Integrating Tabular Data and Satellite Imagery for House Price Prediction with Explainable Techniques

Sheng-Tzong Cheng(✉), Ya-Jin Lyu, Yen-Jung Chang, and Tzu-Yi Chiu

National Cheng Kung University, Tainan, Taiwan
stevecheng1688@gmail.com

Abstract. This study developed a housing price prediction model that integrates tabular data and satellite imagery to enhance prediction accuracy and model interpretability. A dataset named "Taiwan Housing Integrated Dataset (THID)" was constructed, comprising tabular data from Taiwan's real estate registration website and high-resolution satellite imagery from the Google API. In the model design, machine learning techniques were applied to analyze features and predict housing prices based on tabular data. Simultaneously, deep learning models were employed to process high-resolution satellite imagery to capture geographical environmental features. Furthermore, a linear regression model was used to integrate predictions from both tabular data and satellite imagery. To improve model transparency, interpretability techniques were introduced to analyze the decision-making process of the model. Lastly, GPT technology was utilized to transform interpretability analysis results into easily understandable text. Key contributions of this study include 1) the integration of tabular data and satellite imagery for housing price prediction, leveraging a linear regression model to combine predictions from different data sources; 2) emphasizing model interpretability to enhance transparency in the prediction process, and 3) using interpretability techniques to generate explanatory text for housing price predictions.

Keywords: Housing Price Prediction · Satellite Imagery · Interpretability

1 Introduction

House price prediction is a critical area of research with practical applications in personal home-buying decisions, government policy-making, and real estate investments. Traditional methods, such as linear regression and time series analysis, often fall short in handling the complexity and high dimensionality of real estate market data. The advent of data science and machine learning has paved the way for more accurate prediction techniques. Traditionally, these predictions relied on tabular data, including property features, geographic locations, and economic indicators. With the rise of imaging technology, internal and external house photos, and high-resolution satellite imagery have been incorporated into prediction models, enhancing data diversity and accuracy. However, these advanced models often lack transparency and interpretability, which are crucial for user trust and practical application.

D.-J. Deng and J.-C. Chen (Eds.): SGIoT 2024, LNICST 633, pp. 164–182, 2026.
https://doi.org/10.1007/978-3-031-93825-2_15

This study addresses the challenge of improving both curacy and interpretability in house price prediction models by combining tabular data with satellite imagery. Our objectives are to develop a comprehensive prediction model using LightGBM for tabular data and deep learning for satellite imagery, integrate these predictions through linear regression, and provide a detailed interpretability analysis. This includes generating explanatory text to elucidate the prediction mechanisms, thereby enhancing model transparency and user trust. The contributions of this research include proposing a novel, widely applicable prediction method, emphasizing model interpretability, analyzing various data sources and features, and offering detailed explanations for the model's predictions. By achieving these goals, this study aims to provide a reliable and interpretable house price prediction model, contributing valuable insights to the field.

2 Related Work

2.1 Linear Regression Methods

Multiple linear regression models are widely applied due to their ability to reveal correlations between house prices and various influencing factors. Mak, et al. [1] used quantile regression models to describe the relationship between property features and prices in Hong Kong. Mao and Yao [2] extended the application of linear regression models to utilize geographic features for predicting house prices, showing that integrating geographic information such as postal codes can significantly enhance prediction accuracy. Doza and Miller [3] applied linear regression models in California to predict the future median house prices of regions, further exploring the potential of using big data and machine learning techniques for house price predictions. Zhu [4] explored the potential of using KNN to correct biases in linear regression estimates, showing the potential for improving prediction accuracy.

In the real estate market of Ames, Iowa, Abdulhafedh [5] used multiple linear regression analysis combined with random forests to identify the most important predictive variables, resulting in high predictive accuracy. Overall, these studies demonstrate the widespread application and flexibility of linear regression techniques in house price prediction. Whether used alone or in combination with other techniques, they can provide the precise prediction tools needed for real estate market analysis. Each study emphasizes the importance of enhancing model capabilities and integrating new data sources to improve the accuracy of house price predictions, which is an indispensable tool for decision-makers and policymakers in the real estate industry.

2.2 Machine Learning Approaches

In 2020, Thamarai and Malarvizhi [6] modeled house price prediction based on house features (such as the number of bedrooms, house age, and transportation facilities) using decision tree classification, decision tree regression, and multiple linear regression, showing that these models effectively help clients select suitable houses. Supriya et al. [7] introduced a real estate price prediction system using machine learning algorithms. This system aims to predict house prices based on various features such as location, area,

number of bedrooms and bathrooms, etc. It employs multiple machine learning algorithms, including linear regression, decision tree regression, random forest regression, and artificial neural networks, to learn the relationship between these features and corresponding prices by training on a large dataset of house prices. The results indicate that the system can accurately predict house prices, making it a valuable tool for real estate agents, homebuyers, and sellers. Similarly, Fang [8] 2023 utilized various machine learning algorithms to predict house prices, exploring data preprocessing methods, including handling missing values and categorical data. The study employed models such as regression trees, random forests, XGBoost, Gradient Boosting, and LightGBM for house price prediction, showing excellent performance on benchmark datasets, and demonstrating the effectiveness of these methods in house price prediction tasks. In 2023, Li et al. [9] used models like LightGBM, Gradient Boosting, and XGBoost to train data and predict house prices. The results showed that the XGBoost model performed the best in house price prediction, with the lowest root mean square error (RMSE).

