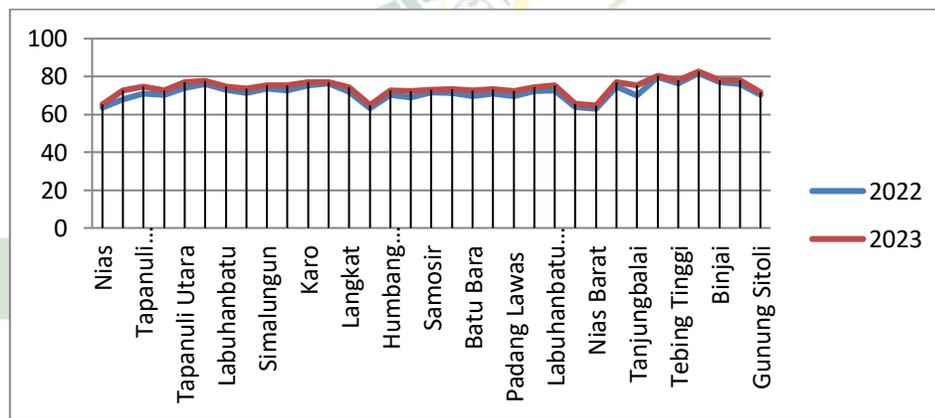


BAB IV HASIL DAN PEMBAHASAN

4.1 Eksplorasi Data

Tujuan dilakukan eksplorasi data untuk memperoleh pemahaman umum tentang karakteristik variabel yang diterapkan pada penelitian ini. Adapun data yang diterapkan yaitu data tahunan. Berikut adalah Eksplorasi data dari semua peubah yang digunakan.

4.1.1 Indeks Pembangunan Manusia

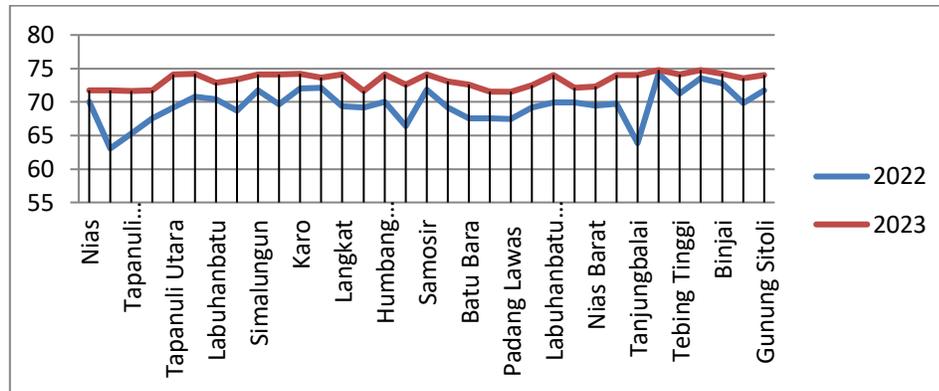


Sumber: (Indeks Pembangunan Manusia di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.1 Grafik Indeks Pembangunan Manusia

Pada Gambar 4.1 di atas diperlihatkan bahwa nilai Indeks Pembangunan Manusia di Sumatera Utara mengalami tren naik dan juga tren turun. Di Sumatera Utara nilai Indeks Pembangunan Manusia yang tertinggi berada di Kota Medan memperoleh nilai IPM sebesar 82,61 di tahun 2023. Sedangkan nilai Indeks Pembangunan Manusia yang terendah berada di Nias Selatan.

4.1.2 Umur Harapan Hidup

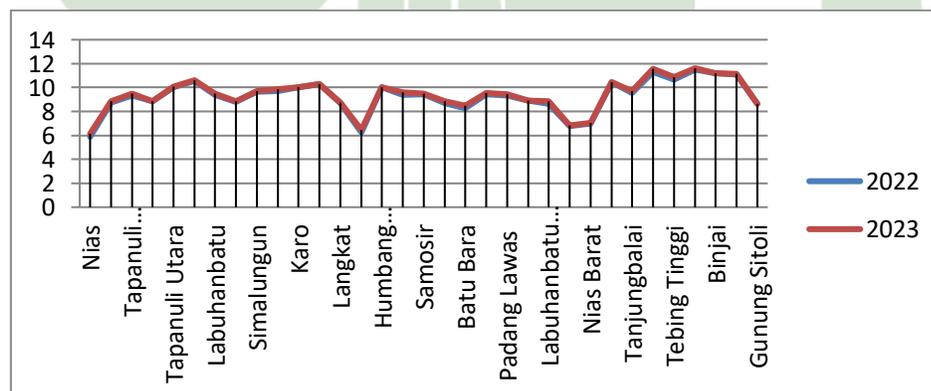


Sumber: (Umur Harapan Hidup saat lahir di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.2 Grafik Umur Harapan Hidup

