International Journal of Information Technology and Applications
Volume 2, Issue 1, March 2025

ISSN 3007-7397

Deep Learning and Knowledge Graph Based Analysis of Property Economic Data in China

Qiaoya Hu*

Foshan University, Foshan, China

792107441@qq.com

Abstract. With the rapid development of China's economy, the real estate market dynamics of Guangdong, Hong Kong and Macao Greater Bay Area (hereinafter referred to as the "Greater Bay Area"), as a national strategic region, have an important impact on the national and global economy. The purpose of this paper is to conduct a comprehensive and in-depth analysis of the real estate economic data of the Greater Bay Area through deep learning and knowledge mapping technologies, and to build a refined data analysis model to dig out the deep laws behind the data, predict market trends, identify investment opportunities, and assess potential risks, so as to provide a scientific basis for governmental decision-making, corporate investment, and academic research.

Keywords: Deep learning, Knowledge graph, , Data analysis, Greater bay area, Real estate economy, House price prediction.

1. Introduction

The Greater Bay Area, which includes nine cities in Guangdong Province and the two special administrative regions of Hong Kong and Macau, is one of the most economically active and open regions in China. Its real estate market is not only concerned with the well-being of people's livelihood, but also an important driver of regional economic development. However, the complexity, volume, and timeliness of real estate economic data pose great challenges to traditional analysis methods. Deep learning and knowledge mapping, as advanced technologies in the field of artificial intelligence, offer new possibilities for deep mining and intelligent analysis of property economic data.

2. Deep Learning Theory Analysis

Deep learning, as one of the most groundbreaking technological directions in the field of artificial intelligence, is centered on the construction of a neural network structure with multiple hidden layers, which realizes deep understanding and intelligent processing of complex data through layer-by-layer feature extraction and representation learning. In the field of real estate economic data analysis, the application of deep learning technology has a unique advantage, which can not only deal with the huge amount of structured and unstructured data in the real estate market, but also automatically learn the potential features and laws embedded in the data through the deep network. Especially when dealing

with high-dimensional non-linear problems such as house price prediction, deep learning models can fully consider the complex association between multi-dimensional factors such as geographic location, housing attributes, surrounding facilities, and transportation conditions, and thus arrive at more accurate prediction results. With the improvement of computing power and optimization of algorithms, the application scenarios of deep learning in real estate market analysis continue to expand, from the initial simple house price prediction to the current market trend analysis, consumer behavioral portrait, investment risk assessment and other aspects, providing a strong technical support for accurate decision-making in the real estate market. It is worth noting that the performance of deep learning models is largely dependent on the quality and quantity of training data, so it is necessary to establish a perfect data collection and pre-processing mechanism in practical applications to ensure that the model can truly reflect the laws of the market.

3. Technical Analysis of the Knowledge Graph

Knowledge graph technology provides a new solution for knowledge management and intelligent services in real estate by building a semantic network of entities, relationships and attributes. The Knowledge Graph uses the triad $G = \{E, R, F\}$ as the unit of knowledge, where, E denotes the set of entities $\{e_1, e_2, \ldots, e_n\}$. Entities e_i are nodes in the Knowledge Graph that can represent concrete entities or abstract knowledge points. R denotes a collection of relations $\{r_1, r_2, \ldots, r_n\}$ Relations r_i denotes edges in the Knowledge Graph, used to represent connections between entities. F represents a set of facts $\{f_1, f_2, \ldots, f_n\}$, a fact f_i can be represented by a concrete small triple $\{h, r, t\}$, where h denotes the head entity, r denotes the relationship, and t denotes the tail entity. The complete knowledge graph is formed by integrating the fact triples.

