ILKOM Jurnal Ilmiah Vol. 14, No. 3, December 2022, pp. 323-328
Accredited 2nd by RISTEKBRIN No. 200/M/KPT/2020; E-ISSN 2548-7779 | P-ISSN 2087-1716

Prediction the Success of the Government's Program of Lomaya (Regional PKH) in Reducing Poverty

Ruhmi Sulaehani [a,1]; Marniyati H. Botutihe [a,2*]
[a] *Universitas Pohuwto, Jl. Trans Sulawesi No. 147 Kec. Marisa, Marisa and 96266, Indonesia*
[1] *ruhmi.sulaehani@gmail.com;* [2*] *marniyati.h.botutihe@gmail.com*
** Corresponding author*

Article history: Received April 18, 2022; Revised June 03, 2022; Accepted November 29, 2022; Available online December 20, 2022

Abstract

Poverty reduction is one indicator of the success of development. The form of support from the Pohuwato Regency Government through the Social Service is to organize PKH-D, which is known as LOMAYA. It is one of the implementations of the Community Movement Towards Independent Prosperity (*Gerakan Masyarakat Menuju Sejahtera Mandiri*). This research was conducted to assist the government in predicting the level of development success indicated by the satisfaction of beneficiaries of lomaya. The method employed was the Naïve Bayes method and forward feature selection. The research data was obtained from a survey of lomaya beneficiaries in the last two years. The accuracy result obtained using the Naïve Bayes algorithm was 94.19%, while Naïve Bayes with the Forward Selection feature was only 94.03%. Therefore, the Naïve Bayes algorithm method is better than the Forward Selection based Naïve Bayes algorithm. Forward selection does not improve accuracy because the selection process causes many attributes to be discarded because they are considered irrelevant. This happened because of the inaccuracy of the data after being selected for its attributes using the Forward Selection feature resulting 1 attribute only as a determinant.

Keywords: Prediction; Naïve Bayes; PKH-D; Lomaya;

Introduction

Reducing the poverty rate is one indicator of the success of development. The poverty rate refers to the percentage of the macro poverty rate issued by the Central Bureau of Statistics which is calculated based on several variables such as people's purchasing power, school enrollment and quality of health. The Pohuwato Regency Government supports efforts to reduce poverty by providing the Program Keluarga Harapan Daerah - PKH-D called LOMAYA (*Layanan Inovasi Sosial menuju Masyarakat Sejahtera Mandiri* - Social Innovation Service Towards an Independent Prosperous Society) through the Social Affairs Agency. LOMAYA or PKH-D is one of the implementations of the community development Program of *Masyarakat Menuju Sejahtera Mandiri* (MAJU-SM). Financial assistances in the form of Conditional Cash Transfers provided are financed by the Regional Revenue and Expenditure Budget (APBD) of Pohuwato Regency [1]. PKH-D is one of the local government's efforts to provide full support and be synchronized with the Program Keluarga Harapan - PKH so that it is more effective, efficient and focused on handling poverty/reducing inequality in the regions. The target recipients of LOMAYA assistance are the people in Karya Nyata Village, including those recorded in the PBDT, but not beneficiaries of the PKD. Priority is given to people in the lowest deciles.

This research was conducted to assist the government in predicting the poverty reduction program success by measuring the level of satisfaction of beneficiaries. The main objective of the PKH-D program is to reduce poverty. Unfortunately, there is no direct data from the community that can be used as a reference to measure the level of success of government programs that have been implemented in reducing poverty. The data is needed to examine and analyze the efficiency and effectiveness LOMAYA program in increasing people's welfare leading to poverty reduction. Therefore, a system is needed to predict the satisfaction level of PKH beneficiaries as a reference for the government to assess the success of the Lomaya program (PKH-D) in reducing the poverty rate.

http://dx.doi.org/10.33096/ilkom.v14i3.1149.323-328

Several studies have been conducted to examine the prediction of poverty reduction [2]-[5]. However, literatures are lack of the impact of the PKH program on poverty reduction. Instead of focusing on beneficiary satisfaction, researches on the PKH program that has been carried out [6]-[9] have focused more on decision support systems for beneficiary eligibility, whether the distribution of the PKH program is on target or not. Therefore, this study examines the government's success in reducing poverty through PKH program based on the level of satisfaction and benefits received by beneficiaries. This study employs Naïve Bayes and Forward Selection instead of Naïve Bayes as many previous researches [10][13].

