Analisis Emosi Melalui Media Sosial Twitter Dengan Menggunakan Metode Naïve Bayes dan Perbandingan Fitur N-gram dan TF-IDF

Fildza Sakinah Alnaz[1], Warih Maharani[2]

[1,2] Universitas Telkom, Bandung
[1]fildzaalnaz@students.telkomuniversity.ac.id, [2]wmaharani@telkomuniversity.ac.id

Abstrak

Media sosial semakin marak digunakan dan diminati oleh banyak orang di seluruh dunia. Twitter merupakan salah satu media sosial yang aktif digunakan oleh masyarakat. Melalui twitter, seseorang dapat mengekspresikan diri mereka dan saling berbagi informasi mengenai kehidupan, pengalaman, pendapat maupun emosi yang dirasakannya. Emosi dikatakan sebagai kondisi secara umum yang disebabkan oleh suatu peristiwa yang dialami seseorang. Keadaan emosional tersebut dapat mempengaruhi peranan penting dalam kehidupan manusia seperti lingkup sosial, bisnis, maupun dalam pengambilan keputusan seperti cara berfikir seseorang mengenai perasaannya sendiri. Maka dari itu, pada penelitian ini dilakukan analisis jenis emosi seseorang melalui cuitan dari media sosial Twitter menggunakan metode *Naïve Bayes* dengan model *Multinomial Naïve Bayes* dan perbandingan fitur ekstraksi, yaitu TF-IDF dan N-gram. Berdasarkan pengujian yang dilakukan, diperoleh nilai akurasi terbaik sebesar 0.6485. Untuk nilai rata-rata *precision* yaitu 0.69, untuk *recall* sebesar 0.64 dan untuk *f1-score* sebesar 0.66.

Kata kunci : twitter, emosi, analisis emosi, naïve bayes, perbandingan fitur

Abstract

Social media is now increasingly used and in demand by many people around the world. Twitter is one of the social media that is actively used by the public. Through Twitter, people can express themselves and share information about their lives, experiences, opinions and emotions. Emotion is said to be a general condition caused by an event experienced by a person. This emotional state can affect important roles in human life such as in the social sphere, in business, as well as in decision making such as the way a person thinks about his own feelings. Therefore, in this study, an analysis of one's emotions was carried out through tweets from social media Twitter uses Naïve Bayes method with the Multinomial Naïve Bayes model and comparison of feature extraction, namely TF-IDF and N-gram. Based on the tests performed, the best accuracy value is 0.6485. For the average precision value is 0.69, for recall is 0.64 and for f1-score is 0.66.

Keywords: twitter, emotion, emotion analysis, naïve bayes, feature comparison

1. Pendahuluan

1.1 Latar Belakang

Media sosial kini semakin marak digunakan dan diminati oleh banyak orang di seluruh dunia. Berdasarkan laporan digital We Are Social, pengguna aktif media sosial tercatat mencapai 3,80 miliar pada bulan Januari 2020. Seiring berjalannya waktu, terjadi peningkatan sebesar 13.2% dalam kurun waktu setahun, sehingga total pengguna aktif media sosial pada bulan Januari 2021 menjadi 4,20 miliar.[1] Semenjak pandemi Covid-19 yang terjadi di Indonesia, banyak kegiatan yang pada awalnya dilaksanakan secara tatap muka, namun harus dilakukan secara daring. Hal ini dapat berdampak pada kondisi emosional seseorang, dikarenakan hal tersebut dapat mempengaruhi cara berpikir seseorang mengenai perasaannya, dan cara mereka bertindak [1]. Masyarakat yang biasanya dapat berkomunikasi dan berkeluh kesah secara langsung, sekarang lebih banyak dilakukan menggunakan media sosial.

Twitter merupakan salah satu media sosial yang aktif digunakan oleh masyarakat. Berdasarkan data dari statista, jumlah pengguna twitter di Indonesia mencapai 15,7 juta sampai dengan bulan Juli 2021.[2] Melalui media sosial twitter, seseorang dapat mengekspresikan opini, pengalaman, ataupun hal lain yang menjadi perhatian mereka. Selain itu, twitter juga dapat mempermudah seseorang untuk saling berbagi informasi tentang kehidupan, pengalaman, maupun aktifitas mereka secara *realtime* [2].

Media sosial twitter sudah menjadi alat komunikasi yang dapat memberikan kebebasan kepada penggunanya untuk mengungkapkan pendapat maupun emosi yang dirasakannya melalui sebuah cuitan [3]. Emosi dapat dikatakan sebagai kondisi secara umum yang disebabkan oleh suatu peristiwa yang dialami seseorang [4].

[1] https://wearesocial.com/digital-2021

[2] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

Di lain hal, keadaan emosional seseorang dapat diartikan sebagai suasana hati, berupa ekspresi dari pikiran, pendapat maupun ide mereka. Hal tersebut dapat mempengaruhi sikap dan perilaku seseorang dalam melakukan sesuatu [1]. Berbeda ketika seseorang berkomunikasi secara langsung, mereka akan lebih mudah untuk mengetahui emosi yang dirasakannya. Namun, ketika seseorang berkomunikasi melalui media sosial twitter, emosi itu tidak dapat tergambarkan dengan jelas. Maka dari itu, perlu dilakukan analisis jenis emosi melalui cuitan dari media sosial twitter untuk dapat memahami dan mengetahui emosi yang terkandung dalam cuitan tersebut [5]. Dikarenakan emosi juga memberikan peranan penting dalam berbagai kehidupan manusia seperti dalam lingkup sosial, bisnis, maupun dalam pengambilan keputusan [6].

Penggunaan *Naïve Bayes* pada penelitian ini dikarenakan *Naïve Bayes* merupakan salah satu teknik klasifikasi yang sederhana dan dapat menghasilkan akurasi cukup tinggi serta tidak memakan waktu yang lama apabila digunakan pada dataset yang besar [7]. Sedangkan untuk *Multinomial Naïve Bayes* digunakan karena model tersebut bekerja dengan baik pada data terdistribusi secara *multinomial* dan dapat digunakan secara luas dalam pengklasifikasian sebuah teks [3].

