The Factors Influencing Market Price Of Sago Starch In Meranti Islands By Linear Regression

Diki Arisandi, Seri Hartati, Bahjatul Murtasidin, Amir Syamsuadi, Liza Trisnawati, Luluk Elvitar, Muhammad Azhari Herli

1 3 4 Departement of Informatics Engineering, Faculty of Engineering, Universitas Abdurrab, Pekanbaru.
2 3 4 Departement of Governmental Science, Faculty of Social and Political Sciences, Universitas Abdurrab, Pekanbaru.
1 Departement of Pharmaceutical and Food Analysts, Faculty of Medical and Health Sciences, Universitas Abdurrab, Pekanbaru

E-mail: diki@univrab.ac.id

Abstract. One of the primary food from Indonesia is sago, and one of the sago producers is the Meranti islands. In addition to processed sago products into foodstuffs, sago products are also sold as sago starch. This study discusses the three factors that influence the market price of sago starch are the number of sago farmers, the area of sago land and annual production. Machine learning approach used is the linear regression. After training data, it can be observed that there are two factors that have a significant influence on the market price of sago starch. The results can be reference material for the industrial and trading offices of Meranti islands in observing the annual sago price so that the farmers and sago entrepreneurs can get the appropriate price of sago starch.

1. Introduction
Indonesia is a country rich in natural resources, especially for primary food, one of the primaries is sago. Sago is a plant producing carbohydrate derived from its stem. Sago also has been consumed as a primary food in Eastern Indonesia. Until now, sago is the highest source of carbohydrates in the world because it can produce 200-400 kg of dried starch per stem [1]. One of the largest sago producers in Indonesia comes from districts of meranti islands. Meranti Islands is an expansion of Bengkalis territory established on December 19, 2008. This districts also included in the triangle of economic growth Indonesia - Malaysia - Singapore (IMS-GT) [2].

The sago harvested every period is used for residents in Meranti Islands is only 10% of the total production, but 90% of the sago produced is sold elsewhere such as Java Island in the form of sago starch. The fixed market price of sago starch in kilogram from the data of last seven years did not change much, but the number of farmers, the amount of sago farm and the amount of production per year are significantly increase [3]. This research will observe and analyze whether the number of sago farmers, the area of the farm and the amount of production remains to influence the price of sago. Researchers used a linear regression approach to analyze the impact of factors affecting the market price determined.

Content from this work may be used under the terms of the CreativeCommons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd
by the local government. Linear regression is an approach used to assess the effect of one or more exogenous to one endogenous variable [4]. The aim of this research can be a consideration of the local government in determining the market price so that sago farmers get a right price, sustainable sago plantation sector, increasing the regional economy and the welfare of sago farmers can be assured.

2. State of The Art

Districts of Meranti islands has nine subdistricts; the total area is 3,707.84 km² with 216,329 population. This area consists of four large island clusters and one of the islands districts in Riau Province. Geographically, this districts is in the international shipping and trade line of Malacca Strait and two neighboring countries, Singapore and Malaysia [5].

Sago plants grown in Meranti islands has become the primary commodity in driving the regional economy. Sago (Metroxylon sago, Rottb.) is a vital source of carbohydrate as a foodstuff besides rice, corn, and tubers. In Indonesia, sago has been processed into both food and non-food products [6]. The type of sago growing in the area of Meranti is sago buni / thorn, sago sangka and sago bemban [7] as shown in Figure 2:

The data generated from the production of agricultural products such as sago in meranti islands and its supporting factors can be processed to deliver modeling, visualization, prediction, and classification.
This process is known as machine learning. In data analysis, machine learning is a method used to design complex models and algorithms that fit the predictions. This analytical model allows researchers, data scientists, engineers, and analysts to "produce reliable and recurring decisions and results" and find "hidden insights" through learning from historical relationships and trends in data [8],[9],[10]. One of the methods is the linear regression. It is a linear approach to modeling the relationship between an endogenous and one or more exogenous variables. Mathematically, the basic formula of linear regression can be written as follows [11]:

\[ y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \]

Linear regression can be described as regression analysis where there is a linear relationship between the exogenous (x) and endogenous (y) variable. Visually, the relation between the x and y variables can be described as follows:

![Figure 3. The Relation Between Variables](image)

One of previous work related to this research is Deforestation driven by urban population growth and agricultural trade in the twenty-first century. This work discusses the population growth and trade in farm products that affect deforestation. They said, reducing atmospheric carbon emissions from tropical deforestation is considered a cost-effective choice for mitigating climate change. They were using satellite-based estimates of forest loss for 2000 to 2005 to assess economic, agricultural and demographic correlates across 41 countries in the humid tropics. The methods used is linear regression and regression tree, show that forest loss is positively correlated with urban population growth and exports of agricultural products for this period. Rural population growth is not associated with forest loss, indicating the importance of urban-based and international demands for farm products as drivers of deforestation [12].

