A NOVEL APPROACH FOR PREDICTING HOUSE PRICE USING K-NEAREST NEIGHBOR ALGORITHM COMPARING ACCURACY PREDICTION WITH ARTIFICIAL NEURAL NETWORK

Z.Sri Sai Swetha[1], R. Dhanalakshmi[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: This work is a comparative study of artificial intelligence Novel K-Nearest Neighbors Algorithm (KNN) and Novel Artificial neural network for the house price prediction system to improve the accuracy of house price prediction.

Materials and methods: From Machine Learning, Novel K-Nearest Neighbors (N=10) and Novel Artificial neural network (N=10)methods are simulated by varying the KNN parameter and Novel Artificial neural network to optimize the pH.The sample size was calculated using the G power of 80% for two groups and there are 40 samples used in this work.

Results: Based on the obtained results KNN has significantly better classification accuracy (67.92) compared to the Novel Artificial neural network algorithm (61.94) Statistical significance difference between long short term memory and Novel Artificial neural network was found to be 0.156 ($p>0.05$) which infers that both groups are insignificant.

Conclusion: The ANN algorithm produces better results in prediction on house price monitoring to improve the price prediction accuracy than the Novel Artificial neural network algorithm.

Keywords: Artificial intelligence, Machine Learning, Novel Artificial Neural Network, Prediction, Classification.

[1]Research Scholar, Department of Computer Science and Engineering,Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

[2*]Project Guide, Department of Computer Science and Engineering,Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

1. Introduction

In this research work, artificial intelligence and machine learning is a good prediction method that had a great impact on house price prediction (Akyuz, Erdogan, and Yildiz, n.d.) which makes the price prediction based on infinite houses and their land zones in the city. The destruction of fake prices acts like perhaps the best danger to the idea of consistent truth using novel classification accuracy (Kaushal and Shankar, n.d.) since the advancement of the fake prices in prediction of the house,there has been a developing joint exertion by the scholarly local area to explore and create the approaches fit for the investigating (Kirkeby and Larsen 2021) and also used to recognize and intercede an exhibition of fraud in house price prediction.The importance of this research study is to identify the logical proof that has effectively uncovered the multiple fake price of a house bogus action in search of houses (Ahtesham, Bawany, and Fatima 2020) the various artificial intelligence applications of this research study is real estate price prediction system and to apply limited features of house price predictions up to an arbitrary likelihood of roughly 62% correctness (Sarip and Hafez 2015) moreover,the battle against the unregistered house dealers the informal community and the plot utilization issues indistinguishable (Chai et al. 2018).

The Literature survey in house price prediction systems and many comparisons have also been done for this prediction. There are several research papers available in IEEE Xplore and Science Direct. Around 300 articles are identified that relate relevant research to house price prediction. Although the accuracy of the price prediction software has remarkably developed and marked over the past two decades (Samukcham 2021), it is important to note that the automatic price prediction in house price prediction systems is not entirely automatic (Kirkeby and Larsen 2021). Identification accuracy can be a bit poor in cases where the image captured conditions are not optimal (Kaushal and Shankar, n.d.). To manage this uncertainty,in many applications algorithms have human users with a 'house list' displaying the topmost matching image which were collected from the database ranked in the order of resemblance to the scrutiny image (Deghi et al. 2020b).Our team has extensive knowledge and research experience that has translated into high quality publications(K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022)

From this literature survey, it can be concluded that this K-nearest neighbor algorithm for house price prediction systems helps for the house price prediction systems to improve the accuracy of house price which brings effective results as the previous results were giving less accuracy as it was detecting some false images. In this system, the algorithm tracks the false of house's price in house price prediction (Zhang et al. 2020) has been implemented in accuracy for artificial intelligence in machine learning algorithm, laboratory of neurogenetics and neuroscience algorithm, K-Neighbor algorithm. The aim of the study is to show that the K-Nearest Neighbors algorithm appears to perform better among all the algorithms. In the previous study efficiency of predicting the house price in a city by using classifications of accuracy with the proposed K-Nearest Neighbor algorithm to improve the classification accuracy. Eventually , the proposed approach validity is estimated utilizing data recovery measurements like affectability,exactness,and precision (Deghi et al. 2020a).

