

# Analisis Emosi Data *Tweet* Berbahasa Indonesia

Annisa Fitria Anwar Damanik  
Departemen Matematika  
Universitas Indonesia  
Depok, Indonesia  
nichadamanik@gmail.com

Nasywa Safira Ardanty  
Departemen Matematika  
Universitas Indonesia  
Depok, Indonesia  
nsysafira17@gmail.com

Kamal Muftie Yafi  
Departemen Matematika  
Universitas Indonesia  
Depok, Indonesia  
kamalmuya@gmail.com

Rifa Nayaka Utami  
Departemen Matematika  
Universitas Indonesia  
Depok, Indonesia  
rifanayaka12@gmail.com

**Abstrak** — Pada era media sosial yang terus berkembang, data *tweet* merupakan sumber informasi yang dapat digunakan untuk memahami sentimen dan emosi masyarakat. Analisis emosi pada data *tweet* menjadi penting untuk mengungkapkan perasaan dan opini masyarakat yang terkandung dalam pesan-pesan tersebut. Interaksi yang terjadi di media sosial ini bersifat secara tidak langsung, sehingga perasaan yang dirasakan oleh seseorang hanya dapat dilihat dalam bentuk teks pada *tweet* yang mereka bagikan. Maka dari itu, dibutuhkan sistem klasifikasi dalam interaksi sosial media untuk mengetahui emosi seseorang yang dikategorikan menjadi lima emosi dasar yaitu *joy*, *fear*, *love*, *sadness*, dan *anger*. Pada penelitian ini, digunakan *machine learning* berupa metode *Naïve Bayes*, metode *Support Vector Machine* (SVM), dan metode *Logistic Regression* untuk melakukan klasifikasi emosi pada *tweet*. Digunakan pula fitur *Natural Language Processing* (NLP) berupa *Bag of Words* (BoW) dan *Term Frequency Inverse Document Frequency* (TF-IDF). Model akan diuji pada data *tweet* yang belum memiliki label emosi, dengan tujuan memprediksi emosi yang mungkin terkandung dalam *tweet-tweet* tersebut. Hasil penelitian ini berpotensi untuk digunakan pada pemantauan sentimen *online*, analisis tren sosial, pengambilan keputusan berdasarkan data media sosial, dan meningkatkan pemahaman terhadap perasaan dan pandangan masyarakat di platform ini.

**Kata kunci**—analisis emosi, text mining, Bahasa Indonesia, klasifikasi

## I. PENDAHULUAN

### A. Latar Belakang

Di era digital saat ini, media sosial telah menjadi salah satu platform utama bagi individu untuk berbagi pemikiran, perasaan, dan pengalaman dengan lebih mudah dan cepat. Di Indonesia, Twitter adalah salah satu platform media sosial yang sangat populer. Bahkan, berdasarkan data dari We Are Social, jumlah pengguna Twitter di Indonesia mencapai 18,45 juta pada 2022. Jutaan *tweet* dibagikan setiap hari oleh pengguna dari Indonesia. *Tweet-tweet* ini mencerminkan berbagai macam emosi dasar yang dialami oleh pengguna, termasuk kebahagiaan (*joy*), kesedihan (*sadness*), kemarahan (*anger*), ketakutan (*fear*), dan cinta (*love*). Analisis emosi dalam *tweet* Indonesia menjadi semakin relevan dalam berbagai konteks, termasuk riset psikologi, analisis emosi, analisis sentimen, pemahaman tren sosial, dan pengambilan keputusan bisnis [1].

Salah satu alasan pentingnya analisis emosi *tweet* di Indonesia adalah untuk memahami perasaan dan sikap masyarakat atas suatu hal. Emosi-emosi dasar tersebut dapat memberikan wawasan mendalam tentang respons Masyarakat mengenai peristiwa atau isu tertentu. Selain itu, analisis emosi

*tweet* juga memiliki implikasi dalam pengambilan keputusan bisnis dan pemasaran. Perusahaan dapat menggunakan data emosi dari *tweet* untuk mengukur respons pelanggan terhadap produk atau layanan mereka.

Dalam konteks penelitian akademis, analisis emosi *tweet* juga dapat digunakan untuk mengidentifikasi tren psikologis dan sosial dalam masyarakat Indonesia. Hal ini dapat membantu peneliti memahami bagaimana faktor-faktor tertentu, seperti berita, peristiwa sosial, atau tren budaya, dapat memengaruhi emosi dan sikap masyarakat secara luas.

