Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 21, No. 3, pp. 359–373, 2025

USED SAILBOAT PRICING RESEARCH BASED ON BP NEURAL NETWORK OPTIMIZED BY ANT COLONY ALGORITHM

Yuhao Piao, Qijia Zhou, Lijun Huang and Guodong Li

Abstract: We developed a sailboat pricing model that takes into account the different effects of different boat features on pricing and the important role of brokers and pricing strategies in the boat sales process. The main points of the approach are: first, we perform feature screening to filter out the main factors affecting pricing, then construct a BP neural network regression model to analyze the impact of various factors on pricing, and finally use a modified ant colony algorithm to optimize the parameters of the BP neural network regression model to produce results. Prior to this, we preprocessed the collected data and introduced dummy variables to improve the data quality, as the data substituted into the model is highly susceptible to noise, missing values, and inconsistent data. Next, we applied the model to evaluate the weights of various factors affecting the pricing of single-hull and double-hull vessels. After substituting different vessel types from different regions into the model, the results show that the weights predicted by our developed model are closer to the actual values than those predicted by the BP neural network model, the random forest model, the XGBoost model, and the LightGBM model, which proves the value of applying the model to sailboat pricing. Finally, we used the model to predict the prices of monohull and catamaran boats in different regions and showed the average prices in a box plot, which showed that the Caribbean was the cheapest among catamarans and the most expensive among monohull sailboats.

Key words: *BP neural network, regression factor analysis, ACO algorithm, pricing mode*

Mathematics Subject Classification: *68T05*

---

1 Introduction

Sailing boats, which use wind power to move across water, are an ancient mode of transportation that came after boats and rafts. Like many luxury items, the value of a sailing boat can fluctuate with age and market conditions, and there are many factors that can affect its pricing, such as the material of the hull, the length of the boat, and the area of the sail. These influencing factors are interconnected. Understanding the composition of boat prices and their influencing factors can help operators more comprehensively assess market trends and objectively measure the true value of sailing boats. Analyzing and predicting the trend of sailing boat prices and understanding the patterns of market fluctuations in specific regions can provide a reliable basis for business decisions. Therefore, determining the pricing of a sailing boat is an important consideration.

As the world's freight traffic increases and markets continue to grow, so does the study of commodity pricing. Lu et al. [20] use the binary tree model to comprehensively consider many factors such as geographical location, supporting facilities, and wear, and finally

 DOI: https:// doi.org/10.61208/pjo-2025-017

evaluate the second-hand house. Zhang et al. [32] introduced BP neural networks in the evaluation process, using actual historical transaction data on clarks on and training models to achieve the simulation of the evaluated boat prices. Xiao et al. [27] developed a BP neural network model to estimate used ship prices based on age, deadweight tonnage (DWT), newbuilding price and one-year charter rate. Roberto Bi Xiaowei et al. [1] based on factor analysis, a ranking method that affects the importance of the target is constructed. Feng Shan et al, [30] Yang Bo et al, [29] use Bp-MIV method to substitute features into the model, and the MIV values are used to filter important variables. However, this screening method is prone to overfitting. Gu et al. [7] Yi et al. [31] use logistics models to evaluate and predict the creditworthiness of businesses. Wu et al. [26] through main influencing factors affecting the characteristic value of second-hand residential buildings are decomposed and classified, and the feature price model is established on this basis, and the regression parameters are predicted by effective generalized least squares (FGLS). Li et al. [14] use XGBoost model to predict changes in second-hand house prices. Liu et al. [16] Wang et al. [25] price using importance sampling and least squares reduces the significance of the method. Meng et al. [11] Zhang et al. [33] use a typical generalized linear model, the impact of various influencing factors on commodity prices is described more accurately. Ding et al. [4] the particle swarm optimization algorithm is used to solve the problem of optimal pricing of goods under random demand. Sun et al. [21] use dynamic programming to model electricity price increases. Ning et al. [13] use the LightGBM (Gradient Boosted Tree) algorithm to predict pricing. Xiao et al. [28]compare the accuracy of multiple linear regression model and BP neural network model on pricing problems. Huang et al. [10] consider structured and unstructured data generation pricing based on deep learning techniques. Liu et al. [18] use the structural equation model to divide into two dimensions, internal and external, to explore the difference in pricing between second-hand and new goods. Zhou et al. [15] based on the time series analysis model, a random model for second-hand housing price prediction is established. M. Davison et al. [9] motivate a new stochastic electricity price model different in that it directly models price. Tang et al. [23] establish a housing pricing model for submarket division based on adversarial generative network is proposed. Sun et al. [22] using the prospect theory option pricing framework, the value function is introduced to describe the prospect value judgment of investors in the face of profit and loss. Fu et al. [6] Based on regression and time series cross-section regression of CAPM model used for stock pricing. He et al. [8] The ant colony algorithm is improved: the heuristic function is improved, and the distance between the optional node and the target point is introduced into the heuristic function. Liu et al. [19] ACO algorithm and GA algorithm are combined to improve ant colony algorithm by ACO-GA hybrid algorithm. Tong et al. [24] A chaotic ant colony algorithm based on dynamic volatile factors is proposed. As mentioned above, the pricing of used sailing boats is influenced by a number of factors, including the make, model, size, year, frequency of use and maintenance of the boat. In addition, prices vary considerably from region to region, and pricing methods vary as well. We used factor analysis to determine the main factors to ensure the accuracy of the study, and combined with neural network algorithms to build a model to obtain more informative results.

