

# Analisis Kinerja Algoritma Machine Learning Untuk Klasifikasi Emosi

Sudianto

Fakultas Informatika, Program Studi Teknik Informatika, Institut Teknologi Telkom, Purwokerto, Indonesia

Email: sudianto@ittelkom-pwt.ac.id

Email Penulis Korespondensi: sudianto@ittelkom-pwt.ac.id

Submitted:11/09/2022; Accepted:27/09/2022; Published: 30/09/2022

**Abstrak**—Media sosial merupakan salah satu wadah untuk berkespresi atau membagikan aktivitas keseharian. Berbagai kejadian baru kerap diperbincangkan di media sosial, seperti halnya pada media sosial Twitter. Acapkali, perbincangan yang dilakukan oleh pengguna Twitter saat memberikan sebuah ulasan maupun opini memiliki berbagai emosi; seperti marah, sedih, takut ataupun gembira. Tantangan yang terjadi, emosi sukar untuk dideskripsikan, terkadang mengakibatkan multitafsir dan salah sangka yang mengakibatkan perdebatan hingga pelaporan pada pihak berwajib. Hal ini menunjukkan, emosi pada ulasan maupun opini penting untuk dilakukan klasifikasi, karena emosi yang bersumber dari teks sukar untuk dimengerti. Selain itu, klasifikasi emosi perlu dilakukan untuk mempercepat identifikasi emosi. Tujuan penelitian ini yaitu mengetahui algoritma yang memiliki kinerja yang optimal dalam klasifikasi emosi. Metode Machine Learning yang digunakan yaitu algoritma Naïve Bayes, Random Forest, dan Support Vector Machines. Hal ini dilakukan untuk mengetahui algoritma yang dominan dalam melakukan klasifikasi emosi. Hasil pemodelan, klasifikasi menggunakan algoritma Random Forest memperoleh akurasi yang dominan dengan nilai akurasi 81,3%, dan diikuti dengan algoritma SVM dengan nilai akurasi 76,6% dan nilai akurasi algoritma Naïve Bayes 79,1%. Selain itu, dari kecepatan waktu dalam menyelesaikan klasifikasi, algoritma Random Forest memiliki waktu tercepat yaitu 1,27 detik.

**Kata Kunci:** Emosi; Kecerdasan Buatan; Klasifikasi; KNN; RF; SVM

**Abstract**—Social media is a place to express or share daily activities. Various new events are often discussed on social media, such as on Twitter. Frequently, the conversations conducted by Twitter users when giving a review or opinion have various emotions, such as anger, sadness, fear, or joy. Emotions are difficult to describe the challenges that occur, sometimes leading to multiple interpretations and misunderstandings leading to debates and reporting to the authorities. So this shows that emotions in reviews and opinions are essential for classification because emotions that come from texts are difficult to understand. In addition, the classification of emotions needs to be done to speed up the identification of emotions. The purpose of this study is to find out which algorithm has optimal performance in the classification of emotions. Machine Learning methods are the Naïve Bayes algorithm, Random Forest, and Support Vector Machines; this is done to determine the dominant algorithm in classifying emotions. The results of the modeling and classification using the Random Forest algorithm obtained a dominant accuracy with an accuracy value of 81.3%, followed by the SVM algorithm with an accuracy value of 76.6% and an accuracy value of 79.1% Naïve Bayes algorithm. In addition, from the speed of time in completing the classification, the Random Forest algorithm has the fastest time of 1.27 seconds.

**Keywords:** Artificial Intelligence; Classification; Emotion; KNN; RF; SVM

## 1. PENDAHULUAN

Pengguna media sosial kian hari semakin meningkat. Media sosial kini tidak hanya sebatas sarana untuk berinteraksi, melainkan media sosial juga kini digunakan untuk berbagai hal, seperti berbagi dan bertukar informasi pada komunitas virtual. Fenomena ini tidak lepas dengan meningkatnya jumlah penggunaan internet yang terus bertambah, serta didukung adanya infrastruktur yang semakin memadai. Salah satu media sosial yang banyak digunakan yaitu Twitter. Twitter merupakan platform media sosial yang bisa digunakan untuk membagikan opini maupun ulasan. Twitter dicirikan dengan adanya #hashtag dan *trending* topik. Selain itu, opini di Twitter cepat menarik perhatian masyarakat. Berdasarkan opini pengguna, acap kali opini di Twitter menjadi sangat bervariasi dan memiliki umpan balik yang beragam sehingga menjadi viral. Sementara, pada media sosial Twitter tidak hanya menyampaikan opini maupun informasi saja, pengguna Twitter juga dapat mengekspresikan emosinya.

