

**OPTIMASI JUMLAH *CLUSTER K-MEANS* PADA PENGELOMPOKAN NILAI
UJIAN NASIONAL TINGKAT SMK DI PROVINSI JAWA TENGAH DENGAN
METODE *ELBOW & SILHOUETTE***

Tugas Akhir disusun untuk memenuhi syarat
mencapai gelar Kesarjanaan Komputer pada
Program Studi Sistem Informasi
Jenjang Program Starta-1



Oleh :

SAVIRA PURNOMO

17.01.55.046

[22944]

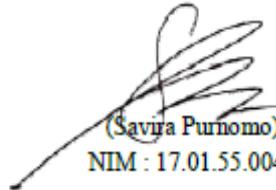
**FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS STIKUBANK (UNISBANK)
SEMARANG**

PERNYATAAN KESIAPAN UJIAN TUGAS AKHIR

Saya, Savira Purnomo dengan ini menyatakan bahw Laporan Tugas Akhir yang berjudul :

Optimasi Jumlah *Cluster K-Means* Pada Pengelompokan Nilai Ujian Nasional Tingkat SMK Provinsi Jawa Tengah Dengan Metode *Elbow & Silhouette*

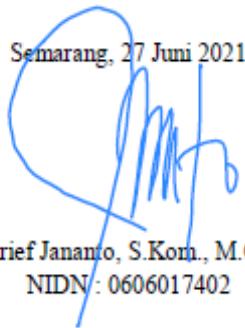
Adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah, sebagian atau seluruhnya, atas nama saya atau pihak lain.



(Savira Purnomo)
NIM : 17.01.55.0046

Disetujui oleh Pembimbing,

Kami setuju laporan tersebut diajukan untuk Ujian Tugas Akhir



Semarang, 27 Juni 2021
(Arief Jananto, S.Kom., M.Cs.)
NIDN : 0606017402

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR/ SKRIPSI

Yang bertanda tangan di bawah ini, saya menyatakan bahwa TUGAS AKHIR / SKRIPSI dengan Judul :

OPTIMASI JUMLAH CLUSTER K-MEANS PADA PENGELOMPOKAN NILAI UJIAN NASIONAL TINGKAT SMK DI PROVINSI JAWA TENGAH DENGAN METODE ELBOW & SILHOUETTE

yang telah diuji di depan tim penguji pada tanggal 14-07-2021, adalah benar hasil karya saya dan dalam TUGAS AKHIR /SKRIPSI ini tidak terdapat keseluruhan atau sebagian tulisan orang lain yang saya ambil dengan cara menyalin, atau meniru dalam bentuk rangkaian kalimat atau simbol yang saya aku seolah-olah sebagai tulisan saya sendiri dan atau tidak terdapat bagian atau keseluruhan tulisan yang saya salin, tiru atau yang saya ambil dari tulisan orang lain tanpa memberikan pengakuan pada penulis aslinya.

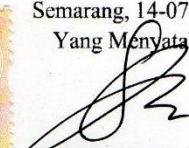
Apabila saya melakukan hal tersebut diatas, baik sengaja maupun tidak, dengan ini saya menyatakan menarik TUGAS AKHIR / SKRIPSI yang saya ajukan sebagai hasil tulisan saya sendiri.

Bila kemudian terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, berarti gelar dan ijazah saya yang telah diberikan oleh Universitas Stikubank (UNISBANK) Semarang batal saya terima.



Semarang, 14-07-2021

Yang Menyatakan


(SAVIRA PURNOMO)

NIM. 17.01.55.0046

SAKSI 1
Tim Penguji



(ARIEF JANANTO, S.Kom., M.Cs.)

SAKSI 3
Tim Penguji



(SUGIYAMTA, S.Kom, M.Kom)

SAKSI 2
Tim Penguji



(KRISTOPHORUS HADIONO, Ph.D)



HALAMAN PENGESAHAN

Tugas Akhir dengan Judul
**OPTIMASI JUMLAH CLUSTER K-MEANS PADA PENGELOMPOKAN NILAI UJIAN
NASIONAL TINGKAT SMK DI PROVINSI JAWA TENGAH DENGAN METODE ELBOW &
SILHOUETTE**

Ditulis oleh

NIM : 17.01.55.0046
Nama : SAVIRA PURNOMO

Telah dipertahankan di depan Tim Dosen Pengaji Tugas Akhir dan diterima sebagai salah satu syarat
guna menyelesaikan Jenjang Program S1 Program Studi pada Fakultas Universitas Stikubank
(UNISBANK) Semarang.

