

Deteksi Emosi Pada Twitter Menggunakan Metode Naïve Bayes Dan Kombinasi Fitur

Fera Fanesia¹, Randy Cahya Wihandika², Indriati³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹fefanesia@student.ub.ac.id, ²rendicahya@ub.ac.id, ³indriati.tif@ub.ac.id

Abstrak

Emosi bersifat umum dan penting dalam kehidupan yang membentuk perilaku manusia. Mendeteksi emosi memberikan peranan penting dalam berbagai aspek karena dapat diterapkan dalam berbagai bidang seperti pengambilan keputusan, memprediksi keadaan emosi manusia, memberikan *review* terhadap kualitas produk, melacak dukungan pada masalah politik, dan mengenali gangguan depresi seseorang. Mengidentifikasi emosi dapat menggunakan data tekstual yaitu berupa teks, karena teks dapat digunakan untuk berkomunikasi dan menyampaikan informasi. Salah satu media sosial untuk bertukar informasi yaitu Twitter. Twitter berisi informasi tentang sikap serta keadaan emosi seseorang. Oleh karena itu, dilakukan deteksi emosi pada Twitter untuk menentukan emosi seseorang dengan menerapkan metode *Naïve Bayes* dan kombinasi fitur. Model klasifikasi *Naïve Bayes* yang digunakan pada penelitian ini yaitu *Multinomial Naïve Bayes* untuk tipe data diskrit dan *Bernoulli Naïve Bayes* untuk tipe data biner, kombinasi fitur yang digunakan pada penelitian yaitu fitur linguistik, fitur ortografik, dan kombinasi fitur *N-gram*. Hasil akurasi terbaik didapatkan pada pengujian kombinasi fitur *N-gram* sebesar 0,555, sedangkan hasil akurasi pada pengujian kombinasi fitur sebesar 0,5317 yang meliputi fitur linguistik, fitur ortografik, dan fitur *N-gram* memiliki nilai akurasi lebih baik dibandingkan pengujian dengan fitur tunggal dan lebih rendah dibandingkan pengujian kombinasi fitur *N-gram*. Hal ini dikarenakan pengaruh dari fitur linguistik, fitur ortografik, dan fitur *N-gram*. Penggabungan kombinasi fitur ini bisa menutupi kelemahan masing-masing fitur yang bisa meningkatkan akurasi walaupun peningkatannya tidak terlalu signifikan.

Kata kunci: deteksi emosi, klasifikasi, kombinasi fitur, naïve bayes, twitter.

Abstract

Emotion shapes human behavior in general and very important in life. Detecting emotions provides an important role in various aspects because it can be applied in various fields such as decision-making, predicting human emotions conditions, providing a review product quality, tracking support for political problems, and recognizing depression disorders. Identifying emotions can use textual data that is text, text can be used to communicate and declare information. The social media that used to exchange information is Twitter. Twitter contains information about human attitude and human emotions. Therefore, emotional detection is needed to determine human emotions using Naïve Bayes method and feature combinations. This research using several Naïve Bayes classification models namely Bernoulli Naïve Bayes for binary data types and Multinomial Naïve Bayes for discrete data types. Feature Combination used in this research is as follows: linguistic features, orthographic features, and N-gram feature combinations. The best accuracy result obtained a value of 0.555 that is in testing N-gram feature combinations. While the combination of features including linguistic features, orthographic features, and N-gram features produced an accuracy value of 0.5317 which means this value was better than testing with a single feature and lower than testing the N-gram feature combinations. This is due to the influence of linguistic features, orthographic features, and N-gram features. Based on these results it can be concluded that by using combination features can cover the weaknesses of each feature that can improve the performance of accuracy even though the increase is not too significant.

Keywords: classification, emotion detection, feature combinations, naïve bayes, twitter.

1. PENDAHULUAN

Emosi bersifat umum dan penting dalam semua aspek kehidupan manusia. Teori ilmiah mengenai emosi telah dikembangkan dalam beberapa penelitian dalam filsafat, psikologi, ilmu kognitif, dan ilmu saraf yang merupakan elemen penting dari sifat manusia. Dalam bidang komputer pun emosi juga menarik perhatian pada beberapa penelitian khususnya pada bidang interaksi manusia komputer. Sistem interaksi manusia komputer yang berhasil harus dapat mengenali, menafsirkan, dan memproses emosi manusia (Canales dan Barco, 2014). Mendeteksi emosi memberikan peranan penting dalam berbagai aspek karena dapat diterapkan dalam berbagai bidang seperti pengambilan keputusan dalam lingkungan sosial maupun lingkungan bisnis. Dalam bidang pendidikan, analisis emosi dapat digunakan untuk *e-learning* sistem cerdas yaitu dapat memprediksi keadaan emosi siswa ketika suatu peristiwa terjadi selama sesi *e-learning*, untuk meningkatkan kualitas belajar online serta kemampuan siswa (Daouas dan Lejmi, 2018). Selain itu, analisis emosi dapat digunakan dalam bisnis yaitu untuk memberikan *review* terhadap kualitas suatu produk, dan analisis konsumen dalam mengidentifikasi keluhan konsumen (Nurlaila, Saptono, & Wiranto, 2017). Dalam bidang politik dapat digunakan untuk untuk melacak dukungan pada masalah politik. Dalam bidang psikologi, untuk melacak kondisi emosional seseorang, serta untuk mengenali gangguan depresi dan kecemasan seseorang (Yuliana dan Faizah, 2018). Dengan demikian, mendeteksi emosi memiliki banyak kegunaan dan menunjukkan pentingnya melakukan deteksi emosi.

