

# Klasifikasi Keluarga Miskin Menggunakan Metode K-Nearest Neighbor Berbasis Euclidean Distance

Fitra Kurnia, S.Kom, MT<sup>1</sup>, Jhoni Kurniawan<sup>2</sup>, Ichsan Fahmi ST, MT<sup>3</sup>,  
Siti Monalisa, ST, M.Kom<sup>4</sup>

<sup>1,2</sup>Program Studi Teknik Informatika Fakultas Sains dan Teknologi

<sup>4</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

JL. HR. Soebrantas No.155 KM. 15 Simpang Baru Panam Pekanbaru 28293

Universitas Nusa Cendana, Jalan Adi Sucipto Penfui Kupang, NTT 85812

<sup>1</sup>fitra.k@uin-suska.ac.id

## Abstrak

Kemiskinan merupakan salah satu tantangan terbesar yang masih terus dihadapi semua bangsa termasuk Indonesia. Sesuai amanat konstitusi yang tertuang dalam pasal 34 Undang Undang Dasar 1945 maka pemerintah terus berupaya melindungi dan memberdayakan masyarakat miskin melalui berbagai program strategis. Salah satu problem utama yang sering menghambat suksesnya program pemerintah tersebut adalah ketersediaan data sasaran masyarakat miskin yang sesuai. Hal ini menyebabkan program pemerintah tidak berjalan secara maksimal karena perlu verifikasi dan validasi ulang. Penelitian ini bertujuan untuk merancang aplikasi yang dapat melakukan klasifikasi keluarga miskin menggunakan Algoritma K-Nearest Neighbor dengan empat indikator utamanya adalah Nomor Kartu Keluarga, jumlah anggota keluarga, jenis pekerjaan dan penghasilan bulanan. Klasifikasi ini menyediakan data keluarga miskin yang sesuai untuk menerima berbagai bantuan program pemerintah. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara data yang dievaluasi dengan data latihnya. Euclidean distance digunakan sebagai referensi mencari jarak terdekat. Penelitian ini menggunakan 100 contoh data keluarga dengan tingkat akurasi tertinggi yang diperoleh dengan perbandingan 90:10 pada nilai K=5, K=7 dan K=9 adalah sebesar 90%. Selain akurasi, digunakan pengujian Black Box dan User Acceptance Test untuk menguji aplikasi yang telah dirancang. Nilai 75% yang berada dalam kategori sangat bagus diperoleh dari pengujian User Acceptance Test menggunakan quisioner dengan skala likert.

**Kata kunci:** Akurasi, Data Mining, Euclidean distance, Klasifikasi, K-Nearest Neighbor

## Abstract

Poverty is one of the biggest challenges that is still being resolved by all nations including Indonesia. In accordance with the constitutional mandate contained in article 34 of the 1945 Constitution, the government continues to protect and empower the poor through various strategic programs. One of the main problems that often contradicts the intended government program is availability appropriate poor people's data. This causes government programs to not run optimally, requiring verification and re-validation. This research aims to design applications that can classify poor families using The K-Nearest Neighbor algorithm with four main indicators is the Family Card Number, the number family members, types of work and monthly income. This classification provides data on poor families that are suitable for receiving various government program assistance. The working principle of K-Nearest Neighbor is to find the closest distance between the data that is evaluated and the training data. Euclidean distance is used as a reference to find the closest distance. This study uses 100 examples of family data with the highest level obtained with 90:10 at a value of K = 5, K = 7 and K = 9 is 90%. Beside accuracy, Black Box testing and User Acceptance Test are used to test applications that have been designed. A value of 75% included in the excellent category was obtained from the User Acceptance Test using a Likert scale questionnaire.