Methods

The method employed was the Naïve Bayes method, a simple method in the form of a network. The Bayesian network can be directed as a table with a combined probability distribution. For feature selection, this study applied Forward Selection which functions to correct inconsistent data and clean datasets from noisy and duplicate data. The **Equation 1** of Naïve Bayes is as follow.

$$P(x|y) = \frac{P(x|y)P(x)}{P(y)} \quad (1)$$

Here, y is the data of an unknown class, x is hypothesis of data y is a specific class, P(x|y) is the probability of hypothesis x based on condition y (posteriori probability), P(x) is the probability of hypothesis x (prior probability), P(y|x) is the probability of y based on condition in hypothesis x, P(y) is the probability of y [10].

This research went through some stages. The first stage was to collect data. The data from questionnaire was used as the dataset to be tested. The dataset was collected to predict the government's level of success in reducing the poverty rate. The second stage was data processing, this stage was carried out by converting the questionnaire data into excel format, then analyzing, identifying, correcting inconsistent data and cleaning the dataset from noisy, and duplicating data. After doing this process, the dataset was ready to be processed into machine learning. The third stage was classification. The classification used Naïve Bayes where the dataset determined the level of success. The fourth stage was the accuracy results produced by Naïve Bayes. Fifth stage, the validation stage used cross validation. This stage was used to predict the accuracy of the test data.

Results and Discussion

At the data collection stage, researchers went directly to the research locations, the Social Service Agency of Pohuwato Regency, the Research and Development Planning Agency (Baperlitbang) of Pohuwato Regency and to the villages receiving LOMAYA financial assistance program. The data obtained at the Social Service Agency was in the financial assistance program beneficiaries, in the the Baperlitbang Agency was poverty data while in the villages were questionnaire responses about the government's success regarding LOMAYA financial assistance by looking at beneficiaries' satisfaction level. Some of the community recipients of research assistance gathered in the hall of the village office to distribute questionnaires and some researchers came directly to the homes of the recipients to do an interview about LOMAYA financial assistance and its impact on reducing poverty.

The data on beneficiaries obtained from this study were 967 records, but the data on beneficiaries who filled out the questionnaire were 346 data from several sub-districts in Pohuwato regency. Researchers only asked for data on beneficiaries who had received LOMAYA financial assistance for more than 2 years. The obtained data would be used as a predictive dataset for the government's level of success in reducing poverty with the PKH-LOMAYA Region program by looking at the satisfaction level of the recipients. The obtained data from the Social Service Agency was in the form of beneficiary/recipient data which consisted of several variablesv, district, village, hamlet, family name, NIK-national family number, type of assistance, the amount of money received and 11 questions symbolized by (P1-P11). The data collected in this study was taken from beneficiary data and data from interviewing recipients in Pohuwato regency. The total recorded data was 367. The total data obtained was divided into 90% training data and 10% testing data. From the data obtained, there were questions left empty, also some people did not fill in all the distributed questionnaire. Therefore, the data preprocessing process must be done to get representative data in accordance with the modeling.

The preprocessing process started with the data selection stage. A process carried out before data preprocessing was to select data. Recipient data became case data in the data mining process. Attributes/variables used in decision making were variable P11 and attributes namely district, village, hamlet, family name, NIK-national family number, type of financial assistance, amount of received money, P1, P2, P3, P4, P5, P6, P7, P8, P9 and P10 . The preprocessing/data cleaning stage was applied to delete empty data, and changed inconsistent data, some data was found to have empty responses. Next was data reduction stage where there may be inconsistency data and duplicate data or the same data would be made into one tuple to represent the tuple. In the experimental stage using the rapid miner, it began with importing data, the data used Ms.Excel. The first experiment used the Naïve Bayes method without feature selection. There were several steps to import data. The stage after the data import was the experiment stage. For experimental data, it was divided into data training and data testing. Data training was 90% and data testing

was 10% using split data. The data was then combined with the method used, namely the Naïve Bayes method and connected to the apply model and performance testing.

Furthermore, experiments using the Naïve Bayes method based on forward feature selection was conducted. The feature selected attributes influencing the decision making. The experimental stage using the Naïve Bayes method based on forward selection was started from the main process stage. After it was completed, it proceeded to the validation process stage, and the training testing stage was carried out after the validation process stage was completed.

