International Journal of Electrical Electronics & Computer Science Engineering
Special Issue - ICSCAAIT-2018 | E-ISSN : 2348-2273 | P-ISSN : 2454-1222
Available Online at www.ijeecse.com

# Predicting Success of Residential Project Using Artificial Neural Network

Manoj Kumar Jain
PhD Scholar, School of Engineering, Poornima University, Jaipur, India
manojjain1972@gmail.com

***Abstract:*** ***Predicting the success of residential projects as well as ranking them has intrigued many scholars and realty industry leaders. It is a difficult and challenging problem. Predicting or forecasting the success of residential projects is very important because the buyers will invest their hard earned money into it, thinking of spending their entire life earning and expecting some long term benefits out of it at the same time developer will also invest lot of money and expectedly will want to reap dividends This paper examines and analyzes the use of neural networks as a prediction and classification tool. Specifically the use of neural network in predicting the success of a residential project before the project is actually launched is explored. We have converted the prediction problem into a classification one, i.e., a residential project based on its total number(%) of bookings during launch offer is classified in one of the seven output categories chosen ,ranging from 'flop' to 'top'. We have used nine different types of independent input parameters - Location, Cost/sq ft, Developer Track record, Completion time of project, Maintenance cost, Type & size of accommodation, Facilities/Common amenities, interest rate for bank loan and interior for this classification. Because our neural network model can predict the success of a residential project in gaining maximum booking before the project is actually launched, it can be used as a powerful decision aid by housing project developer and individual house buyer. We have chosen to use the Error Back Propagation learning algorithm and the Multi Layer Perceptron (MLP) network architecture to formulate our problem It has been concluded that neural networks do have the capacity to classify residential project and predict number of booking before the project is actually launched. If properly trained, lot of people can draw benefits from the use of this prediction tool.***

***Keywords: Prediction, Residential Project, Neural Networks, Multilayer Perceptron, Back Propagation.***

## I. INTRODUCTION

Predicting the success of the residential project has intrigued many scholars and housing industry leaders as a difficult and challenging problem. To some analysts, 'Realty sector is the land of hunch and the wild guess' due to the difficulty and uncertainty associated with predicting the housing demand because it also depend on factors like economic condition of country ,interest rate for loan etc.. Such unpredictability of the product demand makes the real estate business one of the riskiest endeavors for developer to take in today's competitive world. .No one can tell you how a housing project is going to do in the marketplace.

Despite the difficulty associated with the unpredictable nature of the problem domain, several researchers have attempted to develop models for forecasting the success of residential project, primarily using statistics-based forecasting approaches. Most of them have tried to predict the total performance of residential project after the project is launched for booking. However the results obtained were not very accurate for decision support.

In our paper, we explore the use of neural networks in forecasting the performance of residential project before the project is actually launched. We have converted this forecasting problem into a classification one, i.e. , we classify a residential project based upon its total number (%) of booking during launch period in one of the seven categories, ranging from 'flop' to 'very successful Neural networks (NN) are known to be biologically inspired analytical techniques, capable of modeling extremely complex non-linear functions. For many years, linear modeling has been the commonly used technique in capturing and representing functional relationships between dependent and independent variables, largely because of its well-known statistically explainable optimization strategies. In the problem scenarios where the linear approximation of a function was not valid (which was frequently the case), the models suffered accordingly. Now, such cases can easily be modeled with neural networks. Applications of neural networks have been reported in many diverse fields addressing problems in areas such as prediction, classification, and clustering. However, none of these include an application in forecasting success of residential project. This study is one of the first to attempt the use of neural networks for addressing this challenging problem that has drawn the attention of many researchers in such areas of decision support systems and management science.

The remainder of this paper is organized as follows. Section 2 briefly reviews the basics of ANN & also presents the comparison between the statistical techniques & our ANN approach. Section 3 gives the details of our methodology by specifically talking about the data & variable definition and proposed architecture, and finally in the conclusion section we discuss the benefits, this approach has to offer as well as our plan of action that we intend to take in order to implement this problem.