In the economic analysis of properties in the Greater Bay Area, knowledge mapping can systematically organize and correlate property information scattered in various data sources to form a complete knowledge system. For example, knowledge mapping can clearly show the correlation between a property project and its surrounding supporting facilities, transportation routes, educational resources, and also reflect multi-dimensional information such as historical price changes, developer background, property service evaluation, etc. This structured knowledge representation not only improves the efficiency of information retrieval, but also allows for the discovery of potential market opportunities and risks through knowledge-based reasoning. Another important feature of knowledge mapping is its dynamic updating capability, which can continuously supplement and improve the knowledge base with market changes, ensuring that decision-making is always based on the latest and most comprehensive information. In practical applications, knowledge mapping can also be combined with natural language processing technology to realize intelligent Q&A and decision support functions, providing accurate information services and decision-making suggestions for participants in the real estate market.

4. Acquisition and Preprocessing of Real Estate Economic Data in the Greater

Bay Area

4.1. Data Sources and Collection Methods

Real estate transaction data mainly comes from official statistical databases released by government departments, including information on commercial housing transactions, land transaction records and real estate development investment data released by local housing and construction commissions. Meanwhile, the online platforms of real estate agencies are also an important source of data, which record a large amount of second-hand housing transaction information, rental market dynamics and user browsing behavior data. In addition, unstructured data such as user comments and news reports on social media platforms provide important references for analyzing market sentiment and consumer preferences. In the process of data collection, it is necessary to establish an automated data collection system, regularly obtain updated data through distributed crawler technology, and design corresponding data parsing and storage strategies for the characteristics of different data sources, so as to ensure the real-time and integrity of the data.

4.2. Data cleansing and standardization

The cleaning and standardization of real estate economic data is a key link in ensuring the reliability of subsequent analysis. First, in the quality assessment stage of raw data, a multi-dimensional assessment index system needs to be adopted, including data completeness, accuracy, consistency, timeliness and other aspects, and abnormal patterns existing in the data need to be identified through statistical analysis methods, such as outlier detection, duplicate data identification, missing value analysis and so on. In the data preprocessing stage, the Z-score normalization method can be used to deal with numerical type features (see Equation 1). For the detected problem data, it is necessary to adopt appropriate processing strategies according to different types of problems, for example, for the outliers of numerical-type features, a reasonable threshold range can be set by combining the domain knowledge, and the anomaly detection method based on statistical distribution can be used for screening; for the missing data, it is necessary to choose the appropriate filling method according to the type of missing mechanism, which can be the co-filling based on similar samples or the For missing data, it is necessary to choose a suitable filling method based on the type of missing mechanism, which can be based on the collaborative filling of similar samples or the interpolation prediction based on temporal characteristics, to ensure that the filled data meets both statistical laws and business logic. In data standardization, it is necessary to formulate unified data coding specifications and measurement standards, which not only include basic unit conversion (such as the area is unified into square meters, and the price is unified to the same pricing basis), but also need to take into account the differences between different regions, and establish a cross-regional data comparable system, especially when dealing with geographic location information, it is necessary to establish unified rules for the conversion of the coordinate system, to ensure the accuracy and consistency of the spatial data. and consistency.

$$Z = \frac{x-\mu}{\sigma} \tag{1}$$

where x is the original value, μ is the mean and σ is the standard deviation.

4.3. Data feature extraction

In feature extraction, it is necessary to construct the feature system from multiple dimensions, firstly, the basic attribute features of the property project, including hard indicators such as building area, household structure, building age, decoration standard, etc. These features need to be reasonably binned and coded to better capture the nonlinear relationship between the features and the target variables. Next are location-related derived features, which require spatial analysis techniques to calculate the distance relationship between the project and various important facilities, such as accessibility indicators to subway stations, schools, hospitals, and commercial centers, and to take into account changes in the time dimension, such as the extent to which future planned facilities will affect the current property value. When constructing market dynamics features, it is necessary to extract price trend features from historical transaction data, including indicators such as regional price growth rate, volatility, transaction activity, etc. Meanwhile, market sentiment indicators are constructed by combining social media data, and market expectations are quantified by analyzing the emotional tendency of user comments through natural language processing technology. In terms of policy-related features, it is necessary to extract policy-oriented features from policy texts, build a policy strength index, and assess the potential impact of policies on the market.