1.2 Topik dan Batasannya

Topik dari penelitian ini adalah sistem untuk menganalisis klasifikasi jenis emosi seseorang melalui data yang didapatkan dari twitter berupa cuitan. Penelitian ini menggunakan metode *Naïve Bayes* dengan model *Multinomial Naïve Bayes* dengan fitur ekstraksi N-gram dan TF-IDF. Data yang digunakan adalah cuitan berbahasa Indonesia yang berisi 4401 tweet. Kelas emosi tersebut terbagi menjadi lima yaitu *anger*, *fear*, *happy*, *sadness*, dan *love*. Adapun untuk rumusan masalah pada penelitian ini yaitu bagaimana mengimplementasikan fitur N-gram dan TF-IDF dalam menganalisis jenis emosi pada cuitan di media sosial twitter, kemudian bagaimana penggunaan metode *Multinomial Naïve Bayes* dan fitur ekstraksi terhadap nilai akurasi dalam menganalisis jenis emosi pada cuitan di media sosial twitter.

1.3 Tujuan

Tujuan dari penelitian ini adalah untuk menganalisis klasifikasi jenis emosi seseorang melalui data yang didapatkan dari twitter berupa cuitan menggunakan Metode *Multinomial Naïve Bayes* dengan fitur ekstraksi. Kemudian untuk mengimplementasikan metode *Multinomial Naïve Bayes* dengan fitur ekstraksi dalam menganalisis jenis emosi guna mengetahui nilai akurasi yang dihasilkan.

1.4 Organisasi Tulisan

Pada laporan ini, pembahasan dibagi menjadi beberapa bagian. Bagian 1 berisi tentang latar belakang, topik dan batasan, dan tujuan dari dilakukannya penelitian ini. Pada bagian 2, berisi tentang studi terkait yang sebelumnya sudah dilakukan. Pada bagian 3, berisi tentang sistem yang dibangun berupa rancangan maupun produk yang dihasilkan seperti pengumpulan data, preprocessing, ekstraksi fitur, dan proses klasifikasi. Pada bagian 4, berisi tentang hasil dari pengujian yang telah dilakukan dan analisis dari hasil pengujian tersebut. Terakhir pada bagian 5, berisi kesimpulan dan saran dari penelitian yang telah dilakukan.

2. Studi Terkait

Terdapat beberapa penelitian yang telah berhasil menganalisis dan mengklasifikasi emosi seseorang melalui media sosial. Pada penelitian berbahasa Indonesia yang membahas tentang penerapan n-gram untuk sentiment analysis review hotel yang dilakukan oleh Indrayuni, dkk [8] menggunakan metode *Naïve Bayes* menghasilkan akurasi sebesar 82,67%. Kemudian diterapkan fitur N-gram pada tahap preprocessing dan menghasilkan peningkatan nilai akurasi hingga 2% menjadi 84,67%. Dari hal tersebut, dapat disimpulkan bahwa penggunaan N-gram dengan *Naïve Bayes* untuk analisis sentimen review hotel memiliki kinerja yang lebih baik apabila dibandingkan dengan tidak menggunakan N-gram. Selanjutnya, penelitian yang dilakukan oleh Riska Dwi Handayani, dkk [9] melalui facebook untuk membandingkan dua fitur yaitu N-gram dan TF-IDF. Metode *Naïve Bayes* dengan TF-IDF mendapatkan hasil akurasi sebesar 80%, sedangkan untuk *Naïve Bayes* dan N-gram menghasilkan akurasi sebesar 76,4%.

Pada penelitian lain tentang analisa persebaran tweet terhadap isu kenaikan tarif listrik dasar menggunakan metode *Naïve Bayes* dengan fitur N-gram dapat meningkatkan akurasi menjadi 92%, yang sebelumnya menghasilkan akurasi sebesar 89,67% tanpa penggunaan fitur tersebut [10]. Kemudian penelitian yang dilakukan oleh Fera Fanesya, dkk [7] menggunakan metode *Naïve Bayes* dengan kombinasi fitur yaitu fitur linguistik, ortografik, dan N-gram. Fitur N-gram menghasilkan akurasi terbaik sebesar 0,5554. Pada pengujian kombinasi fitur linguistik, ortografik, dan N-gram menghasilkan akurasi sebesar 0,5317. Selain itu, penggunaan metode *Naïve Bayes* memiliki kelebihan yaitu merupakan teknik klasifikasi yang sederhana dan

dapat menghasilkan akurasi cukup tinggi serta tidak memakan waktu yang lama apabila digunakan pada dataset yang besar [7].

2.1 Emosi

Menurut Daniel Goleman (1995), seorang pakar kecerdasan emosional, yang didapatkan dari *Oxford English Dictionary* memaknai emosi sebagai setiap kegiatan atau pergolakan pikiran, perasaan, nafsu, setiap keadaan mental yang hebat dan meluap-luap. Bisa disimpulkan bahwa emosi merujuk kepada suatu perasaan dan pikiran-pikiran yang khas, suatu keadaan biologis dan psikologis, dan serangkaian kecendrungan untuk bertindak [11]. Emosi juga dikatakan sebagai perasaan tertentu yang mencirikan keadaan pikiran seperti bahagia, marah, cinta, dan sebagainya [11].

2.2 Naïve Bayes

Naïve Bayes merupakan salah satu teknik klasifikasi yang banyak digunakan karena pada waktu pemrosesan tidak memakan waktu yang lama apabila menggunakan dataset yang besar. *Naïve Bayes* menggunakan keseluruhan probabilitas, yaitu probabilitas dokumen terhadap kategori (*prior*), kemudian teks akan dikategorikan berdasarkan probabilitas maksimum (*posterior*) (Raschka, 2014) [7].

$$Posterior = \frac{Likelihood \cdot Prior}{Evidence} \quad (1)$$

$$p\,(h|x) = \frac{p\,(x|h).p\,(h)}{p\,(x)} \quad (2)$$

Keterangan:
p(h|x) = nilai *posterior* atau probabilitas kata h dari kelas x
p(x|h) = nilai *likelihood* atau probabilitas kemunculan x untuk kata h p(h)
p(x) = nilai *evidence* atau probabilitas kelas x [9]

