

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR
DALAM PENENTUAN KLASIFIKASI TINGKAT
KEDALAMAN KEMISKINAN PROVINSI JAWA TIMUR**

SKRIPSI

**OLEH
CLARIZA ADELINA RACHMA
NIM. 18610077**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR
DALAM PENENTUAN KLASIFIKASI TINGKAT
KEDALAMAN KEMISKINAN PROVINSI JAWA TIMUR**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Clariza Adelina Rachma
NIM. 18610077**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR
DALAM PENENTUAN KLASIFIKASI TINGKAT
KEDALAMAN KEMISKINAN PROVINSI JAWA TIMUR**

SKRIPSI

**Oleh
Clariza Adelina Rachma
NIM. 18610077**

Telah Diperiksa dan Disetujui Untuk Diuji
Tanggal 27 Juni 2022

Dosen Pembimbing I



Dr. Sri Harini, M.Si
NIP/NIDT. 19731014 200112 2 002

Dosen Pembimbing II



Erna Herawati, M.Pd
NIP/NIDT. 19890813 201903 1 012

Mengetahui,
Ketua Program Studi Matematika



Dr. Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR
DALAM PENENTUAN KLASIFIKASI TINGKAT
KEDALAMAN KEMISKINAN PROVINSI JAWA TIMUR**

SKRIPSI

**Oleh
Clariza Adelina Rachma
NIM. 18610077**

Telah Dipertahankan di Depan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 29 Juni 2022

Ketua Penguji : Ria Dhea Layla Nur Karisma, M.Si
Anggota Penguji I : Angga Dwi Mulyanto, M.Si
Anggota Penguji II : Dr. Sri Harini, M.Si
Anggota Penguji III : Erna Herawati, M.Pd


.....

.....

.....

.....

Mengetahui,
Ketua Program Studi Matematika

Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005



PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Clariza Adelina Rachma

NIM : 18610077

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Implementasi Algoritma *K-Nearest Neighbor* dalam Penentuan
Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur.

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 29 Juni 2022

Yang membuat pernyataan,



Clariza Adelina Rachma

NIM. 18610077

MOTO

“Apa yang sudah ditakdirkan terjadi, maka akan terjadi, pada waktu yang tepat, tidak pernah terlalu cepat dan tidak pernah sangat terlambat. Semua Bermakna.”

“dan Dia tetap bersamamu dimana pun kamu berada”

(Q.S Al-Hadiid :04)

PERSEMBAHAN

Segala puji bagi Allah SWT dengan segala kerendahan hati, penulis
persembahkan skripsi ini untuk:

Ayahanda H. Mashadi dan Mama Eny Novita yang senantiasa memberikan do'a,
dukungan dan motivasi untuk keberhasilan serta kesuksesan penulis.

Kakak tersayang Nadya Vabela Rachma, Umik Hj. Nasucha, Om Dudon, Tante
Indri dan adik tercinta Lovita yang selalu memberi semangat dan doa-doa bagi
penulis dalam menuntut ilmu.

Teman seperjuangan AKSIOMA 18' terkhusus Meliana Nike, Indri Fatikhu,
Aulia Farras, Thalia Amira, Aprila S, Ahmad Faisal, Alya Rahma, Kiranamey dan
teman-teman dekat penulis yang senantiasa ada untuk mendukung, mendoakan
dan memberi semangat kepada penulis.

KATA PENGANTAR

Assalamu 'alaikum Warahmatullahi Wabarakatuh

Segala puji bagi Allah Subhanahu Wa Ta'ala yang telah melimpahkan rahmat, taufik serta hidayahnya, sehingga penulis masih diberi kesehatan dan kesempatan dalam menyelesaikan proposal skripsi yang berjudul “Implementasi Algoritma *K-Nearest Neighbor* dalam Penentuan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur” sebagai salah satu persyaratan dalam memperoleh gelar sarjana Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Shalawat serta salam semoga senantiasa tercurahkan kepada junjungan kita Nabi Muhammad Saw yang telah membawa kita dari jaman jahiliah menuju jaman islamiah.

Dalam penyusunan proposal skripsi ini penulis mendapatkan bimbingan, arahan, dan bantuan dari berbagai pihak, oleh karena itu penulis mengucapkan terimakasih kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang
3. Dr. Elly Susanti, S.Pd., M.Sc, selaku ketua Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Sri Harini, M.Si, selaku dosen pembimbing I yang telah meluangkan waktunya untuk memberikan bimbingan, arahan, nasihat, solusi dan banyak ilmu sehingga penulis dapat menyelesaikan proposal skripsi dengan baik.
5. Erna Herawati, M.Pd, selaku dosen pembimbing II yang telah memberikan bimbingan, arahan, dan banyak ilmu kepada penulis.
6. Ria Dhea Layla Nur Kharisma, M.Si selaku dosen penguji utama dalam ujian skripsi yang telah memberikan saran yang bermanfaat bagi penulis.
7. Angga Dwi Mulyanto, M.Si selaku dosen penguji I dalam ujian skripsi yang telah memberikan saran dan solusi yang bermanfaat bagi penulis.

8. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
9. Ayah Mashadi dan Mama Eny Novita selaku orang tua penulis serta segenap keluarga yang senantiasa memberikan dukungan, semangat, dan doa terbaik.
10. Seluruh mahasiswa program studi matematika angkatan 2018 yang telah memberikan bantuan, semangat, motivasi selama penyelesaian skripsi.
11. Seluruh pihak yang telah membantu dalam menyelesaikan skripsi ini.
Penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca maupun bagi penulis serta dapat dijadikan sebagai penambah wawasan ilmu matematika terutama dalam bidang statistik.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 29 Juni 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR SIMBOL	xiv
DAFTAR LAMPIRAN	xv
ABSTRAK	xvi
ABSTRACT	xvii
مستخلص البحث.....	xviii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	6
1.6 Definisi Istilah	6
BAB II KAJIAN TEORI	9
2.1 Metode K-Nearest Neighbor.....	9
2.1.1 Algoritma K-Nearest Neighbor	11
2.1.2 Transformasi Data	12
2.2 <i>K-Fold Cross Validation</i>	13
2.3 <i>Confusion Matrix</i>	14
2.4 Kemiskinan	15
2.5 Indeks Kedalaman Kemiskinan	18
2.6 Kajian Integrasi Kemiskinan dengan Al-Qur'an/Hadits.....	20
2.7 Kajian Kemiskinan dengan Teori Pendukung	22
BAB III METODE PENELITIAN	24
3.1 Jenis Penelitian	24
3.2 Data dan Sumber Data	24
3.3 Teknik Pengumpulan Data	24
3.4 Instrumen Penelitian	25
3.5 Teknik Analisis Data	25
3.6 Diagram Alur Penelitian	27
BAB IV PEMBAHASAN	28
4.1 Analisis Deskriptif	28
4.2 Analisis Data.....	35
4.2.1 Transformasi Data	35
4.2.2 Implementasi Algoritma <i>K-Nearest Neighbor</i>	36

4.2.3 Hasil Klasifikasi Menggunakan Algoritma <i>K-Nearest Neighbor</i>	38
4.2.4 <i>K-Fold Cross Validation</i>	41
4.2.5 Uji Akurasi Algoritma <i>K-Nearest Neighbor</i>	42
4.3 Konsep Klasifikasi dalam Islam	43
BAB V PENUTUP	45
5.1 Kesimpulan	45
5.2 Saran	46
DAFTAR PUSTAKA	47
LAMPIRAN	50
RIWAYAT HIDUP	58

DAFTAR TABEL

Tabel 2.1	Penentuan Label/Kelas.....	12
Tabel 2.2	<i>Confusion Matrix</i>	14
Tabel 4.1	Analisis Deskriptif Variabel Penelitian.....	30
Tabel 4.2	Penentuan Label/Kelas	35
Tabel 4.3	Data Latih 38 Kabupaten/Kota	38
Tabel 4.4	Data Uji	38
Tabel 4.5	Perolehan Jarak Data Uji	40
Tabel 4.6	Hasil Urutan Jarak.....	41
Tabel 4.7	Perbandingan Akurasi Parameter <i>K-NN</i>	41
Tabel 4.8	<i>Confusion Matrix</i> Hasil Klasifikasi <i>K-NN</i>	42

DAFTAR GAMBAR

Gambar 2.1	<i>10-fold Cross Validation</i>	14
Gambar 3.1	Diagram Alur Penelitian.....	27
Gambar 4.1	Pie Chart Kedalaman Kemiskinan Jatim 2020.....	28
Gambar 4.2	Bar Chart Indeks Kedalaman Kemiskinan Jatim 2020	29
Gambar 4.3	Bar Chart Indeks Pengeluaran Perkapita Jatim 2020	31
Gambar 4.4	Bar Chart Rata-rata Lama Sekolah Jatim 2020	32
Gambar 4.5	Bar Chart Laju PDRB Jatim 2020.....	33
Gambar 4.6	Bar Chart Tingkat Pengangguran Terbuka Jatim 2020.....	34

DAFTAR SIMBOL

$Dist(X, Y)$: Jarak antar objek (<i>Euclidean Distance</i>)
X_i	: Sampel data
Y_i	: Data uji
D	: Dimensi data
i	: Variabel data
x_i'	: Data standarisasi ke- i
x_i	: Data aktual ke- i
x_{max}	: Nilai maksimal dari data x
x_{min}	: Nilai minimal dari data x
P_1	: Indeks kedalaman kemiskinan absolut
n	: Jumlah penduduk
q	: Banyaknya penduduk yang berada dibawah garis kemiskinan
z	: Garis kemiskinan
y_i	: Rata-rata pengeluaran kapita perbulan bagi penduduk yang berada dibawah garis kemiskinan ($i = 1, 2, 3, \dots, q$), $y_i < z$

DAFTAR LAMPIRAN

Lampiran 1. Data Indeks Kedalaman Kemiskinan	50
Lampiran 2. Data Setelah Tranformasi Data.....	52
Lampiran 3. Data Setelah Standarisasi Data	54
Lampiran 4. Hasil Akurasi Klasifikasi Menggunakan <i>Software Rapidminer</i> .	56

ABSTRAK

Rachma, Clariza Adelina. 2022. **Implementasi Algoritma *K-Nearest Neighbor* Dalam Penentuan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Sri. Harini, M.Si (II) Erna Herawati, M.Pd.

Kata Kunci: Klasifikasi, Kemiskinan, Tingkat Kedalaman Kemiskinan, *K-Nearest Neighbor*

Kemiskinan menjadi masalah yang dijumpai diseluruh Negara, khususnya Negara berkembang seperti Indonesia. Rendahnya tingkat ketimpangan dalam pengeluaran antar masyarakat miskin disuatu daerah dapat diketahui dengan melihat turunnya nilai indeks kedalaman kemiskinan yang menunjukkan bahwa rata-rata pengeluaran mendekati garis kemiskinan. Pengklasifikasian tingkat kemiskinan yang tepat akan bermanfaat dalam pengambilan kebijakan pemerintah dalam pemberantasan kemiskinan. Tujuan penelitian ini untuk mendeskripsikan karakteristik tingkat kedalaman kemiskinan provinsi Jawa Timur 2020 dan mendapatkan hasil klasifikasi tingkat kedalaman kemiskinan dengan menggunakan algoritma *K-Nearest Neighbor*. Hasil penelitian ini menunjukkan bahwa terdapat 23 Kabupaten/Kota yang termasuk dalam kategori klasifikasi tingkat kemiskinan kurang dari rata-rata dan 15 Kabupaten/Kota sisanya masuk dalam kategori klasifikasi tingkat kemiskinan lebih dari rata-rata. Semakin tinggi indeks pengeluaran perkapita, laju PDRB dan rata-rata lama sekolah yang dikeluarkan di suatu wilayah maka memperlihatkan adanya peningkatan kesejahteraan masyarakat dan kualitas SDM di wilayah Kabupaten/Kota. Sedangkan semakin tinggi tingkat pengangguran terbuka suatu wilayah maka memperlihatkan penurunan pada tingkat kesejahteraan masyarakat di wilayah Kabupaten/Kota. Hasil akurasi yang dihasilkan dari klasifikasi menggunakan algoritma *K-Nearest Neighbor* menunjukkan akurasi tertinggi sebesar 76.67% dengan nilai parameter k terbaik yaitu $k = 1$ dan $k = 2$.

ABSTRACT

Rachma, Clariza Adelina, 2022. **The Implementation of the K-Nearest Neighbor Algorithm in Determining the Classification of the Gap of Poverty in the Province of East Java.** Thesis. Mathematics Study Program, Faculty of Science and Technology. UIN Maulana Malik Ibrahim Malang, Advisors: (I) Dr. Sri Harini, M.Si (II) Erna Herawati, M.Pd

Keywords: Classification, Poverty, Poverty Gap Level, K-Nearest Neighbor

Poverty is a problem found in all countries, especially developing countries like Indonesia. The low level of inequality in spending among the poor in an area can be seen by looking at the decrease in the poverty gap index value which shows that the average expenditure is close to the poverty line. The right classification of poverty levels will be useful in making government policies in poverty eradication. The purpose of this study was to describe the characteristics of the poverty gap of East Java province in 2020 and obtain the results of the classification of the poverty gap level using the K-Nearest Neighbor algorithm. The results of this study indicate that there are 23 regencies/cities were included in the category of poverty level classification off less than the average and the remaining 15 regencies/cities were included in the category of poverty level classification of more than average. The higher the index of per capita expenditure, the rate of GRDP and average length of schooling issued in an area, the higher the welfare of the people and the quality of human resources in the Regency/City area. Meanwhile, the higher the open unemployment rate in an area, the lower the level of community welfare in the Regency/City. The results of the accuracy of the classification using the K-Nearest Neighbor algorithm show the highest accuracy of 76.67% based on the best k parameter values, namely $k = 1$ and $k = 2$.

مستخلص البحث

رحمة ، كلاريزا أديلينا. ٢٠٢٢. تطبيق خوارزمية (K-NN) في تحديد تصنيف مستوى الفقر في مقاطعة جاوي الشرقية. البحث العلمي قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. لمشرفة : (١) دكتورة سري هارينى ، الماجستير. (٢) آرنا هيراواتي ، الماجستير.

الكلمات الرئيسية: تصنيف , فقر , مستوى الفقر , K- أقرب الجار

الفقر هو مشكلتنا التي توجد في كثير من البلدان ، وخاصة في البلدان النامية مثل إندونيسيا. يمكن رؤية المستوى المنخفض من عدم المساواة في الإنفاق بين الفقراء في المنطقة من خلال النظر إلى الانخفاض في القيمة المؤشر مستوى الفقر مما يدل على أن متوسط الإنفاق قريب من خط الفقر. سيكون التصنيف الصحيح لمستويات الفقر المفيدا في صنع السياسات الحكومية في القضاء على. الفقر الغرض من هذه الدراسة هو تحديد خصائص مستوى الفقر في مقاطعة جاوي الشرقية ٢٠٢٠ والحصول على نتائج تصنيف عمق الفقر باستخدام الخوارزمية *K-Nearest Neighbor*. تشير نتائج هذه الدراسة إلى أن هناك ثلاث وعشرين منطقة/مدينة مدرجة في فئة تصنيف مستوى الفقر أقل من المتوسط وأن المناطق الخمسة عشر/المدن المتبقية مدرجة في فئة تصنيف مستوى الفقر أكثر من المتوسط. كلما ارتفع مؤشر الإنفاق الفردي ارتفع معدل PDRB و يظهر متوسط مدة الدراسة الصادرة في منطقة ما زيادة في رفاهية المجتمع وجودة الموارد البشرية في منطقة ريجنسي/المدينة. وفي الوقت نفسه كلما ارتفع معدل البطالة المفتوحة في منطقة ما انخفض مستوى الرفاهية المجتمعية في منطقة ريجنسي/المدينة. تظهر نتائج الدقة الناتجة عن التصنيف باستخدام خوارزمية *K-Nearest Neighbor* أعلى دقة بنسبة 67.76% مع أفضل قيمة لمعامل k وهي $k = 1$.

