

**PERBANDINGAN AKURASI DAN WAKTU PROSES ALGORITMA
NAÏVE BAYES DENGAN ALGORITMA SUPPORT VECTOR MACHINE
DALAM ANALISIS SENTIMEN TWITTER APLIKASI PEDULI
LINDUNGI**

SKRIPSI



disusun oleh

Maulana Rizky Hidayat

18.11.1816

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

**PERBANDINGAN AKURASI DAN WAKTU PROSES ALGORITMA
NAÏVE BAYES DENGAN ALGORITMA SUPPORT VECTOR MACHINE
DALAM ANALISIS SENTIMEN TWITTER APLIKASI PEDULI
LINDUNGI**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Sistem Informasi



disusun oleh

Maulana Rizky Hidayat

18.11.1816

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

PERSETUJUAN

SKRIPSI

**PERBANDINGAN AKURASI DAN WAKTU PROSES ALGORITMA
NAÏVE BAYES DENGAN ALGORITMA SUPPORT VECTOR MACHINE
DALAM ANALISIS SENTIMEN TWITTER APLIKASI PEDULI
LINDUNGI**

yang dipersiapkan dan disusun oleh

Maulana Rizky Hidayat

18.11.1816

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 2 Oktober 2021

Dosen Pembimbing,

Mulla Sulstiyono, M.Kom
NIK. 190302248

PENGESAHAN

SKRIPSI

**PERBANDINGAN AKURASI DAN WAKTU PROSES ALGORITMA
NAÏVE BAYES DENGAN ALGORITMA SUPPORT VECTOR MACHINE
DALAM ANALISIS SENTIMEN TWITTER APLIKASI PEDULI
LINDUNGI**

yang dipersiapkan dan disusun oleh

Maulana Rizky Hidayat

18.11.1816

telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Februari 2022

Susunan Dewan Penguji

Nama Penguji

Kusnawi, S.Kom, M. Eng.
NIK. 190302112

Anggit Dwi Hartanto, M.Kom
NIK. 190302163

Mulla Sulistyono, M.Kom
NIK. 190302248

Tanda Tangan

Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Mei 2022

DEKAN FAKULTAS ILMU KOMPUTER

HANIF AL FATTAH, M.KOM.
NIK. 190302096

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 19 February 2022



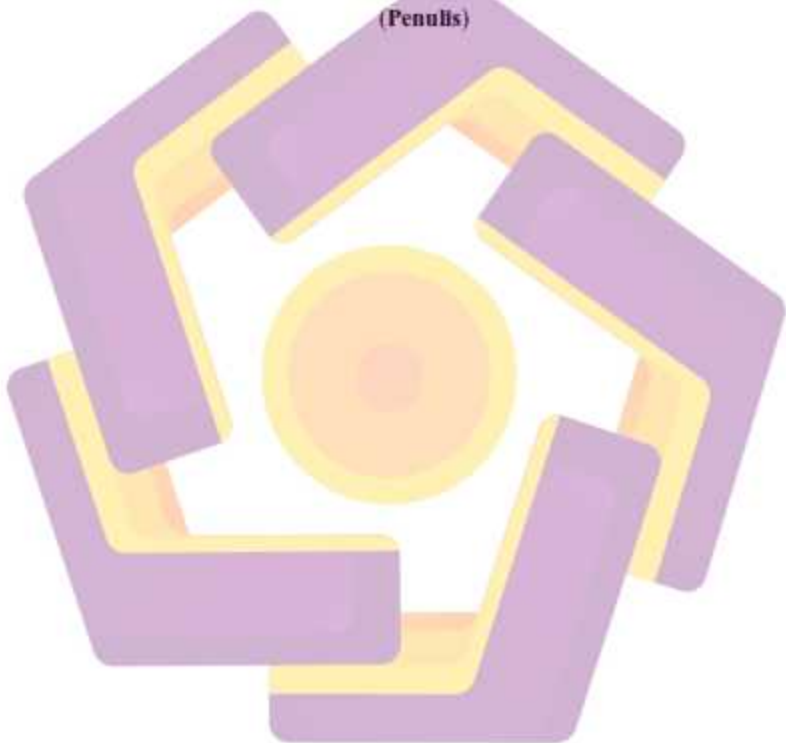
Maulana Rizky Hidayat

NIM. 18.11.1816

MOTTO

"Jangan Membandingkan Dirimu Dengan Orang Lain Karena Semua Punya
Proses Dan Waktu Masing Masing"

