Application of data mining techniques using the K-Means Method on Unmet Need of Health Services by Province in Indonesia

W Wahidin¹, M Mugihartadi¹, Tri Hasanah Bimastari Aviani³, Heny Pratiwi², Yusup Indra Wijaya¹, A Andi⁴, Agus Perdana Windarto⁴*, Ari Waluyo⁶

¹Akademi Keperawatan Pemkab Purworejo, Central Java, Indonesia
²STMIK Widya Cipta Dharma, Indonesia
³Universitas Bina Insan, Indonesia
⁴Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin, Indonesia
⁵STIKOM Tunas Bangsa, Pematangsiantar, Indonesia
⁶Politeknik Dharma Patria, Indonesia

Email: * agus.perdana@amiktunasbangsa.ac.id, henypratiwi@wicida.ac.id

Abstract. The research objective is to cluster the Unmet Need of health services in Indonesia by region by utilizing data mining techniques. The data source was obtained from the Central Statistics Agency (abbreviated BPS) in the health category using Unmet Need data on health services by province (2015-2018) consisting of 34 data records. The variable used is the Unmet Need of health services by province (percent). The data mining technique used is k-means which is part of clustering. Data processing is performed using RapidMiner 5.3 software. The cluster labels used in the study are high cluster (C1) and low cluster (C2). The results state that the application of k-means can be carried out where 9 provinces are in cluster C1 consisting of Riau, West Nusa Tenggara, East Nusa Tenggara, Central Kalimantan, South Borneo, Central Sulawesi, Southeast Sulawesi, Gorontalo and West Sulawesi with final centroids (C ) = 6.96 and 25 provinces are in cluster C2 with final centroid (C) = 4.36. The results of mapping in the form of regional clusters can provide information for the government and should Unmet Need health services in Indonesia must be worth 0% for each region.

1. Introduction
The number of Unmet Needs in Indonesia can reach 10.5 percent, which if adjusted with the number of fertile age couples who can reach 36 million, this includes a high number of unmet need. There are about 4 million couples of childbearing age who are in need but still not served due to several factors. Unmet Need figures in Indonesia according to data from the 2012 IDHS, from 1991 to 2012, the situation declined. In 1991 (17%), 1994 (15.3%), 1997 (13.6%), 2002-2003 (13.2%), 2007 (13.1%), and 2012 (11.4%) [1]. Unmet need is a couple of childbearing age (abbreviated as KB) who do not want children, want children with a distance of 2 years or more but do not use contraceptives [2]. This has led to several impacts, namely maternal death and prone to abortion in early childhood [3]. Around 38% of women of childbearing age who do not do family planning (abbreviated as family planning) and cause pregnancy and die due to pregnancy and unsafe birth. There are still many health services that do not dominate contraceptive needs in each region, thereby causing a decrease in the
quality of human resources (abbreviated as HRD), maternal mortality, abortion at an early age, and many negative impacts caused by the lack of unmet need health services.

These problems make researchers want to dig further information on Unmet Need for health services in Indonesia. In this case the researchers used secondary data from the Indonesian Statistics Agency (abbreviated as BPS) with url: https://www.bps.go.id/. There are several computer science techniques [4]–[9] that can be carried out such as the use of data mining techniques [10]–[13]. Data mining is a technique for obtaining information using several well-known methods such as clustering, classification, estimation, association and prediction [14]. In the case of Unmet Need for health services in Indonesia, researchers want to map clusters in Unmet Need for health services in Indonesia so that the results of the mapping can be useful information. There are many clustering techniques that are known for their fast computational ability [11], [15] such as the k-means [16] method. In addition, the k-means method is also more efficient with limited data [17]. Several previous studies using the ability of the k-means method in solving problems [10] in clean water customers. The results of the study state that the k-means method can be applied in classifying clean water customers based on provinces where 70% of Indonesian people are still low aware of the need for clean water. Based on this, it is expected that the research results can provide a mapping of regional clusters so that each region can improve services to the community. The more 0% the Unmet Need the health services in Indonesia the better.