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kemiskinan menjadi masalah yang dijumpai di seluruh negara, khususnya negara berkembang seperti Indonesia. Individu disebut miskin apabila individu tersebut tidak memiliki kapabilitas guna mencukupi kebutuhan dasar minimal yang layak bagi kehidupannya, seperti kebutuhan akan sandang, pangan, perumahan dan kebutuhan sosial. Kemiskinan juga disebut sebagai ketidakmampuan masyarakat dalam mencukupi kebutuhan primer dan sekunder sebagai standar minimum kebutuhan fundamental yang mesti tercukupi (Bahauddin dkk, 2021). Kemiskinan terjadi dikarenakan sedikitnya asset dan pendapatan untuk mencukupi keperluan dasar seperti pakaian, makanan, perumahan, tingkat kesehatan dan pendidikan yang diterima (Ferezagia, 2018).

Kemiskinan menjadi masalah utama pembangunan ekonomi di Indonesia. Berdasarkan data dari Badan Pusat Statistik (BPS) mengenai angka kemiskinan di Indonesia diperoleh bahwa pada September 2020, jumlah penduduk miskin sebanyak 27.55 juta jiwa atau setara dengan 10.19% dari total penduduk Indonesia. Angka ini mengalami kenaikan sebanyak 1.13 juta orang dari data bulan Maret 2020, serta meningkat sebanyak 2.76 juta orang dibanding dengan bulan September 2019. Sebelumnya jumlah orang miskin di Indonesia pada periode September 2019 mencapai 24.79 juta jiwa. Besarnya peningkatan angka kemiskinan ini diduga karena adanya pandemi Covid-19 (BPS, 2020).

Tingginya tingkat kemiskinan di Indonesia dikarenakan oleh beragam faktor, yaitu pengangguran, inflasi, pertumbuhan ekonomi, dan pendidikan. Selain

itu, tingginya kemiskinan juga memiliki efek yang bisa mengakibatkan banyak anak tidak mampu mengakses pendidikan yang berkualitas, tidak memiliki kemampuan untuk memenuhi biaya kesehatan, masih sangat minimnya jumlah tabungan yang dimiliki, sulitnya mengakses pelayanan publik, tidak memiliki kapabilitas untuk memenuhi perlindungan dan jaminan sosial terhadap keluarga, serta terlihat dari makin tingginya masyarakat yang melakukan urbanisasi (Hendayanti & Nurhidayati, 2020). Permasalahan kemiskinan dapat ditandai dengan rendahnya tingkat pendidikan, kondisi kesehatan yang kurang memadai, serta rendahnya kesejahteraan masyarakat ditambah dengan pendapatan yang kecil hingga menimbulkan ketimpangan.

Jumlah kemiskinan yang tinggi akan meningkatkan beragam masalah disertai dengan tingkat kesejahteraan masyarakat menurun. Oleh karena itu, perlu adanya klasifikasi tingkat kedalaman kemiskinan daerah di provinsi Jawa Timur supaya dapat ditemukan penyelesaian yang tepat sesuai tingkatan masing-masing. Selain diperlukan cara untuk mengklasifikasikan tingkat kedalaman kemiskinan, sebenarnya ada banyak ayat dalam Al-Qur'an yang menganjurkan standar hidup yang lebih baik bagi manusia untuk membantu mereka yang masih terjerat kemiskinan dalam mengurangi kesulitan yang mereka hadapi. Hal ini bertujuan untuk meningkatkan taraf hidup masyarakat miskin. Hal ini sejalan dengan tujuan hukum syariah, yaitu untuk kesejahteraan fisik dan mental umat manusia dalam kehidupan ini dan selanjutnya. Dalam firman Allah SWT Q.S *at-Taubat* ayat 103 (Kemenag, 2022):

خُذْ مِنْ أَمْوَالِهِمْ صَدَقَةً تُطَهِّرُهُمْ وَتُزَكِّيهِمْ بِهَا وَصَلِّ عَلَيْهِمْ ۖ إِنَّ صَلَاتَكَ سَكَنٌ لَهُمْ ۗ وَاللَّهُ سَمِيعٌ عَلِيمٌ

“Ambillah zakat dari harta mereka, guna membersihkan dan menyucikan mereka, dan berdoalah untuk mereka. Sesungguhnya doamu itu (menumbuhkan) ketenteraman jiwa bagi mereka. Allah Maha Mendengar, Maha Mengetahui.”

Ayat ini begitu jelas menyarankan mereka yang memiliki standar hidup yang lebih tinggi untuk membantu mereka yang kurang beruntung daripada mereka. Karena sesungguhnya bahwa harta kekayaan pada seseorang tidak hanya berada pada mereka golongan kaya saja. Salah satu cara dalam membantu mereka yang berhak menerimanya dapat melalui sedekah dan pembayaran zakat. Hal ini bertujuan supaya orang miskin dapat memperoleh manfaat dari kebaikan dan bantuan yang diberikan oleh mereka yang mampu.

Salah satu metode yang dapat digunakan dalam pengklasifikasian tingkat kedalaman kemiskinan daerah di provinsi Jawa Timur adalah algoritma *k-Nearest Neighbor*. *K-Nearest Neighbor* atau *k-NN* adalah metode klasifikasi berdasarkan fakta bahwa objek yang ‘dekat’ satu sama lain juga akan memiliki karakteristik yang serupa. Artinya apabila karakteristik suatu objek diketahui, maka objek lain juga bisa diprediksi berdasarkan tetangga terdekatnya (Wisdayani dkk, 2019). *K-NN* memiliki kelebihan dan kekurangan, beberapa kelebihan di antaranya yaitu secara efisien menangani data pelatihan yang besar dan tahan terhadap data pelatihan yang *noise*, dan dapat memberikan hasil data yang akurat. Algoritma *k-NN* memiliki prinsip kerja mencari jarak terdekat antara data yang akan dievaluasi dan (*k*) tetangga (*neighbor*) terdekatnya dalam data latih berdasarkan metode pembelajaran supervised. Kemudian data tersebut diklasifikasikan ke dalam kelas berdasarkan kategori (*k*) tetangga yang mayoritas. Kelas klasifikasi dapat dilihat berdasarkan kelas yang sering muncul. Penentuan sebuah variabel sangat berpengaruh terhadap hasil dan akurasi dari klasifikasi yang akan dilakukan.

Berdasarkan kajian penelitian mengenai metode *K-Nearest Neighbor* tentang peramalan arus lalu lintas jangka pendek menghasilkan nilai k yang optimal untuk masing-masing kategori kendaraan dengan nilai error RMSE yang kecil (Esesiawati, 2017). Penelitian mengenai perbandingan metode *K-Nearest Neighbor (K-NN)* dan *Learning Vector Quantization (LVQ)* dalam permasalahan klasifikasi tingkat kemiskinan, telah menghasilkan hasil klasifikasi dengan nilai akurasi sebesar 95.5185% untuk *K-NN* dan 75.9259% untuk *LVQ* (Santoso, 2016). Penelitian lainnya tentang membandingkan tiga metode yaitu *Logic Fuzzy*, *Artificial Neural Network*, dan *K-Nearest Neighbor* terhadap prediksi harga jual rumah memperoleh hasil bahwa *Logic Fuzzy* lebih unggul dengan akurasi terbaik 88% (Mukhlisin dkk, 2017). Penelitian lainnya tentang metode *K-NN* berbasis *Euclidian Distance* pada klasifikasi keluarga miskin menghasilkan nilai akurasi 90% dengan parameter $k=5, 7, \text{ dan } 9$ (Kurnia dkk, 2019). Selanjutnya penelitian mengenai algoritma *K-Nearest Neighbor* dalam perhitungan tingkat akurasi berdasarkan data cuaca di Indonesia. Hasil penelitian ini memberikan informasi bahwa terdapat perubahan suhu secara signifikan setiap tahunnya dengan nilai akurasi 89% (Yusuf dkk, 2021). Berdasarkan uraian diatas, maka penelitian ini menggunakan pendekatan Algoritma *K-Nearest Neighbor (KNN)* dalam penentuan klasifikasi tingkat kedalaman kemiskinan di provinsi Jawa Timur.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas dapat dibuat rumusan masalah penelitian sebagai berikut:

1. Bagaimana karakteristik tingkat kedalaman kemiskinan di daerah Jawa Timur?

2. Bagaimana hasil klasifikasi algoritma *K-Nearest Neighbor* pada data indeks kedalaman kemiskinan di daerah Jawa Timur?

1.3 Tujuan Penelitian

Berdasarkan dengan rumusan masalah yang telah diuraikan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mendeskripsikan karakteristik tingkat kedalaman kemiskinan di daerah Jawa Timur.
2. Menganalisis hasil klasifikasi yang dihasilkan dihasilkan algoritma *K-Nearest Neighbor* pada data indeks kedalaman kemiskinan di daerah Jawa Timur.

1.4 Manfaat Penelitian

Manfaat yang diharapkan oleh peneliti yaitu sebagai berikut berdasarkan pada tujuan penelitian diatas yaitu:

1. Bagi penulis
Sebagai tambahan pengetahuan dan wawasan tentang implementasi Algoritma *K-Nearest Neighbor* dalam suatu klasifikasi tingkat kedalaman kemiskinan di daerah Jawa Timur. Serta dapat digunakan sebagai bahan pertimbangan analisis statistik algoritma yang lain.
2. Bagi jurusan
Sebagai bahan pertimbangan bagi mahasiswa lain terkait analisis statistik algoritma yang lain dan sebagai bahan bacaan di perpustakaan.
3. Bagi instansi

Sebagai tambahan informasi terkait faktor yang mempengaruhi perubahan indeks kedalaman kemiskinan. Sehingga dapat menjadi bahan pertimbangan dalam membuat kebijakan upaya menanggulangi tingkat kemiskinan di daerah Jawa Timur.

1.5 Batasan Masalah

Agar pembahasan dalam penelitian ini tidak meluas, maka penulis memberi batasan masalah sebagai berikut:

1. Data yang digunakan adalah data indeks kedalaman kemiskinan di daerah Jawa Timur tahun 2020 dengan 4 variabel indikator tingkat kedalaman kemiskinan.
2. Nilai parameter (k) yang digunakan dalam pemodelan range 1 sampai 10.
3. Perhitungan jarak (k) menggunakan *Euclidean Distance*.
4. Tingkat kedalaman kemiskinan dikategorikan menjadi 2 kelas yaitu kelas 0 (kurang dari rata-rata) dan kelas 1 (lebih dari rata-rata).

1.6 Definisi Istilah

<i>Training</i>	: Set data yang digunakan untuk melatih algoritma
<i>Testing</i>	: Data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi.
<i>K-Nearest Neighbor</i>	: Metode untuk mengklasifikasikan suatu objek berdasarkan tetangga k terdekat
<i>Non Parametric</i>	: Uji statistik yang tidak membutuhkan

	jenis distribusi populasi
<i>Supervised</i>	: Pembelajaran <i>machine learning</i> yang menggunakan data berlabel untuk melatih model
<i>Cross Validation</i>	: Teknik untuk memperkirakan kesalahan kinerja dan akurasi suatu algoritma atau model
<i>K-Fold Cross Validation</i>	: Teknik validasi silang
<i>Confusion Matrix</i>	: Alat berbentuk matriks yang digunakan untuk pengujian tingkat akurasi pada kelas klasifikasi dari suatu algoritma
<i>TP</i>	: Jumlah kelas positif yang diklasifikasikan sebagai positif
<i>TN</i>	: Jumlah kelas negatif yang diklasifikasikan sebagai negatif
<i>FP</i>	: Jumlah kelas negatif yang diklasifikasikan sebagai positif
<i>FN</i>	: Jumlah kelas positif yang diklasifikasikan sebagai negatif
Indeks Kedalaman Kemiskinan	: Ukuran rata-rata kesenjangan pengeluaran masing-masing penduduk miskin terhadap garis kemiskinan
Garis kemiskinan	: Keadaan yang terletak dibawah garis nilai standar kebutuhan minimal

Integer : Tipe data yang merepresentasikan
bilangan bulat

String : Tipe data yang terdiri dari berbagai
karakter

BAB II KAJIAN TEORI

2.1 Metode K-Nearest Neighbor

Metode *K-Nearest Neighbor* merupakan metode non parametrik yang dapat digunakan untuk pengklasifikasian berdasarkan tetangga k terdekat dan regresi. Algoritma *K-Nearest Neighbor* merupakan metode untuk mengklasifikasikan objek yang paling dekat dengan objek berdasarkan data pembelajaran (Harun dkk, 2020). Algoritma *K-Nearest Neighbor* bekerja dengan mencari jarak antara data yang akan dievaluasi (data *training*) dan himpunan terdekat dari k tetangga (*neighbor*) terdekat dalam data baru (data *testing*). Ketika masuk dalam data kategori baru yang tidak diketahui untuk diklasifikasikan, maka kategori data baru tersebut harus ditentukan kategorinya berdasarkan sampel yang lain. Karakteristik data yang akan diklasifikasikan wajib diekstraksi dan dibandingkan dengan karakteristik dari setiap data kategori yang telah diketahui dalam data *testing*, kemudian data tetangga terdekat k harus diambil dari data *testing* untuk menghitung kategori yang mana sebagian besar data berada (Wang dkk, 2021). Paduan nilai terdekat berdasarkan dengan prosedur perkiraan waktu yang diharapkan, yang kemudian digunakan sebagai nilai masa depan seperti yang diharapkan. Prediksi berdasarkan *K-Nearest Neighbor* mengacu pada pola urutan yang diamati berulang dari waktu ke waktu (Esesiawati, 2017). Jika pola perilaku sebelumnya dapat diidentifikasi yang mirip dengan pola perilaku deret waktu saat ini, nilai-nilai selanjutnya berdasarkan nilai sebelumnya dapat digunakan untuk memprediksi pola atau nilai perilaku pada periode mendatang (Meade, 2002).