We developed an enhanced regression model using BP neural networks to analyze the factors influencing pricing. Initially, we pre-processed collected data through factor analysis to identify key factors. Dummy variables were then introduced to enhance data quality. Subsequently, we employed neural network regression and ant colony algorithms to evaluate factor weights impacting the pricing of single-hull and double-hull vessels. Our results indicate that our model is more accurate compared to the BP neural network, random forest, XGBoost, and LightGBM models. This underscores the model's efficacy in pricing sailboats. Additionally, we applied the model to predict prices for monohull and catamaran boats in different regions. Utilizing box plots to display average prices, we found that the Caribbean features the least expensive catamarans and the most expensive monohull sailboats. Our study demonstrates that introducing dummy variables significantly enhances model stability and adaptability.

2 Data Processing

2.1 Data Collection

In order to construct a model, it is necessary to gather data as the problem at hand lacks comprehensive information. To fully analyze the issue, we need to collect detailed data pertaining to the characteristics of a sailing boat. Fortunately, numerous websites dedicated to sailboat pricing offer us access to a wealth of information. For further insight, please refer to Table 1, which outlines our detailed data sources.

Table 1. Data source website

Database Names	Database Websites
Type, Engine_count	www.SailboatData.com
Price, Location, Material, Cabins, Motor_type	www.sailboatlistings.com
Length, Beam, Draft, Year	www.yachtworld.com

2.2 Data cleaning and processing

2.2.1 Removing irrelevant data

First of all, since there are missing values in the collecting class variables, e.g. Country/Region/State, random completions will affect the overall characteristics of the data, so these variables are excluded.

2.2.2 Handling outliers

To improve the accuracy and reliability of the data, it is important to detect and handle outliers using the Median Absolute Deviation (MAD) method. In the case of sailing data, outliers can be caused by measurement errors, equipment malfunctions, or other factors, and must be addressed to ensure the integrity of the data. Therefore, employing MAD detection and processing techniques is crucial.

$$MAD = median(\mid x_i - x_m \mid). \tag{2.1}$$

x_i is the value in the data sample set,x_i is the median of that sample set of data.

2.2.3 Min-max normalization

Considering the subsequent establishment of the intelligent model and factor analysis model, the continuous data such as Year, Length and Beam data, are normalized by min-max to reduce the inconsistency of the dimension. What's more, this accelerates the convergence speed of intelligent algorithms and improves the stability and robustness of the model.

$$x_i^{'} = \frac{x_i - \min(x)}{\max(x) - \min(x)}. \tag{2.2}$$

x_i is the ith sample eigenvalue in the original dataset,$x_i^{'}$ is the corresponding normalized value, $\max(x)$ and $\min(x)$ are the corresponding column maximum and minimum values.

3 Pricing Model Based on BP Neural Network Regression

3.1 Feature screening by factor analysis

First, KMO and Bartlett's tests (Table 2) should be conducted to determine if factor analysis is appropriate.

Table 2. KMO test and Bartlett's test

KMO values	0.786	
	Approximate cardinality	18497.713
Bartlett's test of sphericity	df	45
	P	0.000***
Note: ***, **, * represent 1%, 5%, 10% level of significance respectively		

From Table 2 KMO values between 0.7 and 0.9 indicate that this data is suitable for factor analysis.

Step1: Calculate the simple coefficient matrix R between the variables

$$R_{10\times10} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{110} \\ r_{21} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ r_{101} & \cdots & \cdots & r_{1010} \end{bmatrix}. \tag{3.1}$$

Step2: Find the eigenvalues $\lambda_1, \lambda_2, ...\lambda_{10}$ of R and the corresponding eigenvectors $\mu_1, \mu_2, ...\mu_{10}$

Step3: Calculation of factor loadings, determining the number of factors.

$$\begin{cases} A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{110} \\ a_{21} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{101} & \cdots & \cdots & a_{1010} \end{pmatrix} = \begin{pmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{\lambda_2} & \cdots & u_{110}\sqrt{\lambda_{10}} \\ u_{21}\sqrt{\lambda_1} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ u_{101}\sqrt{\lambda_1} & \cdots & \cdots & u_{1010}\sqrt{\lambda_{10}} \end{pmatrix}, \\ a_k = \sum_{i=1}^{k} S_i^2/p = \sum_{i=1}^{k} \lambda_i \Big/ \sum_{i=1}^{10} \lambda_i. \end{cases} \tag{3.2}$$

We can conclude that $k = 6$ by analyzing the results in Table 3, where a_k represents the cumulative contribution of feature k to the explained variance by d and $R's$ eigenvalues should be greater than 1.