Pada identifikasi emosi terdapat beberapa jenis karakteristik fisiologi yang digunakan untuk mendeteksi emosi yaitu suara, ekspresi wajah, gerakan tangan, gerakan tubuh, detak jantung, tekanan darah, dan informasi dari data tekstual (Consoli, 2009). Data tekstual paling banyak dipelajari karena teks relatif lebih mudah. Mengidentifikasi emosi yang diekspresikan dalam teks memiliki dua alasan. Pertama, emosi bisa implisit dan dipicu oleh peristiwa atau situasi tertentu. Kedua, teks menggambarkan suatu peristiwa atau situasi yang menyebabkan emosi bisa tanpa kata-kata secara eksplisit (Bata et al., 2015).

Teks merupakan suatu media yang dapat digunakan untuk berkomunikasi dan

menyampaikan informasi, serta dapat mengekspresikan emosi. Teks juga media utama dalam komunikasi menggunakan komputer seperti email, blog, dan media sosial untuk bertukar informasi dan berkomunikasi. Media sosial untuk bertukar informasi salah satunya yaitu Twitter dimana pengguna membagikan pesan terbaru, berbagi informasi, berdiskusi suatu hal tertentu dalam bentuk tulisan sepanjang 280 karakter. Twitter juga sebagai tempat mencurahkan hati dan berisi informasi tentang sikap serta keadaan emosi seseorang.

Deteksi emosi dilakukan dengan pendekatan klasifikasi emosi. Klasifikasi adalah proses dengan menggabungkan atau mengelompokkan dua atau lebih data yang memiliki kesamaan pada suatu kriterianya. Mendeteksi emosi dalam klasifikasi emosi diberikan kelas emosi sesuai dengan yang ditentukan. Metode klasifikasi yang biasa digunakan antara lain metode *C4.5*, *Naïve Bayes*, *K-Nearest Neighbor* (KNN), dan metode lainnya. *Naïve Bayes* merupakan algoritme yang didasarkan pada teorema bayes yang menggunakan probabilitas bersyarat untuk mengklasifikasikan data ke dalam kelas yang ditentukan sebelumnya (Kumari, 2014).

Oleh karena itu, penelitian ini menggunakan metode *Naïve Bayes* sebagai metode klasifikasi untuk deteksi emosi Twitter berbahasa Indonesia, dengan klasifikasi ke dalam 5 kelas yaitu senang, sedih, takut, cinta, dan marah. Selain itu digunakan kombinasi beberapa fitur yaitu fitur *N-gram*, fitur linguistik, dan fitur ortografik. Fitur *N-gram* meliputi *unigram* dan *bigram*, fitur Linguistik meliputi jumlah singkatan, jumlah kata kapital, jumlah dari JJ, PRP, NN pada *Part-of-Speech tagging*, sedangkan fitur ortografik meliputi adanya tanda tanya, adanya tanda seru, adanya tanda tanya di akhir twit, adanya tanda seru diakhir twit, dan jumlah huruf kapital. Metode *Naïve Bayes* pada penelitian ini diharapkan mampu melakukan deteksi emosi pada Twitter berbahasa Indonesia dengan tepat dan sesuai, sehingga maksud utama dari penelitian ini terpenuhi yaitu dapat mengetahui pengaruh kombinasi fitur-fitur yang digunakan terhadap metode *Naive Bayes* untuk deteksi emosi Twitter.

2. LANDASAN KEPUSTAKAAN

2.1. Emosi

Dalam Kamus Besar Bahasa Indonesia,

emosi didefinisikan sebagai keadaan dan reaksi psikologis dan fisiologis seperti kegembiraan, kesedihan, keharuan, dan kecintaan. Emosi memiliki peran penting dalam interaksi antar manusia. Emosi menunjukan suatu keputusan yang cerdas serta rasional, membantu untuk saling berhubungan dengan mengekspresikan perasaan.

Sebuah model kategori emosi yang sangat populer untuk digunakan adalah model emosi (Ekman, 1999) Ekman melakukan penelitian di seluruh dunia dan menemukan enam emosi yang sama dengan ekspresi wajah. Enam spesifikasi dasar emosi manusia yaitu *anger* (marah), *disgust* (jijik), *fear* (takut), *happiness* (bahagia), *sadness* (sedih) dan *surprise* (terkejut). Model emosi Ekman telah digunakan oleh beberapa penelitian yang digunakan untuk mengenali keadaan emosi dari data tekstual maupun ekspresi wajah.

2.2. Twitter

Twitter merupakan media sosial dimana pengguna membaca dan menulis pesan singkat tentang berbagai topik, Twitter diluncurkan pada tahun 2006 dan didirikan oleh Jack Dorsey. Jumlah pengguna Twitter aktif melebihi 22% dari pengguna internet di dunia (Kayser & Bierwisch, 2016) sehingga terdapat lebih dari 342 juta pengguna yang aktif setiap bulannya. Twitter memiliki perbedaan dibandingkan dengan media sosial lainnya yaitu jumlah karakter pada Twitter hanya dibatasi 280 karakter, sehingga saat pengguna menulis pesan maka disebut *tweet* atau suatu kicauan.

