JURNAL FASILKOM
Volume 13 No. 2 | Agustus 2023: 334-339

P-ISSN : 2089-3353
E-ISSN : 2808-9162

Analisis Opini Publik Terhadap Undang-Undang KUHP Tahun 2022 Menggunakan Algoritma *Naïve Bayes Classifier*

Febby Apri Wenando[1]
[1]Sistem Informasi, Fakultas Teknologi Informasi, Universitas Andalas
[1]febby.apri@it.unand.ac.id*

Abstract

Sentiment analysis, also called opinion mining, involves an automated process of understanding, extracting, and processing textual data to obtain sentiment information expressed in a person's opinion about a subject or object, usually taking a negative or positive attitude. This research attempts to categorize tweet data into positive and negative sentiments. Using Indonesian language text from the social media platform Twitter, this research utilizes public opinion in these tweets to analyze public sentiment to determine public perceptions regarding the RUU KUHP that has just been passed. The data collection was taken from social media Twitter, totaling 142 tweets data used in this research. This tweet data classification uses the Naïve Bayes Classifier algorithm. Before the analysis, an initial stage is carried out to prepare the data, called the pre-processing step. This stage is carried out to clean the text, including processes such as case folding, tokenization, normalization, and stopword removal. The results of the 142 test data that were classified resulted in 62 positive sentiment data and as many as 80 negative sentiment data. After the evaluation, the performance results of the Naïve Bayes Classifier algorithm were obtained with an accuracy value of 87%.

Keywords: sentiment analysis, undang-undang KUHP, naïve bayes

Abstrak

Analisis sentimen, juga disebut penambangan opini, melibatkan proses otomatis dalam memahami, mengekstraksi, dan memproses data tekstual untuk mendapatkan informasi sentimen yang diungkapkan dalam opini seseorang tentang suatu subjek atau objek, biasanya mengambil sikap negatif atau positif. Penelitian ini berupaya untuk mengkategorikan data tweet menjadi sentimen positif dan negatif. Dengan menggunakan teks berbahasa Indonesia dari platform media sosial Twitter, penelitian ini memanfaatkan opini masyarakat dalam *tweet* tersebut untuk analisis sentimen masyarakat untuk mengetahui persepsi masyarakat terkait Rancangan Undang-Undang KUHP yang baru saja disahkan. Kumpulan data yang digunakan diambil dari social media Twitter, sebanyak 142 data *tweet* yang gunakan pada penelitian ini. Klasifikasi data *tweet* ini menggunakan algoritma *Naïve Bayes Classifier*. Sebelum dilakukan analisis dilakukan tahapan awal untuk mempersiapkan data, disebut dengan tahapan *pre-processing*, tahapan ini dilakukan untuk membersihkan teks, meliputi proses seperti *case folding*, tokenisasi, normalisasi, dan *stopword removal*. Hasil dari 142 data uji yang klasifikasi menghasilkan 62 data bersentimen positif dan sebanyak 80 data sentimen negatif. Setelah dilakukan evaluasi didapat hasil performa algoritma *Naïve Bayes Classifier* dengan nilai akurasi sebesar 87%.

Kata kunci: sentiment analysis, undang-undang KUHP, *naïve bayes*

1. Pendahuluan

Sekretaris Jenderal DPR RI, Indra Iskandar menyatakan bahwa RKUHP disahkan jadi KUHP pada tanggal 6 Desember 2022. Pengesahan ini sesuai dengan keputusan badan musyawarah DPR RI. Disahkannya Rancangan Kitab Undang-Undang Hukum Pidana (RKUHP) menjadi KUHP ini, menuai polemik panjang di masyarakat karena banyak pasal-pasal karet dan kontroversial dalam KUHP tersebut, jika diterapkan di Indonesia[1].

RKUHP banyak mendapat kritik dari berbagai pihak. Bahkan, Koalisi Nasional Reformasi KUHP telah dibentuk dan menolak pengesahan RUU KUHP. Menurut serikat pekerja, proses pembuatan RKUHP dipandang sebagai bentuk pembatasan partisipasi masyarakat dalam penyusunan peraturan perundang-undangan yang sangat penting.