## II. LITERATURE REVIEW

### *A. Review of ANN:*

A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural Network technology stemmed from the desire to develop an artificial system that could perform

'intelligent' tasks similar to those performed by the human brain. Neural networks resemble the human brain in the following two ways:

1. A neural network acquires knowledge through learning.
2. A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights.

Neural Networks are non-linear data driven self adaptive approach as opposed to the traditional model based method. They are powerful tool for modeling especially when the underlying data relationship is unknown. ANNs can identify and learn correlated pattern between input data set and corresponding target values. After training ANNs can be used to predict the outcome of new independent input data.

The true power and advantage of neural networks lies in their ability to represent both linear and non-linear relationships and in their ability to learn these relationships directly from the data being modeled. In solving the problems by ANNs "learning by example" replaces "programming". This feature makes such computational models very appealing in application domain where one has little or incomplete understanding of the problem to be solved but where training data is readily available. Traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics. From a statistical perspective neural Network are interesting because of their potential use in prediction and classification problem

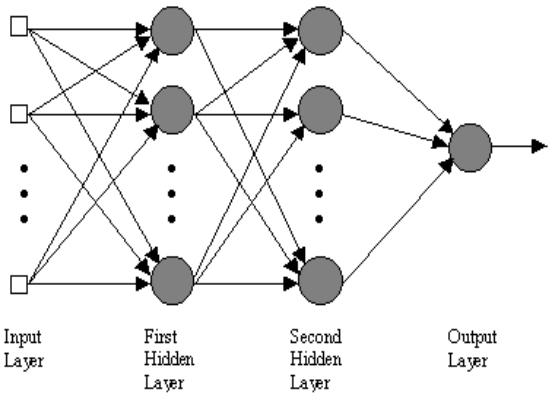
*B. Multilayer Perceptron & Back-Propagation:*

The most common and popular form of neural network model is the multilayer perceptron (MLP). A multilayer perceptron:

- Has any number of inputs
- Has one or more hidden layers with any number of units
- Uses linear combination functions in the input layers
- Uses generally sigmoid activation function in the hidden layers
- Has any number of outputs

.Given enough data, enough hidden units and enough training time , an MLP with just one hidden layer can learn to approximate virtually any function to any degree of accuracy. (A statistical analogy is approximating a function with nth order polynomial). For this reason MLP are known as universal approximator and can be used when you have little prior knowledge of the relationship between inputs and targets. Although one hidden layer is always sufficient provided you have enough data, there are situation where a network with two or more hidden layer may require fewer hidden and weights than a network with one hidden layer, so using a extra hidden layer sometimes can improve generalization [1].

A graphical representation of an MLP with two hidden layer is shown below in fig 1.




.

Fig. 1. Block Diagram of a MLP with 2 Hidden Layer

The MLP and many other neural networks learn using an algorithm called back-propagation. With back-propagation, the input data is repeatedly presented to the neural network. With each presentation the output of the neural network is compared to the desired output and an error is computed. This error is then fed back (back-propagated) to the neural network and used to adjust the weights such that the error decreases with each iteration and the neural model gets closer and closer to producing the desired output. This process is known as "training".

## III. METHODOLOGY USED

*A. Neural Networks and Statistical Modeling:*

Though commonly known as black box approach, in the last decade, artificial neural networks have been studied by statisticians in order to understand their prediction power from a statistical perspective[2] These studies indicate that there are a large number of theoretical commonalities between the traditional statistical methods, such as discriminant analysis, logistic regression, and multiple linear regressions, and their counterparts in neural networks, such as multi-layered perceptron, recurrent networks, and associative memory networks.[15]

Multi layer perceptron (MLP) neural network architecture is known to be a strong function approximator for prediction & classification problems. It has been shown that given the right size and structure, MLP is capable of learning arbitrarily complex non-linear functions to an arbitrary accuracy level. Thus it is a likely candidate for exploring the rather difficult problem of mapping residential project success to the underlying characteristics. The error back-propagation

algorithm presents the best mapping; it is thus used in our approach.

