

**ANALISIS *HYPERPARAMETER TUNING* PADA PERFORMA ALGORITMA  
*K-NEAREST NEIGHBOR* DALAM MEMPREDIKSI CURAH HUJAN.  
(STUDI KASUS: KOTA PEKANBARU)**

**SKRIPSI**

**Oleh :  
FADILLAH SIVA  
NIM. 19650055**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**ANALISIS *HYPERPARAMETER TUNING* PADA PERFORMA  
ALGORITMA *K-NEAREST NEIGHBOR* DALAM MEMPREDIKSI  
CURAH HUJAN. (STUDI KASUS: KOTA PEKANBARU)**

**SKRIPSI**

**Oleh :  
FADILLAH SIVA  
NIM. 19650055**

**Diajukan kepada:  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**HALAMAN PERSETUJUAN**

**ANALISIS *HYPERPARAMETER TUNING* PADA PERFORMA  
ALGORITMA *K-NEAREST NEIGHBOR* DALAM MEMPREDIKSI  
CURAH HUJAN. (STUDI KASUS: KOTA PEKANBARU)**

**SKRIPSI**

Oleh :  
**FADILLAH SIVA**  
NIM. 19650055

Telah Diperiksa dan Disetujui untuk Diuji:  
Tanggal 9 Maret 2023

Pembimbing I

  
(Okta Qomaruddin Aziz, M.Kom)  
NIP. 19911019 201903 1 013

Pembimbing II

  
(Dr. Irwan Budi Santoso, M.Kom)  
NIP. 19770103 201101 1 004

Mengetahui,  
Ketua Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



  
Dr. Fachri Kurniawan, M.MT. IPM  
19771020 200912 1 001

**HALAMAN PENGESAHAN**

**ANALISIS HYPERPARAMETER TUNING PADA PERFORMA ALGORITMA K-NEAREST NEIGHBOR DALAM MEMPREDIKSI CURAH HUJAN. (STUDI KASUS: KOTA PEKANBARU)**

**SKRIPSI**

Oleh :

**FADILLAH SIVA**

**NIM. 19650055**

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Sarjana Komputer (S.Kom) Tanggal: 10 April 2023

**Susunan Dewan Penguji**

**Susunan Dewan Penguji**

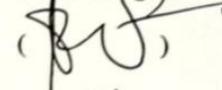
Ketua Penguji : (Prof. Dr. Suhartono, M.Kom)  
NIP. 19680519 200312 1 001

Anggota Penguji I : (Fajar Rohman Hariri, M.Kom)  
NIP. 19890515 201801 1 001

Anggota Penguji II : (Okta Qomaruddin Aziz, M.Kom)  
NIP. 19911019 201903 1 013

Anggota Penguji III : (Dr. Irwan Budi Santoso, M.Kom)  
NIP. 19770103 201101 1 004

**Tanda Tangan**

()  
()  
()  
()

Mengetahui dan Mengesahkan  
Program Studi Teknik Informatika



  
Dr. Fachri Kurniawan, M.MT. IPM  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Fadillah Siva  
NIM : 19650055  
Jurusan : Teknik Informatika  
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 9 Maret 2023

Yang membuat pernyataan,



Fadillah Siva

NIM. 19650055

**MOTTO**

“Be Yourself and Do It by Yourself”

## **HALAMAN PERSEMBAHAN**

**Puji Syukur ke hadirat Allah SWT, shalawat serta salam kepada junjungan  
alam Nabi Muhammad SAW. Penulis persembahkan sebuah karya ini  
kepada:**

Kepada orang tua penulis, Bapak Firman Saputra dan Ibu Sri Emilya yang selalu memberikan segala bentuk dukungan dan motivasi agar penulis mampu menyelesaikan perkuliahan selama empat tahun ini. Kepada keluarga besar yang telah banyak mendukung secara moril maupun materil.

Kepada bapak Okta Qomaruddin Azis, M.Kom selaku dosen pembimbing I dan bapak Dr. Irwan Budi Santoso, M.Kom selaku dosen pembimbing II yang telah dengan sabar dan tabah menghadapi penulis dalam proses bimbingan. Serta juga kepada bapak Prof. Dr. Suhartono, M.Kom dan bapak Fajar Rohman Hariri, M.kom selaku dosen penguji dan sekaligus secara tidak langsung juga sebagai pembimbing penulis. Serta kepada seluruh dosen Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah mengajarkan ilmu-ilmu yang insyaAllah sangat bermanfaat ke depannya.

Kepada istri serta anak dan cucu di masa depan yang telah menjadi motivasi penulis walaupun belum menjumpainya. Serta kepada teman-teman Teknik Informatika angkatan 2019 dan angkatan lainnya yang telah memberi semangat juang untuk dalam menjalankan perkuliahan, saya ucapkan terima kasih

## KATA PENGANTAR

*Assalamu'alaikun Wr. Wb.*

Syukur alhamdulillah penulis haturkan kehadiran Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Skripsi ini dengan baik.

Selanjutnya penulis haturkan ucapan terima kasih seiring do'a dan harapan jazakumullah ahsanal jaza' kepada semua pihak yang telah membantu terselesaikannya Skripsi ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Prof. Dr. HM. Zainuddin MA, selaku rektor UIN Maulana Malik Ibrahim Malang, yang telah banyak memberikan pengetahuan dan pengalaman yang berharga.
2. Dr. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT. IPM, selaku Ketua Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Okta Qomaruddin Aziz, M.Kom dan Dr. Irwan Budi Santoso, M.Kom selaku dosen pembimbing Skripsi, yang telah banyak memberikan pengarahan dan pengalaman yang berharga
5. Segenap sivitas akademika Program Studi Teknik Informatika, terutama seluruh dosen, terima kasih atas segenap ilmu dan bimbingannya.
6. Papa, Mama, serta keluarga tercinta yang senantiasa memberikan semangat, doa dan restunya kepada penulis dalam menuntut ilmu.
7. Teman-teman Alien 19, EkataTech, GabutParah, dan GEMMA yang selalu memberikan semangat kepada penulis untuk menyelesaikan Skripsi ini.
8. Semua pihak yang ikut membantu dalam menyelesaikan Skripsi ini baik berupa materiil maupun moril.

Penulis menyadari bahwa dalam penyusunan Skripsi ini masih terdapat kekurangan dan penulis berharap semoga Skripsi ini bisa memberikan manfaat

kepada para pembaca khususnya bagi penulis secara pribadi. *Amin Ya Rabbal Alamin.*

*Waasalamu'alaikum Wr. Wb.*

Malang, 9 Maret 2023

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGAJUAN</b> .....	ii
<b>HALAMAN PERSETUJUAN</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>HALAMAN PERNYATAAN</b> .....	v
<b>MOTTO</b> .....	vi
<b>HALAMAN PERSEMBAHAN</b> .....	vii
<b>KATA PENGANTAR</b> .....	viii
<b>DAFTAR ISI</b> .....	x
<b>DAFTAR GAMBAR</b> .....	xii
<b>DAFTAR TABEL</b> .....	xiii
<b>ABSTRAK</b> .....	xiv
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Pernyataan Masalah .....	8
1.3 Tujuan Penelitian .....	8
1.4 Manfaat Penelitian .....	8
1.5 Batasan Masalah .....	8
1.6 Sistematika Penulisan .....	8
<b>BAB II STUDI PUSTAKA</b> .....	10
2.1 Penelitian Terdahulu .....	10
2.2 Prediksi .....	13
2.3 Curah Hujan .....	14
2.4 Z Score .....	15
2.5 <i>K-Nearest Neighbor</i> .....	16
2.6 Root Mean Squared Error (RMSE).....	20
<b>BAB III DESAIN DAN IMPLEMENTASI</b> .....	22
3.1 Metode Penelitian .....	22
3.1.1 Studi Literatur .....	22
3.1.2 Pengumpulan Data .....	23
3.1.3 Implementasi.....	23
3.1.4 Evaluasi Kinerja KNN .....	31
3.2 Desain <i>Interface</i> .....	31
<b>BAB IV UJI COBA DAN PEMBAHASAN</b> .....	36
4.1 Skenario Uji Coba .....	36
4.2 Hasil Uji Coba .....	38
4.3 Hasil Analisis.....	40
4.4 Pembahasan .....	45
4.5 Integrasi Prediksi Hujan dengan Islam.....	50
<b>BAB V KESIMPULAN DAN SARAN</b> .....	60
5.1 Kesimpulan.....	60
5.2 Saran .....	60

**DAFTAR PUSTAKA**  
**LAMPIRAN**

## DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian .....	22
Gambar 3. 2 <i>Code</i> proses pengolahan data pada bahasa R .....	23
Gambar 3. 3 Desain Implementasi .....	24
Gambar 3. 4 <i>Code Z Score</i> pada bahasa R .....	25
Gambar 3. 5 <i>Code Euclidean</i> dan <i>Manhattan Distance</i> pada bahasa R.....	26
Gambar 3. 6 Alur untuk mencari tetangga terdekat .....	27
Gambar 3. 7 <i>Code</i> untuk mencari tetangga terdekat pada bahasa R .....	28
Gambar 3. 8 Alur untuk menghitung nilai prediksi KNN pada bahasa R.....	29
Gambar 3. 9 <i>Code</i> untuk menghitung nilai prediksi KNN pada bahasa R .....	30
Gambar 3. 10 <i>Code</i> untuk evaluasi KNN dengan RMSE pada bahasa R .....	31
Gambar 3. 11 Tampilan awal dari interface .....	32
Gambar 3. 12 Tampilan input data acuan .....	32
Gambar 3. 13 Tampilan input data prediksi .....	33
Gambar 3. 14 Tampilan halaman untuk menentukan parameter prediksi .....	34
Gambar 3. 15 Tampilan halaman untuk melihat hasil prediksi .....	34
Gambar 4.1 <i>Line Plot RMSE Euclidean Distance</i> .....	41
Gambar 4.2 <i>Line Plot RMSE Manhattan Distance</i> .....	42
Gambar 4.3 <i>Line Plot RMSE Metode Prediksi Average</i> .....	43
Gambar 4.4 <i>Line Plot RMSE Metode Prediksi Weighted</i> .....	44

## DAFTAR TABEL

Tabel 2. 1 Studi Literatur .....	10
Tabel 2. 2 Contoh Pemberian Bobot .....	20
Tabel 4. 1 Skenario Pembagian Data .....	36
Tabel 4. 2 Sub-percobaan metode perhitungan jarak .....	37
Tabel 4. 3 Sub-percobaan variasi perhitungan prediksi .....	37
Tabel 4. 4 Spesifikasi Perangkat .....	38
Tabel 4. 5 Contoh Perbedaan Hasil Prediksi pada masing-masing skenario .....	38
Tabel 4. 6 Contoh data perbandingan hasil prediksi pada setiap variasi nilai k ...	46
Tabel 4. 7 Contoh perbandingan prediksi variasi perhitungan prediksi.....	48
Tabel 4. 8 Contoh perbandingan target data acuan dan tetangga terdekat .....	48

## ABSTRAK

Siva, Fadillah, 2023. **Analisis Performa Algoritma *K-Nearest Neighbor* Dalam Memprediksi Curah Hujan (Studi kasus Kota Pekanbaru)**. Skripsi, Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (I) Okta Qomaruddin Aziz, M.Kom. (II) Dr. Irwan Budi Santoso, M.Kom.

Kata Kunci : Analisis Performa, K-Nearest-Neighbor, Prediksi, RMSE, Curah Hujan, Analisis, Weighted Distance.

Iklm saat ini bisa dikatakan sangat tidak menentu. Hal ini mengakibatkan curah hujan terkadang tinggi dan terkadang rendah, maka sulit untuk memperkirakan kadar curah hujan. Hal ini dapat diatasi dengan kecerdasan buatan yang membuat curah hujan yang mendatang bisa diprediksi dengan faktor-faktor yang dapat menentukan curah hujan. Beberapa faktor yang dapat mempengaruhi curah hujan adalah temperatur, kelembaban udara, serta kecepatan angin. KNN atau *K-Nearest Neighbor* adalah salah satu algoritma yang dapat digunakan untuk melakukan prediksi. Algoritma ini banyak diimplementasikan untuk kasus prediksi maupun klasifikasi berdasarkan data tetangga terdekat. KNN memiliki perhitungan yang mudah dipahami dan sederhana. Hasil dari pengujian model KNN untuk prediksi curah hujan dengan 3 faktor yang menentukan curah hujan memiliki beragam nilai error yang dihasilkan pada masing-masing parameter yang digunakan. Penelitian dilakukan dengan memvariasikan beberapa parameter yaitu nilai K, metode penghitungan jarak antar data, serta metode untuk penentuan hasil prediksi pada KNN. Skema terbaik dihasilkan pada saat parameter K=23, metode perhitungan jarak menggunakan *Euclidean Distance*, dan penentuan nilai prediksi KNN menggunakan *Average* dengan RMSE sebesar 15,98 dengan data sebagai acuan 50% dan sebagai testing 50%.

## ABSTRACT

Siva, Fadillah, 2023. **Performance Analysis of the K-Nearest Neighbor Algorithm in Predicting Rainfall (Case Study of Pekanbaru City)**. Undergraduate Thesis, Department of Informatics Engineering, Faculty of Science and Technology, State Islamic University of Maulana Malik Ibrahim Malang. Advisor : (I) Okta Qomaruddin Aziz, M.Kom. (II) Dr. Irwan Budi Santoso, M.Kom..

The current climate can be said to be very unpredictable. This results in sometimes high and sometimes low rainfall, making it difficult to predict rainfall levels. Artificial intelligence can help to predict upcoming rainfall by analyzing factors that affect precipitation, including temperature, air humidity, and wind speed. One algorithm that can be implemented for prediction is K-Nearest Neighbor (KNN). This algorithm is widely implemented for prediction and classification based on the nearest neighbor data. KNN has calculations that are easy to understand and simple. The results of testing the KNN model for predicting rainfall with three factors that determine rainfall produced various error values for each parameter used. Research was conducted by varying several parameters, including the value of K, the method of calculating distance between data, and the method for determining KNN prediction results. The best scheme was produced when the parameter K was 23, the distance calculation method used Euclidean Distance, and the determination of KNN prediction values used Average with an RMSE of 15,98 with 50% of the data used as reference and 50% used for testing.

**Key words :** Performance Analysis, K-Nearest-Neighbor, Predicting, RMSE, Rainfall, Analysis, Weighted Distance.

## خلاصة

شفاء, فضيلة, 2023 . تحليل أداء خوارزمية في التنبؤ بهطول الأمطار(دراسة حالة بمينة بيكانبارو) رسالة جامعية ، قسم الهندسة المعلوماتية ، كلية العلوم والتكنولوجيا ، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرفان : (I) أوكتا قمر الدين عزيز، الماجيستر. (II) د. إيروان بودي سانتوسو، الماجيستر

الكلمة المفتاحية : تحليل الأداء, K-Nearest-Neighbor, التنبؤ, RMSE, هطول الامطار, تحليل, Weighted Distance

يمكن القول إن المناخ الحالي غير معين. ينتج عن هذا هطول أمطار غزيرة و منخفضة في بعض الأحيان ,ثم من الصعب تقدير مستوى هطول الأمطار. يمكن تحليل هذه المشكلة عن طريق الذكاء الاصطناعي الذي يجعل هطول الأمطار في المستقبل قابلا للتنبؤ بالعوامل التي يمكن أن تحدد هطول الأمطار ومثل درجة الحرارة ورطوبة الهواء وسرعة الرياح. إحدى الخوارزميات التي يمكن تنفيذها في عمل التنبؤات هي ((K-Nearest neighbor (KNN) يتم تنفيذ هذه الخوارزمية على نطاق واسع لحالات التنبؤ والتصنيف بناء على بيانات من الجيران القريبين. لدى (KNN) حسابات سهلة الفهم وبسيطة. نتائج اختبار نموذج (KNN) للتنبؤ بهطول الأمطار مع 3 عوامل تحدد هطول الأمطار لها قيم خطأ مختلفة تم إنشاؤها في كل معلمة مستخدمة. أجريت الدراسة من خلال تغيير العديد من المعلمات وهي قيم (K)، طريقة حساب المسافات بين البيانات، وكذلك طرق تحديد نتائج التنبؤ على (KNN). يتم إنشاء أفضل مخطط عندما تكون المعلمة (K=23). طريقة حساب المسافة باستخدام مسافة ماثلاتن(Euclidean Distance) ، وتحديد قيمة التنبؤ (KNN) باستخدام المسافة المرجحة مع (RMSE) من 15,98 مع البيانات كمرجع بنسبة 50% وكاختبار 50%.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Hujan adalah elemen krusial untuk aktifitas harian manusia. Informasi mengenai cuaca menjadi sorotan yang kian hari semakin diperhatikan oleh masyarakat seiring naiknya kasus fenomena alam ekstrim. Segala aktifitas dapat terhalang jika terjadi cuaca ekstrim, seperti hujan yang dapat mengganggu jadwal petani menjemur padi, panas terik matahari yang dapat membuat ladang sawah hingga terjadinya gagal panen oleh petani, dan sebagainya. Menurut Mirawati, dampak dari fenomena yang dihasilkan oleh cuaca ekstrim ini dapat di minimalisir jika informasi yang dapat memprediksi peluang dari cuaca ekstrim ini tersedia dengan luas seperti prediksi gelombang tinggi yang dapat membantu nelayan menentukan waktu untuk berlayar, prediksi curah hujan yang dapat membantu petani menentukan waktu panen dan menjemur padi, prediksi gempa bumi, angin kencang, dan prediksi-prediksi cuaca lainnya (Mirawati et al., 2013).

Pada QS. Qaf ayat 9 Allah SWT menjelaskan tentang hujan:

وَنَزَّلْنَا مِنَ السَّمَاءِ مَاءً مُّبْرَكًا فَأَنْبَتْنَا بِهِ جَنَّاتٍ وَحَبَّ الْحَصِيدِ

*"Dan dari langit Kami turunkan air yang memberi berkah lalu Kami tumbuhkan dengan (air) itu pepohonan yang rindang dan biji-bijian yang dapat dipanen."* (Surah Qaf ayat 9).

Ibnu Katsir menjelaskan tafsir ayat ini pada kitab tafsirnya yaitu hujan diturunkan dan menumbuhkan taman-taman, kebun-kebun, dan sebagainya dari hujan tersebut. Lalu disambung dengan kalimat yang memiliki arti tumbuhan menghasilkan sesuatu berupa sumber makanan yang bisa disimpan pada waktu yang cukup lama (Muhammad, 2005).

Selain pada surat Qaf ayat 9, pada surat Al-Baqarah ayat 22 menjelaskan hujan merupakan bukti kekuasaan Allah:

الَّذِي جَعَلَ لَكُمُ الْأَرْضَ فِرَاشًا وَالسَّمَاءَ بِنَاءً وَأَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجَ بِهِ مِنَ الثَّمَرَاتِ رِزْقًا لَكُمْ ۗ فَلَا تَجْعَلُوا لِلَّهِ أُندَادًا وَأَنْتُمْ تَعْلَمُونَ

*"Dialah yang menjadikan Bumi sebagai hamparan bagimu dan langit sebagai atap, dan Dia menurunkan air (hujan) dari langit, lalu Dia menghasilkan dengan hujan itu segala buah-buahan sebagai rezeki untukmu, karena itu janganlah kamu mengadakan sekutu-sekutu bagi Allah, padahal kamu mengetahui." (Surah Al-Baqarah Ayat 22).*

Menurut tafsir yang dijelaskan oleh Ibnu Katsir, ayat ini berisi penjelasan Allah tentang keesaan-Nya dan senantiasa memberikan nikmat kepada manusia berupa menurunkan air dari langit berupa hujan sebagai sumber kehidupan bagi manusia di bumi. Dan dari air itu tumbuh lah tanaman yang memberi hasil berupa buah-buahan untuk rezeki bagi semua manusia (Muhammad, 2005).

Ayat tersebut menggambarkan bumi terbentuk sebagai hamparan luas di mana berbagai makhluk hidup tinggal dengan atap berupa langit. Tuhan memberikan anugerah dengan menurunkan hujan dari langit dan berbagai jenis tanaman yang dapat menghasilkan berbagai macam sayuran dan buahan ditumbuhkan. Qur'an surat An-Nur ayat 43 menjelaskan juga proses terjadinya hujan:

أَلَمْ تَرَ أَنَّ اللَّهَ يُرْسِطُ السَّحَابًا ثُمَّ يُؤَلِّفُ بَيْنَهُ ثُمَّ يَجْعَلُهُ رُكَامًا فَتَرَى الْوَدْقَ يَخْرُجُ مِنْ خِلَالِهِ وَيُنَزِّلُ مِنَ السَّمَاءِ مِنْ جِبَالٍ فِيهَا مِنْ بَرَدٍ فَيُصِيبُ بِهِ مَنْ يَشَاءُ وَيَصْرِفُهُ عَنِ مَنْ يَشَاءُ ۗ يَكَادُ سَنَا بَرْقِهِ يَذْهَبُ بِالْأَبْصَارِ

*"Tidaklah kamu melihat bahwa Allah mengarak awan, kemudian mengumpulkan antara (bagian-bagian)nya, kemudian menjadikannya bertindih-tindih, maka kelihatanlah olehmu hujan keluar dari celah-celahnya dan Allah (juga) menurunkan (butiran-butiran) es dari langit, (yaitu) dari (gumpalan-gumpalan awan seperti) gunung-gunung, maka ditimpakan-Nya (butiran-butiran) es itu kepada siapa yang dikehendaki-Nya dan dipalingkan-Nya dari siapa yang dikehendaki-Nya. Kilauan kilat awan itu hampir-hampir menghilangkan penglihatan." (Surah An-Nur Ayat 43).*

Menurut tafsir pada kitab Zubdatut Tafsir Min Fathil Qadir yang ditulis oleh Mudarris tafsir Universitas Islam Madinah yang bernama Syaikh Dr. Muhammad Sulaiman Al Asyqar, dijelaskan bahwa ayat ini menjelaskan tahapan terjadinya hujan, pertama Allah mengarahkan awan-awan lembut ke arah yang Allah kehendaki, lalu awan dikumpulkan hingga menjadi pekat dan besar dan bertumpuk. Kemudian hujan turun dari arah awan yang merupai gunung-gunung dan juga kilat tercipta dari awan itu.