Pasal-pasal yang kontroversial dalam RUU KUHP 2022 yaitu RKUHP Dinilai Mengancam Kebebasan Pers, Pasal RKUHP Tentang Akses Alat Pencegahan Kehamilan, Pasal RKUHP Tentang Penghinaan Hakim dan masih banyak lagi[2].

Penggunaan jejaring sosial telah menyebar dan sangat cepat menjangkau seluruh lapisan masyarakat. Tidak hanya sebagai sarana bersosialisasi dan berkomunikasi, namun juga menyampaikan aspirasi, mewakili apa yang terjadi dan dirasakan masyarakat. [3]. Di jejaring sosial Twitter, banyak pihak yang mengutarakan pandangannya terhadap pengesahan KUHP pada tahun 2022. Hal ini dapat dijadikan acuan untuk mengetahui pendapat masyarakat terhadap berlakunya KUHP 2022.

Dalam penelitian ini, kita akan membahas langkah-langkah yang diambil untuk menyelesaikan proses analisis sentimen. Dari tahap pra-pemrosesan,

P-ISSN : 2089-3353
E-ISSN : 2808-9162

kemudian lanjut ke Langkah analisis sentimen untuk mengukur kualitas hasil analisis menggunakan beberapa parameter seperti akurasi, presisi, dan recall. Terdapat penelitian terdahulu yang dapat dijadikan referensi. Penulis menyebutkan jurnal penelitian pertama yang Mengimplementasikan Analisis Sentimen Twitter mengenai Opini Masyarakat Terhadap RKUHP tahun 2019 dapat disimpulkan bahwa topik ini tetap perlu untuk dibahas dan dilakukan penelitian lebih lanjut terkait bagaimana pendapat masyrakat terhadap aturan dan kebijakan yang ditetapkan oleh peerintah, sehingga bisa diketahui bagaimana sentiment masyarakat umum terhadap aturan maupun kebijakan tersebut [4][5][6][7].

Menurut Liu (2008), sentiment analysis (analisis sentimen) atau sering disebut juga dengan opinion mining (penambangan opini) adalah studi komputasi untuk mengenali dan mengekspresikan opini, sentimen, evaluasi, sikap, emosi, subjektifitas, penilaian atau pandangan yang terdapat dalam suatu teks[5][6] [7].

Text mining memiliki definisi menambang data yang dalam bentuk tekstual yang sumber datanya biasanya berasal dari data, tujuannya adalah mencari kata-kata yang dapat mewakili isi data sehingga dapat dilakukan analisis hubungan antar data. [8].

Tujuan dari Text Mining adalah untuk memperoleh informasi berguna dari sekumpulan data. Oleh karena itu, sumber data yang digunakan dalam Text Mining adalah kumpulan teks yang berformat tidak beraturan. terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari text mining antara lain yaitu pengkategorisasian teks (*text categorization*) dan pengelompokan teks (*text clustering*). Text mining merupakan penerapan konsep dan teknik data mining Mencari pola dalam teks, khususnya proses menganalisis teks guna menemukan informasi yang berguna untuk tujuan tertentu. Berdasarkan ketidakteraturan struktur data teks, proses text mining memerlukan beberapa langkah awal untuk mempersiapkan dasar agar teks dapat dimodifikasi agar memiliki struktur yang lebih koheren.

Twitter adalah jaringan sosial yang paling cepat berkembang saat ini karena memungkinkan pengguna untuk berinteraksi dengan orang lain kapan saja, di mana saja dari komputer atau perangkat seluler mereka (Handoko and Neneng 2021) [9]. Sejak diluncurkan pada Juli 2006, jumlah pengguna Twitter telah berkembang pesat. Pada September 2010, Twitter diperkirakan memiliki sekitar 160 juta pengguna terdaftar(Alita, Fernando, and Sulistiani 2020). Pengguna Twitter sendiri mungkin terdiri dari berbagai jenis lingkaran yang memungkinkan pengguna berinteraksi dengan teman, keluarga, dan kolega mereka (Fadly and Wantoro 2019).

