

FORECASTING RICE INVENTORY IN INDONESIA USING THE ARIMA ALGORITHM METHOD

^aFaisal Rizki Kurniawan, ^bRudi Sutomo

^{ab}Information Systems Study Program, Faculty of Engineering and Informatics, Multimedia Nusantara University, Jl. Scientia Boulevard, Gading Serpong, Tangerang, Banten – 15811

ABSTRACT

Objective – The world's global development is increasingly developing and complex, accompanied by globalization, making various sectors need to follow these developments. The agricultural sector, the primary sector, especially in need of food for every society, especially in Indonesia, is also touched by technological developments. Planning the supply of rice inventory is needed monthly so that there is no excess or shortage of the required rice inventory stock.

Methodology – Made predictions from rice inventory stock data using the CRISP-DM method to analyze the data and use the ARIMA Algorithm.

Findings – This research predicts the amount of rice inventory production that will be carried out in the next few months by applying the forecasting or prediction method using the CRISP-DM method and using the ARIMA algorithm.

Novelty – This study predicts the amount of stock of an item using RapidMiner tools.

Limitations - This study uses rice inventory data using the total amount for a month in a year and tools such as RapidMiner and tableau.

Keywords — ARIMA, CRISP-DM Method, Forecasting, Research, Rice Inventory.

JEL Classification: D24, E17, Q18

Article info: Received July 2021; Revised August 2021; Accepted August 2021

Article Correspondence:

Recommended Citation: Kurniawan and Sutomo (2021). Forecasting Rice Inventory in Indonesia Using the ARIMA Algorithm Method. *Journal of Multidisciplinary Issues*, 1 (2), 1-12.

I. INTRODUCTION

Indonesia is an agricultural country with an area of 1.9 million km². Agriculture is prevalent, and most of the Indonesian population, especially the agricultural sector. These two things are crucial for the people of Indonesia. Rice is a basic need needed by around 78% of the Indonesian population to meet daily energy intake, especially carbohydrate intake (Java, Winter, & Moving, 2018).

Rice is a basic need in Indonesia. The Central Bureau of Statistics is always on duty and records data on rice inventory production in every region in Indonesia. A decline can indicate annual rice inventory production as in 2018 and 2019, which fell by 7.75%, while in 2019 and 2020, there was an increase in production by 1.02%. The drastic decline in 2019 occurred because the expected amount of production did not reach the desired target, so that there was a buildup of stock. To reduce these errors in 2020, the agricultural center produced rice inventory to a minimum to reduce the problem so that this result was sufficient to reach the production target determined (Badan Pusat Statistik, 2020).

From this problem, every production agency, such as the agricultural sector in Indonesia, needs predictions or forecasts for food stock produced, such as rice inventory. This is necessary so that there are no errors in the production of the stock so that there is no excess or deficiency in the production of rice inventory that is carried out.

Then to be able to determine the availability of stock from rice inventory production, it is necessary to forecast production. This forecasting can be done as with the ARIMA model. The ARIMA model was chosen because it has the advantage of analyzing random situations, trends, seasons, and even cyclical characteristics in the analyzed sales time series (Rianto & Yunis, 2021).

This study aims to determine the production forecasting of rice inventory production from the agricultural sector in Indonesia in predicting the amount of stock from rice inventory supplies so that rice inventory production can be on target. This forecasting will use the ARIMA model with total production data from the relevant agencies.

II. LITERATURE REVIEW

A. Rice Inventory

Rice is the part of the grain that has been separated from the husk. Husk (Javanese marang) is anatomically called 'pala' (covered part) and 'lemma' (covered part). At one stage of processing the rice inventory harvest, the grain is ground with a mortar or milled to separate the outer part (husk of the grain) from its contents. This part of the filling, white, reddish, purple, or even black, is called rice inventory.

Rice is generally grown as an annual crop. Rice plants can grow to a height of 1 - 1.8 m. The leaves are long and slender, with a length of 50-100 cm and a width of 2-2.5 cm. Edible rice inventory is 5 – 12 mm long and 2-3 mm thick.