5. Deep Learning-based Model for Analyzing Real Estate Economic Data

5.1. Model Architecture Design

In the analysis of property and economic data in the Greater Bay Area, the architectural design of the deep learning model needs to take full account of the data characteristics and analysis objectives. For the core task of house price prediction, a hierarchical neural network structure can be constructed, using a multilayer perceptron network as the infrastructure in the bottom layer, and gradually extracting the nonlinear correlation relationship between property features through several well-designed fully-connected layers, with the number of neurons in each layer optimally configured according to the dimensionality and complexity of the input features, and a pyramid-type structural design is usually adopted, i.e., the number of neurons gradually decreases as the layers deepen to achieve effective compression and abstraction of features. deepening of the level, the number of neurons is gradually reduced to achieve effective compression and abstraction of features. Considering the significant temporal characteristics of property data, it is especially important to introduce the long and short-term memory network module in the model. This recursive neural network structure can effectively capture the long-term dependency of house price changes, and filter and retain important historical information through the gating mechanism. When dealing with geographic location-related features, the convolutional neural network module can be introduced to extract features from spatial data through the sliding window mechanism, effectively capturing regional development features and spatial dependencies. The architecture of the deep learning model constructed based on the above design ideas

is shown in Figure 1, which demonstrates the core components of the model, including the input layer, the LSTM layer, the CNN layer and the output layer, and clearly represents the processing flow of the data in the model. In order to improve the robustness and training effect of the model, a batch normalization layer and a random deactivation layer are added between the key layers of the network, which not only accelerates the convergence process of the model, but also effectively prevents the occurrence of overfitting phenomenon. The Long Short-Term Memory (LSTM) network units are updated as follows:

$$\begin{cases} f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t = f_t{*}C_{t-1} + i_t{*}\tilde{C}_t \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t{*}\, tanh(C_t) \end{cases} \tag{2}$$

where f_t, i_t, o_t is the activation vector of the forget gate, input gate and output gate respectively, C_t is the cell state, h_t is the hidden state, σ is the sigmoid function, and $*$ denotes the element-by-element multiplication. CNN convolutional operations:

$$(V{*}X)_{ij} = \sum_m \sum_n V_{mn} X_{i+m,j+n} \tag{3}$$

where V is the convolution kernel, X is the input feature map and i, j is the coordinates of the output feature map.

The output layer is usually a fully connected layer that maps features to the final output space, the computation of which can be expressed as:

$$\hat{y} = softmax(W_{out}H + b_{out}) \tag{4}$$

where $\hat{y}$ is the predicted output of the model, W_{out} is the weight matrix of the output layer, b_{out} is the bias term, and H is the output of the previous layer.

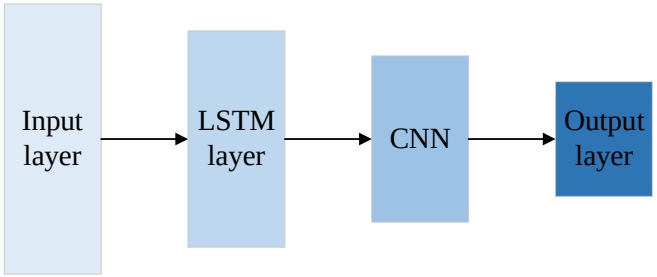

Figure 1. Deep learning model architecture diagram.

5.2. Key Algorithms and Implementation

The training process of the deep learning model involves the synergistic cooperation of several key algorithms. In the choice of optimization algorithm, considering the complexity and non-smooth

characteristics of the property data, the adaptive moment estimation algorithm is a more reasonable choice, which is able to automatically adjust the learning rate of each parameter according to the gradient information, adaptively balancing the training speed and stability during the training process, and is particularly suitable for dealing with sparse data and non-smooth data. Adaptive Moment Estimation (Adam) optimization algorithm update rule:

$$\begin{cases} m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla J(\theta_{t-1}) \\ v_t = \beta_2 v_{t-1} + (1-\beta_2)\left(\nabla J(\theta_{t-1})\right)^2 \\ \widehat{m}_t = \frac{m_t}{1-\beta_1^t} \\ \hat{v}_t = \frac{v_t}{1-\beta_2^t} \\ \theta_t = \theta_{t-1} - \alpha \frac{\widehat{m}_t}{\sqrt{\hat{v}_t}+\int} \end{cases} \tag{5}$$

where m_t and v_t are the first-order and second-order moment estimates, respectively. β_1, β_2 are the decay rates, α is the learning rate, and ò is the numerical stability term.

