

Analisis Sentimen Opini Publik Tentang Undang-Undang Cipta Kerja Pada Twitter

Tamora Nonia Wijaya
Program Studi Sistem Informasi
Universitas Nusantara PGRI Kediri
Kediri, Indonesia
tamoranonia@gmail.com

Rini Indriati
Program Studi Sistem Informasi
Universitas Nusantara PGRI Kediri
Kediri, Indonesia
rini.indriati@unpkediri.ac.id

Muhammad Najibulloh Muzaki
Program Studi Sistem Informasi
Universitas Nusantara PGRI Kediri
Kediri, Indonesia
m.n.muzaki@gmail.com

Diterima : Juni 2021
Disetujui : Juni 2021
Dipublikasi : Juli 2021

Abstrak— *Twitter* menjadi salah satu media sosial yang paling umum digunakan oleh masyarakat Indonesia. Dengan kepopuleran yang dimilikinya, menjadikan *Twitter* salah satu laman untuk mengekspresikan opini publik mengenai isu yang sedang dalam perbincangan. Undang-Undang Cipta Kerja adalah Omnibus Law pertama dalam hukum Indonesia yang disahkan yang mana didalamnya terdapat aturan yang mengatur tentang ketenagakerjaan, penyederhanaan perizinan, persyaratan investasi, hingga administrasi pemerintahan. Opini pro dan opini kontra terhadap Undang-Undang Cipta Kerja banyak dituangkan masyarakat Indonesia pada sosial media, dalam kasus ini *Twitter*. Penelitian yang dilakukan berguna untuk menganalisa opini masyarakat Indonesia di media sosial *Twitter* terhadap Undang-Undang Cipta Kerja dengan mengklasifikasi opini kedalam kelas positif atau negatif. Dengan menggunakan metode *Naive Bayes Classifier*, klasifikasi opini dilakukan oleh peneliti. Beberapa langkah yang dilakukan dalam penelitian yang diteliti kali ini yaitu pengumpulan data, pelabelan manual, *preprocessing*, *term weighting*, pemodelan, pengujian, dan evaluasi performa. Performa terbaik yang diperoleh oleh *Naive Bayes Classifier* adalah akurasi sebesar 89.9%, *precision* sebesar 90%, *recall* sebesar 89.9%, dan *f-1 score* sebesar 89.9%. Hasil yang didapat menunjukkan bahwa masyarakat Indonesia 52.9% kontra dan 47.1% pro terhadap Undang-Undang Cipta Kerja.

Kata Kunci— analisis sentimen; UU Cipta Kerja; *Naive Bayes Classifier*

Abstract— *Twitter* is one of the most familiar social media that used by Indonesian. With that popularity, *Twitter* is one of a page to express public opinions about the recent issues. Job Creation Law is the first Omnibus Law that was enacted in Indonesia and regulates employment, simplification of licensing, investment requirements, and government administration. Therefore, Job Creation Law received pros and cons from Indonesians on *Twitter*. The conducted research is to analyze public opinions towards by classifying the opinions into positive or negative class. The opinions were classified using *Naive Bayes Classifier* method. The stages in this study are data collection, manual labeling, *preprocessing*, *term weighting*, classification modeling, testing, and performance evaluation. The performance obtained by *Naive Bayes Classifier* is 89.9% accuracy, 90% precision, 89.9% recall,

and 89.9% *f-1 score*. The results showed that Indonesians are 52.9% against and 47.1% support Job Creation Law.

Keywords— sentiment analysis; Job Creation Law; *Naive Bayes Classifier*

I. PENDAHULUAN

Twitter sudah tidak asing sebagai salah satu aplikasi media sosial yang umum digunakan masyarakat Indonesia. *Twitter* dapat menjadi laman bagi para penggunanya untuk mengirim juga menerima pesan secara langsung (*real-time*). *Twitter* menempati peringkat kelima sebagai media sosial yang sering diakses oleh masyarakat Indonesia [1], menjadikan laman aplikasi tersebut sebagai tempat publik untuk beropini menyampaikan berbagai pendapatnya mengenai isu yang sedang menjadi perbincangan.

