

Implementasi KNN dan *Decision Tree* Untuk Klasifikasi Kelayakan Bantuan Disabilitas DKI Jakarta

Nathan Ramadhani¹, Dwi Wahyu Maulana², Muhammad Rizky Fahreza^{*3}

^{1,2,3}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana

Email : ¹41521010002@student.mercubuana.ac.id, ²41521010147@student.mercubuana.ac.id,

^{*3}41521010130@student.mercubuana.ac.id

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi layak atau tidaknya penerima bantuan disabilitas menerima bantuan, berdasarkan peraturan Gubernur DKI Jakarta Nomor 44 Tahun 2022 tentang Pemberian Bantuan Sosial Dalam Rangka Perlindungan. Penelitian ini menggunakan pendekatan kualitatif dengan menggunakan python dan mengambil studi kasus nyata menggunakan dataset dari Dinas Sosial DKI Jakarta tahun 2020. Penelitian ini melibatkan dua algoritma machine learning, yaitu *Decision Tree* untuk mengklasifikasikan layak atau tidaknya dan *K-Nearest Neighbors* (KNN) untuk memvisualisasikan penerima bantuan yang termasuk ke dalam kelompok layak dan tidak layak. Banyaknya data penyandang disabilitas yang tercatat, membuat Dinas Sosial Provinsi DKI Jakarta mengalami hambatan dalam mengidentifikasi penyandang disabilitas. Jumlah penduduk yang tinggi dan sumber daya serta dana yang terbatas, menjadi tantangan besar yang sedang dihadapi oleh Dinas Sosial untuk memastikan bahwa bantuan disalurkan dengan tepat dan efektif. Salah satu upaya yang dapat dilakukan oleh Dinas Sosial Provinsi DKI Jakarta adalah memberikan Kartu Penyandang Disabilitas Jakarta (KPDJ). Hasil penelitian ini menunjukkan klasifikasi dengan algoritma *decision tree* mendapatkan skor akurasi 0.78, hal ini menunjukkan bahwa sistem yang dikembangkan mampu memberikan rekomendasi dengan tingkat keakuratan yang tinggi dalam menentukan penerima manfaat bantuan, sehingga dapat meningkatkan efisiensi dalam proses pengambilan keputusan terkait penyaluran bantuan.

Kata Kunci— KNN, *Decision Tree*, Disabilities

Abstract

The purpose of this study is to identify whether recipients of disability assistance are eligible to receive aid based on the Governor of DKI Jakarta Regulation Number 44 of 2022 concerning the Provision of Social Assistance for Protection. This study uses a qualitative approach with Python and takes a real case study using a dataset from the DKI Jakarta Social Service in 2020. The study involves two machine learning algorithms: *Decision Tree* to classify eligibility and *K-Nearest Neighbors* (KNN) to visualize the recipients of assistance who fall into the eligible and non-eligible groups. The large number of recorded disabled individuals has created challenges for the DKI Jakarta Provincial Social Service in identifying them. The high population, along with limited resources and funds, presents a significant challenge for the Social Service to ensure that assistance is distributed accurately and effectively. One of the efforts that the DKI Jakarta Provincial Social Service can undertake is to issue the Jakarta Disability Card (KPDJ). The results of this study show that the classification with the *Decision Tree* algorithm achieved an accuracy score of 0.78, indicating that the developed system is capable of providing recommendations with a high level of accuracy in determining aid recipients, thereby improving efficiency in the decision-making process related to aid distribution.