**Keywords:** Accuracy, Data Mining, Euclidean distance, Classification, K-Nearest Neighbor.

## 1. Pendahuluan

Masalah kemiskinan merupakan hal yang sering dilihat karena menyangkut berbagai macam aspek, seperti hak untuk terpenuhinya pangan, kesehatan, pendidikan dan pekerjaan. Penduduk yang masih berada di bawah garis kemiskinan mencakup mereka yang berpendapatan rendah, tidak berpendapatan tetap atau tidak berpendapatan sama sekali. Jumlah penduduk miskin di suatu wilayah dapat mencerminkan tingkat kemiskinan wilayah tersebut. Besar kecilnya jumlah penduduk miskin sangat dipengaruhi oleh garis kemiskinan, karena penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perbulan di bawah garis kemiskinan. Semakin tinggi garis kemiskinan, semakin banyak penduduk yang tergolong sebagai penduduk miskin. Secara teori kemiskinan dapat dilihat sebagai masalah multi dimensi karena berkaitan dengan ketidak-

mampuan akses secara ekonomi, sosial, budaya, politik dan partisipasi dalam masyarakat. Kemiskinan memiliki arti yang lebih dalam karena berkaitan dengan ketidak-mampuan untuk mencapai aspek di luar pendapatan (*non-income factors*) seperti akses kebutuhan minimum, kesehatan, pendidikan, air bersih.

Salah satu permasalahan yang sering menyebabkan program pemerintah terkait keluarga miskin kurang efektif dijalankan adalah terkait ketersediaan data yang sesuai, data keluarga miskin yang tepat sasaran. Ketersediaan data ini sering tidak dapat diverifikasi dan validasi karena luasnya wilayah dengan jumlah keluarga miskin yang sangat besar dan sering berubah. Status keluarga miskin yang diperoleh sering tidak akurat datanya baik menyangkut jumlah maupun lokasi penyebarannya. Seringkali di suatu wilayah terjadi kekurangan jumlah keluarga penerima bantuan sementara di wilayah lain terjadi over penerima bantuan. Adanya data yang sesuai dapat membantu perencanaan program yang tepat sasaran. Klasifikasi keluarga miskin menggunakan K-NN adalah upaya untuk menyediakan data yang sesuai untuk target program bagi keluarga miskin.

Data *mining* merupakan teknologi yang menggabungkan metoda analisis tradisional dengan algoritma yang canggih untuk memproses data dengan volume besar. Data *mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi didalam *database*. Data *mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat dan tersimpan di dalam database yang besar[1]

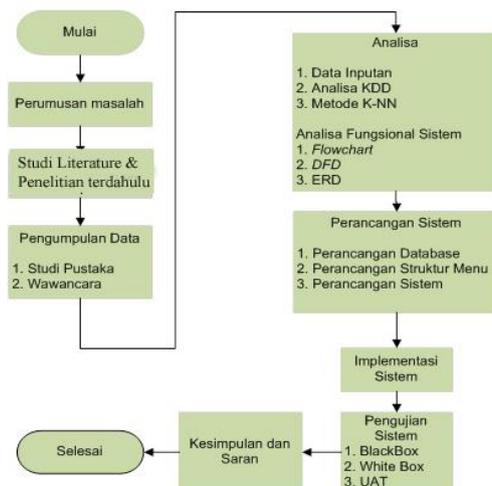
Metode *K-Nearest Neighbor* (K-NN) menjadi salah satu metode yang terkenal, tergabung dalam 10 algoritma top Data Mining[2]. Metode K-NN dapat digunakan untuk penyeleksian masyarakat yang layak menerima dana bantuan keluarga miskin. Metode ini melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Metode K-NN berfungsi untuk mengklasifikasikan atau mengelompokkan kriteria dari data-data yang diuji.

K-NN telah banyak diterapkan dalam berbagai bidang penelitian. Di dalam[3],K-NN digunakan untuk mendeteksi masa ovulasi pada wanita berdasarkan citra pola ferning saliva. Sedangkan [4] menggunakan K-NN untuk mendiagnosis penyakit sapi berbasis android. Sering juga K-NN diperbandingkan dengan metode lainnya atau dikombinasikan dengan metode lainnya untuk suatu tujuan yang spesifik. [5] mencoba membandingkan Metode Naïve Bayes Classifier dan K-NN untuk klasifikasi tugas akhir mahasiswa jurusan teknik informatika.