Sedangkan menurut tafsir yang dijelaskan oleh Ibnu Katsir bahwa ayat ini berisi penjelasan tentang proses terjadinya hujan. Allah mengarak awan dengan kekuasaan-Nya. Allah menyatukan awan-awan yang bertebaran dan menumpuk awan-awan itu menjadi satu, dan turunlah hujan dari awan tersebut. Terdapat dua jenis yang diturunkan Allah dari langit, yaitu berupa cairan maupun butiran es atau salju. Allah menurunkan itu kepada setiap orang yang telah diberi rahmat, namun ditunda dari setiap orang yang telah dikehendaki. Allah memalingkan hujan dan salju bukan berarti tidak memberi rahmat orang-orang tersebut, namun karena jika yang diturunkan berlebih, maka itu akan menjadi bencana bagi mereka. Lalu muncul kilatan petir yang dapat menghilangkan penglihatan manusia jika dilihat (Muhammad, 2005).

Pada zaman maju yang serba modern dan berteknologi ini, dunia digital berkembang sangat pesat. Menurut Muhammad Danuri dalam *paper*-nya yang berjudul “ *Perkembangan dan Transformasi Teknologi Digital* ”, beliau menjelaskan bahwa teknologi berkembang dimulai dari hal sederhana dalam keseharian hingga pada *level* memenuhi kepuasan dan kebutuhan individu dan

kelompok. Seiring perpindahan masa, teknologi terus berkembang ke seluruh sektor, yaitu sektor pertanian, sektor industry, sektor informasi, hingga sektor komunikasi dan informasi (Danuri, 2019).

Salah satu kemajuan teknologi informasi adalah hadirnya *Data Mining* yang merupakan ekstraksi informasi dari suatu data (Susanto & Sudiyatno, 2014). *Data Mining* mencakup empat metode utama, yaitu Klasifikasi (*Classification*), Pengelompokan (*Clustering*), Asosiasi (*Association*), Regresi (*Regression*), dan Peramalan (*Forecasting*). Dalam *Data Mining* juga terdapat berbagai macam algoritma, seperti *SVM*, *KNN*, *Decision Tree*, dan sebagainya.

Karakteristik maupun sifat dari hujan adalah informasi penting bagi kehidupan manusia sehari-hari, karena akan memiliki berdampak terhadap aktifitas ekonomi pada masyarakat dalam banyak sektor. Kebutuhan dalam mengetahui curah hujan yang akan datang menjadikan prediksi curah hujan ini menjadi sangat penting. Hujan yang tidak menentu dapat memiliki resiko yang buruk pada sektor-sektor yang bergantung kondisi cuaca (Adhani, 2013).

Algoritma *KNN* yaitu singkatan dari *K – Nearest Neighbor* adalah suatu algoritma sangat populer sehingga masuk 10 algoritma *data mining* populer (Wu et al., 2008). Telah banyak penelitian yang mengungkapkan prediksi menggunakan algoritma *KNN* memiliki nilai akurasi yang tinggi bahkan melebihi algoritma lainnya seperti *Naïve Bayes*, dan *SVM*.

Penelitian yang melakukan prediksi menggunakan *KNN* banyak dilakukan para peneliti, beberapa contohnya yaitu sebuah yang dilakukan oleh Haerul Fatah dan Agung Subekti yang membahas tentang “*Prediksi harga Cryptocurrency*

*dengan metode KNN*”, yang mana penelitian ini menggunakan data *history* dari harga tahunan *Cryptocurrency* tahun sebelumnya. Penelitian ini menghasilkan *error* rata-rata sebesar -0,00135 pada  $K = 3$  (Fatah & Subekti, 2018).

Pada penelitian komparasi terhadap KNN dilakukan oleh Wilda Imama Salsabila dan Tesa Eranti Putri yang membahas tentang memprediksi kelulusan mahasiswa dengan metode *KNN* dan *Naïve Bayes*, yang mana penelitian ini menggunakan indeks prestasi semester 1 hingga 4 untuk menjadi tolak ukur dalam memprediksi kelulusan mahasiswa. Data *training* berupa data kelulusan . Penelitian ini menghasilkan akurasi menggunakan algoritma *Naïve Bayes* adalah 80.9% dan algoritma KNN lebih unggul dengan angka akurasi 98.7% (Sabilla & Putri, 2017).

Penelitian komparasi lainnya juga dilakukan oleh Shedriko yang membahas tentang “*Perbandingan Algoritma SVM, MLP, Decision Tree dan KNN dalam prediksi harga beras premium*”, yang mana data penelitian ini adalah data harga beras premium Badan Pusat Statistik dan data cuaca dari BMKG. Penelitian ini menghasilkan RMSE menggunakan algoritma *Support Vector Machine* sebesar 579.391, *Decision Tree* sebesar 1980.05, MLP sebesar 3954.792, sedangkan algoritma KNN unggul dengan angka RMSE sebesar 352.040 (Shedriko, 2021).

Menurut Marni dan Muhammad Ishak Jumarang dalam penelitiannya tentang hubungan suhu dan kelembaban udara terhadap hujan menjelaskan bahwa variabel berpengaruh terhadap curah hujan seperti kelembaban udara, kecepatan angin, dan suhu (Jumarang & others, 2016).

Jika faktor-faktor ini memiliki perubahan data, maka hasil dari prediksi curah hujan juga akan berubah. Curah hujan dapat diprediksi menggunakan KNN dengan menjadikan faktor-faktor diatas sebagai *feature* atau *x variables* dan curah hujan sebagai *label* atau *y variables*.

Pada perhitungan jarak terhadap data acuan dan data *testing* pada algoritma KNN, terdapat banyak metode seperti *Euclidean Distance* hingga *Manhattan Distance*. *Euclidian Distance* menghitung jarak antar data dengan menarik garis lurus dari kedua data, sedangkan *Manhattan Distance* menghitung jarak antar data dengan menjumlahkan selisih antara masing-masing variabel. Pada KNN penentuan nilai prediksi dilakukan dengan merata-rata kan data target pada data terdekat sebanyak k. Namun pada KNN ini juga terdapat beberapa kelemahan, diantaranya menurut Dudani pada penelitiannya yang pertama kali menemukan teknik *Weighted-Distance* pada KNN, KNN menganggap sama semua data terdekat yang telah dihimpun sebanyak k, yang mana data terdekat ke-k itu memiliki jarak yang jauh dari data yang akan diprediksi. Oleh karena itu, Dudani menerapkan metode *Weighted-Distance* pada KNN. Metode *Weighted-Distance* memberi pembobotan nilai target pada data terdekat. Bobot terbesar diberikan kepada data yang memiliki jarak terdekat, lalu data terdekat yang memiliki jarak lebih besar akan diberikan bobot lebih kecil. Hal ini terjadi karena diasumsikan bahwa data yang memiliki jarak terkecil seharusnya memiliki nilai target yang tidak jauh dengan data yang akan di prediksi (Dudani, 1976).

KNN memiliki banyak *hyper-parameter* yang dapat mempengaruhi nilai prediksi, seperti nilai K, metode perhitungan jarak antar data, maupun metode

perhitungan nilai prediksi. Oleh karena itu diperlukan lah *tuning* pada *hyper-parameter* untuk meningkatkan performa algoritma dalam memprediksi. *Hyper-parameter tuning* akan dilakukan dengan menganalisa performa KNN pada setiap variasi *hyper-parameter*. Analisis performa akan dilakukan dengan membandingkan hasil prediksi algoritma KNN yang menggunakan penentuan nilai prediksi dengan rata-rata tanpa pembobotan (di dinamakan dengan metode *Average*) dan algoritma KNN yang menggunakan *Weighted-Distance* (dinamakan dengan metode *Weighted*). Untuk pemberian bobot yang digunakan untuk penelitian yang akan dilakukan yaitu metode pembobotan *Ranked Order Centroid* (ROC). *Ranked Order Centroid* adalah teknik pembobotan yang digunakan dalam analisis multi-kriteria. Teknik ini menggunakan peringkat yang diberikan oleh responden untuk setiap kriteria, dan kemudian menghitung titik sentral dari peringkat tersebut agar dapat menentukan bobot relatif setiap kriteria. Variasi nilai K juga akan dilakukan untuk mengetahui K terbaik pada prediksi. Normalisasi menggunakan *Z Score* juga akan digunakan dalam penelitian memprediksi nilai Curah hujan pada Kota Pekanbaru, Provinsi Riau. Data yang akan diolah merupakan data kecepatan angin, kelembapan, dan temperature, yang mana ini diambil dari website resmi BMKG.

Hasil prediksi algoritma KNN akan dievaluasi untuk mengetahui seberapa besar *error* yang dihasilkan. Evaluasi hasil prediksi KNN dapat dilakukan dengan beberapa metode seperti metode *Root Mean Squared Error* (RMSE). RMSE adalah sebuah metode evaluasi dengan mengakar kuadrat kan *mean* selisih antara data sebenarnya dan data hasil prediksi yang dikuadratkan.

## 1.2 Pernyataan Masalah

Bersandar penjelasan latar belakang diatas, maka penelitian ini memiliki rumusan masalah yaitu berapa besar *error* yang dihasilkan antara algoritma KNN dan KNN yang menggunakan metode *Weighted-Distance* dalam memprediksi curah hujan di kota Pekanbaru?

## 1.3 Tujuan Penelitian

Berdasar perumusan masalah di atas, penelitian ini bertujuan menganalisa *error* antara algoritma KNN dan KNN yang menggunakan metode *Weighted-Distance* dalam memprediksi curah hujan pada kota Pekanbaru.

## 1.4 Manfaat Penelitian

Penelitian yang akan dilakukan ini diharap bermanfaat dalam menyajikan suatu informasi kepada setiap elemen masyarakat tentang prediksi curah hujan suatu wilayah di Indonesia sehingga suatu wilayah dapat menjadikan data tersebut sebagai dasar untuk mengambil kebijakan terkait iklim daerahnya.

## 1.5 Batasan Masalah

Agar penelitian bisa berfokus kepada permasalahan, maka penelitian diterapkan beberapa batasan permasalahan berikut:

1. Analisis performa yang dibahas hanya berdasarkan *error* yang dihasilkan pada setiap skema percobaan.
2. Data kelembaban, kecepatan angin, temperature, dan curah hujan yang didapatkan pada website resmi BMKG.
3. Data penelitian sebagai sample adalah data pada tahun 2019 hingga 2021 Kota Pekanbaru, Provinsi Riau.

## **1.6 Sistematika Penulisan**

Sistematika penulisan pada penelitian ini sebagai berikut :

### **Bab I Pendahuluan**

Bab ini menjelaskan bagian-bagian yang melatarbelakangi penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah pada penelitian, dan sistematika penulisan.

### **Bab II Studi Pustaka**

Bab ini menjelaskan penelitian sebelumnya yang berkaitan dengan penelitian yang didefinisikan oleh penulis. Bab ini memuat teori serta konsep dari berbagai sumber yang relevan.

### **Bab III Desain dan Implementasi**

Bab ini menjelaskan rancangan penelitian serta rancangan percobaan yang akan dilakukan, mulai dari studi literatur, pengambilan data, perancangan, serta desain sistem yang digunakan.

### **Bab IV Uji Coba dan Pembahasan**

Bab ini menjelaskan hasil serta pembahasan terkait uji coba yang dilakukan terhadap objek penelitian.

### **Bab V Kesimpulan dan Saran**

Bab ini berisi kesimpulan dari seluruh penelitian dan saran untuk penelitian selanjutnya agar penelitian menjadi lebih baik.

## BAB II

### STUDI PUSTAKA

#### 2.1 Penelitian Terdahulu

Tabel 2. 1 Studi Literatur

<b>Peneliti</b>	<b>Tahun</b>	<b>Algoritma</b>	<b>Target</b>	<b>Sitasi</b>
Widya Fadilah, Yufis Azhar, dan Dewi Agfiannisa	2020	KNN, SVM	Harga Saham	(Fadilah et al., 2020)
Haerul Fatah dan Agus Subekti	2018	KNN, Linear Regression, Random Forest, Decision Tree, Neural Network	Harga Cryptocurrency	(Fatah & Subekti, 2018)
Li Kuang, Xiaoliang Fan , Yujia Zhu, Shenmei Tu dan Han Yan	2019	Weighted-Distance KNN, cost-sensitive Bayesian Network	Durasi Kecelakaan Lalu Lintas	(Kuang et al., 2019)
Immanuel Jhonson Arizona Saragih, Dedi Sucahyono , Ricko Yudistira dan Inlim Rumahorbo	2020	Linear Regression	Curah Hujan	(Saragih et al., 2020)
Nola Ritha, Martaleli Bettiza, Ariel Dufan	2020	Levenberg-Marquardt Backpropagation	Curah Hujan	(Ritha et al., 2016)

Tabel 2.1 menunjukkan beberapa penelitian tentang penerapan metode KNN untuk memprediksi suatu kasus numerik telah dilakukan, diantaranya dilakukan oleh Widya Fadilah, Yufis Azhar, dan Dewi Agfiannisa pada tahun 2020. Penelitian yang dilaksanakan yaitu memprediksi harga dari saham PT. Telekomunikasi Indonesia dengan mengimplementasikan algoritma KNN dan SVM. Data diperoleh pada website yahoo.finance.com berupa harga dari saham PT. Telekomunikasi Indonesia, dengan variabel yaitu data harga dari saham harian mulai tanggal 29 April 2015 hingga data harga saham harian tanggal 28 April 2020 dengan atribut berupa *Date, Volume, Adj Close, Close, High Low, dan Open*. Sebelum diolah oleh algoritma KNN, pertama data di *preprocessing*. Setelah itu data diproses

menggunakan KNN dan SVM yang menghasilkan RMSE KNN sebesar 0.11 dengan akurasi sebesar 0.94 (Fadilah et al., 2020).

Penelitian penerapan metode KNN untuk memprediksi suatu kasus numerik juga dilakukan oleh Haerul Fatah dan Agus Subekti pada tahun 2018. penelitian yang dilakukan adalah memprediksi harga *Cryptocurrency* menggunakan KNN, *Linear Regression*, *Random Forest*, *Decision Tree*, dan *Neural Network*. Penelitian ini menggunakan data harga 3 jenis *Cryptocurrency*, yaitu *Bitcoin*, *Ripple*, dan *Ethereum* dengan atribut berupa *date*, *open*, *high*, *low*, dan *close*. Sebelum diproses, data dilakukan *preprocessing*. Setelah itu data diproses menggunakan KNN, *Linear Regression*, *Random Forest*, *Decision Tree*, dan *Neural Network*. Rata-rata *error* yang dihasilkan oleh KNN pada *Bitcoin* adalah -0,00135, pada *Ethereum* sebesar -0,00041, dan *Ripple* sebesar -0,00135 (Fatah & Subekti, 2018).

Penelitian penerapan metode KNN untuk memprediksi kasus numerik juga dilakukan oleh Li Kuang, Xiaoliang Fan , Yujia Zhu, Shenmei Tu dan Han Yan pada 2019. Penelitian yang dilakukan adalah memprediksi durasi kecelakaan lalu lintas menggunakan *Weighted-Distance KNN*, dan *cost-sensitive Bayesian Network*. Penelitian menggunakan data kecelakaan di kota Xiamen tahun 2015. Sebelum diproses, data dilakukan *preprocessing*. Lalu diproses menggunakan KNN dan *cost-sensitive Bayesian Network*. MAPE yang dihasilkan KNN adalah 0.391 (Kuang et al., 2019).

Tiga penelitian diatas akan dijadikan acuan untuk mengetahui bagaimana proses algoritma KNN dalam memprediksi. Perbedaan penelitian ini dengan tiga penelitian diatas adalah studi kasus prediksi yang dilakukan dan metode evaluasi

yang diterapkan.

Selain penelitian tentang penerapan KNN untuk memprediksi, pada tabel 2.1 terdapat juga penelitian tentang prediksi curah hujan menggunakan algoritma *machine learning* yang telah dilakukan, diantaranya dilakukan oleh Immanuel Jhonson Arizona Saragih, Dedi Sucahyono, Ricko Yudistira dan Inlim Rumahorbo pada tahun 2020. Penelitian yang dilakukan adalah memprediksi curah hujan menggunakan *Linear Regression*. Data yang digunakan yaitu data iklim tahun 1989 hingga 2018 yang didapat dari Stasiun Klimatologi Deli Serdang pada *website* resmi BMKG online (<http://bmksoft.database.bmkg.go.id>). Atribut data berupa rata-rata kelembapan udara bulanan, suhu udara dan total hujan bulanan. Sebelum diolah menggunakan *Linear Regression*, pertama data di *preprocessing*. Setelah itu data diproses menggunakan *Linear Regression* yang menghasilkan RMSE sebesar 80,53 mm/bulan (Saragih et al., 2020).

Penelitian tentang prediksi curah hujan menggunakan algoritma *machine learning* juga dilakukan oleh Nola Ritha, Martaleli Bettiza, Ariel Dufan pada tahun 2016. Penelitian yang dilakukan adalah memprediksi curah hujan mengimplementasikan *Levenberg-Marquardt* serta algoritma *Backpropagation*. Data penelitian ini adalah iklim bulanan mulai tahun 2010 hingga tahun 2014 didapat pada kantor BMKG Kota Tanjungpinang. Variabel berupa data kelembapan udara, suhu, curah hujan, kecepatan angin, dan tekanan udara. Data diproses menggunakan *Levenberg-Marquardt* dan *Backpropagation* yang menghasilkan RMSE *Backpropagation* sebesar 0,07876 dan *Levenberg-Marquardt* sebesar 0,0776 (Ritha et al., 2016).

Dua penelitian diatas akan dijadikan acuan untuk mengetahui variabel atau atribut apa saja yang dibutuhkan untuk memprediksi curah hujan. Algoritma yang diimplementasikan untuk memprediksi dan metode yang diberlakukan untuk evaluasi pada penelitian ini berbeda dengan dua penelitian diatas.

## **2.2 Prediksi**

Pengertian dari prediksi yaitu proses pengolahan informasi yang secara sistematis dapat memperkirakan apa kejadian yang mungkin saja terjadi pada waktu yang belum terjadi atau yang akan datang berdasarkan data-data dan informasi-informasi yang didapatkan pada fenomena telah terjadi pada masa lalu atau di masa lalu. Prediksi tidak selalu harus memberikan hasil pasti akan sesuatu yang akan terjadi, namun prediksi dapat berusaha sebaik mungkin untuk memberikan hasil kemungkinan yang paling mendekati kejadian yang bisa saja terjadi pada saat mendatang (Kafil, 2019).

Kamus Besar Bahasa Indonesia menjelaskan tentang pengertian prediksi yaitu suatu peramalan atau perkiraan suatu nilai di saat mendatang berdasarkan informasi yang lalu sebagai acuan Peramalan ataupun prediksi bertujuan untuk mendapatkan sebuah informasi tentang suatu perubahan atau peristiwa yang kemungkinan bisa saja terjadi di saat mendatang, hal ini mempengaruhi suatu kebijakan pada kehidupan bermasyarakat. Menurut penelitian Mohammad Kafli tahun 2019 tentang Prediksi penjualan menggunakan KNN, disebutkan terdapat tiga bentuk dari ramalan atau prediksi (Kafil, 2019), yaitu:

### 1. Proyeksi

Proyeksi adalah peramalan yang dilakukan berdasarkan ekstrapolasi atas kecenderungan yang terjadi pada masa lalu. Proyeksi memberikan pernyataan berdasarkan argumen dengan tegas berdasar dari metode tertentu dan kasus tertentu.

### 2. Prediksi

Prediksi adalah suatu peramalan yang dilakukan dengan berdasarkan asumsi teoritik. Asumsi yang dimaksud adalah berbentuk hukum teoritis, proposisi teoritis, maupun analogi.

### 3. Perkiraan

Perkiraan adalah suatu proses peramalan yang dilakukan berdasarkan sebuah penelitian informative ataupun berdasarkan dari penilaian yang dilakukan oleh ahli atau pakar yang terkait dengan suatu kondisi masyarakat yang bisa saja terjadi pada saat mendatang.

## **2.3 Curah Hujan**

Hujan merupakan cairan (kadang beku) yang jatuh ke permukaan bumi yang mana cairan tersebut berasal dari uap air yang mengalami perubahan wujud menjadi cair pada atmosfer bumi dengan suhu dibawah titik beku (Prawaka et al., 2016). Terdapat beragam bentuk butiran air yang turun, jika ukuran butiran air melebihi 0.5 milimeter maka akan disebut sebagai hujan, jika butiran air berukuran 0.2 milimeter hingga 0.5 milimeter maka akan disebut hujan gerimis, dan air berukuran tidak melebihi 0.2 milimeter akan menguap sebelum sampai ke permukaan bumi, hal ini disebabkan cairan tersebut akan mengalami penguapan sebelum menyentuh permukaan bumi (Handayani & Adri, 2015).

Curah hujan memiliki pengertian yaitu air yang turun ke atas permukaan bumi dan tidak menguap, meresap, maupun mengalir pada permukaan datar. Tiap 1 mm curah hujan merupakan air hujan setinggi 1 mm yang turun ke permukaan bumi yang datar dengan luas 1 m<sup>2</sup> tanpa mengalami penguapan, peresapan, dan pengaliran (Syaifullah, 2014) . Jika suatu daerah yang memiliki curah hujan tidak melebihi 60 mm per bulan, maka akan digolongkan sebagai daerah yang mengalami kekeringan. Namun jika curah hujan melebihi 100 mm per bulan, maka akan digolongkan sebagai daerah yang mengalami kebasahan. Dan curah hujan dengan ukuran 60 hingga 100 mm per bulan, maka akan digolongkan sebagai daerah yang mengalami kelembaban (Warsito & Sumiyati, 2007). Menurut Marni dan Muhammad Ishak Jumarang, kelembaban udara, suhu, dan kecepatan angin adalah beberapa faktor yang dapat memengaruhi curah hujan. (Jumarang & others, 2016).