Keywords— KNN, *Decision Tree*, Disabilities

1. PENDAHULUAN

Penyandang disabilitas merupakan salah satu kelompok rentan yang memerlukan perhatian khusus dalam upaya meningkatkan kualitas hidup mereka. Dalam konteks ini, penerimaan bantuan dan dukungan menjadi krusial untuk memastikan kebutuhan mereka terpenuhi merupakan kunci untuk membangun kesejahteraan bangsa dan negara. Namun, proses identifikasi dan penyaluran bantuan seringkali menghadapi hambatan yang kompleks, terutama dalam mengelola volume data yang besar dan beragam. Jumlah penduduk yang tinggi dan sumber daya serta dana yang terbatas, menjadi tantangan besar yang sedang dihadapi oleh Dinas Sosial untuk memastikan bahwa bantuan disalurkan dengan tepat dan efektif. Salah satu upaya yang dapat dilakukan oleh Dinas Sosial Provinsi DKI Jakarta adalah memberikan Kartu Penyandang Disabilitas Jakarta (KPDJ). KPDJ merupakan program bantuan yang bertujuan untuk mencegah kerentanan sosial serta upaya mendukung pemenuhan berbagai kebutuhan dasar dan kesejahteraan sosial bagi para penyandang disabilitas. Penerima manfaat KPDJ akan menerima dana sebesar Rp300.000 tiap bulannya melalui Bank DKI. Berdasarkan peraturan Gubernur DKI Jakarta Nomor 44 Tahun 2022 tentang Pemberian Bantuan Sosial Dalam Rangka Perlindungan, kriteria penerima bantuan sosial diantaranya harus memiliki KTP atau KK berdomisili Provinsi DKI Jakarta, telah terdata di dukcapil sebagai disabilitas, status kesejahteraan rumah tangga atau desil 1-4.

Selain itu calon penerima bantuan juga harus terdaftar di DTKS dan bukan merupakan bagian dari TNI, Polri, ASN dan BUMN/BUMD. Pendataan ini bertujuan untuk mengurangi kesalahan dalam penyaluran bantuan dan memastikan bantuan tepat sasaran, dan meminimalkan potensi penyalahgunaan dana.

Proses pendataan yang mendetail memerlukan waktu yang cukup lama dan sumber daya yang banyak. Dalam proses penyaluran bantuan sering kali menimbulkan masalah baru. Salah satu risiko yang muncul adalah kemungkinan adanya korupsi, di mana oknum tertentu bisa memanfaatkan celah dalam sistem yang kompleks untuk keuntungan pribadi. Selain itu, proses yang panjang dan berbelit-belit dapat menyebabkan penundaan dalam penyaluran bantuan, sehingga masyarakat yang benar-benar membutuhkan tidak segera menerima bantuan yang diperlukan. Masalah lain yang sering muncul adalah ketidaksesuaian jumlah penyediaan dengan kebutuhan di lapangan, yang bisa terjadi akibat data yang tidak akurat atau perubahan kebutuhan yang cepat di masyarakat. Diperlukan sistem yang efisien dan transparan untuk mempercepat pendataan dan penyaluran, serta memastikan bantuan mencapai mereka yang benar-benar membutuhkan.

Tujuan utama dari penelitian ini adalah membuat sistem yang mampu melakukan identifikasi penerima manfaat bantuan dengan tingkat keakuratan yang tinggi, serta meningkatkan efisiensi dalam proses pengambilan keputusan terkait penyaluran bantuan. Dengan memanfaatkan algoritma KNN untuk analisis asosiasi dan algoritma Decision Tree untuk pengklasifikasian data diharapkan akan tercipta suatu sistem yang dapat memberikan rekomendasi yang lebih baik dalam menentukan penerima manfaat secara tepat.

Data mining merupakan metode yang bermanfaat untuk memperoleh informasi berharga dari sejumlah data yang dilakukan dengan menggunakan pengetahuan seperti statistik, matematika dan pengenalan pola. Pola tersebut nantinya dapat menghasilkan informasi yang bermanfaat. Menurut Han et al. (2014) data mining memiliki beberapa tahapan yaitu :

1. Data Cleaning
2. Data Integration
3. Data Selection
4. Data Transformation
5. Data mining
6. Pattern Evaluation
7. Knowledge Presentation

Dalam konteks identifikasi penerima manfaat bantuan bagi penyandang disabilitas, data mining digunakan untuk menganalisis dan memproses data demografis, kondisi ekonomi, kebutuhan khusus, serta riwayat bantuan. Tujuan utamanya adalah untuk meningkatkan efisiensi dalam penyaluran bantuan dengan memastikan bahwa bantuan disalurkan kepada mereka yang benar-benar membutuhkannya.