2.2.1 Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan salah satu metode *supervised learning*, dan juga model dari pengembangan algoritma bayes [12][13]. *Multinomial Naïve Bayes* juga cocok digunakan untuk mengklasifikasikan sebuah teks atau dokumen. Pada *Multinomial Naive Bayes*, penentuan kelas dokumen ditentukan dengan kata yang muncul, dan juga jumlah kemunculan sebuah kata [13].

$$C_{MAP} = \arg\max_{c\epsilon\{c_l,c_s\}} P(c|d)$$
$$= \arg\max_{c\epsilon\{c_l,c_s\}} P(c)\prod_{k=1}^{m} P(t_k|c)$$

Parameter P(tk|c) (*probability likelihood*) diestimasikan dengan menghitung kejadian tk pada semua dokumen training di c: [13]

$$P(t_k|c) = \frac{1+N_k}{|V|+N},$$

dimana Nk adalah jumlah kemunculan tk dalam dokumen pelatihan di c dan N adalah jumlah total kejadian kata dalam c.

2.3 TF-IDF

TF – IDF merupakan skema pembobotan yang populer. TF IDF merupakan fitur ekstraksi teks dengan cara mencari bobot suatu kata (*term*) pada dokumen (Robertson, 2004). Dalam TF-IDF terdapat dua rumus yang digabungkan, yaitu TF untuk menghitung frekuensi munculnya kata pada dokumen. Sedangkan IDF merupakan perhitungan kuantitas term (istilah) [9]. Berdasarkan penelitian (Fauzi dkk, 2015) dalam mencari nilai TF-IDF memiliki 2 persamaan yaitu:

$$TF\,(d,t) = f(d,t) \quad (3)$$

$$IDF\,(t) = 1 + \log \frac{Nd}{df(t)} \quad (4)$$

Dimana:
t = term (istilah)
TF = jumlah term
f(d,t) = frekuensi munculnya term t pada dokumen d
Nd = total dokumen

Df (t) = jumlah dokumen yang berisi term t
Maka dari itu, nilai TF-IDF dapat ditemukan dengan rumus pada persamaan 5

$$TF-IDF = TF(d,t).IDF(t) \quad (5)$$

2.4 N-gram

Menurut (Gencosman, Ozmutlu, dan Ozmutlu, 2014) N-gram merupakan sub-urutan n karakter dari kata yang telah diberikan [9]. Fitur N-gram dapat digunakan untuk membantu mengurangi analisis morfem pada setiap kalimat. (Bozkir, 2017) [14]. Dapat dikatakan bahwa, fitur N-gram ini merupakan gabungan dari fitur unigram dan bigram [7]. Pada proses N-gram, akan dilakukan pengambilan jumlah n karakter yang digunakan sebagai *term* dan selanjutnya akan dihitung kemunculan kata tersebut dan probabilitasnya. Ekstraksi fitur ini digunakan untuk mengambil fitur pada suatu cuitan ketika belum diklasifikasikan dengan *Naïve Bayes* model *Multinomial Naïve Bayes* yang diharapkan dapat meningkatkan akurasi dalam melakukan klasifikasi data melalui twitter [10].

2.5 Confusion Matrix

Confusion Matrix merupakan metode yang dapat digunakan untuk mengevaluasi hasil dari metode klasifikasi yang dilakukan. *Confusion matrix* berisi informasi aktual (actual) dan prediksi (predicted) pada sistem klasifikasi [15]. Dapat dikatakan bahwa *confusion matrix* adalah sebuah tabel yang memberikan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan [16]. Setiap kolom yang dimiliki matriks tersebut merupakan contoh dari kelas prediksi, untuk tiap baris yang dimiliki merupakan hasil dari kelas sebenarnya [15]. Evaluasi dengan *confusion matrix* menghasilkan nilai *accuracy, precision, recall* dan *f-measure.*

a. Nilai *accuracy*: persentase jumlah record data yang diklasifikasikan secara benar oleh sebuah algoritma dapat membuat klasifikasi setelah dilakukan pengujian pada hasil klasifikasi tersebut (Han & Kamber, 2006) [17]. Berikut merupakan rumus dari persamaan *accuracy* [18]:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

b. Nilai *precision*: Nilai *precision* merupakan proporsi jumlah kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya [17]. Berikut merupakan rumus dari persamaan *precision* [18]:

$$precision = \frac{TP}{TP+FP} \quad (7)$$

c. Nilai *recall*: Nilai *recall* merupakan proporsi jumlah kasus positif yang sebenarnya yang diprediksi positif secara benar (Powers, 2011) [17]. Berikut merupakan rumus dari persamaan *recall* [18]:

$$recall = \frac{TP}{TP+FN} \quad (8)$$

d. Nilai *f-measure*:

$$f-measure = \frac{2*precision*recall}{presicion+recall} \quad (9)$$

Dimana:
- TP merupakan *True Positive*
- FP merupakan *False Positive*
- TN merupakan *True Negative*
- FN merupakan *False Negative*

Berikut merupakan tabel dari *Confusion Matrix*: [17]

