DYNAMIC FREIGHT PRICING BY ARTIFICIAL INTELLIGENCE AND MATHEMATICAL HYBRID SYSTEM

Tuğçe ELÇİ, Ahmet Yesevi TÜRKER, [1]Hasan GÜNEY, Ahmet USTAOĞLU, Deniz KANTAR

Borusan Logistics, Türkiye, tugce.elci@borusan.com, ahmetyesevi.turker@borusan.com, hasan.guney@borusan.com, ahmet.ustaoglu@borusan.com, deniz.kantar@borusan.com

https://doi.org/10.37904/clc.2023.4859

Abstract

In domestic road operations, the demand for increased efficiency and performance is on the rise. To enhance operational efficiency and minimize costs, it becomes crucial to generate dynamic pricing for routes through digital media. The primary objective, in line with our developed model, is to boost the profit margin by estimating freight prices on our digital platform based on factual data. These estimates are then transmitted to the operational units to reduce current freight costs. Our digital platform serves as a nexus for suppliers and customers in land transportation. In freight transportation, it's vital to provide cost-effective and dependable methods for locating cargo, aligning cargo owners with the right carriers, and quickly determining precise route pricing. We have established a digital learning system for freight pricing estimation using machine learning techniques and a mathematical model. We evaluated 41 regression models to select the artificial intelligence models for our prediction algorithm. Four regression models that demonstrated the best performance were chosen to build our artificial intelligence system. Acknowledging the combination algorithm developed by our company, the system operates within a continuous learning cycle and selects the most effective artificial intelligence models at specific intervals. Additionally, the mathematical model's purpose is to complement the results of artificial intelligence by using actual prices from similar routes to prevent potential errors. The degree to which the mathematical model refines the artificial intelligence's results is determined by our developed decision-making intelligent algorithm. The hybrid model's results were tested on real expeditions, confirming its success.

Keywords: Artificial intelligence, logistics, dynamic pricing, mathematical model, machine learning

1. INTRODUCTION

The logistics industry has a vital importance for the world economy. In recent days, e-commerce and last-mile delivery, technology and automation, supply chain resilience, sustainability, etc. are important subtopics of the logistics sector [1]. While these subtopics are getting more popular, digitalization is beginning to become widespread. In the logistics industry, freight prices are mostly determined based on expert opinions and historical information. The decision of freight price is impressed by similar and past freights in the current marketplace [2]. In the circumstances, it is possible to appear the discussion between suppliers and shippers because of the high cost and commissions. In addition, the preparation of the freight price according to the opinions of authorized people in the market also supports the discussions [3].

This paper proposes a dynamic freight price prediction method based on machine learning algorithms and mathematical models. This hybrid method includes machine learning algorithms such as extreme gradient boosting (XGBoost) regression, light gradient boosting machine (LightGBM) regression, etc. The mathematical model also supports this hybrid system to increase the accuracy score of this hybrid system.

2. METHODOLOGY

Efficiency and performance need in domestic road transportation operations are continuously increasing. In this regard, generating dynamic pricing for routes through digital means a crucial role in improving efficiency and minimizing costs in operations. The primary objective is to predict freight prices within our digital platform, eTA, accurately and convey these predictions to the operations units. This is aimed at reducing the current freight costs in eTA, thereby increasing the buy-sell margin.

The eTA platform is a digital platform that connects cargo carriers and cargo owners in road transportation. Ensuring the economical and reliable provision of cargo finding and finding suitable carrier processes in cargo transportation, along with the fast and accurate pricing of the route, is vital for the platform. Machine learning techniques and mathematical models have been employed to establish a digital learning system for freight prediction. The mathematical model's purpose is to support and prevent possible incorrect results from artificial intelligence by using actual prices from similar shipments on the same route. The extent to which the mathematical model will adjust the results of artificial intelligence is determined by the developed decision-maker smart algorithm.

2.1. General Structure of Algorithm

When a shipment is initiated, it is priced in two different ways based on the characteristics of the shipment:

- **Unit-based pricing:** This pricing method is applied for shipments between locations A and B for the entire shipment.
- **Price per Ton Shipment:** This pricing method corresponds to the price for 1 ton of cargo for shipments between locations A and B.

The system flow as demonstrated in Figure 1 involves the operation of three different artificial intelligence models and a mathematical model. The shipment classification algorithm directs the shipment to the appropriate category based on its characteristics.