Overall, the application of machine learning methods in house price prediction shows significant advantages. These methods not only improve prediction accuracy but also handle large-scale data and multidimensional features, providing powerful tools for participants in the real estate market. Whether simple linear regression or complex neural networks and Gradient Boosting models, machine learning technologies continuously drive the accuracy and practicality of house price prediction.

2.3 Combining Tabular Data with Other Data Sources

In house price prediction research, combining multiple data sources has become a key strategy for enhancing prediction accuracy. Bhagat et al. [10] proposed a web application that combines market trend analysis with multiple linear regression algorithms for house price prediction, aiming to assist clients in investing in real estate without needing intermediaries, thus reducing transaction risks. Subsequently, Cardenas et al. [11] explored the effectiveness of combining structured data with unstructured textual descriptions, using techniques like TF-IDF and Bag of Words (BoW) to process textual descriptions of houses, demonstrating the effectiveness of data combination.

Regarding data fusion strategies, Gao et al. [12] proposed a location-centered multitask learning framework that integrates data sources such as transportation, education, community, and facilities, demonstrating significant advantages in house price prediction. Additionally, Kang et al. [13] proposed a framework combining structural attributes of houses, photos of houses, and local facility data, using machine learning models and geographically weighted regression for prediction, proving its high prediction accuracy. Recent research has increasingly emphasized the importance of visual data. Nouriani and Lemke [14] proposed an innovative method for house price estimation using interior, exterior, and satellite images. This method not only considers textual attributes of houses but also includes visual features extracted through deep learning, and classifies these images by luxury level. Combining this data with textual information for model input demonstrated the effectiveness of multimodal data, significantly improving the practicality and accuracy of estimations. Following this, Ahmed [15] proposed a method for house price prediction by extracting visual features and textual information from real estate websites. Although these data sources may not fully reflect the final actual

selling prices, the method has significantly improved prediction accuracy, increasing the R-value threefold and significantly reducing the mean squared error. These studies indicate that combining multiple data sources can provide various predictive advantages and meet the diverse and complex needs of future house price prediction research.

2.4 Explainability in House Price Prediction

In the field of house price prediction, utilizing machine learning technologies to enhance prediction accuracy while also emphasizing model interpretability is becoming increasingly important. Model interpretability not only improves transparency but also strengthens user trust in the model's predictions. A technique such as LIME (Local Interpretable Model-Agnostic Explanations) [16] has demonstrated its indispensable value. This tool reveals the key factors influencing prediction decisions and precisely quantifies their contributions to the outcome.

Although LIME is well-established in the field of house price prediction, it primarily explains the influence of specific features and does not provide a clear set of rules for model decisions. To further improve model trustworthiness and transparency, the introduction of the Anchor technique is crucial. The Anchor [17] provides a method for generating high-precision model-agnostic explanations by defining locally "sufficient" prediction conditions, enhancing model interpretability. This technique has shown its value in other fields, such as medical diagnosis and financial market forecasting.

The application of these techniques is not limited to house price prediction. For example, in healthcare, El Shawi et al. [18] compared the effectiveness of LIME and other interpretability techniques. These tools helped doctors understand complex models used for disease diagnosis, increasing the models' acceptance and trustworthiness. These studies show that interpretability techniques such as LIME and Anchor play a crucial role in enhancing the transparency and trustworthiness of house price prediction models and models in other fields.

3 Approach

This chapter provides an in-depth look at the research methods and model frameworks used for house price prediction. It begins with an exploration of data sources, from traditional tabular data to satellite imagery. Various machine learning models, including XGBoost, CatBoost, LightGBM, and deep learning models such as CNN, are introduced, along with techniques for combining these models to enhance prediction accuracy and interpretability.

3.1 Taiwan Housing Integrated Dataset (THID)

The THID is a comprehensive dataset constructed by our team, specifically designed for house price prediction. This dataset aggregates rich real estate transaction information from Taiwan, including traditional tabular data such as property features, transaction prices, and geographic locations, combined with emerging satellite imagery data to provide a detailed view of the surrounding environment of properties. The multi-source

data structure of THID not only enhances the accuracy of predictive models but also improves interpretability and transparency, allowing researchers to comprehensively analyze real estate market dynamics. Through this self-constructed dataset, we explore the application of machine learning and deep learning techniques for predicting and analyzing house price trends.

- Tabular Data Sources:

Taiwan's Real Price Registration data is a public platform used by the Taiwanese government to promote market transparency. This data can be obtained through the Real Price Registration website or the government's open data platform. The training data for this research is sourced from the open data platform and is accessible to research institutions, academic units, and commercial companies.

- Tabular Data Content:

The Real Price Registration data utilized in this study encompasses housing transaction records from six major cities in Taiwan: Taipei City, New Taipei City, Taoyuan City, Taichung City, Tainan City, and Kaohsiung City.

- Tabular Data Processing:

To ensure accuracy and reliability in Taiwan's Real Price Registration data, we addressed several challenges. Missing values and erroneous data, such as transaction dates earlier than building completion dates, were corrected. Data formats not directly usable for machine learning, like township names as strings and unstructured information in the "Remarks" field, were pre-processed. The below steps significantly enhance data quality, providing a solid foundation for building reliable prediction models.