Semarang, 02-08-2021
Ketua



(ARIEF JANANTO, S.Kom., M.Cs.)
NIDN. 0606017402

Sekretaris



(KRISTOPHORUS HADIONO, Ph.D)
NIDN. 0622027601

Anggota



(SUGIYAMTA, S.Kom, M.Kom)
NIDN. 0615076702

Mengetahui,
Universitas Stikubank (UNISBANK) Semarang
Fakultas Teknologi Informasi
Dekan



(KRISTOPHORUS HADIONO, Ph.D)
NIDN. 0622027601



HALAMAN MOTO DAN PERSEMBAHAN

MOTTO :

- Takdir Allah lebih baik dari harapan dan rencanamu.
- Jangan menyerah pada kehidupan karena sesungguhnya ada Tuhan yang selalu menyayangimu.

PERSEMBAHAN :

Skripsi ini saya persembahkan kepada :

1. Tuhan Yang Maha Esa yaitu Mahas Suci Allah dengan segala nikmat, hidayah, rahmat dan petunjuknya sehingga penulis dapat menyelesaikan skripsi ini dengan penuh rasa syukur.
2. Kedua orang tua saya yang selalu memberikan support, semangat, wejangan dan doa yang tidak henti-hentinya dipanjangkan untuk kelancaran skripsi saya. Tanpa mereka mungkin saya sudah tidak tahu harus bagaimana dengan skripsi saya ini.
3. Bp. Arief Jananto, S.Kom.,M.Cs. selaku Ketua Program Studi Sistem Informasi dan dosen pembimbing saya yang selalu memberikan saran dan bimbingannya sehingga saya tidak putus asa dengan skripsi saya.
4. Rismawati sahabat saya yang selalu memberikan semangat ketika saya sudah menyerah dengan skripsi saya. Selalu menyemangati kesehatan mental dan pikiran saya sehingga saya tidak menyerah dengan hidup saya.
5. Panji Suci Sugiyanto yang sudah mau menampung keluh kesah saya, memberikan seluruh waktunya untuk membuat saya tidak menyerah dan putus asa.
6. Teman-teman saya yang membantu memberikan informasi dan ilmu mereka yang tidak pelit.
7. *Last but not least, I wanna thank me,I wanna thank me for believing in ,I wanna thank me for doing all this hard work alone,I wanna thank me for no days off, I wanna thank me for not giving up on me, I wanna thank me because I'm still alive and not doing anything stupid.*

KATA PENGANTAR

Assalamu'alaikum Wr.Wb.

Dengan memanjangkan rasa puja dan puji syukur kepada Allah SWT, laporan tugas akhir dengan judul "**Optimasi Jumlah Cluster K-Means pada Pengelompokan Nilai Ujian Nasional Tingkat SMK Provinsi Jawa Tengah dengan Metode Elbow & Silhouette**" yang disusun dengan penuh perjuangan sebagai syarat untuk menyelesaikan program studi S-1 Sistem Informasi UNISBANK Semarang dapat terselesaikan dengan baik karena adanya dukungan dari berbagai pihak. Oleh karena itu penulis mengucapkan terimakasih kepada :

1. Kedua orang tua saya yang selalu memberikan do'a, dukungan, semangat, dan motivasi dalam penggeraan tugas akhir ini sehingga dapat terselesaikan.
2. Dr. Safik Fauzi, SH, M.Hum, selaku Rektor Universitas Stikubank Semarang.
3. Kristophorus Hadiono, Ph.D, selaku Dekan Fakultas Teknologi Informasi.
4. Arief Jananto, S.Kom, M.Cs. selaku Kepala Program Studi Sistem Informasi.
5. Dosen – dosen pengampu di Program Studi Sistem Informasi yang telah memberikan ilmu dan pengalamannya, sehingga penulis dapat mengimplementasikan ilmu yang disampaikannya.

Penulis memohon maaf apabila terdapat banya kesalahan dan kekurangan dalam penulisan tugas akhir ini. Sehingga penulis sangat terbuka dalam menerima kritik, saran, pendapat dan masukan. Semoga dengan disusunnya laporan tugas akhir ini dapat memberikan manfaat bagi pembaca atau bahkan bisa dilanjutkan demi mendapatkan hasil yang lebih maksimal.

Wassalamu'alaikum Wr.Wb.

