APPLICATION OF THE APRIORI ALGORITHM TO DETERMINE THE COMBINATION OF POVERTY INDICATORS

Sri Siswanti[1]; Retno Tri Vulandari[2*)]; Andriani Kusumaningrum[3]; Setiyowati[4]

[1]Informatika; [2]Teknologi Informasi; [1]Sistem Informasi Akuntansi; [1]Sistem Informasi;
STMIK Sinar Nusantara Surakarta
www.sinus.ac.id
[1]syswanty@sinus.ac.id; [2*)]retnotv@sinus.ac.id; [3]andriani@sinus.ac.id; [4]setiyowati@sinus.ac.id
(*) Corresponding Author

Abstract— Poverty is a society that has not been solved until now. The decline in poverty in Laweyan District from 2000 to 2013 was 5.71%, among the five lowest in the reduction in the percentage of poverty in Central Java Province. The problem of poverty is very complex, and the differences in regional characteristics, as well as the techniques used, also influence the indicators of the causes of poverty and the formulation of policies for poverty alleviation. This study uses Principal Component Analysis as part of data preprocessing, followed by applying association rules with the Apriori Algorithm to explore the relationship pattern of poverty indicators. Based on the research that has been conducted on the poverty dataset, which consists of 46 attributes, it is found that the attributes that have passed the preprocessing data are six attributes, namely the Poor Population, ADHB in the Communication Sector, ADHB in the Mining and Excavation Sector, ADHB in the Agriculture and Food Crops Sector, ADHB in the Plantation Sector. and unemployment. These six attributes are transformed into Ascending, Fixed, and Descending categorical data. The fuzzification process for the increase and decrease categories uses the shoulder-type triangle membership function. Applying the Apriori Algorithm to the poverty dataset with a minimum support of 0.4 and a minimum confidence of 0.8 produces 38 rules that show the relationship between indicators and poverty and 134 rules that show the relationship pattern between indicators.

Keywords: Aprioric Rule; Association Rule; Poverty.

Abstrak—*Kemiskinan merupakan masalah yang sampai saat ini belum memiliki solusi yang efektif, meskipun berbagai teori, metode dan langkah-langkah yang harus diambil telah dijelaskan oleh para ahli untuk mengatasi masalah kemiskinan. Penurunan persentase kemiskinan di Kabupaten Laweyan dari tahun 2000 sampai dengan tahun 2013 adalah sebesar 5,71% yang termasuk lima terendah dalam penurunan persentase kemiskinan di Provinsi Jawa Tengah. Masalah kemiskinan sangat kompleks dan perbedaan karakteristik daerah serta teknik yang digunakan turut mempengaruhi indikator penyebab kemiskinan dan perumusan kebijakan penanggulangan kemiskinan. Indikator kemiskinan yang digunakan pada penelitian sebelumnya, perlu dicek indikator mana yang lebih berdampak pada kemiskinan agar pola hubungan yang diperoleh lebih menarik. Dalam penelitian ini, metode yang digunakan adalah Principal Component Analysis sebagai bagian dari preprocessing data yang dilanjutkan dengan penerapan aturan asosiasi dengan Algoritma Apriori untuk menggali pola hubungan indikator kemiskinan. Berdasarkan penelitian yang telah dilakukan terhadap dataset kemiskinan yang terdiri dari 46 atribut, diketahui bahwa atribut yang telah lolos preprocessing data sebanyak 6 atribut yaitu Penduduk Miskin, ADHB Bidang Komunikasi, ADHB Bidang Pertambangan dan Bidang Penggalian, ADHB pada Bidang Pertanian dan Tanaman Pangan, ADHB pada Bidang Perkebunan. dan pengangguran. 6 atribut. Proses fuzzifikasi untuk kategori naik dan turun menggunakan fungsi keanggotaan segitiga tipe bahu. Penerapan Algoritma Apriori pada dataset kemiskinan dengan support minimal 0,4 dan confidence minimal 0,8 menghasilkan 38 rule yang menunjukkan pola hubungan antara indikator dengan kemiskinan dan 134 rule yang menunjukkan pola hubungan antar indikator.*

Kata Kunci*: Algoritma Apriori, Aturan Asosiasi, Kemiskinan*

INTRODUCTION

Poverty is a problem that has yet to have an effective solution to date, although various theories, methods, and steps that must be taken have been described by experts to tackle the problem of poverty. Based on data from the Central Bureau of Statistics, Central Java Province, the percentage of poverty in Indonesia from 2000 to 2013 has decreased by 7.67% from 19.14% to 11.47%. However, this percentage is still far from the hope of this nation to be free from poverty.

