

Klasifikasi Emosi Pada Pengguna Twitter Menggunakan Metode Klasifikasi Decision Tree

Azward Nurfauzan¹, Warih Maharani²

^{1,2}Universitas Telkom, Bandung

¹azwardfauzan@students.telkomuniversity.ac.id, ²wmaharani@telkomuniversity.ac.id

Abstrak

Twitter merupakan salah satu media sosial yang memiliki banyak pengguna aktif di dunia, salah satunya Indonesia. Media sosial ini telah menjadi sarana komunikasi dalam berinteraksi dan sebagai sarana untuk membagikan pendapat, perasaan, dan pemikiran yang dirasakan oleh seseorang. Komunikasi yang terjadi pada twitter ini bersifat secara tidak langsung, sehingga perasaan atau emosi yang dirasakan oleh seseorang tersebut hanya dapat dilihat dalam bentuk teks pada *tweet* yang dipublikasikan. Oleh karena itu, dibutuhkan sistem klasifikasi dalam interaksi sosial media untuk mengetahui emosi atau kondisi seseorang yang dikategorikan menjadi lima jenis emosi dasar yang dikemukakan oleh Shaver yaitu *anger*, *love*, *sadness*, *happy*, dan *fear*. Metode yang digunakan dalam penelitian ini adalah metode *decision tree* untuk melakukan klasifikasi emosi pada *tweet* dan menggunakan metode TF-IDF untuk melakukan pembobotan kata. Pada penelitian ini mendapatkan hasil dengan tingkat akurasi terbaik sebesar 55% dengan rata-rata nilai *precision* 54%, *recall* 53%, dan *f1-score* 53%.

Kata kunci : klasifikasi, emosi, twitter, *decision tree*

Abstract

Twitter is one of the social media that has many active users in the world, one of which is Indonesia. This social media has become a means of communication in interacting and as a means of share opinions, feelings, and thoughts that are felt by someone. The communication that occurs on Twitter is indirectly, so the feelings or emotions that felt by someone can only be seen in the form of text in published tweets. Therefore, it takes a classification system in social media interactions to know the emotions or conditions of a person categorized into five basic types of emotions expressed by Shaver namely *anger*, *love*, *sadness*, *happy*, and *fear*. The method used in this study is decision tree method to do emotional classification on tweet and use TF-IDF method to do word weighting. In this study obtained result with the best accuracy rate of 55% with an average precision value of 54%, recall 53%, and f1-score 53%.

Keywords: classification, emotion. twitter. decision tree

1. Pendahuluan

Latar Belakang

Pada era digital saat ini, perkembangan dunia maya telah berkembang dengan pesat. Berkembangnya dunia maya ini telah memberikan konektivitas pada pengguna media sosial dalam mengekspresikan perasaan, pendapat dan pemikiran mereka mengenai berbagai topik. Pendapat atau opini yang diekspresikan oleh seseorang ini dapat disebut sebagai sentimen [1]. Pendapat atau opini seseorang dalam mengekspresikan perasaan mereka pun beragam. Menurut Shaver [2], terdapat lima emosi dasar yang dirasakan oleh manusia, yaitu *anger*, *happy*, *sadness*, *fear*, dan *love*.

Berdasarkan data dari Hootsui [3], menyatakan bahwa pada tahun 2020 salah satu *platform* yang sering digunakan oleh masyarakat Indonesia adalah Twitter. Terdapat 56% pengguna aktif media sosial di Indonesia yang menggunakan Twitter. Melihat jumlah pengguna aktif Twitter di Indonesia, masyarakat dengan bebas mengekspresikan perasaan emosi mereka yang dituangkan dalam bentuk teks dalam *tweet* [4]. Komunikasi yang terjadi pada media sosial ini merupakan komunikasi yang bersifat secara tidak langsung, sehingga sulit untuk mengetahui perasaan atau kondisi orang tersebut. Berbeda dengan komunikasi secara langsung dimana kita dapat merasakan perasaan atau kondisi dari orang tersebut. Oleh karena itu, dibutuhkan sistem klasifikasi dalam interaksi media sosial Twitter untuk mengetahui emosi atau kondisi seseorang berdasarkan *tweet* yang telah dipublikasikan.