Tabel 1. *Confusion Matrix*

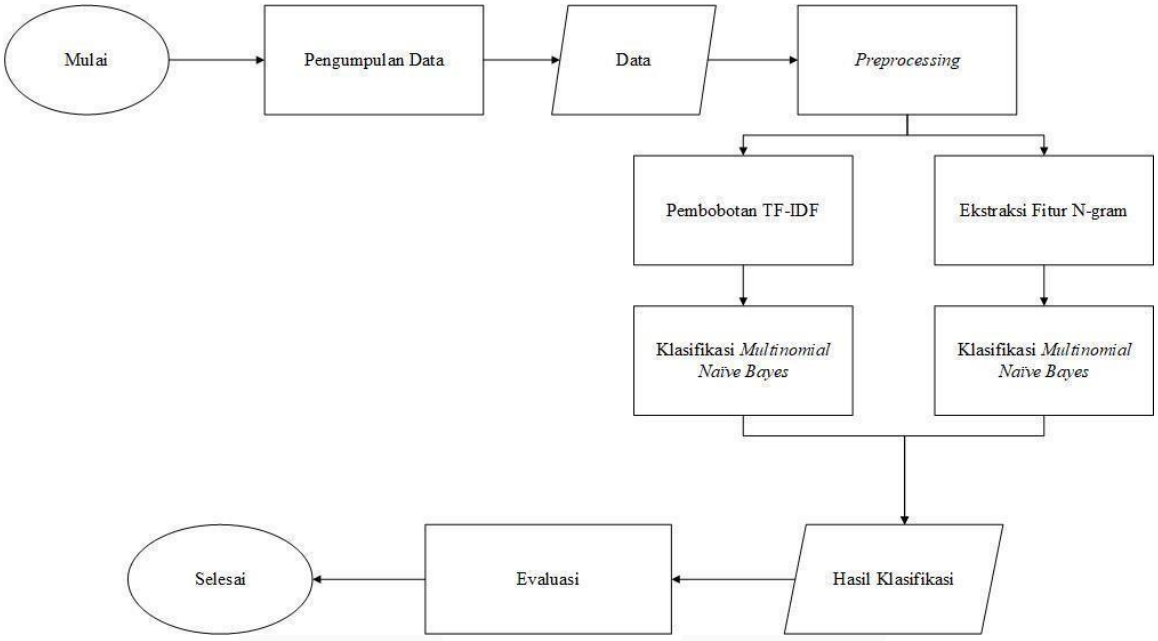
Correct Classification	*Classified as*	
	+	-
+	*True positives*	*False negatives*
-	*False positives*	*True negatives*

Sumber: Han & Kamber 2006

3 Sistem yang Dibangun

3.1 Deksripsi Umum Sistem

Pada penelitian ini, sistem yang dibangun akan menganalisis dan mengklasifikasikan jenis emosi seseorang kedalam beberapa emosi yaitu *anger*, *sadness*, *happy*, *love*, dan *fear* melalui twitter dengan menggunakan Bahasa Indonesia. Kemudian sistem akan melakukan tahap *preprocessing* seperti *case folding*, *convert slangword*, *cleaning*, *tokenization*, *stopword removal*, dan *stemming*. Tahapan tersebut dilakukan untuk membersihkan data cuitan yang telah didapatkan sebelumnya. Selanjutnya, hasil dari *preprocessing* tersebut akan dilakukan fitur ekstraksi pembobotan kata dengan TF-IDF. Selanjutnya akan memasuki tahap klasifikasi. Terakhir yaitu tahap evaluasi. Hasil dari klasifikasi dengan metode *Multinomial Naïve Bayes* akan dievaluasi menggunakan pengujian *confusion matrix*. Berikut merupakan tahapan yang digambarkan:



Gambar 1. Rancangan Sistem

Berdasarkan pada Gambar 1. Berikut merupakan penjelasan tahapan yang dilakukan:

3.2 Pengumpulan Data

Dataset pada penelitian ini diambil melalui penelitian sebelumnya yang dilakukan oleh Mei Silviana Saputri, Rahmad Mahendra, dan Mirna Adriani tentang *Emotion Classification on Indonesian Twitter Dataset* [19]. Dataset yang digunakan pada penelitian ini sebanyak 4401 tweet berbahasa Indonesia. Terdapat lima jenis emosi yang digunakan yaitu *anger* (marah), *fear* (takut), *happy* (bahagia), *sadness* (sedih), *love* (cinta). Berikut merupakan pembagian jumlah data setiap label emosi yang digunakan:

Jumlah data label *anger*: 1101
Jumlah data label *happy*: 1017
Jumlah data label *sadness*: 997
Jumlah data label *fear*: 649
Jumlah data label *love*: 637

Gambar 2. Perbandingan data tiap label

Berikut merupakan contoh cuitan yang terdapat didalam data yang diperoleh:

Tabel 2. Contoh data yang digunakan

Data	Label
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.	*anger*
Kepingin gudeg mbarek Bu hj. Amad Foto dari google, sengaja, biar teman-teman jg membayangkannya. Berbagi itu indah.	*happy*
Pulang udah H-4 lebaran dilema sekali. Seperti tidak bisa melakukan apa2 dirumah sebelum lebaran. Buka puasa bareng cuman 3 hari sama keluarga begitu juga sahur.	*sadness*
Setiap kesempatan yg pernah hadir tuk dapat membuatmu selalu merasa aman dan nyaman, kini jadi suatu kehormatan yg pernah didapat	*love*
Hari ini jadwal presentasi proker di LPPM karna ada praktikum jadinya pak ketua di temani sekretaris 1. Tapi selaku penanggung jawab ada di sekretaris 2 jadi deg deg an ini. Bismillah pak ketua (Btw pen Upload foto pak ketua tapi jangan dulu yaa)	*fear*

3.3 Pelabelan Data

Pada proses ini, data yang diperoleh dengan 5 label yaitu *anger, happy, sadness, love* dan *fear* diubah menjadi label angka. Label *anger* = 0, *happy*=1, *sadness*=2, *love*=3, dan *fear*=4.

3.4 Preprocessing

Pada tahap ini, data yang diperoleh akan dilakukan pemroresesan terlebih dahulu sebelum masuk ke tahap pembobotan kata. Tahap ini dilakukan agar data yang digunakan bersih dan hanya mengandung data yang penting dan dapat disesuaikan dengan kebutuhan untuk klasifikasi emosi dengan *Multinomial Naïve Bayes*. Adapun tahapan *preprocessing* yang dilakukan pada penelitian ini yaitu:

a. *Case Folding*

Pada tahap ini akan dilakukan proses pengubahan semua text menjadi huruf kecil. Kemudian semua karakter yang bukan merupakan huruf, seperti angka dan tanda baca akan dihilangkan. Tahap ini dilakukan agar mempermudah dalam proses pencarian [10].

b. *Cleaning*

Pada tahap ini bertujuan untuk mengurangi dan membersihkan data cuitan seperti tanda baca, *mention*, *unicode*, dan *url*.

c. *Tokenization*

Setelah proses sebelumnya selesai, data tersebut memerlukan proses tokenisasi. Tokenisasi adalah proses untuk memecahkan sebuah kalimat menjadi kata per kata. Hasil dari proses pemecahan kata ini disebut dengan token.

