

## Penerapan Algoritma MKNN pada Data Historis Gempa Bumi yang Berpotensi Tsunami

Muh. Rijal Maulana Hamsar<sup>1</sup>, Firman Tempola<sup>2,\*</sup>, Saiful Do Abdullah<sup>3</sup>, Alfanugrah A Hi. Usman<sup>2</sup>,  
Muhammad Sabri Ahmad<sup>2</sup>

<sup>1,2,3,4,5</sup>Teknik Informatika, Fakultas Teknik, Universitas Khairun

<sup>1,2</sup>Jl. Jati Metro Kampus 3 Universitas Khairun, Ternate 97714

\*Corresponding Author: firman.tempola@unkhair.ac.id

**Abstract**— In implementing the Modified K-Nearest Neighbor method for determining tsunami potential by comparing the calculation of the Euclidean and Manhattan distances, 3 earthquake criteria are used, namely strength, depth and epicenter with 2 classification classes, tsunami potential and no tsunami potential. This algorithm classifies data based on similarity to other data. The results obtained are that the algorithm can classify the status of an earthquake whether it has the potential for a tsunami or not, taking into account the balance of the composition of the training data used. With the highest accuracy value of 90% for  $K = 1$ . The results of the comparison of euclidean and manhattan were obtained from several test scenarios, namely changes in dataset composition, changes in  $K$  values and changes in training data. The average system performance obtained is 80% for the Euclidean distance and 82% for the Manhattan distance. The test results show that the composition of the dataset greatly influences the performance of the system obtained. Then the use of the Manhattan distance has a higher level of accuracy than the Euclidean distance with an average difference of 2%.

**Keywords**— Modified K-Nearest Neighbor, Earthquake, Tsunami, data mining, clasification

**Intisari**— Dalam implementasi metode Modified K-Nearest Neighbor untuk penentuan potensi tsunami dengan membandingkan perhitungan jarak euclidean dan manhattan digunakan 3 kriteria gempa bumi yaitu kekuatan, kedalaman dan pusat gempa dengan 2 kelas klasifikasi, berpotensi tsunami dan tidak berpotensi tsunami. Algoritma ini mengklasifikasikan data berdasarkan similarity terhadap data lainnya. Hasil yang diperoleh yaitu algoritma dapat melakukan klasifikasi status dari sebuah gempa apakah berpotensi tsunami atau tidak berpotensi dengan memperhatikan keseimbangan komposisi data latih yang digunakan. Dengan Nilai akurasi tertinggi yaitu 90% untuk  $K=1$ . Hasil perbandingan euclidean dan manhattan didapati dari beberapa skenario pengujian yaitu perubahan komposisi dataset, perubahan nilai  $K$  dan perubahan data latih. Diperoleh rata-rata kinerja sistem sebesar 80% untuk jarak euclidean dan 82% untuk jarak manhattan hasil pengujian menunjukkan bahwa komposisi dataset sangat berpengaruh terhadap kinerja sistem yang didapat. Maka Penggunaan jarak manhattan memiliki tingkat akurasi yang lebih tinggi dibandingkan jarak euclidean dengan selisih rata-rata 2%.

**Kata kunci**— Modified K-Nearest Neighbor, Gempa bumi, Tsunami, data mining, klasifikasi

### I. PENDAHULUAN

Hampir setiap tahun Indonesia selalu terjadi gempa. Hal ini karena Indonesia secara geologi berada pada pertemuan yaitu lempeng

indoaustralia, lempeng Eurasia dan lempeng pasifik. Pertemuan tiga lempeng besar aktif tersebut mengakibatkan memiliki kontur tanah dan relief di Indonesia beragam dari dataran

rendah ke dataran tinggi, suhu rendah ke suhu tinggi dan dari tanah tekstur kapur hingga tanah gambut sehingga bencana alam kerap kali menerpa bumi pertiwi.

Gempa bumi merupakan proses terjadi pergeseran lapisan tanah atau batuan pada permukaan bumi secara spontan akibat pergeseran lempeng-lempeng tektonik. Gempa bumi yang disebabkan oleh aktivitas pergerakan lempeng tektonik disebut gempa bumi tektonik [1]. Dampak dari gempa bumi tektonik bisa mengakibatkan terjadinya tsunami jika pusat gempa berada di laut.

Di Indonesia gempa hingga mengakibatkan tsunami kerap kali terjadi, sebagaimana yang baru terjadi pada tahun 2018 di Sulawesi tengah. [2] Menyebutkan 90% gempa bumi bawah lautlah yang menyebabkan terjadinya tsunami.