Dengan semakin berkembangnya teknologi pemrosesan bahasa alami dan analisis sentimen, analisis emosi *tweet* Indonesia menjadi lebih dapat diakses dan relevan. Data ini dapat memberikan wawasan yang berharga bagi berbagai pihak, termasuk pemerintah, bisnis, peneliti, dan masyarakat umum, dalam memahami dinamika emosi dan sikap di tengah perkembangan dunia digital yang terus berlanjut [5].

### B. Rumusan Masalah

1. Bagaimana cara efektif untuk mengidentifikasi dan mengklasifikasikan emosi dalam *tweet* berbahasa Indonesia?

### C. Tujuan dan Manfaat

1. Untuk mengetahui cara efektif dalam mengidentifikasi dan mengklasifikasikan emosi dalam *tweet* berbahasa Indonesia.
2. Untuk mengetahui emosi seseorang berdasarkan *tweet* yang telah dipublikasikan.
3. Untuk mengetahui model mana yang memiliki performansi dan akurasi yang baik dalam mengklasifikasikan emosi pada *tweet* di Twitter.

## II. PEMBAHASAN

### A. Landasan Teori

a) *Text Preprocessing*: *Text preprocessing* merupakan tahapan proses awal terhadap teks untuk mempersiapkan teks menjadi data yang akan diolah selanjutnya. Sekumpulan karakter yang bersambungan (teks) harus dipecah-pecah menjadi unsur yang lebih berarti, yang dapat dilakukan dalam tingkatan yang berbeda. Tahapan yang dilakukan dari *text preprocessing* dapat dilihat pada 2.1. Tahapan *text preprocessing* terdiri dari proses pembersihan teks (*text cleaning*), *case folding*, normalisasi kata, dan penghapusan duplikasi [5].

- 1) Pembersihan teks (*text cleaning*) adalah proses membersihkan teks yang melibatkan penghapusan karakter atau elemen yang tidak relevan atau mengganggu dalam teks, seperti tanda baca, karakter khusus, atau angka [3].
- 2) *Case folding* adalah proses mengubah semua karakter dalam teks dari huruf kecil atau huruf besar, tergantung pada preferensi atau aturan tertentu. Pada penelitian ini dilakukan *lowercase* atau merubah huruf-huruf besar menjadi huruf kecil yang bertujuan untuk membuat teks menjadi lebih seragam, sehingga mempermudah perbandingan atau pencocokan teks tanpa memperhatikan apakah karakter-karakter tersebut dalam huruf besar atau huruf kecil [5].
- 3) Normalisasi kata adalah proses mengubah kata-kata yang tidak baku atau tidak standar menjadi bentuk yang baku atau standar dalam bahasa tertentu. Tujuan normalisasi kata adalah untuk menciptakan konsistensi dalam penulisan dan komunikasi [5].
- 4) Penghapusan duplikasi adalah proses membersihkan data dari entri yang sama atau serupa yang muncul lebih dari satu kali dalam dataset. Tujuan dilakukannya penghapusan duplikasi adalah untuk menghindari kesalahan analisis, memastikan konsistensi data, dan meningkatkan kualitas analisis agar mendapat hasil yang lebih valid [8].

b) *Analisis Emosi*: Analisis emosi adalah proses untuk mengidentifikasi dan memahami emosi yang terkandung dalam teks. Tujuan dari analisis emosi adalah untuk mengukur sentimen atau perasaan dari data teks.

c) *Analisis Sentimen*: Analisis sentimen atau *opinion mining* (penambangan opini) merupakan proses untuk mengekstrak suatu opini atau pendapat dari dokumen untuk topik tertentu. Analisis sentimen merupakan bagian penting dari analisis emosi. Tujuannya adalah untuk menentukan sebuah teks memiliki sentimen positif, negatif, atau netral [2].

d) *Natural Language Processing*: *Natural Language Processing* atau biasa disingkat dengan NLP, merupakan manipulasi dari bahasa alami manusia, seperti teks atau percakapan oleh perangkat lunak komputer yang bertujuan untuk mengkaji interaksi antara komputer dengan bahasa (alami) manusia. NLP biasanya digunakan untuk menggambarkan sebuah fungsi dari komponen perangkat lunak atau perangkat keras dalam sistem komputer yang menganalisis atau mensintesis bahasa lisan atau tulisan. Kata '*natural*' atau alami adalah untuk membedakan perkataan dan tulisan manusia dengan bahasa yang lebih formal, seperti notasi matematika atau logika, atau bahasa-bahasa komputer, seperti Java, LISP, dan C++. Sebenarnya, '*Natural Language Understanding*' (NLU) dapat dikaitkan dengan tujuan untuk memiliki sistem komputer yang dapat memahami bahasa alami sebagai kekuatan manusia. NLP membantu pengambilan berperingkat yang berarti bahwa pendataan atau pengindeksan dengan juga mengurutkan bobot sebuah kata [4].

