

**PERBANDINGAN METODE NAÏVE BAYES DAN K-NEAREST
NEIGHBORS DALAM MENGLASIFIKASIKAN INDEKS
PEMBANGUNAN MANUSIA MENURUT KABUPATEN/
KOTA DI INDONESIA TAHUN 2022**

SKRIPSI

*Diajukan sebagai salah satu persyaratan untuk memperoleh gelar
Sarjana Statistika*



**Oleh
RUDI ANGGARA
NIM. 19337014**

**PROGRAM STUDI SARJANA STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2025**

PERSETUJUAN SKRIPSI

PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST NEIGHBORS* DALAM MENGLASIFIKASIKAN INDEKS PEMBANGUNAN MANUSIA MENURUT KABUPATEN/ KOTA DI INDONESIA TAHUN 2022

Nama	Rudi Anggara
NIM	19337014
Program Studi	SI Statistika
Departemen	Statistika
Fakultas	Matematika dan Ilmu Pengetahuan Alam

Padang, 18 Juli 2025

Mengetahui
Kepala Departemen Statistika



Dr. Yenni Kurniawati, S.Si., M.Si.
NIP. 198402232010122005

Disetujui Oleh:
Pembimbing



Tessy Octavia Mukhti, S.Si., M.Stat.
NIP. 199610312022032011

PENGESAHAN LULUS UJIAN SKRIPSI

Nama	Rudi Anggara
NIM	19337014
Program Studi	S1 Statistika
Departemen	Statistika
Fakultas	Matematika dan Ilmu Pengetahuan Alam

PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST NEIGHBORS* DALAM MENGLASIFIKASIKAN INDEKS PEMBANGUNAN MANUSIA MENURUT KABUPATEN/ KOTA DI INDONESIA TAHUN 2022

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Skripsi
Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang

Padang, 18 Juli 2025

Tim Penguji

	Nama
Ketua	Tessy Octavia Mukhti, S.Si., M.Stat.
Anggota	Dr. Yenni Kurniawati, S.Si., M.Si.
Anggota	Dina Fitria, M. Si.

Tanda tangan



SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Rudi Anggara
NIM/TM : 19337014/2019
Program Studi : S1 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan, bahwa Skripsi saya dengan judul "*Perbandingan Metode Naïve Bayes dan K-Nearest Neighbors dalam Mengklasifikasikan Indeks Pembangunan Manusia menurut Kabupaten/Kota di Indonesia Tahun 2022*" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan Negara.

Demikianlah pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005

Saya yang menyatakan,



Rudi Anggara
NIM. 19337014

PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST NEIGHBORS* DALAM MENGLASIFIKASIKAN INDEKS PEMBANGUNAN MANUSIA MENURUT KABUPATEN/ KOTA DI INDONESIA TAHUN 2022

Rudi Anggara

ABSTRAK

Indeks Pembangunan Manusia (IPM) adalah indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia pada suatu wilayah. IPM Indonesia mengalami peningkatan setiap tahun, namun IPM di beberapa wilayah kabupaten/kota di Indonesia masih berada pada kategori rendah. Rendahnya IPM di kabupaten/kota tersebut dikarenakan tidak meratanya pembangunan antar wilayah di Indonesia. Tidak meratanya pembangunan dipengaruhi oleh berbagai faktor, seperti umur panjang dan hidup sehat, pengetahuan dan standar hidup layak maupun faktor-faktor lainnya. Penelitian ini bertujuan untuk mengetahui pengklasifikasian dan tingkat akurasi dalam klasifikasi IPM menurut kabupaten/kota di Indonesia tahun 2022.

Penelitian ini merupakan penelitian terapan. Data yang digunakan merupakan data sekunder yang bersumber dari publikasi Badan Pusat Statistik tahun 2022. Analisis yang digunakan adalah metode *Naïve Bayes* dan *K-Nearest Neighbors*. Variabel dalam penelitian ini terdiri dari variabel prediktor yaitu Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita yang disesuaikan. Sedangkan untuk variabel respons yaitu nilai IPM.

Hasil klasifikasi IPM menurut kabupaten/kota di Indonesia terhadap 103 data *testing* pada metode *Naïve Bayes* berhasil mengklasifikasikan 7 amatan dikategorikan sangat tinggi, 37 amatan dikategorikan tinggi, 44 amatan dikategorikan sedang, dan 8 amatan dikategorikan rendah. Pada metode KNN berhasil mengklasifikasikan 7 amatan dikategorikan sangat tinggi, 39 amatan dikategorikan tinggi, 43 amatan dikategorikan sedang, dan 6 amatan dikategorikan rendah. Akurasi metode *Naïve Bayes* sebesar 93,2%, sedangkan KNN memiliki tingkat akurasi sebesar 92,23. Berdasarkan nilai tersebut disimpulkan bahwa metode *Naïve Bayes* memiliki performa yang lebih baik dalam mengklasifikasikan IPM menurut kabupaten/kota di Indonesia tahun 2022.

Kata Kunci : IPM, KNN, *Naïve Bayes*

COMPARISON OF NAÏVE BAYES AND K-NEAREST NEIGHBORS METHODS IN CLASSIFYING HUMAN DEVELOPMENT INDEX BY DISTRICTS/CITY INDONESIA IN 2022

Rudi Anggara

ABSTRACT

The Human Development Index (HDI) is an important indicator for measuring the success of efforts to build human quality in a region. Indonesia's HDI has increased every year, but the HDI in several districts/cities in Indonesia is still in the low category. The low HDI in these districts/cities is due to unequal development between regions in Indonesia. The inequality of development is influenced by factors such as life expectancy, education, proper living standards, and other factors. This study aims to determine the classification and accuracy level in classifying HDI by district/city in Indonesia 2022.

This type of research is applied research. The data used secondary data sourced from the 2022 publication of the Central Bureau of Statistics. The analysis methods used are the Naïve Bayes and K-Nearest Neighbors methods. The variables consist of predictor variable, Namely Life Expectancy, Average Years of Schooling, Expected Years of Schooling, and Adjusted per Capita Income. The response variable is the HDI value.

The HDI classification results for district/cities in Indonesia on 103 testing data using the Naïve Bayes method classified 7 observation into the very high category, 37 observation into the high category, 44 observation into the medium category, and 8 observation into the low category. Using the KNN method, 7 observation into the very high category, 39 observation into the high category, 43 observation into the medium category, and 6 observation into the low category. The accuracy of the Naïve Bayes method was 93,2%, while the KNN method accuracy was 92,23%. Based on these accuracy, it can be concluded that the Naïve Bayes method has very good performance in classifying HDI by district/city in Indonesia 2022.

Keywords : HDI, KNN, Naïve Bayes

KATA PENGANTAR

Puji dan syukur atas limpahan rahmat, hikmat, hidayah, dan inayah yang telah Allah SWT berikan, sehingga skripsi yang berjudul “**Perbandingan Metode *Naïve Bayes* dan K-Nearest Neighbors dalam Mengklasifikasikan Indeks Pembangunan Manusia menurut Kabupaten/Kota di Indonesia Tahun 2022**” dapat diselesaikan dengan baik.

Skripsi ini merupakan salah satu syarat memperoleh gelar Sarjana Statistika pada Program Studi Statistika, Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang. Skripsi ini dalam penyusunannya tidak terlepas dari berbagai macam kendala, tetapi berkat bantuan, bimbingan, dan dukungan dari berbagai pihak, serta ridha dari Allah SWT sehingga kendala-kendala yang dihadapi dapat diatasi. Oleh sebab itu, ucapan terima kasih disampaikan kepada pihak-pihak yang telah berkontribusi dalam penyusunan skripsi ini.

1. Ibu Dr. Yenni Kurniawati, S.Si., M.Si., Kepala Departemen Statistika dan Dosen Penguji yang telah memberikan arahan dan masukan sehingga skripsi ini dapat diselesaikan dengan baik.
2. Ibu Tessy Octavia Mukhti, M.Stat., Dosen Pembimbing sekaligus Dosen Penasehat Akademik yang telah meluangkan waktu untuk membimbing dan memberikan arahan selama penyusunan skripsi.
3. Ibu Dina Fitria, M.Si., Dosen Penguji yang telah memberikan arahan dan masukan sehingga skripsi ini dapat diselesaikan dengan baik.

4. Bapak dan Ibu Dosen serta Staff Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang.
5. Teristimewa kepada kedua orang tua, Bapak Yofriman dan Almarhumah Ibu Guslina serta saudara penulis atas doa, kasih sayang, dukungan, dan segalanya yang telah diberikan kepada penulis dalam proses menyelesaikan skripsi penulis.
6. Sahabat, teman, serta rekan-rekan yang selalu memberikan dukungan serta semangat kepada penulis.

Semoga skripsi ini memberikan manfaat untuk penulis sendiri dan semua pihak. Skripsi ini tidak terlepas dari kesalahan dan kekeliruan, oleh sebab itu penulis mengharapkan saran dan kritik yang bersifat membangun. Akhir kata penulis ucapkan terima kasih.

Padang, 18 Juli 2025

Rudi Anggara

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
DAFTAR LAMPIRAN	viii
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Batasan Masalah.....	6
C. Rumusan Masalah	6
D. Tujuan Penelitian	6
E. Manfaat Penelitian	7
BAB II KERANGKA TEORITIS.....	8
A. Klasifikasi	8
B. <i>Naïve Bayes</i>	10
C. <i>K-Nearest Neighbors</i>	13
D. Pengukuran Kinerja Klasifikasi	15
E. Indeks Pembangunan Manusia.....	17
BAB III METODOLOGI PENELITIAN.....	20
A. Jenis Penelitian.....	20
B. Jenis dan Sumber Data	20
C. Variabel Penelitian	20
D. Struktur Data	21
E. Tahapan Analisis.....	21
F. Diagram Alir	22
BAB IV HASIL DAN PEMBAHASAN	24
A. Hasil Penelitian	24
B. Pembahasan.....	36
BAB V KESIMPULAN DAN SARAN.....	38
A. KESIMPULAN	38
B. SARAN	38
DAFTAR PUSTAKA	40
LAMPIRAN	42

DAFTAR GAMBAR

Gambar	Halaman
1. Indeks Pembangunan Manusia Indonesia Tahun 2015-2022.....	2
2. Lima Kabupaten/Kota dengan IPM Tertinggi dan Terendah Tahun 2022	3
3. Diagram Alir	23

DAFTAR TABEL

Tabel	Halaman
1. <i>Confusion Matrix</i>	15
2. Klasifikasi Akurasi.....	16
3. Kategori IPM.....	19
4. Variabel Penelitian	20
5. Struktur Data	21
6. Analisis Deskriptif	24
7. Data yang telah distandardisasi	25
8. Tabel Deskriptif Data <i>Training</i>	27
9. <i>Mean</i> dan Standar Deviasi Variabel Prediktor berdasarkan Kategori IPM	27
10. Data <i>Testing</i> Pertama	27
11. <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	30
12. Hasil Jarak <i>Euclidean</i>	32
13. Hasil Urutan Jarak Data <i>Training</i>	33
14. <i>Confusion Matrix</i> Metode KNN	34
15. Perbandingan Akurasi Parameter KNN	35
16. Nilai Akurasi Klasifikasi <i>Naïve Bayes</i> dan KNN	36

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Penelitian	42
2. Data Penelitian yang telah distandardisasi	43
3. <i>Syntax</i> Standardisasi Data menggunakan <i>software</i> Python.....	44
4. Data <i>Training</i>	45
5. Data <i>Testing</i>	46
6. <i>Syntax</i> Pembagian Data <i>Testing</i> dan Data <i>Training</i> menggunakan <i>software</i> Python	47
7. Hasil Klasifikasi Data <i>Testing</i> Metode <i>Naive Bayes</i>	48
8. Hasil Klasifikasi Data <i>Testing</i> Metode KNN.....	49
9. Perhitungan Jarak <i>Euclidean</i> Data <i>Testing</i> Pertama	50
10. Jarak <i>Euclidean</i> KNN antara Data <i>Training</i> dan Data <i>Testing</i>	51