*B. Data and Variable Definition:*

Currently in our study, we are considering n number of residential projects, launched in NCR region and in nearby areas over the last 2 years. The variable of interest in this study is the total number (%) of booking during the launch period. It does not include any booking after the launch period is over. We have converted this prediction problem into a classification one i.e. a residential project based on it's total number (%) of booking is classified in one of the seven categories, ranging from 'FLOP' to 'Successful'. This process of converting a continuous variable in a limited number of classes is commonly called 'discretization' in neural network literature. Discrete values are a limited number of intervals in a continuous spectrum, whereas continuous values can be infinitely many. Many prefer using discrete values as opposed to continuous ones in developing prediction models because (i) discrete values are closer to a knowledge-level representation ; (ii) data can be reduced and simplified through discretization , (iii) for both users and experts, discrete features are easier to understand, use, and explain ; and finally, (iv) discretization makes many learning algorithms faster and more accurate.

In our study, the dependent variable was discretized into seven classes using the following breakpoints:

Table I. Discretization of Dependent Variable

| Class No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| % Booking | 100% | 90 - 100% | 80 - 90% | 70 – 80% | 60 – 70% | 50 - 60% | < 50% |

We plan to use nine different types of independent variables. Our choice of independent variables is based on the feedback that have been received from the realty industry experts and previous studies conducted in this field. Each categorical- independent variable is converted into 1-of-N binary representations. Thus we get a number of pseudo- representations that increases the independent variable count from 9 to 29

.A neural network treats these pseudo variables as different mutually exclusive information channels .In the process of value assignment, all pseudo representations of a categorical variable will be given a value of 0, except the one that holds true for the current case, which will be given the value of 1. For instance, in 1-of-N binary representation, the interest rate for loan variable would be represented with three pseudo variables denoting high, medium and low and for instance if the current interest rate is high, the value of high would be set to 1 and the values of medium and low would both be set to 0.

Following is a short description of these variables with their respective representation schemas.

*1) Location of Property*: The location of the property is, perhaps, one of the most crucial aspects in the reality business. A buyer would like to know how well connected is the piece of land with available infrastructure , what are the development prospects and how far the location is from common places like Railway station, bus stand and airport etc. We have classified location variable into 3 categories- Within City, Outskirt, and Remote

*2) Track Record of Developer*: Another important parameter in deciding the success of project is the developer's or builder's reputation for producing on a timely and economical basis. Buyers are also worried about whether the builders will deliver the quality that have been promised by them. A good track record can be helpful in arranging financing and attracting investors for a new project. We have 4 possible ratings in this category, viz, Excellent, Good, Poor and not Known

*3) Completion Time of Project*: Yet another important criterion that a buyer would like to know is the completion time of project promised by the developer. If the buyer is currently having a rented accommodation, he will have to pay both Loan EMI as well as rent till the time of completion of the project, so he want to invest in the project with less completion time. We have classified this parameter into 3 categories – 1.5 to 2 years, 2 to 3 years and > 3 years

*4) Cost Per Sq. Ft.*: Cost per square ft is an important time dependent parameter that a buyer considers before investing. As this parameter is time dependent, instead of assigning it an absolute value, we specify this parameter by comparing the cost with cost of other project in the same area. We have specified 3 categories for this parameter- Expensive, Reasonable and Economical.

*5) Type & Size of Apartment:* Type and size of apartment is another parameter that a prospective buyer considers before investing as total cost of apartment depends on cost per square ft as well as size of apartment. A buyer would like to invest in an apartment that will suit his requirements as well as his budget. We Categorize this parameter into 4 Categories – 1 BHK, 2 BHK, 3 BHK and >= 4 BHK

*6) Maintenance Cost:* Monthly maintenance cost is another criterion that a buyer considers before investing in an apartment. Buyers also want to know, whether the maintenance cost will be taken in advance or it will be charged on monthly basis. We classify this parameter into three categories – High, fair and low.

*7) Interior Design:* Design and Material used for Wall, Ceiling, Floor, Windows and Doors is the interior provided by builders. Before purchasing a apartment a

buyer also go through interior of the apartment. We have 3 possible rating in this category- Luxurious, good, budget

*Facility/ Common Amenities Provided:* Power backup, Security arrangement, swimming pool; children play area, water softening plant, ample car parking and in-house maintenance facility are some of the common amenities provided in a residential society. Before investing a buyer like to know, how well these common facilities are provided by builders. We have classified this parameter in 3 categories – Excellent, good and fair.