Table 3. Total variance explained

Factor	Characteristic roots	Explanation of variance	Total explanation of variance
Year	3.091	30.913	30.913
Length	1.577	15.774	46.687
Draft	1.067	10.671	57.358
Beam	1.038	10.379	67.737
Location	1.017	9.424	77.161
Material	1.001	8.961	91.122
Motor type	0.602	6.02	94.142

From Table 3, the cumulative contribution of the first 6 factors is 91.122% > 90%, which can reflect the overall information more comprehensively.

Analysis of the Figure 1 gravel map shows that the slope is gentler by the 6th component, so the first 6 components can be selected as the principal components. Indicates that Year, Length, Draft, Beam, Location, Material were selected as the main factors.

Analysis of the Figure 2(factor loading matrix heat map) shows that the Length, Draft and Beam factors have large loading coefficients and can be identified as a particular component.

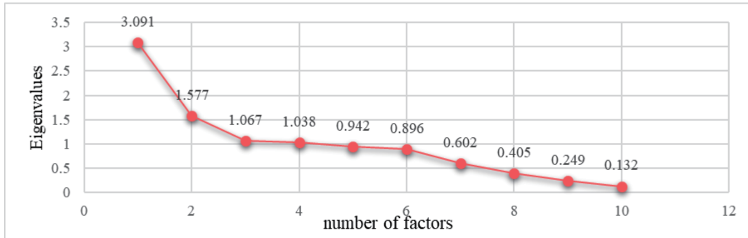

Figure 1: Rubble map

Figure 2: Factor load matrix heat map

3.2 BP neural network regression model

After determining the important parameters, this study opted for a BP neural network regression model due to the complex relationship between the pricing of different types of sailboats and their respective factors. This model has strong non-linear fitting capabilities and is suitable for various types of regression analyses. Additionally, the model automatically adjusts its parameters based on the data characteristics to achieve a better fit.

The flowchart for building a model is shown in Figure 3:

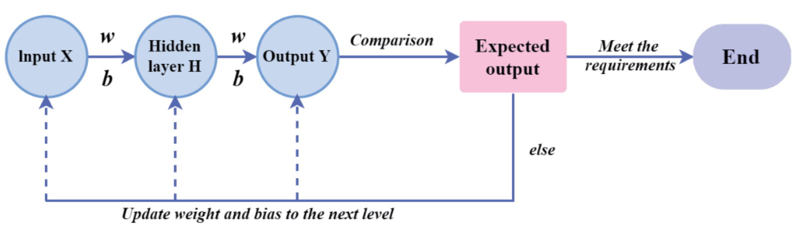

Figure 3: Model flowchart

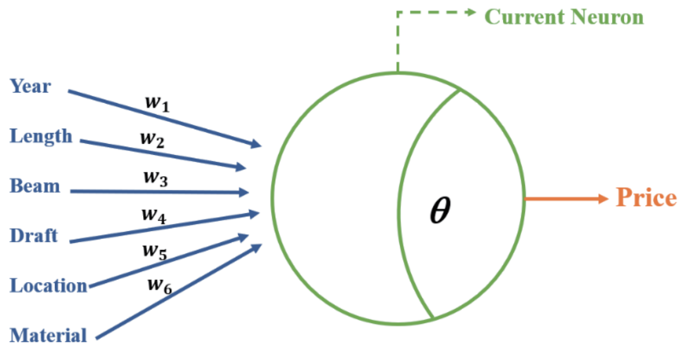

Figure 4: Neuron

3.2.1 Neuron model

From Figure 4 the model takes length, draft, year, beam, location, and material as independent variables and assigns a corresponding weight to each. The values of these variables are multiplied by their respective weights and processed through an activation function to obtain the final output y, which represents the price.

Output formula obtained:

$$y = f(\sum_{i=1}^{n} w_i x_i - \theta). \tag{3.3}$$

The θ is activation function, w_i is connection weight of the ith neuron.

3.2.2 Activation function

We determined the remaining parameters using the method of controlling variables, and obtained the following results by utilizing various activation functions.

Sigmoid:

$$\theta = \frac{1}{1 + e^{-z}}. \tag{3.4}$$

Tanh:

$$\theta = \frac{e^z - e^z}{e^z + e^z}. \tag{3.5}$$

Relu:

$$\theta = \max(0, x). \tag{3.6}$$

The top diagram represents “monohull” and the bottom diagram represents “catamarans”.From Figure 5,6,7 we ultimately decided to use the sigmoid function.Because sigmoid function loss of the function is smaller, the accuracy is higher.It is evident that for both “monohull” and “catamarans” boats, the optimal solution is obtained by iterating through the sigmoid function.

