

Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine

Irma Putri Rahayu*, Ahmad Fauzi, Jamaludin Indra

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Buana Perjuangan Karawang, Karawang, Indonesia

Email: ¹:if19.irmarahayu@mhs.ubpkarawang.ac.id, ²afauzi@ubpkarawang.ac.id, ³jamaludin.indra@ubpkarawang.ac.id

Email Penulis Korespondensi: if19.irmarahayu@mhs.ubpkarawang.ac.id

Submitted: 22/12/2022; Accepted: 29/12/2022; Published: 31/12/2022

Abstrak—Dalam rangka menyiapkan mahasiswa menghadapi pesatnya perkembangan teknologi, perubahan kehidupan kerja dan keterampilan, mahasiswa harus lebih siap menghadapi kemajuan zaman. Perguruan tinggi harus mampu melaksanakan proses pembelajaran yang inovatif agar mahasiswa mencapai hasil belajar yang optimal yang meliputi aspek pengetahuan, keterampilan dan sikap. Maka diluncurkanlah program MBKM untuk menjawab tuntutan tersebut. Namun, MBKM memiliki pro dan kontra dalam pelaksanaannya, maka perlu menganalisis dan mengevaluasi kebijakan untuk meningkatkan kinerja melalui umpan balik dari masyarakat dengan melakukan analisis sentimen terhadap kebijakan MBKM pada pengguna twitter dari tahun 2019 hingga 2022 dengan tagar #kampusmerdeka. Penelitian ini menggunakan algoritma *Naïve Bayes* dan *SVM* untuk menentukan akurasi berdasarkan klasifikasi sentimen. Data yang digunakan 1118 data dengan sentimen positif 618 data dan sentimen negatif 500 data. Penelitian ini menghasilkan akurasi sebesar 86%, presisi sebesar 87% dan recall sebesar 80% dengan data testing menggunakan algoritma *Naïve Bayes*. Kemudian menggunakan algoritma *SVM* kernel linear dengan data testing yang sama menghasilkan akurasi sebesar 93%, presisi sebesar 100% dan recall sebesar 84%. Oleh karena itu, penting dilakukan kajian untuk meningkatkan program MBKM agar pelaksanaannya jelas sesuai dengan prosedur yang ada.

Kata Kunci: Analisis Sentimen; Kampus Merdeka; SVM; Naive Bayes; Klasifikasi

Abstract—In order to prepare students to face the rapid development of technology, changes in work life and skills, students must be better prepared to face the progress of the times. Universities must be able to carry out innovative learning processes so that students achieve optimal learning outcomes which include aspects of knowledge, skills and attitudes. So the MBKM program was launched to answer these demands. However, MBKM has pros and cons in its implementation, so it is necessary to analyze and evaluate policies to improve performance through feedback from the public by conducting sentiment analysis of MBKM policies on twitter users from 2019 to 2022 with the hashtag #kampusmerdeka. This study used the *Naïve Bayes* and *SVM* algorithms to determine accuracy based on sentiment classification. The data used 1118 data with positive sentiment 618 data and negative sentiment 500 data. This study resulted in an accuracy of 86%, precision of 87% and recall of 80% with testing data using the *Naïve Bayes* algorithm. Then using the linear kernel *SVM* algorithm with the same testing data resulted in accuracy of 93%, precision of 100% and recall of 84%. Therefore, it is important to conduct studies to improve the MBKM program so that its implementation is clearly in accordance with existing procedures.

Keywords: Sentiment Analysis; Merdeka Campus; SVM; Naive Bayes; Classification

1. PENDAHULUAN

Proses pembelajaran melalui MBKM merupakan ekspresi esensial dari pembelajaran yang berpusat pada mahasiswa. Mahasiswa dapat memilih dari delapan program MBKM yang telah disetujui oleh Kemendikbud dan bebas memilih mata kuliah non program selama satu semester (20sks)[1]. MBKM menawarkan tantangan dan peluang untuk mengembangkan inovasi, kreativitas dan mengembangkan kemandirian dalam penelitian dan pemecahan masalah melalui realitas dan dinamika lapangan. Melalui program MBKM yang telah dirancang dan dilaksanakan, softskill dan professional skill mahasiswa akan dilatih hingga tingkat tinggi. MBKM akan berhasil implementasinya jika perguruan tinggi mampu dan berani merubah pemikiran menuju kurikulum yang adaptif dan fleksibel [2].