Rice is mainly used for processing the most important staple food for the world's citizens. Rice is also used to make various kinds of snacks and cakes, mainly from sticky rice, including for tapas. In addition, rice is an essential component of herbal medicine for kencur and param. Popular drinks made from processed rice are arak and water tajin. In the food industry, rice is processed into rice flour. The embryo part is also processed into a food supplement called rice eye flour.

B. Data Mining dan Deep Learning

Data mining is also called the process of looking for interesting patterns or information in selected data using specific techniques or methods. Techniques, methods, or algorithms in data mining vary widely. In simple terms, data mining is discovering information by looking for patterns or with certain rules from a very large amount of data (Kusumasari, Rafizan, & Rafizan, 2017).

Deep learning is a learning method that is part of artificial intelligence. The way deep learning works is by using machine learning that uses artificial neural networks (ANN). Deep learning is also called being able to learn its computational methods. This deep learning technology is one of the most popular technologies for recognizing an activity or object with a higher accuracy level than the previous machine method (Kusumasari et al., 2017).

C. Cross-Industry Standard Process (CRISP-DM)

This study uses data mining techniques using the CRISP-DM framework to forecast the production of goods or stocks from rice inventory stocks. The Cross-Industry Standard Process (CRISP-DM) framework is the stage of data mining work, namely the data analysis process to find new insights/knowledge that experts widely use to solve problems from research conducted (Fadillah, 2015). The stages of work carried out using CRISP-DM include Business Understanding, Data Understanding, Data Preparation, Modeling Phase, Evaluation Phase, and Deployment Phase (Ncr et al., 2000).

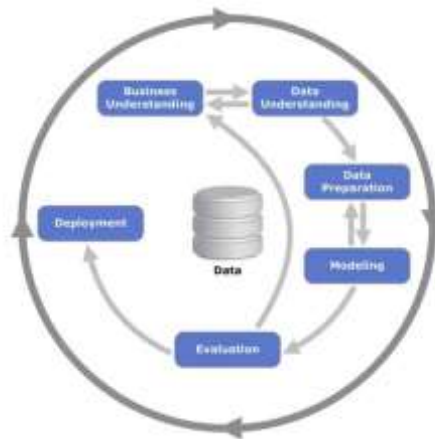


Figure 1. CRISP-DM Lifecycle

CRISP-DM consists of six phases, as follows:

1. *The Business Understanding Phase* is to understand the goals and needs from a business point of view and then translate this knowledge into problem definition in data mining. Furthermore, plans and strategies will be determined to achieve these goals (Ncr et al., 2000).
2. *Data Understanding Phase* is a data collection that a process will follow to gain an in-depth understanding of the data, identify data quality problems, or detect an interesting part of the data that can be used to hypothesize to produce information (Ncr et al., 2000).
3. *Data Preparation Phase* builds the final dataset (data to be processed in the modeling phase) from raw data. This stage also includes selecting tables, records, and data attributes, including the process of cleaning and transforming data to be used as input in the modeling stage (Ncr et al., 2000).
4. *The modeling phase* selects and applies various techniques, and some of the parameters will be adjusted to get the optimal value (Ncr et al., 2000).
5. *The evaluation phase* evaluates the effectiveness and quality of the model before it is used and determines whether the model can achieve the goals set in the initial Phase (Business Understanding) (Ncr et al., 2000).
6. *Deployment Phase*, where the knowledge or information that has been obtained will be organized and presented in a particular form so that users can use it. The deployment stage can be making simple reports or implementing iterative data mining processes within the company (Ncr et al., 2000).

D. Forecasting

In analyzing the company's business activities, it is necessary to predict the future business; as stated by Martiningtyas (2004) in his book, statistics that estimate and indicate what will happen in the future are called forecasting (Nofiyanto, Nugroho, & Kartini, 2015).

E. ARIMA

ARIMA (Auto-Regressive Integrated Moving Average) is also known as the Box Jenkins time series method. ARIMA is well known for time series forecasting. ARIMA has perfect accuracy for short-term forecasting and for non-stationary time series data at linear times. Meanwhile, the accuracy is not suitable for forecasting data that is quite long because it usually tends to be flat. The shorter the

forecasting data, the better the accuracy will be. In addition, ARIMA will experience a decrease in accuracy if there is a non-linear time-series component in the observation data (Nofiyanto et al., 2015).