Central Java is one of Indonesia's provinces that also experiences poverty reduction problems. The decline in the percentage of poverty in Central

Java from 2000 to 2013 was 8.29%. One of the districts in Central Java Province which are among the five lowest in reducing the percentage of poverty is Kecamatan Laweyan, where the percentage of poverty in 2000 was 10.04% while in 2013, it was 4.33% (Badan Pusat Statistik Jawa Tengah, 2022).

Nowadays, poverty is no longer seen as a problem of a person's low income but is caused by various factors that cause poverty. As stated by Nurwati (Nurwati, 2008), poverty is a multidimensional problem because it correlates with education, type of work, gender, access to essential services, and infrastructure.

The opinion that poverty is a multidimensional problem is also supported by Permana (2016a) and Rusdati and Sebayang (Mei Alfianto, Istiyani, & Priyono, 2019). The unemployment factor and Gross Regional Domestic Product (GRDP) were included in the study of poverty in these two studies but gave different results. Income distribution has a positive correlation to economic growth. This means that if the distribution of income increases, then economic growth will also increase (Silva, 2016). On the other hand, Rusdati and Sebayang showed that the unemployment rate did not have a significant effect, while GRDP and public spending significantly affected poverty. Wanto (2018) found that the GRDP of the services sector at constant prices, the number of unemployed, per capita average expenditure per month for durable goods, and per capita, average expenditure for fruit groups are closely related to the number of poor people in Riau.

Gross regional domestic product and human development index affect poverty. This also occurred in 35 districts/cities in the province of Central Java in 2011 - 2015. The growth rate of the gross regional domestic product positively affects the poverty level, meaning that the higher the market price, the higher the poverty level. The human development index hurts the poverty level, meaning that the more the human development index value increases, the poverty level decreases.(Andhykha, Handayani, & Woyanti, 2018) This shows that the problem of poverty is very complex, and the differences in regional characteristics and the techniques used are also influential in determining indicators of the causes of poverty and formulating policies for poverty alleviation.

With the variety of poverty indicators used in previous studies, it is necessary to check which indicators have more impact on poverty so that the pattern of relationships obtained is more attractive. Wanto (2018) uses Principal Component Analysis (PCA) to reduce the poverty variable, where the variables taken are those with an eigenvalue of more than 1. In Wanto (2018), this algorithm can remember and generalize what exists. There are five architectural models. Besides Wanto, other researchers using PCA are Adji et al. (Hendro, Adji, & Setiawan, 2012), to reduce 13 variables into four variables that have more influence on coronary heart disease.

Data mining is one of the fields of science that can explore the relationship between these indicators and poverty. Data mining that combines methods and tools in statistics, databases, information theory, data visualization, and machine learning has been widely used to explore information, patterns, and data characteristics. The main tasks in data mining are summarization, sequence, classification, association, and clustering (Han, Kamber, & Pei, 2012). Association rules are tasks used to find relationships between items in a dataset.

Hakim and Fauzy use association rules with the Apriori algorithm to determine the relationship pattern of traffic accidents with age, type of accident, time, ownership of a driving license (SIM), occupational gender, and accident rate (Hakim & Fauzy, 2015). Sitanggang determines indicators of possible effects of fire events. The indicators studied were the physical environment, socio-economic conditions, weather, and peatlands (Sukaesih Sitanggang, 2013). Arafah & Mukhlash, (2015) use the Apriori algorithm to analyze patterns of stock price movements between companies and between sectors in Indonesia, where the dataset consists of 63 companies from 9 sectors.

Based on the studies conducted, it can be seen that the indicators of the causes of poverty are very diverse. Differences in conditions from one region to another also have a different effect on poverty that occurs in each region. Therefore, we need an effective way to alleviate poverty in each region.

MATERIALS AND METHODS

The research stages are explained as follows (Farida, Chulkamdi, & Wulansari, 2022).

The data in this study are secondary data from the Central Bureau of Statistics, Central Java Province, 2016–2019. The data consists of the number of poor people, average expenditure per capita for the food group, average expenditure per capita for the non-food group, Gross Regional Domestic Product (GRDP) in 9 sectors, inflation, export value, and harvested area for rice and secondary crops. From 7 groups, the number of large and medium manufacturing industries, the number of unemployed, and the number of residents. This dataset is then formed into Dataset 1 and Dataset 2.