Gempa yang bumi yang terjadi di seluruh belahan bumi, tidak seorang pun yang dapat memprediksi kapan dan dimana akan terjadinya gempa bumi, namun untuk memprediksi potensi terjadinya tsunami dapat di lakukan dengan teknologi, dengan beberapa kriteria setelah terjadinya gempa bumi seperti pusat gempa berada dilaut magnitudo gempa minimal 7 dan kedalaman gempa di bawah 100 km atau gempa dangkal.

Penelitian sebelumnya pernah mengkaji penentuan potensi tsunami berdasarkan gempa bawah laut dengan mengimplementasikan metode Modified K-Nearest Neighbor (MKNN), dari penelitian ini di dapati nilai rata-rata akurasi minimum sebesar 69,63% pada 300 *dataset* dengan menggunakan metode perhitungan jarak *euclidean distance*. Tujuan dari modifikasi penelitian tersebut adalah untuk mengatasi kelemahan jarak pada metode KNN. Salah satu tahap pencocokan kesamaan ataupun kemiripan ciri-ciri suatu dalam klasifikasi dengan implementasi algoritma Modified KNearest Neighbor (MKNN) adalah dengan menggunakan pendekatan jarak [3].

Pada penelitian [4] menerapkan metode *naïve bayes classifier* untuk prediksi gempa bumi yang berpotensi tsunami di Indonesia. Dalam penelitian ini menunjukkan bahwa

algoritma NBC mampu mengetahui dampak dari gempa bumi, hanya saja dalam penelitian tidak pengujian kinerja dari algoritma naïve bayes.

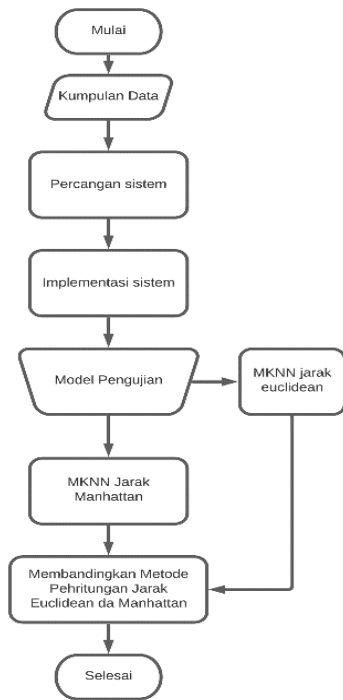
Dengan perkembangan teknologi yang sangat pesat saat ini memungkinkan dapat dimanfaatkan untuk menyelesaikan berbagai permasalahan, data yang terdapat pada historis gempa bumi memiliki informasi yang dapat di analisis dan diolah sehingga bisa menentukan potensi tsunami atau tidak setelah gempa bumi. Analisis dan pengolahan data histori gempa dapat di lakukan dengan proses data mining.

Pada implementasinya algoritma akan mengklasifikasi status dari gempa bumi apakah berpotensi tsunami atau tidak dengan menggunakan 3 parameter gempa bumi yaitu kekuatan gempa, sumber gempa dan kedalaman gempa. Dengan di terapkan perbandingan jarak euclidean dan manhattan pada metode MKNN untuk melihat akurasi tertinggi.

Penelitian dengan membandingkan perhitungan jarak juga pernah dilakukan oleh Rosa dengan penelitian Perbandingan Metode Pengukuran Jarak Pada Algoritma K-Nearest Neighbor Untuk Klasifikasi dengan 4 perhitungan jarak yaitu *euclidean*, *manhattan*, *Chebyshev* dan *cosine distance* hasil akurasi tertinggi adalah dengan menggunakan jarak *cosine distance* sebesar 92,559% [5].

## II. METODOLOGI PENELITIAN

Pada penelitian ini menggunakan metode MKNN. MKNN sendiri merupakan pengembangan dari metode KNN dengan menambahkan beberapa perhitungan seperti validitas dan menghitung *weight voting* untuk menentukan hasil akhir dari klasifikasi. Dilakukan juga perbandingan perhitungan jarak *euclidean* dan *manhattan* untuk melihat hasil akurasi tertinggi dari kedua jarak tersebut. Tahapan metode penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

#### A. Kumpulan Data

Tahap awal adalah mengumpulkan data yang akan dibutuhkan dalam penelitian ini. Data yang digunakan adalah kumpulan data gempa bumi Maluku utara pada tahun 2010 – 2019 dan juga digunakan data gempa bumi tsunami global untuk dilakukan pengujian.