*8) Interest Rate for Loan:* Home loan interest rate is one of the biggest and the most critical factor that drives Indian real estate. Loan interest rate put a deep impact on the pockets of the people intending to buy residential properties by taking home loan, so it is a parameter that a buyer considers first before purchasing a house. We classify this parameter into 3 categories - Low (< 9%), Average (9 to 13%) and High (>13%).

A summary of above-mentioned variables is given in Table 1. In all there are 29 decision variables & 7 output variables that we intend to use in our approach.

Table II. Summary of Independent Variables

| S. No. | Independent Variable Name | No. Of Values | Possible values |
|---|---|---|---|
| 1) | Location | 3 | Within city, Outskirt, Remote |
| 2) | Track Record of developer | 4 | Excellent, Good, Poor, Not-Known |
| 3) | Completion time | 3 | 1.5 to 2 Years, 2 to 3 years, > 3 years |
| 4) | Cost per sq ft. | 3 | Expensive, Reasonable, Economical |
| 5) | Size & Type of Apartment | 4 | 1 BHK, 2 BHK, 3 BHK, >3 BHK |
| 6) | Maintenance Cost | 3 | High, fair , Low |
| 7) | Interior Design | 3 | Luxurious, Good, Budget |
| 8) | Facilities/ Common amenities | 3 | Excellent, Good, Fair |
| 9) | Interest Rate for loan | 3 | Low (< 9%), Average (9 to 13%) High(>13%) |

## IV. CONCLUSION

Our approach of using Neural Networks in this field aims to classify the residential projects in one of the seven output categories & thus predict their success during launch offer, even before the project is actually launched. We've started implementing our problem & would come up with the results soon. This study is the first attempt to use Neural Networks for addressing this challenging problem that combines two different application domains of Predicting & Classification & brings out the much-desired output. This model would be highly beneficial to the:

*1) Residential Project Developer*: They'll be able to know in advance how successful their project will be. They can come to know their projects strong & weak points & can therefore make appropriate changes

*2) House Buyer*: By classifying the residential project in one of the seven output categories, house buyers will be able to decide whether they should invest their hard earned money in a particular residential project or not.

## V. REFERENCES

[1] Zurada, Jacek M. Introduction to Artificial Neural Systems.3rd Ed West Publishing Company, 1992

[2] Igor Alexsander and Helen Morton. Introduction to Neural Computing. International Thomson Computer Press. 1995

[3] Schalkoff, Robert Artificial Neural Networks,2nd Ed McGraw-Hill.1997

[4] White, H. "Some asymptotic results for learning in single hidden layer feed-forward network models". Journal of American Statistical Association, 84(408), 1003–1013.1989.

[5] Breiman, L. M. "some properties of splitting criteria. Machine Learning", 24(1), 41–47. Coskunoglu, O., Hansotia, B., & Muzaffar, S." A New Logic model for decision making and its applications". The Journal of the Operations Research Society, 36(1), 35–41. 1996.

[6] M.Negnevitsky, (2005)"Multi-Layer Neural Networks with Improved Learning Algorithms", Proceedings of the Digital Imaging Computing: Techniques and Applications

[7] A. Ahmed and NI. M. Fahmy, IEEE, Fellow" Application of Multi-layer Neural Networks tcil Image Compression" IEEE International Symposium on Circuits and Systems, June 9-12, 1997,Hong Kong

[8] Roheet Bhatnagar, Vandana Bhattacharjee and Mr inal Kanti Ghose, "Software Development

Effort Esti mation –Neural Network Vs. Regression Modeling Approach", International Journal of Engineering Science and Technology, Vol. 2(7), 2010,page: 2950-2956.

[9] X. Yao, "A review of evolutionary artificial neural networks", Int. J. Intell. Syst., vol. 8, no. 4, pp. 539-567, 1993.

[10] L. Chan, F. Fallside, "An adaptive training algorithm for backpropagation networks", Comput. Speech Language, vol. 2, pp. 205-218, 1987.

[11] H. Izadi, J. Sadri, N.A Mehran, "Intelligent Mineral Identification Using Clustering and Artificial Neural Networks Techniques, Publihed in the Proceedings of The First Iranian Conference on Pattern Recognition and Image Analysis (PRIA 2013), Birjand, March 2013.