Clustering of Childhood Diarrhoeal Diseases using Gaussian Mixture Model

Ashilla Maula Hudzaifa*[1] and Defi Yusti Faidah[2]

[1, 2] Department of Statistics, Universitas Padjadjaran, INDONESIA.
(E-mail:ashilla21001@mail.unpad.ac.id, defi.yusti@unpad.ac.id)

ABSTRACT

Diarrhoea remains one of the leading causes of childhood morbidity and mortality worldwide, particularly in developing countries [1]. Diarrhoea is a digestive disorder characterized by the need to defecate more often than usual [2]. Diarrhoea remains a major public health problem in Indonesia, with 40% of cases occurring in children under the age of five. Bandung City is one of the largest cities in West Java Province with the sixth highest number of diarrhoea cases in West Java. The high number of diarrhoea cases in Bandung City is a serious concern in public health efforts, especially among young children. The incidence of diarrhoea in infants in Bandung City in 2022 was recorded as high as 6376 cases. Therefore, this study examines the grouping of sub-districts in Bandung City based on factors that affect the incidence of diarrhoea. Five factors are used, namely the prevalence of diarrhoea in children under five years of age, clean and healthy living behaviour (PHBS), healthy latrines, population density, and exclusive breastfeeding in infants under six months of age. Based on the results of clustering with the Gaussian mixture model [3], 5 regional groups were obtained based on similarity of characteristics with five factors.

Cluster 1 are areas characterized by a high percentage of exclusively breastfed infants under 6 months of age. Cluster 2 are areas characterized by a high percentage of households with healthy latrines and PHBS facilities. Cluster 3 are areas characterized by a high population density per hectare. Cluster 4 are areas characterized by a high percentage of households with PHBS and a high population density per hectare. Cluster 5 are areas characterized by a high percentage of under-five children with diarrhoea and a low percentage of PHBS households.

The results of this clustering indicate the importance of interventions and strategies by the Bandung city government to prevent the increasing number of diarrhoea cases among children under five. Therefore, it can provide better insight into the distribution of diarrhea cases to achieve the Sustainable Development Goals (SDGs) and improve the health status of children under five in Indonesia, particularly in Bandung City.

Keywords: *Childhood, Clustering,* Diarrhoea*, Gaussian Mixture Model, SDGs.*

References:

[1] H. U. Ugboko, O. C. Nwinyi, S. U. Oranusi, and J. O. Oyewale, "Childhood diarrhoeal diseases in developing countries," Heliyon, vol. 6, no. 4, 2020.

[2] N. J. CaJacob and M. B. Cohen, "Update on Diarrhea," Pediatrics in Review, vol. 37, no. 8, pp. 313-322, 2016.

[3] M. S. Yang, C. Y. Lai, and C. Y. Lin, "A robust EM clustering algorithm for Gaussian mixture models," Pattern Recognition, vol. 45, no. 11, pp. 3950-3961, 2012.