

Implementasi Algoritma SVM Non-Linear Pada Klasifikasi Analisis Sentimen Perkembangan AI di Sektor Pendidikan

Alda Nabila Putri, Aryanti Aryanti*, Sopian Soim

Jurusan Teknik Elektro, Teknik Telekomunikasi, Politeknik Negeri Sriwijaya, Palembang, Indonesia

Email: 062040352205@student.polsri.ac.id, aryanti@polsri.ac.id, sopiansoim@gmail.com

Email Penulis Korespondensi: aryanti@polsri.ac.id

Submitted: 08/07/2024; Accepted: 08/09/2024; Published: 09/09/2024

Abstrak—Seiring kemajuan teknologi, membuat pemanfaatan pada platform X atau yang dulunya Twitter semakin meluas, sehingga memungkinkan pengguna untuk bertukar pendapat mengenai berbagai topik termasuk dampak transformatif AI di sektor Pendidikan. Meskipun AI memiliki potensi yang besar dalam merevolusi kualitas dan aksesibilitas Pendidikan, di sisi lain AI juga dapat mendatangkan tantangan potensial, salah satunya ketergantungan teknologi yang berlebihan. Analisis sentimen merupakan pendekatan komputasi untuk mengidentifikasi, mengekstraksi, serta mengklasifikasi sentimen, opini, dan emosi yang diungkapkan dalam teks. Untuk mengkaji masalah tersebut, penelitian ini mengimplementasikan model Support Vector Machine Non-Linear untuk menganalisis sentimen mengenai AI di sektor pendidikan. Pada penelitian ini membangun empat model SVM dengan fungsi kernel yang berbeda yaitu kernel Linear, RBF, Polinomial, dan Sigmoid. Dengan memanfaatkan 3.000 data tweet yang dikumpulkan dari platform X dengan teknik scraping, model SVM dengan kernel polinomial berhasil menjadi best model, dengan nilai accuracy, precision, recall dan f1-score sebesar 92%. Model ini mampu mengklasifikasikan sebanyak 52,9% datatweet dengan sentimen positif dan 47,1% datatweet dengan sentimen negatif, yang menunjukkan bahwa secara umum pengguna platform X cenderung bersentimen positif terhadap perkembangan AI di sektor pendidikan.

Kata Kunci: Analisis Sentimen; Algoritma SVM-Non Linear; Platform X; Perkembangan AI; Kernel

Abstract—As technology advances, the utilization of the X platform or formerly Twitter is expanding, allowing users to exchange opinions on various topics including the transformative impact of AI in the Education sector. While AI has great potential in revolutionizing the quality and accessibility of education, it can also bring potential challenges, such as over-reliance on technology. Sentiment analysis is a computational approach to identify, extract, and classify sentiments, opinions, and emotions expressed in text. To examine the problem, this research implements a Non-Linear Support Vector Machine model to analyze sentiment about AI in the education sector. This study built four SVM models with different kernel functions, namely linear, RBF, Polynomial, and Sigmoid kernels. By utilizing 3,000 tweet data collected from platform X by scraping technique, the SVM model with polynomial kernel succeeded in becoming the best model, with accuracy, precision, recall and f1-score values of 92%. This model was able to classify 52.9% of the tweet data with positive sentiment and 47.1% of the tweet data with negative sentiment, which shows that in general, users of platform X tend to have a positive sentiment towards the development of AI in the education sector.

Keywords: Sentiment Analysis; SVM Non-Linear Algorithm; Platform X; AI Development; Kernel

1. PENDAHULUAN

Kemajuan teknologi dalam beberapa dekade terakhir, terutama di bidang Artificial Intelligence (AI) telah mengubah cara hidup dan interaksi manusia di berbagai aspek kehidupan [1]. Artificial Intelligence atau kecerdasan buatan berpotensi besar untuk merevolusi berbagai aspek kehidupan manusia, bahkan sekarang Artificial Intelligence (AI) telah mengambil peran yang semakin dominan dalam berbagai sektor, termasuk dalam sektor pendidikan. Keberadaan AI telah memberikan dampak signifikan bagi dunia Pendidikan sehingga dapat menimbulkan sentimen positif atau negatif terhadap perkembangan teknologi ini [2]. Dengan memanfaatkan Artificial Intelligence (AI), kita dapat membekali pelajar dengan pemahaman serta keterampilan yang diperlukan untuk berhasil di era digital ini. Pendidikan yang di dukung dengan Artificial Intelligence (AI) tentunya akan menumbuhkan pemikiran yang kritis, memungkinkan siswa menganalisis informasi yang kompleks serta mempersiapkan mereka secara komprehensif untuk mengatasi tantangan kedepannya [3]. Bersamaan dengan dampak positif perkembangan AI di sektor pendidikan, tentunya terdapat juga dampak negatif yang perlu diperhatikan salah satunya ketergantungan yang berlebihan pada teknologi.

Di zaman digital saat ini, platform media sosial seperti twitter yang kini dikenal sebagai X, menjadi platform yang populer. Melalui fitur hastag pengguna X dapat memantau topik terkini secara real-time [4]. Pengguna X telah signifikan mempengaruhi pandangan masyarakat terhadap berbagai isu, termasuk perkembangan AI dalam sektor pendidikan. Pembahasan pada topik tersebut tentunya memicu kontroversi di antara mereka yang mendukung dan menentang perkembangan Artificial Intelligence (AI) di sektor pendidikan [5]. Dimana sentimen tersebut dapat digunakan untuk pengambilan dataset menggunakan teknik web scraping pada aplikasi X untuk mengumpulkan dataset yang diperlukan [6].