2. Materials and Methods

The research work is carried out in the Programming Lab Of Machine Learning, Department Of Computer Science And Engineering at Saveetha School Of Engineering, Saveetha Institute Of Medical And Technical Sciences, Chennai.The sample size has been calculated using G power software by comparing the process and their results. In each group, Two numbers of groups are selected for comparing the process and their results. In each group, 20 sets of samples and 40 samples in total are selected for this work. The pre-test power values calculated using G power 3.1 software (g power setting the parameters: statistical difference between two independent means, α=0.05, power=0.80,two algorithms (K-Nearest Neighbors and Artificial neural network) are implemented using technical analysis software. In this work, no human and animal samples were used so no required (G and Sundar 2020).

K-Nearest Neighbor

K-nearest neighbor is one in all the only machine learning algorithms supported supervised learning technique. KNN algorithm program assumes the similarity between the new data and market cases into the class, that's the most almost like market classes. The KNN algorithm program stores all the similarities.

The main aim and principle behind nearest method is

- To find a predefined number of training sample close in distance to the new data point
- To predict the new labeo from there number of training samples KNN could be

the non-parametric algorithm program as a result of it doesn't build any assumption on underlying knowledge. It's conjointly referred to as a lazy learner algorithm program as a result of it doesn't learn from the coaching set directly instead it stores the dataset and at the time of classification, it performs an associate on the dataset.

This algorithm is used to take care of grouping model issues. The K-Nearest Neighbor algorithm or KNN algorithm essentially makes an imaginary boundary to classify the data. When new data points come in, the algorithm will try to anticipate that to the nearest to the boundary line.

Accordingly larger k value means smoother covers of separation resulting in less complicated models, Though more modest k value tends to over fit the historical data and predict the future

Input: *Training and testing data*
Output: *Accuracy*
Step 1. Import dataset
Step 2. Preprocess the data
Step 3. x,y(define the features needed)
Step 4. Split the dataset into train and test data
Step 5. Define the model
Step 6. Define the evaluation method
Step 7. Fit the model on the training and test data
Step 8. Evaluate the model on the dataset
Step 9. Report performance

Novel Artificial Neural Network

Novel Artificial Neural Network (ANN) or neural networks are computational algorithms. It intended to simulate the behavior systems composed of "neurons". ANNs are computational models inspired by an animal's central system of interconnected "neurons" which can compute values from inputs. A neural network is an oriented graph. It consists of nodes which in the biological analogy represent neurons, connected by srcs. It corresponds to dendrites and synapses. Neural networks find great applications in data mining used in sectors. For example economics, forensics, etc and for pattern recognition. It can be also used for data classification in a large amount of data after careful training.

A neural network may contain the following 3 layers:

- Input layer : The activity of the input units represents the raw information that can feed into the network.
- Hidden layer : To determine the activity of each hidden unit. The activity of the hidden units and the weights between the hidden and output units
- Output layer :The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units.

ANNs are considered as simple mathematical models to existing data analysis technologies. Although it is not a comparable model to enhance existing data analysis technologies. Although it is not comparable with the power of the human brain, still it is the basic building block of Artificial intelligence.

Input: *Training and Testing data*
output: *Accuracy score*
Step 1. From sklearn import artificial intelligence
Step 2. Import dataset
Step 3. Preprocess the data
Step 4. x,y(Define the features needed)
Step 5. Split the dataset into training and testing sets
Step 6. Fit the ANN algorithm on the training and testing dataset
Step 7. Train the ANN algorithm
Step 8. Predict the result
Step 9. Calculate the accuracy score.

