

**PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DAN SUPPORT
VECTOR MACHINE PADA ANALISIS SENTIMEN TWITTER**

(Studi kasus : Data Tweet Dengan Kata Kunci “Omicron”)

SKRIPSI

Untuk Memenuhi Sebagian Persyaratan Mencapai Derajat Sarjana S-1

Program Studi Teknik Informatika



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Disusun Oleh:

Ahabullah Fakhri Muhammad
18106050040

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UIN SUNAN KALIJAGA YOGYAKARTA**

2022



PENGESAHAN TUGAS AKHIR

Nomor : B-1498/Un.02/DST/PP.00.9/07/2022

Tugas Akhir dengan judul : PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TWITTER (Studi kasus : Data Tweet Dengan Kata Kunci [Omicron])

yang dipersiapkan dan disusun oleh:

Nama : AHABULLAH FAKHRI MUHAMMAD
Nomor Induk Mahasiswa : 18106050040
Telah diujikan pada : Kamis, 23 Juni 2022
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Muhammad Didik Rohmad Wahyudi, S.T., MT.
SIGNED

Valid ID: 62da3f734662d



Penguji I

Ir. Maria Ulfah Siregar, S.Kom., MIT., Ph.D.
SIGNED

Valid ID: 62d9ee1e8af43



Penguji II

Agus Mulyanto, S.Si., M.Kom.
SIGNED

Valid ID: 62bbea7578812



Yogyakarta, 23 Juni 2022
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi
Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 62da44d4f1307



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi
Lamp :

Kepada
Yth. Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga Yogyakarta
di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Ahabullah Fakhri Muhammad
NIM : 18106050040
Judul Skripsi : Perbandingan Metode Naive Bayes Classifier dan Support Vector Machine pada Analisis Sentimen Twitter (Studi Kasus : Data Tweet dengan Kata Kunci : "Omicron")

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 16 Juni 2022
Pembimbing

Muhammad Didik Rohmad Wahyudi, S.T., MT
NIP. 19760812 200901 1 015

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini:

Nama : Ahabullah Fakhri Muhammad

NIM : 18106050040

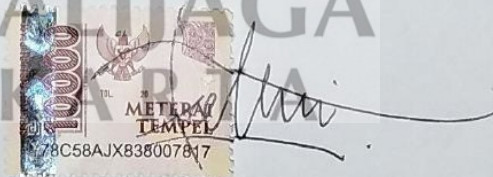
Program Studi : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul **“PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TWITTER (Studi kasus : Data Tweet Dengan Kata Kunci “Omicron”) ”** merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah diajukan untuk memperoleh gelar kesarjana di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Yogyakarta, 17 Juni 2022



Ahabullah Fakhri Muhammad
NIM. 18106050040

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Allah Swt. atas ridhonya saya dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang saya ajukan adalah “ Perbandingan Metode *Naive Bayes Classifier* dan *Support Vector Machine* Pada Analisis Sentimen Twitter (Studi Kasus : Data Tweet Dengan Kata Kunci ‘ Omicron ‘)” dapat terselesaikan dengan baik. Tidak dapat disangkal bahwa butuh usaha yang keras dalam penyelesaian pengerjaan skripsi ini. Namun, skripsi ini tidak akan selesai tanpa orang-orang tercinta di sekeliling saya yang telah membantu dan mendukung. Terima kasih saya sampaikan kepada :

1. Bapak Prof. Dr.Phil. Al Makin, S.Ag., M.A., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Ibu Dr. Dra. Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Ibu Ir.Maria Ulfah Siregar, S.Kom., MIT., Ph.D., selaku Ketua Program Studi S1 Teknik Informatika.
4. Bapak Muhammad Didik Rohmad Wahyudi, S.T., MT., selaku Dosen Pembimbing yang telah sabar membimbing, mengarahkan, dan memberi nasihat selama penyusunan skripsi.
5. Kedua Orang tua saya Bapak Ir. Susanto, Ibu Prof Alimatul Qibtiyah, S.Ag., M.Si., Ph.D. dan keluarga yang telah memberikan doa dan dukungan selama proses pembuatan skripsi.