#### 1) Metode *Bag of Words* (BoW)

Metode *Bag of Words* (BoW) merupakan representasi sederhana yang digunakan pada *Natural Language Processing* (NLP) dan *Information Retrieval* (IR). Model ini juga dikenal sebagai model *vector space*.

Model *Bag of Words* (BoW) adalah salah satu cara untuk mengekstraksi fitur-fitur dari sebuah teks untuk digunakan dalam pemodelan, seperti untuk algoritma pembelajaran mesin (*machine learning*). Prinsip dasar BoW adalah menghitung frekuensi kemunculan kata-kata dalam sebuah teks tanpa memperhatikan urutan kata atau struktur gramatikalnya [7].

#### 2) Metode *Term Frequency Inverse Document Frequency* (TF-IDF)

Metode *Term Frequency Inverse Document Frequency* (TF-IDF) merupakan salah satu bagian *information retrieval system* dalam mengubah data non-terstruktur menjadi bentuk yang dapat memenuhi kebutuhan informasi. Metode ini akan memberikan bobot pada suatu kata atau term sehingga kata-kata yang sering muncul dan unik terhadap jumlah dokumen memiliki bobot yang besar.

$$IDF_i = \log\left(\frac{D}{df_i}\right) \quad (1)$$

$$W_{ij} = TF_{ij} \times IDF_i \quad (2)$$

Keterangan.

$TF_{ij}$  = frekuensi dari term ke- $i$  yang muncul pada dokumen ke- $j$

$IDF_i$  = *inverse document frequency* dari suatu term ke- $i$  terhadap keseluruhan dokumen

$D$  = jumlah keseluruhan dokumen

$df_i$  = jumlah dokumen yang mengandung term

Prinsip utama metode ini adalah bekerja dengan meningkatkan proposional skor akhir sebuah kata dengan berapa kali sebuah kata muncul dalam dokumen, tetapi juga diimbangi dengan jumlah dokumen yang mengandung kata tersebut [2].

e) *Machine Learning*: *Machine Learning* adalah teknik untuk melakukan inferensi (menitikberatkan ranah hubungan variabel) terhadap data dengan penekatan matematis. *Machine learning* berfokus pada pengembangan algoritma komputer yang memungkinkan sistem untuk belajar dari data dan pengalaman, tanpa perlu diprogram secara eksplisit. Tujuannya adalah untuk membuat model (matematis) yang merefleksikan pola-pola data yang dapat digunakan untuk melakukan prediksi atau membuat keputusan berdasarkan data yang diberikan, tanpa harus memprogram aturan-aturan secara khusus [9].

#### 1) Metode *Naïve Bayes*

Metode *Naïve Bayes* adalah salah satu algoritma klasifikasi yang berdasarkan teorema Bayes. Teorema Bayes adalah teorema yang menggambarkan perhitungan probabilitas suatu peristiwa berdasarkan probabilitas peristiwa lain yang terkait. Dalam konteks analisis emosi, program akan menghitung probabilitas label tertentu berdasarkan fitur-fitur (atribut) dari data [10].

#### 2) Metode *Support Vector Machine* (SVM)

*Support Vector Machine* (SVM) adalah algoritma yang digunakan untuk mencari *hyperplane* terbaik yang memisahkan dua atau lebih kelas data dengan jelas. *Hyperplane* ini adalah batas keputusan yang memaksimalkan *margin* antara kelas-kelas tersebut.

*Margin* adalah jarak terdekat antara *hyperplane* dan titik-titik pelatihan terdekat dari setiap kelas [2].