Penentuan nilai k dalam algoritma klasifikasi *K-Nearest Neighbor* dapat dicari berdasarkan nilai sampel k terdekat k_1, k_2, \dots, k_s . Semakin banyak data yang ada semakin kecil jumlah k yang dipilih, sedangkan jika ukuran dimensi data yang ada lebih besar, jumlah k yang dipilih harus lebih tinggi. Dalam menentukan nilai k , lebih baik menggunakan angka ganjil seperti $k = 1, 3, 5, \dots$, dst. Nilai k harus memenuhi syarat yaitu $k < N$ dimana N merupakan jumlah dari dataset latih, karena nilai k digunakan untuk mencari jumlah mayoritas dari kelas/label pada data latih maka nilai k tidak boleh lebih dari jumlah dataset latih (Lall & Sharma, 1996). Mencari tetangga terdekat atau jarak pada algoritma *K-NN* terdapat 5 cara yaitu Jarak *Euclidean*, Jarak *Manhattan*, Jarak *Cosine*, Jarak *Correlation*, Jarak *Hamming*. Jarak antara dua tetangga k terdekat berdasarkan nilai kemiripan dapat dihitung menggunakan jarak *Euclidean* yang didefinisikan sebagai berikut (Sreemathy, 2012):

$$Dist(X, Y) = \sqrt{\sum_{i=1}^D (X_i - Y_i)^2} \quad (2.1)$$

dengan:

$Dist(X, Y)$: jarak antar objek (*Euclidean Distancing*)

X_i : sampel data

Y_i : data uji

D : dimensi data

i : variabel data

Kinerja dari setiap model tetangga k terdekat dapat dicari dengan menggunakan optimasi parameter, misalnya dengan menggunakan *cross-validation*. *Cross validation* merupakan teknik untuk memvalidasi suatu kinerja

dan akurasi. *K-Nearest Neighbor* memiliki kelebihan dan kekurangannya sebagai berikut:

1. Kelebihan *K-Nearest Neighbor*
 - a. Mudah diimplementasikan
 - b. Tidak perlu membangun model, menentukan banyak parameter atau membuat asumsi tambahan
 - c. Tangguh terhadap *training* data yang *noise*
 - d. Efektif apabila data latihnya besar
2. Kekurangan *K-Nearest Neighbor*
 - a. *K-Nearest Neighbor* perlu menentukan nilai yang tepat dari parameter k (jumlah dari tetangga terdekat).
 - b. Pembelajaran berdasarkan jarak. Namun jarak dan atribut yang digunakan tidak jelas dalam memperoleh hasil yang terbaik.
 - c. Biaya cukup tinggi karena diunakan untuk menghitung jarak setiap sampel uji keseluruhan sampel *training*.

2.1.1 Algoritma K-Nearest Neighbor

Algoritma *K-Nearest Neighbor* dapat dituliskan sebagai berikut (Asha Gowda Karegowda dkk, 2012):

1. Tentukan nilai k tetangga terdekat
2. Hitung jarak data baru dengan data *training*
 Ukuran jarak yang digunakan adalah jarak *Euclidean* persamaan (2.1)
3. Urutkan jarak dari yang terdekat
4. Periksa kelas k tetangga terdekat
5. Kelas data baru = kelas mayoritas tetangga terdekat

2.1.2 Transformasi Data

Transformasi data merupakan suatu proses perubahan atribut data kedalam atribut data baru bentuk lain yang sesuai (Meilina, 2015). Dalam penentuan kelas/label dari data indeks kedalaman kemiskinan dapat menggunakan proses transformasi data. Kriteria klasifikasi yang digunakan dalam penentuan kelas/label pada indeks kedalaman kemiskinan sebagai berikut (Hendayanti & Nurhidayati, 2020) :

Tabel 2. 1 Penentuan Label/Kelas

Label/Kelas	Indeks Kedalaman Kemiskinan (IKK)
Kelas 0	IKK < Rata-rata
Kelas 1	IKK >= Rata-rata

Berikut merupakan rumus yang dapat digunakan dalam menentukan atau mencari nilai rata-rata:

$$Rata - rata = \frac{\text{total indeks kedalaman kemiskinan}}{\text{banyak data}} \quad (2.2)$$

2.1.3 Standarisasi Data

Satuan pengukuran pada data yang digunakan dapat mempengaruhi analisis data. Misalnya, jika variabel pengeluaran perkapita dalam persen dan variabel rata-rata lama sekolah dalam tahun maka dapat menyebabkan hasil yang sangat berbeda. Untuk menghindari masalah tersebut, maka perlu dilakukan proses standarisasi terhadap data yang digunakan. Standarisasi data dilakukan dengan menyesuaikan atribut data skala yang sama dengan interval [0,1] (Han & Kamber, 2013). Berikut merupakan rumus yang digunakan dalam perhitungan standarisasi data:

$$x_{i'} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2.3)$$

dengan:

$x_{i'}$: data standarisasi ke-i

x_i : data aktual ke -i

x_{max} : nilai maksimal dari data x

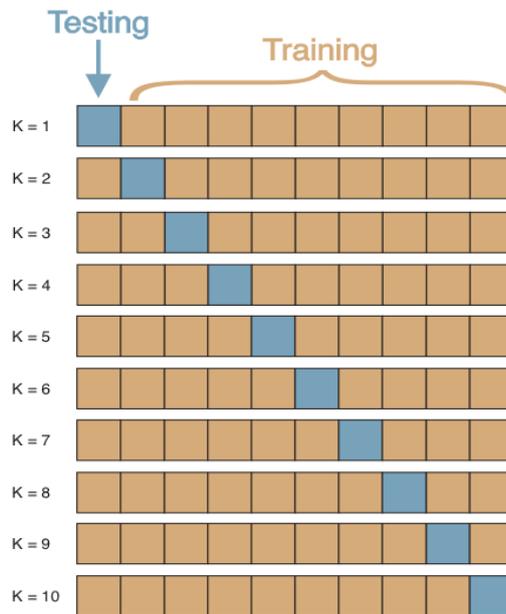
x_{min} : nilai minimal dari data x

2.2 *K-Fold Cross Validation*

Cross Validation merupakan standar pemeriksaan yang dilakukan untuk memperkirakan kesalahan kinerja suatu algoritma atau model (Sumarlin, 2018). *K-Fold Cross Validation* merupakan teknik validasi silang yang umum digunakan dalam menentukan suatu nilai k terbaik dalam klasifikasi. *K-Fold Cross Validation* dilakukan pada proses klasifikasi data awal secara acak dibagi menjadi k subset berukuran sama sesuai kebutuhan. Pada *cross validation*, proses *training* dan *testing* dikerjakan sebanyak k kali. Dalam perulangan ke-i, partisi yang disediakan menjadi data *testing* dan partisi sisanya menjadi data *training* (Han & Kamber, 2013). Tingkat akurasi tertinggi digunakan menjadi nilai k dalam penentuan k terbaik atau k -optimal (Banjarsari dkk, 2015). Perhitungan akurasi pada *k-fold cross validation* menggunakan persamaan (2.4).

Contoh pada pengujian data 38 Kabupaten/Kota di Provinsi Jawa Timur data akan dibagi menjadi *10-fold cross validation* yang sama rata. Kemudian data tersebut dieksekusi 10 kali untuk setiap subset data yang akan memiliki peluang untuk dijadikan sebagai data *testing* atau data *training*. Pada pengujian dengan data yang telah dipartisi dan diulang sebanyak 10 kali dengan posisi data *testing*

yang berbeda untuk setiap iterasi, posisi data *testing* pada iterasi pertama berada di awal, kemudian pada iterasi kedua data *testing* berada di posisi kedua begitu seterusnya hingga iterasi ketiga. Berikut merupakan gambaran model penggunaan *10-fold cross validation*.



Gambar 2. 1 *10-fold Cross Validation*

2.3 Confusion Matrix

Confusion matrix adalah alat berbentuk matriks yang digunakan untuk pengujian tingkat akurasi pada kelas klasifikasi dari suatu algoritma (Sasongko, 2016). *Confusion matrix* disajikan dalam tabel matriks yang berisi informasi terkait kelas klasifikasi data asli dan kelas klasifikasi hasil prediksi. Adapun bentuk contoh *confusion matrix* sebagai berikut:

Tabel 2. 2 *Confusion Matrix*

<i>Confusion Matrix</i>		Nilai Prediksi	
		TRUE	FALSE
Nilai Asli	TRUE	TP (True Positive)	FN (False Negative)
	FALSE	FP (False Positive)	TN (True Negative)

Keterangan tabel:

1. *True Positive* (TP) adalah jumlah kelas positif yang diklasifikasikan sebagai positif.
2. *False Positive* (FP) adalah jumlah kelas negatif yang diklasifikasikan sebagai positif.
3. *False Negative* (FN) adalah jumlah kelas positif yang diklasifikasikan sebagai negatif.
4. *True Negative* (TN) adalah jumlah kelas negatif yang diklasifikasikan sebagai negatif.

Akurasi klasifikasi menentukan seberapa tepat model klasifikasi yang dibentuk secara keseluruhan. Klasifikasi berupaya membentuk model dengan akurasi yang tinggi, karena semakin tinggi tingkat akurasi maka semakin baik model yang dihasilkan. Berikut formulasi dari nilai yang dihasilkan melalui *confusion matrix* untuk menghitung *accuracy*, *precision*, dan *recall* (Han & Kamber, 2013):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2.4)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2.5)$$

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

2.4 Kemiskinan

Kemiskinan salah satu persoalan serius dari persoalan-persoalan lain yang sulit dihadapi oleh pemerintah. Persoalan kemiskinan masih sering menjadi persoalan yang berulang di beberapa negara, terutama negara berkembang seperti Indonesia. Beberapa ahli berpendapat bahwa kemiskinan adalah sedikitnya akses

masyarakat terhadap pelayanan seperti kesehatan, pendidikan, informasi, politik, dan partisipasi pembangunan. Badan pusat statistik (BPS) mendefinisikan kemiskinan sebagai ketidakmampuan masyarakat dalam memenuhi kebutuhan dasar minimal, seperti kebutuhan primer dan sekunder. Suatu keadaan yang terletak dibawah garis nilai standar kebutuhan minimal yaitu garis kemiskinan (*poverty line*) atau batas kemiskinan (*poverty threshold*). Masyarakat dapat dikategorikan miskin jika pengeluarannya berada dibawah garis kemiskinan tersebut (BPS, 2022a). Garis kemiskinan yang dimaksud adalah batas minimal pengeluaran konsumsi masyarakat penduduk dalam memenuhi kebutuhan pangan dan non pangan seperti sandang, pangan, rumah, pendidikan dan kesehatan. Garis kemiskinan dapat diartikan sebagai garis pembatas yang membedakan tingkat kesejahteraan antar penduduk (Sarjono dkk, 2019). Berdasarkan garis kemiskinan, tingkat kemiskinan masyarakat dapat dikelompokkan menjadi sangat miskin, miskin, hampir miskin, dan rentan miskin lainnya.

Menurut Badan Pusat Statistik terdapat beberapa faktor yang mempengaruhi tingkat kemiskinan yaitu indeks pengeluaran perkapita, rata-rata lama sekolah, laju PDRB, dan tingkat pengangguran terbuka. Indeks pengeluaran perkapita merupakan biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan yang dibagi dengan banyaknya anggota rumah tangga yang telah disesuaikan. Rata-rata lama sekolah menunjukkan jumlah tahun yang digunakan oleh penduduk dalam menempuh pendidikan formal pada umur lebih dari 25 tahun dengan batas minimum 0 tahun dan maksimum 15 tahun. Produk Domestik Regional Bruto merupakan salah satu indikator yang digunakan untuk mengukur kinerja perekonomian suatu daerah dalam satu periode (Bintang & Woyanti,

2018). Laju pertumbuhan produk domestik regional bruto menunjukkan presentase pertumbuhan produksi barang dan jasa di suatu wilayah perekonomian. Oleh karena itu, pertumbuhan ekonomi mengindikasikan pertumbuhan produksi barang dan jasa di suatu wilayah perekonomian dalam selang waktu tertentu. Tingkat pengangguran terbuka merupakan presentase jumlah pengangguran terhadap jumlah angkatan kerja. Pengangguran terbuka terdiri dari masyarakat yang tak punya pekerjaan dan mencari pekerjaan, masyarakat yang tak punya pekerjaan dan mempersiapkan usaha, masyarakat yang tidak memiliki pekerjaan dan tidak mencari pekerjaan karena merasa tidak akan mendapatkan pekerjaan, dan yang terakhir yaitu masyarakat yang sudah punya pekerjaan, tetapi belum mulai bekerja.

Kemiskinan memiliki 2 macam ukuran secara umum yaitu kemiskinan absolut dan kemiskinan relatif (Nasution dkk, 2017):

1. Kemiskinan Absolut

Kemiskinan absolut atau mutlak merupakan kemiskinan yang berhubungan dengan standar hidup minimum dan tingkat pendapatan suatu individu masyarakat yang sifatnya tetap tanpa dipengaruhi oleh keadaan ekonomi suatu masyarakat. Miskin absolut dapat dihitung dengan membandingkan tingkat pendapatan yang dibutuhkan individu untuk bertahan hidup dengan tingkat pendapatan individu tersebut. Ketika pendapatan pribadi berada di bawah garis kemiskinan, maka dapat dikatakan berada dalam kemiskinan absolut, sehingga mengakibatkan ketidakmampuan dalam memenuhi kebutuhan pokoknya. Tingkat pendapatan minimum merupakan pembatas antara keadaan miskin dengan tidak miskin.

2. Kemiskinan Relatif

Kemiskinan yang dilihat dari segi ketimpangan sosial dan sangat berhubungan dengan persebaran pendapatan disebut dengan kemiskinan relatif. Ketika suatu individu telah mampu dalam mencukupi kebutuhan dasar minimalnya, akan tetapi masih terbelang rendah dibanding masyarakat sekitarnya, maka individu tersebut dikategorikan miskin relatif. Semakin besar kesenjangan pendapatan antara kelas atas dan kelas bawah maka semakin besar jumlah penduduk miskin.

2.5 Indeks Kedalaman Kemiskinan

Permasalahan kemiskinan bukan hanya tentang jumlah dan persentase penduduk miskin. Salah satu indikator lain yang perlu untuk ditinjau adalah tingkat kedalaman kemiskinan. Tingkat kedalaman kemiskinan menggambarkan ukuran rata-rata kesenjangan pengeluaran dari garis kemiskinan untuk setiap orang miskin (Wardani dkk, 2021). Semakin tinggi nilai tingkat kedalaman kemiskinan maka akan berpengaruh pada semakin jauh rata-rata pengeluaran penduduk dari garis kemiskinan. Sedangkan semakin rendah nilai tingkat kedalaman kemiskinan, maka semakin dekat pengeluaran rata-rata penduduk dengan garis kemiskinan. Dengan demikian hal ini akan berefek pada semakin mudahnya masyarakat miskin keluar dari kemiskinan.

Berdasarkan data dari BPS diketahui tingkat kedalaman kemiskinan meningkat jadi 1.61% pada Maret 2020 dari posisi Maret 2019 yang sebesar 1.55%. Kondisi tersebut menunjukkan bahwa rata-rata pengeluaran masyarakat miskin cenderung mendekati garis kemiskinan (BPS, 2020). Tingkat kedalaman kemiskinan tertinggi di Pulau Jawa yaitu Jawa Timur dengan tingkat kedalaman

kemiskinan sebesar 2.07% pada tahun 2018 dan 1.61% pada tahun 2019. Di tahun 2020, tingkat kedalaman kemiskinan tertinggi ditempati oleh Daerah Istimewa Yogyakarta, namun angka kedalaman kemiskinan di provinsi Jawa Timur masih cenderung tinggi. Untuk mengatasi kemiskinan tersebut pemerintahan provinsi Jawa Timur merencanakan berbagai program seperti program elektrifikasi, program Bansos Usaha Ekonomi Produktif Kelompok Usaha Bersama (KUBE), Renovasi Rumah Tinggal Layak Huni (Rutilahu), program Desa Berdaya, PKH-plus, serta program pemberdayaan usaha perempuan (JATIM PUSPA).