#### 2.4 Z Score

*Z Score* merupakan salah satu teknik untuk standarisasi atau normalisasi suatu data. *Z Score* melakukan normalisasi data dengan cara membagi selisih antara data dan rata-rata data dengan standar deviasi data tersebut. Berikut merupakan persamaan dari *Z Score* (Radhitya & Sudipa, 2020):

$$Z = \frac{x - \mu}{\sigma} \quad (2.1)$$

Keterangan:

$Z$  = Nilai hasil perhitungan *Z Score*.

$x$  = Nilai awal yang akan dilakukan normalisasi

$\mu$  = Rata-rata dari seluruh data awan

$\sigma$  = Standar deviasi

## 2.5 *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) adalah salah satu jenis algoritma pembelajaran berbasis contoh (*instance-based learning*) atau non-parametrik. Dalam KNN, prediksi dari suatu titik data baru ditentukan dengan mengambil rata-rata dari  $k$  tetangga terdekat dari titik data baru dalam data set pelatihan.  $K$  adalah parameter yang ditentukan oleh pengguna dan dapat diubah sesuai kebutuhan. Pemilihan tetangga terdekat dilakukan dengan mengukur jarak pada titik data *testing* dengan titik data acuan dengan menggunakan perhitungan metrik seperti metode *Euclidean Distance* dan metode *Manhattan Distance*. KNN sering digunakan dalam klasifikasi, di mana tujuan adalah untuk menetapkan titik data terhadap beberapa kelas data yang sudah ditentukan berdasarkan nilai fitur nya. Namun, juga dapat digunakan untuk regresi, di mana tujuan adalah untuk memprediksi nilai numerik dari titik data baru berdasarkan nilai-nilai dari tetangga terdekat (Cover & Hart, 1967).

KNN biasa diperuntukkan dalam melakukan proses klasifikasi suatu objek dengan menghitung jarak terdekat antara data pembelajaran dengan objek tersebut. KNN memiliki tingkat akurasi yang cukup tinggi serta dengan efisiensi yang tinggi pula dalam beberapa kasus pengklasifikasian (Utomo et al., 2019). Namun, selain untuk klasifikasi, KNN juga digunakan untuk prediksi pada beberapa dekade terakhir (Hasmawati & Muchtar, 2017).

Terdapat beberapa tahapan dalam KNN (Fatah & Subekti, 2018), yaitu:

1. Menentukan parameter K, K yang dimaksudkan adalah banyaknya jumlah tetangga yang terdekat.
2. Menghitung selisih atau jarak antara data *testing* dan seluruh data acuan.

Untuk menentukan jarak antara data baru dan data lama memiliki banyak cara, namun yang paling populer digunakan yaitu *Euclidian distance* (Leidiana, 2013). *Euclidean distance* adalah metrik jarak yang digunakan dalam berbagai bidang, seperti pembelajaran mesin, analisis data, dan visi komputer. Pada dasarnya, *Euclidean Distance* menghitung jarak dengan cara mencari akar kuadrat dari selisih koordinat pada dua titik. Didefinisikan sebagai akar dari kuadrat dari perbedaan koordinat antara dua titik, jarak ini sering digunakan dalam berbagai bidang, termasuk analisis data dan pembelajaran mesin. Berikut persamaan dari *Euclidean Distance*:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (2.2)$$

Keterangan:

$d_{ij}$  = Jarak.

$n$  = Nomor vektor .

$x_{ik}$  = Titik tujuan.

$x_{jk}$  = Titik asal.

Selain *Euclidean Distance*, juga terdapat metode perhitungan jarak *Manhattan Distance* (Suwanda et al., 2020). *Manhattan distance* atau *L1 norm distance* adalah salah satu cara untuk pengukuran jarak dua titik dalam bidang kartesian.

Perhitungan jarak dilakukan dengan menambahkan selisih jarak antara setiap koordinat titik. Dalam pembelajaran mesin, metode ini dapat digunakan dalam metrik pembelajaran, di mana algoritma mencari fitur yang membuat jarak *Manhattan* antara titik-titik dari kelas yang berbeda sebesar mungkin (Deza & Deza, 2009). Berikut merupakan persamaan dari *Manhattan Distance*:

$$d_{man}(x, y) = \sum_{k=1}^d |x_j - y_j| \quad (2.3)$$

Keterangan:

$d_{man}$  = Hasil perhitungan *Manhattan*

$x_j$  = Titik tujuan

$y_j$  = Titik asal

3. Mengurutkan jarak antara objek dengan masing-masing sampel data pelatihan dan mengambil data tetangga dengan jarak terdekat dari data uji.
4. Mengumpulkan atau menentukan kategori tetangga terdekat.
5. Menghitung nilai prediksi

Terdapat dua metode perhitungan nilai prediksi yang diterapkan pada penelitian yang akan dilakukan ini, yaitu metode *Average* dan metode *Weighted*. Metode *Average* merupakan penentuan nilai prediksi KNN dengan merata-rata kan nilai target dari objek data terdekat sebanyak nilai  $k$  :

$$Y' = \frac{\sum_n^k Y_n}{k} \quad (2.4)$$

Keterangan :

$Y'$  = Nilai prediksi

$Y_n$  = Nilai target pada atribut yang masuk dalam jangkauan  $K$ .

$k$  = Banyak objek yang dijadikan acuan.

Metode *Weighted* merupakan penentuan nilai prediksi KNN dengan pemberian bobot pada nilai target dari objek data terdekat akan memiliki bobot paling besar sebanyak nilai  $k$ . Pemberian bobot menggunakan metode ROC (Gou et al., 2012). *Ranked Order Centroid* adalah teknik pembobotan yang digunakan dalam analisis multi-kriteria. Teknik ini menggunakan peringkat yang diberikan oleh responden untuk setiap kriteria, dan kemudian menghitung titik sentral dari peringkat tersebut untuk mendapatkan nilai bobot relatif dari setiap kriteria. Bobot dari setiap kriteria dihitung menggunakan rumus yang sesuai dengan metode yang digunakan untuk menghitung titik sentral. Setelah bobot relatif ditentukan, bobot tersebut digunakan untuk mengevaluasi alternatif yang berbeda. Setiap alternatif dinilai dengan menggunakan skor yang dihitung dari bobot relatif kriteria dan nilai alternatif untuk setiap kriteria. *Ranked Order Centroid* dianggap sebagai metode yang baik karena tidak memerlukan asumsi tentang distribusi data, mudah digunakan, dan dapat digunakan dengan berbagai jenis kriteria dan alternatif. Namun, metode ini memiliki beberapa keterbatasan seperti tidak mengakomodasi perbedaan dalam tingkat pentingnya kriteria dan tidak memperhitungkan interaksi antar kriteria. Berikut persamaan ROC :

$$Y' = Y_n \left( \frac{1}{k} \sum_{i=1}^k \left( \frac{1}{i} \right) \right) \quad (2.5)$$

Keterangan:

$Y'$  = Nilai prediksi

$Y_n$  = Nilai target pada atribut yang masuk dalam jangkauan K.

$k$  = Banyak objek yang dijadikan acuan.

Contoh pemberian bobot pada  $k = 3$ :

Tabel 2. 2 Contoh Pemberian Bobot

Index	Bobot
1	$\frac{\left( \frac{1}{1} + \frac{1}{2} + \frac{1}{3} \right)}{3}$
2	$\frac{\left( \frac{1}{2} + \frac{1}{3} \right)}{3}$
3	$\frac{\left( \frac{1}{3} \right)}{3}$

## 2.6 Root Mean Squared Error (RMSE)

RMSE atau singkatan dari *Root Mean Squared Error* adalah teknik untuk menjumlahkan kesalahan kuadrat nilai aktual dan nilai prediksi. (Larena, 2015). Hasil prediksi akan dikatakan baik ketika nilai dari RMSE yang dihasilkan semakin kecil (Lasulika, 2017).

RMSE adalah ukuran kesalahan yang diimplementasikan dalam statistik regresi. RMSE dihitung sebagai akar dari kuadrat *mean* selisih kuadrat nilai yang diharapkan dengan yang sebenarnya. Hal ini mengukur seberapa jauh nilai prediksi dari model dari nilai aktual. Semakin kecil RMSE, semakin baik model tersebut dalam memprediksi nilai aktual. RMSE digunakan dalam banyak aplikasi, termasuk

evaluasi model regresi linear dan analisis performa sistem pembelajaran mesin.

*Root Mean Squared Error* (RMSE) dapat digunakan untuk mengevaluasi performa model *K-Nearest Neighbors* (KNN). Dalam KNN, model memprediksi nilai untuk suatu titik data baru dengan mencari k titik data terdekat dalam dataset yang diberi (disebut sebagai tetangga) dan menghitung nilai rata-rata dari tetangga-tetangga tersebut. RMSE digunakan untuk mengukur seberapa jauh nilai prediksi dari model dari nilai aktual.

Sebuah model akan dinilai baik dalam memprediksi jika menghasilkan nilai RMSE yang kecil. Namun perlu diingat bahwa RMSE hanyalah salah satu cara untuk mengevaluasi performa model, dan tidak selalu menjadi ukuran yang baik untuk semua jenis masalah Berikut persamaan RMSE :

$$RMSE = \sqrt{\sum \frac{|Y' - Y|^2}{N}} \quad (2.6)$$

Keterangan:

$Y'$  = Nilai prediksi

$Y$  = Nilai sebenarnya

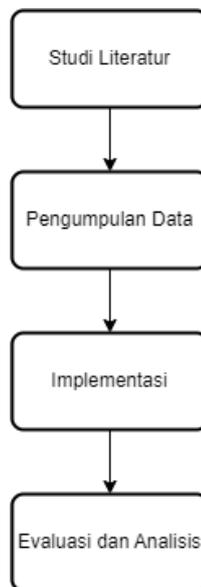
$N$  = Banyak data

## BAB III

### DESAIN DAN IMPLEMENTASI

#### 3.1 Metode Penelitian

Penjelasan tahap-tahap pada penelitian memerlukan sebuah desain tahapan pada penelitian yang berguna untuk penjelasan secara urai.



Gambar 3. 1 Tahapan Penelitian

Beberapa langkah yang akan dilaksanakan di penelitian, dimulai dengan studi literatur, pengumpulan data, implementasi, serta diakhiri dengan evaluasi dan analisis data.

##### 3.1.1 Studi Literatur

Pengumpulan serta mempelajari penelitian-penelitian terdahulu untuk dijadikan rujukan pada penelitian yang dilakukan pada tahapan ini. Terdapat lima penelitian yang dijadikan rujukan seperti yang telah disajikan pada BAB II bagian sub-bab penelitian terdahulu.

### 3.1.2 Pengumpulan Data

Data iklim harian di Pekanbaru pada tahun 2019 – 2021 akan digunakan pada penelitian ini. Data berupa kelembaban, temperatur, kecepatan angin dan curah hujan kota Pekanbaru. Terdapat 892 data yang akan digunakan. Data didapat dari situs resmi BMKG Indonesia.

### 3.1.3 Implementasi

Setelah data dikumpulkan, maka pada tahap implementasi ini data akan dilakukan suatu proses pengolahan hingga mengeluarkan hasil prediksi. Pada tahap implementasi, terdapat dua tahapan yang akan dilakukan yaitu tahap *preprocessing* dan tahap pengolahan data menggunakan algoritma KNN. Gambar 3.2 akan menjelaskan program, serta Gambar 3.3 akan menjelaskan alur implementasi.

```
knn = function(near, x_fit, x_pred, y, k, weighted_pred = F, p = NULL){
  # variabel yang akan menampung nilai prediksi
  predictions = c()

  # Proses pengolahan data prediksi menggunakan KNN
  for(i in 1:nrow(x_pred)){

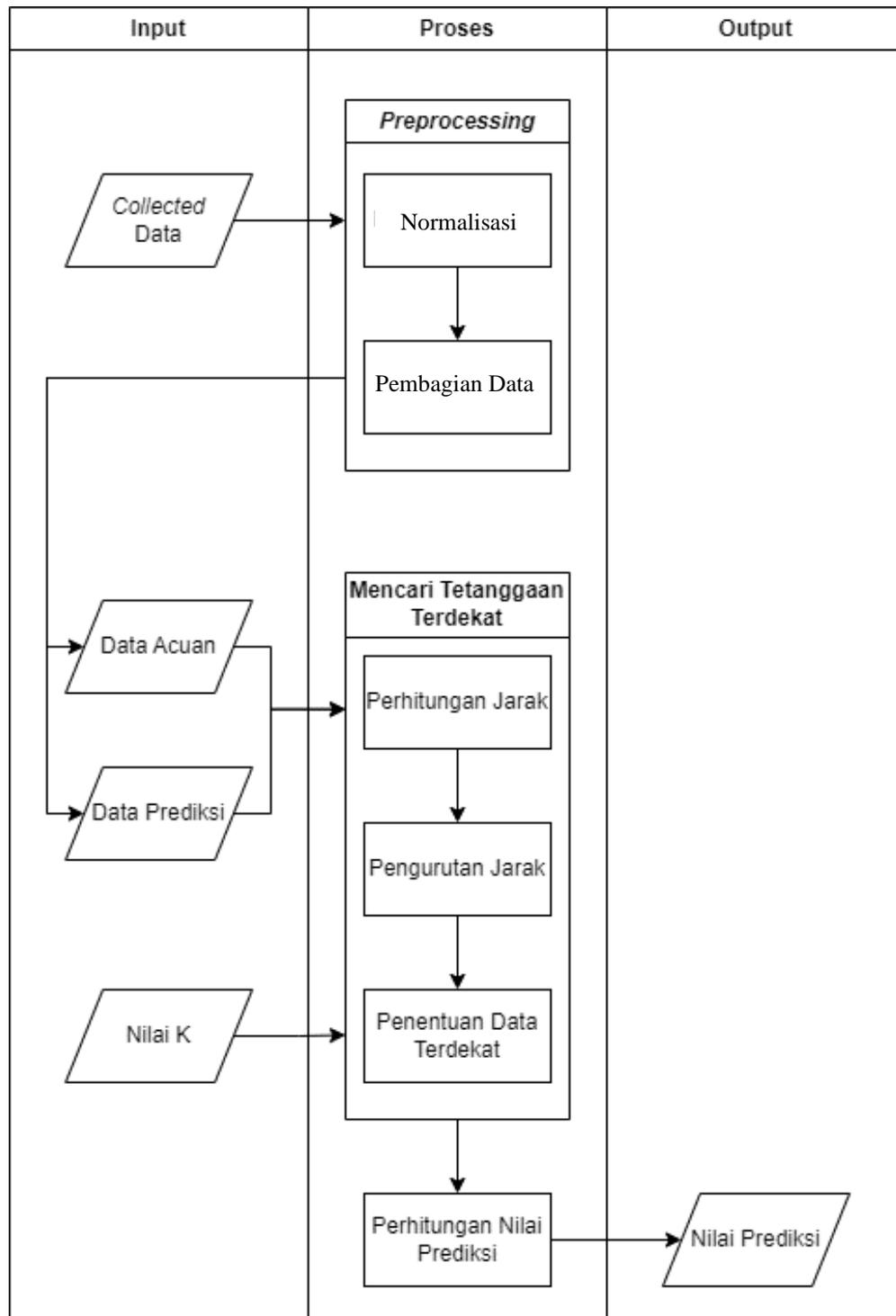
    ind = as.numeric(unlist(strsplit(near$nearest_ind[[i]], ",")))
    nn = list(ind)
    neighbors = nn[[1]]

    if(weighted_pred == T){
      pred = knn_prediction(x_fit[neighbors[1:k], ],y, w = T, z = k)
    } else{
      pred = knn_prediction(x_fit[neighbors[1:k], ],y, w = F)
    }

    # If more than 1 predictions, make prediction with 1 more k
    if(length(pred)>1){
      pred = knn(x_fit, x_pred[i,],y, k = k+1,
                func = func, weighted_pred = weighted_pred, p == p)
    }

    predictions[i] = pred
  }
  return(predictions)
}
```

Gambar 3. 2 Code proses pengolahan data pada bahasa R



Gambar 3. 3 Desain Implementasi

*Preprocessing* adalah tahapan yang bertujuan untuk mempersiapkan data agar dapat melakukan proses pembelajaran mesin dengan baik pada penelitian. Pembagian data menjadi data *testing* dan data acuan juga dilakukan pada tahapan *preprocessing*. Data didapat akan dibagi untuk penelitian, yaitu data acuan, dan data *testing* (Fadilah et al., 2020). Setelah dibagi, data acuan maupun data *testing* akan dilakukan normalisasi dengan tujuan menghilangkan ketimpangan pada data. Normalisasi data akan dilakukan menggunakan *Z Score* (Persamaan 2.1).

```
#Menghitung Normalisasi menggunakan Zscore
zS = function(data) {
  #data akan dikurangi dengan rata-rata data
  #tersebut lalu dibagi dengan stadr deviasi data
  z <- (data-mean(data))/sd(data)
  return(z)
}
```

Gambar 3. 4 *Code Z Score* pada bahasa R

Pada gambar 3.4 merupakan *code Z Score* pada bahasa pemrograman R. Dapat dilihat bahwa fungsi untuk menghitung nilai *Z Score* diinisialisasikan dengan *zS* dengan parameter *x* yang akan diisi dengan data yang akan dinormalisasi, yang mana fungsi ini nantinya akan menjadi fungsi yang menjalankan proses sesuai rumus normalisasi *Z Score* itu sendiri.

Data diolah menggunakan algoritma KNN dengan beberapa tahapan yaitu, penentuan nilai *K*, perhitungan jarak, pengurutan jarak data, penentuan data terdekat, dan perhitungan nilai prediksi.

Menentukan parameter *K*, *K* yang dimaksudkan adalah banyaknya jumlah tetangga yang terdekat. Jika yang ingin digunakan untuk prediksi hanya 3 data terdekat, maka nilai *K* adalah 3. Data yang telah dilakukan tahapan *preprocessing*

di input.

Menghitung selisih atau jarak antara data *testing* dan seluruh data acuan menggunakan *Euclidian Distance* (persamaan 2.2) maupun *Manhattan Distance* (persamaan 2.3). Selisih jarak masing-masing data acuan dengan data *testing* akan dihitung untuk menentukan kedekatan tiap objek acuan dengan objek *testing*.

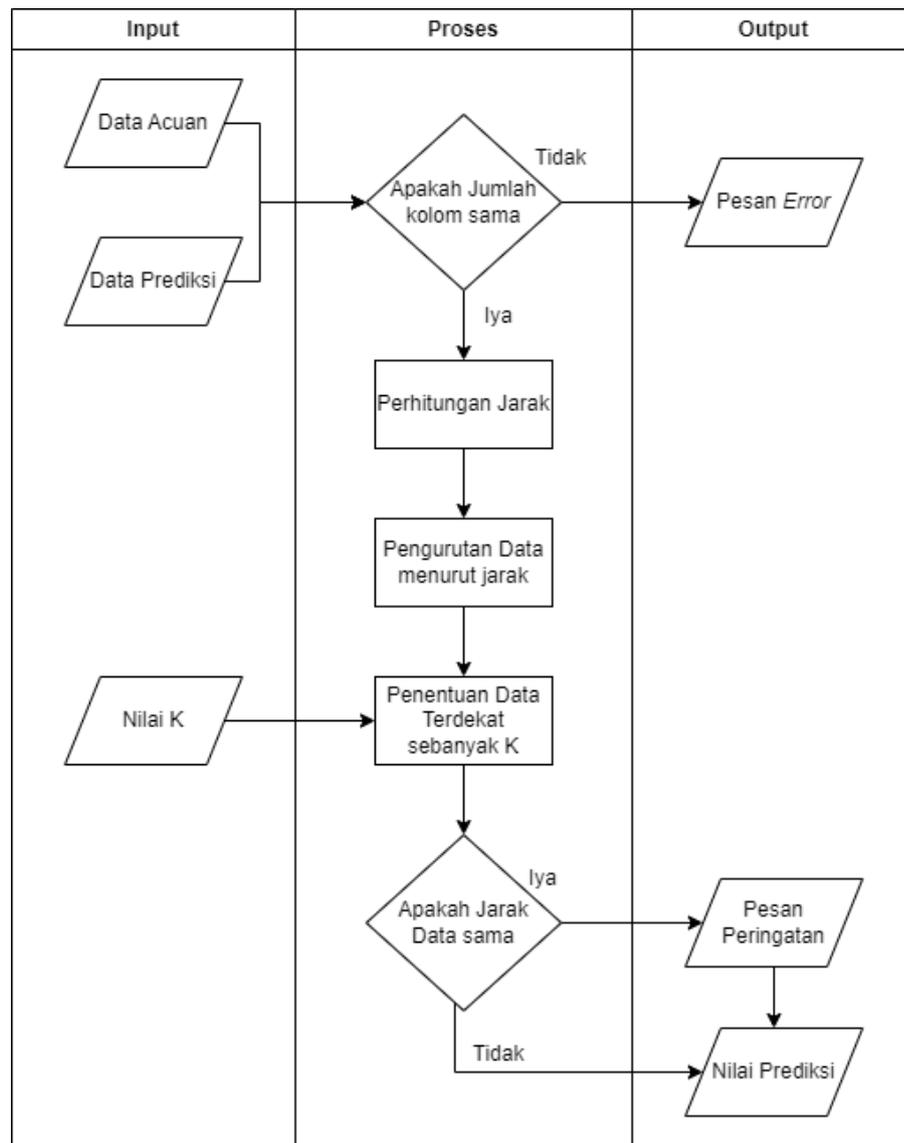
```
euclidean_distance = function(a, b){
  # Melihat apakah data a dan b memiliki panjang yang sama
  if(length(a) == length(b)){
    sqrt(sum((a-b)^2))
  } else{
    stop('vectors harus memiliki panjang yang sama')
  }
}

manhattan_distance = function(a, b){
  # Melihat apakah data a dan b memiliki panjang yang sama
  if(length(a) == length(b)){
    sum(abs(a-b))
  } else{
    stop('vectors harus memiliki panjang yang sama')
  }
}
```

Gambar 3.5 Code *Euclidean* dan *Manhattan Distance* pada bahasa R

Pada gambar 3.5 merupakan *code Euclidean Distance dan Manhattan Distance* pada bahasa pemrograman R. Pada gambar 3.5 dapat dilihat bahwa fungsi untuk menghitung selisih jarak antar data diinisialisasikan dengan nama dari masing-masing metode itu sendiri dengan parameter a dan b yang masing-masing akan diisi dengan data asal dan data tujuan, yang mana kedua fungsi ini awalnya akan memeriksa apakah jumlah kolom antara data asal dan tujuan sama atau tidak, lalu jika sama maka masing-masing metode akan menjalankan perhitungan jarak sesuai dari rumus masing-masing metode tersebut.

