

ABSTRAK

Badan Pusat Statistik (BPS) melaksanakan Survei Potensi Desa (Podes) untuk mendukung pembangunan daerah. Namun, survei ini sering menghadapi kendala non-respon, yaitu kondisi responden yang dipilih untuk berpartisipasi dalam survei tidak memberikan respon atau tidak mengisi kuesioner dengan lengkap sehingga menyebabkan data menjadi bias. Untuk mengatasi hal ini, maka dilakukan pengelompokan desa berdasarkan karakteristiknya menggunakan *K-Means*. Hanya saja, inisialisasi *centroid* pada *K-Means* pada dasarnya menggunakan cara acak yang menyebabkan hasilnya kurang optimal. Untuk mengatasi hal tersebut maka digunakan Algoritma *Pillar*. Tujuan penelitian yaitu untuk menguji kinerja kombinasi *K-Means* + Algoritma *Pillar* dalam menentukan *centroid* awal. Penelitian ini menggunakan data Podes dari BPS Banyumas. Evaluasi yang digunakan yaitu *Sum of Square Error (SSE)*, *Variance Between Cluster*, *Silhouette Score*, dan *Variance Within Cluster*. Semakin rendah nilai *SSE* dan *Variance Within Cluster*, maka dapat dinyatakan bahwa klaster semakin homogen dan titik-titik data dalam klaster lebih berdekatan. Sedangkan semakin tinggi nilai *Silhouette Score* dan *Variance Between Cluster*, maka dapat dinyatakan bahwa jarak antar klaster semakin jelas. Hasil evaluasi menunjukkan bahwa *K-Means* lebih unggul dalam menjaga homogenitas dan kedekatan titik dalam klaster, dengan *SSE* lebih rendah dengan selisih 4328194,60 dan *Variance Within Cluster* lebih kecil dengan selisih 208,15. Sebaliknya, *K-Means* + Algoritma *Pillar* lebih baik dalam membentuk klaster yang lebih terpisah, dibuktikan dengan *Variance Between Cluster* lebih tinggi dengan selisih 217,54 dan *Silhouette Score* lebih besar dengan selisih 0,02. Kesimpulannya, *K-Means* lebih optimal untuk klaster yang terpusat dan homogen, sedangkan *K-Means* + Algoritma *Pillar* lebih unggul dalam membentuk klaster dengan batas yang lebih jelas.

Kata Kunci: Algoritma *Pillar*, *K-Means*, *Silhouette Score*, *Sum of Square Error*, *Variance Between Cluster*, *Variance Within Cluster*