Indeks kedalaman kemiskinan dapat memperlihatkan kualitas kemiskinan disuatu daerah. Rendahnya tingkat ketimpangan dalam pengeluaran antar masyarakat miskin disuatu daerah dapat diketahui dengan melihat turunnya nilai indeks kedalaman kemiskinan yang menunjukkan bahwa rata-rata pengeluaran mendekati garis kemiskinan. Indeks kedalaman kemiskinan dapat juga disebut sebagai ukuran kemiskinan yang menghitung seberapa jauh masyarakat miskin berada dibawah garis kemiskinan (Sarjono dkk 2019). Adapun persamaan dalam perhitungan indeks kedalaman kemiskinan menurut BPS sebagai berikut (BPS, 2022a):

$$P_1 = \frac{1}{n} \sum_{i=1}^q \left[\frac{z - y_i}{z} \right] \quad (2.5)$$

dimana:

P_1 = Indeks kedalaman kemiskinan absolut

n = Jumlah penduduk

q = Banyaknya penduduk yang berada dibawah garis kemiskinan

z = Garis kemiskinan

y_i = Rata-rata pengeluaran kapita sebulan bagi penduduk yang berada dibawah garis kemiskinan ($i = 1, 2, 3, \dots q$), $y_i < z$

2.6 Kajian Integrasi Kemiskinan dengan Al-Qur'an/Hadits

Alqur'an yang sesuai dengan pembahasan ini adalah (Kemenag, 2022):

خُذْ مِنْ أَمْوَالِهِمْ صَدَقَةً تُطَهِّرُهُمْ وَتُزَكِّيهِمْ بِهَا وَصَلِّ عَلَيْهِمْ ۖ إِنَّ صَلَاتَكَ سَكَنٌ لَهُمْ ۗ وَاللَّهُ سَمِيعٌ عَلِيمٌ

“Ambillah zakat dari harta mereka, guna membersihkan dan menyucikan mereka, dan berdoalah untuk mereka. Sesungguhnya doamu itu (menumbuhkan) ketenteraman jiwa bagi mereka. Allah Maha Mendengar, Maha Mengetahui.”

Menurut tafsir Al-Mishbah pada ayat 103 menjelaskan cara membersihkan diri, yang untuknya Allah SWT memerintahkan Nabi SAW untuk mengambil sebagian harta mereka untuk sedekah dengan harapan Allah SWT mengampuni. Salah satu cara dalam pengampunan-Nya yaitu dengan bersedekah dan membayar zakat kepada yang berhak menerimanya. Yakni dengan Nabi SAW mengambil harta mereka atas nama Allah SWT, engkau membersihkannya dan menyucikan jiwa mereka serta mengembangkan harta mereka, dan berdo'alah memohon keselamatan dan kesejahteraan bagi mereka. Sungguh doa'mu dapat membawa ketenangan pikiran bagi mereka yang selama ini resah karena takut akibat dosanya selama ini. Dan katakanlah kepada mereka bahwa Allah Maha Mendengar dan Maha Mengetahui (Riwayati & Hidayah, 2018).

Pada ayat 103 juga dapat dipahami bahwa harta kekayaan pada seseorang tidak hanya berada pada golongan orang kaya saja, namun lebih baik disalurkan untuk membantu orang miskin yang tidak seberuntung mereka. Karena orang yang beriman merupakan orang-orang yang sadar bahwa didalam harta kekayaan

mereka terdapat terdapat hak-hak orang lain. Islam mengajarkan umatnya untuk memberikan sebagian harta kekayaannya dengan melakukan sedekah dan pembayaran zakat. Golongan yang berhak menerima yaitu fakir dan miskin. Fakir merupakan suatu golongan yang mempunyai keadaan ekonomi amat buruk seperti tidak ada penghasilan, tidak ada usaha, tidak ada alat dan kemampuan dalam bekerja. Sedangkan miskin merupakan suatu keadan yang memiliki kemampuan dalam sumber penghasilan, namun masih tidak cukup dalam memenuhi kebutuhan dasar minimalnya seperti kebutuhan primer. Maka dari hal ini bertujuan supaya orang miskin menerima bantuan dan itikad baik dari si kaya/mampu. Sehingga permasalahan kemiskinan dapat diselesaikan dengan pengelolaan zakat yang adil. Selain itu, islam memberikan jawaban dan solusi terkait permasalahan kemiskinan yang ada di masyarakat, sebagaimana yang tertuang pada firman Allah SWT Q.S al-Nisā' ayat 9 (Kemenag, 2022):

لِيَخْشَ الَّذِينَ لَوْ تَرَكُوا مِنْ خَلْفِهِمْ ذُرِّيَّةً ضِعَافًا خَافُوا عَلَيْهِمْ فَلْيَتَّقُوا اللَّهَ وَلْيَقُولُوا قَوْلًا سَدِيدًا

“Dan hendaklah takut (kepada Allah) orang-orang yang sekiranya mereka meninggalkan keturunan yang lemah di belakang mereka yang mereka khawatir terhadap (kesejahteraan)nya. Oleh sebab itu, hendaklah mereka bertakwa kepada Allah, dan hendaklah mereka berbicara dengan tutur kata yang benar.”

Pesan yang disampaikan ayat ini adalah hendaklah masyarakat yang saat ini hidupnya lebih beruntung daripada mereka masyarakat miskin merasa takut untuk meninggalkan generasi yang lemah. Namun, kelemahan disini bukan hanya kelemahan ekonomi, tetapi juga kelemahan intelektual, serta kelemahan spiritual sosial yang penting untuk diperhatikan oleh masyarakat. Hal ini bertujuan untuk menciptakan generasi yang memiliki rasa kepedulian yang tinggi dan menjaga kesejahteraan bersama.

2.7 Kajian Kemiskinan dengan Teori Pendukung

Penelitian ini memiliki beberapa tahapan. Tahapan pertama yaitu melakukan pengumpulan data. Tahapan kedua yaitu analisis statistika deskriptif untuk memberikan gambaran umum karakteristik tingkat kedalaman kemiskinan pada 38 Kabupaten/Kota di provinsi Jawa Timur. Tahapan ketiga yaitu melakukan transformasi data dengan mengubah data indeks kedalaman kemiskinan menjadi bentuk label/kelas dengan kategori klasifikasi tingkat kedalaman kelas 0 (kurang dari rata-rata) dan kelas 1 (lebih dari rata-rata). Tahapan keempat melakukan proses standarisasi data dengan interval $[0,1]$. Tahapan selanjutnya yaitu perhitungan data menggunakan algoritma *k-Nearest Neighbor* untuk memperoleh hasil klasifikasi tingkat kedalaman kemiskinan daerah di provinsi Jawa Timur berdasarkan indikator indeks pengeluaran per kapita, rata-rata lama sekolah, laju PDRB, dan tingkat pengangguran terbuka.

Dalam tahapan perhitungan menggunakan algoritma *k-Nearest Neighbor* perlu menentukan parameter nilai k guna untuk menentukan jarak tetangga terdekat pada objek. Nilai k yang digunakan dalam proses *k-Nearest Neighbor* adalah $k = 1$ sampai $k = 10$. Setelah menentukan nilai parameter k maka urutkan jarak terdekat menggunakan rumus jarak *Euclidean* seperti pada persamaan (2.1). Kemudian mencari jumlah label/kelas yang mayoritas untuk menetapkan label/kelas sebagai nilai prediksi yang akan dievaluasi. Tahapan selanjutnya yaitu melakukan pengujian *k-fold cross validation* agar hasil akurasi yang dihasilkan lebih valid sebanyak 10 kali pengujian. Setelah itu melakukan uji tingkat akurasi dengan tujuan untuk mengetahui keakuratan suatu algoritma menggunakan *confusion matrix*. Tahap terakhir yaitu deskripsikan dan analisis hasil klasifikasi

indeks kedalaman kemiskinan di 38 Kabupaten/Kota Provinsi Jawa Timur berdasarkan tahapan-tahapan di atas.

BAB III METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian pada penelitian ini adalah penelitian deskriptif kuantitatif. Pada penelitian ini akan dilakukan analisis terhadap hasil dari algoritma *k-nearest neighbor* yang diperoleh dari subjek data yang bersifat kuantitatif atau numerik yaitu data tingkat kedalaman kemiskinan pada 38 Kabupaten/Kota di provinsi Jawa Timur tahun 2020.

3.2 Data dan Sumber Data

Data yang digunakan adalah data sekunder. Data sekunder adalah data yang diperoleh melalui media perantara seperti data yang telah tersedia pada laman resmi. Data yang digunakan adalah data indeks kedalaman kemiskinan (IKK), indeks pengeluaran per kapita, rata-rata lama sekolah, laju pertumbuhan produk domestik regional bruto serta tingkat pengangguran terbuka pada 38 Kabupaten/Kota di provinsi Jawa Timur tahun 2020 yang diakses melalui laman resmi Badan Pusat Statistik.

3.3 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah teknik dokumenter. Teknik dokumenter merupakan teknik pengumpulan data dengan mencari data atau variabel yang dibutuhkan melalui sumber-sumber informasi yang ada, baik tertulis maupun sumber lisan (Nilamsari, 2014). Seperti data yang telah tersimpan atau terpublikasi pada laman resmi.

3.4 Instrumen Penelitian

Instrumen penelitian yang digunakan dalam penelitian ini berupa *software Rapidminer* dan *tools Microsoft. Excel*. Berikut variabel yang akan digunakan dalam penelitian yaitu:

1. Variabel *Dependent* (Y)

Variabel terikat yaitu label/kelas output dalam penelitian ini berdasarkan tingkat kedalaman kemiskinan dengan label/kelas kelas 0 (kurang dari rata-rata) dan kelas 1 (lebih dari rata-rata).

2. Variabel *Independent* (X)

a. Indeks pengeluaran perkapita (X_1)

Persentase biaya yang dikeluarkan dalam satu bulan yang kemudian dibagi dengan jumlah keluarga yang ada.

b. Rata-rata lama sekolah (X_2)

Persentase jumlah tahun yang digunakan oleh penduduk dalam menempuh pendidikan formal pada umur lebih dari 25 tahun dengan batas minimum 0 tahun dan maksimal 15 tahun.

c. Laju pertumbuhan Produk Domestik Regional Bruto (X_3)

Persentase pertumbuhan suatu nilai yang dihasilkan dari barang dan jasa di seluruh unit ekonomi pada suatu wilayah.

d. Tingkat Pengangguran Terbuka (X_4)

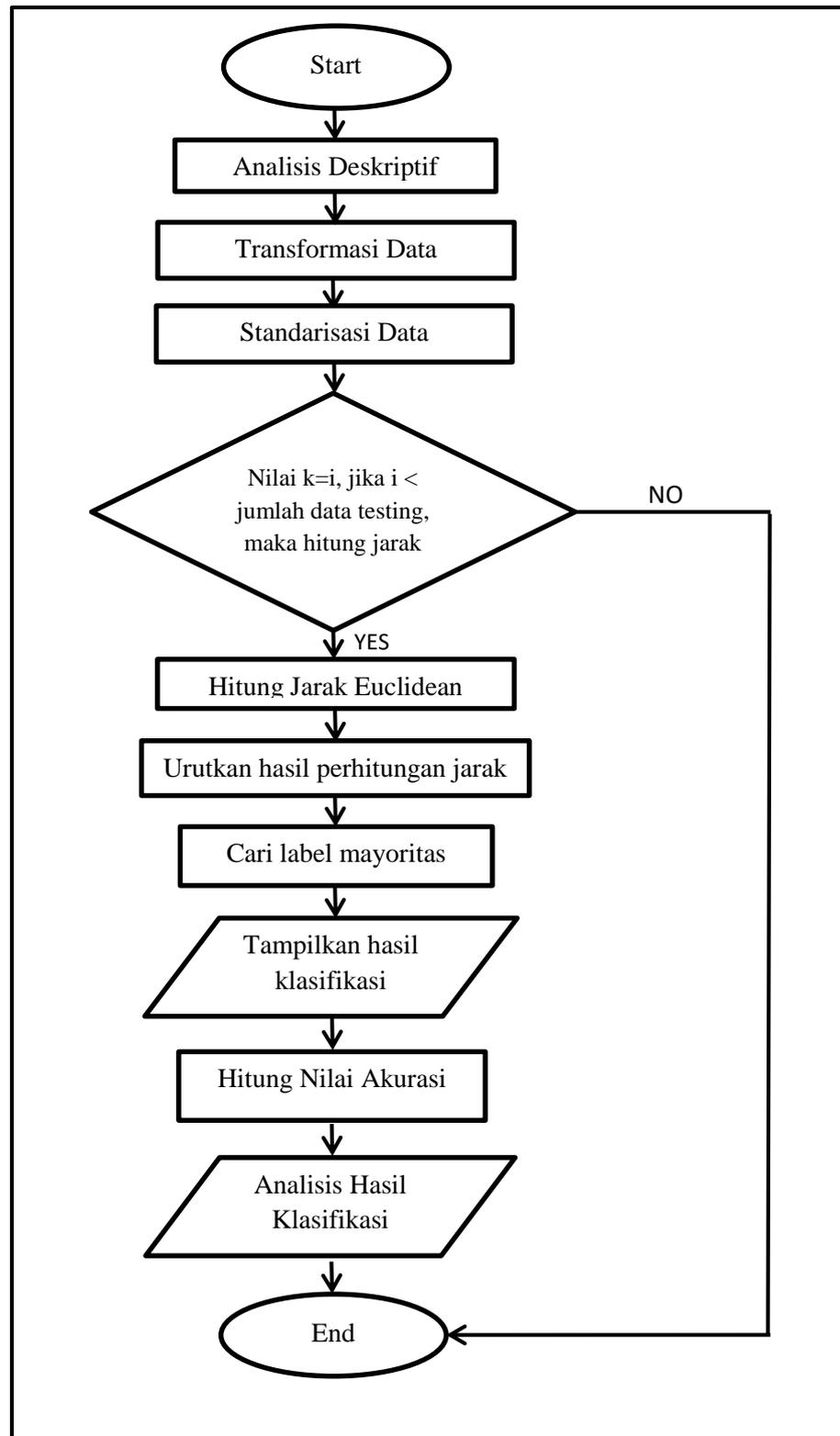
Persentase jumlah pengangguran terhadap jumlah angkatan kerja.

3.5 Teknik Analisis Data

Adapun langkah-langkah yang dilakukan dalam penelitian ini yakni sebagai berikut:

1. Pengumpulan data klasifikasi kedalaman kemiskinan meliputi variabel *independent* dan variabel *dependent*.
2. Melakukan analisis deskriptif terhadap data tingkat kedalaman kemiskinan untuk mendapatkan gambaran umum karakteristik tingkat kedalaman kemiskinan.
3. Melakukan transformasi data dengan mengubah data pada indeks kedalaman kemiskinan untuk mendapatkan kelas/label tingkat kedalaman kemiskinan dengan kelas 0 (kurang dari rata-rata) dan kelas 1 (lebih dari rata-rata).
4. Melakukan standarisasi data yang akan digunakan dengan interval $[0,1]$.
5. Perhitungan menggunakan algoritma *k-nearest neighbor*.
6. Menetapkan nilai parameter k tetangga terdekat yang akan digunakan. Nilai parameter k yang digunakan dalam proses *k-NN* antara 1 sampai 10.
7. Hitung jarak dan kemudian urutkan jarak yang terbentuk dari yang terkecil ke terbesar berdasarkan nilai inputan k .
8. Tentukan jarak terdekat sesuai nilai inputan k .
9. Cari jumlah label mayoritas dan tetapkan label tersebut sebagai nilai prediksi klasifikasi yang akan di analisis.
10. Melakukan evaluasi akurasi untuk mengetahui keakuratan algoritma *k-nearest neighbor* menggunakan *confusion matrix*.
11. Memvalidasi hasil akurasi klasifikasi menggunakan *k-fold cross validation*.
12. Analisis hasil dari prediksi klasifikasi yang sudah terbentuk dan uji akurasi pada *confusion matrix*.

