

Komparasi Metode Naive Bayes dan SVM pada Sentimen Twitter Mengenai Persoalan Perppu Cipta Kerja

Nur Muhammad Farhan¹, Bayu Setiaji²

farhan@students.amikom.ac.id, bayusetiaji@amikom.ac.id

Universitas Amikom Yogyakarta

Informasi Artikel

Diterima : 31 Ags 2023

Direview : 18 Sep 2023

Disetujui : 23 Okt 2023

Kata Kunci

Analisis Sentimen, Perpu Cipta Kerja, Naïve Bayes, Support Vector Machine, Komparasi Algoritma

Abstrak

Salah satu yang baru ini ramai diperbincangkan adalah persoalan UU Cipta Kerja yang banyak ragam orang berbincangkan mengenai dampak positif dan negatif mengenai hal tersebut. Oleh karena itu analisa sentimen pada persoalan UU Cipta Kerja untuk bisa mengetahui berapa banyak orang yang menolak atau mendukung hal tersebut. Penelitian ini menggunakan data tweet sebanyak 622 data tweet yang berbahasa Indonesia. Kemudian ulasan tersebut di katagorikan ke beberapa sentimen dan algoritma, sentimen positif mendapatkan sebanyak 224 tweet, sebanyak 332 tweet yang bersentimen negative, dan sebanyak 66 bersentimen netral. Kemudian data tersebut dimasukkan kedalam algoritma Naïve Bayes dan SVM untuk menentukan tingkat akurasi yang didapatkan. Algoritma Naïve Bayes mendapatkan akurasi 73% dengan data akurasi training 87% dan SVM mendapatkan 78% untuk data testing dengan akurasi 99% dengan data training. Dari hasil tersebut menunjukka bahwa tingkat akurasi algoritma SVM lebih tinggi daripada akurasi algoritma Naïve Bayes.

Keywords

Sentiment Analysis, Job Creation Regulation, Naïve Bayes, Support Vector Machine, Algorithm Comparison

Abstrak

One that has recently been discussed is the issue of the Job Creation Law, which many people are discussing about the positive and negative impacts on this matter. Therefore, sentiment analysis on the issue of the Job Creation Law is to be able to find out how many people reject or support it. This study uses tweet data as much as 622 tweet data in Indonesian. Then the review is categorized into several sentiments and algorithms, positive sentiment gets as many as 224 tweets, as many as 332 tweets with negative sentiments, and as many as 66 with neutral sentiments. Then the data is entered into the Naïve Bayes and SVM algorithms to determine the level of accuracy obtained. The Naïve Bayes algorithm gets 73% accuracy with 87% training accuracy data and SVM gets 78% for testing data with 99% accuracy with training data. These results show that the accuracy of the SVM algorithm is higher than the accuracy of the Naïve Bayes algorithm.

A. Pendahuluan

Pendahuluan Twitter adalah salah satu media sosial yang berperan sebagai wadah layanan komunikasi di kalangan masyarakat. Menggunakan media Twitter, masyarakat di seluruh dunia dapat berinteraksi dengan keluarga, teman dan orang-orang tersayang melalui komputer atau ponsel yang terhubung dengan internet[1]. Twitter menjadi media layanan sosial yang paling dikenal di dunia dengan jumlah lebih dari 200 juta pengguna aktif dan lebih dari 10,6 milyar tweet yang sudah dipublikasikan[2]. Menurut laporan berjudul "Digital 2021" diterbitkan oleh sumber manajemen konten Hootsuite dan agen pemasaran media sosial We Are Social, pengguna internet di Indonesia pada awal 2021 telah mencapai titik senilai 202,6 juta dan hingga 170 juta penduduk Indonesia adalah pengguna internet, aktif di jejaring sosial [3]. Menurut laporan HootSuite dan We Are Social, Twitter masuk dalam lima besar media sosial yang paling banyak dikunjungi di Indonesia.