2. Methodology
2.1. Data Mining
Data Mining is a computerized technique for analyzing and extracting knowledge automatically. Data mining has many settlement techniques including clustering which divides data in a set into several groups that have the same data [10].

2.2. Research Method
Sources of research data obtained from secondary data. The data is from the government website (Badan Pusat Statistik) with the health category where the health data taken is the Unmet Need data for health services in Indonesia which consists of 34 data records (2015-2018). The variable used is the province (output in the form of mapping) and the average percentage of Unmet Need for health services (Input). The process of processing research data is assisted by RapidMiner software 5.3. The Klastering method used is K-Means. Here is a research diagram Flowchart Figure:

![Flowchart diagram](image)

3. Results and Discussion
Before carrying out the clustering process using the k-means method, the data used were the results of the average Unmet Need for health services in Indonesia (percent) in 2015-2018. The following is the original data and processed data which are used as a reference in carrying out the clustering process.
Table 1. Original data

| The province   | 2015    | 2016    | 2017    | 2018    |
|----------------|---------|---------|---------|---------|
| Aceh           | 3.09    | 2.86    | 3.62    | 4.53    |
| North Sumatra  | 3.16    | 3.47    | 4.19    | 3.8     |
| West Sumatra   | 3.49    | 3.19    | 4.22    | 3.31    |
| Riau           | 5.15    | 5.36    | 6.25    | 6.11    |
| Jambi          | 4.57    | 4.66    | 4.62    | 4.66    |
| South Sumatra  | 4.33    | 3.6     | 4.65    | 4.05    |
| Bengkulu       | 5.52    | 4.92    | 4.58    | 5.65    |
| Lampung        | 5.49    | 4.03    | 5.64    | 5.27    |
| Bangka Belitung| 3.26    | 3.76    | 3.57    | 3.99    |
| Riau           | 3.12    | 3.1     | 3.88    | 3.46    |
| DKI Jakarta    | 3.4     | 3.08    | 3.56    | 3.2     |
| West Java      | 4.01    | 4.4     | 5.12    | 4.69    |
| Central Java   | 4.71    | 4.44    | 5.26    | 4.84    |
| DI Yogyakarta  | 4.61    | 4.41    | 4.37    | 4.31    |
| East Java      | 5.24    | 4.11    | 5.75    | 4.74    |
| Banten         | 4.45    | 4.22    | 4.74    | 5.98    |
| Bali           | 4.35    | 3.44    | 4.36    | 3.19    |
| West Nusa Tenggara| 6.06   | 6.49    | 6.44    | 7.31    |
| East Nusa Tenggara | 7.0     | 7.06    | 7.66    | 8.09    |
| West Kalimantan| 4.91    | 4.74    | 6.01    | 5.72    |
| Central Kalimantan| 6.05  | 5.62    | 7.11    | 6.76    |
| South Borneo   | 7.8     | 6.35    | 6.41    | 5.51    |
| East Kalimantan| 2.99    | 3.41    | 4.18    | 4.04    |
| North Kalimantan| 3.14   | 3.2     | 7.55    | 4.81    |
| North Sulawesi | 4.55    | 4.58    | 5.44    | 5.46    |
| Central Sulawesi| 7.85   | 7.06    | 9.32    | 8.57    |
| South Sulawesi | 5.15    | 4.53    | 5.51    | 5.49    |
| Southeast Sulawesi| 8.21  | 7.89    | 7.52    | 7.58    |
| Gorontalo      | 8.51    | 6.23    | 8.11    | 7.46    |
| West Sulawesi  | 7.16    | 6.2     | 6.78    | 5.93    |
| Maluku         | 5.37    | 5.05    | 6.28    | 5.57    |
| North Maluku   | 4.68    | 3.67    | 5.05    | 5.5     |
| West Papua     | 3.92    | 4.01    | 5.46    | 4.33    |
| Papua          | 2.78    | 2.63    | 3.28    | 3.01    |