Beberapa penelitian yang disajikan ini menunjukkan bahwa demikian luasnya area yang dapat diselesaikan dengan K-NN. Karena itu K-NN menjadi salah satu bagian dari algoritma yang paling sering digunakan dalam DM[2].

## 2. Metodologi Penelitian

Metodologi penelitian disajikan dengan mendeskripsikan masalah yang dilengkapi dengan penyajian diagram alur pelaksanaan penelitian. Gambar 1. merupakan tahapan metode penelitian yang digunakan dalam penelitian ini.



Gambar 1. Tahapan Metodologi Penelitian

### A. Perumusan Masalah

Tahap ini merupakan awal dari tahapan metodologi penelitian dengan mencari dan mempelajari tentang permasalahan dalam penelitian. Kemudian dilanjutkan dengan pencarian solusi dalam mengatasi permasalahan tersebut. Dalam tahapan ini juga ditentukan ruang lingkup dan latar belakang dari topik penelitian.

### B. Studi Literature dan Penelitian Terdahulu

Studi literatur adalah mencari referensi teori yang relevan dengan kasus atau permasalahan yang ditemukan. Data Mining merupakan landasan utama teori yang digunakan dalam penelitian ini termasuk di dalamnya adalah algoritma K-NN untuk proses klasifikasi.

#### 1. Data Mining(DM)

Data *mining* merupakan teknologi yang menggabungkan metoda analisis tradisional dengan algoritma yang canggih untuk memproses data dengan volume besar. DM adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam *database*. DM merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat dan tersimpan di dalam database besar[1].

#### 2. K-Nearest Neighbor

*K-Nearest Neighbor*(K-NN) merupakan suatu pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama[6], yaitu berdasarkan pada pencocokan bobot dari jumlah fitur yang ada. K-NN merupakan salah satu metode yang digunakan dalam pengklasifikasian. Prinsip kerja K-NN adalah mencari jarak terdekat antara data yang dievaluasi dengan K-NN terdekatnya dalam data pelatihan. Berikut rumus pencarian jarak menggunakan rumus jarak *Euclidian*[7].

1. Menentukan Nilai N buah data awal sebagai basis pengetahuan awal sistem.
2. Menentukan Nilai K (tetangga) terdekat.
3. Mempersiapkan data training berupa nilai kriteria suatu data baru yang belum diketahui statusnya.
4. Menentukan Status setiap data training berdasarkan suatu aturan tertentu untuk menghasilkan basis pengetahuan sistem.
5. Menghitung jarak setiap sampel data training terhadap data yang diuji (data uji) berdasarkan persamaan:

$$d(x, y) = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

d = jarak antara titik pada *data training* x dan titik data testing y yang diklasifikasi, dimana  
 $x = x_1, x_2, \dots, x_i$  dan  $y = y_1, y_2, \dots, y_i$   
 i = merepresentasikan nilai atribut

$n$  = merupakan dimensi atribut.

6. Menetapkan status data uji berdasarkan nilai rata-rata  $K$  buah sampel data training terdekat.

Tahapan selanjutnya adalah tahapan untuk melakukan pencarian informasi mengenai penelitian-penelitian yang telah dilakukan sebelumnya yang berkaitan dengan data *mining* menggunakan metode K-NN. Tahap ini meliputi pencarian referensi yang berupa jurnal-jurnal (jurnal internasional maupun nasional), buku-buku yang berupa *e-book*, dan referensi-referensi lainnya yang berkaitan dengan penelitian ini.

### C. Pengumpulan Data

Pada tahap ini terjadi proses pengumpulan data yang diperlukan untuk merancang dan membangun sistem untuk menentukan penerimaan dana bantuan. Data yang dikumpulkan dalam tahap ini bersumber dari sejumlah referensi dan wawancara yang dilakukan kepada pihak terkait.