2.3. Text Mining

Text mining adalah proses mendapatkan pembelajaran yang intensif dimana mengumpulkan dokumen dengan menggunakan perangkat analisis tertentu. *Text mining* merupakan bidang ilmu yang baru, peneliti-peneliti mencoba menyelesaikan kelebihan informasi, jika dibandingkan dengan teknik dari data mining, pemrosesan Bahasa alami, dan *information retrieval* (pencarian informasi), *text mining* mengekstrak informasi yang berguna dari berbagai sumber data dengan melakukan identifikasi serta eksplorasi suatu pola data dari kumpulan dokumen, dan pola-pola yang ditemukan tidak diformalisasikan ke dalam *record* di database melainkan dalam data teksual yang tidak terstruktur (Feldman & Sanger, 2007)

2.4. Preprocessing Text

Pre-processing text merupakan bagian yang paling penting dari *text mining*. *Preprocessing text* merupakan proses dimana dokumen dilakukan menyesuaikan dengan proses yang akan dilalui. (Kannan dan Gurusamy, 2014).

2.4.1. Tokenisasi

Tahapan ini dilakukan pemisahan deretan kata atau *tweet* menjadi token dan

juga menghilangkan karakter-karakter tertentu seperti tanda baca, angka, tag HTML, link URL, *username*, dan karakter lain.

2.4.2. Filtering

Tahapan ini merupakan proses penghapusan *stopwords* yang dilakukan untuk memilih kata-kata yang dianggap penting. *Stopwords* merupakan kumpulan kata yang tidak memiliki makna atau tidak mengandung informasi yang signifikan. Sehingga dengan dihapusnya *stopword* hanya terdapat kata-kata yang penting dalam dokumen.

2.4.3. Stemming

Tahapan ini merupakan proses pemotongan imbuhan atau pengembalian kata menjadi kata dasar.

2.5. Kombinasi Fitur

Tabel 1 Kombinasi Fitur

Tipe		Id	Deskripsi Fitur
Fitur Linguistik		F1	Jumlah singkatan.
		F2	Jumlah kata kapital.
		F3	Jumlah kata sifat atau <i>adjective</i> (JJ) yang ada pada <i>tweet</i> .
		F4	Jumlah kata ganti orang atau <i>pronoun</i> (PRP) yang ada pada <i>tweet</i> .
		F5	Jumlah kata benda atau noun(NN) yang ada pada <i>tweet</i> .
Fitur Ortografik		F6	Adanya tanda tanya.
		F7	Adanya tanda seru.
		F8	Adanya tanda tanya diakhir <i>tweet</i> .
		F9	Adanya tanda seru diakhir <i>tweet</i> .
		F10	Jumlah huruf kapital.
Fitur N-Gram		F11	Unigram.
		F12	Bigram.

2.6. Naïve Bayes

Naïve Bayes merupakan salah satu algoritme klasifikasi karena kemudahan penggunaannya serta memakan waktu pemrosesan yang cepat bila diterapkan pada dataset yang besar. Model probabilitas pada klasifikasi Naïve Bayes ini berdasarkan teorema Bayesian. Algoritme Naïve Bayes menggunakan keseluruhan probabilitas, yaitu probabilitas dokumen terhadap kategori (prior), kemudian teks akan dikategorikan berdasarkan probabilitas maksimum (posterior) (Raschka, 2014). Tahap pertama yang dilakukan yaitu dilakukan perhitungan probabilitas dokumen dalam kategori (*prior*):

$$P(w_j) = \frac{N_{wj}}{N_c} \quad (1)$$

N_{wj} adalah jumlah data latih yang termasuk jenis emosi w_j dan N_c adalah jumlah seluruh data latih. Berikut persamaan probabilitas posterior.

$$P(W_j | X_i) = \frac{P(X_i | W_j) \cdot P(W_j)}{P(X_i)} \quad (2)$$

Dimana $P(X_i | W_j)$ adalah probabilitas *likelihood*, $P(W_j)$ adalah prior, $P(X_i)$ adalah peluang kemunculan kata. Dalam persamaan 3 merupakan penyederhaan perhitungan posterior pada proses klasifikasi dengan peluang kemunculan kata dihilangkan, karena tidak akan berpengaruh pada perbandingan hasil klasifikasi pada tiap kategori.

$$P(W_j | X_i) = P(X_i | W_j) \cdot P(W_j) \quad (3)$$

Probabilitas *likelihood* yang berawal dari $i=1$ hingga $i=k$ sehingga perhitungan nilai posterior dapat dilakukan dalam persamaan 4.

$$P(X | W_j) = P(X_1 | W_j) \times P(X_2 | W_j) \times \dots \times P(X_k | W_j) \quad (4)$$

2.6.1. Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan salah satu metode spesifik dari metode *Naïve Bayes* yang menggunakan *conditional probability*. *Conditional probability* dapat dilakukan dengan menggunakan frekuensi kemunculan suatu kata pada suatu kelas (*raw term frequency*). *Multinomial Naïve Bayes* ini juga merupakan salah satu metode *supervised learning* dalam *machine learning* pada proses pengklasifikasian teks dengan menggunakan nilai probabilitas suatu kelas dalam suatu dokumen (Raschka, 2014). Berikut persamaan *Multinomial Naïve Bayes* menggunakan *conditional probability* menggunakan *add one*

dan *Laplace smoothing*.

$$P(X_i | W_j) = \frac{N_{xi,wj} + 1}{N_{wj} + |V|} \quad (5)$$

$N_{xi,wj} + 1$ = Jumlah dari suatu kata query yang muncul dalam satu kelas, penambahan angka 1 digunakan untuk menghindari nilai 0.