ARIMA has attributes (p, d, q) where the meaning of these attributes is as follows:

- P is the order of the autoregressive model (AR value).
- D is the level of data difference.
- Q is order data from moving – average value model (MA).

F. ARIMA Model Selection

AR (Autoregressive model), MA (Moving-Average model), and ARMA (Autoregressive Moving Average) models use the assumption that the resulting time series data is stationary. Time series data is mostly non-stationary. If the data is not stationary, then the method used to make the data stationary is different. Then for the non-stationary data in the average with the transformation process for the part of the data that is not stationary in the variance [7].

The best model selection is made using the ARIMA model with the smallest MSE (Mean Squared Error) or RMSE (Root Mean Squared Error) value resulting from the following equation (Nofiyanto et al., 2015):

$$MSE = \frac{\sum_{i=1}^j (x_t - x'_t)^2}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^j (x_t - x'_t)}{n}}$$

Where n is the amount of data, and the variable x is:

$$X_t = \text{Data in period } t$$
$$x'_t = \text{Forecasting data in period } t$$

G. Tableau Software

Tableau Software is an interactive data visualization software. Tableau can also use and create relational databases, online analytical processing cubes, cloud databases, spreadsheets to produce graphical data visualizations. The Tableau software can also extract, store, and retrieve data from an in-memory data engine.

H. RapidMiner Studio

RapidMiner is a data mining software platform that provides an integrated data mining process environment for data preparation, data mining, deep learning, text mining, and predictive analytics. It is used for business and commercial use, as well as for research, education, training, rapid prototyping, and application development, and supports all steps in the data mining process, including data preparation, visualization results, model validation, and optimization.

III. RESEARCH METHODOLOGY

The methodology used in this study uses data mining techniques with the CRISP-DM framework in this study into 6 phases as follows (Fadillah, 2015).

A. Business Understanding

The purpose of this study is to forecast the production of rice inventory supplies using the ARIMA model. Then the flow of this research is as follows.

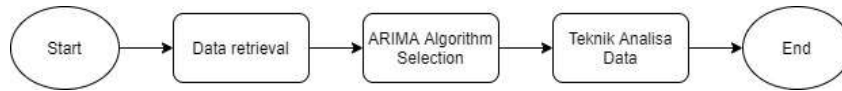


Figure 2. Research Flow

B. Data Understanding

Data Understanding this study uses data from the total stock of rice inventory produced from the agricultural sector with the title "DataBeras," which was downloaded from the Central Bureau of Statistics (Badan Pusat Statistik, 2020). In this data, there is some information related to production data from rice inventory as follows:

- DATE, in this data, describes the date of each month per year.
- Total Ton, in the data, describes the total production of rice inventory by the number of tons in months per year.

From an understanding of the data used in this study, implementing the forecasting of rice inventory data is divided into two processes in this study. The use of Rapidminer tools in forecasting or predicting using the ARIMA algorithm can carry out two processes, namely carrying out the validation process of forecasting and making predictions from data for the future.