### D. Analisa

Pada tahapan analisa ini penelitian dilakukan menggunakan berkas data yang berhubungan dengan pemberian dana bantuan yang didapat dan selanjutnya dilakukan analisa. Analisa dibagi menjadi dua tahap yaitu analisa proses dan analisa sistem. Tahapan analisa yang tersaji dalam pada gambar 1 dapat diuraikan sebagai berikut :

#### a. Data Selection

*Data Selection* adalah data variabel mentah kemudian diseleksi menjadi beberapa variabel. Variabel yang digunakan ada 4, yaitu Nomor KK, Pekerjaan Kepala Keluarga, Penghasilan Perbulan, Jumlah Anggota Keluarga. Untuk setiap variabel diberi inialisasi sebagaimana tersaji dalam tabel 1 berikut ini:

Tabel 1. Inialisasi Variabel

Variabel	Keterangan
A1	Nomor KK
A2	Jumlah Anggota Keluarga
A3	Pekerjaan Kepala Keluarga
A4	Penghasilan Perbulan

Setelah data diinputkan, maka dilakukan proses pembagian data untuk pelatihan dan pengujian. Pembagian data yang dilakukan untuk proses pelatihan dan pengujian dengan menggunakan metode K-NN dengan membagi data latih (*training*) dan data uji (*testing*). Adapun jumlah keseluruhan data yang digunakan pada penelitian ini adalah 100 data kepala keluarga.

#### b. Transformasi

Transformasi data adalah untuk mengubah skala pengukuran data asli menjadi bentuk lain sehingga data dapat digunakan. Tabel 2 menunjukkan data yang mengalami proses transformasi.

Tabel 2. Data Transformasi

No	A1	A2	A3	A4
1	1407101201130000	5	1	1
2	1407102310100400	3	1	1
3	1407100312070150	10	1	1
4	1407101611070600	8	1	1
5	1407102605870000	5	1	1
6	1407101611070280	3	2	2
7	1407102411740000	3	2	2
8	140710040908008	5	2	2
9	1407102004800000	6	3	3
10	1407102110100030	5	3	3
11	1407101201130000	2	3	3

...	...	...	...	...
100	1407100712120000	9	5	4

Nilai bobot yang ditentukan untuk mengatur tingkatan pekerjaan disajikan dalam tabel 3.

Tabel 3. Data Pekerjaan

No.	Data Pekerjaan	Nilai Bobot
1.	Buruh Harian	1
2.	Guru Honorer	2
3.	Bidan	3
4.	Petani	4
5.	Wiraswasta	5
6.	Pegawai Negeri Sipil (PNS)	6

Sedangkan nilai bobot penghasilan yang ditentukan disajikan dalam tabel 4.

Tabel 4. Data Penghasilan

No.	Data Penghasilan	Nilai Bobot
1.	Rp. 0 < Penghasilan ≤ 500.000	1
2.	Rp.500.000 < Penghasilan ≤ 1.000.000	2
3.	Rp. 1.000.000 < Penghasilan ≤ 1.500.000	3
4.	Rp. 1.500.000 < Penghasilan ≤ 3.500.000	4
5.	Rp. 3.500.000 < Penghasilan ≤ 4.000.000	5
6.	Rp. 4.000.000 < Penghasilan ≤ 5.000.000	6

Kelas yang digunakan hanya dua yaitu Ya (berhak menerima bantuan) dan Tidak (tidak berhak menerima bantuan). Tabel 5 menunjukkan kelas yang digunakan.

Tabel 5. Data Kelas

No.	Kelas
1.	Ya
2.	Tidak

### c. Data Mining

Data Mining merupakan tahapan inti dari penelitian ini, pada tahapan ini dilakukan *klasifikasi* data penerima dana bantuan keluarga miskin dengan menggunakan algoritma *K-Nearest Neighbor*. Pada penelitian ini perhitungan jarak yang digunakan adalah *Euclidean Distance* yang merupakan perhitungan jarak dari 2 titik dalam *Euclidean Space*. Setelah didapatkan jarak maka ditentukan jarak paling minimal.