3.6 Diagram Alur Penelitian

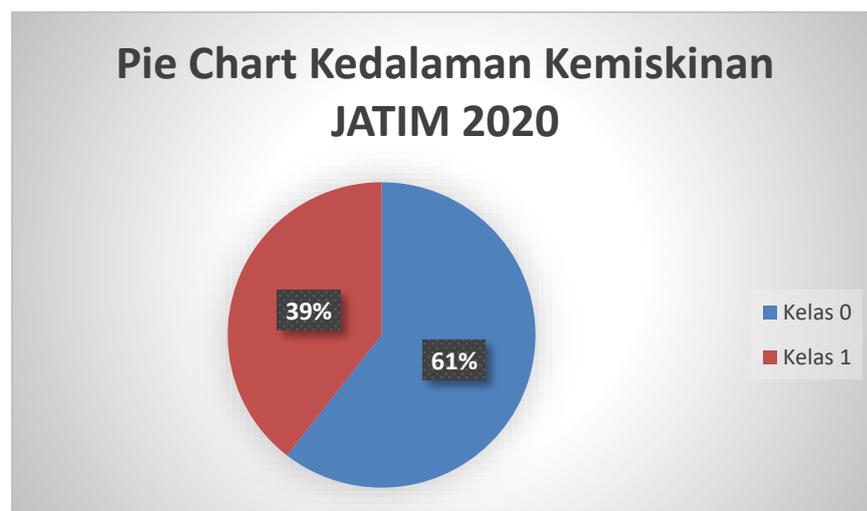


Gambar 3. 1 Diagram Alur Penelitian

BAB IV PEMBAHASAN

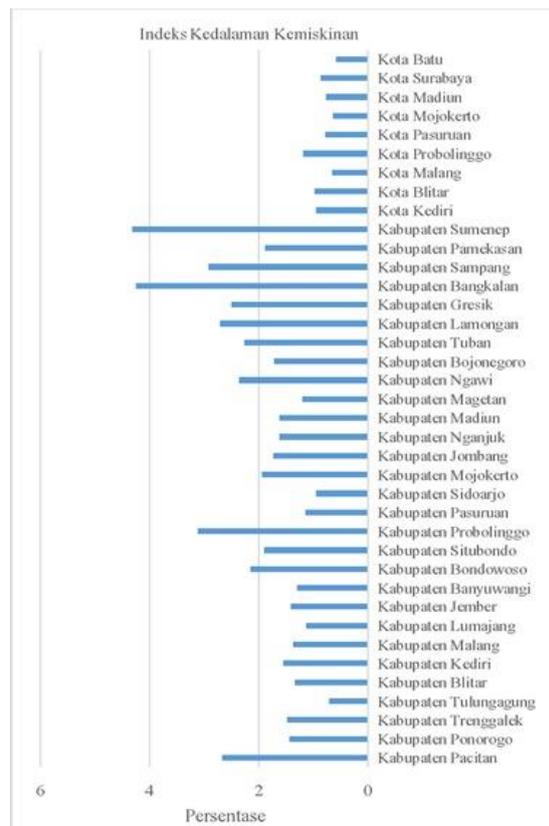
4.1 Analisis Deskriptif

Tahapan awal pada penelitian ini adalah analisis deskriptif yang bertujuan untuk memperoleh sedikit gambaran umum tentang karakteristik tingkat kedalaman kemiskinan pada 38 Kabupaten/Kota di provinsi Jawa Timur tahun 2020 berdasarkan algoritma *k-Nearest Neighbor*. Pada penelitian ini menggunakan variabel *dependent Y* yaitu tingkat kedalaman kemiskinan di Jawa Timur tahun 2020. Nilai variabel Y digambarkan dalam bentuk *pie chart* sebagai berikut:



Gambar 4. 1 Pie Chart Kedalaman Kemiskinan Jatim 2020

Pada Gambar 4.1 menggambarkan presentase 38 Kabupaten/Kota di provinsi Jawa Timur 2020. Terdapat 23 Kabupaten/Kota yang termasuk dalam kategori klasifikasi kedalaman kemiskinan kelas 0 (kurang dari rata-rata) sebesar 61%. Sementara itu, 15 Kabupaten/Kota lainnya sebesar 39% yang merupakan Kabupaten/Kota yang termasuk dalam kategori klasifikasi kedalaman kemiskinan kelas 1 (lebih dari rata-rata).



Gambar 4.2 Bar Chart Indeks Kedalaman Kemiskinan Jatim 2020
Sumber : (BPS, 2022b)

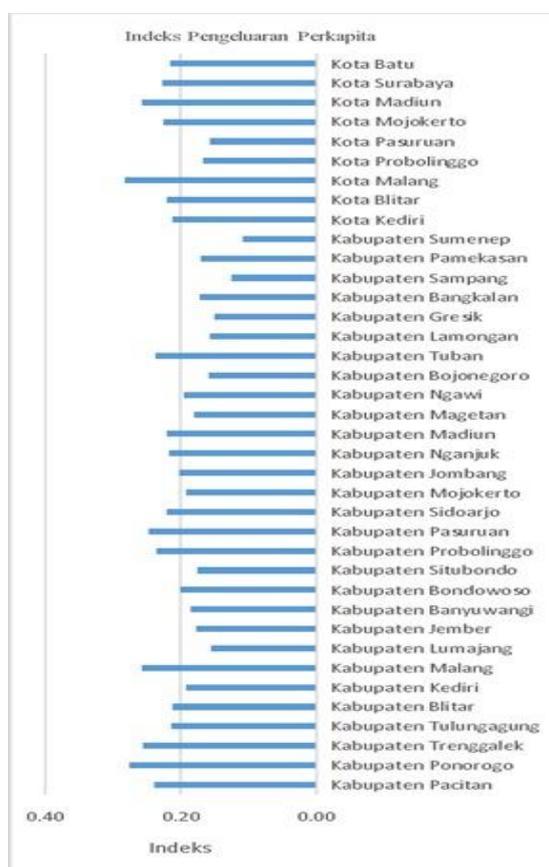
Pada Gambar 4.2 menggambarkan karakteristik variabel Y indeks kedalaman kemiskinan 38 Kabupaten/Kota provinsi Jawa Timur tahun 2020. Hasil pada Gambar 4.2 menunjukkan bahwa daerah yang menduduki indeks kedalaman kemiskinan tertinggi yaitu Kabupaten Sumenep dengan nilai indeks kedalaman kemiskinan sebesar 4.33%. Kemudian untuk daerah yang menduduki indeks kedalaman kemiskinan terendah yaitu Kota Batu dengan nilai sebesar 0.59%. Suatu Kabupaten/Kota dapat dikatakan memiliki kedalaman kemiskinan kelas 0 apabila indeks kedalaman kemiskinannya berada dibawah rata-rata. Dan jika suatu Kabupaten/Kota tingkat kedalaman kemiskinannya kelas 1 maka indeks kedalaman kemiskinan berada diatas rata-rata. Terdapat empat variabel *dependent* (X) yang diduga menjadi faktor indeks kedalaman kemiskinan di provinsi Jawa Timur dapat dilihat pada tabel 4.1 berikut:

Tabel 4. 1 Analisis Deskriptif Variabel Penelitian

Analisis Deskriptif					
Variabel	N	Mean	Std. Deviasi	Minimum	Maximum
Y	38	1,69	0,9013	0,59%	4,33%
X ₁	38	0,20	0,0399	0,11%	0,28%
X ₂	38	7,94	1,5097	4,85%	11,14%
X ₃	38	-2,77	1,5024	-6,46%	-0,29%
X ₄	38	5,62	1,9769	2,28%	10,97%

Berdasarkan tabel 4.1 dapat dilihat bahwa variabel Y yaitu mengenai tingkat kedalaman kemiskinan yang mempunyai nilai rata-rata sebesar 1,69%, standar deviasi sebesar 0,9013% yang artinya variabel Y bersifat homogen karena standar deviasi yang dihasilkan lebih kecil dari nilai rata-rata, nilai minimum daerah yang memiliki tingkat kedalaman kemiskinan sebesar 0,59% dan nilai maximum daerah yang memiliki tingkat kedalaman kemiskinan sebesar 4,33%. Pada variabel X₁ yaitu mengenai indeks pengeluaran perkapita diperoleh nilai rata-rata sebesar 0,20%, standar deviasi 0,0399% yang artinya variabel X₁ bersifat homogen karena nilai standar deviasi yang dihasilkan lebih kecil dari nilai rata-rata dan menunjukkan rentang variasi data indeks pengeluaran perkapita antar Kabupaten/Kota di provinsi Jawa Timur tahun 2020, serta nilai minimum sebesar 0,11% dan nilai maksimum sebesar 0,28%. Hal ini menunjukkan bahwa indeks pengeluaran perkapita berada pada kisaran 0,11%–0,28%. Selanjutnya pada variabel X₂ yaitu mengenai rata-rata lama sekolah diperoleh nilai rata-rata sebesar 7,94 tahun dengan standar deviasi sebesar 1,5097 yang artinya variabel X₂ bersifat homogen, dan nilai minimum sebesar 4,85 tahun serta nilai maksimum yaitu 11,14 tahun. Hal ini menunjukkan bahwa nilai rata-rata lama sekolah berada pada kisaran 4,85-11,14 tahun. Selain itu, dapat dilihat pada variabel X₃ yaitu mengenai laju Produk Domestik Regional Bruto diperoleh nilai rata-rata sebesar -2,77%

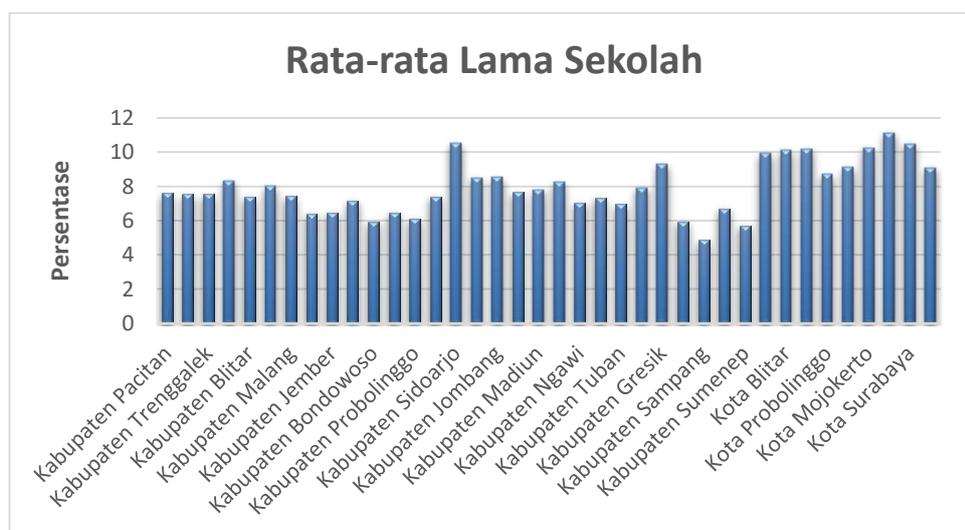
dengan standar deviasi yang bernilai 1,5204% yang artinya variabel X_3 bersifat heterogen karena sebaran pada variabel X_3 tidak merata, serta nilai minimum sebesar -6,46% dan nilai maximum sebesar -0.29%. Dapat juga dilihat untuk variabel X_4 yaitu mengenai tingkat pengangguran terbuka diperoleh nilai rata-rata sebesar 5,62% dengan standar deviasi sebesar 1,9769% yang artinya variabel X_4 bersifat homogen dan menunjukkan jumlah angka pengangguran terbuka yang cukup rendah serta nilai minimum sebesar 2,28% dan nilai maximum sebesar 10,97%.



Gambar 4.3 Bar Chart Indeks Pengeluaran Perkapita Jatim 2020
Sumber : (BPS, 2022c)

Pada Gambar 4.3 menunjukkan karakteristik variabel X_1 yaitu indeks pengeluaran perkapita berdasarkan Kabupaten/Kota di provinsi Jawa Timur tahun 2020. Pengeluaran perkapita merupakan pengeluaran yang dihasilkan dari satu

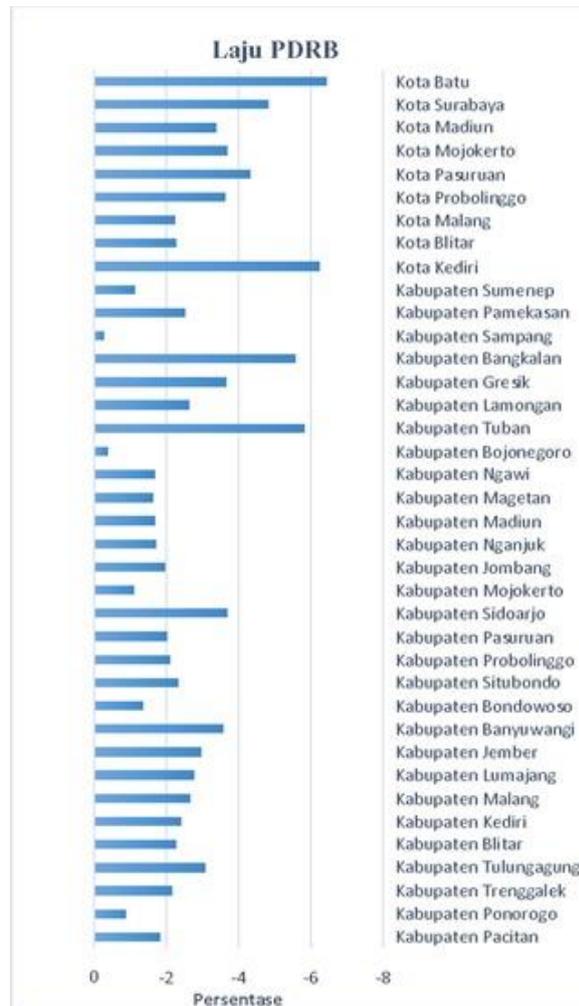
keluarga dalam satu bulan. Berdasarkan Gambar 4.3 dapat diketahui bahwa Kota Malang dan Kabupaten Ponorogo merupakan 2 daerah yang memiliki nilai pengeluaran perkapita tertinggi. Nilai pengeluaran perkapita yang dikeluarkan oleh penduduk Kota Malang dan Kabupaten Ponorogo selama satu tahun sebesar 0,28%. Kemudian untuk daerah yang memiliki nilai pengeluaran perkapita terendah yaitu Kabupaten Sumenep dengan nilai sebesar 0,11%.