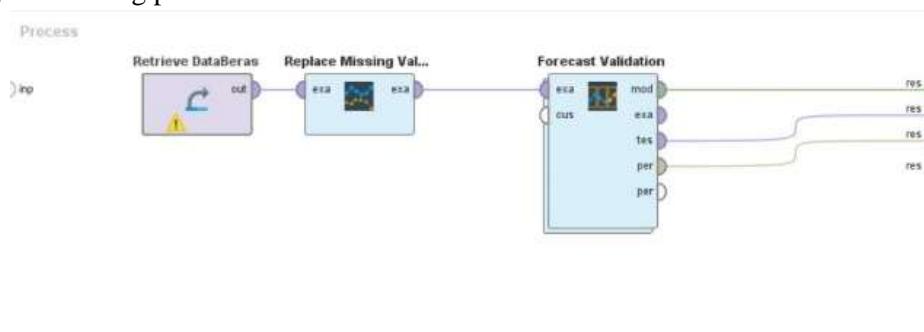


Figure 3. Forecast Validation Process

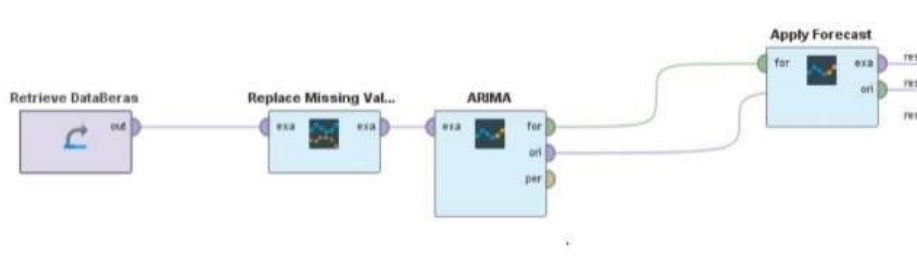


Figure 4. Rice Inventory Data Prediction Process

C. Data Preparation

Data preparation in this study, several stages were used to produce modeling with the ARIMA model, namely as follows:

1. ARIMA modeling requires indexes and attributes. The index required is in the form of a serialized time or date series, while for the attribute rules, it is in the form of an integer or real.
2. Retrieve data or describe data sets by inputting "DataBeras" data into rapidminer tools to process data and visualizations using tableau tools (Sari, Wanto, & Windarto, 2018).

3. Replace Missing Value, which is used to overcome missing or missing data in retrieving research data so as not to affect the next data process.
4. Forecast Validation wherein it performs ARIMA modeling by dividing the training data in each data split (window) with a size smaller or equal to the size of the data, broken down again in the subprocess with several steps to get the value of the RMSE.
5. In this Phase, ARIMA Process modeling from ARIMA determines each value of AR, MA, or ARMA with its attributes, namely (p, d, q), to get the results of each matter. Then determine the attributes of the time series by using the table for the Total Tons section in numeric or real data, which is compared (Indices) with the time index using the DATE table.

D. Modeling Phase

This research uses RapidMiner software to generate a value using the ARIMA model in the modeling phase. The ARIMA model is defined by its three order parameters, p, d, q. P determines the number of Autoregressive terms in the model. d determines the amount of difference that is applied to the time series values. q determines the number of Moving Average terms in the model that serves to get the value result from RMSE.

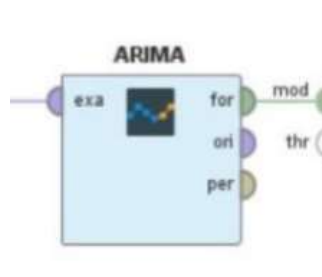


Figure 5. ARIMA Model

```
<operator activated="true" class="time_series:arima_trainer" compatibility="9.9.000"
  <parameter key="time_series attribute" value="Total (Ton)"/>
  <parameter key="has indices" value="false"/>
  <parameter key="indices attribute" value="" />
  <parameter key="p: order of the autoregressive model" value="2"/>
  <parameter key="d: degree of differencing" value="0"/>
  <parameter key="q: order of the moving-average model" value="1"/>
  <parameter key="estimate constant" value="true"/>
  <parameter key="main criterion" value="aic"/>
</operator>
<connect from_port="training set" to_op="ARIMA" to_port="example set"/>
<connect from_op="ARIMA" from_port="forecast model" to_port="model"/>
<portSpacing port="source training set" spacing="0"/>
<portSpacing port="sink model" spacing="0"/>
<portSpacing port="sink through 1" spacing="0"/>
</process>
```

Figure 6. Syntax from p, d, q

E. Evaluation Phase

In the evaluation phase of this study, two types of evaluation were used as follows:

1. Forecast Validation, at this stage, is the process of forecast validation that validates the model carried out in the modeling phase, namely the model from ARIMA namely, to obtain results from forecasting forecast results with actual data.