### 3) Metode *Logistic Regression*

*Logistic Regression* adalah algoritma klasifikasi yang digunakan untuk memodelkan hubungan antara variabel dependen biner (dalam kasus analisis emosi, kategori emosi) dengan satu atau lebih variabel independen (fitur-fitur teks). Meskipun biasanya digunakan untuk klasifikasi biner, *logistic regression* dapat digunakan untuk kasus dengan lebih dari dua kategori (multi-kelas). Dalam konteks klasifikasi multi-kelas, *logistic regression* dapat diadaptasi dengan dua pendekatan utama, yaitu

- *One-vs-Rest (OvR)* atau *One-vs-All (OvA)*: Dalam pendekatan ini, beberapa *model logistic regression* dilatih, masing-masing untuk satu kelas, dan kelas lain dianggap sebagai kelas "lainnya." Ketika program mengklasifikasikan sampel baru, program akan menjalankan semua model dan memilih kelas dengan probabilitas tertinggi. Ini adalah pendekatan yang umum digunakan dalam Regresi Logistik untuk masalah multi-kelas.
- *Multinomial Logistic Regression*: Dalam pendekatan ini, *logistic regression* digunakan untuk menangani masalah multi-kelas secara langsung tanpa menggunakan skema *One-vs-Rest*. Model ini memodelkan distribusi probabilitas langsung untuk setiap kelas dan menggunakan fungsi *softmax* untuk menghitung probabilitas kelas. Pendekatan ini dapat diterapkan ketika data multi-kelas lebih bersih [11].

f) *Confusion Matrix*: *Confusion matrix* merupakan alat ukur standar yang digunakan untuk mengetahui seberapa akurat hasil perkiraan dari sistem klasifikasi. Pada kasus klasifikasi dengan lima label emosi, matriks yang terbentuk berukuran  $5 \times 5$ . *Confusion matrix* terdiri dari empat jenis nilai, yaitu:

- 1) *True Positives (TP)*  
Jumlah sampel yang benar-benar diklasifikasikan dengan benar sebagai positif (dalam hal ini, emosi tertentu).
- 2) *True Negatives (TN)*  
Jumlah sampel yang benar-benar diklasifikasikan dengan benar sebagai negatif (dalam hal ini, bukan emosi tertentu).
- 3) *False Positives (FP)*  
Jumlah sampel yang salah diklasifikasikan sebagai positif, padahal seharusnya negatif (dalam hal ini, bukan emosi tertentu, tetapi salah diklasifikasikan sebagai emosi tersebut).
- 4) *False Negatives (FN)*  
Jumlah sampel yang salah diklasifikasikan sebagai negatif, padahal seharusnya positif (dalam hal ini, emosi tersebut, tetapi salah diklasifikasikan sebagai bukan emosi tersebut).

Dalam kasus ini, matriks memiliki lima baris dan lima kolom yang mewakili setiap kombinasi label emosi.

TABLE I. ILUSTRASI *CONFUSION MATRIX*

Aktual	Prediksi				
	0	1	2	3	4
0	TP	FN	FP	FP	FP
1	FN	TP	FN	FP	FN
2	FP	FN	TP	FP	FN
3	FP	FP	FP	TP	FN
4	FP	FN	FN	FN	TP

Berdasarkan Tabel 1, angka 0 menyatakan *sadness*, angka 1 menyatakan *joy*, angka 2 menyatakan *love*, angka 3 menyatakan *fear*, dan angka 4 menyatakan *anger*. Pada *confusion matrix* tersebut, dapat dihitung metrik-metrik untuk mengevaluasi kinerja klasifikasi model pada data. Metrik-metrik tersebut adalah

- *Akurasi*: Akurasi mengukur sejauh mana model klasifikasi benar dalam memprediksi semua kelas dengan benar dibandingkan dengan jumlah total sampel. Rumus akurasi adalah sebagai berikut.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

- *Presisi*: Presisi mengukur seberapa akurat model dalam mengklasifikasikan kasus sebagai positif. Rumus presisi adalah sebagai berikut.

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (4)$$

- *Recall*: *Recall* mengukur sejauh mana model mampu menemukan semua kasus positif yang sebenarnya. Rumus *recall* adalah sebagai berikut.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

- *F1-Score*: *F1-score* menggabungkan *presisi* dan *recall* menjadi satu nilai yang menyiratkan kualitas prediksi model secara keseluruhan.

$$F1 = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (6)$$

Makin besar nilai keempat skor tersebut, maka makin baik model yang digunakan [2].

### B. Metode Penelitian

1. Metode Pengumpulan Data  
Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh melalui perlombaan DAC IFest.
2. Langkah Penelitian  
Langkah-langkah penelitian dapat dilihat pada gambar berikut ini.