Pada sebelumnya telah banyak dilakukan penelitian mengenai klasifikasi emosi seseorang menggunakan metode SVM dan KNN. Berdasarkan jurnal penelitian Jaishree [5], metode tersebut tidak mendapatkan hasil yang baik dalam menentukan kelas emosi seperti *fear*. Pada penelitiannya tersebut menggunakan dataset dengan bahasa Inggris yang telah dikumpulkan dan diberi label kelas emosi secara otomatis menggunakan *NRC-Lexicon*. Penelitiannya mendapatkan tingkat akurasi terbaik menggunakan *decision tree* sebesar 99,6%. Pada

penelitian Fera, et al. [6] mengenai deteksi emosi pada Twitter menggunakan *naïve bayes* dengan menggunakan data yang berbahasa Indonesia, penelitiannya menggunakan kombinasi fitur *N-gram* dan mendapatkan hasil akurasi sebesar 53%. Pada penelitian M. Hasan, et al. [7] mengenai deteksi emosi pada Twitter dengan data yang berbahasa Inggris mendapatkan hasil dengan menggunakan *decision tree* dengan nilai *precision* 90,1%, *recall* 89,9%, dan *F-measure* 90%. Pada penelitian Jaishree dan penelitian M. Hasan mendapatkan hasil akurasi yang tinggi dengan menggunakan *decision tree* dan menggunakan data bahasa Inggris. Namun untuk klasifikasi emosi dengan menggunakan data yang berbahasa Indonesia masih belum optimal. Maka dalam penelitian ini dilakukan klasifikasi emosi seseorang pada Twitter menggunakan metode *decision tree* untuk dapat mengetahui emosi yang dirasakan oleh seseorang berdasarkan *tweet* yang berbahasa Indonesia, metode ini digunakan karena memiliki tingkat akurasi yang lebih baik pada kasus klasifikasi emosi pada Twitter.

Topik dan Batasannya

Topik yang diangkat dalam penelitian tugas akhir ini, yaitu bagaimana mengimplementasikan metode *decision tree* dalam mengklasifikasikan emosi pada *tweet* di Twitter dan bagaimana performansi dan akurasi metode *decision tree* dalam mengklasifikasikan emosi pada *tweet* di Twitter. Sedangkan batasan masalah yang terdapat pada penelitian tugas akhir ini, yaitu penelitian ini hanya dapat mengklasifikasikan emosi pada *tweet* di Twitter secara eksplisit dan hanya dapat mengklasifikasikan satu jenis emosi pada setiap *tweet*. Pada penelitian ini terdapat lima jenis emosi yang digunakan, yaitu *anger*, *happy*, *sadness*, *love*, dan *fear*. Data yang digunakan dalam penelitian ini menggunakan dataset dengan Bahasa Indonesia yang berjumlah 4401 *tweet* data.

Tujuan

Tujuan dari penelitian tugas akhir ini, yaitu untuk mengetahui emosi seseorang berdasarkan *tweet* yang telah dipublikasikan menggunakan metode klasifikasi *decision tree* dan untuk mengetahui performansi metode *decision tree* dengan menggunakan proses *stemming* pada tahap *preprocessing*.

Organisasi Tulisan

Pada laporan Tugas Akhir ini berisi mengenai beberapa bagian, yaitu bagian Studi Terkait yang menjelaskan studi yang mendukung penelitian ini. Selanjutnya bagian Sistem Yang Dibangun yang menjelaskan mengenai gambaran sistem yang diterapkan pada penelitian. Pada bagian Evaluasi menjelaskan mengenai hasil pengujian dan analisis hasil pengujian. Pada bagian Kesimpulan berisi mengenai kesimpulan dan saran untuk penelitian selanjutnya.

2. Studi Terkait

2.1 Sentiment Analysis

Sentiment Analysis atau disebut juga sebagai *opinion mining* adalah bidang studi yang mempelajari dan menganalisis mengenai opini, kritik, ataupun emosi seseorang yang diluapkan dalam bentuk teks [8]. *Sentiment Analysis* ini banyak digunakan untuk mengetahui dan menganalisis ulasan, pendapat ataupun emosi seseorang pada suatu topik. Seseorang dapat mengekspresikan emosi atau perasaan mereka dalam bentuk perilaku tertentu atau mengekspresikan perasaan mereka melalui tulisan. Perilaku masyarakat dalam menuangkan opini mereka memiliki pengaruh yang besar untuk individu ataupun untuk organisasi bisnis dalam mengambil suatu keputusan.

2.2 TF-IDF

Term Frequency – Inverse Document (TF-IDF) adalah suatu metode pembobotan kata untuk mengetahui seberapa pentingnya kata dalam suatu dokumen atau korpus [9]. *Term Frequency* (TF) adalah frekuensi kemunculan suatu kata dalam dokumen, sedangkan *Inverse Document Frequency* (IDF) adalah nilai pembobotan suatu kata berdasarkan frekuensi kemunculan kata tersebut pada beberapa dokumen. Berikut merupakan rumus pada algoritma TF-IDF:

$$IDF(w) = \log \left(\frac{N}{DF(w)} \right) \quad (1)$$

$$TF - IDF(w, d) = TF(w, d) \times IDF(w) \quad (2)$$

Keterangan:

W merupakan *word* (kata)

N merupakan banyaknya dokumen

DF(w) merupakan banyaknya kata dalam dokumen

TF-IDF(w,d) merupakan nilai pembobotan suatu kata
 D merupakan dokumen

2.3 Decision Tree

Algoritma *decision tree* adalah salah satu metode klasifikasi *supervised learning* yang menyelesaikan masalah dengan merepresentasikan permasalahan tersebut ke dalam struktur pohon [10]. Berikut merupakan rumus pada algoritma *decision tree* yang digunakan dalam penelitian ini:

$$Gini(D) = 1 - \sum_{i=1}^m p_j^2 \tag{3}$$

Keterangan:

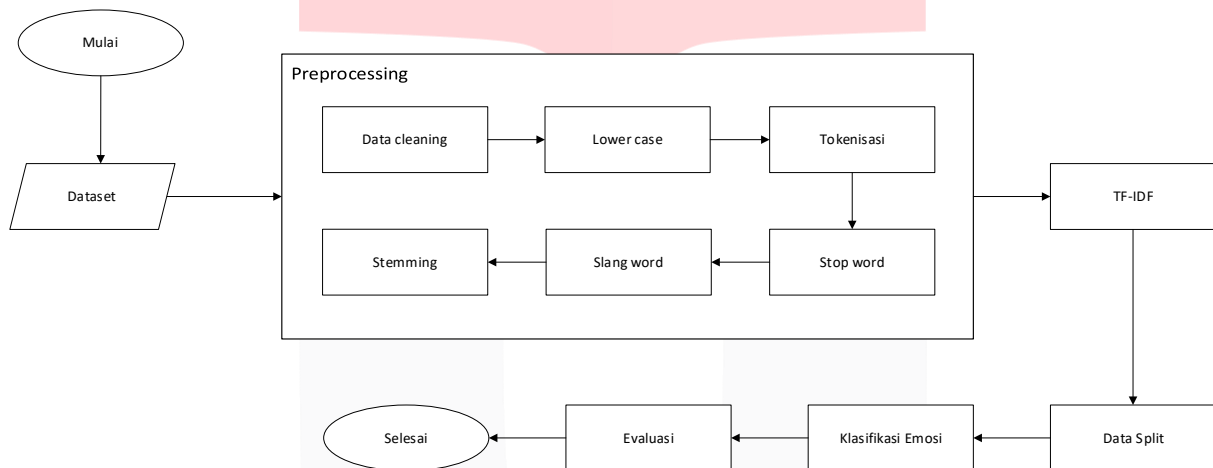
Gini(D) merupakan impurity dari suatu partisi D

m merupakan jumlah indeks

p_i merupakan peluang tuple D pada indeks ke i

3. Sistem yang Dibangun

Pada bagian ini akan menjelaskan mengenai sistem klasifikasi emosi menggunakan klasifikasi *decision tree*. Alur pada sistem yang dibangun ini dapat dilihat pada Gambar 1. Alur Sistem yang Dibangun, sebagai berikut:



Gambar 1. Alur Sistem yang Dibangun

3.1 Dataset

Dataset yang digunakan dalam penelitian ini berasal dari penelitian yang telah dilakukan oleh Mei Silviana Saputri dan Rahmad Mahendra mengenai *Emotion Classification on Indonesian Twitter Dataset* [11]. Dataset ini terdiri dari 4401 *tweet* yang berbahasa Indonesia dengan lima jenis emosi yang terbagi menjadi 997 kelas *sadness*, 1017 kelas *happy*, 649 kelas *fear*, 637 kelas *love*, dan 1101 kelas *anger*.

3.2 Preprocessing Data

Pada tahap preprocessing data, data akan dibersihkan untuk menghilangkan data yang masih terdapat *noise* untuk dapat diproses pada tahap selanjutnya [12]. Tahap preprocessing ini terbagi menjadi beberapa bagian, yaitu *data cleaning*, *lowercase*, *tokenisasi*, *stop word removal*, *slang handling* dan *stemming*.

3.2.1 Data Cleaning

Data Cleaning adalah proses yang bertujuan untuk membersihkan data yang terdapat bagian yang tidak diperlukan (*noise*) [12].

Tabel 1. Data Cleaning

Input	Output
"Pagi yg selalu mengajarkan kita, bahwa ada harapan di setiap langkah kehidupan... . ."	Pagi yg selalu mengajarkan kita bahwa ada

	harapan di setiap langkah kehidupan
--	-------------------------------------

3.2.2 Lowercase

Lowercase adalah proses yang bertujuan untuk merubah semua huruf menjadi huruf kecil [13].

Tabel 2. Lowercase

Input	Output
Pagi yg selalu mengajarkan kita bahwa ada harapan di setiap langkah kehidupan	pagi yg selalu mengajarkan kita bahwa ada harapan di setiap langkah kehidupan

3.2.3 Tokenisasi

Tokenisasi adalah proses yang berfungsi untuk membagi kalimat menjadi beberapa kata atau token [9].

Tabel 3. Tokenisasi

Input	Output
pagi yg selalu mengajarkan kita bahwa ada harapan di setiap langkah kehidupan	'pagi', 'yg', 'selalu', 'mengajarkan', 'kita', 'bahwa', 'ada', 'harapan', 'di', 'setiap', 'langkah', 'kehidupan'

3.2.4 Stop Word Removal

Stop Word Removal adalah proses yang berfungsi untuk menghapus kata yang tidak penting dalam dokumen [9].

