



**CLUSTERING KETERAMPILAN TEKNOLOGI INFORMASI DAN KOMPUTER  
UNTUK MENGURANGI KESENJANGAN PENGUASAAN TIK MENGGUNAKAN  
METODE K-MEANS PADA SELURUH PROVINSI DI INDONESIA**

**SKRIPSI**

**BARIROTUN NAJAH**

**NIM. 1910511119**

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**2023**



**CLUSTERING KETERAMPILAN TEKNOLOGI INFORMASI DAN KOMPUTER  
UNTUK MENGURANGI KESENJANGAN PENGUASAAN TIK MENGGUNAKAN  
METODE K-MEANS PADA SELURUH PROVINSI DI INDONESIA**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer**

**BARIROTUN NAJAH**

**NIM. 1910511119**

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**2023**

## PERNYATAAN ORISINALITAS

Tugas akhir ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Barirotun Najah

NIM : 1910511119

Tanggal : 3 Juli 2023

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya, maka saya bersedia di tuntutan dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 3 Juli 2023

Yang menyatakan,

  
Barirotun Najah

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK  
KEPENTINGAN AKADEMIS**

Sebagai civitas akademis Universitas Pembangunan Nasional Veteran Jakarta,  
saya yang bertanda tangan di bawah ini:

Nama : Barirotun Najah

NIM : 1910511119

Fakultas : Ilmu Komputer

Program Studi : S-1 Informatika

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan pelayanan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas karya ilmiah saya yang berjudul :

**CLUSTERING KETERAMPILAN TEKNOLOGI INFORMASI DAN  
KOMPUTER UNTUK MENGURANGI KESENJANGAN PENGUASAAN  
TIK MENGGUNAKAN METODE K-MEANS PADA SELURUH  
PROVINSI DI INDONESIA**

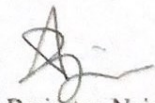
Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya selama mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat : Jakarta

Pada tanggal : 3 Juli 2023

Yang menyatakan,



Barirotun Najah

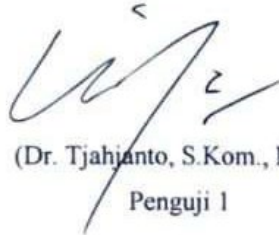


## LEMBAR PENGESAHAN

Yang bertanda tangan di bawah ini:

Nama : Barirotun Najah  
NIM : 1910511119  
Program Studi : S-1 Informatika  
Judul Skripsi/TA : Clustering Keterampilan Teknologi Informasi Dan Komputer Untuk Mengurangi Kesenjangan Penguasaan TIK Menggunakan Metode K-Means Pada Seluruh Provinsi Di Indonesia

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



(Dr. Tjahjanto, S.Kom., M.M.)  
Penguji 1



(Anita Muliawati, S.Kom., M.TI.)  
Penguji 2



(Dr. Widya Cholil, M.I.T.)  
Pembimbing



(Dr. Ermatita, M.Kom.)  
Dekan



(Dr. Widya Cholil, M.I.T.)  
Ketua Program Studi

Ditetapkan di : Jakarta  
Tanggal Ujian : 3 Juli 2023



## Abstrak

Penelitian ini dilakukan untuk mengetahui pengelompokan keterampilan teknologi informasi dan komputer (TIK) berdasarkan provinsi di Indonesia menggunakan metode k-means yang nantinya dapat membantu pemerintah untuk mengetahui pengelompokan tingkat keterampilan TIK menurut provinsi. Teknologi Informasi dan Komputer (TIK) merupakan seluruh kegiatan yang berhubungan dengan pemrosesan, pengelolaan dan penyampaian atau pemindahan informasi antar sarana atau media. Penelitian ini merupakan penerapan algoritma kmeans dan evaluasi teknik *Davies Bouldin Index* (DBI) serta *Calinski Harabasz Index*. Penelitian menggunakan metode elbow untuk mengetahui jumlah cluster yang terbaik. Penelitian ini memperoleh cluster terbaik dengan k sebesar 3 yaitu, cluster 0 dengan kategori cukup terampil sebanyak 22 provinsi, cluster 1 dengan kategori sangat terampil sebanyak 8 provinsi, dan cluster 2 dengan kategori kurang terampil sebanyak 4 provinsi.

**Kata Kunci :** Teknologi Informasi dan komputer (TIK), K-Means, Cluster, *Davies Bouldin Index* (DBI), *Calinski Harabasz Index*.