6. Ibu Febriyanti Lestari, MA selaku dosen ahli bahasa yang telah bersedia untuk membantu dalam memvalidasi data.
7. Bapak-Ibu Dosen Teknik Informatika UIN Sunan Kalijaga yang telah memberikan banyak ilmu bagi penulis.
8. Seluruh staf dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
9. Staf dan karyawan Perpustakaan UIN Sunan Kalijaga Yogyakarta yang telah membantu dalam pencarian referensi penulisan skripsi.
10. Teman-teman seperjuangan Teknik Informatika 2018 UIN Sunan Kalijaga Yogyakarta.
11. Teman-teman tongkrongan Afifi, Aman, Fajar, Agive, Alim, Resa, Ambar, Rita, Uqi, Nawwab, Irfan, Naufan
12. Teman-teman kontrakan dan UKM Basket.
13. Teman-teman KKN Kelompok 169 Abhinaya Muda.
14. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu memberikan dukungan.

Penulis menyadari masih banyak kekurangan dalam skripsi ini, sehingga kritik dan saran yang membangun dari pembaca sangat penulis harapkan. Semoga skripsi ini dapat memberikan manfaat untuk mendorong penelitian penelitian selanjutnya.

Yogyakarta, Maret 2022

Penyusun

Ahabullah Fakhri Muhammad

18106050040



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Skripsi ini penulis persembahkan untuk almamater, kedua orang tua, keluarga, dan teman-teman seperjuangan yang selalu memberikan support dan semangat kepada penulis serta semua pihak yang turut andil dalam proses penulisan ini yang tidak dapat disebutkan satu persatu.

*Terimakasih
Semuanya.*



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN MOTTO

“Always think that something good is about happen”

-Ahabullah Fakhri M -

“A little progress each day is better than no progress at all ”



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

PENGESAHAN TUGAS AKHIR.....	i
SURAT PERSETUJUAN TUGAS AKHIR.....	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
KATA PENGANTAR	iv
HALAMAN PERSEMBAHAN	vii
HALAMAN MOTTO	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	4
1.5 Manfaat Penelitian	4
1.6 Kebaruan Penelitian.....	5
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Tinjauan Pustaka.....	7
2.2 LANDASAN TEORI.....	14
2.2.1 Data Mining	14
2.2.2 Machine Learning.....	15
2.2.3 Sentimen Analisis	16
2.2.4 Klasifikasi	17
2.2.5 Pembobotan TF-IDF.....	17
2.2.6 Naive Bayes	19
2.2.7 Support Vector Machine.....	20

2.2.8 Twitter.....	22
2.2.9 Python.....	24
BAB III METODE PENELITIAN	26
3.1 Metode Penelitian	26
3.2 Tahap-tahap penelitian.....	26
3.3 Studi Literatur	28
3.4 Pengumpulan Data.....	28
3.5 Preprocessing Data.....	28
3.6 Analisa Data dan Pelabelan Data.....	29
3.6.1 Analisa Data.....	29
3.6.2 Pelabelan Data.....	29
3.7Analisa Hasil.....	30
3.8 Implementasi.....	30
BAB IV	32
HASIL DAN PEMBAHASAN.....	32
4.1 Pengumpulan Data.....	32
4.2 Preprocessing.....	35
4.2.1 Cleansing.....	35
4.2.2 Case Folding.....	38
4.2.3 Tokenisasi	40
4.2.4 Stopword Removal.....	42
4.2.5 Stemming.....	43
4.3 Analisa Data dan Pelabelan Data.....	45
4.3.1 Analisa Data.....	45
4.3.2 Pelabelan Data.....	46
4.4 Analisa dan Evaluasi.....	49
4.4.1 Analisa.....	49
Term Frequency - Inverse Document Frequency (TF-IDF)	50
<i>Naive Bayes Classifier</i>	52
<i>Support Vector Machine</i>	58
4.4.2 Evaluasi Model Klasifikasi.....	63

4.5 Implementasi dan Hasil.....	64
BAB V	71
KESIMPULAN DAN SARAN.....	71
5.1 Kesimpulan	71
5.2 Saran	73
DAFTAR PUSTAKA	74
LAMPIRAN.....	77