Analisis sentimen adalah salah satu metode meneliti dalam menyampaikan opini ataupun sentimen yang dicurahkan lewat kalimat ataupun text[5]. Proses pelatihan pada analisis sentimen lebih sulit daripada bidang *machine learning* lain karena data pada analisis sentimen bersifat subjektif yaitu berupa opini yang nilainya tidak konkrit[4].

Sebagian besar masyarakat memberikan opini mereka mengenai pengesahan perpu cipta kerja di suatu platform media sosial Twitter sehingga sering menjadi trending untuk belakangan ini. Opini yang diberikan oleh masyarakat berupa tweet yang merupakan sebuah data teks dimana suatu data tersebut bisa dinilai dan bisa menjadi sebuah kesimpulan seberapa banyak orang yang belum memahami tentang dampak dari perpu cipta kerja dari hal yang positif dan hal negatif.

Berdasarkan tweet yang diberikan oleh netizen atau masyarakat Indonesia mengenai perpu cipta kerja diharapkan bahwa pihak pemerintah bisa lebih tegas dalam memperinci sebuah aturan atau perpu yang terbaru agar tidak terjadi kesalahpahaman antara masyarakat dengan pihak lain. Hal tersebut juga bisa menjadi salah satu pencegahan perpecahan di negara Indonesia karena ketidaksepahaman mengenai perpu tersebut. Dengan adanya penelitian ini juga bisa memberikan pandangan kepada pihak terkait mengenai pandangan masyarakat tentang perpu cipta kerja ini.

Dari opini dari permasalahan tersebut akan dibagi menjadi tiga bagian yaitu positif, negatif, dan netral[6][7]. Penelitian ini bertujuan untuk mengetahui sentimen yang diberikan pada UU cipta kerja yang disahkan menjadi Perpu dan bisa menghasilkan sebuah pandangan masyarakat yang masih minim dan bisa menjadi bumerang bagi negara Indonesia ini. Pada proses klasifikasi di penelitian ini menggunakan algoritma *machine learning* yaitu dengan *Naïve Bayes* dan *Support Vector Machine* untuk mengetahui hasil klasifikasi yang terjadi pada sentimen yang sudah dilontarkan dari masyarakat melalui Twitter. Selain itu untuk mengetahui seberapa tinggi nilai sentimen tersebut digunakan metode *lexicon*, dimana sebuah kata-kata yang dituliskan sudah dibobotkan oleh sebuah metode ini. Penelitian ini diharapkan bisa memberikan akurasi yang tinggi sehingga bisa menjadi tolak ukur kepada pemerintah tentang pengetahuan masyarakat yang masih terbatas.

Beberapa penelitian yang meneliti atau mempelajari tentang *text mining* diantaranya penelitian dengan topik [8], pada penelitian tersebut melakukan teks sentimen dengan pembagian dua variabel dan menggunakan TF-IDF dalam metode

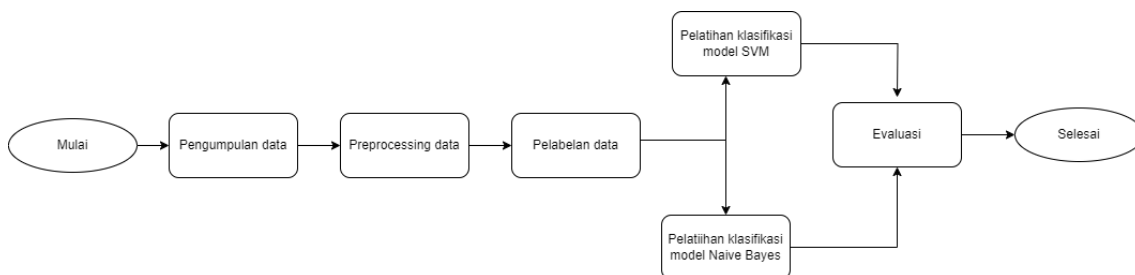
LSTM dengan hasil akurasi 77% sedangkan untuk perbandingan dengan algoritma *Naïve Bayes* mendapatkan akurasi sebesar 76%, hal tersebut menunjukkan bahwa algoritma LSTM lebih tinggi daripada *Naïve Bayes* pada penelitian tersebut.

