

ISSN 1840-4855

e-ISSN 2233-0046

Original scientific article

<http://dx.doi.org/10.70102/afts.2026.1836.560>

PREDICTING INVESTMENT PREFERENCES FOR GOLD AND REAL ESTATE AMONG INDIAN SALARIED EMPLOYEES USING GRAPH NEURAL NETWORKS

Twinkle¹, S. Nazim Sha^{2*}

¹Research Scholar, Faculty of Management, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, Tamil Nadu, India. e-mail: tv6112@srmist.edu.in, orcid: <https://orcid.org/0009-0007-8869-4930>

^{2*}Assistant Professor, Faculty of Management, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, Tamil Nadu, India. e-mail: nazims@srmist.edu.in, orcid: <https://orcid.org/0000-0001-5769-067X>

Received: May 25, 2026; Revised: July 03, 2026; Accepted: August 14, 2026; Published: September 29, 2026

SUMMARY

The decisions made in investing among salaried individuals are guided by financial objectives, risk appetite, and demographics, with gold and real estate continuing to be some of the most favored ones in India. In this study, investment preference for gold and real estate among salaried individuals in India and a prediction model for predicting preferred investment types based on demographics are explored. Data was obtained using a questionnaire that was filled out online through Google Forms by 300 salaried individuals from different sectors. Label correlation guided borderline oversampling (LCGBO) was used to solve class imbalance, and an explicit feature interaction-aware graph neural network (EFIGNN) was developed to predict investment preference. One-way ANOVA, T-test, and Chi-square were used to determine the influence of age, income level, and gender on investment choice. According to ANOVA results, there was a significant age effect on the preference for real estate ($p = 0.012$) and gold ($p = 0.021$). On the other hand, the Chi-square showed a significant relationship between income and investment preference ($\chi^2 = 8.13$ for gold and $\chi^2 = 2.95$ for real estate). T-test showed a significant difference in gender ($p = 0.001$ for real estate and $p = 0.005$ for gold). EFIGNN achieved 98% accuracy and an RMSE of 0.1%, outperforming ANN, LSTM, and GNN. Real estate emerged as the most preferred investment, with demographic factors significantly shaping investment behavior and EFIGNN demonstrating strong predictive performance for investment preference modeling.

Key words: gold investment, real estate investment, graph neural network, investment prediction, demographic factors, salaried individuals.

INTRODUCTION

Saving is appreciated by every citizen of the world, and it also has a great influence on the way the economy of a country is formed. Saving part of your income is a universal dream, as people hope to get a better payoff and get financial security in case of any unpredictable problems and future ambitions [1]. Beating inflation is one of the reasons why it is best to invest, and as the purchasing power of money is eroded over time, there is a need to acquire high-quality goods and services with the same amount of money [2]. When making opinions about investments, it is important to take into account such variables

as the potential of returns, the time of lock-in, risk tolerance, and flexibility of investments. Those investors who ignore these aspects risk making huge losses and unexpected risks [3]. Two of the traditional and popular investment opportunities in India are gold and real estate. Gold, being the most firmly rooted symbolically and emotionally in Indian culture, is an adaptable investment choice [4]. Its flexibility enables people to invest depending on their financial ability and investment interest, so it is accessible to a very wide group of investors [5]. Gold is simple to purchase in small amounts, and its liquidity allows it to be converted to cash comparatively fast in case of necessity [6]. Moreover, gold is a safe haven during unpredictable economic times, and it has been used as a hedge against inflation and currency variation. However, investments in real estate are believed to be larger and demand a bigger starting capital [7]. The real estate is a secure long-term investment regardless of increased prices. There are some benefits offered by real estate including tax advantages, diversified investment, and the opportunity for rental income. Besides, real estate is an investment that can grow over time [8]. Furthermore, owning property gives a person a movable asset and constant source of income through rent. In general, deciding on whether to invest in real estate or gold depends on personal financial objectives and preferences [9]. Real estate offers long-term stability and income opportunities, but gold provides mobility and liquidity. These are both strong investments which can complement each other within a balanced portfolio [10].

However, the complexities of gold investment are different from those involved with real estate investment among individuals who earn salaries. While the acquisition of gold is relatively simple and cheap to start with, the value of gold can be easily influenced by market speculations and appreciates highly. Alternatively, real estate investment is characterized by expensive entry costs and maintenance costs. It is less liquid and involves complicated legal procedures, but it is expected that in the long-run it offers stability and rental incomes. Besides, the gold investment may involve storage and security costs, while real estate investment is associated with regulation and taxation issues. A balance between all these factors will be necessary to make an informed decision based on critical analysis of personal commercial intentions and risk appetite.

- Which demographic factors like age, income, and gender play a vital role in the choice between gold and real estate investment among salaried Indians?
- How accurately does EFIGNN predict investment choice against other methods?

Research Objectives

The objectives of this research include the study of the financial decisions taken by the salaried individuals of Chennai concerning gold and property investments, along with other related factors like financial decisions, tax planning, and investments.

The above objectives of the study give insight into the following points:

- a) Study the financial management approaches adopted by salaried employees of Chennai and how investment decisions are made.
- b) Make comparisons between the investment choices in real estate and gold by them.

Summary of Contributions of this Research Work:

- Investment preferences of salaried individuals in Chennai, particularly gold and property.
- A new EFIAGNN approach used for enhancing investment preference prediction.
- The research work makes use of a questionnaire survey done through Google forms to collect the required information from 300 salaried individuals in Chennai.
- There is a combination of different statistics, such as One-Way ANOVA, T-Test, and Chi-Square in order to examine possible relations between demographic attributes (age, income, gender) and investments.

- For the evaluation of the proposed EFIGANN method there was carried out a comparative analysis with other methods that existed before.

The rest of the paper is structured in the following way. Literature review will be discussed in Section 2. The section will include the discussion of current research and research gap. Section 3 will include Research Methodology and will describe data collection, oversampling with LCGBO and new EFIGANN model. Results will be presented in Section 4 and will discuss demographic analysis and statistical analysis. Discussion will be provided in Section 5 and will include practical application, limitations and future scope of the research. Finally, Section 6 will be devoted to the Conclusion.