After data collection is complete, data preprocessing is divided into four processes (Edastama, Bist, & Prambudi, 2021).

- Data cleaning, namely removing data objects that are incomplete on the factor.
- Data interpolation, namely converting annual data to monthly data.
- Data selection, namely eliminating factors that do not significantly influence poverty using Principal Component Analysis (PCA).
- Data transformation, namely converting data into categorical data using fuzzification.

Pattern Determination using Association Rules At this stage, pattern determination from categorical data is carried out using association rules with the Apriori algorithm.

Pattern Analysis and Knowledge Representation The pattern obtained from applying the Apriori algorithm association rules is then represented by identifying exciting patterns.

Preparation of Reports and Publications: report writing and paper preparation are carried out for publication at this stage.

RESULTS AND DISCUSSION

1. Data Collection

The dataset is secondary data from the Central Bureau of Statistics, Central Java Province, in 2016-2019. The dataset consists of 46 attributes. The attributes contained in the dataset are as follows.

- Several poor people (A1).
- Average per capita expenditure on food group (A2).
- Average per capita expenditure for the non-food group (A3).
- Gross Regional Domestic Product (GRDP) is divided into 14 sectors based on the Current Price (ADHB) and Constant Price (ADHK) in each sector, namely GRDP in the Building Sector; PDRB in the Manufacturing Sector; PDRB in Services Sector, PDRB in the Financial Sector; PDRB of the Rental and Company Services Sector; PDRB of the Electricity and Gas Sector; PDRB of Clean Water Sector; PDRB in Transportation Sector; GDP in the Communication Sector; PDRB of Trade, Hotel and Restaurant Sector; PDRB in Mining and Excavation Sector; GRDP of Agriculture and Food Crops Sector; GRDP of the Plantation Sector; GDP of the Animal Husbandry Sector; Forestry Sector PDRB; as well as GRDP of the Fisheries Sector (A4 to A35).
- Inflation (A36).
- Export value (A37).
- The harvested area for rice and secondary crops consists of 7 groups: rice; corn; soy; peanuts; green beans; cassava; and sweet potato (A38 to A43).
- Several large and medium manufacturing industries (A44).
- Several are unemployed (A45).
- Total population (A46).

2. Preprocessing Data

This stage is divided into five stages: data cleaning, data interpolation, data selection, data fuzzification, and data transformation. Following are the preprocessing stages of the poverty dataset.

Data cleaning is performed to eliminate inconsistent data. ADHB (A6) and ADHK (A22) Attributes The Services Sector for 2009-2010 and 2011-2014 have different data collection services, so the two attributes must be removed from the dataset so that the current dataset has 44 attributes.

Data on attributes are annual, and some are monthly, so it is necessary to interpolate annual data into monthly data. Data interpolation was carried out on 35 attributes other than the attributes of inflation and rice plant area, corn plant area, soybean plant area, peanut plant area, green bean plant area, cassava plant area, and sweet potato plant area. Interpolation was carried out using the Matlab program. The results of interpolating annual data into monthly data can be seen in Table 1.

Table 1. Interpolation Result Dataset

Year	Month	A1	A2	⋯	A45	A46
2009	Jan	15969.85	286776.8	⋯	6917.538	36021.38
	Feb	15793.69	288581.5	⋯	7007.077	55896.77
	Mar	15617.54	290386.3	⋯	7096.615	75772.15
	⋮	⋮	⋮	⋮	⋮	⋮
	Nov	14208.31	304824.5	⋯	7812.923	234775.2
	Dec	14032.15	306629.2	⋯	7902.462	254650.6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2014	Jan	13683.62	474038.5	⋯	4603.308	146138.3
	Feb	13713.23	474048.9	⋯	4617.615	160551.6
	Mar	13742.85	474059.4	⋯	4631.923	174964.9
	⋮	⋮	⋮	⋮	⋮	⋮
	Nop	13979.77	474143.1	⋯	4746.385	290271.4
	Dec	14009.38	474153.5	⋯	4760.692	304684.7

PILAR Nusa Mandiri: Journal of Computing and Information System
Vol. 19, No. 1 March 2023 | DOI: 10.33480/pilar.v18i1.4161

3. Data Selection

The Apriori algorithm is quite wasteful in memory usage and spends the most time during the scanning process.(Fitrina, Kustanto, & Vulandari, 2018) The interpolated dataset consisting of 44 attributes with 72 records will make applying the Apriori Algorithm less effective because the large number of rules generated will make it difficult to understand the patterns between indicators. Therefore, it is necessary to perform data selection using Principal Component Analysis (PCA) to obtain the most related attributes.