1. Data Cleaning: Addressing missing values and outliers to ensure data integrity.
2. Feature Engineering: Encoding categorical features and normalizing continuous variables for better model processing.
3. Data Splitting: Divide the entire dataset into training, validation, and test sets using a ratio of 7:2:1. This step is crucial for assessing the generalization ability of machine learning models, ensuring that performance on unseen data meets expectations.
4. Outlier Handling: Using the Interquartile Range (IQR) method to detect and remove extreme values.

Satellite Image

In this study, residential addresses from the real estate transaction data were converted into geographical coordinates using Google's Geocoding API. Subsequently, satellite imagery for the corresponding locations was obtained using Google's Maps Static API.

Figure 1 displays nearby schools, indicated by visible playgrounds, adding neighborhood value.

3.2 Machine Learning

In this study, we employed multiple machine learning models for house price prediction. These models include XGBoost, CatBoost, and LightGBM.

Fig. 1. THID - Satellite Image. Revealing the Presence of Schools Near the Houses

- **XGBoost:** An optimized distributed Gradient Boosting library known for its high performance and speed, particularly with large datasets. It offers parallel tree boosting that efficiently solves many data science problems.
- **CatBoost:** A Gradient Boosting algorithm that excels in handling categorical features, reducing the need for extensive preprocessing. It is fast, accurate, and robust, with features to prevent overfitting.
- **LightGBM:** A highly efficient and scalable Gradient Boosting framework using tree-based learning algorithms. It handles large datasets with low memory usage and supports complex data structures, featuring histogram-based decision tree learning and leaf-wise growth.

3.3 Convolutional Neural Networks (CNN)

This study also applied CNN for house price prediction. CNNs excel in processing image data, so we chose to use the ResNet18 model to analyze satellite imagery. Additionally, we used Multilayer Perceptron (MLP) and Kolmogorov-Arnold Network (KAN) [19] for data fusion and further predictive analysis.

ResNet18

ResNet18 is a variant within the ResNet (Residual Network) family, featuring 18 layers. ResNet was introduced by He et al. [20] and represents a significant breakthrough in the field of deep learning. The core idea of ResNet is the introduction of residual blocks, which use skip connections to alleviate the vanishing gradient problem in deep networks. ResNet18 has the following characteristics:

- **Fully Connected Layers:** Each layer in an MLP is a fully connected layer, meaning each neuron in one layer is connected to every neuron in the next layer. This gives the MLP a strong learning capability to handle various complex nonlinear relationships.
- **Activation Functions:** MLPs typically use nonlinear activation functions (such as ReLU, Sigmoid, or Tanh). These activation functions enable the network to capture nonlinear features in the data.
- **Backpropagation:** MLPs are trained using the back-propagation algorithm, which updates the weights by computing the gradient of the loss function to minimize prediction error.

MLPs excel in handling structured data (such as tabular data) and are commonly used for classification and regression tasks. According to the universal approximation theorem, MLPs can approximate any continuous function, making them widely adaptable to various applications.

KAN

The KAN is a neural network structure based on the Kolmogorov-Arnold representation theorem. This network enhances performance and interpretability by incorporating learnable activation functions and a special structural design. The main features of KAN are as follows:

- **Kolmogorov-Arnold Representation Theorem:** This theorem states that any continuous function can be represented as a combination of multiple inner products and nonlinear transformations. Based on this theory, KAN can represent and learn complex mathematical relationships.
- **Learnable Activation Functions:** Unlike traditional neural networks that use fixed activation functions, KAN introduces learnable activation functions, allowing the network to adaptively adjust the form of the activation functions, thereby enhancing the model's flexibility and performance.
- **Special Structural Design:** The structural design of KAN allows summation operations at each layer and applies learnable activation functions on the edges. This design enables the network to more effectively learn and represent complex relationships in the data.

KAN excels in applications requiring high interpretability and flexibility, such as complex system modeling, data analysis, and prediction.

Explainable Artificial Intelligence (XAI)

In machine learning models, explaining the decision-making process is crucial for trusting and using these models. XAI aims to make the behavior and decision-making process of models transparent and understandable. In this study, we used three common interpretability techniques: LIME, and Anchor.

LIME

LIME is a versatile, model-agnostic method for local interpretability. It generates perturbed data around an input and uses a simple linear model to approximate the complex model's behavior locally. LIME can be applied to any model, including black-box models like deep neural networks and ensemble models, and handles various data types such as text, images, and tabular data. Its main advantage is its flexibility and intuitiveness, providing easy-to-understand explanations for individual predictions and helping users comprehend the model's behavior on specific data points.

Anchor

Anchor is a rule-based model explanation method that generates "anchors"—specific rules that, when satisfied, keep the model's prediction consistent in most scenarios. These rules aim for high coverage, ensuring that the model's predictions remain stable over a broad range when the rules apply. The simplicity and clarity of the rules generated

by Anchor allow users to easily understand their impact on model predictions. Anchor provides local explanations, making the model's behavior within a specific region predictable and consistent. This method offers clear, rule-based insights that enhance user trust and understanding of machine learning models.