LITERATURE REVIEW

Asemi et al., (2023) [11] have presented research unveiling the influence of managerial traits on investor decision prediction. The "adaptive neuro-fuzzy inference system" (ANFIS) to create a personalised investment recommendation system to examine the influence of managerial qualities on investment decisions was used. Information on the investment kinds, habits, and management attributes of potential investors was obtained through a survey. An ANFIS technique was created using survey data and was a hybrid model that included the benefits of fuzzy logic systems and ANNs. A validation set was used to assess the model's performance after it was trained on 1542 survey data pairs.

Gu et al., (2022) [12] have presented an empirical study for investment method under the VMD-LSTM technique. For a period of data analysis, the K-Nearest Neighbour, Artificial Neural Network, Grey, and LSTM (Long Short-Term Memory Networks) prediction models were compared. In order to qualitatively analyse the market price, it then introduces a few indicators, including the risk coefficient and the convergence-divergence ratio.

Li et al., (2024) [13] have presented graph neural networks with deep mutual learning for designing multi-modal recommendation systems. A brand-new multi-modal recommendation system that combines deep mutual learning methods with GNNs was introduced. It cooperatively trains several uni-modal bipartite user-item graphs using the mutual knowledge distillation technique. To produce unit-modal embeddings, each GNN was explicitly trained on the uni-modal user-item bipartite graph, which was isolated from the original multi-modal user-item bipartite network [20].

Setty et al., (2024) [14] presented a cost-sensitive machine learning to support start up investment decisions. Using various machine learning methods, the empirical analyses forecast the performance of 10,000 Israeli businesses. Models are transformed into cost-sensitive versions using the MetaCost algorithm, which minimises total cost rather than total error.

Hu et al., (2023) [15] have presented an Investor preference analysis for an online optimization method with missing data. In order to ascertain investors' preferences for more commercial product suggestions, an online optimisation methodology was utilized. The suggested approach addresses the scenario in which there were numerous missing values in investors' records and updates the investor's favourite for recently arrived data sets.

Fossen et al., (2023) [16] have presented risk preference and entrepreneurial investment at the top of the wealth distribution [. Included in the statistics were typical measures of risk tolerance and balance sheets of private wealth, specifically the worth of personal private company assets.

Nyhus et al., (2024) [17] examined the mediating effects of personality factors and financial overconfidence in investigating gender variations in crypto investing intentions. Regulations to shield consumers from less-than-ideal choices were informed by research on the effects of consumer gender, financial overconfidence, and the Big Five personality factors on cryptocurrency investing intentions.

There are various gaps in the current body of research on investment preferences. One method may ignore the impact of various investor preferences and outside variables, even though it makes use of ANFIS to forecast investor decisions based on management attributes. A VMD-LSTM model was used

in another study to predict investments; however, it may not completely account for complicated feature interactions or recent advances in machine learning. Furthermore, while research on graph neural network-based multi-modal recommendation systems offers insightful information, it might not directly address the unique dynamics of investing preferences. Models that incorporate a wider range of variables, take advantage of the most recent technological advancements, and provide deeper insights into the unique behaviours of individual investors are therefore required. Addressing these gaps motivated the development of a model that incorporates EFIAGNN method. It improves representation learning and performance by modelling complicated feature interactions effectively. Better predictions result from its ability to incorporate both local and global interdependence, particularly in cases using structured data. The EFIAGNN framework can be applied in various contexts including graph-based recommendation systems, link prediction, and node classification.

From the above literature review, it is clear that most of the research carried out previously was based on traditional machine learning and deep learning approaches such as ANFIS, VMD-LSTM, and classical GNN-based recommendation systems, without much attention being paid to the modelling of interactions between features to predict the investment preference. Prior researches also did not consider the role of demographic factors in making decisions on specific asset classes such as gold and property. There is clearly a research gap here that this study attempts to fill by applying the framework of EFIAGNN.

RESEARCH METHODOLOGY

In the proposed system, predicting investment preferences using the EFIAGNN method is discussed. Data were collected, preprocessed, and interpreted for further analysis. The final step involves employing the EFIAGNN method to predict investment preferences based on various factors, including age, income, and gender. The block diagram of the proposed EFIAGNN is given below in figure 1.

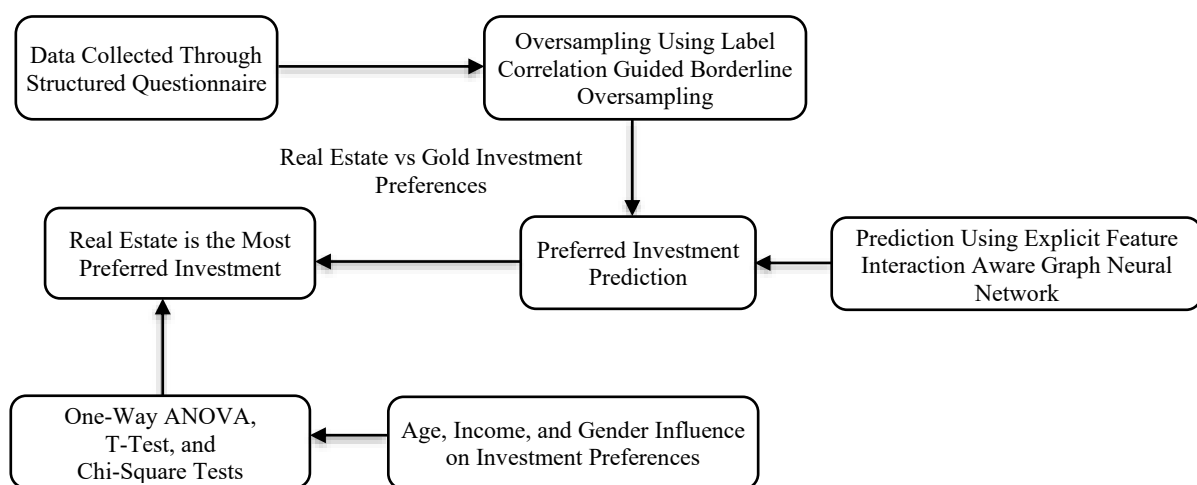


Figure 1. EFIAGNN framework for investment preference prediction

Data Acquisition

An organized survey was used to collect primary data through a questionnaire method. The sample population consisted of salaried employees, with a total of 300 samples collected in Chennai, Tamil Nadu. The questionnaire aimed to assess the general capabilities and technical skills of wellbeing experts by a five-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5). Data were gathered once between January and May 2024. To ensure the reliability and accuracy of the survey, a pilot study with a model size of 100 was conducted, leading to adjustments in the survey, which are given in table 1. For data analysis, statistical tools like descriptive data, inferential data, and reliability testing were employed to evaluate and interpret the responses.

