

ANALISIS SENTIMEN TERHADAP PROGRAM MERDEKA BELAJAR – KAMPUS MERDEKA PADA TWITTER MENGGUNAKAN *SUPPORT VECTOR MACHINE (SVM)*

Ni'ma Kholila¹⁾

¹⁾Fakultas Teknologi Informasi, Universitas Islam Balitar Blitar
e-mail: lilaae9@gmail.com¹⁾

Abstrak : Program Merdeka Belajar - Kampus Merdeka sebagai salah satu kebijakan publik oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, tidak lepas dari opini publik. Arah opini terbagi dalam tiga kategori, opini positif, opini negatif, dan opini netral. Publik menyampaikan opini melalui berbagai platform media sosial. Sebagai salah satu media sosial yang sederhana, Twitter justru memiliki pengaruh terhadap proses pembentukan dan pengarahannya opini publik. Memanfaatkan data tweet dengan keyword Merdeka Belajar - Kampus Merdeka antara Agustus 2020 sampai dengan Februari 2021, analisis sentimen dilakukan untuk mengenali opini publik terhadap Program Merdeka Belajar - Kampus Merdeka. Melalui penelitian ini diharapkan opini, faktor opini, hingga permasalahan yang mungkin timbul dari penerapan kebijakan Merdeka Belajar – Kampus Merdeka dapat dikenali sedini mungkin. Support Vector Machine digunakan untuk melakukan klasifikasi arah sentimen seseorang terhadap Program Merdeka Belajar – Kampus Merdeka, baik sentimen positif maupun sentimen negatif. Support Vector Machine dipilih karena mampu memberikan solusi yang konstan dengan performa yang baik. Nilai precision yang tinggi menunjukkan kecilnya nilai False Positive (terprediksi positive, sebenarnya negative). Hasilnya nilai recall yang tinggi menunjukkan kecilnya nilai False Negative (terprediksi negative, sebenarnya positive). Nilai F-measure yang baik mengindikasikan bahwa model klasifikasi sentimen mempunyai precision dan recall yang baik. Secara umum dapat disimpulkan bahwa nilai akurasi untuk masing-masing kelas dilihat dari nilai F-measure, kelas positif sebesar 94.8% dan kelas negatif sebesar 95%.

Kata Kunci—Analisis Sentimen, Merdeka Belajar - Kampus Merdeka, Support Vector Machine, Twitter

Abstract : Merdeka Belajar - Kampus Merdeka Program as one of the public policies by the Ministry of Education, Culture, Research, and Technology, cannot be separated from public opinion. Opinions are divided into three categories, positive opinions, negative opinions, and neutral opinions. The public expresses opinions through various social media platforms. As a simple social media, Twitter actually has an influence on the process of forming and directing public opinion. Tweet data with the keyword Merdeka Belajar - Kampus Merdeka between August 2020 and February 2021, is used to conduct sentiment analysis to identify the direction of public opinion towards Merdeka Belajar - Kampus Merdeka Program. Through this research, hoped that opinions, opinion factors, and problems that might happen from the implementation of Merdeka Belajar - Kampus Merdeka Program can be identified as soon as possible. The Support Vector Machine is used to classify the direction of a person's sentiment towards Merdeka Belajar - Kampus Merdeka Program, both positive and negative sentiments. The Support Vector Machine was chosen because it is able to provide a constant solution with a good performance. The final, a high precision value indicates a small False Positive value (predicted positive, actually negative). A high recall value indicates a small False Negative value (predicted negative, actually positive). A good F-measure value indicates that the sentiment classification model has good precision and recall. In general, it can be concluded that the accuracy value for each class is seen from the F-measure value, the positive class is 94.8% and the negative class is 95%.