### d. Interpretasi

Tahapan ini merupakan proses dari suatu model *klasifikasi*. Hasil yang didapatkan kemudian ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

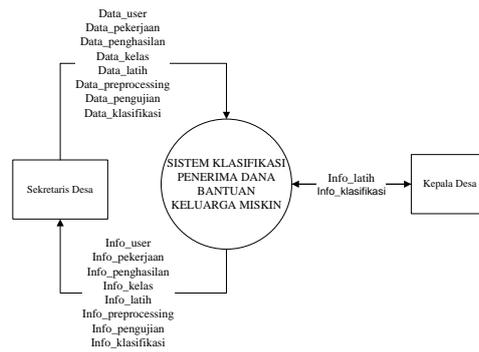
### e. Perancangan

Setelah dilakukan analisa, selanjutnya dilakukan perancangan aplikasi yang dibuat berdasarkan analisa sebelumnya.

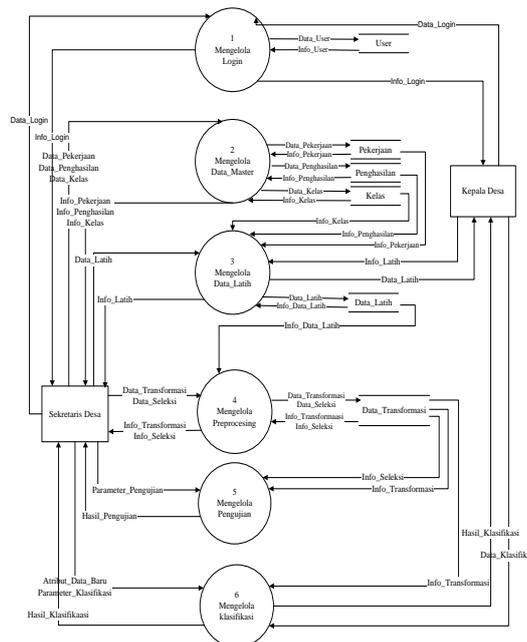
*Context Diagram* merupakan analisa fungsional sistem secara keseluruhan yang menggambarkan kegiatan yang dilakukan oleh sistem sekretaris desa serta kepala desa yang berperan menggunakannya. Gambar 2 merupakan *context diagram* sistem klasifikasi penerima dana bantuan keluarga miskin.

Data *flow diagram* (DFD) merupakan bagan yang menggambarkan pergerakan aliran data dari mana asal dan tujuan data tersebut, data tersebut mengalir melalui proses-proses yang ada dan *stakeholder* yang berinteraksi dengan sistem. Dari DFD ini dapat ditemukan data-data apa saja yang disimpan kedalam *database* sistem.

DFD *Level 1* sistem klasifikasi penerima dana bantuan keluarga miskin yang menampilkan *Entitas*, *Proses*, *Data store* dan aliran data yang menunjukkan alur jalannya data di dalam sistem. Gambar 3 berikut menunjukkan DFD level 1.



Gambar 2. Context Diagram Klasifikasi Penerima Dana Bantuan Keluarga Miskin



Gambar 3. DFD Level 1 Sistem Klasifikasi Penerima Dana Bantuan Keluarga Miskin

### 3. Hasil Dan Analisis

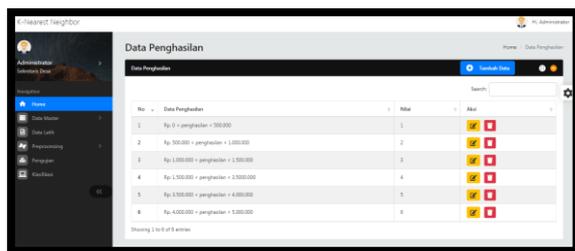
Perancangan konseptual yang telah dilakukan kemudian diterapkan dalam bentuk suatu aplikasi beserta menu pendukungnya. Halaman menu utama adalah halaman paling awal ketika pengguna mengakses aplikasi Klasifikasi penerima dana bantuan keluarga miskin menggunakan metode K-NN. Tampilan Halaman utama dapat dilihat dalam gambar 4.