Undang-Undang Cipta Kerja merupakan Omnibus Law dalam hukum Indonesia yang disahkan oleh pemerintah, yang mana didalamnya terdapat aturan yang mengatur masalah mengenai ketenagakerjaan, penyederhanaan perizinan, persyaratan investasi, hingga administrasi pemerintahan. Dengan keputusan yang ada mengenai pengesahannya, banyak menuai beragam opini publik baik pro maupun kontra karena menilai kurangnya transparansi terhadap publik dan juga merugikan para buruh kerja. Opini pro dan kontra mengenai Undang-Undang Cipta Kerja banyak dituangkan oleh masyarakat Indonesia di media sosial *Twitter*. Oleh karena itu, dibutuhkan analisa opini pada *tweet* (pesan yang dikirim di *Twitter*) untuk menentukan opini yang bersifat positif atau negatif. Informasi ini dapat digunakan untuk mengetahui sentimen dan sudut pandang masyarakat Indonesia terhadap Undang-Undang Cipta Kerja.

Sebelumnya telah banyak yang meneliti tentang analisis sentimen di media sosial *Twitter*. Hamdana, dkk [2] telah melakukan penelitian analisis sentimen terhadap Omnibus Law dengan membandingkan akurasi *Naive Bayes* dan *K-Nearest Neighbor (K-NN)* dalam sistem berbasis java. Tingkat akurasi yang didapatkan menggunakan *Naive Bayes* adalah 75%. Sedangkan tingkat akurasi pada metode *K-NN* adalah 88%. Hasan, dkk [3] telah melakukan penelitian dengan melakukan perbandingan antara sentimen *lexicon (W-WSD, SentiWordNet, TextBlob)* dan menggunakan

perbandingan antara *Naive Bayes* dengan *Support Vector Machine (SVM)* dalam melakukan analisis sentimen. Hasil yang didapatkan menunjukkan bahwa hasil dari *Naive Bayes* menggunakan *W-WSD* memiliki akurasi yang paling tinggi sebesar 79%. Kemudian Fitri, dkk [4] melakukan perbandingan metode *Naive Bayes* dan *Support Vector Machine (SVM)* mengenai *tweet* yang berhubungan dengan Universitas Muhammadiyah Malang dan menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk pembobotan fitur atau kata. Hasil perbandingan menunjukkan *Naive Bayes* memiliki hasil akurasi yang lebih baik dibandingkan *SVM* yaitu 73.65%.

Naive Bayes Classifier dipilih sebagai metode pada penelitian ini karena dapat mengklasifikasikan pendapat ke dalam klasifikasi yang telah ditentukan yaitu sentimen positif, negatif ataupun netral. *Naive Bayes Classifier* sering digunakan sebagai metode klasifikasi karena kemudahannya dalam proses *training* dan klasifikasi [5]. Meskipun *Naive Bayes* merupakan metode yang sederhana dibandingkan metode klasifikasi lain, *Naive Bayes* efektif dan memiliki keakuratan tinggi dalam mengklasifikasi dokumen dengan banyak data [6].

Penelitian ini menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)* sebagai metode *term weighting* atau pembobotan kata untuk menunjukkan pengaruh suatu kata dalam isi dokumen. *TF-IDF* dipilih karena mudah diterapkan dalam menghitung keakuratan sebuah dokumen dibanding metode lain [7]. Penelitian terkait pernah dilakukan oleh Susandi, dkk [8] dengan melakukan komparasi antara *TF-IDF* dan *Weighted Inverse Document Frequency (WIDF)* dalam melakukan klasifikasi teks Berbahasa Indonesia. Hasil pengujian menunjukkan bahwa *TF-IDF* memiliki ketepatan (*precision*) sebesar 70.7% dan pembobotan *WIDF* memiliki ketepatan (*precision*) sebesar 63.1%.

II. METODE

A. Pengumpulan Data

Data dikumpulkan menggunakan teknik *scraping* data *tweet* pada *Twitter* menggunakan library *Twint* [9]. Data yang dikumpulkan merupakan data *tweet* yang mengandung kata dan *hashtag* “Undang-Undang Cipta Kerja” dan “Omnibus Law”. Data yang diambil adalah sebanyak 1000 *tweet* yang dikirim pada bulan Oktober 2020 hingga Nopember 2020.