Pada Gambar 4.2 di atas diperlihatkan bahwa nilai Umur Harapan Hidup Manusia saat lahir pada Sumatera Utara mengalami tren, baik tren naik maupun tren turun. Nilai UHH yang tertinggi di Sumatera Utara pada tahun 2022 berada di Kota Medan dengan nilai UHH sebesar 74,32, dengan kata lain pada tahun 2022 bayi yang lahir memiliki rata-rata untuk dapat hidup mencapai 74,32 tahun. Sedangkan, pada tahun 2023 di Kota Medan nilai UHH mengalami kenaikan dengan nilai UHH sebesar 74,76.

4.1.3 Rata-rata Lama Sekolah

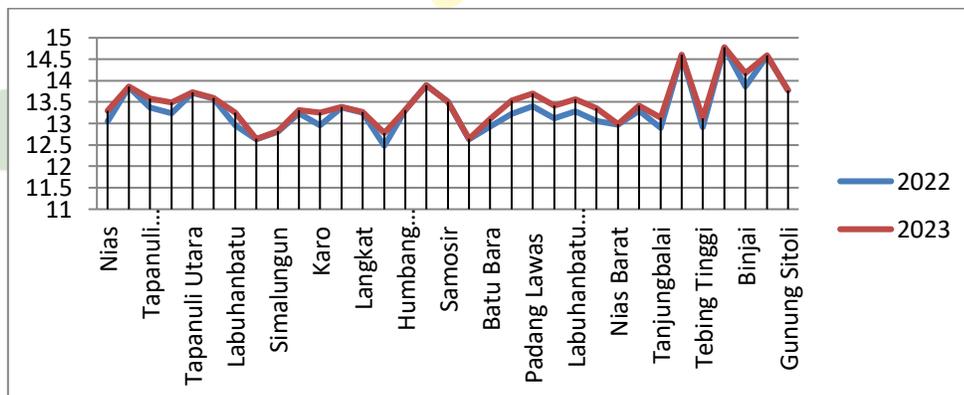


Sumber: (Rata-rata Lama Sekolah di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.3 Grafik Rata-rata Lama Sekolah

Pada Gambar 4.3 di atas diperlihatkan nilai Rata-rata Lama Sekolah di Sumatera Utara mengalami tren, baik tren naik maupun tren turun. Indeks Pendidikan sangat berperan dalam meningkatkan kualitas individu setiap manusia. Rata-rata Lama Sekolah serta Harapan Lama Sekolah merupakan gambaran dalam dimensi pendidikan. Dapat dilihat dari Grafik 4.3 diatas bahwa di Provinsi Sumatera Utara yang tertinggi Rata-rata Lama sekolahnya berada pada Kota Medan dengan nilai 11,5 pada tahun 2022 dan 11,62 pada tahun 2023. Hal ini berarti bahwasannya masyarakat di Kota Medan yang berusia tersebut telah menempuh pendidikan Sekolah Menengah Atas (SMA).

4.1.4 Harapan Lama Sekolah

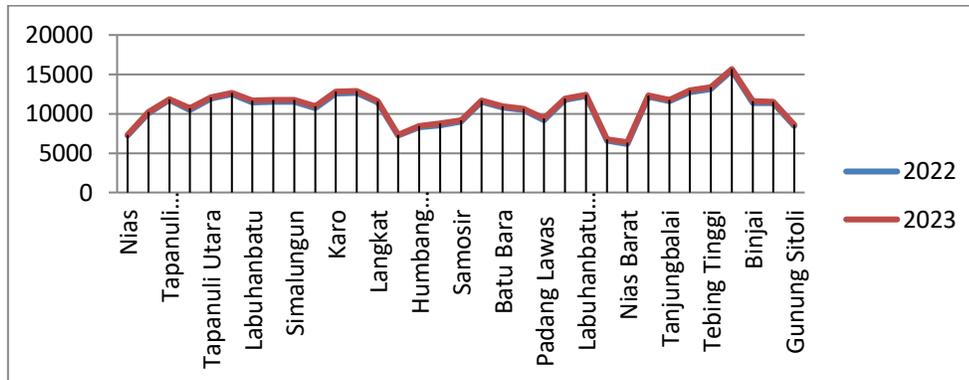


Sumber: (Harapan Lama Sekolah di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.4 Grafik Harapan Lama Sekolah