#### B. Perancangan Sistem

Setelah data-data dikumpulkan selanjutnya dilakukan perancangan sistem. Sistem ini dibangun dengan menggunakan Bahasa pemrograman PHP dan MYSQL. Adapun metode pengembangan sistem yang digunakan yaitu metode *Waterfall*. Metode *waterfall* memiliki 4 tahapan [6].

#### C. Implementasi Sistem

Selanjutnya dilakukan implementasi sesuai dengan perancangan yang telah dilakukan sebelumnya untuk mengimplmentasikan algoritma pada bahasa pemrograman PHP.

#### D. Model Pengujian

Untuk mengukur dan melihat implementasi algoritma pada studi kasus penentuan potensi tsunami akan dilakukan model pengujian dengan menggunakan metode

MKNN jarak *euclidean* dan *manhattan*. Penelitian ini juga akan diuji dengan metode *Black Box Testing* [7].

#### E. Perbandingan Metode Jarak

Setelah dilakukan pengujian dengan metode MKNN menggunakan jarak *euclidean* dan *manhattan*, kemudian akan dilihat perbandingan penggunaan jarak mana yang menghasilkan nilai akurasi tertinggi. Untuk lebih tahapan dalam penelitian ini dapat lihat pada Gambar 1.

Data mining adalah adalah suatu proses pencarian pengetahuan atau pola dengan mengsktrak data pada jumlah yang besar. [8]. Metode MKNN salah satu algortima yang dikelompokkan pada *supervised learning*. *Supervised learning* yaitu proses pembelajaran berdasarkan pedoman atau guru. Sering juga diartikan menemukan pola dari data dengan menggunakan beberapa kriteria dari data-data sebelumnya. Salah satu teknik yang terdapat dalam *supervised learning* adalah klasifikasi [9].

*Modified K-Nearest Neighbor* (MKNN) merupakan pengembangan dari metode *K-Nearest Neighbor*. Pada proses MKNN terdapat penambahan dua proses perhitungan, yaitu perhitungan validitas dan proses *weight voting*. Adapun langkah-langkah MKNN adalah sebagai berikut [10]:

1) *Menentukan nilai K* : Pertama kita perlu menentuka nilai K terlebih dahulu.

2) *Mencari jarak*: Selanjutnya mencari jarak antar data *training* menggunakan dua perhitungan yaitu *euclidean* dan *manhattan*.

$$d_{(x,y)} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (1)$$

Keterangan:

$x_i$  = Sampel data latih.

$y_i$  = Data uji.

$d_{(x,y)}$  = Jarak antara titik pada data latih x dan titik data latih y.

P = Dimensi data.

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

Keterangan:

d = jarak antara y dan x

x = kasus baru  
 y = kasus lama  
 i = setiap data  
 n = banyak data

3) *Menghitung validitas*: Langkah ketiga menghitung validitas data *training* dengan menggunakan persamaan 3.

$$Validitas_{(i)} = \frac{1}{k} \sum_{i=1}^k S(lbl(x), lbl Ni(x)) \quad (3)$$

Keterangan

k = Jumlah titik terdekat.

Lbl (x) = kelas x

Lbl Ni (x) = label kelas titik terdekat x

Dimana S digunakan menghitung similaritas antara titik a dan data ke b pada *nearest neighbour* dengan menggunakan persamaan 4 dibawah ini dimana a merupakan label a pada data latih dan b merupakan label selain a pada data latih.

$$S = a, b = \begin{cases} 1 & a=b \\ 0 & a \neq b \end{cases} \quad (4)$$

4) *Weight voting (pembobotan)*:

Selanjutnya divalidasi setiap data latih yang akan dikalikan dengan *weight voting* sesuai jarak pada setiap tetangganya. Rumus *weight voting* seperti Persamaan 5.

$$W_{(i)} = Validitas_{(x)} x \frac{1}{de+a} \quad (5)$$

Berikut ini adalah contoh penerapan MKNN pada status gempa bumi untuk penentuan potensi tsunami dengan menggunakan dua perhitungan jarak *euclidean* dan *manhattan*. Disimulasikan dengan menggunakan 6 data latih dan 1 data uji seperti pada tabel 1. Status untuk BS = Berpotensi tsunami dan TBS = Tidak berpotensi tsunami.