Kebijakan program MBKM ditanggapi beragam oleh beberapa perguruan tinggi di Indonesia, terdapat beberapa masalah dan kendala yang dihadapi mahasiswa saat mengikuti program ini adalah pelaksanaan program masih belum jelas, belum adanya pedoman berupa tata tertib, petunjuk pelaksanaan dan instruksi prosedur selama operasi. Sebagian besar mitra yang mengerjakan program ini berlokasi di ibu kota, yang memengaruhi mahasiswa dari perguruan tinggi daerah yang kesulitan dengan biaya transportasi, dukungan pendapatan, dan tempat tinggal [1]. Kebijakan masing-masing mitra berbeda dalam hal jam operasi, tugas dan tunjangan atau biaya hidup. Tanggapan atau respon dari berbagai kalangan masyarakat, disampaikan dalam bentuk komentar dan opini yang diberikan melalui sosial media. Salah satu sosial media yang digunakan untuk menanggapi kebijakan ini adalah melalui twitter. Masyarakat umum maupun akademik menyampaikan respon terkait masih adanya kebingungan terhadap pelaksanaan kebijakan ini dalam nada positif maupun negatif. Oleh karena itu, dilakukan analisis sentimen terkait kepuasan program MBKM melalui opini masyarakat, sehingga diharapkan program MBKM dapat terlaksana dengan jelas dan sesuai prosedur.

Penelitian terdahulu terkait analisis sentimen dilakukan oleh Silitonga (2022) mengenai Analisis Sentimen Kampus Merdeka menggunakan Machine Learning dengan menggunakan sebanyak 351 data dari twitter menghasilkan sentimen positif sebanyak 267 tweet dan sentimen negatif sebanyak 85 tweet. Data tersebut dibagi menjadi 70% data latih dan 30% untuk data uji, sehingga akurasi model SVM adalah 80,75%. Penelitian berikut

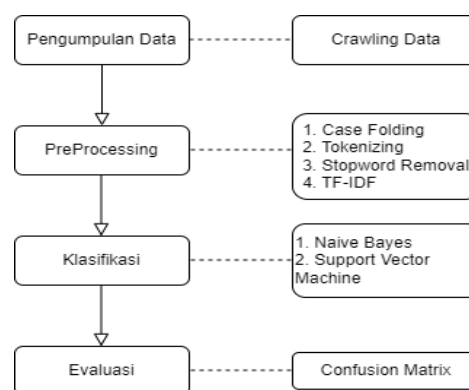
adalah Analisis Sentimen terhadap Program Merdeka Belajar - Kampus Merdeka di twitter dengan menggunakan algoritma SVM. Didapatkan nilai akurasi untuk setiap kelas yang ditunjukkan oleh nilai F-measure, kelas positif sebesar 94.8% dan kelas negatif sebesar 95%. Penelitian oleh Gunawan (2022) mengenai K-Nearest neighbor (KNN) untuk menganalisis sentimen terhadap kebijakan merdeka kampus merdeka pada komentar twitter menggunakan data sebanyak 700 dengan 350 data positif dan 350 data negatif. Nilai akurasi tertinggi dengan 700 set data dibandingkan 90 % data training, 10 % data testing adalah nilai k-8 yaitu 84,28 %, precision 84,42 %, recall 84,28 % dan f-measure 84,35 %. Analisis opini independen terhadap politik kampus menggunakan *Naive Bayes* dan TF-IDF berdasarkan komentar youtube oleh Zhafira (2021) menggunakan data API sebanyak 1000 data komentar, mendapatkan hasil terbaik sebesar 97% didapat dari iterasi ke-4 pada *10-fold cross validation* dengan komposisi data latih sebesar 900 dengan 100 data uji dan penerapan pembobotan TF-IDF. Rata-rata *score accuracy* yang didapat dari uji validitas data menggunakan *k-fold cross validation* yaitu 91.8% dengan nilai *precision, recall, f-measure* sejumlah 90.35%, 93.6%, 91.95%. Penelitian mengenai Analisis sentimen implementasi program merdeka belajar kampus merdeka menggunakan *Naive Bayes, K-Nearest Neighbors* dan *Decision Tree* mendapatkan nilai akurasi sebesar 99,22% untuk *Naive Bayes*, 96,90% untuk *KNN* dan *Decision tree* sebesar 37,21% [6].

Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berskala Besar Menggunakan Algoritma SVM, hasil analisis sentimen terhadap 725 dokumen data uji menggunakan algoritma SVM dan 9-fold cross validation mengungkapkan sentimen positif sebesar 28%, sentimen negatif sebesar 27% dan sentimen netral sebesar 45% [7]. Kajian Analisis Publik terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma SVM, hasil pengujian TF-IDF didapatkan accuracy 74%, precision 75%, recall 92% dan F1-Score 83%.

Berdasarkan penjelasan di atas, maka metode yang memberikan nilai akurasi terbaik adalah metode *Naive Bayes*. Namun, penelitian ini dilakukan dengan menggunakan dua algoritma yaitu *Naive Bayes* dan *SVM*. Studi ini dilakukan untuk menunjukkan tingkat akurasi dari algoritma *Naive Bayes* dan *SVM* dalam hal sentimen masyarakat terhadap kepuasan program MBKM dengan menggunakan metode yang efektif dan akurat untuk membantu pengambilan keputusan terhadap program MBKM yang dilaksanakan.

2. METODOLOGI PENELITIAN

Terdapat 4 langkah dalam penelitian ini yaitu pengumpulan data, preprocessing, klasifikasi dan evaluasi. Gambar 1. menunjukkan langkah-langkah dalam proses penelitian.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Data yang digunakan yaitu data yang tersedia untuk umum di platform Twitter. Teknik akuisisi data yang digunakan adalah teknik scraping. Scraping adalah tahap mengumpulkan data umpan balik dan opini dari tweet media sosial. Informasi tersebut diambil dari twitter pada tanggal 28 Oktober 2022 menggunakan Python dengan tagar #kampusmerdeka dan menghasilkan 1120 data.

2.2 Preprocessing

Adalah langkah pembersihan data sebelum diklasifikasikan. Preprocessing menghindari pengocokan data, data tidak sempurna dan data yang tidak konsisten. Pengolahan data yaitu pembersihan data mentah yang diperoleh selama proses crawling agar siap untuk digunakan pada langkah selanjutnya seperti case folding, tokenizing, stopword removal dan TF-IDF [9].

1. Case Folding : proses mengubah huruf besar menjadi huruf kecil dan menghilangkan tanda baca seperti koma, titik dan tanda baca lainnya, agar teks konsisten[10].
2. Tokenizing : proses memotong kalimat menjadi kata per kata, hal ini dilakukan dengan cara memotong suatu kalimat dengan acuan tiap spasi[11].

3. Stopword Removal : proses menghapus kata-kata tidak penting dalam pelatihan [12].
4. TF-IDF digunakan untuk mengetahui frekuensi dari istilah tertentu yang relatif terhadap sebuah kata dalam kumpulan dokumen dan melihat seberapa umum atau tidak umum sebuah kata yang ada diantara corpus (sekumpulan teks yang terstruktur) atau memberikan bobot pada setiap kata yang ada menggunakan TF-IDF[13]. Rumus TF-IDF sebagai berikut :

$$TF(t) = f_{t,d} / \sum_t t, d \tag{1}$$

Dimana $f_{t,d}$ adalah frekuensi sebuah kata dalam dokumen, $\sum_t t, d$ adalah total kata yang terdapat didalam dokumen. Kemudian hitung IDF (Inverse Document Frequency) dari dokumentasi tersebut persamaan :

$$IDF(t) = \log(|D| / f_{t,D}) \tag{2}$$

$|D|$ adalah jumlah dokumen dalam koleksi, sedangkan $f_{t,D}$ adalah jumlah dokumen dimana t ada di D . Setelah nilai TF dan IDF nilainya didapatkan, langkah selanjutnya adalah menghitung TF-IDF menggunakan persamaan berikut:

$$TF-IDF = TF(t) * IDF(t) \tag{3}$$

2.3 Klasifikasi

Proses untuk mengkategorikan komentar yang sentimennya tidak diketahui. Proses klasifikasi menggunakan algoritma SVM dan Naive Bayes yang terbagi menjadi dua set data yaitu training dan data testing. Kemudian data akan diklasifikasikan berdasarkan ketetapan Confusion Matrix[8].