Emosi merupakan perasaan yang saat itu dirasakan. Berbagai emosi muncul dalam diri dengan kondisi yang berbeda; seperti cinta, gembira, takut, marah hingga sedih. Kondisi yang dialami tentunya memengaruhi cara berpikir dan bertindak sehingga menghasilkan emosi [1]. Sejauh ini, deteksi emosi dapat dilakukan melalui suara dan ekspresi wajah. Namun emosi dengan ekspresi tertulis, justru sering membuat multitafsir serta mengakibatkan perdebatan. Selain itu, emosi dalam tulisan sukar untuk dijabarkan. Meskipun struktur kosa kata emosi dapat didefinisikan pada lima kelas yaitu marah (*anger*), takut (*fear*), senang (*happiness*), cinta (*love*), dan sedih (*sadness*) [2]. Akan tetapi, identifikasi emosi di media sosial Twitter tidak hanya menggunakan teknik *text processing* saja, karena *tweet* emosi pada media Twitter yang singkat dan tata bahasa yang tidak teratur, menambah tingkat kesulitan dalam memahami emosi yang disampaikan. Sehingga dibutuhkan teknik *text processing* lebih lanjut, yaitu dengan teknik *text mining* untuk mengekstrak kata-kata yang tidak terstruktur untuk membuat model Machine Learning yang dapat melakukan klasifikasi maupun pengelompokan emosi lebih cepat.

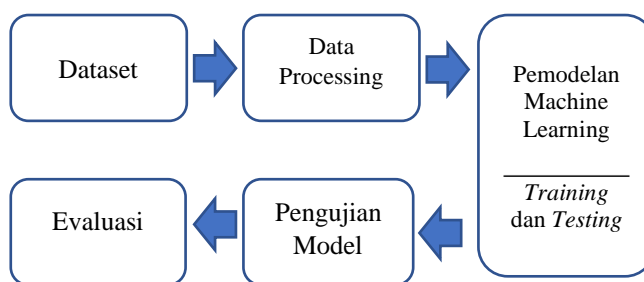
Penelitian sebelumnya, klasifikasi emosi telah banyak dilakukan diantaranya; seperti penelitian perbandingan algoritma pada klasifikasi emosi *tweet* berbahasa Indonesia menggunakan K-Nearest Neighbor (KNN) dan Logistic Regression [3], deteksi emosi pada *tweet* berbahasa Indonesia menggunakan Latent Dirichlet Allocation (LDA) dan konversi ekspresi simbol [4], klasifikasi *text* berbasis ekstraksi fitur dengan algoritma K-Nearest Neighbor (KNN) [5], deteksi emosi dengan multi kelas untuk data *tweet* Bahasa Indonesia yang berimbang [6], implementasi *text mining*

untuk deteksi emosi menggunakan metode Leksikon (studi kasus: *tweet* tentang Covid-19 [7]). Namun berbeda dengan penelitian sebelumnya, penelitian ini melakukan klasifikasi dengan beberapa algoritma Machine Learning dengan skenario data berimbang. Hal ini dilakukan untuk memperoleh model yang optimal dalam melakukan klasifikasi emosi.

Oleh karena itu, tujuan penelitian ini yaitu mengetahui algoritma yang memiliki akurasi yang optimal dalam klasifikasi emosi. Metode Machine Learning yang digunakan seperti algoritma Naïve Bayes, Random Forest, Support Vector Machines. Pertimbangan menggunakan ketiga algoritma karena memiliki akurasi yang dominan pada klasifikasi berbasis teks.