Based on the results of attribute reduction using PCA, six attributes were obtained, namely the Poor (A1), ADHB for the Communication Sector (A12), ADHB for the Mining and Excavation Sector (A14), ADHB for the Agriculture and Food Crops Sector (A15), ADHB for the Plantation Sector (A16), and unemployment (A45) which can be seen in Table 2.

Table 2. PCA Results Dataset

Year	Month	A1	A12	A14	A15	A16	A45
2009	Jan	15969.8	9803.8	301620.6	35800.6	9803.8	6917.5
	Feb	15793.6	9838.9	304930	36012.5	9838.9	7007.0
	Ma	15617.5	9873.9	308239.5	36224.4	9873.95	7096.6
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	Nop	14208.3	10154.2	334715.4	37919.4	10154.2	7812.9
	Dec	14032.1	10189.2	338024.8	38131.3	10189.24	7902.4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2014	Jan	13683.6	233560	3310111	502004.5	233560	4603.3
	Feb	13713.2	236211.2	3311692	505047.3	236211.2	4617.6
	Ma	13742.8	238862.4	3313274	508090.1	238862.4	4631.9
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	Nop	13979.7	260072	3325924	532432.6	260072	4746.3
	Dec	14009.3	262723.2	3327506	535475.4	262723.2	4760.6

4. Data Transformation

The data transformation process begins by calculating the difference from the sorted data and looking for the percentage of the data. The results of the percentage of data are then divided into three parts, namely, increase, stay, and decrease.

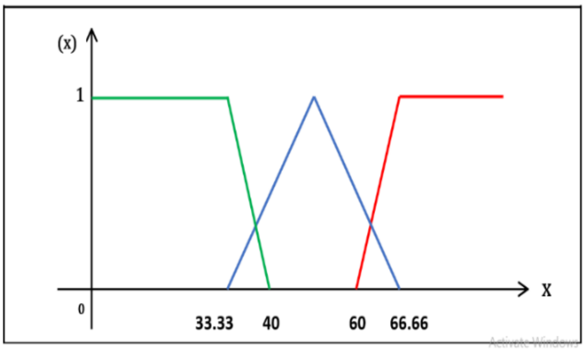

Figure 1. Representation of Membership Functions for Increments

Based on Figure 1 and Figure 2 explains that the movement does not change. A fixed category is given. Meanwhile, to determine the category of increase and decrease, the data was fuzzified using the shoulder-type triangle membership function for the percentage increase in value and the shoulder-type triangle membership function for the percentage decrease in value. Each increase and decrease membership function is divided into three categories, namely Low, Medium, and High, which are described as follows. The representation of the membership function for the percentage increase in value is presented in Figure 1. The representation of the membership function for the percentage reduction value is presented in Figure 2.

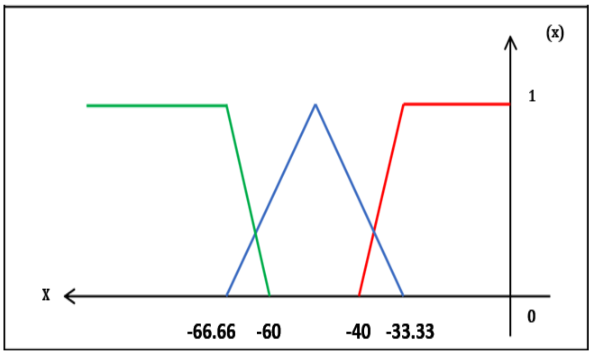

Figure 2. Representation of Membership Functions for Decrease

5. Association Rules Process with Apriori Algorithm

Here is an example of applying the Apriori Algorithm on six attributes for ten records. The dataset in Table 3 is part of the actual dataset that has gone through data preprocessing. So that Table 3 has been in the form of categorical data.

