

**ANALISIS *CLUSTERING* DATA PENYANDANG
DISABILITAS MENGGUNAKAN METODE
AGGLOMERATIVE *HIERARCHICAL*
CLUSTERING DAN *K-MEANS***

SKRIPSI

SILVIA NOER

20190040104



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK KOMPUTER DAN DESAIN**

JULI

2023

**ANALISIS *CLUSTERING* DATA PENYANDANG
DISABILITAS MENGGUNAKAN METODE
AGGLOMERATIVE *HIERARICAL*
CLUSTERING DAN *K-MEANS***

SKRIPSI

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh
Gelara Sarjana kompuer*

SILVIA NOER

20190040104



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK KOMPUTER DAN DESAIN**

JULI

2023

PERNYATAAN PENULIS

JUDUL : ANALISIS *CLUSTERING* DATA PENYANDANG DISABILITAS
MENGUNAKAN METODE AGGLOMERATIVE
HIERARCHICAL CLUSTERING DAN *K-MEANS*

NAMA : SILVIA NOER

NIM 20190040104

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti- bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.



Materai

SILVIA NOER

Penulis

PERSETUJUAN SKRIPSI

JUDUL : ANALISIS *CLUSTERING* DATA PENYANDANG
DISABILITAS MENGGUNAKAN METODE
AGGLOMERATIVE *HIERARCHICAL CLUSTERING* DAN *K-*
MEANS
NAMA : SILVIA NOER
NIM 20190040104

Skripsi Ini telah diperiksa dan disetujui
Sukabumi, Juli 2023

Kaprodi Teknik Informatika,

Pembimbing,

Anggun Fergina, M.Kom
NIDN. 0407029301

Gina Purnama Insany, S.ST., M.Kom
NIDN. 0417077908



PENGESAHAN SKRIPSI

JUDUL : ANALISIS *CLUSTERING* DATA PENYANDANG DISABILITAS
MENGUNAKAN METODE AGGLOMERATIVE *HIERARCHICAL*
CLUSTERING DAN *K-MEANS*

NAMA : SILVIA NOER

NIM 20190040104

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 10 juli 2023 Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

Sukabumi, ... 2023

Pembimbing I

Gina Purnama Insany, S.ST., M.Kom
NIDN. 0417077908



Pembimbing II

Alun Sujjada, S.Kom., M.T
NIDN. 0718108001

Ketua Penguji

Kepala Program Studi

Kamdan, M.Kom
NIDN. 0401107401

Anggun Fergina, M.Kom
NIDN. 0407029301

Dekan Fakultas Teknik Komputer dan Desain

Ir. Paikun, S.T., M.T., IPM Asean Eng.
NIDN. 0402037401

HALAMAN PERSEMBAHAN

Rasa syukur kepada Allah SWT. karena telah memberikan saya kekuatan dan makna kesabaran. Atas karunia-Mu. Akhirnya, skripsi sederhana ini dapat terselesaikan. Shalawat dan salam selalu terlimpah curahkan kepada Nabi Muhammad SAW. Untuk mereka yang saya sayangi, saya persembahkan karya sederhana ini. Ayah dan Mamah, karya sederhana ini saya persembahkan untuk Mamah saya (Namiroh) dan ayah (Deni Noerman) sebagai tanda terima kasih dan rasa hormat saya atas cinta dan dukungan mereka yang tiada henti, yang tidak akan pernah bisa saya balas hanya dengan selembar kertas bertuliskan kata-kata. "persembahan." Karena sampai saat ini saya belum bisa berbuat lebih banyak, semoga ini menjadi langkah awal untuk membuat Mamah dan Ayah bangga. untuk orang tua saya. untuk Mamah dan ayah. Terima kasih telah memberikan dukungan dan kepercayaan dalam menyelesaikan Tugas Akhir ini. Saya berharap doa kalian dan semua hal baik yang kalian berikan akan membantu saya menjadi orang yang baik juga.