N_{wj} = Jumlah kata yang ada pada kelas c_j

V = Jumlah seluruh kata unik yang ada pada semua kelas

2.6.2. Bernoulli Naïve Bayes

Bernoulli Naïve Bayes menerapkan algoritme pelatihan dan klasifikasi, untuk klasifikasi “dengan menggunakan binary (0 dan 1) pada fitur biner. Nilai 1 pada Bernoulli Naïve Bayes menyatakan fitur tersebut terdapat dalam dokumen dan 0 menyatakan fitur tidak terdapat dalam dokumen” (Raschka, 2014).

$$P(X | W_j) = P(X_i | W_j)^b \times (1 - P(X_i | W_j))^{(1-b)} \quad (b \in 0,1) \quad (6)$$

Keterangan:

$P(X_i | W_j)$ = Probabilitas kata yang terdapat dalam dokumen kelas W_j

Bit = Kemunculan “kata x_i yang terdapat dalam dokumen, Jika terdapat kata dalam dokumen maka bit = 1 dan jika tidak terdapat dalam” dokumen maka bit = 0.

$$P(x_i | w_j) = \frac{df_{xi,y} + 1}{df_{y+2}} \quad (7)$$

Keterangan:

$df_{xi,y}$ = Jumlah “dokumen dalam dataset training yang mengandung fitur x_i dan termasuk dalam kelas” w_j .

df_C = Jumlah “dokumen dalam dataset training yang termasuk dalam” kelas x_i .

+1 dan +2 = Parameter *Laplace Smoothing*

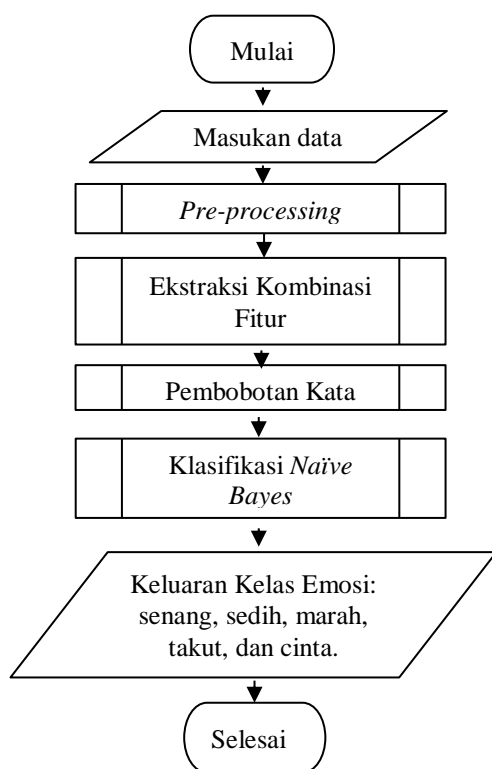
3. METODOLOGI PENELITIAN

3.1. Teknik Pengumpulan Data

Data pada penelitian ini yaitu menggunakan data sekunder dari penelitian sebelumnya mengenai klasifikasi emosi pada dataset Twitter oleh (Saputri et al., 2018). Pada penelitian tersebut terdapat 4403 twit dalam Bahasa

Indonesia dengan lima label kelas emosi yaitu cinta, marah, senang, sedih, dan takut. Data tersebut didapatkan dengan menggunakan Twitter *Streaming* API selama dua minggu, dimulai dari 1 Juni 2018 sampai 14 Juni 2018. Data yang diambil berupa kalimat dari status Twitter pengguna. Dari data tersebut diambil sejumlah 4396 twit yang dibagi menjadi dua tipe data yaitu data latih dan data uji dimana 80% data dijadikan data latih dan 20% data dijadikan data uji. Data latih yang digunakan berjumlah 811 twit berlabel senang, 797 twit berlabel sedih, 509 twit berlabel cinta, 515 twit berlabel takut, dan 882 twit berlabel marah yang diambil dari 3514 data latih. Data uji yang digunakan berjumlah 202 twit berlabel senang, 199 twit berlabel sedih, 134 twit berlabel cinta, 127 twit berlabel takut, dan 220 twit berlabel marah yang diambil dari 882 data uji.

3.2. Gambaran Umum Sistem



Gambar 1 Diagram Alir Sistem

4. PENGUJIAN DAN HASIL ANALISIS

Pada bagian ini menjelaskan mengenai pengujian yang dilakukan dan menganalisis hasil dari pengujian tersebut. Hasil analisis yang dilakukan menggunakan metode *Confusion Matrix*. Hasil evaluasi berupa nilai, akurasi, *recall*, *precision*, dan *f-measure* yang ditampilkan dalam bentuk tabel dan diagram.