Pada Gambar 4.4 di atas diperlihatkan Harapan Lama Sekolah di Sumatera Utara mengalami tren naik maupun tren turun. Dapat dilihat pada grafik 4.4 diatas bahwa di Provinsi Sumatera Utarayang tertinggi nilai Harapan Lama Sekolahnya berada pada Kota Medan dengan nilai 14,77 pada tahun 2022 dan 14,78 pada tahun 2023. Kenaikan nilai HLS menunjukkan bahwa kualitas pendidikan masyarakat di Sumatera Utara telah meningkat, sebagaimana tercermin dari bertambahnya nilai Harapan Lama Sekolah.

4.1.5 Pengeluaran Per Kapita Yang Disesuaikan

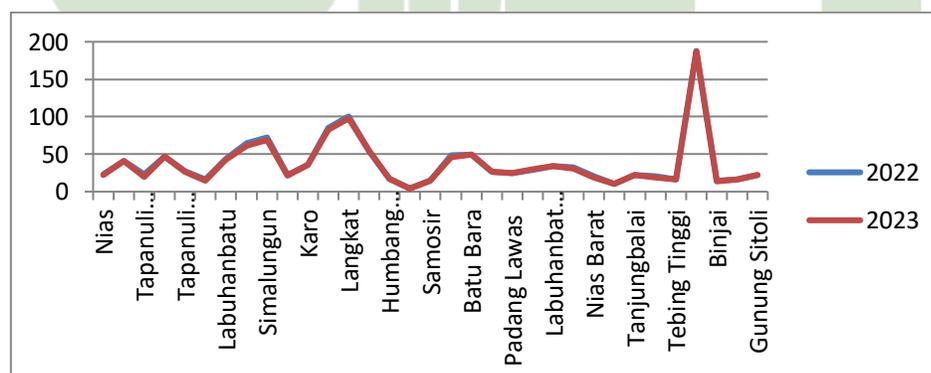


Sumber: (Pengeluaran Per Kapita Yang Disesuaikan di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.5 Pengeluaran perkapita yang disesuaikan

Pada Gambar 4.5 di atas dapat ditinjau bahwa pada Provinsi Sumatera Utara yang tertinggi nilai Pengeluaran perkapita yang disesuaikan berada pada Kota Medan dengan nilai 15503 pada tahun 2022 dan 15674 pada tahun 2023. Ini menunjukkan bahwa daya beli penduduk di Provinsi Sumatera Utara mengalami peningkatan antara tahun 2022 dan 2023. Peningkatan ini berkaitan dengan kemampuan penduduk untuk memperoleh pendapatan serta masalah ketenagakerjaan.

4.1.6 Jumlah Penduduk Miskin



Sumber: (Jumlah Penduduk Miskin di Provinsi Sumatera Utara Tahun 2022-2023)

Gambar 4.6 Jumlah Penduduk Miskin

Pada Gambar 4.6 di atas dapat ditinjau bahwa di Provinsi Sumatera Utara yang tertinggi Jumlah Penduduk Miskin berada pada Kota Medan dengan 187,74 ribu jiwa

pada tahun 2022 dan 187,28 ribu jiwa pada tahun 2023. Pada tahun 2022-2023 di Sumatera Utara Jumlah penduduk miskin mengalami penurunan itu artinya masyarakat sudah mengalami kesadaran dalam memenuhi kebutuhannya terutama dalam mengakses pelayan kesehatan, peningkatan pendapatan juga akses pendidikan dimana hal tersebut berkaitan dalam mempengaruhi jumlah penduduk miskin.

4.2 Normalisasi Data

Proses ini bertujuan untuk menyelaraskan skala nilai dari variabel prediktor (X) yang diterapkan pada penelitian ini perlu berada pada rentang tertentu. Normalisasi yang diterapkan adalah metode normalisasi data Min-Max. Berdasarkan persamaan (2.8) perhitungan hasil dari normalisasi menggunakan variabel prediktor (X) adalah sebagai berikut:

Normalisasi data untuk Umur Harapan Hidup pada data ke-1

Nilai X untuk data ke-1 : 70,06

Nilai X_{min} Umur Harapan Hidup : 63,05

Nilai X_{max} Umur Harapan Hidup : 74,76

$$\begin{aligned}
 X_{normalisasi} &= \frac{X - X_{min}}{X_{max} - X_{min}} \\
 &= \frac{70,06 - 63,05}{74,76 - 63,05} \\
 &= 0,5986
 \end{aligned}$$