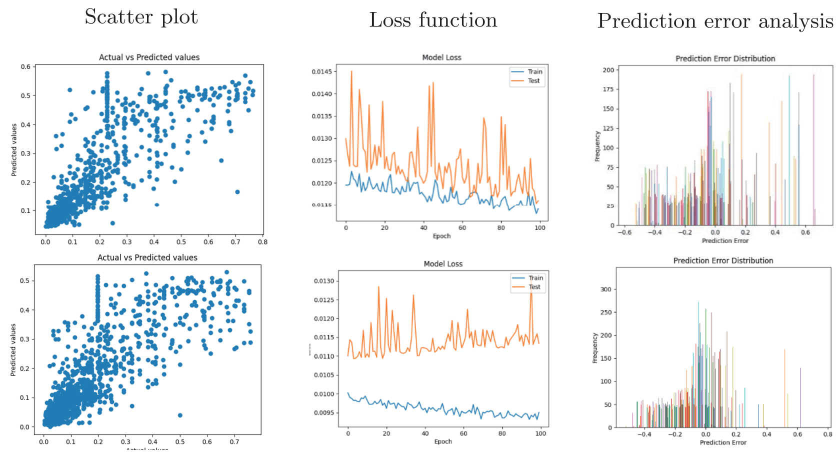

Figure 5: Comparison chart of sigmoid

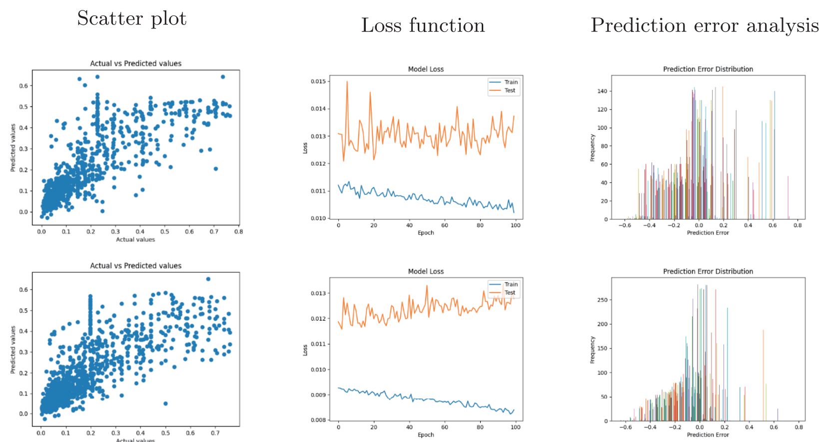

Figure 6: Comparison chart of Tanh

3.2.3 Basic framework

This model consists of three layers.
(1)Input layer: This layer serves as the input terminal for information, including variables such as length, draft, year, brand, location, and material.
(2)Hidden layer: This layer is responsible for processing the information and contains a total of “q” neurons. The formula for calculating the value of “q” is as follows:

$$N_s = \frac{N_s}{(\alpha^*(N_i + N_o))}. \tag{3.7}$$

Which N_i is number of neurons,N_o stand for output layer’s number of neurons,N_s represent training set’s number of samples,N_b is number of samples in the output set.

α: Arbitrary parameter, the range of ‘ l’ value is 2-10.

In this model, we take the value of q as 50.

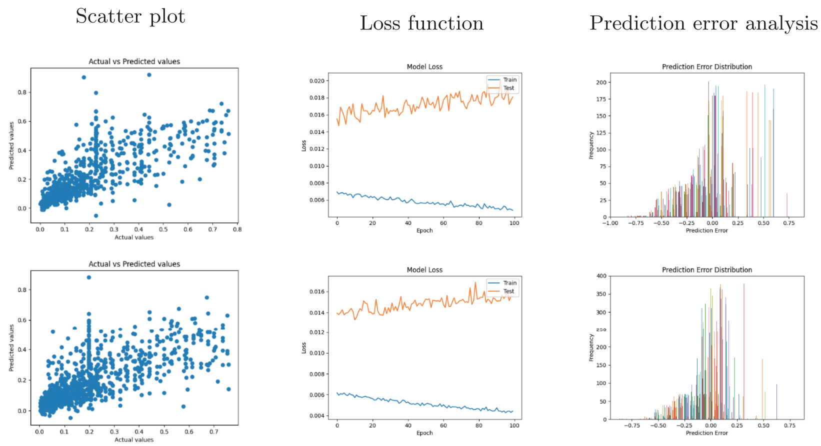

Figure 7: Comparison chart of Relu

From the above comparison, the activation function used this time is still sigmoid function.