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi dari sentimen analisis	16
Gambar 2.2 Hyperplane Support Vector Machine (SVM)	21
Gambar 3.1 Tahap-tahap Penelitian	27
Gambar 4.1 Proses Analisis Sentimen	32
Gambar 4.2 Hasil Implementasi Model Klasifikasi Pada Data Uji	65
Gambar 4.3 Hasil Implementasi Model Klasifikasi Pada Data Uji	66
Gambar 4.4 Word Cloud Sentimen positif Klasifikasi Naive Bayes	67
Gambar 4.5 Word Cloud Sentimen negatif Klasifikasi Naive Bayes	67
Gambar 4.6 Word Cloud Sentimen netral Klasifikasi Naive Bayes	68
Gambar 4.7 Word Cloud Sentimen positif Klasifikasi SVM.....	68
Gambar 4.8 Word Cloud Sentimen negatif Klasifikasi SVM.....	69
Gambar 4.9 Word Cloud Sentimen netral Klasifikasi SVM.....	70

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	10
Tabel 2.1 Tinjauan Pustaka.....	10
Tabel 4.1 Contoh data tweet.....	31
Tabel 4.2 Contoh cleansing pada data tweet.....	34
Tabel 4.3 Contoh Casefolding pada tweet.....	37
Tabel 4.4 Contoh tokenisasi pada tweet.....	38
Tabel 4.5 Contoh stopword removal pada tweet.....	40
Tabel 4.6 Contoh stemming pada data tweet.....	42
Tabel 4.7 Contoh data tweet yang telah diberi label.....	44
Tabel 4.8 Contoh data tweet.....	47
Tabel 4.9 Nilai TF dan IDF.....	49
Tabel 4.10 Hasil perhitungan probabilitas <i>term</i> positif.....	52
Tabel 4.11 Hasil perhitungan probabilitas <i>term</i> negatif.....	53
Tabel 4.12 Hasil perhitungan probabilitas <i>term</i> netral.....	54
Tabel 4.13 Contoh Klasifikasi 3 Kelas dengan OVR.....	57
Tabel 4.14 Contoh Data Latih One Versus Rest.....	58
Tabel 4.15 Contoh dari nilai X.....	59
Tabel 4.16 Contoh dari nilai Y.....	59
Tabel 4.17 Jumlah Data Setiap Kelas Sentimen.....	61

PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TWITTER

(Studi kasus : Data Tweet Dengan Kata Kunci “Omicron”)

Ahabullah Fakhri Muhammad

18106050040

INTISARI

Pandemi COVID-19 telah mengubah berbagai aspek kehidupan manusia saat ini, baik di bidang kesehatan, pendidikan dan bahkan finansial. Salah satu pengendalian peningkatan COVID-19 saat ini adalah vaksinasi.

Namun banyak pro kontra yang diperbincangkan terhadap keefektifan vaksin ini. Belum tuntas pembicaraan mengenai hal itu, saat ini muncul varian baru COVID-19 yaitu omicron. Pemberitaan melalui kanal youtube juga dicari dan dilihat orang untuk mengetahui seperti apa virus omicron ini dan efektifitas vaksin yang sudah dilakukan masyarakat terhadap virus baru ini. Dalam penelitian ini kita melihat bagaimana sudut pandang masyarakat twitter mengenai COVID-19 varian baru yaitu omicron.

Penelitian ini bertujuan untuk menganalisis terhadap opini atau sentimen yang berkembang di masyarakat terhadap COVID-19 omicron menggunakan metode *Naive Bayes Classifier* dan *Support Vector Machine*.

Hasil implementasi pada data uji sebanyak 5.967 data tweet menggunakan model *Naive Bayes Classifier* menghasilkan klasifikasi sentimen positif sebanyak 12,7% dengan 756 data, kelas sentimen netral sebanyak 8,4% dengan 504 data dan sentimen negatif sebanyak 78,9% dengan 4.707 data sedangkan menggunakan model *Support Vector Machine* menghasilkan klasifikasi kelas sentimen positif sebanyak 16,7% dengan 996 data, kelas sentimen netral sebanyak 18,6% dengan 1.111 data dan kelas sentimen negatif sebanyak 64,7% dengan 3.860 data. Data yang digunakan adalah sebanyak 10.110 tweets dari twitter, yang mana 4.143 digunakan sebagai data latih dengan label dan 5.967 data uji tanpa label. Proses pembelajaran terhadap data latih menghasilkan nilai akurasi sebesar 74,21% dengan metode *Naive Bayes Classifier* dan 80,58% dengan metode *Support Vector Machine*