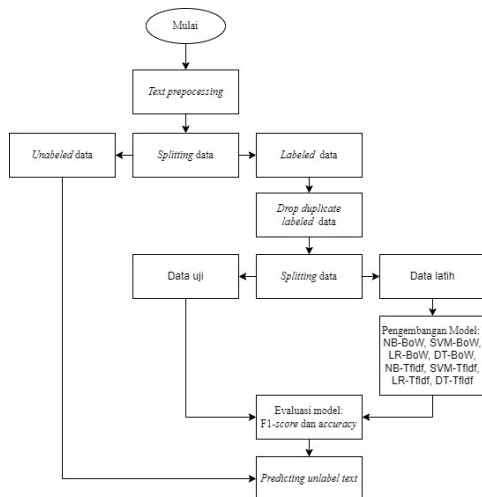


Fig. 1. Flowchart Langkah Penelitian

Pada langkah *text preprocessing*, prosesnya dapat dilihat dari gambar berikut ini.

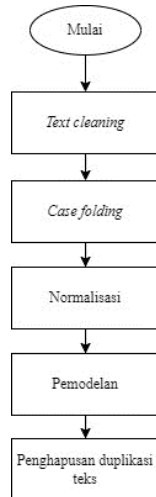


Fig. 2. Flowchart Text Preprocessing

### C. Hasil dan Pembahasan

#### 1. Dataset

Data berisikan 5151 tweets, dengan 4154 tweets telah diberi label emosi. Selanjutnya, akan dilihat karakteristik pada data yang sudah diberi label emosi (*labelled data*).

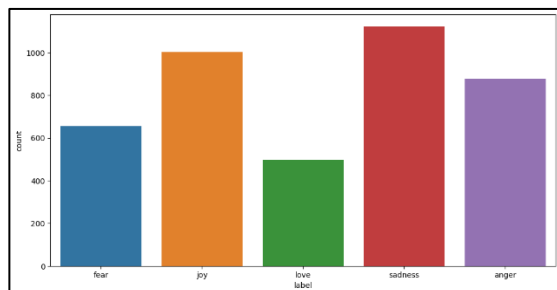


Fig. 3. Grafik Perbandingan Label pada *Labelled Data*

Berdasarkan Gbr. 3, *labelled data* terdiri dari 1123 label *sadness*, 1002 label *joy*, 876 label *anger*, 654 label *fear*, dan 498 label *love*. Hal itu menunjukkan bahwa *tweet* pada *labelled data* lebih banyak mengarah kepada label emosi *sadness*, yaitu 23% dari *labelled data*. Selain itu,

perbandingan antara *tweet* pada setiap label emosi cenderung tidak seimbang. Intipan data dapat dilihat pada tabel berikut.

TABLE II. INTIPAN ISI RAW DATA

index	label	tweet
0	fear	Lebih menyeramkan kalo punya grup WA keluarga yg isinya sharing2 hadist & ayat Al-Qur'an DAN grup WA lingkungan paroki gereja.... Serem kalo salah reply
1	joy	Hepibesdey canteeeekk [USERNAME] panjang umur, jadi pribadi yang jauh lebih baik, semoga apa yang di cita-citakan terwujud dan langgeng sama [USERNAME]
2	joy	Jumat lalu nonton ini! Selain tertarik settingnya yg di Jogja, aku penasaran stlh baca bahaya laten acak2 rambut di review Kak [USERNAME] Dan emang bagus banget! Mata sembab gak karuan. Nonton yuk mumpung masih tayang.
3	joy	Nyante dulu.. Sruput kopi biar tambah bijak...
4	love	Happy annyversarry yg pertama kalinya .. Trsrah org mau blng apa, aku gk pdli.. Ini yg aku rasain slma 1blan.. Aku cukup bhagia sm km smuanya baik" aja, jgn ada mslh yy sayang, cemburu krna org lain wajarkan.. Dan smg kita juga makin mngrti satu sama lain.
5	sadness	beli kaos banyak dengan ukuran yg udah downgrade, kucariin kok ngga ada ternyata masuk ke lemari Bapak, mbak ART ku kayaknya ga notice aku udah kurusan.

#### 2. Text Preprocessing

*Text preprocessing* dilakukan untuk menormalkan data teks sehingga menjadi data terstruktur. Proses ini dilakukan agar dapat meningkatkan kualitas data teks yang akan digunakan dalam proses pemodelan.

##### - Cleaning

*Cleaning* merupakan tahapan untuk membersihkan data teks dari karakter khusus, simbol, dan format yang tidak diinginkan. Proses ini dilakukan agar model tidak terganggu oleh format-format tersebut. Contoh hasil proses ini adalah sebagai berikut.

