

SKRIPSI

**ANALISIS *SILHOUETTE* PADA PERHITUNGAN PENGUKURAN
JARAK ANTARA DATA OBJEK MENGGUNAKAN
METODE *AVERAGE LINKAGE CLUSTERING***

***SILHOUETTE ANALYSIS ON THE CALCULATION OF DISTANCE
MEASURE BETWEEN OBJECT DATA USING THE
AVERAGE LINKAGE CLUSTERING METHOD***



SYUROYA

24010118120044

**DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2022

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS *SILHOUETTE* PADA PERHITUNGAN PENGUKURAN
JARAK ANTARA DATA OBJEK MENGGUNAKAN
METODE *AVERAGE LINKAGE CLUSTERING*

Telah dipersiapkan dan disusun oleh :

SYUROYA

24010118120044

Telah dipertahankan didepan Tim Penguji

Pada tanggal 21 Juli 2022

Susunan Tim Penguji

Pembimbing II/Penguji

Penguji



Dr. Dra. Titi Udjiani S.R.R.M. M.Si.

NIP. 196402231991022001



Drs. Y.D Sumanto. M.Si.

NIP. 19690918993031002

Mengetahui,

Ketua Departemen Matematika

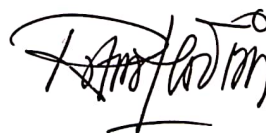
Universitas Diponegoro



Dr. Susilo Harivanto. S.Si. M.Si.

NIP. 197410142000121001

Pembimbing I/Penguji



Ratna Herdiana. M.Sc., Ph.D

NIP. H. 7. 196411242019092001

ABSTRAK

ANALISIS *SILHOUETTE* PADA PERHITUNGAN PENGUKURAN JARAK ANTARA DATA OBJEK MENGGUNAKAN METODE *AVERAGE LINKAGE CLUSTERING*

Oleh

Syuroya

24010118120044

Average Linkage merupakan salah satu metode *clustering* data yang mengelompokkan sekumpulan objek-objek berdasarkan jarak rata-rata antar objek. Perhitungan ukuran jarak antar objek yang berbeda-beda akan menghasilkan hasil *clustering* yang berbeda. Dalam menentukan hasil *clustering* yang baik berdasarkan perhitungan ukuran jarak diperlukan evaluasi dan salah satu analisis evaluasi hasil *clustering* yang populer yaitu analisis *Silhouette*. Pada penelitian ini dilakukan analisis *Silhouette* terhadap perhitungan lima jenis ukuran jarak yaitu *Euclidean distance*, *Manhattan distance*, *Minkowski distance*, *Chebysev distance*, dan *Squared Euclidean distance* menggunakan metode *Average Linkage clustering*. Berdasarkan hasil pengujian analisis *Silhouette* pada *clustering* dataset *Healthy Lifestyle Cities Record 2021* yang berisi 44 kota teratas di seluruh dunia yang mengungkap gaya hidup sehat berdasarkan lima variabel, diperoleh nilai rata-rata *Silhouette Coefficient* dari *Euclidean distance* sebesar 0,36548 (*Weak Structure*), *Manhattan distance* sebesar 0,35783 (*Weak Structure*), *Minkowski distance* sebesar 0,36812 (*Weak Structure*), *Chebysev distance* sebesar 0,28299 (*Weak Structure*) dan *Squared Euclidean distance* sebesar 0,53802 (*Good Structure*). *Squared Euclidean distance* terpilih menjadi ukuran jarak terbaik dengan hasil *clustering* yaitu *cluster* 1 terdiri dari 32 kota dengan tingkat gaya hidup sehat yang tinggi, *cluster* 2 terdiri dari 9 kota dengan tingkat gaya hidup sehat yang sedang, *cluster* 3 terdiri dari 1 kota dengan tingkat gaya hidup sehat sangat tinggi dan *cluster* 4 terdiri dari 2 kota dengan tingkat gaya hidup sehat yang rendah.

Kata Kunci : *Average Linkage*, Ukuran jarak, Analisis *Silhouette*

ABSTRACT

SILHOUETTE ANALYSIS ON THE CALCULATION OF DISTANCE MEASURE BETWEEN OBJECT DATA USING THE AVERAGE LINKAGE CLUSTERING METHOD

by

Syuroya

24010118120044

Average Linkage is a data clustering method that groups a set of objects based on the average distance between objects. The calculation of distance measures between different objects will produce different clustering results. In determining good clustering results based on the calculation of distance measures an evaluation is required and one of the popular clustering result evaluation analyses is silhouette analysis. In this study, Silhouette's analysis was carried out on the calculation of five types of distance measures, namely Euclidean distance, Manhattan distance, Minkowski distance, Chebysev distance, dan Squared Euclidean distance using the Average Linkage clustering method. Based on the results of Silhouette's analysis test on the clustering of the Healthy Lifestyle Cities Record 2021 dataset which contains the top 44 cities worldwide that reveal healthy lifestyles based on five variables, obtained the average Silhouette Coefficient value of Euclidean distance of 0.36548 (Weak Structure), Manhattan distance of 0.35783 (Weak Structure), Minkowski distance of 0.36812 (Weak Structure), Chebysev distance of 0.28299 (Weak Structure) and Squared Euclidean distance of 0.53802 (Good Structure). Squared Euclidean distance was chosen to be the best distance measure with clustering results, namely cluster 1 consisting of 32 cities with a high level of healthy lifestyle, cluster 2 consisting of 9 cities with moderate levels of healthy lifestyle, cluster 3 consisting of 1 city with a very high level of healthy lifestyle and cluster 4 consisting of 2 cities with a low level of healthy lifestyle.

Keywords : *Average Linkage, Distance measure, Silhouette Analysis*