4.1. Pengujian Terhadap Fitur *Unigram*

Pengujian ini melakukan klasifikasi dengan menerapkan metode *Naïve Bayes* tanpa melalui proses kombinasi fitur. Hasil pengujian fitur *Unigram* dihasilkan tingkat akurasi sebesar 0,5124, *precision* sebesar 0,5554, *recall* sebesar 0,5326, dan *F-measure* sebesar 0,5068.

Setelah dilakukan analisis dari dokumen uji yang diklasifikasikan dengan salah, terdapat beberapa twit memiliki kalimat yang pendek yaitu hanya memiliki beberapa kata pada sebuah twit, hal ini menyebabkan dalam pengambilan nilai fitur pada dokumen uji hanya bergantung pada beberapa kata tersebut. Misal pada dokumen uji “*Alhamdulillah lancar acaranya semoga berkah dunia-akhirat Amin*” dokumen uji tersebut diklasifikasikan salah karena hanya menyimpan sedikit nilai fitur kata *Unigram*. Faktor selanjutnya yang menyebabkan klasifikasi dokumen uji yang salah yaitu antara kelas satu dan kelas lainnya memiliki kata-kata yang hampir sama, hal tersebut menyebabkan dalam pengambilan nilai fitur dari kelas satu dan kelas lainnya memiliki nilai yang hampir sama. Sehingga, dalam pengklasifikasian dokumen uji akan sulit dikenali yang berdampak pada hasil klasifikasi.

4.2. Pengujian Terhadap Fitur Tunggal

4.2.1. Pengujian Terhadap Fitur Linguistik

Hasil pengujian fitur linguistik dihasilkan tingkat akurasi sebesar 0,29, *precision* sebesar 0,288, *recall* sebesar 0,275, dan *F-measure* sebesar 0,259.

Pada pengujian ini terdapat penurunan sebesar 0,222 terhadap pengujian *Unigram*. Setelah dilakukan analisis dari dokumen uji memiliki penurunan akurasi dikarenakan dalam pengambilan nilai fitur pada fitur linguistik yang jarang muncul pada dokumen uji dan pada pengambilan fitur dengan *PoS Tagging* terdapat kata yang didapatkan tidak ada dalam kamus sehingga nilai fiturnya 0. Sehingga dalam nilai fitur pada dokumen uji tidak mencerminkan pada pengklasifikasian dokumen ujinya dan berdampak pada hasil klasifikasi.

4.2.2. Pengujian Terhadap Fitur Ortografik

Hasil pengujian fitur ortografik dihasilkan tingkat akurasi sebesar 0,212, *precision* sebesar

0,174, *recall* sebesar 0,194, dan *F-measure* sebesar 0,125.

Pada pengujian ini terdapat penurunan 0,30 terhadap pengujian tanpa kombinasi fitur. Setelah dilakukan analisis dari dokumen uji memiliki penurunan akurasi dikarenakan dalam pengambilan nilai fitur pada fitur ortografik yang jarang muncul pada dokumen uji, hal ini menyebabkan dalam pengambilan nilai fitur pada dokumen uji tidak mencerminkan pada pengklasifikasian dokumen ujinya. Sehingga dalam pengklasifikasian dokumen uji akan sulit dikenali yang berdampak pada hasil klasifikasi.

4.2.3. Pengujian Terhadap Fitur N-gram

Hasil pengujian fitur *N-gram* dihasilkan tingkat akurasi sebesar 0,555, *precision* sebesar 0,559, *recall* sebesar 0,530, dan *F-measure* sebesar 0,517.

Pada pengujian ini terdapat peningkatan akurasi sebesar 0,002 terhadap pengujian tanpa kombinasi fitur. Fitur *N-gram* merupakan kombinasi dari fitur *Unigram* dan fitur *Bigram*. Pada fitur *Unigram* mengandung kata tunggal yang berada pada data latih dan fitur *Bigram* mengandung kata tunggal yang bersebelahan di dapat dari fitur *Unigram*, setiap dokumen latih memiliki kata yang berbeda dan akan memiliki kata tunggal yang bervariasi pada setiap *tweet*. Pada fitur *Unigram*, pengambilan nilai fitur antar kelasnya memiliki kata tunggal yang hampir sama. Sehingga akan menyebabkan susah melakukan klasifikasi kelas karena perbedaan antara masing-masing kelas yang nilainya hampir sama. Sedangkan pada fitur *Bigram* yang merupakan kata ganda dari *unigram* sehingga hal tersebut akan berpengaruh pada kemunculan fitur dalam dokumen, kemungkinan kata tunggal yang merupakan perbedaan antar kelas pada dokumen latih muncul bersebelahan akan sedikit ditemukan pada dokumen uji.

4.3. Pengujian Terhadap Kombinasi Dua Fitur

4.3.1. Pengujian terhadap fitur linguistik dan fitur ortografik

Hasil pengujian fitur linguistik dan fitur ortografik dihasilkan tingkat akurasi sebesar 0,3049, *precision* sebesar 0,2958, *recall* sebesar 0,3020, dan *F-measure* sebesar 0,2765.