TABLE III. HASIL *CLEANING*

Input	Output
Lebih menyeramkan kalo punya grup WA keluarga yg isinya sharing2 hadist & ayat Al-Qur'an DAN grup WA lingkungan paroki gereja.... Serem kalo salah reply	Lebih menyeramkan kalo punya grup WA keluarga yg isinya sharing2 hadis ayat AL Qur an dan grup WA lingkungan paroki gereja Serem kalo salah reply

##### - Normalisasi

Normalisasi teks adalah proses mengubah teks dalam bentuk tertentu menjadi bentuk standar atau lebih umum. Tujuannya adalah untuk menghasilkan teks yang lebih seragam dan mudah diproses secara komputasional.

TABLE IV. HASIL NORMALISASI

Input	Output
Lebih menyeramkan kalo punya grup WA keluarga yg isinya sharing2 hadis ayat AL Qur an dan grup WA lingkungan paroki gereja Serem kalo salah reply	Lebih menyeramkan kalau punya grup WA keluarga yang isinya sharing2 hadis ayat AL Qur an dan grup WA lingkungan paroki gereja Seram kalau salah reply

Pada contoh di atas, proses normalisasi teks mengubah kata 'kalo' menjadi 'kalau' dan 'yg' menjadi 'yang'. Pada penelitian ini, kamus

normalisasi diperoleh dari <https://www.kaggle.com/code/aldonistan/preprocessing-the-indonesian-twitter-toxic-comment/notebook>

#### - Case Folding

*Case folding* melibatkan konversi semua huruf dalam teks menjadi huruf kecil (*lowercase*) atau huruf besar (*uppercase*), tergantung pada aturan yang ditentukan. Proses ini dilakukan untuk menghilangkan ambiguitas dalam teks yang disebabkan oleh variasi huruf besar-kecil. Dalam kasus ini, dilakukan *case folding* menjadi huruf kecil. Contoh hasil proses ini adalah sebagai berikut.

TABLE V. HASIL CASE FOLDING

Input	Output
Lebih menyeramkan kalau punya grup WA keluarga yang isinya sharing2 hadis ayat AL Qur an dan grup WA lingkungan paroki gereja Seram kalau salah reply	lebih menyeramkan kalau punya grup wa keluarga yang isinya sharing2 hadis ayat al qur an dan grup wa lingkungan paroki gereja seram kalau salah reply

#### 3. Splitting Data (Data berlabel dan data tidak berlabel)

Pada tahapan ini, baris-baris yang memiliki label dipisahkan dengan baris-baris yang belum memiliki label. Selanjutnya, dilakukan proses pembersihan pada data yang telah berlabel dengan tahapan sebagai berikut.

##### - Penanganan Data Duplikat

Terdapat 138 data duplikat yang dihapus dalam data yang sudah memiliki label. Pada tahapan ini, data duplikat dalam data berlabel dihapus.

##### - Penanganan Missing Value

Terdapat 1 record dari data berlabel yang tidak berisi *tweet* (NaN) setelah *preprocessing*. Karena persentase *missing value* dari data kurang dari 0,001%, maka *missing value* dapat dianggap tidak penting dan dapat dihapus.

##### - Penanganan Data yang Hanya Berisi 1 Kata

Pada penelitian ini, jika data yang dihasilkan setelah *preprocessing* hanya berisi 1 kata, terutama jika kata tersebut adalah kata yang tidak memiliki makna atau konteks yang jelas, maka ini dapat menjadi masalah dalam analisis emosi. Informasi yang dapat diekstrak dari satu kata seringkali sangat terbatas dan mungkin tidak memberikan insight yang berguna. Maka, data yang berisi 1 kata akan dihapus.

##### - Splitting Data (Data Latih dan Data Uji)

Setelah itu, data yang sudah memiliki label dibagi menjadi data latih dan data uji dengan perbandingan jumlah data 8:2. Data latih digunakan untuk membangun model prediksi, sementara data uji digunakan untuk menguji akurasi model. Tujuan pembagian data ini adalah untuk mengukur kinerja model secara obyektif dan menghindari masalah *overfitting*.

TABLE VI. HASIL SPLITTING DATA (DATA LATIH DAN DATA UJI)

Jenis Data	Jumlah
------------	--------

Data Latih	3330
Data Uji	807
Total	4037

#### 4. Pemodelan

Proses pemodelan dilakukan menggunakan *Natural Language Processing* (NLP) berupa fitur *Bag of Words* (BoW) dan fitur *Term Frequency Inverse Document Frequency* (TF-IDF), serta *Machine Learning* berupa metode *Naïve Bayes*, metode *Support Vector Machine* (SVM), dan metode *Logistic Regression*. Dengan demikian, model yang digunakan adalah NB-BoW, SVM-BoW, LR-BoW, DT-BoW, NB-TFIDF, SVM-TFIDF, LR-TFIDF dan DT-TFIDF.