## 2. METODOLOGI PENELITIAN

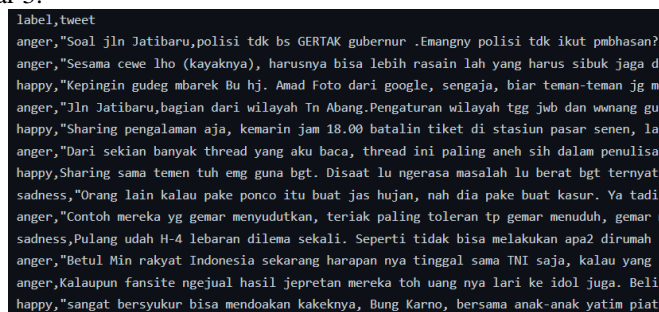
Pembuatan model klasifikasi pada penelitian ini, dilakukan melalui beberapa proses: (1) *Data Collection*; (2) *preprocessing* data; (3) pembuatan model Machine Learning; (4) pengujian model; (5) evaluasi. Pada Gambar 1 alir penelitian yang dilakukan yaitu pertama pengumpulan *dataset* yang diperoleh dari *tweet* pada media sosial Twitter. Kemudian dilakukan *preprocessing*; seperti *cleaning*, *case folding*, normalisasi dan ekstraksi fitur. Setelah itu, data dibagi menjadi data *training* dan *testing* untuk pembuatan model. Terakhir, model dievaluasi akurasi yang optimal menggunakan *confusion matrix*.



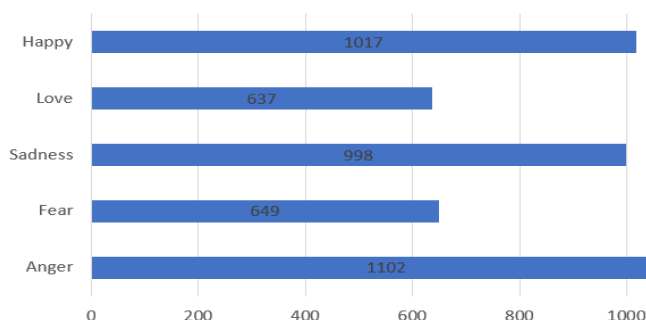
Gambar 1. Alir penelitian

### 2.1 Dataset

*Dataset* pada penelitian ini berupa *dataset* teks. *Dataset* diperoleh dari *tweet* media sosial Twitter. *Dataset* telah dikumpulkan melalui penelitian sebelumnya yaitu di Indonesian-Twitter-Emotion-Dataset [8]. *Dataset* diperoleh melalui *crawling data* komen dengan *streaming API* dari 1 Juni 2018 hingga 14 Juni 2018. Jumlah *tweet* pada *dataset* terdiri dari 4403 *tweets* berbahasa Indonesia. Selain itu, *dataset* juga telah dilengkapi dengan label yang terdiri dari lima kelas emosi yaitu: “*anger*”, “*fear*”, “*happy*”, “*love*”, dan “*sadness*”. Kumpulan dan jumlah *dataset* dapat dilihat pada Gambar 2 dan Gambar 3.



Gambar 2. Dataset Emosi Berbahasa Indonesia



Gambar 3. Jumlah Masing-Masing Kategori Emosi Dari Dataset

Pada Gambar 2 dan Gambar 3, *dataset* telah memiliki label. Hal ini sesuai dengan hasil penelitian sebelumnya. Jumlah masing-masing kategori terdiri dari; *happy* 1017 label, *love* 637 label, *sadness* 998 label, *fear* 649 label, *anger* 1102 label. Namun untuk kebutuhan *training* dan *testing* pembuatan model, penelitian ini membagi jumlah label masing-masing kategori 630 label. Hal ini bertujuan agar data disetiap kategori berimbang.