Sebagai situs jejaring sosial, Twitter menyediakan akses bagi penggunanya untuk mengirim pesan singkat (disebut Tweets) hingga 140 karakter (Darwis, Siskawati, and Abidin 2021). Tweet itu sendiri dapat terdiri dari pesan teks dan foto(Putra, Darwis, and Priandika 2021). Tweet ini memungkinkan pengguna Twitter untuk berinteraksi lebih dekat dengan pengguna Twitter lainnya dengan memposting tentang pemikiran(Widodo and Ahmad 2017), tindakan, peristiwa terkini, berita terkini, dan banyak lagi (Rusliyawati, Putri, and Darwis 2021).

Metode Term Frequency Inverse Document Frequency (TF-IDF) ialah teknik penentuan seberapa term mewakili konten dalam dokumen dengan memberi bobot ke masing-masing kata yang terkandung di dalamnya [10]. Nilai TF-IDF didapat dari perkalian TF dan IDF.

Nilai TF didapat dari perhitungan Persamaan 1. Nilai IDF didapat dari perhitungan Persamaan berikut :

$$tf(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \quad (1)$$

$$\mathrm{idf}(t,D) = \log \frac{N}{|\{d \in D : t \in d\}|} \quad (2)$$

Dengan *tf*(*t*,*d*) merupakan nilai term frequency *t* di dokumen *d*, N*t*,*d* merupakan jumlah munculnya term *t* di dokumen *d*, N*d* merupakan total term yang terdapat pada dokumen *d*, *idf t* merupakan nilai IDF dari term *t*, *n* merupakan jumlah koleksi dokumen, dan *n-k* merupakan banyaknya dokumen yang memuat *term t*.

Pengklasifikasi Naive Bayes adalah algoritma pembelajaran mesin yang termasuk kedalam supervised classification [11]. Pengklasifikasi Naive Bayes berasal Teorema Bayes, yang mengasumsikan bahwa data tidak berkorelasi secara statistik, seperti yang ditunjukkan dalam persamaan. (3).

Naive Bayes banyak digunakan untuk klasifikasi teks dimana salah satu penggunaannya adalah menyaring spam Email, Analisis Sentimen, Sistem Rekomendasi, dan lain-lain [12]. Terlepas dari kesederhanaannya, pengklasifikasi Naive Bayesian biasanya bekerja dengan baik dan banyak digunakan karena sering mengungguli metode klasifikasi yang lebih canggih [13] .

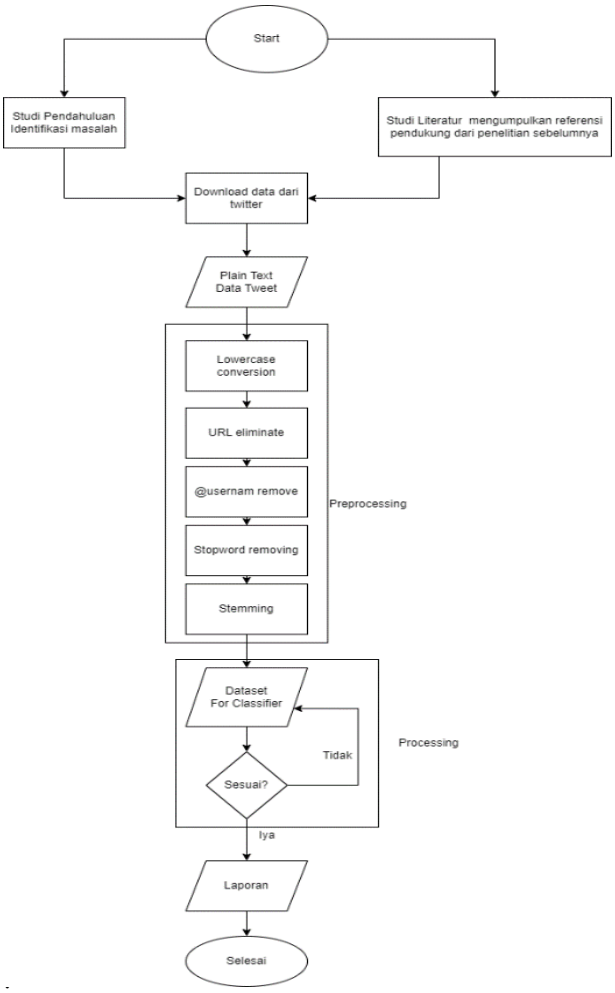
$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad (3)$$

dimana,
P(A|B) : probabilitas hipotesis A untuk data B
P(B|A) : probabilitas data B benar jika hipotesis A benar
P(A) : probabilitas bahwa hipotesis A benar (terlepas dari datanya)
P(B) : data probabilistik (terlepas dari hipotesisnya)