Mengurutkan jarak antara objek dengan masing-masing sampel data acuan. Setelah jarak selisih jarak antara data acuan dan data *testing* dihitung, maka data diurutkan dari jarak data yang memiliki jarak yang terkecil hingga jarak terbesar. Setelah mengetahui objek yang memiliki selisih jarak yang terkecil, maka diambil data target pada objek terdekat sebanyak nilai K.



Gambar 3. 6 Alur untuk mencari tetangga terdekat

```

nearest_neighbors = function(x,obs, k, FUN){
  # Melihat apakah variabel antara prediksi dan acuan sama
  if(ncol(x) != ncol(obs)){
    stop('Data harus memiliki jumlah variabel yang sama')
  }
  dist = apply(x,1, FUN,obs)

  # Mencari Tetangga Terdekat
  distances = sort(dist)[1:k]
  neighbor_ind <- c()
  for (i in 1:k) {
    aa <- which(dist %in% sort(dist)[i])
    neighbor_ind <- c(neighbor_ind, aa)
  }

  if(length(neighbor_ind) != k){
    warning(
      paste('Terdapat kemiripan jarak.| Used k:',length(neighbor_ind))
    )
  }
  neighbor_ind = neighbor_ind[!duplicated(neighbor_ind)]
  ret = list(neighbor_ind, distances)
  return(ret)
}

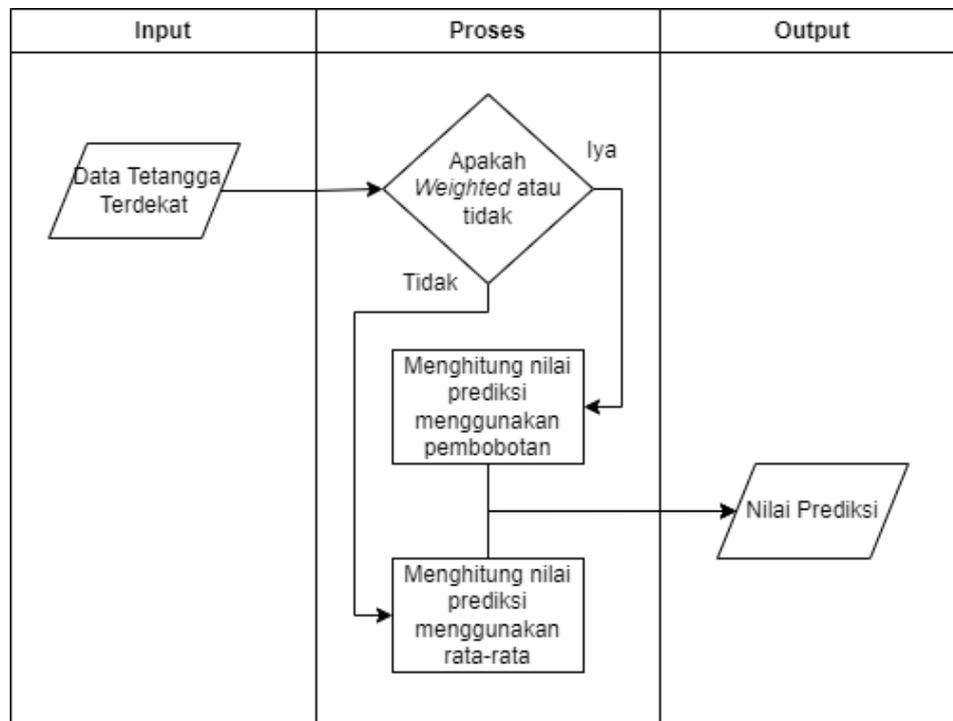
```

Gambar 3.7 Code untuk mencari tetangga terdekat pada bahasa R

Gambar 3.6 maupun Gambar 3.7 menjelaskan fungsi untuk mencari data tetangga terdekat diinisialisasikan dengan nama dari *nearest\_neighbors*, dengan parameter *x* sebagai data acuan, *obs* sebagai data yang ingin dicari tetanggaannya, *k* adalah banyak data tetangga terdekat yang akan dicari, dan *fun* adalah metode yang digunakan untuk mencari data tetangga terdekat. Fungsi ini awalnya akan memeriksa apakah jumlah kolom antara data asal dan tujuan sama atau tidak, lalu jika sama maka fungsi ini akan menjalankan perhitungan jarak dan ditampung dengan variabel *dist*. Setelah dihitung, maka jarak akan diurutkan dan diambil sebanyak *k* dan ditampung dengan variabel *distances*. *Neighbor\_ind* merupakan variabel yang menampung index dari data terdekat. Lalu index data dan jaraknya akan dikembalikan pada fungsi ini.

Terdapat dua cara untuk menghitung nilai prediksi KNN pada penelitian ini, yaitu dengan metode *average* atau *weighted*. Menghitung nilai prediksi menggunakan metode *average* adalah dengan cara menghitung *mean* dari nilai seluruh data terdekat menurut jangkauan K (Menggunakan persamaan 2.4). Setelah mengambil data target, maka rata-rata data target dari objek terdekat sebanyak nilai K dihitung untuk dijadikan sebagai nilai prediksi.

Menghitung nilai prediksi KNN menggunakan metode *weighted* adalah menghitung nilai prediksi KNN dengan pemberian bobot pada nilai target mulai dari objek data terdekat akan memiliki bobot paling besar sebanyak nilai k. Pemberian bobot pada penelitian ini menggunakan metode ROC (Menggunakan persamaan 2.5).



Gambar 3. 8 Alur untuk menghitung nilai prediksi KNN pada bahasa R

```

knn_prediction = function(x,y,z,w = F){
  x = as.matrix(x)
  if(is.factor(x[,y]) | is.character(x[,y])){
    groups = table(x[,y])
    pred = names(groups[groups == max(groups)])
  }
  if(is.numeric(x[,y])){
    if (w == T) {
      bobot <- 0
      pred <- 0
      st <- 0
      for (i in 1:length(x[,y])) {
        bobot <- 0
        st <- st+1
        for (j in st:z) {
          bobot <- bobot + (1/j)
        }
        bobot <- bobot/z
        pred <- pred + (x[i,y]^bobot)
      }
    } else if (w == F) {
      pred = mean(x[,y])
    }
  }
  # If no pred, then class is not correct
  if(try(class(x[,y])) == 'try-error'){
    stop('Y should be factor or numeric.')
  }
  return(pred)
}

```

Gambar 3.9 Code untuk menghitung nilai prediksi KNN pada bahasa R

Fungsi untuk menghitung nilai prediksi KNN dijelaskan pada gambar 3.8 maupun Gambar 3.9 yaitu dengan menginisialisasikan dengan nama `knn_prediction` dengan parameter `x` sebagai data acuan, `y` sebagai data yang akan diprediksi, `z` sebagai banyak data, dan `w` untuk menentukan apakah prediksi yang digunakan *Weighted* atau tidak (*Average*). Fungsi ini akan melihat apakah prediksi yang diminta menggunakan *Weighted* atau tidak, jika `w` adalah `True` maka perhitungan yang dijalankan menggunakan metode *Weighted* dengan perhitungan pembobotan menggunakan metode *Rank Order Centroid*. Jika `w` adalah `false`, maka metode perhitungan yang digunakan adalah *Average*.

### 3.1.4 Evaluasi Kinerja KNN

Hasil prediksi dan data aktual akan dibandingkan untuk mengetahui akurasi dari prediksi menggunakan RMSE (persamaan 2.6). Hasil *error* yang dihasilkan oleh perhitungan menggunakan RMSE akan menjadi tolak ukur seberapa akurat memprediksi curah hujan pada kota Pekanbaru menggunakan metode KNN.

```
#Menghitung evaluasi error dengan Root Mean Squared Error
rmse = function(y_pred, y_real){
  #Mencari akar kuadrat dari rata-rata
  #selisih data dipangkat dua
  sqrt(mean((y_pred-y_real)^2))
}
```

Gambar 3. 10 Code untuk evaluasi KNN dengan RMSE pada bahasa R

Pada gambar 3.10 merupakan *code* untuk evaluasi KNN menggunakan RMSE pada bahasa pemrograman R. Pada gambar 3.9 dapat dilihat bahwa fungsi untuk menghitung nilai RMSE diinisialisasikan dengan *rmse* dengan parameter *y\_pred* sebagai nilai prediksi dan *y\_real* sebagai nilai aktual. Fungsi ini nantinya akan menjadi fungsi yang menjalankan proses sesuai rumus normalisasi RMSE itu sendiri.

## 3.2 Desain Interface

Penelitian ini juga merancang antarmuka yang dapat memudahkan pengguna dalam melakukan percobaan prediksi curah hujan dengan menggunakan metode K-Nearest Neighbor (KNN).



Gambar 3. 11 Tampilan awal dari *interface*

Pada gambar 3.11 adalah tampilan awal *interface* berupa *page* Home yang berisi penjelasan tentang aplikasi ini. Dapat dilihat juga terdapat *navigator* yang ketika di klik maka akan berpindah tampilan ke masing-masing *page* sesuai judul yang telah diberikan pada *navigator* tersebut.



Gambar 3. 12 Tampilan input data acuan

Tampilan dari menu *Import Data* yang berisi halaman untuk *input* data acuan dapat dilihat pada gambar 3.12. Setelah data di *import*, maka halaman akan menampilkan data yang telah ditambahkan berupa tabel pada bagian sebelah kanan

halaman. Data dapat di *import* dengan klik tombol “*Browse*”, setelah itu pilihlah file data yang akan di *import*. Data akan tampil di sebelah kan setelah proses *import* data berhasil dilakukan tanpa ada *error*.



Gambar 3. 13 Tampilan input data prediksi

Tampilan dari menu *Import Data Prediksi* yang berisi halaman untuk *input* data yang akan diprediksi dapat dilihat pada gambar 3.13. Setelah data di *import*, maka halaman akan menampilkan data yang telah ditambahkan berupa tabel pada bagian sebelah kanan halaman. Data dapat di *import* dengan klik tombol “*Browse*”, setelah itu pilihlah file data yang akan di *import*. Data akan tampil di sebelah kan setelah proses *import* data berhasil dilakukan tanpa ada *error*.



Gambar 3. 14 Tampilan halaman untuk menentukan parameter prediksi

Tampilan menu *Select Variable* pada gambar 3.14 berisi halaman untuk menentukan parameter yang akan digunakan untuk memprediksi, seperti menentukan nilai K dan metode perhitungan jarak yang akan digunakan. Parameter yang ingin digunakan cukup dengan cara klik pilihan parameter yang ingin digunakan, seperti nilai K dan metode perhitungan jarak yang ingin digunakan.

	real	Average	Weighted
1	10.4	3.633333333333333	5.982333333333333
2	1.2	0.666666666666667	0.553333333333333
3	11.7	0	0

Gambar 3. 15 Tampilan halaman untuk melihat hasil prediksi

Tampilan dari menu *Result* yang berisi halaman untuk menampilkan hasil prediksi berupa tabel dapat dilihat pada gambar 3.15. Tabel akan menampilkan hasil prediksi menggunakan *average* maupun *weighted* dan juga dengan perhitungan jarak dengan metode *Euclidean* maupun *Manhattan*.

## BAB IV

### UJI COBA DAN PEMBAHASAN

#### 4.1 Skenario Uji Coba

Skenario percobaan bertujuan untuk melihat nilai prediksi dan seberapa besar *error* yang dihasilkan pada kondisi yang bervariasi. Dari seluruh data yang ada, akan dilakukan pembagian data, yaitu data sebagai acuan dan data sebagai *testing* dengan persentase divariasikan. Tabel 4.1 merupakan iterasi percobaan.

Tabel 4. 1 Skenario Pembagian Data

Skenario	Persentase Data Acuan	Persentase Data <i>Testing</i>	Keterangan
1	80%	20%	80% data acuan sedangkan 20% data <i>testing</i> .
2	70%	30%	70% data acuan sedangkan 30% data <i>testing</i> .
3	60%	40%	60% data acuan sedangkan 40% data <i>testing</i> .
4	50%	50%	50% data acuan sedangkan 50% data <i>testing</i> .

Selain pembagian data, terdapat juga sub-iterasi pada skenario percobaan ini yaitu variasi metode pengukuran jarak. Metode pengukuran jarak digunakan untuk mencari objek terdekat pada metode KNN. Tabel 4.2 menjelaskan metode pengukuran jarak yang akan digunakan .

Tabel 4. 2 Sub-percobaan metode perhitungan jarak

Sub-skenario	Metode	Keterangan
1	<i>Euclidean Distance</i>	<i>Euclidean Distance</i> merupakan suatu cara menghitung jarak antar titik dihitung dari garis lurus dua titik.
2	<i>Manhattan Distance</i>	<i>Manhattan Distance</i> merupakan suatu cara menghitung jarak antar titik yang akan dihitung dari langkah minimal yang dapat ditempuh antara satu titik dengan titik lainnya.

Selain variasi metode perhitungan jarak, pada penelitian ini juga terdapat sub-iterasi dengan melakukan variasi nilai K yang ditentukan untuk mengetahui nilai K yang menghasilkan prediksi terbaik.

Selain itu, pada penelitian ini juga terdapat percobaan dengan melakukan variasi metode penentuan nilai prediksi yang ditentukan untuk mengetahui metode penentuan prediksi yang menghasilkan prediksi terbaik pada KNN. Tabel 4.3 menjelaskan variasi metode yang akan digunakan.

Tabel 4. 3 Sub-percobaan variasi perhitungan prediksi

Sub-skenario	Metode	Keterangan
1	<i>Average</i>	Metode <i>Average</i> merupakan penentuan nilai prediksi KNN dengan merata-rata kan nilai target dari objek data terdekat sebanyak nilai k
2	<i>Weighted menggunakan Ranked Order Centroid (ROC)</i>	Metode <i>Weighted</i> merupakan penentuan nilai prediksi KNN dengan pemberian bobot pada nilai target dari objek data terdekat yang memiliki bobot paling besar sebanyak nilai k.

		Pemberian bobot menggunakan metode ROC.
--	--	---

Setelah hasil prediksi didapatkan dari percobaan pada setiap iterasi dan sub iterasi dari skenario percobaan yang telah dijelaskan sebelumnya, maka akan dihitung seberapa *error* yang akan dihasilkan menggunakan metode RMSE dengan persamaan 2.3.

Tabel 4. 4 Spesifikasi Perangkat

<i>Brand</i>	Asus VivoBook 14/15
<i>Processor</i>	Intel Core i7-10510U
Bahasa Pemrograman	R
<i>Software</i>	R Studio
RAM	8 GB
OS	Microsoft Windows 11 Home Single Language

## 4.2 Hasil Uji Coba

Percobaan dilakukan menurut skenario percobaan yang dijelaskan pada sub-bab 4.1 sebelumnya. Hasil uji coba akan dikelompokkan dalam sebuah bentuk tabel. Pada tabel 4.5 dapat dilihat contoh perbedaan hasil prediksi pada variasi skenario dan sub-skenario yang dilakukan.

Tabel 4. 5 Contoh Perbedaan Hasil Prediksi pada masing-masing skenario

Tanggal	Real	Skenario	Metode	RMSE					
				<i>Euclidean Distance</i>			<i>Manhattan Distance</i>		
				K=3	K=5	K=7	K=3	K=5	K=7
19/11/2019	8,90	1	<i>Average</i>	0,47	1,96	2,57	0,47	1,96	2,96
			<i>Weighted</i>	0,00	0,00	0,00	0,00	0,00	0,00
		2	<i>Average</i>	4,77	3,68	4,34	4,00	3,40	3,26
			<i>Weighted</i>	4,77	2,78	2,26	4,30	2,78	2,26

		3	<i>Average</i>	0,77	0,76	3,70	0,20	2,56	3,70
			<i>Weighted</i>	0,00	0,00	0,00	0,00	0,00	0,00
		4	<i>Average</i>	3,90	19,20	17,51	3,90	19,20	17,51
			<i>Weighted</i>	6,28	4,70	3,81	6,28	4,70	3,81
11/02 /2020	1,40	1	<i>Average</i>	6,27	7,38	6,24	6,27	7,38	5,70
			<i>Weighted</i>	0,85	0,18	0,15	0,85	0,18	0,15
		2	<i>Average</i>	1,27	4,98	5,39	1,27	4,98	5,39
			<i>Weighted</i>	2,32	1,73	1,41	2,32	1,73	1,41
		3	<i>Average</i>	0,73	6,84	10,24	0,73	6,84	7,47
			<i>Weighted</i>	0,00	0,00	0,00	0,00	0,00	0,00
		4	<i>Average</i>	12, 53	8,64	8,76	12,53	8,64	8,76
			<i>Weighted</i>	10, 44	0,05	0,04	10,44	0,05	0,04
06/04 /2020	35, 80	1	<i>Average</i>	1,17	2,98	2,16	1,17	3,80	3,63
			<i>Weighted</i>	0,97	0,00	0,00	0,97	0,00	0,00
		2	<i>Average</i>	4,33	2,80	2,00	4,33	2,80	4,60
			<i>Weighted</i>	6,76	4,33	3,52	6,76	4,33	3,52
		3	<i>Average</i>	4,33	3,42	4,96	4,33	4,66	3,54
			<i>Weighted</i>	6,76	4,33	3,52	6,76	4,33	3,52
		4	<i>Average</i>	1,67	4,16	8,56	1,67	4,16	8,56
			<i>Weighted</i>	0,97	0,00	0,00	0,42	0,00	0,00

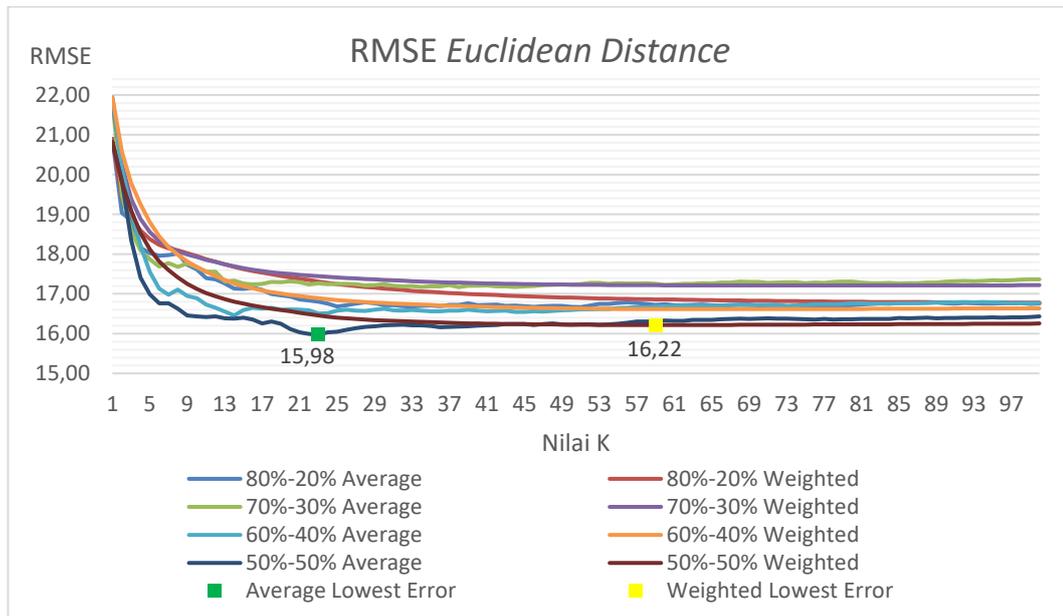
Pada Tabel 4.5 dapat dilihat perbedaan hasil prediksi dari tiga sampel data, yaitu data tanggal 30/08/2019, 19/11/2019, dan 29/08/2020. Pada data tanggal 19/11/2019, seluruh skenario dan sub-skenario menghasilkan nilai prediksi yang tidak sama dengan nilai *real*. Nilai prediksi yang paling dekat dengan nilai *real* terjadi pada kondisi skenario 4 dengan perhitungan prediksi menggunakan *Weighted*, perhitungan jarak *Euclidean* maupun *Manhattan* dan nilai  $k = 3$  menghasilkan nilai prediksi 6.28. Nilai prediksi yang paling jauh dengan nilai *real* terjadi pada kondisi skenario 1 dan 3 dengan perhitungan prediksi menggunakan *Weighted*, perhitungan jarak *Euclidean* maupun *Manhattan* dan keseluruhan variasi nilai  $K$ , menghasilkan nilai prediksi 0.00.

Pada data tanggal 11/02/2020, seluruh skenario dan sub-skenario menghasilkan nilai prediksi yang tidak sama dengan nilai *real*. Nilai prediksi yang paling dekat dengan nilai *real* terjadi pada dua kondisi yaitu skenario 2 dengan perhitungan prediksi menggunakan *Weighted*, perhitungan jarak *Euclidean* maupun *Manhattan* dan nilai  $k = 7$  menghasilkan nilai prediksi 1.41. Nilai prediksi yang paling jauh dengan nilai *real* terjadi pada kondisi skenario 4 dengan perhitungan prediksi menggunakan *Average*, perhitungan jarak *Euclidean* dan nilai  $k = 3$ , menghasilkan nilai prediksi 12.53.