### III. HASIL DAN PEMBAHASAN

#### A. Dataset

Dataset yang digunakan pada penelitian ini adalah menggunakan *dataset* Maluku Utara dan *dataset* global untuk dilakukan uji coba algoritma untuk melihat pengaruh dari setiap *dataset* yang digunakan. Untuk *dataset* nya dapat lihat pada tabel 1 dan tabel 2.

**Tabel 1.** Dataset Gempa Maluku Utara

No	Tgl	Loc	Pusat	Depth	Mag	Status
1	16/11/2019	Northern Molucca Sea	Laut	5,7	42	TBS
2	14/11/2019	Northern Molucca Sea,	Laut	5,9	10	TBS
3	14/11/2019	Laut maluku	Laut	7,4	73	BS
4	11/11/2019	Halbar Halut	Laut	5,7	119	TBS
5	25/10/2019	Jailolo	Laut	4,7	28	TBS
....	....	....	....	....	....	....
320.	11/06/2013	Laut maluku	Laut	5,8	50	TBS

**Tabel 2.** Dataset Gempa Global

No	Tgl	Loc	Pusat	Depth	Mag	Status
1	22/07/2020	USA	Laut	7,8	28	BS
2	23/06/2020	Mexico	Laut	7,4	20	BS
3	18/06/2020	South of the Kermadec Island	Laut	7,4	37	BS
4	25/03/2020	Rusia	Laut	7,5	57	BS
5	28/01/2020	Cuba Region	Laut	7,7	23	BS
....	....	....	....	....	....	....
320.	18/11/1929	Kanada	Laut	7,2	10	BS

Penelitian dilakukan dengan membandingkan metode perhitungan jarak pada Algoritma MKNN. Sistem yang dibangun ini berbasis web Adapun hasil implementasinya ditunjukkan pada Gambar 2.



**Gambar 2.** Implementasi Algoritma Pada PHP

#### B. Implementasi Algoritma MKNN

Berikut penjelasan dari implementasi penggunaan Algoritma MKNN diantaranya ialah:

1) Implementasi MKNN Pada Penentuan Potensi Tsunami Maluku Utara: *Implementasi MKNN* Dari keseluruhan *dataset* Maluku Utara

pada tabel 1 akan di bagi menjadi 310 data latih dan 3 data uji dengan komposisi *dataset* BS sebanyak 6 data dan TBS sebanyak 304 data. Data uji yang digunakan dapat dilihat pada tabel 3.

**Tabel 3.** Data Uji Potensi Tsunami Maluku Utara

No	Tgl	Loc	Pusat	Depth	Mag	Status
1	22/07/2020	Laut	7,1	73	Laut	?
2	23/06/2020	Laut	7,1	49	Laut	?
3	18/06/2020	Laut	7,3	10	Laut	?

Hasil pengujian menunjukkan bahwa ketika terjadi sebuah gempa yang memenuhi kriteria berpotensi tsunami, saat diterapkan pada pengujian sistem algoritma MKNN hasil yang dapat diklasifikasi oleh sistem adalah tidak berpotensi tsunami.

6	05/04/2021	Laut Maluku	Laut	7,3	10	TIDAK BERPOTENSI TSUNAMI	Salah	1	
7	05/04/2021	Ternate	Laut	7,1	49	TIDAK BERPOTENSI TSUNAMI	Salah	1	
8	05/04/2021	Laut Maluku	Laut	7,1	73	TIDAK BERPOTENSI TSUNAMI	Salah	1	

**Gambar 3.** Hasil Pengujian Potensi Tsunami Maluku Utara

Hal ini berarti hasil klasifikasi yang dihasilkan bernilai salah. Hal tersebut dikarenakan pengaruh data latih yang tidak seimbang yang hanya dominan pada satu kelas saja sehingga mempengaruhi nilai kedekatan jarak antara data latih dan data uji.

2) Skenario Pengujian Algoritma : Adapun beberapa skenario pengujian yang dilakukan pada proses pengujian ini adalah sebagai berikut:

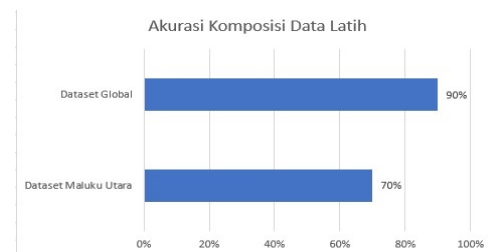
a. Analisis Perubahan Komposisi Data Latih Terhadap Akurasi

Pada pengujian ini digunakan dua pengujian *dataset* yaitu pengujian *dataset* pertama gempa Maluku utara dan *dataset* kedua adalah data gempa Global. Untuk *dataset* Maluku dapat dilihat pada tabel 1 dan *dataset* global dilihat pada tabel 2. *Dataset* tersebut dibagi dua menjadi 310 data latih dan 10 data uji yang akan digunakan pada semua skenario pengujian. Lihat tabel 4 untuk data uji.

