REAL ESTATE LOCATION PREDICTION BASED ON DATA ANALYTICS

PROF. SUPRIYA SHRIVASTAV[1], SWATI PRIYA[2], TEJASWINI C[3], SAHIL PRIYADARSHI[4], ROHIT BHARADWAJ[5]

[1] *Assistant Professor, CSE, AMCEC, Karnataka, India*

[2] *Student, CSE, AMCEC, Karnataka, India*

[3] *Student, CSE, AMCEC, Karnataka, India*

[4] *Student, CSE, AMCEC, Karnataka, India*

[5] *Student, CSE, AMCEC, Karnataka, India*

ABSTRACT

Machine Learning (ML) excels at most predictive tasks but its complex nonparametric structure renders it less useful for inference and out of sample predictions. This article aims to elucidate and enhance the analytical capabilities of ML in real estate through Interpretable ML (IML). Specifically, we compare a hedonic ML approach to a set of model-agnostic interpretation methods. Our results suggest that IML methods permit a peek into the black box of algorithmic decision making by showing the web of associative relationships between variables in greater resolution. In our empirical applications, we confirm that size and age are the most important rent drivers. Further analysis reveals that certain bundles of hedonic characteristics, such as large apartments in historic buildings with balconies located in affluent neighbourhoods, attract higher rents than adding up the contributions of each hedonic characteristic. Building age is shown to exhibit a U-shaped pattern in that both the youngest and oldest buildings attract the highest rents. Besides revealing valuable distance decay functions for spatial variables, IML methods are also able to visualise how the strength and interactions of hedonic characteristics change over time, which investors could use to determine the types of assets that perform best at any given stage of the real estate investment cycle.

1. INTRODUCTION

When a new technology matures, it is typically adopted in business operations by firms in order to differentiate themselves from their competitors. There is a growing trend among private companies to advertise and sell their products and services via internet-based technologies. To compete, most large real estate-related firms have created and maintained their own websites to provide value-added and comprehensive services that assist customers in completing property transactions at a lower cost. Real estate marketing and brokerage, real estate appraisal, auctions, tenders, and mortgage brokerage are all available online. The goal of [1] is to propose the use of transaction data, hedonic models, and internet-based technologies by real estate related firms to provide potential home buyers and sellers with instant and online property appraisal services. The hedonic price model is used to calculate the price index of many housing estates individually. Because the values of the attribute coefficients may change in response to changes in the environment, it may be required to make a professional decision on time intervals to re-run the hedonic price models, if necessary. Authenticated users can connect to the system via SSL after the algorithms have been updated. They can browse and search the valuation reports by entering search criteria for properties into the system, which instantly filters the results based on the users' requests, and a detailed asset valuation report will be displayed in the browser.

A new wave of technology innovation, namely artificial intelligence (AI), has nowadays been being put to practical use in various business fields, especially in recent years, due to improvements in hardware performance and increases in the collection and use of big data [2]. Machine learning (a subset of AI) is a very powerful tool to collect, analyse, and interpret big data for predicting outcomes. It has been extensively deployed in many industries, including the real estate market. Using machine learning in the real estate market can help improve decision–making, reduce risk, and increase efficiency in property valuation, management, and investment. First, machine learning algorithms can analyse historical sales data and other relevant factors such as demographics, location, size, and amenities to accurately predict the value of a property [3]. They also automatically categorize properties, ranking search results and suggesting comparable properties. Machine learning can make real estate transactions simpler. This can also aid in the decision-making process for both buyers and sellers. Second, machine learning algorithms can locate properties that are anticipated to appreciate in value or yield a high rental income by using historical data and recent market patterns. They can be used to analyse market trends, property data, and economic indicators to assess the risk associated with investing in a particular property or market. Third, analysis of data on occupancy rates, rental rates, and tenant behaviour can be used to optimize property management operations, such as lease renewal, rent collection, and maintenance scheduling. Fourth, machine learning algorithms can analyse data to detect potential fraud, such as mortgage fraud. Fifth, it can analyse energy consumption data from buildings and identify patterns to optimize energy use and reduce costs.