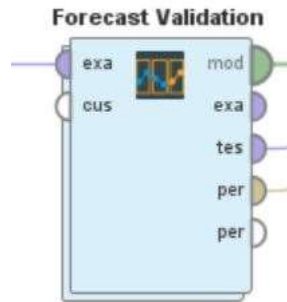


Figure 7. Forecast Validation Stages

```
<operator activated="true" class="time_series:forecast_validation" compatibility="9.7.002"
<parameter key="time_series_attribute" value="Total (Ton)"/>
<parameter key="has_indices" value="true"/>
<parameter key="indices_attribute" value="Date"/>
<parameter key="expert_settings" value="false"/>
<parameter key="unit" value="example based"/>
<parameter key="windows_defined" value="from start"/>
<parameter key="custom_start_point" value="5"/>
<parameter key="custom_end_point" value="100"/>
<parameter key="window_size" value="20"/>
<parameter key="custom_start_time" value="2000-01-01 00:00:00"/>
<parameter key="custom_end_time" value="2030-01-01 00:00:00"/>
```

Figure 8. Syntax from forecast validation in Rapidminer

2. Apply Forecast, in this process using a model from ARIMA to predict the value of inventory data in the next predetermined time.



Figure 9. Stages of Applying Forecast

```
<operator activated="true" class="time_series:apply_forecast"
<parameter key="forecast_horizon" value="10"/>
<parameter key="add_original_time_series" value="true"/>
<parameter key="add_combined_time_series" value="true"/>
```

Figure 10. Syntax from Apply Forecast stage in Rapidminer

F. Deployment Phase

In this study, the deployment phase is to get the results of the first process, namely the validation of the forecasting results to see the results of the production of rice inventory supplies that have been carried out, namely by looking at the comparison of actual data and forecast data and getting the best RMSE value. Then the second process is to forecast the rice inventory to get the production value for the future. This process functions to predict the production of rice inventory supplies that will be produced so that it can be following the target to be achieved.

IV. RESULTS AND DISCUSSION

In this study, two processes were carried out to obtain results and discussions on forecasting rice inventory supplies. These two processes consist of a validation process of forecasting rice inventory supplies and forecasting rice inventory supplies that will be carried out in the future. The following are the results of two processes from the results and discussion in this study.

A. Forecast Validation

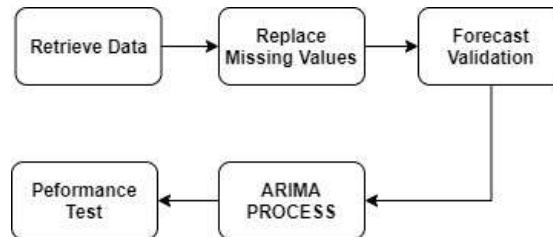


Figure 11. Forecast validation process flow

In the first process flow, validate the sales of existing rice inventory stocks and then get the best RMSE results. Here, the ARIMA process model is carried out with attributes (p, d, q) which use the ARIMA(1,0,1), ARIMA(1,0,2), and ARIMA(2,0,1) models whose results are as follows.

Table 1. Performance Test results for RMSE values

ARIMA Model	RMSE Result
ARIMA(1,0,1)	1157798,42
ARIMA(1,0,2)	1229369,41
ARIMA(2,0,1)	1136971,08

The search for the value of the RMSE is that the smaller the value, the better it will be, which is useful for the forecasting process at the next stage. Then from these results, it can be concluded that ARIMA(2,0,1) has a smaller RMSE value than other ARIMA models, so that this model will be used until the forecasting process stage.

In the forecasting validation carried out with the previous rice inventory stock production from 2018 to early 2020, it turned out that a significant decline was found at the end of 2019, so that it did not reach the production target.



Figure 12. Comparison of Forecasting with Actual Data

From these results, it can be found that the decline in 2019 then becomes a step to determine the production of rice inventory supplies in 2020; this can be seen from the production in 2020, which balances the production that previously experienced a decline so that production for rice inventory supplies can be produced stably and for achieving the target. After the production has stabilized, the rice inventory production will carry out more production than the previous plan to avoid excessive demand for rice inventory ingredients.