Gambar 4.4 Bar Chart Rata-rata Lama Sekolah Jatim 2020
Sumber : (BPS, 2022d)

Pada Gambar 4.4 menunjukkan karakteristik variabel X_2 yaitu rata-rata lama sekolah berdasarkan Kabupaten/Kota di provinsi Jawa Timur tahun 2020. Semakin tinggi nilai rata-rata lama sekolah di suatu daerah maka rata-rata lama sekolah yang dihabiskan oleh penduduk usia lebih dari 25 tahun untuk bersekolah tinggi. Sehingga banyak penduduk di daerah tersebut beranggapan bahwa sekolah itu penting sehingga pendidikan yang diselesaikan tinggi. Berdasarkan Gambar 4.4 dapat diketahui bahwa daerah yang memiliki rata-rata lama sekolah paling tinggi yaitu Kota Madiun sebesar 11,14 tahun yang artinya penduduk usia lebih dari 25 tahun Kota Madiun telah menempuh pendidikan selama 11,14 tahun. Selanjutnya disusul oleh Kota Surabaya sebesar 10,49 tahun dan Kota Mojokerto

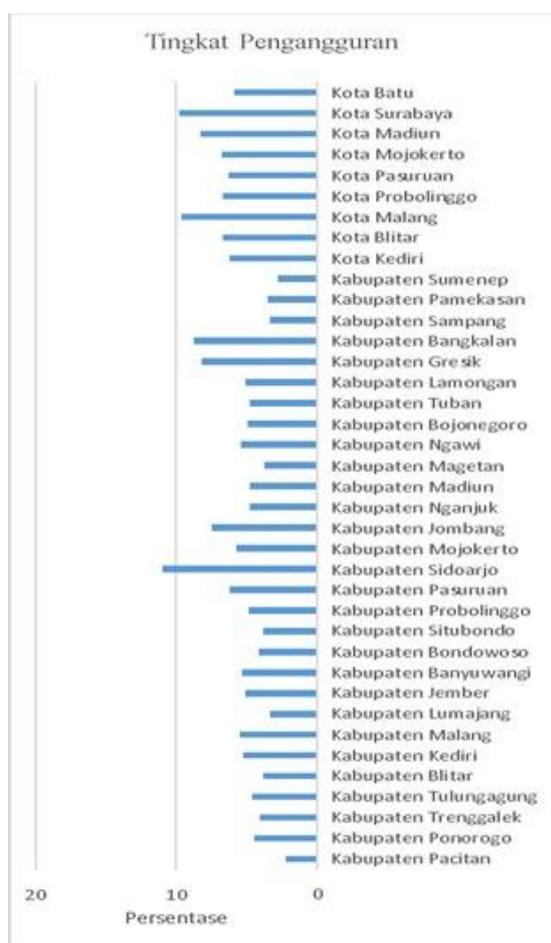
sebesar 10,25 tahun. Sementara itu, daerah yang memiliki rata-rata lama sekolah terendah yaitu Kabupaten Sampang dengan nilai sebesar 4,85 tahun.



Gambar 4.5 Bar Chart Laju PDRB Jatim 2020
Sumber : (BPS, 2022e)

Pada Gambar 4.5 menunjukkan karakteristik variabel X_3 yaitu laju Produk Domestik Regional Bruto (PDRB) berdasarkan Kabupaten/Kota di provinsi Jawa Timur tahun 2020. Kemampuan sumber daya ekonomi yang besar ditunjukkan dengan laju PDRB yang tinggi. Peningkatan pada PDRB menggambarkan perkembangan produksi barang dan jasa masyarakat. Hal tersebut akan mengakibatkan tingkat pengangguran menjadi menurun dan tingkat pendapatan menjadi meningkat sehingga menunjukkan meningkatnya tingkat kesejahteraan

masyarakat. Berdasarkan Gambar 4.5 dapat diketahui bahwa daerah Kabupaten Sampang memiliki perkembangan laju PDRB tertinggi sebesar -0,29% tahun dibandingkan dengan daerah lainnya. Kemudian disusul daerah Kabupaten Bojonegoro sebesar -0,4% dan Kabupaten Ponorogo sebesar -0,9%. Sementara itu, daerah yang memiliki laju PDRB terendah sebesar -6,46% yaitu Kota Batu.



Gambar 4. 6 Bar Chart Tingkat Pengangguran Terbuka Jatim 2020
Sumber : (BPS, 2022f)

Pada Gambar 4.6 menunjukkan karakteristik variabel X_4 yaitu tingkat pengangguran terbuka berdasarkan Kabupaten/Kota di provinsi Jawa Timur tahun 2020. Tingkat pengangguran terbuka mengidentifikasi besarnya persentase angkatan kerja yang termasuk dalam pengangguran. Semakin tinggi tingkat pengangguran terbuka maka makin banyak angkatan kerja yang tidak terserap

pada pasar kerja. Berdasarkan gambar diatas dapat diketahui bahwa daerah yang memiliki persentase tingkat pengangguran terbuka tertinggi sebesar 10,97% yaitu Kabupaten Sidoarjo. Kemudian disusul daerah Kota Surabaya dengan persentase sebesar 9,79% dan Kota Malang dengan persentase sebesar 9,61%. Sementara itu, daerah Kabupaten Pacitan memiliki persentase tingkat pengangguran terbuka terendah yaitu 2,28%.

4.2 Analisis Data

4.2.1 Transformasi Data

Pada proses tranformasi data akan dilakukan proses untuk mengubah suatu atribut menjadi bentuk/format yang sesuai dengan program yang akan dijalankan. Transformasi data akan dilakukan pada data indeks kedalaman kemiskinan (lampiran 1) untuk dijadikan sebagai label/kelas. Dalam mendapatkan label/kelas pada data indeks kedalaman kemiskinan dapat dilakukan dengan mencari nilai rata-rata dari indeks kedalaman kemiskinan. Adapun perhitungan dalam mencari nilai rata-rata sebagai berikut:

$$\text{Rata - rata} = \frac{\text{total indeks kedalaman kemiskinan}}{\text{banyak data}}$$

$$\text{Rata - rata} = \frac{64,3}{38} = 1,69$$

Setelah diperoleh rata-rata dari data indeks kedalaman kemiskinan, penentuan label/kelas dapat dilakukan dengan kriteria sebagai berikut:

Tabel 4. 2 Penentuan Label/Kelas

Label/Kelas	Indeks Kedalaman Kemiskinan (IKK)
Kelas 0	IKK Rata-rata < 1,69
Kelas 1	IKK Rata-rata >= 1,69

Tabel 4.2 menyajikan bahwa nilai indeks kedalaman kemiskinan di suatu Kabupaten/Kota $< 1,69$ artinya wilayah tersebut termasuk dalam kriteria klasifikasi kelas 0. Dan apabila nilai indeks kedalaman kemiskinan di suatu Kabupaten/Kota $\geq 1,69$ artinya kedalaman kemiskinan termasuk dalam kriteria klasifikasi kelas 1. Sehingga diperoleh hasil yang telah dilakukan transformasi data menjadi label/kelas kelas 0 dan kelas 1 pada data indeks kedalaman kemiskinan (data lengkapnya terdapat pada lampiran 2) dimana terdapat 23 wilayah di Provinsi Jawa Timur termasuk dalam kategori klasifikasi kemiskinan kelas 0 dan 15 wilayah lainnya termasuk dalam kategori klasifikasi kemiskinan kelas 1.

4.2.2 Standarisasi Data

Data yang akan digunakan harus dilakukan proses standarisasi. Standarisasi berbasis jarak dapat membantu dalam mencegah variabel yang awalnya memiliki interval tinggi dari variabel yang awalnya memiliki interval rendah. Dengan menggunakan rumus pada persamaan (2.3), maka semua variabel akan memiliki rentang nilai yang sama dengan interval $[0,1]$. Hasil perhitungan standarisasi data (lampiran 3 untuk hasil lengkapnya) sebagai berikut:

Hasil standarisasi untuk variabel X_1 :

$$X_{11}' = \frac{x_1 - x_{min}}{x_{max} - x_{min}} = \frac{0.24 - 0.11}{0.28 - 0.11} = 0.742381$$

$$X_{12}' = \frac{x_2 - x_{min}}{x_{max} - x_{min}} = \frac{0.28 - 0.11}{0.28 - 0.11} = 0.959041$$

$$X_{13}' = \frac{x_3 - x_{min}}{x_{max} - x_{min}} = \frac{0.26 - 0.11}{0.28 - 0.11} = 0.844288$$

⋮

$$X_{138}' = \frac{x_{38} - x_{min}}{x_{max} - x_{min}} = \frac{0.22 - 0.11}{0.28 - 0.11} = 0.612939$$

Hasil standarisasi untuk variabel X_2 :

$$X_{21}' = \frac{x_1 - x_{min}}{x_{max} - x_{min}} = \frac{7.6 - 4.85}{11.14 - 4.85} = 0.437202$$

$$X_{22}' = \frac{x_2 - x_{min}}{x_{max} - x_{min}} = \frac{7.54 - 4.85}{11.14 - 4.85} = 0.427663$$

$$X_{23}' = \frac{x_3 - x_{min}}{x_{max} - x_{min}} = \frac{7.55 - 4.85}{11.14 - 4.85} = 0.429253$$

⋮

$$X_{238}' = \frac{x_{38} - x_{min}}{x_{max} - x_{min}} = \frac{7.6 - 4.85}{11.14 - 4.85} = 0.670906$$

Hasil standarisasi untuk variabel X_3 :

$$X_{31}' = \frac{x_1 - x_{min}}{x_{max} - x_{min}} = \frac{(-1.84) - (-6.46)}{(-0.29) - (-6.46)} = 0.748784$$

$$X_{32}' = \frac{x_2 - x_{min}}{x_{max} - x_{min}} = \frac{(-0.9) - (-6.46)}{(-0.29) - (-6.46)} = 0.901135$$

$$X_{33}' = \frac{x_3 - x_{min}}{x_{max} - x_{min}} = \frac{(-2.17) - (-6.46)}{(-0.29) - (-6.46)} = 0.6953$$

⋮

$$X_{338}' = \frac{x_{38} - x_{min}}{x_{max} - x_{min}} = \frac{(-6.46) - (-6.46)}{(-0.29) - (-6.46)} = 0$$

Hasil standarisasi untuk variabel X_4 :

$$X_{41}' = \frac{x_1 - x_{min}}{x_{max} - x_{min}} = \frac{2.28 - 2.28}{10.97 - 2.28} = 0$$

$$X_{42}' = \frac{x_2 - x_{min}}{x_{max} - x_{min}} = \frac{4.45 - 2.28}{10.97 - 2.28} = 0.249712$$

$$X_{43}' = \frac{x_3 - x_{min}}{x_{max} - x_{min}} = \frac{4.11 - 2.28}{10.97 - 2.28} = 0.210587$$

⋮

$$X_{438}' = \frac{x_{38} - x_{min}}{x_{max} - x_{min}} = \frac{5.93 - 2.28}{10.97 - 2.28} = 0.420023$$

4.2.3 Implementasi Algoritma *K-Nearest Neighbor*

Pengklasifikasian dengan algoritma *K-Nearest Neighbor* (*K-NN*) dibagi menjadi dua proses, yaitu proses *training* dan *testing*. Pada proses *training*, *K-NN* menggunakan data sampel yang terdiri dari variabel-variabel dan kelas target yang diambil dari banyaknya kelas klasifikasi sebagai *input*. Sedangkan pada proses *testing*, *K-NN* menggunakan nilai perhitungan jarak untuk atribut-atribut dari setiap data uji terhadap seluruh atribut pada data *training* dengan rumus jarak *Euclidean* (persamaan 2.1). Hasil pengklasifikasian berdasarkan nilai parameter *k* akan dibandingkan untuk mendapatkan hasil akurasi terbaik. Selanjutnya dihasilkan sejumlah nilai *k* tetangga terdekat, dimana hasil dari data uji yang baru diklasifikasikan berdasarkan mayoritas dari kelas kategori pada *k* tetangga terdekat. Berikut merupakan implementasi algoritma *K-NN* dengan menggunakan 38 dataset Kabupaten/Kota di provinsi Jawa Timur sebagai data latih yang telah di standarisasi.

Tabel 4. 3 Data Latih 38 Kabupaten/Kota

No	Wilayah	X1	X2	X3	X4	Kelas/Label
1	Kabupaten Pacitan	0.74	0.44	0.75	0.00	Kelas 1
2	Kabupaten Ponorogo	0.96	0.43	0.90	0.25	Kelas 0
3	Kabupaten Trenggalek	0.84	0.43	0.70	0.21	Kelas 0
4	Kabupaten Tulungagung	0.60	0.55	0.55	0.27	Kelas 0
5	Kabupaten Blitar	0.59	0.40	0.68	0.18	Kelas 0
...
...
36	Kota Madiun	0.85	1.00	0.50	0.70	Kelas 0
37	Kota Surabaya	0.68	0.90	0.26	0.86	Kelas 0
38	Kota Batu	0.61	0.67	0.00	0.42	Kelas 0

Tabel 4. 4 Data Uji

No	Wilayah	X1	X2	X3	X4	Kelas/Label
1	Kota Malang	1.00	0.85	0.68	0.84	?

Sebelum melakukan perhitungan jarak, perlu menentukan nilai k tetangga terdekat yang akan digunakan. Pada permasalahan ini, nilai parameter k yang akan digunakan yaitu $k = 1$ sampai $k = 10$. Tahap berikutnya yaitu menghitung jarak berdasarkan *Euclidean Distance*.

Untuk data latih 1 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.74 - 1.00)^2 + (0.44 - 0.85)^2 + (0.75 - 0.68)^2 \\ + (0.00 - 0.84)^2 \end{matrix}} = 0.97273$$

Untuk data latih 2 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.96 - 1.00)^2 + (0.43 - 0.85)^2 + (0.90 - 0.68)^2 \\ + (0.25 - 0.84)^2 \end{matrix}} = 0.75796$$

Untuk data latih 3 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.84 - 1.00)^2 + (0.43 - 0.85)^2 + (0.70 - 0.68)^2 \\ + (0.21 - 0.84)^2 \end{matrix}} = 0.77414$$

Untuk data latih 4 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.60 - 1.00)^2 + (0.55 - 0.85)^2 + (0.55 - 0.68)^2 \\ + (0.27 - 0.84)^2 \end{matrix}} = 0.76929$$

Untuk data latih 5 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.59 - 1.00)^2 + (0.40 - 0.85)^2 + (0.68 - 0.68)^2 \\ + (0.18 - 0.84)^2 \end{matrix}} = 0.89789$$

Untuk data latih 36 dengan data uji:

$$D = \sqrt{\begin{matrix} (0.85 - 1.00)^2 + (1.00 - 0.85)^2 + (0.50 - 0.68)^2 \\ + (0.70 - 0.84)^2 \end{matrix}} = 0.31145$$

Untuk data latih 37 dengan data uji:

$$D = \sqrt{(0.68 - 1.00)^2 + (0.90 - 0.85)^2 + (0.26 - 0.68)^2 + (0.86 - 0.84)^2} = 0.53075$$

Untuk data latih 38 dengan data uji:

$$D = \sqrt{(0.61 - 1.00)^2 + (0.67 - 0.85)^2 + (0.00 - 0.68)^2 + (0.42 - 0.84)^2} = 0.90736$$

Tabel 4. 5 Perolehan Jarak Data Uji

Data ke-	Euclidean Distance	Ranking	Kelas/Label
1	7.78277	28	Kelas 1
2	5.95356	10	Kelas 0
3	6.09179	14	Kelas 0
4	5.39594	13	Kelas 0
5	6.4276	22	Kelas 0
...	
...	
36	1.96553	2	Kelas 0
37	2.61528	6	Kelas 0
38	5.69377	24	Kelas 0

Tabel 4.5 menunjukkan perolehan jarak data uji ke setiap data latih, tahap selanjutnya adalah mengurutkan setiap nilai jarak tersebut dari yang terkecil (terdekat) ke jarak terbesar. Kemudian dilakukan pemeriksaan ke setiap kelas berdasarkan inputan nilai k (tetangga terdekat). Karena pada penelitian ini nilai parameter k yang digunakan yaitu $k = 1$ sampai $k = 10$, maka yang diambil 10 terkecil. Berdasarkan Tabel 4.6 kelas yang mayoritas muncul saat $k = 1$ yaitu kelas 0, $k = 2$ yaitu kelas 0, $k = 3$ yaitu kelas 0, $k = 4$ yaitu kelas 0 dan seterusnya. Sehingga diperoleh hasil klasifikasi untuk data uji Kota Malang tergolong dalam kategori klasifikasi tingkat kemiskinan kelas 0 yang artinya kurang dari rata-rata.