P-ISSN : 2089-3353
E-ISSN : 2808-9162

2. Metode Penelitian

Tahapan pada penelitian ini dijelaskan pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Ada berbagai cara untuk melakukan pengumpulan data melalui media sosial Twitter, tentu saja dengan adanya kelebihan dan kekurangannya masing-masing [14]. Dari twitter API[15], tweepy, TWINT, dan snscrape [16], penulis memilih snscrape yang berbasis bahasa python untuk melakukan crawling data cuitan twitter. Dalam proses ini (dan proses selanjutnya), data disimpan dalam file yang diformat csv. Penulis membuat word cloud untuk menemukan kata-kata yang sering muncul di tweet.

2.2. Tahap Preprocessing

Penelitian ini memerlukan preprocessing data yang meliputi pengubahan suatu dokumen yang tidak terstruktur menjadi data terstruktur untuk kebutuhan analisis lebih lanjut agar tidak menurunkan kinerjanya sendiri karena model data yang tidak terstruktur cocok untuk sistem. [17]. Pada tahap preprocessing, yang digunakan hanya variabel data komentar tweet saja

2.2.1. Tahapan Cleaning

Agar kata cloud dan analisis sentimen mencerminkan hasil nyata, beberapa pembersihan data harus dilakukan terlebih dahulu. Pembersihan data dilakukan dengan menghapus URL, mention, hashtag, emoji, tanda baca, dan angka., melakukan tahapan case folding guna menyamakan huruf pada data tweet menjadi huruf kecil (lowercase), juga melakukan tokenisasi, yang dimana proses memisahkan setiap kata pada suatu data tweet[18].

2.2.2. Tahapan Stopword

Pada tahapan Stopword, menghapus kata yang dianggap tidak memberikan dampak terhadap informasi yang ada pada suatu data tweet. Pembersihan data Stopwords pada penelitian ini menggunakan library python pySastrawi [19].

2.2.3. Tahapan Stemming

Dilakukan agar pada suatu data tweet yang memiliki imbuhan menjadi kata dasar dan menjadi suatu pola untuk dapat melakukan klasifikasi kata lain yang memiliki kata dasar maupun yang memiliki arti serupa namun berbeda karena memiliki imbuhan yang beda[17]. Pada tahap stemming ini menggunakan stemmer Tala, dimana tala memilih menggunakan komputasi dalam pencarian kata dasar dengan menggunakan algoritma berbasis aturan [20] .

2.2.4. Pelabelan Sistem

Dari data crawling cuitan twitter, dipilih 142 Data tweet Twitter menunjukkan contoh data tweet Twitter dengan sentimen positif (62 data) dan negatif (80 data). 142 buah data ini kemudian akan menjadi dataset pelatihan (berlabel) dan dataset pengujian (data tidak berlabel yang sama). Untuk memberi label kelas ini pada setiap timeline, penulis menggunakan kode dengan layout:0 untuk perasaan negatif dan 1 untuk perasaan positif. Pelabelan positif (kode=1) tidak berarti kalimat tersebut mendukung pengesahan KUHP secara eksplisit, tetapi pelabelan positif juga dilakukan untuk cuitan-cuitan yang membela pemerintah yang dalam hal ini mengenai pengesahan Undang-Undang KUHP.

2.3 Tahap Pemodelan

Saat ini, pengklasifikasi Naïve Bayes diterapkan sebagai metode klasifikasi dalam analisis sentimen. Pengklasifikasi Naïve Bayes yang diperkenalkan ke dalam penambangan data pembelajaran terawasi adalah sebuah pendekatan terhadap guru, yang diberi label kumpulan data (kelas) [21].

2.4. Klasifikasi

Klasifikasinya menggunakan pengklasifikasi Naïve Bayes dan dapat diklasifikasikan ke dalam kategori tertentu berdasarkan kata-kata yang terdapat dalam dokumen.

