

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1. Klasifikasi

Klasifikasi adalah suatu proses dalam pengelompokan data. Klasifikasi dapat digunakan untuk menemukan fungsi pembeda kelas data. Klasifikasi merupakan suatu teknik dengan melihat pada kriteria dan atribut dari kelompok yang telah didefinisikan pada data baru dengan memanipulasi hasil data yang ada untuk diberikan sejumlah aturan (Simanjuntak dkk, 2013).

Ada beberapa algoritma dalam pengklasifikasian yang sudah dikembangkan oleh penelitian diantaranya adalah *K-Nearest Neighbor* (K-NN), *Modified K-Nearest Neighbor* (MK-NN), *Decision Tree Classifier*, *Rule-Based Classifier*, *Neural-Network*, *Support Vector Machine*, *Naive Bayes Classifier*, dan *Artificial Neural Network* (ANN).

2.2. Transformasi Data

Transformasi pada penelitian ini digunakan mempersempit *range* data latih. Transformasi yang digunakan pada penelitian ini menggunakan transformasi min-max. Pada proses ini dihasilkan nilai dari data latih yang berkisar antara 0 sampai 1. Adapun proses transformasi yang dilakukan dapat dilihat pada persamaan 2.1 berikut :

$$v^i = \frac{v - \min_a}{\max_a - \min_a} (\text{new_max}_a - \text{new_min}_a) + \text{new_min}_a \quad (2.1)$$

Dimana :

v^i : Data baru setelah transformasi

v : Data sebelum transformasi

new_max_a : Batas nilai max baru adalah 1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

new_min_a : Batas nilai min baru adalah 0

max_a : Nilai maximum pada kolom

min_a : Nilai minimum pada kolom.

2.3. Metode K-NN

Metode K-Nearest Neighbor (K-NN) adalah sebuah metode yang dilakukan untuk mengklasifikasi objek berdasarkan data pembelajaran yang jaraknya paling dekat dari objek tersebut. *Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada (Ndaumanu dkk 2014).

Metode ini sama dengan teknik *clustering* yaitu dengan mengelompokan data baru berdasarkan jarak dari data baru itu ke beberapa data tetangga terdekat.

Jarak *Euclidean*

Untuk menjelaskan jarak antara dua titik yaitu titik pada data *training* (x) dan titik pada data *testing* (y) maka digunakan rumus *Euclidean* berikut (Krisandi dkk 2013) :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{2.2}$$

Keterangan :

- yi= sampel data
- xi= data uji/testing
- d= jarak
- i= variabel data
- n= dimensi

Dengan d adalah jarak antara titik pada data *training* (x) dan titik data *testing* (y) yang akan diklasifikasikan, dimana $x = x_1, x_2, \dots, x_i$ dan $y = y_1, y_2, \dots, y_i$ dan I representasi nilai atribut serta n merupakan dimensi atribut.

2.4. Metode MKNN

Modified K-Nearest Neighbor (MKNN) adalah menempatkan label kelas data sesuai dengan k validasi titik data yang sudah didapat dari perhitungan K-NN.

Dalam proses algoritma MKNN terdapat proses-proses tambahan dari metode K-NN, yaitu *validasi data training* dan *weight voting* (Zainuddin dkk, 2013).

2.4.1. Validitas Data Latih

Pada proses ini setiap data *training* harus melalui validasi, validasi setiap data tergantung dari setiap tetangganya. Setelah perhitungan validitas tiap data maka nilai validitas tersebut akan digunakan sebagai informasi lebih mengenai data *training*. Validasi digunakan untuk menghitung jumlah titik dengan label yang sama untuk data *training*.