Tabel 4. 6 Hasil Urutan Jarak

Data ke-	Euclidean Distance	Ranking	Kelas/Label
32	0.00	1	Kelas 0
36	1.96553	2	Kelas 0
15	2.00017	3	Kelas 0
31	2.18368	4	Kelas 0
35	2.61528	5	Kelas 0
...
...
14	3.20781	8	Kelas 0
7	4.02279	9	Kelas 0
2	3.54634	10	Kelas 0

4.2.4 Hasil Klasifikasi Menggunakan Algoritma *K-Nearest Neighbor*

Dalam penelitian ini menggunakan 38 data uji dengan nilai parameter $k = 1$ sampai $k = 10$ dengan pengujian *10-fold cross validation* untuk evaluasi algoritma *K-NN*. Sehingga diperoleh perbandingan hasil akurasi algoritma *K-NN* berdasarkan nilai parameter $k = 1$ sampai $k = 10$ ditampilkan pada Tabel 4.7 berikut:

Tabel 4. 7 Perbandingan Akurasi Parameter *K-NN*

Parameter k	Akurasi (%)
$k = 1$	76.67%
$k = 2$	76.67%
$k = 3$	60.83%
$k = 4$	66.67%
$k = 5$	65.83%
$k = 6$	63.33%
$k = 7$	60.83%
$k = 8$	66.67%
$k = 9$	64.17%
$k = 10$	64.17%

Pada Tabel 4.7 dapat diketahui bahwa akurasi terbaik dihasilkan oleh nilai parameter $k = 1$ dan $k = 2$ dengan nilai akurasi yang dihasilkan yaitu sebesar 76.67%.

4.2.5 Uji Akurasi Algoritma *K-Nearest Neighbor*

Untuk mengetahui performa atau kinerja dari algoritma *K-Nearest Neighbor* dalam melakukan klasifikasi terhadap suatu kelas/label yang telah ditentukan, maka akan dilakukan pengujian pada hasil akurasi. Pada pengujian ini, terdapat data yang berjumlah 38 Kabupaten/Kota. Hasil dari pengujian akurasi akan dimunculkan dalam bentuk *confusion matrix* 3 x 3 pada tabel 4.9.

Tabel 4. 8 *Confusion Matrix* Hasil Klasifikasi *K-NN*

<i>Confusion Matrix</i>		Nilai Aktual	
		POSITIF (Kelas Tinggi)	NEGATIF (Kelas Rendah)
Nilai Prediksi	POSITIF (Kelas Tinggi)	7	1
	NEGATIF (Kelas Rendah)	8	22

Untuk nilai akurasi berdasarkan Tabel *confusion matrix* hasil klasifikasi *K-Nearest Neighbor* diperoleh akurasi klasifikasi sebesar 76.32%. Berdasarkan nilai parameter $k = 1$, dan $k = 2$ kemudian untuk hasil akurasi parameter k lainnya dapat dilihat pada (lampiran 4). Tabel tersebut menunjukkan bahwa dapat mengklasifikasikan 7 wilayah kelas “positif” yang terprediksi benar sebagai kelas “positif” dan 22 wilayah kelas “negatif” terprediksi benar seperti kelas aslinya yaitu kelas “negatif”. Kemudian terdapat 8 wilayah kelas “positif” yang terprediksi salah yaitu kelas “negatif” dan 1 wilayah kelas “negatif” yang terprediksi tidak sesuai kelas aslinya yaitu “positif”. Perhitungan nilai akurasi

diperoleh dari jumlah data tepat hasil klasifikasi dibagi dengan total data dengan perhitungan sebagai berikut:

$$\begin{aligned}
 Akurasi &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \\
 &= \frac{7 + 22}{7 + 22 + 8 + 1} \times 100\% \\
 &= 76.32\%
 \end{aligned}$$

4.3 Konsep Klasifikasi dalam Islam

Seperti yang telah dijelaskan pada sub bab 2.6, bahwa “*Ambillah zakat dari harta mereka guna membersihkan dan menyucikan mereka*”. Yang dimaksudkan yaitu bahwasannya zakat telah disyariatkan untuk membersihkan diri dari harta yang bisa saja diperoleh dengan cara yang kurang tepat, sehingga mendorong pemilik harta agar slalu bersyukur kepada Allah SWT atas rezeki yang telah diberikan-Nya. Zakat merupakan salah satu bentuk hubungan sesama manusia dengan saling mengasihi, saling peduli dan saling menyayangi yang harus direalisasikan.

Dalam penerimaan zakat tidak semua orang atau golongan berhak dalam menerima zakat. Berikut telah diklasifikasikan golongan siapa saja yang berhak dalam menerima zakat yang telah dijelas dalam firman Allah SWT Q.S At-Taubah ayat 60 (Kemenag, 2022) yang artinya:

“Sesungguhnya zakat itu hanyalah untuk orang-orang fakir, orang miskin, amil, zakat, yang dilunakkan hatinya (mualaf), untuk memerdekakan hamba sahaya, untuk membebaskan orang yang berhutang, untuk yang berada di jalan Allah dan untuk orang yang sedang di dalam perjalanan sebagai kewajiban dari Allah. Allah Maha Mengetahui lagi Maha Bijaksana.”

Berdasarkan tafsir dari as-Sa'di bahwasannya golongan yang berhak mendapatkan zakat diklasifikasikan menjadi 8 golongan. Pertama dan kedua yaitu fakir dan miskin. Sebenarnya keduanya adalah golongan yang berbeda dimana fakir lebih membutuhkan daripada miskin, karena Allah telah menyebutkan dari awal bahwasannya fakir merupakan orang yang tidak memiliki apapun untuk memenuhi kebutuhannya, sedangkan miskin dia memiliki setengah lebih tetapi belum mencukupi. Golongan ketiga yaitu amil zakat, dia merupakan orang yang bekerja dan sibuk pada segala urusan zakat, sehingga dia berhak mendapatkan zakat sebagai gaji dari pekerjaan mengurus zakat. Golongan keempat yaitu orang yang sedang dibujuk hatinya. Kelima yaitu riqab, memerdekakan hamba sahaya. Keenam Gharim, jenis ini terbagi menjadi dua yaitu orang yang mengeluarkan harta demi mendamaikan dua kelompok yang berselisih dan orang yang berhutang harta untuk dirinya sendiri. Golongan ketujuh yaitu orang yang berperang di jalan Allah. Dan golongan yang kedelapan yaitu Ibnu Sabil, orang asing yang kehabisan bekal dan sedang tidak dinegerinya.

Berdasarkan penjelasan yang telah dipaparkan, telah menunjukkan bahwa konsep klasifikasi yang digunakan telah berdasarkan pada ajaran Islam. Sehingga diharapkan golongan orang kaya dapat membayarkan zakat dengan tepat waktu sesuai dengan ketentuan syar'i, pasti tidak ada lagi masyarakat atau golongan fakir dan miskin. Sehingga akan membantu pemerintah untuk memberikan solusi terbaik dalam mengurangi tingkat kemiskinan.

BAB V PENUTUP

5.1 Kesimpulan

Bedasarkan rumusan masalah dan hasil pembahasan pada bab sebelumnya, dapat ditarik kesimpulan sebagai berikut:

1. Tingkat kedalaman kemiskinan tertinggi di provinsi Jawa Timur pada tahun 2020 adalah Kabupaten Sumenep. Sedangkan tingkat kedalaman kemiskinan terendah berada pada wilayah Kota Batu. Terdapat 23 Kabupaten/Kota di provinsi Jawa Timur yang termasuk dalam kategori klasifikasi tingkat kemiskinan kelas 0 (kurang dari rata-rata) dan 15 Kabupaten/Kota lainnya termasuk dalam kategori klasifikasi tingkat kemiskinan kelas 1 (lebih dari rata-rata). Semakin tinggi indeks pengeluaran perkapita, laju PDRB dan rata-rata lama sekolah yang dikeluarkan di suatu wilayah maka memperlihatkan adanya peningkatan kesejahteraan masyarakat dan kualitas SDM di wilayah Kabupaten/Kota tersebut. Sedangkan semakin tinggi tingkat pengangguran terbuka di suatu wilayah maka memperlihatkan penurunan pada tingkat kesejahteraan masyarakat di wilayah Kabupaten/Kota tersebut.
2. Hasil klasifikasi berdasarkan algoritma *K-Nearest Neighbor* menunjukkan bahwa algoritma *K-Nearest Neighbor* telah berhasil diimplementasikan dengan baik untuk permasalahan klasifikasi tingkat kedalaman kemiskinan pada wilayah provinsi Jawa Timur tahun 2020. Hasil dari klasifikasi algoritma *KNN* menunjukkan akurasi tertinggi sebesar 76.67% berdasarkan nilai parameter k terbaik yaitu $k = 1$ dan $k = 2$.

5.2 Saran

Dalam rangka pengembangan dan penyempurnaan penelitian ini lebih lanjut, dengan demikian penulis memberikan saran untuk pengembangan selanjutnya sebagai berikut:

1. Menggunakan metode klasifikasi yang lain sehingga dapat membandingkan hasil uji akurasi yang dilakukan agar dapat menemukan metode klasifikasi yang terbaik.
2. Penelitian dapat dikembangkan dengan menambah jumlah data dan beberapa variabel dependen lainnya, sehingga dapat diperoleh hasil akurasi algoritma yang lebih baik.
3. Pada penelitian selanjutnya dapat dilakukan dengan mengubah jenis jarak (*distance*) yang digunakan serta parameter k .

DAFTAR PUSTAKA

- Asha Gowda Karegowda, M.A. Jayaram, & A.S. Manjunath. (2012). Cascading k-means clustering and k-nearest neighbor classifier for categorization of diabetic patients. *International Journal of Engineering and Advanced Technology*, 1(3), 147–151. <http://www.galaxy.gmu.edu/interface/I01/I2001Proceedings/Jbreault>
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Banjarsari, M. A., Budiman, H. I., & Farmadi, A. (2015). Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4. 2(2), 50–64.
- Bintang, A. B. M., & Woyanti, N. (2018). Pengaruh PDRB, Pendidikan, Kesehatan, Dan Pengangguran Terhadap Tingkat Kemiskinan Di Jawa Tengah (2011-2015). *Media Ekonomi Dan Manajemen*, 33(1), 20–28. <https://doi.org/10.24856/mem.v33i1.563>
- BPS. (2022a). *Badan Pusat Statistik*. <https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html>
- BPS. (2022b). *Badan Pusat Statistik*. <https://malangkota.bps.go.id/indicator/23/66/1/indeks-kedalaman-kemiskinan-p1-dan-indeks-keparahan-kemiskinan-p2-menurut-kab-kota-di-jawa-timur.html>
- BPS. (2022c). *BPS Kabupaten Lumajang*. <https://lumajangkab.bps.go.id/indicator/23/57/1/indeks-l-pengeluaran-perkapita-menurut-kabupaten-kota-di-jawa-timur-2018-2020.html>
- BPS. (2022d). *BPS Provinsi Jawa Timur*. <https://jatim.bps.go.id/indicator/26/32/1/rata-rata-lama-sekolah.html>
- BPS. (2022e). *BPS Provinsi Jawa Timur*. <https://jatim.bps.go.id/statictable/2021/05/28/2145/laju-pertumbuhan-produk-domestik-regional-bruto-atas-dasar-harga-konstan-2010-menurut-kabupaten-kota-di-provinsi-jawa-timur-persen-2017-2020.html>
- BPS. (2022f). *BPS Provinsi Jawa Timur*. <https://jatim.bps.go.id/indicator/6/54/1/tingkat-pengangguran-terbuka-tpt-provinsi-jawa-timur.html>

- Esesiawati, O. (2017). *Peramalan Arus Lalu Lintas Jangka Pendek Menggunakan Metode K-Nearest Neighbor (Studi Kasus : Jalan Basuki Rahmat Surabaya)*.
- Ferezagia, D. V. (2018). *Analisis Tingkat Kemiskinan. Jurnal Sosial Humaniora Terapan. 1(1)*, 1–6.
- Han, J., & Kamber, M. (2013). *Data Mining : Concepts and Techniques Third Edition, Morgan Kaufmann Publishers, San Fransisco*.
- Harun, R., Pelangi, K. C., & Lasena, Y. (2020). Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma K-Nearest Neighbor (KNN). *Jurnal Manajemen Informatika Dan Sistem Informasi, 3(1)*, 8–15. <https://doi.org/10.36595/MISI.V3I1.125>
- Hendayanti, N. P. N., & Nurhidayati, M. (2020). Regresi Logistik Biner dalam Penentuan Ketepatan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi-Provinsi di Indonesia. *Sainstek : Jurnal Sains Dan Teknologi, 12(2)*, 63–70. <https://doi.org/10.31958/js.v12i2.2483>
- Kemenag. (2022). *Qur'an Kemenag*. <https://quran.kemenag.go.id/>
- Kurnia, F., Kom, S., Kurniawan, J., & St, I. F. (2019). *Klasifikasi Keluarga Miskin Menggunakan Metode K- Nearest Neighbor Berbasis Euclidean Distance. 11(1)*, 230–239.
- Lall, U., & Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research, 32(3)*, 679–693. <https://doi.org/10.1029/95WR02966>
- Meade, N. (2002). A comparison of the accuracy of short term foreign exchange forecasting methods. *International Journal of Forecasting, 18(1)*, 67–83. [https://doi.org/10.1016/S0169-2070\(01\)00111-X](https://doi.org/10.1016/S0169-2070(01)00111-X)
- Meilina, P. (2015). Penerapan Data Mining dengan Metode Klasifikasi Menggunakan Decision Tree dan Regresi. *Jurnal Teknologi Universitas Muhammadiyah Jakarta, 7(1)*, 11–20. jurnal.ftumj.ac.id/index.php/jurtek
- Mukhlisin, M. F., Saputra, R., & Wibowo, A. (2017). Predicting house sale price using fuzzy logic, Artificial Neural Network and K-Nearest Neighbor. *Proceedings - 2017 1st International Conference on Informatics and Computational Sciences, ICICoS 2017, 17(1)*, 171–176. <https://doi.org/10.1109/ICICOS.2017.8276357>
- Nasution, D. Q., Yulmardi, Y., & ... (2017). Studi Tentang Kemiskinan di Kabupaten Batanghari dan Kabupaten Muaro Jambi Periode 2002-2016. *E-Jurnal Ekonomi ...*, 7(2), 79–90. <https://online-journal.unja.ac.id/JSEL/article/view/11923>

- Nilamsari, N. (2014). Memahami Studi Dokumen Dalam Penelitian Kualitatif. *Wacana*, 13(2), 177–181.
- Riwayati, S., & Hidayah, N. B. (2018). Zakat Dalam Telaah QS. At-Taubah: 103 (Penafsiran Enam Kitab). *Al Furqan: Jurnal Ilmu Al Quran Dan Tafsir*, 1(2), 77 – 91.
- Santoso. (2016). *Perbandingan Metode K-Nearest Neighbor (K-Nn) Dan Learning Vector Quantization (Lvq) Untuk Permasalahan Classification Problem.*
- Sarjono, S., Hartoyo, S., & Hakim, D. B. (2019). Strategi Penanggulangan Kemiskinan Di Kota Jakarta Timur. *Jurnal Manajemen Pembangunan Daerah*, 9(1), 25–29. https://doi.org/10.29244/jurnal_mpd.v9i1.27542
- Sasongko, T. B. (2016). *Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM.* 2(2), 244–253.
- Sreemathy, J. (2012). *an Efficient Text Classification Using Knn and Naive Bayesian.* 4(3), 392–396.
- Sumarlin, D. A. (2018). *Implementasi K-Nearest Neighbord Pada Rapidminer Untuk Prediksi Kelulusan Mahasiswa.* *Jurnal Teknologi Infoemasi.* 10(1), 35–41.
- Wang, H., Xu, P., & Zhao, J. (2021). Improved KNN algorithms of spherical regions based on clustering and region division. *Alexandria Engineering Journal*, 1–15. <https://doi.org/10.1016/j.aej.2021.09.004>
- Yusuf, M., Rangkuti, R., Alfansyuri, V., Gunawan, W., Informatika, T., Komputer, I., & Mercu Buana, U. (2021). *Penerapan Algoritma K-Nearest Neighbor (Knn) Dalam Memprediksi Dan Menghitung Tingkat Akurasi Data Cuaca Di Indonesia.* 2(2), 11–16.