(Penulis)



KATA PENGANTAR

Puji dan syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir / Skripsi dengan topic penelitian “Perbandingan Akurasi Dan Waktu Proses Algoritma Naïve Baiyes Dengan Algoritma Support Vector Machine Dalam Analisis Sentimen Twitter Aplikasi Peduli Lindungi” ini sebagai salah satu syarat dalam menyelesaikan program S-1 Teknik Informatika. Di Fakultas Ilmu Komputer Universitas Amikom Yogyakarta. Penulis mengucapkan terima kasih kepada pihak pihak yang telah membantu dalam penyelesaian Tugas Akhir ini, terutama kepada:

1. Teristimewa kepada kedua orang tua dan seluruh keluarga besar yang telah banyak memberikan dukungan moral dan material.
2. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta
3. Ibu Windha Mega Pradnya D, M.Kom selaku ketua jurusan S1 Informatika Universitas Amikom Yogyakarta
4. Bapak Mulia Sulistiyono, M.Kom selaku Pembimbing Tugas Akhir
5. Seluruh Dosen dan Staff Jurusan Teeknik Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta
6. Dan semua pihak yang telah mensupport saya yang tidak bisa saya sebutkan satu persatu. Serta teman seperjuangan IF 01 2018.
7. Teman saya Patriot Kusuma Sejati, Dede Permana. Yang telah mensupport skripsi saya.

Penulis menyadari bahwa dalam penulisan ini yang jauh dari sempurna, baik dari segi penyusunan, bahasa, ataupun penulisannya. Oleh karena itu penulis mengharapkan kritik dan saran yang sifatnya membangun.

Akhir kata penulis mengucapkan terima kasih dan semoga Tugas Akhir ini bermanfaat terutama untuk penulis sendiri, perusahaan serta bagi yang membaca.

Yogyakarta, 2 Februari 2022

Penulis

DAFTAR ISI

JUDUL.....	I
PERSETUJUAN	II
PENGESAHAN	III
PERNYATAAN	IV
MOTTO	VI
KATA PENGANTAR.....	VII
DAFTAR ISI.....	VIII
DAFTAR TABEL.....	X
ABSTRACT.....	XII
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH	1
1.3 BATASAN MASALAH.....	4
1.4 MAKSUD DAN TUJUAN PENELITIAN	4
1.5 MANFAAT PENELITIAN.....	4
1.6 METODE PENELITIAN	4
1.7 SISTEMATIKA PENULISAN.....	5
BAB II LANDASAN TEORI.....	7
2.1 KAJIAN PUSTAKA.....	7
2.2 DASAR TEORI.....	11
2.2.1 Data Mining	11
2.2.2 Sentiment Analysis.....	11
2.2.3 Twitter.....	11
2.3 METODE ANALISIS.....	12
2.3.1 Text Mining.....	12
2.3.2 Preprocessing	12

2.3.3 Naive Bayes Classifier.....	12
2.3.4 Support Vector Machine.....	13
2.3.5 Evaluasi.....	14
2.3.6 Validasi.....	16
BAB III METODE PENELITIAN.....	17
3.1 ALAT DAN BAHAN PENELITIAN.....	17
3.2 ALUR PENELITIAN.....	17
3.2.1 Pengumpulan Data.....	19
3.2.2 Text Preprocessing.....	19
3.2.3 Pembobotan Term.....	20
3.2.4 Klasifikasi.....	20
3.2.5 Validasi dengan K-FOLD CROSS Validation.....	23
3.2.6 Evaluasi.....	23
BAB IV HASIL DAN PEMBAHASAN.....	24
4.1 IMPLEMENTASI PENGAMBILAN DATA & TEXT PREPROCESSING.....	24
4.2 IMPLEMENTASI ALGORITMA KLASIFIKASI SERTA PENGUJIAN.....	26
4.3 PERBANDINGAN ALGORITMA KLASIFIKASI.....	31
BAB V PENUTUP.....	33
5.1 KESIMPULAN.....	33
5.2 SARAN.....	34
DAFTAR PUSTAKA.....	35

