



# Sentimen Analisis Masyarakat Indonesia Terhadap Presiden Rusia Pada Komentar Media Berita *Online*

Ihud Hafid <sup>1\*</sup>, Windu Gata <sup>2</sup>, Khairunisa Hilyati <sup>3</sup>, Valianda Farradillah Hakim <sup>4</sup>, Sri Rahayu <sup>5</sup>

<sup>1,2,3,4,5</sup> Fakultas Teknologi Informasi, Universitas Nusa Mandiri.

## article info

### Article history:

Received 30 June 2022

Received in revised form

19 October 2022

Accepted 20 November 2022

Available *online* January 2023

### DOI:

<https://doi.org/10.35870/jti.k.v7i1.698>

### Keywords:

Text Mining; Vladimir Putin; Sentiment Analysis; NB; SVM; PSO.

### Kata Kunci:

Text Mining; Vladimir Putin; Sentimen Analisis; NB; SVM; PSO.

## abstract

Russia's invasion of Ukraine was criticized by various parties, including from Indonesia. The attitude shown by the Indonesian government is not the same as the response of the Indonesian people based on various comments on online news media pages. Comments by online news readers are used as an assessment of the Russian President who is involved in the conflict between Russia and Ukraine in the form of sentiment analysis. This study succeeded in obtaining data as many as 352 comments from one of the online news media, the data had previously gone through the cleansing stage to eliminate duplication. To get basic information on comments, Text mining and Text Pre-Processing become an important part of the process. The algorithm used in this research is the Naive Bayes (NB) and Support Vector Machine (SVM) classification algorithm which is optimized using Particle Swarm Optimization (PSO). The two algorithms were tested and gave the result that PSO-based SVM got the best accuracy, which was 79.90% and AUC 0.901.

## abstrak

Tindakan Invasi Rusia ke Ukraina dikecam oleh berbagai pihak termasuk dari Indonesia. Sikap yang ditunjukkan oleh pemerintah Indonesia ternyata tidak sama dengan respon masyarakat Indonesia berdasarkan berbagai macam komentar pada halaman media berita online. Komentar yang pembaca berita online digunakan sebagai penilaian terhadap tokoh Presiden Rusia yang terlibat dalam konflik antara Rusia dan Ukraina dalam bentuk analisis sentimen. Penelitian ini berhasil mendapatkan data sebanyak 352 komentar yang berasal dari salah satu media berita online, data tersebut sebelumnya telah melalui tahap cleansing untuk menghilangkan duplikasi. Untuk mendapatkan informasi dasar pada komentar, Text mining dan Pre-Processing Text menjadi bagian penting dalam prosesnya. Algoritma klasifikasi yang digunakan oleh peneliti yaitu Support Vector Machine (SVM) dan Naive Bayes (NB) yang dioptimalkan menggunakan Particle Swarm Optimization (PSO). Kedua algoritma itu diuji dan memberikan hasil bahwa SVM berbasis PSO mendapatkan hasil akurasi terbaik yaitu sebesar 79.90% dan AUC 0.901.

\*Corresponding author. Email: 14210208@nusamandiri.ac.id <sup>1\*</sup>.

## 1. Latar Belakang

Sehubungan dengan keinginan Ukraina yang ingin bergabung ke NATO diiringi dengan Rusia yang menolaknya menjadikan Rusia dan Ukraina menjadi perhatian dunia selama beberapa bulan terakhir ini. Alasan Rusia menolak adalah jika Ukraina bergabung dengan NATO sebagai *existential threat* akan mengancam Rusia dan eksistensi *Collective Security Treaty Organization* - aliansi militer yang menaungi Rusia dan 8 negara sekutunya. Tindakan Rusia yang menempatkan pasukan di jalur perbatasannya dengan Ukraina membuat suasana semakin tegang, terlebih ketika Rusia mengakui kemerdekaan Luhansk dan Donetsk kemudian pada Kamis 24 Februari 2022 yang lalu membuat pengumuman perintah operasi militer di timur Ukrain [1].