LAMPIRAN

Lampiran 1. Data Indeks Kedalaman Kemiskinan

Wilayah	Indeks Pengeluaran Perkapita (X1)	Rata-rata Lama Sekolah (X2)	Laju PDRB (X3)	Tingkat Pengangguran (X4)
Kabupaten Pacitan	0.24	7.6	-1.84	2.28
Kabupaten Ponorogo	0.28	7.54	-0.9	4.45
Kabupaten Trenggalek	0.26	7.55	-2.17	4.11
Kabupaten Tulungagung	0.21	8.33	-3.09	4.61
Kabupaten Blitar	0.21	7.39	-2.29	3.82
Kabupaten Kediri	0.19	8.02	-2.41	5.24
Kabupaten Malang	0.26	7.42	-2.68	5.49
Kabupaten Lumajang	0.16	6.4	-2.79	3.36
Kabupaten Jember	0.18	6.48	-2.98	5.12
Kabupaten Banyuwangi	0.19	7.16	-3.58	5.34
Kabupaten Bondowoso	0.20	5.93	-1.36	4.13
Kabupaten Situbondo	0.18	6.46	-2.33	3.85
Kabupaten Probolinggo	0.24	6.11	-2.12	4.86
Kabupaten Pasuruan	0.25	7.4	-2.03	6.24
Kabupaten Sidoarjo	0.22	10.5	-3.69	10.97
Kabupaten Mojokerto	0.19	8.51	-1.11	5.75
Kabupaten Jombang	0.20	8.54	-1.98	7.48
Kabupaten Nganjuk	0.22	7.64	-1.71	4.80
Kabupaten Madiun	0.22	7.81	-1.69	4.80

Wilayah	Indeks Pengeluaran Perkapita (X1)	Rata-rata Lama Sekolah (X2)	Laju PDRB (X3)	Tingkat Pengangguran (X4)
Kabupaten Magetan	0.18	8.24	-1.64	3.74
Kabupaten Ngawi	0.20	7.06	-1.69	5.44
Kabupaten Bojonegoro	0.16	7.33	-0.4	4.92
Kabupaten Tuban	0.24	6.95	-5.85	4.81
Kabupaten Lamongan	0.16	7.92	-2.65	5.13
Kabupaten Gresik	0.15	9.3	-3.68	8.21
Kabupaten Bangkalan	0.17	5.95	-5.59	8.77
Kabupaten Sampang	0.13	4.85	-0.29	3.35
Kabupaten Pamekasan	0.17	6.69	-2.54	3.49
Kabupaten Sumenep	0.11	5.71	-1.13	2.84
Kota Kediri	0.21	9.93	-6.25	6.21
Kota Blitar	0.22	10.11	-2.28	6.68
Kota Malang	0.28	10.18	-2.26	9.61
Kota Probolinggo	0.17	8.7	-3.64	6.70
Kota Pasuruan	0.16	9.12	-4.33	6.33
Kota Mojokerto	0.23	10.25	-3.69	6.74
Kota Madiun	0.26	11.14	-3.39	8.32
Kota Surabaya	0.23	10.49	-4.85	9.79
Kota Batu	0.22	9.07	-6.46	5.93

Lampiran 2. Data Setelah Tranformasi Data

Wilayah	Indeks Pengeluaran Perkapita (X1)	Rata-rata Lama Sekolah (X2)	Laju PDRB (X3)	Tingkat Pengangguran (X4)	Kelas/Label
Kabupaten Pacitan	0.24	7.6	-1.84	2.28	Kelas 1
Kabupaten Ponorogo	0.28	7.54	-0.9	4.45	Kelas 0
Kabupaten Trenggalek	0.26	7.55	-2.17	4.11	Kelas 0
Kabupaten Tulungagung	0.21	8.33	-3.09	4.61	Kelas 0
Kabupaten Blitar	0.21	7.39	-2.29	3.82	Kelas 0
Kabupaten Kediri	0.19	8.02	-2.41	5.24	Kelas 0
Kabupaten Malang	0.26	7.42	-2.68	5.49	Kelas 0
Kabupaten Lumajang	0.16	6.4	-2.79	3.36	Kelas 0
Kabupaten Jember	0.18	6.48	-2.98	5.12	Kelas 0
Kabupaten Banyuwangi	0.19	7.16	-3.58	5.34	Kelas 0
Kabupaten Bondowoso	0.20	5.93	-1.36	4.13	Kelas 1
Kabupaten Situbondo	0.18	6.46	-2.33	3.85	Kelas 1
Kabupaten Probolinggo	0.24	6.11	-2.12	4.86	Kelas 1
Kabupaten Pasuruan	0.25	7.4	-2.03	6.24	Kelas 0
Kabupaten Sidoarjo	0.22	10.5	-3.69	10.97	Kelas 0
Kabupaten Mojokerto	0.19	8.51	-1.11	5.75	Kelas 1
Kabupaten Jombang	0.20	8.54	-1.98	7.48	Kelas 1
Kabupaten Nganjuk	0.22	7.64	-1.71	4.80	Kelas 0
Kabupaten Madiun	0.22	7.81	-1.69	4.80	Kelas 0
Kabupaten Magetan	0.18	8.24	-1.64	3.74	Kelas 0

Wilayah	Indeks Pengeluaran Perkapita (X1)	Rata-rata Lama Sekolah (X2)	Laju PDRB (X3)	Tingkat Pengangguran (X4)	Kelas/Label
Kabupaten Ngawi	0.20	7.06	-1.69	5.44	Kelas 1
Kabupaten Bojonegoro	0.16	7.33	-0.4	4.92	Kelas 1
Kabupaten Tuban	0.24	6.95	-5.85	4.81	Kelas 1
Kabupaten Lamongan	0.16	7.92	-2.65	5.13	Kelas 1
Kabupaten Gresik	0.15	9.3	-3.68	8.21	Kelas 1
Kabupaten Bangkalan	0.17	5.95	-5.59	8.77	Kelas 1
Kabupaten Sampang	0.13	4.85	-0.29	3.35	Kelas 1
Kabupaten Pamekasan	0.17	6.69	-2.54	3.49	Kelas 1
Kabupaten Sumenep	0.11	5.71	-1.13	2.84	Kelas 1
Kota Kediri	0.21	9.93	-6.25	6.21	Kelas 0
Kota Blitar	0.22	10.11	-2.28	6.68	Kelas 0
Kota Malang	0.28	10.18	-2.26	9.61	Kelas 0
Kota Probolinggo	0.17	8.7	-3.64	6.70	Kelas 0
Kota Pasuruan	0.16	9.12	-4.33	6.33	Kelas 0
Kota Mojokerto	0.23	10.25	-3.69	6.74	Kelas 0
Kota Madiun	0.26	11.14	-3.39	8.32	Kelas 0
Kota Surabaya	0.23	10.49	-4.85	9.79	Kelas 0
Kota Batu	0.22	9.07	-6.46	5.93	Kelas 0

Lampiran 3. Data Setelah Standarisasi Data

Wilayah	Indeks Pengeluaran Perkapita (X1)	Rata-rata Lama Sekolah (X2)	Laju PDRB (X3)	Tingkat Pengangguran (X4)	Kelas/Label
Kabupaten Pacitan	0.74	0.44	0.75	0.00	Kelas 1
Kabupaten Ponorogo	0.96	0.43	0.90	0.25	Kelas 0
Kabupaten Trenggalek	0.84	0.43	0.70	0.21	Kelas 0
Kabupaten Tulungagung	0.60	0.55	0.55	0.27	Kelas 0
Kabupaten Blitar	0.59	0.40	0.68	0.18	Kelas 0
Kabupaten Kediri	0.48	0.50	0.66	0.34	Kelas 0
Kabupaten Malang	0.86	0.41	0.61	0.37	Kelas 0
Kabupaten Lumajang	0.27	0.25	0.59	0.12	Kelas 0
Kabupaten Jember	0.39	0.26	0.56	0.33	Kelas 0
Kabupaten Banyuwangi	0.44	0.37	0.47	0.35	Kelas 0
Kabupaten Bondowoso	0.52	0.17	0.83	0.21	Kelas 1
Kabupaten Situbondo	0.38	0.26	0.67	0.18	Kelas 1
Kabupaten Probolinggo	0.73	0.20	0.70	0.30	Kelas 1
Kabupaten Pasuruan	0.79	0.41	0.72	0.46	Kelas 0
Kabupaten Sidoarjo	0.65	0.90	0.45	1.00	Kelas 0
Kabupaten Mojokerto	0.47	0.58	0.87	0.40	Kelas 1
Kabupaten Jombang	0.54	0.59	0.73	0.60	Kelas 1
Kabupaten Nganjuk	0.62	0.44	0.77	0.29	Kelas 0
Kabupaten Madiun	0.64	0.47	0.77	0.29	Kelas 0
Kabupaten Magetan	0.41	0.54	0.78	0.17	Kelas 0

Kabupaten Ngawi	0.50	0.35	0.77	0.36	Kelas 1
Kabupaten Bojonegoro	0.28	0.39	0.98	0.30	Kelas 1
Kabupaten Tuban	0.74	0.33	0.10	0.29	Kelas 1
Kabupaten Lamongan	0.27	0.49	0.62	0.33	Kelas 1
Kabupaten Gresik	0.23	0.71	0.45	0.68	Kelas 1
Kabupaten Bangkalan	0.36	0.17	0.14	0.75	Kelas 1
Kabupaten Sampang	0.09	0.00	1.00	0.12	Kelas 1
Kabupaten Pamekasan	0.36	0.29	0.64	0.14	Kelas 1
Kabupaten Sumenep	0.00	0.14	0.86	0.06	Kelas 1
Kota Kediri	0.60	0.81	0.03	0.45	Kelas 0
Kota Blitar	0.64	0.84	0.68	0.51	Kelas 0
Kota Malang	1.00	0.85	0.68	0.84	Kelas 0
Kota Probolinggo	0.33	0.61	0.46	0.51	Kelas 0
Kota Pasuruan	0.27	0.68	0.35	0.47	Kelas 0
Kota Mojokerto	0.67	0.86	0.45	0.51	Kelas 0
Kota Madiun	0.85	1.00	0.50	0.70	Kelas 0
Kota Surabaya	0.68	0.90	0.26	0.86	Kelas 0
Kota Batu	0.61	0.67	0.00	0.42	Kelas 0

Lampiran 4. Hasil Akurasi Klasifikasi Menggunakan *Software Rapidminer*

Parameter $k = 1$

accuracy: 76.67% +/- 18.76% (micro average: 76.32%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	7	1	87.50%
pred. Kode 0	8	22	73.33%
class recall	46.67%	95.65%	

Parameter $k = 2$

accuracy: 76.67% +/- 18.76% (micro average: 76.32%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	7	1	87.50%
pred. Kode 0	8	22	73.33%
class recall	46.67%	95.65%	

Parameter $k = 3$

accuracy: 60.83% +/- 19.66% (micro average: 60.53%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	4	4	50.00%
pred. Kode 0	11	19	63.33%
class recall	26.67%	82.61%	

Parameter $k = 4$

accuracy: 66.67% +/- 16.67% (micro average: 65.79%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	4	2	66.67%
pred. Kode 0	11	21	65.62%
class recall	26.67%	91.30%	

Parameter $k = 5$

accuracy: 65.83% +/- 16.41% (micro average: 65.79%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	5	3	62.50%
pred. Kode 0	10	20	66.67%
class recall	33.33%	86.96%	

Parameter $k = 6$

accuracy: 63.33% +/- 11.92% (micro average: 63.16%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	4	3	57.14%
pred. Kode 0	11	20	64.52%
class recall	26.67%	86.96%	

Parameter $k = 7$

accuracy: 60.83% +/- 11.82% (micro average: 60.53%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	3	3	50.00%
pred. Kode 0	12	20	62.50%
class recall	20.00%	86.96%	

Parameter $k = 8$

accuracy: 66.67% +/- 16.67% (micro average: 65.79%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	4	2	66.67%
pred. Kode 0	11	21	65.62%
class recall	26.67%	91.30%	

Parameter $k = 9$

accuracy: 64.17% +/- 17.15% (micro average: 63.16%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	3	2	60.00%
pred. Kode 0	12	21	63.64%
class recall	20.00%	91.30%	

Parameter $k = 10$

accuracy: 64.17% +/- 17.15% (micro average: 63.16%)

	true Kode 1	true Kode 0	class precision
pred. Kode 1	3	2	60.00%
pred. Kode 0	12	21	63.64%
class recall	20.00%	91.30%	



BUKTI KONSULTASI SKRIPSI

Nama : Clariza Adelina Rachma
NIM : 18610077
Fakultas/Jurusan : Sains dan Teknologi/Matematika
Judul Skripsi : Implementasi Algoritma *K-Nearest Neighbor* Dalam Penentuan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur
Pembimbing I : Dr. Sri Harini M.Si
Pembimbing II : Erna Herawati, M.Pd

No	Tanggal	Hal	TandaTangan
1.	11 Januari 2022	Konsultasi BAB I	1.
2	24 Januari 2022	Revisi BAB I dan Konsultasi BAB III	2.
3.	27 Januari 2022	ACC BAB I dan BAB III	3.
4.	09 Februari 2022	Konsultasi Kajian Agama BAB I dan BAB II	4.
5.	23 Februari 2022	Konsultasi BAB II	5.
6.	10 Maret 2022	ACC Kajian Agama BAB I dan BAB II	6.
7.	29 Maret 2022	Konsultasi BAB II	7.
8.	07 April 2022	ACC BAB I, II, III	8.
9.	28 April 2022	Konsultasi BAB IV	9.
10.	20 Mei 2022	Konsultasi Kajian Agama BAB IV	10.
11.	24 Mei 2022	Konsultasi BAB IV	11.
12.	30 Mei 2022	Revisi BAB IV dan V	12.
13.	02 Juni 2022	ACC Kajian Agama BAB IV	13.
14.	10 Juni 2022	ACC Keseluruhan	14.

Malang, 29 Juni 2022

Mengetahui,
Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc
NIP.197411292000122005