Algoritma K-Nearest Neighbors (KNN) adalah salah satu algoritma pembelajaran berbasis instansi yang umum digunakan untuk mengelompokkan data. Algoritma ini mengelompokkan data berdasarkan mayoritas dari kategori k-nearest neighbors atau tetangga terdekatnya dalam ruang atribut. Dalam konteks penelitian ini, algoritma KNN digunakan untuk memprediksi kelayakan penerima manfaat bantuan berdasarkan karakteristik individu penyandang disabilitas. Dengan menggunakan atribut-atribut yang relevan, algoritma KNN akan mengklasifikasikan individu ke dalam kelompok penerima manfaat atau bukan penerima manfaat.

Decision Tree adalah algoritma pembentukan pohon keputusan yang digunakan untuk pengklasifikasian data. Algoritma ini membangun pohon keputusan berdasarkan atribut-atribut data dan memilih atribut terbaik untuk memisahkan data ke dalam kelas-kelas yang berbeda berdasarkan nilai Entropy dan Gain. Menurut Amin, et al., (2015) sebelum mengekstrak data dalam bentuk tree terdapat beberapa langkah yang harus diperhatikan sebagai berikut :

1. Melakukan perhitungan nilai entropy untuk semua anggota atribut.
2. Melakukan perhitungan nilai Gain pada setiap nilai atribut.
3. Memilih root berdasarkan nilai Gain terbesar.
4. Membuat node cabang berdasarkan nilai Gain terbesar kecuali nilai Gain untuk root.

Dalam penelitian ini, algoritma Decision Tree digunakan untuk mengklasifikasikan jenis disabilitas penerima manfaat bantuan berdasarkan karakteristik dari data individu penyandang disabilitas.

Penggunaan data mining, khususnya algoritma KNN dan Decision Tree, dalam proses identifikasi penerima manfaat bantuan diharapkan dapat meningkatkan efisiensi penyaluran bantuan. Dengan menganalisis pola-pola hubungan antar atribut data dan mengklasifikasikan kelayakan penerima manfaat, sistem dapat memberikan rekomendasi yang lebih akurat dan tepat dalam menentukan penerima manfaat. Hal ini diharapkan dapat mengurangi kesalahan dalam penyaluran bantuan, memastikan bantuan tepat sasaran, dan meningkatkan efisiensi dalam proses pengambilan keputusan terkait penyaluran bantuan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan menggunakan python dan mengambil studi kasus nyata menggunakan dataset dari Dinas Sosial DKI Jakarta tahun 2020. Penelitian ini melibatkan dua algoritma machine learning, yaitu Decision Tree dan K-Nearest Neighbors (KNN), untuk melakukan analisis data.

1. Decision Tree

Algoritma machine learning yang digunakan untuk tugas klasifikasi dan regresi. Algoritma ini bekerja dengan membagi dataset ke dalam subset yang lebih kecil berdasarkan fitur tertentu [7]. Struktur pohon terbentuk dengan node yang merepresentasikan fitur dari dataset, cabang yang merepresentasikan aturan keputusan, dan daun yang merepresentasikan hasil (kelas) [7].

Nilai entropy dapat dihitung menggunakan persamaan 1

$$Entropy(S) = \sum_{i=1}^n -p_i \star \log_2 p_i \quad (1)$$

Keterangan :

S : Himpunan kasus

n : Jumlah partisi S

p_i : Jumlah sampel pada kelas i

Nilai gini index dapat dihitung dengan persamaan 2

$$\text{Gini index} = 1 - \sum_{i=1}^n p_i^2 \quad (2)$$

Keterangan :

n : Jumlah partisi S

p_i : Jumlah sampel pada kelas i

2. K-Nearest Neighbors (KNN)

Algoritma K-Nearest Neighbor(KNN) adalah metode untuk mengklasifikasikan objek berdasarkan data training yang paling dekat dengan objek tersebut[11]. KNN bekerja dengan mencari titik data terdekat (neighbors) dari titik data baru yang akan diklasifikasikan. Jarak terdekat biasanya dihitung menggunakan metrik Euclidean, meskipun metrik lainnya juga bisa digunakan. Kelas dari titik data baru ditentukan berdasarkan mayoritas kelas dari K tetangga terdekat.