Opinion mining atau biasa dikenal dengan analisis sentimen merupakan proses mengidentifikasi dan mengekstraksi opini atau perasaan dari sebuah text. Bertujuan untuk mengidentifikasi persepsi ataupun pandangan publik terhadap suatu topik, peristiwa, maupun isu tertentu. Dalam tugasnya analisis sentimen akan melakukan proses klasifikasi teks yang akan menempatkan teks dalam kategori positif atau negatif berdasarkan opini yang terkandung di dalamnya [7]. Analisis sentimen masyarakat terhadap kemajuan Artificial Intelligence (AI) dalam sektor pendidikan menjadi semakin krusial untuk memahami pandangan, harapan, dan kekhawatiran yang mungkin timbul

sehubungan dengan penerapan teknologi Artificial Intelligence (AI) dalam sektor pendidikan. Dengan melakukan analisis sentimen, kita dapat mengidentifikasi tanggapan masyarakat terhadap inovasi dan perubahan yang dihadirkan oleh perkembangan Artificial Intelligence (AI) di sektor pendidikan. Dalam hal ini tentunya tidak hanya dapat membantu pemangku kepentingan membuat keputusan, namun juga dapat memberikan wawasan yang mendalam mengenai hal-hal yang perlu diperhatikan dalam implementasi Artificial Intelligence (AI) di sektor pendidikan.

Algoritma SVM merupakan salah satu dari beberapa jenis algoritma yang populer dalam melakukan analisis sentimen, pada machine learning, algoritma SVM adalah metode klasifikasi yang menggunakan pola data pengajaran untuk memprediksi kelas [8]. Dikembangkan oleh Vladimir Vapnik, SVM memisahkan data ke dalam kelas yang berbeda menggunakan hyperplane. Hyperplane ini berfungsi sebagai pemisah untuk mendapatkan margin maksimum antara dua kelas data (positif dan negatif). Algoritma SVM dengan kernel linear tidak efektif untuk memisahkan data yang tidak terpisah secara linear atau memiliki hubungan kompleks [9]. Sehingga digunakan teknik non-linear untuk mengatasi keterbatasan ini. Umumnya SVM dengan teknik non-linear memanfaatkan beberapa jenis kernel untuk meningkatkan kemampuan SVM dalam memisahkan data yang sebelumnya tidak dapat dipisahkan secara linear [10].

Dalam konteks klasifikasi analisis sentimen, sudah beberapa penelitian sebelumnya diantaranya penelitian oleh Ahmad Al K, Hilda Rachmi, dan Suparni pada tahun 2023 [11] mengenai klasifikasi analisis sentimen mengenai klaim manfaat jaminan hari tua yang diuji menggunakan tiga jenis algoritma yaitu K-NN, NB, dan SVM serta penerapan PSO atau *Particle Swarm Optimization* yang memiliki nilai akurasi, presisi, recall dan AUC sebesar 82,92%, 84,53%, 80,90% dan 0,892 secara berurutan.

Pada tahun 2023 Dedy Atmajaya, Anissa Febrianti, dan Herdianti Darwis [12] melakukan penelitian mengenai analisis sentimen mengenai ChatGPT di Twitter dengan mengimplementasikan algoritma SVM dan Naive Bayes. Penelitian ini bertujuan untuk mengklasifikasikan sentimen pengguna Twitter mengenai ChatGPT ke dalam kelas positif dan negatif. Dalam penelitian ini algoritma SVM berhasil mengungguli algoritma NB dengan nilai akurasi sebesar 59%.

Pada penelitian [13] di tahun 2020 oleh Valentino Kevin Sitanayah Q, Ade Iriani, dan Hendriyanto Dwi P, melakukan klasifikasi analisis sentimen dan perbandingan akurasi algoritma SVM standar dan SVM-PSO. Berdasarkan hasil penelitian menunjukkan bahwa kombinasi algoritma SVM-PSO berhasil mengungguli SVM standar dengan nilai akurasi sebesar 96,04%.

Selanjutnya pada tahun 2023 dilakukan penelitian oleh Salsabila Rabbani, Dea Safitri, Nadila Rahmadhani, Al Amin Fadillah S, M. Khairul A [14] mengenai analisis sentimen mengenai kenaikan harga BBM serta melakukan perbandingan evaluasi kernel pada algoritma SVM. Berdasarkan hasil penelitian menunjukkan bahwa sebagian besar sentimen masuk ke dalam kelas negatif, yang berarti bersentimen negatif terhadap kenaikan harga BBM. Evaluasi pada kernel menghasilkan bahwa kombinasi algoritma SVM dan kernel RBF menjadi model terbaik dengan tingkat akurasi sebesar 87%.