Table 3. Sample Dataset

T	A1	A12	A14	A15	A16	A45
1	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
2	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
3	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
4	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
5	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
6	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
7	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45
8	NR_A1	NT_A12	NS_A14	TR_A15	NT_A16	TR_A45
9	TR_A1	NS_A12	NR_A14	TR_A15	NS_A16	NR_A45
10	TR_A1	NR_A12	NR_A14	TR_A15	NR_A16	NR_A45

The application of the Apriori Algorithm to Table 3. is as follows. Select k = 1 so that the 1-itemset result is shown in Table 4. Based on Table 4 explains low increase category (NR) for the number of poor people (A1) has 80% support. This means that the effect of the number of poor people on poverty indicators is 80%. The low increase category (NR) in the communications sector (A12) has 80% support. This means that the influence of the communication sector on poverty indicators is 80%. As well as the explanation for the value of support for each poverty indicator.

Table 4. One-itemset

No	Itemset	Support
1	NR_A1	80%
2	TR_A1	20%
3	NR_A12	80%
4	NS_A12	10%
5	NT_A12	10%
6	NR_A14	90%
7	NS_A14	10%
8	NR_A15	70%
9	TR_A15	30%
10	NR_A16	80%
11	NS_A16	10%
12	NT_A16	10%
13	NR_A45	20%
14	TR_A45	80%

Table 5. One-itemset Selection Process

No	Itemset	Support
1	NR_A1	80%
2	TR_A1	15%
3	NR_A12	80%
4	NS_A12	10%
5	NT_A12	10%
6	NR_A14	90%
7	NS_A14	15%
8	NR_A15	70%
9	TR_A15	30%
10	NR_A16	80%
11	NS_A16	5%
12	NT_A16	15%
13	NR_A45	10%
14	TR_A45	80%

Determine minimum support = 30%, then select itemset that has support ≥ minimum support.

In Table 5, the coloured data is not selected because it does not meet the minimum support threshold, so the selected data is presented in Table 6.

Table 6. Results of 1-itemset Selection

No	Itemset	Support
1	NR_A1	80%
2	NR_A12	80%
3	NR_A14	90%
4	NR_A15	70%
5	TR_A15	30%
6	NR_A16	80%
7	TR_A45	80%

Next, select k = k + 1 to determine the 2-item set. Then choose a 2-item set that meets the minimum support threshold. Similar to the work in Step 2, the coloured data was not selected because it did not meet the minimum support threshold, so the selected data was presented in Table 7.

Table 7. 2nd-itemset selection Data

No	Itemset		Support
1	NR_A1	NR_A12	70%
2	NR_A1	NR_A14	70%
3	NR_A1	NR_A15	70%
4	NR_A1	NR_A16	70%
5	NR_A1	TR_A45	80%
6	NR_A12	NR_A14	80%
7	NR_A12	NR_A15	70%
8	NR_A12	NR_A16	80%
9	NR_A12	TR_A45	70%
10	NR_A14	NR_A15	70%
11	NR_A14	NR_A16	80%
12	NR_A14	TR_A45	70%
13	NR_A15	NR_A16	70%
14	NR_A15	TR_A45	70%
15	NR_A16	TR_A45	70%

The 3-itemset determination can be seen in Table 8.

Vol. 19, No. 1 March 2023 | DOI: 10.33480/pilar.v18i1.4161

Table 8. 3rd – itemset Selection Data

No	Itemset			Support
1	NR_A1	NR_A12	NR_A14	70%
2	NR_A1	NR_A12	NR_A15	70%
3	NR_A1	NR_A12	NR_A16	70%
4	NR_A1	NR_A12	TR_A45	70%
5	NR_A1	NR_A14	NR_A15	70%
6	NR_A1	NR_A14	NR_A16	70%
7	NR_A1	NR_A14	TR_A45	70%
8	NR_A1	NR_A15	NR_A16	70%
9	NR_A1	NR_A15	TR_A45	70%
10	NR_A1	NR_A16	TR_A45	70%
11	NR_A12	NR_A14	NR_A15	70%
12	NR_A12	NR_A14	NR_A16	80%
13	NR_A12	NR_A14	TR_A45	70%
14	NR_A12	NR_A15	NR_A16	70%
15	NR_A12	NR_A15	TR_A45	70%
16	NR_A12	NR_A16	TR_A45	70%
17	NR_A14	NR_A15	NR_A16	70%
18	NR_A14	NR_A15	TR_A45	70%
19	NR_A14	NR_A16	TR_A45	70%
20	NR_A15	NR_A16	TR_A45	70%

The 6-itemset determination can be seen in Table 9.

