

**SKRIPSI**

**KLASIFIKASI TINGKAT KEMISKINAN DI  
PROVINSI SUMATERA UTARA TAHUN 2021  
MENGUNAKAN ANALISIS DISKRIMINAN DAN  
*K-NEAREST NEIGHBOR***



*Unipa Surabaya*

**Khusnul Khotimah  
NIM. 192400024**

**PROGRAM STUDI STATISTIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS PGRI ADI BUANA SURABAYA  
2023**

**KLASIFIKASI TINGKAT KEMISKINAN DI  
PROVINSI SUMATERA UTARA TAHUN 2021  
MENGUNAKAN ANALISIS DISKRIMINAN DAN  
*K-NEAREST NEIGHBOR***

**SKRIPSI**

Diajukan Kepada Fakultas Sains dan Teknologi  
Universitas PGRI Adi Buana Surabaya  
Untuk Memenuhi Sebagian Persyaratan Dalam Memperoleh  
Gelar Sarjana Statistika Pada Program Studi Statistika

**Khusnul Khotimah  
NIM. 192400024**

**PROGRAM STUDI STATISTIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS PGRI ADI BUANA SURABAYA  
2023**

## LEMBAR PERSETUJUAN

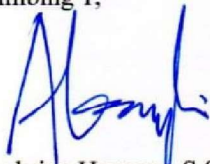
Yang bertanda tangan di bawah ini menyatakan bahwa :

Nama : Khusnul Khotimah  
NIM : 192400024  
Pembimbing 1 : Alfisyahrina Hapsery, S.Si., M.Si  
Pembimbing 2 : Muhammad Athoillah, S.Si., M.Si  
Judul Skripsi : Klasifikasi Tingkat Kemiskinan di Provinsi Sumatera  
Utara Tahun 2021 Menggunakan Analisis  
Diskriminan dan *K-Nearest Neighbor*

Disetujui untuk diseminarkan pada seminar skripsi di Program Studi  
Statistika Universitas PGRI Adi Buana Surabaya.

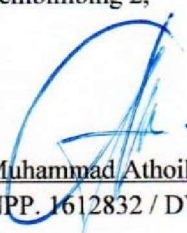
Surabaya, 15 Juni 2023

Pembimbing 1,



Alfisyahrina Hapsery, S.Si., M.Si.  
NPP. 1804856 / DY

Pembimbing 2,



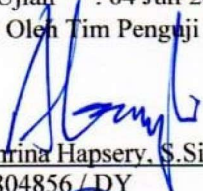
Muhammad Athoillah, S.Si., M.Si.  
NPP. 1612832 / DY

## LEMBAR PENGESAHAN

Skripsi disusun untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Statistika (S.Stat)  
di Universitas PGRI Adi Buana Surabaya  
Oleh:

Nama : Khusnul Khotimah  
NIM : 192400024  
Pembimbing 1 : Alfisyahrina Hapsery, S.Si., M.Si  
Pembimbing 2 : Muhammad Athoillah, S.Si., M.Si  
Judul Skripsi : Klasifikasi Tingkat Kemiskinan di Provinsi Sumatera  
Utara Tahun 2021 Menggunakan Analisis  
Diskriminan dan *K-Nearest Neighbor*


Tanggal Ujian : 04 Juli 2023  
Disetujui Oleh Tim Penguji Skripsi :

  
Alfisyahrina Hapsery, S.Si., M.Si  
NPP. 1804856 / DY

(Pembimbing 1)

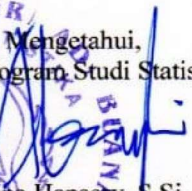
  
Muhammad Athoillah, S.Si., M.Si  
NPP. 1612832 / DY

(Pembimbing 2)

  
Artanti Indrasetianingsih, S.Si., M.Si  
NPP. 0609466 / DY

(Penguji)

Mengetahui,  
Ketua Program Studi Statistika,

  
Alfisyahrina Hapsery, S.Si., M.Si  
NPP. 1804856 / DY

## BERITA ACARA UJIAN SKRIPSI

Skripsi ini diterima dan disetujui oleh panitia ujian skripsi Sarjana Statistika Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.