The production of an item such as rice inventory or other goods requires a forecasting process in the production to achieve the right target and as a benchmark for the development trend of the production of the item itself. In addition to forecasting, it is also seen from the trend of demand for these goods to be used as a benchmark within a certain time such as the amount of production per year in total production can experience a decrease or increase in demand from the production of these goods.

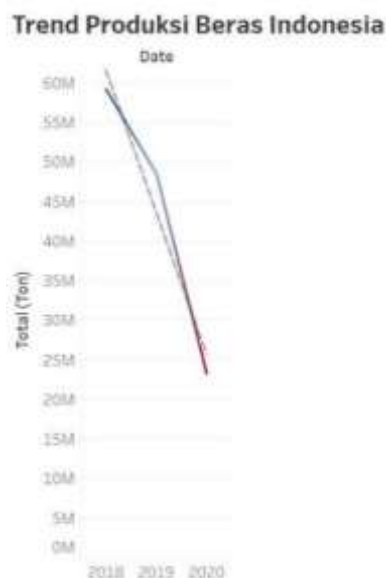


Figure 13. Indonesian Rice inventory Production Trend (Tableau)

When viewed from the trend of rice inventory production carried out, the trend has decreased every year. This decline in production is since in 2019 rice inventory production decreased due to reduced

production and price inventory that became more expensive so that the demand for rice inventory decreased.

Because of this problem, the agricultural sector needs to forecast from their rice inventory stock inventory to get the right production results. According to the target, there are no excess and shortages in the rice inventory stock they have. After validating the forecasting, they entered the next stage: making predictions or forecasting rice inventory stock inventories.

B. Forecasting Process

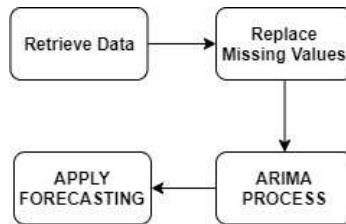


Figure 14. Forecasting Process Flow

In this stage, the forecasting or prediction process can be seen from the results of the best RMSE value, namely the ARIMA model (2,0,1), to get the results from the correct forecast to be used for rice inventory production in agriculture in the future so that the production target can always be correct.



Figure 15. Forecasting Indonesian Rice Inventory Production (RapidMiner)

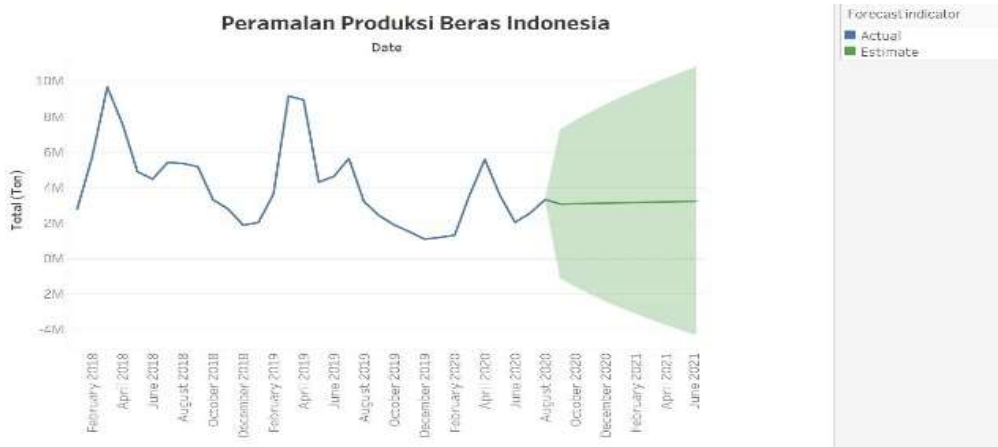


Figure 16. Forecasting Indonesian Rice Inventory Production (Tableau)

If you look at the two forecasts above, both using the Rapidminer tools and the Tablue tools, ARIMA modeling forecasts the production of rice inventory supplies for the future. This study using forecasting for the next ten months. The ARIMA model can predict the amount of rice inventory supply for the next ten months, which can be seen in the table below.