The hardware configuration was an intel core i5 processor with a RAM size of 8GB, the system type used was a 64bit OS, X64 based processor with an HDD of ITB. The operating system used was Apple ios and the tool used for implementing was jupiter Notebook with python programming language.

Statistical Analysis

SPSS software is used for statistical analysis of KNN and the ANN algorithm between house price predicting groups was done through SPSS version 2. The independent variable is KNN accuracy and the dependent variable is efficiency. The independent T-test analyses Novel K nearest neighbor algorithm approach gives accuracy of the KNN (Samukcham 2021) for both methods. The proposed Novel K nearest neighbor algorithm approach gives accuracy 67.92% which is significantly better in classification accuracy when compared to the ANN algorithm which has less accuracy 61.94%.

3. Results

Table 1 shows the simulation results of the proposed algorithm Novel K-Nearest Neighbors and the existing system Novel Artificial neural network were run at different intervals in the apple os and mean accuracy of the KNN algorithm was 67.92% and the Novel Artificial Neural Network was 61.94%.

Table 2 represents the T-test along with the Novel Artificial neural network algorithm.The mean,

standard deviation and standard error mean were calculated by taking an independent variable T-test along with the Novel Artificial Neural Network algorithm with a p value of 0.156 and effect size = 1.313.

This represents the Mean of the KNN algorithm which is better compared with the Artificial neural (ANN) algorithm with a standard deviation of 1.10 and 0.894 respectively. From the results, the KNN algorithm (67.92%) gives better accuracy than the Novel Artificial neural network (ANN) algorithm (61.94%). Figure 1 gives the comparison chart of KNN of Novel Artificial neural network (ANN) algorithms in terms of mean and accuracy. The mean accuracy of the KNN algorithm is better than the Novel Artificial neural network (ANN).

4. Discussion

The KNN and Novel Artificial neural network (ANN) algorithms are implemented and compared the prices for prediction to improve the accuracy of house prices. From obtained results, it is concluded that the KNN algorithm provides better accuracy results compared to the Novel Artificial neural network (ANN) algorithm.The Data evaluation was performed using KNN SPSS tool (Kaushal and Shankar, n.d.). The experiments were conducted among the study groups K-Nearest Neighbors and Novel Artificial neural network by varying sample size.

In the recent surveys, the proposed (Kamencay et al. 2016) Novel K-nearest neighbors has increased throughput value with 67.92 and significance valueless with 0.059 (Kamencay et al. 2016). In this research work, Novel Artificial neural network and the K-nearest neighbors algorithms were analyzed for increasing the throughput percentage of student access privilege. It is observed that the K-Nearest Neighbors appear to increase the throughput compared to the Novel Artificial neural network algorithm (61.94%). The sigmoid function maps the predictions to the probabilities of the K-Nearest Neighbor based on the access privilege.The main advantage of using KNN algorithm is fast processing, high quality of output. The disadvantage of using ANN algorithms is less accuracy, more time complexity. The experimental results of machine learning and artificial intelligence show evidence there is a significant difference between the Novel Artificial neural network and K-nearest neighbors algorithms (Kaushal and Shankar, n.d.). In this paper , the Novel Artificial neural network algorithm was implemented with an accuracy percentage of 61.94%. From the above discussion, it was observed that the factors which affect the current research study were due to attributes that affect the throughput percentage of house price databases in the research work are title, text, subject, data (G and Sundar 2020). The attributes that mainly concentrated to increase the throughput were to predict the correct price of a house.

The limitations of the proposed work are one of the attributes in the price prediction dataset used for predicting the house price in the residential area . The easy access and exponential growth of the information available on the residential database have made it intricate to distinguish between original and fake price. House price prediction has limited paper prediction ability based on future prices of the house, which makes better house price prediction in the future. The Novel Artificial neural network algorithm can address future predictions of house price's.