P-ISSN : 2089-3353
E-ISSN : 2808-9162

2.5. Evaluasi

Evaluasi kinerja dilakukan untuk memeriksa hasil klasifikasi dengan mengukur kebenaran nilai sistem. [22]. Parameter yang digunakan untuk mengukur nilai sebenarnya adalah akurasi. Akurasi adalah persentase dokumen yang diklasifikasikan dengan benar oleh sistem. Perhitungan akurat menggunakan metode matriks konfusi. Hitung keakuratannya menggunakan persamaan (4).

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (4)$$

2.6. Hasil Analisis Sentimen

Hasil dari keseluruhan penelitian yang telah dilaksanakan akan disajikan pada program. Setelah itu dapat dilakukan prediksi sentimen dengan menginputkan teks pada program lalu dihasilkan penentuan apakah teks tersebut merupakan sentimen positif ataupun negatif

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Dataset yang berhasil diakuisisi terdiri dari 501 data. Dari 501 catatan ini, diambil sebanyak 146 data untuk digunakan pada tahap selanjutnya. Dataset yang berhasil dikumpulkan ini dapat dilihat pada Gambar 2.

	text	username	label
0	@tribunnews Lah situ ngomong begitu SETELAH - ...	AshariSetiabud1	0
1	Setelah KUHP disahkan, para turis khawatir bil...	detikcom	0
2	Kemenkumham tengah menyiapkan petunjuk pelaksa...	korantempo	0
3	Kitab pidana itu resmi menggantikan regulasi l...	temponewsroom	0
4	Begini Isi RUU KUHP yang Baru Disahkan Pemerin...	infoka_id	0
...	...	...	...
496	KUHP (BARU) DISAHKAN, PASAL " KUMPUL KEBO " ...	jateng_time	0
497	Pengen deh pasal zina uu kuhp yg baru disahkan...	nugietriatmojo	0
498	4| Yasonna menjelaskan KUHP yang baru saja dis...	SekayuLapas	0
499	1| RUU KUHP resmi disahkan menjadi Undang-Unda...	SekayuLapas	0
500	Dalam Pasal 603 Kitab Undang-Undang Hukum Pida...	idntimes	0

Gambar 2. Hasil Crawling Data

3.2. Preprocessing

Setelah dilakukan crawling/pengumpulan data, maka selanjutnya dilakukan tahap preprocessing terhadap dataset. Pada tahap ini, akan dilakukan beberapa hal, yaitu membersihkan dataset dengan menghilangkan karakter-karakter yang tidak dibutuhkan (emoji, tanda baca, angka, url), case folding, serta menghilangkan tweet duplikat. Setelah dilakukan pembersihan ini, dataset yang ada menjadi 142 data. Data yang telah dibersihkan ini selanjutnya akan diberi label secara manual dengan ketentuan 0 adalah sentimen negatif dan 1 adalah sentimen positif, seperti yang dapat dilihat pada Gambar 3 berikut.

	text	username	label
0	@tribunnews Lah situ ngomong begitu SETELAH - ...	AshariSetiabud1	0
1	Setelah KUHP disahkan, para turis khawatir bil...	detikcom	0
2	Kemenkumham tengah menyiapkan petunjuk pelaksa...	korantempo	0
3	Kitab pidana itu resmi menggantikan regulasi l...	temponewsroom	0
4	Begini Isi RUU KUHP yang Baru Disahkan Pemerin...	infoka_id	0
...	...	...	...
496	KUHP (BARU) DISAHKAN, PASAL " KUMPUL KEBO " ...	jateng_time	0
497	Pengen deh pasal zina uu kuhp yg baru disahkan...	nugietriatmojo	0
498	4| Yasonna menjelaskan KUHP yang baru saja dis...	SekayuLapas	0
499	1| RUU KUHP resmi disahkan menjadi Undang-Unda...	SekayuLapas	0
500	Dalam Pasal 603 Kitab Undang-Undang Hukum Pida...	idntimes	0

Gambar 3. Dataset yang Telah Bersih

Selanjutnya, data yang telah bersih akan ditokenisasi dimana pada tahap ini tweet akan dipisahkan kata per kata seperti pada Gambar 3 berikut ini.