## ***Abstract***

*This research was conducted to determine the grouping of information technology and computer (ICT) skills by province in Indonesia using the k-means method which can later help the government to determine the grouping of ICT skill levels by province. Information Technology and Computers (ICT) are all activities related to processing, managing and delivering or transferring information between facilities or media. This research is the application of the Kmeans algorithm and the evaluation of the Davies Bouldin Index (DBI) technique and the Calinski Harabasz Index. This research uses the elbow method to determine the best number of clusters. This study obtained the best cluster with k of 3, namely, cluster 0 with the moderately skilled category in 22 provinces, cluster 1 in the highly skilled category in 8 provinces, and cluster 2 in the less skilled category in 4 provinces.*

**Keywords :** *Information and Computer Technology (ICT), K-Means, Cluster, Davies Bouldin Index (DBI), Calinski Harabasz Index.*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-nya, sehingga tugas akhir ini dapat berhasil diselesaikan untuk memenuhi syarat untuk memperoleh gelar sarjana. Pada penelitian untuk tugas akhir ini banyak pihak yang membantu penulis dalam menyelesaikan tugas akhir. Untuk itu penulis ingin mengucapkan terima kasih kepada:

1. Orang tua dan keluarga saya yang selalu memberikan dukungan baik dukungan dari sisi material maupun dari sisi moral sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Ibu Dr. Widya Cholil, S.Kom., M.I.T. selaku Ketua Program Studi Informatika dan dosen pembimbing yang sudah membimbing penulis dalam melakukan penyusunan tugas akhir.
3. Ibu Dr. Ermatita, M.Kom. selaku Dekan Fakultas Ilmu Komputer.
4. Ibu Helena Nurramdhani Irmanda, S.Pd., M.Kom. selaku Dosen Pembimbing Akademik.
5. Bapak, Ibu Dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta yang sudah mengajarkan saya berbagai ilmu yang bermanfaat semasa perkuliahan.
6. Teman-teman satu server Discord kelas-c informatika yang selalu memberikan dukungan dan semangat dalam perkuliahan.
7. Serta seluruh pihak yang belum disebutkan diatas yang membantu dan mendukung penulis dalam mengerjakan tugas akhir.

Semoga tugas akhir ini dapat bermanfaat dan dapat dijadikan referensi untuk penelitian selanjutnya.

Jakarta, 6 Juni 2023

Penulis



Barirotun Najah



## DAFTAR ISI

PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS .....	iv
LEMBAR PENGESAHAN .....	v
Abstrak.....	vi
<i>Abstract</i> .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
BAB 1 PENDAHULUAN .....	13
1.1 Latar Belakang .....	13
1.2 Rumusan Masalah .....	15
1.3 Tujuan Penelitian.....	16
1.4 Manfaat Penelitian.....	16
1.5 Ruang Lingkup.....	16
1.6 Luaran Yang Diharapkan .....	17
1.7 Sistematika Penulisan.....	17
BAB 2 TINJAUAN PUSTAKA .....	19
2.1 Data Mining.....	19
2.2 Algoritma K-Means.....	20
2.3 Clustering .....	21
2.4 Keterampilan .....	21
2.5 Teknologi Informasi dan Komunikasi (TIK) .....	21
2.6 Data Warehouse .....	22
2.7 Metode Elbow .....	22
2.8 Evaluasi .....	23
2.8.1 Davies Bouldin Index (DBI).....	23
2.8.2 Calinski Harabasz Index.....	25
2.9 Hipotesis Penelitian.....	25

2.10 Studi Literatur .....	26
<b>BAB 3 METODOLOGI PENELITIAN .....</b>	<b>30</b>
3.1 Kerangka Berpikir .....	30
3.1.1 Identifikasi Masalah .....	31
3.1.2 Studi Literatur .....	31
3.1.3 Pengumpulan Data .....	31
3.1.4 Praproses Data .....	31
3.1.5 Metode Elbow .....	32
3.1.6 Metode K-Means .....	32
3.1.7 Arsitektur Datawarehouse .....	32
3.1.8 Evaluasi .....	33
3.2 Sarana Pendukung .....	33
3.3 Waktu Kegiatan .....	33
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>	<b>35</b>
4.1 Pengumpulan data .....	35
4.2 Pra-proses Data .....	36
4.3 Metode Elbow .....	37
4.4 Metode Kmeans .....	39
4.5 evaluasi .....	44
<b>BAB 5 SIMPULAN DAN SARAN .....</b>	<b>47</b>
5.1 Simpulan .....	47
5.2 Saran .....	48
<b>DAFTAR PUSTAKA .....</b>	<b>49</b>
<b>LAMPIRAN .....</b>	<b>52</b>

## DAFTAR TABEL

Tabel 3.1. Waktu Kegiatan .....	34
Tabel 4.2. Sampel Data Setelah Pra-proses .....	36
Tabel 4.3. Hasil Selisih SSE .....	38
Tabel 4.4. Centroid Awal.....	39
Tabel 4.5. Perhitungan K-Means dengan $k = 3$ .....	40
Tabel 4.6. Hasil Centroid Baru .....	42
Tabel 4.7. Hasil Cluster Akhir .....	42
Tabel 4.8. Nilai Davies Bouldin Index .....	44
Tabel 4.9. Nilai Calinski Harabasz Index .....	45

## DAFTAR GAMBAR

Gambar 2.1. Metode KDD (Knowledge Discovery in Database) .....	19
Gambar 2.2. Arsitektur Data Warehouse .....	22
Gambar 3.1. Kerangka Berfikir .....	30
Gambar 4.1. Grafik Metode Elbow.....	38
Gambar 4.2. Visualisasi Persebaran Cluster Kmeans .....	43
Gambar 4.3. Hasil Evaluasi Menggunakan Davies Bouldin Index .....	44
Gambar 4.4. Hasil Evaluasi Menggunakan Calinski Harabasz Score .....	45