Pada data tanggal 06/04/2020, seluruh skenario dan sub-skenario menghasilkan nilai prediksi yang tidak sama dengan nilai *real*. Nilai prediksi yang paling dekat dengan nilai *real* terjadi pada 2 kondisi yaitu skenario 4 dengan perhitungan prediksi menggunakan *Average*, perhitungan jarak *Euclidean* maupun *Manhattan* dan nilai  $k = 7$  menghasilkan nilai prediksi 8.56. Nilai prediksi yang paling jauh dengan nilai *real* terjadi pada kondisi skenario 4 dan skenario 1 dengan perhitungan prediksi menggunakan *Weighted*, perhitungan jarak *Euclidean* maupun *Manhattan* dan nilai  $k = 5$  dan 7, menghasilkan nilai prediksi 0.00.

### **4.3 Hasil Analisis**

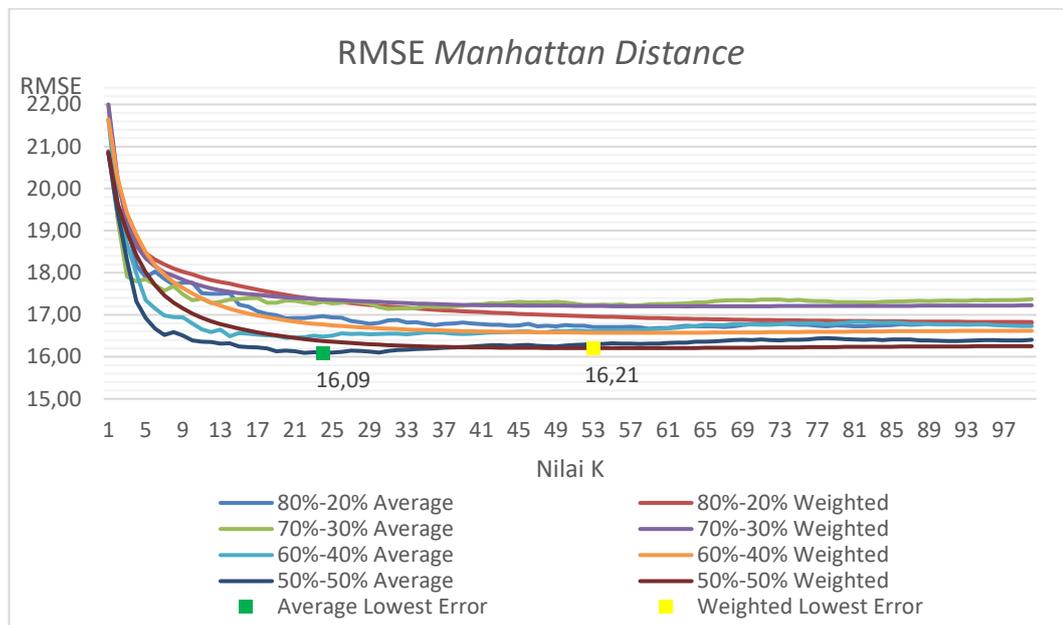
Percobaan pada penelitian ini dilaksanakan menurut skenario percobaan yang dijelaskan pada sub-bab 4.1 sebelumnya. Hasil uji coba akan dikelompokkan dan disajikan berbentuk tabel. Tinjauan hasil RMSE tiap iterasi disajikan Lampiran 1 dan Gambar 4.1 hingga Lampiran 2 dan Gambar 4.4.



Gambar 4.1 Line Plot RMSE Euclidean Distance

Hasil *error* pada percobaan dengan menggunakan metode perhitungan jarak menggunakan *Euclidean Distance* dapat dilihat pada Lampiran 1 dan Gambar 4.1. Pada metode perhitungan jarak menggunakan *Euclidean Distance*, algoritma KNN dengan menggunakan metode *Average* menghasilkan *error* yang lebih rendah dari *error* yang dihasilkan algoritma KNN yang menggunakan metode *Weighted*.

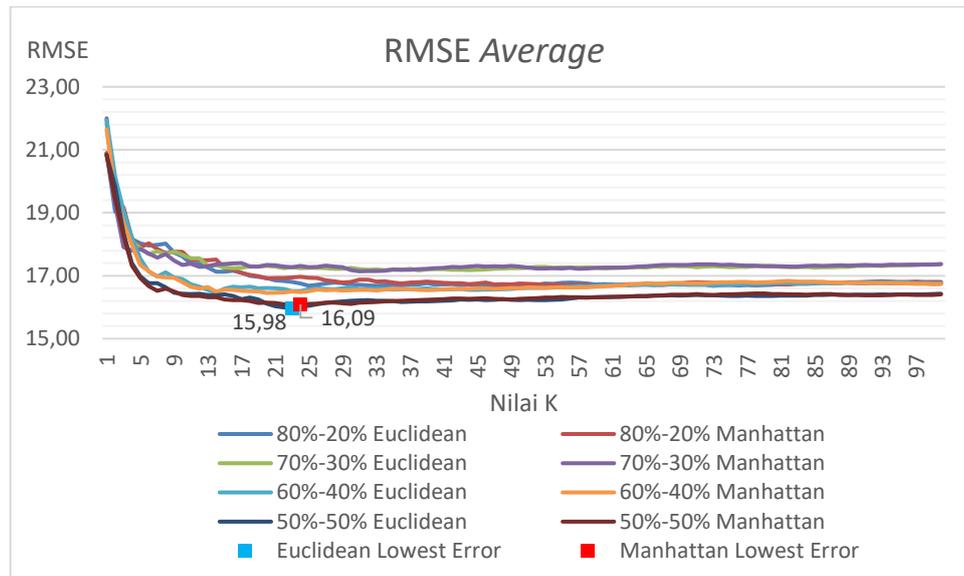
*Error* terkecil yang dihasilkan pada metode perhitungan jarak menggunakan *Euclidean Distance* dan metode perhitungan nilai prediksi menggunakan *Average* adalah 15,98 pada kondisi  $K = 23$ . Sedangkan *error* terkecil yang dihasilkan pada metode perhitungan jarak menggunakan *Euclidean Distance* dan metode perhitungan nilai prediksi menggunakan *Weighted* adalah 15,22 pada kondisi  $K = 59$ .



Gambar 4.2 Line Plot RMSE Manhattan Distance

Hasil *error* pada percobaan dengan menggunakan metode perhitungan jarak menggunakan *Manhattan Distance* dapat dilihat pada Lampiran 2 dan Gambar 4.2. Pada metode perhitungan jarak menggunakan *Manhattan Distance*, algoritma KNN dengan menggunakan metode *Average* menghasilkan *error* yang lebih rendah dari *error* yang dihasilkan algoritma KNN yang menggunakan metode *Weighted*.

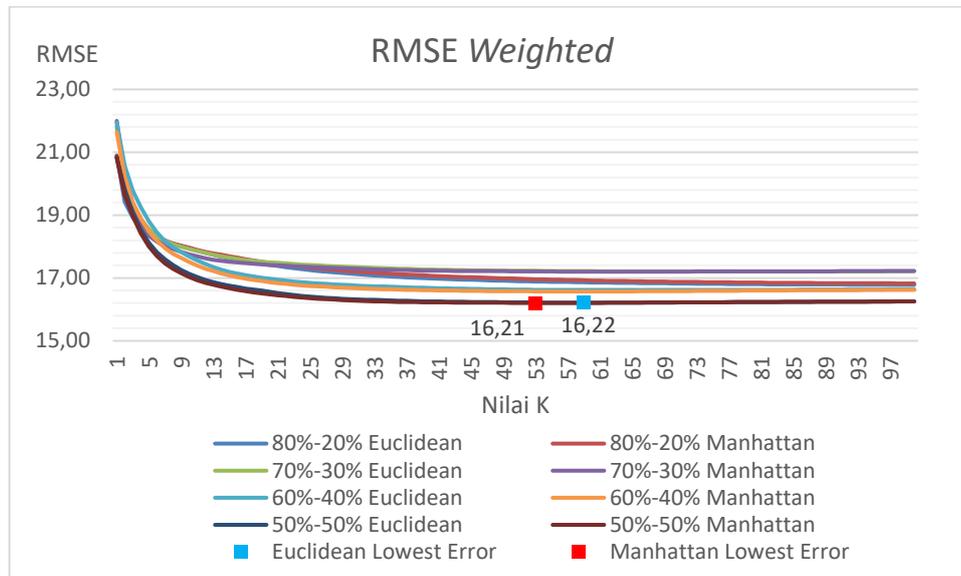
*Error* terkecil yang dihasilkan pada metode perhitungan jarak menggunakan *Manhattan Distance* dan metode perhitungan nilai prediksi menggunakan *Average* adalah 16,09 pada kondisi  $K = 24$ . Sedangkan *error* terkecil yang dihasilkan pada metode perhitungan jarak menggunakan *Manhattan Distance* dan metode perhitungan nilai prediksi menggunakan *Weighted* adalah 16,22 pada kondisi  $K = 53$ .



Gambar 4.3 Line Plot RMSE Metode Prediksi Average

Hasil *error* pada percobaan dengan menggunakan metode perhitungan prediksi menggunakan *Average* dapat dilihat pada Lampiran 3 dan Gambar 4.3. Pada metode perhitungan prediksi menggunakan *Average*, terdapat beberapa kondisi perhitungan jarak *Euclidean Distance* menghasilkan *error* yang lebih kecil dari *Manhattan Distance*, namun juga terdapat beberapa kondisi perhitungan jarak *Euclidean Distance* menghasilkan *error* yang lebih besar dari *Manhattan Distance*.

*Error* terkecil yang dihasilkan pada metode perhitungan prediksi menggunakan *Average* dan metode perhitungan jarak menggunakan *Euclidean Distance* adalah 15,98 pada kondisi  $K = 23$ . Sedangkan *error* terkecil yang dihasilkan pada metode perhitungan prediksi menggunakan *Average* dan metode perhitungan jarak menggunakan *Manhattan Distance* adalah 16,09 pada kondisi  $K = 24$ .



Gambar 4.4 Line Plot RMSE Metode Prediksi *Weighted*

Hasil *error* pada percobaan dengan menggunakan metode perhitungan prediksi menggunakan *Weighted* dapat dilihat pada Lampiran 4 dan Gambar 4.4. Pada metode perhitungan prediksi menggunakan *Weighted*, terdapat beberapa kondisi perhitungan jarak *Euclidean Distance* menghasilkan *error* yang lebih kecil dari *Manhattan Distance*, namun juga terdapat beberapa kondisi perhitungan jarak *Euclidean Distance* menghasilkan *error* yang lebih besar dari *Manhattan Distance*.

*Error* terkecil yang dihasilkan pada metode perhitungan prediksi menggunakan *Weighted* dan metode perhitungan jarak menggunakan *Euclidean Distance* adalah 16,22 pada kondisi  $K = 59$ . Sedangkan *error* terkecil yang dihasilkan pada metode perhitungan prediksi menggunakan *Weighted* dan metode perhitungan jarak menggunakan *Manhattan Distance* adalah 16,21 pada kondisi  $K = 53$ .

Secara keseluruhan skenario, algoritma KNN dengan menggunakan metode *Average* umumnya menghasilkan *error* yang lebih rendah dari *error* yang dihasilkan algoritma KNN yang menggunakan metode *Weighted*. Pada perbandingan hasil perhitungan prediksi dan menghitung jarak dengan menggunakan antara *Euclidean Distance* dan *Manhattan Distance*, terdapat kondisi dimana *error* yang dihasilkan menggunakan *Euclidean Distance* lebih kecil dari *error* yang dihasilkan menggunakan *Manhattan Distance*. Namun terdapat juga beberapa kondisi dimana *error* yang dihasilkan menggunakan *Euclidean Distance* lebih besar dari *error* yang dihasilkan menggunakan *Manhattan Distance*.

*Error* terkecil dihasilkan pada skenario 4 (pembagian data acuan dan data *testing* masing-masing 50% - 50%) dengan metode perhitungan jarak *Euclidean Distance* dan metode perhitungan prediksi menggunakan *Average* pada  $K = 23$  dengan *error* sebesar 15,98.

#### **4.4 Pembahasan**

Setelah dilakukan percobaan pada penerapan metode KNN untuk memprediksi curah hujan kota Pekanbaru, maka dapat dilihat metode perhitungan hasil prediksi, metode perhitungan jarak, variasi nilai  $K$ , dan pembagian data acuan dan data *testing* mempengaruhi hasil dari prediksi menggunakan KNN. Hasil uji coba akan dibahas secara detail pada paragraf-paragraf selanjutnya.

KNN merupakan metode *Lazy Learning*, yaitu metode yang melakukan pembelajaran sangat sedikit ataupun sama sekali tidak melakukan pembelajaran. Pada studi kasus curah hujan, data harus dilakukan proses normalisasi. Proses ini bertujuan untuk menyamakan rentang nilai terhadap variabel-variabel yang

mempengaruhi curah hujan.

Terjadi perbedaan *error* yang dihasilkan pada variasi nilai K yang berbeda. Hal ini tentu saja bisa terjadi dikarenakan nilai K menentukan berapa objek data terdekat yang akan menjadi acuan yang mana nilai targetnya akan dihitung untuk menjadi nilai prediksi pada objek data yang akan diprediksi. Secara umum, pada Lampiran 1 hingga Lampiran 4 umumnya nilai k yang semakin besar akan menghasilkan *error* yang semakin kecil. Hal ini dapat terjadi dikarenakan semakin banyak nilai yang ikut di dijadikan acuan untuk menghitung nilai prediksi. Semakin banyak data yang digunakan sebagai acuan dalam prediksi, maka hasil prediksi cenderung menjadi lebih akurat atau baik. Namun terdapat titik lembah yang dapat dilihat pada visualisasi Gambar 4.1 hingga Gambar 4.2 yang mana terjadi titik terendah *error*. Hal ini menjelaskan bahwa *error* tidak hanya menurun jika terjadi penambahan nilai K, namun juga terjadi kenaikan *error* meskipun tidak terlalu tinggi.

Tabel 4. 6 Contoh data perbandingan hasil prediksi pada setiap variasi nilai k

Tanggal	Nilai Aktual	<i>Euclidean Distance</i>			<i>Manhattan Distance</i>		
		k = 3	k = 5	k = 7	k = 3	k = 5	k = 7
11/01/2019	0,60	2,73	5,30	4,11	6,80	5,30	14,19
12/01/2019	7,50	16,43	13,04	12,94	16,43	15,00	12,94
17/01/2019	0,00	5,13	4,02	9,66	3,03	4,02	9,66
19/01/2019	2,00	2,83	3,58	8,83	2,83	3,58	6,03
03/02/2019	0,00	10,87	10,86	11,26	10,87	10,86	11,26
04/02/2019	5,20	6,63	8,40	6,60	6,63	8,40	6,71
24/02/2019	0,00	1,47	6,24	4,46	1,47	0,88	4,46

Pada tabel 4.6 merupakan contoh yang diambil dari prediksi pada kondisi data acuan 80% dengan perhitungan prediksi menggunakan metode *average*. Terdapat beberapa data yang semakin besar nilai k maka akan semakin dekat nilai

prediksi dari data aktual.

Perbedaan nilai *error* yang dihasilkan juga terlihat ketika metode perhitungan jarak divariasikan. Pada tabel 4.6 pada penelitian ini dapat dilihat bahwa perhitungan KNN menggunakan antara metode *Euclidean Distance* maupun *Manhattan Distance* menghasilkan nilai *error* yang berbeda. Hal ini terjadi dikarenakan proses perhitungan yang berbeda antara *Manhattan Distance* dan *Euclidean Distance* melakukan perhitungan jarak dengan cara menarik garis lurus antar satu data dengan data lainnya. Proses perhitungan ini sama dengan perhitungan *Pythagoras* pada perhitungan matematika. Sedangkan metode *Manhattan Distance* melakukan perhitungan yang hampir mirip seperti menghitung keliling pada suatu bangun ruang atau bangun datar, metode ini menghitung semua selisih atribut.

Perbedaan nilai *error* yang dihasilkan juga terlihat ketika metode penentuan nilai prediksi divariasikan. Pada tabel 4.7 menunjukkan bahwa perhitungan KNN dengan menggunakan metode *Average* dan *Weighted* menghasilkan nilai *error* yang berbeda. Hal ini dapat terjadi dikarenakan proses penentuan nilai prediksi yang berbeda antara *Average* dan *Weighted*. Metode *Average* melakukan penentuan nilai prediksi dengan merata-ratakan nilai target pada objek data terdekat. Sedangkan metode *Weighted* melakukan penentuan nilai prediksi dengan memberi bobot pada nilai target pada objek data terdekat, semakin dekat data objek semakin besar bobot yang diberikan. Pada tabel 4.7 akan diperlihatkan contoh data yang memiliki hasil prediksi yang berbeda pada setiap variasi metode perhitungan prediksi yang diterapkan.

Tabel 4. 7 Contoh perbandingan hasil prediksi pada variasi metode perhitungan prediksi

Tanggal	Data Aktual	Hasil Prediksi					
		Average			Weighted		
		k = 3	k = 5	k = 7	k = 3	k = 5	k = 7
23/11/2021	0,00	17,20	16,24	11,76	13,73	0,09	0,07
27/11/2021	0,00	0,00	0,00	0,00	0,00	0,00	0,00
06/12/2021	31,40	25,77	24,14	18,74	21,69	1,37	1,11
09/12/2021	0,10	10,53	6,76	7,84	9,99	7,39	5,99
11/12/2021	23,10	22,63	19,00	19,36	10,51	0,00	0,00
20/12/2021	0,00	22,43	13,90	16,07	20,13	15,05	12,21
24/12/2021	3,10	0,17	0,26	2,64	0,00	0,00	0,00

Pada tabel 4.7 merupakan contoh data yang diambil pada kondisi data acuan 80% dengan perhitungan jarak terdekat menggunakan metode *Manhattan Distance*. Dapat dilihat bahwa akan terjadi perbedaan hasil prediksi yang dihasilkan jika metode untuk perhitungan prediksi yang diterapkan berbeda.

Hasil nilai *error* yang besar juga terjadi dikarenakan selisih antara nilai target pada data yang akan diprediksi dengan nilai target pada data tetangga terdekat memiliki selisih yang sangat jauh. Hal ini terjadi dikarenakan data curah hujan yang tidak biasa. Pada tabel 4.8 akan diperlihatkan contoh data acuan yang memiliki data target yang jauh dari data tetangganya:

Tabel 4. 8 Contoh perbandingan nilai target data acuan dan tetangga terdekat

Tanggal	Curah Hujan	Data Terdekat		
		Tanggal	Curah Hujan	Jarak
23-11-2019	92,30	04-11-2020	15.1	0.13
		14-04-2021	59.4	0.15
		24-04-2021	0	0.20
		13-11-2021	2.0	0.22
		13-05-2019	52.5	0.29
		26-11-2020	1.5	0.29
		22-04-2021	105	0.29
01-11-2021	85,50	12-02-2019	0	0.07
		12-03-2019	8.4	0.07
		03-04-2020	80.4	0.07
		27-05-2020	15.8	0.07

		10-07-2020	4.4	0.07
		14-02-2021	12.6	0.07
		29-03-2021	37.4	0.07
19-12-2019	68,00	12-02-2019	0	0.07
		12-03-2019	8.4	0.07
		03-04-2020	80.4	0.07
		27-05-2020	15.8	0.07
		10-07-2020	4.4	0.07
		14-02-2021	12.6	0.07
		29-03-2021	37.4	0.07

Dapat ditinjau pada tabel 4.8 bahwa ada data yang mempunyai target dengan nilai jauh dari nilai target yang dimiliki oleh data tetangga terdekatnya. Pada data acuan tanggal 23-11-2019 memiliki nilai target sebesar 92.3. Sedangkan, dari data tetangga terdekat yang memiliki nilai target yang paling mendekati adalah data tanggal 04-11-2020 dengan nilai target 15.1 dan memiliki selisih sebesar 77.2 dari data target yang dimiliki data acuan.

Pada data acuan tanggal 01-11-2021 memiliki nilai target sebesar 85,50. Sedangkan, dari data tetangga terdekat yang memiliki nilai target yang paling mendekati adalah data tanggal 12-02-2019 dengan nilai target 0 dan memiliki selisih sebesar 85,50 dari data target yang dimiliki data acuan.

Pada data acuan tanggal 19-12-2019 memiliki nilai target sebesar 68,00. Sedangkan, dari data tetangga terdekat yang memiliki nilai target yang paling mendekati adalah data tanggal 12-02-2019 dengan nilai target 0 dan memiliki selisih sebesar 68,00 dari data target yang dimiliki data acuan.

Secara keseluruhan, RMSE pada prediksi KNN menggunakan metode *Average* dari percobaan diatas masih menghasilkan RMSE yang lebih baik dari prediksi KNN menggunakan metode *Weighted*.

#### 4.5 Integrasi Prediksi Hujan dengan Islam

Prediksi adalah istilah yang bisa digunakan untuk meramalkan suatu keadaan maupun kondisi yang akan datang. Prediksi juga sering dijadikan acuan untuk mengambil keputusan terhadap apa yang akan dilakukan, seperti nelayan yang akan mengatur kapan waktu untuk berlayar, petani yang akan mengatur kapan waktu untuk bercocok tanam dan kapan waktu untuk memanen, dan juga bagi pemerintah untuk menentukan kebijakan terhadap kondisi yang akan datang.

Diharapkan penelitian ini dapat membantu masyarakat lain untuk saling membantu dan melindungi diri dari cuaca buruk. Potongan QS Al-Maidah pada ayat 2 telah memberi suatu penjelasan tentang hal itu:

وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ ۖ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ

*”Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan pelanggaran.”* ( QS. Al-Maidah : 2).

Ibnu Katsir memberi penjelasan bahwa Allah memerintahkan semua hamba yang telah mengaku beriman untuk tolong-menolong antar manusia untuk melakukan hal kebaikan dan menjauhi setiap hal yang berujung pada perbuatan dosa. Allah sangat melarang manusia untuk tolong-menolong ketika melakukan keburukan, maksiat, dan perilaku haram lainnya.

