

Application of the Naïve Bayes Algorithm and Simple Exponential Smoothing for Food Commodity Prices Forecasting

Muhammad Lutfi¹, Hidayatul Muttaqien², Aulia Apriliani³, Hazriani Zainuddin⁴, Yuyun⁵
^{1,2,3,4,5}Department of Computer System, Faculty of Engineering, Sekolah Tinggi Manajemen
Informatika dan Komputer Handayani, Indonesia, 90321

{sainteclutfi@handayani.ac.id¹, hidayatul@student.handayani.ac.id²,
auliaapriliani@student.handayani.ac.id³, hazriani@handayani.ac.id⁴, yuyunwabula@handayani.ac.id⁵}

Abstract. Inconstancy of the market prices can affect society's purchasing power. One effort to anticipate the price uncertainty is by conducting commodity price forecasting. In the concept of forecasting, the commodity prices can be predicted by studying sales data in the previous period. This study aims to implement a decision support system in predicting food commodity prices trend. In data collection, the authors used list of food commodities provided by Industry and Trade Service of Gowa Regency. For data analysis, we use Naive Bayes algorithm to predict the food commodity prices in the future and Simple Exponential Smoothing to find out the price trend in a certain period. As a result, both methods can predict commodity prices and market tendency in a given time completely.

Keywords: *Forecasting, Commodity, Price.*

1 Introduction

A market is a place where buyers and sellers together to make transaction of good and service. Besides that, the market is an important part of life to guarantee the fulfillment of needs. However, food commodity becomes difficult if the price stability is uncertain. In general, prices are formed due to the interaction between supply and demand. If tender is high and demand is low, then price will go down. Conversely, if supply is low while demand is high then price will rise [1]. This also applies to food commodity prices. Food commodities are one of the government's main priorities to ensure the availability of stock in the future. From the economic aspect, food prices are an important sector to control price stabilization in order to maintain people's purchasing power. Lack of information on commodity prices causes fluctuation which is feared to affect economic stability.

There are many ways to anticipate fluctuations of commodity prices, one of which is prediction method. The purpose of prediction is to estimate the food commodity cost that occurs in the future. In addition, predictions also provide an overview of the cost situation for a certain period [2]. The need for forecasting to facilitate the stakeholders involved in decision maker related to cost stability and availability stock. The main objective of this research is to predict food commodity prices. This study uses the Naïve Bayes algorithm and Simple Exponential Smoothing to forecast food commodity prices and price trends on a particular period.

Various works have been studied to understand the application of Naïve Bayes algorithm. In this section, the authors explain how Naive Bayes algorithm can be used to predict

problems as a reference for readers. Many applications and methods have been proposed such as used this algorithm for stock market prediction [3]. In this paper was to predict future stock price using prediction concept. In that Parse Records then calculate predicted value and send to user. And automatically perform operations like purchase and sale shares using Automation concept. [4] Described the data classification for prediction of new student graduation. Evaluation results show classified data correctly (correct classified instances) in accordance with the grouping of choice graduated. [5] Used data number of order to predict number of production. [6] Presented cooperative data to predict bad credit. They found that there are relation job criteria, living status, and monthly income to predict bad credit. [7][8][9] Proposed Naïve Bayesian classification to predict performance of a student. They collected student's data from 100 private graduate institutions from a private educational. As concluded that their work can help educational institutions to predict student performance. [10] Demonstrated the questionnaire data for students' success prediction that cover the fields affecting students' performance the most. Then [11], [12], [13] proposed same method for breast cancer detection, stomach cancer, and purchase prediction. They conclude that the Naïve Bayes algorithm is selected as an optimal choice to conduct prediction.

There are already a large number of research papers discussing the naive Bayes algorithm. To improve the accurate result, authors add a Exponential Smoothing to analyze the prices trend of food commodity.

2 Material and Methods

2.1 Dataset Structure

The study took place in Gowa Regency, South of Sulawesi, Indonesia. The dataset consists of two mounts days with eight food commodities. For the simulation stage, authors used granulated sugar commodity dataset starting from February 1st to the 9th, 2017.