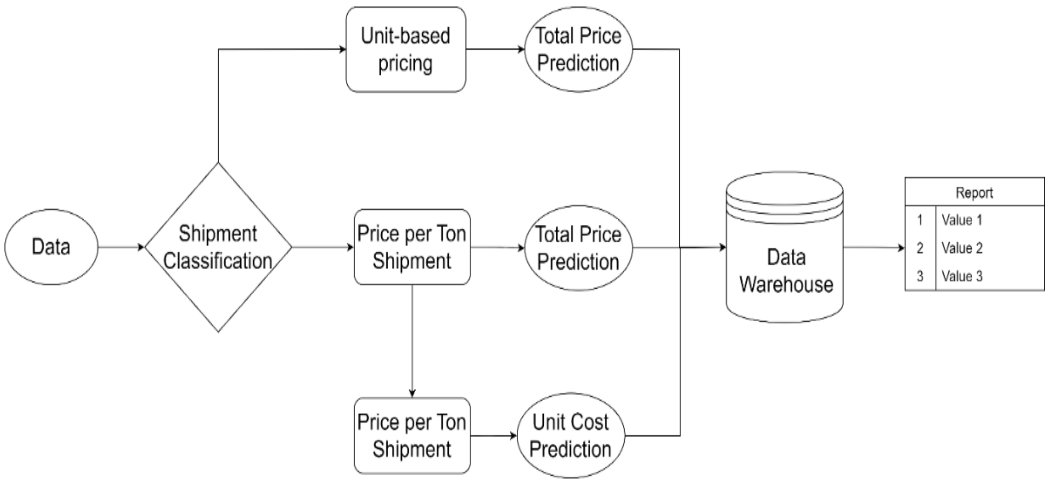

Figure 1 General structure of the system

Workflow:

The freight pricing system with detail structure as shown in Figure 2 consists of three main blocks:

1. Artificial Intelligence Models and Combinator
2. Mathematical Model and Calibrator
3. Decision Maker

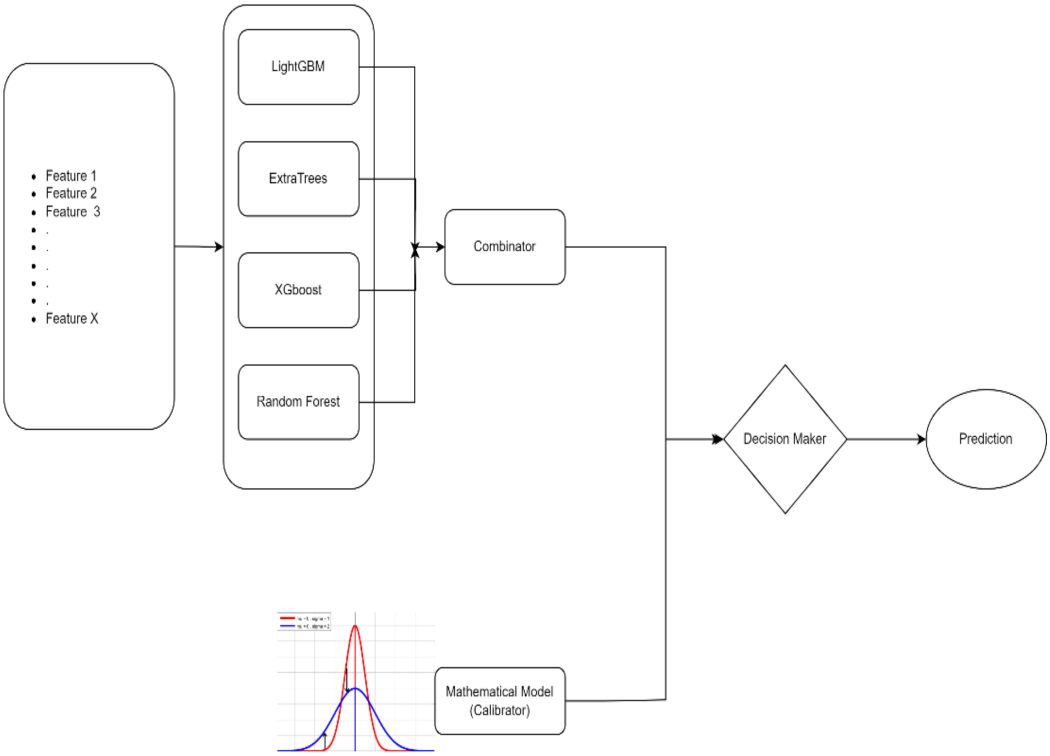

Figure 2 Detail structure and workflow of the system

2.2. Feature Selection

The freight price is associated with logistics parameters such as distance, fuel, highway toll, etc [4]. In logistics, the data have different relationships with each other. In this case, feature selection is the efficient and fertile method of data preprocessing operations for machine learning problems [5]. In this paper, data in eTA are analyzed. At the start point, 77 different features exist in the main cluster. In feature selection methods, filter methods are evaluated in this research. Correlation analysis is applied to limit the number of features in this cluster [6]. After this application, the number of features decreased to 22. These features are information on location (arrival/departure), distance, fuel price, dollar exchange rate, type of load, type of vehicle, etc.