Tolong menolong tidak hanya dilakukan kepada sesama muslim saja, namun tolong-menolong juga diperintahkan untuk dilakukan kepada sesama manusia, bahkan kepada orang zholim sekalipun, hadis lainnya dari Anas bin Malik riwayat Imam Ahmad yang membahas hal itu:

*”Tolonglah saudaramu baik yang dalam keadaan berbuat zalim atau di zalimi”* (HR. Imam Ahmad).

Rasulullah juga pernah bersabda hal serupa pada hadis yang diriwayatkan oleh Ibnu Majah:

*“Orang mukmin yang bergaul dengan manusia lainnya dan bersabar atas tindakan yang menyakiti diri mereka lebih baik daripada orang mukmin yang tidak bergaul dengan manusia dan tidak bersabar atas apa yang menyakiti mereka”*(HR Ibnu Majah).

Terdapat beberapa keutamaan dari tolong-menolong, diantaranya dijelaskan pada potongan hadis riwayat Muslim:

*“Barangsiapa yang menghilangkan kesusahan dari kesusahan-kesusahan dunia orang mukmin, maka Allah akan menghilangkan kesusahan dari kesusahan-kesusahan hari kiamat. Barangsiapa yang memberi kemudahan orang yang kesulitan (utang), maka Allah akan memberi kemudahan baginya di dunia dan akhirat.”*(HR Muslim).

Pada hadis itu dijelaskan bahwa orang yang senantiasa tolong-menolong, pertolongan diturunkan oleh Allah untuk orang tersebut di dunia hingga akhirat kelak. Dijelaskan juga pada hadis lain yaitu hadis yang diriwayatkan Bukhari sebagai berikut:

*“Rasulullah shallallahu ‘alaihi wa sallam bersabda “Muslim yang satu adalah saudara muslim yang lain, oleh karena itu ia tidak boleh menganiaya dan mendiamkannya. Barang siapa memperhatikan kepentingan saudaranya, maka Allah akan memperhatikan kepentingannya. Barang siapa membantu kesulitan seorang muslim, maka Allah akan membantu kesulitannya dari beberapa kesulitannya nanti pada hari kiamat. Dan barangsiapa menutupi (aib) seorang muslim, maka Allah akan menutupi (aibnya) pada hari kiamat.” ”* (HR Bukhari).

Pada hadis lainnya terdapat pula Allah akan menolong seseorang selagi orang itu menolong orang lain. Hadis tersebut telah diriwayatkan oleh Imam Muslim:

*“Sesungguhnya Allah akan menolong seorang hamba-Nya selama hamba itu menolong orang lain.”* (HR. Muslim).

Imam Bukhari meriwayatkan hadis yang dengan penjelasan bahwa Allah akan memberi ampunan yang melimpah untuk hamba yang memberi penolongan

terhadap orang lain:

*“Barang siapa menolong orang yang sangat membutuhkan, maka Allah mencatatnya sebanyak 73 ampunan. Satu ampunan terdapat kebaikan semua masalahnya, yang 72 (menaikkan) derajatnya pada hari kiamat.”* (HR. Bukhari).

Oleh karena itu, sikap tolong-menolong akan menjadikan manusia saling bermanfaat bagi manusia lainnya, dan orang yang bermanfaat kepada sesama manusia adalah manusia yang terbaik. Hadis yang diriwayatkan Imam Ahmad telah menjelaskan hal ini:

*“Sebaik-baik manusia adalah yang paling bermanfaat bagi manusia.”*(HR. Ahmad).

Menurut Syaikh Ali bin Hasan Al-Halabi dalam kitab *Ahkamusy Syita’ Fis Sunnatil Muthahharah* dijelaskan bahwa Melakukan dan mempercayai prediksi hujan itu boleh, namun terdapat dua hal yang perlu diperhatikan, boleh mempercayai dengan alasan untuk mempersiapkan diri terhadap apa yang akan terjadi, namun tidak boleh mempercayai serta menyebarluaskan dengan menganggap prediksi tersebut bersifat mutlak atau pasti terjadi (Al-Halabi, n.d.)

Namun, kondisi apapun yang akan terjadi pada dunia ini, segala yang terjadi itu merupakan ujian dari Allah dan sebagai jalan untuk ampunan dan rahmat dari Allah serta pembelajaran bagi setiap manusia yang telah merasa beriman. QS Al-Baqarah ayat 155-157 menjelaskan:

وَلَنَبْلُوَنَّكُمْ بِشَيْءٍ مِّنَ الْخَوْفِ وَالْجُوعِ وَنَقْصٍ مِّنَ الْأَمْوَالِ وَالْأَنْفُسِ وَالثَّمَرَاتِ وَبَشِّرِ الصَّابِرِينَ (155) الَّذِينَ إِذَا  
 أَصَابَتْهُمُ مُصِيبَةٌ قَالُوا إِنَّا لِلَّهِ وَإِنَّا إِلَيْهِ رَاجِعُونَ (156) أُولَئِكَ عَلَيْهِمْ صَلَوَاتٌ مِّن رَّبِّهِمْ وَرَحْمَةٌ وَأُولَئِكَ هُمُ  
 الْمُهْتَدُونَ (157)

*“Dan sungguh akan Kami berikan cobaan kepada kalian dengan sedikit ketakutan, kelaparan, kekurangan harta, jiwa, dan buah-buahan. Dan berikanlah berita gembira kepada orang-orang yang sabar (yaitu) orang-orang yang apabila ditimpa musibah, mereka mengucapkan, "Inna lillahi wainna ilaihi raji'un." Mereka itulah yang mendapat keberkatan yang sempurna dan rahmat dari*

*Tuhannya, dan mereka itulah orang-orang yang mendapat petunjuk” ( QS. Al-Baqarah : 155-157).*

Ibnu Katsir menjelaskan pada kitab tafsirnya, ayat tersebut terdapat penjelasan bahwa Allah SWT pasti menimpakan cobaan pada setiap hambanya dengan tujuan melatih dan menguji keimanan hambanya. Sebagaimana dijelaskan Allah SWT pada QS. Muhammad ayat 3 sebagai berikut:

وَلَنَبْلُوَنَّكُمْ حَتَّىٰ نَعْلَمَ الْمُجَاهِدِينَ مِنكُمْ وَالصَّابِرِينَ وَتَبْلُوَ أَخْبَارَكُمْ

*“Dan sungguh, Kami benar-benar akan menguji kamu sehingga Kami mengetahui orang-orang yang benar-benar berjihad dan bersabar di antara kamu; dan akan Kami uji perihal kamu.” ( QS. Muhammad : 3).*

Namun bagi hambanya yang beriman pastilah ia akan bersabar dan mengatakan *“Innalillahi wa inna ilaihi raji’un”*. Dan orang beriman serta sabar ketika mendapat ujian akan dilimpahkan rahmat serta keberkahan dari Allah SWT. Terkadang Allah memberi ujian kepada manusia dengan kebahagiaan, dan ketika masa yang lain Allah memberi ujian kepada manusia dengan kesusahan seperti bencana, rasa takut, dan kelaparan. Orang yang diuji dengan merasakan kesusahan dan lapar serta ketakutan akan terlihat keimanan sesungguhnya.

Manusia yang mengucapkan *“Innalillahi wa inna ilaihi raji’un”* sewaktu ditimpakan musibah terhadap mereka, maka mereka sesungguhnya mengetahui bahwasanya setiap yang ada di dunia hanyalah milik Allah. Mereka mengetahui bahwa segala yang telah dilakukan tidak akan sia-sia di hadapan Allah. Mereka mengakui bahwa mereka hanyalah seorang hamba yang lemah dan akan kembali kepada Allah saat di akhirat kelak. Bagian akhir ayat terdapat penjelasan bahwa Allah memberikan keberkahan yang sempurna serta mendapat petunjuk. Allah akan memberikan ganjaran terhadap kesabaran yang telah dilalui seorang hamba,

ganjaran berupa nikmat yang tidak terduga. Selain itu, Allah juga akan memberi petunjuk kepada hamba tersebut untuk kembali ke jalan Allah serta petunjuk dalam menjalani kehidupan yang baik. Selain dua keutamaan tersebut, Allah akan memberi pahala atas musibah yang telah menimpanya atas kesabarannya dalam menghadapi musibah tersebut serta keselamatan dari azab. Salah satu hadis yang menyampaikan keutamaan dari mengucapkan *“Innalillahi wa inna ilaihi raji’un”* sewaktu mendapat suatu musibah terdapat pada hadis Shahih Muslim, dari Ummu Salamah:

*“Tidaklah seorang hamba ditimpa musibah, lalu ia mengucapkan “innalillahi wa inna ilaihi raji’un. Ya Allah, berikanlah pahala dalam musibahku ini dan berikanlah ganti kepadaku yang lebih baik darinya”, melainkan Allah akan memberikan pahala kepadanya dalam musibah itu dan memberikan ganti kepadanya dengan yang lebih baik darinya”*(HR. Muslim).

Hadis lainnya dari Abu Sinan riwayat Imam Ahmad:

*“Allah berfirman “ Hai malaikat maut, apakah engkau sudah mencabut nyawa anak hambaku?”, malaikat menjawab “Iya”, lalu Allah bertanya “Apa yang ia ucapkan?”, lalu malaikat menjawab “Ia memujimu dan mengucapkan kalimat istirja’(innalillahi wa inna ilaihi raji’un).”, Maka Allah berfirman kepada para malaikat : “Buatkan untuknya sebuah rumah di surga, dan namailah rumah itu dengan Baitul Hamdi (Rumah Pujian)”.”*(HR. Imam Ahmad).

Namun Allah tidak akan pernah memberikan cobaan kepada manusia dengan ujian yang melebihi kemampuannya. Setiap ujian yang diberikan kepada manusia adalah ujian yang pasti dapat mampu dilalui oleh manusia tersebut. QS.

Al-Baqarah pada ayat 286 menjelaskan sebagai berikut:

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا ۗ لَهَا مَا كَسَبَتْ وَعَلَيْهَا مَا اكْتَسَبَتْ ۗ رَبَّنَا لَا تُؤَاخِذْنَا إِنْ نَسِينَا أَوْ أَخْطَأْنَا ۗ رَبَّنَا وَلَا تَحْمِلْ عَلَيْنَا إَصْرًا كَمَا حَمَلْتَهُ ۗ عَلَى الَّذِينَ مِنْ قَبْلِنَا ۗ رَبَّنَا وَلَا تُحَمِّلْنَا مَا لَا طَاقَةَ لَنَا بِهِ ۗ وَاعْفُ عَنَّا ۗ وَارْحَمْنَا ۗ أَنْتَ مَوْلَانَا فَانصُرْنَا عَلَى الْقَوْمِ الْكَافِرِينَ

*“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. Dia mendapat (pahala) dari (kebajikan) yang dikerjakannya dan dia mendapat (siksa) dari (kejahatan) yang diperbuatnya. (Mereka berdoa), “Ya Tuhan kami, janganlah*

*Engkau hukum kami jika kami lupa atau kami melakukan kesalahan. Ya Tuhan kami, janganlah Engkau bebani kami dengan beban yang berat sebagaimana Engkau bebani kepada orang-orang sebelum kami. Ya Tuhan kami, janganlah Engkau pikulkan kepada kami apa yang tidak sanggup kami memikulnya. Maafkanlah kami, ampunilah kami, dan rahmatilah kami. Engkaulah pelindung kami, maka tolonglah kami menghadapi orang-orang kafir.”” (QS. Al-Baqarah : 286)*

Ibnu Katsir menjelaskan pada kitab tafsir nya, ayat tersebut memberi penjelasan Allah tidak menimpa suatu cobaan terhadap manusia yang tidak sanggup menerima cobaan yang akan diberikan. Hal ini adalah suatu kelembutan, serta kasih sayang Allah ke setiap umat-Nya. Adapun seseorang yang tidak mampu untuk menolak segala godaan dan bisikan jiwa, maka cobaan tidak akan diberikan terhadapnya. Dan rasa benci kepada hal buruk itu merupakan sebagian dari iman. Bagi manusia yang dapat melewati cobaan itu, Allah akan pahala atas kebaikan yang dilakukan. Bagi manusia yang melakukan kejahatan, maka Allah akan memberi siksa atas kejahatan yang telah dilakukannya.

Pada ayat ini Allah juga telah menyematkan doa agar dilancarkan dalam menghadapi cobaan. Terdapat doa untuk memohon kepada Allah untuk tidak diberi hukuman karena meninggalkan kewajiban ataupun mengerjakan hal haram karena lupa atau tidak mengetahui hal yang benar secara syariat. Pada doa ini juga terdapat permohonan agar manusia tidak diberi cobaan berat meskipun mampu untuk ditunaikan. Dan ayat ini ditutup dengan doa agar Allah mengampuni dan memberi rahmat kepada manusia dan permohonan untuk pertolongan Allah kepada orang beriman ketika mendapat cobaan yang datang melalui orang kafir (Muhammad, 2005).

Namun kembali lagi bahwa penelitian ini dilakukan salah satunya sebagai sebuah pembelajaran untuk manusia lainnya. QS. Al-Baqarah ayat 31 menjelaskan:

وَعَلَّمَ آدَمَ الْأَسْمَاءَ كُلَّهَا ثُمَّ عَرَضَهُمْ عَلَى الْمَلَائِكَةِ فَقَالَ أَنْبِئُونِي بِأَسْمَاءِ هَؤُلَاءِ إِنْ كُنْتُمْ صَادِقِينَ

“Dan Dia ajarkan kepada Adam nama-nama (benda) semuanya, kemudian Dia perlihatkan kepada para malaikat, seraya berfirman, “Sebutkan kepada-Ku nama semua (benda) ini, jika kamu yang benar!” (QS. Al-Baqarah : 31)

Ibnu Katsir memberi penjelasan ayat tersebut menunjukkan Allah menurunkan ilmu serta pemahaman untuk mengetahui dan memahami kepada manusia dengan perkara yang tidak diketahuinya sebelumnya. Hal ini dilakukan agar manusia dapat memanfaatkan ilmu dan kemampuan tersebut untuk meningkatkan kualitas hidup dan mengatasi masalah-masalah yang dihadapi.

Ibnu Katsir juga menyatakan bahwa Allah menurunkan terhadap manusia kesempatan dalam memperoleh ilmu dan kemampuan yang diperlukan untuk mendapatkan kebahagiaan kehidupan di dunia hingga akhirat. Ibnu Katsir menjelaskan ayat itu menunjukkan manusia diwajibkan untuk berusaha dalam mencari ilmu dan meningkatkan kemampuannya, karena ilmu dan kemampuan yang dimiliki seseorang akan menentukan kualitas hidupnya (Muhammad, 2005).

QS. AL-Baqarah pada ayat 269 menjelaskan tentang keutamaan orang yang memiliki ilmu:

يُؤْتِي الْحِكْمَةَ مَنْ يَشَاءُ ۚ وَمَنْ يُؤْتَ الْحِكْمَةَ فَقَدْ أُوتِيَ خَيْرًا كَثِيرًا ۗ وَمَا يَذَّكَّرُ إِلَّا أُولُو الْأَلْبَابِ

“Allah menganugerahkan al hikmah (kefahaman) kepada siapa yang dikehendaki-Nya. Dan barangsiapa yang dianugerahi hikmah, ia benar-benar telah dianugerahi karunia yang banyak. Dan hanya orang-orang yang berakallah yang dapat mengambil pelajaran.” (QS. Al-Baqarah : 269)

Menurut kitab tafsir Ibnu Katsir, ayat ini menunjukkan hikmah (kebijaksanaan) adalah sesuatu yang sangat berharga dan bermanfaat bagi

seseorang. Orang yang diberi hikmah akan menerima banyak kebaikan, karena hikmah akan membantunya dalam mengambil keputusan dan menentukan tindakan yang tepat dalam hidup.

Ibnu Katsir juga menyatakan bahwa hikmah tidak bermanfaat untuk duniawi saja, akan tetapi juga bermanfaat untuk agama. Seseorang yang diberi hikmah akan dapat memahami ajaran agama dengan baik dan mengaplikasikan ajaran tersebut dalam kehidupannya.

Sedangkan orang yang tidak diberi hikmah, tidak akan dapat mengambil pelajaran dari setiap peristiwa, sehingga akan kesulitan dalam mengambil keputusan yang tepat dalam hidup. (Muhammad, 2005).

Namun perlu diketahui juga bahwa bencana yang terjadi di bumi ini tidak hanya karena faktor alam saja, namun terdapat beberapa merupakan akibat oleh manusia itu sendiri, contohnya adalah banjir yang diakibatkan oleh saluran air yang tersumbat oleh sampah yang dihasilkan manusia dan penebangan ataupun pembakaran hutan sehingga kurangnya resapan air oleh tanah. Lalu ada juga longsor yang diakibatkan oleh penambangan ataupun pengikisan batu bukit oleh manusia. Banyak ayat-ayat didalam Qur'an menjelaskan hal ini, Q.S Ar-Rum ayat 41 dan 42 diantara ayat-ayat itu berbunyi:

ظَهَرَ الْفَسَادُ فِي الْبَرِّ وَالْبَحْرِ بِمَا كَسَبَتْ أَيْدِي النَّاسِ لِيُذِيقَهُمْ بَعْضَ الَّذِي عَمِلُوا لَعَلَّهُمْ يَرْجِعُونَ (41) قُلْ  
سِيرُوا فِي الْأَرْضِ فَانظُرُوا كَيْفَ كَانَ عَاقِبَةُ الَّذِينَ مِنْ قَبْلُ ۚ كَانُوا أَكْثَرَهُمْ مُشْرِكِينَ (42)

“*Telah nampak kerusakan di darat dan di laut disebabkan karena perbuatan tangan manusia, supaya Allah merasakan kepada mereka sebahagian dari (akibat) perbuatan mereka, agar mereka kembali (ke jalan yang benar). Katakanlah: “Adakanlah perjalanan di muka bumi dan perhatikanlah bagaimana kesudahan orang-orang yang terdahulu. kebanyakan dari mereka itu adalah orang-orang yang mempersekutukan (Allah).”* (QS. Ar Rum 41-42).

Menurut penjelasan tafsir Ibnu ayat tersebut Allah memberi penegasan bahwa manusia telah melakukan kerusakan diatas bumi. Lalu Allah mendatangkan akibat kerusakan itu berupa bencana kepada manusia agar manusia itu menyesal dan kembali kepada Allah. Dan kerusakan dilakukan oleh manusia di bumi telah terjadi sejak dahulu, serta telah Allah beri juga dampaknya kepada mereka. Maka jika dilakukan sebuah perjalanan diatas bumi, maka peninggalan-peninggalan kaum terdahulu akan dapat dilihat untuk dijadikan pembelajaran (Muhammad, 2005).

Pada ayat lain, yaitu pada QS. Al-Baqarah ayat 60 Allah terdapat penjelasan lain bahwa kerusakan di atas bumi adalah akibat dari manusia:

وَإِذِ اسْتَسْقَىٰ مُوسَىٰ لِقَوْمِهِ ۖ فَقُلْنَا اضْرِبْ بِعَصَاكَ الْحَجَرَ ۖ فَانفَجَرَتْ مِنْهُ اثْنَتَا عَشْرَةَ عَيْنًا ۗ قَدْ عَلِمَ كُلُّ  
 أُنَاسٍ مَّشْرَبَهُمْ ۗ كُلُوا وَاشْرَبُوا مِنْ رِزْقِ اللَّهِ وَلَا تَعْتُوا فِي الْأَرْضِ مُفْسِدِينَ

*“Dan (ingatlah) ketika Musa memohon air untuk kaumnya, lalu Kami berfirman, “Pukullah batu itu dengan tongkatmu!” Maka memancarlah daripadanya dua belas mata air. Setiap suku telah mengetahui tempat minumnya (masing-masing). Makan dan minumlah dari rezeki (yang diberikan) Allah, dan janganlah kamu melakukan kejahatan di bumi dengan berbuat kerusakan.”(QS. Al-Baqarah : 60).*

Ibnu Katsir menjelaskan pada kitabnya bahwa ayat tersebut menceritakan kisah bani Israil yang tidak mensyukuri nikmat yang diturunkan kepada kaum itu. Kaum itu bahkan sampai melakukan kerusakan yang mereka anggap lebih baik dibandingkan yang telah Allah berikan. Ini kembali sebagai bukti bahwa manusia dapat melakukan kerusakan untuk memenuhi hasrat ketamakan duniawinya (Muhammad, 2005).

Pada QS. Al-A'raf pada ayat 56 telah diberitakan dengan jelas agar tidak melakukan kerusakan di atas permukaan bumi:

وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ خَوْفًا وَطَمَعًا ۚ إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ

*“Dan janganlah kamu berbuat kerusakan di bumi setelah (diciptakan) dengan baik. Berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat*

*Allah sangat dekat kepada orang yang berbuat kebaikan.”(QS. Al-A’raf : 56).*

Menurut Ibnu Katsir pada ayat ini dengan jelas dilarang merusak di atas muka bumi ketika bumi diperbaiki Hal telah berjalan dengan baik atas izin Allah lalu dirusak oleh manusia, maka itu akan membahayakan bagi umat manusia itu sendiri (Muhammad, 2005).

Pada Q.S Shod ayat 28 dijelaskan setiap orang melakukan kerusakan dan zhalim akan dibedakan dengan setiap orang yang telah beriman dan telah melakukan kebajikan:

أَمْ نَجْعَلُ الَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ كَالْمُفْسِدِينَ فِي الْأَرْضِ أَمْ نَجْعَلُ الْمُتَّقِينَ كَالْفُجَّارِ

*“Pantas kah Kami memperlakukan orang-orang yang beriman dan mengerjakan kebajikan sama dengan orang-orang yang berbuat kerusakan di bumi? Atau pantas kah Kami menganggap orang-orang yang bertakwa sama dengan orang-orang yang jahat?” (QS. Shad : 28)*

Pada kitab tafsir Ibnu Katsir, ayat tersebut menerangkan manusia yang telah beriman dan telah mengerjakan suatu kebajikan itu berbeda dengan hambanya yang senantiasa merusak di atas muka bumi. Setiap hamba telah merusak di atas bumi demi kepentingan duniawinya atas dasar ketamakan dan kufur nikmat terhadap dunia. Maka orang-orang ini sesungguhnya telah berbuat suatu kezaliman terhadap orang lain dan dirinya sendiri (Muhammad, 2005).