Stratified random sampling method was used for selecting the sample since it would enable proper representation among the various demographic categories like age, gender, income and employment sector. The respondents that would form the sample included salaried employees from Chennai since the population under consideration includes people whose decisions of investing are based on stable and regular income. Chennai has been selected as the research location due to the diversified economic environment which includes industries like IT, health, educational sector and government sector. The inclusion criteria required respondents to be salaried employees having at least one year of work experience as the sample would be exposed enough for making financial decisions. Those who were not in the employment of an organization as well as retired and self-employed individuals were excluded since their behavior in terms of investments and income is quite different from that of the salaried employees. A sample size of 300 was found to be adequate for the test statistics used (ANOVA, T-test, Chi-square). In addition, a pilot study with 100 respondents was done to test the validity of the questionnaire through Cronbach's Alpha (As shown in table 1).

Table 1. Cronbach's alpha test

Cronbach's Alpha	Total Items
0.973 and 0.823	100

Oversampling Using Label Correlation Guided Borderline Oversampling (LCGBO)

In this sector, oversampling of the input dataset using LCGBO [18] is discussed. By carefully positioning fresh data points in high-informative locations, LCGBO reduces noise in contrast to typical oversampling, enhancing model generalization. LCGBO guarantees superior feature representation by preserving the structural integrity of minority classes, which enhances prediction accuracy and stability in investment preference modelling. The minority and majority investment classes are identified by analyzing the dataset's class distribution, ensuring that underrepresented investment preferences, such as gold, are detected for targeted oversampling in equation (1):

$$\hat{A}^{(b)} = \sigma(CD + E\hat{A}^{(b-l)}) \quad (1)$$

Here, $\hat{A}^{(b)}$ indicates the minority class, C represents the count of categories, b indicates the samples and E indicates the features. The label correlation matrix is computed by analysing the relationships between investment preferences, such as gold and real estate, to identify the strength of correlation and guide the over sampling process as given in equation (2)

$$n_p^o(x_i) = \frac{\text{dist}(x_i, s_i^p)}{\text{dist}(x_i, s_i^p)} \quad (2)$$

In this case, x and s are the category counts of two features, $n_p^o(x_i)$ is the correlation matrix. The synthetic samples for the minority class are generated by interpolating between existing instances near the decision boundary, ensuring a more balanced representation of investment preferences in the dataset in equation (3). The generated synthetic samples into the original dataset, enhancing the minority class representation and ensuring a balanced distribution of investment preferences for training the EFIGANN model as given in equation (4)

$$G^O = |B^O| \times p \quad (3)$$

$$G^I = |\bar{w}^I > 0| \quad (4)$$

Where, p is the nearest neighbour, B^O represents the class label, $\bar{w}^I$ indicates the target variable. Finally, LCGBO effectively addresses class imbalance by generating informative synthetic samples, thereby enhancing the model's capacity to predict investment preferences with higher accuracy and robustness.

Prediction Using Explicit Feature Interaction Aware Graph Neural Network

Explainability of the Model: In order to explain from a financial point of view, the feature importance can be calculated using the permutation or SHAP values for the EFIGANN model. This will help in determining how demographics (age, income, gender, and savings rate) affect the preference for gold vs. real estate.

In this section Prediction of the EFIGANN approach is described. EFIGANN is used to predict investment preferences such as gold or real estate. By explicitly modelling interactions between features, it can capture complicated relationships between them efficiently, which results in more accurate predictions. EFIGANN is the best at processing high-dimensional, diversified data, which makes it perfect for deciphering complex investing behaviours. Additionally, its prediction performance is improved by its capacity to combine both organised and unstructured data.