(3)Output layer: This layer is used to output, which is the Listing Price we are looking for, since the output layer does not need to change state, we set the activation function to itself. The function is as follow:

$$W = \mathrm{x}'. \tag{3.8}$$

Which x_i is input from ith Neuron.

The first step is the forward propagation of information, followed by the second step, which is the backward propagation of information. During each propagation, the weights and biases from the hidden layer to the output layer, as well as those from the input layer to the hidden layer, are adjusted in sequence. The respective input formulas are as follows:

The input of the jth input neuron:

$$W_j = \sum_{h=1}^{q} w_{hj} b_h. \tag{3.9}$$

The input of the hth hidden layer neuron:

$$W_{\mathrm{h}} = \sum_{i=1}^{d} v_{ih} x_i. \tag{3.10}$$

3.2.4 Forward propagation

This is the process of entering information from the input layer into the network, and obtaining the final input layer result through the calculation of each layer. We multiply the value with the weight and bring it into the activation function, the formula is as follows:

From input layer to hidden layer:

$$W_{\mathrm{h}} = \sum_{i=1}^{d} v_{ih} x_i + \theta_h. \tag{3.11}$$

From hidden layer to output layer:

$$W_{\mathrm{j}} = \sum_{h=1}^{q} v_{hj} b_h + \theta_j. \tag{3.12}$$

3.2.5 Back propagation

Adjust the network parameters by calculating the error between the output layer and the expected value, so that we can becomes the error smaller. When the loss function is an error value, the formula for calculating the error is as follows:

$$E = \frac{1}{2} \sum_{k=1}^{2} \left(y_k - T_k\right)^2. \tag{3.13}$$

y_k represents the value of the output layer and T_k represents the expected value.For the weight update of backpropagation, in this model we use the adam algorithm to optimize the weight. This algorithm combines Momentum and Adaptive Learning Rate, Combines the average of past gradients with the learning rate in each direction. Finally, we get the updated weights:

$$W_t = W_{t-1} - \eta \nabla L(W_{t-1}). \tag{3.14}$$

W_t represents weight, L represents learning rate, and η represents momentum.

3.3 Ant Colony Algorithm Optimization

One disadvantage of the BP neural network is that it can easily stray into the local optimal solution because of its sensitive training response to starting weight and threshold setting. A population optimization-based search technique is the ant colony algorithm. It offers the benefits of strong global optimization, parallel computing, and positive and negative feedback mechanisms mixed with distributed computing. These may be readily coupled with the BP neural network technique to increase the prediction value's accuracy.

The optimized BP neural network's parameters, or the total number of weights and thresholds, are initially set to be calculated as m. Set these parameters as $p_i(1 <= i <= m)$, each of which has a random non-zero value of N, to form a set S_{pi}, and then we have k ants at the initial position of the anthill, and the ants a select any set of weights and thresholds in the set S_{pi}, and then select the next set in accordance with the probability formula below, until each ant has selected a combination of weights and thresholds in each of the m sets.

$$P_\alpha(\tau_j(S_{pi})) = \frac{\tau_j(S_{pi})}{\sum_{j=1}^{N} \tau_j(S_{pi})}. \tag{3.15}$$

Then we take the weights and thresholds chosen by each ant as the weights and thresholds of the BP neural network, input the training samples, calculate the error between the actual output and the desired output, sort the errors, find the smallest error, and the corresponding ants are the optimal solution, if the optimal solution satisfies the end conditions of the ant colony algorithm, then the optimal solution corresponds to the optimal combination of weights and thresholds for the BP neural network. If it does not satisfy it means that the pheromone allocation τ is unreasonable, in order to make the pheromone allocation more reasonable.The update rule for the pheromone is improved as Eq. (3.16).

$$\tau_j(S_{pj})(t+1) = (1-\rho)\tau_j(S_{pj})(t) + \rho \Delta \tau_j(S_{pj}). \tag{3.16}$$

When all ants complete one cycle of the full set,the pheromone is updated according to Eq. (3.17).

$$\Delta\tau_j(S_{pj}) = \sum_{a=1}^{k} \Delta\tau_j^{a}(S_{pj}) + \Delta\tau_j^{best}(S_{pj}). \tag{3.17}$$

Until the number of iterations N of the ant colony algorithm meets the maximum number of iterations, end the cycle, and finally get the combination of weights and thresholds, and continuously adjust the BP neural network connection weights and thresholds until the set error requirements or the maximum number of iterations are reached, and get the best BP neural network model.

The following are the stages for integrating the model for simplicity of understanding:

Step1: After initializing the BP neural network's weights and biases, the ant colony algorithm's parameters—such as the number of ants and the rate of pheromone volatilization —are set.

Step 2: The ant colony algorithm's search space is mapped to the neural network's parameters. Each ant's fitness is then determined, and the neural network's error value serves as the fitness index.

Step 3: The fitter ants release more pheromones to lead other ants, and the ant colony algorithm adjusts pheromones based on fitness.