Semarang, 14 Juni 2021

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KESIAPAN UJIAN TA.....	ii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
ABSTRAK.....	vi
<i>ABSTRAC</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Tujuan dan Manfaat	2
1.4 Metode Penelitian	4
1.4.1 Obyek Penelitian	4
1.4.2 Metode Pengumpulan Data	4
1.4.3 Metode Analisis Data	5
1.5 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1 Studi Literatur	8
2.2 Landasan Teori	10
BAB III METODE	15
3.1 Metode Penelitian	15
3.1.1 Fase Pemahaman Bisnis (<i>Bussiness Understanding Phase</i>)	15
3.1.2 Fase Pemahaman Data (<i>Data Understanding Phase</i>)	15
3.1.3 Fase Persiapan Data (<i>Data Preparation</i>)	16
3.1.4 Fase Pemodelan (<i>Modeling Phase</i>)	18
3.1.5 Fase Evaluasi	20
3.1.6 Fase Penyebaran	20
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	21
4.1 Perhitungan Manual Algoritma Data Mining	21
4.1.1 Algoritma <i>K-Means</i>	21
4.1.2 Metode <i>Elbow</i>	24
4.1.3 Metode <i>Silhouette</i>	26
4.2 Implementasi Algoritma <i>K-Means</i> dengan RStudio	27
4.2.1 <i>Install</i> dan <i>Load Packages</i> RStudio	27
4.2.2 <i>Data Preparation</i>	28

	4.2.3 Mencari K Optimal <i>Cluster</i>	30
	4.2.4 Eksekusi <i>K-Means</i>	32
BAB V	4.3 Pembahasan Hasil Analisa	39
	4.3.1 Analisa Rekap <i>Cluster</i>	41
	KESIMPULAN DAN SARAN	43
	5.1 Kesimpulan	43
	5.2 Saran	44
	DAFTAR PUSTAKA	46
	LAMPIRAN	47

DAFTAR GAMBAR

Gambar	
1.1 Fase <i>CRISP-DM</i>	5
2.1 Diagram Optimasi <i>Elbow</i>	12
2.2 Diagram Optimasi <i>Silhouette</i>	13
3.1 <i>Flowchart Cluster K-Means</i>	18
4.1 Grafik Optimasi <i>Elbow</i>	24
4.2 Diagram Optimasi <i>Silhouette</i>	26
4.3 Dataset Nilai UN	29
4.4 Data Numerik	29
4.5 Data Scale	30
4.6 Diagram Optimasi <i>Elbow</i>	31
4.7 Diagram Optimasi <i>Silhouette</i>	31
4.8 Hasil dari <i>clustering k-means</i>	32
4.9 Hasil dari <i>clustering k-means</i>	33
4.10 Hasil dari optimasi <i>silhouette</i>	33
4.11 <i>clustering vector</i>	34
4.12 <i>sum of square</i>	34
4.13 Visualisasi optimasi <i>elbow</i>	35
4.14 Visualisasi optimasi <i>silhouette</i>	36
4.15 <i>Pie Chart</i> Dengan 3 Kluster	37
4.16 <i>Pie Chart</i> Dengan 2 Kluster	38
4.17 Rata-rata tiap <i>cluster</i>	39
4.18 <i>Vector Cluster</i> Optimasi <i>Elbow</i>	40
4.19 <i>Vector Cluster</i> Optimasi <i>Silhouette</i>	40
4.20 Rekap kluster metode optimasi <i>elbow</i>	41
4.21 Rekap kluster metode optimasi <i>silhouette</i>	42

DAFTAR TABEL

Tabel	
3.1 Potongan Dataset Nilai Ujian Nasional	15
3.2 Penjelasan Atribut Tabel Nilai Ujian Nasional	16
3.3 Hasil Seleksi Atribut	17
3.4 Penjelasan <i>flowchart clustering k-means</i>	19
4.1 Potongan Dataset Nilai UN	21
4.2 Iterasi 1	23
4.3 Hasil <i>Cluster</i> dengan Metode Optimasi <i>Elbow</i>	25
4.4 Rata-rata hasil <i>cluster</i>	25
4.5 Hasil <i>cluster</i> dengan optimasi <i>silhouette</i>	26
4.6 Rata-rata hasil <i>cluster</i> dengan optimasi <i>silhouette</i>	27
4.7 Hasil Rekap <i>Cluster</i>	39

DAFTAR LAMPIRAN

LISTING PROGRAM RSTUDIO.....	47
<i>Clustering K-Means</i> dengan Metode Optimasi <i>Elbow</i>	47
<i>Clustering K-Means</i> dengan Metode Optimasi <i>Silhouette</i>	50
POTONGAN DATASET	53
LEMBAR BIMBINGAN	54
BERITA ACARA REVISI	55
SURAT KETERANGAN BEBAS PLAGIARISME TURNITIN	56
HASIL TURNTIN BAB 1	57
HASIL TURNITIN ABSTRAK	58
HASIL TURNITIN SKRIPSI	59