3.4 Interpretable Text Generation with GPT

In this study, we utilize the GPT (Generative Pre-trained Transformer) API provided by OpenAI [21] to enhance model interpretability. By generating easy-to-understand text, GPT helps explain the decision-making process of machine learning models, thereby improving model transparency and interpretability. This allows end-users to better understand how the model makes its predictive decisions.

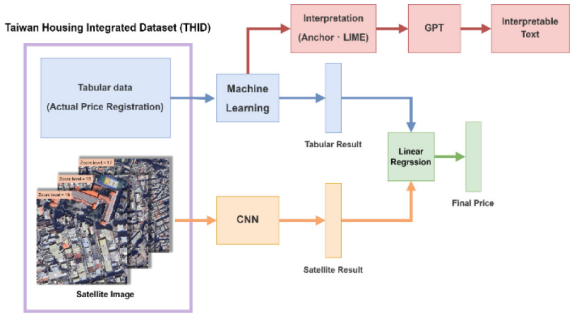

Fig. 2. Our Architecture

GPT is an advanced language generation model developed by OpenAI, which learns the deep semantics and structure of language by training on vast amounts of text data. In our implementation, GPT is used to generate text based on inputs provided by other interpretability techniques, such as Anchor and LIME. These techniques reveal which features have the most significant impact on the model's predictions, and GPT utilizes this information to construct narratives that explain the model's behavior.

The process of generating explanatory text involves feature explanation extraction and text generation. First, we use Anchor and LIME techniques to analyze the machine learning model, identifying the features and decision logic that significantly impact the prediction results. Then, GPT receives these analysis results and generates easy-to-understand text that details the model's decision-making process. These texts are reviewed and iterated as necessary to ensure their accuracy and relevance.

3.5 Architecture

In this study, we adopt a comprehensive machine learning architecture to predict the final prices of real estate properties by combining structured tabular data and unstructured satellite imagery data to enhance prediction accuracy and model interpretability. The data sources mainly include tabular data from the Real Price Registration and satellite images

at different zoom levels. The tabular data encompasses information such as transaction prices, locations, and building types, while the satellite images provide visual information about the geographical environment of the properties.

The design and execution process of the model involves three main parts, as shown in Fig. 2. First, the tabular data is input into a machine learning model specifically designed to analyze and learn important features within the data to predict house prices. Second, the satellite images are fed into the deep convolutional neural network ResNet18, which is designed to capture spatial hierarchical features in image data and identify visual elements that influence house prices, such as the surrounding environment and location. Finally, the outputs from the tabular data model and the CNN model are combined and analyzed using a linear regression model to derive the final house price prediction. To enhance model interpretability and transparency, we employ Anchor and LIME techniques to explain the predictive behavior of the machine learning model. These techniques help researchers and users understand the factors considered by the model when making predictions and identify the features that have the most significant impact on the prediction results. Additionally, we use the GPT model to generate explanatory text based on the explanations provided by Anchor and LIME, further enhancing the transparency of the model's decision-making process. This allows end-users to understand the model output in a more intuitive manner, increasing trust in the model's predictions. This multi-layered, multimodal architecture design enables the entire system to provide not only accurate predictions but also in-depth explanations, which is highly valuable for decision-makers in the real estate market.

4 Implementation and Experiments

4.1 Datasets

In this study, our primary data source is the THID, a dataset specifically compiled with real estate transaction data from the year 2023 (Republic Era 112). This dataset includes approximately 100,000 transaction records after thorough data cleaning, covering different regions and types of real estate across Taiwan's six major cities, reflecting the market's diversity. The tabular data component of this dataset is derived from the government-provided Real Price Registration [22], ensuring accurate and official transaction records. In addition to traditional tabular data, this study also utilized Google API to convert addresses from the real estate registry into geographical coordinates, obtaining high-resolution satellite imagery of the corresponding locations. These images capture the natural landscapes and urban layouts surrounding the properties, providing additional visual information for analyzing factors affecting housing prices. By using images from various zoom levels, we can more accurately assess the impact of geographical features on housing prices. These rich data sources are not only used for predicting real estate prices but also provide key information needed to enhance model explainability. The real-time update of this data ensures the timeliness and practicality of the research, making it an essential resource for analyzing and predicting housing prices.

4.2 Training Procedure and Evaluation Metrics

Training Procedure

The model training process starts with data preprocessing, including filling in missing values, re-moving anomalies, and standardizing data. Categorical features are encoded, and outlier analysis using the IQR is performed to enhance data quality. Advanced machine learning models like XGBoost, CatBoost, and LightGBM are trained, followed by feature importance evaluation to remove less important features. Hyperparameter tuning is conducted using Bayesian optimization and cross-validation, with an early stopping mechanism to prevent overfitting. This comprehensive workflow ensures optimal predictive performance and generalization, providing a robust foundation for house price prediction.

Evaluation Metrics

To assess the performance of the prediction models, we used several precise evaluation metrics:

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

MAE represents the average absolute difference between predicted and actual values.

- Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

MSE calculates the average squared differences between predicted and actual values, with lower values indicating higher accuracy.

- Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

RMSE provides the error scale in original units for easier comparison with actual data.

- R-squared (R^2):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

R^2 measures the proportion of variance in the dependent variable predictable from the independent variables, with values closer to 1 indicating stronger model performance. These metrics were used during cross-validation and final testing phases to ensure comprehensive and objective model evaluation, analyzing performance across multiple datasets for stability and reliability.