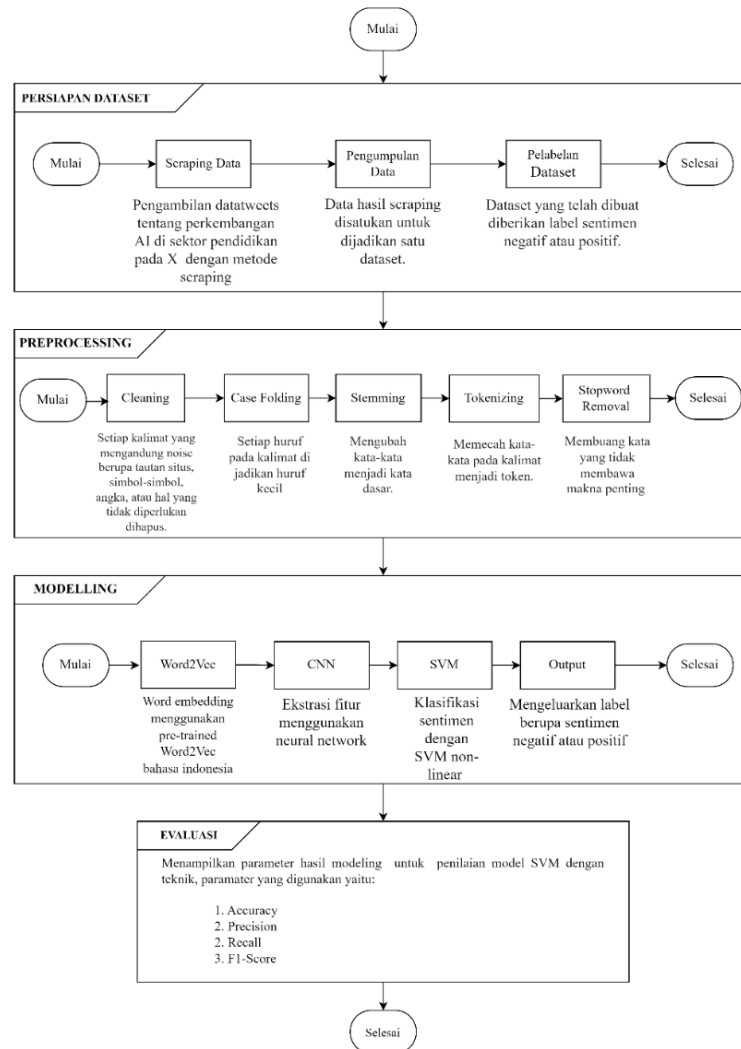
Lalu pada penelitian [15] oleh Risa Wati dan Siti Ernawati pada tahun 2021 mengimplementasikan algoritma SVM menggunakan python untuk menganalisis sentimen persepsi publik terhadap PPKM. Selain melakukan klasifikasi analisis sentimen, penelitian ini juga membandingkan kernel pada algoritma SVM yaitu dengan menggunakan jenis kernel RBF dan linear. Berdasarkan hasil penelitian algoritma SVM dengan kernel linear berhasil menjadi algoritma pengklasifikasi terbaik dengan tingkat akurasi sebesar 86%.

Dari konteks permasalahan dan analisis yang relevan, masyarakat telah berhasil memahami kecenderungan sentimen yang terjadi. Didorong oleh semakin pentingnya peran Artificial Intelligence (AI) dalam dunia pendidikan, penelitian ini bertujuan untuk mengimplementasikan model SVM non-linear untuk menganalisis sentimen pengguna pada platform X mengenai perkembangan AI di sektor pendidikan dengan menggunakan dataset yang di dapatkan dari platform X menggunakan teknik scraping. Pemilihan algoritma SVM non-linear dikarenakan kemampuannya yang tinggi dalam menangani data yang tidak dapat dipisahkan secara linear, sehingga dapat menghasilkan tingkat akurasi yang lebih akurat. Selain mengimplementasikan model SVM non-linear penelitian ini juga bertujuan untuk mengembangkan dan mengevaluasi model SVM non-linear dengan menggunakan empat jenis kernel untuk menentukan kombinasi yang memberikan kinerja terbaik dalam hal akurasi, presisi, recall, dan F1-score.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian pada algoritma SVM non-linear dalam mengklasifikasi analisis sentimen pengguna platform X terhadap perkembangan AI di sektor pendidikan tentunya melalui beberapa rangkaian proses, untuk proses lengkap mulai dari pengumpulan data hingga dengan evaluasi hasil model dapat dilihat pada Gambar 1. Pada analisis sentimen ini dilakukan pengelompokan data menjadi dua kelas, yaitu kelas dengan sentimen positif dan kelas dengan sentimen negatif. Dataset penelitian diperoleh melalui teknik scraping data tweet dengan kata kunci “AI di sektor pendidikan” pada platform X.



Gambar 1. Tahapan Penelitian

2.2 Persiapan Dataset

Tahapan awal dalam penelitian ini adalah persiapan dataset yang dilakukan dengan menggunakan teknik scraping melalui platform X atau Twitter dengan mengumpulkan tweet yang membahas mengenai AI di sektor pendidikan. Setelah proses scraping, datateks yang diperoleh akan disatukan untuk membentuk sebuah dataset mentah yang komprehensif dalam bentuk CSV (*Comma Separated Values*). Lalu dataset tadi akan diberikan label kelas positif dan negatif dengan memanfaatkan teknik otomatisasi seperti *Natural Language Processing* (NLP).