Pada Hari : Senin  
Tanggal : 4 Juli  
Tahun : 2023

Panitia Ujian Skripsi:

1. Ketua



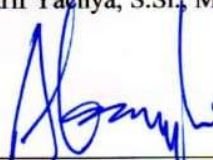
Dr. Dian Kartunia Binawati, M.Si

2. Sekretaris :

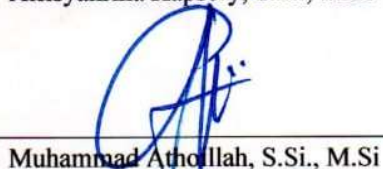


Dr. Arif Yachya, S.Si., M.Si

3. Anggota :



Alfisyahrina Hapsery, S.Si., M.Si



Muhammad Athollah, S.Si., M.Si



Artanti Indrasetianingsih, S.Si., M.Si

## KATA PENGANTAR

Puji dan syukur penulis haturkan kehadirat Allah SWT yang mana atas segala rahmat serta hidayah-Nya penulis dapat menyelesaikan penyusunan skripsi yang berjudul **“Klasifikasi Tingkat Kemiskinan di Provinsi Sumatera Utara Tahun 2021 Menggunakan Analisis Diskriminan dan *K-Nearest Neighbor*”** dengan baik dan tepat waktu. Penyusunan skripsi ini dimaksudkan untuk memenuhi salah satu syarat dalam menyandang gelar Sarjana Statistika pada Program Studi S1 Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.

Ucapan banyak terima kasih penulis sampaikan terhadap semua pihak terkait yang telah memberikan semangat, dukungan serta bantuan baik itu secara langsung maupun tidak langsung. Karena dengan adanya dorongan tersebut penulis dapat menyelesaikan penyusunan skripsi ini dengan baik dan tepat waktu. Penulis ucapkan kepada :

1. Ibu Dra. Diah Karunia Binawati, M.Si. selaku Dekan Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya.
2. Ibu Alfisyahrina Hapsery, S.Si., M.Si. selaku Ketua Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya dan selaku Dosen Pembimbing 1 Skripsi yang telah senantiasa meluang waktu dalam memberikan arahan, bimbingan dan masukan-masukan positifnya sehingga penulis dapat menyelesaikan penyusunan skripsi ini.
3. Bapak Muhammad Athoillah M.Si. selaku Dosen Pembimbing 2 Skripsi yang telah senantiasa meluang waktu dalam memberikan arahan, bimbingan dan masukan-masukan positifnya sehingga penulis dapat menyelesaikan penyusunan skripsi ini.
4. Ibu Artanti Indrasetianingsih, S.Si., M.Si. selaku ibu Dosen Penguji Skripsi yang telah senantiasa memberikan arahan, bimbingan dan masukan-masukan positifnya sehingga penulis dapat menyelesaikan penyusunan skripsi ini.
5. Seluruh Bapak/Ibu Dosen Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya yang telah memberikan banyak ilmu selama penulis belajar di kampus ini.
6. Bapak, Ibu serta Keluarga yang senantiasa memberikan kasih sayang, doa dan dukungan untuk kelangsungan dalam pengerjaan skripsi ini.

7. Sahabat seperjuangan dan seluruh teman-teman Statistika angkatan 2019 yang senantiasa saling memberikan dukungan maupun dorongan dalam menyelesaikan skripsi ini.
8. Semua pihak terkait yang telah membantu dalam penyusunan skripsi ini.

Penulis menyadari bahwa skripsi yang disusun ini masih jauh dari kesempurnaan sehingga kritik dan saran akan sangat bermanfaat dalam menyempurnakan langkah-langkah lebih lanjut menjadi lebih baik lagi. Harapan dari penulis semoga skripsi ini dapat bermanfaat bagi semua pihak khususnya dalam bidang Statistika.