Tabel 4. Stop Word Removal

Input	Output
'pagi', 'yg', 'selalu', 'mengajarkan', 'kita', 'bahwa', 'ada', 'harapan', 'di', 'setiap', 'langkah', 'kehidupan'	'pagi', 'yg', 'mengajarkan', 'harapan', 'langkah', 'kehidupan'

3.2.5 Slang Word

Slang Word adalah proses yang berfungsi untuk merubah kata yang tidak formal menjadi kata yang sesuai dengan KBBI [14].

Tabel 5. Slang Word

Input	Output
'pagi', 'yg', 'mengajarkan', 'harapan', 'langkah', 'kehidupan'	'pagi', 'yang', 'mengajarkan', 'harapan', 'langkah', 'kehidupan'

3.2.6 Stemming

Stemming adalah proses yang berfungsi untuk menghapus imbuhan pada suatu kata untuk mengurangi ukuran kamus yang berisikan kata dari koleksi dokumen [9].

Tabel 6. Stemming

Input	Output
'pagi', 'yang', 'mengajarkan', 'harapan', 'langkah', 'kehidupan'	'pagi', 'yang', 'ajar', 'harapan', 'langkah', 'hidup'

3.3 TF-IDF

Pada tahap *Term Frequency-Inverse Document Frequency* (TF-IDF) ini, dilakukan proses pembobotan pada setiap kata yang terdapat pada dokumen. Berikut contoh pembobotan TF-IDF pada lima sampel *tweet* sebagai berikut:

D1 = terima kasih senang (*happy*)

D2 = laku cinta (*love*)

D3 = bikin orang kesal (*anger*)

D4 = lelah bosan jenuh (*sadness*)

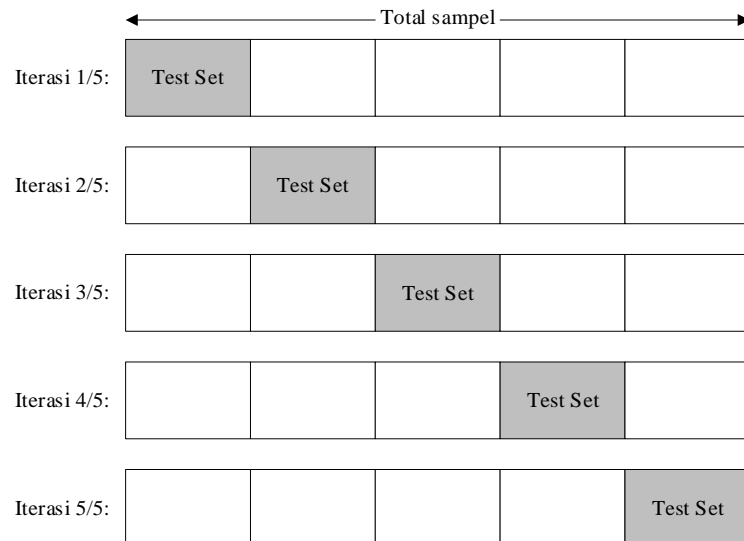
D5 = jatuh cinta pandang mata (*love*)

Tabel 7. TF-IDF

Term	TF					DF	IDF	TF-IDF				
	D1	D2	D3	D4	D5			D1	D2	D3	D4	D5
terima	1	0	0	0	0	1	$\log(5/1)=0,7$	0,7	0	0	0	0
kasih	1	0	0	0	0	1	$\log(5/1)=0,7$	0,7	0	0	0	0
senang	1	0	0	0	0	1	$\log(5/1)=0,7$	0,7	0	0	0	0
laku	0	1	0	0	0	1	$\log(5/1)=0,7$	0	0,7	0	0	0
cinta	0	1	0	0	1	2	$\log(5/2)=0,4$	0	0	0,4	0	0,4
bikin	0	0	1	0	0	1	$\log(5/1)=0,7$	0	0	0,7	0	0
orang	0	0	1	0	0	1	$\log(5/1)=0,7$	0	0	0,7	0	0
kesal	0	0	1	0	0	1	$\log(5/1)=0,7$	0	0	0,7	0	0
lelah	0	0	0	1	0	1	$\log(5/1)=0,7$	0	0	0	0,7	0
bosan	0	0	0	1	0	1	$\log(5/1)=0,7$	0	0	0	0,7	0
jenuh	0	0	0	1	0	1	$\log(5/1)=0,7$	0	0	0	0,7	0
jatuh	0	0	0	0	1	1	$\log(5/1)=0,7$	0	0	0	0	0,7
pandang	0	0	0	0	1	1	$\log(5/1)=0,7$	0	0	0	0	0,7
mata	0	0	0	0	1	1	$\log(5/1)=0,7$	0	0	0	0	0,7

3.4 Data Split

Pada tahap ini dataset dibagi menjadi dua bagian, yaitu menjadi data *train* dan data *test*. Pembagian dataset ini menggunakan metode *K-Fold Cross Validation*. Dataset dibagi rata sebanyak k bagian dan melakukan iterasi data training dan data testing sebanyak k, sehingga seluruh dataset pernah dijadikan sebagai data *train* dan data *test* [15]. Berikut merupakan gambaran mengenai *K-Fold Cross Validation* dengan menggunakan nilai k=5.



