Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

FEATURE LEARNING AND PRICE ESTIMATION USING DEEP LEARNING IN REAL ESTATE

[#1]**Kuna Neelima,** *Department of MCA,*
[#2]**Dr. T. Ravikumar,** *Professor, Department of CSE,*
Vaageswari College of Engineering(Autonomous), Karimnagar, TG.

ABSTRACT: Taiwanese houses, flats, and other types of urban homes should be looked at very carefully. Real estate researchers and forecasters frequently implement surveys and statistical instruments to accumulate data. Nevertheless, when confronted with multidimensional data, these methodologies become inefficient and time-consuming. The objective of this initiative was to develop a real estate forecasting model that could adapt to the current environment and circumstances. The data was standardized to ensure that it was categorized correctly after being extracted from public government records. In order to guarantee that the appropriate clustering method was implemented, cross-statistical analysis was implemented. In order to enhance the precision of classification, an automatic encoder for deep learning was implemented. Double-bottom map ppaper swarm optimization (DBM-PSO) was implemented to identify the optimal clustering solution. We sought to determine the factors that contributed to the increase in property prices in Taiwan over the past decade by employing cluster analysis and deep learning on publicly available data. The findings indicate that the average unit price, the frequency of real estate transactions, and the building material and construction index all significantly influence the prices of property in Taiwan. These discoveries can assist policymakers and researchers in addressing specific aspects of real estate development and improving market regulation to prevent excessive growth. The framework that was proposed is a novel approach to analyzing and forecasting real estate trends.

Keywords: *Machine learning; real estate; particle swarm optimization algorithm; economy*

1. INTRODUCTION

Owning a residence is the primary objective of the majority of Taiwanese families, and the adage "own the land, own the wealth" is profoundly ingrained in their culture. Real estate encompasses land and constructed structures, including residences, apartments, and fences. A residence is both a secure residential environment and a symbol of security and protection; however, real estate serves both functions. Real estate is a prudent investment and a means of safeguarding oneself from inflation, as its value and rental income increase in tandem with inflation rates and economic expansion.

Through open databases, researchers can obtain housing information for further investigation. A wide range of information is available in the Federal Reserve Economic Data (FRED) database, including housing inventory, starts, sales, the affordability index, the price index, transactions, interest rates, and more. However, due to the complexity and frequent fluctuations of real estate data, the majority of studies concentrate solely on trends and projections regarding the value of properties. It is crucial to be aware of the number of rooms, the price per unit, and the registration process. Machine learning is employed to generate

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

long-term predictions due to risks such as market volatility, pandemics, and economic issues, which render events less predictable.

Previous research has examined real estate factors, including the deconstruction of office vacancy rates, the use of methods from various disciplines to determine prices, and the prediction of prices based on trends. Government policy, population growth, economic development, and natural disasters are among the numerous factors that contribute to changes in Taiwanese housing. Property values are significantly influenced by psychological and macroeconomic factors, including income, mortgage debt, and limited space.

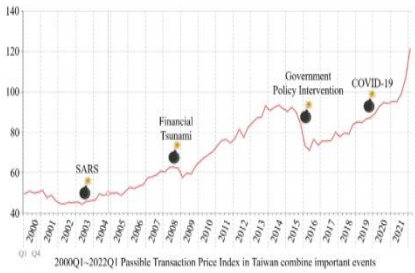

Figure1. Transaction price index in Taiwan combines actual

The cost of homes in Taiwan has been consistently increasing, but this has not deterred individuals from purchasing home. Sociodemographic data and stock market indices are extremely beneficial for real estate research. The paper of real estate has experienced a significant increase in the wake of the Great Financial Crisis. It is evident that the connections between financial markets, real estate, and social factors have become significantly more significant. Accurate forecasts are indispensable for policymakers who are responsible for maintaining market equilibrium and ensuring that businesses can compete.

2. METHODOLOGY

This section provides an explanation of the proposed grouping paradigm. The reconstruction error and the distance between data points and their corresponding groups in the code layer are reduced by a four-layer autoencoder network in the initial phase. The nonlinear mapping method is more effective in displaying the data layer. The process consists of four primary steps: (1) standardizing data for various features and attributes; (2) reducing the number of dimensions by utilizing an autoencoder to extract low-dimensional data from multidimensional datasets; (3) grouping data into low-dimensional spaces; and (4) examining the groups that were defined in terms of real estate markets. Following this, we will conduct a thorough examination of the selected autoencoder network and demonstrate its potential for aggregation.