2.3 Tahap Preprocessing

Setelah dataset terkumpul dan diberikan label, tahap selanjutnya yaitu preprocessing yang berfungsi untuk membersihkan dan mempersiapkan data yang akan digunakan agar lebih terstruktur. Adapun beberapa tahapan dalam proses preprocessing adalah sebagai berikut:

a. Cleaning

Pada tahap ini dilakukan proses pembersihan teks seperti menghapus HTML dan URL, simbol-simbol, angka, mentions & hastags, dan kata-kata atau elemen yang tidak diperlukan untuk memastikan teks siap untuk tahap pemodelan.

b. Case Folding

Setiap karakter huruf dalam teks akan diubah menjadi huruf kecil guna untuk meningkatkan konsistensi.

c. Stemming

Pada tahapan ini akan menghilangkan imbuhan pada kata sehingga akan menghasilkan kata dasar.

d. Tokenizing

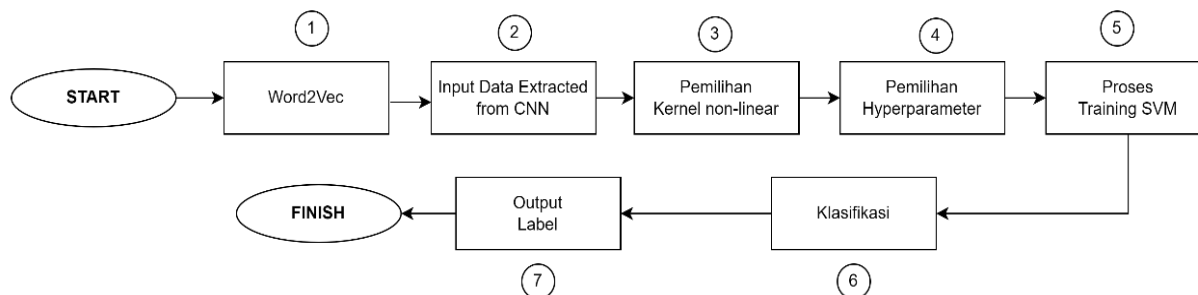
Kalimat akan dipisah menjadi kata-kata (token) yang lebih ringkas. Proses tokenizing ini dilakukan dengan mengidentifikasi setiap spasi dalam kalimat, sehingga memungkinkan pemisahan kata-kata secara efektif.

e. Eliminasi Stopword

Pada tahap ini, kata-kata yang tidak memberikan kontribusi tambahan terhadap analisis teks akan dihilangkan.

2.4 Modelling

Tahap modelling merupakan tahapan inti dari proses analisis sentimen ini, beberapa tahapan dilakukan untuk memastikan bahwa dataset yang sudah diproses ditahapan sebelumnya dapat digunakan dengan efektif untuk membangun model klasifikasi yang akurat.

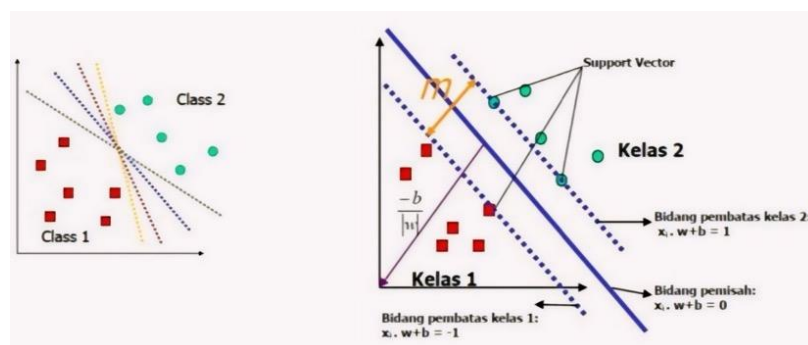


Gambar 2. Tahapan Modelling

Pada tahap ini, Word2Vec digunakan untuk mengubah kata-kata dalam dataset menjadi vektor numerik yang mewakili kata-kata tersebut [16]. Vektor-vektor ini menjadi input untuk Convolutional Neural Network (CNN) yang bertugas mengekstraksi fitur penting dari teks. Fitur-fitur ini kemudian diklasifikasi oleh Support Vector Machine (SVM). Pemilihan kernel non-linear bertujuan untuk memindahkan data ke dimensi yang lebih ringkas atau memisahkan data yang tidak dapat dipisahkan secara linear [17]. Penelitian ini menggunakan empat jenis kernel: linear, radial basis function, polinomial, dan sigmoid sedangkan hyperparameter akan dipilih sesuai dengan kernel yang digunakan. Setelah proses pelatihan model SVM digunakan untuk mengklasifikasi sentimen sehingga menghasilkan output berupa label positif dan negatif.

2.5 Support Vector Machine

Algoritma SVM merupakan algoritma yang berkerja dengan cara mencari hyperplane atau bidang pembatas yang optimal untuk memisahkan dua kelas secara maksimal dalam ruang berdimensi tinggi [18]. Dalam model SVM, support vector merujuk pada objek data yang paling dekat dengan hyperplane atau bidang pembatas. SVM mengoptimalkan hyperplane ini dengan mempertimbangkan hanya support vector, yang memungkinkan algoritma bekerja lebih efisien karena mengabaikan sebagian besar objek data yang tidak terlibat langsung dalam menentukan hyperplane optimal. Melalui pendekatan ini memungkinkan SVM akan bekerja dengan lebih efisien.