	Text	Label
0	[bukti, kecerobohan, dpr, pembuatan, kuhp, rekomendasi, lagi, mewakili, rakyat]	0
1	[tau, kah, kuhp, baru, korban, pelecehan, bisa, jadi, pelaku, nah, ini, bisa, jadi, implementasi...	0
2	[kuhp, luar, biasa, kuhp, modernisasi, dan, melepas, sistem, kolonialisme, di, indonesia]	1
3	[apanya, yang, salah, justru, wamenkumham, menjelasikan, bahwa, fasal, perzinahan, di, kuhp, tid...	1
4	[wah, nih, org, kena, pasal, hati, hati, bro, ntar, kena, kuhp, penghinaan, terhadap, kepala, ne...	1
...	...	...
137	[untung, gue, ga, jadi, masuk, fh, kuhp, nya, bikin, pusing]	1
138	[hukuman, mati, ternyata, bertentangan, dengan, ham, seperti, penjelasan, pasal, dan, rkuhp, tap...	1
139	[kalian, tau, ga, pasal, tentang, hukum, di, masyarakat, dianggap, bisa, membuka, ruang, perseku...	1
140	[rkuhp, udh, disahkan, dah, ketebak, indonesia, sebentar, lagi, gimana]	0
141	[yang, bener, aja, lagi, lagi, mengesahkan, pasalpasal, karet]	0

Gambar 4. Dataset Hasil Tokenisasi

Kemudian akan dibuat fungsi menghapus stopword pada seluruh tweet di dataset, dimana stopword ini sebelumnya telah dibuat pada file txt.

	Text	Label
0	[bukti, kecerobohan, dpr, pembuatan, kuhp, rekomendasi, lagi, mewakili, rakyat]	0
1	[tau, kah, kuhp, baru, korban, pelecehan, bisa, jadi, pelaku, nah, ini, bisa, jadi, implementasi...	0
2	[kuhp, luar, biasa, kuhp, modernisasi, dan, melepas, sistem, kolonialisme, di, indonesia]	1
3	[apanya, yang, salah, justru, wamenkumham, menjelasikan, bahwa, fasal, perzinahan, di, kuhp, tid...	1
4	[wah, nih, org, kena, pasal, hati, hati, bro, ntar, kena, kuhp, penghinaan, terhadap, kepala, ne...	1
...	...	...
95	[menkumham, yasonna, semprot, hotman, paris, soal, kuhp, seenak, udelnya]	0
96	[wagub, puji, pasal, kuhp, turis, asing, justru, berbondongbondong, ke, bali]	1
97	[banyak, pasal, fundamental, kuhp, baru, dinilai, lebih, bermanfaat]	1
98	[apalagi, buzzerrp, jual, diri, gak, laku, jual, martabat, hanya, untuk, receh, kehidupan, apa, ...	0
99	[sandiaga, uno, jamin, privasi, wisatawan, asing, terjaga, meski, ada, kuhp]	0

Gambar 5. Dataset Hasil *Stopwords Removal*

Setelah itu, dibuat fungsi stemming untuk menghilangkan seluruh kata berimbuhan sehingga kata-kata tersebut kembali ke kata dasar. Hasilnya dapat dilihat pada Gambar 5.

	Text	Label
0	bukti ceroboh dpr buat kuhp rekomendasi lagi w...	0
1	tau kah kuhp baru korban leceh bisa jadi laku ...	0
2	kuhp luar biasa kuhp modernisasi dan lepas sis...	1
3	apa yang salah justru wamenkumham menjelasikan...	1
4	wah nih org kena pasal hati hati bro ntar kena...	1
...	...	...
137	untung gue ga jadi masuk fh kuhp nya bikin pusing	1
138	hukum mati nyata tentang dengan ham seperti je...	1
139	kalian tau ga pasal tentang hukum di masyaraka...	1
140	rkuhp udh sah dah tebak indonesia sebentar lag...	0
141	yang bener aja lagi lagi kesah pasalpasal karet	0

Gambar 6. Dataset Hasil Stemming

3.3. Tahap Pemodelan

3.3.1 Evaluasi

Langkah terakhir setelah pra-pemrosesan adalah klasifikasi untuk menentukan apakah data yang diuji positif atau negatif. Dari proses pelatihan dengan menggunakan 142 data berlabel diperoleh hasil matriks evaluasi klasifikasi yaitu hasil pengukuran akurasi, *recall* dan F seperti terlihat pada Gambar 4. Untuk