1. Rumu perhitungan naïve bayes :

$$P(H | X) = \frac{P(H | X) P(H)}{P(X)} \tag{4}$$

Keterangan :

X = Data kelas tidak diketahui

H = data hipotesis X adalah kelas yang terpisah

$P(H|X)$ = Probabilitas hipotesis kondisional H

$P(H)$ = Probabilitas hipotesis H

$P(X|H)$ = Probabilitas X berdasarkan kondisi pada hipotesis

$P(X)$ = Probabilitas X

2. Rumus untuk perhitungan SVM :

Konsep operasional SVM adalah mencari garis batas optimal yang berfungsi untuk memisahkan dua kelas. Untuk mendapatkan hyperline di SVM menggunakan persamaan (5) :

$$(w \cdot x_i) + b = 0 \tag{5}$$

Pada data x_i , yang termasuk dalam kategori -1 dapat dirumuskan seperti pada persamaan (6).

$$(w \cdot x_i + b) \leq 1, y_i = -1 \tag{6}$$

Sedangkan data x_i yang termasuk dalam kategori +1 dapat dirumuskan seperti pada persamaan (7).

$$(w \cdot x_i + b) \geq 1, y_i = 1 \tag{7}$$

Adapun kernel yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Persamaan Kernel

Kernel	Persamaan
Linear	$K(x_i, x) = x_i^T x$
Sigmoid	$K(x_i, x) = \tanh(y_i^T x + r)$

2.4 Evaluasi

Evaluasi termasuk untuk memastikan bahwa tes itu benar. Tes dilakukan untuk mengukur keakuratan hasil dan menemukan hasil tes yang terbaik. Pengukuran akurasi terhadap model menggunakan confusion matrix, perhitungannya dengan akurasi, presisi, recall. Berikut Tabel dari confusion matrix.

Tabel 2. Confusion Matrix

	Class prediction	
	1	0
Actual class 1	True positive	False positive
Actual class 0	False negative	True negative

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+FN+TN} \tag{8}$$

$$\text{Presisi} = \frac{TP}{TP+FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{10}$$

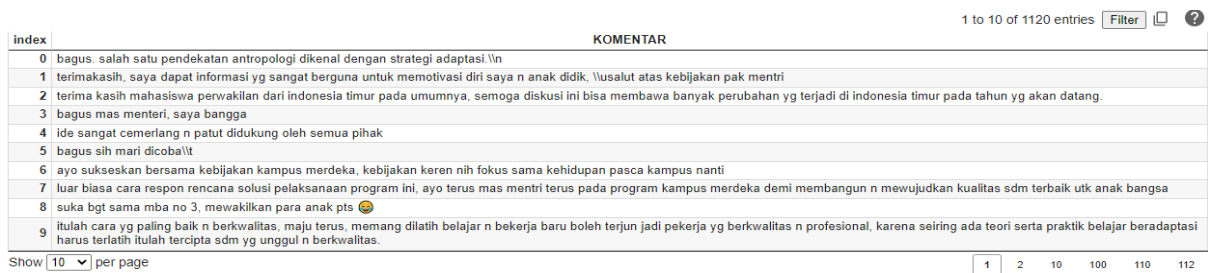
Keterangan:

1. Akurasi: perbandingan jumlah prediksi yang benar
2. Presisi: proporsi kasus dengan hasil positif yang benar
3. True Positive : jumlah record positif yang diklasifikasikan oleh classifier sebagai positif.
4. True Negative : jumlah record negatif yang diklasifikasikan sebagai negatif oleh classifier.
5. False positive : jumlah record negatif yang dianggap positif oleh classifier.
6. False Negative : jumlah record positif yang diklasifikasikan sebagai negatif oleh classifier.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data yang digunakan adalah data crawling twitter yang diterima dalam format .csv hingga 1120 data dalam 2 kolom. Isi data dapat ditunjukkan pada Gambar 2.