Kata kunci : Covid-19, Omicron, Analisis Sentimen, Naive Bayes Classifier, Support Vector Machine

COMPARISON OF NAIVE BAYES CLASSIFIER AND SUPPORT VECTOR MACHINE METHODS IN TWITTER SENTIMENT ANALYSIS

(Case Study: Tweet Data With Keywords ‘Omicron’)

Ahabullah Fakhri Muhammad

18106050040

ABSTRACT

The COVID-19 pandemic has changed various aspects of human life today, both in the fields of health, education and even finance. One way to control the current increase in COVID-19 is vaccination. However, many pros and cons were discussed regarding the effectiveness of this vaccine. The discussion about this has not yet been completed, now a new variant of COVID-19 has emerged, namely Omicron. The news through the YouTube channel is also sought after and seen by people to find out what this omicron virus looks like and the effectiveness of the vaccine that has been carried out by the community against this new virus. In this study, we look at the perspective of the Twitter community regarding the new variant of COVID-19, namely Omicron.

This study aims to analyze the opinions or sentiments that develop in the community towards COVID-19 omicron using the Naive Bayes Classifier and Support Vector Machine methods.

The results of the implementation on test data of 5,967 tweet data using the Naive Bayes Classifier model resulted in a classification of positive sentiment as much as 12.7% with 756 data, neutral sentiment class as much as 8.4% with 504 data and negative sentiment as much as 78.9% with 4,707 data. Using the Support Vector Machine model, the positive sentiment class classification is 16.7% with 996 data, the neutral sentiment class is 18.6% with 1,111 data and the negative sentiment class is 64.7% with 3,860 data. The data used are 10,110 tweets from twitter, of which 4,143 are used as training data with labels and 5,967 test data without labels. The learning process on the training data produces an accuracy value of 74.21% with the Naive Bayes Classifier method and 80.58% with the Support Vector Machine method.

Keywords : Covid-19, Omicron, Sentiment Analysis, Naive Bayes Classifier, Support Vector Machine.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Banyak warga Indonesia yang masih acuh terhadap isu COVID-19 ini, khususnya varian Omicron. Padahal sudah memakan banyak korban baik untuk kaum dewasa maupun anak-anak. Dengan perkembangan teknologi kita dapat dengan mudah mencari informasi terkait COVID-19 Omicron ini di internet baik melalui media sosial, website berita online bahkan lewat Youtube.

Media sosial adalah media daring yang digunakan untuk kebutuhan komunikasi jarak jauh, proses interaksi antara user satu dengan user lain, serta mendapatkan sebuah informasi melalui perangkat aplikasi khusus menggunakan jaringan internet. Tujuan dari adanya social media sendiri adalah sebagai sarana komunikasi untuk menghubungkan antar pengguna dengan cakupan wilayah yang sangat luas.

Agar pengguna media sosial (medsos) lebih mudah dan cepat, dibutuhkan koneksi internet yang stabil dan cepat. Anda tidak perlu lagi menghubungi orang lain melalui kabel telepon atau alat komunikasi tradisional. Cukup dengan mengakses media sosial, anda dapat terhubung dengan banyak orang, membuat forum, diskusi bersama, mengunggah aktivitas keseharian anda, dan lain sebagainya.

Salah satu platform media sosial yang populer khususnya di Indonesia yakni Twitter. Twitter merupakan salah satu media sosial yang didirikan oleh Jack Dorsey yang umumnya berfungsi untuk mengirimkan pesan yang disebut dengan kicauan atau (tweet). Penggunaan twitter umumnya diakses para pengguna di Indonesia mencapai 59% dan menduduki peringkat ke-5 media sosial yang sering digunakan pada tahun 2020. Dari data yang ada menjadikan platform twitter menjadi media sosial yang cukup memiliki pengaruh bagi para pengguna di Indonesia.