Surabaya, 10 Februari 2023  
Penulis,

Khusnul Khotimah

## **MOTTO**

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya”  
(Q.S Al-Baqarah Ayat 286)

“Maka sesungguhnya bersama kesulitan ada kemudahan”  
(Q.S Al-Insyirah ayat 5)

“Siapa yang menempuh jalan untuk mencari ilmu, maka Allah akan mudahkan baginya jalan menuju surga”  
(HR. Muslim)

## **PERSEMBAHAN**

Skripsi ini dipersembahkan untuk orang-orang yang telah memberikan bantuan, dukungan, do'a serta harapan dalam menyelesaikan skripsi ini:

1. Kedua orang tua yang saya cintai ibu Sopiya dan bapak Jamil
2. Saudara - saudara kandung yang saya sayangi Muhammad Soim, Siti Nur Hidayah, Fitri Lina Wati dan Abdul Muntholib
3. Dosen pembimbing 1 ibu Alfisyahrina Hapsery, S.Si., M.Si
4. Dosen pembimbing 2 bapak Muhammad Athoillah, S.Si., M.Si
5. Dosen penguji ibu Artanti Indrasetyaningih, S.Si., M.Si
6. Teman – teman seperjuangan Suara Muslimah
7. Teman – teman Statistika angkatan 2019



## SURAT PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Khusnul Khotimah

NIM : 192400024

Pembimbing 1 : Alfisyahrina Hapsery, S.Si., M.Si

Pembimbing 2 : Muhammad Athoillah, S.Si., M.Si

Judul Skripsi : Klasifikasi Tingkat Kemiskinan di Provinsi Sumatera Utara Tahun 2021 Menggunakan Analisis Diskriminan dan *K-Nearest Neighbor*

Menyatakan bahwa skripsi ini adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian sumber informasinya dicantumkan.

Pernyataan ini dibuat sebenar – benarnya secara sadar dan bertanggung jawab dan saya bersedia menerima sanksi pembatalan skripsi apabila terbukti melakukan duplikasi terhadap skripsi atau karya ilmiah lain yang sudah ada.

Surabaya, 29 Juli 2023



Khusnul Khotimah

## DAFTAR ISI

	<b>Halaman</b>
HALAMAN JUDUL .....	i
LEMBAR PERSETUJUAN .....	ii
LEMBAR PENGESAHAN .....	iii
BERITA ACARA UJIAN SKRIPSI.....	iv
KATA PENGANTAR.....	v
HALAMAN PERSEMBAHAN .....	vii
SURAT PERNYATAAN .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL .....	xi
DAFTAR GAMBAR.....	xii
DAFTAR SIMBOL .....	xiii
DAFTAR LAMPIRAN.....	xvi
ABSTRAK.....	xvii
<i>ABSTRACT</i> .....	xviii
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah.....	5
BAB 2 TINJAUAN PUSTAKA .....	7
2.1 Klasifikasi .....	7
2.2 Analisis Diskriminan.....	8
2.2.1 Pengujian Asumsi Pada Analisis Diskriminan .....	8
2.2.2 Pembentukan Fungsi Diskriminan.....	10