Source: Indonesian Statistics Center

Table 2. Data that has been processed

| The province   | Average Unmet Need for Health Services (Percent) |
|----------------|-----------------------------------------------|
| Aceh           | 3.53                                          |
| North Sumatra  | 3.66                                          |
| West Sumatra   | 3.55                                          |
| Riau           | 5.72                                          |
| Jambi          | 4.63                                          |
| South Sumatra  | 4.16                                          |
| Bengkulu       | 5.17                                          |
| Lampung        | 5.11                                          |
| Bangka Belitung| 3.65                                          |
| Riau           | 3.39                                          |
| DKI Jakarta    | 3.31                                          |
| West Java      | 4.56                                          |
| Central Java   | 4.81                                          |
| DI Yogyakarta  | 4.43                                          |
| East Java      | 4.96                                          |
| Banten         | 4.85                                          |
| Bali           | 3.84                                          |
| West Nusa Tenggara | 6.58  |
| East Nusa Tenggara | 7.20  |
| West Kalimantan| 5.35                                          |
| Central Kalimantan| 6.39  |
| South Borneo   | 6.52                                          |
| East Kalimantan| 3.66                                          |
| North Kalimantan| 4.68                                         |
| North Sulawesi | 5.01                                          |
| Central Sulawesi| 8.20                                         |
| South Sulawesi | 5.17                                          |
| Southeast Sulawesi| 7.80  |
| Gorontalo      | 7.75                                          |
| West Sulawesi  | 6.52                                          |
| Maluku         | 5.57                                          |
| North Maluku   | 4.73                                          |
| West Papua     | 4.43                                          |
| Papua          | 2.93                                          |

Source: Processed Data

The data in table 2 is used to carry out the clustering process by using 2 cluster mapping labels according to needs namely C1: high cluster label for Unmet Need for health services (percent) and C2: low cluster label for Unmet Need for health services (percent). In the process of determining the initial centroid value randomly carried out namely:

Table 3. Initial Centroids

| Attribute | C1: high cluster | C2: low cluster |
|-----------|------------------|-----------------|
| Average Unmet Need for Health Services (Percent) | 8.2   | 2.93 |

Using the initial centroids in table 3, the complete results of grouping in the first iteration can be seen in the following table:

Table 4. K-means calculation results (first iteration)

| The province   | Average Unmet Need for Health Services (Percent) | C1       | C2       | Shortest Distance | First iteration |
|----------------|-----------------------------------------------|----------|----------|-------------------|-----------------|
| Aceh           | 3.53                                          | 4.675    | 0.6      | 0.6               | C2              |
| North Sumatra  | 3.66                                          | 4.545    | 0.73     | 0.73              | C2              |
| West Sumatra   | 3.55                                          | 4.6475   | 0.6275   | 0.6275            | C2              |
| Riau           | 5.72                                          | 2.4825   | 2.7925   | 2.4825            | C1              |
| Jambi          | 4.63                                          | 3.5725   | 1.7025   | 1.7025            | C2              |
| South Sumatra  | 4.16                                          | 4.0425   | 1.2325   | 1.2325            | C2              |
In table 4, the results of the iteration will repeat continuously and stop if the condition of the last iteration has the same cluster as the previous iteration. In this process, iteration stops at stage 3 (third iteration), which means having the same cluster as the second iteration. Here is the final centroid and the final cluster mapping as shown in the following table:

| Province                | Average Unmet Need for Health Services (Percent) | C1     | C2     | Shortest Distance | First iteration |
|-------------------------|--------------------------------------------------|--------|--------|-------------------|-----------------|
| Bengkulu                | 5.17, 3.0325, 2.2425                              |        |        | 2.2425            | C2              |
| Lampung                 | 5.11, 3.0925, 2.1825                              |        |        | 2.1825            | C2              |
| Bangka Belitung         | 3.65, 4.555, 0.72                                |        |        | 0.72              | C2              |
| Riau                    | 3.39, 4.81, 0.465                               |        |        | 0.465             | C2              |
| DKI Jakarta             | 3.31, 4.89, 0.385                                |        |        | 0.385             | C2              |
| West Java               | 4.56, 3.645, 1.63                                |        |        | 1.63              | C2              |
| Central Java            | 4.81, 3.3875, 1.8875                             |        |        | 1.8875            | C2              |
| DI Yogyakarta           | 4.43, 3.775, 1.5                                 |        |        | 1.5               | C2              |
| East Java               | 4.96, 3.324, 2.035                               |        |        | 2.035             | C2              |
| Banten                  | 4.85, 3.3525, 1.9225                             |        |        | 1.9225            | C2              |
| Bali                    | 3.84, 4.365, 0.91                                |        |        | 0.91              | C2              |
| West Nusa Tenggara      | 6.58, 1.625, 3.65                                |        |        | 3.65              | C2              |
| East Nusa Tenggara      | 7.20, 0.9975, 4.2775                             |        |        | 0.975             | C1              |
| West Kalimantan         | 5.35, 2.855, 2.42                                |        |        | 2.42              | C2              |
| Central Kalimantan      | 6.39, 1.815, 3.46                                |        |        | 1.815             | C1              |
| South Borneo            | 6.52, 1.6825, 3.5925                             |        |        | 1.6825            | C1              |
| East Kalimantan         | 3.66, 5.454, 0.73                                |        |        | 0.73              | C2              |
| North Kalimantan        | 4.68, 3.525, 1.75                                |        |        | 1.75              | C2              |
| North Sulawesi          | 5.01, 3.1925, 2.0825                             |        |        | 2.0825            | C2              |
| Central Sulawesi        | 8.20, 0.5, 2.572                                |        |        | 0              | C1              |
| South Sulawesi          | 5.17, 3.03, 2.245                                |        |        | 2.245             | C2              |
| Southeast Sulawesi      | 7.80, 2.3, 4.875                                |        |        | 4.875             | C1              |
| Gorontalo               | 7.75, 0.4475, 4.8275                             |        |        | 0.4475            | C1              |
| West Sulawesi           | 6.52, 1.6825, 3.5925                             |        |        | 1.6825            | C1              |
| Maluku                  | 5.57, 2.6325, 2.6425                             |        |        | 2.6325            | C1              |
| North Maluku            | 4.73, 3.475, 1.8                                 |        |        | 1.8               | C2              |
| West Papua              | 4.43, 3.77, 1.505                               |        |        | 1.505             | C2              |
| Papua                   | 2.93, 5.275, 0                                  |        |        | 0                | C2              |

| Table 5. last centroid | | |
|------------------------|--------------------------|
| Attribute              | C1: high cluster | C2: low cluster |
| Average Unmet Need for Health Services (Percent) | 6.96 | 4.37 |