B. Pelabelan Manual

Data selanjutnya diseleksi untuk mendapatkan data yang sesuai berdasarkan kebutuhan. Proses seleksi dilakukan dengan menyaring data yang tidak mengandung opini [10]. Data yang diperoleh dari hasil seleksi adalah sebanyak 687 data.

Pemberian label pada data dilakukan secara manual ke dalam klasifikasi kelas yang telah diputuskan, yakni negatif dan juga positif. Hasil dari pelabelan secara manual diperoleh 357 *tweet* dengan kelas positif dan 330 *tweet* dengan kelas negatif.

C. Preprocessing

Data selanjutnya akan dilakukan *preprocessing* [11] meliputi :

1. *Cleaning*, adalah proses menghapus karakter yang dianggap kurang penting seperti URL, *username*, *mention*, *hashtag*, tanda baca, angka, dan emotikon yang ada pada *tweet*.

2. *Case folding*, adalah proses menyeragamkan teks dengan mengubah huruf *uppercase* menjadi *lowercase*.
3. *Tokenization*, adalah proses pemecahan suatu teks berdasarkan setiap penyusun katanya.
4. Normalisasi, adalah proses menyesuaikan kata yang tidak sesuai dengan KBBI.
5. *Filtering*, adalah proses menghapus kata yang tidak mengandung makna (*stopword*) dan mengambil kata-kata penting hasil dari normalisasi.
6. *Stemming*, adalah proses pemetaan kata menjadi bentuk kata dasarnya.

D. Term Weighting

Term Weighting atau pembobotan kata adalah proses perhitungan bobot tiap kata pada dokumen sehingga dapat diketahui kemiripan suatu kata pada dokumen tersebut [12]. *Term Weighting* pada penelitian yang dilakukan yakni dengan penggunaan *Term Frequency-Inverse Document Frequency (TF-IDF)* sebagai metode penelitiannya.

Berikut adalah rumus *TF-IDF* [13] :

$$TF = \frac{\text{jumlah kemunculan kata}}{\text{jumlah kata pada dokumen}} \quad (1)$$

$$IDF = \log \cdot \frac{\text{jumlah dokumen}}{\text{jumlah suatu kata muncul}} \quad (2)$$

$$TFIDF = TF \cdot IDF \quad (3)$$

E. Pemodelan

Naive Bayes Classifier digunakan sebagai metode guna mengklasifikasi data yang ada dalam penelitian ini. *Naive Bayes Classifier* yaitu metode klasifikasi dengan penggunaan nilai probabilitas sebagai metodenya, yang mana dengan asumsi benar untuk label yang juga merupakan frekuensi dari setiap label yang terdapat pada *training set* dan kontribusi dari tiap fitur [14]. *Naive Bayes Classifier* juga digunakan dalam memprediksi adanya probabilitas dalam kelas data yang mana bahwa data tersebut akan masuk dalam kelas data tertentu [15].

Teorema *Naive Bayes* dapat dirumuskan sebagai berikut [15]:

$$P(H|X) = \frac{P(X|H)}{P(X)} \cdot P(H) \quad (4)$$

Dimana:

X : Data dengan kelas yang belum diketahui

H : Hipotesis data merupakan suatu kelas spesifik

P(H|X) : Probabilitas hipotesis H berdasarkan syarat X

P(H) : Probabilitas hipotesis H

P(X|H) : Probabilitas X berdasarkan syarat hipotesis H

P(X) : Probabilitas X

F. Pengujian Model

Pengujian dilakukan untuk menguji performa model dalam melakukan klasifikasi. Pada proses pengujian, dilakukan *split validation* yakni data dibagi secara acak sebagai data *training* dan data *testing* [16]. Data *training* merupakan data yang telah didapatkan dan digunakan untuk pembentukan model yang dapat memprediksi apakah *tweet* tersebut masuk ke kategori positif atau negatif. Sedangkan data *testing* merupakan data yang dipakai untuk menguji

keakuratan model klasifikasi. Kedua data akan dilakukan pembagian rasio sesuai yang telah ditentukan sebelumnya [17].