Oleh karena itu, penelitian ini akan melakukan analisis terhadap sentimen atau opini masyarakat terhadap varian COVID-19 Omicron. Data yang akan dianalisis adalah data twitter berbahasa Inggris dengan kata kunci “omicron”. Data tersebut kemudian diklasifikasikan menjadi tiga jenis sentimen, yaitu positif, negatif dan netral. Metode yang akan digunakan dalam penelitian ini adalah *Naive bayes Classifier* dan *Support Vector Machine* dikarenakan kedua metode tersebut mudah untuk dipahami, cepat dalam penghitungan dan memiliki nilai akurasi yang tinggi. Dengan adanya penelitian ini kami harapkan mengubah sudut pandang warga Indonesia terhadap COVID-19 Omicron ini dan membuat warga Indonesia lebih waspada.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang akan diselesaikan dalam penelitian ini yaitu:

1. Belum adanya penelitian analisis sentimen twitter khususnya mengenai COVID 19 Omicron dengan metode *Naive bayes Classifier* dan *Support Vector Machine*.
2. Belum adanya akurasi yang jelas mengenai penelitian analisis sentimen khususnya mengenai COVID 19 Omicron dengan metode *Naive bayes Classifier* dan *Support Vector Machine*.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini yaitu sebagai berikut.

1. Untuk mengetahui sentimen analisis twitter dari COVID 19 Omicron dengan metode *Naive bayes Classifier* dan *Support Vector Machine*.
2. Mengetahui akurasi dan gambaran yang diperoleh dari metode *Naive bayes Classifier* dan *Support Vector Machine* untuk melakukan analisis sentimen mengenai COVID 19 Omicron agar masyarakat waspada terhadap bahayanya COVID 19 Omicron.

1.4 Batasan Masalah

Pada penelitian ini didefinisikan beberapa batasan sebagai berikut:

1. Media yang digunakan untuk mengambil data *tweet* pada *Twitter* adalah *Twitter API*
2. *Tweet* yang diambil dan dianalisis menggunakan *tweet* berbahasa Inggris dikarenakan *tweet* berbahasa Inggris lebih banyak dibanding menggunakan *tweet* berbahasa Indonesia
3. Durasi pengambilan data dari twitter selama 1 bulan lebih 2 minggu dengan rentan waktu mulai dari Desember 2021 hingga Maret 2022.

1.5 Manfaat Penelitian

Penelitian ini dapat menjadi referensi bagi peneliti mendatang bila akan melakukan penelitian dengan tema yang serupa. Adapun manfaat dari penelitian ini adalah :

1. Dapat mengetahui algoritma mana yang memiliki akurasi yang lebih tinggi dalam mengklasifikasikan sentimen terhadap data tweet
2. Bagi UIN Sunan Kalijaga, dapat digunakan sebagai tambahan referensi terkait dengan peningkatan penggunaan sosial media seiring kemajuan teknologi.
3. Bagi penulis, dapat memberikan pengetahuan dan wawasan yang baru, bahwa data tweet dapat dijadikan sebagai sumber data yang bisa digunakan

untuk menganalisis opini terhadap suatu topik.

4. Bagi pembaca, dengan adanya penelitian ini dapat digunakan sebagai sumber pembelajaran untuk penelitian selanjutnya dan menambah wawasan baru.

1.6 Kebaruan Penelitian

Penelitian mengenai sentimen analisis menggunakan metode *Naive Bayes* dan *Support Vector Machine* memang sudah banyak yang dilakukan. Namun belum ada yang meneliti dengan tema efektivitas vaksin terhadap COVID 19 Omicron secara khusus.

1.7 Sistematika Penulisan

Berikut ini adalah sistematika penulisan yang dilakukan dalam penelitian ini :

BAB I PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian dan sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Menjelaskan penelitian yang telah dilakukan sebelumnya serta landasan teori yang berhubungan dengan topik yang akan dibahas dalam penelitian ini.

BAB III METODE PENELITIAN

Membahas metode yang dilakukan dalam penelitian serta langkah-langkah yang harus dilakukan untuk mencapai tujuan dalam menyelesaikan penelitian tugas akhir.

BAB IV HASIL DAN PEMBAHASAN

Bagian ini menjelaskan proses yang dilakukan dalam penelitian mulai dari pengumpulan data hingga hasil yang didapatkan dari penelitian.