Table 9. 6th – item set Selection Data

No	Itemset						Support
1	NR_A1	NR_A12	NR_A14	NR_A15	NR_A16	TR_A45	70%

Based on the itemsets obtained, 574 rules are generated, presented in Table 10.

Table 10. Best Rule in 6th-itemset Selection Data

No	Best Rules	Conf
1	A45=TR_A45 ==> A1=NR_A1	100%
2	A1=NR_A1 ==> A45=TR_A45	100%
3	A12=NR_A12 ==> A14=NR_A14	100%
⋮	⋮	⋮
572	A16=NR_A16 ==> A1=NR_A1 A12=NR_A12 A14=NR_A14 A15=NR_A15 A45=TR_A45	88%
573	A12=NR_A12 ==> A1=NR_A1 A14=NR_A14 A15=NR_A15 A16=NR_A16 A45=TR_A45	88%
574	A1=NR_A1 ==> A12=NR_A12 A14=NR_A14 A15=NR_A15 A16=NR_A16 A45=TR_A45	88%

The application of the Apriori Algorithm to the poverty dataset consisting of 6 attributes and 71 records (data attached) with a minimum support of 40% and a minimum confidence of 80% in Java is presented in Table 11 and Table 12. Based on Table 11 explains the best rules between the indicator of poverty. Based on Table 12, applying the Apriori algorithm to the poverty dataset produces 38 rules. This shows the relationship between indicators and poverty. Rule 1 on the relationship between indicators and poverty can be seen if the number of poor people is low, then it is related to the unemployment sector

Table 11. Rules Between Indicators

No	Best Rules	Conf
1	A12=NR_A12 ==> A14=NR_A14	100%
2	A16=NR_A16 ==> A12=NR_A45	100%
3	A12=NR_A12 ==> A14=NR_A14	100%
⋮	⋮	⋮
132	A14=NR_A14 ==> A12=NR_A12 A15=NR_A15	83%
133	A14=NR_A14 ==> A15=NR_A15 A16=NR_A16	83%
134	A14=NR_A14 ==> A12=NR_A12 A15=NR_A15 A16=NR_A16	83%

Table 12. Rules Between Indicators and Poverty

No	Best Rules	Conf
1	A1=TR_A1 ==> A14=NR_A14	100%
2	A1=TR_A1 A12=NR_A12 ==> A14=NR_A14	100%
3	A1=TR_A1 A16=NR_A16 ==> A12=NR_A12	100%
⋮	⋮	⋮
36	A1=TR_A1 ==> A14=NR_A14 A16=NR_A16	98%
37	A1=TR_A1 A14=NR_A14 ==> A12=NR_A12 A16=NR_A16	98%
38	A1=TR_A1 ==> A12=NR_A12 A14=NR_A14 A16=NR_A16	98%

6. Results Analysis

Based on Table 11, applying the Apriori algorithm to the poverty dataset produces 134 rules showing the relationship between indicators. Rule 1 on the relationship between indicators shows that when the ADHB in the Communication Sector experiences a low rise, the ADHB for the Mining and Excavation Sector will also experience a low rise. Meanwhile, based on Table 12, applying the Apriori algorithm to the poverty dataset resulted in 38 rules showing the relationship between indicators and poverty. Rule 1 on the relationship between indicators and poverty shows that if the number of poverty decreases is low, it is associated with ADHB in the Mining and Excavation Sector, which has a low rise. In other words, ADHB in the Mining and Excavation Sector will have an impact on reducing the number of poverty on a low scale.

CONCLUSION

The results of this study can be concluded as follows. The preprocessing process, which includes data cleaning and data selection using Principal Component Analysis, reduces the attributes to 6 attributes, namely the number of poor people;

ADHB Communication Sector; ADHB Mining and Quarrying Sector, ADHB Agriculture and Food Crops Sector; ADHB Plantation Sector; and unemployment. The preprocessing process for data transformation uses fuzzy variables with a shoulder-type membership function for increasing and decreasing categories and fixed categories if there is no change in data (t + 1) from data (t). Based on the results of applying the Apriori algorithm to the poverty dataset, it is found that the association between indicators and poverty is 38 rules, with the lowest confidence value is 98%. Based on the result of applying the Apriori algorithm to the poverty dataset, it is found that the association relationship between the indicators is 134 rules, with the lowest confidence value is 83%. The application of the Apriori Algorithm can be optimal, so in further research, several things can be done as follows. The fuzzification process is calculated with different membership functions for each attribute by considering the data distribution. The application of the Apriori Algorithm to find association patterns should be developed in other fields with shorter time intervals, for example, daily, to obtain more detailed patterns.