Gambar 2. K-Fold Cross Validation dengan k=5

3.5 Klasifikasi Emosi

Pada tahap ini dilakukan proses klasifikasi emosi dengan menggunakan metode *decision tree* untuk menentukan prediksi kelas label emosi. Pada metode ini melakukan hasil klasifikasi dengan menggunakan struktur pohon keputusan, dimana setiap node yang ada menggambarkan atribut dan cabang menggambarkan nilai dari atribut. Pada pembentukan struktur pohon keputusan ini dilakukan dengan perhitungan berdasarkan hasil dari pembobotan kata dengan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) [16]. Pada tahap TF-IDF akan diambil 1000 kata dengan nilai tertinggi atau kata yang memiliki frekuensi kemunculan terbanyak pada dokumen. Setelah itu, dilakukan perhitungan pembobotan pada atribut atau kata yang telah diboboti dengan TF-IDF dengan menggunakan rumus *gini*. Atribut yang memiliki nilai terbesar akan dijadikan sebagai *root* pada *decision tree*. Setelah itu, dilakukan pemisahan node berdasarkan nilai *gini* yang tertinggi hingga pembentukan *decision tree* selesai. Pada penelitian ini, *decision tree* diimplementasikan dengan menggunakan *Scikit-Learn*, dimana *Scikit-Learn* ini sudah teruji dan memiliki dokumentasi yang lengkap.

3.6 Evaluasi Sistem

Pada tahap evaluasi ini dilakukan perhitungan nilai *precision*, *recall* dan *F1-Score*. Perhitungan ini digunakan untuk mengukur tingkat akurasi dan performansi dari model klasifikasi yang telah dibangun pada penelitian ini [17]. Perhitungan tingkat akurasi ini didapatkan berdasarkan hasil dari *Confusion Matrix*, berikut merupakan table *Confusion Matrix* pada penelitian ini:

Tabel 8. Confusion Matrix

A	B	C	D	E	Kelas
x	x	x	x	x	A-Sadness
x	x	x	x	x	B-Happy
x	x	x	x	x	C-Fear
x	x	x	x	x	D-Love
x	x	x	x	x	E-Anger

- Rumus Precision:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

- Rumus Recall:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

- Rumus F1 Score:

$$F1 \text{ Score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \quad (7)$$

4. Evaluasi

Pada bagian ini akan dijelaskan mengenai hasil pengujian dari sistem yang telah dibangun serta performansi yang telah didapat.

4.1 Hasil Pengujian

4.1.1 Kata yang paling berpengaruh untuk setiap kelas

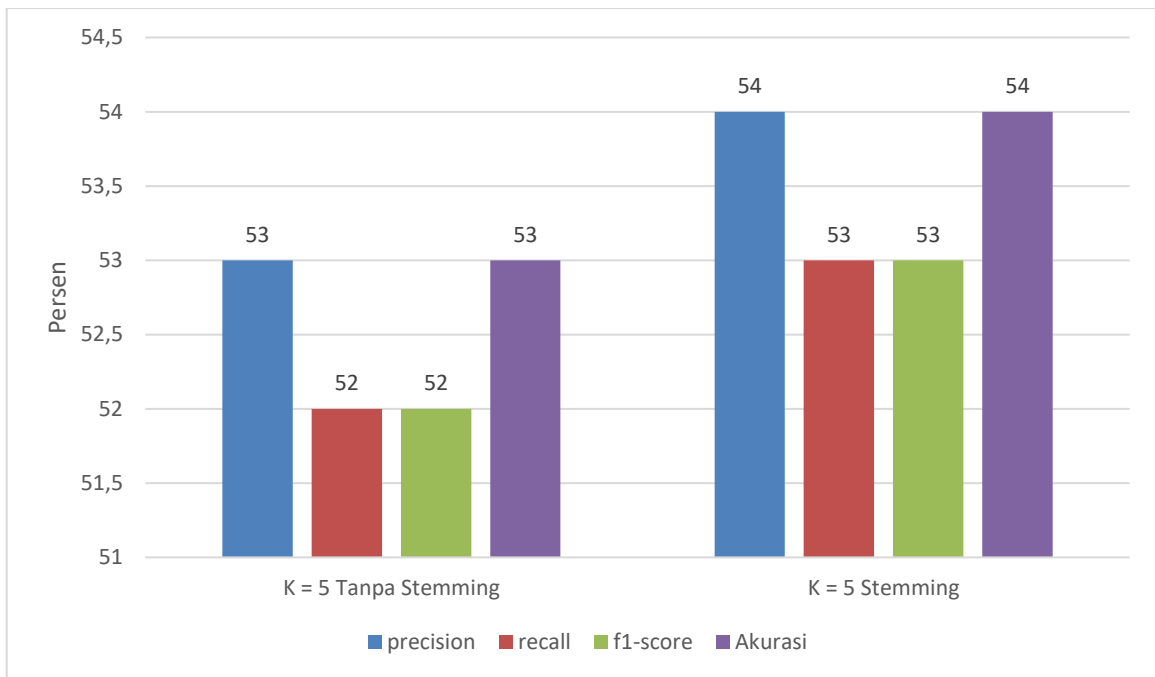
Pada tahap pembobotan kata menggunakan metode TF-IDF, diurutkan kata atau *term* dengan nilai tertinggi pada setiap kelasnya. Pengekstrakan fitur ini bertujuan untuk dapat mengetahui kata atau *term* yang berpengaruh untuk setiap kelas atau label.