BAB I

PENDAHULUAN

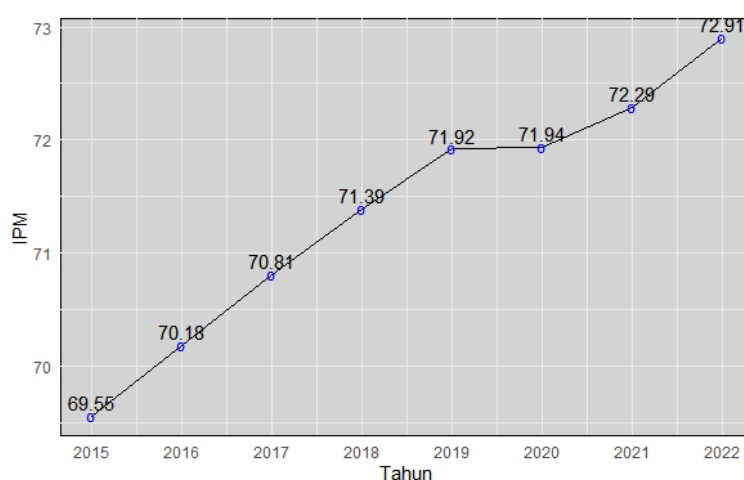
A. Latar Belakang

Pembangunan merupakan suatu upaya yang dilakukan untuk menciptakan perubahan agar tercapainya kehidupan di masa mendatang yang lebih baik. Tujuan utama dari pembangunan adalah mengupayakan kesejahteraan serta mewujudkan kemakmuran manusia. Manusia memegang peranan penting dalam menciptakan pembangunan, khususnya pembangunan untuk menciptakan lingkungan masyarakat yang memiliki kehidupan yang layak.

Dalam melaksanakan pembangunan, konsep pembangunan manusia memiliki cakupan dimensi yang sangat luas. Dengan konsep pembangunan manusia tersebut, pemerintah harus membawa penduduk sebagai pusat perhatian, artinya di samping meningkatkan kualitas manusia tetapi juga memanfaatkan kemampuan manusia tersebut untuk berperan aktif dalam mengisi pembangunan. Pembangunan manusia diperkenalkan oleh *United Nations Development Programme* (UNDP). Menurut UNDP (1990), pembangunan manusia merupakan model pembangunan yang ditujukan untuk memperluas pilihan yang dapat ditumbuhkan melalui upaya pemberdayaan penduduk. Untuk mengetahui tingkat keberhasilan proses pembangunan manusia, UNDP telah mengembangkan sebuah indikator yaitu Indeks Pembangunan Manusia (IPM).

IPM merupakan indikator paling penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia pada suatu wilayah. IPM dibangun dari tiga dimensi dasar, yaitu umur yang panjang dan hidup sehat, pendidikan, serta standar hidup layak (BPS, 2022). IPM dapat menggambarkan

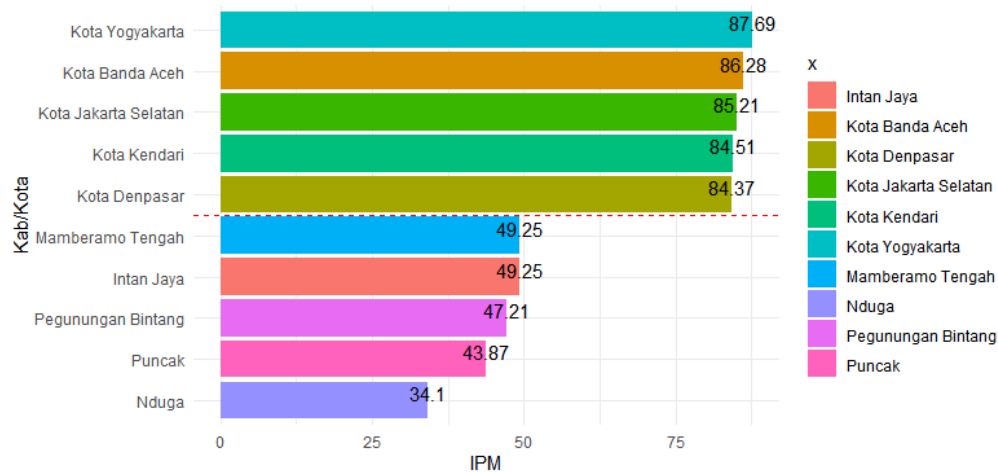
bagaimana perkembangan atau kondisi masyarakat di suatu wilayah secara terukur dan representatif. IPM juga dapat menjelaskan bagaimana keadaan masyarakat dapat mengakses pembangunan yang sudah dilakukan oleh pemerintah dalam memperoleh kesehatan, pendidikan, pengeluaran dan lain sebagainya. Gambar 1 berikut merupakan grafik pertumbuhan IPM Indonesia dari tahun 2015 hingga tahun 2022.



Gambar 1. Indeks Pembangunan Manusia Indonesia Tahun 2015-2022

Berdasarkan Gambar 1, terlihat bahwa IPM Indonesia terus mengalami peningkatan yang signifikan selama 8 tahun terakhir. Pada tahun 2015, IPM Indonesia yang hanya sebesar 69,55 poin terus meningkat hingga pada tahun 2022 angka IPM Indonesia berada pada angka 72,91 poin. Peningkatan IPM dari tahun ke tahun merupakan indikasi positif bahwa kualitas pembangunan manusia di Indonesia semakin membaik.

Walaupun terjadi peningkatan setiap tahun, IPM Indonesia masih ditemukan permasalahan yaitu ketidakmerataan IPM di setiap kabupaten/kota. Gambar 2 berikut menyajikan IPM kabupaten/kota yang memiliki IPM tertinggi dan terendah di Indonesia.



Gambar 2. Lima Kabupaten/Kota dengan IPM Tertinggi dan Terendah Tahun 2022

Berdasarkan grafik di atas, dimunculkan kabupaten/kota dengan IPM tertinggi dan kabupaten/kota dengan IPM terendah pada tahun 2022. Terlihat bahwa IPM tertinggi terdapat pada Kota Yogyakarta dengan IPM sebesar 87,69 poin, sedangkan IPM terendah terdapat pada Kabupaten Nduga sebesar 34,1 poin. Dilihat dari 5 kabupaten/kota yang memiliki IPM terendah, angka IPM masih berada sangat jauh di bawah rata-rata IPM Indonesia yaitu 72,91 poin. Rendahnya IPM di kabupaten/kota tersebut karena tidak meratanya pembangunan antar wilayah di Indonesia. Tidak meratanya pembangunan yang terjadi dipengaruhi oleh berbagai faktor, baik faktor yang berasal dari 3 aspek dasar seperti umur panjang dan hidup sehat, pengetahuan dan standar hidup layak maupun faktor-faktor lainnya.

Tinggi rendahnya IPM tidak terlepas dari program pembangunan yang dilaksanakan pemerintah pusat dan daerah untuk meningkatkan kesejahteraan masyarakat. Namun pelaksanaan program yang dilaksanakan masih belum efektif karena lemahnya koordinasi antar tingkat pemerintahan. Dikarenakan

pembangunan manusia melibatkan perubahan yang hasilnya terlihat dari waktu ke waktu, maka program pembangunan harus terus dilaksanakan dan dipantau secara terus-menerus. Selain itu, penentuan program pembangunan yang dilaksanakan harus tepat sasaran dan sejalan dengan prioritas wilayah berdasarkan kategori IPM wilayah kabupaten/kota (Kemala & Wijayanto, 2021).

Untuk menyikapi masalah tidak meratanya pembangunan antar wilayah, diperlukan sebuah sistem keputusan yang dapat dengan cepat dan akurat dalam menentukan kategori IPM di setiap kabupaten/kota, guna meningkatkan efektivitas kinerja pemerintah dalam menganalisis kategori IPM di setiap wilayah. Penentuan kategori IPM dapat dipermudah dengan penerapan metode klasifikasi. Hal ini juga dapat dijadikan acuan pemerintah agar bisa lebih memperhatikan dan mengupayakan program-program pembangunan manusia, untuk memaksimalkan tingkat pembangunan manusia di masa yang akan datang.

Klasifikasi merupakan proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (Han & Kamber, 2006). Pada penelitian ini metode klasifikasi yang digunakan berfokus pada perbandingan tingkat akurasi metode *Naïve Bayes Classifier* dan *K-Nearest Neighbors* (KNN) dengan menggunakan *confusion matrix*.

Naïve Bayes merupakan sebuah metode klasifikasi yang memanfaatkan teori probabilitas untuk memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Kelebihan klasifikasi *Naïve Bayes* yaitu

memiliki akurasi dan kecepatan yang tinggi apabila di aplikasikan ke dalam data yang besar. Selain itu, hanya memerlukan *training* data yang kecil untuk mengestimasi parameter yang dibutuhkan untuk klasifikasi (Kusrini & Luthfi, 2009).

K-Nearest Neighbors (KNN) adalah metode berbasis jarak yang mengklasifikasikan titik data baru berdasarkan mayoritas kategori atau nilai K tetangga terdekatnya. Metode KNN memiliki keunggulan yaitu paling umum, paling sederhana dan non parametrik yang tidak mendasarkan pada distribusi data tertentu. Selain itu, metode ini banyak digunakan peneliti karena mempunyai nilai akurasi yang tinggi, *insentive* terhadap *outlier* dan tidak ada asumsi terhadap data (Harrington, 2012).

Beberapa penelitian terdahulu mengenai *Naïve Bayes* dan KNN antara lain, Putri *et al* (2022) menerapkan klasifikasi *Naïve Bayes* dan KNN pada analisis data penyakit jantung. Dari penelitian ini, diketahui nilai akurasi pada klasifikasi *Naïve Bayes* yaitu sebesar 86,17%. Sedangkan nilai akurasi pada klasifikasi KNN dengan menggunakan nilai parameter K adalah 9 yaitu sebesar 85,11%. Kemudian Lubis (2021), melakukan klasifikasi *Decision Tree* dan *Naïve Bayes* pada Indeks Pembangunan Manusia di Provinsi Sumatera Utara. Dari penelitian ini, diketahui nilai akurasi *Naïve Bayes* yaitu sebesar 91%. Sedangkan nilai akurasi pada klasifikasi *Decision Tree* yaitu sebesar 88%.

Oleh karena itu, peneliti ingin melakukan sebuah penelitian dengan membandingkan metode *Naïve Bayes* dan KNN menggunakan data Indeks Pembangunan Manusia kabupaten/kota di Indonesia. Berdasarkan latar belakang

di atas, peneliti tertarik untuk melakukan penelitian dengan judul **“Perbandingan Metode *Naïve Bayes* dan *K-Nearest Neighbors* dalam Mengklasifikasikan Indeks Pembangunan Manusia menurut Kabupaten/Kota di Indonesia Tahun 2022”**.

B. Batasan Masalah

Berdasarkan latar belakang di atas, batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan dalam penelitian merupakan data Indeks Pembangunan Manusia kabupaten/kota di Indonesia tahun 2022.
2. Variabel yang digunakan berjumlah 5 variabel yaitu Indeks Pembangunan Manusia (Y), Angka Harapan Hidup (X_1), Rata-rata Lama Sekolah (X_2), Harapan Lama Sekolah (X_3), dan Pengeluaran per Kapita (X_4).
3. Analisis yang digunakan yaitu *Naïve Bayes* dan KNN.

C. Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, rumusan masalah yang terdapat pada penelitian ini adalah sebagai berikut.

1. Bagaimana klasifikasi Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia dengan menggunakan metode *Naïve Bayes* dan metode KNN?
2. Bagaimana perbandingan akurasi metode *Naïve Bayes* dan KNN dalam mengklasifikasikan data Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2022?

D. Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini sebagai berikut.

1. Untuk mengetahui klasifikasi Indeks Pembangunan Manusia (IPM) kabupaten/kota dengan menggunakan metode *Naïve Bayes* dan metode KNN.
2. Untuk mengetahui perbandingan akurasi metode *Naïve Bayes* dan KNN dalam mengklasifikasikan data Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2022.

E. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut.

1. Bagi penulis, penelitian ini dapat membantu menambah wawasan dan memperluas pengetahuan yang berhubungan dengan klasifikasi *Naïve Bayes* dan KNN.
2. Bagi Pemerintahan, hasil penelitian ini dapat digunakan jika terdapat data baru, bisa menjalankan hasil penelitian yang sudah tersedia.
3. Bagi pembaca, hasil penelitian ini dapat digunakan dalam pengembangan penelitian lebih lanjut, serta dapat menjadi referensi atau sumber baca.

BAB V

KESIMPULAN DAN SARAN

A. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan sebelumnya, maka dapat disimpulkan beberapa hal sebagai berikut.

1. Hasil pengklasifikasian IPM menurut kabupaten/kota di Indonesia, metode *Naïve Bayes* berhasil mengklasifikasikan 7 amatan dikategorikan sangat tinggi, 37 amatan dikategorikan tinggi, 44 amatan dikategorikan sedang, dan 8 amatan dikategorikan rendah. Pada metode KNN dengan $K = 7$, metode ini berhasil mengklasifikasikan 7 amatan dikategorikan sangat tinggi, 39 amatan dikategorikan tinggi, 43 amatan dikategorikan sedang, dan 6 amatan dikategorikan rendah.
2. Metode *Naïve Bayes* memiliki tingkat akurasi sebesar 93,2%, sedangkan metode KNN dengan $K = 7$, memiliki tingkat akurasi sebesar 92,23%. Berdasarkan nilai tersebut dapat dikatakan bahwa metode *Naïve Bayes* memiliki performa yang sangat baik dalam mengklasifikasikan IPM menurut kabupaten/kota di Indonesia tahun 2022.