Step 4: The neural network's weights and biases are updated using the pheromones that the ant colony algorithm provided.

4 Analysis of Results

Due to the numerous weight values between neural networks, we need to list the sum of weight vectors. The specific weight values can be found in Figure 8.

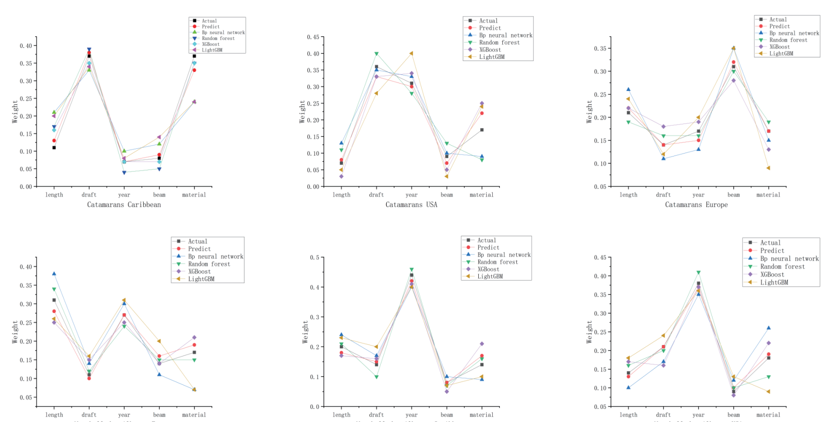

Figure 8: Model Accuracy Comparison

As can be seen in Figure 8 that the accuracy of our model is higher compared to BP neural network model, random forest model, XGBoost model, LightGBM model.It is proved that our model works better in pricing

And we made a comparison of the accuracy of the model before and after the introduction of dummy variables. The resultant images are shown in Figure 9:

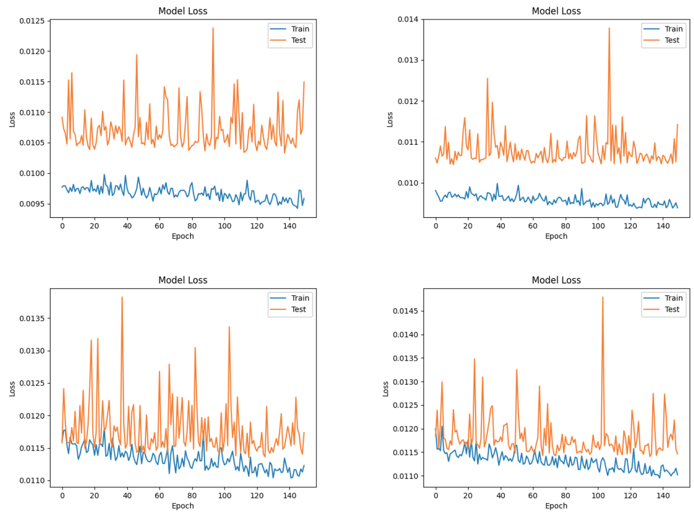

Figure 9: Model Loss

The top two pictures represent “Monohulled Sailboats” and the bottom two represent “Catamarans”. Figure 9 Comparison of accuracy before and after the introduction of dummy variables It can be concluded that from Figure 9.The number of outliers is reduced, the final converged loss function is smaller, and the model convergence speed is significantly faster by introducing the region dummy variables. Therefore, the model accuracy is higher after adding the region dummy variables.

Finally, we forecast the prices of Catamarans and Monohulled Sailboats for the three regions of Caribbean, Europe, and USA, and the results are shown in Figure 10: of dummy variables. The resultant images are shown in Figure 9:

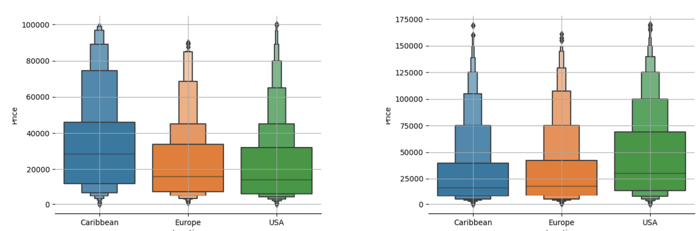

Figure 10: Boxplots show differences in prices across different regions

The results show that among catamarans, the Caribbean is the cheapest, while among monohulls, the Caribbean is the most expensive. People in the Caribbean value speed and manoeuvrability more, which is closely related to the Caribbean’s geographic location near the ocean, and this validates the reasonableness of our model from a practical point of view.

USA Monohulled Sailboats are the cheapest and Catamarans are the most expensive, indicating that people in the USA focus more on the stability of their boats, which allows them to sail faster downwind and be used for transporting lighter cargo at sea, which also indicates that people in the USA value economic development and justifies our conclusions.

5 Sensitivity Analysis

For our sailboat pricing model, although we have identified the important parameters through factor analysis, we still need to determine which of the six factors is most important to potential buyers.