2.3. The Combinator

Artificial intelligence models are in a continuous learning cycle and update themselves at certain time intervals with new incoming data. This feature allows the artificial intelligence to quickly converge towards accuracy, addressing seasonality as well as sharp inflation, fuel prices, and exchange rate changes, preventing an increase in error rates. In the artificial intelligence component, there are four independent artificial intelligence models, including Extratrees, LightGBM, Random Forest, and XGBoost Regressor [7,8,9]. All artificial intelligence inputs go to all four artificial intelligence models, and each artificial intelligence model produces results. These results are subjected to an error calculation based on the data sets of the most recent days. This calculation uses Root Mean Square Error (RMSE) because RMSE is a standard statistical metric that can be used to measure model performances [10].

The working system of The Combinator is as follows:

$$Lowest\ RMSE \leq Models \leq \ Lowest\ RMSE + Std\ Dev\ of\ Models$$

The Combinator algorithm creates RMSE scores based on the specified data set and calculates the standard deviation of the generated scores. According to these calculations, it selects the models with the lowest RMSE scores, that is, the models that work best, as shown in the equation. Then, the results of the selected models are averaged to obtain the final result. In this way, this process is renewed in each learning cycle, and the chosen model changes according to the success of the artificial intelligence algorithms in each cycle.

2.4. The Calibrator

In addition to The Combinator, the mathematical model called The Calibrator is designed to give robustness to the system and compensate if there is any problem caused by wrong data or any sudden change in fuel price or inflation. This system design is as follows. First, when the new freight is created to be valued, The Combinator takes "Departure City", "Departure County", "Arrival City", "Arrival County", "Freight Type" and "Vehicle Type" as 6 parameters and filter previous freights which are identical in terms of these parameters. The interval is 30 days, and the past recent days have more impact on the result. With these filtered identical freights, the time-based average is taken after anomaly situations (freights having beyond +2 standard deviation in price) are trimmed.

However, in some cases, freight of a less common type may occur on a specific route. Due to the application of strict filters in such situations, the calibration process does not work properly due to a lack of data and does not correct the artificial intelligence algorithm. In these cases, as a solution, filters are gradually reduced. If the number of freights is less than 3, filters are gradually renewed depending on the type of freight. The Vehicle Type and Freight Type are removed from the filter, and the process is rerun.

2.5. The Decision Maker

The decision-maker algorithm gives a weighted average of the result of the previously calculated combination and calibrator algorithms. The working logic of the algorithm is to calculate how much the mathematical model will correct the error rate of artificial intelligence and it simply takes the weighted average of the Kal (Calibrator) and Kom (Combinator) as shown in the Equation 1. This process step occurs as a result of the mathematical formulation developed within the company. This formulation is as follows:

$$K(Len) = \frac{(\alpha\frac{Len.Kal}{10} + Kom)}{(\alpha\frac{Len}{10} + 1)} \tag{1}$$

Where:

Len - The number of data that occurred previously on this route.

Kal - Result of the Calibrator algorithm.

Kom - The result of the Combinator algorithm.

α - Constant

The constant α is calculated using the number of trips that have previously occurred on that route (Len) and is represented as follows:

$$a = \begin{cases} 4 & , \ Len < 10 \\ 5 & , \ 10 < Len < 50 \\ 6 & , \ Len > 50 \end{cases} \tag{2}$$

The constant α shown in Equation 2 changes according to the Len variable is due to the increase in confidence in the prices given if there are many trips on that route.

The Decision Maker is an algorithm that operates based on the results of the Combinator and the Calibrator algorithms. If the difference between the results of the combination and calibration algorithms exceeds the specified threshold value, the result is calculated using the equation above. If there is a difference below the specified threshold value between the results, meaning the results are close to each other, then at this stage, the average of the two results is directly used as the final result. Thus, by minimizing errors and deviations, a price forecast closer to reality is achieved.

3. EXPERIMENTAL PART AND RESULT

After the information of the newly created shipment is entered, firstly, The Combinator which uses trained AI models and combinates them, gives an estimated shipment price. The overall performance on 4056 shipments can be seen in Table 1. The combinator gives a 95% approximation for more than half of the shipment. And at 80% convergence in estimation, almost 90% of total shipments fit in it. These performances are significantly higher than the Manuel System. Additionally, when the calibrator was run, the estimated freight price became closer to the actual price. In this direction, the success rate of 72.57% with 95% accuracy before the calibrator increased to 95% success rate with 80% accuracy.

Table 1 Accuracy Score Table

	%95 Accuracy	**%90 Accuracy**	**%85 Accuracy**	**%80 Accuracy**	**Number of Shipment**
Manuel System (ETA)	%51.5	%72.51	%82.61	%87.35	4056
The Combinator	%60.32	%81.78	%87.25	%89.49	4056
The Decision Maker	%72.57	%86.57	%92.45	%95.95	4056

To illustrate that, when the price is 100 dollars, in both cases we generate 110 or 90 dollars, we have a 10% percent divergence in this estimation. Therefore, the performance is also looked at in this way.