A. Experiment Result of Naïve Bayes

The testing on the successful rate prediction result of the government using the Naïve Bayes method was carried out with K-Fold Cross Validation. K-Fold Validation was started by dividing the desired number of k-fold data. In this study, the Confusion Matrix was chosen as the evaluation measuring tool to make it easier to analyze the performance of the algorithm because the Confusion Matrix provided information in the form of numbers so that the success ratio can be calculated. Model evaluation was carried out based on tests to estimate true and false objects. The following Confusion Matrix used the Naïve Bayes method. The obtained accuracy results on the beneficiary data set using the Naïve Bayes algorithm are shown in table 4 that shows an accuracy rate of 94.19%. The results of the confusion matrix are as shown in **Table 1**, the performance is shown in **Figure 1**.

Table 1. The Result of Confusion Matrix Using Algorithm Naïve Bayes

	true (Very Satisfied)	True Satisfied	True Quite Satisfied	class precision
Prediction of Very Satisfied	244	7	6	94.94%
Prediction of satisfied	3	47	2	90.38%
Prediction of quite satisfied	0	0	1	100.00%
class recall	98.79%	87.04%	11.11%	

PerformanceVector

```
PerformanceVector:
accuracy: 94.19%
ConfusionMatrix:
True:    Sangat Puas     Puas    Cukup Puas
Sangat Puas:    244      7       6
Puas:   3       47       2
Cukup Puas:     0        0       1
```

Figure 1. Performance Vector Naïve Bayes

This experiment used 310 training records. Based on the confusion matrix, 244 records were predicted to be Very Satisfied in the Very Satisfied data group, 7 records were predicted to be Very Satisfied in the Satisfied data group, 6 records were predicted to be Very Satisfied in the Fairly Satisfied group. Furthermore, 2 records were predicted to be Satisfied but they were in the Fairly Satisfied group, 47 records were predicted to be Satisfied in the Satisfied group, 3 records were predicted to be Satisfied in the Very Satisfied group and 1 record was Prediction Quite Satisfied in the Fairly Satisfied group.

Precision: The proportion of the number of samples with a true value that was correctly predicted.

$$Precision = \frac{TP}{TP + FP} = \frac{244}{244 + 7 + 6} \times 100\% = 94.94\%$$

Recall: The proportion of samples with true values that were correctly predicted.

$$Recall = \frac{TP}{TP + FN} = \frac{244}{244 + 3 + 0} \times 100\% = 98.79\%$$

From these results, the accuracy value can be calculated.

$$Accuraccy = \frac{The\ number\ of\ correct\ prediction}{Total\ numer\ of\ prediction} \times 100\%$$

$$= \frac{244 + 47 + 1}{244 + 7 + 6 + 43 + 47 + 2 + 0 + 0 + 1} \times 100\%$$

$$= \frac{292}{310} \times 100\% = 94.19\%$$

B. Experiment Result of Naïve Bayes and Forward Selection

In the experiment using feature selection, the selection feature first selected attributes considered to be used as attributes that can be very influential in predicting the satisfaction level of the communities. In the experiment using the Forward Selection feature, many attributes were discarded because considerable they did not have any effect and only one attribute became very influential in determining the prediction results, namely attribute P6 (Question 6). It helps reduce the burden of poor family needs? a. Very Helpful (Point 3), Helpful (Point 2), Fairly Helpful (Point 1), Not Helpful (Point 0). The following table results of the Confusion Matrix Using the Naïve Bayes and Forward Selection Algorithms. The results of the confusion matrix were shown in **Table 2**, and the results of the performance vector were shown in **Figure 2**.

Table 2. Confusion Matrix Result Using Naïve Bayes Algorithms and Forward Selection

	True very satisfied	**true satisfied**	**true quite satisfied**	**class precision**
Prediction of Very Satisfied	260	3	3	97.74%
Prediction of Satisfied	7	55	7	79.71%
Prediction of quite satisfied	0	0	0	0.00%
class recall	97.38%	94.83%	0.00%	

PerformanceVector

```
PerformanceVector:
accuracy: 94.03%
ConfusionMatrix:
True:   Sangat Puas     Puas    Cukup Puas
Sangat Puas:    260     3       3
Puas:   7       55      7
Cukup Puas:     0       0       0
```

Figure 2. Performance Vector Using Naïve Bayes and Forward Selection

Precision: The proportion of the number of samples with true values that was correctly predicted.

$$Precision = \frac{TP}{TP + FP} = \frac{260}{260 + 3 + 3} = \times 100\% = 97.74\%$$

Recall: The proportion of samples with true values that were correctly predicted.