Gambar 3. Hyperplane Support Vector Machine (SVM)

Dalam proses klasifikasi menggunakan model SVM, terkadang kernel linear tidak dapat memisahkan data secara optimal, yang mengakibatkan kinerja klasifikasi yang kurang baik. Hal tersebut biasanya disebabkan karena teknik linier tidak mampu menangani data yang memiliki pola kompleks atau tidak teratur. Untuk mengatasi permasalahan tersebut penggunaan teknik non-linear dapat menjadi solusi. Dengan menggunakan kombinasi kernel dalam model SVM, memungkinkan untuk mengubah data ke ruang dimensi yang lebih tinggi, sehingga dapat meningkatkan kemampuan pemisahan antar kelas dan meningkatkan kinerja keseluruhan model SVM.

2.6 Evaluasi

Evaluasi dilakukan dengan menggunakan metode *confusion matrix* untuk menilai kinerja model dalam mengklasifikasikan sentimen. *Confusion matrix* memberi kan informasi detail tentang hasil klasifikasi yang dihasilkan oleh model, termasuk jumlah prediksi yang benar dan kesalahan yang terjadi untuk setiap kelas. [19]. Dalam pengujian, kelas yang diidentifikasi sebagai positif terdiri dari True Positive (TP) dan False Negative (FN), sementara kelas yang diidentifikasi sebagai negatif terdiri dari True Negative (TN) dan False Positive (FP). Berdasarkan nilai-

nilai inilah, *confusion matrix* digunakan untuk menghitung *accuracy*, *precision*, *recall*, dan *f1-score* sebagai metrik evaluasi kinerja model [20].

Untuk menghitung nilai akurasi, dilakukan pengukuran proporsi prediksi benar secara keseluruhan. Presisi mengukur seberapa baik model dalam mengidentifikasi data yang sebenarnya positif dari semua hasil yang diprediksi sebagai positif. Recall digunakan untuk seberapa baik model dalam mengidentifikasi semua kasus positif yang sebenarnya ada. Sedangkan, F1-score digunakan untuk mengukur rata-rata harmonik dari presisi dan recall, memberikan ukuran kinerja yang seimbang antara keduanya.. Berikut ini rumus-rumus untuk perhitungan dari *confusion matrix*, sebagai berikut:

- a. *Precision* dihitung dengan membagi jumlah True Positive oleh jumlah True Positive ditambah jumlah False Positive dan dikali 100%.

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{1}$$

- b. *Recall* dihitung dengan membagi jumlah True Positive oleh jumlah True Positive ditambah jumlah False Negative dan dikali 100%.

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{2}$$

- c. *F1-Score* dihitung dengan mengkali dua hasil kali presisi dan recall, kemudian dibagi dengan jumlah presisi dan recall yang kemudian dikalikan dengan 100%.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \tag{3}$$

3. HASIL DAN PEMBAHASAN

Pada bagian ini berisikan pemaparan dari hasil penelitian di setiap tahapannya dimulai dari hasil persiapan dataset hingga ke tahap evaluasi yang akan dipresentasikan dengan menggunakan gambar, tabel, dan narasi.

3.1 Persiapan Data

Setelah data berhasil terkumpul melalui teknik scraping pada platform X, kemudian data akan di simpan dengan format CSV. Data tweet yang berhasil diperoleh yaitu sebanyak 3000 tweet dengan pembagian 80% untuk data training (2400 sample), 10% untuk data validation (300 sample), dan 10% untuk data testing (300 sample) dengan rentang waktu januari 2022 - januari 2024. Berikut tabel 1 yang menampilkan contoh datatweet mengenai sentimen perkembangan AI.

Tabel 1. Data Tweet

Username	Teks Tweet
XXX	SINDOnews@SINDOnews·27 AprKecerdasan Buatan Siap Berintegrasi dalam Dunia PendidikanKecerdasan Buatan Siap Berintegrasi dalam Dunia PendidikanDari tekno.sindonews.com150
XXX	Kompas Klasika@KompasKlasika·25 AprKecerdasan buatan terus berkembang dan diadopsi dalam pendidikan. Pemanfaatannya tetap dalam tujuan pendidikan. http://dlvr.it/T5zx4g #Humaniora #AdadiKompasPendidikan Tinggi Didorong Adopsi dan Integrasikan Kecerdasan BuatanDari kompas.id34

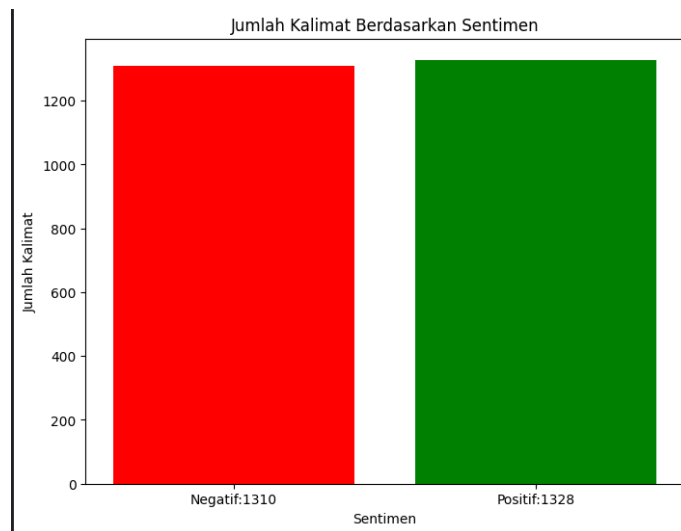
Setelah data dikumpulkan, data akan diberikan label positif dan negatif sesuai dengan kelasnya. Berikut tabel 2 yang menampilkan beberapa contoh data yang telah diberikan label “positif” dan “negatif”.