ABSTRAK

Persoalan disabilitas masih menjadi isu yang menarik persepsi publik terkait dengan diskriminasi yang mereka temui sehari-hari. Masyarakat maupun lingkungan masih menstigmatisasi atau memperlakukan penyandang disabilitas secara tidak adil. Padahal banyak penyandang disabilitas memiliki kemampuan yang setara dengan orang-orang yang tidak memiliki keterbatasan fisik. Melalui studi kasus ini, peneliti berharap dapat membantu memberikan kontribusi dalam meningkatkan kesadaran agar hak-hak penyandang disabilitas dapat terpenuhi, khususnya terkait dengan akses pekerjaan dan pendidikan yang layak dan menggali informasi baru terkait penyandang disabilitas. dengan mengolah dataset penyandang disabilitas di Indonesia dengan mengambil sampel data di 7 Provinsi di Indonesia berdasarkan 137 kota dan kabupaten dengan 3 jenis variabel disabilitas yaitu cacat fisik, cacat netra atau buta, Cacat runtu dan wicara. Menggunakan analisis *clustering* algoritma *K-means* dan agglomerative hierarchical *clustering*. untuk mengidentifikasi *outlier* dan anomali peneliti menggunakan metode EDA (*Exploratory Data Analysis*) agar dapat membantu dalam memperkaya hasil analisis data. saat menentukan *centroid* peneliti menggunakan metode hirarki agglomerative yaitu *Average Linkage* berdasarkan Uji validasi nilai korelasi *cophenetic* yang tertinggi hal ini menghasilkan hasil yang sama pada penelitian sebelumnya, Jumlah *cluster* yang ditentukan dengan metode *K-means* sebanyak 3 *cluster* dengan hasil *cluster* 1 merupakan tingkat tinggi dengan jumlah 62 kota dan kabupaten, *cluster* 2 merupakan *cluster* dengan kategori tingkat sedang jumlah 37 kota dan kabupaten, *cluster* 3 merupakan *cluster* dengan kategori tingkat rendah berjumlah 27 kota dan kabupaten. hasil evaluasi terbaik menggunakan metode DBI *Davies bouldin-Index* dengan menghasilkan 2 *cluster* karena jumlah *cluster* dengan nilai *Davies Bouldin-Index* yang lebih rendah dianggap menunjukkan tingkat kualitas lebih baik dalam pembagian *cluster*.

Kata kunci: Penyandang disabilitas , agglomerative Hierarchical clustering, K-means , EDA (Exploratory Data Analysis), DBI Davies bouldin-Index

ABSTRACT

The issue of disability is still an issue that attracts public perception in relation to the discrimination they encounter every day. Society and the environment still stigmatize or treat persons with disabilities unfairly. Even though many people with disabilities have abilities that are equivalent to people who do not have physical limitations. Through this case study, the researcher hopes to help contribute to raising awareness so that the rights of persons with disabilities can be fulfilled, particularly in relation to access to proper employment and education and to explore new information regarding persons with disabilities. by processing datasets of persons with disabilities in Indonesia by taking data samples in 7 provinces in Indonesia based on 137 cities and regencies with 3 types of disability variables namely physical disability, blind or deaf disabilities, deaf and speech disabilities. Using K-means algorithm clustering analysis and agglomerative hierarchical clustering. to identify outliers and anomalies researchers use the EDA (Exploratory Data Analysis) method in order to assist in enriching the results of data analysis. when determining the centroid the researcher used the agglomerative hierarchical method, namely the Average Linkage based on the validity test of the highest cophenetic correlation value, this produced the same results as in the previous study. a total of 62 cities and regencies, cluster 2 is a cluster with a moderate level category totaling 37 cities and regencies, cluster 3 is a cluster with a low level category totaling 27 cities and regencies. the best evaluation results used the DBI Davies bouldin-Index method by producing 2 clusters because the number of clusters with a lower Davies Bouldin-Index value was considered to indicate a better level of quality in cluster division.

Keywords: Persons with disabilities, agglomerative Hierarchical clustering, K-means , EDA (Exploratory Data Analysis), DBI Davies bouldin-Index

KATA PENGANTAR

Puji syukur kami panjatkan kehadirat Allah SWT, akhirnya penulis dapat menyelesaikan skripsi yang berjudul *Analisis Clustering Data Penyandang Disabilitas dengan Menggunakan Metode Agglomerative Hierarical Clustering*. Skripsi ini ditulis dengan tujuan untuk memenuhi persyaratan terakhir untuk mencapai gelar Sarjana Komputer di Universitas Nusa Putra. Sehubungan dengan itu, penulis menyampaikan terima kasih dan penghargaan yang sebesar-besarnya kepada:

1. Orang tua dan keluarga yang senantiasa selalu memberikan dukungan dan do'a.
2. Bapak Dr. Kurniawan ST, M.Si, MM Rektor Universitas Nusa Putra Sukabumi
3. Ketua Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi Ibu Anggun Fergina, M.Kom
4. Dosen Pembimbing I Ibu Gina Purnama Insany, S.ST., M.Kom
5. Dosen Pembimbing II Bapak Alun Sujjada, S.Kom, M.T
6. Dosen Penguji Bapak Kamdan, M.Kom
7. Para Dosen Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi
8. Teman-teman mahasiswa Teknik Informatika 2019
9. Last but not least, I want to thank myself. I want to thank myself for believing in me, for working so hard, for never giving up, for always trying to give more than I receive, for always trying, starting and accepting for everything, thank you so much for me.



Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, oleh karena itu peneliti sangat mengharapkan evaluasi dan feedback dari berbagai pihak.

Sukabumi, juli 2023

Silvia Noer

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas NUSA PUTRA, saya yang bertanda tangan dibawah ini :

Nama : Silvia Noer

NIM 20190040104

Program Studi : Teknik Informatika

Jenis Karya : Tugas Akhir

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Nusa Putra **Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty- Free Right*)** atas karya ilmiah saya yang berjudul : “ Analisis *Clustering* Data Penyandang Disabilitas Menggunakan Metode *Agglomerative Hierarical Clustering* Dan *K-Means* “ beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non Eksklusif ini Universitas Nusa Putra berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.



Dibuat di : Sukadumi

Pada Tanggal : Juli 2023

Yang Menyatakan,

Silvia Noer

DAFTAR ISI

JUDUL	i
PERNYATAAN PENULIS.....	ii
PERSETUJUAN SKRIPSI.....	iii
PENGESAHAN SKRIPSI	iv
HALAMAN PERSEMBAHAN.....	v
ABSTRAK.....	vi
KATA PENGANTAR	viii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terkait.....	6
2.2 Penyandang disabilitas.....	12
2.3 Analisis Cluster.....	15
2.4 Python Programming Language	15
2.5 Machine Learning.....	15
2.6 EDA (Exploratory Data Analys)	27
2.7 Google Colab	27
2.8 Kerangka Berfikir	27
BAB III METODELOGI PENELITIAN.....	29



3.1 Tahapan Penelitian.....	29
3.2 Diagram Penelitian	29
3.3 Metode Pengumpulan Data.....	31
3.4 Metode Pengembangan sistem EDA (Exploratory Data Analysis)	31
3.5 Analisis kebutuhan.....	34
BAB IV HASIL DAN ANALISIS	36
4.1 Analisis Masalah.....	36
4.2 Data Akusisi.....	36
4.3 EDA & Praproses	36
4.4 Clustering.....	43
4.5 Web Visualisasi Hasil Analisis.....	50
4.6 Perbandingan hasil.....	52
BAB V KESIMPULAN DAN SARAN	53
5.1 Kesimpulan.....	53
5.2 Saran	53
DAFTAR PUSTAKA	54
LAMPIRAN.....	57



DAFTAR TABEL

Tabel Gambar 2. 1 Penelitian Terkait.....	6
Tabel Gambar 3. 1 Perangkat Kebutuhan Software	35
Tabel Gambar 3. 2 Perangkat Keras (<i>Hardware</i>).....	35
Tabel Gambar 4. 1 Evaluasi <i>Cophenetic correlation coefficient</i>	45
Tabel Gambar 4. 2 <i>Cluster</i> dengan <i>Average linkage</i>	46
Tabel Gambar 4. 3 Jumlah <i>cluster</i>	47
Tabel Gambar 4. 4 Nilai Inertia setiap <i>cluster</i>	48
Tabel Gambar 4. 5 Jumlah <i>cluster</i> dan nilai DBI.....	49
Tabel Gambar 4. 6 Perbandingan hasil	52



DAFTAR GAMBAR

Gambar 2. 1 Kerangka Berfikir.....	28
Gambar 3. 1 Diagram Penelitian	29
Gambar 3. 2 EDA (<i>Exploratory Data Analysis</i>)	31
Gambar 3. 3 Tahapan EDA pada penelitian.....	32
Gambar 4. 1 <i>Dataset</i>	36
Gambar 4. 2 <i>Import Library</i>	37
Gambar 4. 3 <i>Import Data</i>	37
Gambar 4. 4 <i>Missing Value Imputation</i>	38
Gambar 4. 5 <i>Descriptive Statics</i>	38
Gambar 4. 6 <i>Source code boxplot</i>	39
Gambar 4. 7 <i>visualisasi box-plot</i>	39
Gambar 4. 8 <i>Source code</i> menghitung IQR	40
Gambar 4. 9 Menghitung <i>lower</i> dan <i>upper limit</i> untuk <i>outlier</i>	41
Gambar 4. 10 <i>Source code</i> menampilkan data <i>outlier</i>	41
Gambar 4. 11 data <i>outlier</i>	41
Gambar 4. 12 <i>Source code</i> data tanpa <i>outlier</i>	42
Gambar 4. 13 Visualisasi <i>Box-plot</i> tanpa <i>outlier</i>	42
Gambar 4. 14 Dendogram <i>Complete Linkage</i>	43
Gambar 4. 15 Dendogram <i>Single Linkage</i>	44
Gambar 4. 16 Dendogram <i>Average Linkage</i>	44
Gambar 4. 17 <i>Cophenetic correlation coefficient</i>	45
Gambar 4. 18 Hasil <i>Cluster K-means</i> dan <i>agglomerative</i> dengan metode <i>average linkage</i>	46
Gambar 4. 19 <i>Elbow method</i>	47
Gambar 4. 20 Hasil <i>cluster</i> dengan metode DBI.....	49
Gambar 4. 21 Visualisasi hasil analisis ke dalam sistem	50
Gambar 4. 22 Visualisasi Batas <i>Confidence</i>	51