2.2 Methods

2.2.1 Naive Bayes Algorithm

Naive Bayes Classifier (NBC) is a probabilistic based prediction technique with consideration of independence [14]. All attributes in the dataset are considered has no relationship. This algorithm uses a probability approach by summing the average for each attribute. The formula of Naive Bayes is as follows:

$$P(Y|X) = \frac{P(Y)P(X|Y)}{P(X)} \dots\dots\dots(1)$$

Where:

- P (Y | X) = Probability of data with vector X in class Y
- P (Y) = Initial probability of class Y
- P (| Y) = Probability X based on the condition of hypothesis H
- P (X) = Probability X

2.2.2 Single Exponential Smoothing

Simple Exponential Smoothing is used for short-term forecasting. In the previous study [11] used $\alpha = 0.7$, with the value of Mean Squared Error = 13.2267 and Mean Absolute Percentage Error (MAPE) = 0.0387% [15]. To compare the forecasting result, writers use $\alpha = 0.9$. The formulas of the Simple Exponential Smoothing are as follows:

$$F_{t+1} = \alpha D_t + (1 - \alpha) F_t \dots\dots\dots (2)$$

Description:
 Ft+1 = Forecasting result
 α = Initialization
 Dt = Current data
 Ft = Previous forecasting

3 Result and Discussion

The forecasting algorithm used in this system is a simple exponential smoothing. The method works based on the average value of data in the past (smoothing) in an exponential manner. The purpose of this method is to determine the minimum α value of means absolute percentage error (MPE).

Following is the implementation of the simple exponential smoothing in the PHP programming.

$$F_{t+1} = \alpha D_t + (1 - \alpha) F_t \dots\dots\dots(3)$$

Implementation:

```
<?php
$no = 1;
foreach($data_komoditi as $index => $data) {
  $alpha=0.9; if($index==0){ $s1[$index]=$data->harga; }else{
  $s1[$index]=$alpha*$data->harga+(1-$alpha)*$s1[$index-1]; }
  $hasil = $s1[$index]; if ($hasil > $data->harga) { $style = "danger";
  $icon = "<i style='color:red;' class='icon-arrow-up'></i>";
  }elseif ($hasil < $data->harga) {$style = "success";
  $icon = "<i style='color:green;' class='icon-arrow-down'></i>";
  }elseif ($hasil == $data->harga) { $style = ""; $icon = "(Stabil) "; }
  $a[$index]=$s1[$index]; $e = $data->harga - $hasil;
  $percenteror = $e/$data->harga; ?>
```

$$M = \frac{A_t - F_t}{A_t} \dots\dots\dots(4)$$

Implementation :

```
$e = $data->harga - $hasil; $percenteror = $e/$data->harga;
```

The following are the results of the prediction of the price of granulated sugar with (value α = 0.9).

Table 3.1 Forecasting of granulated sugar prices

Date	Actual prices	Forecasting prices	Margin of error
2017-02-01	14.000	14.000	0
2017-02-02	14.000	14.000	0
2017-02-03	14.000	14.000	0
2017-02-04	14.000	14.000	0
2017-02-06	14.000	14.000	0
2017-02-07	12.500	12.650	0.0120

2017-02-08	12.500	12.515	0.0012
2017-02-09	12.500	12.502	0.0001

Following is implementation of Naïve Bayes in the PHP programming.

```
$bahan_pokok = $this->input->post('kondisi_bahan_pokok'); $cuaca = $this->input->post('cuaca');
$skondisi_kendaraan = $this->input->post('kondisi_kendaraan'); $persediaan = $this->input->post('persediaan');
```

Step 1; count of total

$P(Y = \text{go up}) = 5/11$ ‘number of data “go up” in the column “prediction result” divided with number of data

$P(Y = \text{down}) = 6/11$ ‘number of data “down” in the column “prediction result” divided with number of data

Program codes:

```
$naik=$this->Naive_bayes_model->Count("WHERE keterangan='Naik'")->num_rows();
$turun=$this->Naive_bayes_model->Count("WHERE keterangan='Turun'")->num_rows();
$total=$this->Naive_bayes_model->Count("")->num_rows();
```

Step 2; calculate number of data with same case

$P(\text{condition of granulated sugar} = \text{Bad} \mid Y = \text{go up}) = 1/5$

‘Total of granulated sugar “bad” and prediction “go up” divided with number of data go up

$P(\text{condition of granulated sugar} = \text{bad} \mid Y = \text{down}) = 4/6$

‘Total of granulated sugar “bad” and prediction “down” divided by number of data down.