index	KOMENTAR
0	bagus. salah satu pendekatan antropologi dikenal dengan strategi adaptasi.\n
1	terimakasih, saya dapat informasi yg sangat berguna untuk memotivasi diri saya n anak didik, \nusalut atas kebijakan pak menteri
2	terima kasih mahasiswa perwakilan dari indonesia timur pada umumnya, semoga diskusi ini bisa membawa banyak perubahan yg terjadi di indonesia timur pada tahun yg akan datang.
3	bagus mas menteri, saya bangga
4	ide sangat cemerlang n patut didukung oleh semua pihak
5	bagus sih mari dicoba!\n
6	ayo sukseskan bersama kebijakan kampus merdeka, kebijakan keren nih fokus sama kehidupan pasca kampus nanti
7	luar biasa cara respon rencana solusi pelaksanaan program ini, ayo terus mas menteri terus pada program kampus merdeka demi membangun n mewujudkan kualitas sdm terbaik utk anak bangsa
8	suka bgt sama mba no 3, mewakili para anak pts 🤗
9	itulah cara yg paling baik n berkualitas, maju terus, memang dilatih belajar n bekerja baru boleh terjun jadi pekerja yg berkualitas n profesional, karena seiring ada teori serta praktik belajar beradaptasi harus terlatih itulah tercipta sdm yg unggul n berkualitas.

Gambar 2. Isi Dataset

3.2 Preprocessing Data

Pada tahap preprocessing sebelum data tweet digunakan, perlu mendapatkan data bersih dan melabelinya. Langkah-langkah preprocessing yaitu :

3.2.1 Case Folding

Proses merubah huruf menjadi kecil dan menghapus tanda baca, angka, dan URL. Bisa dilihat pada Tabel 3.

Tabel 3. Cuplikan Case Folding

Sebelum	Sesudah
Bagus. Salah satu pendekatan antropologi dikenal dengan STRATEGI ADAPTASI.\n	bagus salah satu pendekatan antropologi dikenal dengan strategi adaptasi

3.2.2 Tokenizing

Yaitu fungsi konversi teks ke token. Berikut adalah contoh hasil proses tokenisasi yang diurutkan berdasarkan abjad seperti pada Tabel 4.

Tabel 4. Cuplikan Tokenisasi

Sebelum	Sesudah
bagus salah satu pendekatan antropologi dikenal dengan strategi adaptasi	[bagus, salah, satu, pendekatan, antropologi, dikenal, dengan, strategi, adaptasi]

3.2.3 Stopword

Kata-kata yang jarang muncul, memiliki sedikit informasi dan tidak penting, sehingga proses stopwords dapat diterapkan dan menghapus dari dokumen

Tabel 5. Hasil Stopword

(adalah, akan, akhir, aku, saya, antara, antaranya, apabila atau bahwa)

3.2.4 TF-IDF

TF-IDF digunakan untuk mengetahui frekuensi dari istilah tertentu yang relatif terhadap sebuah kata dalam kumpulan dokumen dan melihat seberapa umum atau tidak umum sebuah kata yang ada diantara corpus (sekumpulan teks yang terstruktur).

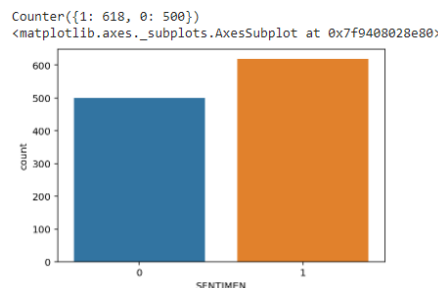
term	TF	TF-IDF
gini	0.07692307692307693	0.3221438938640711
nih	0.07692307692307693	0.2582709475507303
baru	0.07692307692307693	0.2563234238596311
kuliah	0.07692307692307693	0.24557250521538812
memang	0.07692307692307693	0.3177470927994597
baik	0.07692307692307693	0.27348831242226257
bukan	0.07692307692307693	0.30231088545621726
tahun	0.07692307692307693	0.30231088545621726
ga	0.07692307692307693	0.22513380015900306
jelas	0.07692307692307693	0.3317718279374562
cuma	0.07692307692307693	0.3317718279374562
ilmu	0.07692307692307693	0.3058893481973629
aja	0.07692307692307693	0.23614155736213177