#### 5. Analisis Hasil Pemodelan

Dari seluruh model yang dihasilkan, seluruh modelnya *overfit*. Hal ini terlihat dari akurasi data latih yaitu sekitar 90%, sementara data uji hanya sekitar 60%. *Overfit* dapat terjadi karena ketidakseimbangan jumlah data dari masing-masing label. Untuk mengurangi efek *overfit*, dapat dilakukan *oversampling* atau augmentasi data. Namun, setelah dilakukan *oversampling*, akurasi data uji hanya naik sedikit, tapi tetap *overfit*. Selanjutnya, evaluasi model dilakukan dengan menggunakan *confusion matrix*, dengan memperhatikan nilai *F1-score* dan akurasi.

TABLE VII. SKOR EVALUASI MODEL

Model	F1-Score	Akurasi
NB-BoW	64.36	64.36
SVM-BoW	62.25	62.25
LR-BoW	64.48	64.48
DT-BoW	49.13	49.13
NB-TFIDF	49.50	49.50
SVM-TFIDF	65.22	65.22
LR-TFIDF	66.96	66.96
DT-TFIDF	49.75	49.75

Berdasarkan skor evaluasi, model LR-TFIDF (*Logistic Regression - Term Frequency Inverse Document Frequency*) memiliki *F1-score* dan *accuracy* tertinggi bila dibandingkan dengan model lainnya. Nilai *accuracy* dari model tersebut cukup tinggi, yaitu sebesar 65,22%. Akan ditinjau lebih lanjut *confusion matrix* dan skor evaluasi untuk model LR-TFIDF.

Predicted Actual	0	1	2	3	4
0	144	34	6	5	25
1	36	139	4	3	15
2	17	11	67	1	3
3	29	10	0	83	9
4	41	16	0	2	108
precision recall f1-score support					
0	0.54	0.67	0.60	214	
1	0.66	0.71	0.68	197	
2	0.87	0.68	0.76	99	
3	0.88	0.63	0.74	131	
4	0.68	0.65	0.66	167	
accuracy			0.67	808	
macro avg	0.73	0.67	0.69	808	
weighted avg	0.69	0.67	0.67	808	
Accuracy X_train: 91.79					
Accuracy X_test: 66.96					
F1 Score: 66.96					

Fig. 4. Hasil Evaluasi Skor Model LR-TFIDF

Pada Gbr. 5, angka 0 menyatakan *sadness*, angka 1 menyatakan *joy*, angka 2 menyatakan *love*, angka 3 menyatakan *fear*, dan angka 4 menyatakan *anger*. Model ini memiliki rata-rata nilai *precision* 72,6%, *recall* 66,8%, dan *F1-score* 68,8%.

### III. PENUTUP

#### 6. Prediksi

Berdasarkan model yang telah dipilih, yaitu model LR-TFIDF (*Logistic Regression - Term Frequency Inverse Document Frequency*), akan diprediksi label-label untuk data yang tidak memiliki label. Intipan data hasil prediksi dapat dilihat pada tabel berikut.