Keywords—Sentiment Analysis, Merdeka Belajar - Kampus Merdeka, Support Vector Machine, Twitter

I. PENDAHULUAN

KEMENTERIAN Pendidikan, Kebudayaan, Riset, dan Teknologi meluncurkan kebijakan Merdeka Belajar. Merdeka Belajar ditujukan bagi pendidikan tinggi dengan tajuk Kampus Merdeka. Melalui kebijakan Kampus Merdeka, Nadiem berharap Perguruan Tinggi terdorong untuk lebih adaptif, inovatif, dan menghilangkan paradigma bahwa pendidikan hanya tanggung jawab satuan pendidikan [1]. Setiap kebijakan publik tidak lepas dari opini publik. Arah opini terbagi menjadi tiga kategori, opini positif, opini negatif, dan opini netral [2]. Demikian penerapan kebijakan Merdeka Belajar

– Kampus Merdeka tidak lepas dari opini pro, opini kontra, dan opini netral. Sebagai salah satu media sosial yang sederhana, *twitter* justru memiliki pengaruh yang tinggi terhadap proses pembentukan dan pengarahannya opini publik. *Twitter* dianggap sebagai tempat yang relatif ideal untuk menyampaikan opini karena banyak orang berpengaruh di dunia menggunakan *twitter*. Bahkan pemimpin negara hingga pemimpin keagamaan memiliki akun resmi *twitter* [3].

Analisis sentimen diantaranya digunakan terhadap kebijakan penggunaan transportasi umum darat dalam kota. Selain untuk mengetahui opini publik terhadap pelayanan transportasi umum, analisis sentimen juga digunakan untuk mengetahui faktor opini yang sering muncul. Selanjutnya analisis opini dapat digunakan untuk proses evaluasi penggunaan transportasi umum darat dalam kota [4].

Ada pun analisis sentimen terkait kenaikan tarif BPJS Kesehatan. Sebagai salah satu Badan Usaha Milik Negara dengan jumlah pengguna yang besar, tentu diperlukan layanan umpan balik untuk mengetahui respon atau opini pengguna terhadap layanan BPJS [5]. Dengan mengenali arah opini publik terhadap suatu kebijakan sekaligus mengetahui faktor opininya, diharapkan permasalahan yang mungkin timbul dari kebijakan publik dapat dikenali sedini mungkin. Melalui penelitian ini diharapkan opini, faktor opini, hingga permasalahan yang mungkin timbul dari penerapan kebijakan Merdeka Belajar – Kampus Merdeka dapat dikenali sedini mungkin. Meskipun tergolong relatif baru, *Support Vector Machine* (SVM) merupakan teknik klasifikasi yang populer, selalu mencapai solusi yang konstan untuk setiap *running*, dan yang tidak kalah penting performa yang baik selalu ditampilkan SVM dalam melakukan proses klasifikasi [6]. Sehingga untuk proses analisis sentimen terhadap program Merdeka Belajar – Kampus Merdeka, peneliti menggunakan SVM untuk melakukan klasifikasi sentimen.

II. TINJAUAN PUSTAKA

A. Merdeka Belajar – Kampus Merdeka

Merdeka Belajar – Kampus Merdeka merupakan salah satu kebijakan baru yang dicetuskan oleh Menteri Pendidikan dan Kebudayaan, Nadiem Makariem. Salah satu program dari kebijakan Merdeka Belajar – Kampus Merdeka adalah Hak Belajar Tiga Semester di Luar Program Studi. Program tersebut merupakan amanah dari berbagai regulasi/landasan hukum pendidikan tinggi dalam rangka peningkatan mutu pembelajaran dan lulusan pendidikan tinggi [7].

Kebijakan Merdeka Belajar - Kampus Merdeka diharapkan dapat menjadi jawaban atas tuntutan tersebut. Kampus Merdeka merupakan wujud pembelajaran di perguruan tinggi yang otonom dan fleksibel sehingga tercipta kultur belajar yang inovatif, tidak mengekang, dan sesuai dengan kebutuhan mahasiswa [8].