Gambar 3. Hasil perhitungan TF-IDF

Pada tahap preprocessing, data mengalami penyusutan dari 1120 data menjadi 1118 data. Hal ini terjadi karena terdapat data yang duplicate. Data duplicate harus dihilangkan untuk meningkatkan akurasi.

3.3 Klasifikasi

Setelah tahap preprocessing data selanjutnya yaitu proses klasifikasi. Klasifikasi dilakukan di Platform Google Colab dengan bahasa pemrograman Python menggunakan algoritma Naïve Bayes dan SVM. Klasifikasi tersebut dibagi menjadi 2 kelas, yaitu positif dan negatif. Dataset dibagi menjadi data training sebanyak 783 dan data testing sebanyak 335 menggunakan ratio 0.3. Artinya, 70% data digunakan untuk proses training dan 30% data testing. Pada tahap klasifikasi sentimen menampilkan hasil evaluasi dari model klasifikasi yang dibuat. Selain itu menampilkan hasil kelas sentimen positif dan negatif, dapat dilihat pada Gambar 4



Gambar 4. Hasil Klasifikasi Sentimen Positif dan Negatif

Gambar diagram diatas menjelaskan bahawa yang berlabel 0 itu merupakan sentimen negatif yang berjumlah 500 data dan yang berlabel 1 sentimen positif berjumlah 618 data. Klasifikasi menggunakan algoritma Multinomial Naïve Bayes dengan data training 783 data mendapatkan akurasi 97.95% dan menggunakan data test 335 mendapatkan akurasi sebesar 86.30%. Untuk proses klasifikasi dengan algoritma SVM menggunakan dua fungsi kernel yaitu linear dan sigmoid. Hasil klasifikasi dapat dilihat pada Tabel 6.

Tabel 6. Hasil klasifikasi untuk kernel SVM

Pembagian Data	Kernel	Akurasi
783 Data Training	Linear	99.74%
335 Data Testing	Linear	86.01%
783 Data Training	Sigmoid	53.32%
335 Data Testing	Sigmoid	59.82%

Tabel diatas menjelaskan bahawa menggunakan data training dengan menggunakan kernel linear sebanyak 783 data dan testing sebanyak 335 data mendapatkan akurasi sebesar 99.74% dan 86.01%. Sedangkan menggunakan kernel sigmoid dengan data traing dan testing yang sama mendapatkan hasil sebesar 53.32% dan 59.82%. Jadi, dapat disimpulkan bahwa kernel linear dalam algoritma SVM lebih baik dari kernel sigmoid.

3.4 Evaluasi

Evaluasi menggunakan metode confusion matrix untuk mengetahui akurasi, presisi, dan recall. Pada tabel dibawah menjelaskan bahawa untuk mengevaluasi model menggunakan data training dan testing dengan dua algoritma yaitu algoritma Naïve Bayes dan SVM. Data training berjumlah 783 data dan data testing berjumlah 335 data.

Tabel 7. Hasil Confusion Matrix

	Akurasi	Presisi	Recall
Algoritma Naïve Bayes			
Training	97%	98%	97%
Testing	86%	87%	80%
Algoritma SVM			
Training	99%	99%	100%
Testing	93%	100%	84%

Tabel diatas menjelaskan bahwa hasil evaluasi menggunakan dua algoritma yaitu Naïve Bayes dan SVM dengan menggunakan data training dan testing mendapatkan hasil akurasi sebesar 97%, presisi 98% dan recall sebesar 97%. Sedangkan algoritma SVM mendapatkan akurasi sebesar 99%, presisi 99% dan recall 100%.

4. KESIMPULAN

Penerapan algoritma *Naïve Bayes* dan *SVM* dapat mengklasifikasi sentimen positif dan negatif pada program MBKM dengan menggunakan 1118 data tweet. Berdasarkan hasil penelitian yang dilakukan pada sentimen analisis MBKM, penelitian ini menghasilkan akurasi sebesar 86%, presisi sebesar 87% dan recall sebesar 80% dengan data testing menggunakan algoritma Naïve Bayes. Kemudian menggunakan algoritma SVM kernel linear dengan data testing yang sama menghasilkan akurasi sebesar 93%, presisi sebesar 100% dan recall sebesar 84%. Dari hasil yang diperoleh terdapat beberapa hal yang perlu dianjurkan untuk pengembangan lanjutan dari penelitian ini. Peneliti menganjurkan bahwa langkah data *preprocessing* memainkan peran yang sangat penting dalam proses klasifikasi teks, di mana tata cara yang dilakukan harus menyesuaikan dengan karakteristik dari dataset tersebut. Hasil penelitian ini dapat menjadi bahan evaluasi tim ke depan mengingat tidak hanya mata kuliah pilihan wajib mahasiswa yang belum dibenahi, tetapi juga proses konversi kredit masing-masing jurusan masih kurang memadai dan sebagian belum terlaksana.

REFERENCES

- [1] J. S. Komputer *et al.*, “Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka,” vol. 6, no. September, pp. 916–930, 2022.
- [2] S. J. Pipin and H. Kurniawan, “Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM,” vol. 23, no. 2, pp. 197–208, 2022.
- [3] P. D. P. Silitonga, I. Sri Morina, M. Hasibuan, and U. Lestari, “Analisis Sentimen Kampus Merdeka Menggunakan Machine Learning,” *J. ICT Inf. Commun. Technol.*, vol. 21, no. 1, pp. 39–43, 2022.
- [4] R. Gunawan, R. Septiadi, F. Apri Wenando, H. Mukhtar, and Syahril, “K-Nearest Neighbor (KNN) untuk Menganalisis Sentimen terhadap Kebijakan Merdeka Belajar Kampus Merdeka pada Komentar Twitter,” *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 3, no. 2, pp. 152–158, 2022, doi: 10.37859/coscitech.v3i2.3841.
- [5] D. F. Zhafira, B. Rahayudi, and I. Indriati, “Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube,” *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
- [6] A. Rozaq, Y. Yunitasari, K. Sussolaikah, E. Resty, N. Sari, and R. I. Syahputra, “Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naïve Bayes , K-Nearest Neighbors Dan Decision Tree,” *J. Media Inform. Budidarma*, vol. 6, no. April, pp. 746–750, 2022, doi: 10.30865/mib.v6i2.3554.
- [7] H. Tuhuteru, “Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine,” vol. 4, no. 1, 2020.
- [8] M. D. Alkhussayid, “Implementasi Algoritma K-Nearest Neighbors Pada Penentuan Jurusan Siswa,” vol. 4, no. September, pp. 25–36, 2022, doi: 10.30865/json.v4i1.4759.
- [9] E. H. Harahap, L. Muflikhah, and B. Rahayudi, “Implementasi Algoritma Support Vector Machine (SVM) Untuk Penentuan Seleksi Atlet Pencak Silat,” vol. 2, no. 10, pp. 3843–3848, 2018.
- [10] B. Lampung, “DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA,” vol. 7, no. 1, pp. 1–11, 2020.
- [11] F. Reviatika *et al.*, “Analisis Klasifikasi SMS Spam Menggunakan Logistic Regression,” vol. 04, no. 03, pp. 155–160, 2021.
- [12] N. L. Chusna and M. S. Arif, “Klasifikasi SMS Spam Berbahasa Indonesia Menggunakan Algoritma Multinomial Naïve Bayes,” vol. 5, pp. 1316–1325, 2021, doi: 10.30865/mib.v5i4.3119.
- [13] F. D. Pramakrisna, F. D. Adhinata, N. Annisa, and F. Tanjung, “Aplikasi Klasifikasi SMS Berbasis Web Menggunakan Algoritma Logistic Regression Web-based Classifying SMS Application Using Logistic Regression Algorithm,” vol. 11, no. 2, pp. 90–97, 2022, doi: 10.34148/teknika.v11i2.466.