

LEMBAR PERSETUJUAN PEMBIMBING

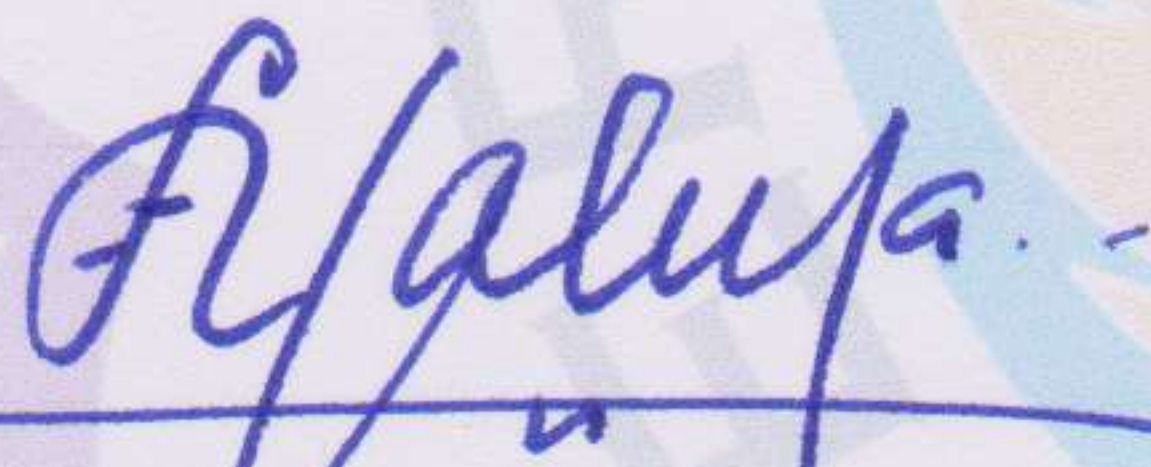
Skripsi yang berjudul "*Penerapan K-Means dan Agglomerative Hierarchical Clustering Untuk Pengelompokan Data Indikator Pendidikan*"
(Studi Kasus Kabupaten/Kota di Wilayah Indonesia Timur)

Oleh

MUZDALIFA DEARIVANY Z. MAMU
NIM. 413416003


Telah diperiksa dan disetujui untuk diuji

Pembimbing I



Dra. Lailany Yahya, M.Si
NIP. 196812191994032001

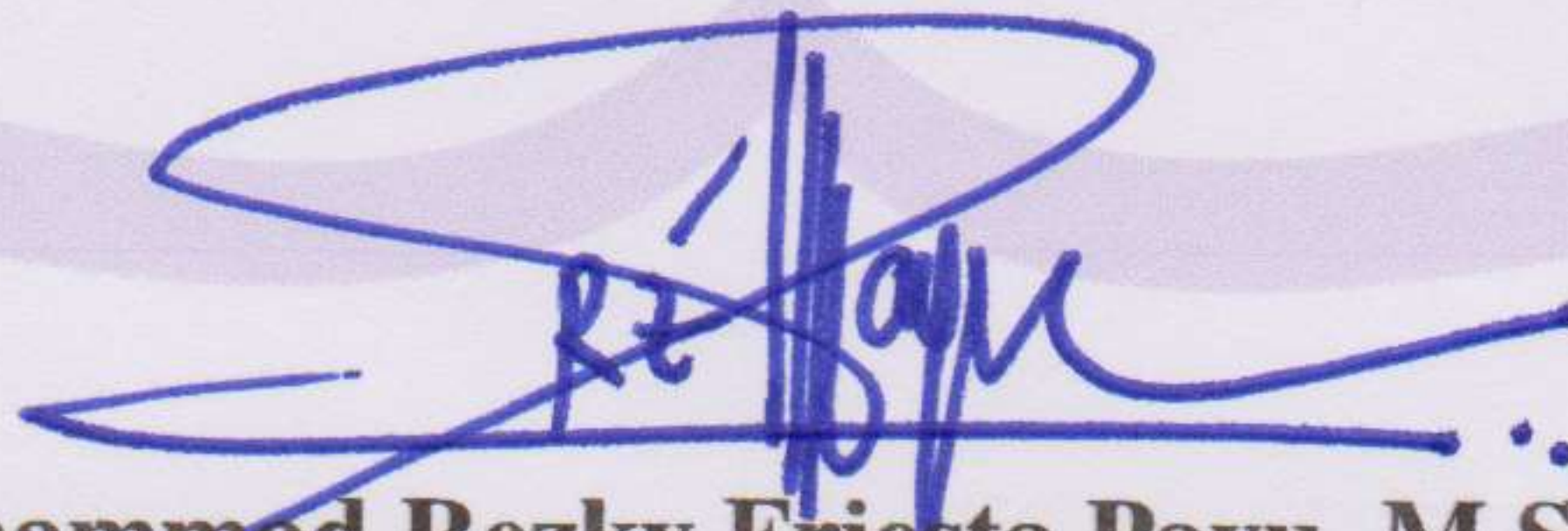
Pembimbing II,



Muhammad Rezky Friesta Payu, M.Si
NIP. 198912152018031003

Mengetahui,

Ketua Program Studi Statistika



Muhammad Rezky Friesta Payu, M.Si.
NIP.198912152018031003

LEMBAR PENGESAHAN

Skripsi yang berjudul "**Penerapan *K-Means* dan *Agglomerative Hierarchical Clustering* Untuk Pengelompokan Data Indikator Pendidikan**"
(Studi Kasus Kabupaten/Kota di Wilayah Indonesia Timur)