Tabel 9. Lima Term yang Memiliki Pengaruh Tertinggi Pada Setiap Kelas

Anger	Happy	Sadness	Fear	Love
enggak	enggak	enggak	takut	cinta
orang	orang	orang	enggak	sayang
sudah	banget	sudah	kalo	enggak
kalo	sudah	kalo	banget	banget
begitu	kasih	banget	sudah	kasih

4.1.2 Pengujian menggunakan nilai *k-fold* sebesar 5 dengan proses *stemming* dan tanpa *stemming*

Pada pengujian ini dilakukan dengan cara membagi jumlah dataset menjadi data *train* dan data *test* dengan nilai *k-fold* sebesar 5 dengan menggunakan proses *stemming* dan tanpa menggunakan proses *stemming*. Pada proses ini dataset dibagi menjadi 5 bagian dan dilakukan pengulangan data *train* dan data *test* secara bergantian. Berikut merupakan hasil dari pengujian dengan nilai *k-fold* sebesar 5:



Gambar 3. Hasil Pengujian K=5

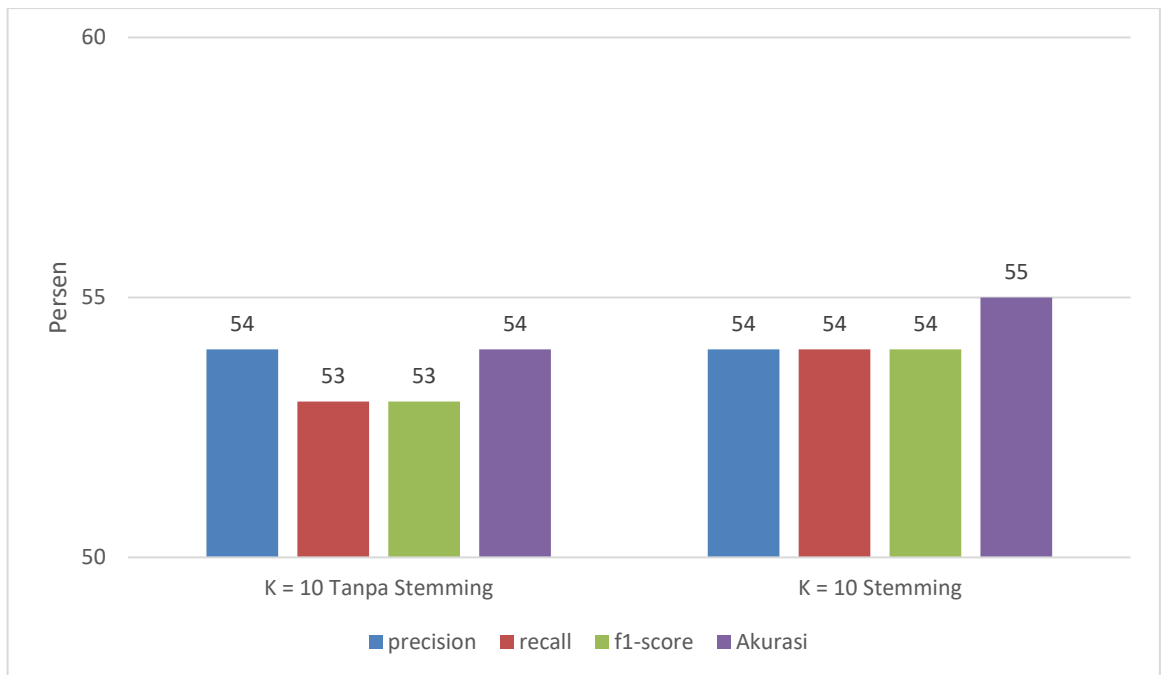
Tabel 10. Akurasi Data Validasi Setiap *fold* pada K = 5

Fold	K = 5 Tanpa Stemming	K = 5 Stemming
Fold 1	44%	44%
Fold 2	55%	54%
Fold 3	52%	53%
Fold 4	48%	51%
Fold 5	49%	49%
Rata-rata	50%	50%

Hasil pengujian klasifikasi emosi menggunakan metode *decision tree* dengan nilai $k=5$ dengan menggunakan proses *stemming* menghasilkan nilai akurasi 53%, *precision* 54%, *recall* 52%, dan *f1-score* 52%. Sedangkan hasil tanpa menggunakan proses *stemming* menghasilkan nilai akurasi 54%, *precision* 54% *recall* 53%, dan *f1-score* 54%.