4.3 Experimental Results

Tabular Data Prediction

This study utilized three mainstream machine learning models, LightGBM, XGBoost, and CatBoost, to explore the impact of tabular data on housing price prediction. Experiments were conducted under four data processing scenarios: using "All Data," "Only Real Price Registration Data," "All Data excluding price per square meter," and "Only Real Price Registration Data excluding price per square meter." We also compared model performance with and without outlier handling.

The results, presented in Tables 1 and 2, show that without outlier handling, models performed well on metrics like MAE, MSE, and RMSE but had weaker R^2 scores, indicating limitations in predicting variability. In contrast, with outlier handling, all evaluation metrics improved, highlighting the importance of proper outlier handling for enhancing prediction accuracy and reliability.

Table 1. Housing price prediction performance: no outlier handling

Data	Model	MAE	MSE	RMSE	R^2
All	CatBoost	0.00010	7.79E-07	0.00088	0.91413
	LightGBM	0.00015	2.20E-06	0.00148	0.75771
	XGBoost	0.00015	1.40E-06	0.00118	0.84528
All (excl. Price/m^2)	CatBoost	0.00033	1.91E-06	0.00138	0.78906
	LightGBM	0.00038	3.65E-06	0.00191	0.59752
	XGBoost	0.00044	4.66E-06	0.00216	0.48666
Only Real Price Registration	CatBoost	0.00010	1.19E-06	0.00109	0.86845
	LightGBM	0.00013	1.75E-06	0.00132	0.80699
	XGBoost	0.00017	2.56E-06	0.00160	0.71827
Only Real Price Registration (excl. Price/m^2)	CatBoost	0.00036	3.05E-06	0.00175	0.66349
	LightGBM	0.00038	3.34E-06	0.00183	0.63164
	XGBoost	0.00048	6.27E-06	0.00250	0.30943

Despite adding external features such as schools, convenience stores, and metro stations, these did not significantly improve predictive performance, possibly due to poor representation in the dataset or low correlation with housing prices. Figures 3 and 4 illustrate model performance, showing that outlier treatment leads to better alignment of predicted and actual values, as seen in Q-Q and scatter plots.

Ultimately, the LightGBM model trained on "All Data excluding price per square meter" with outlier handling was selected for the stacked model to achieve more stable and accurate predictions.

Satellite Image Prediction

In the satellite image-based prediction model experiments depicted in Tables 3 and 4,

Table 2. Housing price prediction performance: with outlier handling

Data	Model	MAE	MSE	RMSE	R^2
All	CatBoost	0.00010	7.79E-07	0.00088	0.91413
	LightGBM	0.00015	2.20E-06	0.00148	0.75771
	XGBoost	0.00015	1.40E-06	0.00118	0.84528
All (excl. Price/m^2)	CatBoost	0.00033	1.91E-06	0.00138	0.78906
	LightGBM	0.00038	3.65E-06	0.00191	0.59752
	XGBoost	0.00044	4.66E-06	0.00216	0.48666
Only Real Price Registration	CatBoost	0.00010	1.19E-06	0.00109	0.86845
	LightGBM	0.00013	1.75E-06	0.00132	0.80699
	XGBoost	0.00017	2.56E-06	0.00160	0.71827
Only Real Price Registration (excl. Price/m^2)	CatBoost	0.00036	3.05E-06	0.00175	0.66349
	LightGBM	0.00038	3.34E-06	0.00183	0.63164
	XGBoost	0.00048	6.27E-06	0.00250	0.30943

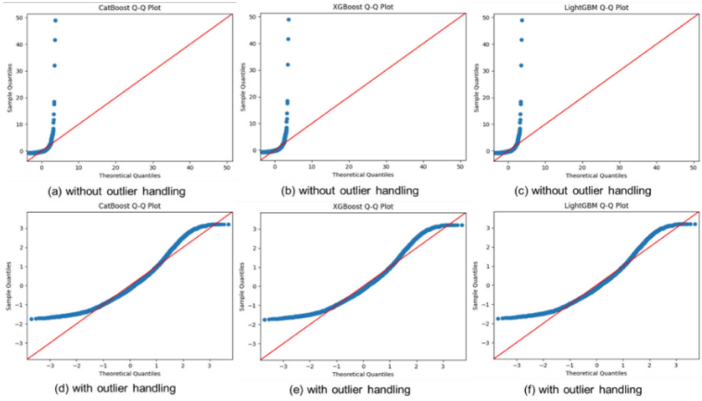

Fig. 3. Q-Q Plot: Comparison of CatBoost, XGBoost, and LightGBM

the complexity of images significantly influenced model performance. Table 3 shows that under a single zoom level (zoom level 18), models like AlexNet and GoogleNet performed poorly, indicated by their negative R^2 values, possibly because these older architectures struggled with the complex image details. In contrast, ResNet18 and its modified version, ModifiedResNet18, were better suited for handling such complexity due to the resilience provided by Residual Blocks, which effectively mitigate the vanishing gradient problem in deep networks.