2.3	<i>K-Nearest Neighbor (K-NN)</i> .....	11
2.4	Data <i>Training</i> dan Data <i>Testing</i> .....	12
2.5	Pengukuran Ketepatan Klasifikasi .....	13
2.6	Kemiskinan .....	15
2.7	Penelitian Terdahulu .....	18
BAB 3 METODOLOGI PENELITIAN .....		21
3.1	Sumber Data.....	21
3.2	Variabel Penelitian .....	21
3.3	Struktur Data .....	22
3.4	Langkah Analisis.....	22
3.5	Diagram Alir Penelitian .....	24
BAB 4 ANALISIS DAN PEMBAHASAN .....		27
4.1	Analisis Deskriptif Variabel Dependen .....	27
4.2	Analisis Deskriptif Variabel Independen.....	28
4.3	Klasifikasi Tingkat Kemiskinan Menggunakan Analisis Diskriminan.....	31
4.3.1	Pengujian Asumsi Analisis Diskriminan .....	32
4.3.2	Pembentukan Fungsi Diskriminan.....	34
4.3.3	Pengukuran Ketepatan Klasifikasi.....	36
4.4	Klasifikasi Tingkat Kemiskinan Menggunakan <i>K-Nearest Neighbor (K-NN)</i> .....	38
4.5	Perbandingan Klasifikasi Tingkat Kemiskinan Menggunakan Analisis Diskriminan dan K-NN .....	43
BAB 5 KESIMPULAN DAN SARAN .....		45
5.1	Kesimpulan .....	45
5.2	Saran.....	46
DAFTAR PUSTAKA .....		47
LAMPIRAN .....		51

## DAFTAR TABEL

	<b>Halaman</b>
<b>Tabel 2. 1</b> Matriks Konfusi Pada Klasifikasi Biner .....	8
<b>Tabel 2. 2</b> Kategori Standarisasi <i>Classifier</i> Berdasarkan Nilai AUC.....	15
<b>Tabel 2. 3</b> Kategori Indeks Kearifan Kemiskinan .....	17
<b>Tabel 3. 1</b> Variabel Penelitian.....	21
<b>Tabel 3. 2</b> Struktur Data Kemiskinan di Provinsi Sumatera Utara.....	22
<b>Tabel 4. 1</b> Analisis Deskriptif Variabel Independen .....	28
<b>Tabel 4. 2</b> Dataset Kemiskinan Sumatera Utara.....	31
<b>Tabel 4. 3</b> Uji <i>Pearson Correlation</i> Jarak Mahalanobis dengan <i>Chi-Square Quantile</i> .....	32
<b>Tabel 4. 4</b> Hasil Pengujian Multikolinearitas.....	33
<b>Tabel 4. 5</b> Hasil Pengujian Homoskedastisitas dengan Box'M.....	33
<b>Tabel 4. 6</b> Hasil Pengujian <i>Wilk's Lambda</i> .....	34
<b>Tabel 4. 7</b> Perhitungan <i>Z Score</i> pada Data <i>Testing</i> .....	35
<b>Tabel 4. 8</b> <i>Confusion Matrix</i> dengan Analisis Diskriminan .....	37
<b>Tabel 4. 9</b> Urutan Jarak <i>Euclidean</i> Antara Data <i>Training</i> dan Data <i>Testing</i> Pertama.....	40
<b>Tabel 4. 10</b> Nilai Akurasi Berdasarkan Nilai k pada K-NN.....	41
<b>Tabel 4. 11</b> <i>Confusion Matrix</i> dengan k = 3.....	41
<b>Tabel 4. 12</b> Hasil Klasifikasi K-NN dengan k = 3 .....	42
<b>Tabel 4. 13</b> Performa Klasifikasi Analisis Diskriminan dan K-NN .....	43

## DAFTAR GAMBAR

	<b>Halaman</b>
<b>Gambar 1. 1</b> Persentase Penduduk Miskin Provinsi Sumatera Utara Tahun 2010-2021 .....	1
<b>Gambar 2. 1</b> Klasifikasi Biner dan Multi Kelas Klasifikasi .....	7
<b>Gambar 2. 2</b> Ilustrasi Pembagian Data <i>Training</i> dan Data <i>Testing</i> .....	13
<b>Gambar 3. 1</b> Diagram Alir Secara Umum .....	24
<b>Gambar 3. 2</b> Diagram Alir Analisis Diskriminan.....	25
<b>Gambar 3. 3</b> Diagram Alir <i>K-Nearest Neighbor</i> .....	26
<b>Gambar 4. 1</b> Peta Tingkat Kemiskinan di Sumatera Utara Berdasarkan IKK.....	27
<b>Gambar 4. 2</b> Hasil Pengujian Normal Multivariat.....	32
<b>Gambar 4. 3</b> Nilai AUC Analisis Diskriminan dan K-NN .....	43