3. The Proposed Method
The required data comes from the industrial and trading offices of Meranti. The data collected are sago farmer's data, sago land area, annual production and market price per year. The data obtained from 2011 to 2018 is because the relevant office established independently since separated from Bengkalis district and become the new district in 2011. The number of sago farmers, the area of sago land and the annual production was the exogenous (x), while the market price per year was the endogenous (y). The relationship between variables seen in Figure 4 as follows:
As a visualization, the relationship between variables will generate results in the form of linear regression graphs to visualize the distribution of the data, also regression tree as a rule generator. Later, it can be used as consideration for the decision maker. The red line in the linear regression graph is referred to as the best fit straight line as seen in figure 3 below [13]:

![Figure 3. Linear Regression Graph](image)

**Figure 3. Linear Regression Graph**

**Figure 4. The Variables Influencing The Market Price**

Based on obtained data, as the first is to analyze the relationship between variables. The aim was to determine the influence of exogenous to endogenous. The results can be seen in the following table:

| Model | Unstandardized Coefficients | Standardized Coefficients | t     | Sig. |
|-------|-----------------------------|---------------------------|-------|------|
|       |                             |                            |       |      |
| 1     | (Constant)                  | 5193,570                  | 111,976 | 46,381 | .000 |
| num_of_farmers | .047 | .100 | .062 | .473 | .638 |
| area_sago_land | .007 | .039 | .068 | .173 | .864 |
| ann_production | -.001 | .008 | -.069 | -.175 | .862 |

a. Dependent Variable: marketprice

From table 1 in the sig. column (marked in red) that exogenous does affect endogenous. This can be seen as the value on the red mark beyond 0.05 (α) which is the minimal trust value or maximum error rate of a variable. Also, based on the analysis results in table 1, Annual production and area of sago land
have the more significant influence on market price than the number of farmers. After obtained the amount of influence between variables, we created a visualization of linear regression; it can be seen in the following figure:

Based on Figure 6 it can be seen that the distribution of dots on the graph follows the exogenous variables. However, if we analyzed the market price as a class variable, it can be seen that the distribution of dots at 4800 and 5800 is relatively the same. This indicates that the shift of sago price follows the factors such as the number of sago farmers, annual production and area of sago land. In other words, If there was an escalation in these factors, then the price of sago will be adjusted as well. The results seen from the graph in figure 6 are also supported by the rule tree as follows:
Figure 7. Regression Tree

Figure 7 is a tree generated from linear regression based on input data from 2011. These three factors influence the market price of sago starch in the range of 2011 to 2018. The results indicated that there was a change in market price if the number of farmers, land area and production results has changed. Besides, the tree can deliver the average market price, so it is not fixed on the highest price or the lowest price only.

5. Conclusion
In this study based on field observations and data from the relevant source, the factors that influence the price of sago starch are the number of sago farmers, annual production and area of sago farm. In addition, to analyze the effect of exogenous to endogenous variables, using linear regression can also be delivered graphical visualization to see the distribution of data on which factors have a significant influence on the market price of sago starch. Besides, the resulting rule tree can be a reference price to avoid selling at the highest price or the lowest price only. The results of this study can indeed be a reference to the industry and trade offices meranti district in viewing the development of sago starch prices so that farmers and sago entrepreneurs gain an appropriate rate in selling their sago starch.

After the research was completed by mid-year of 2018, there was no other study to examines the factors that influence the market price of sago starch by machine learning. The method, linear regression, is a simple machine learning method. There should be a review using other ways to get more comparison. Also, several other factors that could be considered material for further research. There are the number of sago refineries and the distance between sago farms to sago refineries and available containers, which may affect the market price of sago starch. By using the additional factors is expected to enrich the research related to sago starch.

References
[1] Dewi RK, Bintoro MH. Karakter Morfologi dan Potensi Produksi Beberapa Aksesi Sagu (Metroxylon spp.) di Kabupaten Sorong Selatan, Papua Barat Morphological Characteristics and Yield Potential of Sago Palm (Metroxylon spp.) Accessions in South Sorong District, West Papua. Agron Indones. 2016;44:91–7.
[2] Pemerintah Kabupaten Kepulauan Meranti. Naskah Akademik Kabupaten Meranti. 2016.
[3] Badan Pusat Statistik Kabupaten Kepulauan Meranti. Kepulauan Meranti Regency in Figures 2016. Selat Panjang; 2016. Report No.: 1410.2016.01.

[4] Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. 5th Ed. Balding DJ, Cressie NA., Fitzmaurice GM, editors. New Jersey: John Wiley & Sons, Inc.; 2012.

[5] Syahza A. Model Pengembangan Daerah Tertinggal Dalam Upaya Percepatan Pembangunan Ekonomi Pedesaan Di Kabupaten Kepulauan Meranti Propinsi Riau [Internet]. Pekanbaru; 2012. Available from: http://repository.unri.ac.id

[6] Singhal RS, Kennedy JF, Gopalakrishnan SM, Kaczmarek A, Knill CJ, Akmar PF. Industrial production, processing, and utilization of sago palm-derived products. Carbohydr Polym. 2008;72:1–20.

[7] Bupati Kepulauan Meranti. Pengembangan sagu di kepulauan meranti. Selat Panjang; 2017.

[8] Müller A, Guido S. Introduction to Machine Learning with Python. Blanchette M, Roumeliotis R, editors. 1005 Gravenstein Highway North, Sebastopol, CA 95472: O’Reilly Media, Inc.; 2016.

[9] Arisandi D, Nofriandi. Analisa Pola Pemilihan Program Studi Bagi Calon Masiswa Di Universitas Abdurrah Menggunakan Association Rule. RABIT (Jurnal Teknol dan Sist Inf Univrab). 2016;1:8–12.

[10] Indra Z, Trisnawati L. Pengembangan intelligent data collector untuk analisis big data artikel berita online. RABIT (Jurnal Teknol dan Sist Inf UNIVRAB). 2018;3:49–59.

[11] Alpaydin E. Introduction to Machine Learning. 3rd Ed. Massachusetts: The MIT Press; 2014.

[12] Defries RS, Rudel T, Uriarte M, Hansen M. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. Nat Geosci [Internet]. Nature Publishing Group; 2010;3:178–81. Available from: http://dx.doi.org/10.1038/ngeo756

[13] Kubat M. An Introduction to Machine Learning [Internet]. Florida: Springer US; 2017. Available from: http://link.springer.com/10.1007/978-3-319-63913-0