Ketika Rusia menyerang sejumlah kota di Ukraina secara besar-besaran, ketika itulah invasi dimulai. Tindakan invasi itu dikecam oleh berbagai pihak, salah satunya adalah Indonesia yang mendukung kecaman terhadap agresi Rusia melalui Resolusi PBB. Akan tetapi reaksi warganet Indonesia tidak sesuai dengan sikap pemerintah. Masyarakat Indonesia memberikan reaksi yang berbeda dalam menanggapi tindakan invasi Rusia ke Ukraina, ada yang cenderung mendukung invasi tersebut, ada yang malah menyukai sosok pemimpin kuat dan tegas seperti Presiden Rusia Vladimir Putin yang dinilai sebagai tokoh yang gagah, tegas, kuat serta nasionalis, hal itu lah yang menjadikan sosok Presiden Rusia itu menjadi salah satu tokoh dunia yang ramai dipicarakan pada media sosial. Topik ini menarik untuk mengklasifikasikan sentiment masyarakat Indonesia [2].

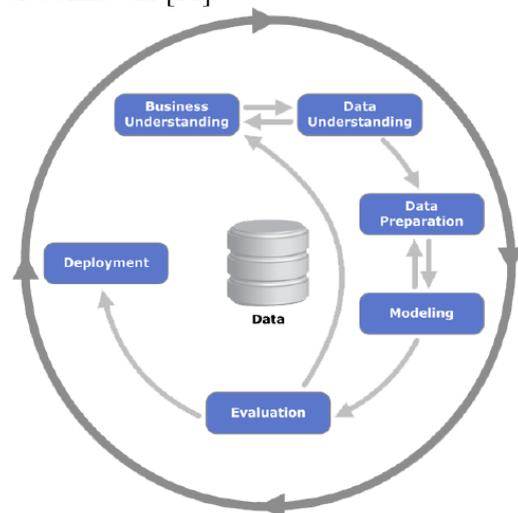
Analisis sentiment merupakan proses perbandingan untuk mengkategorikan pendapat terhadap suatu topik apakah positif, negatif atau netral, selain itu juga merupakan juga untuk menganalisis pendapat, sentiment dan penilaian emosional pada topik yang dibahas [3]. Penelitian ini menerapkan *Term Frequency Inverse Document Frequency* (TF-IDF) sebagai pembobot berdasarkan jumlah kemunculan istilah pada data [4][5]. Analisis sentiment yang dilakukan pada penelitian ini menggunakan Teknik *Machine Learning* yang menggunakan data dari teks komentar pada media berita online di Indonesia. Text Mining adalah penemuan pengetahuan baru di database

dalam bentuk tekstual [6]. *Text mining* merupakan penerapan konsep dan teknik *data mining* yang bertujuan untuk memproses penganalisisan teks guna memperoleh informasi dengan tujuan tertentu kemudian akan diolah dengan berbagai metode [7].

Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes* (NB) lalu untuk pengklasifikasian serta menerapkan *Particle Swarm Optimization* (PSO) untuk mengoptimalkannya. PSO sendiri memiliki konsep sederhana, konvergensi cepat, mudah diimplementasikan dan mudah di terapkan diberbagai bidang untuk pengoptimalisasian [8]. Penelitian ini menggunakan SVM karena memiliki konsep yang lebih matematis dan menggunakan NB karena dikenal dengan kelebihanannya yaitu cepat, sederhana dan memiliki akurasi yang tinggi [9][10]. Tujuan dari penelitian ini adalah untuk mendapatkan metode terbaik untuk menganalisis sentiment dengan kombinasi tahapan *preprocessing* dengan menggunakan *stopword* yang diharapkan menghasilkan peningkatan akurasi pada algoritma yang digunakan. Selain itu untuk melihat sentimen masyarakat Indonesia terhadap tindakan Vladimir Putin sebagai Presiden Rusia yang melakukan tindakan invasi Rusia ke Ukraina pada komentar pembaca di halaman media berita *online Kompas.com*.

## 2. Metode Penelitian

Peneliti menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai metode penelitian yang terdiri dari 6 tahapan seperti pada gambar berikut in [11]:



Gambar 1. Metode Penelitian

### *Business Understanding*

Pada tahap *Business Understanding*, diawali dengan mencari informasi melalui situs Kompas.com dengan objek yang ditujukan untuk Vladimir Putih. Media *online* tidak hanya sebagai media membaca berita, namun juga dapat digunakan untuk melihat masalah yang terjadi dan untuk menganalisis sentimen terhadap tokoh tertentu [12]. Penelitian ini menggunakan *feature selection* untuk meningkatkan akurasi klasifikasi. Untuk pengambilan data komentar menggunakan teknik *Web Scraping. Tools* yang digunakan untuk proses ini yaitu WebHarvy [13].