Codes:

```
$bahan_pokok_naik = $this->Naive_bayes_model->Count("WHERE kondisi_bahan_pokok like '$bahan_pokok' and keterangan='Naik'")->num_rows();
$bahan_pokok_turun = $this->Naive_bayes_model->Count("WHERE kondisi_bahan_pokok like '$bahan_pokok' and keterangan='Turun'")->num_rows();
```

$P(\text{weather} = \text{good} \mid Y = \text{go up}) = 1/5$

Number of weather data “good” divided by number of data go up

$P(\text{weather} = \text{good} \mid Y = \text{down}) = 4/6$

Number of weather data “good” divided by number of down data

Codes:

```
$cuaca_naik= $this->Naive_bayes_model->Count("WHERE cuaca like '$cuaca' and keterangan='Naik'")->num_rows();
```

```
$cuaca_turun= $this->Naive_bayes_model->Count("WHERE cuaca like '$cuaca' and keterangan='Turun'")->num_rows();
```

$P(\text{Stock} = \text{many} \mid Y = \text{go up}) = 2/5$

Numbers of data with stock “many” divided by number of data go up

$$P(\text{Stock} = \text{many} \mid Y = \text{down}) = 4/6$$

Numbers of data with stock “many” divided by number of data down

Codes:

```
$kondisi_kendaraan_naik=$this->Naive_bayes_model->Count("WHERE
kondisi_kendaraan like '$kondisi_kendaraan' and keterangan='Naik'")->num_rows();
```

```
kondisi_kendaraan_turun=$this->Naive_bayes_model->Count("WHERE
kondisi_kendaraan like '$kondisi_kendaraan' and keterangan='Turun'")->num_rows();
```

$$P(\text{Transportation condition} = \text{good} \mid Y = \text{go up}) = 2/5$$

Numbers of data with Transportation condition “good” divided by number of data down

$$P(\text{Transportation condition} = \text{good} \mid Y = \text{down}) = 3/6$$

Numbers of data with Transportation condition “good” divided by number of data down

Codes:

```
$persediaan_naik= $this->Naive_bayes_model->Count("WHERE persediaan like
'$persediaan' and keterangan='Naik'")->num_rows();
```

```
$persediaan_turun= $this->Naive_bayes_model->Count("WHERE persediaan like
'$persediaan' and keterangan='Turun'")->num_rows();
```

Step 3: Multiply all variable results go up and Down

$$P(\text{Commodity condition} = \text{bad}, (\text{weather} = \text{good}), (\text{stock} = \text{many}), (\text{transportation condition} = \text{good}) \mid \text{go up})$$

$$= \{P(P(\text{Commodity condition} = \text{bad} \mid Y = \text{go up}), P(\text{weather} = \text{good} \mid Y = \text{go up}), P(\text{stock} = \text{many} \mid Y = \text{go up}), P(\text{transportation condition} = \text{good} \mid Y = \text{go up}))\}$$

$$1/5 * 1/5 * 2/5 * 2/5 * 5/11 = 0,0028$$

$$P(\text{Commodity condition} = \text{bad}, (\text{weather} = \text{good}), (\text{stock} = \text{many}), (\text{transportation condition} = \text{good}) \mid \text{down})$$

$$= \{P(P(\text{Commodity condition} = \text{bad} \mid Y = \text{down}), P(\text{weather} = \text{good} \mid Y = \text{down}), P(\text{stock} = \text{many} \mid Y = \text{down}), P(\text{transportation condition} = \text{good} \mid Y = \text{down}))\}$$

$$4/6 * 4/6 * 4/6 * 3/6 * 6/11 = 0,077$$

Codes:

```
$p=($naik/$total)*($bahan_pokok_naik/$naik)*($cuaca_naik/$naik)*($kondisi_kendaraan_naik/$naik)*($persediaan_naik/$naik);
```

```
$p2=($turun/$total)*($bahan_pokok_turun/$turun)*($cuaca_turun/$turun)*($kondisi_kendaraan_turun/$turun)*($persediaan_turun/$turun);
```

Step 4: Comparison result of down and go up

Due to (P|down) greater than (P| go up) the decision is “down”.

```
if ( $p >= $p2 ) {$data['keterangan'] = "go up";} //naik=go up
```

else {\$data['keterangan'] = "Down";} // turun=down
 Result of training data from step 1 to step 4 as shown in table 3.2

Table 3.2 Data training of Naïve Bayes

No	Condition of sugar	Weather	Stock	Transportation condition	Forecasting result
1	Good	Good	Many	Good	Down
2	Good	Good	Many	Bad	Down
3	Good	Good	Many	Bad	Go up
4	Bad	Bad	Few	Bad	Down
5	Bad	Good	Many	Good	Down
6	Bad	Bad	Few	Good	Down
7	Good	Bad	Few	Bad	Go up
8	Good	Bad	Many	Good	Go up
9	Good	Bad	Many	Bad	Go up
10	Bad	Good	Many	Bad	Down
11	Bad	Bad	Few	Good	Go up