Pada pengujian ini terdapat peningkatan akurasi sebesar 0,01 terhadap pengujian fitur

linguistik dan terdapat peningkatan akurasi sebesar 0,09 terhadap pengujian fitur ortografik. Analisis dari dokumen uji yang diklasifikasikan dengan salah dari penggabungan fitur linguistik dan fitur ortografik. Dalam pengambilan nilai fitur pada fitur linguistik dan fitur ortografik yang jarang muncul pada dokumen uji. Sehingga dalam pengambilan nilai fiturnya pada dokumen uji tidak mencerminkan klasifikasi dokumen ujinya dan akan berdampak pada hasil klasifikasi.

4.3.2. Pengujian terhadap fitur linguistik dan fitur N-gram.

Hasil pengujian fitur linguistik dan fitur *N-gram* dihasilkan tingkat akurasi sebesar 0,3458, *precision* sebesar 0,2919, *recall* sebesar 0,3244, dan *F-measure* sebesar 0,3244.

Pengujian ini terdapat peningkatan akurasi sebesar 0,05 terhadap pengujian fitur linguistik dan terdapat penurunan akurasi sebesar 0,32 terhadap pengujian fitur *N-gram*. Pada pengujian gabungan fitur linguistik dan fitur *N-gram* memiliki kelemahan yang sama seperti pada pengujian fitur linguistik dan pengujian fitur *N-gram*. Dalam pengambilan nilai fitur pada fitur linguistik yang jarang muncul pada dokumen uji, sehingga dalam pengambilan nilai fiturnya tidak mencerminkan klasifikasi dokumen ujinya. Sedangkan pengambilan nilai fitur pada fitur *N-gram* terdapat kata-kata dari fitur *bigram* lebih banyak terjadi kemunculan kata umum antar kelasnya daripada kata yang menjadi perbedaan dari semua kelasnya.

4.3.3. Pengujian terhadap fitur ortografik dan fitur N-gram.

Hasil pengujian fitur ortografik dan fitur *N-gram* dihasilkan tingkat akurasi sebesar 0,5374, *precision* sebesar 0,5683, *recall* sebesar 0,5376, dan *F-measure* sebesar 0,5296.

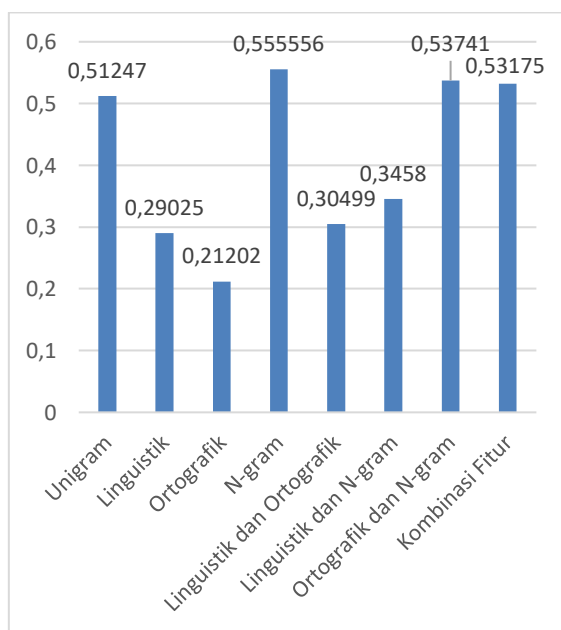
Pada pengujian ini terdapat peningkatan akurasi sebesar 0,32 terhadap pengujian fitur ortografik dan terdapat penurunan akurasi sebesar 0,02 terhadap fitur *N-gram*. Pada pengujian gabungan fitur ortografik dan fitur *N-gram* memiliki kelemahan yang sama seperti pada pengujian fitur ortografik dan pengujian fitur *N-gram*. Dalam pengambilan nilai fitur pada fitur ortografik yang jarang muncul pada dokumen uji, sehingga dalam pengambilan nilai fiturnya tidak mencerminkan pada pengklasifikasian dokumen ujinya. Sedangkan

pengambilan nilai fitur pada fitur *N-gram* terdapat kata-kata dari fitur *bigram* lebih banyak terjadi kemunculan kata umum antar kelasnya daripada kata yang menjadi pembeda dari semua kelasnya.

4.4. Pengujian terhadap fitur linguistik, fitur ortografik, dan fitur *N-gram*.

Hasil pengujian kombinasi fitur dihasilkan tingkat akurasi sebesar 0,5317, *precision* sebesar 0,5472, *recall* sebesar 0,5416, dan *F-measure* sebesar 0,5040.

Penggunaan gabungan kombinasi fitur terdapat peningkatan akurasi sebesar 0,02 terhadap pengujian unigram dan mempunyai nilai akurasi lebih baik dibandingkan pengujian fitur tunggal, hal ini karena pengaruh dari fitur linguistik, fitur ortografik dan fitur *N-gram*. Dapat disimpulkan bahwa kombinasi fitur ini akan menutupi kelemahan dari masing-masing fitur akan tetapi peningkatannya tidak terlalu signifikan.



Gambar 2 Grafik Perbandingan Akurasi Pengujian

Pada Gambar 2 ditunjukkan hasil akurasi pada setiap pengujian. Pada hasil akurasi tertinggi terdapat pada pengujian fitur *N-gram* sebesar 0,5554. Sehingga fitur *N-gram* tersebut memberikan nilai fitur yang paling berpengaruh dibandingkan fitur linguistik dan fitur ortografik. Hasil akurasi pada pengujian kombinasi fitur sebesar 0,5317 mempunyai nilai akurasi yang lebih baik dibandingkan pengujian dengan fitur tunggal, berbeda dengan menggunakan fitur tunggal yang hanya mencapai 0,29 pada fitur

linguistik dan 0,212 pada fitur ortografik.