We therefore performed sensitivity analyses for both the model before and after the introduction of dummy variables .The results are shown in Figure 11:

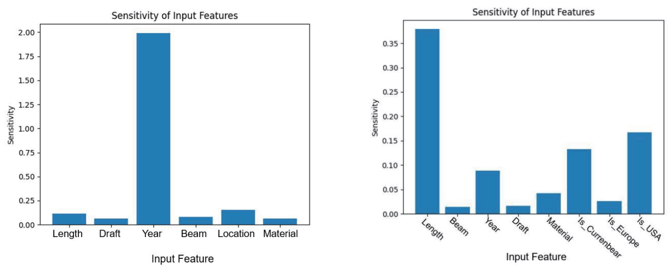

Figure 11: Sensitivity Analysis chart

In the first model, the sailboat's price is highly sensitive to the year of manufacture. This suggests that the year of manufacture is a crucial factor for potential buyers, as older boats are more likely to have maintenance issues and may exceed their budget for upkeep. This finding indicates that the model aligns well with real-world scenarios.

For the second model, the introduction of dummy variables has reduced the sensitivity to the year of manufacture, as well as the sensitivity of all parameters to below 0.4. This indicates that the model has achieved its intended goal of stability. Additionally, the sensitivity of Europe is lower than the other two regions, which is consistent with the box plot analysis mentioned earlier, where Europe appears to be in the middle position in all scenarios.

6 Conclusion

In this study, we successfully developed a sailing ship pricing model that takes into account the different effects of different ship characteristics on pricing, as well as the important roles of brokers and pricing strategies in the ship sales process. Our research methodology consists of the following steps: first, filtering out the main factors affecting pricing through a feature screening strategy; second, constructing a BP neural network regression model to analyze the effects of various factors on pricing; and finally, optimizing the parameters of the BP neural network regression model using an improved ant colony algorithm to produce the final results. Before constructing the model, we preprocessed the collected data and introduced dummy variables to improve the data quality, as the data used in the model is susceptible to noise, missing values, and inconsistent data. We then applied the model to assess the differences in the weights of factors affecting the pricing of monohulls and catamarans. By incorporating ship types from different regions into the model, the results show that the weight values predicted by our developed model are closer to the actual values compared to the BP neural network model, the random forest model, the XGBoost model, and the LightGBM model, thus validating the model's value for sailing ship pricing. Finally, we used the model to predict the prices of monohulls and catamarans in different regions and showed the average prices in a box-and-line plot. The results show that catamarans (catamarans) have the lowest prices in the Caribbean, while monohulled sailboats (Monohulled Sailboats) have the highest prices. In addition, we demonstrate that the introduction of dummy variables greatly improves the stability of the model and makes it highly adaptable. In conclusion, this study provides useful references and lessons for the development and optimization of ship pricing models.

References

[1] X.W. Bi, M.M. Liu, D.S. Fang and Y. Li, Target threat ranking method based on factor analysis, *Electron. Inf. Countermeas. Technol.* 38 (2023) 1–7.

[2] D. Di Caprio, A. Ebrahimnejad, H. Alrezaamiri and F.J. Santos-Arteaga, A novel ant colony algorithm for solving shortest path problems with fuzzy arc weights, *Alexandria Eng. J.* 61 (2022) 3403–3415.

[3] M. Davison, C.L. Anderson, B. Marcus and K. Anderson, Development of a hybrid model for electrical power spot prices, *IEEE Trans. Power Syst.* 17 (2002) 257–264.

[4] T.Q. Ding, P. Tian and Z.Y. Tian, Perishable commodity pricing model and particle swarm solution method, *Comput. Eng. Appl.* 4 (2005) 230–232.

[5] M. Dorigo and T. Stützle, Ant Colony Optimization: Overview and Recent Advances, Springer, Cham, 2019.

[6] Z.G. Fu, An empirical study on CAPM model: Based on Chinese stock market data, *Hebei Enterp.* 11 (2023) 62–64.

[7] Z.L. Gu, Research on comprehensive evaluation method of enterprise public credit based on logistic model, *Inf. Syst. Eng.* 353 (2023) 132–134.

[8] M.Y. He, X.Y. Hu, J.H. Zhang, P.S. Yao, F.D. Liu and Y.X. Men, Tower crane path planning based on improved ant colony algorithm, *J. Meas. Sci. Instrum.* 15 (2023) 1–10.

[9] W. Hou, Z. Xiong, C. Wang and H. Chen, Enhanced ant colony algorithm with communication mechanism for mobile robot path planning, *Robotics and Autonomous Systems* 148 (2022): 103949.

[10] H. Huang, Research on commodity automatic pricing model based on deep learning, *Mod. Commer. Trade Ind.* 40 (2019) 188–190.