**Tabel 4.** Data Uji Pengujian Algoritma

No	Tgl	Loc	Pusat	Depth	Mag	Status
1	11/06/2013	Laut Maluku	Laut	5,8	50	?
2	28/04/2017	Ternate	Laut	6,8	48	?
3	14/11/2019	Laut Maluku	Laut	7,1	73	?
4	07/07/2019	Ternate Malut	Laut	7,1	49	?
5	15/11/2014	Laut Maluku	Laut	7,3	10	?
6	11/03/2010	Northern Molucca Sea	Laut	3,6	10	?
7	10/03/2010	Southern Molucca Sea	Laut	4	10	?
8	06/01/2011	Northern Molucca Sea	Laut	4,7	20	?
9	17/2/2016	Ternate	Laut	3,8	10	?
10	17/03/2015	Ternate	Laut	6,2	10	?

Pada pengujian *dataset* pertama didapati nilai akurasi yang sama pada jarak *euclidean* dan *manhattan* sebesar 70% dengan menggunakan  $K=1$ . Komposisi *dataset* pertama adalah TBS sebanyak 311 dan BS sebanyak 9 data. Dalam proses klasifikasi algoritma cenderung menghasilkan klasifikasi yang Tidak Berpotensi Tsunami, saat dilakukan pengujian dengan menggunakan data uji yang Berpotensi Tsunami algoritma salah dalam melakukan klasifikasi. Hal ini dikarenakan komposisi *dataset* yang dominan hanya pada satu label saja.

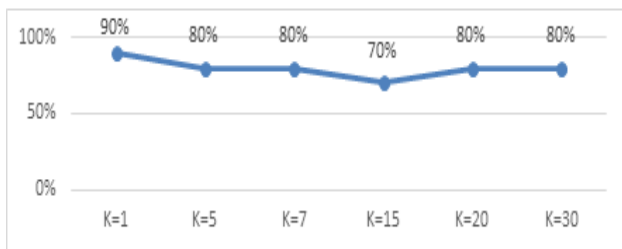


**Gambar 4.** Akurasi Pengujian Komposisi Data Latih

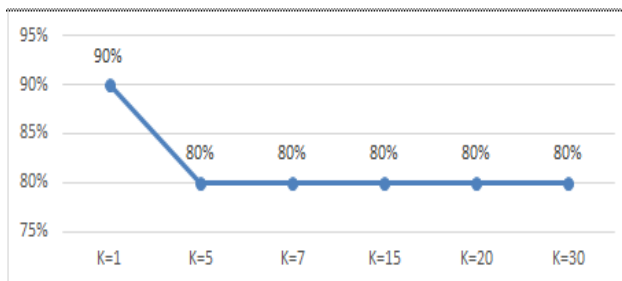
Namun pada pengujian dengan menggunakan *dataset* global, hasil akurasi meningkat menjadi 90% dengan jarak *euclidean* dan *manhattan*  $K=1$ . Komposisi *dataset* global adalah TBS sebanyak 195 dan BS sebanyak 125. Yang berarti komposisi data latih yang digunakan sebagai pengenalan pola klasifikasi pada algoritma MKNN perlu diperhatikan. Karena algoritma akan menghasilkan klasifikasi yang salah jika komposisi data latih dari pengujian Algoritma tidak seimbang.

b. Analisis Skenario Pengujian Pengaruh Nilai K Terhadap Akurasi

Digunakan beberapa pengujian nilai K yaitu K=1, K=5, K=7 K=15 K=20 dan K=30. *Dataset* yang digunakan menggunakan *dataset* yang memiliki nilai terbaik pada pengujian sebelumnya yaitu *dataset* global, dengan data uji yang sama pada pengujian sebelumnya pada tabel 9 hasil pengujian pengaruh nilai K dapat lihat pada gambar 5 dan gambar 6.