4. DISCUSSION

Nowadays, researchers generally use machine learning algorithms such as Random Forest Regressor, XGBoost Regressor, Artificial Neural Networks, etc. to predict freight prices [11,12]. These models generally work but, in some cases, the constant model could not give the high accuracy score or low RMSE value. Also, this paper proposes a combination of machine learning models and mathematical models. The model selection system ensures that the best model suitable for the data works without being dependent on a single model, thus producing predictions that are closer to reality. Furthermore, the mathematical model comes into play when artificial intelligence encounters a deadlock and suggests to the user the pricing closest to the real values.

5. CONCLUSION

In conclusion, the development of a new algorithm aligned with the company's needs and structure, utilizing Borusan's internal resources, has revolutionized the way pricing and driver-customer pairing are achieved through the ETA platform. By offering a reliable and route-based stable pricing mechanism, this algorithm has introduced transparency for both drivers and customers.

This system, powered by artificial intelligence, is remarkably resilient to sudden inflation and currency fluctuations, dynamically adapting to current market prices. Furthermore, its seamless integration with the operational shipment creation system used within the business unit, driven by real-time data, positions it as a valuable asset that grants a competitive advantage against imported counterparts. It not only streamlines data preparation and execution but also significantly reduces response times to customer demands.

As a result, this innovation is poised to reduce the company's workload and enhance its responsiveness to customer requests. The AI-driven algorithm has not only met but exceeded expectations, proving to be a game-changer in the industry, ultimately setting a new standard for pricing, reliability, and operational efficiency.

ACKNOWLEDGEMENTS

Tuğçe Elçi, Deniz Kantar, Ahmet Yesevi Türker, Hasan Güney and Ahmet Ustaoğlu are members of Data and Artificial Intelligence Services in Borusan Logistics.

REFERENCES

[1] LEE, T.-H.; HEO, J.-S. 2021 Logistics Industry Outlook; Issue Paper 2021-02; The Korea Transport Institute: Sejong, Korea, 2021.

[2] DO, K.-H. A Study on the Problems and Improvement of the THC, CAF & BAF in Korea Ocean Freight; Konkuk University Graduate School: Seoul, Republic of Korea, 2009.

[3] JANG, H. S., CHANG, T. W., & KIM, S. H. (2023, January 6). Prediction of Shipping Cost on Freight Brokerage Platform Using Machine Learning. Sustainability, 15(2), 1122. https://doi.org/10.3390/su15021122

[4] KOVÁCS, G. (2017, April 26). First Cost Calculation Methods for Road Freight Transport Activity. Transport and Telecommunication Journal, 18(2), 107–117. https://doi.org/10.1515/ttj-2017-0010

[5] LI, J., CHENG, K., WANG, S., MORSTATTER, F., TREVINO, R. P., TANG, J., & LIU, H. (2017, December 6). Feature Selection. ACM Computing Surveys, 50(6), 1–45. https://doi.org/10.1145/3136625

[6] JOVIC, A., BRKIC, K., & BOGUNOVIC, N. (2015, May). A review of feature selection methods with applications. 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). https://doi.org/10.1109/mipro.2015.7160458

[7] MARTIELLO MASTELINI, S., NAKANO, F. K., VENS, C., & DE LEON FERREIRA DE CARVALHO, A. C. P. (2023). Online extra trees regressor. IEEE Transactions on Neural Networks and Learning Systems, 34(10), 6755–6767. https://doi.org/10.1109/TNNLS.2022.3212859

[8] ZHANG, P., JIA, Y., & SHANG, Y. (2022). Research and application of XGBoost in imbalanced data. International Journal of Distributed Sensor Networks, 18(6), 155013292211069. https://doi.org/10.1177/15501329221106935

[9] COULSTON, J. W., BLINN, C. E., THOMAS, V. A., & WYNNE, R. H. (2016). Approximating prediction uncertainty for random forest regression models. Photogrammetric Engineering and Remote Sensing, 82(3), 189–197. https://doi.org/10.14358/pers.82.3.189

[10] CHAI, T., & DRAXLER, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250

[11] KULKARNI, P., GALA, I., & NARGUNDKAR, A. (2023). Freight cost prediction using machine learning algorithms. In Lecture Notes in Electrical Engineering (pp. 507–515). Springer Nature Singapore.

[12] YANG, Z., & MEHMED, E. E. (2019). Artificial neural networks in freight rate forecasting. Maritime Economics & Logistics, 21(3), 390–414. https://doi.org/10.1057/s41278-019-00121-x