$$Recall = \frac{TP}{TP + FP} = \frac{260}{260 + 7 + 0} \times 100\% = 97.38\%$$

From these results, the accuracy value can be calculated.

$$Accuracy = \frac{The\ number\ of\ correct\ prediction}{Total\ number\ of\ prediction} \times 100\%$$

$$= \frac{260 + 55 + 0}{260 + 3 + 3 + 7 + 55 + 7 + 0 + 0 + 0} \times 100\% = \frac{294}{310} \times 100\% = 94.03\%$$

C. Comparison of the results of Naïve Bayes method with Naïve Bayes based on Forward Selection method

From the experimental results, it can be compared that the Naïve Bayes algorithm method was better than using the Naïve Bayes algorithm based on Forward Selection. Causes Forward Selection cannot improve accuracy results due to inaccurate data after the data was selected for its attributes using the Forward Selection feature. Many attributes were discarded because they were labelled irrelevant and only used 1 attribute as a determinant in predicting the success rate of the government with the PKH-Regional program (LOMAYA) by looking at the level of satisfaction of the beneficiaries. Accuracy result obtained using the Naïve Bayes algorithm was 94.19%, while Naïve Bayes based on Forward Selection Features was 94.03%.

The assessment of the government's level of success in reducing the poverty rate with the PKH-Regional Program can be seen from satisfaction level of the beneficiaries. This system can be used to predict the successful level by looking at the level of satisfaction from the predicted data using the Naïve Bayes method. The percentage of local government success rates with LOMAYA financial assistance program by looking at the level of satisfaction of the beneficiaries can be seen in the PKH-D (LOMAYA) Program Success Rate Graph. As seen in the graph below the level of satisfaction with very satisfied was 83%, satisfied was 17%, sufficient was 0% while dissatisfied was not included in the graph because none of the beneficiaries was dissatisfied with the PKH-D Assistance Program (LOMAYA).

Conclusion

Based on the experimental results, the conclusion of this study is that the Naïve Bayes algorithm produces the best accuracy compared to Naïve Bayes with Forward Selection. Forward Selection did not improve accuracy because when attributes were selected using the Forward Selection feature, many of them were discarded because they were considered irrelevant. This occurred due to the inaccurate data, and it used only 1 attribute as a determinant in predicting the government success rate of implementing the PKH-Regional (LOMAYA) program. This research can be used by the government in predicting the government's success in reducing poverty with the PKH-Regional (LOMAYA) program. This research can also help find out whether the PKH-Regional (LOMAYA) program has been distributed properly and appropriately or not. In addition, this research can help the government determine the satisfaction level of the PKH-Regional (LOMAYA) program beneficiaries.

References

[1] Dinas Sosial Kabupaten Pohuwato “juknis lomaya 2020”. 2020

[2] Syaharuddin, E. Pujiana, I. P. Sari, V. M. Mardika, and M. Putri, “Analisis algoritma Back Propagation dalam prediksi angka kemiskinan di Indonesia,” J. Pendidik. Berkarakter, vol. 3, no. 1, pp. 11–17, 2020.

[3] H. S. Pakpahan, Y. Basani, and R. R. Hariani, “Prediksi jumlah penduduk miskin kalimantan timur menggunakan Single dan Double Exponential Smoothing,” Inform. Mulawarman J. Ilm. Ilmu Komput., vol. 15, no. 1, pp. 47–51, 2020. 5, pp. 930-936, Oct. 2020.

[4] A. Karim, U. Islam, and N. Walisongo, “Perbandingan prediksi kemiskinan di Indonesia Menggunakan Support Vector Machine (SVM) dengan Regresi Linear,” vol. 6, no. 1, pp. 107–112, 2020.

[5] R. H. Purba, M. Zarlis, and I. Gunawan, “TIN : Terapan Informatika Nusantara algoritma Backpropagation dalam memprediksi jumlah angka kemiskinan di Provinsi Sumatera Utara TIN : Terapan Informatika Nusantara,” vol. 1, no. 1, pp. 55–63, 2020.

[6] S. A. Suleman and R. Resnawaty, “Program Keluarga Harapan (Pkh): antara Perlindungan Sosial dan pengentasan kemiskinan,” Pros. Penelit. dan Pengabdi. Kpd. Masy., vol. 4, no. 1, p. 88, 2017. Mojokerto) The Policy Implementation of

[7] D. Wintana, H. Hikmatulloh, N. Ichsan, J. J. Purnama, and A. Rahmawati, “Klasifikasi penentuan penerima manfaat program keluarga harapan (pkh) menggunakan algoritma C5.0 (Studi kasus: Desa Sukamaju, Kec.Kadudampit),” Klik - Kumpul. J. Ilmu Komput., vol. 6, no. 3, p. 254, 2019.