**Gambar 5.** Pengujian Pengaruh Nilai K Jarak *Euclidean*

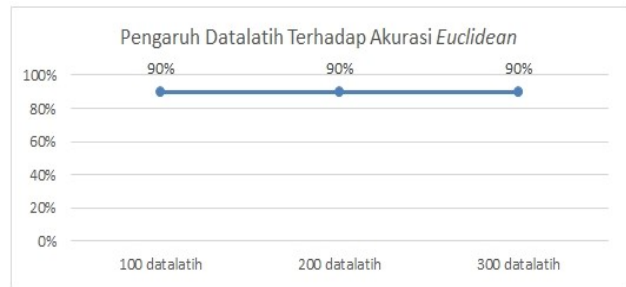


**Gambar 6.** Pegujian Pengaruh Nilai K Jarak *Manhattan*

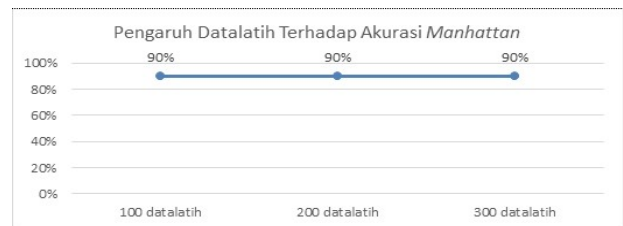
Hasil uji inialisasi nilai K yang bervariasi menghasilkan kinerja algoritma yang berbeda-beda. Sehingga dapat disimpulkan bahwa penentuan nilai K dapat mempengaruhi kinerja dari konsep kedekatan Manhattan distance pada klasifikasi status gempa yang berpotensi tsunami.

**c. Skenario Pengujian Pengaruh Perubahan Data Latih Terhadap Akurasi**

Ada beberapa skenario pengujian yaitu menggunakan 100 data latih, 200 data latih dan 300 data latih yang dipilih secara acak dengan menggunakan 10 data *testing* yang sama pada pengujian sebelumnya. Di gunakan nilai K terbaik pada pengujian sebelumnya yaitu K=1. Hasil pengujian dapat di lihat pada gambar 7 dan gambar 8.



**Gambar 7.** Pengaruh Data Latih Terhadap Akurasi Jarak *Euclidean*

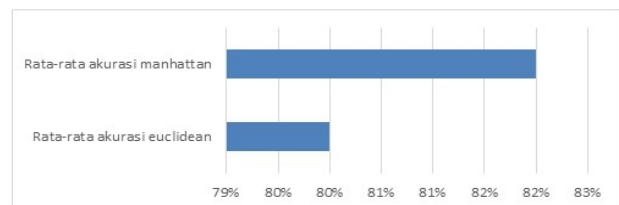


**Gambar 8.** Pengaruh Data Latih Terhadap Akurasi Jarak *Manhattan*

Nilai akurasi yang di dapat cenderung sama yaitu sebesar 90%. Hal ini dikarenakan atribut yang tidak banyak dan juga hanya mempunyai dua kelas klasifikasi sehingga algoritma dapat dengan mudah mengenali pola yang ada pada data latih dengan menghitung nilai kedekatannya pada setiap atribut.

**d. Skenario Pengujian Perbandingan Jarak Euclidean dan Manhattan**

Perbandingan akurasi didapatkan dari hasil beberapa skenario pengujian sebelumnya. Dari hasil pengujian menunjukkan bahwa jarak *manhattan* memiliki nilai rata-rata akurasi yang lebih baik dari pada jarak *euclidean*. Yaitu 80% untuk *euclidean* dan 82% untuk *manhattan*. Agar lebih jelas lihat pada gambar 9.



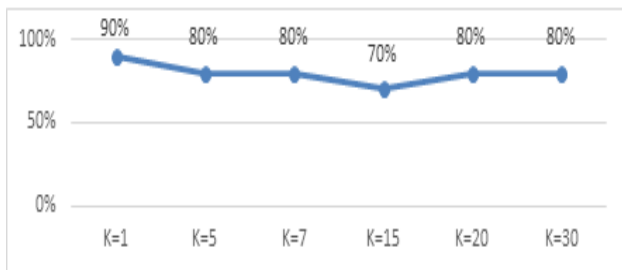
**Gambar 9.** Rata-rata Akurasi Jarak *Euclidean* dan *Manhattan*

Namun pada pengujian dengan menggunakan dataset global, hasil akurasi meningkat menjadi 90% dengan jarak euclidean dan manhattan K=1. Komposisi dataset global adalah TBS sebanyak 195 dan BS sebanyak 125.

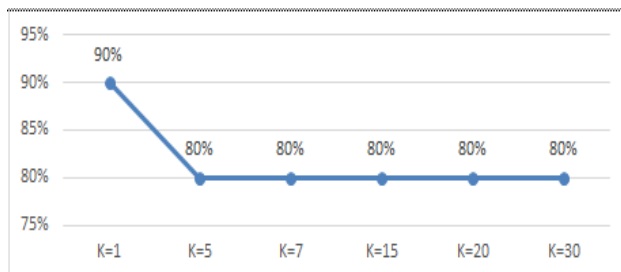
Yang berarti komposisi data latih yang digunakan sebagai pengenalan pola klasifikasi pada algoritma MKNN perlu diperhatikan. Karena algoritma akan menghasilkan klasifikasi yang salah jika komposisi data latih dari pengujian Algoritma tidak seimbang.

e. Analisis Skenario Pengujian Pengaruh Nilai K Terhadap Akurasi

Digunakan beberapa pengujian nilai K yaitu K=1, K=5, K=7 K=15 K=20 dan K=30. *Dataset* yang digunakan menggunakan *dataset* yang memiliki nilai terbaik pada pengujian sebelumnya yaitu *dataset* global, dengan data uji yang sama pada pengujian sebelumnya pada tabel 9 hasil pengujian pengaruh nilai K dapat lihat pada gambar 10 dan gambar 11.



Gambar 10. Pengujian Pengaruh Nilai K Jarak *Euclidean*



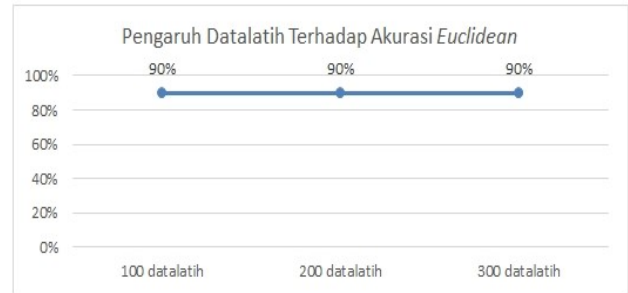
Gambar 11. Pegujian Pengaruh Nilai K Jarak *Manhattan*

Hasil pengujian dengan memvariasikan nilai K mendapatkan nilai akurasi tingkat akurasi yang berbeda-beda. Dengan demikian variasi nilai K pada *Euclidean distance* dapat mempengaruhi kinerja dari algoritma MKNN dalam klasifikasi status gempa bumi yang berpotensi tsunami.

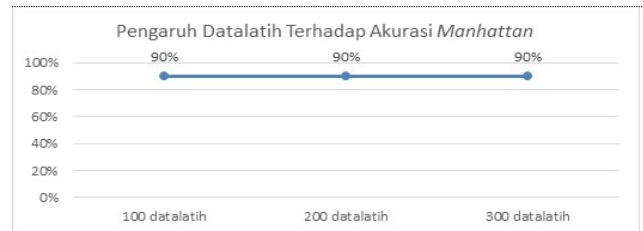
f. Skenario Pengujian Pengaruh Perubahan Data Latih Terhadap Akurasi

Ada beberapa skenario pengujian yaitu menggunakan 100 data latih, 200 data latih dan

300 data latih yang dipilih secara acak dengan menggunakan 10 data *testing* yang sama pada pengujian sebelumnya. Di gunakan nilai K terbaik pada pengujian sebelumnya yaitu K=1. Hasil pengujian dapat di lihat pada gambar 12 dan gambar 13.



Gambar 12. Pengaruh Data Latih Terhadap Akurasi Jarak *Euclidean*

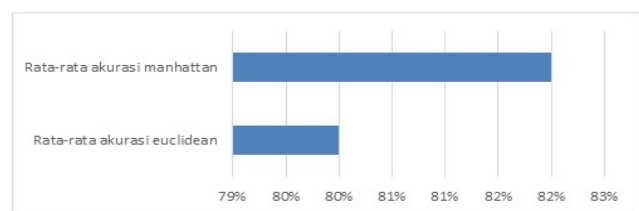


Gambar 13. Pengaruh Data Latih Terhadap Akurasi Jarak *Manhattan*

Nilai akurasi yang di dapat cenderung sama yaitu sebesar 90%. Hal ini dikarenakan atribut yang tidak banyak dan juga hanya mempunyai dua kelas klasifikasi sehingga algoritma dapat dengan mudah mengenali pola yang ada pada data latih dengan menghitung nilai kedekatannya pada setiap atribut.