Tabel 2. Contoh Data Yang di Beri Label

Text	Label
AI dapat membantu guru untuk mempersonalisasi pembelajaran dan memberikan umpan balik yang lebih cepat kepada siswa	Positif
Pengguna AI di pendidikan dapat meningkatkan akses pendidikan ke daerah terpencil	Positif
Namun kekhawatiran AI dapat menyebabkan hilangnya pekerjaan bagi guru	Negatif

Berikut ini histogram yang menampilkan distribusi jumlah kalimat sentimen berdasarkan dengan kategori kelas yang berbeda. Sumbu horizontal (X) merepresentasikan jenis sentimen, sementara sumbu vertikal (Y) menunjukkan

jumlah kalimat yang termasuk dalam masing-masing kategori. Histogram ini menggunakan dua warna untuk membedakan antara kedua kategori sentimen, yaitu merah untuk sentimen negatif dan hijau untuk sentimen positif.



Gambar 4. Histogram Jumlah Data Berdasarkan Kelasnya

Dari histogram ini, terlihat bahwa terdapat 1.310 kalimat yang tergolong dalam kategori sentimen negatif dan 1.328 kalimat yang tergolong dalam kategori sentimen positif. Meskipun jumlah kalimat dalam kedua kategori ini hampir sama, kalimat dengan sentimen positif sedikit lebih banyak daripada yang negatif, dengan selisih 18 kalimat. Hal ini menunjukkan bahwa data yang dianalisis memiliki distribusi sentimen yang relatif seimbang antara positif dan negatif.

3.2 Preprocessing

Tahap preprocessing dilakukan untuk mengubah data mentah atau belum diproses menjadi dataset yang siap digunakan. Dalam tahap ini, data bisa dimurnikan, diubah bentuknya, atau diolah agar dapat dianalisis lebih lanjut. Preprocessing terdiri dari beberapa tahapan, diantaranya pembersihan (cleaning), normalisasi huruf (case folding), stemming, tokenisasi (tokenizing), dan penghapusan stopwords (stopword removal). Pada tabel 3 menampilkan setiap tahapan dan hasil dari preprocessing.

Tabel 3. Tahapan Preprocessing

Proses	Output
<i>Teks Asli</i>	Antara Bengkulu@ANTARABengkulu-20 MarPerusahaan pelopor pemanfaatan teknologi kecerdasan buatan dalam pendidikan bahasa Indonesia http://Vokal.ai menyediakan aplikasi http://ngaji.ai untuk membantu pengguna belajar membaca Al Quran secara mandiri.Ngaji.ai bantu upaya belajar baca Al Quran secara mandiri - ANTARA News BengkuluDari bengkulu.antaranews.com90
<i>Cleaning</i>	Antara Bengkulu 20 Mar Perusahaan pelopor pemanfaatan teknologi kecerdasan buatan dalam pendidikan bahasa Indonesia menyediakan aplikasi untuk membantu pengguna belajar membaca Al Quran secara mandiri Ngajiai bantu upaya belajar baca Al Quran secara mandiri ANTARA News Bengkulu Dari bengkuluantaranewscom90
<i>Case Folding</i>	antara bengkulu 20 mar perusahaan pelopor pemanfaatan teknologi kecerdasan buatan dalam pendidikan bahasa indonesia menyediakan aplikasi untuk membantu pengguna belajar membaca al quran secara mandiri ngajiai bantu upaya belajar baca al quran secara mandiri antara news bengkulu dari bengkuluantaranewscom90
<i>Stemming</i>	antara bengkulu 20 mar usaha pelopor manfaat teknologi cerdas buat dalam didik bahasa indonesia sedia aplikasi bantu guna ajar baca al quran cara diri ngajiai bantu upaya ajar baca al quran cara diri antara news bengkulu dari bengkuluantaranewscom90
<i>Tokenizing</i>	'antara', 'bengkulu', '20', 'mar', 'usaha', 'pelopor', 'manfaat', 'teknologi', 'cerdas', 'buat', 'dalam', 'didik', 'bahasa', 'indonesia', 'sedia', 'aplikasi', 'bantu', 'guna', 'ajar', 'baca', 'al', 'quran', 'cara', 'diri', 'ngajiai', 'bantu', 'upaya', 'ajar', 'baca', 'al', 'quran', 'cara', 'diri', 'antara', 'news', 'bengkulu', 'dari', 'bengkuluantaranewscom90'

Proses	Output
Stopword Removal	'bengkulu', '20', 'mar', 'usaha', 'pelopor', 'manfaat', 'teknologi', 'cerdas', 'buat', 'didik', 'bahasa', 'indonesia', 'sedia', 'aplikasi', 'bantu', 'guna', 'ajar', 'baca', 'quran', 'cara', 'diri', 'ngajiai', 'bantu', 'upaya', 'ajar', 'baca', 'quran', 'cara', 'diri', 'news', 'bengkulu', 'bengkuluantaraneewscom90'

Setelah itu dilakukan proses encoding yang bertujuan untuk mengubah teks atau karakter menjadi representasi digital yang dapat diproses oleh komputer. Pada tabel 3 menyajikan hasil dari proses encoding.