BAB I

PENDAHULUAN

1.1 Latar Belakang

Persoalan disabilitas masih menjadi isu yang menarik persepsi publik terkait dengan diskriminasi yang mereka temui sehari-hari, hal ini banyak disampaikan melalui berbagai tulisan, riset, dan laporan di berbagai belahan dunia.

Masyarakat maupun lingkungan masih menstigmatisasi atau memperlakukan penyandang disabilitas secara tidak adil. Padahal banyak penyandang disabilitas memiliki kemampuan yang setara dengan orang-orang yang tidak memiliki keterbatasan fisik.

Dalam bidang pekerjaan dan pendidikan. Banyak penyandang disabilitas tidak memiliki akses ke pekerjaan dan pendidikan yang sesuai. Bahkan mereka yang telah menyelesaikan studinya dengan sukses sering menghadapi tantangan dalam menemukan profesi yang terhormat dan dapat diterima. kondisi sosial dan ekonomi penyandang disabilitas di Indonesia dipengaruhi oleh kurangnya pendidikan profesional dan *soft skill*. Penyediaan infrastruktur yang tidak memadai adalah masalah besar lain yang harus diatasi.



Meskipun secara legal hak-hak penyandang disabilitas di Indonesia dilindungi undang-undang No. 8 Tahun 2016, memiliki disabilitas masih berbenturan dengan paradigma berpikir yang seringkali diskriminatif terhadap kalangan ini. Mungkin alasan penerapannya yang belum maksimal adalah karena masih sangat baru. Secara internasional, Konvensi PBB yang dibuat pada tahun 2006 dan disetujui pada tahun 2011 ini menjadi titik awal diakuinya hak-hak penyandang disabilitas. Konvensi PBB ini menjadi landasan bagi peran negara dalam mengakui hak-hak penyandang disabilitas[1].

Pada tahun 2021 dari total 7 miliar orang di dunia , 15% di antaranya mengalami disabilitas. 80% dari 15% itu tinggal di negara-negara berkembang. Menurut informasi yang diberikan oleh PBB untuk tahun 2021 [2]. Di Indonesia sendiri Menurut Gufroni, sekitar 34 juta orang memiliki disabilitas, sebagaimana ditentukan dari temuan Survei Sosial Nasional (Susenas) 2021 dan angka ini masih

semakin meningkat setiap tahun nya. Bila dinyatakan sebagai persentase, penyandang disabilitas merupakan 12% dari populasi umum [3].

Melalui studi kasus ini, peneliti berharap dapat membantu memberikan kontribusi dalam meningkatkan kesadaran agar hak-hak penyandang disabilitas dapat terpenuhi, khususnya terkait dengan akses pekerjaan dan pendidikan yang layak. Untuk itu, peneliti akan mengkaji data baru dengan mengolah dataset penyandang disabilitas di Indonesia dengan mengambil sampel data dari tujuh provinsi di Indonesia diantaranya: Provinsi Jawa Barat, Provinsi Jawa Tengah, Provinsi Jawa Timur, Provinsi Bali, Provinsi Jambi, Provinsi Bengkulu, dan Provinsi D.I Yogyakarta.

Terkait topik dataset yang diambil dari beberapa Website opendata setiap Provinsi dan BPS.go.id yang dikeluarkan dalam periode 1 tahun sekali.

Dari studi kasus ini peneliti mengharapkan bisa mengetahui mengenai tingkat disabilitas di Provinsi yang menjadi sampel pada penelitian ini dengan mengelompokkan kota dan kabupaten berdasarkan 3 variabel yaitu jumlah disabilitas dengan cacat fisik, cacat netra atau buta, cacat runtu dan wicara. menggunakan analisis *clustering* algoritma *K-means* dan agglomerative *hierarchical clustering* kemudian *Davies–Bouldin Index* sebagai algoritma evaluasi dari hasil *clustering*. Selanjutnya hasil *cluster* ini bisa menjadi bahan acuan maupun informasi baru untuk terwujudnya hak penyandang disabilitas di berbagai bidang kehidupan dalam merancang dan menetapkan kebijakan terkait. Dilansir dari situs media informasi menurut KOMNAS disabilitas, negara indonesia belum sepenuhnya memenuhi banyak hak penyandang disabilitas. Khususnya, individu dengan keterbatasan dijamin 22 hak berdasarkan undang-undang tentang orang dengan keterbatasan. Beberapa dari hak-hak ini dikelompokkan menjadi 6 dimensi yaitu: eksistensi, politik, ekonomi, masyarakat, budaya, dan layanan publik.