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan untuk mendapatkan performa terbaik algoritma KNN dalam memprediksi curah hujan di kota Pekanbaru. Pengujian yang dilakukan pada bab sebelumnya dilakukan untuk mendapatkan *hyperparameter* yang paling optimal digunakan pada algoritma KNN berdasarkan nilai *error* berupa RMSE yang dihasilkan pada setiap percobaan. Setelah dilakukan pengujian, maka hasil yang paling optimal dihasilkan pada skenario 4 (pembagian data acuan dan data *testing* masing-masing 50% - 50%) dengan metode perhitungan jarak *Euclidean Distance* dan metode perhitungan prediksi menggunakan *Average* pada  $K = 23$  dengan RMSE sebesar 15,98. Hasil yang telah disajikan pada bab sebelumnya juga menunjukkan perbandingan hasil dari variasi *hyperparameter* yang digunakan. Algoritma KNN dengan metode perhitungan jarak *Euclidean Distance* dan *Manhattan Distance* unggul dan rendah pada kondisi tertentu. Algoritma KNN dengan metode perhitungan nilai prediksi menggunakan *Average* lebih baik dibandingkan dengan *Weighted*. Parameter  $K$  yang tinggi menghasilkan nilai *error* yang lebih baik walaupun terjadi kenaikan, namun tidak besar.

#### 5.2 Saran

Berdasar dari eksperimen yang telah dilakukan, terdapat beberapa saran yang diharapkan pada penelitian selanjutnya sebagai berikut :

1. Penelitian ini algoritma KNN yang diimplementasikan untuk menentukan nilai prediksi yaitu algoritma KNN dengan perhitungan nilai prediksi menggunakan metode *Average* dan *Weighted*, penelitian berikutnya diharapkan untuk dapat menggunakan metode KNN yang lainnya seperti PKNN, LKNN, dan sebagainya.
2. Penelitian ini untuk perhitungan jarak hanya menggunakan cara *Euclidean Distance* maupun *Manhattan Distance*, pada penelitian selanjutnya diharapkan untuk bisa menggunakan metode lainnya untuk menghitung jarak antar data seperti *Chebyshev Distance*, *Hamming Distance*, *Canberra Distance*, dan *Minkowski Distance*.

## DAFTAR PUSTAKA

- Adhani, G. (2013). *Pendugaan Curah Hujan Musim Kemarau Menggunakan Data Southern Oscillation Index dan Suhu Permukaan Laut NINO3. 4 dengan Metode Support Vector Regression*.
- Al-Halabi, A. H. (n.d.). *Ahkam asy-Syita' fi as-Sunnah al-Muthahharoh*.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Danuri, M. (2019). Perkembangan dan transformasi teknologi digital. *Jurnal Ilmiah Infokam*, 15(2). <https://doi.org/10.53845/infokam.v15i2.178>
- Deza, M. M., & Deza, E. (2009). Encyclopedia of distances. In *Encyclopedia of distances* (pp. 1–583). Springer. <https://doi.org/10.1007/978-3-642-30958-8>
- Dudani, S. A. (1976). The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, 4, 325–327. <https://doi.org/10.1109/TSMC.1976.5408784>
- Fadilah, W. R. U., Agfiannisa, D., & Azhar, Y. (2020). Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine. *Fountain of Informatics Journal*, 5(2), 45. <https://doi.org/10.21111/fij.v5i2.4449>
- Fatah, H., & Subekti, A. (2018). Prediksi Harga Cryptocurrency Dengan Metode K-Nearest Neighbours. *Jurnal Pilar Nusa Mandiri*, 14(2), 137–144. <https://doi.org/10.33480/pilar.v14i2.30>
- Gou, J., Yi, Z., Du, L., & Xiong, T. (2012). A local mean-based k-nearest centroid neighbor classifier. *The Computer Journal*, 55(9), 1058–1071. <https://doi.org/10.1093/comjnl/bxr131>
- Handayani, L., & Adri, M. (2015). Penerapan JST (Backpropagation) untuk Prediksi curah hujan (Studi kasus: Kota Pekanbaru). *Seminar Nasional Teknologi Informasi Komunikasi Dan Industri*.
- Hasmawati, J. N., & Muchtar, M. (n.d.). *Aplikasi prediksi penjualan barang menggunakan metode k-nearest neighbor (knn)(studi kasus tumaka mart)*. <https://doi.org/10.55679/semantik.v3i2.3658>
- Jumarang, M. I., & others. (n.d.). Analisis Hubungan Kelembaban Udara dan Suhu Udara Terhadap Parameter Tebal Hujan di Kota Pontianak. *PRISMA FISIKA*, 4(3). <https://doi.org/10.26418/pf.v4i3.16830>
- Kafil, M. (2019). Penerapan Metode K-Nearest Neighbors Untuk Prediksi

- Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(2), 59–66. <https://doi.org/10.36040/jati.v3i2.860>
- Kuang, L., Yan, H., Zhu, Y., Tu, S., & Fan, X. (2019). Predicting duration of traffic accidents based on cost-sensitive Bayesian network and weighted K-nearest neighbor. *Journal of Intelligent Transportation Systems*, 23(2), 161–174. <https://doi.org/10.1080/15472450.2018.1536978>
- Lareno, B. (2015). Analisa dan Perbandingan akurasi model prediksi rentet waktu arus lalu lintas jangka pendek. *CSRID (Computer Science Research and Its Development Journal)*, 6(3), 148–158. <https://doi.org/10.22303/csrid.6.3.2014.148-158>
- Lasulika, M. E. (2017). Prediksi Harga Komoditi Jagung Menggunakan K-Nn Dan Particle Swarm Optimazation Sebagai Fitur Seleksi. *ILKOM Jurnal Ilmiah*, 9(3), 233–238. <https://doi.org/10.33096/ilkom.v9i3.148.233-238>
- Leidiana, H. (2013). Penerapan algoritma k-nearest neighbor untuk penentuan resiko kredit kepemilikan kendaraan bermotor. *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, 1(1), 65–76.
- Mirawati, T. D., Yasin, H., & Rusgiyono, A. (2013). Prediksi curah hujan di Kota Semarang dengan Metode Kalman Filter. *Prosiding Seminar Nasional Statistika Universitas Diponegoro 2013*, 561–570.
- Muhammad, A. (2005). *Tafsir Ibnu Katsir*. Pustaka Imam Asy-Syafi'i.
- Prawaka, F., Zakaria, A., & Tugiono, S. (2016). Analisis Data Curah Hujan yang Hilang Dengan Menggunakan Metode Normal Ratio, Inversed Square Distance, dan Cara Rata-Rata Aljabar (Studi Kasus Curah Hujan Beberapa Stasiun Hujan Daerah Bandar Lampung). *Jurnal Rekayasa Sipil Dan Desain*, 4(3), 397–406.
- Radhitya, M. L., & Sudipa, G. I. (2020). PENDEKATAN Z-SCORE DAN FUZZY DALAM PENGUJIAN AKURASI PERAMALAN CURAH HUJAN. *SINTECH (Science and Information Technology) Journal*, 3(2), 149–156. <https://doi.org/10.31598/sintechjournal.v3i2.567>
- Ritha, N., Bettiza, M., & Dufan, A. (2016). Prediksi Curah Hujan dengan Menggunakan Algoritma Levenberg-Marquardt dan Backpropagation. *Jurnal Sustainable: Jurnal Hasil Penelitian Dan Industri Terapan*, 5(2), 11–16. <https://doi.org/10.31629/sustainable.v5i2.358>
- Sabilla, W. I., & Putri, T. E. (2017). Prediksi Ketepatan Waktu Lulus Mahasiswa dengan k-Nearest Neighbor dan Naïve Bayes Classifier. *Jurnal Komputer Terapan*, 3(2), 233–240. <https://doi.org/10.35143/jkt>

- Saragih, I. J. A., Rumahorbo, I., Yudistira, R., & Sucahyono, D. (2020). Prediksi Curah Hujan Bulanan Di Deli Serdang Menggunakan Persamaan Regresi Dengan Prediktor Data Suhu Dan Kelembapan Udara. *Jurnal Meteorologi Klimatologi Dan Geofisika*, 7(2), 6–14.
- Shedriko, S. (2021). Perbandingan Algoritma SVM dan KNN dalam Mengklasifikasi Kelulusan Mahasiswa pada Suatu Mata Kuliah. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 6(2), 115–122. <https://doi.org/10.30998/string.v6i2.9160>
- Susanto, H., & Sudiyatno, S. (2014). Data mining untuk memprediksi prestasi siswa berdasarkan sosial ekonomi, motivasi, kedisiplinan dan prestasi masa lalu. *Jurnal Pendidikan Vokasi*, 4(2). <https://doi.org/10.21831/jpv.v4i2.2547>
- Suwanda, R., Syahputra, Z., & Zamzami, E. M. (2020). Analysis of euclidean distance and manhattan distance in the K-means algorithm for variations number of centroid K. *Journal of Physics: Conference Series*, 1566(1), 12058. <https://doi.org/10.1088/1742-6596/1566/1/012058>
- Syaifullah, M. D. (2014). Validasi data TRMM terhadap data curah hujan aktual di tiga DAS di Indonesia. *Jurnal Meteorologi Dan Geofisika*, 15(2). <https://doi.org/10.31172/jmg.v15i2.180>
- Utomo, P. B., Utami, E., & Raharjo, S. (2019). IMPLEMENTASI METODE K-NEAREST NEIGHBOR DAN REGRESI LINEAR DALAM PREDIKSI HARGA EMAS. *Informasi Interaktif*, 4(3), 155–159.
- Warsito, B., & Sumiyati, S. (2007). Prediksi Curah Hujan kota Semarang dengan Feedforward Neural Network Menggunakan Algoritma Quasi Newton BFGS dan Levenberg-Marquardt. *Jurnal Presipitasi*, 3(2), 46–52. <https://doi.org/10.14710/presipitasi.v3i2.46-52>
- Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., & others. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1–37. <https://doi.org/10.1007/s10115-007-0114-2>

## LAMPIRAN

### Lampiran 1

#### RMSE Euclidean Distance

K	RMSE							
	Euclidean Distance							
	80%-20%		70%-30%		60%-40%		50%-50%	
	Average	Weighted	Average	Weighted	Average	Weighted	Average	Weighted
1	20,90	20,90	21,83	21,83	21,94	21,94	20,81	20,81
2	19,03	19,42	19,49	20,24	20,15	20,58	19,84	19,82
3	18,85	18,93	18,59	19,39	19,06	19,78	18,36	19,09
4	18,16	18,60	18,10	18,88	18,22	19,25	17,40	18,53
5	18,03	18,38	17,87	18,56	17,55	18,80	16,99	18,12
6	17,96	18,23	17,68	18,33	17,13	18,45	16,76	17,82
7	17,98	18,15	17,78	18,19	16,98	18,17	16,76	17,60
8	18,02	18,10	17,68	18,07	17,11	17,98	16,63	17,41
9	17,73	18,02	17,77	17,99	16,95	17,82	16,46	17,26
10	17,62	17,95	17,67	17,92	16,89	17,68	16,43	17,13
11	17,39	17,87	17,56	17,86	16,73	17,55	16,42	17,02
12	17,37	17,80	17,56	17,80	16,65	17,45	16,43	16,94
13	17,26	17,74	17,31	17,74	16,55	17,35	16,39	16,87
14	17,13	17,68	17,33	17,69	16,46	17,26	16,38	16,81
15	17,12	17,62	17,26	17,65	16,59	17,19	16,40	16,76
16	17,16	17,58	17,24	17,61	16,65	17,14	16,35	16,71
17	17,10	17,53	17,25	17,57	16,63	17,09	16,25	16,67

18	17,00	17,49	17,30	17,54	16,65	17,05	16,30	16,63
19	16,97	17,45	17,29	17,52	16,60	17,01	16,25	16,60
20	16,93	17,42	17,31	17,50	16,61	16,98	16,11	16,56
21	16,86	17,38	17,29	17,48	16,60	16,94	16,03	16,52
22	16,83	17,34	17,24	17,46	16,58	16,92	15,99	16,49
23	16,80	17,31	17,27	17,44	16,51	16,89	15,98	16,45
24	16,75	17,28	17,24	17,43	16,51	16,86	16,03	16,43
25	16,68	17,25	17,25	17,41	16,57	16,84	16,05	16,40
26	16,71	17,22	17,25	17,40	16,60	16,83	16,10	16,38
27	16,75	17,20	17,24	17,39	16,58	16,81	16,13	16,37
28	16,79	17,18	17,21	17,37	16,57	16,79	16,17	16,35
29	16,77	17,16	17,22	17,36	16,60	16,78	16,19	16,34
30	16,71	17,13	17,25	17,35	16,63	16,77	16,21	16,33
31	16,72	17,12	17,20	17,34	16,58	16,76	16,22	16,32
32	16,69	17,10	17,18	17,33	16,58	16,75	16,23	16,31
33	16,68	17,08	17,20	17,32	16,60	16,74	16,21	16,30
34	16,70	17,06	17,17	17,31	16,58	16,73	16,21	16,29
35	16,71	17,05	17,19	17,30	16,56	16,72	16,19	16,29
36	16,69	17,03	17,19	17,29	16,57	16,71	16,16	16,28
37	16,72	17,02	17,22	17,29	16,58	16,70	16,17	16,27
38	16,73	17,01	17,17	17,28	16,58	16,70	16,18	16,27
39	16,76	17,00	17,20	17,27	16,61	16,69	16,18	16,26
40	16,71	16,99	17,21	17,27	16,58	16,68	16,20	16,25
41	16,72	16,98	17,21	17,26	16,56	16,68	16,21	16,25
42	16,73	16,97	17,19	17,26	16,57	16,67	16,22	16,25

43	16,70	16,96	17,19	17,25	16,58	16,66	16,24	16,24
44	16,71	16,95	17,17	17,25	16,55	16,66	16,24	16,24
45	16,69	16,94	17,19	17,24	16,55	16,65	16,24	16,24
46	16,67	16,93	17,20	17,24	16,56	16,65	16,22	16,24
47	16,69	16,92	17,22	17,24	16,56	16,64	16,24	16,23
48	16,70	16,92	17,23	17,24	16,57	16,64	16,26	16,23
49	16,70	16,91	17,24	17,23	16,59	16,64	16,23	16,23
50	16,68	16,90	17,24	17,23	16,60	16,63	16,22	16,23
51	16,67	16,90	17,25	17,23	16,61	16,63	16,23	16,22
52	16,71	16,89	17,28	17,23	16,61	16,63	16,23	16,22
53	16,75	16,89	17,28	17,23	16,62	16,63	16,23	16,22
54	16,74	16,88	17,25	17,22	16,62	16,62	16,24	16,22
55	16,77	16,88	17,27	17,22	16,64	16,62	16,25	16,22
56	16,79	16,87	17,26	17,22	16,65	16,62	16,27	16,22
57	16,77	16,87	17,26	17,22	16,67	16,62	16,30	16,22
58	16,75	16,87	17,27	17,22	16,66	16,62	16,31	16,22
59	16,72	16,86	17,25	17,22	16,69	16,62	16,31	16,22
60	16,74	16,86	17,23	17,22	16,69	16,62	16,33	16,22
61	16,72	16,86	17,24	17,22	16,69	16,62	16,32	16,22
62	16,71	16,85	17,25	17,21	16,71	16,62	16,32	16,22
63	16,72	16,85	17,26	17,21	16,70	16,62	16,34	16,22
64	16,73	16,84	17,27	17,21	16,71	16,62	16,35	16,22
65	16,72	16,84	17,26	17,21	16,71	16,62	16,35	16,22
66	16,71	16,84	17,29	17,21	16,71	16,62	16,37	16,22
67	16,71	16,83	17,28	17,21	16,72	16,62	16,37	16,22

68	16,73	16,83	17,31	17,21	16,72	16,62	16,37	16,22
69	16,73	16,83	17,30	17,21	16,71	16,62	16,37	16,22
70	16,71	16,83	17,30	17,21	16,72	16,62	16,38	16,22
71	16,72	16,82	17,27	17,21	16,73	16,62	16,39	16,22
72	16,72	16,82	17,29	17,21	16,72	16,62	16,38	16,23
73	16,69	16,82	17,30	17,21	16,70	16,62	16,38	16,23
74	16,70	16,82	17,29	17,21	16,71	16,62	16,37	16,23
75	16,70	16,81	17,27	17,21	16,72	16,62	16,36	16,23
76	16,69	16,81	17,28	17,21	16,74	16,62	16,35	16,23
77	16,70	16,81	17,28	17,21	16,74	16,62	16,37	16,23
78	16,69	16,81	17,30	17,21	16,74	16,62	16,35	16,23
79	16,70	16,80	17,31	17,21	16,74	16,62	16,36	16,23
80	16,72	16,80	17,31	17,21	16,76	16,62	16,36	16,23
81	16,73	16,80	17,28	17,21	16,76	16,62	16,37	16,23
82	16,75	16,80	17,27	17,21	16,76	16,62	16,37	16,23
83	16,77	16,80	17,27	17,21	16,75	16,62	16,37	16,24
84	16,76	16,80	17,28	17,21	16,76	16,62	16,37	16,24
85	16,78	16,79	17,26	17,21	16,77	16,63	16,39	16,24
86	16,79	16,79	17,27	17,21	16,76	16,63	16,39	16,24
87	16,78	16,79	17,27	17,21	16,76	16,63	16,40	16,24
88	16,78	16,79	17,28	17,21	16,77	16,63	16,40	16,24
89	16,79	16,79	17,28	17,21	16,78	16,63	16,39	16,24
90	16,76	16,79	17,31	17,21	16,78	16,63	16,40	16,24
91	16,76	16,79	17,32	17,21	16,78	16,63	16,40	16,24
92	16,77	16,79	17,32	17,21	16,79	16,63	16,40	16,24

93	16,76	16,79	17,31	17,21	16,79	16,63	16,40	16,25
94	16,75	16,79	17,33	17,21	16,80	16,63	16,40	16,25
95	16,76	16,78	17,34	17,21	16,79	16,63	16,41	16,25
96	16,76	16,78	17,33	17,22	16,78	16,64	16,40	16,25
97	16,76	16,78	17,34	17,22	16,77	16,64	16,41	16,25
98	16,77	16,78	17,36	17,22	16,77	16,64	16,41	16,25
99	16,75	16,78	17,37	17,22	16,77	16,64	16,42	16,25
100	16,76	16,78	17,36	17,22	16,77	16,64	16,43	16,25

**Lampiran 2**

**RMSE Manhattan Distance**

K	Manhattan Distance							
	80%-20%		70%-30%		60%-40%		50%-50%	
	Average	Weighted	Average	Weighted	Average	Weighted	Average	Weighted
1	20,89	20,89	22,00	22,00	21,65	21,65	20,85	20,85
2	19,32	19,60	19,27	20,22	19,62	20,23	19,43	19,64
3	19,18	19,14	17,91	19,16	18,68	19,41	18,29	18,96
4	18,17	18,75	17,79	18,63	17,98	18,90	17,32	18,42
5	17,90	18,46	17,85	18,34	17,34	18,48	16,92	18,01
6	18,03	18,31	17,71	18,15	17,15	18,18	16,67	17,70
7	17,85	18,19	17,57	18,01	16,98	17,95	16,52	17,46
8	17,71	18,10	17,68	17,92	16,94	17,77	16,59	17,29
9	17,77	18,03	17,48	17,83	16,94	17,63	16,50	17,15
10	17,76	17,96	17,34	17,75	16,79	17,51	16,39	17,03

11	17,52	17,89	17,39	17,69	16,66	17,39	16,36	16,93
12	17,50	17,83	17,28	17,63	16,59	17,30	16,35	16,85
13	17,50	17,78	17,30	17,58	16,64	17,22	16,31	16,78
14	17,51	17,74	17,36	17,54	16,49	17,14	16,32	16,72
15	17,24	17,69	17,38	17,51	16,56	17,08	16,25	16,67
16	17,20	17,65	17,39	17,49	16,55	17,03	16,23	16,62
17	17,08	17,60	17,40	17,47	16,52	16,98	16,23	16,58
18	17,02	17,55	17,29	17,45	16,51	16,94	16,20	16,54
19	16,99	17,51	17,29	17,43	16,50	16,90	16,13	16,51
20	16,91	17,47	17,34	17,42	16,45	16,87	16,15	16,48
21	16,92	17,44	17,33	17,40	16,46	16,84	16,13	16,45
22	16,92	17,41	17,29	17,39	16,46	16,81	16,09	16,42
23	16,94	17,38	17,27	17,38	16,50	16,79	16,11	16,40
24	16,96	17,35	17,31	17,36	16,48	16,77	16,09	16,38
25	16,94	17,33	17,27	17,35	16,51	16,75	16,10	16,36
26	16,93	17,30	17,28	17,34	16,56	16,73	16,11	16,34
27	16,85	17,28	17,32	17,33	16,54	16,72	16,15	16,33
28	16,82	17,26	17,29	17,33	16,55	16,70	16,14	16,31
29	16,79	17,23	17,27	17,32	16,53	16,69	16,13	16,30
30	16,81	17,21	17,18	17,31	16,54	16,68	16,10	16,29
31	16,87	17,20	17,14	17,29	16,55	16,67	16,14	16,28
32	16,87	17,18	17,15	17,28	16,55	16,66	16,16	16,27
33	16,82	17,17	17,15	17,27	16,53	16,65	16,17	16,26
34	16,82	17,15	17,16	17,27	16,56	16,64	16,19	16,25
35	16,78	17,14	17,20	17,26	16,58	16,64	16,20	16,25