### *Data Understanding*

Pada tahapan ini, data yang telah di peroleh akan melalui tahapan *Pre-Processing Data*. Data diambil mulai dari tanggal 24 Februari 2022 hingga 31 Mei 2022 dan berhasil mendapatkan 644 data dari [www.kompas.com](http://www.kompas.com), setelah itu akan dilakukan proses *cleansing* dan menyisakan 352 data dengan bahasa indonesia yang siap digunakan. Tahapan *Pre-Processing* antara lain adalah *Transform Case, Tokenize, Filter Tokenize dan Stopword* [14].

### *Data Preparation*

Data yang siap digunakan akan di proses menggunakan Rapid Miner, aplikasi tersebut adalah aplikasi *freeware* yang terdapat berbagai macam metode pengolahan data yang siap untuk digunakan secara mudah [15]. Pada Rapid Miner, tahapan *Transform Case* adalah mengubah huruf kecil agar mempermudah proses *Tokenize* nantinya. Kemudian tahapan *Tokenize* dimana akan dilakukan proses pemotongan string input berdasarkan setiap kata. Setelah itu ada proses *Filter Tokenize* dimana akan menghilangkan karakter tanda baca dan dibatasi panjang teks pada data. Terakhir adalah tahapan *Stopword*, dimana data sebelum di klasifikasikan dihilangkan terlebih dahulu teks yang tidak berhubungan dengan topik penelitian tanpa mengurangi bobot kalimat tersebut [16].

### *Modeling*

Penelitian ini menggunakan 2 algoritma klasifikasi yaitu, *NB* dan *SVM* dengan optimasi PSO dan untuk pemodelan menggunakan Rapid Miner versi 9.10. Setiap model akan melalui pengujian pengklasifikasian positif dan negatif bertujuan

mendapat nilai akurasi terbaik masing-masing algoritma yang digunakan.

### *Evaluation*

Pada tahapan *Evaluation*, penelitian ini menggunakan *10 Fold-Cross Validation* yang terdiri dari dua *subproses* yaitu data pelatihan (*training set*) dan data pengujian (*testing set*). Pada *subproses training set*, data digunakan untuk menguji model algoritma yang ditentukan dengan *data training* yang tersedia. Setelah seluruh model algoritma melalui proses *data training*, akan dilanjutkan proses *data testing*. Hal ini bertujuan untuk mendapatkan model algoritma yang sesuai dengan tujuan penelitian.

### *Deployment*

Pada tahapan ini, data akan di implementasikan menggunakan tools, yaitu Rapid Miner. Hal ini mewakili proses yang telah dilakukan sebelumnya, yaitu pemodelan dengan menggunakan Rapid Miner yang digunakan untuk penelitian.

## 3. Hasil dan Pembahasan

Dataset yang digunakan pada penelitian ini bersumber dari komentar tentang Vladimir Putin, proses pengambilan data mentah berdasarkan atribut yang diperlukan. Pengumpulan data dimulai pada tanggal 24 Februari 2022, hingga 31 Mei 2022 dan menghasilkan data primer sebanyak 644 data komentar dari media berita online [www.kompas.com](http://www.kompas.com). Proses selanjutnya adalah pembersihan data dan menyisakan 352 komentar. Pada penelitian ini, bahasa yang digunakan adalah Bahasa Indonesia, maka komentar yang diambil adalah komentar yang menggunakan Bahasa Indonesia.

Dalam proses mengolah data untuk mendapatkan model yang sesuai dengan kasus penelitian ini yaitu, sentimen analisis Vladimir Putin dengan menggunakan algoritma klasifikasi *NB* dan *SVM* digunakanlah sebuah *tools* Rapidminer versi 9.10. Penelitian ini termasuk dalam bagian *text mining* maka dari itu terdapat tahapan yang perlu di lakukan sebelum mendapatkan model yang terbaik.

### 1) *Pre-Processing*

Pada tahap *Pre-Processing* merupakan awal pengolahan data sebelum diklasifikasikan dengan algoritma *NB* dan *SVM* kemudian dioptimalkan dengan PSO. Berikut adalah tahapan *Pre-Processing*

yang diterapkan pada penelitian ini:

2) *Transform Cases*

Tahapan ini merupakan proses mengubah teks menjadi huruf kecil sehingga saat dilakukan proses klasifikasi terdapat keseragaman huruf untuk menghindari kesalahan dalam proses *Tokenize*.