Jarak euclidean dapat dihitung dengan persamaan

$$\text{jarak} = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + (x_3 - x'_3)^2} \quad (3)$$

Keterangan :

x_n, x_m : record ke-n , record ke-m

x'_n, x'_m : record ke-n' , record ke-m'

3. Populasi dan Sampel

Data yang digunakan dalam penelitian ini adalah data individu dari suatu dataset yang berisi informasi mengenai nama, desa, usia, jenis kelamin, dan jenis disabilitas warga DKI Jakarta dan Kepulauan Seribu. Data ini diambil dari website Satu Data Indonesia (data.go.id). Dataset tersebut terdiri dari 10.000 data individu penyandang disabilitas yang diambil dari tahun 2020.

Tabel 1. Sample Dataset disabilitas warga DKI Jakarta

Jenis Kelamin	Usia	Jenis Disabilitas	Desil
Perempuan	71	Cacat Fisik & Mental	3
Laki-laki	68	Tuna Netra	0
Perempuan	32	Cacat Mental Retardasi	3
Perempuan	70	Tuna Netra	3
Perempuan	29	Tuna Rungu, Wicara & Daksa	3
...	...	...	...

2.1 Teknik Pengumpulan Data dan Pengembangan Instrumen

1. Preprocessing Data:

Dalam dataset asli, terdapat beberapa kolom yang tidak relevan atau tidak diperlukan untuk analisis lebih lanjut. Menghapus kolom-kolom ini membantu menyederhanakan data dan fokus pada fitur yang relevan. Untuk mempermudah pemahaman dan analisis data, nama kolom diubah menjadi lebih deskriptif. Penggantian ini dilakukan agar peneliti dan pembaca lebih mudah memahami isi dan konteks data. Beberapa kolom dalam dataset mengandung teks yang mungkin memiliki karakter khusus atau spasi yang tidak perlu. Pembersihan ini penting untuk menghindari kesalahan saat melakukan encoding atau analisis.

2. One-Hot Encoding:

Digunakan untuk mengubah data kategorik menjadi bentuk numerik yang bisa diproses oleh model Decision Tree.

- a. Label Encoding: Digunakan untuk mengubah data kategorik menjadi bentuk numerik untuk model KNN.
- b. Pembagian Data: Data dibagi menjadi fitur (x_{train}) dan target (y_{train}) untuk masing-masing model.

2.2 Teknik Analisis Data

1. Deskripsi Data

Dalam penelitian ini, atribut yang menjadi sasaran dalam pengujian menggunakan desil adalah variabel-variabel kunci yang mencerminkan karakteristik dan kondisi individu penyandang disabilitas di DKI Jakarta dan Kepulauan Seribu. Atribut-atribut ini meliputi informasi demografis seperti umur, jenis kelamin, dan jenis disabilitas. Tujuan dari analisis ini adalah untuk memahami distribusi dan pola dalam data, serta untuk mengidentifikasi kelompok-kelompok tertentu yang mungkin memerlukan perhatian khusus dalam kebijakan sosial dan program bantuan. Data utama yang digunakan dalam penelitian ini berasal dari dataset tahun 2020, yang berisi informasi mengenai penyandang disabilitas di wilayah tersebut. Dataset ini mencakup 14.937 individu dengan berbagai jenis disabilitas, termasuk cacat fisik, mental, tuna netra, dan kombinasi dari beberapa jenis disabilitas. Informasi ini mencakup detail alamat, nama, umur, dan jenis disabilitas yang dihadapi oleh masing-masing individu.

2. Decision Tree

Berikut langkah-langkah Decision Tree:

- a. Inisialisasi Model: Menggunakan `DecisionTreeClassifier` dari Scikit-Learn.
- b. Pelatihan Model: Model dilatih menggunakan data yang telah di encode dengan One-Hot Encoding.
- c. Visualisasi Model: Model yang terlatih divisualisasikan dalam bentuk tree menggunakan `export graphviz` dan

3. K-Nearest Neighbors(KNN)

Berikut langkah-langkah KNN :

- a. Inisialisasi Model: Menggunakan `Neighbors Classifier` dari Scikit Learn dengan $K=3$.
- b. Pelatihan Model: Model dilatih menggunakan data yang telah di encode dengan Label Encoding.
- c. Prediksi: Model digunakan untuk mengklasifikasikan jenis disabilitas berdasarkan data baru.
- d. Visualisasi: Visualisasi dilakukan dengan scatter plot untuk melihat distribusi data dan hasil klasifikasi.

