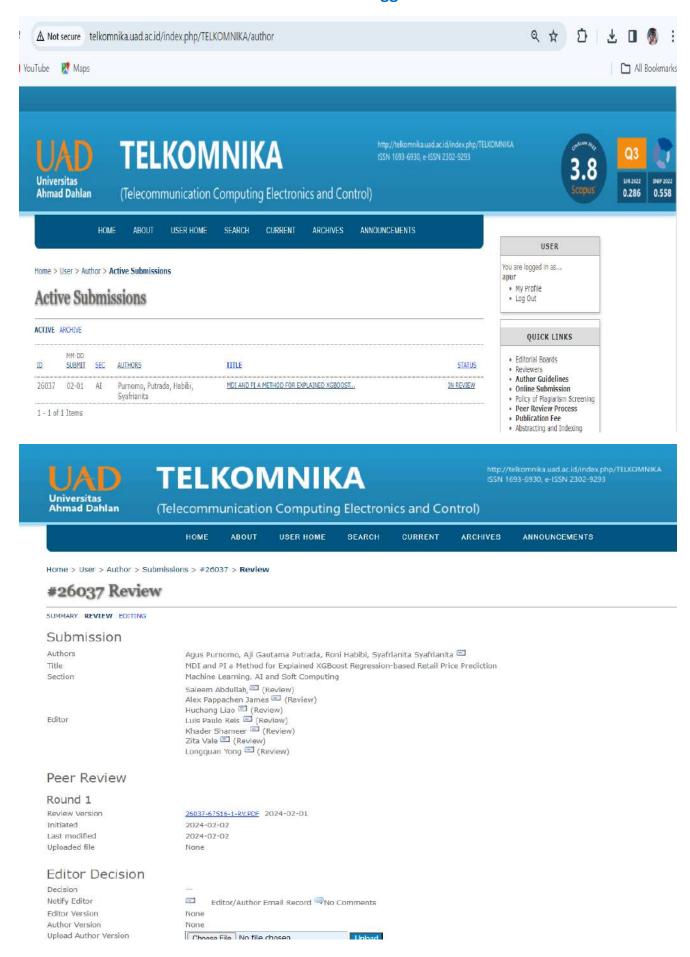
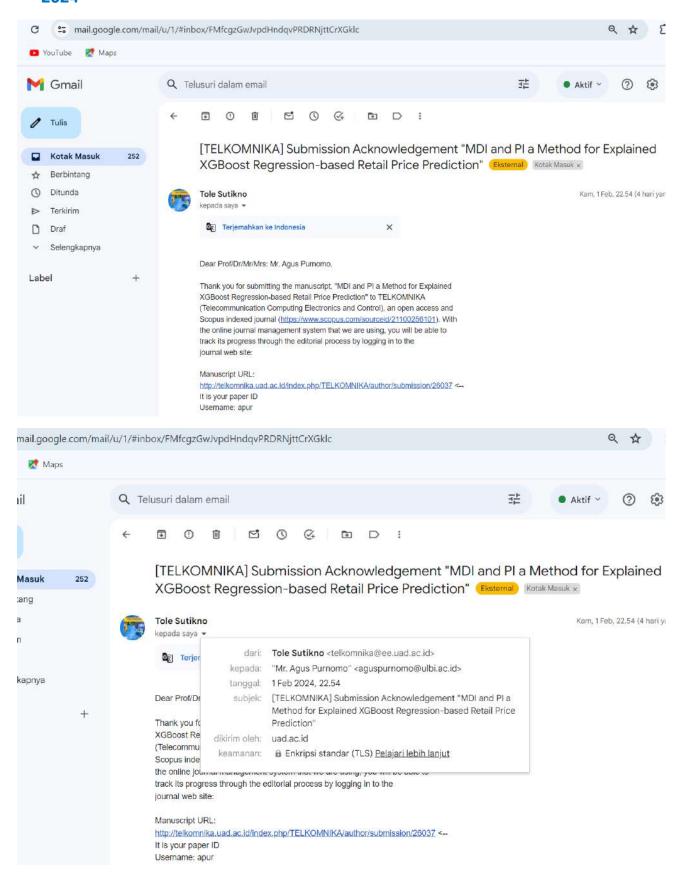
# KOMUNIKASI DENGAN EDITOR JURNAL TELKOMNIKA, Scopus Q2

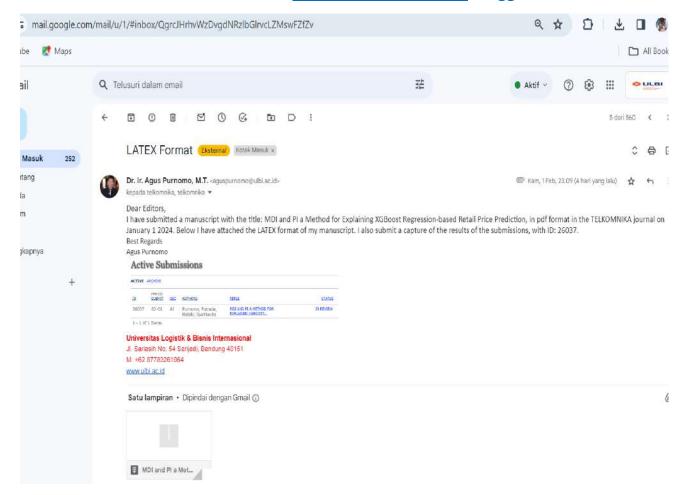
1. Submissions ke OJS Jurnal TELKOMNIKA tanggal 1 Februari 2024



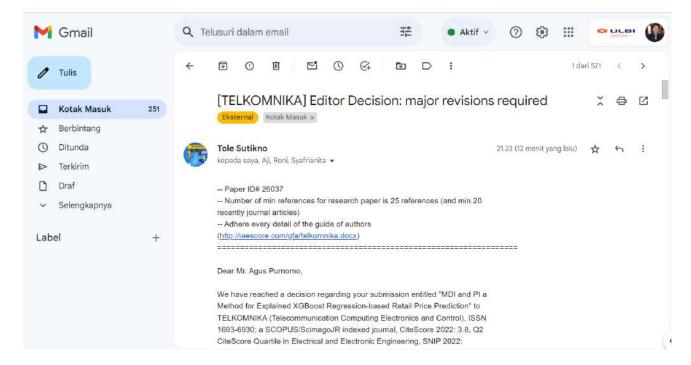
# 2. Notifikasi ke email tentang Submissions ke OJS TELKOMNIKA, tanggal 1 Februari 2024

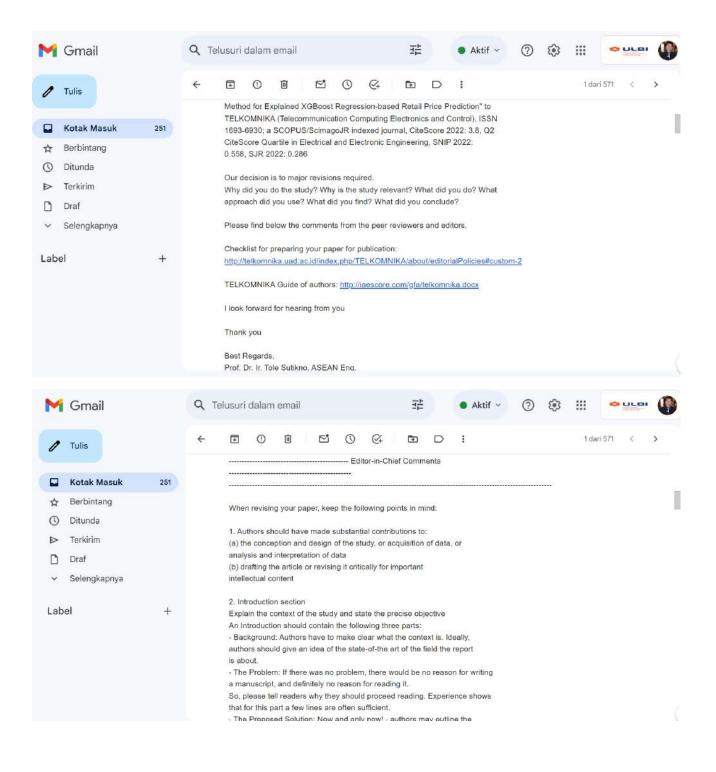


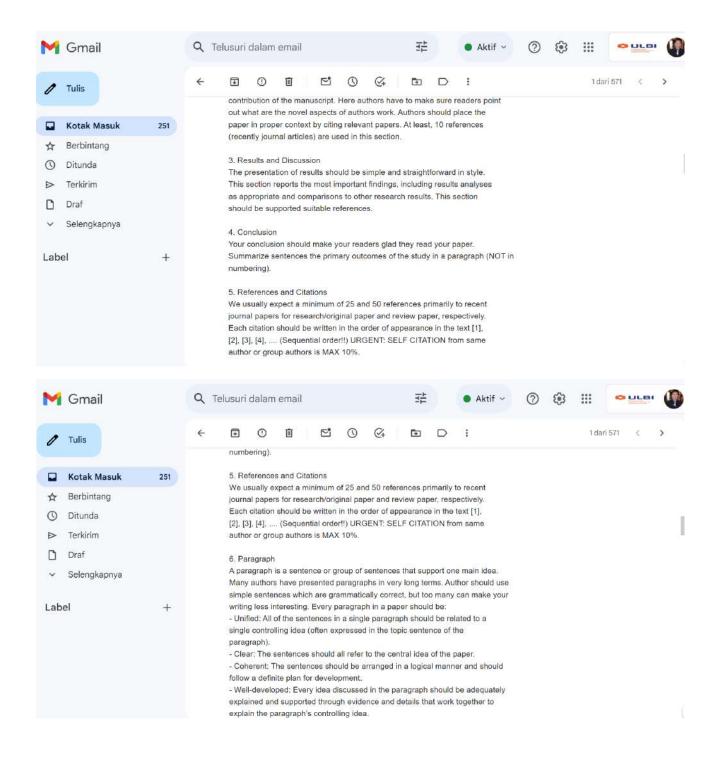
# 3. Submit LATEX Format ke email: telkomnika@uad.ac.id tanggal 1 Februari 2024

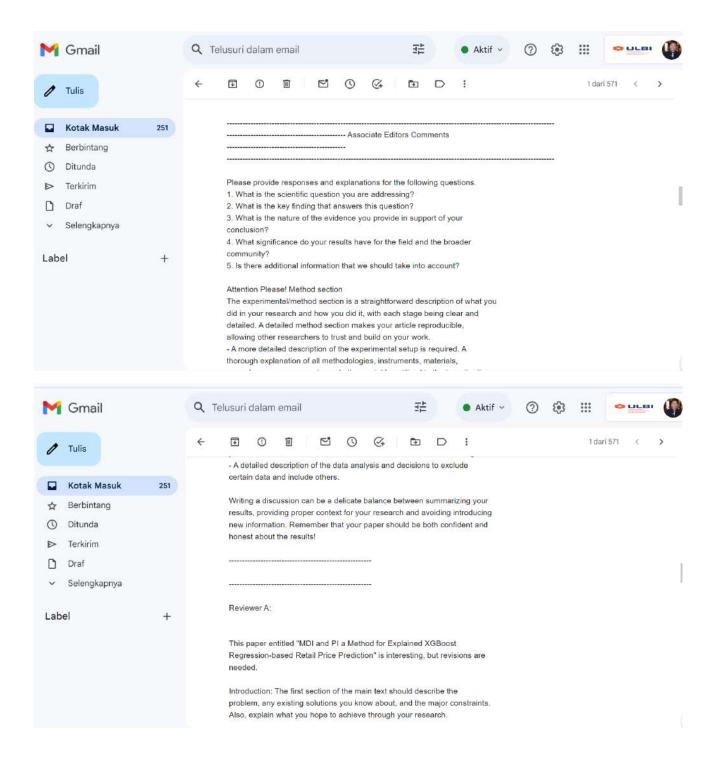


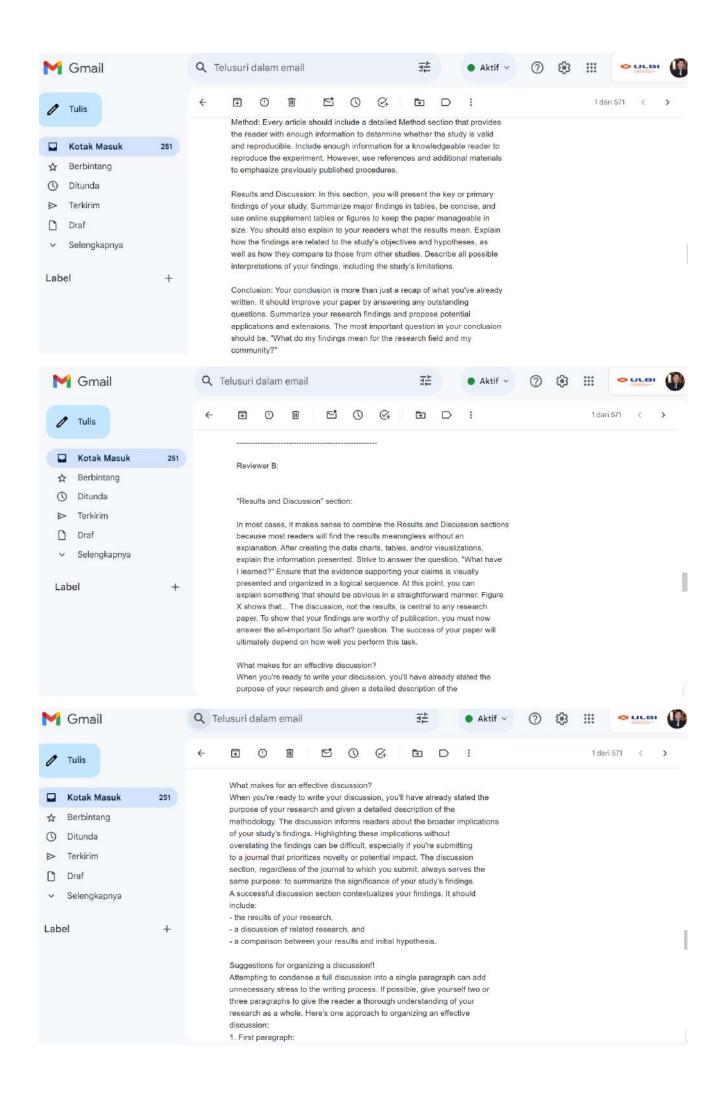
# 4. Email dari Editor Decision: major revisions required, tgl 24 Februari 2024

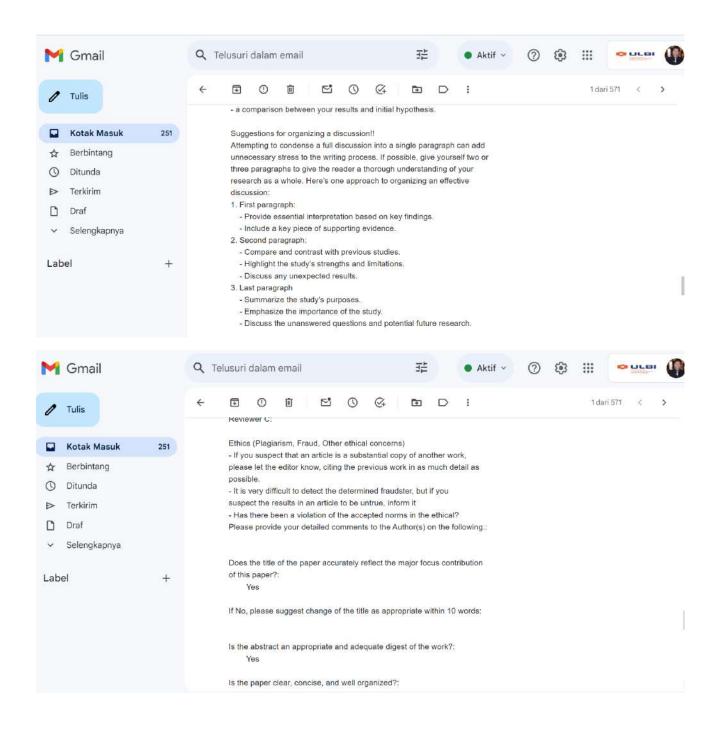


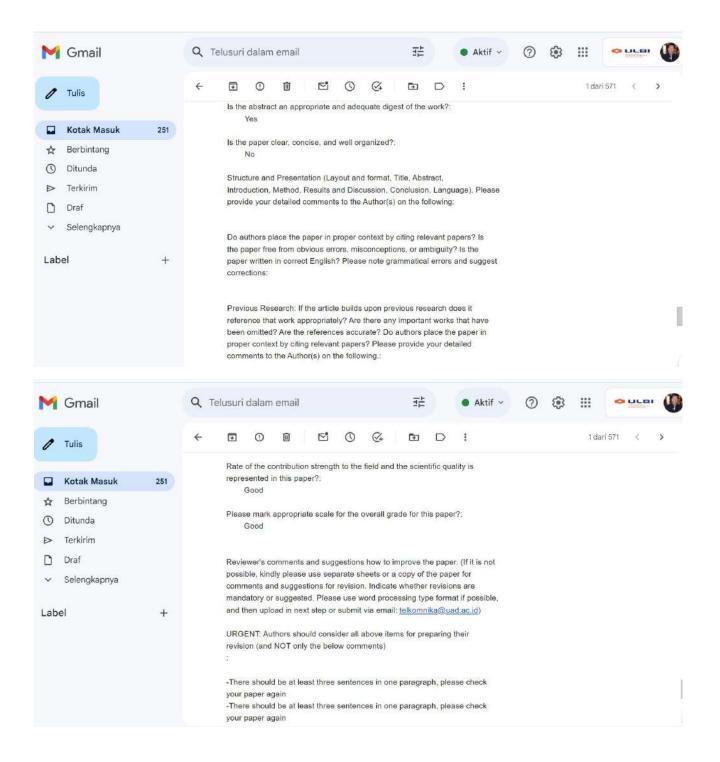


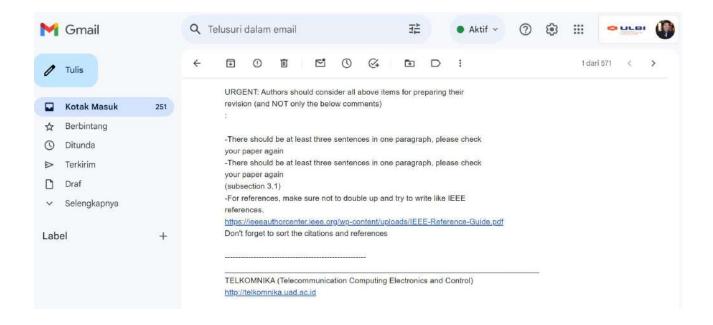




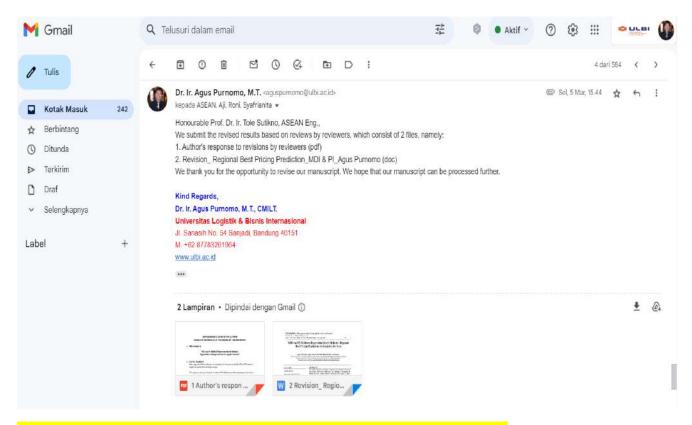








5. Author mengemail Editor untuk menyampaikan hasil revisi pada tgl 5 Maret 2024



Author respone terhadap revisi dari para reviewer adalah sebagai berikut:

# IMPROVEMENTS MADE BY THE AUTHOR BASED ON THE RESULTS OF THE REVIEW BY THE REVIEWERS

1. Title changed to:

"MDI and PI XGBOOST Regression-Based Methods:

# **Regional Best Pricing Prediction for Logistics Services**"

# 2. Abstract changed to:

Referring to the Telkomnika journal template, the Abstract should be 100 to 200 words in length; we made the following changes:

"The logistics industry in Indonesia, where PT Pos Indonesia is the main player, faces fierce price competition. The challenge is determining the optimal price for regional logistics services in each region to gain a competitive advantage and increase revenue. This complex task involves local market conditions, competition, customer preferences, operational costs, and economic factors. This research proposes using machine learning to overcome the complexity of price prediction. The price prediction model developed uses the Extreme Gradient Boosting Regression (XGBR), Support Vector Machine, Random Forest, and Logistics Regression algorithms. This research contributes to using MDI (Mean Decrease in Impurity) and Permutation Importance (PI) to explain how machine learning models make the best price predictions. The results can help company management better understand how to make optimal pricing decisions. The test results show 0001, 0.005, 0.458, 0.009, and 0.9998. Using machine learning techniques and explanatory models, PT Pos Indonesia can more effectively determine optimal prices in each region, increase profits, and compete in the growing regional market."

# 3. Introduction:

# a. Summary of review results:

"The first part of the main text should explain the problem, any solutions you know of, and the main constraints. Also, explain what you hope to achieve through your research."

# b. Author response:

The improvements made are as follows:

"The best pricing of PT Pos Indonesia's logistics services in different regions could be more optimal, so it loses sales competition with other logistics providers. The non-optimal price is because PT Pos Indonesia has not used the best price prediction method in determining the regional best prices, which includes relevant factors such as the variable total price of competitors, the number of customers, the freight price, the number of competitors and the product score given by the customer. As a result, the price of logistics services set from period to period becomes uncompetitive and loses competition with the prices of other logistics service providers. Therefore, the problem in this study is how to create the best logistics service price prediction model by including relevant factors that can be used for each region of PT Pos Indonesia to compete with other logistics service providers. Inspired by the gap that has been explained, our research aims to create the best regional best price prediction for logistics services by including relevant factors with a high level of explanation so that PT Pos Indonesia can be competitive with other logistics providers.

We suggest using MDI and PI to provide the expected level of explanation based on relevant factors. We used XGBoost Regression (XGBR) to predict the price of logistics services and compared it with eight other regression models. The model performance is then evaluated using R-squared, MSE, RMSE, or MAPE. Furthermore, the best machine learning model is used to make optimal price predictions for each region based on relevant factors. This research uses local consumer behavior data analysis and machine learning approaches to help companies such as PT Pos Indonesia understand consumer preferences and behavior in different regions. Finally, we use MDI and PI methods to improve the interpretation of PT Pos Indonesia's local consumer

behavior analysis. To the best of our knowledge, no research uses machine learning to analyze local consumer behavior for optimal regional pricing, especially for logistics service providers. The contributions of this research are, therefore, as follows:

- 1. To create an optimal price prediction model for logistics services using XGBR that can be applied in different regions so that the company can be competitive with its competitors in terms of price.
- 2. Produce a logistics services price forecasting model that can be explained by MDI and PI, illustrating the sensitivity of the model to various relevant factors.

The remainder of this paper is presented in several parts. Section 2 contains a review of previous research that supports our findings and the research design and methodology. The results of our research are presented in Section 3. Finally, section 4 summarises and highlights the main points of our contribution."

# 4. Method:

# a. Summary of review results:

"Every article should include detailed Method section provides a that the reader with enough information to determine whether the study is valid and reproducible. Include enough information for knowledgeable reader to additional reproduce the experiment. However, use references and materials to emphasize previously published procedures."

# b. Author response:

"We have presented the detailed methodology in full in section 3. Research design and we even explained the proposed methodology in section 2. The other four reviewers did not object to the method presented by the author, so we concluded that the method was clear enough to solve the problem and achieve the research objectives. However, we combine the Research design and the proposed methodology into "Method" referring to the Telkomnika journal template."

# 5. Results and Discussion:

# a. Summary of review results:

Summarize the main findings in tables, be concise, and use online supplement tables or figures to keep the paper a manageable size. You should also explain to your readers what the results mean. Explain how the findings relate to the research objectives and hypotheses, as well as how they compare with other research.

Suggestions for organizing a discussion:

- 1. First paragraph:
  - Provide essential interpretation based on key findings.
  - Include a key piece of supporting evidence.
- 2. Second paragraph:
  - Compare and contrast with previous studies.
  - Highlight the study's strengths and limitations.
  - Discuss any unexpected results.
- 3. Last paragraph
  - Summarize the study's purposes.
  - Emphasize the importance of the study.
  - Discuss the unanswered questions and potential future research.

# b. Author response:

The improvements made are as follows:

# 3.1. Result

In the first test, we analyzed the retail price periodically. Here, the total price is aggregated monthly by summing every value. We plot a regression line between total price and aggregate customers (also summed up monthly) from the dataset. This can model the relationship between these two variables. Figure 2 shows the linear relationship between total price and customers. We can interpret the linear relationship objectively with the r-squared value, which is 0.98. That number is considered very high because it approximates the best value of r-squared, 1.0. The p-value of the regression line is 0.01, meaning the null hypothesis is rejected. In regression analysis, rejecting the null hypothesis means there is a significance in the slope of the regression line and that the two variables are strongly related. In normative terms, the increase in total price is related to the increase of customers that visit the store.



Figure 2. Linear regression and customer per-month analysis; Total price vs customers

In the second test, we plot a regression line between the total price and the number of weekends per month from the dataset to observe the relationship between these two variables. Figure 3 shows the linear relationship between total price and customers. The r-squared value of the regression line is 0.86. That number is considered high, however, it is not as high as the previous result. On the other hand, the p-value of the regression line is 0.01, meaning the null hypothesis is also rejected, which leads to the conclusion that the null hypothesis is rejected. There is still a significance in the slope of the regression line, while the two variables are strongly related. In normative terms, the increase in total price is correlated to the increase in the number of weekends per month of retail.



Figure 3. Linear regression and customer per-month analysis; Weakly analysis of total

Analysis of the customer per-month bar chart can be useful for several things, including trend analysis, churn, market effectiveness, customer planning, and prediction. Figure 4 shows that the most customers were in November 2017, and the fewest were in January 2017. After conducting data exploration, we carry out the data pre-processing stage. At this stage, we group data between average and total. The data grouped to calculate the average is 'product id,' 'month year,' 'comp1 diff,' 'comp2 diff,' 'comp3 diff,' 'fp1 diff,' 'fp2 diff,' 'fp3 diff,' 'product score,' and 'unit price.' Meanwhile, the data that is grouped to calculate the total amount is 'product id,' 'month year,' 'total price,' 'freight price,' and 'customers.'

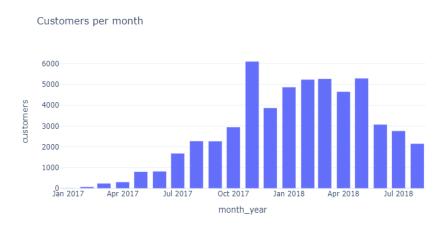


Figure 4. Linear regression and customer per-month analysis; Customer per-month

After the data has been grouped into average and total, the next step is to calculate the average and total of the two data groups. The results of these calculations are stored in two variables, namely, product mean and product sum. Next, after getting the average and total results, the two are combined into one data frame, which contains information about the average and total based on 'product id.' The final stage in the feature engineering process is calculating the logarithm of the variable to be predicted, namely 'unit price,' and the results will be stored in the variable y log, which contains the logarithm values from 'unit price.' The next step is Modelling. At this stage, eight regression models are compared to get the best prediction value. The table 1 displays the evaluation value of each regression model. The XGBR model has the best results compared to other regression models, with an MSE value of 0.0001, MAE of 0.005, MAPE of 0.458, RMSE of 0.009, and R-Square of 0.9998.

Table 1. I	Regression	model	perfo	rmance	compariso	n

Model	Evaluation Metrics						
	MSE	MAE	MAPE	RMSE	R2		
LR	0.1121	0.258	28.193	0.335	0.7243		
RR	0.1127	0.263	28.082	0.336	0.7229		
Lasso	0.1525	0.333	35.971	0.390	0.6250		
RFR	0.0149	0.101	10.212	0.122	0.9633		
GBR	0.0016	0.031	3.117	0.040	0.9961		
ABR	0.0148	0.097	9.728	0.122	0.9645		
XGBR	0.0001	0.005	0.458	0.009	0.9998		
KNR	0.1008	0.253	25.286	0.318	0.7520		
SVR	0.1503	0.294	34.944	0.388	0.6302		

The explainability model stage is to explain and describe why the XGBR model produces certain decisions and results. The MDI graph in Figure 5 shows the relationship between feature values and their impact on predicted values. MDI values on the y-axis (vertical) play a crucial role in understanding the significance of features in the prediction-making process. The elevation of MDI values indicates the level of importance each feature holds in influencing predictions. Visualized as bars on the graph, each feature's bar height signifies the magnitude of its impact. The emphasis should be placed on bars with the highest

MDI values, as these features are deemed the most pivotal in shaping the model's predictions. Features characterized by elevated MDI values play a substantial role in minimizing impurity during the construction of decision trees. Furthermore, a positive MDI value signifies a positive correlation between the feature and the prediction outcome. In simpler terms, higher values of the feature generally support higher predictions. This analysis aids in comprehending the pivotal features that contribute significantly to the accuracy of the model's predictions.

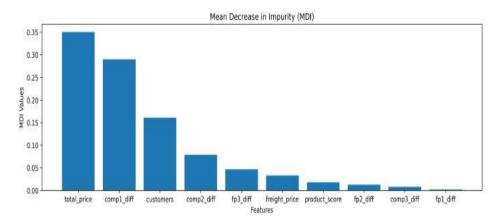


Figure 5. MDI Summary Plot

The Permutation Importance (PI) value serves as a valuable metric for understanding the impact of features on a model's performance when their values are randomly permuted. In Figure 6 and Table 2, we present the PI results, wherein feature names are arranged based on their weight magnitudes. Subsequently, we compute the maximum and minimum error values derived from the PI analysis. It is noteworthy that the features of the highest PI weight correspond with the MDI, specifically the 'total price.' This consistent alignment indicates that 'total price' significantly influences both MDI and PI, underscoring its importance in predicting outcomes.

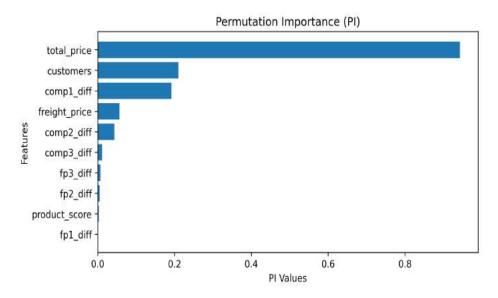


Figure 6. PI Summary Plot

However, disparities arise when comparing the lowest weight in PI with the MDI score. While PI identifies 'fp1 diff' as the feature with the lowest weight, MDI designates 'freight price' as having the lowest MDI value. This discrepancy highlights the nuanced nature of PI, which is inherently model-specific. In In this particular instance, the model under examination is XGBoost (XGBR), revealing that PI results can be influenced by the intricacies of the underlying model. A comprehensive understanding of these

differences enhances our insight into how features contribute to model performance, taking into account both MDI and PI perspectives.

Table 2. The PI Result

Weight	Feature	
$0.9896 \pm 0.2898$	'total'	
$0.1006 \pm 0.0450$	'comp2'	
$0.0966 \pm 0.0172$	'comp1'	
$0.0845 \pm 0.0341$	'customers'	
$0.0302 \pm 0.0231$	'comp3'	
$0.0123 \pm 0.0025$	'product'	
$0.0119 \pm 0.0057$	'fp2'	
$0.0084 \pm 0.0013$	'fp3'	
$0.0048 \pm 0.0022$	Freight'	

### 3.2. Discussion

Several studies have carried out price predictions with various regression models. Shahrel et al. [34] proved that SVR is better than linear regression in price prediction. Durganjali et al. [35] proposed ABR for house price prediction. Finally, Bonamutial et al. [36] demonstrated that RFR is better than KNR in smartphone price prediction. Our research shows that XGBR has better overall performance than LR, RR, Lasso, RFR, ABR, KNR, and SVR in predicting retail prices. Our research contribution is an optimum retail price prediction using XGBR.

In our research, we highlight the different interpretations offered by Mean Decrease in Impurity (MDI) and Permutation Importance (PI) techniques in analyzing features. These interpretations, when combined, contribute to a more comprehensive Explainable Artificial Intelligence (XAI). While both methods score features based on their contribution to predictive power, MDI goes a step further by providing insight into the positive/ negative impact of each feature. In contrast, PI offers error values that indicate the sensitivity of features and their influence on overfitting. Our dual contribution is to improve the explanation of retail price prediction models through PI techniques and to leverage MDI and PI to analyze influential features, especially 'total price'. Our analysis underscores the important role of 'total price,' 'comp1 diff,' and 'customer' in model development, with 'total price' being the most influential. The choice between MDI and PI depends on the research needs. Overall, the findings emphasize the importance of 'total price' in forming an optimized pricing model. Our research contributions are summarised in Table 3 by comparing them with state-of-the-art research in retail price prediction.

Table 3. A comparison of state-of-the-art research on the retail price prediction

Reference	Prediction Model	R-Squared	Explainability Method	
			MDI	Permutation Importance
Shahrel et al. [34]	SVR	0.6302	×	*
Durganjali et al. [35]	ABR	0.9645	×	*
Bonamutial et al. [36]	RFR	0.9633	×	×
Proposed Method	XGBR	0.9998	✓	✓

# 6. Conclusion:

# a. Summary of review results:

"Your conclusion is more than just a recap of what you've already written. It should improve your paper by answering any outstanding questions. Summarize your research findings and propose potential applications and extensions. The most important question in your conclusion should be, "What do my findings mean for the research field and my community?"

### b. Author response:

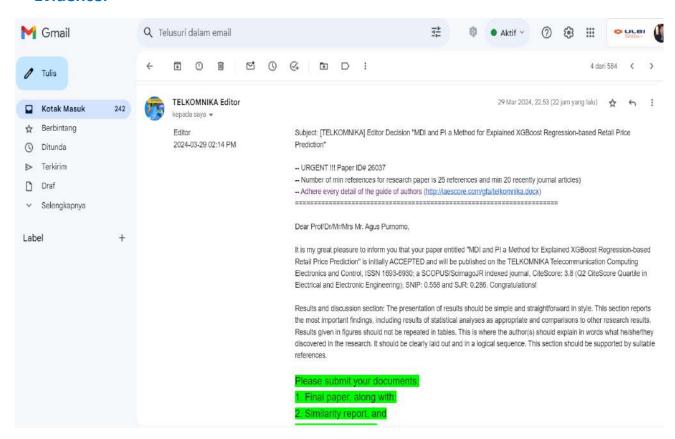
The improvements made are as follows:

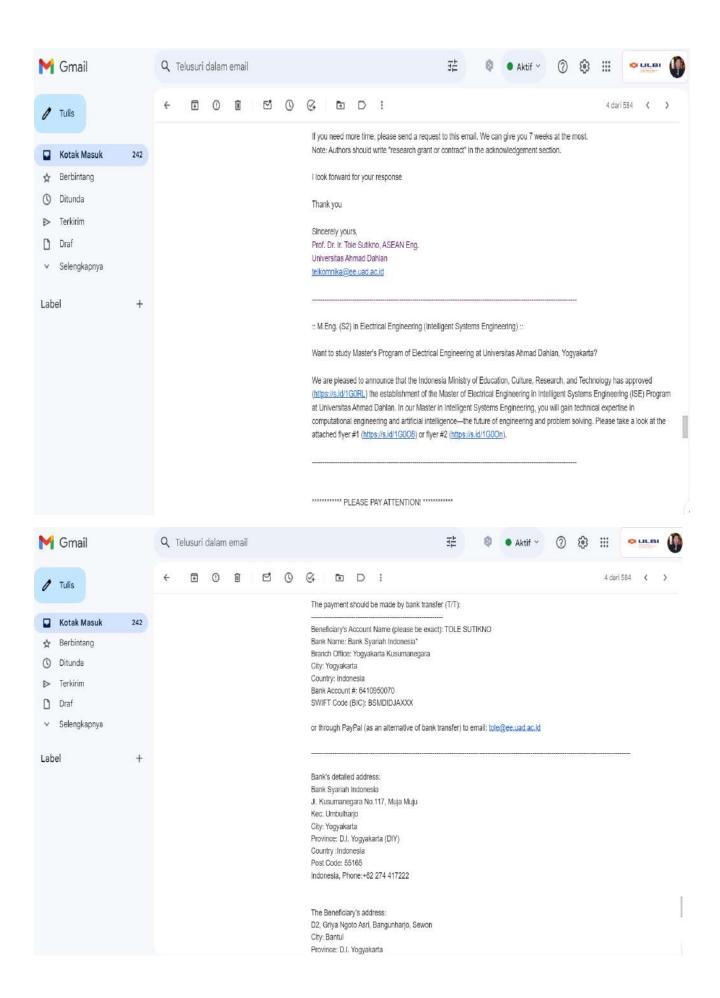
"In this study, we successfully built a regional best price prediction model for logistics services using XGBoost Regressor (XGBR) and its explanation model to improve the interpretation

of price prediction. To compare the analysis of the relevant factors influencing the model, we tested XGBR with LR, RR, Lasso, RFR, GBR, ABR, KNR, and SVR, and also with two eXplainable Artificial Intelligence (XAI) methods, namely Mean Decrease in Impurity (MDI) and Permutation Importance (PI). The test results show that XGBR is superior to the benchmark method, as evidenced by the MSE, MAE, MAPE, RMSE, and r-squared values of 0.0001, 0.005, 0.458, 0.009, and 0.9998, respectively. In addition, based on the MDI and PI explanatory models, total price emerged as the most influential relevant factor in predicting the optimal regional best price prediction for logistics services. This study demonstrates that XGBR outperforms eight other regression models, making it the most effective method. Furthermore, it has practical implications for the logistics industry, as companies can use it to determine the optimal Regional Best Pricing Prediction for Logistics Services, giving them a competitive edge over their competitors and boosting sales and profits.

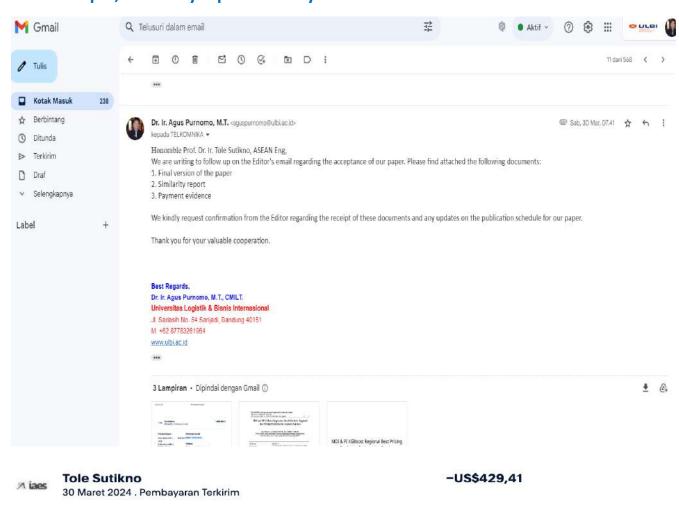
This study suggests two avenues for future research to improve the accuracy of Regional Best Pricing Prediction as a price competitive advantage strategy in the logistics industry. Firstly, we recommend using datasets directly from the company to improve the accuracy and relevance of the model to the actual conditions of the company. For model development, we propose incorporating constraint functions to support complex decision-making strategies in determining the optimal Regional Best Pricing Prediction for Logistics Services. This can be achieved through the application of evolutionary algorithms, allowing the model to adapt to changing market dynamics."

# 6. Editor mengemail Author pada tanggal 29 Maret 2024, tentang Accepted paper dan meminta Author untuk mengirim Final Paper, Similarity report dan Payment Evidence.





# 7. Author mengemail Editor pada tanggal 30 Maret 2024, dengan melampirkan Final Paper, Similarity report dan Payment Evidence.



# Dibayar dengan

### Informasi kontak

Kartu Debit VISA x-4448

US\$429,41Pesan Tole Sutikno

Anda akan melihat "PAYPAL \*tole" pada laporan kartu Anda.

**ID Transaksi** 

6V7075943B396163V

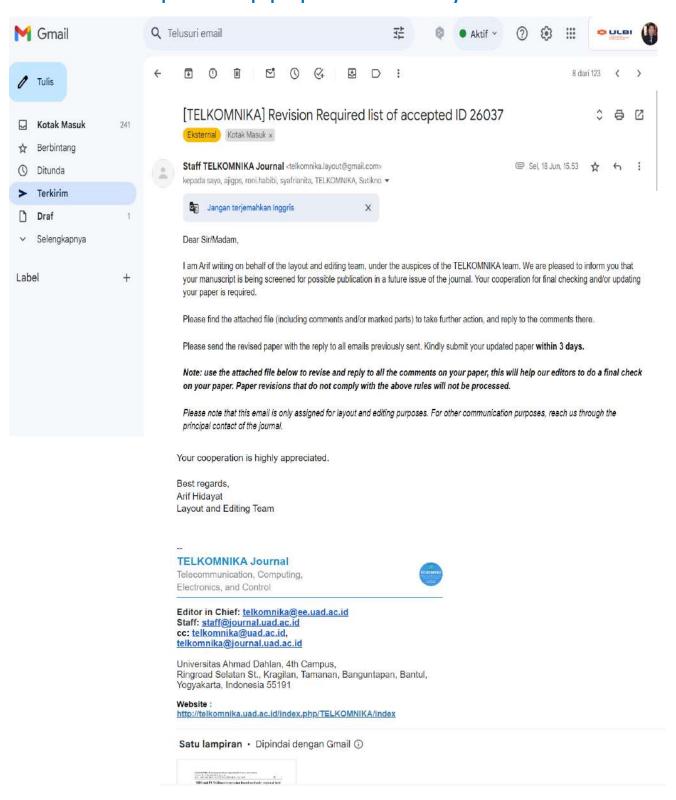
### Catatan

Payment of publication fee for paper (ID# 26037, title: MDI and PI XGBoost Regression-Based Methods: Regional Best Pricing Prediction for Logistics Services) to the TELKOMNIKA Journal. Authors: Agus Purnomo, Aji Gautama Putrada, Roni Habibi, Syafrianita

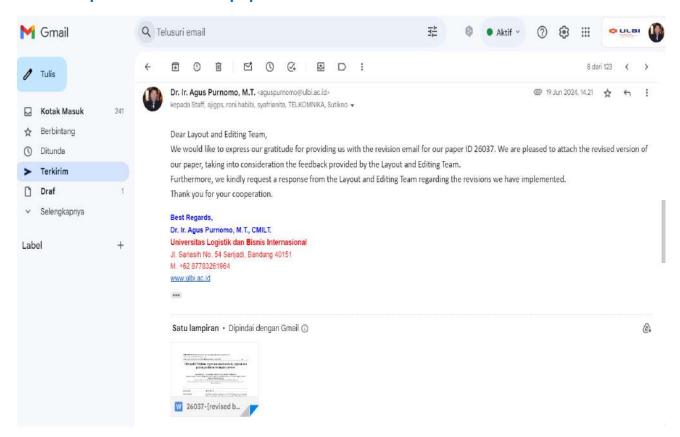
# Perincian

Dikirim ke Tole US\$415,00 Sutikno Biaya US\$14,41 Total US\$429,41

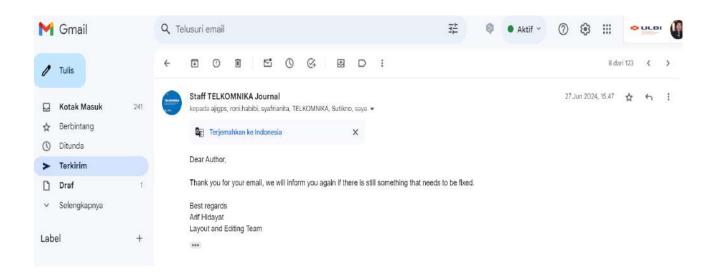
# 8. Staf Telkomnika Journal mengemail Author pada tanggal 18 Juni 2024, tentang final revisi untuk penerbitan paper pada issue berikutnya.



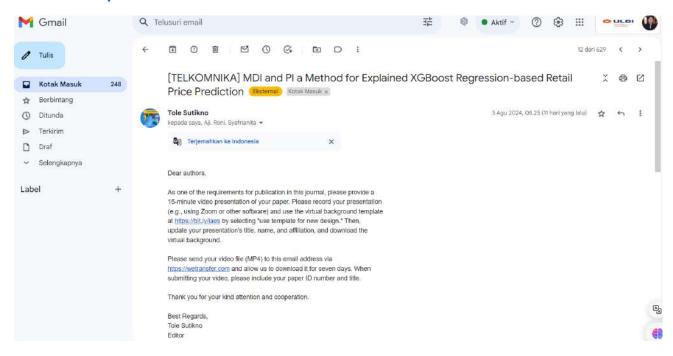
9. Author mengemail Staf Telkomnika Journal pada tanggal 19 Juni 2024, dengan melampirkan Final revisi paper.



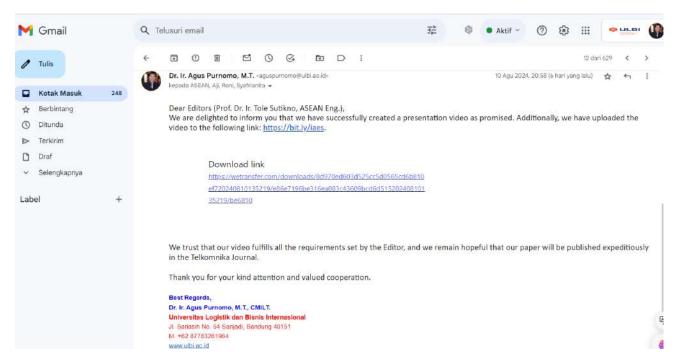
10. Staf Telkomnika Journal mengemail Author pada tanggal 27 Juni 2024, tentang email final revisi telah diterima dan akan diinformaikan jika ada yang perlu diperbaiki lagi.



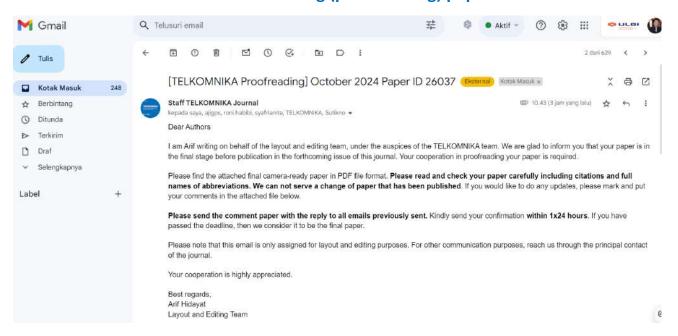
11. Editor mengemail Author pada tanggal 5 Agustus 2024, tentang kewajiban membuat video presentasi dari paper sepanjang 15 menit sebagai syarat akhir untuk dipublikasikan.



12. Author mengemail Editor pada tanggal 10 Agustus 2024, dengan menyampaikan link upload file presentasi.



# 13. Staf Layout Telkomnika mengemail Author pada tanggal 16 Agustus 2024, untuk meminta Author meneliti ulang (proofreading) paper sebelum diterbitkan .



# 14. Author mengemail Staf Layout Telkomnika pada tanggal 16 Agustus 2024, dengan menyampaikan sedikit perbaikan tentang afiliasi.