3) *Tokenizing*

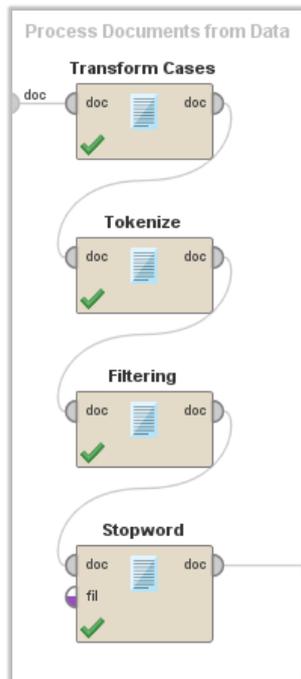
Tahapan ini adalah proses pemecahan kalimat menjadi potongan karakter sesuai dengan kebutuhan atau biasa disebut sebagai token, kata yang diambil adalah kata-kata dengan minimal 3 karakter dan tidak lebih dari 25 karakter.

4) *Filtering*

Tahapan *Filtering* merupakan tahapan membuang kata-kata yang kurang penting atau merupakan kata-kata tidak deskriptif seperti “yang”, “dan”, “di”, “dari”.

5) *Stopword*

Pada proses ini, kata-kata yang kurang relevan dihilangkan menggunakan *Indonesian Stopword Removal*.

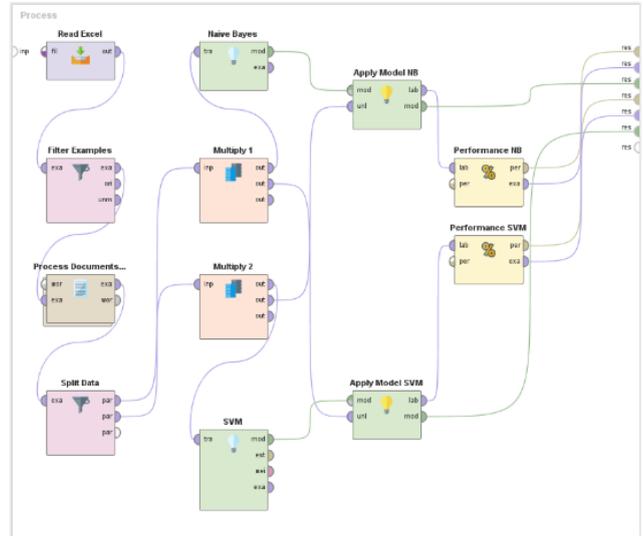


Gambar 2. Model Processing Data

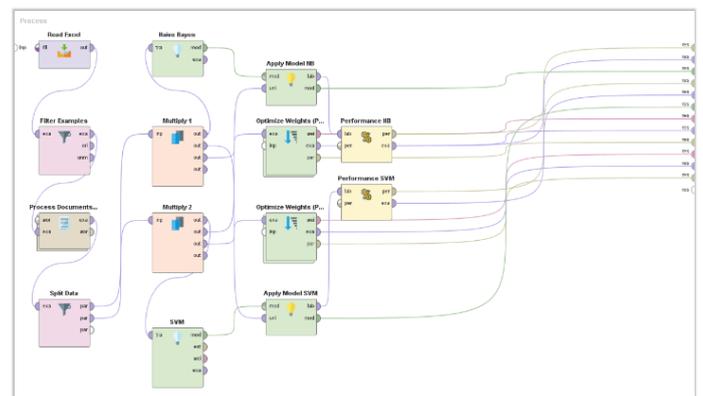
3.1 *Pembuatan Model Klasifikasi*

Pada tahapan pertama ini dibuat model menggunakan algoritma klasifikasi berdasarkan dataset yang telah melalui tahap *data preparation* dan *pre-processing* sebelumnya. Pembuatan model ini menggunakan algoritma klasifikasi NB dan SVM untuk mendapatkan hasil awal seperti yang bisa

dilihat pada gambar 3. Setelah itu dilakukan perbandingan kedua algoritma tersebut dengan menggunakan PSO. Tujuan memanfaatkan PSO dalam proses modelling adalah untuk meningkatkan nilai akurasi pada dua klasifikasi tersebut seperti yang bisa dilihat pada gambar 4.