Algoritma *K-means clustering* dilakukan untuk pengolahan data dan cukup efektif pada saat pengelompokan data dengan ukuran besar algoritma ini merupakan salah satu metode *Data mining* yang sangat populer di industri pemrosesan data, Namun hasil *clustering* menggunakan *K-means* sangat bergantung pada penentuan pusat awal *cluster (centroid)*, dengan begitu maka hasil *clustering* menggunakan

metode *K-means* akan baik. Pada penelitian ini, peneliti menentukan pusat awal *cluster (centroid)* dengan menggunakan Algoritma *agglomerative hierarchical clustering*. Berdasarkan saran dari penelitian yang sudah ada, metode *hierarchical* yang digunakan dengan mengevaluasi *koefisien cophenetic*, metode pengelompokan terbaik dihasilkan oleh *average linkage clustering*. Penelitian ini menggunakan pembagian data berdasarkan jumlah kabupaten atau kota yang memiliki tingkat penyandang disabilitas dengan kategori dari jenis disabilitas yaitu cacat fisik, cacat netra atau buta, cacat rungu dan wicara. Di akhir agar hasil semakin efektif peneliti menggunakan algoritma *Davies–Bouldin index* mengukur evaluasi *cluster* yang didasarkan pada nilai kohesi dan separasi.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka rumusan masalah yang diajukan penulis, yaitu:

1. Bagaimana menganalisis *clustering* data penyandang disabilitas pada 7 Provinsi di Indonesia menggunakan metode *K-means* dan *agglomerative hierarchical clustering*
2. Bagaimana cara menangani *outlier* dengan metode pengembangan sistem EDA (*Exploratory Data Analysis*)
3. Bagaimana cara menentukan jumlah *cluster* yang optimal berdasarkan metode *Elbow* dan *Davies–Bouldin Index*?

1.3 Batasan Masalah

Penulis membatasi penelitian studi kasus pada 7 provinsi Indonesia kategori jenis penyandang disabilitas di 134 kota dan kabupaten dengan 3 variabel cacat fisik, cacat netra atau buta, cacat rungu agar penelitian ini fokus dan terarah. menuju tercapainya tujuan yang diinginkan peneliti fokus pada analisis data, dan tahapannya belum siap sampai ke *deployment machine learning*.

1.4 Tujuan

Tujuan akhir yang ingin dicapai dengan dilakukannya penelitian ini adalah:

1. Mengelompokkan jenis penyandang disabilitas dengan algoritma *K-means* dan *agglomerative hierarchical clustering*.
2. Menghilangkan *outlier* pada data disabilitas dengan metode pengembangan sistem EDA (Exploratory Data Analysis)
3. Menentukan metode evaluasi yang tepat untuk mengetahui Jumlah *cluster* yang optimal.
4. Membantu memberikan kontribusi dalam meningkatkan kesadaran agar hak-hak penyandang disabilitas dapat terpenuhi, khususnya terkait dengan akses pekerjaan dan pendidikan yang layak.

1.5 Manfaat Penelitian

1.5.1 Manfaat Bagi Peneliti

1. Untuk mengeksplorasi secara mendalam pemahaman tentang topik tersebut dan mengembangkan keahlian peneliti dalam bidang *data science*.
2. Memberikan pengalaman praktis bagi peneliti dalam mengumpulkan, menganalisis, dan menafsirkan data.



1.5.2 Manfaat Bagi Masyarakat

penelitian ini dapat memperoleh gambaran pengelompokan kota dan kabupaten yang memiliki penyandang disabilitas. Selanjutnya hasil *cluster* ini dapat menjadi bahan acuan untuk terwujudnya hak penyandang disabilitas mental di berbagai bidang kehidupan dalam merancang dan menetapkan kebijakan terkait.

1.6 Sistematika Penulisan

Penulis menyusun menjadi beberapa bab untuk memudahkan pembaca dalam mempelajari dan memahami hasil penelitian yang dilakukannya:

BAB I: PENDAHULUAN

Dalam bab ini dibahas mengenai: Latar Belakang, Batasan Masalah, Rumusan Masalah, Tujuan dan Manfaat Penelitian, Sistematika Penulisan.