| The province                | Average Unmet Need for Health Services (Percent) | C1     | C2     | Shortest Distance | Last iteration |
|-----------------------------|--------------------------------------------------|--------|--------|-------------------|----------------|
| Aceh                        | 3.53, 3.4381, 0.7924                              |        |        | 0.792             | C2              |
| North Sumatra               | 3.66, 3.3081, 0.6624                              |        |        | 0.662             | C2              |
| West Sumatra                | 3.55, 3.4106, 0.7649                              |        |        | 0.765             | C2              |
| Riau                        | 5.72, 1.2456, 1.4001                              |        |        | 1.246             | C1              |
| Jambi                       | 4.63, 2.3356, 0.3101                              |        |        | 0.310             | C2              |
| South Sumatra               | 4.16, 2.8056, 0.1599                              |        |        | 0.160             | C2              |
| Bengkulu                    | 5.17, 1.7956, 0.8501                              |        |        | 0.850             | C2              |
| Lampung                     | 5.11, 1.8556, 0.7901                              |        |        | 0.790             | C2              |
| Bangka Belitung             | 3.65, 3.3181, 0.6724                              |        |        | 0.672             | C2              |
| Riau                        | 3.39, 3.5731, 0.9274                              |        |        | 0.927             | C2              |
| DKI Jakarta                 | 3.31, 3.6531, 1.0074                              |        |        | 1.007             | C2              |
| West Java                   | 4.56, 2.4081, 0.2376                              |        |        | 0.238             | C2              |
| Central Java                | 4.81, 2.1506, 0.4951                              |        |        | 0.495             | C2              |
| DI Yogyakarta               | 4.43, 2.5381, 0.1076                              |        |        | 0.108             | C2              |
| East Java                   | 4.96, 2.0031, 0.6426                              |        |        | 0.643             | C2              |
| Banten                      | 4.85, 2.1156, 0.5301                              |        |        | 0.530             | C2              |
| Bali                        | 3.84, 3.1281, 0.4824                              |        |        | 0.482             | C2              |
| West Nusa Tenggara          | 6.58, 0.3881, 2.2576                              |        |        | 0.388             | C1              |
| East Nusa Tenggara          | 7.20, 0.2394, 2.8851                              |        |        | 0.239             | C1              |
| West Kalimantan             | 5.35, 1.6181, 1.0276                              |        |        | 1.028             | C2              |
| Central Kalimantan          | 6.39, 0.5781, 2.0676                              |        |        | 0.578             | C1              |
| South Borneo                | 6.52, 0.4456, 2.2001                              |        |        | 0.446             | C1              |
| East Kalimantan             | 3.66, 3.3081, 0.6624                              |        |        | 0.662             | C2              |
| North Kalimantan            | 4.68, 2.2881, 0.3576                              |        |        | 0.358             | C2              |
| North Sulawesi              | 5.01, 1.9556, 0.6901                              |        |        | 0.690             | C2              |
The province | Average Unmet Need for Health Services (Percent) | C1 | C2 | Shortest Distance | Last Iteration
--- | --- | --- | --- | --- | ---
Central Sulawesi | 8.20 | 1,2369 | 3,8826 | 1.237 | C1
South Sulawesi | 5.17 | 1,7931 | 0,8526 | 0,853 | C2
Southeast Sulawesi | 7.80 | 0,8369 | 3,4826 | 0,837 | C1
Gorontalo | 7.75 | 0,7894 | 3,4351 | 0,789 | C1
West Sulawesi | 6.52 | 0,4456 | 2,2001 | 0,446 | C1
Maluku | 5.57 | 1,3956 | 1,2501 | 1,250 | C2
North Maluku | 4.73 | 2,2381 | 0,4076 | 0,408 | C2
West Papua | 4.43 | 2,5331 | 0,1126 | 0,113 | C2
Papua | 2.93 | 4,0381 | 1,3924 | 1,392 | C2

Figure 2. Graph of Unmet Need cluster results on Health Services in Indonesia

In table 6, the cluster results for C1: high cluster label for Unmet Need for health services (percent) consist of 9 provinces namely: Riau, West Nusa Tenggara, East Nusa Tenggara, Central Kalimantan, South Borneo, Central Sulawesi, Southeast Sulawesi, Gorontalo and West Sulawesi and C2: low cluster label for Unmet Need for health services (percent) consisting of 25 provinces. The results of manual calculations will be tested with RapidMiner 5.3 software. Following are the results of processing with the RapidMiner 5.3 software.

Figure 3. Cluster Results from the RapiMiner software 5.3
Based on the results of the cluster using the RapidMiner 5.3 software in figures 3 and 4, the results are the same as the manual calculations performed. Where the number of regions in cluster C1 (cluster_1) is 9 provinces and cluster C2 (cluster_0) is 25 provinces. Where the final centroid value at C1 (cluster_1) = 6.96 and C2 (cluster_0) = 4.36.

4. Conclusion

Based on the results of the study it can be concluded that the k-means method can be applied to Unmet Need of health services in Indonesia. The final result is in the form of a cluster mapping of provincial regions in Indonesia. The results show that there are 26% of regions in Indonesia that have Unmet Need health services in C1 cluster and 74% in C2 cluster. This information can be input for the government that there are still many areas in Indonesia where Unmet Need for health services is still low (as evidenced by the results of the C1 cluster). Because after all Unmet Need 0% health services (Zero unmet need).