g. Skenario Pengujian Perbandingan Jarak *Euclidean* dan *Manhattan*

Perbandingan akurasi didapatkan dari hasil beberapa skenario pengujian sebelumnya. Dari hasil pengujian menunjukkan bahwa jarak *manhattan* memiliki nilai rata-rata akurasi yang lebih baik dari pada jarak *euclidean*. Yaitu 80% untuk *euclidean* dan 82% untuk *manhattan*. Agar lebih jelas lihat pada gambar 14.



**Gambar 14.** Rata-rata Akurasi Jarak *Euclidean* dan *Manhattan*

#### IV. KESIMPULAN

Berdasarkan implementasi algoritma MKNN pada data historis gempa bumi untuk penentuan potensi tsunami dengan menggunakan dataset gempa Maluku Utara pada tahun-tahun sebelumnya didapati hasil bahwa algoritma hanya mampu mengklasifikasi gempa yang tidak berpotensi tsunami sehingga status dalam penentuan potensi tsunami di Maluku Utara adalah tidak berpotensi tsunami. Algoritma MKNN tidak bisa mengenali pola proses klasifikasi potensi tsunami di Maluku Utara karena komposisi dan penggunaan data latih gempa bumi yang tidak seimbang.

Dalam melakukan perbandingan jarak yang diterapkan pada metode MKNN yaitu jarak euclidean dan manhattan untuk mengklasifikasi status gempa bumi berdasarkan data historis gempa, hasil dari perbandingan di dapati dari beberapa skenario pengujian sebelumnya yaitu pengaruh komposisi data latih, perubahan nilai K dan perubahan data latih. Nilai akurasi tertinggi yang didapati adalah sebesar 90% untuk kedua perhitungan jarak menggunakan K=1. Dengan demikian diperoleh rata-rata kinerja sistem sebesar 80% untuk jarak euclidean dan 82% untuk jarak manhattan. Berdasarkan rekapitulasi hasil pengujian untuk setiap skenario pengujian yang dilakukan, komposisi dataset sangat berpengaruh terhadap kinerja sistem yang didapat. Sehingga Penggunaan jarak manhattan memiliki tingkat akurasi yang lebih tinggi dibandingkan jarak euclidean dengan selisih rata-rata 2%.

#### REFERENSI

[1] Sunarjo M TAUFIK GUNAWAN SUGENG

- PRIBADI, *Gempabumi edisi populer.* .
- [2] N. A. Mustofa, "Gempa Bumi, Tsunami Dan Mitigasinya," *Gempa Bumi, Tsunami Dan Mitigasinya*, vol. 7, no. 1, 2010, doi: 10.15294/jg.v7i1.92.
- [3] Farisa Adelia, Indriati, Achmad Ridok. (2013). Penentuan Potensi Tsunami Akibat Gempa Bumi Bawah Laut dengan Metode Modified K-Nearest Neighbor (MKNN)
- [4] D. P. Utomo and B. Purba, "Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia," no. September, pp. 846–853, 2019.
- [5] T. Rosa, "Perbandingan Metode Pengukuran Jarak Pada Algoritma K-Nearest Neighbor Untuk Klasifikasi," 2019.
- [6] Hipolitus, "Sistem Penunjang Keputusan Potensi Tsunami Menggunakan K-Nearest Neighbor," 2017.
- [7] Cholifah, W. N., Yulianingsih, Y., & Sagita, S. M. (2018). Pengujian Black Box Testing pada Aplikasi Action & Strategy Berbasis Android dengan Teknologi Phonegap. *String (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 3(2), 206.
- [8] P. P. A. Jojo Jennifer Sianipar, M. Tanzil Furqon, "Identifikasi Diagnosis Gangguan Autisme Pada Anak Menggunakan Metode Modified K-Nearest Neighbor ( MKNN )," vol. 1, no. 9, pp. 825–831, 2017.
- [9] Y. Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4 . 5," vol. V2.i2(213-, 2019.
- [10] S. I. Fernanda, D. E. Ratnawati, and P. P. Adikara, "Identifikasi Penyakit Diabetes Mellitus Menggunakan Metode Modified K-Nearest Neighbor ( MKNN )," vol. 1, no. 6, pp. 507–513, 2017.