Gambar 4. Halaman Menu Utama

Halaman data *master* data pekerjaan merupakan halaman dimana data jenis pekerjaan ditampilkan. Tampilan data pekerjaan dapat dilihat dalam Gambar 5 berikut menunjukkan Halaman data pekerjaan.

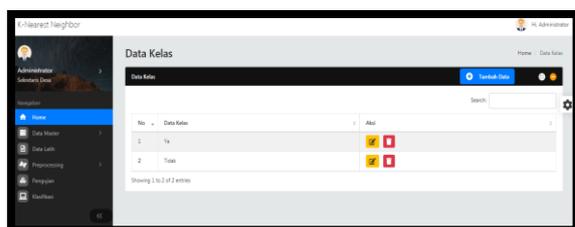
Gambar 5. Halaman Data pekerjaan

Halaman data *master* data penghasilan merupakan halaman dimana data jenis penghasilan ditampilkan. Gambar 6 berikut menampilkan halaman data penghasilan.



Gambar 6. Halaman Data penghasilan

Halaman data *master* data kelas merupakan halaman dimana data kelas. Tampilan data kelas dapat dilihat dalam gambar 7.



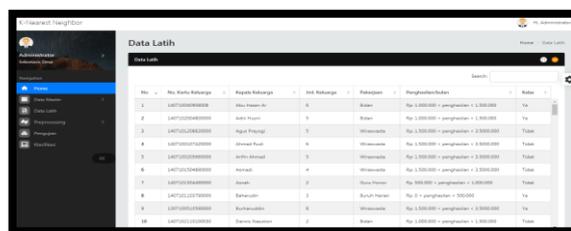
Gambar 7. Halaman Data Kelas

Halaman data *master* data user merupakan halaman dimana data *user*. Tampilan data *user* dapat dilihat dalam gambar 8.



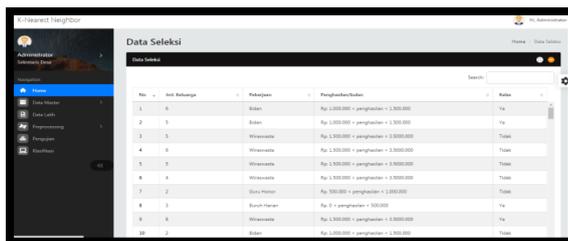
Gambar 8. Halaman Menu User

Halaman data *latih* merupakan halaman dimana data latihan dilakukan. Tampilan data latihan dapat dilihat dalam gambar 9.



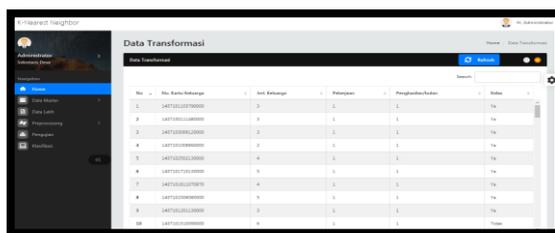
Gambar 9. Halaman Data Latih

Halaman data *preprocessing* merupakan halaman dimana data seleksi dilakukan. Tampilan data seleksi dapat dilihat dalam gambar 10.



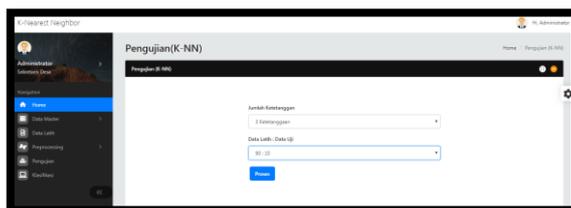
Gambar 10. Halaman seleksi

Halaman data *transformasi* merupakan halaman dimana data *transformasi* dilakukan. Tampilan data *transformasi* dapat dilihat dalam gambar 11.



