

Implementation of Algorithm K-Means for Grouping District/City in Central Java Based on Macro Economic Indicators

Authors : Nur Aziza Luxfiati

Abstract : Clustering is partitioning data sets into sub-sets or groups in such a way that elements certain properties have shared property settings with a high level of similarity within one group and a low level of similarity between groups. . The K-Means algorithm is one of the algorithms clustering as a grouping tool that is most widely used in scientific and industrial applications because the basic idea of the algorithm is very simple. In this research, applying the technique of clustering using the k-means algorithm as a method of solving the problem of national development imbalances between regions in Central Java Province based on macroeconomic indicators. The data sample used is secondary data obtained from the Central Java Provincial Statistics Agency regarding macroeconomic indicator data which is part of the publication of the 2019 National Socio-Economic Survey (Susenas) data. score and determine the number of clusters (k) using the elbow method. After the clustering process is carried out, the validation is tested using the methods Between-Class Variation (BCV) and Within-Class Variation (WCV). The results showed that detection outlier using z-score normalization showed no outliers. In addition, the results of the clustering test obtained a ratio value that was not high, namely 0.011%. There are two district/city clusters in Central Java Province which have economic similarities based on the variables used, namely the first cluster with a high economic level consisting of 13 districts/cities and the cluster second with a low economic level consisting of 22 districts/cities. And in the cluster second, namely, between low economies, the authors grouped districts/cities based on similarities to macroeconomic indicators such as 20 districts of Gross Regional Domestic Product, with a Poverty Depth Index of 19 districts, with 5 districts in Human Development, and as many as Open Unemployment Rate. 10 districts.

Keywords : clustering, K-Means algorithm, macroeconomic indicators, inequality, national development

Conference Title : ICAMFE 2022 : International Conference on Applied Mathematical Finance and Economics

Conference Location : Istanbul, Türkiye

Conference Dates : September 27-28, 2022