B. SARAN

Berdasarkan penelitian yang dilakukan, adapun saran yang diberikan oleh penulis sebagai berikut.

1. Bagi pemerintah, dapat dimanfaatkan oleh Badan Pusat Statistik (BPS) untuk mengidentifikasi kabupaten/kota dengan klasifikasi IPM tertentu dengan cepat.

2. Bagi peneliti selanjutnya, diharapkan menggunakan metode yang klasifikasi lain seperti metode *Random Forest*, *Neural Network*, atau *Support Vector Machine*, untuk memprediksi kasus yang sama ataupun dengan kasus yang berbeda.

DAFTAR PUSTAKA

- Badan Pusat Statistik (BPS). 2022. "Indeks Pembangunan Manusia", diakses pada 15 Januari 2023 pukul 08.00.
- Badan Pusat Statistik (BPS). 2022. "Metode Baru Indeks Pembangunan Manusia", diakses pada 15 Januari 2023 pukul 09.00
- Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An Improved Method to Construct Basic Probability Based on the Confusion Matrix for Classification Problem. *Information Sciences*, 250-261.
- Gorunescu, F. (2011). *Data Mining: Concepts and Techniques*. Berlin: Springer.
- Han, J., & Kamber, M. (2006). *Data Mining Concepts and Techniques, second edition*. California: Morgan Kaufman.
- Han, J., Kamber, M., & Pei, J. (2006). *Data Mining: Concepts and Techniques*. Waltham: Morgan Kauffman.
- Hardani, d. (2020). *Metode Penelitian Kualitatif & Kuantitatif*. Yogyakarta: CV Pustaka Ilmu.
- Harrington, P. (2012). *Machine Learning in Action*. United States of America: Manning Publications Co.
- Irina, F. (2017). *Metode Penelitian Terapan*. Yogyakarta: Parama Ilmu.
- Isman, Andani, A., & Latief, A. (2021). Perbandingan Metode KNN dan LBPH pada Klasifikasi Daun Herbal. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(3), 557-564.
- Jananto, A. (2013). Algoritma Naïve Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Jurnal Teknologi Informasi DINAMIK*, 18(1), 9-16.
- Jiang, L., Cai, Z., & Wang, D. (2010). Improving naive bayes for classification. *International Journal of Computers and Applications*, 32(3), 328-332.
- Kusrini & Luthfi, E.T. (2009). *Algoritma Data Mining*. Yogyakarta: Andi Offset.
- Larose, D. T. (2005). *Discovering Knowledge In Data: An Introduction to Data*. United State of America: John Wiley & Sons.
- Lubis, R. M. (2021). Analisis Metode Decision Tree Algorithm Dan Metode Naive Bayes Mengidentifikasi Pertumbuhan Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara. *JUTISAL (Jurnal Teknik Informatika Komputer Universal)*, 1(2), 9-16.

- Mustika, & dkk. (2021). *Data Mining dan Aplikasinya*. Bandung: Widina Bhakti Persada Bandung.
- Musu, W., Ibrahim, A., & Heriadi. (2021). Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5. *Prosiding Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 10(1), 186-195.
- Nair, A., Kuban, B. D., Obuchowski, N., & Vince, D.G. (2001). Assessing Spectral Algorithms to Predict Atherosclerotic Plaque Composition with Normalized and Raw Intravascular Ultrasound Data. *Ultrasound in Medicine and Biology*, 27(10), 1319-1331.
- Putri, R. W., Ristyawan, A., & Muzaki, M. N. (2022). Perbandingan Kinerja Algoritma K-NN dan NBC Untuk Klasifikasi Penyakit Jantung. *JTECS : Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem & Komputer*, 2(2), 143-154.
- Soofi, A. A., & Awan, A. (2017). Classification Techniques in Machine Learning: Applications and Issues. *Journal of Basic & Applied Sciences*, 13, 459-465.
- Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
- Suntoro, J. (2019). *Data Mining : Algoritma dan Implementasi dengan Pemrograman PHP*. Jakarta: Elex Media Komputindo.
- UNDP. (1990). *Global Human Development Report*. New York: Human Resources Department.
- Wu, X., & Kumar, V. (2009). *The Top Ten Algorithms in Data Mining*. London: CRS Press Taylor & Francis Group.