Untuk penelitian [9] penelitian ini berisikan tentang sentimen pemindahan ibu kota, data yang di peroleh dari data Twitter dan mempunyai hasil 1236 data dengan memperoleh hasil akurasi sebesar penggunaan algoritma SVM 96.68%. Pada penelitian [10] yang meneliti dengan membandingkan dua metode yaitu SVM dan *Naïve Bayes*, menghasilkan nilai akurasi SVM sebesar 76,2% sedangkan untuk metode 62,47 pada penelitian sentimen terhadap Perusahaan Listrik Negara.

Selain itu juga terdapat penelitian[11] yang membahas tentang sentimen analisis mengenai Informasi *Covid-19* menggunakan algoritma SVM dan *Naïve Bayes*. Pada penelitian tersebut memperoleh hasil sebesar 80,23% SVM dan *Naïve Bayes* sebesar 78.02%. Penelitian tersebut menyatakan bahwa akurasi algoritma SVM lebih tinggi dibandingkan dengan *Naïve Bayes*.

B. Metode Penelitian

Metode penelitian pada penelitian ini menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes* untuk mengklasifikasikan sentimen dan kemudian dikomparasikan atau dibandingkan mengenai kedua akurasi algoritma. Dalam membuat model dilakukan dengan pemrograman *Python*. Dibawah ini alur penelitian yang diimplementasikan.



Gambar 1. Alur Penelitian

1. Pengumpulan Data

Tahap pertama yang dilakukan pada penelitian ini yaitu mengumpulkan data. Data diambil dari dataset Twitter dengan cara *crawling*. Data diambil dari dataset Twitter dengan cara *crawling*. Data yang didapatkan pada pengambilan data sebanyak 622 data sentimen yang terdiri dari tanggal dan isi dari tweet tersebut. Data tersebut di ambil dari tanggal 30 Desember 2022 hingga 21 Maret 2022, dari data tersebut kemudian disimpan kedalam format .csv. Pada awal pengumpulan data sudah menggunakan kode pemrograman *Python* untuk melakukan *case folding* dan *cleansing* pada awal pengambilan data Contoh data yang diambil dari Twitter pada tabel 1.

Tabel 1. Contoh Data

No	Date	Tweet
1	2023-03-20 14:25:12+00:00	Aksi dari berbagai elemen mahasiswa menyuarakan penolakan perpu ciptakerja semoga DPRRI telinganya difungsikan untuk di dengar
2	2023-03-20 12:47:21+00:00	massa aksi tolak perpu cipta kerja masih bertahan di lokasi hingga pukul 1834 WIB
3	2023-03-20 12:28:49+00:00	Berkali masyarakat menolak perpu cipta kerja namun pemimpin abaikan penolakan dari rakyat padahal pemimpin sekarang hanya taunya membuat kebijakan tanpa melihat langsung nasib rakyatnya serbaserbi mmc islamkaffah
4	2023-03-20 12:19:45+00:00	Menilik perpu cipta kerja untk buruh sangat buruk para buruh seperti sapi perah yg dimnfaatkn tenaga nya tapi upah yg diberikn tidak sesuai bahkn kerap kali trjadi pemotongn upah karna hal trtentu nasib rakyat prlu diprjuangkn serbaserbi mmc islamkaffah
5	2023-03-20 12:09:45+00:00	Nasib buruh kian di ujung tanduk setelah disahkannya perpu cipta kerja aturan yg tertuang benar berpijak kepada perusahaanpemilik modal bagaimana rakyat bisa sejahtera jika tenaga diperas tanpa ampun serbaserbi mmc islamkaffah

2. Preprocessing Data

Preprocessing data bertujuan agar data yang di eksekusi dapat mendapatkan hasil yang maksimal. Tahap ini melaukan pemrosesan sebanyak 622 data yang sudah didapatkan melalui *crawling* Twitter. Metode atau teknik yang digunakan untuk *preprocessing* data menggunakan teknik *Case Folding*, *Cleansing*, *Tokenizing*, *Stemming*.