Table 2. Forecasting Results in the Next 10 Months

DATE (Date / Year)	Forecasting Results Stock Production (TON)
September 2020	3813254.69
Oktober 2020	4015670.44
November 2020	4072663.35
December 2020	4087467.03
Januari2021	4091229.46
DATE (Date / Year)	Forecasting Results Stock Production (TON)
February 2021	4092179.74
Maret 2021	4092419.30
April 2021	4092479.67
Mei 2021	4092494.87
Juni 2021	4092498.71

The results from the table above can be found in the rice inventory supply increases in production starting from mid-2020 when compared to the previous early 2020. Still, it is not far off the target as in 2019, which is the production target of rice inventory supplies. From the table above, it can also be found that if rice inventory supplies enter 2021, there is an increase in production, which tends to be relatively stable in production from the rice inventory stock produced. Then it can be seen if there are results that tend to be relatively the same from production forecasting in March to June 2021.

From these results, the ARIMA model can also forecast rice inventory inventories. The forecasting results are the targets to produce rice inventory supplies so that rice inventory production can always reach the expected target and there are no shortages and excesses of rice inventory production stocks.

V. CONCLUSION

This study applies an application of forecasting methods to rice inventory supplies in Indonesia which aims to predict the amount of production value that will be carried out in the future and re-validation of production forecasts that have been carried out previously. The use of RapidMiner Tools using the CRISP-DM method can forecast by applying the model from the ARIMA algorithm, namely by carrying out two processes with each function. The process of Forecast Validation serves to validate the forecasting of data that has already occurred and other processes to forecast from the data to get a value with a predetermined future time.

ARIMA Time Series modeling can predict the forecast of the amount of rice inventory to measure the amount of rice inventory needed for the future. This model can also be used for prediction purposes from other types of goods according to the needs of the production of goods carried out. Based on the results, it was concluded that forecasting the amount of inventory or stock of goods is very necessary, especially in the large-scale and long-term sales business, to reduce errors and losses from the inaccurate number of stock items needed.

REFERENCES

- Badan Pusat Statistik. (2020). *Latar Belakang*. Bps.Go.Id.
- Fadillah, A. P. (2015). Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). *Jurnal Teknik Informatika Dan Sistem Informasi*, 1(3), 260–270. <https://doi.org/10.28932/jutisi.v1i3.406>
- Java, E., Winter, U., & Moving, A. I. (2018). *Peramalan Persediaan Beras Provinsi Jawa Timur Menggunakan Winter Dan Autoregressive Integrated Moving Average Indonesia . Bertambahnya jumlah penduduk Jawa Timur mengakibatkan kebutuhan beras Jawa Timur , pemerintah membentuk Badan Urusan Logistik (BULOG. 1, 323–329.*
- Kusumasari, D., Rafizan, O., & Rafizan, D. O. (2017). Studi Implementasi Sistem Big Data Untuk Mendukung Kebijakan Komunikasi dan Informatika Studi Implementasi Sistem Big Data Untuk Mendukung Kebijakan Komunikasi dan Informatika. *Jurnal Masyarakat Telematika Dan Informasi*, 8 No. 2, 81–96.
- Ncr, P. C., Spss, J. C., Ncr, R. K., Spss, T. K., Daimlerchrysler, T. R., Spss, C. S., & Daimlerchrysler, R. W. (2000). Step-by-step data mining guide. *SPSS Inc*, 78, 1–78.
- Nofiyanto, A., Nugroho, R. A., & Kartini, D. (2015). Peramalan Permintaan Paving Blok Dengan Metode ARIMA. *Proceedings Konferensi Nasional Sistem Dan Informatika (KNS\&I)*, 9, 54–59.
- Rianto, M., & Yunis, R. (2021). Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest. *Paradigma - Jurnal Komputer Dan Informatika*, 23(1). <https://doi.org/10.31294/p.v23i1.9781>
- Sari, R. W., Wanto, A., & Windarto, A. P. (2018). Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi). *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 2(1), 224–230. <https://doi.org/10.30865/komik.v2i1.930>