G. Evaluasi Performa

Performa dari prediksi yang dilakukan oleh algoritma *Naive Bayes* dan *TF-IDF* dievaluasi menggunakan teknik *Confusion Matrix* [18][19]. *Confusion Matrix* memberikan informasi perbandingan nilai aktual dengan nilai prediksi. Pada *Confusion Matrix*, terdapat empat nilai yang dihasilkan.

		Nilai Aktual	
		1 (Positive)	0 (Negative)
Nilai Prediksi	1 (Positive)	<p>TP (True Positive)</p>	<p>FP (False Positive)</p>
	0 (Negative)	<p>FN (False Negative)</p>	<p>TN (True Negative)</p>

Gambar 1. Confusion Matrix

Empat nilai tersebut dapat dijelaskan sebagai berikut;

1. *True Positive (TP)*, merupakan data aktual yang bernilai positif diprediksi benar.
2. *True Negative (TN)*, merupakan data aktual yang bernilai negatif diprediksi benar.
3. *False Positive (FP)*, merupakan data aktual yang bernilai negatif diprediksi positif.
4. *False Negative (FN)*, merupakan data aktual yang bernilai positif diprediksi negatif.

Matriks evaluasi yang digunakan, yaitu [20] [21] :

1. *Accuracy*, adalah rasio prediksi benar (positif dan negatif) dari jumlah seluruh data.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

2. *Precision*, adalah rasio prediksi benar yang positif dengan perbandingannya terhadap jumlah seluruh hasil prediksi positif.

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

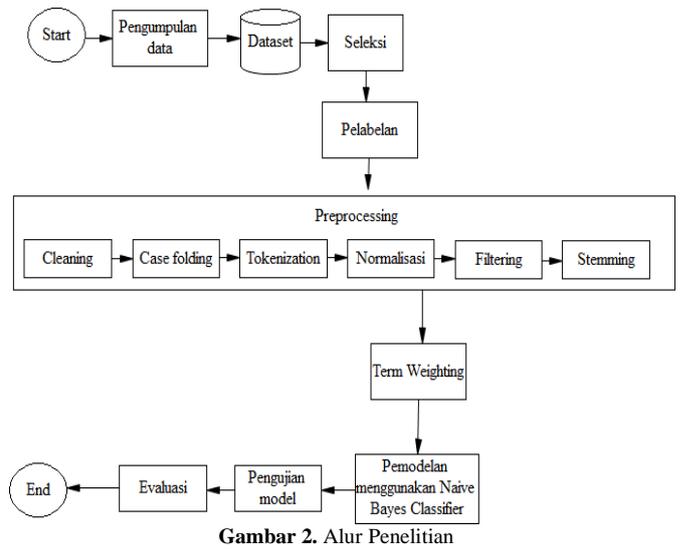
3. *Recall*, adalah rasio prediksi benar positif dengan perbandingannya terhadap seluruh data yang sudah benar positif.

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

4. *F1 Score*, adalah perhitungan rerata antara *precision* dan *recall*.

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision+recall} \quad (8)$$

Alur pada penelitian ini digambarkan pada Gambar 2.



Gambar 2. Alur Penelitian

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data dikumpulkan dengan teknik *web scraping* menggunakan *library Twint*. Data yang dikumpulkan merupakan data *tweet* yang mengandung kata dan *hashtag* “Undang-Undang Cipta Kerja” dan “Omnibus Law”. Berikut merupakan sampel *tweet* yang diperoleh.

	id	conversation_id	created_at	date	timezone	place	tweet	language	hashtags	cashtags	geo	source	user_rt_id
0	1324137994913421377	1324137994913421377	1.604534e+12	2020-11-05 06:54:33	700	NaN	Cavelkot Akhyar Tolak Omnibus Law https://t.c...	in			NaN	NaN	NaN
1	1324137112340441985	1324137112340441985	1.604534e+12	2020-11-05 06:51:03	700	NaN	Siapkan dirimu dan rapatkan barisan perlawanan...	in	[#tolakomnibuslaw, #pangumurperjuangan]		NaN	NaN	NaN
2	1324136805252231168	1324136805252231168	1.604534e+12	2020-11-05 06:49:50	700	NaN	Yg paling bertanggungjawab itu pasti bos bobot...	in	[#omnibuslaw]		NaN	NaN	NaN

Gambar 3. Sampel Tweet Hasil Web Scraping.