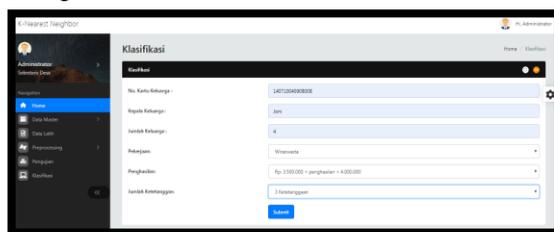
Gambar 11. Halaman Transformasi

Halaman data pengujian merupakan halaman dimana data diuji. Tampilan data pengujian dapat dilihat dalam gambar 12.



Gambar 12. Halaman Pengujian

Halaman data klasifikasi merupakan halaman dimana data diklasifikasikan. Tampilan data klasifikasi dapat dilihat dalam gambar 13.



Gambar 13. Halaman klasifikasi

Aplikasi yang telah selesai dikembangkan kemudian dilakukan pengujian. Proses pengujian bertujuan untuk melihat kesesuaian dengan hasil analisa perancangan konseptualnya. Pengujian yang dilakukan terdiri dari pengujian Black Box, Pengujian User Acceptance Test dan Pengujian Akurasi K-NN.

#### A. Pengujian *Black Box*

Setelah tahap implementasi telah selesai dilaksanakan, maka dilanjutkan dengan pengujian aplikasi dari implementasi yang telah dibuat. Hasil pengujian memperlihatkan aplikasi yang dikembangkan telah sesuai dengan apa yang diharapkan. Pengujian pertama dilakukan pada Halaman Menu Login, kemudian secara berurutan dilakukan pengujian pada Halaman Menu Data

Master, Halaman Menu Data Latih, Halaman Menu Preprocessing, Halaman Menu Pengujian dan Halaman Menu Klasifikasi.

### B. Pengujian *User Acceptance Test* (UAT)

Pengujian UAT dalam sistem ini ditujukan kepada 1 administrator dan 1 user, agar dapat mengetahui apakah sistem sudah sesuai dengan kebutuhan yang diharapkan oleh pengguna. Dengan menggunakan quisioner dalam skala likert diketahui bahwa nilai rata rata pengujian UAT adalah 75%.

### C. Pengujian Akurasi K-NN

Pada penelitian ini dibangun sebuah sistem untuk mengklasifikasikan penerima dana bantuan keluarga miskin menggunakan metode K-NN. Data yang digunakan sebagai inputan adalah data kepala keluarga tahun 2017-2018. Dengan data yang berjumlah 100 data kepala keluarga (KK) dengan sampel 4 data latih dan 1 data uji dengan nilai  $K = 3$ . Selanjutnya data masukan tersebut diproses untuk mengetahui mana yang menerima dan tidak menerima dana bantuan. Terdapat pada tabel 6 di bawah ini:

Tabel 6. Proses K-NN

Latih	AT 1	AT 2	AT 3
Latih 1	5	1	1
Latih 2	3	1	1
Latih 3	10	1	1
Latih 4	8	1	1
Uji 1	3	1	1

Tabel 6 di atas merupakan data yang diproses menggunakan metode K-NN. Proses penghitungan jarak menggunakan rumus persamaan 1 sebagai berikut:

$$\begin{aligned} \text{latih 1- uji 1} &= \sqrt{(5 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} \\ &= \sqrt{4 + 0 + 0} \\ &= \sqrt{4} \\ &= 2 \end{aligned}$$

$$\begin{aligned} \text{latih 2- uji 1} &= \sqrt{(3 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} \\ &= \sqrt{0 + 0 + 0} \\ &= \sqrt{0} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{latih 3- uji 1} &= \sqrt{(10 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} \\ &= \sqrt{49 + 0 + 0} \\ &= \sqrt{49} \\ &= 7 \end{aligned}$$

$$\begin{aligned} \text{latih 4- uji 1} &= \sqrt{(8 - 3)^2 + (1 - 1)^2 + (1 - 1)^2} \\ &= \sqrt{25 + 0 + 0} \\ &= \sqrt{25} \\ &= 5 \end{aligned}$$