BAB V KESIMPULAN DAN SARAN

Menjelaskan kesimpulan dari seluruh tahapan penelitian serta saran, sehingga dapat menjadi bahan pertimbangan bagi pihak-pihak yang berkepentingan serta pengembangan selanjutnya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari penelitian diatas kita dapat menyimpulkan bahwa, Data yang digunakan adalah data yang diambil dari *Twitter* dengan cara *crawling* menggunakan *API twitter* dengan kata kunci 'Omicron'. Data yang didapatkan sebanyak 10.110 data tweet, kemudian dilakukan seleksi data sehingga mendapatkan 4.143 data latih dan 5.967 data uji. Data diolah menggunakan metode *Naive Bayes Classifier* dan *Support Vector Machine* dengan pembobotan *TF-IDF* serta dicari akurasiya menggunakan *k-fold cross validation*.

Berdasarkan penelitian yang telah dilakukan terhadap 4.143 data latih dengan pembobotan *TF-IDF* menggunakan metode *Naive Bayes Classifier* didapatkan nilai rata-rata akurasi saat menggunakan *16-fold cross validation* sebesar 74,21%. Sedangkan nilai rata-rata akurasi saat menggunakan *16-fold cross validation* pada metode *Support Vector Machine* adalah sebesar 80,58% .

Dari perbandingan hasil akurasi diatas dapat disimpulkan bahwa kombinasi penggunaan metode untuk mencari model klasifikasi menggunakan *Naive Bayes Classifier* dan *Support Vector Machine*, metode SVM lah yang memiliki akurasi yang lebih tinggi dibanding metode NBC.

Selanjutnya, mengimplementasikan hasil tersebut pada 5.967 data uji yang belum memiliki label. Untuk model *Naive Bayes Classifier* menghasilkan sentimen positif sebanyak 12,7 % dengan 756 data, kelas sentimen netral sebanyak 8,4% dengan data 504, dan kelas sentimen negatif sebanyak 78,9% dengan data 4.707. Sedangkan untuk model *Support Vector Machine* dengan parameter $C=3$, $\gamma=0,5$ menghasilkan kelas sentimen positif sebanyak 16,7% dengan 996 data, kelas sentimen netral 18,6% dengan 1.111 data, dan kelas sentimen negatif 64,7% dengan 3.860 data. Dari kedua model tersebut sama-sama menjadikan kelas sentimen negatif sebagai nilai tertinggi dan nilai yang rendah untuk kedua kelas sentimen lainnya.

Dengan adanya hasil sentimen diatas, kita dapat menyimpulkan bahwa sudut pandang masyarakat mayoritas menganggap Omicron ini adalah suatu hal yang negatif. Diharapkan dengan adanya hasil dari penelitian ini masyarakat lebih waspada terhadap COVID-19 Omicron ini, salah satu cara adalah dengan mematuhi protokol kesehatan yang sudah diterapkan oleh pemerintah, seperti menggunakan masker, *social distancing*, selalu membawa *hand sanitizer* dan lain-lain.

5.2 Saran

Pada penelitian ini tentunya tidak terlepas dari kelemahan dan kekurangan. Oleh karena itu peneliti menyarankan adanya perbaikan-perbaikan yang dapat dijadikan masukan untuk peneliti selanjutnya, adapun sebagai berikut :

1. Penelitian selanjutnya diharapkan menggunakan data latih dengan jumlah yang lebih banyak agar mendapatkan model yang lebih akurat.
2. Penelitian selanjutnya diharapkan bisa menggunakan tahapan preprocessing yang berbeda
3. Penelitian selanjutnya diharapkan dapat melakukan proses labeling data dengan menggunakan dua atau lebih personel agar dapat memberi sudut pandang yang berbeda.
4. Untuk data yang menggunakan bahasa asing, harus menggunakan *expert* atau ahli bahasa untuk memvalidasi data tersebut.
5. Penelitian selanjutnya diharapkan menggunakan pengembangan dari algoritma Naive Bayes Classifier.