Persamaan dari validitas setiap data latih adalah sebagai berikut :

$$validitas(x) = \frac{1}{z} \sum_{i=1}^z s(label(x), label(N_i(x))) \quad (2.3)$$

Keterangan :

- z = jumlah titik terdekat
- label = kelas x
- N_i = label kelas titik terdekat x

Fungsi S digunakan untuk menghitung kesamaan titik a data ke- b tetangga terdekat. Untuk mendefinisikan fungsi S dapat dijelaskan dengan persamaan berikut :

$$s(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \quad (2.4)$$

keterangan :

- a = kelas a pada data training
- b = kelas lain selain a pada data training

2.4.2. Weight Voting

Pada proses ini, untuk bobot masing-masing tetangga dihitung dengan menggunakan $1/(d_e + 0,5)$. Validitas dari tiap data untuk data training dikalikan dengan *bobot* berdasarkan pada jarak *euclidian*. Sehingga metode MKNN, didapat persamaan *weight voting* tiap tetangga sebagai berikut :

$$w(i) = validitas(i) \times \frac{1}{d_e + 0,5} \quad (2.5)$$

Keterangan :

$w(i)$ = perhitungan *weight voting*

$validitas(i)$ = nilai validitas

d_e = jarak *Euclidean*

2.5. Perhitungan Akurasi

Untuk melihat akurasi pada pengujian klasifikasi, pada penelitian ini menggunakan *confusion matrix*. Menurut (Gurunesu, 2011), Jika data yang akan diuji memiliki *class* positif dan negatif, maka dapat dibuatkan tabel seperti tabel 2.1 berikut:

Tabel 2.1 Tabel *Confusion matrix 2 class*

Classification Observed Class	Predicted Class		
		Class = Yes	Class = No
	Class = Yes	(True Positive - TP)	(False Negative - FN)
	Class = No	(False Positive - FP)	(True Negative - TN)

Keterangan:

True Positive – TP = Jumlah data kelas benar yang diklasifikasikan sebagai kelas benar

True Negative – TN = Jumlah data kelas salah yang diklasifikasikan sebagai kelas salah

False Positive – FP = Jumlah data kelas salah yang diklasifikasikan sebagai kelas benar

False Negative – FN = Jumlah data kelas benar yang diklasifikasikan sebagai kelas salah

Berdasarkan tabel 2.1 diatas, dapat dilakukan perhitungan akurasi seperti persamaan 2.6 berikut:

$$Akurasi = \frac{TP+TN}{TP+FN+FP+TN} \times 100\% \quad (2.6)$$

2.6. User Acceptance Test (UAT)

Menurut (Betha, 2006) yang dikutip melalui (Mutiara dkk, 2014), *User Acceptance Test* adalah proses pengujian oleh *user* dan menghasilkan dokumen untuk dijadikan bukti bahwa aplikasi yang dikembangkan dapat diterima *user* dan hasil pengujiannya dianggap memenuhi kebutuhan pengguna. *User Acceptance*

Test menggunakan angket atau kuisoner yang berisi pertanyaan seputar sistem yang telah dibangun yang disebarakan kepada responden untuk menemukan *feedback* dari *user*. Pertanyaan dalam angket berbentuk objektif dan disesuaikan dengan tujuan dari pengujian menggunakan UAT dimana para responden dapat memilih salah satu dari beberapa alternatif jawaban yang telah diberi bobot atau skoring.

Setelah pertanyaan disusun langkah selanjutnya adalah penentuan skoring ilmiah yang secara umum berpedoman pada aturan *Likert* dan *Gutman*. Kedua metode ini memenuhi kaidah ilmiah dalam penentuan dan penilaian skoring suatu instrumen penelitian. Pada penelitian ini akan digunakan aturan *likert* dalam penentuan skoring.