Hasil dari penjumlahan perhitungan manual dengan contoh 4 data latih dan 1 data uji dengan nilai ketetangaan adalah 3 ditunjukkan pada tabel 7.

Pengujian Akurasi K-NN dilakukan dengan beberapa pembagian data latih dan data uji dimana pemilihan data uji dan data latih dilakukan secara acak oleh sistem

Tabel 7. Jarak Data Latih Dan Uji

Data latih dan uji	Hasil
Latih 1- uji 1	2
Latih 2- uji 1	0
Latih 3- uji 1	7
Latih 4- uji 1	5

Perbandingan data latih, data uji dan jumlah ketetangaan dapat dilihat pada tabel 8.

Tabel 8. Hasil Akurasi Pengujian Sistem

A	B	C	D	E	F
K=3	70 %	85 %	76 %	70 %	76 %
K=5	90 %	65 %	66 %	70 %	72 %
K=7	90 %	65 %	66 %	70 %	72 %
K=9	90 %	90 %	66 %	67 %	72 %
K=11	80 %	65 %	66 %	67 %	72 %

Keterangan:

A:Ketetapan

B:Perbandingan Data latih : uji (90:10)

C:Perbandingan Data latih : uji (80:20)

D:Perbandingan Data latih : uji (70:30)

E:Perbandingan Data latih : uji (60:40)

F:Perbandingan Data latih : uji (50:50)

#### 4. Kesimpulan

Aplikasi yang telah dirancang diuji menggunakan metode Black Box dan UAT. Hasil pengujian UAT yang dilakukan menggunakan kuisioner dengan skala likert mendapatkan nilai 75% yang berada dalam kategori sangat bagus. Pengujian K-NN untuk mengklasifikasikan keluarga miskin pada 100 contoh data diperoleh nilai akurasi tertinggi dengan perbandingan 90:10 pada nilai K=5, K=7 dan K=9 sebesar 90%.

#### Daftar Pustaka

- [1] Turban, Efraim, et al. 2005. Decision Support Systems and Intelligent Systems 7th Ed. New Jersey : Pearson Education.
- [2] Wu, X, etc. Top 10 algorithms in data mining, 2008. J.Knowledge Inf Syst 14:1– 37. Springer-Verlag.
- [3] Subairi, Rahmadwati, dan Erni Yudaningtyas, Implementasi Metode k Nearest Neighbor pada Pengenalan Pola Tekstur Citra Saliva untuk Deteksi Ovulasi, Jurnal EECCIS Vol. 12, No. 1, April 2018, hlm 9 -14.
- [4] Idham Triatmaja, Nurul Hidayat, Moch. Cholil Mahfud, Sistem Diagnosis Penyakit Sapi Menggunakan Metode Neighbor Weighted K-Nearest Neighbor Berbasis Android, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol. 2, No. 8, Agustus 2018, hlm. 2944-2946.
- [5] Yusra, Dhita Olivita, Yelfi Vitriani, Perbandingan Klasifikasi Tugas Akhir Mahasiswa Jurusan Teknik Informatika Menggunakan Metode Naïve Bayes Classifier dan K-Nearest Neighbor, Jurnal Sains, Teknologi dan Industri, Vol. 14, No. 1, Desember 2016, hlm 79 – 85.
- [6] Kusriani & Emha, T. (2009). Algoritma Data Mining. Yogyakarta: Andi Offset.
- [7] Agusta, Y. (2007). K-Means – Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika Volume.3*, Hal 47-60.