a. Case Folding

Case Folding adalah salah satu metode klasifikasi dengan cara kerjanya memproses kata dalam sebuah data dokumen menjadi bentuk standar atau huruf kecil. Sedangkan karakter yang lain dianggap sebagai pembatas.[12]. Berhubung pengambilan data sudah melalui *case folding*, maka *case folding* pada penelitian ini tidak dilakukan di pertengahan *preprocessing*.

b. Cleansing

Cleansing adalah salah satu tahapan dalam runtutan metode *preprocessing* yang dimaksudkan untuk menghilangkan karakter maupun simbol *link url* (<https://situs.com>), *username* ataupun *mention* @nama aku user, hastag(#), retweet, dan juga *emocticon*. Namun pada penelitian kali ini pengambilan data sudah melalui *cleansing* sehingga data yang disajikan pada tabel 1 sudah melalui *cleansing*.

c. Tokenizing

Tokenizing adalah sebuah tahap pemisahan antar kalimat atau sebuah kata yang sendiri-sendiri. Pada penelitian ini untuk *tokenizing* menggunakan *library* Python yaitu *nltk* dengan *word_tokenize* dan *FreqDisk* Hasil *tokenazing* dapat dilihat pada tabel dibawah yaitu tabel 2.

Tabel 2. Hasil Dari Tokenizing

Tweet	Tweet_tokens
aksi dari berbagai elemen mahasiswa menyuarakan penolakan perpu ciptakerja	['aksi', 'dari', 'berbagai', 'elemen', 'mahasiswa', 'menyuarakan', 'penolakan', 'perpu', 'ciptakerja', 'semoga', 'dprri',

semoga DPRRI telinganya difungsikan untuk di dengar	'telinganya', 'difungsikan', 'untuk', 'di', 'dengar']
massa aksi tolak perpu cipta kerja masih bertahan di lokasi hingga pukul wib	['massa', 'aksi', 'tolak', 'perpu', 'cipta', 'kerja', 'masih', 'bertahan', 'di', 'lokasi', 'hingga', 'pukul', 'wib']
berkali masyarakat menolak perpu cipta kerja namun pemimpin abaikan penolakan dari rakyat padahal pemimpin sekarang hanya taunya membuat kebijakan tanpa melihat langsung nasib rakyatnya serbaserbi mmc islamkaffah	['berkali', 'masyarakat', 'menolak', 'perpu', 'cipta', 'kerja', 'namun', 'pemimpin', 'abaikan', 'penolakan', 'dari', 'rakyat', 'padahal', 'pemimpin', 'sekarang', 'hanya', 'taunya', 'membuat', 'kebijakan', 'tanpa', 'melihat', 'langsung', 'nasib', 'rakyatnya', 'serbaserbi', 'mmc', 'islamkaffah']
menilik perpu cipta kerja untk buruh sangat buruk para buruh seprti sapi perah yg dimnfaatkn tenaga nya tapi upah yg diberikn tidak sesuai bahkn kerap kali trjadi pemotongn upah karna hal trtentu nasib rakyat prlu diprjuangkkn serbaserbi mmc islamkaffah	['menilik', 'perpu', 'cipta', 'kerja', 'untk', 'buruh', 'sangat', 'buruk', 'para', 'buruh', 'seprti', 'sapi', 'perah', 'yg', 'dimnfaatkn', 'tenaga', 'nya', 'tapi', 'upah', 'yg', 'diberikn', 'tidak', 'sesuai', 'bahkn', 'kerap', 'kali', 'trjadi', 'pemotongn', 'upah', 'karna', 'hal', 'trtentu', 'nasib', 'rakyat', 'prlu', 'diprjuangkkn', 'serbaserbi', 'mmc', 'islamkaffah']
nasib buruh kian di ujung tanduk setelah disahkannya perpu cipta kerja aturan yg tertuang benar berpijak kepada perusahaanpemilik modal bagaimana rakyat bisa sejahtera jika tenaga diperas tanpa ampun serbaserbi mmc islamkaffah	['nasib', 'buruh', 'kian', 'di', 'ujung', 'tanduk', 'setelah', 'disahkannya', 'perpu', 'cipta', 'kerja', 'aturan', 'yg', 'tertuang', 'benar', 'berpijak', 'kepada', 'perusahaanpemilik', 'modal', 'bagaimana', 'rakyat', 'bisa', 'sejahtera', 'jika', 'tenaga', 'diperas', 'tanpa', 'ampun', 'serbaserbi', 'mmc', 'islamkaffah']