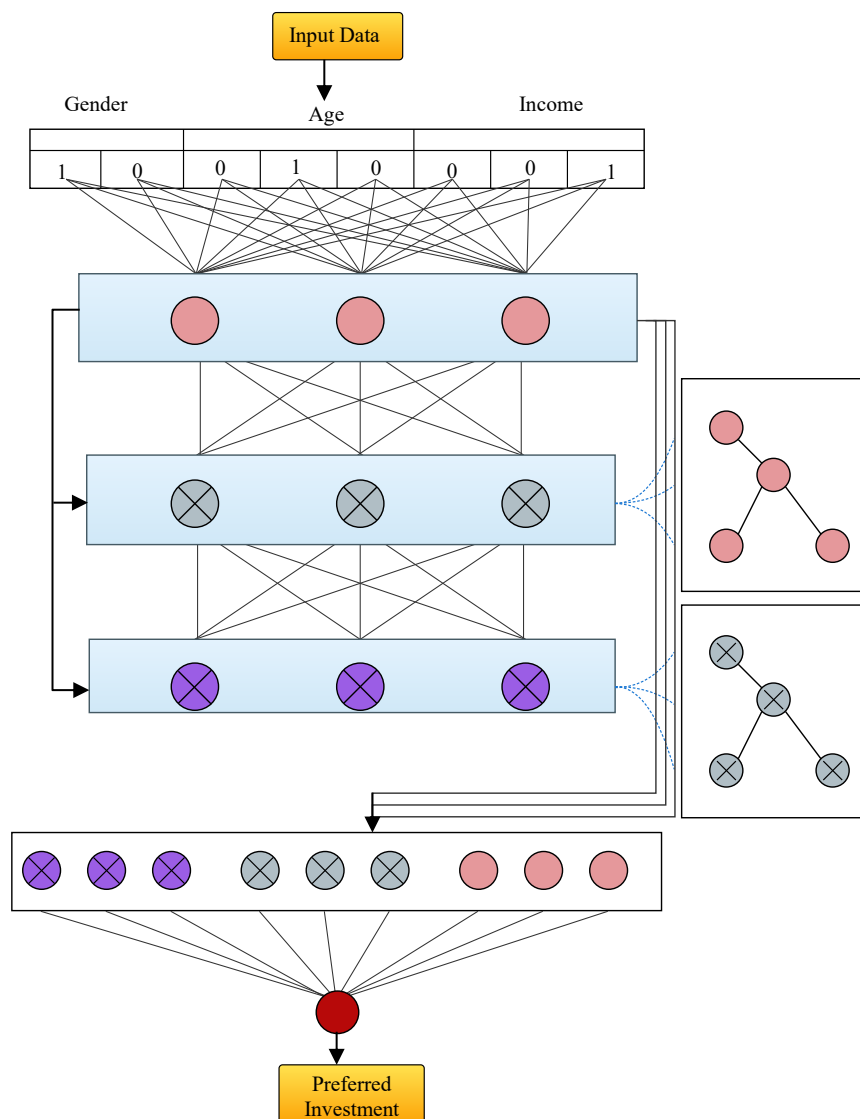


Figure 2. Architecture of EFIGANN

Figure 2 illustrates the EFIGANN architecture for predicting preferred investments based on input data like gender, age, and income. The model uses fully connected layers to work with this input data. These processed features are then fed to independent embedding layers, which may encode the multiple relationships in each set of features. The results of the embedding layers are then aggregated and then into a last layer that forecasts the desirable investment. This architecture implies a multi-layer perceptron model with dedicated embedding modules to extract features and make individual

predictions. The normalisation assures that the features are on the same scale in equation (5). This is essential to EFIGANN, because the neural networks are sensitive to the sizes of input features.

$$Y^l = Y^{(0)}V^{(l)} \quad (5)$$

Where Y^l signifies the linear conversion of input features, $Y^{(0)}$ denotes raw feature matrix, $V^{(l)}$ indicates trainable weight matrix. Categorical variables such as gender and income groups are converted into continuous embeddings to enable the model to learn deeper associations as given in equation (6).

$$Y^l = \hat{B}Y^{(l-1)}V^{(l)}\theta Y^{(l)} \quad (6)$$

Where Y^l stands for the linear transformation, $\hat{B}$ represents the adjacency matrix, $Y^{(l-1)}$ is the feature matrix, θ indicates the Hadamard product. A graph is constructed where each node represents a sample from the dataset, and edges represent relationships between features. At each layer of the EFIGANN, interactions between features are explicitly modeled. It is given in equation (7)

$$\hat{X} = [Y^{(1)} \parallel Y^{(2)} \parallel \dots \parallel Y^{(L)}]V^{(out)} \quad (7)$$

Wherever $\parallel$ signifies concatenation operator; L represents amount of layers; $V^{(out)}$ indicates trainable output matrix. The power of graph-based neural networks is utilized to propagate feature information across the graph, improving prediction accuracy by considering both individual features and their relationships within the graph as given in equation (8)

$$X_u^{(l)} = \sum_{n=1}^N y_{u,n}^{(0)} V_n^{(l)} \quad (8)$$

Where $X_v^{(l)}$ represents weighted amount of the input features of node, u denotes the node index, N is the number of features. The predictions for investment preferences as gold or real estate is generated based on the learned feature interactions and graph-based feature propagation. It is given in equation (9)

$$b_{m,i}^{(l)} = y_{m,i}^{(0)} V_i^{(l)} \quad (9)$$

Where $b_{m,i}^{(l)}$ signifies representation of vector for l st order feature i of the node m , $V_i^{(l)}$ indicates weight vector for feature i . Finally EFIGANN has predicted that Real estate is the most preferred investment.

Hypothesis

Hypothesis 1

H0 - There is no important difference in the preference for Real Estate and Gold investments across different age groups.

H1 - There is important difference in the preference for Real Estate and Gold investments across various age groups.

Hypothesis 2

H0 - There is no important impact of income levels on investment preference.

H1 - Income levels have a substantial effect on investment preference.

Hypothesis 3

H0: There is no important difference in the preference for Real Estate and Gold investments between males and females.

H1: There is an important difference in the preference for Real Estate and Gold investments between males and females.

RESULT & DISCUSSION

This section demonstrates the efficiency of the proposed technique according to the simulation outcome. This work proposes an approach to predict the preferred investment. The proposed technique was implemented in MATLAB [version, R2023a] on a system with [processor, Intel Core i7-12700H], [RAM, 16 GB], and running on [OS, Windows 11].

Data Analysis and Interpretation

In this section, the collected data were analyzed and interpreted. Table 2 presents the demographic profile of employees, including key details on age, income, and gender distributions. The dataset presents a range of financial information and demographic features. 60% were males, with the age distribution in 26–35 age period (45%). Of the participants, 39 % have completed high school, and 53% have one to five years of experience. 80% of them work in the private sector, and 50% in IT/technology. Most people save 10% of their salary (70%), while the majority have annual incomes of less than 500,000.

Table 2. Demographic profile of the employees

Variables	Sub-variables	Frequency	Percentage
Age	18-25	94	30%
	26-35	114	45%
	35 & Above	92	25%
Gender	Male	189	60%
	Female	111	40%
Education	High School	139	39%
	Graduate	59	24%
	Post Graduate	55	19%
	Others	47	18%
Years of Experience	< 1 Year	39	15%
	1-5 Years	161	53%
	6-10 Years	79	24%
	>10 Years	21	8%
Organization	IT/Technology	209	50%
	Healthcare	61	40%
	Education	30	10%
Employment	Government	100	20%
	Private	200	80%
Annual Income	Less than 500000	161	70%
	500000-750000	79	20%
	>750000	21	10%
Annual Savings	10%	209	70%
	20%	61	20%
	30% and Above	30	10%