DAFTAR TABEL

Tabel 2. 1 Matrik Literatur Review dan Posisi Penelitian.....	9
Tabel 4. 1 Confusion Matrix Klasifikasi Naïve Bayes Dengan Split 8020.....	26
Tabel 4. 2 Perhitungan Klasifikasi Naive Bayes Dengan Split 8020.....	26
Tabel 4. 3 Hasil Pengujian Naïve Bayes	27
Tabel 4. 4 Hasil Pengujian K Fold Cross Validation	28
Tabel 4. 5 Waktu Proses Naïve Bayes	28
Tabel 4. 6 Hasil Klasifikasi SVM Dengan Split 8020	29
Tabel 4. 7 Perhitungan Klasifikasi SVM Dengan Split 8020	29
Tabel 4. 8 Hasil Pengujian Naïve Bayes	30
Tabel 4. 9 Hasil Pengujian K Fold Croos.....	31
Tabel 4. 10 Waktu Proses SVM.....	31
Tabel 4. 11 Hasil Perbandingan Algoritma SVM.....	31

DAFTAR GAMBAR

Gambar 2 . 1 Ilustrasi SVM [2].....	14
Gambar 2 . 2 Confusion Matrix (sumber : https://ksnugroho.medium.com/)	15
Gambar 3 . 1 WorkFlow Penelitian.....	19
Gambar 3 . 2 Diagram alir tahapan Naïve Bayes.....	21
Gambar 3 . 3 Diagram Alir Support Vector Machine	23
Gambar 4 . 1 Dataset Twitter.....	24
Gambar 4 . 2 Tahap Text Preprocessing.....	25



INTISARI

Aplikasi Peduli Lindungi merupakan program pemerintah untuk memantau serta menanggulangi lonjakan Covid-19 di Indonesia. Namun aplikasi ini juga menuai kontroversi di masyarakat sehingga banyak mengundang berbagai macam opini di kalangan masyarakat salah satunya di Twitter. Twitter telah menjadi media berbasis web dan smartphone yang paling populer dengan lebih dari 200 juta pengguna dan lebih dari 10,6 miliar tweet yang telah dihasilkan, dari informasi tersebut dapat digunakan untuk menganalisis pendapat dan opini masyarakat terhadap aplikasi Peduli Lindungi melalui analisis sentimen dengan menggunakan klasifikasi pendapat ke dalam 2 kelas yaitu positif dan negatif. Analisis sentimen diperlukan untuk mengetahui penilaian umum pada suatu objek.

Algoritma support vector machine mendefinisikan hyperplane yang baik menjadikan algoritma SVM memiliki tingkat akurasi yang baik dibandingkan algoritma lainnya, sedangkan algoritma Naive Bayes merupakan algoritma klasifikasi machine learning dengan penalaran probabilitas yang tak kalah baik dengan algoritma lainnya.

Dari hasil penelitian ini akurasi menunjukkan bahwa metode Naive Bayes lebih unggul dengan 80% tanpa menggunakan pengujian validasi menggunakan K-Fold Cross Validation dan 85% dengan menggunakan K-Fold Cross Validation pada fold - 3. Sedangkan perhitungan waktu proses Naive Bayes juga unggul dengan mendapatkan waktu 0.009365 detik, namun algoritma Support Vector Machine dilakukan pengujian K-Fold – Cross Validation mendapatkan 86% pada iterasi ke 3.

Kata Kunci: Machine Learning, Naive Bayes, Support Vector Machine, Validasi

ABSTRACT

The Peduli Protect application is a government program to monitor and cope with the spike in Covid-19 in Indonesia. but this application also reaps controversy in the community so that it invites various kinds of opinions among the community, one of which is on Twitter. Twitter has become the most popular web and smartphone-based media with more than 200 million users and more than 10.6 billion tweets that have been generated, from this information can be used to analyze public opinion and opinion on the Peduli Protect application through sentiment analysis using classification income into 2 classes, namely positive and negative. Sentiment analysis is needed to find out the general assessment of an object.

The support vector machine algorithm defines a good hyperplane that makes the SVM algorithm has a good level of accuracy compared to other algorithms, while the Naive Bayes algorithm is a machine learning classification algorithm with probability reasoning that is not inferior to other algorithms.

From the results of this study, the accuracy shows that the Naïve Bayes method is superior by 80% without using validation testing using K - Fold Cross-Validation and 85% by using K-Fold Cross Validation on the fold - 3. While the calculation of Naïve Bayes processing time is also superior by the getting time of 0.009365 seconds the Support Vector Machine algorithm is tested for K-Fold - Cross Validation getting 86% in the 3rd iteration.

Keywords: *Machine Learning, Nave Bayes, Support Vector Machine, Validation*