d. *Stemming*

Stemming merupakan salah satu tahapan akhir yang digunakan untuk merubah kata imbuhan menjadi kata dasar. *Stemming* yang digunakan yaitu *library* sastrawi, dalam penelitian ini juga menggunakan swifter dan juga menggunakan *stopword* yang menggunakan bahasa Indonesia. Untuk hasil data yang sudah memlalui proses ini sebagai berikut pada tabel 3.

Tabel 3. Contoh Hasil *Stemming*

Tweet	tweet_tokens_stemmed
Aksi dari berbagai elemen mahasiswa menyuarakan penolakan perpu ciptakerja semoga DPRRI telinganya difungsikan untuk di dengar	['aksi', 'elemen', 'mahasiswa', 'suara', 'tolak', 'perpu', 'ciptakerja', 'moga', 'dprri', 'telinga', 'fungsi', 'dengar']
massa aksi tolak perpu cipta kerja masih bertahan di lokasi hingga pukul 1834 WIB	['massa', 'aksi', 'tolak', 'perpu', 'cipta', 'kerja', 'tahan', 'lokasi', 'wib']
Berkali masyarakat menolak perpu cipta kerja namun pemimpin abaikan penolakan dari rakyat padahal pemimpin sekarang hanya taunya membuat kebijakan tanpa melihat langsung nasib rakyatnya serbaserbi mmc islamkaffah	['kali', 'masyarakat', 'tolak', 'perpu', 'cipta', 'kerja', 'pimpin', 'abai', 'tolak', 'rakyat', 'pimpin', 'tau', 'bijak', 'langsung', 'nasib', 'rakyat', 'serbaserbi', 'mmc', 'islamkaffah']
Menilik perpu cipta kerja untk buruh sangat buruk para buruh seprti sapi perah yg dimnfaatkn tenaga nya tapi upah yg diberikn tidak sesuai bahkn kerap kali trjadi pemotongn upah karna hal trtentu nasib	['tilik', 'perpu', 'cipta', 'kerja', 'buruh', 'buruk', 'buruh', 'seprti', 'sapi', 'perah', 'dimnfaatkn', 'tenaga', 'upah', 'diberikn', 'sesuai', 'bahkn', 'kerap', 'kali', 'pemotongn', 'upah', 'trtentu', 'nasib', 'rakyat', 'diprjuangkkn', 'serbaserbi', 'mmc', 'islamkaffah']

rakyat prlu diprjuangkkn serbaserbi mmc
islamkaffah

Nasib buruh kian di ujung tanduk setelah disahkannya perpu cipta kerja aturan yg tertuang benar berpijak kepada perusahaanpemilik modal bagaimana rakyat bisa sejahtera jika tenaga diperas tanpa ampun serbaserbi mmc islamkaffah

['nasib', 'buruh', 'kian', 'ujung', 'tanduk', 'sah', 'perpu', 'cipta', 'kerja', 'atur', 'tuang', 'pijak', 'perusahaanpemilik', 'modal', 'rakyat', 'sejahtera', 'tenaga', 'peras', 'ampun', 'serbaserbi', 'mmc', 'islamkaffah']

3. Pelabelan Data

Pelabelan data, pada proses atau metode ini menggunakan *lexicon* untuk melakukan pelabelan sentimen secara otomatis. Untuk mendapatkan file atau data *lexicon* men-*download* pada github yang sudah menyediakan data matang. Dari hasil *lexicon* didapatkan data sejumlah 332 data negatif, 224 data positif, dan sebanyak 66 data netral. *Lexicon* sendiri dihitung dari nilai *polarity* pada kata atau sentimen yang diambil. Untuk contoh hasil pelabelan menggunakan *lexicon* bisa dilihat pada tabel 4.