Setelah dilakukan analisis dengan melihat Gambar 6.1, ketika fitur linguistik dengan akurasi 0,29 dan fitur ortografik dengan akurasi 0,212 dilakukan pengujian kombinasi antara keduanya maka menghasilkan akurasi yang juga rendah yaitu 0,3. Begitupun pada fitur linguistik maupun fitur ortografik yang dilakukan pengujian kombinasi dengan fitur *N-gram* yang juga menghasilkan akurasi yang cukup rendah, hal ini dikarenakan fitur linguistik dan fitur ortografik yang memberikan nilai fitur tidak terlalu signifikan, sedangkan kombinasi fitur *N-gram* memberikan pengaruh lebih signifikan. Terbukti bahwa pada dokumen uji "*Selamat Siang Sahabat Pos. Kiriman tsb sudah diterima ya sahabat. Terima kasih telah mempercayakan dan menggunakan layanan kami :)*". Pada pengujian fitur linguistik diklasifikasikan salah, pada pengujian fitur ortografik diklasifikasikan salah, pada fitur *N-gram* diklasifikasikan benar dan pada kombinasi fitur dari ketiganya diklasifikasi benar, sehingga ketika digabungkan semua fiturnya menghasilkan klasifikasi benar, hal ini dikarenakan pada fitur *N-gram* memberikan nilai fitur yang lebih baik dibandingkan fitur linguistik dan fitur ortografik. Sehingga penggabungan kombinasi fitur menunjukkan akan menutupi kekurangan dari masing-masing fitur.

Namun nilai akurasi pengujian *N-gram* lebih tinggi 0,02 dari pengujian kombinasi fitur, karena terdapat beberapa faktor yang menyebabkan selisih tingkat akurasinya. Terlihat pada dokumen uji "*Setuju, ajarkan kedamaian dan sampaikan pemahaman islam sebagai agama sempurna dalam mengatur hidup di dunia hingga akhirat. baik dalam adab individu, masyarakat dan bernegara. Allahu a'lam*". Pada dokumen uji tersebut hasil pengujian fitur *N-gram* diklasifikasikan benar namun hasil pengujian kombinasi fitur diklasifikasikan salah, hal ini disebabkan karena dokumen uji tersebut pada fitur linguistik dan fitur ortografik tidak memberikan nilai yang signifikan, terbukti pada pengujian yang terdapat fitur linguistik dan fitur ortografik diklasifikasikan salah, tetapi pada pengujian unigram, pengujian *N-gram*, dan pengujian fitur ortografik & fitur *N-gram* diklasifikasikan benar. Hal tersebut dikarenakan dokumen uji lebih banyak memberikan nilai pada fitur unigram dan bigram dibandingkan dengan fitur linguistik dan fitur ortografik. Sehingga faktor

tersebut berpengaruh ketika dilakukan pengujian kombinasi fitur.

Tetapi fitur linguistik dan fitur ortografik juga berpengaruh pada pengujian kombinasi fitur. Terbukti pada dokumen uji *“Terkadang kita harus merasakan sakitnya mencintai Ketika rasa takut kehilangan itu ada Ketika kebersamaan yang telah terjalin Membuat kita tak ingin berpisah Dan kamu adalah salah satu alasan Mengapa aku takut kehilangan dan tak ingin berpisah”*. Pada dokumen uji tersebut hasil pengujian *Unigram*, *N-gram*, dan pengujian fitur Ortografik & fitur *N-gram* diklasifikasikan dengan salah, sedangkan hasil pengujian lainnya diklasifikasikan dengan benar, termasuk juga pengujian kombinasi fitur diklasifikasikan dengan benar. Hal tersebut disebabkan bahwa fitur linguistik dan fitur ortografik juga berpengaruh pada pengambilan nilai fiturnya, sehingga ketika dilakukan pengujian kombinasi fitur bisa menutupi kelemahan masing-masing fitur yang bisa meningkatkan akurasi walaupun peningkatannya tidak terlalu signifikan.

Dari hasil pengujian dan analisis tersebut didapatkan kesimpulan bahwa dalam mendeteksi emosi Twitter dengan kombinasi fitur dapat mendapatkan klasifikasi dengan lebih baik, walaupun nilai akurasi tertinggi hanya mencapai 0,5559, hal ini dikarenakan oleh beberapa faktor yang dapat menurunkan tingkat akurasinya. Pada deteksi emosi ini menggunakan klasifikasi *multi class* yaitu terdapat 5 kelas, dengan banyaknya kelas akan mempengaruhi hasil akurasi dikarenakan peluang dari masing-masing kelas lebih kecil, berbeda dengan *binary class* yang peluangnya 50%. Dan terdapat faktor lain yang memengaruhi hasil akurasi yaitu terdapat beberapa dokumen twit yang berbeda kelas tetapi memiliki kata-kata yang hampir sama.