The design of the loss function needs to be customized according to the characteristics and requirements of the specific task, and for the house price prediction task, a composite loss function based on mean square error can be constructed.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the sample size.

In addition to the basic prediction error term, a constraint term on the distribution of the predicted values is added to ensure that the prediction results are consistent with the actual market laws. In order to improve the generalization ability of the model, a variety of regularization techniques need to be introduced, and this paper uses weight decay to control the model complexity:

$$L = J(\theta) + \frac{\lambda}{2}\sum_i \theta_i^2 \tag{7}$$

where $J(\theta)$ is the loss function, λ is the regularization coefficient, and θ_i is the model parameters.

The early stopping strategy is used to avoid overfitting, while the adversarial training technique is introduced to enhance the robustness of the model to noise. In the actual implementation process, it is necessary to make full use of parallel computing resources to improve the training efficiency by means of data parallelism or model parallelism, especially when dealing with large-scale datasets, it is necessary to design efficient data loading and preprocessing processes to ensure the full use of computing resources. The optimization of the discriminator is as follows:

$$\min_D \max_G V(D,G) = \mathbb{E}_{x\sim p_{data}}[log\, D\,(x)] + \mathbb{E}_{z\sim p_z}[log(1 - D(G(z)))] \tag{8}$$

where p_{data} is the distribution of the real data, z is a random variable obeying the prior distribution p_z ($z \sim N(0,1)$), D is the discriminator function, and G is the generator function. The generator is optimized as follows:

$$\min_G V(G, D^*) = \mathbb{E}_{z \sim p_z}\left[log\left(1 - D^*(G(z))\right)\right] \tag{9}$$

where D^* is the optimal discriminator.

5.3. Model Training and Optimization

Model training is a process that requires continuous optimization and adjustment. First of all, in the stage of dataset division, it is necessary to give full consideration to the characteristics of time series data, and use the time sliding window to divide the data, to ensure the consistency of the distribution of the training set, validation set, and test set, and also to consider the balance of samples from different cities and different types of properties. During the training process, it is necessary to monitor the trend of multiple key indicators in real time, including the training loss, validation loss, and various types of evaluation indicators, etc. The training process is intuitively demonstrated through visualization tools, so as to identify and solve problems in training in a timely manner. The changes of training loss are as follows:

$$L_{train} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{10}$$

where L_{train} denotes the training loss, N is the number of samples, y_i is the true value while $\hat{y}_i$ is the predicted value.

The tuning of hyperparameters is a key aspect to improve the performance of the model, and the optimal parameter combinations can be searched automatically using Bayesian optimization, which is more efficient compared to the traditional grid search, and is able to intelligently select the next set of parameters to be evaluated based on the results of the historical attempts. The posterior probability update in Bayesian optimization is as follows:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{11}$$

where $P(\theta|D)$ is the posterior probability of the parameter θ given the data D, $P(D|\theta)$ is the likelihood function, and $P(\theta)$ is the prior probability of the parameter θ.

For the data characteristics of different cities in the Greater Bay Area, a migration learning strategy can be used to migrate the model trained in one city to other cities to improve the adaptability of the model. The weight adjustment in migration learning is as follows:

$$W_{new} = \alpha W_{old} + (1 - \alpha)W_{scratch} \tag{12}$$

where W_{new} is the new weights after migration learning, W_{old} is the weights of the source task, $W_{scratch}$ is the weights trained from zero, and α is a hyperparameter between 0 and 1 to control the mixing ratio of the source task weights and the new task weights.