Begitu seterusnya, sampai hasil perhitungan pada data ke-66

Tabel 4.1 Normalisasi Data

No.	Kabupaten/Kota	IPM	UHH	RLS	HLS	PPD	JPM
1	Nias	63,69	0,5986	0	0,2435	0,1096	0,1046
2	Mandailing Natal	68,05	0	0,5017	0,5957	0,4105	0,2012
3	Tapanuli Selatan	70,92	0,1904	0,6028	0,387	0,5855	0,1036
4	Tapanuli Tengah	70,31	0,3868	0,5192	0,3304	0,4561	0,2344
5	Tapanuli Utara	74,14	0,5218	0,7178	0,5391	0,6028	0,1277
6	Toba	75,96	0,6584	0,8188	0,4783	0,664	0,0679

7	Labuhanbatu	72,92	0,6251	0,6132	0,2043	0,5589	0,2137
8	Asahan	71,13	0,4851	0,5122	0,0652	0,5632	0,3292
9	Simalungun	73,77	0,7455	0,6533	0,1435	0,5642	0,3726
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
66	Gunung Sitoli 2	71,55	0,9377	0,4826	0,5652	0,2608	0,0981

Sumber: (data diolah menggunakan Microsoft Excel)

Menurut tabel 4.1, nilai dari variabel X yang dipakai dalam penelitian ini memiliki skala yang seragam dan berada pada rentang 0 hingga 1. Secara lengkap hasil perhitungan normalisasi data dapat dilihat pada lampiran 3.

4.3 Jarak Euclidean

Jarak *euclidean* digunakan untuk mengukur jarak terdekat antara satu data dengan data lainnya. Teknik ini adalah yang paling umum dipakai dalam metode *K-Nearest Neighbor*. Rumus dari jarak *euclidean* dapat dilihat di persamaan (2.9). Berikut adalah penyelesaian dari jarak *euclidean*.

$$d_{x_{train-i}, x_{test-j}} = \sqrt{\sum_{i,j=1}^n (X_{train-i} - X_{test-j})^2}$$

$$d_{Nias, Nias} = \sqrt{(0,5986 - 0,5986)^2 + (0 - 0)^2 + (0,2435 - 0,2435)^2 + (0,1096 - 0,1096)^2 + (0,1046 - 0,1046)^2}$$

$$= 0$$

$$d_{Nias, Mandailing Natal} = \sqrt{(0,5986 - 0)^2 + (0 - 0,5017)^2 + (0,2435 - 0,5957)^2 + (0,1096 - 0,4105)^2 + (0,1046 - 0,2012)^2}$$

$$= \sqrt{(0,5986)^2 + (-0,5017)^2 + (-0,3522)^2 + (-0,3009)^2 + (-0,0966)^2}$$

$$= \sqrt{0,8339973}$$

$$= 0,9132345$$

Begitu seterusnya, sampai hasil perhitungan gunung sitoli 2, gunung sitoli 2.

$$\begin{aligned}
 & d_{\text{Gunung Sitoli2, Gunung Sitoli2}} \\
 &= \sqrt{(0,9377 - 0,9377)^2 + (0,4826 - 0,4826)^2 + (0,5652 - 0,5652)^2 +} \\
 &\quad (0,2608 - 0,2608)^2 + (0,0981 - 0,0981)^2} \\
 &= 0
 \end{aligned}$$

Tabel 4.2 Jarak Euclidean

No.	Kabupaten/Kota	Nias	Mandailing Natal	Tapanuli Selatan	...	Gunung Sitoli 2
1	Nias	0	0.9132345	0.8814728	...	0.6886349
2	Mandailing Natal	0.9132345	0	0.3607845	...	0.9558062
3	Tapanuli Selatan	0.8814728	0.3607845	0	...	0.8426427
⋮	⋮	⋮	⋮	⋮	⋮	⋮
66	Gunung Sitoli 2	0.6886349	0.9558062	0.8426427	...	0

Sumber: (data diolah menggunakan Microsoft Excel)

Hasil perhitungan jarak *euclidean* secara lengkap dapat ditemukan pada lampiran 4.

4.4 Prediksi *K-Nearest Neighbor* (K-NN)

Tujuan dilakukan penelitian ini untuk memprediksi Indeks Pembangunan Manusia berdasarkan jarak terdekat antara data *testing* dengan data *training* dengan jumlah k tertentu. Data *training* mencakup periode 1 hingga 33, sementara data *testing* meliputi periode 34 hingga 66. Pada penelitian ini menggunakan nilai $k = 3, 4, 5, 10, 13, 15, 18,$ dan 24 . Langkah pertama yang dilakukan adalah menghitung jarak *Euclidean*. Setiap jarak terdekat dari setiap data *training* untuk data *testing* pada setiap k yang sudah ditentukan dapat ditemukan pada tabel dibawah ini.