Tabel 3 Contoh Dokumen Twit

Kelas	Dokumen Twit
Senang	Setelah mendaki kelaparan, gak lupa kita makan siang dulu karena, makan siang kali ini kita di temenin nasi ayam kampung.
Sedih	Ada drama drama cryingnya juga. Apalagi kalo lapar biasa makanan ada tersedia tinggal ambil. Kalo dikos musti cari makan sendiri pas gaada teman nah ini bisa bisa nangis. Jadi kangen ibu.
Takut	Teman ngajak main ke Thailand. Kebetulan saya suka wisata kuliner, tapi disana banyak yg gak halal, nanti malah gak makan apa2.

Cinta	selamat siang cinta calon teman hidupku belum makan siang kah? ini aku buatin makan siang, dimakan ya sampe habis jangan dibuang2, capek tau masaknya
Marah	Namanya juga kerja harus profesional, kalo gak profesional mau makan apa ? Cinta? Gak kenyang lebih suka makan daging apalagi yg bakarnya

Dari Tabel 3 yang menunjukkan dokumen twit pada masing-masing kelas yang diklasifikasikan salah. Terdapat beberapa kata yang hampir mirip di setiap kelasnya yaitu kata “makan, lapar, suka, siang, teman, dan cinta” hal tersebut menyebabkan dalam pengambilan nilai fitur antara kelas satu dan kelas yang lainnya memiliki nilai yang hampir sama. Sehingga dalam pengklasifikasian dokumen uji akan sulit dikenali yang akan berdampak pada hasil klasifikasi yang dapat menurunkan hasil akurasi. Terdapat beberapa twit yang tidak merepresentasikan emosi sebenarnya terlihat pada Tabel 3 contoh dokumen twit dengan kelas takut yang tidak merepresentasikan emosi takut, melainkan lebih merepresentasikan emosi senang. Dapat disimpulkan bahwa emosi seseorang tidak terlalu nampak pada twit seseorang sehingga akurasi yang didapatkan cukup rendah.

5. KESIMPULAN

Pada penelitian ini dilakukan deteksi emosi pada Twitter dengan menggunakan metode *Naïve Bayes* serta kombinasi fitur yang meliputi fitur linguistik, fitur ortografik, dan fitur *N-gram*. Berdasarkan hasil pengujian didapatkan hasil akurasi tertinggi yaitu pada pengujian kombinasi fitur *N-gram* sebesar 0,5554. Fitur ini merupakan kombinasi dari fitur *Unigram* dan fitur *Bigram*, kombinasi fitur *N-gram* tersebut memberikan nilai fitur yang paling berpengaruh dibandingkan fitur linguistik dan fitur ortografik. Hasil akurasi pada pengujian kombinasi fitur sebesar 0,5317 yang meliputi fitur linguistik, fitur ortografik, dan fitur *N-gram* memiliki nilai akurasi lebih baik dibandingkan pengujian dengan fitur tunggal dan lebih rendah dibandingkan pengujian kombinasi fitur *N-gram*. Hal ini karena pengaruh dari fitur linguistik dan fitur ortografik. Dapat disimpulkan bahwa penggabungan kombinasi fitur bisa menutupi kelemahan masing-masing fitur yang bisa meningkatkan akurasi walaupun peningkatannya tidak terlalu signifikan dan emosi seseorang tidak terlalu nampak pada twit sehingga akurasi yang didapatkan cukup rendah.

DAFTAR PUSTAKA

- Bata, J., Suyoto, Pranowo, 2015. Leksikon Untuk Deteksi Emosi Dari Teks Bahasa Indonesia 8.
- Binali, H., Potdar, V., 2012. Emotion detection state of the art. CUBE International Information Technology Conference, p. 501.
- Canales, L., Martínez-Barco, P., 2014. Emotion Detection from text: A Survey. JISIC, pp. 37–43.
- Consoli, D., 2009. Emotions That Influence Purchase Decisions And Their Electronic Processing. Annales Universitatis Apulensis Series Oeconomica.
- Crannell, W.C., Clark, E., Jones, C., James, T.A., Moore, J., 2016. A pattern-matched Twitter analysis of US cancer-patient sentiments. Journal of Surgical Research 206, 536–542.
- Daouas, T., Lejmi, H., 2018. Emotions recognition in an intelligent elearning environment. Interactive Learning Environments 26, 991–1009.
- Ekman, P., 1999. Basic emotions. Handb. Cogn. Emot..
- Feldman, R., Sanger, J., 2007. The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press.
- Kannan, D.S., Gurusamy, V., 2014. Preprocessing Techniques for Text Mining 7.
- Kayser, V. & Bierwisch, A., 2016. Using Twitter for Foresight: An Opportunity?. Futures, pp. 50-63.
- Kumari, A., 2014. Study on Naive Bayesian Classifier and its relation to Information Gain. International Journal on Recent and Inovation Trends in Computing and Communication Vol II, pp. 601-603.
- Nurlaila, A., Saptono, R., Wiranto, W., 2017. Classification Of Customers Emotion Using Naïve Bayes Classifier (Case Study: Natasha Skin Care). ITSMART Vol. 6 No.2 6, 92–97.
- Ozturk, N. & Ayvaz, S., 2017. Analysis on Twitter: A Text Mining Approach to the Syrian Refugee. Telematics and Informatics, pp. 136-147.
- Saputri, M.S., Mahendra, R., Adriani, M., 2018.