2. PROBLEM STATEMENTS

1. Manual data analysis: Time-consuming, labour-intensive, and prone to biases and errors.
2. Subjectivity and bias: Reliance on subjective opinions, experiences, and preferences, leading to suboptimal decision-making.
3. Limited insights: Limited access to comprehensive and up-to-date data, difficulty in processing complex patterns and correlations.
4. Scalability and efficiency: Inefficiency in analysing data for multiple locations, difficulty in keeping up with changing market dynamics.
5. Lack of predictive power: Reliance on intuition or past experiences, which may not accurately forecast future trends and outcomes.

3. OBJECTIVE

Accuracy: The model aims to achieve high accuracy in predicting the most suitable location for a business by leveraging comprehensive and up-to-date real estate data, minimizing biases, and reducing human errors in data analysis.

Predictive Power: The model aims to leverage historical data and machine learning algorithms to provide accurate and reliable predictions of future trends and outcomes in real estate markets, aiding in making informed decisions about the best business location.

Scalability and Efficiency: The model aims to efficiently process and analyse large datasets for multiple locations, providing scalable and timely insights that can adapt to changing market dynamics and support efficient decision-making processes.

4. LITERATURE SURVEY

This paper is a survey of the literature on Real Estate Investment Trusts, common as REITs. The literature is separated into three major research topics: investment financing decisions, and return and risk issues. The central papers addressing each optics are described and their results are summarized. Suggestions for further also are provided. Since their beginnings in 1960, shares of Real Estate Investment Trusts, commonly as REITs, have been publicly traded. The general availability of daily returns data f which hold only real estate-related assets, have given researchers unique opportune study issues in real estate financial economics. For example, access to returns d allowed researchers to 1) perform event studies that show how the value of RE react to significant public announcements, 2) study the return and risk relationship retired real estate and draw comparisons to the performance of securitized rea 3) study the performance of REITs relative to other asset classes, 4) study the eve inclusion of securitized and securitized real estate in investment portfolios, and Ine agency problems in real estate and REITs. Additionally, the unique institute tax characteristics of REITs provide opportunities to isolate and study issues centre porate finance, such as dividend policy and capital structure. The dimensions of the REIT literature have expanded greatly in recent years. T of this paper is to organize this literature and summarize the main findings. T of the task is narrowed by concentrating on papers written since 1980 that relate to financial economics that introduce REITs as special cases. An exception is ma pre- 1980 paper provides a foundation for the current literature. We review both and unpublished research and discuss those papers that are the most relevant and dative of the major topical areas of research identified for this review. The REIT literature is divided into three major topical areas: investment decisions, in decisions, and return and risk issues affecting REITs. To better relate the one another, we further divide the literature into subtopics. The remaining section paper follow the organizational structure.

4.1. Institutional background and a brief history of REITs1

REITs are closed-end investment companies with publicly traded stock. They serve as financial intermediaries to facilitate the flow of funds from investors to the real estate sector of the economy. Some of the many investors who demand real estate use REITs as a way to invest in the real estate class for return and portfolio objectives while retaining the investment liquidity provided by the secondary market for REIT shares. The REIT organizational form was authorized in 1960 when federal legislation provided tax exemptions for "qualified REITs" that satisfy certain requirements. To be qualified, REITs must meet the following criteria: 1) 75% of all assets must consist of mortgages, real estate equities, case or government securities; 2) at least 95% of taxable income must be distributed to REIT shareholders each year; 3) at least 75% of gross income must be derived from rents, mortgages, and gains from the sale of real estate; and 4) real property must not be held primarily for sale in the ordinary course of business.2 Qualified REITs pass through income, untaxed at the organization level, to the owners in much the same way as partnerships, although owners do not receive the benefit of pass through of tax losses. Boards of directors or trustees administer REITs on behalf of the owners, while day-to-day management is performed either internally or externally. Internally administered REITs employ their own acquisition and asset management staffs. When REITs are externally man- aged, outside advisors perform these services and bill the organization for expenses. The National Association of Real Estate Investment Trusts (NAREIT) classifies REITs into three categories: 1) equity REITs that have direct ownership of income producing real estate, principally office buildings, shopping centers, warehouses, and apartments; 2) mort- gage REITs that consist of investments in debt instruments secured by mortgages; and 3) hybrid REITs that combine both direct ownership of real estate and mortgage debt.3 Equity REITs hold 45%, mortgage REITs hold 46%, and hybrid REITs hold 9% of total REIT assets (NAREIT, 1993).4 The growth of REIT assets was slow from 1960 to 1968 after which organizers and in- vestors began to understand the advantages of REITs. With REIT assets of less than $1 billion in 1968, assets grew to $21 billion by 1975. Most of the growth during this period was due to investment by mortgage REITs; however, during the mid-1970s many of these REITs went into bankruptcy and liquidation due to poor loan underwriting and rising in- terest rates. At present each framework may be moved towards innovation for the simplicity from claiming operations. The training framework will be moving towards e-taking. Individuals tend to move from the manual to robotized methodology. That primary goal of the this will be will anticipate that lodging cost with admiration to the plan of the clients. Those exhibit strategies may be a long procedure in which those customers necessities to contact the land operator. The land operators give acceptable A suggestive on the lodging costs prediction. This strategy includes high hazard a direct result the land operator might furnish the bad data of the clients. They employments those straight relapse calculations should figure the cost. This analyses likewise utilized to foresee the best area for the clients to purchasing the houses. The information here utilized is from those Mumbai lodging board since 2009. Eventually, Tom's perusing utilizing this straight relapse he predicted the rate for every square foot. This prediction indicates the square feet of the house will be raised Eventually Towards 2018. (Bhagat et al.; 2016)