## DAFTAR SIMBOL

- $f_{ij}$  : jumlah data di kelas ke-i pada kelas ke-j
- $\mu$  atau  $\bar{x}$  : *mean* atau rata-rata
- $\Sigma_i$  atau  $S_i$  : matriks varians kovarians
- $d_{ij}$  : jarak mahalanobis
- $x_i$  : vektor dimensi dari nilai pengukuran ke-i
- $x_j$  : vektor dimensi dari nilai pengukuran ke-j
- $S^{-1}$  : matriks invers dari S
- $r_{ij}$  : *pearson correlation*
- $VIF$  : *Variance Inflation Factor*
- $R_j$  : koefisien korelasi variabel independen ke-j
- $c_1$  : pendekatan  $\chi^2$
- $v_i$  :  $n_i - 1$  (asumsi sampel independen ke-i)
- $S_{pl}$  : matriks varians kovarians gabungan
- $\Lambda$  : *wilk's lambda*
- $E$  : matriks jumlah kuadrat dan hasil kali data dalam suatu kelompok
- $H$  : matriks jumlah kuadrat dan hasil kali data dalam antarkelompok

$\hat{y}$	: fungsi diskriminan
$a_i$	: koefisien diskriminan dari variabel atribut ke- $i$
$\hat{m}$	: <i>cutting score</i>
$\bar{y}_1$	: rata-rata $y$ yang diperoleh dari rata-rata $x$ yang termasuk dalam grup 1
$\bar{y}_2$	: rata-rata $y$ yang diperoleh dari rata-rata $x$ yang termasuk dalam grup 2
$dist(X_1, X_2)$	: jarak <i>euclidean</i> (antara objek data $x_{1i}$ dan $x_{2i}$ )
$x_{1i}$	: data <i>training</i>
$x_{2i}$	: data <i>testing</i>
$n$	: jumlah fitur
$k$	: nilai parameter dari banyaknya tetangga
$TN$	: ( <i>true negative</i> ) jumlah data pada kelas negatif dengan klasifikasi negatif
$TP$	: ( <i>true positive</i> ) jumlah data pada kelas positif dengan klasifikasi positif
$FN$	: ( <i>false negative</i> ) jumlah data pada kelas positif diklasifikasikan negatif
$FP$	: ( <i>false positive</i> ) jumlah data pada kelas negatif diklasifikasikan positif
$AUC$	: <i>Under The Curve</i>

*IKK* : Indeks Keparahan Kemiskinan

*X* : variabel independen

*Y* : variabel dependen

*1* : kategori tinggi

*0* : kategori rendah



## DAFTAR LAMPIRAN

	<b>Halaman</b>
<b>Lampiran A.</b> Data Penelitian Tingkat Kemiskinan di Provinsi Sumatera Utara Tahun 2021 .....	51
<b>Lampiran B.</b> Kategori Tingkat Kemiskinan Berdasarkan IKK .....	52
<b>Lampiran C.</b> Kabupaten/Kota Berdasarkan Tingkat Kemiskinan.....	53
<b>Lampiran D.</b> Pembagian Data <i>Training</i> dan Data <i>Testing</i> .....	53
<b>Lampiran E.</b> Urutan Jarak <i>Euclidean</i> Antara Data <i>Training</i> dan Data <i>Testing</i> .....	55
<b>Lampiran F.</b> <i>Syntax</i> Impor Data dan Pembagian Data Melalui R-Studio .....	64
<b>Lampiran G.</b> <i>Syntax</i> Analisis Diskriminan .....	64
<b>Lampiran H.</b> <i>Syntax K-Nearest Neighbor</i> .....	65
<b>Lampiran I.</b> Output Hasil Klasifikasi Analisis Diskriminan .....	66
<b>Lampiran J.</b> Output Hasil Klasifikasi <i>K-Nearest Neighbor</i> .....	69