[11] S.W. Meng, Application of generalized linear model in automobile insurance pricing, *Math. Pract. Theory* 45 (2015) 100–105.

[12] C. Miao, G. Chen, C. Yan and Y. Wu, Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm, *Computers & Industrial Engineering* 156 (2021): 107230.

[13] K.N. Ning, Recommendation Analysis of Second-Hand Commodity Price of e-Mall Based on LightGBM, Tianjin, 2019.

[14] J.X. Li, Research on Second-Hand Housing Transaction Price Prediction Based on Machine Learning Model, Yanji, 2022.

[15] Z.T. Li, Dynamic pricing of second-hand cars based on artificial neural network, *Int. J. Comput. Sci. Eng.* 10 (2022) 113–118.

[16] G. Liu and G.D. Gu, Research on American-style option pricing based on Monte Carlo importance sampling method, *Stat. Decis.* 9 (2014) 162–164.

[17] H. Liu, Research on cloud computing adaptive task scheduling based on ant colony algorithm, *Optik.* 257 (2022): 168677.

[18] Z.H. Liu and S. Chen, Structural equation model of used laptop pricing, *Ergonomics* 16 (2010) 27–29+73.

[19] Z.Y. Liu, In the application of TSP problem based on improved ant colony algorithm, *Sci. Technol. Innov.* 23 (2023) 166–171.

[20] T.Y. Lu, Binary tree option pricing model for second-hand house valuation, *Market Weekly* 33 (2020) 56–59.

[21] B. Sun and Q.S. Shi, Analysis of discharge electricity price pricing model for electric vehicle grid, *East China Electr. Power* 39 (2011) 1029–1032.

[22] Y.F. Sun and W.Y. Peng, Behavioral option pricing based on prospect theory framework and Heston model, *J. Guangdong Univ. Technol.* 40 (2023) 1–8.

[23] C.J. Tang, P. Zhang, X.L. Bao and F.X. Xu, GAN-BPM: A submarket segmentation housing pricing model based on GAN, *J. Chongqing Univ. (Nat. Sci. Ed).* 46 (2023) 1–11.

[24] L. Tong, J. Yang, X.S. Gan, D. Shen, W. Yang and C. Chen, More chaos ant colony algorithm based on improved machine conflict free simulation, *J. Syst. Simul.* 37 (2023) 155–166.

[25] J.A. Wang, Application of importance sampling in option pricing, *J. Jimei Univ. (Philos. Soc. Sci).* 4 (2007) 35–39.

[26] H.Y. Wu. Research on North-South Difference of South-South Premium of Second-hand Residential Houses Based on Feature Price Model, Shijiazhuang, 2022.

[27] Q.J. Xiao and Y. M. Zhang, Based on BP neural network of secondhand ship price assessment study, *J. Marine Eng.* 35 (2013) 100–103.

[28] J.L. Xiao, Modern manufacturing technology and equipment, *Mod. Manuf. Technol. Equip.* 3 (2015) 92–94.

[29] B.Y. Zhang, Y.G. Xiao and Y.Z. Zeng, A comparative study on the importance measurement of risk factors in auto insurance pricing—Based on ensemble learning method and generalized linear regression model, *Insur. Res.* 10 (2019) 73–83.

[30] Z.Y. Zhang, A pricing model for used vehicles based on genetic algorithms, *J. Automobil. Technol.* 12 (2017) 89–92.

[31] Y. Zheng, L.W. Chen, D.L. Sun and A.X. Zhang, Study on the influencing factors of China's architectural heritage protection and management based on BP-MIV, *New Archit.* 6 (2022) 82–86.

[32] H.W. Zhou, Pricing model of second-hand housing in Nanjing based on time series, *J. Sci. Technol. Teach. (Zhongxin J)*. 29 (2016) 141–142.

[33] Z.W. Zhou, Research on second-hand housing transaction pricing model based on big data, *J. Comput. Appl. Math.* 43 (2022) 209–215.

Manuscript received 18 August 2023
revised 3 March 2024
accepted for publication 6 May 2024

Yuhao Piao
School of Mathematics and Computing Science
Guilin University of Electronic Technology
Guilin, Guangxi 541001, China
E-mail address: pyh0523@126.com

Qijia Zhou
School of Mathematics and Computing Science
Guilin University of Electronic Technology
Guilin, Guangxi 541001, China
E-mail address: 913278533@qq.com

Lijun Huang
School of Mathematics and Computing Science
Guilin University of Electronic Technology
Guilin, Guangxi 541001, China
E-mail address: 1493836607@qq.com

Guodong Li
School of Mathematics and Computing Science
Guilin University of Electronic Technology
Guilin, Guangxi 541001, China;

Guangxi Colleges and Universities Key Laboratory of
Data Analysis and Computation;

Center for Applied Mathematics of Guangxi (GUET)
Guilin University of Electronic Technology
Guilin 541002, China
E-mail address: lgdzhy@126.com