For classification problems, there is a way to find out the accuracy percentage with the help of the confusion matrices we can find out the accuracy percentage but the regression there is only one possibility to calculate the RMSE root mean squared error here the author says about the error indices an average error of a model calculated using the mean squared error (MSE) or the root mean squared error (RMSE). There is a problem in using the correlation coefficient as the significance test, not an appropriate one so that we prefer the RMSE.

Following a period of consolidation, REIT assets recovered to nearly $17 billion by 1985 NAREIT (1993) is approximately $ total U.S. real estate wealth.7 The and traded REITs have an equity represent 70% of the total (Gilbe Table 1 shows that REIT perform returns on all REITs were negative performance was recorded between formed the S&P 500 and substantitized real estate.

Table 1. Historic average annual returns (percentage): REITs, S&P 500, and unsecuritized real estate investments.

	REITs				Equity w/o Health Care	S&P 500	Russell-NCREIF
	All	Equity	Mortgage	Hybrid			
1972	11.19	8.01	12.17	11.41	8.01	18.90	—
1973	−27.22	−15.52	−36.26	−23.37	−15.52	−14.77	—
1974	−42.23	−21.40	−45.32	−52.22	−21.40	−26.39	—
1975	36.34	19.30	40.79	49.92	19.30	37.16	—
1976	48.97	47.59	51.71	48.19	47.59	23.57	—
1977	19.08	22.42	17.82	17.44	22.42	−7.41	—
1978	−1.64	10.34	−9.97	−7.29	10.34	6.39	—
1979	30.53	35.86	16.56	33.81	35.86	18.20	—
1980	28.02	24.37	16.80	42.46	24.37	32.27	18.07
1981	8.58	6.00	7.07	12.23	6.00	−5.01	16.86
1982	31.64	21.60	48.64	29.56	21.60	21.94	9.44
1983	25.47	30.64	16.90	29.90	30.64	22.39	13.31
1984	14.82	20.93	7.26	17.25	20.93	6.10	13.04
1985	5.92	19.10	−5.20	4.23	19.10	31.07	10.10
1986	19.18	19.16	19.21	18.75	16.41	18.56	6.53
1987	−10.67	−3.64	−15.67	−17.58	−4.48	5.10	5.67
1988	11.36	13.49	7.30	6.60	15.75	16.83	7.04
1989	−1.81	8.84	−15.90	−12.14	4.64	31.37	6.21
1990	−17.35	−14.84	−18.37	−28.21	−23.62	−3.27	1.47
1991	35.68	35.70	31.83	39.16	29.42	30.40	−6.08
1992	12.18	14.59	1.92	16.59	20.66	8.42	−5.03
1993[1]	11.38	12.14	8.84	10.50	12.89	2.21	—

1. Through February 26, 1993.

Sources: NAREIT, Chase Investment Performance Digest, Russell-NCREIF, and *Realty Stock Review*, March 29, 1993.