d. *Convert Slangword*

Slang word atau yang dikenal dengan "kata gaul" merupakan bahasa yang saat ini sering digunakan untuk berkomunikasi. Proses ini dilakukan untuk menghilangkan kata gaul tersebut dan mengubah kata yang tidak baku menjadi baku.

e. *Stopword Removal*

Pada tahap ini dilakukan penghapusan kata yang tidak relevan atau tidak berhubungan dengan semua kalimat seperti kata "dan", "pada", "pula", "saat","bahwa" ,"lah" dan kata-kata lainnya [10]. Penghapusan *stopword* dilakukan dengan bantuan *libray* nltk. Untuk kata "tidak" dalam list kata yang ada di *stopword* akan dihapus, dikarenakan kata tersebut dapat mempengaruhi sebuah kalimat.

f. *Stemming*

Stemming merupakan proses untuk mencari kata dasar dengan cara menghapus imbuhan. Pada tahap ini setiap kata akan dibagi kedalam beberapa kelompok [10].

Berikut merupakan contoh dari hasil *preprocessing* yaitu:

Tabel 3. Hasil *Preprocess* Data

Tahap	Hasil
Data Asli	Soal jln Jatibaru, polisi tdk bs GERTAK gubernur, Emangny polisi tdk ikut pmbhasan? Jgn berpolitik. Pengaturan wilayah, hak gubernur. Persoalan Tn Abang soal turun temurun. Pelik. Perlu kesabaran. [USERNAME] [USERNAME] [URL]
Case Folding	soal jln jatibaru, polisi tdk bs gertak gubernur. emangny polisi tdk ikut pmbhasan? jgn berpolitik. pengaturan wilayah, hak gubernur. persoalan tn abang soal turun temurun. pelik. perlu kesabaran. [username] [username] [url]
Slangword	soal jalan jatibaru, polisi tidak bisa gertak gubernur. emangny polisi tidak ikut pembahasan? jangan berpolitik. pengaturan wilayah, hak gubernur. persoalan tanah abang soal turun temurun. pelik. perlu kesabaran. [username] [username] [url]
Cleaning	soal jalan jatibaru polisi tidak bisa gertak gubernur emangny polisi tidak ikut pembahasan jangan berpolitik pengaturan wilayah hak gubernur persoalan tanah abang soal turun temurun pelik perlu kesabaran username username url
Tokenization	'soal', 'jalan', 'jatibaru', 'polisi', 'tidak', 'bisa', 'gertak', 'gubernur', 'emangny', 'polisi', 'tidak', 'ikut','pembahasan','jangan','berpolitik','pengaturan', 'wilayah', 'hak', 'gubernur', 'persoalan', 'tanah', 'abang', 'soal','turun','temurun', 'pelik', 'perlu', 'kesabaran', 'username','username','url'
Stopword Removal	'jalan', 'jatibaru', 'polisi', 'gertak', 'gubernur', 'polisi', 'pembahasan', 'berpolitik', 'pengaturan', 'wilayah', 'hak', 'gubernur', 'tanah', 'abang', 'turun', 'temurun', 'pelik', 'kesabaran'
Stemming	'jalan', 'jatibaru', 'polisi', 'gertak', 'gubernur', 'polisi', 'bahas', 'politik', 'atur', 'wilayah', 'hak', 'gubernur', 'tanah', 'abang', 'turun', 'turun', 'pelik', 'sabar'

3.5 Pembobotan TF-IDF

Pada tahap ini dilakukan pembobotan terhadap kata yang sudah diperoleh dari hasil preprocessing sebelumnya. Pembobotan tiap kata menggunakan TF-IDF (*Term Frequency-Index Term Frequency*). Berikut merupakan contoh hasil pembobotan kata yaitu kata "jalan" memiliki bobot =0.131, "jatibaru" memiliki bobot=0.264, "polisi" memiliki bobot=0.399, "gertak" memiliki bobot=0.277, "gubernur" memiliki bobot=0.388, "bahas" memiliki bobot=0.201, "politik" memiliki bobot=0.179, "atur" memiliki bobot=0.194, "wilayah" memiliki bobot=0.237, "hak" memiliki bobot=0.196, "gubernur" memiliki bobot=0.388, "tanah" memiliki bobot=0.199, "abang" memiliki bobot=0.195, "turun" memiliki bobot=0.367, "pelik" memiliki bobot=0.277, "sabar" memiliki bobot=0.170

3.6 N-gram

Pada tahap ini, dilakukan proses N-gram yaitu proses pembagian kalimat sebesar nilai N. Seperti contoh N = 1 (*unigrams*) merupakan pembagian kalimat menjadi satu kata, N = 2 (*bigrams*), merupakan pembagian kalimat menjadi dua kata dan N = 3 (*trigrams*) merupakan pembagian kalimat menjadi tiga kata. Berikut merupakan penerapan N-gram pada kata "teman".

Tabel 4. Contoh penerapan karakter N-gram

Jumlah n	Contoh penerapan karakter N-gram
Unigram(n=1)	t-e-m-a-n
Bigram(n=2)	Te-em-ma-an
Trigram(n=3)	Tem-ema-man

3.7 Klasifikasi

Pada tahap ini akan dilakukan proses klasifikasi dengan menggunakan metode *Multinomial Naïve Bayes*. Pada penelitian ini menggunakan model klasifikasi *Multinomial Naïve Bayes*, dikarenakan data yang dimiliki merupakan data yang memiliki banyak label/kelas. Klasifikasi digunakan untuk mendapatkan kelas pengenalan yang sesuai. Fitur yang digunakan pada klasifikasi ini yaitu TF-IDF dan N-gram. Data yang telah diperoleh akan dibagi menjadi *data train* (data latih), dan *data test* (data uji).

3.8 Evaluasi Model

Setelah proses klasifikasi selesai, pada tahap terakhir dilakukan evaluasi. Hasil dari klasifikasi tersebut akan dievaluasi terhadap sistem dengan menggunakan pengujian *confusion matrix*. Evaluasi ini bertujuan untuk menghitung nilai performansi dari hasil yang telah diperoleh berupa *precision*, *recall*, *accuracy*, dan *f1-score*. Tujuan dari penggunaan nilai akurasi dan *f1-score* dalam evaluasi, dikarenakan nilai akurasi digunakan untuk menunjukkan seberapa baik model yang telah dibangun dalam melakukan klasifikasi dengan benar. Sedangkan *f1-score* digunakan untuk menunjukkan seberapa akurat hasil prediksi per tiap kelas nya dalam sistem klasifikasi.