### 2.2 Preprocessing

*Preprocessing* dilakukan untuk memastikan *dataset* yang tersedia bisa digunakan untuk membangun model Machine Learning. Beberapa proses yang dilakukan yaitu melalui tujuh tahap yaitu:

- Data Cleansing*; dilakukan untuk menghilangkan data yang tidak penting atau *noise* pada kalimat, serta simbol maupun karakter spesial. Misalnya *emoji*, hastag dan lainnya. Selain itu, proses ini juga dilakukan untuk mengganti semua kata menjadi huruf kecil semua, atau proses ini disebut *case folding*.
- Tokenizing*; dilakukan untuk membagi atau memotong kalimat menjadi beberapa bagian kata, proses ini menghasilkan kata yang disebut sebagai *token*. Sehingga hasil dari setiap *token* dapat mengetahui frase maupun entitasnya.
- Remove Stopword*; dilakukan untuk menghilangkan kata tidak unik pada kalimat. Kata yang cenderung sebagai kata penghubung akan dihilangkan seperti “yang”, “di” dan “pada”. Proses ini bertujuan untuk meningkatkan hasil analisis saat pembuatan model klasifikasi.
- Stemming*; dilakukan untuk menguraikan kata menjadi bentuk dasarnya. Contohnya “berjalan” menjadi “jalan”.
- Normalisasi; digunakan untuk mengubah kata yang tidak baku menjadi baku.
- Labeling*; digunakan untuk menandai dan membagi kategori *tweet* terhadap beberapa *class*.
- Fitur ekstrasi; ekstrasi dilakukan dengan TF-IDF, fitur diperoleh berdasarkan pembobotan yang terintegrasi antara *term frequency* dengan *inverse dokumen frequency*, sebagai hasil ringkasan [9].

### 2.4 Machine Learning

Penyelesaian dalam klasifikasi teks dapat dilakukan dengan berbagai pendekatan, seperti halnya klasifikasi menggunakan Machine Learning [10], maupun klasifikasi dengan Deep Learning [11]. Pada penelitian ini, klasifikasi dibangun dengan pendekatan Machine Learning. Pada tahap klasifikasi, pembuatan model menggunakan tiga algoritma Machine Learning yaitu Naïve Bayes (NB), Random Forest (RF) dan Support Vector Machines (SVM). Ketiga algoritma ini dipilih untuk perbandingan model yang optimal. Algoritma Naïve Bayes merupakan klasifikasi biner dan *multiclass*. Naïve Bayes menerapkan teknik *supervised learning* dengan menetapkan label menggunakan probabilitas bersyarat. Probabilitas ini diperoleh dari peluang peristiwa berdasarkan peristiwa lain yang telah terjadi [12] [13]. Menghitung kemungkinan yang terjadi pada algoritma Naïve Bayes dapat menggunakan Persamaan 1.

$$P(A|B) = P(B|A)P(A)P(B) \tag{1}$$

Keterangan:

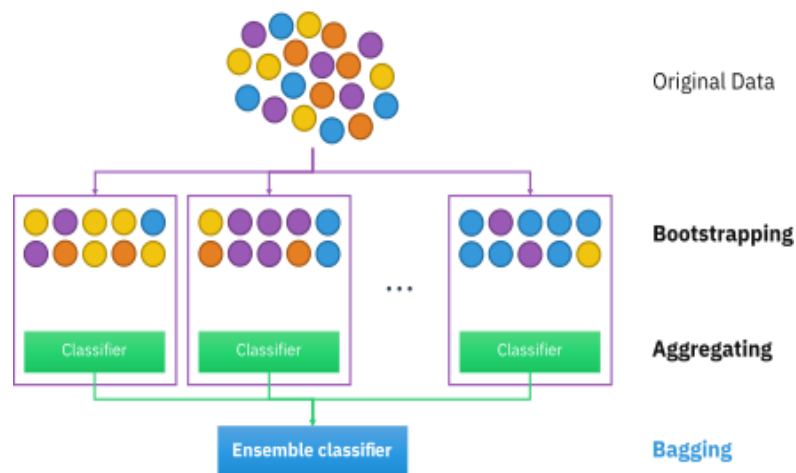
$P(A|B)$  = Kemungkinan A terjadi dengan bukti bahwa B telah terjadi (probabilitas superior)