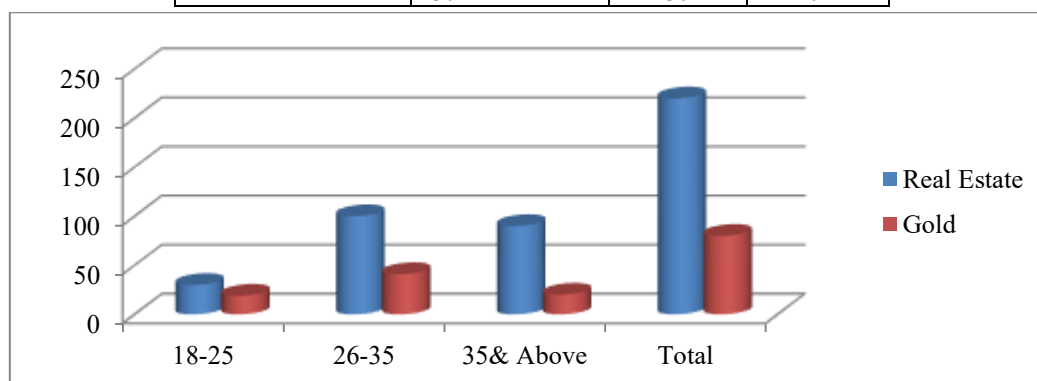


Figure 3. Investment preferences by age group

Table 3. One-way anova test (hypothesis 1)

Investment Avenue	Age Group	Count	Mean	F Value	P Value	Significance
Real Estate	18-25	30	13.64%	4.56	0.012	Significant
	26-35	100	45.45%			
	35& Above	90	40.91%			
Gold	18-25	19	23.75%	3.89	0.021	Significant
	26-35	41	51.25%			
	35&Above	20	25%			

Age and investment preferences are illustrated in figure 3. Table 3 presents the results of the ANOVA test, in which H_0 is rejected at the 0.05 significance level for Real Estate and Gold investments, and the difference in preferences among different age groups is statistically significant [19]. In case of Real Estate investments, the 26-35 age category is the most preferred one (45.45% of people of this age category prefer this type of investment), next comes the 35 and above age category (40.91%), and then the 18-25 age category (13.64%). As for Gold, the most preferable age category again is 26-35 (51.25%), then 35&above (25.00%) and 18-25 (23.75%).

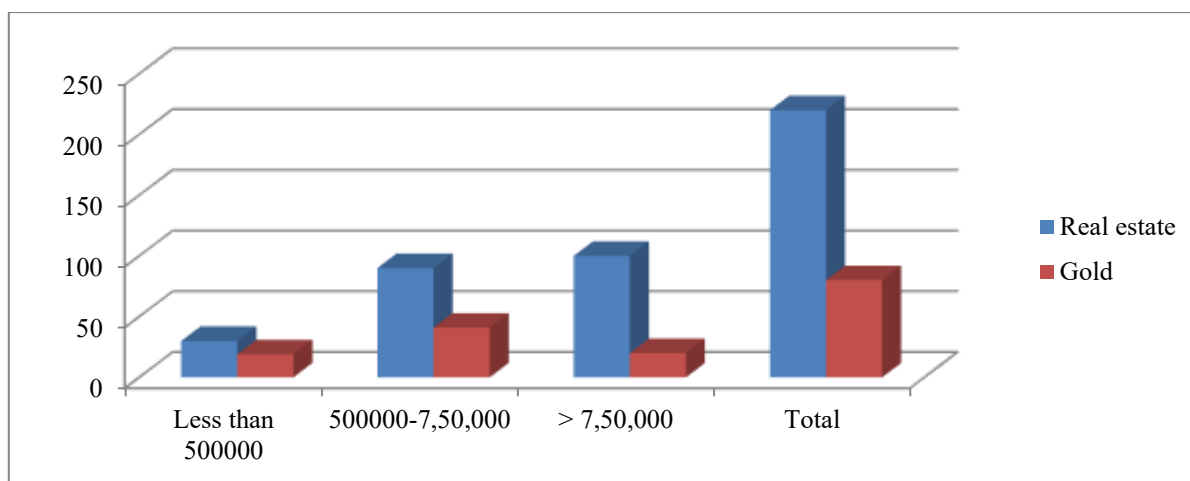


Figure 4. Investment preferences by income level

Table 4. Chi-Square test (hypothesis 2)

Income	Real Estate (Observed)	Gold (Observed)	Real Estate (Expected)	Gold (Expected)	Chi-Square value (Real Estate)	Chi-Square value (Gold)
Less than 500000	30	19	35.87	13.13	0.96	2.62
500000-750000	90	41	96.07	34.93	0.38	1.05
More than 750000	100	20	88.07	31.93	1.61	4.46
Total					2.95	8.13

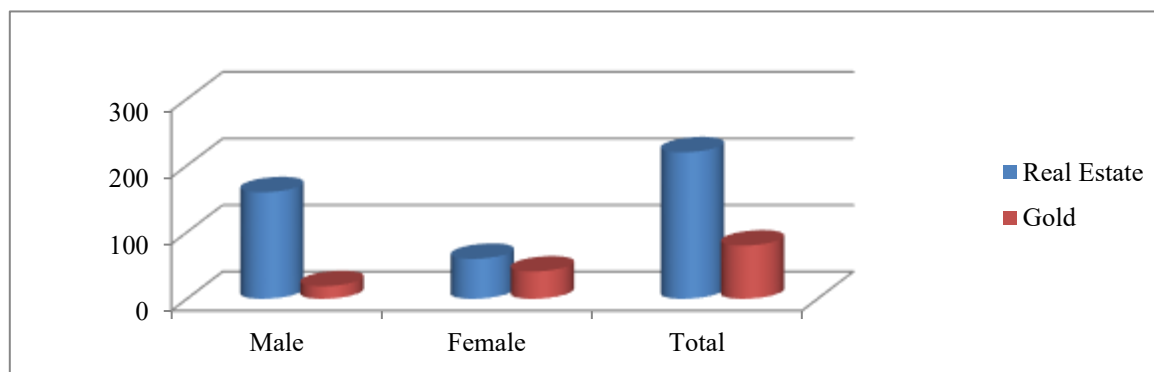


Figure 5. Gender-Based investment preferences

Figure 4 shows the investment preference by people of varying income groups. The results of the Chi-Square test are shown in table 4. Investment preference and income have been found to be significantly related from the study. From the Chi-Square test, the investment preference of employees of Chennai depends significantly on income. Hence, the H_0 is rejected; there exists significant relation between income and investment preference.