Data Standardization

The data contained a variety of sources, characteristics, and information. Our objective was to ensure that the independent variables (x) had identical ranges when we implemented mean normalization. The source file containing the variables is located in the Supplementary Files (98index.xlsx).

The independent variable's mean value is represented by the symbol ±. The raw data and descriptive figures for a variety of variables are included in the supplementary files (98index.xlsx).

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

Reducing High-Dimensional Models Using Auto-Encoders

An autoencoder is a unique type of neural network that has the ability to reconstruct its input by locating a concealed copy of it. With a single layer, an autoencoder network is composed of an encoder and a decoder. The encoder employs a nonlinear mapping function f(x) to transform an input xi into its concealed version hi.

The encoding bias vector is denoted by b(1), while the encoding weight matrix is denoted by W(1). It operates similarly to PCA by employing a one-layer autoencoding network with linear activation functions.

The bias vector b(2) and the decoding weight matrix W(2) are specified. The condition W(2) = W(1)T functions as a regularizer by reducing the number of parameters that require optimization. The coding procedure is used to determine the optimal values for W and b.

Assume that x′i is determined in the same manner as before, and that θ = {W(1), W(2), b(1), b(2)}. A deep autoencoder network is constructed by stacking numerous single-layer autoencoders on top of one another. The concealed form of each layer is communicated to the subsequent layer. The training of deep autoencoders is more challenging due to the high sensitivity of gradient descent to the initial weights. To address this issue, a pre-training method is implemented to establish robust features in the model prior to its fine-tuning. Denoising autoencoders and limited Boltzmann machines are capable of reconstructing inputs that have been distorted. Denoising autoencoders are instructed to restore the original input by assigning random 0 values to certain inputs. Statistical connections between sources are recorded in this manner by learning.

After all of its levels have been pre-trained, the autoencoder is fine-tuned. Within an autoencoder, there are four layers: an encoder with two dense levels that compresses data into a 36-dimensional latent vector, and a decoder with two dense layers that reconstructs the original data. Twenty-five percent of all features were generated by the feature generation layer. In order to observe the remaining 25%, we implemented a 75% and 50% feature reduction. The input vector of our dataset comprises 144 characteristics. Four layers of neurons comprise the autoencoder. 72 neurons are present in Layer 1, while 108 neurons are present in Layer 2. The input data is composed of independent factors that are indexed by time.

3. LITERATURE REVIEW

Peterson & Malik (2021) This investigation introduces a deep learning-based approach to feature extraction and real estate market price estimation. By utilizing characteristics of properties from images, geographic data, and building types, Convolutional Neural Networks (CNN) can accurately predict home prices. The framework evaluates the quality of the area, the proximity of the property to public transportation, and the scale of the property in order to enhance its worth performance. The results of the experiments indicate that the predictions are more precise and accurate than those of conventional regression models.

Garcia & Hoffman (2022) The research demonstrates a sophisticated approach to estimating the value of a home through the use of deep learning and data-driven feature extraction. In order to ascertain the current value of a property, Deep Neural Networks (DNN) analyze economic factors, market dynamics, and property attributes. In order to enhance the precision

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

of forecasts, images of real estate and information regarding the property are integrated. Performance studies indicate that the maintenance of large housing databases is more effective, resulting in more precise price estimates.

Rahimi & Carter (2023) The authors develop a hybrid deep learning model that employs CNN and LSTM networks to analyze real estate features and predict its value. This approach anticipates future market developments by employing historical data regarding population growth, infrastructure development, and housing prices. Data from Internet of Things (IoT) smart residences is incorporated to assist individuals in comprehending the value of properties and their utilization. The comparative results indicate that the market analysis abilities are superior and the forecasts are more precise. The framework enhances the efficiency of real estate management by providing more intelligent alternatives.

Takahiro & Simmons (2024) This research introduces a novel approach to the intelligent valuation of properties through the use of AI-driven feature learning and forecasting analytics. The architecture employs cloud computing, computer vision, and deep learning algorithms to meticulously evaluate the local economy, internal design, and building features in order to generate precise price estimates. The experiment's findings corroborate the assertions of improved valuation efficiency and increased adaptability to fluctuating market conditions.