4.1.3 Pengujian menggunakan nilai *k-fold* sebesar 10 dengan proses *stemming* dan tanpa *stemming*

Pada pengujian ini dilakukan dengan cara membagi jumlah dataset menjadi data train dan data test dengan nilai *k-fold* sebesar 10 dengan menggunakan proses *stemming* dan tanpa menggunakan proses *stemming*. Pada proses ini dataset dibagi menjadi 10 bagian dan dilakukan pengulangan data train dan data test secara bergantian. Berikut merupakan hasil dari pengujian *k-fold* sebesar 10:



Gambar 4. Hasil Pengujian 2 K = 10

Tabel 11. Akurasi Data Validasi Setiap *fold* pada K = 10

Fold	K = 5 Tanpa Stemming	K = 5 Stemming
Fold 1	48%	43%
Fold 2	45%	48%
Fold 3	55%	52%
Fold 4	57%	56%
Fold 5	57%	58%
Fold 6	55%	52%
Fold 7	50%	51%
Fold 8	53%	50%
Fold 9	51%	54%
Fold 10	49%	47%
Rata-rata	52%	52%

Hasil pengujian klasifikasi emosi menggunakan metode *decision tree* dengan nilai $k = 10$ memberikan hasil yang lebih tinggi dari hasil nilai $k = 5$. Hasil nilai $k = 10$ tanpa menggunakan proses *stemming* memberikan hasil akurasi sebesar 54% dengan rata-rata *precision* 54%, *recall* 53%, dan *f1-score* 53%. Sedangkan hasil nilai $k = 10$ dengan menggunakan proses *stemming* memberikan hasil akurasi 55%, *precision* 54%, *recall* 54%, dan *f1-score* 54%.

4.2 Analisis Hasil Pengujian

Pada penelitian ini, komposisi jumlah data pada setiap kelasnya terbagi secara tidak merata. Dataset ini terbagi menjadi 997 kelas *sadness*, 1017 kelas *happy*, 649 kelas *fear*, 637 kelas *love*, dan 1101 kelas *anger*. Tetapi kelas yang memiliki komposisi data yang besar tidak mendapatkan hasil yang lebih baik dibandingkan

dengan kelas yang memiliki komposisi data yang lebih sedikit. Hal ini disebabkan karena kurangnya *knowledge* kata yang dapat menggambarkan nilai pada kelasnya. Hal ini dapat dilihat berdasarkan hasil dari penentuan term tertinggi pada setiap kelasnya. Pada hasil tersebut, terdapat beberapa kata yang sama dengan frekuensi kemunculan kata terbanyak pada kelas yang berbeda. Pada kelas *fear* dan kelas *love* memiliki komposisi data yang lebih sedikit, tetapi memiliki beberapa kata yang dengan frekuensi kemunculan terbanyak yang berbeda. Sehingga kata-kata tersebut dapat menggambarkan nilai pada setiap kelasnya dengan lebih baik.

Pada hasil pengujian yang didapat, pembagian dataset pada klasifikasi emosi menggunakan *decision tree* dengan nilai $k = 10$ menghasilkan performansi yang lebih baik. Hal ini sebabkan oleh metode yang digunakan, yaitu *decision tree* yang merupakan salah satu metode klasifikasi *supervised learning*. Metode *supervised learning* tersebut membutuhkan data train yang besar untuk mendapatkan tingkat akurasi yang baik, hal ini dikarenakan metode *supervised learning* membutuhkan *knowledge* pada setiap kata untuk setiap kelasnya. Sehingga pada nilai $k = 10$ memberikan hasil yang lebih baik dikarenakan memiliki *knowledge* kata yang lebih banyak.

Proses *stemming* pada tahap *preprocessing* memberikan pengaruh terhadap klasifikasi emosi menggunakan *decision tree* ini. Hal ini disebabkan karena proses *stemming* merupakan proses yang bertujuan untuk menghapus imbuhan yang terdapat pada suatu kata. Sehingga proses *stemming* pada *preprocessing* ini berpengaruh pada hasil dari pembobotan kata dari proses TF-IDF yang menyebabkan adanya beberapa kata yang memiliki makna yang sama.