Tabel 4.3 Jarak terdekat untuk data *Testing* 1

Jumlah k	Data Jarak Terdekat
$k = 3$	1, 24 dan 25
$k = 4$	1, 24, 25 dan 14

$k = 5$	1, 24, 25, 14 dan 33
$k = 10$	1, 24, 25, 14, 33, 17, 19, 4, 21 dan 15.
$k = 13$	1, 24, 25, 14, 33, 17, 19, 4, 21, 15, 23, 22 dan 18
$k = 15$	1, 24, 25, 14, 33, 17, 19, 4, 21, 15, 23, 22, 18, 10 dan 7
$k = 18$	1, 24, 25, 14, 33, 17, 19, 4, 21, 15, 23, 22, 18, 10, 7, 20, 13 dan 16
$k = 24$	1, 24, 25, 14, 33, 17, 19, 4, 21, 15, 23, 22, 18, 10, 7, 20, 13, 16, 8, 9, 5, 11, 3 dan 26.

Sumber: data diolah menggunakan RStudio

Begitu seterusnya, sampai data jarak terdekat untuk data *testing* 33

Tabel 4.4 Jarak terdekat untuk data *Testing* 33

Jumlah k	Data Jarak Terdekat
$k = 3$	33, 17 dan 15
$k = 4$	33, 17, 15 dan 10
$k = 5$	33, 17, 15, 10 dan 31
$k = 10$	33, 17, 15, 10, 31, 23, 5, 7, 21 dan 6
$k = 13$	33, 17, 15, 10, 31, 23, 5, 7, 21, 6, 22, 24 dan 11
$k = 15$	33, 17, 15, 10, 31, 23, 5, 7, 21, 6, 22, 24, 11, 9 dan 26,
$k = 18$	33, 17, 15, 10, 31, 23, 5, 7, 21, 6, 22, 24, 11, 9, 26, 4, 20 dan 25
$k = 24$	33, 17, 15, 10, 31, 23, 5, 7, 21, 6, 22, 24, 11, 9, 26, 4, 20, 25, 12, 16, 1, 13, 32 dan 19.

Sumber: data diolah menggunakan RStudio

Secara Lengkap data jarak terdekat untuk data testing dapat dilihat pada lampiran 5. Untuk memperoleh hasil prediksi, variabel prediktor yang dipilih pada setiap nilai k dirata-ratakan menggunakan rumus dalam persamaan (2.10). Perhitungan hasil prediksi dilakukan secara manual. Adapun perhitungannya sebagai berikut:

$$Y = \frac{1}{k} \sum_{i=1}^k y_i$$

$$\begin{aligned}
 Y_{pred1} &= \frac{63,69 + 63,75 + 62,93}{3}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{190,37}{3} \\
 &= 63,46
 \end{aligned}$$

Begitu seterusnya, sampai Y prediksi ke 33

$$\begin{aligned}
 &Y_{pred33} \\
 &= \frac{70,23 + 71,67 + 70,32 + 72,56 + 76,95 + 72,77 + \dots + 76,05 + 69,51}{24} \\
 &= \frac{1717,2}{24} \\
 &= 71,55
 \end{aligned}$$

Menurut hasil perhitungan menggunakan rumus pada persamaan (2.10) dapat dilihat bahwa hasil prediksi data *testing* 1 pada $k = 3$ yaitu sebesar 63,46 dan hasil prediksi untuk data *testing* 33 pada $k=24$ yaitu sebesar 71,55. Hasil prediksi pada data *testing* dan jumlah k lainnya dapat ditemukan pada tabel 4.5 dibawah ini .