Teks	Label
dwik rizkidwika jan menurut gue sih ya kita itu nggak perlu ndakikndakiik metaverse robot ai inainu kucrut kalo yang fundamental di bidang pendidikan formal kek begini aja masih kusut	0
Harian metrohmetromyjkualalumpur pixlr group menjalin kerjasama dengan kementerian pendidikan kpm untuk menyediakan kursus kecerdasan buatan ai percuma kepada orang guru sempena sambutan hari guru pixlr sedia kursus ai percuma kepada guru harian metrodari hmetrocommy rb	1

Dalam proses encoding memungkinkan teks untuk direpresentasikan dalam bentuk biner (0 dan 1) dimana 0 untuk label pada sentimen negatif dan 1 untuk label pada sentimen positif yang dapat disimpan, diproses, atau dikirimkan oleh komputer. Dengan adanya proses encoding ini dapat memastikan data dapat disimpan dan diproses dengan cara yang lebih efisien, mengurangi kesalahan dan meningkatkan kinerja sistem.

3.3 Hasil Evaluasi

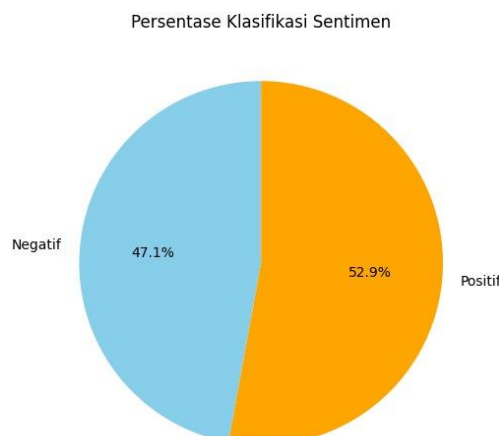
Pada tabel 4 menyajikan hasil dari pengukuran performa klasifikasi terhadap algoritma SVM dengan tiap-tiap jenis kernel (Linear, RBF, Polinomial, dan Sigmoid) meliputi empat buah parameter yaitu *accuracy*, *precision*, *recall* dan *f1-score*.

Tabel 4. Hasil Pengukuran Performa Algoritma SVM

Parameter Performa	Linear	RBF	Polinomial	Sigmoid
<i>Accuracy</i>	89%	90%	92%	89%
<i>Precision</i>	89%	90%	92%	89%
<i>Recall</i>	89%	90%	92%	89%
<i>F1-Score</i>	89%	90%	92%	89%

Berdasarkan hasil pengujian algoritma SVM dalam mengklasifikasi analisis sentimen perkembangan AI pada sektor pendidikan yang ditampilkan pada tabel 2, didapatkan hasil bahwa algoritma SVM dengan teknik non linear melalui kernel polinomial menjadi model terbaik dalam penelitian ini dibandingkan dengan kernel lainnya. Model ini berhasil mengklasifikasi sentimen terhadap perkembangan AI dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* sebesar 92%. Secara keseluruhan, hasil penelitian ini cukup menegaskan bahwa penggunaan teknik non-linear pada algoritma SVM memberikan hasil yang optimal dalam analisis sentimen perkembangan AI di sektor pendidikan.

Selanjutnya, klasifikasi sentimen terhadap perkembangan AI di sektor pendidikan ditentukan menggunakan model terbaik yaitu algoritma SVM dengan kernel polinomial. Hasil klasifikasi ini akan divisualisasikan dengan bentuk diagram lingkaran (*pie chart*) seperti pada Gambar 4 untuk memudahkan pemahaman terhadap distribusi sentimen.



Gambar 5. Pie Chart Hasil Persentase Klasifikasi Sentimen

Pada Gambar 4 menyajikan bahwa sebanyak 52,9% dari dataset bersentimen positif terhadap perkembangan AI di sektor pendidikan, hal tersebut berarti bahwa mayoritas data *tweet* dalam analisis ini memberikan tanggapan atau opini yang mendukung terhadap topik perkembangan AI di sektor pendidikan. Sedangkan pada bagian berwarna biru menunjukkan sebanyak 47,1% dari dataset bersentimen negatif terhadap perkembangan AI di sektor pendidikan.

4. KESIMPULAN

Setelah dilakukan analisis terhadap algoritma SVM dapat disimpulkan bahwa algoritma SVM dengan teknik non-linear mampu meningkatkan performa klasifikasi dan akurasi. Hasil menunjukkan nilai akurasi tertinggi dihasilkan oleh algoritma SVM dengan kernel polinomial yaitu sebesar 92%, selain itu nilai untuk precision, recall, dan f1-score juga masing-masing sebesar 92%, hal ini menandakan bahwa algoritma SVM dengan kernel polinomial tidak hanya unggul dalam hal keakuratan, tetapi juga mampu menjaga keseimbangan antara data yang diklasifikasikan dengan benar dan kesalahan yang minimal. Dengan menggunakan model terbaik yaitu algoritma SVM dengan kernel polinomial mampu melakukan prediksi sentimen terhadap 52,9% teks *tweet* dari dataset bersentimen positif serta sebanyak 41,7% teks *tweet* dari dataset bersentimen negatif. Hasil tersebut menunjukkan bahwa sebagian besar dari pengguna aplikasi X cenderung bersentimen positif terhadap perkembangan Artificial Intelligence (AI) di sektor pendidikan. Penelitian ini tidak hanya memberikan gambaran tentang bagaimana algoritma SVM dapat digunakan secara efektif dalam analisis sentimen, tetapi juga memberikan kontribusi penting dalam memahami bagaimana AI mempengaruhi pandangan publik di sektor yang krusial seperti pendidikan. Diharapkan, hasil penelitian ini dapat menjadi referensi bagi pengambil kebijakan, akademisi, serta pengembang teknologi dalam mengoptimalkan penerapan AI di bidang pendidikan, sekaligus memperkaya literatur tentang pemanfaatan teknik machine learning dalam analisis sentimen sosial.