Tabel 4. Contoh Pelabelan

tweet_tokens_stemmed	polarity_score	polarity
['aksi', 'elemen', 'mahasiswa', 'suara', 'tolak', 'perpu', 'ciptakerja', 'moga', 'dprri', 'telinga', 'fungsi', 'dengar']	2	Positif
['massa', 'aksi', 'tolak', 'perpu', 'cipta', 'kerja', 'tahan', 'lokasi', 'wib']	-3	Negatif
['kali', 'masyarakat', 'tolak', 'perpu', 'cipta', 'kerja', 'pimpin', 'abai', 'tolak', 'rakyat', 'pimpin', 'tau', 'bijak', 'langsung', 'nasib', 'rakyat', 'serbaserbi', 'mmc', 'islamkaffah']	-6	Negatif
['tilik', 'perpu', 'cipta', 'kerja', 'buruh', 'buruk', 'buruh', 'seprti', 'sapi', 'perah', 'dimnfaatkn', 'tenaga', 'upah', 'diberikn', 'sesuai', 'bahkn', 'kerap', 'kali', 'pemotongn', 'upah', 'trtentu', 'nasib', 'rakyat', 'diprjuangkkn', 'serbaserbi', 'mmc', 'islamkaffah']	-3	Negatif
['nasib', 'buruh', 'kian', 'ujung', 'tanduk', 'sah', 'perpu', 'cipta', 'kerja', 'atur', 'tuang', 'pijak', 'perusahaanpemilik', 'modal', 'rakyat', 'sejahtera', 'tenaga', 'peras', 'ampun', 'serbaserbi', 'mmc', 'islamkaffah']	-2	Negatif

4. Klasifikasi Naïve Bayes

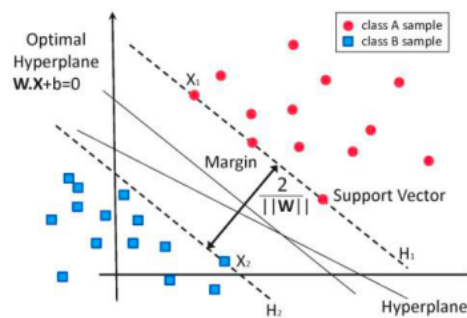
Naïve Bayes Classifier merupakan sebuah metode klasifikasi yang berdasarkan teori *bayes*. Salah satu contoh ciri metode ini yaitu dimana asumsi kondisi sangatlah kuat[14]. Dimana algoritma ini memanfaatkan teori probabilitas yang bisa memprediksi masa depan berdasarkan data latih sebelumnya.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

5. Klasifikasi Support Vector Machine (SVM)

Support Vector Machine sebuah algoritma untuk menentukan batas optimal antar kelompok data. SVM digunakan menemukan *hyperplane* terbaik dengan memaksimalkan gap antar kelompok data. Representasi SVM adalah ditunjukkan pada Gambar 2. *Hyperplane* adalah fungsi yang dapat digunakan untuk membagi antar kelas. Rumus optimal untuk nilai *hyperplane* dapat dilihat pada persamaan (2).

