International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

# The Analysis of Price Optimization Implementation of Cement Product Through Machine Learning

Ade Hendra Wiguna[1], Ravi A. Salim[2]
[1, 2]Business Information System, Gunadarma University, Jakarta, Indonesia, 10430
Email Address: [1] adehendra0372(at)gmail.com, [2] ravisalim(at)gmail.com

***Abstract***— *Price is such a crucial aspect as it is a form of value agreement that is required for exchange in a transaction. Price can also be interpreted as something that must be spent by the buyer to receive the product (Harjanto, 2009). Besides that, price is also part of a marketing strategy which greatly influences the company's income (revenue). As time goes by, technological advancements are so rapid that all business fields highly demand technology to get an accurate and precise information to facilitate decision making. The technology used recently is machine learning which is a sub-field of artificial intelligence that is widely researched and used to solve various problems. One of the largest manufacturing companies in Indonesia engaged in cement production has a vision to be the largest provider of building material solutions in the region with excellent performance and continues to aggressively develop its business. With increasingly fierce competition, companies are trying to maintain market share, oversupply conditions are putting pressure on prices in the market so they need a solution that has reliable capabilities in terms of data analytics that can provide optimal price recommendations using Machine Learning in an effort to increase company profitability. The purpose of this research is to analyze the implementation of pricing for cement products using various machine learning algorithm models such as Random Forest, Decision tree to provide optimal price recommendations.*

***Keywords***— *Price recommendation, Random Forest, Decision Tree.*

## I. INTRODUCTION

One of the largest manufacturing companies in Indonesia engaged in cement production has a vision to be the largest provider of building material solutions in the region with excellent performance and continues to develop its business aggressively.

With increasingly fierce competition, companies are trying to maintain market share, oversupply conditions put pressure on prices in the market so they need a solution that has reliable capabilities in terms of data analytics that can provide optimal price recommendations using Machine Learning in an effort to increase company profitability

Price has a positive significant influence on buying interest, where high buying interest can be formed by the prices offered by service or product providers (Sutrisno & Haryani, 2017). According to Kotler and Armstrong (2001), price is the amount of money that is charged for a product (goods or services), or the value that must be paid to get the benefits of the product. Machine learning is a sub-field of artificial intelligence that is widely researched and used to solve various problems. Reviews from various fields are presented in the form of problem solving and algorithms and are divided into three categories in machine learning, including supervised learning, unsupervised learning, and reinforcement learning (Roihan et al., 2020)

## II. LITERATURE REVIEW

There have been several previous studies conducted by several researchers regarding price determination or prediction using several machine learning algorithms with a high level of accuracy. (Lubis et al., 2005)

Another research as a reference is regarding "Stock closing price prediction using Machine Learning technique" which explains that the stock market has very dynamic and unpredictable characteristics. Predicting stock prices is a very challenging task because it depends on factors including but not limited to political conditions, the global economy, financial reports and company performance. (Vijh et al., 2020)

Another study is regarding "Land Price Research using the Linear Regression Algorithm" which concludes that the Linear Regression Algorithm can provide accurate results or measure the accuracy of predicting land prices in the city of Banjarmasin properly and of course it must go through the correct stages. (Pratama, 2016)

Random Forest is a method introduced by Breiman. As the development and combination of many Decision Trees. If the Decision Tree is a single classification tree, then in the Random Forest many trees are made to determine the prediction results. The combination of bootstrap aggregating and random feature selection in random forests can be used to reduce overfitting problems in small data trains (Breiman, 2001). Because the random forest is an ensemble method from CART, the random forest also has no assumptions or is good for use in nonparametric cases. development, software that is released gradually, reduces process overhead, and produces high-quality code and in the process of development directly engages customers

CART (Classification and Regression Tree) is an exploratory method of classification and regression data based on decision tree techniques. Classification trees are generated when the response variables are in the form of categorical data, while regression trees are generated when the response variables are in the form of numerical data (Breiman & Cutler, 2003).

## III. METHODOLOGY

The research was carried out through interviews regarding the current process of pricing and observing the data, which then continued by studying the patterns and characteristics of the data.



Ade Hendra Wiguna and Ravi A. Salim, "The Analysis of Price Optimization Implementation of Product Cement Through Machine Learning" *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 2, pp. 12-16, 2023.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

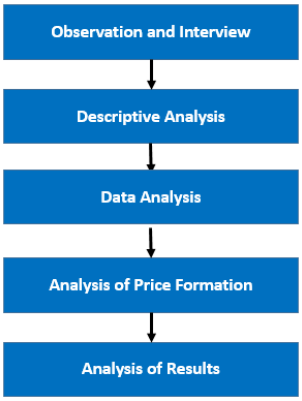
### A. *Observation*

In simple terms, it can be interpreted as observing an object or subject matter, so the observation method can be interpreted as a learning method for scholars by observing an object or subject matter raised or conveyed by the lecturer through the media or the surrounding environment (Pujiyanto, 2021)

### B. *Interview*

According to Anwar Sanusi (2014: 105), "the interview method is a data collection technique that uses questions orally to the research subject". Direct interviews are being conducted with employees. This method is used to find out about the marketing mix carried out at the New Ferry Workshop.

### C. *Research Stage*




## IV. RESULT

### A. *Observation and Interview*

The results of observations regarding business processes carried out by researchers show that this price optimization application is the process of determining the ideal price position for a certain product, a certain area, and a certain brand. This price determination process uses historical data as a reference for determining market position and the amount of quantitative data. The system created using the machine learning method will check the price movements based on the period (monthly), brand, district and price disparity of each competitor on weekly data or what is known as WPM (Weekly Price Monitoring).

The movement of downward prices and also high price disparities will be responded automatically by the system by taking into account the profitability and potential market share that will be obtained so that machine learning will provide optimum price recommendations.

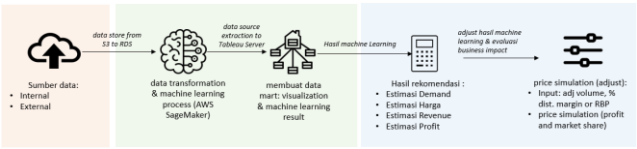



Figure 1. Price Decision-Making Process

### B. *Descriptive Analysis*

The descriptive analysis of this study is based on the data studied that there are 12 companies engaged in cement production that have brands spreaded throughout Indonesia. Competition for each brand occurs in certain areas or districts. Measurement of competition is based on the total volume sold in the market with a particular brand or also known as Market Share. Below is the total national cement volume by brand. Nationally, the Semen Indonesia Group is still in the highest ranking based on sources from the Indonesian Cement Association (ASI).

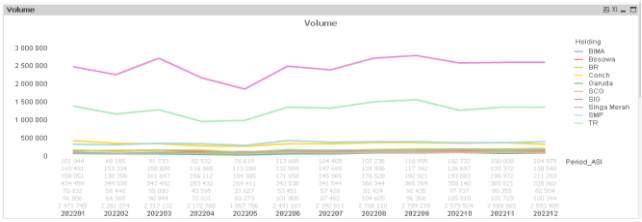



Figure 2. Price Decision-Making Process

### C. *Data Analysis*

Based on the author's analysis, the data sources (RET, WPM, ASI, GPM, Profitability) were merged and converted into fact tables so that they became a data mart that would be used in the machine learning modeling process. There are 2 parts to the data mart, namely the data mart used for BI Tools (Tableau) and the data mart used for machine learning modeling. Fact tables containing retail data can be seen directly from BI tools after the data source upload process, even though the modeling process has not been carried out. For the modeling process, demand and price predictions use the retail demand fact table and price elasticity, which then results of the modeling process can be seen from BI Tools.

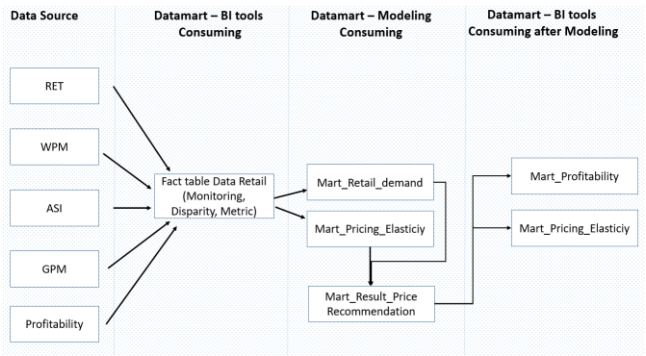



Figure 3. Data Mapping

### D. *Analysis of Price Formation*

Based on the researcher's analysis, the process of establishing the price of cement products is carried out in several stages, including

*a)* Demand Prediction based on Price

This process is to predict the demand that will be obtained when there is a price change. In this process, each model predicts demand by dividing the dataset between training data and test data, 75% of training data is determined and 25% of testing data. There are 4 models used in this testing process, namely Linear Regression, Gradient Boosting, Random Forest



Ade Hendra Wiguna and Ravi A. Salim, "The Analysis of Price Optimization Implementation of Product Cement Through Machine Learning" *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 2, pp. 12-16, 2023.

and Decision Tree, of these four models the results that show the level of accuracy of the results of the validation are the Random Forest because the MAE Test has the smallest value

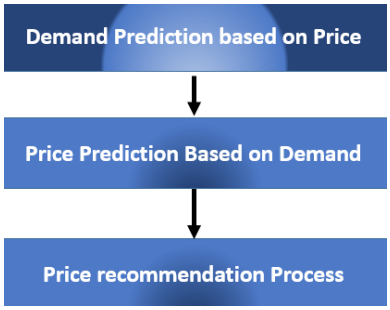



Figure 4. Stages of Price Formation

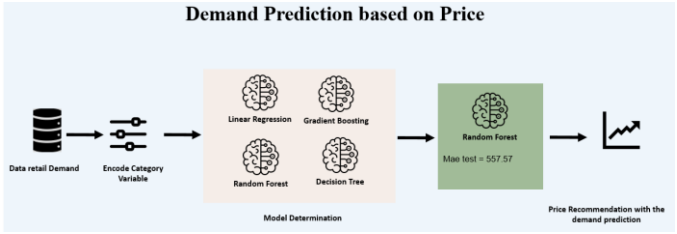



Figure 5. Process Flow of Demand Prediction Based on Price

Error testing is being done to evaluate the results of demand predictions from each model using the Mean Absolute Error (MAE). Figure 6 is the result of validating demand predictions from each machine learning model

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|A_i - F_i|$$

| Gradiant Bossting | Random Forest |
|---|---|
| MAE_TEST = 612.59 | MAE_TEST = 557.57 |
| **Linear Regression** | **Decision Tree** |
| MAE_TEST = 840.989275 | MAE_TEST = 794.74 |

Figure 6. Result of Machine Learning Validation Process

*b)* Price Prediction based on Demand

This process is to predict the price that will be obtained when there is a change in demand or volume. In this process, each model predicts prices by dividing the dataset between data training and data test, comprises of 80% data training and 20% data test. There are 4 models used in this testing process, namely Linear Regression, Gradient Boosting, Random Forest and Decision Tree, from these four models the results that show the level of accuracy are Random Forest because the MAE Test has the smallest value

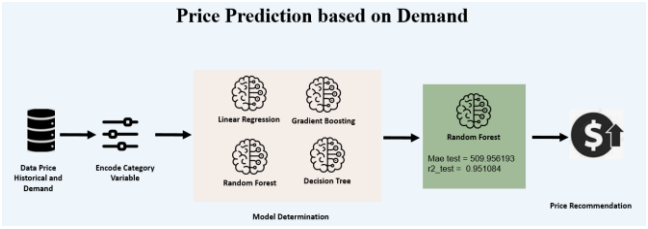



Figure 7. Process Flow of Price Prediction Based on Demand

| Gradiant Bossting | Random Forest |
|---|---|
| MAE_TEST = 623.69 | MAE_TEST = 509.95 |
| **Linear Regression** | **Decision Tree** |
| MAE_TEST = 1018.94 | MAE_TEST = 702.83 |

Figure 8. Result of Machine Learning Validation Process

*c)* Analysis of how Price Prediction works with Random Forest

From the two predictions above, the researcher will explain the analysis of how price prediction works based on demand using the random forest machine learning model which is the result of the process above. The following is a data simulation using a random forest consisting of District variables, brand name (brand), packaging and RBP. The information that will be labeled is what will be predicted, namely the RBP (Retail Buying Price).

| DISTRICT | BRAND | PACKAGING | RBP |
|---|---|---|---|
| NABIRE | Brand 1 | 50 KG | 34,000 |
| BANYUWANGI | Brand 1 | 40 KG | 25,500 |
| PEKALONGAN | Brand 4 | 50 KG | 27,000 |
| SINTANG | Brand 1 | 40 KG | 25,500 |
| KAB. INDRAGIRI HILIR | Brand 2 | 50 KG | 31,500 |
| PANDEGLANG | Brand 2 | 50 KG | 14,750 |
| KAB. BENER MERIAH | Brand 2 | 40 KG | 30,000 |
| KAB. LAMPUNG SELATAN | Brand 4 | 50 KG | 26,075 |
| SALATIGA | Brand 1 | 40 KG | 23,750 |

Figure 9. Example of data

As explained above, a random forest is a collection of several decision trees (decision trees) that randomly select samples from the available dataset, then the results of each tree are voted or selected for the results that appear frequently for each final prediction result. In this case the author uses rapidminer tools to describe up to the 3rd decision tree. The following are some randomly selected decision trees.

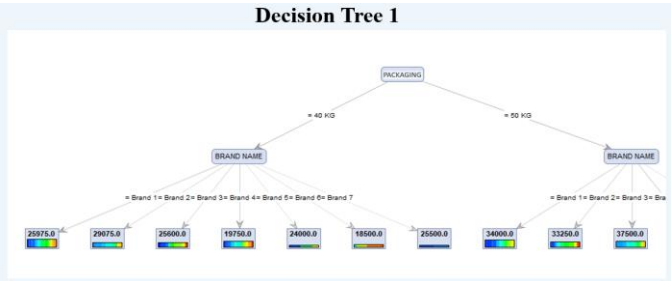



Figure 10. Decision Tree 1



Ade Hendra Wiguna and Ravi A. Salim, "The Analysis of Price Optimization Implementation of Product Cement Through Machine Learning" *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 2, pp. 12-16, 2023.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Figure 10 above shows the first decision tree which consists of brand name and packaging nodes where each node has a price prediction value (RBP) that comes out of the decision tree.

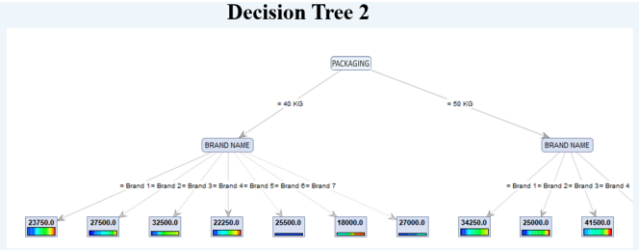



Figure 11. Decision Tree 2

Figure 11 above shows the 2nd decision tree which only consists of one brand name node where each brand has a price prediction value

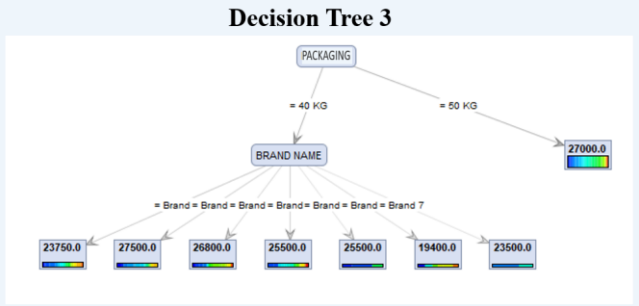



Figure 12. Decision Tree 3

Figure 12 above shows the 3rd decision tree which only consists of 2 nodes, namely packaging and brand name where each brand has a price prediction value

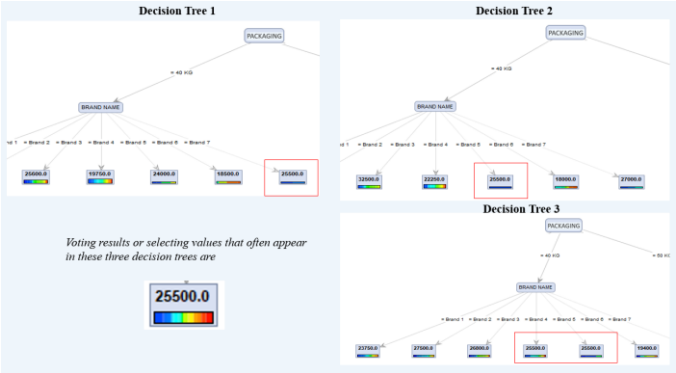



Figure 13. Random Forest

Figure 13 above shows that the three decisions are put together and then voting is carried out or the predicted value that often appears in each tree, and that is the final predicted value of a random forest workings. This data shows the last value of the random forest based on the district and brand name.

| DISTRICT | BRAND | PACKAGING | RBP | RBP PREDICTION |
|---|---|---|---|---|
| BANYUWANGI | Brand 1 | 40 KG | 25500 | 25500 |
| SINTANG | Brand 1 | 40 KG | 25500 | 25500 |
| PEKALONGAN | Brand 4 | 50 KG | 27000 | 27000 |
| KAB. INDRAGIRI HILIR | Brand 2 | 50 KG | 31500 | 31500 |
| NABIRE | Brand 1 | 50 KG | 34000 | 34000 |

Figure 14. Random Forest Prediction Result based on example of data

*d) Price Recommendation Process*

The result of the researcher's analysis is that the final result of the process of forming price predictions is not the final decision to get the optimum value in this price recommendation process, because the optimum results show that when a price change occurs it will have an impact on sales volume, Profit Margin. In terms of business (business sense) that if the price is lowered then the demand or demand will increase so that the market share increases and will provide an optimum profit margin. The results of the researcher's analysis show that the process flow for this price recommendation.

The result of the researcher's analysis is that the final result of the process of forming price predictions is not the final decision to get the optimum value in this price recommendation process, because the optimum results show that when a price change occurs it will have an impact on sales volume, Profit Margin. In terms of business (business sense) that if the price is lowered then the demand will increase so that the market share increases and will provide an optimum profit margin. The results of the researcher's analysis show that the process flow for this price recommendation:

a. Price predictions issued from the results above using a machine learning model are then carried out with an adjustment process with a range from – 5000 to 5000
b. The algorithm will simulate price changes with a range from -5000 to 5000 with multiples of 500.
c. The volume prediction obtained using the demand model is based on price.
d. The algorithm will calculate price and volume profitability using cost tables to get maximum profit
e. The price that gives the maximum profit will be used as the final price recommendation.
f. Price recommendations will come out based on the district, packaging and brand name

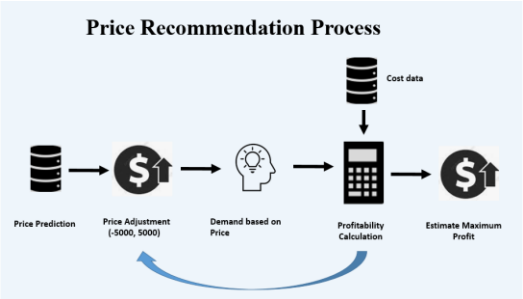



Figure 15. Price Recommendation Process

*E. Analysis of Results*

The results of the price recommendation process using machine learning algorithms have very reliable computing, provided that the data source used has valid data quality so that the data read by the machine learning algorithm provides