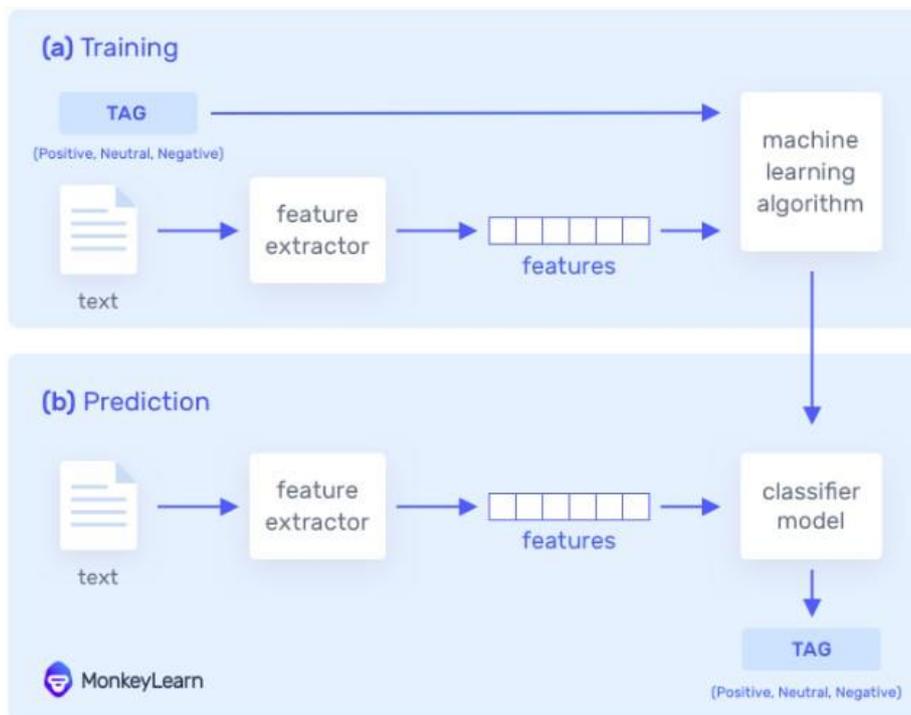
Pihak-pihak yang terkait dalam kebijakan ini antara lain Perguruan Tinggi, Fakultas, Program Studi, Mahasiswa, dan Mitra. Bentuk kegiatan pembelajaran sesuai dengan Permendikbud No 3 Tahun 2020 Pasal 15 ayat 1 dapat dilakukan di dalam Program Studi dan di luar Program Studi meliputi pertukaran pelajar, magang/praktik kerja, asistensi mengajar di satuan pendidikan, penelitian/riset, proyek kemanusiaan, kegiatan wirausaha, studi/proyek independen, dan membangun desa/kuliah kerja nyata tematik [7].

B. Analisis Sentimen

Analisis Sentimen (*sentiment analysis*) atau dikenal sebagai *opinion mining* adalah studi komputasi mengenai sikap, emosi, pendapat, dan penilaian dari sekumpulan teks yang berfokus pada ekstraksi, identifikasi, atau penemuan karakteristik sentimen menggunakan pendekatan *Natural Language Processing* (NLP), statistik, maupun *machine learning*.

Analisis sentimen dapat dilakukan dalam tiga pendekatan:

1. *Rule-based*, menggunakan aturan yang dibuat oleh manusia untuk membantu proses identifikasi subjektivitas, polaritas, atau subjek opini.
2. *Automatic*, tidak bergantung pada aturan yang dibuat secara manual, melainkan menggunakan teknik *machine learning* yang dimodelkan sebagai masalah klasifikasi (positif, negatif, atau netral).
3. *Hybrid*, menggabungkan elemen *rule-based* dan *automatic* dalam satu sistem.



Gambar 1. Analisis sentimen dengan teknik *machine learning*

Gambar 1 menunjukkan proses analisis sentimen dengan pendekatan *automatic* dengan teknik *machine learning* [9]:

1. *Training*, dimulai dengan model mempelajari hubungan antara input berupa teks dengan output berbentuk *tag* yang sesuai berdasarkan sampel pengujian yang digunakan untuk *training*. *Feature extractor* mengirim input teks ke dalam *feature vector*. Pasangan *feature vector* dan *tag* (positif, negatif, atau netral) diproses menggunakan *machine learning algorithm* untuk menghasilkan model.
2. *Prediction*, dimulai dengan *feature extractor* yang digunakan untuk mengubah input teks yang tidak terlihat menjadi *feature vector*. *Feature vector* ini kemudian dimasukkan ke dalam model untuk menghasilkan output berupa prediksi *tag* (positif, negatif, atau netral).
3. *Feature Extraction*, dilakukan dengan pendekatan *bag-of-words* atau *bag-of-ngrams*. Dapat pula diterapkan berdasarkan *word embeddings* (*word vectors*). Pendekatan lain yang bisa dilakukan adalah dengan TF-IDF (*term frequency – inverse document frequency*).
4. *Classification Steps*, melibatkan penggunaan model statistik yang umum digunakan seperti *Naïve Bayes*, *Linear Regression*, *Support Vector Machines*, atau *Neural Networks*.

C. *Term Frequency – Inverse Document Frequency (Tf-idf)*

Tf-idf merupakan salah satu implementasi *feature extraction* untuk mengekstrak kata dalam suatu kalimat dan dokumen [10]. *Tf-idf* merupakan skema pembobotan kata dalam suatu kalimat. *Tf-idf* digunakan untuk menggantikan nilai kemunculan kata dalam kalimat. Tujuannya untuk mencegah kemungkinan munculnya kata berulang dan justru mengurangi nilai informatif dari suatu kalimat. *Tf-idf* menghitung berapa kali sebuah kata muncul dalam teks (*term frequency*) dan membandingkannya dengan seberapa umum kata tersebut muncul di seluruh dokumen (*inverse document frequency*). Nilai *tf-idf* dapat dihitung menggunakan persamaan (1):

$$tfidf(t, d, D) = tf(t, d) \cdot idf(t, D) \quad (1)$$

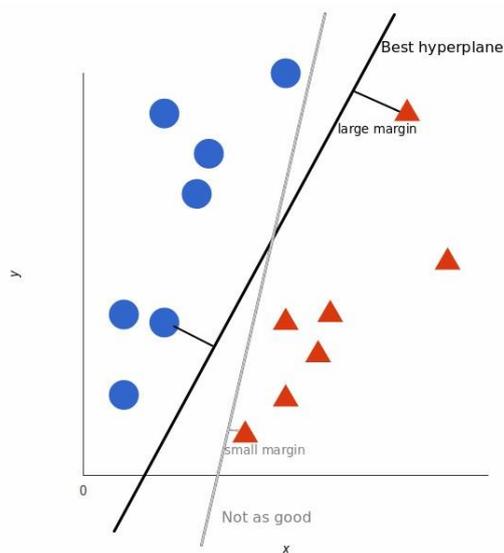
Di mana *t* adalah *term* atau kata, *d* adalah kalimat, dan *D* adalah kumpulan kalimat atau dokumen.

Semakin tinggi nilai *tf-idf*, maka semakin relevan kata tersebut terhadap suatu dokumen. Misal, pada kasus analisis sentimen, kata “baik” memiliki nilai *tf-idf* yang tinggi untuk dokumen berlabel *positive*,

maka kata “baik” dianggap paling relevan dan dapat digunakan sebagai kata kunci dalam dokumen berlabel *positive*. Demikian halnya untuk kata “kecewa” yang memiliki nilai *tf-idf* yang tinggi untuk dokumen berlabel *negative*, maka kata “kecewa” dianggap paling relevan dan dapat digunakan sebagai kata kunci dalam dokumen berlabel *negative* [10].

D. Support Vector Machine

Support Vector Machine (SVM) merupakan teknik *machine learning* yang relatif baru untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. SVM masuk dalam kategori *supervised learning* yang proses klasifikasinya menggunakan fungsi-fungsi linear [6]. SVM bekerja pada set obyek yang *linearly separable*. SVM bekerja dengan mencari dan menentukan nilai *hyperplane* (fungsi pemisah kelas) terbaik. Nilai *hyperplane* terbaik adalah nilai yang terletak di tengah-tengah antara dua set obyek dari dua kelas. Mencari *hyperplane* terbaik berbanding lurus dengan memaksimalkan *margin*. *Margin* adalah jarak tegak lurus antara *hyperplane* dengan obyek terdekat sebagaimana ditunjukkan oleh Gambar 2:



Gambar 2. Support Vector Machine

Sebagai salah satu fungsi linear [6], bentuk SVM secara umum dinyatakan dalam persamaan (2):

$$f(x) = w^T x + b \quad (2)$$

Di mana x adalah vektor input, w adalah parameter bobot, dan b adalah bias. Fungsi linear dengan nilai 0 disebut sebagai *decision boundary*. Fungsi linear dengan nilai +1 disebut sebagai *hyperplane positive* dan fungsi linear dengan nilai -1 disebut sebagai *hyperplane negative*. Berdasarkan fungsi hyperplane tersebut, margin dapat dihitung menggunakan persamaan (3):

$$\text{margin} = \frac{\frac{1}{2} * 2}{\|w\|} = \frac{1}{\|w\|} \quad (3)$$

Guna memaksimalkan nilai *margin*, dapat dilakukan dengan meminimalkan pembaginya yaitu $\|w\|$, sehingga bentuk optimasi *margin* sebagaimana dinyatakan dalam persamaan (4):

$$\underset{w,b}{\operatorname{argmin}} \frac{1}{2} \|w\|^2 \quad (4)$$

E. *F-measure*

F-measure pada prinsipnya digunakan pada *retrieval system* untuk mengukur klasifikasi pencarian dokumen serta performa *query* klasifikasi. Sebelumnya, *F-measure* hanya difokuskan pada perhitungan nilai. Namun seiring perkembangannya *F-measure* lebih menekankan pada kinerja *precision* dan *recall* [6]. *F-measure* dinyatakan dalam persamaan (5):

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

Precision dinyatakan dalam persamaan (6):

$$Precision = \frac{TruePositive}{FalsePositive + TruePositive} \quad (6)$$

Recall dinyatakan dalam persamaan (7):

$$Recall = \frac{TruePositive}{FalseNegative + TruePositive} \quad (7)$$

True Positive jika output kelas positif diprediksi sebagai kelas positif. *False Positive* jika output kelas negatif diprediksi sebagai kelas positif. *False Negative* jika output kelas positif diprediksi sebagai kelas negatif [6].

III. METODE PENELITIAN

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data *tweet* yang diambil menggunakan *keyword* “merdeka belajar”, “kampus merdeka”, “merdeka belajar kampus merdeka”, “mbkm”, dan “program mbkm”. *Scrapping data* dilakukan dengan mengambil data *tweet* antara bulan Agustus 2020 sampai dengan Februari 2021. Penelitian ini hanya menggunakan data *tweet* berbahasa Indonesia. *Scrapping data* dilakukan pada <https://twitter.com/> menggunakan *wrapper* melalui *Puppeteer library* dan *test request* yang dikirim melalui *Postman*.

B. Data Preprocessing

Tahap *preprocessing* dilakukan setelah tahap pengumpulan data melalui proses *scrapping*. *Data preprocessing* meliputi *cleaning* dan *labeling*. *Cleaning* untuk menghilangkan *tweet* berita, *tweet* pengumuman, *tweet* pertanyaan, dan *tweet* mengandung *link*. *Labeling* untuk memberikan label pos untuk sentimen bernilai positif dan label neg untuk sentimen bernilai negatif.

C. Analisis Sentimen

Analisis sentimen dilakukan berdasarkan tahapan-tahapan sesuai Gambar 1, yaitu *training*, *prediction*, *feature extraction*, dan *classification*. Dalam penelitian ini, tahap *classification* menggunakan algoritma *support vector machine*. Prinsip kerja *support vector machine* yang akan digunakan pada penelitian ini mengacu pada Gambar 2. Analisis sentimen dilakukan menggunakan bahasa pemrograman *python*. Adapaun *tools* yang digunakan untuk analisis sentimen adalah *python3*, *jupyter notebook*, dan *scikit-learn library* (berserta *dependency library*).

D. Pengujian

Akurasi hasil analisis sentimen dinilai menggunakan *F-measure*. *F-measure* menekankan penilaian akurasi pada nilai *precision* dan nilai *recall*. