. October, p. 101322, 2022, doi: 10.2139/ssrn.4211073.

[7] M.Cekic, K. N.Korkmaz,H. Mukus,A. A. Hameed, A. Jamil, and F. Soleimani, "Artificial Intelligence Approach for Modeling House Price Prediction," pp. 1–5, 2022, doi: 10.1109/icmi55296.2022.9873784.

[8] D. Kulikauskas, "The user cost of housing in the Baltic states," J. Eur. Real Estate Res., vol. 10, no. 1, pp. 17–34, 2017, doi: 10.1108/JERER-11-2015-0042.

[9] G.J.Rangel,J.W.J.Ng,T.T. Murugasu, and W. C. Poon, "Measuring Malaysian housing affordability: thelifetime income approach," Int. J. Hous. Mark.Anal.,vol.12,no.5,pp.966–984, 2019,doi:10.1108/IJHMA-02-2019-0023.

[10] L. Hu, S. He, and S. Su, "A novel approach to examining urban housing market segmentation: Comparing the dynamics between sales submarkets and rental submarkets," Comput. Environ.UrbanSyst.,vol.94,no.January,p.

101775, 2022, doi: 10.1016/j.compenvurbsys.2022.101775.

[11] F.XueandE.Yao,“Adoptinga random forest approach to model household residential relocation behavior,” Cities, vol. 125,no.May2021,p.103625,2022,doi: 10.1016/j.cities.2022.103625.