B. Pelabelan Manual

Data yang dikumpulkan kemudian dirapikan dengan menghapus kolom-kolom yang tidak diperlukan. Data selanjutnya disaring dengan menghapus *tweet* yang tidak berisi opini. Sampel *tweet* yang telah dilabeli ditampilkan pada Tabel 1 berikut.

Tabel 1. Data yang Telah Diberi Label

No	Tweet	Sentimen
1	Pemerintah membentuk pengelola investasi guna kemajuan UMKM dan peningkatan jumlah lapangan kerja #UUCiptaKerja #OmnibusLaw https://t.co/Tik5mK5bwS	Positif
2	@Yusrilihza_Mhd Sepertinya UU Celaka/omnibus law merupakan UU yg di sahkan & di tangani Secara cacat Hukum , karena ; pengesahan & penandatanganan nya pada UU yg belum siap/belum selesai dan juga UU tersebut tidak di terima lebih dari separuh lebih dari pada Bangsa Indonesia	Negatif
3	UU ini bertujuan mewujudkan masy.Indonesia makmur, adil dan sejahtera. Poin mana saja yg kalian tolak ? Sementara disisi lain sudah berkontribusi sampai dimana kalian untuk mewujudkan masyarakat yg kita dambakan. Tetap optimis dan lihat UU Omnibus Law Cipta Kerja ini bekerja	Positif
4	@republikaonline @OssyDermawan Kenapa Presiden tanda tangan juga sdh tau	Negatif

	temukan pasal 6 UU Cipta Kerja Bermasalah, Aneh begini yah apakah sdh ada titipan dari investor ! sejak diketok palu Azis yg pimpin rapat UU Cipta Kerja dan mattiin Mic oleh Puan tdk memberi peluang masukan dari peserta rapat ! aneh	
5	UU Ciptakerja sudah ditandatangani, Jrx divonis penjara 3 tahun, padahal dia dengan mempertaruhkan 'panggung besar' nya, berani bersuara banyak tentang isu lingkungan yang dianggap 'sepele' : #Balitolakreklamasi #Jogjaoradidol #savepetanikendeng Yakin negeri ini baik baik saja??	Negatif

C. Preprocessing

Setelah data diberi label, selanjutnya dilakukan *preprocessing*. Tahapan *preprocessing* yang diterapkan adalah *cleaning*, *case folding*, *tokenization*, normalisasi, *filtering*, dan *stemming*. Peneliti mengumpulkan daftar kata-kata baku maupun tidak baku yang disesuaikan dengan KBBI untuk digunakan pada proses normalisasi. Pada tahap *filtering* peneliti menggunakan daftar *stopword* yang telah ditentukan sebelumnya. Pada proses *stemming*, digunakan *library Sastrawi* Bahasa Indonesia. Hasil akhir *preprocessing* ditampilkan pada Tabel 2 berikut.

Tabel 2. Hasil *Preprocessing*

No	Tweet
1	perintah bentuk kelola investasi maju usaha mikro kecil dan tengah tingkat lapang kerja
2	undang undang celaka omnibus law undang undang sahkan tangan cacat hukum kesah penandatanganan undang undang belum belum selesai undang undang tidak terima paruh bangsa indonesia
3	undang undang tuju wujud masyarakat indonesia makmur adil sejahtera poin tolak di sisi kontribusi mana wujud masyarakat damba optimis undang undang omnibus law cipta kerja kerja
4	presiden tanda tangan tau temu pasal undang undang cipta kerja masalah aneh titip investor ketok palu azis pimpin rapat undang undang cipta kerja mattiin mic puan tidak peluang masuk serta rapat aneh
5	undang undang ciptakerja ditandatangani jrx vonis penjara tahun taruh panggung berani suara isu lingkung anggap sepele negeri

D. TF-IDF

Penentuan bobot setiap kata pada data *tweet* menggunakan algoritma *TF-IDF*. Hasil dari pembobotan menghasilkan matriks data 1760 atribut dari 687 data.

	term	rank
1681	undang	87.451668
1096	omnibus	47.982613
869	law	47.954733
753	kerja	45.710860
304	cipta	39.278443
...
980	mencret	0.154042
665	jebak	0.153656
430	entas	0.153656
1397	sejati	0.153656
408	edisi	0.153656

1760 rows × 2 columns

Gambar 4. Data yang Telah Diberi Bobot Tiap Kata

E. Pengujian

Proses pengujian dilakukan dengan melihat performa *Naive Bayes Classifier* dan pembobotan kata *TF-IDF* dalam

melakukan klasifikasi. Pada proses ini dilakukan *split validation* untuk mengetahui akurasi terbaik dari setiap rasio data *training* dan data *testing* yang berbeda. Jumlah total data adalah 687 data.