In addition, a model update mechanism should be put in place to retrain the model with new data on a regular basis to ensure that the forecasts always reflect the latest market conditions.

6. Construction of the Greater Bay Area Real Estate Economy Knowledge Graph

6.1. Ontology Design and Construction

The ontology design of the Greater Bay Area Property Economy Knowledge Graph needs to comprehensively cover the core concepts and relationships in the real estate domain. Firstly, it is necessary to define the basic entity types, including property projects, developers, property companies, geographic areas, supporting facilities, etc., and design detailed attribute characteristics for each type of entity. In terms of relationship design, it is necessary to clarify various types of associations between entities, such as the development relationship between developers and property projects, the spatial relationship between property projects and supporting facilities, and the administrative affiliation between different regions. The hierarchical structure of the ontology should fully reflect the system of professional knowledge in the field of real estate, and a clear conceptual classification system should be constructed through the inheritance relationship between subclasses and parent classes. In the choice of ontology description language, it is recommended to adopt a language with strong expressive ability, such as the web ontology language, which not only describes complex conceptual relationships, but also supports logical reasoning and constraint definitions, providing a basis for subsequent knowledge reasoning and querying.

6.2. Entity Relationship Extraction

Entity-relationship extraction is a key aspect of knowledge graph construction in the real estate and economic domain of the Greater Bay Area. For structured data sources, entities and relationships can be extracted directly by rule mapping. For unstructured texts such as news reports and policy documents, deep learning named entity recognition techniques need to be applied to recognize entities such as property project names, developer names, and place names in the text. In terms of relationship extraction, a remote supervision method can be used to utilize known entity relationships as training data, and the learning extraction model automatically discovers new entity relationships [1]. Especially for real estate policy documents, it is necessary to design a specialized policy element extraction model to accurately capture the relationships among policy subjects, policy objects and specific measures. In terms of entity disambiguation, it is necessary to consider issues such as aliases of real estate projects, developer abbreviations, etc., and to establish an entity linking mechanism to ensure the consistency of knowledge.

6.3. Knowledge Reasoning and Integration

Knowledge reasoning is an important means of mining implicit knowledge. In the Greater Bay Area Real Estate Economy Knowledge Graph, simple transfer reasoning can be performed based on ontology rules, such as deducing the city to which the project belongs through the relationship between the project location and administrative division. For complex inference tasks, statistical learning methods can be combined, such as predicting possible relationships between entities through path ordering algorithms. Knowledge fusion, on the other hand, needs to deal with the conflict and complementarity of knowledge from different sources, and filter and integrate conflicting knowledge by designing a credibility assessment mechanism. In the fusion process, special attention should be paid to the problem of timeliness, and a knowledge updating mechanism should be established to delete outdated knowledge and supplement new knowledge in a timely manner. In addition, the issue of multi-language representation of knowledge should be considered to support the alignment and interoperability of multi-language knowledge, such as Cantonese and English, to meet the needs of the multi-language environment in the Greater Bay Area.

7. Integration of Deep Learning and Knowledge Graph Applications

7.1. House Price Forecast and Trend Analysis

In the overall fusion architecture of Deep Learning and Knowledge Graph (Figure 1), Deep Learning is responsible for learning complex feature patterns and prediction laws from massive data, while Knowledge Graph provides structured domain knowledge support, and the combination of the two is able to produce more comprehensive and accurate analysis results. In the analysis of property market forecasting and trends in the Greater Bay Area, the application of this method can simultaneously utilize the statistical laws of data and the logical relationships of domain knowledge. The property knowledge network constructed by knowledge mapping can provide all-round access to the multi-dimensional factors affecting house prices, which include not only the static attributes of the property project (e.g., basic features such as floor area, house type structure, decoration standards, building age, etc.), but also the location This includes not only the static attributes of the property project (such as basic features like floor area, house type structure, decoration standard, building age, etc.), but also the location (such as spatial features like distance to subway stations, supporting school districts, and coverage of commercial facilities, etc.), as well as the dynamic market information (such as changes in supply and demand, policy regulation, regional planning adjustments, and other temporal and serial features). The deep learning model, on the other hand, builds a prediction model on the basis of these rich features by means of complex nonlinear mapping relationships. Especially when dealing with time-series prediction, it can incorporate the attention mechanism to automatically identify the influence weights of different periods and factors on the trend of house prices, so as to improve the accuracy and interpretability of the prediction [2]. In terms of market trend analysis, the fusion model can make multi-scenario forecasts of future market trends by analyzing historical data patterns in the knowledge graph and the current market state, combining the predictive capability of deep learning, and explaining the key factors that may lead to different trends through the knowledge-based reasoning mechanism, so as to provide a more