Table 5. T-test (hypothesis 3)

Investment Avenue	Gender	Count	Mean	T Value	P Value	Significance
Real Estate	Male	160	72.73%	3.45	0.001	Significant
	Female	60	27.27%			
Gold	Male	19	31.67%	2.89	0.005	Significant
	Female	41	68.33%			

The investment preferences among different gender are shown in figure 5. The findings from the t-test are given in table 5, where the data shows significant differences among preferences for investments in real estate and gold between males and females, thus the null hypothesis (H_0) is rejected at the 0.05 level of significance. In specific, men show more preferences for real estate (72.73%) than women (27.27%), while women show more choices for gold (68.33%) than men (31.67%).

Table 6. Rank of investment patterns and options among the salaried employees

Investment Schemes	1	2	3	4	5	6	7	8	9	10	Total	Rank
Gold	16(10)	17(9)	43(8)	8(7)	7(6)	1(5)	2(4)	23(3)	1(2)	7(1)	820	2
Real-estate	19(10)	42(9)	22(8)	6(7)	1(6)	4(5)	3(4)	0(3)	24(2)	9(1)	830	1

With a focus on gold and real estate, table 6 presents the ranking of salaried employees' investment tendencies. The ratings of each investment plan by employees are displayed on a scale of 1 to 10, with the scores for each rank position indicated by the numbers in parenthesis. Gold came in second overall with a total score of 820, while Real Estate was ranked first with a higher total score of 830. The table demonstrates that employees have a larger preference for Real Estate as an investment option than Gold, as evidenced by its higher overall score.

Performance Analysis

Evaluation Protocol (Leakage-Free): The dataset was first split into training and test sets. LCGBO oversampling was applied only to the training data to correct class imbalance. Model results are reported using stratified 5-fold cross-validation on the training data and a final hold-out test check.

Figures 6 to 8 show the simulation results of the EFIGANN method. The performance metrics are analysed with existing methods such as ANN, LSTM, and GNN. Table 7 shows the hyperparameter settings.

Table 7. Hyperparameter settings

Parameter	Value
Epochs	200
Learning rate	0.001
Weight Decay	Le-2
Dropout	0.9
Activation	leaky relu

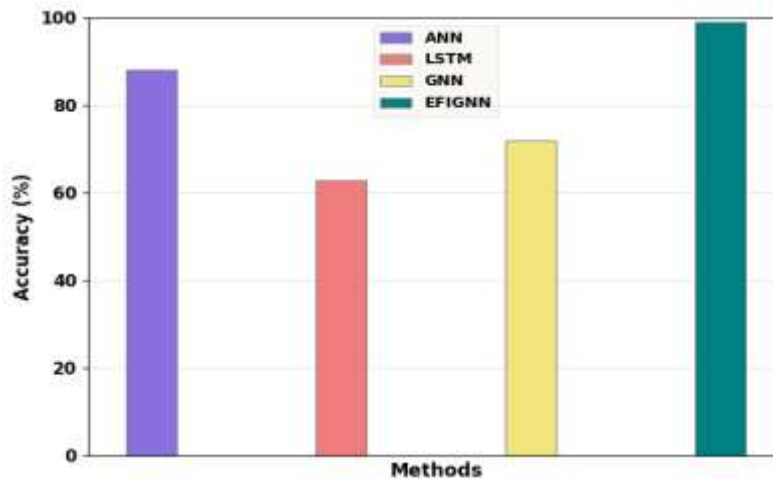


Figure 6. Performance analysis of accuracy

Figure 6 depicts an accuracy analysis. The accuracy of several approaches to forecasting preferences for investments, such as gold or real estate, is contrasted in the graph. ANN, LSTM, GNN, and EFIGANN (Proposed) are among the techniques. EFIGANN outperforms existing approaches, including ANN with approximately 70% accuracy, LSTM with about 75%, and GNN with around 80%. Investors and financial planners are able to make more informed decisions because of this increased precision, which translates into more dependable and accurate projections.

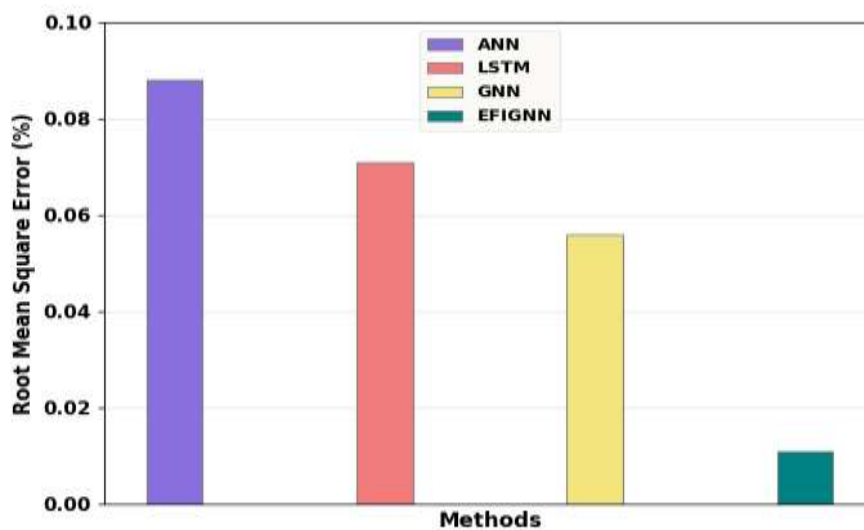


Figure 7. Performance analysis of RMSE

Figure 7 depicts RMSE analysis. The Root Mean Square Error (RMSE) percentages of many techniques for predicting preferences for investments like gold or real estate are contrasted in the graph. EFIGANN has the smallest RMSE percentage of 0.01 compared to ANN, LSTM and GNN that have higher RMSE percentages of 0.1, 0.08 and 0.06 respectively. Its error rate is very low and, therefore, predictions are very accurate and reduce the chances of error and enhance reliability of investment advice. It is due to this high level of precision that financial planners and investors can make more informed judgements on whether to invest in real estate or gold.