Gambar 3. *Design Model Menggunakan Algoritma NB dan SVM*

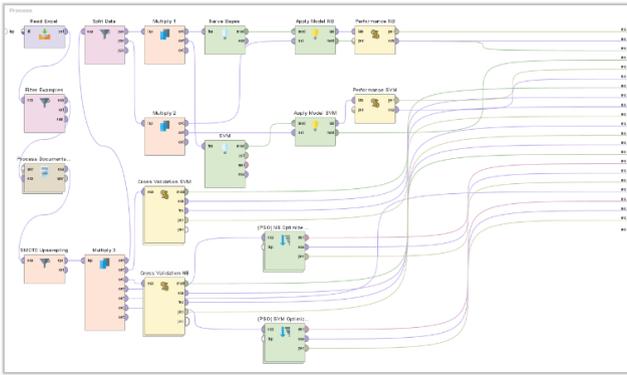


Gambar 4. *Design Model Menggunakan Algoritma NB dan SVM dengan PSO*

3.2 *Evaluasi*

Tujuan dari tahapan evaluasi adalah untuk menentukan nilai model yang sudah dibuat sebelumnya, tahapan ini menggunakan *10 fold-cross validation*. Data yang digunakan adalah data yang telah melalui tahapan *pre-processing* yang berasal dari operator *Read Excel*, untuk mengkonversi file menjadi dokumen maka perlu di tambahkan *Process Documents From Files*. Pada proses validasi, penelitian ini menggunakan *Set Role* yang memiliki fungsi menentukan field pada kelas, kemudian memanfaatkan PSO untuk mendapatkan nilai akurasi yang lebih tinggi dari sebelumnya. Desain model

klasifikasi yang telah menerapkan PSO dapat dilihat pada gambar 5 berikut ini.



Gambar 5. Design Evaluasi 10-Fold Cross-Validation untuk Algoritma NB dan SVM dengan PSO

Pada tahap ini juga menggunakan *Role Set* pada PSO untuk mendapatkan hasil akurasi yang lebih tinggi. Kategori klasifikasi memiliki nilai :

- 1) 0.90-1.00 = sangat baik.
- 2) 0.80-0.90 = baik.
- 3) 0.70-0.80 = adil.
- 4) 0.60-0.70 = buruk.
- 5) 0.50-0.40 = gagal.

Perbandingan *Accuracy* dan *Area Under Curve* (AUC) dari hasil kedua algoritma tersebut dapat dilihat pada tabel 1.

Tabel 1. Perbandingan Nilai Akurasi dan AUC

Algoritma	Accuracy	AUC
SVM	70.00%	0.735
SVM + PSO	79.90%	0.901
NB	67.50%	0.684
NB + PSO	75.38%	0.768

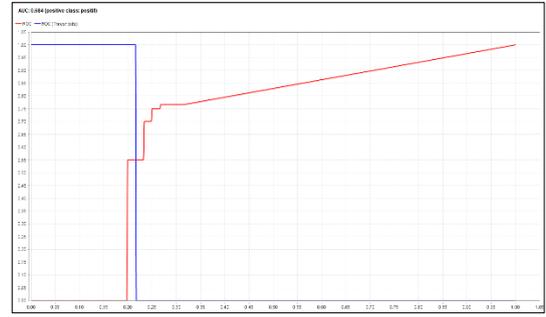
*Hasil Klasifikasi Sentimen Masyarakat*

Berikut ini adalah hasil klasifikasi sentiment masyarakat terkait topik Invansi Rusia ke Ukraina:

- 1) Hasil akurasi dengan menggunakan algoritma NB menghasilkan presentase sebesar 67,50% dan AUC 0.684 seperti pada gambar berikut ini:

accuracy: 67.50%			
	true negatif	true positif	class precision
pred negatif	48	27	64.90%
pred positif	12	33	73.33%
class recall	60.00%	55.00%	