Tabel 3. Hasil Pengujian

No	Rasio	Akurasi
1.	90%:10%	88.4%
2.	80%:20%	89.9%
3.	70%:30%	87.9%
4.	60%:40%	88.4%

Pada Tabel 3 memperlihatkan bahwa besarnya data *training* dan data *testing* mempengaruhi hasil akurasi. Akurasi tertinggi didapatkan pada rasio data *training* 80% dan data *testing* 20% sebesar 89.9%. Yakni dengan jumlah data *training* sebesar 549 data dan data *testing* sebesar 138 data. Hasil klasifikasi label sebenarnya dan prediksi ditunjukkan pada Gambar 5, dimana label 0 merupakan sentimen negatif dan label 1 merupakan sentimen positif.

	Label Sebenarnya	Prediksi
0	1	1
1	0	0
2	0	0
3	1	1
4	0	0
...
133	1	1
134	1	1
135	1	1
136	0	0
137	1	1

138 rows × 2 columns

Gambar 5. Hasil Klasifikasi

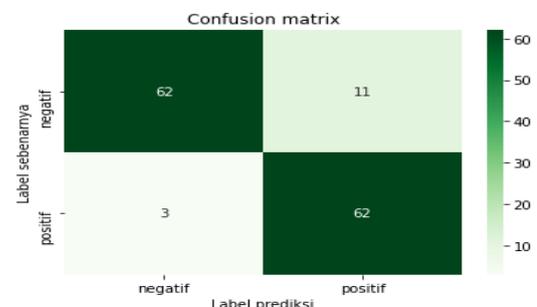
F. Evaluasi

Evaluasi performa dilakukan dengan menggunakan rasio terbaik yang telah dilakukan yaitu data *training* sebesar 80% dan data *testing* sebesar 20%.

```
test time: 0.002s
accuracy : 0.899
precision: 0.905
recall : 0.899
F1 Score : 0.899
```

	precision	recall	f1-score	support
negatif	0.95	0.85	0.90	73
positif	0.85	0.95	0.90	65
accuracy			0.90	138
macro avg	0.90	0.90	0.90	138
weighted avg	0.90	0.90	0.90	138

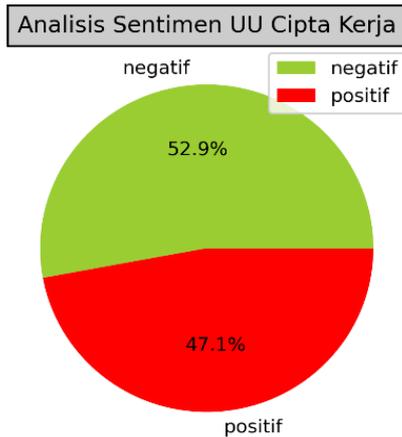
Text(33.0, 0.5, 'Label sebenarnya')



Gambar 6. Hasil Evaluasi Naive Bayes

Dari Gambar 6 diperoleh skor *Naive Bayes* dalam melakukan klasifikasi yaitu akurasi sebesar 89.9%, *precision*

sebesar 90%, *recall* sebesar 89.9%, dan *f-1 score* sebesar 89.9%.



Gambar 7. Sentimen terhadap Undang-Undang Cipta Kerja

Hasil sentimen ditunjukkan pada Gambar 7 yang didapat dari data *testing* sebesar 138 data ditunjukkan pada Gambar 7 yakni 52.9% memiliki sentimen negatif dan 47.1% bersentimen positif terhadap Undang-Undang Cipta Kerja.

G. Word Cloud

Analisis *word cloud* dilakukan untuk mengetahui kata yang sering diungkapkan tentang Undang-Undang Cipta Kerja. *Word cloud* masing-masing sentimen dibandingkan untuk menunjukkan kata-kata yang sering digunakan pada *tweet* yang memiliki sentimen positif atau negatif.