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

Where in equation (10), *TP* (True Positive) and *TN* (True Negative) represent correctly predicted investment preferences, while *FP* (False Positive) and *FN* (False Negative) represent incorrectly predicted preferences.

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

Where in equation (11), n is the number of samples, y_i is the actual value, and $\hat{y}_i$ is the predicted value for the sample i .

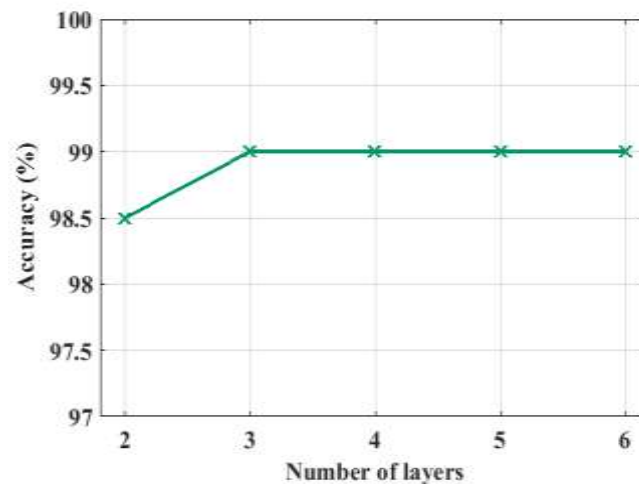


Figure 8. Study of ablation on the number of layers

Figure 8 shows how the number of layers affects the accuracy of EFI-GNN in predicting investment preferences. The accuracy begins to reach 98.5% with two layers, reaches 99 % with three layers and stays constant in deeper configurations that are layers 4 to 6. This pattern indicates that EFI-GNN has the best predictive results at three layers, which is able to fully capture feature interactions without excessively smoothing the results. The loss of accuracy in the deeper layers shows that the model does not lose discriminative power but does not over-aggregate features.

DISCUSSION

The statistical findings provide an essential framework of understanding with respect to the EFIGNN model predictions. The significance of age as a differentiating factor in terms of investment preferences (ANOVA p-value = 0.012 for real estate and p-value = 0.021 for gold) was consistent with the feature embeddings learned through EFIGNN where age was one of the strongest features contributing to preference predictions. In a similar vein, the Chi-Square significance of income in terms of investment preference choices (Chi-Squared value = 2.95 for real estate and 8.13 for gold) supported the fact that EFIGNN utilized the income feature as one of the key nodes for feature propagation. Finally, the T-Test significance of gender in terms of preferences (p-value = 0.001 for real estate and p-value = 0.005 for gold) verified the fact that EFIGNN embeds the gender feature and uses it in its preference prediction model. All in all, the above-mentioned statistical findings show that the factors influencing the EFIGNN prediction process are statistically significant rather than being artifacts of the proposed method.

Investments offer various opportunities to reduce taxable income while supporting long-term financial growth, making them an essential part of tax planning. Besides the reduction of tax liabilities, prudent tax-saving investments help in enhancing long-term wealth and financial security. This paper will

discuss investment preferences of salaried people in Chennai with a particular focus on their decisions between gold and real estate. A structured questionnaire was adopted to gather data via Google Forms, and 300 employees in different industries in Chennai responded to the questionnaire. LCGBO is used to solve the problem of data imbalance. The EFIGANN is suggested to forecast investment preferences. One-Way ANOVA, T-Test, and Chi-Square tests were used to analyze the relationship between demographic variables such as age, income, and gender and investment preferences among residents of Chennai, revealing a significant relationship between these variables. The results of ANOVA showed that both age-differences in the investment preferences of gold (p-value of 0.021) and real estate (p-value of 0.012) were significant at the 0.05 level. The Chi-Square test showed that there was a significant relationship between income and investment decisions of Real Estate and Gold, with values of 2.95 and 8.13, respectively. The results of the T-test revealed a significant difference between genders with p-values of 0.001 in Real Estate and 0.005 in Gold, which were both significant at the 0.05 level.

These results show how important demographic factors are on investment behaviour. The proposed EFIGANN has shown outstanding results, in terms of forecasting Real Estate as the most popular investment with 98 delicacy and 0.1RMSE, which is better than the current approaches. Also, the model emphasizes the benefits of investment decision in real estate, which can bring benefits in terms of tax reduction, making it even more attractive. To provide solid generalization in the presence of class imbalance, stratified 5-fold cross-validation was used. In order to prevent data leakage, oversampling (LCGBO) was applied only to the training folds, while the validation folds remained unaltered.

Practical implications to Risk and Financial management:

- (1) Liquidity and flexibility: Gold is still popular among those investors who are willing to invest in a highly liquid and low-entry cost investment due to its quick conversion to cash in case of an emergency.
- (2) Wealth accumulation and tax planning: The Real estate preference is consistent with long-term wealth acquisition, possible rental income, and tax advantages which are associated with housing finance and capital gains provisions.
- (3) Demographic segmentation: The results of age, income, and gender can be used by financial advisors to create custom product packages (e.g., gold savings plans would be offered to short-horizon investors compared to real estate-based products to long-horizon investors).
- (4) Portfolio risk balancing: A combination of gold (inflation hedge and crisis asset) and real estate (tangible asset and income potential) may enhance diversification of salaried households.

Some limitations associated with this study need to be mentioned. First, the sample of 300 salaried employees from Chennai, collected during a short period of time (January–May 2024), restricts the generalization of the results beyond that region or time frame, and the consideration of gold and real estate as the only investment opportunities prevents the inclusion of other asset classes like stocks and mutual funds. Even though EFIGNN has proved to be highly accurate in predicting future behavior, it still has low levels of interpretability in comparison with the traditional statistical models despite the use of permutation importance and SHAP explanations. The conclusions made here have their practical meaning because the identified demographic factors, such as preference of real estate investments for higher income people and preference of gold for females, may be used for the development of certain financial products. It would be useful to expand the sample size and the geographical area, add other types of investments, use longitudinal data collection, and conduct repeated stratified k-fold cross-validation in the future studies.