Ade Hendra Wiguna and Ravi A. Salim, "The Analysis of Price Optimization Implementation of Product Cement Through Machine Learning" *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 2, pp. 12-16, 2023.

a very high level of data accuracy. Figure 16 is the result of a price recommendation that is formed as a result of computing the algorithm and also provides information on estimated volume and profit margin to be obtained. In Figure 17 is a monitoring dashboard that can see actual market price data compared to the predicted price recommended by the system

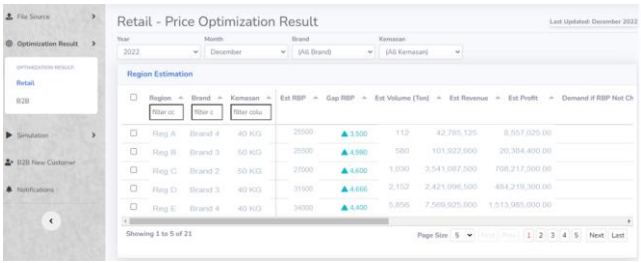



Figure 16. the application of Price Optimization in the regional level


Figure 17. the application of Price Optimization in the province level

## V. CONCLUTION & RECOMMENDATION

Machine learning is a computational method that is very popular nowadays and is very reliable in determining predictive values.

The results of the researcher's analysis that this application involves several models in making predictions so that the machine will be flexible in choosing a suitable model and have better accuracy

In analyzing and testing implementation results, you need some statistical tools such as rapidminer and also have a little knowledge about Jupiter notebooks in order to simplify the process of analyzing how the application works

## REFERENCES

[1] Pratama, S. (2016). Land Price Prediction Using Linear Regression Algorithm. Technology, 7(2), 122–130.Azdy, RA., Azhar SN. (2012). Scrum Implementation in Distributed Software Development. National Seminar on Informatics 2012, ISSN: 1979-2328.

[2] Lubis, C., Sutedjo, E., & Setiadi, B. (2005). Stock Price Prediction Using Hybrid Neural Network Algorithm (Vol. 2005, Issue Snati, pp. 17–19).

[3] Ramadhan, A., Susetyo, B., & Indahwati, I. (2019). Application of the Random Forest Classification Method in Identifying Important Factors in Assessing Education Quality. Journal of Education and Culture, 4(2),

[4] Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Utilization of Machine Learning in Various Fields: Review paper. IJCIT (Indonesian Journal on Computer and Information Technology), 5(1), 75–82.

[5] Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. *Procedia Computer Science*, *167*(2019), 599–606. https://doi.org/10.1016/j.procs.2020.03.326

[6] Supriyadi, R., Gata, W., Maulidah, N., & Fauzi, A. (2020). Application of the Random Forest Algorithm to Determine the Quality of Red Wine. E-Business: Scientific Journal of Economics and Business, 13(2), 67–75.

[7] Baihaqi, W. M., Dianingrum, M., & Ramadhan, K. A. N. (2019). Simple Linear Regression to Predict Patient Visits In Hospitals Based On Type Of Service and Patient Age. Symmetrical Journal, 10(2), 671–680.

[8] Saragih, H. (2015). Analysis Of Pricing Strategy And Its Effect On Sales Volume At Pt. Nutricia Indonesia Prosperous Medan. In Scientific Journal of Methonomi (Vol. 1, Issue 2).

[9] Saputra, I. N. Y., Saadah, S., & Yunanto, P. E. (2021). Analysis of Random Forest, Multiple Regression, and Backpropagation Methods in Predicting Apartment Price Index in Indonesia. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, *7*(2), 238. https://doi.org/10.26555/jiteki.v7i2.20997

[10] Lewis, R. J., Ph, D., & Street, W. C. (2000). An Introduction to Classification and Regression Tree (CART) Analysis. *2000 Annual Meeting of the Society for Academic Emergency Medicine*, *310*, 14p. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1 95.4103&rep=rep1&type=pdf