DAFTAR PUSTAKA

- Ahmad P. A. R.,(2021). Studi Komparasi Algoritma Naive Bayes Classifier dan Support Vector Machine Dalam Analisis Sentimen Tweet (Studi Kasus : Data tweet dengan kata kunci 'psbb' dan 'pandemi')
- Alfikri, M. Z. (2020). Analisis Sentimen Twitter terhadap Kartu Prakerja di tengah Pandemi COVID-19 menggunakan Algoritma Pencocokan String dan library TextBlob.
- Didik G Nu, Yulison H, Agung W (2021). Analisis Sentimen pada Jasa Ojek Online Menggunakan Metode Naive Bayes Jur. Inform. Fak. Mat. dan Ilmu Pengetah. Alam Univ. Jenderal Achmad Yani, pp. 156–161, 2015.
- Fajar R(2018). Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film pada Twitter JURNAL INOVTEK POLBENG - SERI INFORMATIKA, VOL. 3, NO. 1 , JUNI 2018
- Fransiska V. S., Arief W (2019). Analisis Sentimen Pelanggan Toko Online JD.ID menggunakan metode Naive Bayes Classifier berbasis Konversi Ikon Emosi, Jurnal SIMETRIS, Vol. 10 No. 2 November 2019
- Februariyanti, H., & Zuliarso, E. (2012). Klasifikasi Dokumen Berita Teks Bahasa Indonesia menggunakan Ontologi. Teknologi Informasi DINAMIK, 17 <http://www.unisbank.ac.id/ojs/index.php/fti1/article/view/1612/594>
- Handayani, F., & Pribadi, S. (2015). Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110. Jurnal Teknik Elektro, 7(1), 19–24. <https://doi.org/10.15294/jte.v7i1.8585>

- Nugroho, G. A. (2020). Analisis big data penggunaan internet pada pengguna layanan ubiqu
- Muhammad G. T. A., Taufik A., Dwi B. S., Analisa Sentimen Efektivitas terhadap varian COVID-19 Omicron Berbasis Leksikon, *Journal of Information and Information Security (JIFORTY)* Vol. 2, No. 2, Desember 2021, 251–258 e-ISSN : 2722-4058
- Muljono, Dian P. A., Abdul S., Adi P., De R, Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naive Bayes, *Konferensi Nasional Sistem Informasi 2018 STMIK Atma Luhur Pangkalpinang*, 8 – 9 Maret 2018
- Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*, 7(2), 156. <https://doi.org/10.51211/biict.v7i2.1422>
- Rauhan, A. (2019). Pengolahan Data Menggunakan Machine Learning. *Student Paper Pertamina University*, 021, 1–4. [https://library.universitaspertamina.ac.id/xmlui/bitstream/handle/123456789/162/Jurnal Ilmiah.pdf?sequence=1&isAllowed=y](https://library.universitaspertamina.ac.id/xmlui/bitstream/handle/123456789/162/Jurnal%20Ilmiah.pdf?sequence=1&isAllowed=y)
- Rian T., Agung T., Ira D. (2020). Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine, *Jurnal media informatika budidarma* volume 4, nomor 3, Juli 2020 halaman 650-658
- Samsir, Ambiyar, Unung V., Firman E., Ronal W., Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes, *JURNAL MEDIA INFORMATIKA BUDIDARMA* Volume 5, Nomor 1, Januari 2021, Page 157-163

- Septian, J. A., Fahrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF - IDF dan K - Nearest Neighbor. *Journal of Intelligent Systems and Computation*, September, 43–49.
- Sevsa, B. A., & R Wahyudi, M. D. (2019). Analisis Sentimen pada Indeks Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier. *Jurnal Buana Informatika*, 10(2), 112. <https://doi.org/10.24002/jbi.v10i2.2250>
- Tempola, F., Muhammad, M., & Khairan, A. (2018). Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 5(5), 577. <https://doi.org/10.25126/jtiik.201855983>
- Umi R., Rizal S. P., M. Ali F., Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Features, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* e-ISSN: 2548-964X Vol. 1, No. 12, Desember 2017, hlm. 1725-1732 <http://j-ptiik.ub.ac.id>
- Wino M. A. P., Susanti, Erlin, Herwin, Analisis Sentimen Dompot Elektronik Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier, *IT Journal Research and Development (ITJRD)* Vol.5, No.1, Agustus 2020, pp. 72 – 86, E-ISSN : 2528-4053 | P-ISSN: 2528-4061 DOI : 10.25299/itjrd.2020.vol5(1).5159