Oleh

MUZDALIFA DEARIVANY Z. MAMU
NIM. 413 416 003

Program Studi Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam

Telah dipertahankan di depan dewan penguji

Hari, tanggal : Kamis, 30 Juli 2020

Waktu : 10.00-11.30 WITA

Mekanisme Ujian : DARING

Dewan Penguji

- | | |
|--|---------------|
| 1. Dra. Lailany Yahya, M.Si NIP. 196812191994032001 | Penguji Utama |
| 2. Muhammad Rezky Friesta Payu, M.Si NIP. 198912152018031003 | Anggota |
| 3. Novianita Achmad, S.Si., M.Si NIP. 197411171999032003 | Anggota |
| 4. Fahrezal Zubedi, S.Pd, M.Si NIP. 199406062019031012 | Anggota |
| 5. Sri Lestari Mahmud, S.Pd., M.Si NIP. 198910032019032018 | Anggota |

Tanda Tangan

(.....)
(.....)
(.....)
(.....)
(.....)

Mengetahui,

Dekan Fakultas Matematika dan IPA



Prof. Dr. Astin Lukum, M.Si

NIP.196303271988032002

ABSTRAK

Muzdalifa Dearivany Z. Mamu, 2020. *Penerapan K-Means dan Agglomerative Hierarchical Clustering Untuk Pengelompokan Data Indikator Pendidikan (Studi Kasus Kabupaten/Kota di Wilayah Indonesia Timur).* **Skripsi.** Gorontalo. Program Studi Statistika. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Gorontalo.

Pembimbing : (1) **Dra. Lailany Yahya, M.Si,** (2) **Muhammad Rezky Friesta Payu, M.Si**

K-Means adalah metode populer dalam analisis klaster yang mempartisi himpunan data kedalam beberapa klaster berdasarkan jarak. Namun, hasil pengelompokannya sangat bergantung pada pemilihan nilai centroid awal sehingga solusi yang dihasilkan dapat bersifat local optimum. Dalam penelitian ini dilakukan penentuan nilai centroid awal menggunakan metode *Agglomerative Hierarchical Clustering*. Pengelompokan dilakukan pada kab/kota di 13 provinsi yang ada di wilayah Indonesia timur berdasarkan indikator pendidikan tahun 2019 dengan tujuan untuk mengetahui perbandingan kinerja pengelompokan *K-Means* dengan dan tanpa inisiasi *centroid* dengan metode AHC. Hasil pengelompokan dievaluasi menggunakan nilai *Davies Bouldien Index* dan menunjukkan bahwa pengelompokan *K-Means* dengan *Agglomerative Hierarchical Clustering* memberikan hasil yang lebih baik dibandingkan *K-Means* sederhana. Kab/kota dikelompokkan kedalam 3 klaster untuk pengelompokan berdasarkan variabel Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM) serta 5 klaster untuk pengelompokan berdasarkan variabel Angka Mengulang, Angka Putus Sekolah dan Rasio Siswa Guru.

Kata Kunci: *K-Means, Agglomerative Hierarchical Clustering, Davies Bouldien Index, Pendidikan*

ABSTRACT

Muzdalifa Dearivany Z. Mamu, 2020. *The Application of K-Means and Agglomerative Hierarchical Clustering for Grouping the Indicator Data of Education (A Case Study at Districts/Cities in Eastern Indonesia Region)*. Skripsi. Gorontalo. Study Program of Statistics, Department of Mathematics. Faculty of Mathematics and Natural Sciences. State University of Gorontalo.

The Supervisors: **(1) Dra. Lailany Yahya, M.Si, (2) Muhammad Rezky Friesta Payu, M.Si**

K-means is a popular method in analyzing clusters that partition the data set into several clusters based on distance. However, the clustering results depend on the selection of the initial centroid value so that the generated solution can be local optimum. In this research, the determination of the initial centroid value employs the Agglomerative Hierarchical Clustering method. The clustering or grouping is performed at Districts/ Cities in 13 provinces in Eastern Indonesia Region based on the education indicator in 2019 with the aim to find out the comparison of K-Means clustering performance with and without initial centroid with AHC method. Furthermore, the clustering result is evaluated by using the Davies Bouldin Index and indicates that K-Means clustering using Agglomerative Hierarchical Clustering give better result compared to simple K-Means. In addition, the Districts/Cities are grouped into three clusters for grouping based on the variable of Gross Enrollment Rate (GER) and Net Enrollment Rate (NER), and five clusters for grouping based on the variable repetition rate, dropout rate, and pupil-teacher ratio.

Keywords: *K-Means, Agglomerative Hierarchical Clustering, Davies Bouldien Index, Education*