TABLE VIII. INTIPAN DATA HASIL PREDIKSI

index	label	tweet	text_clean
0	love	Saya taKan raguKan besarNya kasih sayang Tuhan yg telah diberiKanNya kepada saya,walau momen sulit yg harus saya hadapi dalam hidup ini. Karna saat cahaya dalam hati kita mulai redup,hanya kasih Sayang dari mereka yg mencintai kita, yg mampu menjadi lentera tuk menerangiNya kembali	saya tak akan raguKan besarnya kasih sayang tuhan yang telah diberikannya kepada saya walau momen sulit yang harus saya hadapi dalam hidup ini karena saat cahaya dalam hati kita mulai redup hanya kasih sayang dari mereka yang mencintai kita yang mampu menjadi lentera untuk meneranginya kembali
1	anger	Cc in ke cebong2 dungu... Maksd hati pengen cari kesalahan [USERNAME] [USERNAME] lwat pohon plastik malah kebongkar semua skandalnya.... Sekali lagi kasian Ahok.... #2019GantiPresiden [URL]	cc in ke cebong cebong dungu maksud hati ingin cari kesalahannamename lewat pohon plastik malah terbongkar semua skandalnya sekali lagi kasihan ahok ganti presiden
2	anger	Nah, yg keterlalu adl kalo ada pjbt eselon yg menyalah2kan pelaksananya atas suatu kesalahan dokumen yg dia ttd/paraf. Istilahnya gak mau disalahkan. Pdhl tugasnya pejabat itu y neliti dokumen sbml dittd/paraf dan tgjawab atas isinya.	nah yang keterlalu adl adalah kalau ada pjbt eselon yang menyalah2kan pelaksananya atas suatu kesalahan dokumen yang dia tertanda paraf istilahnya tidak mau disalahkan padahal tugasnya pejabat itu neliti dokumen sebelum dittd paraf dan tgjawab atas isinya
3	anger	Melody masih membatasi diri ala member ya? Yang dibales mentionnya pun yang dikenal aja, hahaha Ya kalau seperti ini sih kisah Jeketi nda jauh beda dengan sebelum-sebelumnya	melody masih membatasi diri ala member ya yang dibalas mentionnya pun yang dikenal saja haha ya kalau seperti ini sih kisah jakarta tidak jauh beda dengan sebelumnya
4	anger	Rasa amarah ini membuatku merasa seperti aku akan melupakan semuanya.	rasa amarah ini membuatku merasa seperti aku akan melupakan semuanya

#### A. Kesimpulan

Pada penelitian ini, data *tweet* Indonesia diklasifikasikan berdasarkan emosi-emosi dasar manusia menggunakan fitur *Natural Language Processing* (NLP) berupa fitur *Bag of Words* (BoW) dan fitur *Term Frequency Inverse Document Frequency* (TF-IDF), serta *Machine Learning* berupa metode *Naïve Bayes*, metode *Support Vector Machine* (SVM), dan metode *Logistic Regression*. Dengan demikian, model yang dapat digunakan adalah NB-BoW, SVM-BoW, LR-BoW, DT-BoW, NB-TFIDF, SVM-TFIDF, LR-TFIDF dan DT-TFIDF. Model yang dipilih adalah LR-TFIDF (*Logistic Regression - Term Frequency Inverse Document Frequency*) karena mendapatkan hasil dengan nilai akurasi data uji terbaik sebesar 65,22% dengan rata-rata nilai *precision* 72,6%, *recall* 66,8%, dan *F1-score* 68,8%.

#### B. Saran

Dari seluruh model yang dihasilkan, seluruh modelnya *overfit*. Hal ini terlihat dari akurasi data latih yaitu sekitar 90%, sementara data uji hanya sekitar 60%. *Overfit* dapat terjadi karena ketidakseimbangan jumlah data dari masing-masing label. Untuk mengurangi efek *overfit*, dapat dilakukan *oversampling* atau augmentasi data. Untuk penelitian selanjutnya, disarankan untuk memastikan bahwa jumlah data dari masing-masing label tidak berbeda secara signifikan.

#### REFERENSI

- [1] C. Anang, "Pengaruh Media Sosial terhadap Perubahan Sosial Masyarakat di Indonesia", vol.9 No.1, 2016.
- [2] A. I Made, "Pengaruh Stemmer Bahasa Indonesia terhadap Peforma Analisis Sentimen Terjemahan Ulasan Film", 2018.
- [3] A. Jones, "Perancangan Sistem Identifikasi dan Basis Data Kata Kunci Minat Dosen untuk Ujian Pendadaran", 2020.
- [4] P. Jasen dan B. Akhmad, "Penerapan Natural Language Processing dalam Aplikasi Chatbot sebagai Media Pencarian Informasi dengan Menggunakan React".
- [5] S. Mei, M. Rahmad, dan A. Mirna, "*Emotion Classification on Indonesian Twitter Dataset*", November 2018.
- [6] N. Azward dan M. Warih, "Klasifikasi Emosi pada Pengguna Twitter Menggunakan Metode Klasifikasi Decision Tree".
- [7] P. Wahyuningdiah dan H. Retno, "Penggalian Teks dengan Model *Bag of Words* terhadap Data Twitter", Vol.2, No.1, April 2018, hlm 129-398.
- [8] S. Craig, "*Data Cleansing (Data Cleaning, Data Scrubbing)*", January 2022.
- [9] D. Yuki, "Implementasi *Machine Learning* pada Aplikasi Penjualan Produk Digital (Studi pada Grabkios).
- [10] W. Mochammad, "Algortima *Naïve Bayes*", 2019.
- [11] R. Budi, "Pembelajara Mesin (*Machine Learning*)", Desember 2021.