P-ISSN : 2089-3353
E-ISSN : 2808-9162

sensor persepsi negatif diperoleh hasil akurasi 98%, 100 % recall, 99% skor f1. Untuk emosi positif, dapatkan akurasi 96,1%, presisi 100%, dan skor f1 98%.

	precision	recall	f1-score	support
0	0.82	1.00	0.90	9
1	1.00	0.67	0.80	6
accuracy			0.87	15
macro avg	0.91	0.83	0.85	15
weighted avg	0.89	0.87	0.86	15

Gambar 7. Hasil Classifier Evaluation Matrics

3.3.2 Hasil Analisis Sentimen

Dalam program, terdapat bagian untuk menentukan prediksi suatu kalimat dan jenis sentimen yang dihasilkan yakni sentimen positif atau negatif seperti yang disajikan pada Gambar 7.

```
In [117]: for tweet in ['Pemerintah macam anjing cuihhh', 'kuhp bikin indonesia tidak bebas berpendapat lagi']:
              p = naive_bayes_predict(tweet, logprior, loglikelihood)
              print(f'{tweet} -> {p:.2f}')

Pemerintah macam anjing cuihhh -> -0.75
kuhp bikin indonesia tidak bebas berpendapat lagi -> 1.00
```

Gambar 8. Hasil Prediksi Sentimen

Pada pengujian untuk prediksi diatas, kalimat "pemerintah macam anjing cuihhh" menghasilkan output -0,75 yang berarti kalimat tersebut termasuk sentimen negatif. Kemudian kalimat "kuhp bikin indonesia tidak bebas berpendapat lagi" menghasilkan output 1.00 yang berarti kalimat tersebut termasuk sentimen positif.

4. Kesimpulan

Kajian ini dilakukan untuk melihat bagaimana reaksi masyarakat terhadap pengesahan KUHP yang baru-baru ini dilakukan. Data menunjukkan, masyarakat cenderung bereaksi negatif terhadap hal ini. Analisis dilakukan dengan menggunakan metode Naïve Bayes untuk mengklasifikasikan dataset tweet. Metode Naïve Bayes merupakan model yang akurat dengan nilai akurasi sebesar 87%. Oleh karena itu, penulis berharap penelitian ini dapat memberikan pencerahan kepada masyarakat. Berdasarkan hasil pengujian yang dilakukan, kedepannya perlu dilakukan perbandingan kecepatan prediksi sentimen antara Naive Bayes Classifier dengan metode lainnya (SVM, KNN, dll).

Daftar Rujukan

[1] "Tuntut Cabut Pengesahan KUHP di Gedung DPR RI, Ketua BEM UI: Ada Pasal Karet dan Kontroversial - Wartakotalive.com."https://wartakota.tribunnews.com/2022/12/15/tuntut-cabut-pengesahan-kuhp-di-gedung-dpr-ri-ketua-bem-ui-ada-pasal-karet-dan-kontroversial (accessed Dec. 19, 2022).

[2] "Apa Isi RKUHP, Kenapa Ditolak, dan Daftar Pasal Kontroversial." https://tirto.id/apa-isi-rkuhp-kenapa-ditolak-dan-daftar-pasal-kontroversial-gzt6 (accessed Dec. 19, 2022).

[3] M. Z. Ansari, M. B. Aziz, M. O. Siddiqui, H. Mehra, and K. P. Singh, "Analysis of Political Sentiment Orientations on Twitter," in Procedia Computer Science, 2020, vol. 167, pp. 1821–1828. doi: 10.1016/j.procs.2020.03.201.

[4] Astiningrum, M., & Batubulan, K. S. (2020, September). Implementasi Analisis Sentimen Twitter Mengenai Opini Masyarakat Terhadap Rkuhp Tahun 2019. In Seminar Informatika Aplikatif Polinema.

[5] Wenando, F. A., Hayami, R., & Anggrawan, A. J. (2020). Analisis Sentimen Pada Pemerintahan Terpilih Pada Pilpres 2019 Ditwitter Menggunakan Algoritme Naïvebayes. JURTEKSI (Jurnal Teknologi dan Sistem Informasi), 7(1), 101-106.

[6] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," Jurnal Teknoinfo, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.