5. Conclusion

The proposed K-nearest neighbors algorithm to find the house price prediction.This research study applied the K-Nearest Neighbors algorithm using ANN algorithm to find the house price prediction. The Proposed KNN proved with better accuracy of 67.92% when compared to Novel Artificial neural network accuracy is 61.94% for predicting the price of a house in a locality.

Declaration

Conflict of Interests

No conflicts of interest in these manuscripts

Authors Contributions

Author ZSSS was involved in data collection, data analysis, manuscript writing. Author RD was involved in conceptualization, data validation, and critical review of the manuscript

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

1. Railsdata Software pvt Ltd
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering

5. References

Ahtesham, Maida, Narmeen Zakaria Bawany, and Kiran Fatima. 2020. "House Price Prediction Using Machine Learning Algorithm - The Case of Karachi City, Pakistan." *2020 21st International Arab Conference on Information Technology (ACIT)*. https://doi.org/10.1109/acit50332.2020.9300

074.

Akyuz, Sureyya, Birsen Eygi Erdogan, and Ozlem Yildiz. n.d. "A Novel Hybrid House Price Prediction Model." https://doi.org/10.22541/au.160284063.37909495/v1.

Chai, Zhilei, Wei Song, Qinxin Bao, Feng Ding, and Fei Liu. 2018. "Taking Advantage of Hybrid Bioinspired Intelligent Algorithm with Decoupled Extended Kalman Filter for Optimizing Growing and Pruning Radial Basis Function Network." *Royal Society Open Science* 5 (9): 180529.

Deghi, Andrea, Mitsuru Katagiri, Mr Sohaib Shahid, and Nico Valckx. 2020a. *Predicting Downside Risks to House Prices and Macro-Financial Stability*. International Monetary Fund.

Deghi, Andrea, Mitsuru Katagiri, Sohaib Shahid, and Nico Valckx. 2020b. "Predicting Downside Risks to House Prices and Macro-Financial Stability." *IMF Working Papers*. https://doi.org/10.5089/9781513525839.001.

G, Sundar, and G. Sundar. 2020. "Improvement of Stock Price Prediction by Synthesizing Recurrent ANN Back Track Solver." *International Journal of Psychosocial Rehabilitation*. https://doi.org/10.37200/ijpr/v24i1/pr200297.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." *Energy*. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." *Environmental Research* 212 (Pt A): 113153.

Kamencay, Patrik, Tibor Trnovszky, Miroslav Benco, Robert Hudec, Peter Sykora, and Andrej Satnik. 2016. "Accurate Wild Animal Recognition Using PCA, LDA and LBPH." *2016 ELEKTRO*. https://doi.org/10.1109/elektro.2016.7512036.

Kaushal, Anirudh, and Achyut Shankar. n.d. "House Price Prediction Using Multiple Linear Regression." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3833734.

Kirkeby, Sara J., and Vegard Høghaug Larsen. 2021. *House Price Prediction Using Daily News Data*.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." *Computers and Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2022.107902.

Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. "Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database." *Korean Journal of Chemical Engineering*. https://doi.org/10.1007/s11814-022-1120-5.

Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. "Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113136.

Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. "Green and Eco-Friendly Approaches for the Extraction of Chitin and Chitosan: A Review." *Carbohydrate Polymers* 287 (July): 119349.

Samukcham, Millee. 2021. "House Price Prediction Using Machine Learning." *International Journal for Research in Applied Science and Engineering Technology*. https://doi.org/10.22214/ijraset.2021.37451.

Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. "Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies." *Chemosphere* 298 (July): 134341.

Sarip, Abdul G., and Muhammad Burhan Hafez. 2015. "Fuzzy Logic Application for House Price Prediction." *International Journal of*

Property Sciences. https://doi.org/10.22452/ijps.vol5no1.3.

Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. "Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route." *Chemosphere* 298 (July): 134265.

Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. "Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene." *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-022-06636-5.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123814.

Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. "Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113114.