4.2. Review of the literature on investment issues

The objective of this section is to review the literature on investment decisions, which i cludes studies in which REIT data are used, that focus on prices and values without rega to financing decisions of REITs. The literature on financing decisions and REITs is review in the next section. This section of the review is organized by separating the literature in four categories: 1) REITs as real estate and common stocks, 2) REIT asset acquisitio and dispositions, 3) restricting of REITs, and 4) asset market information and REIT price. A good starting point for reviewing the REIT financial economics literature is to appropri ately characterize REITs as financial assets. This appears to be an unambiguous task. The returns on REITs should behave as do returns on real estate because legally qualified REIT must hold high percentages of real estate-related assets. Yet, REITs are securitized claims to real estate, which introduces a low-transaction-cost, trading-market dimension not present in the unsecuritized (i.e., real asset) market. Even before October 1987 when average REIT prices declined in one month by 14 per cent,8 there was general skepticism that REITs are not real estate, but instead hybrid finan- cial assets that embody the economic characteristics of the underlying real asset market coupled with the volatility of the stock market. During the same year as the stock marke crash, Goldman Sachs published a multi-part study of real estate returns and risks (Ross and Zisler, 1987a, 1987b).9 One aspect of the study was to select an appropriate return index upon which to base the analysis. The NAREIT, equity REIT index (EREIT) is con- sidered along with alternative, unsecuritized market indexes. Because of its high volatility Ross and Zisler conclude, "While EREIT is a true return series, it is not a true measure of the returns on the underlying assets in the fund's equity real estate" (1991, p. 181). The appraisal-based return series also are found to be unsatisfactory measures of the "true return on real estate. Ross and Zisler conclude that the true return index for real estate lies somewhere between available securitized and unsecuritized real estate return indexes. Others (Goetzman and Ibbotson, 1990) conjecture that smooth series of appraisal-based returns are closer to the evolution of true property values than those based on transactions or liquidations, such as REIT return indexes. An alternative view is that indexes of REIT returns are true returns on real estate if the market for REIT shares is efficient. Inefficiencies (i.e., REIT share prices not reflecting the values of the underlying real estate-related assets) are exploited through stock market trading, individual asset sales, and liquidation. Ennis and Burik (1991) cite several studies reviewed below (Allen and

Sirmans, 1987, Shilling, Do, and Sirmans, 1989; Gyourko and Keim, 1992; Giliberto, 1989; and Chan, Hendershott, and Sanders, 1990) to support the notion that REIT shares are efficiently priced. A market efficiency argument, together with their claim that the observed volatility of REIT returns is consistent with generally accepted expectations about volatility in the true real estate return series, lead Ennis and Burik to select REIT return indexes over appraisal-based return indexes as proxies for the true returns on real estate. One approach to resolving this controversy is to find evidence of market segmentation across return series for REITs, stock, and unsecuritized real estate.10 Ambrose, Ancel, and Griffiths (1992) test for segmentation among REITs and stocks during 1962 through 1989 using methods of fractal geometry. They conclude that both return series follow random walks that suggest no segmentation. Further, they conclude that REIT returns may not be good proxies for real estate returns. Liu, Hartzell, Greig, and Grissom (1990) find that equity REITs are integrated with the stock market, but they are unable to discern whether real estate is integrated with or segmented from the stock market (an important issue dis- cussed in the pages to follow.

5. METHODOLOGY

i. Data collection: The first step is to collect data on real estate transactions, such as sale price, location, size, and number of bedrooms. This data can be obtained from public records or real estate websites.

ii. Data pre-processing: Next, the collected data will need to be cleaned and pre-processed. This may include handling missing values, removing outliers, and encoding categorical variables

iii. Data exploration: Once the data has been cleaned, it can be explored to gain insights and identify patterns. This may involve creating plots and charts, calculating summary statistics, and performing statistical tests.

iv. Model training: After exploring the data, a machine learning model can be trained to predict real estate prices based on location and other relevant factors. This may involve splitting the data into training and test sets, selecting a model type and hyperparameters, and training the model on the training data.

v. Model evaluation: The trained model can then be evaluated on the test data to assess its performance. This may involve calculating metrics such as mean squared error or root mean vi. Model deployment: Once the model has been trained and evaluated, it can be deployed in a production environment, where it can be used to make real-time price predictions for new properties.

vi. Ongoing model maintenance: To ensure that the model continues to perform well, it may be necessary to monitor its performance over time and make updates as needed. This may involve retraining the model on new data or adjusting its hyperparameters.