comprehensive reference basis for investment decisions.

Table 1. The results of the four models in predicting and analyzing house prices.

Evaluation indicators	MSE	MAE
Random forest	0.0135	0.0118
Support vector machine	0.014 2	0.0115
Deep learning	0.0123	0.0099
Deep Learning + Knowledge Graph	0.0087	0.0075

Finally, this section evaluates the effectiveness of the architecture in forecasting and trend analysis of the property market in the Greater Bay Area through experiments using Mean Square Error (MSE) and Mean Absolute Error (MAE) to assess forecasting accuracy. The experimental results are shown in Table 1:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{14}$$

Experimental results show that the fusion model significantly outperforms the deep learning model alone in terms of prediction accuracy, while providing good model interpretability through the attention mechanism.

7.2. Investment Opportunity Identification and Risk Assessment

In the field of property investment opportunity identification and risk assessment in the Greater Bay Area, the fusion of deep learning and knowledge graph analysis methods provides an all-encompassing decision support system. Through knowledge mapping technology, the knowledge environment of the project to be evaluated can be quickly constructed, which includes a knowledge network composed of information in multiple dimensions, such as the basic information of the project itself, the background of the developer, the development plan of the region in which it is located, the situation of the neighboring projects under construction, and the layout of future infrastructure. The deep learning model, on the other hand, can be based on historical investment case data to learn the characteristic patterns of successful investment projects and construct an investment value assessment model. This model not only takes into account the traditional price-return indicators, the net present value (NPV) is as follows:

$$NPV = \sum_{t=0}^{n}\frac{R_t}{(1+r)^t} - C \tag{15}$$

where R_t is the cash flow at time t, r is the discount rate, C is the initial investment cost, and n is the duration of the project.

At the same time, comprehensive factors such as regional development potential, policy support, and market competition pattern should also be included [3]. In terms of risk assessment, the knowledge graph can sort out various types of risk factors related to the project through relational network analysis, including the qualification risk of the developer, the policy risk of the nature of the land, and the operational risk of market competition, etc., while the deep learning model quantifies the degree of the impact of these risk factors through the learning of historical data and generates a dynamic risk scoring system. The risk score S can be expressed as:

$$S = \sum_{i=1}^{m} w_i \cdot R_i \tag{16}$$

where w_i is the weight of the i-th risk factor,R_i is the score of the i-th risk factor, and m is the total number of risk factors.

This fusion analysis can not only identify potential investment opportunities in a timely manner, but also warn of possible investment risks and provide investors with comprehensive decision support [4], with specific results shown in Figure 2.