Skala Likert

Skala *Likert* adalah suatu skala *psikometrik* yang digunakan dalam kuesioner dan merupakan salah satu teknik yang dapat digunakan dalam evaluasi suatu program atau kebijakan perencanaan. Bobot *likert* ditentukan untuk memberi nilai pada setiap alternatif jawaban dimana pada umumnya bobot *likert* ini memiliki nilai maksimal sebanyak jumlah alternatif jawaban dan paling rendah adalah 1. Rumus penilaian dengan skala likert yaitu :

1. Menentukan jumlah kategori
2. Penentuan Total Skor

Untuk mendapatkan rangkuman hasil penilaian dapat dilakukan dengan rumus sebagai berikut :

$$T \times P_n \tag{2.7}$$

Keterangan :

- T = Frekuensi jawaban yang dipilih
- P_n = Bobot skor likert

3. Interval

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Untuk mendapatkan interval dan interpretasi persen dari kategori digunakan rumus sebagai berikut :

$$Interval (I) = \frac{100\%}{Jumlah\ Kategori} \quad (2.8)$$

4. Interpretasi Skor Perhitungan

Untuk mendapatkan skor perhitungan interpretasi harus diketahui skor tertinggi dan skor terendah dengan rumus sebagai berikut :

$$X = \text{Skor Terendah Likert} \times \text{Jumlah Pertanyaan}$$

$$Y = \text{Skor Tertinggi Likert} \times \text{Jumlah Pertanyaan}$$

Selanjutnya ditentukan indeks persen untuk mengetahui hasil akhir dengan rumus sebagai berikut :

$$Indeks \% = \frac{Total\ Skor}{Y * jumlah\ responden} \times 100 \quad (2.9)$$

2.7. Pengertian Tempat Tinggal

Tempat tinggal adalah tempat seseorang yang selalu hadir dalam hubungannya dengan pelaksanaan hak dan memenuhi kewajiban. Bukan hanya manusia perorangan yang memiliki tempat tinggal, orang-orang berkelompok juga memiliki tempat tinggal. Tempat tinggal dibagi kedalam beberapa bentuk, yaitu: perumahan, rumah tunggal (bulatan), bahkan kos-kosan.

Tempat tinggal merupakan suatu kebutuhan utama (*primer*) dalam kehidupan manusia. Tempat tinggal itu sendiri berguna sebagai suatu wadah perlindungan terhadap gangguan lingkungan alam dan binatang (Zubaidi, 2009).

Syarat perumahan yang baik memiliki kriteria seperti (Kumalasari dkk, 2014) :

1. Harga sesuai dengan kebutuhan
2. Lokasi perumahan tidak banjir
3. Model rumah sesuai keinginan
4. Sumber air bersih
5. Fasilitas

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Untuk menentukan lokasi perumahan yang baik perlu diperhatikan hal-hal berikut ini (Eko Budihardjo 1998, dikutip oleh Alfikri 2007) :

1. Ditinjau dari segi teknis pelaksanaannya
 - a. Mudah mengerjakannya dalam arti tidak banyak pekerjaan *cut* dan *fill*, pembongkaran dan tonggak kayu
 - b. Bukan daerah banjir, bukan daerah gempa, angin ribut dan daerah rayap
 - c. Mudah dicapai tanpa hambatan yang berarti
 - d. Tanahnya baik sehingga konstruksi bangunan yang ada dapat direncanakan dengan sistem yang semurah mungkin
 - e. Mudah mendapat sumber air bersih, listrik, pembuangan air limbah dan lain-lain
 - f. Mudah mendapat bahan-bahan bangunan
 - g. Mudah mendapatkan tenaga-tenaga pekerja dan lain
2. Dilihat dari segi tata guna tanah
 - a. Bukan daerah persawahan
 - b. Bukan daerah kebun
 - c. Bukan daerah usaha seperti pertokoan, perkantoran, hotel, pabrik dan lain-lain
3. Dari segi kemudahan dan kesehatan
 - a. Lokasi sebaiknya jauh dari lokasi pabrik yang dapat mendatangkan polusi misalnya debu pabrik, buangan sampah-sampah dan limbah industri
 - b. Lokasi sebaiknya tidak terlalu terganggu dengan kebisingan
 - c. Lokasi sebaiknya dipilih yang udaranya masih sehat
 - d. Lokasi sebaiknya dipilih mudah mendapatkan air minum, sekolah, pasar, puskesmas dan lain-lain kebutuhan keluarga
 - e. Lokasi sebaiknya mudah dicapai dari tempat kerja penghuni

Area-area yang harus dihindari sebagai lokasi rumah tinggal (Imelda 2003, dikutip oleh Alfikri 2007) :

1. Area Pabrik dan Industri
Pabrik mengeluarkan gas-gas yang sering kali beracun dan berbahaya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Selain itu buangan limbahnya juga sering mencemari air, tanah dan lingkungan.

