

**TUGAS AKHIR**

**PERBANDINGAN PERFORMA *NAÏVE BAYES* & *K-NEAREST NEIGHBOR* DALAM MENGLASIFIKASI STATUS PASIEN *CORONAVIRUS DISEASE* (*COVID-19*)**



**DANANG WICAKSANA**

**17102005**

**PROGRAM STUDI S1 INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2021**

**TUGAS AKHIR**

**PERBANDINGAN PERFORMA *NAÏVE BAYES* & *K-NEAREST NEIGHBOR* DALAM MENGLASIFIKASI STATUS PASIEN *CORONAVIRUS DISEASE* (*COVID-19*)**

***COMPARISON OF NAÏVE BAYES & K-NEAREST NEIGHBOR PERFORMANCE IN CLASSIFICATION CORONAVIRUS DISEASE PATIENT STATUS (COVID-19)***

Disusun Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Komputer



**DANANG WICAKSANA**

**17102005**

**PROGRAM STUDI S1 INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2021**

**HALAMAN PERSETUJUAN PEMBIMBING**

**PERBANDINGAN PERFORMA *NAÏVE BAYES* & *K-NEAREST NEIGHBOR* DALAM MENINGKLASIFIKASI STATUS PASIEN *CORONAVIRUS DISEASE* (*COVID-19*)**

***COMPARISON OF NAÏVE BAYES & K-NEAREST NEIGHBOR PERFORMANCE IN CLASSIFICATION CORONAVIRUS DISEASE PATIENT STATUS (COVID-19)***

Dipersiapkan dan Disusun Oleh:

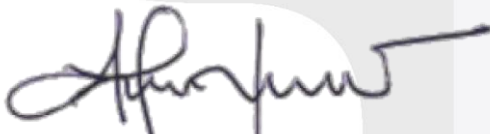
**DANANG WICAKSANA**

**17102005**

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas Akhir  
Pada hari Selasa, 02 Maret 2021

Pembimbing I,

Pembimbing II,



**(Apri Junaidi, S.Kom., M.Kom., MCS)**  
**NIDN. 0407047403**



**(Amalia Beladina Arifa, S.Pd., M.Cs)**  
**NIDN. 0606019201**

Tugas Akhir ini diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 02 Maret 2021

Kaprodi,



**(Fahrudin Mukti Wibowo, S. Kom., M.Eng)**  
**NIDN. 0622018403**

**HALAMAN PENETAPAN PENGUJI**

**PERBANDINGAN PERFORMA *NAÏVE BAYES* & *K-NEAREST NEIGHBOR* DALAM MENINGKLASIFIKASI STATUS PASIEN *CORONAVIRUS DISEASE* (*COVID-19*)**

***COMPARISON OF NAÏVE BAYES & K-NEAREST NEIGHBOR PERFORMANCE IN CLASSIFICATION CORONAVIRUS DISEASE PATIENT STATUS (COVID-19)***

Dipersiapkan dan Disusun Oleh:

**DANANG WICAKSANA**

**17102005**

**Tugas Akhir Telah diuji dan Dinilai Panitia Penguji Program**

**Studi S1 Informatika**

**Fakultas Informatika**

**Institut Teknologi Telkom Purwokerto**

**Pada Tanggal: 02 Maret 2021**

**Ketua  
Penguji**

**(Tri Ginanjar Laksana, S.Kom., M.C.S., M.Kom.)**

**NIDN. 0407088502**

**Anggota  
Penguji I**

**Anggota  
Penguji II**

**(Tri Ginanjar Laksana, S.Kom., M.C.S., M.Kom.)**

**NIDN. 0407088502**

**(Atika Ratna Dewi, S.Si., M.Sc)**

**NIDN. 0615128703**

## HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

### HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama Mahasiswa : **Danang Wicaksana**  
NIM : **17102005**  
Program Studi : **SI Informatika**

Menyatakan bahwa Tugas Akhir dengan judul berikut:

**PERBANDINGAN PERFORMA *NAÏVE BAYES* & *K-NEAREST NEIGHBOR* DALAM MENGLASIFIKASI STATUS PASIEN *CORONAVIRUS DISEASE (COVID-19)***

Dosen Pembimbing Utama : **Apri Junaidi, S.Kom., M.Kom., MCS**  
Dosen Pembimbing Pendamping : **Amalia Beladinna Arifa, S.Pd., M.Cs**

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

**Bekasi, 11 Februari 2021,**  
**Yang Menyatakan,**



**(Danang Wicaksana)**

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT, karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan laporan Tugas Akhir ini dengan baik.

Selama penulis mengerjakan hingga menyelesaikan laporan Tugas Akhir selama dua semester telah berjalan dengan lancar sesuai harapan yang diinginkan. Oleh karena itu, dengan segala kerendahan hati pada kesempatan ini penulis ingin mengucapkan rasa terima kasih kepada:

1. Kedua orang tua yang telah memberikan doa dan dukungan selama proses pembuatan laporan.
2. Bapak Apri Junaidi, S.Kom., M.Kom., MCS. selaku pembimbing satu yang selalu memberikan arahan dan sabar dalam melakukan bimbingan.
3. Ibu Amalia Beladinna Arifa, S.Pd., M.Cs. selaku pembimbing dua yang selalu memberikan arahan dan sabar dalam melakukan bimbingan.
4. Bapak Dr. Ali Rokhman, M.Si selaku Rektor Institut Teknologi Telkom Purwokerto.
5. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Dekan Fakultas Informatika Institut Teknologi Telkom Purwokerto.
6. Bapak Fahrudin Mukti Wibowo, S.Kom., M.Eng. selaku Kepala Program Studi S1 Informatika Institut Teknologi Telkom Purwokerto
7. Para Dosen Program Studi S1 Informatika Institut Teknologi Telkom Purwokerto yang telah memberikan bekal ilmu kepada penulis.
8. Rekan seperjuangan, grup *Ga Jelas* yang telah banyak membantu perihal tugas akhir, baik masukan, kritikan, maupun ancaman.
9. Rekan penyemangat, Kanjeng Shinta yang selalu melarang bermain game sampai ada *progress* Tugas Akhir.
10. Semua Pihak yang tidak dapat disebutkan satu persatu.

Dalam penyusunan laporan Tugas Akhir ini penulis menyadari bahwa terdapat banyak kekurangan dalam penulisan dan penyajian yang telah dituangkan dalam tulisan ini, untuk itu penulis mengharapkan saran dan kritik dari pembaca untuk membangun dan menjadi masukan penulis untuk kedepannya

Bekasi, 19 Januari 2021

DANANG WICAKSANA

## DAFTAR ISI

<b>HALAMAN PERSETUJUAN PEMBIMBING .....</b>	<b>iii</b>
<b>HALAMAN PENETAPAN PENGUJI.....</b>	<b>iv</b>
<b>HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR .....</b>	<b>v</b>
<b>KATA PENGANTAR .....</b>	<b>vi</b>
<b>DAFTAR ISI .....</b>	<b>viii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xiii</b>
<b>DAFTAR GLOSARIUM.....</b>	<b>xiv</b>
<b>ABSTRAK .....</b>	<b>xvi</b>
<b>BAB I.....</b>	<b>1</b>
<b>PENDAHULUAN .....</b>	<b>1</b>
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	4
1.3    Pertanyaan Penelitian .....	4
1.4    Tujuan Penelitian.....	4
1.5    Batasan Masalah.....	5
1.6    Manfaat Penelitian.....	5
<b>BAB II.....</b>	<b>6</b>
<b>TINJAUAN PUSTAKA .....</b>	<b>6</b>
2.1    Tinjauan Pustaka .....	6
2.2    Landasan Teori .....	13
2.2.1    Pengertian <i>COVID-19</i> .....	13
2.2.2    Pengertian Model .....	13
2.2.3    Pengertian <i>Data Mining</i> .....	13
2.2.4    Pengertian Klasifikasi .....	14
2.2.5    Pengertian <i>Python</i> .....	15
2.2.6    Pengertian <i>Matplotlib</i> .....	15
2.2.7    Pengertian <i>Scikit-Learn</i> .....	15
2.2.8    Pengertian <i>Pandas</i> .....	15
2.2.9    Pengertian <i>Numpy</i> .....	16



2.2.10	Pengertian <i>Naïve Bayes</i> .....	16
2.2.11	Pengertian <i>K-Nearest Neighbor</i> .....	16
2.2.12	<i>Confusion Matrix</i> .....	17
<b>BAB III .....</b>		<b>19</b>
<b>METODELOGI PENELITIAN .....</b>		<b>19</b>
3.1	Subjek dan Objek Penelitian .....	19
3.2	Alat dan Bahan .....	19
3.2.1	Perangkat Keras .....	19
3.2.2	Perangkat Lunak .....	19
3.2.3	Dataset.....	20
3.3	Diagram Alur Penelitian.....	20
3.3.1	Identifikasi dan merumuskan masalah.....	21
3.3.2	Studi literatur .....	21
3.3.3	Pengumpulan Data .....	21
3.3.4	<i>Data Preprocessing</i> .....	25
3.3.5	Membuat Model.....	29
3.3.6	Pengujian.....	31
3.3.6.1	Perhitungan <i>Naïve Bayes</i> .....	31
3.3.6.2	Perhitungan <i>K-Nearest Neighbor</i> .....	37
<b>BAB IV .....</b>		<b>41</b>
<b>HASIL PENELITIAN DAN ANALISIS .....</b>		<b>41</b>
4.1	Hasil Klasifikasi <i>Naïve Bayes</i> Menggunakan <i>Python</i> .....	41
4.2	Hasil Klasifikasi <i>K-Nearest Neighbor</i> Menggunakan <i>Python</i> .....	44
4.3	Hasil Perbandingan Performa <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i> ....	47
<b>BAB V .....</b>		<b>50</b>
<b>KESIMPULAN DAN SARAN .....</b>		<b>50</b>
5.1	Kesimpulan.....	50
5.2	Saran.....	50
<b>DAFTAR PUSTAKA .....</b>		<b>52</b>
<b>LAMPIRAN .....</b>		<b>58</b>

## DAFTAR GAMBAR

Gambar 1.1. Grafik kasus <i>COVID-19</i> Global 18 Juli 2020.....	2
Gambar 1.2. Grafik kasus <i>COVID-19</i> Indonesia 18 Juli 2020.....	2
Gambar 3.1. Diagram alur penelitian .....	20
Gambar 3.2. Contoh data yang belum dilakukan <i>cleansing</i> .....	25
Gambar 3.3. Contoh data yang sudah dilakukan <i>cleansing</i> .....	25
Gambar 3.4. Contoh data yang belum dilakukan <i>integration</i> .....	26
Gambar 3.5. Contoh data yang sudah dilakukan <i>integration</i> .....	26
Gambar 3.6. Contoh data yang belum dilakukan <i>selection</i> .....	27
Gambar 3.7. Contoh data yang sudah dilakukan <i>selection</i> .....	27
Gambar 3.8. Cara kerja algoritma <i>Naïve Bayes</i> .....	29
Gambar 3.9. Cara kerja algoritma <i>K-Nearest Neighbor</i> .....	30
Gambar 4.1. Klasifikasi <i>Naïve Bayes</i> Menggunakan Python.....	41
Gambar 4.2. <i>Confusion Matrix Naïve Bayes</i> .....	41
Gambar 4.3. Klasifikasi <i>K-NN</i> Menggunakan Python .....	44
Gambar 4.4. <i>Confusion Matrix K-NN</i> .....	44
Gambar 4.5. <i>Chart</i> Performa Algoritma .....	47
Gambar 4.6. Visualisasi Akurasi <i>K-NN</i> berdasarkan K1-K15 .....	48