References

[1] A. Huda, “Faktor-faktor yang berhubungan dengan Kejadian Unmeet Need Keluarga Berencana di Puskesmas Bandarharjo Kecamatan Semarang Utara,” Fak. Kesehat. Masy. Univ. Muhammadiyah Semarang, 2016.

[2] P. Kecamatan, G. Kabupaten, D. Nugroho, and Y. Dharmawan, “Hubungan Beberapa Faktor Pasangan Usia Subur (Pus) Dengan Unmet Need Keluarga Berencana (Kb) Di Desa Penungkulan Kecamatan Gebang Kabupaten Purworejo Tahun 2016,” J. Kesehat. Masy., vol. 5, no. 4, pp. 265–273, 2017.

[3] E. Ratnaningsih, “Analisis Dampak Unmet Need Keluarga Berencana Terhadap Kehamilan Tidak Diinginkan Di Rumah Sakit Panti Wilasa Citarum Semarang,” J. Kebidanan, vol. 7, no. 2, pp. 80–94, 2018.

[4] I. G. I. Sudipa, C. Astria, K. F. Irnanda, and A. Perdana, “Application of MCDM using PROMETHEE II Technique in the Case of Social Media Selection for Online Businesses. Application of MCDM using PROMETHEE II Technique in the Case of Social Media Selection for Online Businesses.,” 2020.

[5] N. Nasution et al., “Application of ELECTRE Algorithm in Skincare Product Selection,” J. Phys. Conf. Ser., vol. 1471, no. 1, 2020.

[6] H. Pratiwi et al., “Sigmoid Activation Function in Selecting the Best Model of Artificial Neural Networks,” J. Phys. Conf. Ser., vol. 1471, no. 1, 2020.

[7] W. M. Sari et al., “Improving the Quality of Management with the Concept of Decision Support Systems in Determining Factors for Choosing a Cafe based on Consumers,” J. Phys.
[8] M. B. Rahmat, A. Z. Arfianto, E. Setijadi, A. Mauludiyanto, and V. Y. P. Ardhana, “Measurement and Simulation of Microwave Absorber from Burned Rice Husk,” 2019, pp. 99–102.

[9] O. Oktariyana and O. Oktariyani, “Application of Electronic Module Teaching Materials to Improve Student Learning Outcomes in Rhythmic Gymnastics Subject,” Kinestetik J. Ilm. Pendidik. Jasm., vol. 4, no. 2, pp. 122–127, Sep. 2020.

[10] A. P. Windarto et al., “Analysis of the K-Means Algorithm on Clean Water Customers Based on the Province,” J. Phys. Conf. Ser., vol. 1255, no. 1, 2019.

[11] F. Rahman, I. I. Ridho, M. Muflih, and S. Pratama, “Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination C,” 2020.

[12] W. Katrina, H. J. Damanik, F. Parhusip, D. Hartama, A. P. Windarto, and A. Wanto, “C4.5 Classification Rules Model for Determining Students Level of Understanding of the Subject,” J. Phys. Conf. Ser., vol. 1255, no. 012005, pp. 1–7, 2019.

[13] S. Sriyanto, A. Buchori, A. Handayani, P. T. Nguyen, and H. Usman, “Implementation multi factor evaluation process (MFEP) decision support system for choosing the best elementary school teacher,” Int. J. Control Autom., 2020.

[14] B. Supriyadi, A. P. Windarto, T. Soemartono, and Mungad, “Classification of natural disaster prone areas in Indonesia using K-means,” Int. J. Grid Distrib. Comput., vol. 11, no. 8, pp. 87–98, 2018.

[15] R. H. S. Siburian, R. Karolina, P. T. Nguyen, E. L. Lydia, and K. Shankar, “Leaf disease classification using advanced SVM algorithm,” Int. J. Eng. Adv. Technol., vol. 8, no. 6 Special Issue, pp. 712–718, Aug. 2019.

[16] A. P. Windarto, “Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method,” Int. J. Artif. Intell. Res., vol. 1, no. 2, pp. 26–33, 2017.

[17] P. Alkhairi and A. P. Windarto, “Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara,” Semin. Nas. Teknol. Komput. Sains, pp. 762–767, 2019.