Stacked Model Prediction

In this study, we chose to train the LightGBM model using the "All Data (excluding Price per(sq.m)" dataset to predict house prices, aiming to explore the impact of external

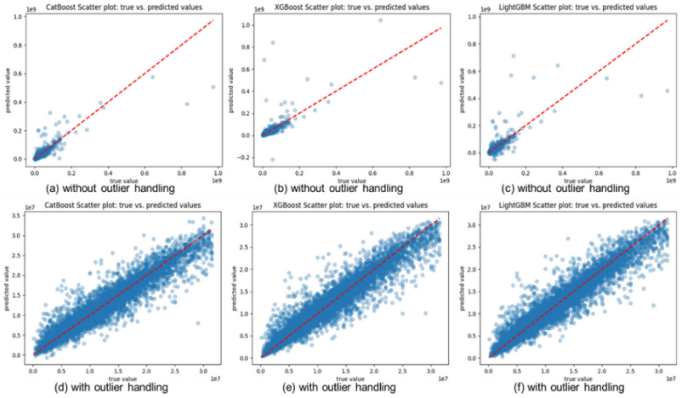

Fig. 4. Scatter Plot: Comparison of CatBoost, XGBoost, and LightGBM

features on model performance. LightGBM outperformed the other two models, hence it was selected as the primary prediction tool for tabular data. For satellite image-based house price prediction, we employed ModifiedResNet18 trained at three different scales, demonstrating excellent performance. The outputs of these two models were used as inputs to a stacked model for final house price prediction. As shown in Table 5, the stacked model's performance surpassed that of any single model, confirming the effectiveness of integrating multiple prediction sources.

However, the incremental improvement with the stacked model came with a considerably higher computational cost. This suggests that future efforts could focus on improving the practical applicability of this method. For instance, annotating satellite images to enrich the dataset or developing techniques to enhance the interpretability of satellite image data could potentially increase the practical value of this methodology. Implementing such adjustments may provide a better balance between performance gains and computational expenses, thereby justifying the use of sophisticated models like the stacked model in real-world applications.

Table 3. Model performance comparison

Model	Zoom level	MAE	MSE	RMSE	R^2
AlexNet [23]	18	0.157046	0.039960	0.199901	−0.000030
ConvNeXt [24]	18	0.156516	0.039773	0.199430	0.015626
GoogleNet [25]	18	0.158843	0.040417	0.201040	−0.000327
ResNet18	18	0.156897	0.038667	0.196641	0.042968
ModifiedResNet18	18	0.150382	0.034554	0.185889	0.144762

Table 6 depicts the distribution of per square meter price prediction errors across these cities, with categories such as <10 k, 10 k–50 k, 50 k–100 k, 100 k–150 k, 150 k–200 k, and >200 k. This table helps to understand the specifics of per-square-meter price

Table 4. Model comparison at different zoom levels

Model	Zoom level	MAE	MSE	RMSE	R^2
ResNet18	18	0.156897	0.038667	0.196641	0.042968
	17, 18, 19	0.165012	0.040348	0.200868	0.001377
ModifiedResNet18	18	0.150382	0.034554	0.185889	0.144763
	17, 18, 19	0.146043	0.033717	0.183625	0.165476

Table 5. Comparison of stacked model prediction performance

Model	Data Type	MAE	MSE	RMSE	R^2
LightGBM	All (excl. Price / m^2)	0.038590	0.003450	0.058740	0.914610
Modified ResNet18	Zoom level: 17, 18, 19	0.146044	0.033718	0.183625	0.165476
Stack model		0.038551	0.003446	0.058699	0.914722

prediction errors in different cities. From the data, most properties in most cities have errors of less than 10 k, signifying high accuracy in predictions. For instance, New Taipei City has 1483 properties with errors under 10 k, while only 50 properties fall within the 50 k–100 k error range. This study uses error per square meter to assess the accuracy and consistency of the housing price prediction models across different counties, providing a more precise measure. This approach allows for a more accurate determination of prediction accuracy. For instance, while the total price error might appear significant, it could be due to the larger area of the property, yet the per-square-meter error might be less than 10,000, indicating relative precision in the model's prediction of price per square meter.

Table 6. Per square meter price prediction errors by city

City	<10 k	10 k–50 k	50 k–100 k	100 k–150 k	150 k–200 k	>200 k
New Taipei City	1483	1080	50	2	0	0
Taoyuan City	1167	455	9	0	0	0
Taichung City	1328	611	21	1	0	0
Taipei City	303	407	34	2	1	0
Tainan City	428	188	3	0	0	0
Kaohsiung City	1450	649	15	3	0	0

Explainability Analysis

In the interpretability analysis phase, we focused on the LightGBM model's predictions, using LIME, Anchor, and explainable text generation to interpret and validate results. These techniques provided unique insights into the model's decision-making process, allowing us to accurately assess performance and validate the reliability of the predictions.