Tabel 5. *Confusion Matrix*

Class Emotion	0	1	2	3	4
0	TP	x	x	x	x
1	x	TP	x	x	x
2	x	x	TP	x	x
3	x	x	x	TP	x
4	x	x	x	x	TP

4 Evaluasi

Pada bagian ini dijelaskan mengenai hasil yang diperoleh dari sistem yang telah dibangun dan diimplementasikan. Terdapat beberapa percobaan yang telah dilakukan, meliputi:

1. Pengujian dengan melakukan pembagian *data test* dan *data train* yang berbeda beda antara fitur N-gram dan TF-IDF

2. Pengujian dengan membandingkan hasil performansi dari pengujian 1, dan menambahkan nilai parameter random_state = 42

4.1 Hasil Pengujian 1

Pada pengujian 1, dilakukan perhitungan akurasi menggunakan metode *Multinomial Naïve Bayes* dengan perbandingan fitur esktraksi. Hasil dari perhitungan antara fitur ekstraksi yang digunakan yaitu TF-IDF dan N-gram, selanjutnya akan dilakukan perbandingan. Pada skenario pertama, dilakukan pembagian data uji yaitu *data test* dan *data train* sebesar 10:90, 20:80, 30:70, 40:60, dan 50:50, guna mengetahui nilai akurasi tertinggi dari perbandingan *data test* dan *data train* yang dilakukan. Berikut nilai akurasi yang dihasilkan:

Tabel 6. Hasil Perbandingan Nilai Akurasi Klasifikasi

Perbandingan data test dan *data train*	Fitur Ekstraksi	
	TF-IDF	N-gram
10:90	0.5600 0.5079 0.5963 0.5623 **0.6031**	0.6009 **0.6326** 0.6303 0.6145 0.5759
20:80,	0.5709 0.5607 0.5698 0.5471 **0.5981**	0.5947 0.5993 0.5981 0.5822 **0.6208**
30:70	0.5851 **0.5927** 0.5594 0.5616 0.5691	0.5783 **0.6010** 0.5942 0.5942 0.6003
40:60	0.5775 **0.5792** 0.5496 0.5752 0.5553	0.5854 0.5689 0.5797 **0.5922** 0.5769
50:50	0.5515 0.5674 **0.5738** 0.5374 0.5583	0.5651 0.5574 0.5592 **0.5683** 0.5615

Pada pengujian pertama untuk fitur TF-IDF dan N-gram, dilakukan lima kali percobaan dan menghasilkan nilai akurasi tertinggi pada perbandingan *data test* dan *data train* 10:90. Pada fitur TF-IDF menghasilkan nilai akurasi sebesar 0.6031 sedangkan untuk fitur N-gram diperoleh nilai akurasi tertinggi dengan nilai 0.6326. Maka dari itu, dilakukan skenario pengujian 2 untuk meningkatkan nilai akurasi dari hasil perhitungan sebelumnya dengan menambahkan parameter nilai random_state = 42.

4.2 Hasil Pengujian 2

Pada skenario kedua, dilakukan perbandingan terhadap hasil pengujian sebelumnya. Berikut nilai akurasi yang dihasilkan setelah dilakukan perubahan nilai random_state = 42

Tabel 7. Hasil Nilai Akurasi setelah diberi random_state = 42

Perbandingan *data test* dan *data train*	Fitur Ekstraksi	
	TF-IDF	N-gram
10:90	0.6054	**0.6485**
20:80	**0.6106**	0.6276
30:70	0.6033	0.6116
40:60	0.5848	0.5826
50:50	0.5592	0.5656

Pada skenario pertama, nilai akurasi tertinggi yang dihasilkan untuk fitur TF-IDF, diperoleh pada perbandingan *data test* dan *data train* 10:90 dengan hasil 0.6031. Kemudian pada skenario kedua mengalami peningkatan sebesar 0,0075 pada perbandingan *data test* dan *data train* 20:80, sehingga nilai akurasi tertinggi menjadi 0.6106. Sedangkan untuk fitur N-gram juga mengalami hal yang sama dengan fitur TF-IDF, yaitu peningkatan pada nilai akurasi yang dihasilkan. Hasil pada pengujian pertama diperoleh nilai sebesar 0.6326, kemudian meningkat sebesar 0.0159, sehingga hasil yang diperoleh setelah diterapkan nilai random_state = 42 yaitu 0.6485 pada perbandingan data yang sama dengan skenario pertama yaitu 10:90. Maka dari itu dapat dibuktikan bahwa penggunaan parameter nilai random_state = 42 mempengaruhi nilai akurasi yang dihasilkan, yaitu dapat meningkatkan nilai akurasi.

Tahap selanjutnya yaitu dilakukan proses evaluasi dengan menggunakan *confusion matrix*. Dengan *confusion matrix*, akan diperoleh nilai *precision*, *recall*, dan *f1-score*. Berikut hasil perhitungan nilai *confusion matrix* pada perbandingan *data test* dan *data train* 10:90 pada fitur N-gram dengan random_state = 42

Tabel 8 Hasil Confusion Matrix 10:90

Label	0	1	2	3	4
0	**86**	6	23	2	2
1	16	**76**	14	4	1
2	24	9	**51**	8	2
3	1	7	11	**36**	0
4	11	5	9	0	**37**

Dapat dilihat pada tabel, model klasifikasi *Multinomial Naïve Bayes* dapat mengklasifikasikan kelas *anger*, *happy*, *sadness*, *love*, dan *fear*. Penggunaan *confusion matrix* ini, dapat digunakan sebagai evaluasi dari klasifikasi yang dilakukan. Berikut hasil perhitungan *confusion matrix* pada perbandingan *data test* dan *data train* 10:90 pada fitur N-gram.