$P(B|A)$  = Kemungkinan B terjadi dengan bukti bahwa A telah terjadi

$P(A)$  = Peluang terjadinya A

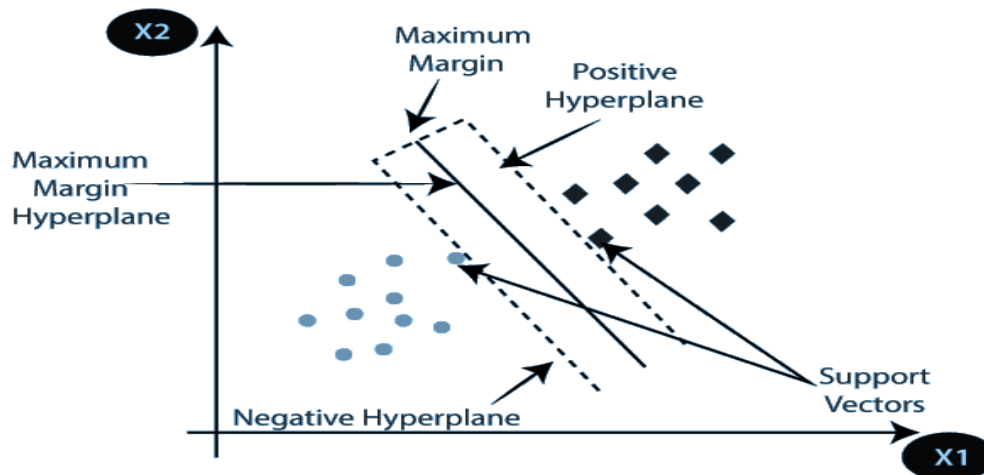
$P(B)$  = Peluang terjadinya B

Selanjutnya, algoritma Machine Learning yang digunakan yaitu algoritma Random Forest. Random Forest merupakan algoritma klasifikasi yang memiliki proses agregasi *decision tree*. Prinsip kerja Random Forest menghasilkan luaran yang dapat berbeda-beda yang disebabkan *decision tree*. Selain itu, Random Forest bekerja dengan *Bagging*, kemudian *voting* dengan tujuan menentukan hasil yang dominan atau mayoritas dari semua proses *decision tree* yang telah dilakukan [14] [15]. Ilustrasi kerja Random Forest bisa dilihat seperti pada Gambar 4.



**Gambar 4.** Ilustrasi kerja Random Forest

Algoritma terakhir, yaitu algoritma Support Vector Machine. SVM merupakan algoritma *supervised learning*. SVM populer digunakan untuk permasalahan klasifikasi, terutama dalam klasifikasi Machine Learning. Cara kerja algoritma SVM, yaitu SVM yaitu membuat garis atau batas keputusan terbaik yang dapat memisahkan ruang dimensi  $n$  ke dalam kelas-kelas sehingga dapat dengan mudah menempatkan titik data baru dalam kategori yang benar. Batas keputusan terbaik ini disebut *hyperplane* [16][17]. Pada algoritma SVM, terdapat titik atau vektor ekstrim yang dapat membantu dalam menciptakan *hyperplane*. Kasus ekstrim ini disebut *support vector*. Ilustrasi kerja Support Vector Machine seperti pada Gambar 5.



Gambar 5. Ilustrasi kerja Support Vector Machine

## 2.4 Evaluasi

Evaluasi pada pengujian model dilakukan dengan *confusion matrix*. Evaluasi ini bertujuan untuk mengetahui kinerja algoritma yang telah diperoleh dari ketiga algoritma Machine Learning. Evaluasi menggunakan *confusion matrix* dilakukan pada setiap kolom matriks yang mewakili setiap kelas yang diprediksi, dan matriks baris mewakili setiap kelas yang sebenarnya. Evaluasi *confusion matrix* memerlukan perhitungan empat ukuran yaitu kondisi *True Positive*, kondisi *True Negative*, kondisi *False Positive*, dan *False Negative*, untuk mewakili setiap hasil klasifikasi yang diperoleh. Selain itu, perhitungan pada proses evaluasi akan menghasilkan nilai akurasi, *recall*, maupun *precision* [16], ilustrasi penggunaan *confusion matrix* seperti pada Gambar 6.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 6. Confusion Matrix

## 3. HASIL DAN PEMBAHASAN

### 3.1 Hasil Preprocessing

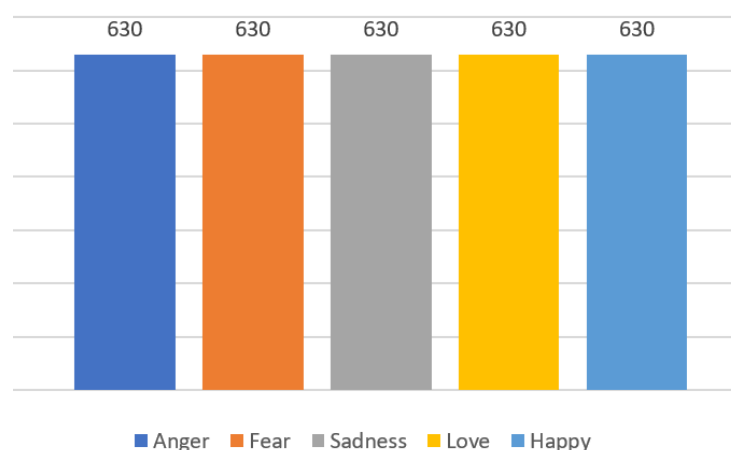
*Preprocessing* dilakukan dengan tujuan agar data yang digunakan sebagai *training* dan *testing* sesuai dengan kebutuhan pembuatan model. Penelitian ini, proses *text processing* dilakukan dengan menghapus simbol maupun karakter yang tidak diperlukan menggunakan *library* Sastrawi. Hasil dari *text processing* dapat dilihat pada Gambar 7.

Tanpa Preprocessing	Dari sekian banyak thread yang aku baca, thread ini paling aneh sih dalam penulisan. Sumpah aneh bgt, mau ngatain ""lebay"" aja segala bikin thread hadeh. Aku juga ga jago nulis, tapi tulisan aku ga seberantakan thread mbaknya."
Cleaning	Dari sekian banyak thread yang aku baca thread ini paling aneh sih dalam penulisan Sumpah aneh bgt mau ngatain lebay aja segala bikin thread hadeh Aku juga ga jago nulis tapi tulisan aku ga seberantakan thread mbaknya
Case Folding	dari sekian banyak thread yang aku baca thread ini paling aneh sih dalam penulisan sumpah aneh bgt mau ngatain lebay aja segala bikin thread hadeh aku juga ga jago nulis tapi tulisan aku ga seberantakan thread mbaknya
Normalisasi	dari sekian banyak thread yang aku baca thread ini paling aneh sih dalam penulisannya sumpah aneh banget mau ngatain lebay saja bikin thread hadeh aku juga tidak jago nulis tapi tulisan aku tidak seberantakan thread mbaknya

**Gambar 7.** Proses *Text Processing*

Pada Gambar 7 merupakan salah satu *tweet* yang dilakukan *text processing*. Data *tweet* yang masih tidak terstruktur diubah menjadi baku dan terstruktur. Proses ini dilakukan dengan *cleaning* untuk menghilangkan karakter spesial, mengganti semua kata menjadi huruf kecil dengan *case folding*, melakukan normalisasi agar data *tweet* menjadi lebih terstruktur. Kemudian, proses *labeling* menggunakan pelabelan yang tersedia pada *dataset*.

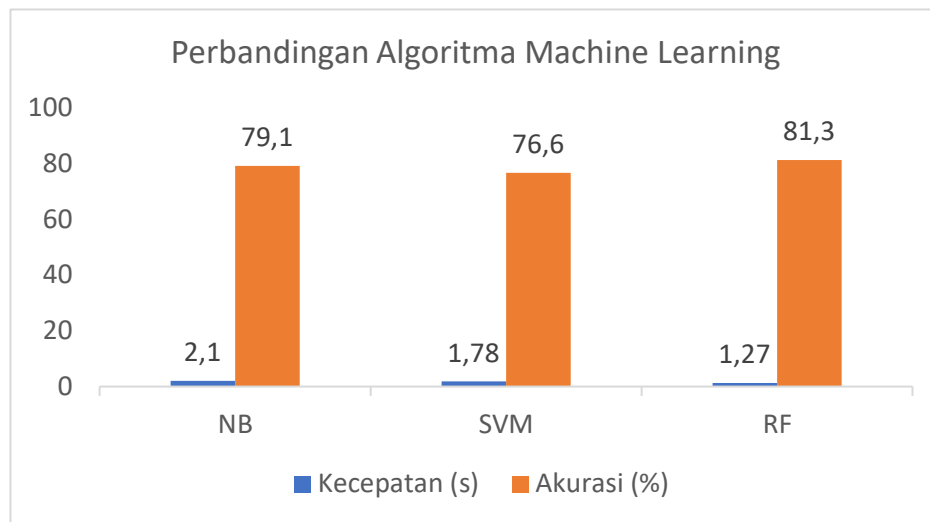
Selanjutnya, kategori *class* atau label terdapat lima *class* yaitu: *anger*, *fear*, *sadness*, *love* dan *happy*. Selain itu, penelitian ini membagi distribusi *dataset* secara berimbang dengan masing-masing *class* memiliki 630 *tweets*, seperti pada Gambar 8. Hal ini bertujuan agar proses *supervised learning* tidak ada *class* yang dominan saat dilakukan *training*.