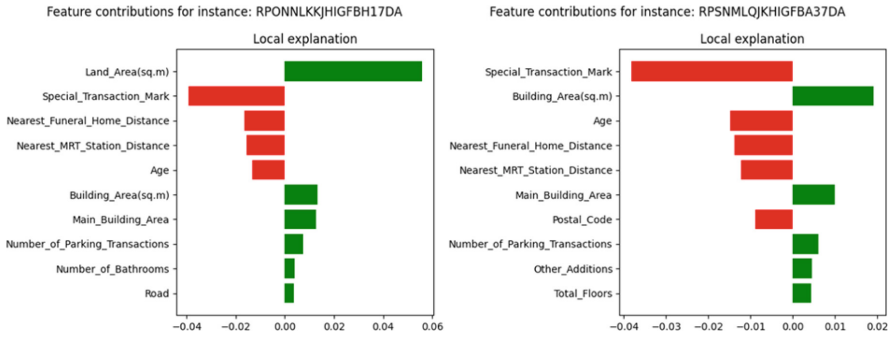

Fig. 5. LIME Analysis Examples - Impact of Real Estate Transaction Features

LIME

In our house price prediction model, LIME is used to explain the reasons behind specific predictions and assess the impact of different features on prediction outcomes. This is achieved by simulating small perturbations in input features and observing how prediction results change. The experimental results, shown in Fig. 5, provide insights into the model's behavior, allowing us to better understand which features are crucial factors affecting house price predictions under given circumstances.

Anchor

In our house price prediction model, we use Anchor to identify the combination of features that have the most significant impact on the model's predictions. This method is particularly useful in pinpointing which conditions are decisive during prediction, thus providing more specific explanations. For example, Anchor may identify features of a particular location or specific conditions of a house, such as area and age, that provide clear reasons for predicting higher or lower house prices. The experimental results, shown in Table 7, demonstrate how the model relies on these key features to make decisions in various scenarios.

Interpretable Text Generation

Finally, we utilized GPT technology to transform the results of LIME and Anchor interpretability analyses into easily understandable text. This enhanced transparency not only clarified the model's decision-making process for all stakeholders but also provided intuitive explanations of how the model integrates different features to make predictions.

Table 7. Anchor rules for predictive analysis

Anchor Rule	Precision	Coverage
• Financial_Institution_Count $<=$ 0.00 • Parking_Area(sq.m) $>$ 14.60 • Attached_Building_Area $>$ 0.00 • Nearest_MRT_Station_Distance $>$ 0.65 • Transaction_Object $>$ 0.00	0.407867	0.0556
• Financial_Institution_Count $>$ 0.00 • Parking_Area(sq.m) $>$ 0.00 • Attached_Building_Area $>$ 0.00 • Nearest_MRT_Station_Distance $>$ 0.65 • Medical_Institution_Count $>$ 1.00 • Number_of_Parking_Transactions $>$ 0.00	0.568339	0.0381

The integration of these interpretability methods not only enhanced the model's transparency and credibility on a technical level but also allowed us to accurately assess and improve the model's predictive performance. This ensures the reliability and effectiveness of the model in real-world applications. As demonstrated in Table 8, these insights significantly contribute to understanding the nuanced factors that drive house price predictions.

Table 8. Explainable text - generated from anchor rules and lime analysis

Explainable technique	Explainable Analysis
LIME	• Special_Transaction_Mark $= -0.040$ • Age $= -0.018$ • Nearest_MRT_Station_Distance $= -0.016$ • Building_Area(sq.m) $= 0.015$ • Nearest_Funeral_Home_Distance $= -0.014$ • Number_of_Parking_Transactions $= 0.008$ • Parking_Area(sq.m) $= -0.005$ • Nearest_Junior_High_School_Distance $= 0.004$ • Transaction_Object $= 0.001$ • ATM_Count $= 0.001$

(*continued*)

Table 8. (*continued*)

Explainable technique	Explainable Analysis
Anchor	• Financial_Institution_Count $< = 0.00$ • Parking_Category > 0.00 • Attached_Building_Area $< = 0.00$ • Transaction_Object > 0.00 • Parking_Area(sq.m) $< = 0.00$
Explainable Text Generated	The house pricing market is a complex one, with numerous factors playing a crucial role in determining the value of a property. From the anchor data provided, we can observe that certain features have a significant impact on the pricing of houses. For instance, having zero or fewer financial institutions in close proximity to the property can result in a decrease in house price. This could be due to limited access to banking services and convenience, making it less desirable for potential buyers Similarly, having ample parking spaces and areas can also positively influence house prices. This could be because of the convenience and ease of access for car owners, making it an attractive feature for those who own vehicles Delving deeper into the LIME analysis results, we can see that various factors such as age, nearest MRT station distance, and nearest funeral home distance have a negative impact on house prices. Older properties tend to depreciate over time, while being too close to public transportation or funeral homes may result in noise pollution and discomfort for potential buyers On the other hand, features like building area and number of parking transactions have a positive influence on house prices. Larger building areas provide more living space while having more parking transactions indicates high demand for parking spaces in that particular area Furthermore, other aspects such as transaction object type and ATM count may also play a minor role in determining house prices In conclusion, it is evident that various factors such as location amenities, accessibility, convenience, noise pollution levels and demand for certain features all contribute to shaping the housing market's pricing landscape. Potential buyers must carefully consider these factors before making any property purchase decisions

5 Conclusions and Future Work

In this section, we will discuss the experimental results of the models we used, and draw conclusions in Sect. 5.1. In Sect. 5.2, we will address our future prospects.