3. HASIL DAN PEMBAHASAN

3.1 Decision Tree

Dalam mengembangkan sistem monitoring dan Dimulai dengan memilih atribut yang paling informatif untuk menjadi akar pohon. Kita akan menggunakan data yang telah diolah dalam code yang diberikan.

1. Node Akar

Kondisi: Usia ≤ 14.5

Samples: 10455

Value: [8116, 2339] (Layak, Tidak Layak)

Di mana p_i adalah adalah proporsi kelas i :

Untuk node ini:

$$p_{Layak} = \frac{8116}{10455} = 0.776$$

$$p_{Tidak Layak} = \frac{2339}{10455} = 0.224$$

$$Gini = 1 - (0.776^2 + 0.224^2)$$

$$Gini = 1 - (0.602 + 0.050)$$

$$Gini = 1 - 0.652$$

$$Gini = 0.348$$

2. Node Kiri (Child Node dari Root)

Kondisi: Usia ≤ 11.5

Samples: 1231

Value: [860, 371] (Layak, Tidak Layak)

Di mana p_i adalah adalah proporsi kelas i :

Untuk node ini:

$$p_{Layak} = \frac{860}{1231} = 0.698$$

$$p_{Tidak Layak} = \frac{371}{1231} = 0.302$$

$$Gini = 1 - (0.698^2 + 0.302^2)$$

$$Gini = 1 - (0.487 + 0.091)$$

$$Gini = 1 - 0.578$$

$$Gini = 0.422$$

3. Node Kanan (Child Node dari Root)

Kondisi: Usia > 14.5

Samples: 9224

Value: [7256, 1968] (Layak, Tidak Layak)

Di mana p_i adalah adalah proporsi kelas i :

Untuk node ini:

$$p_{Layak} = \frac{7256}{9224} = 0.787$$

$$p_{Tidak Layak} = \frac{1968}{9224} = 0.213$$

$$Gini = 1 - (0.787^2 + 0.213^2)$$

$$Gini = 1 - (0.619 + 0.045)$$

$$Gini = 1 - 0.664$$

$$Gini = 0.336$$

4. Node Anak dari Node Kiri

Kondisi: Usia <= 5.5

Samples: 732

Value: [498, 234] (Layak, Tidak Layak)

Di mana p_i adalah proporsi kelas i :

Untuk node ini:

$$p_{Layak} = \frac{498}{732} = 0.68$$

$$p_{Tidak Layak} = \frac{234}{732} = 0.32$$

$$Gini = 1 - (0.462 + 0.102)$$

$$Gini = 1 - 0.564$$

$$Gini = 0.436$$

5. Node Anak dari Node Kanan

Kondisi: Jenis Kelamin <= 0.5

Samples: 9044

Value: [7093, 1951] (Layak, Tidak Layak)

Di mana p_i adalah proporsi kelas i :

Untuk node ini:

$$p_{Layak} = \frac{7093}{9044} = 0.784$$

$$p_{Tidak Layak} = \frac{1951}{9044} = 0.216$$

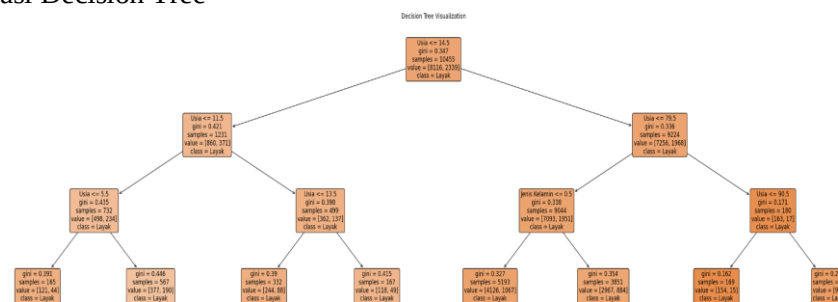
$$Gini = 1 - (0.784^2 + 0.216^2)$$

$$Gini = 1 - (0.615 + 0.047)$$

$$Gini = 1 - 0.662$$

$$Gini = 0.338$$

6. Visualisasi Decision Tree



Gambar 1 visualisasi *Decision Tree*

Visualisasi ini menunjukkan model decision tree membagi dataset berdasarkan atribut Usia dan Jenis Kelamin untuk menentukan apakah seseorang layak menerima bantuan. Setiap node dipisahkan oleh nilai Gini, jumlah sampel, distribusi kelas, dan kelas mayoritas di node tersebut. Semakin dalam node berada, semakin spesifik kondisi pemisahan yang digunakan untuk memprediksi kelayakan penerima bantuan.

Berikut adalah kode program untuk menentukan Layak atau Tidak Layak penyandang disabilitas menerima bantuan:

```
df['Status'] = df['Desil'].apply(lambda x: 'Layak' if x <= 4 else 'Tidak Layak')
```

7. K-Nearest Neighbor

Algoritma K-Nearest Neighbor (KNN) adalah metode klasifikasi yang mengandalkan data yang sudah diklasifikasikan sebelumnya untuk mempelajari dan mengelompokkan data baru. Sebagai bagian dari supervised learning, hasil dari instance baru diklasifikasikan berdasarkan kedekatan jarak dengan kategori yang sudah ada. Algoritma ini berfungsi dengan mengukur jarak terpendek dari sampel uji ke sampel latih untuk menemukan K tetangga terdekat (KNN). Setelah mengidentifikasi K tetangga terdekat, keputusan klasifikasi dilakukan berdasarkan mayoritas kategori dari tetangga tersebut. Umumnya, jarak antara tetangga dihitung menggunakan jarak Euclidean. Langkah-langkah dalam menghitung metode K-Nearest Neighbor meliputi:

a) Persiapan Data

Data menggunakan 5 data terakhir dari 100 Total Data Sampel.

b) Encoding Data

Encode data secara manual:

Jenis Disabilitas: Tuna Daksa = 0, Tuna Rungu = 1, Cacat Fisik & Mental = 2

Status Kelayakan: Layak = 0, Tidak Layak = 1

Sehingga data ter-encode akan menjadi sesuai dengan Tabel 4

Tabel 2. Sampel data ter-encode

Nama	Usia	Jenis Disabilitas	Desil	Status Kelayakan
TURINA	46	2	7	0
ARMAH	58	1	2	1
MASKAH	48	4	7	0
FAHRI AZZAHRA N PUTRA GUSTIFAH	16	0	4	1
MOCHAM AD SOLEH	69	3	2	1

c) Hitung jarak Euclidean antara data baru dan setiap data dalam dataset

Misalkan kita ingin memprediksi jenis disabilitas untuk seseorang dengan:

- Usia: 50
- Jenis Disabilitas: Tuna Daksa (1)
- Desil: 5

Jarak Euclidean dapat dihitung menggunakan rumus [3]:

Perhitungan Jarak Euclidean::

1) Jarak ke TURINA:

$$\begin{aligned}
 Jarak &= \sqrt{(50 - 46)^2 + (1 - 2)^2 + (5 - 7)^2} \\
 Jarak &= \sqrt{4^2 + 1^2 + 2^2} \\
 Jarak &= \sqrt{16 + 1 + 4} \\
 Jarak &= \sqrt{21} \\
 Jarak &= 4.58
 \end{aligned}$$

2) Jarak ke ARMAH:

$$\begin{aligned} \text{Jarak} &= \sqrt{(50 - 58)^2 + (1 - 1)^2 + (5 - 2)^2} \\ \text{Jarak} &= \sqrt{8^2 + 0^2 + 3^2} \\ \text{Jarak} &= \sqrt{64 + 0 + 9} \\ \text{Jarak} &= \sqrt{73} \\ \text{Jarak} &= 8.54 \end{aligned}$$