Word cloud pada sentimen positif ditunjukkan pada Gambar 8.



Gambar 8. *Word Cloud* pada Sentimen Positif

Pada *word cloud tweet* yang mengandung sentimen positif, ditunjukkan bahwa kata-kata yang sering digunakan adalah usaha mikro, mikro menengah, jokowi terpercaya, dukung, presiden, ekonomi, dan sebagainya.

Word cloud pada sentimen negatif ditunjukkan pada Gambar 9.



Gambar 9. *Word Cloud* pada Sentimen Negatif

Pada *word cloud tweet* yang mengandung sentimen negatif, ditunjukkan bahwa kata-kata yang sering digunakan adalah presiden, demo, tolak, pemerintah, aksi, dewan perwakilan, salah ketik, buruh, salah, dan sebagainya.

Pada masing-masing *word cloud* dapat dilihat bahwa pada tiap sentimen memiliki kata-kata yang bermakna berseberangan. Misal pada *tweet* dengan sentimen positif memiliki kata-kata dengan makna positif seperti dukung, kemudahan, manfaat, dan mendorong. Sedangkan pada *tweet* dengan sentimen negatif memiliki kata-kata dengan makna negatif seperti tolak, salah ketik, cacat, dan ugal-ugalan.

IV. KESIMPULAN

A. Kesimpulan

Sesuai dengan penelitian yang telah dilakukan, penerapan metode *Naive Bayes Classifier* dan *TF-IDF* dalam melakukan analisis sentimen tentang Undang-Undang Cipta Kerja menghasilkan akurasi sebesar 89.9%, *precision* sebesar 90%, *recall* sebesar 89.9%, dan *f-1 score* sebesar 89.9% dengan data *training* sebesar 80% dan data *testing* sebesar 20%. Sentimen masyarakat Indonesia terhadap Undang-Undang Cipta di media sosial *Twitter* menggunakan data *testing* sebesar 138 data. Hasil klasifikasi menunjukkan bahwa masyarakat Indonesia 52.9% kontra dan 47.1% pro terhadap Undang-Undang Cipta Kerja; Hasil analisis terhadap *word cloud* masing-masing sentimen menunjukkan bahwa terdapat persamaan dan perbedaan. Persamaannya adalah banyaknya kata jokowi, pemerintah, pekerja, buruh yang sering muncul pada kedua sentimen tersebut. Hal ini menunjukkan pihak yang paling sering muncul pada *tweet* tentang Undang-Undang Cipta Kerja. Sedangkan perbedaan dari kedua sentimen tersebut adalah kata-kata yang memiliki makna berseberangan dan bertolak belakang. Hal ini menunjukkan sudut pandang pengguna *Twitter* terhadap Undang-Undang Cipta Kerja.

B. Saran

Penelitian selanjutnya dapat melakukan pembobotan terhadap emotikon agar batas sentimen positif dan negatif menjadi jelas. Serta dapat menggunakan lebih banyak data *training* daripada penelitian yang telah dilakukan, dengan demikian dapat mendapatkan hasil dengan tingkat keakuratan yang lebih tinggi.