**Gambar 8.** Distribusi *Dataset* Yang Digunakan Untuk Pembuatan Model

Pengelompokan *class* juga dicirikan dengan kata yang dominan, kata-kata yang dominan ini sesuai dengan label atau *class* yang telah ditentukan. Penelitian ini menggunakan WordCloud untuk mengetahui representatif kata yang dominan pada setiap *class*. WordCloud merupakan visualisasi daftar kata-kata yang ada pada sebuah teks, semakin banyak kata yang ada maka semakin besar kemungkinan munculnya. Detail kata-kata yang muncul pada setiap *class* yang dapat dilihat pada Gambar 9 dan Gambar 10.





**Gambar 11.** Hasil Grafik Perbandingan Akurasi Dan Kecepatan

Pada Gambar 11 dan Tabel 2 menunjukkan bahwa masing-masing algoritma Machine Learning yang telah diuji (NB, SVM dan RF) memiliki rerata akurasi sebesar 79%. Selain itu, rerata waktu komputasi dari ketiga algoritma yaitu 1,71 detik. Sedangkan untuk algoritma yang memiliki nilai paling dominan dalam akurasi maupun waktu komputasi yaitu algoritma RF dengan nilai akurasi 81,3% dan waktu komputasi 1,27 detik.

**Tabel 2.** Hasil Perbandingan algoritma Machine Learning

Algoritma	Akurasi (%)	Waktu Komputasi (s)
NB	79.1	02.10
SVM	76.6	01.78
RF	81.3	01.27

## 4. KESIMPULAN

Klasifikasi emosi menggunakan pendekatan Machine Learning sangat membantu untuk mengidentifikasi pengelompokan emosi pada kalimat. Pada penelitian ini, klasifikasi emosi berhasil mengidentifikasi emosi pada kalimat. Sementara itu, dari ketiga algoritma Machine Learning yang dicoba. Algoritma Random Forest memperoleh akurasi yang dominan dibandingkan kedua algoritma lainnya. Algoritma Random Forest memperoleh nilai akurasi sebesar 81,3%. Sedangkan algoritma Naïve Bayes dan Support Vector memperoleh nilai akurasi sebesar 79,1% dan 76,6%. Selanjutnya, dari kecepatan waktu dalam melakukan klasifikasi dari ketiga algoritma. Random Forest tercatat waktu tercepat dengan lama komputasi 01,27 detik, Support Vector Machine tercatat waktu komputasi 01,78 detik, dan algoritma Naïve Bayes tercatat waktu komputasi 02,10 detik.