6. RESULTS

Table 3 presents our results associated with Extra Trees, *k*–Nearest Neighbours, Random Forest, and OLS. In machine learning, we normally do not use R22 as a principal performance metric to evaluate the accuracy of a model, but its value still conveys some useful information. In ET, the R22 is as high as 0.96 in the training set and 0.91 in the test set. The negligible difference indicates no evidence of overfitting or underfitting. The results are then evaluated by the *MSE*, *RMSE*, and *MAPE* criteria. These three performance metrics are estimated to be 0.14405, 0.37953, and 6.49588%, demonstrating that ET fits our training data set very well. For our test set, *MSE*, *RMSE*, and *MAPE* are estimated to be 0.30561, 0.55282, and 9.04653%, respectively, demonstrating that ET also fits our test data set very well.

Table 3. Estimated results based on Random Forest and Ordinary Least Squares.

In *k*–Nearest Neighbors, the R22 is as high as 0.93 in the training set and 0.90 in the test set. The negligible difference indicates no evidence of overfitting or underfitting. The results are then evaluated by the *MSE*, *RMSE*, and *MAPE* criteria. These three performance metrics are estimated to be 0.23986, 0.48976, and 8.49793%, demonstrating that KNN fits our training data set very well. For our test set, *MSE*, *RMSE*, and *MAPE* are estimated to be 0.36211, 0.60176, and 10.39521%, respectively, demonstrating that KNN also fits our test data set very well.

In Random Forest, the R22 is as high as 0.96 in the training set and 0.92 in the test set. The negligible difference indicates no evidence of overfitting or underfitting. The results are then evaluated by the *MSE*, *RMSE*, and *MAPE* criteria. These three performance metrics are estimated to be 0.13155, 0.36270, and 6.22301%, respectively, demonstrating that RF fits our training data set very well. For our test set, *MSE*, *RMSE*, and *MAPE* are estimated to be 0.27918, 0.52837, and 8.88930% respectively, demonstrating that RF also fits our test data set very well.

However, although R22 is estimated to be reasonably good at 0.814 in our OLS model, this value is less than R22 of the test set associated with three algorithms by 6.62~12.90%. Such a difference is remarkable by any standard. Moreover, its three performance metrics are also worse than those of the test set associated with the three algorithms. In terms of *MSE*, its value for OLS is higher than those of these three algorithms by 76.44~128.85%. In terms of *RMSE*, its value for OLS is higher than those of these three algorithms by 32.83~51.28%. For *MAPE*, its value for OLS is higher than those of these three algorithms by 39.90~63.60%. Hence, we can surely confirm that Extra Trees, *k*–Nearest neighbours, and Random Forest outperforms OLS in terms of prediction and error minimization.

Based on the results of our RF estimation, the scatterplot of real estate prices and the residuals for the training set are shown in **Figure 4**. It demonstrates that RF typically matches the data quite well. The relationship between actual prices and their expected values is further illustrated in **Figure 5**. It is noticeable that one dot (whose property price is larger than 40 million) is lying far away from the clustering. With such an exception, almost all our predicted values closely follow the red line, showing that our model adequately fits our training data.

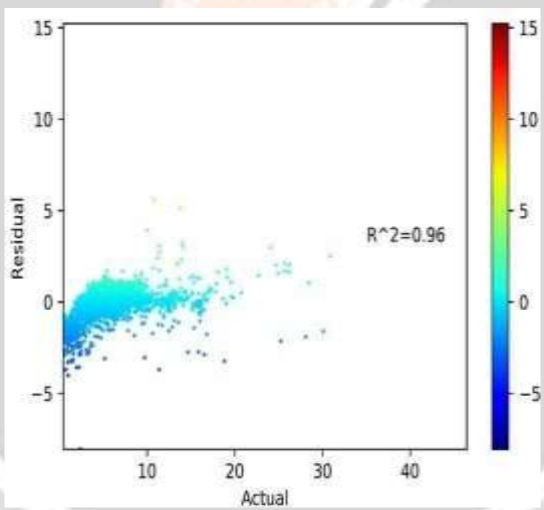

Figure 4. Property prices and residuals based on training set (RF).