If a food commodity (granulated sugar) with data as follows:s

Table 3.3 Result of forecasting data training

Condition of sugar	Weather	Stock	Transportation condition	Forecasting result
Bad	Good	Many	Good	Down

4 Conclusion

Predicting Food Commodity Prices can be useful for the government to make controlling, especially to estimate the food commodity prices in the future. From results of analysis, we found that simple exponential smoothing method with value ($\alpha=0.9$) has the smallest Margin Error. Thus, the value ($\alpha=0.9$) is better than 0.7. Therefore, this value is the most suitable for short-term prices forecasting.

References

- [1] M. A. Rasyidi, "Prediksi Harga Bahan Pokok Nasional Jangka Pendek Menggunakan ARIMA," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 3, no. 2, p. 107, 2017.
- [2] M. E. Lasulika, "Prediksi Harga Komoditi Jagung Menggunakan K-Nn Dan Particle Swarm Optimazation Sebagai Fitur Seleksi," *Ilk. J. Ilm.*, vol. 9, no. 3, p. 233, 2018.
- [3] P. Bhatia, B. Banarasi, D. University, L. D. Syed, H. Ali, and K. M. Chishti, "International Journal on Recent and Innovation Trends in Computing and Communication," *Int. Jounral Recent Innov. trends Comput. Commun.*, vol. 7, no. 3, pp. 4858–4868, 2014.
- [4] S. Syarli and A. A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus : Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 1–5, 2016.
- [5] E. Manalu, F. A. Sianturi, and M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv . Papadan Mama Pastries," *Effrida Manalu, Fricles*

- Ariwisanto Sianturi, Mamed Rofendy Manalu*, vol. 1, no. 2, pp. 16–21, 2017.
- [6] D. Puspitasari, S. S. Al Khautsar, and W. P. Mustika, “Algoritma Naïve Bayes Untuk Memprediksi Kredit Macet Pada Koperasi Simpan Pinjam,” *J. Inform. Upgris*, vol. 4, no. 2, 2019.
- [7] B. Classifier, “Analysis on Students Performance Using Naïve,” *J. Theor. Appl. Inf. Technol.*, vol. 31, no. 16, pp. 3993–4000, 2017.
- [8] A. P. Tribhuvan, P. P. Tribhuvan, and J. G. Gade, “Applying Naive Bayesian Classifier For Predicting Performance Of A Student Using Weka,” *J. Advances in Computational Research*, vol. 7, no. 1, pp. 239–242, 2015.
- [9] A. Hamoud, A. Humadi, W. A. Awadh, and A. S. Hashim, “Students’ Success Prediction Based on Bayes Algorithms,” *Ssrn*, no. November, 2017.
- [10] H. Shaziya, R. Zaheer, and G. Kavitha, “Prediction of Students Performance in Semester Exams using a Naïve bayes Classifier,” *India Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 4, no. 10, pp. 9823–9829, 2015.
- [11] S. Kharya, S. Agrawal, and S. Soni, “Naive Bayes Classifiers: A Probabilistic Detection Model for Breast Cancer,” *Int. J. Comput. Appl.*, vol. 92, no. 10, pp. 26–31, 2014.
- [12] F. Harahap, A. Y. N. Harahap, E. Ekadiansyah, R. N. Sari, R. Adawiyah, and C. B. Harahap, “Implementation of Naïve Bayes Classification Method for Predicting Purchase,” *2018 6th Int. Conf. Cyber IT Serv. Manag.*, no. August 2018, pp. 1–5, 2018.
- [13] D. P. I. M. Meena and D. V. Perumal, “Performance of C4.5 and Naïve Bayes Algorithm to Predict Stomach Cancer - An analysis,” *Ijarcce*, vol. 5, no. 11, pp. 482–487, 2016.
- [14] H. Wasiati and D. Wijayanti, “Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes (Studi Kasus: Di P.T. Karyatama Mitra Sejati Yogyakarta),” *IJNS - Indones. J. Netw. Secur.*, vol. 3, no. 2, pp. 45–51, 2014.
- [15] M. R. Falevy, “Average Dengan Brew Platform Sebagai Mobile Interfaces . Dosen Politeknik Elektronika Negeri Surabaya Institut Teknologi Sepuluh Nopember,” *Inst. Teknol. Sepuluh Nop.*, 2011.