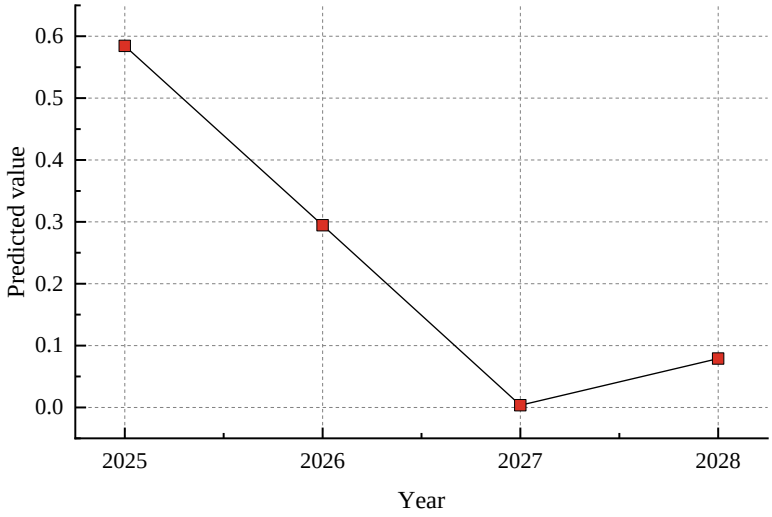

Figure. Investment risk projection map.

As can be seen from the figure, there is a greater investment risk in 2025, which is mainly due to the heightened downside risk of the global economy, which may be characterized by a simultaneous fall in aggregate demand and aggregate supply. Subsequently, the risk decreases year by year and reaches a minimum in 2027, thanks to the gradual recovery of the global economy and the positive adjustment of policies. However, as protectionism becomes the biggest risk to the growth of global trade and the pace of interest rate cuts by the Federal Reserve faces uncertainty, the investment risk tends to pick up in 2028. Therefore, diversified investment strategies should be adopted to spread risks, and changes in global monetary policies should be closely monitored, risk management measures should be put in place, and a risk reserve should be established to cope with potential market volatility, so as to ensure the safety and stability of investments.

7.3. Intelligent Advisory Services and Decision Support

In the Greater Bay Area real estate market, the fusion of deep learning and knowledge graph can also provide market participants with intelligent consulting services and decision-making support. This fusion application is first reflected in the construction of an intelligent Q&A system. Through the structured knowledge system provided by the knowledge graph, the system is able to accurately understand the user's intention to consult and, based on the natural language processing technology of deep learning, transform complex professional questions into query statements of the knowledge graph. For example, when a user inquires about the investment value of a certain region, the system can automatically analyze various related information of the region in the knowledge graph, including planning policies, transportation construction, education and medical care, and other supporting facilities, and combined with the results of the deep learning model's analysis of the historical data, it will give an investment proposal that meets the user's specific needs. In terms of decision support, the convergence system can provide more personalized solutions. Through the deep learning analysis of the user's historical interaction data, the system can build a user profile to understand their investment preferences, risk tolerance and investment objectives. At the same time, with the reasoning ability of knowledge graph, the system can recommend the most suitable investment portfolio in combination with the user profile and provide detailed decision-making basis. This recommendation not only takes into account the traditional return on investment, but also evaluates the match between the project and the user's specific needs, such as commuting convenience, accessibility to educational resources and other personalized factors.

8. Conclusion

The analysis of property economic data in the Greater Bay Area based on deep learning and knowledge mapping provides new perspectives and methods for research and decision-making in the real estate market. By deeply mining the laws behind the data, predicting the market trends, identifying investment opportunities, and evaluating the potential risks, it can provide more accurate and comprehensive decision-making support for the government, enterprises, and investors. In the future, with the continuous progress of technology and the increasing abundance of data, the application of deep learning and knowledge graph in real estate economic data analysis will have a broader prospect.

References

[1] H. S. Munawar, S. Qayyum, F. Ullah, and S. Sepasgozar, “Big data and its applications in smart real estate and the disaster management life cycle: A systematic analysis,” Big Data and Cognitive Computing, vol. 4, no. 2. 4, 2020.
[2] Z. Jiang, “A Novel Method of Data Element Trading and Asset Value Appreciation,” Highlights in Business, Economics and Management, vol. 16, pp. 576-583, 2023.
[3] S. Guo, and R. Zhai, “E-commerce precision marketing and consumer behavior models based on IoT clustering algorithm,” Journal of Cases on Information Technology (JCIT), vol. 24, no. 5, pp. 1-21, 2022.
[4] Y. Fan, Z. Yang, and A. Yavas, “Understanding real estate price dynamics: The case of housing prices in five major cities of China,” Journal of Housing Economics, vol. 43, pp. 37-55, 2019.