[8] E. Fitriani, “Perbandingan algoritma C4.5 dan Naïve Bayes untuk menentukan kelayakan penerima bantuan program keluarga harapan,” Sistemasi, vol. 9, no. 1, p. 103, 2020.

[9] A. A. Sidiq and F. W. Christanto, “Algoritma Naive Bayes Untuk Penentuan Pkh (Program Keluarga Harapan) Berbasis Sistem Pendukung Keputusan,” Riptek, vol. 14, no. 1, pp. 65–71, 2020.

[10] T. H. Apandi, C. A. Sugianto, and C. R. Service, “Algoritma Naive Bayes untuk prediksi kepuasan pelayanan perekaman e-KTP (Naive Bayes Algorithm for Satisfaction Prediction of e-ID,” vol. 7, no. November, pp. 125–128, 2019.

[11] Noviantoro, Tri, and Jen-Peng Huang. "Investigating airline passenger satisfaction: Data mining method." Research in Transportation Business & Management 43 (2022): 100726.

[12] S. F. Shetu, I. Jahan, M. M. Islam, R. Ara Hossain, N. N. Moon and F. Narin Nur, "Predicting Satisfaction of Online Banking System in Bangladesh by Machine Learning," 2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST), 2021, pp. 223-228, doi: 10.1109/ICAICST53116.2021.9497796.

[13] M. Siddik, Y. Desnelita, and Gustientiedina, "Penerapan Naïve Bayes untuk memprediksi tingkat kepuasan mahasiswa terhadap pelayanan akademis," J. Infomedia, vol. 2, no. 4, pp. 89–93, 2019.

[14] K. A. Aeni, "Prediksi kepuasan layanan akademik menggunakan Algoritma Naïve Bayes," JATISI (Jurnal Tek. Inform. dan Sist. Informasi), vol. 7, no. 3, pp. 601–609, 2020.

[15] M. G. Sadewo, A. P. Windarto, and I. S. Damanik, "Algoritma Naïve Bayes dalam memprediksi kepuasan nasabah," Pros. Semin. Nas. Ris. Inf. Sci., vol. 1, no. September, p. 318, 2019.

[16] A. R. Damanik, S. Sumijan, and G. W. Nurcahyo, "Prediksi tingkat kepuasan dalam pembelajaran daring menggunakan algoritma Naïve Bayes," J. Sistim Inf. dan Teknol., vol. 3, pp. 88–94, 2021.

[17] J. A. Baba, G. Yanti, K. Sari, S. Pahu, and R. H. Saputra, "Penerapan algoritma Decision Tree dan Feature Selection Forward untuk analisis kepuasan pemasang iklan terhadap pelayanan iklan," vol. 3, no. 3, pp. 133–139, 2018.

[18] A. Prasetyo, "Algoritma Naïve Bayes berbasis forward selection pada prediksi kelulusan tepat waktu" November 2017, 2018.

[19] D. Berrar, "Bayes' theorem and naive bayes classifier," Encycl. Bioinforma. Comput. Biol. ABC Bioinforma., vol. 1–3, no. January 2018, pp. 403–412, 2018.

[20] M. Guntur, J. Santony, and Y. Yuhandri, "Prediksi harga emas dengan menggunakan Metode Naïve Bayes dalam Investasi untuk meminimalisasi resiko," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 2, no. 1, pp. 354–360, 2018.

[21] M. R. Fanani, "Algoritma Naïve Bayes berbasis Forward Selection untuk prediksi bimbingan konseling siswa," J. DISPROTEK, vol. 11, no. 1, pp. 13–22, 2020.

[22] J. Nasional and S. Informasi, "Implementasi Forward Selection dan Bagging untuk prediksi kebakaran hutan menggunakan Algoritma Naïve Bayes," vol. 01, pp. 1–8, 2022.

[23] F. Nuraeni, Y. H. Agustin, S. Rahayu, D. Kurniadi, Y. Septiana and S. M. Lestari, "Student Study Timeline Prediction Model using Naïve Bayes Based Forward Selection Feature," 2021 International Conference on ICT for Smart Society (ICISS), 2021, pp. 1-5, doi: 10.1109/ICISS53185.2021.9532502.