Gambar 6. Akurasi Algoritma NB

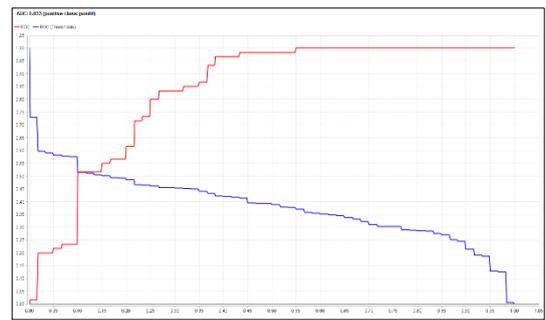


Gambar 7. Hasil AUC Algoritma NB

- 2) Hasil akurasi dengan menggunakan algoritma SVM menghasilkan presentase sebesar 70.00% dan AUC 0.735 seperti pada gambar berikut ini:

accuracy: 70.00%			
	true negatif	true positif	class precision
pred negatif	51	27	65.20%
pred positif	9	33	78.57%
class recall	85.00%	55.00%	

Gambar 8. Hasil akurasi Algoritma SVM

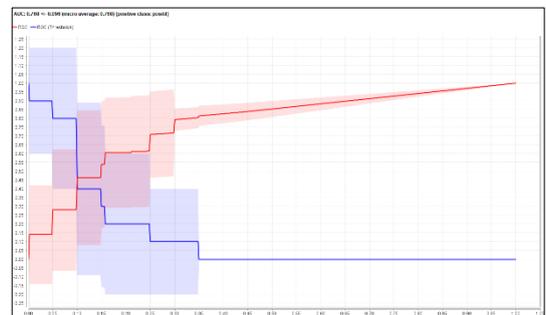


Gambar 9. AUC Algoritma SVM

- 3) Hasil akurasi dengan menggunakan algoritma NB yang telah dioptimalisasikan menggunakan PSO menghasilkan presentasi sebesar 75.38% dan AUC 0.768 seperti pada gambar berikut ini:

accuracy: 75.38% +/- 0.84% (micro average: 75.38%)			
	true negatif	true positif	class precision
pred negatif	176	75	70.12%
pred positif	23	124	84.25%
class recall	68.44%	62.31%	

Gambar 10. Akurasi Algoritma NB + PSO

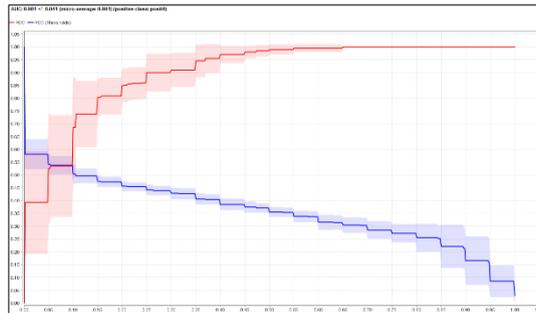


Gambar 11. AUC Algoritma NB + PSO

4) Hasil akurasi dengan menggunakan algoritma SVM yang telah dioptimalisasikan menggunakan PSO menghasilkan presentasi sebesar 79.98% dan AUC 0.901 seperti pada gambar berikut ini:

accuracy: 79.98% +/- 0.03% (micro average: 79.98%)			
	True negatif	True positif	Overall precision
pred negatif	100	01	74.09%
pred positif	19	130	87.00%
class recall	93.45%	69.35%	

Gambar 12. akurasi Algoritma SVM + PSO



Gambar 13. AUC Algoritma SVM + PSO.

#### 4. Kesimpulan

Penelitian ini menghasilkan klasifikasi sentimen analisis yang positif untuk Vladimir Putih pada komentar media berita online, penggunaan Algoritma SVM yang telah dioptimalisasikan menggunakan PSO memiliki hasil akurasi yang lebih baik dibanding dengan algoritma NB yang juga telah dioptimalisasikan dengan PSO. PSO sendiri memiliki peran penting dalam meningkatkan algoritma SVM sehingga algoritma SVM + PSO dapat menjadi solusi pengklasifikasian dalam sentimen analisis Vladimir Putin. Dengan dilakukannya penelitian ini, peneliti menyarankan untuk dikembangkan kembali dengan menggunakan algoritma yang berbeda seperti Decision Tree, K-Means, Neural Network dan algoritma lainnya serta bisa dengan topik yang berbeda.

#### 5. Daftar Pustaka

[1] Eko, S. 2022. *Invasi Rusia dan Dampaknya Terhadap Geopolitik Global*, CNBC Indonesia. Available at: <https://www.cnbcindonesia.com/opini/20220307124740-14-320589/invasi-rusia-dan-dampaknya-terhadap-geopolitik-global> (Accessed: 23 June 2022).