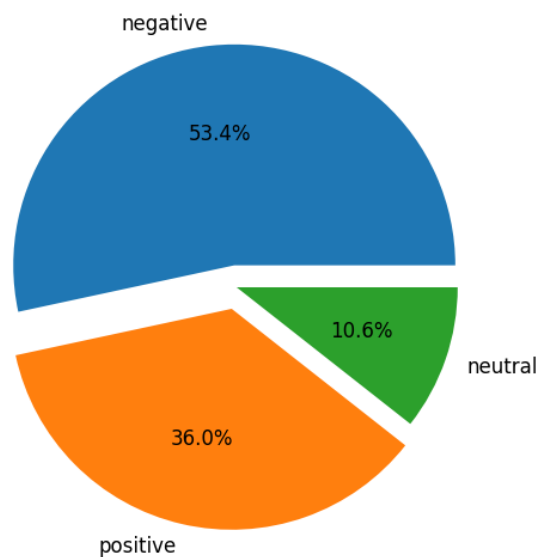
$$w \cdot x + b = 0 \quad (2)$$



Gambar 2. Hyperplane SVM

C. Hasil dan Pembahasan

Hasil dari proses *preprocessing* merupakan tweet dan pembobotan sentimen pada masing-masing tweet. Pembobotan kata dibedakan menjadi 3 bagian, yang pertama ada positif sentimen, negatif, dan juga sentimen netral. Dari masing-masing sentimen mendapatkan jumlah yang bermacam-macam. Negatif mendapatkan sebesar 53.4% dari 622, Positif mendapatkan 36% dari 622 dataset yang diambil, dan netral sebesar 10.6% dari total dataset 622. Masing-masing katagori diperoleh 332 sentimen negatif, 224 sentimen positif, dan 66 sentimen netral. Untuk melihat pembagian yang diperoleh bisa di lihat pada gambar 3.



Gambar 3. Diagram Lingkaran Pembagian Sentimen

a. Pengujian Algoritma

Hasil dari pengujian algoritma yang akan dilakukan adalah algoritma *Naïve Bayes* dan SVM. Berdasarkan data yang sudah dikumpulkan dan diujikan kedalam algoritma *Naïve Bayes* mendapatkan hasil untuk akurasi *training* sebesar 86.72 dengan akurasi testing 72.8. Untuk melihat secara detail hasil dari *Naïve Bayes* bisa dilihat pada tabel berikut.

Tabel 5. Naive Bayes Training Klasifikasi Report

	Precision	Recall	F1-score	support
Negatif	0.89	0.88	0.89	269
Netral	1.00	0.45	0.62	55
Positif	0.82	0.97	0.89	173
Accuracy			0.87	497
Macro avg	0.90	0.77	0.80	497
Weighted avg	0.88	0.87	0.86	497

Tabel 6. Naïve Bayes Testing Klasifikasi Report

	Precision	Recall	F1-score	support
Negatif	0.73	0.83	0.78	63
Netral	0.08	0.09	0.08	11
Positif	0.93	0.75	0.83	51
Accuracy			0.73	125
Macro avg	0.58	0.55	0.56	125
Weighted avg	0.75	0.73	0.74	125

Berdasarkan perhitungan algoritma SVM yang didapatkan perolehan tingkat akurasi pada *training score* sebesar 99.19. Untuk akurasi pada data *testing* diperoleh nilai akurasi sebesar 77.6. Untuk melihat lebih detail dari apa yang diperoleh maka akan ditampilkan pada tabel berikut.

Tabel 7. SVM Training Klasifikasi Report

	Precision	Recall	F1-score	support
Negatif	0.99	1.00	0.99	269
Netral	1.00	0.95	0.97	55
Positif	1.00	0.99	1.00	173
Accuracy			0.99	497
Macro avg	1.00	0.98	0.99	497
Weighted avg	0.99	0.99	0.99	497

Tabel 8. SVM Testing Klasifikasi Report

	Precision	Recall	F1-score	support
Negatif	0.77	0.89	0.82	63
Netral	0.31	0.36	0.33	11
Positif	0.95	0.73	0.82	51
Accuracy			0.78	125
Macro avg	0.63	0.66	0.66	125
Weighted avg	0.80	0.78	0.78	125

D. Simpulan

Berdasarkan hasil dari pelatihan data yang didapatkan sentimen Twitter sebanyak 622 data. Dari data yang diperoleh tersebut kemudian dibagikan menjadi beberapa jenis sentimen atau *polarity* yaitu positif, negatif, dan netral. Klasifikasi sentimen tersebut memperoleh nilai 332, 224, dan 66 sentimen. Dari terjadinya keseimbangan data yang didapatkan tersebut bisa saja menyebabkan kurang tingginya akurasi yang didapatkan. Nilai yang didapatkan dari algoritma *Naïve Bayes* sebesar 73% dan SVM sebesar 78%. Data tersebut menyatakan bahwa SVM mempunyai tingkat ketepatan yang lebih tinggi daripada algoritma *Naïve Bayes*.