2. Kawat tetangga tinggi

Jangan sekali-kali memilih lokasi yang dekat dengan aliran listrik tegangan tinggi, antena radio maupun menara besi. Tegangan tinggi memiliki radiasi yang berdampak negatif terhadap kesehatan manusia.

3. Bandar udara

Bandar udara memiliki kebisingan yang sangat tinggi, yang dapat mengakibatkan sulit berkonsentrasi, serta tidur tidak nyenyak.

4. Jalan lalu lintas regional dan jalan tol

Selain menimbulkan kebisingan jalan raya juga membawa debu.

5. Rel kereta api

Selain bisingnya yang mengganggu pendengaran, kereta api juga dapat mengganggu kesehatan karena memiliki efek elektromagnetik.

6. Pengolahan limbah industri dan pembuangan sampah

Sampah membawa penyakit, tidak hanya melalui serangga seperti lalat tetapi juga gas-gas beracun oleh sampah yang membusuk.

Menurut (Peraturan Daerah No.14 Tahun 2000 dan Imelda Akmal 2003, dikutip oleh Alfikri 2007) terdapat 13 kriteria untuk memilih perumahan yang sesuai dengan kebutuhan konsumen. Adapun 13 kriteria tersebut adalah sebagai berikut :

1. Jarak perumahan terhadap sarana pendidikan
2. Jarak perumahan terhadap sarana kesehatan
3. Jarak perumahan terhadap sarana ibadah
4. Jarak perumahan terhadap jalan utama
5. Jarak perumahan terhadap sarana perdagangan
6. Jarak perumahan terhadap sarana pemerintahan
7. Jarak perumahan terhadap sarana sosial budaya
8. Jarak perumahan terhadap sarana olahraga
9. Jarak perumahan terhadap sarana pelayanan umum

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

10. Jarak perumahan terhadap jalan lintas antar daerah
11. Jarak perumahan terhadap bandara
12. Jarak perumahan terhadap pabrik industri
13. Jarak perumahan terhadap jaringan listrik tegangan tinggal

Pada penelitian ini menggunakan 6 kriteria jarak perumahan yang sesuai dengan perumahan di Pekanbaru berdasarkan kebutuhan konsumen. Adapun 6 kriteria tersebut adalah sebagai berikut :

1. Sarana pendidikan

Sarana pendidikan yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap sarana pendidikan yaitu SD, SMP dan SMA.

2. Sarana kesehatan

Sarana kesehatan yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap sarana kesehatan yaitu Klinik, Bidan atau Rumah Sakit.

3. Sarana ibadah

Sarana ibadah yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap sarana ibadah yaitu Mesjid dan Gereja.

4. Sarana perdagangan

Sarana perdagangan yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap sarana perdagangan yaitu Pasar Tradisional.

5. Terhadap jalan umum

Terhadap jalan umum yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap jalan umum.

6. Terhadap jalan lintas antar daerah

Terhadap jalan lintas antar daerah yang menjadi kriteria pada penelitian ini adalah jarak tempuh terdekat perumahan terhadap jalan lintas antar daerah.