## DAFTAR TABEL

Tabel 2.1. Penelitian Terdahulu .....	8
Tabel 2.2. Model <i>Confusion Matrix</i> .....	17
Tabel 3.1. Contoh dataset pasien Indonesia.....	22
Tabel 3.2 Contoh dataset pasien India .....	23
Tabel 3.3. Inisialisasi Jenis Kelamin/ <i>Gender</i> .....	28
Tabel 3.4. Inisialisasi Kebangsaan/ <i>Nationality</i> .....	28
Tabel 3.5. Inisialisasi Status Pasien .....	28
Tabel 3.6. Hasil Preprocessing Inisialisasi.....	28
Table 3.7. Probabilitas Kelas .....	31
Tabel 3.8. Atribut Jenis Kelamin .....	32
Tabel 3.9. Atribut Umur.....	32
Tabel 3.10. Atribut Kebangsaan.....	34
Tabel 3.11. Contoh Pasien .....	34
Tabel 3.12. Tabel Ranking .....	37
Tabel 3.13. Contoh Kasus Perhitungan <i>K-Nearest Neighbor</i> .....	37
Tabel 3.14. Perhitungan Distance <i>K-Nearest Neighbor</i> .....	38
Tabel 3.15. Urutan <i>Distance</i> Perhitungan <i>K-Nearest Neighbor</i> .....	38
Tabel 3.16. Prediksi Nilai K Dari <i>K-Nearest Neighbor</i> .....	39
Tabel 3.17. Contoh Soal Dari <i>K-Nearest Neighbor</i> .....	40
Tabel 4.1. Perbandingan performa <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i> .....	47
Tabel 4.2. Jumlah K pada <i>KNN</i> .....	48

## DAFTAR LAMPIRAN

Lampiran 1 Dataset Raw Individual Detail Patient COVID-19 Indonesia.....	58
Lampiran 2 Dataset Raw Individual Detail Patient COVID-19 India .....	60
Lampiran 3 Klasifikasi Naïve Bayes .....	62
Lampiran 4 Klasifikasi K-Nearest Neighbor .....	64

## DAFTAR GLOSARIUM

<i>Class</i>	: Nilai yang digunakan untuk label
<i>Test Set</i>	: Dataset yang digunakan untuk menguji model setelah dilakukan proses pemeriksaan pada validasi set
<i>Training Set</i>	: Dataset yang digunakan untuk melatih model
<i>Confusion Matrix</i>	: Suatu metode yang digunakan untuk melakukan perhitungan akurasi pada data mining
Akurasi	: Dimana rasio dari data yang diprediksi benar (positif dan negatif) dengan keseluruhan data
<i>Precision</i>	: Dimana rasio data prediksi benar positif dibandingkan dengan keseluruhan hasil data yang diprediksi positif.
<i>Recall</i>	: Dimana rasio prediksi data benar positif dibandingkan dengan keseluruhan data yang benar diprediksi positif.
TP ( <i>True</i> Positif)	: Dimana data diprediksi benar (Positif) dan memang benar ( <i>True</i> ).
TN ( <i>True</i> Negatif)	: Dimana data diprediksi salah (Negatif) dan memang salah ( <i>True</i> ).
FP ( <i>False</i> Positif)	: Dimana data diprediksi benar (Positif) dan ternyata salah ( <i>False</i> ).
FN ( <i>False</i> Negatif)	: Dimana data diprediksi salah (Negatif) dan ternyata benar ( <i>True</i> ).
<i>Naïve Bayes</i>	: Pengklasifikasian dengan metode probabilitas dan statistik yang memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai teorema bayes.

- K-Nearest Neighbor (K-NN)* : Algoritma yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) suatu data dengan data yang lain.
- Data Cleaning* : Untuk membersihkan nilai kosong, tidak konsisten atau tupel kosong (*missing value* dan *noisy*).
- Data Selection* : Untuk menyeleksi data karena tidak semuanya akan digunakan, maka dari itu hanya beberapa data yang akan diambil dan sesuai untuk dianalisis.
- Data Integration* : Untuk menggabungkan/menyatukan data dari berbagai database yang berkorelasi ke dalam satu *database* baru.
- Data Transformation* : Untuk mengubah data atau digabung ke dalam bentuk yang tepat untuk diproses dalam *data mining*.
- Scikit-Learn* : Sebuah *library* yang terdapat di dalam *python*. *Scikit-Learn* dibangun di atas beberapa data umum dan pustaka matematika *python*.
- Numpy* : *Library* pada *python* yang berfungsi melakukan proses komputasi numerik.
- Pandas* : *Library* analisis data dari *python* yang memiliki struktur data yang kita perlukan untuk membersihkan data mentah ke dalam sebuah bentuk yang cocok untuk analisis (yaitu tabel).