REFERENCES

- [1] L. Gkinko and A. Elbanna, "The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users," *Int J Inf Manage*, vol. 69, Apr. 2023, doi: 10.1016/j.ijinfomgt.2022.102568.
- [2] S. Liza Zahara, Z. Ula Azkia, and M. Minan Chusni Program Studi Pendidikan Fisika UIN Sunan Gunung Djati Bandung, "Implementasi Teknologi Artificial Intelligence (AI) dalam Bidang Pendidikan," *JPSP: Jurnal Penelitian Sains dan Pendidikan*, vol. 3, 2023.
- [3] G. Yulianti, Benardi, N. Permana, and F. A. K. Wijayanti, "Transformasi Pendidikan Indonesia: Menerapkan Potensi Kecerdasan Buatan (AI)," *JISMA: Journal of Information Systems and Management*, vol. 2, no. 6, Des. 2023.
- [4] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *SMATIKA*, vol. 10, no. 2, Des. 2020.
- [5] T. W. Putra, A. Triayudi, and Andrianingsih, "View of Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, dan Decision Tree," *JTIK: Jurnal Teknologi Informasi dan Komunikasi*, vol. 6, no. 1, 2022.
- [6] M. Khoirul, U. Hayati, and O. Nurdianan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna di Google Play Menggunakan Algoritma Naive Bayes," *JATI: Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, Feb. 2023.
- [7] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.
- [8] Q. A. Chairunnisa, Y. Herdiyeni, M. Kusuma, D. Hardhienata, and J. Adisantoso, "Analisis Sentimen Pengguna Twitter Terhadap Program Vaksinasi Covid-19 di Indonesia Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Twitter Users on COVID-19 Vaccination Program in Indonesia using Support Vector Machine Algorithm," *JIKA: Jurnal Ilmu Komputer Agri-Informatika*, vol. 9, no. 1, 2022.
- [9] M. Azimi-Pour, H. Eskandari-Naddaf, and A. Pakzad, "Linear and non-linear SVM prediction for fresh properties and compressive strength of high volume fly ash self-compacting concrete," *Constr Build Mater*, vol. 230, Jan. 2020, doi: 10.1016/j.conbuildmat.2019.117021.
- [10] F. Nie, W. Zhu, and X. Li, "Decision Tree SVM: An extension of linear SVM for non-linear classification," *Neurocomputing*, vol. 401, pp. 153–159, Aug. 2020, doi: 10.1016/j.neucom.2019.10.051.
- [11] A. Al Kaafi, H. Rachmi, and T. Informasi, "Algoritma Klasifikasi pada Analisis Sentimen Klaim Manfaat Jaminan Hari Tua Comparison of Classification Algorithms toward Sentiment Analysis on Claim of Old Age Guaranteed," *SISTEMASI: Jurnal Sistem Informasi Komparasi*, vol. 12, no. 1, Jan. 2023.
- [12] D. Atmajaya, A. Febrianti, H. Darwis, "View of Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *IJCS: Indonesian Journal of Computer Science*, vol. 12, no. 4, 2023.
- [13] V. K. S. Que, A. Iriani, and H. D. Purnomo, "Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization (Online Transportation Sentiment Analysis Using Support Vector Machine Based on Particle Swarm Optimization)," *Jurnal Nasional Teknik Elektro dan Teknik Informasi*, vol. 9, no. 2, Mei 2020.
- [14] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, Oct. 2023, doi: 10.57152/malcom.v3i2.897.
- [15] R. Wati and S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," *JTIUST: Jurnal Teknik Informatika Unika St. Thomas*, vol. 6, no. 2, Des. 2021.
- [16] T. S. Sabrila, V. R. Sari, and A. E. Minarno, "Analisis Sentimen Pada Tweet Tentang Penanganan Covid-19 Menggunakan Word Embedding Pada Algoritma Support Vector Machine Dan K-Nearest Neighbor," *Fountain of Informatics Journal*, vol. 6, no. 2, p. 69, Jul. 2021, doi: 10.21111/fij.v6i2.5536.



- [17] T. Meisya, P. Aulia, N. Arifin, and R. Mayasari, “Perbandingan Kernel Support Vector Machine (SVM) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19”, *SINTECH: Science And Information Technology*, vol. 4, no. 2, Okt. 2021, doi: <https://doi.org/10.31598>.
- [18] H. C. Husada and A. S. Paramita, “Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *Teknika*, vol. 10, no. 1, pp. 18–26, Feb. 2021, doi: 10.34148/teknika.v10i1.311.
- [19] T. Safitri, Y. Umaidah, and I. Maulana, “Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine,” *JAIC: Journal of Applied Informatics and Computing*, vol. 7, no. 1, Jul. 2023.
- [20] N. Fitriyah, B. Warsito, D. Asih, and I. Maruddani, “Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM),” *JURNAL GAUSSIAN*, vol. 9, no. 3, pp. 376–390, 2020.