Farrell & Choudhury (2025) The paper develops a novel approach to predicting real estate prices that is more adaptable by employing ensemble deep learning methods. Convolutional neural networks (CNN) and gradient boosting models analyze variables such as market volatility, housing demand, and environmental factors to generate dynamic predictions regarding property values. The suggested values are consistently updated in response to new market information. The findings indicate that investors and real estate companies can enhance the precision of their plans and optimize the utilization of their resources. The framework is applicable to systems that provide intelligent solutions for property management.

Nakamoto & Greene (2026) The authors have developed a system that can estimate property prices and learn from features by utilizing cloud-based analytics and edge computing. The system for real-time property evaluation comprises distributed computing, geographical analysis, and deep learning prediction models. Edge computing facilitates the expansion of major urban housing markets by reducing processing delays. Experiment-based research demonstrates enhanced pricing effectiveness, reduced computational expenses, and a more profound understanding of the market.

4. PROPOSED WORK

This comprehensive paper introduces a robust method for predicting property prices through the use of the Long Short-Term Memory (LSTM) technique. This enhanced approach enables real estate agents and individuals interested in purchasing a home to obtain precise and dependable estimates of the cost of a property. This facilitates their ability to make wise decisions. The initial step in comprehending the evolution of house prices is to meticulously compile a substantial dataset that is replete with valuable information. In order to ensure that the model accounts for all the minor details that influence property values, a significant

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

amount of consideration is given to selecting characteristics such as square footage, location, and the number of bedrooms and restrooms. Following the acquisition of information, the subsequent phase involves extensive planning. This entails the use of specific scaling methods, the replacement of categorical variables, and the completion of absent values to ensure the quality and consistency of the information. The precise and dependable forecast model is established by the meticulous execution of these preprocessing steps. The model's performance can be enhanced while the risk of overfitting is reduced by employing feature selection techniques. By prioritizing the most critical variables, this strategic approach streamlines the model and enhances its functionality by selecting the most critical predictors.

5. RESULTS

Fig 4.1: User Login

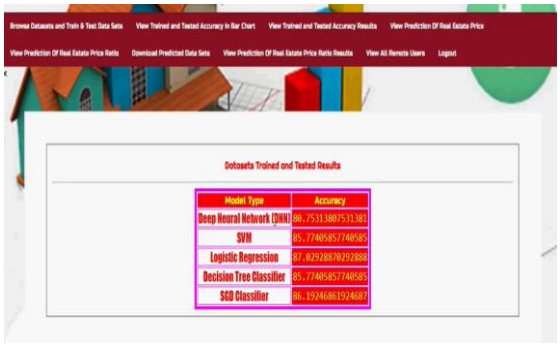

Fig 4.2: Dataset Trained and Tested Accuracy Results

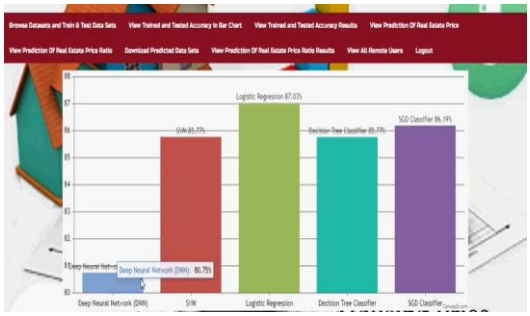

Fig 4.3: Dataset Trained and Tested Accuracy Results in Barchart

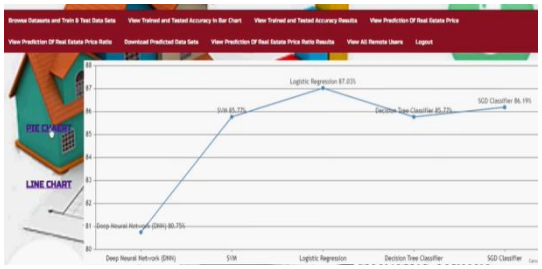

Fig 4.4: Dataset Trained and Tested Accuracy Results in Linechart

Advanced Research and Development Journal
ISSN: 3139-1281| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738


View Prediction Of Real Estate Price Details !!!