Zhang, Hang, Kehan Wang, Mengchu Li, Xinchen He, and Ruibo Zhang. 2020. "House Price Prediction with An Improved Stack Approach." *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1693/1/012062.

S. Hasan, "Determinants of house prices in Turkey: Hedonic regression versus Novel artificial neural network," Expert systems with Applications, pp. 2843-2852, March 2009.

Kannan P, Rabanal P, Scott AM, "Monetary and macroprudential policy rules in a model with house price booms," The BE Journal of Macroeconomics, June 2012.

Andrew, M. and Meen, G., "House price appreciation, transactions and structural change in the British housing market: a macroeconomic perspective," Real Estate Economics, pp.99-116, March 2003.

Beltratti, A. and Morana, C., "International house prices and macroeconomic fluctuations," Journal of Banking & Finance, pp.533-545, March 2010.

L. Visit, "House price prediction: hedonic price model vs.Novel artificial neural network," In New Zealand Agricultural and Resource Economics Society Conference, pp. 25-26. June 2004.

Mason C, Quigley JM, "Non-parametric hedonic housing prices," Housing studies, pp.373-85, July 1996.

TABLES AND FIGURES

Table 1. Group Statistics results Mean of K-Nearest Neighbors Algorithm is 67.921% is more compared to Novel artificial neural network 61.94% and standard error mean for KNN is 0.32 and ANN is 61.94.

	Algorithms	**N**	**Mean**	**Standard Deviation**	**Standard Error Mean**
Accuracy	KNN	10	67.9210	0.97602	0.30864
	ANN	10	61.9470	0.76178	0.24090

Table 2: This Independent sample t-test obtained the significance as 0.156 (p>0.05), mean difference, std. error difference, and lower and upper interval difference. Independent samples t-test is applied for comparison of KNN and ANN.

		Leven's test for equality of variance		**T-Test for Equality of Means**						
		f	**sig**	**t**	**df**	**sig**	**Difference**	**std**	**95% Confidence Interval Of Difference**	
									Lower	**upper**
Accuracy	Equal variance assumed	2.190	0.156	15.258	18	0.003	5.97400	0.39153	5.15144	6.79656
	Equal Variance not assumed			15.258	16.997	0.003	5.97400	0.39153	5.14794	6.80006

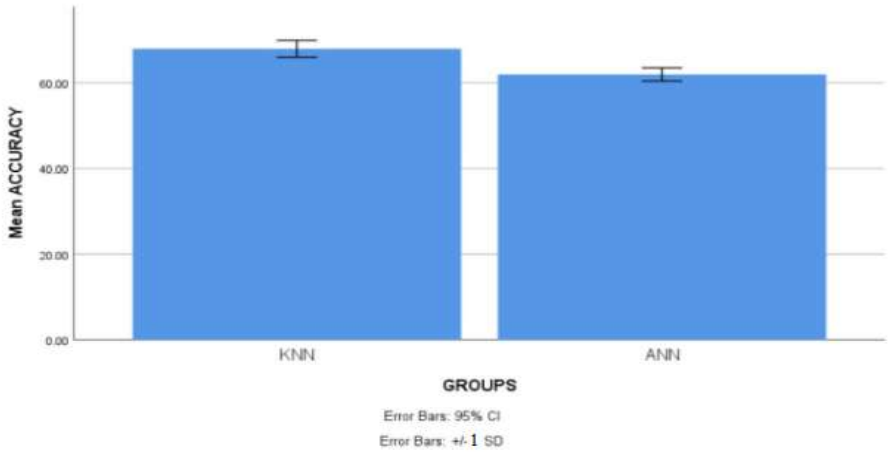

Fig. 1. Comparison of K-Nearest Neighbors algorithm and Novel Linear Regression in terms of mean and accuracy. The mean accuracy of the KNN algorithm is better than the ANN algorithm. X-axis: KNN vs ANN, Y-axis: Mean accuracy of detection ± 1 SD