BAB II: TINJAUAN PUSTAKA

Dalam bab ini menjelaskan teori yang digunakan untuk melengkapi penelitian ini, yaitu: Penelitian Terkait, Analisis *Cluster*, *EDA (Exploratory Data Analysis)*, *Data mining*, Normalisasi Data, *Agglomerative Hierarchical Clustering*, *Single Linkage*, *Average Linkage*, *complete Linkage*, *K-means Clustering*, *Elbow methode*, *Euclidean Distance*, *Davies–Bouldin index*, Google Colab, *Phyton*.

BAB III: METODOLOGI PENELITIAN

Menjelaskan Sistem Dalam Bab Ini: Metode Penelitian, Jenis Penelitian, Pengembangan Sistem, Pengumpulan Data sekunder, Studi Pustaka



BAB IV: HASIL DAN PEMBAHASAN

Analisis Kebutuhan Sistem dan Pengguna, Perancangan, Dataset, (Normalisasi Data Dengan Metode *Z-score*, Perhitungan Algoritma *Agglomerative Hierarchical Clustering*, Perhitungan *K-means Clustering*, Perhitungan *Clustering* Kombinasi *Agglomerative Hierarchical Clustering* dan *K-means* , *Davies–Bouldin index*.

BAB V: PENUTUP

Kesimpulan, Saran.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini telah berhasil melakukan *clustering* mulai dari data selection, Data mining sampai evaluasi dan membuat web visualisasi hasil analisis. maka dapat disimpulkan hasil penelitian sebagai berikut:

Berdasarkan nilai korelasi cophenetic, diperoleh metode terbaik dalam kasus *clustering* untuk data *non outlier* adalah metode *average* maka peneliti mengambil *centroid* dari *cluster* menggunakan *average linkage* yang akan digunakan pada algoritma *K-means*. Dengan menentukan 3 anggota *cluster* dengan algoritma *K-means* untuk kota dan kabupaten dengan kategori disabilitas yaitu *cluster* 1 merupakan tingkat tinggi dengan jumlah 62 kota dan kabupaten, *cluster* 2 merupakan *cluster* dengan kategori tingkat sedang jumlah 37 kota dan kabupaten, *cluster* 3 merupakan *cluster* dengan kategori tingkat rendah berjumlah 27 kota dan kabupaten. menggunakan 2 metode evaluasi untuk mengetahui jumlah *cluster* yang tepat untuk digunakan dalam algoritma *K-means* yaitu *elbow* dan DBI (*Davies bouldin-Index*). hasil evaluasi terbaik menggunakan metode DBI *Davies bouldin-Index* dengan menghasilkan 2 *cluster* karena jumlah *cluster* dengan nilai *Davies Bouldin-Index* yang lebih rendah dianggap menunjukkan tingkat kualitas lebih baik dalam pembagian *cluster*.

5.2 Saran

Saran-saran terhadap penelitian ini adalah sebagai berikut:

- a. Metode *elbow* tidak dapat menghasilkan solusi klaster yang jelas karena beberapa klaster memiliki titik-titik data yang sama dalam hal atribut atau karakteristik akibatnya penurunan inersia mungkin tidak berubah secara signifikan seiring bertambahnya jumlah *cluster*.
- b. Peneliti menyadari bahwa hasil penelitian ini masih jauh dari sempurna ,maka untuk penelitian selanjutnya bisa mencapai sampai tahapan *deploymen machine learning*.