Tabel 4.5 Hasil Prediksi *K-Nearest Neighbor* Tunggal

No.	$k = 3$	$k = 4$	$k = 5$	$k = 10$	$k = 13$	$k = 15$	$k = 18$	$k = 24$
1	63,46	63,39	64,75	67,52	68,56	69,12	69,36	70,35
2	72,01	72,2	72,27	72,23	72,35	72,57	72,91	72,72
3	74,22	73,86	74,03	73,92	73,99	73,54	73,15	72,81
4	72,33	72,48	72,03	72,59	72,64	72,74	72,59	72,51
5	76,09	76,12	75,23	75,34	74,57	74,35	74,03	73,17
6	76,09	76,11	75,84	75,34	74,84	74,74	74,03	73,17
7	74,02	74,56	74,2	73,95	74,26	73,9	73,34	72,96
8	74,02	73,32	72,88	73,35	72,87	72,82	72,75	71,97
9	75,11	74,56	74,88	73,39	73,27	73,69	73,34	72,3
10	74,66	74,99	74,5	74,63	73,82	73,73	73,24	73,17
11	75,91	75,92	75,49	74,74	74,26	73,79	73,34	73,17

12	75,11	75,32	74,84	74,61	74,27	73,89	73,87	73,17
13	73,94	73,69	74,02	72,9	72,47	72,82	72,93	72,66
14	63,54	63,39	64,61	67,76	68,81	69,18	69,64	70,41
15	70,74	72,29	72,35	73,45	73,61	73,29	73,15	72,21
16	70,74	72,29	72,35	72,5	72,67	72,66	73,07	72,4
17	70,74	72,29	72,35	73,35	73,35	73,29	73,15	72,84
18	74,02	73,32	72,88	73,42	72,87	72,82	72,75	71,97
19	73,15	73,71	73,4	72,54	72,47	72,13	72,43	71,96
20	72,79	73,58	73,42	72,75	73,13	72,71	73,15	72,51
21	70,74	71,2	71,78	72,51	72,69	72,27	72,75	72,71
22	73,68	73,3	73,15	73,61	73,74	73,79	73,33	72,51
23	74,77	75,07	74,39	74,03	73,8	73,83	73,29	73,17
24	63,46	65,15	64,75	67,52	68,67	69,1	69,36	70,35
25	63,46	63,39	64,75	67,52	68,6	69,12	69,59	70,35
26	76,09	76,11	75,84	74,67	74,26	74,36	74,03	73,17
27	74,82	74,56	74,84	74,21	74,26	73,79	73,29	72,96
28	77,57	77,17	76,56	75,69	74,81	74,38	74,63	73,58
29	75,83	75,56	75,84	74,88	74,26	74,42	73,81	73,17
30	79,22	78,43	78,13	76,17	75,8	75,39	74,72	73,69
31	77,57	77,17	76,56	75,1	74,62	74,38	74,03	73,14
32	77,57	77,17	76,56	74,82	74,62	74,11	73,68	73,58
33	70,74	71,2	72,35	72,71	72,18	72,46	71,73	71,55

Sumber: (data diolah menggunakan Microsoft Excel)

Menurut Tabel 4.5 didapat bahwa hasil prediksi $k = 4$ memiliki nilai dengan akurasi yang paling tinggi. Hal ini dibuktikan dengan menemukan nilai MAPE. Adapun perhitungannya adalah sebagai berikut:

$$\text{MAPE} = \sum_{i=1}^n \frac{1}{m} \cdot \frac{|x_i - y_i|}{x_i} \times 100\%$$

Untuk $k = 3$ Nilai MAPE-nya sebagai berikut:

$$\begin{aligned} \text{MAPE} &= \frac{1}{33} \cdot \left(\frac{|63,64 - 65,15|}{65,15} + \frac{|72,01 - 72,65|}{72,65} + \frac{|74,22 - 74,58|}{74,58} + \dots \right. \\ &\quad \left. + \frac{|70,74 - 71,55|}{71,55} \right) \times 100\% \\ &= \frac{1}{33} \cdot (0,5067) \times 100\% \\ &= 1,54 \% \end{aligned}$$

$= 1,54 \%$

Begitu seterusnya sampai $k = 24$. Untuk hasil perhitungan nilai MAPE secara lengkap dapat ditemukan pada Tabel 4.6.

Tabel 4.6 Evaluasi Hasil Prediksi *K-Nearest Neighbor* Tunggal

K-NN	MAPE
$k = 3$	1,54 %
$k = 4$	1,29 %
$k = 5$	1,35 %
$k = 10$	2,23 %
$k = 13$	2,69 %
$k = 15$	2,89 %
$k = 18$	3,21 %
$k = 24$	3,83 %
Rata-rata	2,38 %

Sumber: (data diolah menggunakan Microsoft Excel)