REFERENSI

- [1] We Are Social & Hootsuite, "Indonesia Digital Report 2020," 2020. [Daring]. Tersedia pada: <https://datareportal.com/reports/digital-2020-indonesia>.
- [2] E. N. Hamdana, "Pengembangan Sistem Analisis Sentimen Berbasis Java Pada Data Twitter Terhadap Omnibus Law Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor (K-NN)," *J. Inform. Polinema*, vol. 7, no. 2, hal. 79–84, 2021, doi: 10.33795/jip.v7i2.688.
- [3] A. Hasan, S. Moin, A. Karim, dan S. Shamsirband, "Machine Learning-Based Sentiment Analysis for Twitter Accounts," *Math. Comput. Appl.*, vol. 23, no. 1, hal. 11, 2018.
- [4] M. I. Fikri, T. S. Sabrila, dan Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Smatika J.*, vol. 10, no. 02, hal. 71–76, 2020.
- [5] S. Chakrabarti, S. Roy, dan M. V. Soundalgekar, "Fast and accurate text classification via multiple linear discriminant projections," dalam *VLDB '02: Proceedings of the 28th International Conference on Very Large Databases*, Elsevier, 2002, hal. 658–669.
- [6] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creat. Inf. Technol. J.*, vol. 2, no. 3, hal. 207–217, 2015.
- [7] M. A. Rofiqi, A. C. Fauzan, A. P. Agustin, dan A. A. Saputra, "Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 1, no. 2, hal. 58–64, 2019.
- [8] D. Susandi dan U. Sholahudin, "Pemanfaatan Vector Space Model pada Penerapan Algoritma Nazief Adriani, KNN dan Fungsi Similarity Cosine untuk Pembobotan IDF dan WIDF pada Prototipe Sistem Klasifikasi Teks Bahasa Indonesia," *J. ProTekInfo*, vol. 3, no. 1, hal. 22–29, 2016.
- [9] M. Nurmalasari, N. A. Temesvari, dan S. N. Maulana, "Analisis Sentimen terhadap Opini Masyarakat dalam Penggunaan Mobile-JKN untuk Pelayanan BPJS Kesehatan Tahun 2019," *Indones. Heal. Inf. Manag. J.*, vol. 8, no. 1, hal. 35–44, 2020.
- [10] A. Annisa Raudya Wibowo, Nuke Nidya, Aisyah Firdausi Rahma, "ANALISIS SENTIMEN HASHTAG 'DIRUMAHAJA' SAAT PANDEMI COVID-19 DI INDONESIA MENGGUNAKAN NLP," *J. Inform. dan Sist. Inf.*, vol. 1, no. 2, hal. 343–353, 2020.
- [11] H. Jiawei, M. Kamber, dan P. Jian, *Data Mining: Concepts and Techniques*, 3rd ed. Burlington, New Jersey: Morgan Kaufmann Publisher, 2011.
- [12] W. Parasati, F. A. Bachtiar, dan N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*, vol. 4, no. 4, hal. 1090–1099, 2020.
- [13] R. Feldman dan J. Sanger, *The Text Mining Handbook*. New York: Cambridge University Press, 2006.
- [14] M. W. Berry dan J. Kogan, *Text Mining: Applications and Theory*. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2010.
- [15] F. Handayani dan S. Pribadi, "Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110," *J. Tek. Elektro*, vol. 7, no. 1, hal. 19–24, 2015.
- [16] F. Nurhuda, S. Widya Sihwi, dan A. Doewes, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *J. Teknol. Inf. ITSmart*, vol. 2, no. 2, hal. 35, 2016.
- [17] G. Berliana, S. Sa, S. Prodi, T. Informatika, F. Informatika, dan U. Telkom, "Klasifikasi Posting Tweet mengenai Kebijakan Pemerintah Menggunakan Naive Bayesian Classification," *e-Proceeding Eng.*, vol. 5, no. 1, hal. 1562–1569, 2018.
- [18] F. Rozy, S. Rangkuti, M. A. Fauzi, Y. A. Sari, E. Dewi, dan L. Sari, "Analisis Sentimen Opini Film Menggunakan Metode Naive Bayes dengan Ensemble Feature dan Seleksi Fitur Pearson Correlation Coefficient," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 12, hal. 6354–6361, 2018.
- [19] Y. R. Silitonga, J. Arjuna, U. No, dan K. Jeruk, "SISTEM PENDETEKSI BERITA HOAX DI MEDIA SOSIAL DENGAN TEKNIK DATA MINING SCIKIT LEARN," *J. Ilmu Komput.*, vol. 4, hal. 173–179, 2019.
- [20] D. A. Fauziah, A. Maududie, dan I. Nuritha, "Klasifikasi Berita Politik Menggunakan Algoritma K-nearest Neighbor (Classification of Political News Content using K-Nearest Neighbor) Abstrak," *J. Sist. Inf.*, vol. 6, no. 2, hal. 8, 2018.
- [21] Z. Y. Lamasigi, "DCT Untuk Ekstraksi Fitur Berbasis GLCM Pada Identifikasi Batik Menggunakan K-NN," *Jambura J. Electr. Electron. Eng.*, vol. 3, no. 1, hal. 1–6, 2021.