

## ABSTRAK

Wulandari, Noeralita Vefiliana Audya. 2025. Analisis Pada Tingkat Penyelesaian Pendidikan Di 34 Provinsi Indonesia Menggunakan Metode Algoritma *K-Means Clustering*. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

**Pembimbing:** (1) Guruh Wijaya, S.T, M.Kom.; (2) Habibatul Azizah Al Faruq, M.Pd.

Pendidikan merupakan faktor utama dalam pembangunan suatu negara, tetapi terdapat disparitas tingkat penyelesaian pendidikan di berbagai provinsi di Indonesia. Penelitian ini bertujuan untuk menganalisis tingkat penyelesaian pendidikan berdasarkan jenjang SD, SMP, dan SMA di 34 provinsi Indonesia menggunakan algoritma *K-Means Clustering*. Data yang digunakan bersumber dari Badan Pusat Statistik (BPS) periode 2021–2023. Proses analisis dilakukan menggunakan *RapidMiner* untuk menentukan jumlah *cluster* optimal berdasarkan nilai *Davies-Bouldin Index (DBI)*. Hasil klusterisasi menunjukkan bahwa provinsi dengan tingkat penyelesaian pendidikan tinggi cenderung berada di bagian barat Indonesia, sedangkan provinsi dengan tingkat penyelesaian rendah mayoritas berada di bagian timur. Berdasarkan hasil analisis menggunakan metode *K-Means Clustering*, ditemukan bahwa tingkat penyelesaian pendidikan di berbagai provinsi Indonesia memiliki variasi yang cukup signifikan. Jenjang SMA memiliki tingkat penyelesaian paling rendah dibandingkan SD dan SMP, dengan beberapa provinsi yang menunjukkan nilai paling rendah, seperti Papua Barat (*range* 32,95-39,01), Nusa Tenggara Barat (*range* 38,47-44,88), Sulawesi Tenggara (*range* 45,12-46,19), Nusa Tenggara Timur (*range* 54,27-58,4), Gorontalo (*range* 54,79-56,22). Penelitian ini diharapkan dapat menjadi dasar bagi pengambilan kebijakan pendidikan untuk mengurangi disparitas tingkat pendidikan di Indonesia.

**Kata Kunci:** *Data Mining*, *K-Means Clustering*, Tingkat Penyelesaian Pendidikan, *Cluster*.

## **ABSTRACT**

Wulandari, Noeralita Vefiliana Audya. 2025. *Analysis of Education Completion Rates in 34 Provinces of Indonesia Using the K-Means Clustering Algorithm. Final Project. Undergraduate Program. Informatics Engineering Study Program. University Muhammadiyah of Jember.*

**Advisor:** (1) Guruh Wijaya, S.T, M.Kom.; (2) Habibatul Azizah Al Faruq, M.Pd.

*Education is a key factor in a country's development, yet there are disparities in education completion rates across various provinces in Indonesia. This study aims to analyze education completion rates at the elementary (SD), junior high (SMP), and senior high school (SMA) levels in 34 provinces of Indonesia using the K-Means Clustering algorithm. The data used was obtained from the Indonesian Central Bureau of Statistics (BPS) for the 2021–2023 period. The analysis process was conducted using RapidMiner to determine the optimal number of clusters based on the Davies-Bouldin Index (DBI) value. The clustering results indicate that provinces with high education completion rates are mostly located in western Indonesia, while provinces with lower completion rates are predominantly in the eastern regions. Based on the analysis using the K-Means Clustering method, it was found that education completion rates across various provinces in Indonesia vary significantly. The senior high school (SMA) level has the lowest completion rate compared to elementary (SD) and junior high (SMP) levels, with some provinces showing the lowest values, such as West Papua (range 32.95–39.01), West Nusa Tenggara (range 38.47–44.88), Southeast Sulawesi (range 45.12–46.19), East Nusa Tenggara (range 54.27–58.4), and Gorontalo (range 54.79–56.22). This research is expected to serve as a foundation for education policy-making to reduce disparities in education levels across Indonesia.*

**Keywords:** *Data Mining, K-Means Clustering, Education Completion Rate, Cluster.*