ABSTRAK

Firmansyah, Moh Yunus. 2026. Implementasi *Clustering* untuk Identifikasi Faktor Risiko *Stunting* pada Anak Menggunakan Algoritma *K-Means*. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

Pembimbing: (1) Dudi Irawan, ST., M.Kom.; (2) Dewi Lusiana Peter, MT.

Stunting merupakan gangguan pertumbuhan pada balita yang disebabkan oleh kekurangan gizi kronis sehingga tinggi badan anak tidak sesuai dengan usianya. Data pertumbuhan balita di Desa Mayangan, Kecamatan Gumukmas belum dikelompokkan berdasarkan karakteristik pertumbuhan, sehingga diperlukan metode yang dapat membantu proses identifikasi pola data. Penelitian ini bertujuan menerapkan algoritma *K-Means* untuk mengelompokkan data balita berdasarkan umur, berat badan, dan tinggi badan, serta menganalisis karakteristik hasil pengelompokan. Penelitian menggunakan kerangka kerja CRISP-DM dengan 609 data primer balita periode Januari–Desember 2025 yang melalui tahapan *praproses* data dan normalisasi Min-Max. Proses *clustering* dilakukan dengan tiga *cluster* pada setiap bulan, kemudian dievaluasi menggunakan *Silhouette Score, Purity*, dan *Adjusted Rand Index* (ARI) dengan acuan kategori *Z-score* TB/U metode LMS WHO. Hasil penelitian menunjukkan bahwa nilai *Silhouette Score* berada pada kisaran 0,475–0,490, yang menunjukkan kualitas *cluster* cukup. Berdasarkan standar WHO, mayoritas balita berada pada kategori Normal (85,22%–90,80%), diikuti Berisiko (7,55%–12,64%), dan Bermasalah (1,31%–3,12%). Meskipun nilai *Purity* tergolong tinggi, nilai ARI yang mendekati nol menunjukkan bahwa hasil *clustering* lebih menggambarkan kemiripan karakteristik pertumbuhan balita daripada kategori risiko *stunting*. Oleh karena itu, algoritma *K-Means* dapat dimanfaatkan sebagai sistem pendukung untuk mengelompokkan karakteristik pertumbuhan balita, sedangkan penentuan kategori risiko *stunting* tetap mengacu pada perhitungan *Z-score* TB/U berdasarkan standar WHO.

Kata kunci: *Adjusted Rand Index, Clustering, K-Means, Silhouette Score, Stunting, Z-score WHO*

ABSTRACT

Firmansyah, Moh Yunus. 2026. *Implementation of Clustering for Identifying Stunting Risk Factors in Children Using the K-Means Algorithm. Undergraduate Thesis. Bachelor's Program in Informatics Engineering. University of Muhammadiyah Jember.*

Advisors: (1) Dudi Irawan, ST., M.Kom.; (2) Dewi Lusiana Peter, MT.

Stunting is a growth disorder in children under five caused by chronic nutritional deficiencies, resulting in a child's height being below the expected level for their age. Growth data of children under five in Mayangan Village, Gumukmas District, have not yet been grouped based on growth characteristics, creating a need for a method that can assist in identifying patterns within the data. This study aims to apply the K-Means algorithm to cluster child growth data based on age, weight, and height, as well as analyze the characteristics of the resulting clusters. The study employs the CRISP-DM framework using 609 primary data records of children under five collected from January to December 2025, which underwent data preprocessing and Min-Max normalization. The clustering process was performed using three clusters for each month and evaluated using the Silhouette Score, Purity, and Adjusted Rand Index (ARI), with the TB/U Z-score categories calculated using the WHO LMS method as a reference. The results show that the Silhouette Score ranged from 0.475 to 0.490, indicating moderate cluster quality. Based on WHO standards, the majority of children under five were classified as Normal (85.22%–90.80%), followed by At Risk (7.55%–12.64%) and Problematic (1.31%–3.12%). Although the Purity values were relatively high, ARI values close to zero indicate that the clustering results primarily represent similarities in children's growth characteristics rather than stunting risk categories. Therefore, the K-Means algorithm can be utilized as a supporting system for grouping children's growth characteristics, while the determination of stunting risk categories should remain based on TB/U Z-score calculations according to WHO standards.

Keywords: *Adjusted Rand Index*, *Clustering*, *K-Means, Silhouette Score*, *Stunting*, WHO Z-*score*