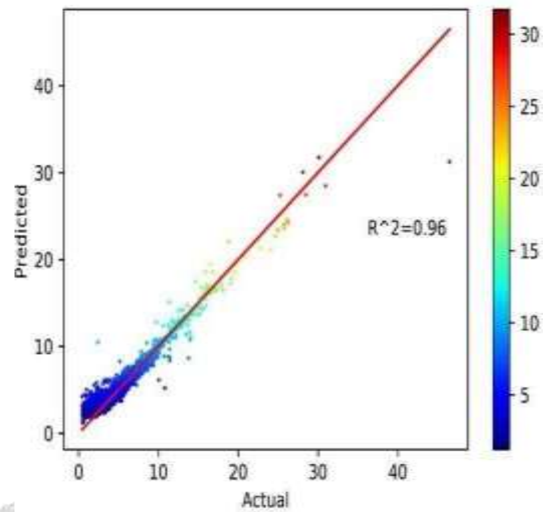

Figure 5. Actual and predicted property prices based on training set (RF).

Figure 6 displays the scatterplot of real estate prices and the residuals for the test set based on the findings of our RF estimation. It proves that RF usually closely matches the data. **Figure 7** also shows the relationship between actual prices and their expected values. Because almost all our predicted values closely follow the red line, with the exception of a few outliers, our model fits the test data set very well.

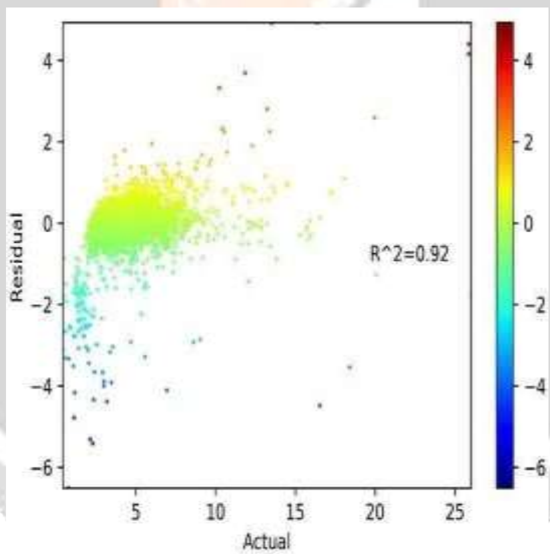

Figure 6. Property prices and residuals based on test set (RF).

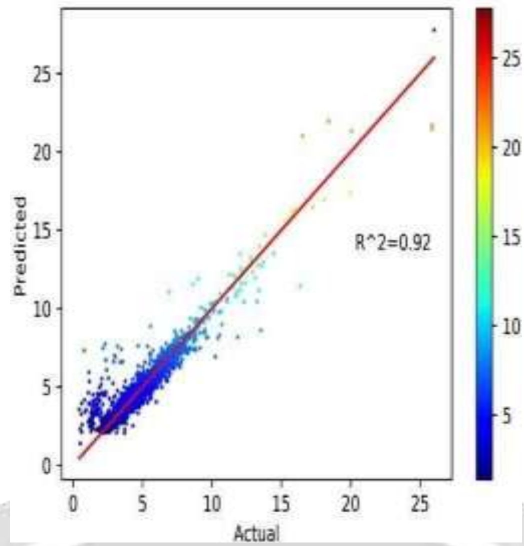

Figure 7. Actual and predicted property prices based on test set (RF)

7. CONCLUSION

This paper attempted to illustrate how machine learning can provide more accurate price predictions than traditional statistical technique, using the real estate market as an example. Extra Trees, *k*-Nearest neighbours, and Random Forest have been found to outperform the hedonic price model in terms of explanatory power and cost minimization. The increase in R square ranges between 6.62% and 12.9%. Accurate price signals in the property market predicted by machine learning algorithms play an important role in promoting sustainable production and consumption patterns. The government can incentivize homeowners to choose sustainable options and encourage developers to invest in sustainable practices by identifying where energy-efficiency improvements are needed. These contribute to the development of a more sustainable real estate market that benefits both the environment and society. With more accurate price information, buyers can identify properties that are overpriced and not worth the investment. This can help to reduce waste from unnecessary property development.