## REFERENCES

- [1] A. Chaer, *Pengantar Semantik Bahasa Indonesia*, Edisi Revi. Jakarta: Rineka Cipta, 2019.
- [2] A. Chaer, *Morfologi bahasa Indonesia*. Jakarta: Rineka Cipta, 2008.
- [3] A. B. P. Negara, H. Muhandi, and F. Sajid, "Perbandingan Algoritma Klasifikasi terhadap Emosi Tweet Berbahasa Indonesia," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 2, p. 242, 2021, doi: 10.26418/jp.v7i2.48198.
- [4] R. M. Cahyaningtyas, R. Kusumaningrum, Sutikno, Suhartono, and D. E. Riyanto, "Emotion detection of tweets in Indonesian language using LDA and expression symbol conversion," *Proc. - 2017 1st Int. Conf. Informatics Comput. Sci. ICICoS 2017*, vol. 2018-Janua, pp. 253–257, 2017, doi: 10.1109/ICICOS.2017.8276371.
- [5] M. Azam, T. Ahmed, F. Sabah, and M. I. Hussain, "Feature Extraction based Text Classification using K-Nearest Neighbor Algorithm," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 18, no. 12, pp. 95–101, 2018, [Online]. Available: [http://paper.ijcsns.org/07\\_book/201812/20181213.pdf](http://paper.ijcsns.org/07_book/201812/20181213.pdf)
- [6] L. Farsiah, Y. S. Chen, and A. Misbullah, "Multi-Classes Emotion Detection for Unbalanced Indonesian Tweets," *Proc. Int. Conf. Electr. Eng. Informatics*, vol. 2020-Octob, 2020, doi: 10.1109/ICELTICs50595.2020.9315403.
- [7] A. S. Aribowo and S. Khomsah, "Implementation Of Text Mining For Emotion Detection Using The Lexicon Method (Case Study: Tweets About Covid-19)," *Telematika*, vol. 18, no. 1, p. 49, 2021, doi: 10.31315/telematika.v18i1.4341.
- [8] M. S. Saputri, R. Mahendra, and M. Adriani, "Emotion Classification on Indonesian Twitter Dataset," *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, no. November, pp. 90–95, 2019, doi: 10.1109/IALP.2018.8629262.
- [9] F. S. Yerzi and Y. Sibaroni, "Analisis Sentimen Terhadap Kebijakan Pemerintah Dalam Menangani Covid-19 Dengan Pendekatan Lexicon Based," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 11354–11366, 2021.
- [10] S. Sudianto, A. D. Sripamuji, I. Ramadhanti, R. R. Amalia, J. Saputra, and B. Prihatnowo, "Penerapan Algoritma Support



- Vector Machine dan Multi-Layer Perceptron pada Klasifikasi Topik Berita,” vol. 11, no. 2, pp. 84–91, 2022.
- [11] W. Afandi, S. N. Saputro, A. M. Kusumaningrum, H. Ardiansyah, M. H. Kafabi, and S. Sudioanto, “Klasifikasi Judul Berita Clickbait menggunakan RNN-,” *J. Pengemb. IT*, vol. 7, no. 2, pp. 85–89, 2022.
- [12] H. A. Santoso, E. H. Rachmawanto, and U. Hidayati, “Fake Twitter Account Classification of Fake News Spreading Using Naïve Bayes,” *Sci. J. Informatics*, vol. 7, no. 2, pp. 2407–7658, 2020, [Online]. Available: <http://journal.unnes.ac.id/nju/index.php/sji>
- [13] E. Fitri, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” *J. Transform.*, vol. 18, no. 1, p. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [14] L. Breiman, “Random forests,” *Random For.*, pp. 1–122, 2001, doi: 10.1201/9780429469275-8.
- [15] M. A. Fauzi, “Random forest approach fo sentiment analysis in Indonesian language,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 1, pp. 46–50, 2018, doi: 10.11591/ijeecs.v12.i1.pp46-50.
- [16] C. Science, “Efficient Classification Algorithms using SVMs for Large Datasets Master of Technology,” no. June, 2007.
- [17] I. Muhammad and Z. Yan, “Supervised Machine Learning Approaches: a Survey,” *ICTACT J. Soft Comput.*, vol. 05, no. 03, pp. 946–952, 2015, doi: 10.21917/ijsc.2015.0133.
- [18] S. Sudioanto, P. Wahyuningtias, H. W. Utami, U. A. Raihan, and H. N. Hanifah, “Comparison Of Random Forest And Support Vector Machine Methods On Twitter Sentiment Analysis ( Case Study : Internet Selebgram Rachel Vennya Escape From Quarantine ) Perbandingan Metode Random Forest Dan Support Vector Machine Pada Analisis Sentimen Twitt,” *Jutif*, vol. 3, no. 1, pp. 141–145, 2022.
- [19] S. Chandra Ayunda Apta, N. Trivetisia, N. A. Winanti, D. P. Martiyarningsih, T. W. Utami, and S. Sudioanto, “Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis,” *J. Pengemb. IT*, vol. 7, no. 2, pp. 80–84, 2022.