Pada penelitian ini mendapatkan hasil terbaik dengan akurasi sebesar 55% dengan rata-rata *precision* 54%, *recall* 54%, dan *f1-score* 54%. Hasil tersebut didapatkan menggunakan proses *stemming* dan menggunakan nilai $k = 10$. Tetapi penelitian ini tidak menghasilkan performansi yang lebih baik dibandingkan dengan performansi penelitian yang dilakukan oleh M. S. Saputri, et al. [11] terkait dataset yang sama. Pada penelitiannya tersebut mendapatkan hasil terbaik dengan nilai *f1-score* 69,73% dengan menggunakan kombinasi fitur. Hal ini disebabkan karena terdapat beberapa data yang bersifat implisit dan kurangnya *knowledge* kata yang dapat menggambarkan nilai dari tiap kelas labelnya.

5. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan terhadap klasifikasi emosi pada pengguna twitter menggunakan metode klasifikasi *decision tree*, telah mendapatkan hasil dengan nilai akurasi terbaik sebesar 55% dengan rata-rata nilai *precision* 54%, *recall* 54%, dan *f1-score* 54%. Hasil pengujian tersebut didapatkan dengan menggunakan nilai $k = 10$ dan menggunakan proses *stemming* pada tahap *preprocessing*. Pada tahap *preprocessing*, penggunaan proses *stemming* memberikan hasil yang lebih baik pada sistem. Tanpa menggunakan proses *stemming* pada *preprocessing* menyebabkan adanya beberapa kata yang memiliki makna yang sama, tetapi berbeda term karena memiliki imbuhan sehingga berpengaruh pada proses pembobotan kata. Pemodelan klasifikasi *decision tree* ini sangat bergantung pada dataset yang digunakan dalam membangun model klasifikasi emosi. Maka penulis memberikan saran untuk penelitian selanjutnya untuk menggunakan dataset berbahasa Indonesia yang lain dengan pembagian kelas label yang lebih merata dan menggunakan proses *stemming* pada tahap *preprocessing*.

Referensi

- [1] R. Bose, R. K. Dey, S. Roy, and D. Sarddar, "Sentiment Analysis on Online Product Reviews," *Adv. Intell. Syst. Comput.*, vol. 933, pp. 559–569, 2020.
- [2] P. R. Shaver, U. Murdaya, and R. C. Fraley, "Structure of the Indonesian emotion lexicon," *Asian J. Soc. Psychol.*, vol. 4, no. 3, pp. 201–224, 2001.
- [3] S. Kemp, "Digital 2020: Indonesia," 18-Feb-2020. [Online]. Available: <https://datareportal.com/reports/digital-2020-indonesia>. [Accessed: 24-Dec-2020].
- [4] M. Purver and S. Battersby, "Experimenting with distant supervision for emotion classification," *EACL 2012 - 13th Conf. Eur. Chapter Assoc. Comput. Linguist. Proc.*, pp. 482–491, 2012.
- [5] J. Ranganathan, A. S. Irudayaraj, and A. A. Tzacheva, "Automatic Detection of Emotions in Twitter Data - A Scalable Decision Tree Classification Method," no. July, 2018.
- [6] F. Fanesya, R. C. Wihandika, and Indriati, "Deteksi Emosi Pada Twitter Menggunakan Metode Naïve Bayes Dan Kombinasi Fitur," vol. 3, no. 7, pp. 6678–6686, 2019.
- [7] M. Hasan, E. Rundensteiner, and E. Agu, "EMOTEX : Detecting Emotions in Twitter Messages," pp. 1–10, 2014.
- [8] B. Liu, *Sentiment Analysis: A Fascinating Problem*. 2012.
- [9] A. Mishra and S. Vishwakarma, "Analysis of TF-IDF Model and its Variant for Document Retrieval," pp. 772–776, 2015.
- [10] B. Gupta, A. Rawat, A. Jain, A. Arora, and N. Dhama, "Analysis of Various Decision Tree Algorithms

- for Classification in Data Mining,” *Int. J. Comput. Appl.*, vol. 163, no. 8, pp. 15–19, 2017.
- [11] M. S. Saputri, R. Mahendra, and M. Adriani, “Emotion Classification on Indonesian Twitter Dataset,” no. January 2019, 2018.
- [12] V. Srividhya and R. Anitha, “Evaluating Preprocessing Techniques in Text Categorization,” *Int. J. Comput. Sci. Appl.*, pp. 49–51, 2010.
- [13] S. Mujilawati, “Pre-Processing Text Mining Pada Data Twitter,” *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 2089–9815, 2016.
- [14] N. Aliyah Salsabila, Y. Ardhitto Winatmoko, A. Akbar Septiandri, and A. Jamal, “Colloquial Indonesian Lexicon,” *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, pp. 226–229, 2019.
- [15] S. Yadav, “Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification,” no. Cv, 2016.
- [16] Y. Astuti, “Klasifikasi Posting Twitter Cuaca Provinsi Diy,” no. November, pp. 1–9, 2015.
- [17] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.