CONCLUSION

This research aimed at examining investment preferences among salaried people in India with an emphasis on investments in gold and property. A survey questionnaire was used in the study that targeted 300 employees across different sectors. The data was analyzed through One-Way ANOVA, T-test, and Chi-square test to determine the impact of age, income, and gender on investment preferences in order to propose the use of Explicit Feature Interaction Aware Graph Neural Network (EFIGNN) in predicting preferred investment type. The ANOVA test proved age to be significantly different in terms of

preferences for both gold ($p = 0.012$) and real estate ($p = 0.021$). Chi-square test showed that there exists a significant relationship between income and investment preference ($\chi^2 = 8.13$ for gold, $\chi^2 = 2.95$ for real estate). The T-test revealed that there were considerable differences between genders, since males had a higher preference for real estate ($p=0.001$), while females had a preference for gold ($p=0.005$). It was found that the suggested EFIGNN framework performed better than the previously used ANN, LSTM, and GNN approaches with an accuracy of 98% and RMSE at 0.1%, making real estate a popular investment choice. It is possible to see that the proposed methodology has proven effective in terms of the effect of demographic features on investment decisions. However, the future research should be based on the expanded sample of different geographically located investors, and the number of investment vehicles should be expanded to include stock markets and mutual funds.

REFERENCES

- [1] Jiang Z, Peng C, Yan H. Personality differences and investment decision-making. *Journal of Financial Economics*. 2024 Mar 1;153:103776. <https://doi.org/10.1016/j.jfineco.2023.103776>
- [2] Wei L, Chengshu W. Company ESG performance and institutional investor ownership preferences. *Business Ethics, the Environment & Responsibility*. 2024 Jul;33(3):287-307. <https://doi.org/10.1111/beer.12602>
- [3] Shi Y, Eremina E, Long W. Machine learning models for early-stage investment decision making in startups. *Managerial and Decision Economics*. 2024 Apr;45(3):1259-79. <https://doi.org/10.1002/mde.4072>
- [4] Qasim AW, Aslam AM. Do real estate investments yield the highest returns in Pakistan? *Journal of Real Estate Portfolio Management*. 2025 Jan 2;31(1):73-97. <https://doi.org/10.1080/10835547.2024.2400460>
- [5] Batten JA, Kinatader H, Szilagyi PG. Should you buy gold stocks or paper gold? *Finance Research Letters*. 2024 Nov 1;69:106202. <https://doi.org/10.1016/j.frl.2024.106202>
- [6] Sorensen M, Yasuda A. Stakeholder impact of private equity investments. In *Handbook of the economics of corporate finance 2023 Jan 1* (Vol. 1, No. 1, pp. 299-341). North-Holland. <https://doi.org/10.1016/bs.hecf.2023.02.007>
- [7] Wu B, Gu Q, Liu Z, Liu J. Clustered institutional investors, shared ESG preferences and low-carbon innovation in family firm. *Technological Forecasting and Social Change*. 2023 Sep 1;194:122676. <https://doi.org/10.1016/j.techfore.2023.122676>
- [8] Fateye TB, Peiser RB, Ajayi CA. Behavioral factors influencing investment decision-making in the Nigerian real estate stock market. *Scientific African*. 2024 Jun 1;24:e02238. <https://doi.org/10.1016/j.sciaf.2024.e02238>
- [9] Filippini M, Leippold M, Wekhof T. Sustainable finance literacy and the determinants of sustainable investing. *Journal of Banking & Finance*. 2024 Jun 1;163:107167. <https://doi.org/10.1016/j.jbankfin.2024.107167>
- [10] Puszer B, Czech M. Gold as an alternative investment in times of turmoil. In *Analytics in finance and risk management 2023 Dec 26* (pp. 56-82). CRC Press. <https://doi.org/10.1201/9780367854690-4>
- [11] Asemi A, Asemi A, Ko A. Unveiling the impact of managerial traits on investor decision prediction: ANFIS approach. *Soft Computing*. 2023 Apr 18:1-21. <https://doi.org/10.1007/s00500-023-08102-2>
- [12] Gu A, Wang Z, Yin Z, Zhou M, Li S, Xun Q, Dong J. Empirical research for investment model based on VMD-LSTM. *Mathematical Problems in Engineering*. 2022;2022(1):4185974. <https://doi.org/10.1155/2022/4185974>
- [13] Li J, Yang C, Ye G, Nguyen QV. Graph neural networks with deep mutual learning for designing multi-modal recommendation systems. *Information Sciences*. 2024 Jan 1;654:119815. <https://doi.org/10.1016/j.ins.2023.119815>
- [14] Setty R, Elovici Y, Schwartz D. Cost-sensitive machine learning to support startup investment decisions. *Intelligent Systems in Accounting, Finance and Management*. 2024 Mar;31(1):e1548. <https://doi.org/10.1002/isaf.1548>
- [15] Hu X, Chen Y, Ren L, Xu Z. Investor preference analysis: An online optimization approach with missing information. *Information Sciences*. 2023 Jul 1;633:27-40. <https://doi.org/10.1016/j.ins.2023.03.066>
- [16] Fossen FM, König J, Schröder C. Risk preference and entrepreneurial investment at the top of the wealth distribution: FM Fossen et al. *Empirical Economics*. 2024 Feb;66(2):735-61. <https://doi.org/10.1007/s00181-023-02475-x>
- [17] Nyhus EK, Frank DA, Król MK, Otterbring T. Crypto cravings: Gender differences in crypto investment intentions and the mediating roles of financial overconfidence and personality. *Psychology & Marketing*. 2024 Mar;41(3):447-64. <https://doi.org/10.1002/mar.21921>
- [18] Zhang K, Mao Z, Cao P, Liang W, Yang J, Li W, Zaiane OR. Label correlation guided borderline oversampling for imbalanced multi-label data learning. *Knowledge-Based Systems*. 2023 Nov 4;279:110938. <https://doi.org/10.1016/j.knosys.2023.110938>

- [19] Rayat CS. Variance-ratio test and analysis of variance (ANOVA). In Statistical Methods in Medical Research 2018 Aug 24 (pp. 95-109). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-0827-7_12
- [20] Kim M, Choi HS, Kim J. Explicit feature interaction-aware graph neural network. IEEE Access. 2024 Jan 24;12:15438-46. <https://doi.org/10.1109/ACCESS.2024.3357887>