5.1 Conclusions

In this study, we successfully developed an innovative housing price prediction model that integrates tabular data and satellite imagery, effectively enhancing the accuracy and interpretability of predictions. By establishing the THID, we not only enriched the data sources but also enhanced the data processing and analysis capabilities using advanced machine learning technologies and deep learning models, such as LightGBM and ResNet18. Additionally, this study's methodological contribution lies in providing a reliable model design framework that can effectively integrate and analyze information from multiple data sources in housing price prediction. We particularly focused on the interpretability of the model, using advanced explanatory techniques such as LIME and Anchor to elucidate the decision-making process of the model. This transparency enhancement is an important expansion of existing technologies, enhancing the trustworthiness of the model and providing a robust decision-support tool for potential policymakers and real estate professionals. Overall, this study not only offers a highly accurate and interpretable housing price prediction model but also significantly enhances the model's application value and practicality by integrating various data sources and explanatory techniques. These achievements provide significant theoretical extensions to the academic community and have practical application value for the analysis and policy-making of the real estate market, demonstrating how technological advancements can enhance the predictive capabilities and explanatory power of economic models, bringing new perspectives and methods to the field of housing price prediction.

5.2 Future Work

Future research will focus on expanding the housing price prediction model by integrating additional data sources such as social media comments, economic indicators, and environmental factors to capture more subtle influences on housing prices. We will also explore emerging machine learning techniques like transfer learning and deep reinforcement learning to enhance prediction accuracy and model adaptability. Improving model interpretability remains a priority, with plans to address the limitations of the Anchor technique in handling continuous values and develop advanced visualization tools and automated explanation generation technologies. Additionally, applying the model to real-time real estate market conditions will help validate and optimize its practicality and effectiveness, enhancing its value and impact in real-world applications.

References

1. Mak, S., Choy, L., Ho, W.: Quantile regression estimates of hong kong real estate prices. Urban Studies **47**(11), 2461–2472 (2010)
2. Mao, Y., Yao, R.: A Geographic Feature Integrated Multivariate Linear Regression Method for House Price Prediction. Proceedings of the 2020 3rd International Conference on Humanities Education and Social Sciences (ICHESS 2020) (2020)
3. Doza, L., Miller, J.R.: Forecasting california housing prices using a linear regression model. Proc. W. Va. Acad. Sci. **92**(1) (2020)
4. Zhu, L.: Optimization of linear regression in house price prediction. ACE **6**, 684–691 (2023)

5. Abdulhafedh, A.S.: Incorporating Multiple Linear Regression in Predicting the House Prices Using a Big Real Estate Dataset with 80 Independent Variables. OALib (2022)
6. Thamarai, M., Malarvizhi, S.P.: House price prediction modeling using machine learning. Int. J. Info. Eng. Electr. Bus. (IJIEEB) **12**(2), 15–20 (2020)
7. Supriya, M.S., Vinayak, G.S., Patgar, V.R., Mahajan, V.: House price prediction system using machine learning algorithms and visualization. IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), pp. 1–6. Bangalore, India (2023)
8. Fang, L.: Machine learning models for house price prediction. ACE **4**, 409–415 (2023)
9. Li, H.: House price prediction based on machine learning. ACE **4**, 623–628 (2023)
10. Bhagat, N., Mohokar, A., Mane, S.: House price forecasting using data mining. International Journal of Computer Applications **152**, 23–26 (2016)
11. Cardenas, E., Shorten, C., Khoshgoftaar, T.M., Furht, B.: A Comparison of House Price Classification with Structured and Unstructured Text Data. The International FLAIRS Conference Proceedings **35** (2022)
12. Gao, G., Bao, Z., Cao, J., Qin, A.K., Sellis, T.: Location-Centered House Price Prediction: A Multi-Task Learning Approach. ACM Trans. Intell. Syst. Technol., New York, NY, USA **13**(2), Article 32, 1–25 (2022)
13. Kang, Y., et al.: Understanding house price appreciation using multisource big geo-data and machine learning. Land Use Policy **111** (2021)
14. Ali, N., Lemke, L.: Vision-based housing price estimation using interior, exterior & satellite images. Intel. Sys. Appl. **14** (2022)
15. Ahmed, H.E., Moustafa, M.: House price estimation from visual and textual features. In: Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016) - NCTA, pp. 62–68 (2016)
16. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should i trust you?: explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), pp. 1135–1144 (2016)
17. Ribeiro, M.T., Singh, S., Guestrin, C.: Anchors: High-Precision Model-Agnostic Explanations. Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 32 No. 1 (2018)
18. El Shawi, R., Sherif, Y., Al-Mallah, M., Sakr, S.: Interpretability in HealthCare A Comparative Study of Local Machine Learning Interpretability Techniques, IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), pp. 275–280. Cordoba, Spain (2019)
19. Liu, Z., et al.: Kan: Kolmogorov-arnold networks. arXiv preprint arXiv:2404.19756 (2024)
20. He, K., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)
21. OpenAI: OpenAI GPT-3 API [gpt-3.5-turbo-instruct] (2024). https://platform.openai.com/docs/models/gpt-3-5-turbo
22. Ministry of the Interior, Taiwan (2023). [Real Estate Transaction Data Release 2023]. https://plvr.land.moi.gov.tw/DownloadOpenData
23. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12), pp. 1097–1105. Curran Associates Inc., Red Hook, NY, USA (2012)
24. Liu, Z., et al.: A ConvNet for the 2020s. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11966–11976. New Orleans, LA, USA (2022)
25. Szegedy, C., et al.: Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9. Boston, MA, USA (2015)