2.8. Penelitian Terkait

Adapun penelitian-penelitian yang terkait dengan penelitian yang akan dilakukan adalah sebagai berikut :

2.8.1. Tempat Tinggal

Adapun beberapa penelitian mengenai Tempat Tinggal dapat dilihat pada Tabel 2.2 berikut:

Tabel 2.2 Penelitian terkait tempat tinggal

No	Nama Peneliti	Judul Penelitian	Tahun Penelitian	Hasil yang Didapatkan
1	Kuncoro, Arsito Ari	Sistem pendukung Keputusan penentuan rumah/tanah multi kriteria kepada calon konsumen dengan metode <i>fuzzy logic</i>	2013	Pada penelitian ini dilakukan pengujian terhadap <i>software</i> kepada calon konsumen properti dengan membagi berdasarkan variable, untuk membantu calon konsumen dalam pengambilan keputusan pemilihan properti multi kriteria menggunakan metode <i>fuzzy logic</i>
2	I wayan Supriana	Sistem pendukung keputusan pemilihan tempat kos dengan metode pembobotan	2012	Pada penelitian ini dilakukan analisis perhitungan kriteria pada koresponden terpilih dengan menentukan kriteria prioritas hingga pemberian nilai kriteria menggunakan metode pembobotan guna membantu dalam pengambilan keputusan untuk memilih tempat tinggal berdasarkan kriteria prioritas yang paling sesuai.

2.8.2. Metode *K-Nearest Neighbor* (KNN)

Adapun beberapa penelitian mengenai metode *K-Nearest Neighbor* (K-NN) dapat dilihat pada Tabel 2.3 berikut:

Tabel 2.3 Penelitian terkait *K-Nearest Neighbor* (K-NN)

No	Nama Peneliti	Judul Penelitian	Tahun Penelitian	Hasil yang Didapatkan
1	Ricky Imanuel Ndaumanu, dkk	Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa dengan Metode <i>K-Nearest Neighbor</i>	2014	Pada penelitian ini dilakukan pengujian terhadap prediksi pengunduran diri mahasiswa menggunakan sistem pendukung keputusan dan nilainya dibandingkan secara manual menggunakan KNN dengan 4 variabel yaitu: IPK, pekerjaan orang tua, jurusan dan semester. Dalam penelitian ini menghasilkan kesesuaian sebesar 79%.
2	Nobertus Krisandi, dkk	Algoritma k-nearest Neighbor dalam klasifikasi data hasil produksi kelapa sawit pada PT. MINAMAS kecamatan Parindu	2013	Pada penelitian ini diaplikasikan algoritma KNN untuk mengklasifikasikan data hasil produksi kelapa sawit, yang menghasilkan 6 kelas klasifikasi data dari 50 sampel hasil produksi kelapa sawit tani dengan ketetapan nilai k=7 yaitu sebesar 34%.

2.8.3. Metode Modified K-Nearest Neighbor (MKNN)

Adapun beberapa penelitian mengenai metode *Modified K-Nearest Neighbor* (MKNN) dapat dilihat pada Tabel 2.4 berikut:

Tabel 2.4 Penelitian terkait *Modified K-Nearest Neighbor* (MKNN)

No	Nama Peneliti	Judul Penelitian	Tahun Penelitian	Hasil yang Didapatkan
1	Andi Nasution	Penerapan Algoritma <i>Modified K-Nearest Neighbor</i> (MKNN) untuk pengklasifikasian penyakit <i>Attention Deficit Hiperactive Disorde</i> (ADHD) pada Anak	2015	Pada penelitian ini dilakukan pengklasifikasian penyakit ADHD dilihat dari gejala yang ada kemudian dibuat dalam bentuk kuisioner dengan menerapkan metode MKNN untuk menghasilkan apakah seorang anak berpotensi penyakit ADHD atau tidak berpotensi penyakit ADHD.
2	Siti Mutrofin, dkk	Optimasi Teknik Klasifikasi		Pada penelitian ini dari hasil algoritma genetika yang dilakukan pada MKNN untuk

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama Peneliti	Judul Penelitian	Tahun Penelitian	Hasil yang Didapatkan
		Modified K-Nearest Neighbor Menggunakan Algoritma Genetika.		memberikan optimasi nilai K pad KNN. Hasil kebenaran tingak evaluasi didasarkan pada nilai akurasi baik menggunakan algoritam KNN, MKNN maupun GMKNN menggunakan data UCI mechine learning.