Tabel 9. Hasil Precision, Recall, dan F1-Score

label	Precision	Recall	F1-score
0	0.62	**0.72**	0.67
1	0.74	0.68	**0.71**
2	0.47	0.54	0.50
3	0.72	0.65	0.69
4	**0.88**	0.60	**0.71**
Rata-rata	0.69	0.64	0.66

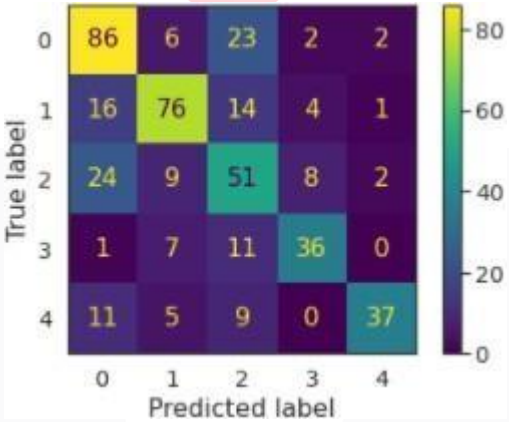
Berdasarkan tabel *confusion matrix* yang diperoleh diatas, bahwa nilai *precision* tertinggi diperoleh label 4 dengan nilai 0.88. Untuk nilai *recall* tertinggi diperoleh pada label 0 dengan nilai 0.72, sedangkan untuk nilai *f1-score* tertinggi diperoleh pada label 1 dan 4 dengan nilai 0.71.

4.3 Analisis Semua Hasil Pengujian

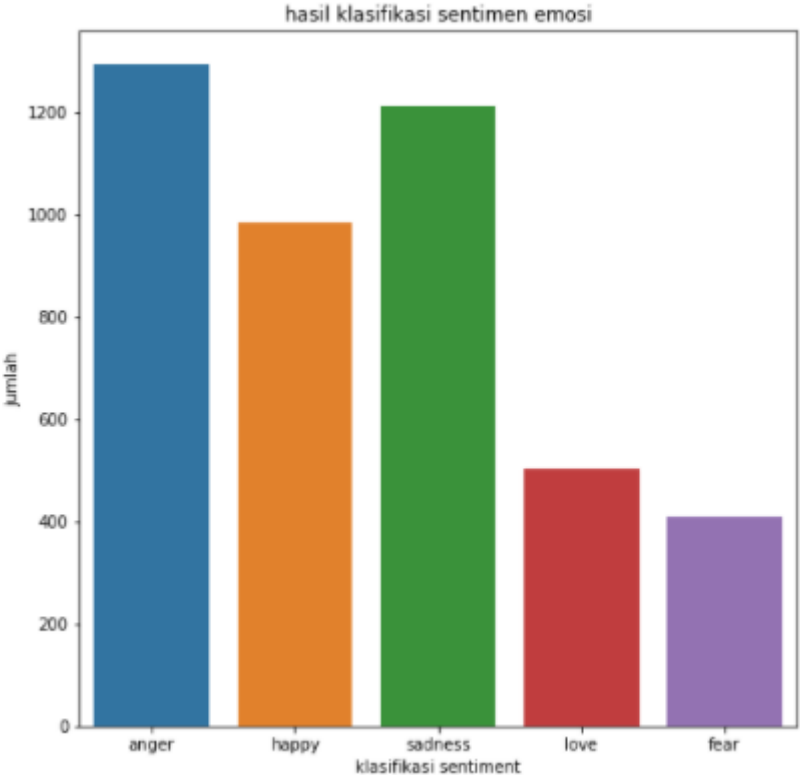
Pengujian pertama dilakukan dengan membagi *data test* dan *data train* yang berbeda beda antara fitur N-gram dan TF-IDF. Pada pengujian pertama yang dilakukan berhasil memperoleh nilai akurasi tertinggi pada perbandingan *data* test dan *data train* 10:90. Untuk fitur TF-IDF sebesar 0.6031 sedangkan untuk fitur N-gram sebesar 0.6326. Kemudian dilakukan pengujian kedua untuk meningkatkan hasil akurasi yang diperoleh dengan menggunakan parameter nilai random_state = 42. Selain itu, penggunaan nilai random_state = 42 dilakukan agar hasil dari setiap pengujian menghasilkan nilai yang tetap. Setelah diterapkan nilai random_state = 42, terjadi peningkatan nilai akurasi pada kedua fitur. Untuk fitur TF-IDF menghasilkan nilai akurasi 0.6106 pada perbandingan *data test* dan *data train* 20:80, sedangkan untuk fitur N-gram menghasilkan nilai 0.6485 pada perbandingan *data test* dan *data train* yang sama seperti pengujian pertama yaitu 10:90.

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa nilai akurasi terbaik didapatkan pada fitur N-gram dengan perbandingan *data test* dan data *train* 10:90 dengan menggunakan nilai random_state = 42. Nilai yang dihasilkan dari perbandingan data uji tersebut yaitu sebesar 0.6485. Sedangkan pada saat sebelum menggunakan nilai random_state=42, nilai akurasi tertinggi pada fitur N-gram didapatkan dengan nilai 0.6326. Pengujian tersebut membuktikan bahwa penggunaan parameter nilai random_state = 42 mempengaruhi nilai akurasi yang dihasilkan, yaitu dapat meningkatkan nilai akurasi.

Setelah didapatkan model dengan nilai akurasi terbaik, dilakukan proses perhitungan nilai *precision*, *recall*, dan *f1-score* menggunakan *confusion matrix*. Berikut merupakan hasil dari *confusion matrix* dari fitur N-gram dengan perbandingan *data tes*t dan *data train* 10:90.