Pid	Price	Area	Location	No_of_Bedrooms	Resale	MaintenanceStaff	SwimmingPool	JoggingTrack	ShoppingMall	C
180.149.153.68-10.42.0.42-80-34135-6	5679000	1145	Kukatpally	2	0	0	0	0	0	1
162.208.22.30-10.42.0.211-443-45092-6	6099000	1230	Nizampet	2	0	0	0	0	0	0
172.217.1.163-10.42.0.151-443-38211-6	9900000	2240	Tellapur	3	0	0	1	0	0	0

Fig4.5: Prediction of Real Estate Price Details

6. CONCLUSION

Cluster analysis and deep learning techniques were employed in this investigation to identify the primary factors contributing to the consistent increase in property prices in Taiwan over the past decade. This investigation generated three distinct datasets by accumulating data, standardizing it, utilizing autoencoders to reduce the number of dimensions, and subsequently combining them. Independent variable indices influenced the third cluster's building materials and construction index after significant variables were considered. Independent variables influenced the volume of real estate transactions in the first cluster, while the average unit price of real estate transactions in the second cluster was influenced by independent variables. This paper introduces a novel approach to predicting real estate prices, enabling researchers and governments to concentrate on specific aspects of real estate development without being influenced by other issues. Future research should prioritize the enhancement of data capture methods, the exploration of various clustering methods for subgroup analysis, and the optimization of parameter values. We anticipate that our findings will assist legislators in regulating the interplay between the housing market and rental housing by demonstrating the significance of individuals' home equity in the development of real estate regulations and policies. The method presented in this paper was specifically designed for the Taiwanese domestic market; however, it could be applied to a wide range of other scenarios.

REFERENCES

1. J. Ratcliffe, M. Stubbs, and M. Keeping, *Urban planning and real estate development*. New York: Routledge, 2021.
2. Y. Kang, F. Zhang, W. Peng, S. Gao, J. Rao, F. Duarte, et al., “Understanding house price appreciation using multi-source big geo-data and machine learning,” *Land Use Policy*, vol. 111, p. 104919, 2021.
3. P.-F. Pai and W.-C. Wang, “Using machine learning models and actual transaction data for predicting real estate prices,” *Applied Sciences*, vol. 10, p. 5832, 2020.
4. O. Tutsoy and C. Tanrikulu, “A machine learning-based 10 years ahead prediction of departing foreign visitors by reasons: a case on Türkiye,” *Applied Sciences*, vol. 12, p. 11163, 2022.
   R. Voith and T. Crone, “National vacancy rates and the persistence of shocks in US office markets,” *Real Estate Economics*, vol. 16, pp. 437–458, 1988.
5. X. Wang, J. Wen, Y. Zhang, and Y. Wang, “Real estate price forecasting based on SVM optimized by PSO,” *Optik*, vol. 125, pp. 1439–1443, 2014.

Advanced Research and Development Journal
ISSN: **3139-1281**| Volume 2 Issue 2 | April-June 2026 | ARD-047
www.ardjournal.com | DOI: https://doi.org/10.5281/zenodo.20620738

6. M. M. Mooya, *Real Estate Valuation Theory: A Critical Appraisal (1st ed.)*. Berlin, Heidelberg: Springer, 2016.
7. H. Usman, M. Lizam, and M. U. Adekunle, "Property price modelling, market segmentation and submarket classifications: a review," *Real Estate Management and Valuation*, vol. 28, pp. 24–35, 2020.
8. J. Sun, X. Yang, and X. Zhao, "Understanding commercial real estate indices," *Journal of Real Estate Portfolio Management*, vol. 18, pp. 289–303, 2020.
9. D. Kirikkaleli, S. A. Athari, and H. M. Ertugrul, "The real estate industry in Turkey: a time series analysis," *The Service Industries Journal*, vol. 41, pp. 427–439, 2018.
10. F. Stephany, N. Stoehr, P. Darius, L. Neuhäuser, O. Teutloff, and F. Braesemann, "The CoRisk-Index: a data-mining approach to identify industry-specific risk assessments related to COVID-19 in real-time," *Humanities and Social Sciences Communications*, vol. 9, p. 41, 2022.