[2] CNN Indonesia. 2022. *Mengapa Banyak Warga Indonesia Dukung Putin Invasi Ukraina?* Available at: <https://www.cnnindonesia.com/internasional/20220311071348-106-769708/mengapa-banyak-warga-indonesia-dukung-putin-invasi-ukraina> (Accessed: 23 June 2022).

[3] Jannati, R., Mahendra, R., Wardhana, C.W. and Adriani, M., 2018, November. Stance classification towards political figures on blog writing. In *2018 International Conference on Asian Language Processing (IALP)* (pp. 96-101). IEEE. DOI: 10.1109/IALP.2018.8629144.

[4] Rofiqi, M.A., Fauzan, A.C., Agustin, A.P. and Saputra, A.A., 2019. Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 1(2), pp.58-64. DOI: 10.28926/ilkomnika.v1i2.18.

[5] Baskoro, B.B., Susanto, I. and Khomsah, S., 2021. Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF (Studi Kasus: Ulasan Pelanggan Pada Situs TRIPADVISOR). *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, 3(2), pp.21-29. DOI: 10.20895/INISTA.V3I2.

[6] Firdaus, A.F. and Firdaus, W.I., 2021. Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi:(Sebuah Ulasan). *JUPITER (Jurnal Penelitian Ilmu dan Teknik Komputer)*, 13(1), pp.66-78.

[7] Munthe, C.J.E., Hasibuan, N.A. and Hutabarat, H., 2022. Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace. *Resolusi: Rekayasa Teknik Informatika dan Informasi*, 2(3), pp.110-115. DOI: <https://doi.org/10.30865/resolusi.v2i3.309>.

- [8] Faisal, A., Alkhalifi, Y., Rifai, A. and Gata, W., 2020. Analisis Sentimen Dewan Perwakilan Rakyat Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization. *JOINTECS (Journal of Information Technology and Computer Science)*, 5(2), pp.61-70. DOI: 10.31328/jointecs.v5i2.1362.
- [9] Ulfah, A.N. and Anam, M.K., 2020. Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM). *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 7(1), pp.1-10. DOI: <https://doi.org/10.35957/jatisi.v7i1.196>.
- [10] Sinaga, S., Sembiring, R.W. and Sumarno, S., 2022. Penerapan Algoritma Naive Bayes untuk Klasifikasi Prediksi Penerimaan Siswa Baru. *Journal of Machine Learning and Data Analytics*, 1(1), pp.55-64.
- [11] Damayanti, S.E. and Kuswayati, S., 2006. Analisis Dan Implementasi Framework Crisp-Dm (Cross Industry Standard Process for Data Mining) Untuk Clustering Perguruan Tinggi Swasta. *Jurnal STT Bandung*.
- [12] Kurniawan, S., Gata, W., Puspitawati, D.A., Tabrani, M. and Novel, K., 2019. Perbandingan Metode Klasifikasi Analisis Sentimen Tokoh Politik Pada Komentar Media Berita Online. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(2), pp.176-183. DOI: 10.29207/resti.v3i2.935.
- [13] Pratama, Y. T., Bachtiar, F. A. and Setiawan, N. Y. 2018. Analisis Sentimen Opini Pelanggan Terhadap Aspek Pariwisata Pantai Malang Selatan Menggunakan TF-IDF dan Support Vector Machine, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(12), pp. 6244–6252.
- [14] Satrio, R.H. and Fauzi, M.A., 2019. Klasifikasi Tweets Pada Twitter Menggunakan Metode K-Nearest Neighbour (K-NN) Dengan Pembobotan TF-IDF. *vol, 3*, pp.8293-8300.
- [15] Prasetyo, V.R., Lazuardi, H., Mulyono, A.A. and Lauw, C., 2021. Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier. *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, 7(1), pp.8-17. DOI: 10.25077/teknosi.v7i1.2021.8-17.
- [16] Fadilah, E., 2018. Implementasi Metode Profile Matching Terhadap Sistem Pendukung Keputusan Penerimaan Dana Zakat pada Badan Amil Zakat Pertamina (BAZMA). *Jurnal MATICS Vol, 10(2)*. DOI: 10.18860/mat.v10i2.5745.