In conclusion, machine learning is expected to play a growing role in shaping our future. It has already been utilized in a variety of industries, ranging from healthcare to finance, and is having a significant impact on how we live and work. Although machine learning has the ability to significantly advance civilization, it also raises certain ethical issues that need to be resolved. For machine learning algorithms to work properly, a significant amount of data is needed. This may give rise to questions regarding the privacy of the people whose data is being exploited. In addition, there is a chance that private information will accidently leak or be misused. Machine learning algorithms can significantly affect people's life by influencing things such as loan or employment approval rates. It is crucial that these decisions are made equitably, openly, and without unduly disadvantaging any particular age group, gender, or race. Furthermore, employment displacement occurs when tasks that were previously carried out by humans are automated via machine learning. It is crucial to take into account how machine learning will affect the workforce and to make sure that employees have access to the training and assistance they need to adjust to these changes. Therefore, it is critical to pay close attention to the ethical implications of machine learning and to make sure that technology is applied responsibly and ethically. To create proper standards and laws to control the use of machine learning, it is necessary for researchers, legislators, and industry stakeholders to work together.

8. REFERENCES

1. Mak, S.W.K.; Choy, L.H.T.; Ho, W.K.O. Hedonic models, internet-based technologies and the provision of online property appraisal. *Constr. Innov.* **2008**, *8*, 92–105. [**Google Scholar**] [**CrossRef**]
2. Isada, F. The impact of inter-organisational network structures on research outcomes for artificial intelligence technologies. *Int. J. Econ. Sci.* **2022**, *11*, 1–18. [**Google Scholar**] [**CrossRef**]
3. Baldominos, A.; Blanco, I.; Moreno, A.J.; Iturrarte, R.; Bernárdez, Ó.; Afonso, C. Identifying real estate opportunities using machine learning. *Appl. Sci.* **2018**, *8*, 2321. [**Google Scholar**] [**CrossRef**][**Green Version**]

4. Sun, D.; Du, Y.; Xu, W.; Zuo, M.; Zhang, C.; Zhou, J. Combining online news articles and web search to predict the fluctuation of real estate market in big data context. *Pac. Asia J. Assoc. Inf. Syst.* **2015**, *6*, 19–37. [Google Scholar] [CrossRef]
5. Scholastica (Gay) Cororaton. Single-Family Homeowners Typically Accumulated $225,000 in Housing Wealth over 10 Years. National Association of Realtors. 2022. Available online: https://www.nar.realtor/blogs/economists-outlook/single-family-homeownerstypically-accumulated-225K-in-housing-wealth-over-10-years (accessed on 14 March 2023).
6. Pojar, J.; Macek, D.; Heralová, R.S.; Vitásek, S. Advances in costs optimization methods—Key study of maintenance and restoration of culture heritage. *Int. J. Econ. Sci.* **2022**, *11*, 163–178. [Google Scholar] [CrossRef]
7. National Association of Home Builders. Housing's Contribution to Gross Domestic Product. 2023. Available online: https://www.nahb.org/news-and-economics/housingeconomics/housings-economic-impact/housings-contribution-to-gross-domesticproduct (accessed on 14 March 2023).
8. Tsertekidis, G. Migrating from Greece to Germany after 2010: A qualitative approach. *Int. J. Econ. Sci.* **2022**, *11*, 73–92. [Google Scholar] [CrossRef]
9. Morano, P.; Guarini, M.R.; Tajani, F.; Liddo, F.D.; Anelli, D. Incidence of different types of urban green spaces on property prices—A case study in the Flamino District of Rome (Italy). In Proceedings of the Computational Science and Its Applications—ICSA2019—19th International Conference, Saint Petersburg, Russia, 1–4 July 2019; Lecture Notes in Computer Science. Misra, S., Gervasi, O., Murgante, B., Stankova, E.N., Korkhov, V.
10. Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Tarantino, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2019. Part IV. Volume 11622, pp. 23–34. [Google Scholar]
11. Hromada, E.; Cermakova, K.; Pecha, M. Determinants of House Prices and Housing Affordability Dynamics in the Czech Republic. *Eur. J. Interdiscip. Stud.* **2022**, *14*, 119–132. [Google Scholar] [CrossRef]
12. Ho, W.K.O. Modelling speculative activity in the Hong Kong residential property market. *Rev. Urban Reg. Dev. Stud.* **2000**, *12*, 137–148. [Google Scholar] [CrossRef]
13. Borgersen, T.A. A housing market with Cournot competition and a third housing sector. *Int. J. Econ. Sci.* **2022**, *11*, 13–27. [Google Scholar] [CrossRef]
14. Propersi, A.; Mastrilli, G.; Gundes, S. The third sector and social housing in Italy case study of a profit and non–profit public private partnership. In Proceedings of the 10th International Conference ISTR, Siena, Italy, 10–13 July 2012. [Google Scholar]
15. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [Google Scholar] [CrossRef][Green Version]
16. Mora-Garcia, R.T.; Cespedes-Lopez, M.F.; Perez-Sanchez, V.R. Housing Price Prediction Using Machine Learning Algorithms in COVID-19 Times. *Land* **2022**, *11*, 2100. [Google Scholar] [CrossRef]
17. Babu, A.; Chandran, A.S. Literature review on real estate value prediction using machine learning. *Int. J. Comput. Sci. Mob. Appl.* **2019**, *7*, 8–15. [Google Scholar]
18. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. KNN model-based approach in classification. In Proceedings of the On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE (OTM 2003), Catania, Italy, 3–7 November 2003; Lecture Notes in Computer Science. Meersman, R., Tari, Z., Schmidt, D.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; Volume 2888. [Google Scholar] [CrossRef]
19. Yildirim, H. Property value assessment using artificial neural networks, hedonic regression and nearest neighbors regression methods. *Selcuk Univ. J. Eng. Sci. Technol.* **2019**, *7*, 387–404. [Google Scholar] [CrossRef]
20. Mukhlishin, M.F.; Saputra, R.; Wibowo, A. Predicting house sale price using fuzzy logic, artificial neural network and k–nearest neighbor. In Proceedings of the 2017 1st International Conference on Informatics and Computational Sciences, Semarang, Indonesia, 15–16 November 2017; pp. 171–176. [Google Scholar]
21. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; Chapman & Hall/CRC Press: New York, NY, USA, 1984. [Google Scholar]
22. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [Google Scholar] [CrossRef][Green Version]
23. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [Google Scholar] [CrossRef][Green Version]
24. Koktashev, V.; Makee, V.; Shchepin, E.; Peresunko, P.; Tynchenko, V.V. Pricing modeling in the housing market with urban infrastructure effect. *J. Phys. Conf. Ser.* **2019**, *1353*, 012139. [Google Scholar] [CrossRef]
25. Varian, H.R. Big data: New tricks for econometrics. *J. Econ. Perspect.* **2014**, *28*, 3–28. [Google Scholar] [CrossRef][Green Version]

26. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Trees vs neurons: Comparison between random forest and ANN for high–resolution prediction of building energy consumption. *Energy Build.* **2017**, *147*, 77–89. [**Google Scholar**] [**CrossRef**]
27. Wang, C.C.; Wu, H. A new machine learning approach to house price estimation. *New Trends Math. Sci.* **2018**, *6*, 165–171. [**Google Scholar**] [**CrossRef**]
28. Ho, W.K.O.; Tang, B.S.; Wong, I.S.W. Predicting property prices with machine learning algorithms. *J. Prop. Res.* **2021**, *38*, 48–70. [**Google Scholar**] [**CrossRef**]
29. Diewert, W.E.; de Haan, J.; Hendriks, R. Hedonic regressions and the decomposition of a house price index into land and structure components. *J. Econom. Rev.* **2015**, *34*, 106–126. [**Google Scholar**] [**CrossRef**]
30. Rambaldi, A.N.; McAllister, R.R.; Fletcher, C.S. Decoupling land values in residential property prices: Smoothing methods for hedonic imputed price indices. In Proceedings of the 34th IARIW General Conference, Dresden, Germany, 21–27 August 2016. [**Google Scholar**]
31. De Aquino Afonso, B.K.; Melo, L.C.; de Oliveira, W.D.G.; Da Silva Sousa, S.B.; Berton, L. Housing prices prediction with a deep learning and random forest ensemble. In Proceedings of the Anais do Encontro Nacional de Inteligencia Artificial e Computacion, Rio Grande, Brazil, 20–23 October 2020, unpublished manuscript. [**Google Scholar**]
32. Alpaydin, E. *Introduction to Machine Learning*; MIT Press: Cambridge, MA, USA, 2009. [**Google Scholar**]
33. Rating and Valuation Department. Hong Kong Property Review Monthly Supplement. Government of Hong Kong Special Administrative Region. April 2022. Available online: **https://www.rvd.gov.hk/doc/en/statistics/full.pdf** (accessed on 10 September 2022).
34. Cornell University School of Hotel Administration The Scholarly Commons John B. Corgel Cornell University School of Hotel Administration, jc81@cornell.edu Willard McIntosh MIG Realty Advisors, Inc. Steven H. Ott University of Kentucky