Setelah semua pengujian dilakukan, didapatkan hasil klasifikasi dari 5 jenis emosi. Untuk label *anger* sebanyak 1322, label *fear* sebanyak 316, label *happy* sebanyak 813, label *love* sebanyak 417 dan label *sadness* sebanyak 1533. Berikut hasil diagram yang dihasilkan:



5 Kesimpulan

Pada penelitian ini dilakukan analisis emosi pada media sosial twitter menggunakan metode *Naïve Bayes* dengan model Multinomial *Naïve Bayes* dengan perbandingan Fitur N-gram dan TF-IDF. Untuk kelas emosi yang digunakan yaitu *anger, happy, sadness, love,* dan *fear*. Berdasarkan hasil penelitian yang telah dilakukan, diperoleh nilai akurasi terbaik sebesar 0.6485. Untuk nilai rata-rata *precision* yaitu 0.69, untuk *recall*sebesar 0.64 dan untuk *f1-score* sebesar 0.66. Hasil tersebut diperoleh dengan melakukan tahapan sesuai dengan yang ada dalam penelitian ini, dengan menerapkan metode *Multinomial Naïve Bayes* dengan fitur ekstraksi TF-IDF dan N-gram serta penggunaan nilai random_state = 42. Penggunaan nilai random_state = 42 digunakan agar hasil dari setiap pengujian menghasilkan nilai yang tetap, selain itu juga dapat meningkatkan nilai akurasi. Untuk hasil yang lebih maksimal, penulis menyarankan agar menggunakan data yang lebih banyak dan juga seimbang untuk setiap kelasnya. Hal ini dilakukan agar mendapatkan nilai akurasi yang lebih baik. Selain itu, ketika melakukan proses *preprocessing* dapat menggunakan kamus bahasa yang lebih lengkap agar kata-kata yang diproses menjadi lebih baik.

REFERENSI

[1] Badugu, Srinivasu, dan Matla Suhasini. 2017. Emotion detection on twitter data using knowledge base approach. International Journal of Computer Applications, 162(10).

[2] Hadna, N. M. S., Paulus Insap Santosa, dan Wing Wahyu Winarno. 2016. Studi literatur tentang perbandingan metode untuk proses analisis sentimen di Twitter. Seminar Nasional Teknoogi Informasi dan Komunikasi

[3] Deshpande, M., dan Vignesh Rao. 2017. Depression detection using emotion artificial intelligence. International Conference on Intelligent Sustainable.

[4] Arifin, dan Ketut Eddy Purnama. 2012. Classification of emotions in indonesian texts using K-NN method. International Journal of Information and Electronics Engineering, 2(6), 899-903.

[5] L, Adinda Dwi, Budi Harijanto, dan Faisal Rahutomo. 2020. Implementasi Deep Learning Untuk Deteksi Ekspresi Emosi Pada Twitter. Seminar Informatika Aplikatif Polinema (SIAP).

[6] Rohman, Arif Nur, Ema Utami, dan Suwanto Raharjo. 2019. Deteksi Kondisi Emosi pada Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing. Jurnal Eksplora Informatika, 9(1), 70-76.

[7] Fanesya, Fera. 2019. Deteksi Emosi pada Twitter Menggunakan Metode Naïve Bayes dan Kombinasi Fitur. Doctoral dissertation, Universitas Brawijaya.

[8] Indrayuni, Elly, dan Mochamad Wahyudi. 2015. PENERAPAN CHARACTER N-GRAM UNTUK SENTIMENT ANALYSIS REVIEW HOTEL MENGGUNAKAN ALGORITMA NAÏVE BAYES

[9] Handayani, R. D., Kusrini, K., dan Hanif Al Fatta. 2020. Perbandingan Fitur Ekstraksi Untuk Klasifikasi Emosi Pada Sosial Media. Jurnal Ilmiah SINUS, 18(2), 21-30.

[10] Nugroho, Agung. 2018. Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstrasi Fitur N-Gram. Jurnal Sains Komputer dan Informatika (J-SAKTI), 2(2), 200-209.

[11] Simanungkalit, Chrishartanto, dan Irfan Setia Zega. 2019. HUBUNGAN SOSIALISASI VISI MISI EFEKTIF TERHADAP TINGKAT IKATAN EMOSIONAL STAF DENGAN INSTITUSI DI PUSKESMAS PERAWATAN PLUS AWAAI KEC. SITOLU ORIKAB. NIAS UTARA. Indonesian Trust Health Journal, 2(2), 226-236.

[12] Yulianto, A., Anisa Herdiani., dan Indra Lukmaha Sardi. 2019. Klasifikasi Keberpihakan Tweet Menggunakan Multinomial Naïve Bayes (studi Kasus: Pemilihan Presiden 2019). eProceedings of Engineering, 6(2).

[13] Kalokasari, D. H., I. Shofi, dan A. H. Setyaningrum. 2017. Implementasi Algoritma Multinomial Naive Bayes Classifier Pada Sistem Klasifikasi Surat Keluar (Studi Kasus: Diskominfo Kabupaten Tangerang). Jurnal Teknik Informatika.

[14] Bozkir, A. Selman, Esra Sahin, Murat Aydos, Ebru Akcapinar Sezer, dan Fatih Orhan. 2017. Spam E-Mail Classification by Utilizing N-Gram Features of Hyperlink Texts. 2017 IEEE 11th International Conference on Application of Information and Communication Technologies (AICT) (pp. 1-5). IEEE.

[15] Setiawan, A. B. 2018. Kajian Komparasi Penerapan Algoritma C4. 5 Dan Naive Bayes Sebagai Penunjang Keputusan Pinjaman Uang (Studi Kasus Di Koperasi Karyawan PT. Karyamitra Budisentosa Pandaan Pasuruan) (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).

[16] Indriani, Aida. 2014. Klasifikasi Data Forum dengan menggunakan Metode Naive Bayes Classifier. Seminar Nasional Aplikasi Teknologi Informasi (SNATI) (Vol. 1, No. 1).

[17] Andriani, Anik. 2013. Sistem Pendukung Keputusan Berbasis Decision Tree Dalam Pemberian Beasiswa Studi Kasus: Amik "Bsi Yogyakarta.". Seminar Nasional Teknologi Informasi Dan Komunikasi (Vol. 2013, pp. 163-168).

[18] Muljono, Nurul Anisa Sri Winarsih, dan Catur Supriyanto. 2016. Evaluation of Classification Methods for Indonesian Text Emotion Detection. 2016 International seminar on application for technology of information and communication (ISemantic) (pp. 130-133). IEEE.

[19] Saputri, M.S., R. Mahendra, dan M. Adriani, 2018. Emotion classification on indonesian twitter dataset. In 2018 International Conference on Asian Language Processing (IALP) (pp. 90-95). IEEE.