Dari banyaknya data yang diperoleh terbukti sentimen yang paling besar atau paling banyak berpendapat tentang UU cipta kerja adalah negatif.

Setelah mengetahui topik yang diangkat mendapatkan responden atau sentimen yang terbanyak adalah negatif hal ini bisa menjadi salah satu penyebab tidak mendukungnya masyarakat mengenai UU cipta kerja dan bisa saja menjadi salah satu upaya pemerintah memberi kesepahaman mengenai pendapat tentang UU cipta kerja tersebut.

Penelitian ini masih terdapat kekurangan pada data penelitian ini masih kurang banyak untuk mendukung menjadikan hasil akurasi yang maksimal dan pada pembagian data sentimen pada penelitian ini masih tidak seimbang, sehingga data *training* dan data *testing* pada pembagian setiap sentimen kurang memberikan hasil yang maksimal, kemungkinan yang akan terjadi apabila pembagian data sentimen tersebut sama rata dari semua sentimen maka akan memberikan hasil akurasi algoritma yang lebih baik.

E. Referensi

- [1] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter".
- [2] M. Rangga, A. Nasution, and M. Hayaty, "Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter," *JURNAL INFORMATIKA*, vol. 6, no. 2, pp. 212–218, 2019, [Online]. Available: <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>
- [3] I. Yunanto and S. Yulianto, "TWITTER SENTIMENT ANALYSIS PEDULILINDUNGI APPLICATION USING NAÏVE BAYES AND SUPPORT VECTOR MACHINE," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 4, pp. 807–814, Aug. 2022, doi: 10.20884/1.jutif.2022.3.4.292.
- [4] R. Wati and S. Ernawati, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," 2021. [Online]. Available: <https://netlytic.org>
- [5] C. F. Hasri and D. Alita, "PENERAPAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN TERHADAP DAMPAK VIRUS CORONA DI TWITTER," *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, vol. 3, no. 2, pp. 145–160, 2022, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/informatika>
- [6] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," vol. 6, no. 3, 2021, [Online]. Available: <http://situs.com>

-
- [7] S. Zahoor and R. Rohilla, "Twitter Sentiment Analysis Using Machine Learning Algorithms: A Case Study," in *Proceedings - 2020 International Conference on Advances in Computing, Communication and Materials, ICACCM 2020*, Institute of Electrical and Electronics Engineers Inc., Aug. 2020, pp. 194–199. doi: 10.1109/ICACCM50413.2020.9213011.
- [8] Y. Romadhoni, K. Fahmi, and H. Holle, "Analisis Sentimen Terhadap PERMENDIKBUD No.30 pada Media Sosial Twitter Menggunakan Metode Naive Bayes dan LSTM," vol. 7, no. 2, 2022.
- [9] P. Arsi and R. Waluyo, "ANALISIS SENTIMEN WACANA PEMINDAHAN IBU KOTA INDONESIA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)," vol. 8, no. 1, pp. 147–156, 2021, doi: 10.25126/jtiik.202183944.
- [10] H. Tuhuteru and A. Iriani, "Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine dan Naive Bayes Classifier," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 3, no. 3, pp. 394–401, Oct. 2018, doi: 10.30591/jpit.v3i3.977.
- [11] N. Hafidz, S. Anggraeni, W. Gata, M. Ilmu Komputer STMIK Nusa mandiri Jakarta, and T. Komputer STMIK Nusa Mandiri Jakarta, "Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naive Bayes."
- [12] D. Atika, A. Ari Aldino, S. Informasi, J. Pagar Alam No, L. Ratu, and K. Kedaton, "TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP TEKANAN MENTAL PADA MEDIA SOSIAL TWITTER," 2022. [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [13] A. P. Wibawa, M. Guntur, A. Purnama, M. Fathony Akbar, and F. A. Dwiyanto, "Metode-metode Klasifikasi," *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, vol. 3, no. 1, 2018.