

Clustering Analysis Using Link-Based Cluster Ensemble Method on Poverty Data in Indonesia = Analisis Clustering dengan Metode Link-Based Cluster Ensemble pada Data Kemiskinan di Indonesia

Maria Angelica Christabel, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920551988&lokasi=lokal>

Abstrak

Poverty is still a serious challenge for all countries in the world, including Indonesia. The Central Statistics Agency (BPS) noted that 25.9 million people were still living in poverty as of March 2023. The challenge of inefficient and poorly targeted budget distribution in some programs hampers poverty alleviation efforts. In line with that, the objective of this study is to categorize provinces based on poverty indicators to provide an understanding of the characteristics and patterns of poverty in various Indonesian provinces. This research uses the Link-based Cluster Ensemble method, which combines various solutions from different clustering methods, namely K-medoids and Agglomerative Hierarchical Clustering to obtain more optimal results. The final result of ensemble clustering is obtained through a similarity matrix by applying the Connected-Triple-based Similarity (CTS) algorithm, which utilizes pairwise similarity relationships between all data points. The data for this study comes from BPS regarding Poverty Data in Indonesia in 2023, with the selection of variables that can reveal the characteristics of poverty in Indonesia, such as aspects of education, employment, expenditure, and household facilities. The results showed that the Link-based Cluster Ensemble method, which was formed through a three-member ensemble from the results of K-Medoids and Agglomerative Hierarchical Clustering, successfully outperformed the single clustering method based on cluster evaluation values using Silhouette, Davies-Bouldin, and Dunn Index. The results grouped the provinces into four clusters: Cluster 1 reveals poverty conditions with low education participation and sanitation facilities, Cluster 2 with low education and high expenditure, Cluster 3 with high unemployment, and Cluster 4 with low clean water facilities.

..... Kemiskinan masih menjadi tantangan serius bagi seluruh negara di dunia, termasuk Indonesia. Badan Pusat Statistik (BPS) mencatat sebanyak 25,9 juta masyarakat masih hidup dalam kemiskinan per Maret 2023. Tantangan dalam penyaluran anggaran yang tidak efisien dan tepat sasaran pada beberapa program menghambat upaya pengentasan kemiskinan. Sejalan dengan hal itu, tujuan dari penelitian ini adalah mengelompokkan provinsi berdasarkan indikator kemiskinan untuk memberikan pemahaman tentang karakteristik dan pola kemiskinan di berbagai provinsi Indonesia. Penelitian ini menggunakan metode Link-based Cluster Ensemble, yang menggabungkan berbagai solusi dari metode pengelompokan berbeda, yaitu K-medoids dan Agglomerative Hierarchical Clustering untuk mendapatkan hasil yang lebih optimal. Hasil akhir pengelompokan ensemble diperoleh melalui similarity matrix dengan menerapkan algoritma Connected-Triple-based Similarity (CTS), yang memanfaatkan hubungan kesamaan berpasangan antara seluruh titik data. Data penelitian ini berasal dari BPS mengenai Data Kemiskinan di Indonesia tahun 2023, dengan pemilihan variabel yang dapat mengungkapkan karakteristik kemiskinan di Indonesia, seperti aspek pendidikan, ketenagakerjaan, pengeluaran konsumsi, dan fasilitas rumah tangga. Hasil penelitian menunjukkan bahwa metode Link-based Cluster Ensemble, yang dibentuk melalui tiga anggota ensemble dari hasil pengelompokan K-Medoids dan Agglomerative Hierarchical Clustering, berhasil mengungguli metode pengelompokan tunggal berdasarkan nilai evaluasi cluster menggunakan Silhouette, Davies-

Bouldin, dan Dunn Index. Hasil penelitian mengelompokkan provinsi ke dalam empat cluster: Cluster 1 mengungkapkan kondisi kemiskinan dengan aspek partisipasi pendidikan dan fasilitas sanitasi yang rendah, Cluster 2 dengan kondisi aspek pendidikan yang rendah dan pengeluaran konsumsi yang tinggi, Cluster 3 dengan kondisi pengangguran yang tinggi, dan Cluster 4 dengan kondisi fasilitas air bersih yang rendah.