DAFTAR PUSTAKA

- [1] U. Siregar, “Kaum Disabilitas Masih ‘Dianaktirikan’ di Tanah Air?,” *dw.com*, Mar. 12, 2021. <https://www.dw.com/id/disabilitas-dan-tantangannya/a-55625999> (accessed May 08, 2023).
- [2] admin plpp, “Hari Disabilitas Internasional 2021: Libatkan Penyandang Disabilitas PascaCovid-19,” Dec. 03, 2021. <https://puslapdik.kemdikbud.go.id/hari-disabilitas-internasional-2021-libatkan-penyandang-disabilitas-pascacovid-19/> (accessed Jun. 05, 2023).
- [3] M. P. Akbar, “Kesetaraan Akses Bagi Penyandang Disabilitas,” Dec. 04, 2021. <https://www.republika.id/posts/22808/kesetaraan-akses-bagi-penyandang-disabilitas#:~:text=Menurut%20Gufroni%2C%20berdasarkan%20hasil%20Survei%20Sosial%20Nasional%20%28Susenas%29,jumlah%20penduduk%20penyandang%20disabilitas%20sekitar%2034%20juta%20jiwa.> (accessed Jun. 05, 2023).
- [4] Ade Indah Sari, Heru Satria Tambun, Widodo Saputra, Irfan Sudahri Damanik, Ilham Syahputra Saragih (2020) “Implementasi Algoritma K-Means Dalam Pengelompokan”.
- [5] D. Permata Sari, F. Sembiring, D. Putri sulisdianto, and Y. Jentiner, “Jurnal Restikom : Riset Teknik Informatika dan Komputer IMPLEMENTASI ALGORITMA FUZZY C-MEANS UNTUK MEMPREDIKSI AKHIR SISWA MEMBOLOS (STUDI KASUS : FAKTOR SISWA MEMBOLOS DI SMPN 1 PARAKANSALAK),” vol. 2, no. 1, pp. 1–7, 2020, [Online]. Available: <https://restikom.nusaputra.ac.id>
- [6] M. Radhi, D. Ryan Hamonangan Sinurat, S. Hamonangan Sinurat, and E. Indra, “ANALISIS BIG DATA DENGAN METODE EXPLORATORY DATA ANALYSIS (EDA) DAN METODE VISUALISASI MENGGUNAKAN JUPYTER NOTEBOOK,” *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 4, no. 2, 2021.
- [7] E. Widodo, N. Sari, I. Hidayati, F. Yubinas, M. Yuniarti, and R. D. Novyantika, “Program Studi Pendidikan Matematika FKIP UMS” 2018, [Online]. Available: www.kependudukan.jogjaprovo.go.id.
- [8] Fadhil M Basysyar, Yudhistira Arie Wijaya, Irfan Ali, Saeful Anwar 2021. “CLUSTERING DATA DISABILITAS MENGGUNAKAN ALGORITMA K-MEANS DI KABUPATEN CIREBON”.
- [9] J. Ilmu and K. Sosial, “PENYANDANG DISABILITAS DI INDONESIA: PERKEMBANGAN ISTILAH DAN DEFINISI,” 2019.
- [10] Y. Prinada, “Profil Provinsi Jambi: Sejarah, Peta, Letak Wilayah & Topografi Baca selengkapnya di artikel ‘Profil Provinsi Jambi: Sejarah, Peta, Letak Wilayah & Topografi’, <https://tirto.id/gz2i>,” *Tirto.id*, Jan. 03, 2023. <https://tirto.id/profil-provinsi-jambi-sejarah-peta-letak-wilayah-topografi-gz2i> (accessed Jun. 06, 2023).
- [11] Indrawan, “Penyandang Difabel Harus Dapatkan Layanan Optimal Read more at: <https://jambione.com/read/2022/03/04/22760/penyandang-difabel-harus-dapatkan-layanan-optimal> Baca Terus Berita Ter Update jambione.com Follow Instagram https://www.instagram.com/jambione.com_official dan Like Fans Page

- <https://www.facebook.com/Jambione.Online>,” *Jambione.com*, Mar. 04, 2022. <https://jambione.com/read/2022/03/04/22760/penyandang-difabel-harus-dapatkan-layanan-optimal> (accessed Jun. 06, 2023).
- [12] I. M. Adnan, “Profil Provinsi Bali: Sejarah, Geografi dan Peta Wilayah ,” *Tirto.id*, Dec. 01, 2022. <https://tirto.id/profil-provinsi-bali-sejarah-geografi-dan-peta-wilayah-gzhy> (accessed Jun. 06, 2023).
- [13] suaradewata, “12.086 Penduduk Penyandang Disabilitas Ada di Bali,” *suaradewata.com*, Aug. 23, 2021. <https://www.suaradewata.com/read/202108230020/12-086-penduduk-penyandang-disabilitas-ada-di-bali.html> (accessed Jun. 06, 2023).
- [14] N. Ramadhani, A. F. Rahman, and D. Riskiyati, “ANALISIS CLUSTER DATA REGISTER PERKARA LALU LINTAS MENGGUNAKAN ALGORITMA K-MEANS,” 2017.
- [15] M. Ferdyandi, N. Y. Setiawan, and F. Abdurrachman Bachtiar, “Prediksi Potensi Penjualan Makanan Beku berdasarkan Ulasan Pengguna Shopee menggunakan Metode Decision Tree Algoritma C4.5 dan Random Forest (Studi Kasus Dapur Lilis),” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [16] G. P. Insany, M. A. Ayu, and T. Mantoro, “Using Machine Learning Techniques and Wi-Fi Signal Strength for Determining Indoor User Location,” in *2021 IEEE 7th International Conference on Computing, Engineering and Design (ICCED)*, IEEE, 2021, pp. 1–5.
- [17] I. Cholissodin and A. Soebroto, *Buku Ajar AI, Machine Learning & Deep Learning*. 2019.
- [18] “KLASTERISASI DATA DISABILITAS MENGGUNAKAN ALGORITMA K-MEANS”.
- [19] Dinas kependudukan dan catatan sipil, “Jumlah Penduduk Penyandang Disabilitas Berdasarkan Kategori Disabilitas di Jawa Barat,” Sep. 30, 2022. <https://opendata.jabarprov.go.id/id/dataset/jumlah-penduduk-penyandang-disabilitas-berdasarkan-kategori-disabilitas-di-jawa-barat> (accessed Jun. 14, 2023).
- [20] Dinas Sosial Provinsi Jawa Tengah, “Rekap Data Penyandang Masalah Kesejahteraan Sosial (PMKS) Provinsi Jawa Tengah, 2021,” Mar. 22, 2021. <https://jateng.bps.go.id/statictable/2022/03/22/2607/rekap-data-penyandang-masalah-kesejahteraan-sosial-pmks-provinsi-jawa-tengah-2021.html> (accessed Jun. 14, 2023).
- [21] Dinas Sosial Provinsi Jawa Timur, “Penyandang Masalah Kesejahteraan Sosial Menurut Kabupaten/Kota di Provinsi Jawa Timur,” Oct. 16, 2019. <https://jatim.bps.go.id/statictable/2019/10/16/2044/penyandang-masalah-kesejahteraan-sosial-menurut-kabupaten-kota-di-provinsi-jawa-timur-2017.html> (accessed Jun. 14, 2023).
- [22] portal resmi daerah istimewa yogya, “data disabilitas D.I Yogyakarta.” <https://jogjaprov.go.id/> (accessed Jun. 16, 2023).

