

ABSTRAK

Kesejahteraan rakyat pada suatu daerah dapat dilihat dari indikator-indikator yang mampu mengukur kesejahteraan rakyat. Kesejahteraan rakyat pada masing-masing daerah berbeda-beda. Oleh karena itu dapat dilakukan pengelompokan daerah di Indonesia untuk melihat kemiripan kondisi kesejahteraan rakyat di suatu daerah dengan daerah lain sehingga dapat membantu pemerintah dalam menyusun dan menentukan prioritas pembangunan. Data indikator kesejahteraan berupa data campuran. Metode pengelompokan yang dapat digunakan untuk menangani data campuran yaitu metode *cluster ensemble* dengan menggunakan algoritma CEBMDC (*Cluster Ensemble Based Mixed Data Clustering*) dengan melakukan pengelompokan untuk data numerik dan data kategorik secara terpisah, selanjutnya hasil pengelompokan digabungkan. Objek pengamatan pada penelitian ini adalah 34 provinsi di Indonesia. Pengelompokan untuk data numerik menggunakan metode hirarki *agglomerative*, sedangkan pengelompokan data kategorik menggunakan metode ROCK (*Robust Clustering Using Links*). Metode ROCK ini selanjutnya akan digunakan lagi pada tahap penggabungan. Hasil pengelompokan terbaik untuk data numerik dengan nilai CP terkecil diperoleh pada metode *Ward* yang menghasilkan 2 *cluster*. Untuk pengelompokan data kategorik dengan nilai CP* terbesar diperoleh pada $\theta = 0.30$ yang menghasilkan 4 *cluster*. Untuk penggabungan hasil numerik dan kategorik dengan *cluster ensemble* diperoleh pada $\theta = 0.30$ dengan nilai CP* terbesar yang menghasilkan 4 *cluster* yang terdiri dari 1 provinsi pada *cluster* 1, 19 provinsi pada *cluster* 2, 10 provinsi pada *cluster* 3, dan 4 provinsi pada *cluster* 4.

Kata Kunci : Data Campuran, Data Numerik, Data Kategorik, *Cluster Ensemble*, CEBMDC, ROCK, CP, CP*.