3) Jarak ke MASKAH:

$$\begin{aligned} \text{Jarak} &= \sqrt{(50 - 48)^2 + (1 - 4)^2 + (5 - 7)^2} \\ \text{Jarak} &= \sqrt{2^2 + 3^2 + 2^2} \\ \text{Jarak} &= \sqrt{4 + 9 + 4} \\ \text{Jarak} &= \sqrt{17} \\ \text{Jarak} &= 4.12 \end{aligned}$$

4) Jarak ke ORAHIM:

$$\begin{aligned} \text{Jarak} &= \sqrt{(50 - 16)^2 + (1 - 0)^2 + (5 - 4)^2} \\ \text{Jarak} &= \sqrt{34^2 + 1^2 + 1^2} \\ \text{Jarak} &= \sqrt{1156 + 1 + 1} \\ \text{Jarak} &= \sqrt{1158} \\ \text{Jarak} &= 34.03 \end{aligned}$$

5) Jarak ke MOCHAMAD SOLEH:

$$\begin{aligned} \text{Jarak} &= \sqrt{(50 - 69)^2 + (1 - 3)^2 + (5 - 2)^2} \\ \text{Jarak} &= \sqrt{19^2 + 2^2 + 3^2} \\ \text{Jarak} &= \sqrt{361 + 4 + 9} \\ \text{Jarak} &= \sqrt{374} \\ \text{Jarak} &= 19.34 \end{aligned}$$

d) Pilih K Tetangga Terdekat:

Dengan K=3, kita memilih 3 tetangga terdekat berdasarkan jarak terkecil:

ROCHIM (3.32)

MOCH ARIEF SYAEFUDDIN (3.61)

MASKAH (4.12)

Hasil Klasifikasi:

Mayoritas kelas dari ketiga tetangga terdekat adalah:

ROCHIM: Layak

MOCH ARIEF SYAEFUDDIN: Tidak Layak

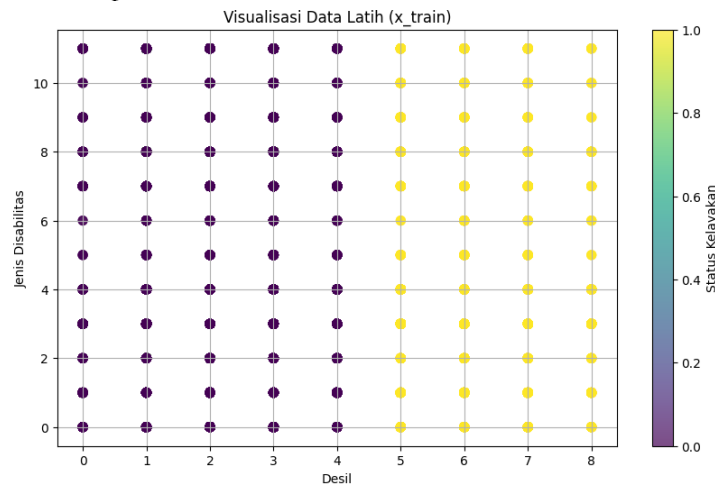
MASKAH: Tidak Layak

Dari 3 tetangga terdekat, 2 mengatakan "Tidak Layak" dan 1 menyatakan "Layak".

Oleh karena itu, prediksi untuk data baru (Usia: 50, Jenis Disabilitas: Tuna Daksa, Desil:

5) adalah Tidak Layak.

- e) Visualisasi KNN dengan atribut Jenis Disabilitas dan Desil dan K=3
Visualisasi KNN dapat dilihat di Gambar 2



Gambar 2. Visualisasi KNN untuk atribut Desil dan Jenis Disabilitas

Metode K-Nearest Neighbor (KNN) menggunakan data tetangga terdekat untuk menentukan status kelayakan berdasarkan keputusan mayoritas. Dengan K=3, kita dapat memperoleh prediksi untuk Usia = 50, Jenis_Disabilitas = 1, Desil = 5 menggunakan Kode Program :

```
from sklearn.neighbors import KNeighborsClassifier
```

```
K=3
```

```
model = KNeighborsClassifier(n_neighbors=K)
```

```
model.fit(x_train_knn, y_train_status)
```

```
y_new = model.predict(x_train_knn)
```

```
y_new
```

```
y_pred = model.predict(x_new)
```

```
print(y_pred)
```