- [23] BPS (Hasil Podes 2018), “Ada Penyandang Cacat dan Jenis Kecacatan 2018,” 2018. <https://jambi.bps.go.id/indicator/12/1693/1/-ada-penyandang-cacat-dan-jenis-kecacatan.html> (accessed Jun. 15, 2023).
- [24] BPS provinsi bengkulu, “BANYAKNYA DESA/KELURAHAN MENURUT KEBERADAAN PENYANDANG DISABILITAS,” Aug. 22, 2022. <https://bengkulu.bps.go.id/statictable/2022/08/25/1494/banyaknya-desa-kelurahan-menurut-keberadaan-penyandang-disabilitas.html> (accessed Jun. 15, 2023).
- [25] Satu Data Indonesia Provinsi Bali, “DATA PENYANDANG DIISABILITAS 2019,” 2019. <https://balisatudata.baliprov.go.id/dokumen/data-penyandang-disabilitas-2019-448?year=2019> (accessed Jun. 15, 2023).
- [26] E. Sadar Wasih, S. Rahmatullah, and S. Mintoro, “Implementasi Data Mining Pada Data Pelanggaran Lalu Lintas di Lampung Utara Menggunakan Algoritma K-Means (Studi Kasus Kejaksaan Negeri Lampung Utara).”
- [27] W. Budiaji, Y. L. A. Salampey, J. Agribisnis Fakultas Pertanian Universitas Sultan Ageng Tirtayasa Jalan Raya Jakarta Km, and P. Serang Banten, “PEDUGAAN STANDAR DEVIASI UNTUK SAMPEL KECIL DALAM PENELITIAN PERTANIAN,” *Jurnal Ilmu Pertanian dan Perikanan Desember*, vol. 1, no. 1, pp. 37–42, 2012.
- [28] S. Nurlaela, A. Primajaya, T. Nurmanjaya, U. H. Singaperbangsa Karawang Jl Ronggowaluyo, and T. Jambe Timur, “INFORMATIKA ALGORITMA K-MEDOIDES UNTUK CLUSTERING PENYAKIT MAAG DI KABUPATEN KARAWANG,” *Jurnal Informatika, Manajemen dan Komputer*, vol. 12, no. 2, 2020.
- [29] I. W. Septiani, Abd. C. Fauzan, and M. M. Suda, “Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, no. 4, p. 556, Jul. 2022, doi: 10.30865/json.v3i4.4055.
- [30] Sahir hafni syafri, *METODOLOGI PENELITIAN*, Cetakan pertama. Banguntapan, Bantul-Jogjakarta (Kantor I) : KBM INDONESIA, 2021.
- [31] P. Pola Untuk Mengetahui Jumlah Target Pengunjung Mall Berdasarkan Usia, P. Pertahun, T. Untuk Mempermudah Mengetahui, and D. Setiawan Suparno, “Target Pasar Menggunakan Metode EDA, K-Means, Hierarchial Clustering, Confusion Matrix,” 2021.
- [32] E. Koutroulis and K. Kalaitzakis, “Development of an integrated data-acquisition system for renewable energy sources systems monitoring,” 2003. [Online]. Available: www.elsevier.com/locate/renene
- [33] E. D. Wahyuni, A. A. Arifiyanti, and M. Kustyani, “Exploratory Data Analysis dalam Konteks Klasifikasi Data Mining,” pp. 263–269, 2019, [Online]. Available: <http://journal.itny.ac.id/index.php/ReTII>