Hasil prediksinya yang baik dapat dilihat dengan nilai MAPE yang kecil. Berdasarkan Tabel 4.6 dapat dilihat bahwa $k = 4$ mendapatkan nilai MAPE yang relatif kecil yaitu sebesar 1,29 %. Sementara itu nilai MAPE yang relatif besar terdapat pada $k = 24$ dengan nilai 3,83 %. Setelah melakukan prediksi menggunakan

metode K-Nearest Neighbor Tunggal selanjutnya dilakukan pengoptimalan kinerja K-Nearest Neighbor dengan suatu teknik yaitu teknik *Ensemble*. Prinsip dari teknik *ensemble* ini adalah dengan menerapkan pembobotan, sehingga tidak diperlukan lagi mencari nilai k yang optimal agar mendapatkan nilai akurasi yang tinggi. Adapun dilakukan teknik *ensemble* ini bertujuan untuk melihat manakah yang jauh lebih baik dalam melakukan prediksi Indeks Pembangunan Manusia diantara metode *K-Nearest Neighbor* Tunggal dengan *Ensemble K-Nearest Neighbor*.

4.5 Prediksi *Ensemble K-Nearest Neighbor* (K-NN)

Adapun tujuan dari Teknik *Ensemble* yaitu untuk menggabungkan hasil prediksi dari K-NN tunggal menjadi suatu prediksi akhir. *Weighted mean* (rata-rata terboboti) merupakan teknik *ensemble* yang diterapkan pada penelitian ini. Bobot untuk teknik ini yaitu korelasi antara data sebenarnya dengan data hasil prediksi K-NN tunggal, yang dihitung dengan menerapkan rumus pada persamaan (2.13). Adapun perhitungannya adalah sebagai berikut:

$$r_i = \frac{n \sum \bar{y}_i \hat{y}_t - (\sum \bar{y}_i)(\sum \hat{y}_t)}{\sqrt{(n \sum \bar{y}_i^2 - (\sum \bar{y}_i)^2)(n \sum \hat{y}_t^2 - (\sum \hat{y}_t)^2)}}$$

Untuk $k = 3$ perhitungannya adalah sebagai berikut

$$r_i = \frac{\{33 ((63,46 \times 65,15) + (72,01 \times 72,65) + \dots + (77,57 \times 78,1) + (70,74 \times 71,55))\} - \{(63,46 + 72,01 + \dots + 77,57 + 70,74)\}\{(65,15 + 72,65 + \dots + 78,1 + 71,55)\}}{\sqrt{\left[\{33 ((63,46)^2 + (72,01)^2 + \dots + (77,57)^2 + (70,74)^2)\} - \{(63,46 + 72,01 + \dots + 77,57 + 70,74)\}^2 \right] \times \left[\{33 ((65,15)^2 + (72,65)^2 + \dots + (78,1)^2 + (71,55)^2)\} - (65,15 + 72,65 + \dots + 78,1 + 71,55)^2 \right]}}$$

$$r_i = \frac{(33 \times 178784,81) - (2408,17 \times 2442,41)}{\{(33 \times 176296,9) - 2408,17^2\}\{(33 \times 181335,8) - 2442,41^2\}}$$

$$r_i = \frac{18144,006}{\sqrt{345929552}}$$

$$r_i = 0,975527$$

Begitu seterusnya sampai perhitungan korelasi untuk $k = 24$. Adapun hasil perhitungan korelasi product moment dapat ditemukan pada tabel 4.7 dibawah ini

Tabel 4.7 Korelasi antara data aktual dengan hasil prediksi menggunakan metode *K-Nearest Neighbor*

r_h	Korelasi
$r_{k=3}$	0,975527
$r_{k=4}$	0,971242
$r_{k=5}$	0,9732998
$r_{k=10}$	0,964590
$r_{k=13}$	0,959216
$r_{k=15}$	0,961953
$r_{k=18}$	0,95621
$r_{k=24}$	0,940805
Jumlah	7,702843

Sumber: (data diolah menggunakan Microsoft Excel)

Setelah dilakukan perhitungan nilai korelasi dari data aktual terhadap data hasil prediksi K-NN tunggal maka selanjutnya dilakukan perhitungan pembobot *ensemble* dengan menggunakan rumus pada persamaan (2.12). Adapun perhitungannya adalah sebagai berikut:

$$W_i = \frac{r_i}{\sum_{j=1}^s r_j}$$

$$W_1 = \frac{0,975527}{7,702843}$$

$$= 0,126645$$

Begitu seterusnya sampai perhitungan w_8 . Adapun hasil perhitungan secara lengkapnya dapat dilihat pada tabel 4.8 dibawah ini.

Tabel 4.8 Pembobot *Ensemble*

w_i	Hasil Pembobot <i>ensemble</i>
w_1	0,126645
w_2	0,1260888
w_3	0,126356
w_4	0,125225
w_5	0,1245276
w_6	0,1248828
w_7	0,1241372
w_8	0,1221373
Jumlah	1

Sumber: (data diolah menggunakan Microsoft Excel)

Setelah dilakukan perhitungan terhadap pembobot *ensemble* maka selanjutnya melakukan prediksi *Ensemble* K-NN menggunakan rumus dari persamaan (2.11)

$$\mu = \frac{\sum_{i=1}^s w_i y_i}{\sum_{i=1}^s w_i}$$

$$\mu_1 = \frac{(0,126645 \times 63,46) + (0,126089 \times 63,39) + \dots + (0,122137 \times 70,35)}{1}$$

$$= 67,04$$

Adapun hasil perhitungan secara lengkapnya dapat ditemukan pada tabel 4.9 dibawah ini.

Tabel 4.9 Hasil Prediksi *Ensemble* K-NN

No.	Kabupaten/Kota	<i>Ensemble</i> K-NN
1	Nias	67,04
2	Mandailing Natal	72,41
3	Tapanuli Selatan	73,69
4	Tapanuli Tengah	72,49
5	Tapanuli Utara	74,87

6	Toba	75,03
7	Labuhanbatu	73,9
8	Asahan	73
9	Simalungun	73,83
10	Dairi	74,1
11	Karo	74,59
12	Deli Serdang	74,39
13	Langkat	73,18
14	Nias Selatan	67,14
15	Humbang Hasundutan	72,63
16	Pakpak Bharat	72,33
17	Samosir	72,66
18	Serdang Bedagai	73,01
19	Batu Bara	72,73
20	Padang Lawas Utara	73,01
21	Padang Lawas	72,07
22	Labuhanbatu Selatan	73,39
23	Labuhanbatu Utara	74,05
24	Nias Utara	67,27
25	Nias Barat	67,07
26	Sibolga	74,83
27	Tanjungbalai	74,1
28	Pematangsiantar	75,56
29	Tebing Tinggi	74,73
30	Medan	76,46
31	Binjai	75,34
32	Padang Sidempuan	75,28
33	Gunung Sitoli	71,86
	Rata-rata	72,97

Sumber: (data diolah menggunakan Microsoft Excel)

Berdasarkan Hasil Prediksi Indeks Pembangunan Manusia di Provinsi Sumatera Utara menggunakan metode *Ensemble* K-NN ini dinilai cukup baik jika dibandingkan hanya menggunakan metode K-NN Tunggal saja. Hasil prediksi menggunakan metode *Ensemble K-Nearest Neighbor* ini memiliki rata-rata sebesar 72,97. Hal

tersebut juga dapat dibuktikan dengan menghitung evaluasi hasil prediksi *Ensemble* K-NN dibawah ini.

$$\begin{aligned} \text{MAPE} &= \sum_{i=1}^n \frac{1}{m} \cdot \frac{|x_i - y_i|}{x_i} \times 100\% \\ &= \frac{1}{33} \cdot \left(\frac{|67,04 - 65,15|}{65,15} + \frac{|72,41 - 72,65|}{72,65} + \frac{|73,69 - 74,58|}{74,58} + \dots \right. \\ &\quad \left. + \frac{|71,86 - 71,55|}{71,55} \right) \times 100\% \\ &= \frac{1}{33} \cdot (0,6928) \times 100\% \\ &= 2,1 \% \end{aligned}$$

Tabel 4.10 Evaluasi Hasil Prediksi *K- Nearest Neighbor Tunggal dan Ensemble*

K-NN

KNN	MAPE %	Akurasi	Rata-Rata
$k = 3$	1,54 %	Tinggi	2,38 %
$k = 4$	1,29 %	Tinggi	
$k = 5$	1,35 %	Tinggi	
$k = 10$	2,23 %	Tinggi	
$k = 13$	2,69 %	Tinggi	
$k = 15$	2,89 %	Tinggi	
$k = 18$	3,21 %	Tinggi	
$k = 24$	3,83 %	Tinggi	
<i>Ensemble</i> K-NN	2,1 %	Tinggi	2,1 %

Sumber: (data diolah menggunakan Microsoft Excel)

Perolehan nilai MAPE yang kecil, maka akan semakin tinggi tingkat akurasi hasil prediksi. Tabel diatas menunjukkan bahwa teknik *ensemble* memperoleh hasil nilai MAPE yang lebih kecil yaitu sebesar 2,1 %. Hal ini lebih baik dibandingkan hanya menerapkan metode K-NN Tunggal, yang memiliki nilai MAPE sebesar 2,38 %.