Dari data testing di atas, maka outputnya sebagai berikut:

```
[1]
```

yang dimana 1 adalah Tidak Layak

4. KESIMPULAN

Berdasarkan hasil dan pembahasan, Algoritma Decision Tree dan KNN dapat diterapkan untuk mengklasifikasi bantuan penerima penyandang disabilitas. Algoritma Decision Tree dapat mengklasifikasikan data penyandang disabilitas berdasarkan beberapa atribut dengan skor akurasi sebesar 0.78.

Berdasarkan hasil penelitian ini, hasil perhitungan manual decision tree dan KNN kami mendapatkan hasil yang sama dengan program python.

DAFTAR PUSTAKA

- [1] et al. Kurniadi, Y U., "Penyandang Disabilitas di Indonesia," Nusan. J. Ilmu Pengetah. Sos., vol. 7, no. 2, pp. 408–420, 2020.

-
- [2] E. R. Susanto, A. S. Puspaningrum, and N. Neneng, "Model Rekomendasi Penerima Bantuan Sosial Berdasarkan Data Kesejahteraan Rakyat," *J. Tekno Kompak*, vol. 15, no. 1, p. 1, 2021, doi: 10.33365/jtk.v15i1.915.
- [3] J. K. Propiona, "Implementasi Aksesibilitas Fasilitas Publik Bagi Penyandang Disabilitas," *J. Anal. Sosial.*, vol. 10, pp. 1–18, 2021, doi: 10.20961/jas.v10i0.47635.
- [4] I. S. Latif and I. A. Pangestu, "Problematisasi Penyalahgunaan Bantuan Sosial Pada Masa Pandemi," *Justisi*, vol. 8, no. 2, pp. 95–107, 2022, doi: 10.33506/js.v8i2.1612.
- [5] Z. Rahman and S. H. Malawat, "Efektivitas Program Bantuan Sosial Tunai Bagi Masyarakat (Studi di Kelurahan Pelambuan Kota Banjarmasin)," *ePrints Univ. Islam Kalimantan*, pp. 1–9, 2021.
- [6] D. Irawan et al., "Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Mengklasifikasi Jenis Penerimaan Bantuan Studi Kasus Desa Madukoro Lampung Utara," *J. Teknol. Komput. dan Sist. Inf.*, vol. 6, no. 2, pp. 204–207, 2023, [Online]. Available: <https://jurnal.ftikomibn.ac.id/>
- [7] N. I. Nella, N. Y. Setiawan, and D. E. Dian, "Klasifikasi Penerima Bantuan Program Keluarga Harapan menggunakan Algoritma Decision Tree C4. 5 (Studi Kasus: Desa Mlirip Kabupaten Mojokerto)," ... *Teknol. Inf. dan ...*, vol. 6, no. 3, pp. 1332–1339, 2022, [Online]. Available: <https://jptiik.multi.web.id/index.php/jptiik/article/view/10801%0Ahttps://jptiik.multi.web.id/index.php/jptiik/article/download/10801/4769>
- [8] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [9] P. Lestari, N. Suarna, and W. Prihartono, "Implementasi Data Mining Clustering Dalam Mengelompokkan Penduduk Penyandang Disabilitas Menggunakan Algoritma K-Means," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 958–966, 2024, doi: 10.36040/jati.v8i1.8445.
- [10] Y. A. Setianto, K. Kusriani, and H. Henderi, "Penerapan Algoritma K-Nearest Neighbour Dalam Menentukan Pembinaan Koperasi Kabupaten Kotawaringin Timur," *Creat. Inf. Technol. J.*, vol. 5, no. 3, p. 232, 2019, doi: 10.24076/citec.2018v5i3.179.
- [11] P. Sokibi and V. Dwi Kartika, "Sistem Klasifikasi Penentuan Penerima Bantuan Langsung Tunai Menggunakan Metode K-Nearest Neighbor," pp. 145–155, 2023, doi: 10.30864/eksplora.v13i2.907.