REFERENCE

Andhykha, R., Handayani, H. R., & Woyanti, N. (2018). Analisis Pengaruh PDRB, Tingkat Pengangguran, dan IPM Terhadap Tingkat Kemiskinan di Provinsi Jawa Tengah. *Media Ekonomi Dan Manajemen*, *33*(2), 113–123. https://doi.org/10.24856/mem.v33i2.671

Arafah, A. A., & Mukhlash, I. (2015). The Application of Fuzzy Association Rule on Co-movement Analyze of Indonesian Stock Price. *Procedia Computer Science*, *59*(Iccsci), 235–243. https://doi.org/10.1016/j.procs.2015.07.541

Badan Pusat Statistik Jawa Tengah. (2022). *Provinsi Jawa Tengah Dalam Angka 2022 Jawa Tengah Province in Figures 2022*. Badan Pusat Statistik Provinsi Jawa Tengah.

Edastama, P., Bist, A. S., & Prambudi, A. (2021). Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm. *International Journal of Cyber and IT Service Management*, *1*(2), 159–172. https://doi.org/10.34306/ijcitsm.v1i2.46

Farida, N., Chulkamdi, M. T., & Wulansari, Z. (2022). Application of Data Mining By Using a Priori Algorithm To Improve Customer Purchasing Decisions At Mikamart Blitar Store. *International Journal of Multidisciplinary Research and Literature*, *1*(5), 526–534. https://doi.org/10.53067/ijomral.v1i5.58

Fitrina, N., Kustanto, K., & Vulandari, R. T. (2018). Penerapan Algoritma Apriori Pada Sistem Rekomendasi Barang Di Minimarket Batox. *Jurnal Teknologi Informasi Dan Komunikasi (TIKomSiN)*, *6*(2), 21–27. https://doi.org/10.30646/tikomsin.v6i2.376

Hakim, L., & Fauzy, A. (2015). Menggunakan Metode Association Rules. *University Research Colloquium*, 73–81.

Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. In *Data Mining: Concepts and Techniques* (3rd ed.). Amsterdam: Elsevier. https://doi.org/10.1016/C2009-0-61819-5

Hendro, G., Adji, T. B., & Setiawan, N. A. (2012). Penggunaan Metodologi Analisa Komponen Utama (PCA) untuk Mereduksi Faktor-Faktor yang Mempengaruhi Penyakit Jantung Koroner. *Seminar Nasional ScrETec*, 1–5.

Mei Alfianto, D. B., Istiyani, N., & Priyono, T. H. (2019). Faktor-Faktor Yang Mempengaruhi Tingkat Kemiskinan di Provinsi Jawa Timur. *E-Journal Ekonomi Bisnis Dan Akuntansi*, *6*(1), 85. https://doi.org/10.19184/ejeba.v6i1.11108

Nurwati, N. (2008). Kemiskinan : Model Pengukuran , Permasalahan dan Alternatif Kebijakan. *Jurnal Kependudukan Padjadjaran*, *10*(1), 1–11.

Permana, R. (2016). Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan di Provinsi Kalimantan Timur. *Journal FEB UNMUL: Forum Ekonomi: Jurnal Ekonomi, Manajemen Dan Akuntansi*, *18*(2), 111–129.

Silva, M. B. (2016). Percepção da população assistida sobre a inserção de estudantes de medicina na Unidade Básica de Saúde. *Trabalho de Conclusão de Curso*, *1*(9), 1–10. https://doi.org/10.1017/CBO9781107415324.004

Sukaesih Sitanggang, I. (2013). Spatial Multidimensional Association Rules Mining in Forest Fire Data. *Journal of Data Analysis and Information Processing*, *01*(04), 90–96. https://doi.org/10.4236/jdaip.2013.14010

Wanto, A. (2018). Penerapan Jaringan Saraf Tiruan Dalam Memprediksi Jumlah Kemiskinan. *Klik - Kumpulan Jurnal Ilmu Komputer*, *5*(1), 61–74.

This sheet is intentionally left blank