36	16,75	17,12	17,19	17,25	16,58	16,63	16,20	16,24
37	16,79	17,11	17,18	17,25	16,57	16,62	16,21	16,23
38	16,79	17,09	17,21	17,24	16,55	16,62	16,22	16,23
39	16,81	17,08	17,22	17,24	16,54	16,61	16,23	16,23
40	16,79	17,07	17,24	17,23	16,55	16,61	16,24	16,22
41	16,78	17,06	17,25	17,23	16,57	16,60	16,26	16,22
42	16,77	17,05	17,28	17,23	16,58	16,60	16,27	16,22
43	16,77	17,04	17,27	17,22	16,58	16,60	16,27	16,22
44	16,74	17,03	17,29	17,22	16,58	16,59	16,26	16,22
45	16,75	17,02	17,31	17,22	16,60	16,59	16,27	16,22
46	16,79	17,01	17,29	17,22	16,60	16,59	16,29	16,21
47	16,73	17,01	17,30	17,22	16,58	16,58	16,26	16,21
48	16,74	17,00	17,30	17,22	16,59	16,58	16,25	16,21
49	16,72	16,99	17,31	17,22	16,61	16,58	16,25	16,21
50	16,75	16,98	17,29	17,21	16,60	16,58	16,27	16,21
51	16,74	16,98	17,26	17,21	16,63	16,58	16,28	16,21
52	16,74	16,97	17,23	17,21	16,62	16,58	16,29	16,21
53	16,71	16,96	17,23	17,21	16,61	16,57	16,31	16,21
54	16,71	16,95	17,24	17,21	16,63	16,57	16,31	16,21
55	16,71	16,95	17,23	17,21	16,63	16,57	16,32	16,21
56	16,71	16,94	17,25	17,20	16,63	16,57	16,31	16,21
57	16,72	16,94	17,22	17,20	16,63	16,57	16,31	16,21
58	16,70	16,93	17,23	17,20	16,63	16,57	16,30	16,21
59	16,67	16,92	17,25	17,20	16,65	16,57	16,31	16,21
60	16,69	16,92	17,26	17,20	16,67	16,57	16,32	16,21

61	16,69	16,91	17,26	17,20	16,68	16,57	16,33	16,21
62	16,71	16,91	17,26	17,20	16,69	16,57	16,34	16,21
63	16,74	16,90	17,27	17,20	16,72	16,57	16,34	16,21
64	16,72	16,90	17,29	17,20	16,73	16,57	16,36	16,21
65	16,73	16,90	17,30	17,20	16,76	16,57	16,36	16,21
66	16,73	16,89	17,32	17,20	16,76	16,58	16,37	16,21
67	16,71	16,89	17,34	17,20	16,76	16,58	16,38	16,21
68	16,73	16,88	17,34	17,20	16,77	16,58	16,39	16,22
69	16,75	16,88	17,34	17,20	16,78	16,58	16,40	16,22
70	16,78	16,88	17,34	17,20	16,77	16,58	16,40	16,22
71	16,80	16,88	17,36	17,20	16,76	16,58	16,42	16,22
72	16,78	16,87	17,36	17,20	16,76	16,58	16,40	16,22
73	16,78	16,87	17,36	17,21	16,78	16,59	16,39	16,22
74	16,78	16,87	17,34	17,21	16,80	16,59	16,40	16,22
75	16,77	16,87	17,35	17,21	16,79	16,59	16,41	16,23
76	16,76	16,86	17,33	17,21	16,80	16,59	16,41	16,23
77	16,74	16,86	17,32	17,21	16,80	16,59	16,43	16,23
78	16,73	16,86	17,33	17,21	16,78	16,59	16,44	16,23
79	16,75	16,86	17,30	17,21	16,79	16,60	16,44	16,23
80	16,74	16,85	17,30	17,21	16,81	16,60	16,42	16,23
81	16,73	16,85	17,30	17,21	16,83	16,60	16,41	16,24
82	16,72	16,85	17,29	17,21	16,83	16,60	16,41	16,24
83	16,74	16,85	17,29	17,21	16,82	16,60	16,41	16,24
84	16,74	16,85	17,31	17,21	16,81	16,60	16,40	16,24
85	16,76	16,84	17,32	17,21	16,82	16,61	16,41	16,24

86	16,77	16,84	17,32	17,21	16,81	16,61	16,42	16,24
87	16,76	16,84	17,32	17,21	16,79	16,61	16,41	16,24
88	16,77	16,84	17,33	17,21	16,79	16,61	16,39	16,24
89	16,79	16,84	17,32	17,21	16,77	16,61	16,39	16,25
90	16,79	16,84	17,33	17,21	16,77	16,61	16,38	16,25
91	16,80	16,83	17,34	17,22	16,77	16,61	16,37	16,25
92	16,81	16,83	17,33	17,22	16,77	16,61	16,38	16,25
93	16,82	16,83	17,33	17,22	16,76	16,62	16,38	16,25
94	16,82	16,83	17,35	17,22	16,77	16,62	16,39	16,25
95	16,81	16,83	17,34	17,22	16,77	16,62	16,39	16,25
96	16,81	16,83	17,35	17,22	16,75	16,62	16,40	16,25
97	16,81	16,83	17,35	17,22	16,75	16,62	16,39	16,25
98	16,81	16,83	17,35	17,22	16,74	16,62	16,39	16,25
99	16,80	16,83	17,35	17,22	16,74	16,62	16,39	16,25
100	16,81	16,83	17,37	17,22	16,73	16,62	16,41	16,25

### Lampiran 3

#### RMSE Average

K	RMSE							
	Average							
	80%-20%		70%-30%		60%-40%		50%-50%	
	Euclidean	Manhattan	Euclidean	Manhattan	Euclidean	Manhattan	Euclidean	Manhattan
1	20,90	20,89	21,83	22,00	21,94	21,65	20,81	20,85
2	19,03	19,32	19,49	19,27	20,15	19,62	19,84	19,43

3	18,85	19,18	18,59	17,91	19,06	18,68	18,36	18,29
4	18,16	18,17	18,10	17,79	18,22	17,98	17,40	17,32
5	18,03	17,90	17,87	17,85	17,55	17,34	16,99	16,92
6	17,96	18,03	17,68	17,71	17,13	17,15	16,76	16,67
7	17,98	17,85	17,78	17,57	16,98	16,98	16,76	16,52
8	18,02	17,71	17,68	17,68	17,11	16,94	16,63	16,59
9	17,73	17,77	17,77	17,48	16,95	16,94	16,46	16,50
10	17,62	17,76	17,67	17,34	16,89	16,79	16,43	16,39
11	17,39	17,52	17,56	17,39	16,73	16,66	16,42	16,36
12	17,37	17,50	17,56	17,28	16,65	16,59	16,43	16,35
13	17,26	17,50	17,31	17,30	16,55	16,64	16,39	16,31
14	17,13	17,51	17,33	17,36	16,46	16,49	16,38	16,32
15	17,12	17,24	17,26	17,38	16,59	16,56	16,40	16,25
16	17,16	17,20	17,24	17,39	16,65	16,55	16,35	16,23
17	17,10	17,08	17,25	17,40	16,63	16,52	16,25	16,23
18	17,00	17,02	17,30	17,29	16,65	16,51	16,30	16,20
19	16,97	16,99	17,29	17,29	16,60	16,50	16,25	16,13
20	16,93	16,91	17,31	17,34	16,61	16,45	16,11	16,15
21	16,86	16,92	17,29	17,33	16,60	16,46	16,03	16,13
22	16,83	16,92	17,24	17,29	16,58	16,46	15,99	16,09
23	16,80	16,94	17,27	17,27	16,51	16,50	15,98	16,11
24	16,75	16,96	17,24	17,31	16,51	16,48	16,03	16,09
25	16,68	16,94	17,25	17,27	16,57	16,51	16,05	16,10
26	16,71	16,93	17,25	17,28	16,60	16,56	16,10	16,11
27	16,75	16,85	17,24	17,32	16,58	16,54	16,13	16,15

28	16,79	16,82	17,21	17,29	16,57	16,55	16,17	16,14
29	16,77	16,79	17,22	17,27	16,60	16,53	16,19	16,13
30	16,71	16,81	17,25	17,18	16,63	16,54	16,21	16,10
31	16,72	16,87	17,20	17,14	16,58	16,55	16,22	16,14
32	16,69	16,87	17,18	17,15	16,58	16,55	16,23	16,16
33	16,68	16,82	17,20	17,15	16,60	16,53	16,21	16,17
34	16,70	16,82	17,17	17,16	16,58	16,56	16,21	16,19
35	16,71	16,78	17,19	17,20	16,56	16,58	16,19	16,20
36	16,69	16,75	17,19	17,19	16,57	16,58	16,16	16,20
37	16,72	16,79	17,22	17,18	16,58	16,57	16,17	16,21
38	16,73	16,79	17,17	17,21	16,58	16,55	16,18	16,22
39	16,76	16,81	17,20	17,22	16,61	16,54	16,18	16,23
40	16,71	16,79	17,21	17,24	16,58	16,55	16,20	16,24
41	16,72	16,78	17,21	17,25	16,56	16,57	16,21	16,26
42	16,73	16,77	17,19	17,28	16,57	16,58	16,22	16,27
43	16,70	16,77	17,19	17,27	16,58	16,58	16,24	16,27
44	16,71	16,74	17,17	17,29	16,55	16,58	16,24	16,26
45	16,69	16,75	17,19	17,31	16,55	16,60	16,24	16,27
46	16,67	16,79	17,20	17,29	16,56	16,60	16,22	16,29
47	16,69	16,73	17,22	17,30	16,56	16,58	16,24	16,26
48	16,70	16,74	17,23	17,30	16,57	16,59	16,26	16,25
49	16,70	16,72	17,24	17,31	16,59	16,61	16,23	16,25
50	16,68	16,75	17,24	17,29	16,60	16,60	16,22	16,27
51	16,67	16,74	17,25	17,26	16,61	16,63	16,23	16,28
52	16,71	16,74	17,28	17,23	16,61	16,62	16,23	16,29

53	16,75	16,71	17,28	17,23	16,62	16,61	16,23	16,31
54	16,74	16,71	17,25	17,24	16,62	16,63	16,24	16,31
55	16,77	16,71	17,27	17,23	16,64	16,63	16,25	16,32
56	16,79	16,71	17,26	17,25	16,65	16,63	16,27	16,31
57	16,77	16,72	17,26	17,22	16,67	16,63	16,30	16,31
58	16,75	16,70	17,27	17,23	16,66	16,63	16,31	16,30
59	16,72	16,67	17,25	17,25	16,69	16,65	16,31	16,31
60	16,74	16,69	17,23	17,26	16,69	16,67	16,33	16,32
61	16,72	16,69	17,24	17,26	16,69	16,68	16,32	16,33
62	16,71	16,71	17,25	17,26	16,71	16,69	16,32	16,34
63	16,72	16,74	17,26	17,27	16,70	16,72	16,34	16,34
64	16,73	16,72	17,27	17,29	16,71	16,73	16,35	16,36
65	16,72	16,73	17,26	17,30	16,71	16,76	16,35	16,36
66	16,71	16,73	17,29	17,32	16,71	16,76	16,37	16,37
67	16,71	16,71	17,28	17,34	16,72	16,76	16,37	16,38
68	16,73	16,73	17,31	17,34	16,72	16,77	16,37	16,39
69	16,73	16,75	17,30	17,34	16,71	16,78	16,37	16,40
70	16,71	16,78	17,30	17,34	16,72	16,77	16,38	16,40
71	16,72	16,80	17,27	17,36	16,73	16,76	16,39	16,42
72	16,72	16,78	17,29	17,36	16,72	16,76	16,38	16,40
73	16,69	16,78	17,30	17,36	16,70	16,78	16,38	16,39
74	16,70	16,78	17,29	17,34	16,71	16,80	16,37	16,40
75	16,70	16,77	17,27	17,35	16,72	16,79	16,36	16,41
76	16,69	16,76	17,28	17,33	16,74	16,80	16,35	16,41
77	16,70	16,74	17,28	17,32	16,74	16,80	16,37	16,43

78	16,69	16,73	17,30	17,33	16,74	16,78	16,35	16,44
79	16,70	16,75	17,31	17,30	16,74	16,79	16,36	16,44
80	16,72	16,74	17,31	17,30	16,76	16,81	16,36	16,42
81	16,73	16,73	17,28	17,30	16,76	16,83	16,37	16,41
82	16,75	16,72	17,27	17,29	16,76	16,83	16,37	16,41
83	16,77	16,74	17,27	17,29	16,75	16,82	16,37	16,41
84	16,76	16,74	17,28	17,31	16,76	16,81	16,37	16,40
85	16,78	16,76	17,26	17,32	16,77	16,82	16,39	16,41
86	16,79	16,77	17,27	17,32	16,76	16,81	16,39	16,42
87	16,78	16,76	17,27	17,32	16,76	16,79	16,40	16,41
88	16,78	16,77	17,28	17,33	16,77	16,79	16,40	16,39
89	16,79	16,79	17,28	17,32	16,78	16,77	16,39	16,39
90	16,76	16,79	17,31	17,33	16,78	16,77	16,40	16,38
91	16,76	16,80	17,32	17,34	16,78	16,77	16,40	16,37
92	16,77	16,81	17,32	17,33	16,79	16,77	16,40	16,38
93	16,76	16,82	17,31	17,33	16,79	16,76	16,40	16,38
94	16,75	16,82	17,33	17,35	16,80	16,77	16,40	16,39
95	16,76	16,81	17,34	17,34	16,79	16,77	16,41	16,39
96	16,76	16,81	17,33	17,35	16,78	16,75	16,40	16,40
97	16,76	16,81	17,34	17,35	16,77	16,75	16,41	16,39
98	16,77	16,81	17,36	17,35	16,77	16,74	16,41	16,39
99	16,75	16,80	17,37	17,35	16,77	16,74	16,42	16,39
100	16,76	16,81	17,36	17,37	16,77	16,73	16,43	16,41

**Lampiran 4**

**RMSE Weighted**

K	Wighted							
	80%-20%		70%-30%		60%-40%		50%-50%	
	Euclidean	Manhattan	Euclidean	Manhattan	Euclidean	Manhattan	Euclidean	Manhattan
1	20,90	20,89	21,83	22,00	21,94	21,65	20,81	20,85
2	19,42	19,60	20,24	20,22	20,58	20,23	19,82	19,64
3	18,93	19,14	19,39	19,16	19,78	19,41	19,09	18,96
4	18,60	18,75	18,88	18,63	19,25	18,90	18,53	18,42
5	18,38	18,46	18,56	18,34	18,80	18,48	18,12	18,01
6	18,23	18,31	18,33	18,15	18,45	18,18	17,82	17,70
7	18,15	18,19	18,19	18,01	18,17	17,95	17,60	17,46
8	18,10	18,10	18,07	17,92	17,98	17,77	17,41	17,29
9	18,02	18,03	17,99	17,83	17,82	17,63	17,26	17,15
10	17,95	17,96	17,92	17,75	17,68	17,51	17,13	17,03
11	17,87	17,89	17,86	17,69	17,55	17,39	17,02	16,93
12	17,80	17,83	17,80	17,63	17,45	17,30	16,94	16,85
13	17,74	17,78	17,74	17,58	17,35	17,22	16,87	16,78
14	17,68	17,74	17,69	17,54	17,26	17,14	16,81	16,72
15	17,62	17,69	17,65	17,51	17,19	17,08	16,76	16,67
16	17,58	17,65	17,61	17,49	17,14	17,03	16,71	16,62
17	17,53	17,60	17,57	17,47	17,09	16,98	16,67	16,58

18	17,49	17,55	17,54	17,45	17,05	16,94	16,63	16,54
19	17,45	17,51	17,52	17,43	17,01	16,90	16,60	16,51
20	17,42	17,47	17,50	17,42	16,98	16,87	16,56	16,48
21	17,38	17,44	17,48	17,40	16,94	16,84	16,52	16,45
22	17,34	17,41	17,46	17,39	16,92	16,81	16,49	16,42
23	17,31	17,38	17,44	17,38	16,89	16,79	16,45	16,40
24	17,28	17,35	17,43	17,36	16,86	16,77	16,43	16,38
25	17,25	17,33	17,41	17,35	16,84	16,75	16,40	16,36
26	17,22	17,30	17,40	17,34	16,83	16,73	16,38	16,34
27	17,20	17,28	17,39	17,33	16,81	16,72	16,37	16,33
28	17,18	17,26	17,37	17,33	16,79	16,70	16,35	16,31
29	17,16	17,23	17,36	17,32	16,78	16,69	16,34	16,30
30	17,13	17,21	17,35	17,31	16,77	16,68	16,33	16,29
31	17,12	17,20	17,34	17,29	16,76	16,67	16,32	16,28
32	17,10	17,18	17,33	17,28	16,75	16,66	16,31	16,27
33	17,08	17,17	17,32	17,27	16,74	16,65	16,30	16,26
34	17,06	17,15	17,31	17,27	16,73	16,64	16,29	16,25
35	17,05	17,14	17,30	17,26	16,72	16,64	16,29	16,25
36	17,03	17,12	17,29	17,25	16,71	16,63	16,28	16,24
37	17,02	17,11	17,29	17,25	16,70	16,62	16,27	16,23
38	17,01	17,09	17,28	17,24	16,70	16,62	16,27	16,23
39	17,00	17,08	17,27	17,24	16,69	16,61	16,26	16,23
40	16,99	17,07	17,27	17,23	16,68	16,61	16,25	16,22
41	16,98	17,06	17,26	17,23	16,68	16,60	16,25	16,22
42	16,97	17,05	17,26	17,23	16,67	16,60	16,25	16,22

43	16,96	17,04	17,25	17,22	16,66	16,60	16,24	16,22
44	16,95	17,03	17,25	17,22	16,66	16,59	16,24	16,22
45	16,94	17,02	17,24	17,22	16,65	16,59	16,24	16,22
46	16,93	17,01	17,24	17,22	16,65	16,59	16,24	16,21
47	16,92	17,01	17,24	17,22	16,64	16,58	16,23	16,21
48	16,92	17,00	17,24	17,22	16,64	16,58	16,23	16,21
49	16,91	16,99	17,23	17,22	16,64	16,58	16,23	16,21
50	16,90	16,98	17,23	17,21	16,63	16,58	16,23	16,21
51	16,90	16,98	17,23	17,21	16,63	16,58	16,22	16,21
52	16,89	16,97	17,23	17,21	16,63	16,58	16,22	16,21
53	16,89	16,96	17,23	17,21	16,63	16,57	16,22	16,21
54	16,88	16,95	17,22	17,21	16,62	16,57	16,22	16,21
55	16,88	16,95	17,22	17,21	16,62	16,57	16,22	16,21
56	16,87	16,94	17,22	17,20	16,62	16,57	16,22	16,21
57	16,87	16,94	17,22	17,20	16,62	16,57	16,22	16,21
58	16,87	16,93	17,22	17,20	16,62	16,57	16,22	16,21
59	16,86	16,92	17,22	17,20	16,62	16,57	16,22	16,21
60	16,86	16,92	17,22	17,20	16,62	16,57	16,22	16,21
61	16,86	16,91	17,22	17,20	16,62	16,57	16,22	16,21
62	16,85	16,91	17,21	17,20	16,62	16,57	16,22	16,21
63	16,85	16,90	17,21	17,20	16,62	16,57	16,22	16,21
64	16,84	16,90	17,21	17,20	16,62	16,57	16,22	16,21
65	16,84	16,90	17,21	17,20	16,62	16,57	16,22	16,21
66	16,84	16,89	17,21	17,20	16,62	16,58	16,22	16,21
67	16,83	16,89	17,21	17,20	16,62	16,58	16,22	16,21

68	16,83	16,88	17,21	17,20	16,62	16,58	16,22	16,22
69	16,83	16,88	17,21	17,20	16,62	16,58	16,22	16,22
70	16,83	16,88	17,21	17,20	16,62	16,58	16,22	16,22
71	16,82	16,88	17,21	17,20	16,62	16,58	16,22	16,22
72	16,82	16,87	17,21	17,20	16,62	16,58	16,23	16,22
73	16,82	16,87	17,21	17,21	16,62	16,59	16,23	16,22
74	16,82	16,87	17,21	17,21	16,62	16,59	16,23	16,22
75	16,81	16,87	17,21	17,21	16,62	16,59	16,23	16,23
76	16,81	16,86	17,21	17,21	16,62	16,59	16,23	16,23
77	16,81	16,86	17,21	17,21	16,62	16,59	16,23	16,23
78	16,81	16,86	17,21	17,21	16,62	16,59	16,23	16,23
79	16,80	16,86	17,21	17,21	16,62	16,60	16,23	16,23
80	16,80	16,85	17,21	17,21	16,62	16,60	16,23	16,23
81	16,80	16,85	17,21	17,21	16,62	16,60	16,23	16,24
82	16,80	16,85	17,21	17,21	16,62	16,60	16,23	16,24
83	16,80	16,85	17,21	17,21	16,62	16,60	16,24	16,24
84	16,80	16,85	17,21	17,21	16,62	16,60	16,24	16,24
85	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,24
86	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,24
87	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,24
88	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,24
89	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,25
90	16,79	16,84	17,21	17,21	16,63	16,61	16,24	16,25
91	16,79	16,83	17,21	17,22	16,63	16,61	16,24	16,25
92	16,79	16,83	17,21	17,22	16,63	16,61	16,24	16,25

93	16,79	16,83	17,21	17,22	16,63	16,62	16,25	16,25
94	16,79	16,83	17,21	17,22	16,63	16,62	16,25	16,25
95	16,78	16,83	17,21	17,22	16,63	16,62	16,25	16,25
96	16,78	16,83	17,22	17,22	16,64	16,62	16,25	16,25
97	16,78	16,83	17,22	17,22	16,64	16,62	16,25	16,25
98	16,78	16,83	17,22	17,22	16,64	16,62	16,25	16,25
99	16,78	16,83	17,22	17,22	16,64	16,62	16,25	16,25
100	16,78	16,83	17,22	17,22	16,64	16,62	16,25	16,25