[7] S. Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen di Twitter and U. Gadjah Mada, "INCORBIZ (Indonesian Closed Corpus for Business) View project Semantic E-Learning View project Nurrun Muchammad Shiddieqy Hadna." [Online]. Available: https://www.researchgate.net/publication/292831965

[8] S. Analisis Pada Twitter CommuterLine, A. Pratama Putra, Y. Pratama, E. Kharisma Krisnadi, I. Purnamasari, and D. Dwi Saputra, "Text Mining untuk Sentimen Analisis dengan Metode Naïve Bayes, SMOTE, N-Gram dan AdaBoost Pada Twitter CommuterLine," 2022.

[9] Y. Pratiwi and A. Yaqin, "Klasifikasi Tweet Tidak Senonoh Twitter dengan Naïve Bayes Classifier," 2022.

[10] Wenando, F. A., Septiadi, R., Gunawan, R., & Mukhtar, H. (2022). K-Nearest Neighbor (KNN) untuk Menganalisis Sentimen terhadap Kebijakan Merdeka Belajar Kampus Merdeka pada Komentar Twitter. Jurnal CoSciTech (Computer Science and Information Technology), 3(2), 152-158."

[11] M. H. Alsharif, A. H. Kelechi, K. Yahya, and S. A. Chaudhry, "Machine learning algorithms for smart data analysis in internet of things environment: Taxonomies and research trends," Symmetry (Basel), vol. 12, no. 1, 2020, doi: 10.3390/SYM12010088.

[12] S. Sawla, "Introduction to Naive Bayes for Classification."

[13] B. Wagh, J. v Shinde, and N. R. Wankhade, "Sentimental Analysis on Twitter Data using Naive Bayes," International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, pp. 316–319, 2016.

[14] A. Troya, R. G. Pillai, C. R. Rivero, Z. Genç, S. Kayal, and D. Araci, "Aspect-Based Sentiment Analysis of Social Media Data With Pre-Trained Language Models," 2021 5th International Conference on Natural Language Processing and Information Retrieval (NLPIR), 2021.

[15] P. E. Akilandeswari, R. Harshita, and Sumanth. Ko.M, "Sentiment Analysis using Machine Learning through Twitter Streaming API," International Journal of Engineering & Technology, 2018.

[16] R. and S. H. and K. V. and M. E. Chandrasekaran Ranganathan and Desai, "Examining Public Sentiments and Attitudes Toward COVID-19 Vaccination: Infoveillance Study Using Twitter Posts," JMIR Infodemiology, vol. 2, no. 1, p. e33909, Apr. 2022, doi: 10.2196/33909.

[17] F. S. Mufidah, S. Winarno, F. Alzami, E. D. Udayanti, and R. R. Sani, "Analisis Sentimen Masyarakat Terhadap Layanan Shopeefood Melalui Media Sosial Twitter Dengan Algoritma Naïve Bayes Classifier," JOINS (Journal of Information System), vol. 7, no. 1, pp. 14–25, May 2022, doi: 10.33633/joins.v7i1.5883.

[18] A. Perdana, A. Hermawan, and D. Avianto, "Analisis Sentimen Terhadap Isu Penundaan Pemilu di Twitter Menggunakan Naive Bayes Clasifier," Jurnal Sisfokom (Sistem Informasi dan

P-ISSN : 2089-3353
E-ISSN : 2808-9162

Komputer), vol. 11, no. 2, pp. 195–200, Jul. 2022, doi: 10.32736/sisfokom.v11i2.1412.

[19]A. E. Budiman and A. Widjaja, "Analisis Pengaruh Teks Preprocessing Terhadap Deteksi Plagiarisme Pada Dokumen Tugas Akhir," Jurnal Teknik Informatika dan Sistem Informasi, vol. 6, no. 3, Dec. 2020, doi: 10.28932/jutisi.v6i3.2892.

[20]F. Z. Tala, "A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia."

[21]D. Alita, Y. Fernando, and H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," Jurnal TEKNOKOMPAK, vol. 14, no. 2, p. 86, 2020.

[22]M. A. Djamaludin, A. Triayudi, and E. Mardiani, "Analisis Sentimen Tweet KRI Nanggala 402 di Twitter menggunakan Metode Naïve Bayes Classifier," Jurnal Teknologi Informasi dan Komunikasi), vol. 6, no. 2, p. 2022, 2022, doi: 10.35870/jti.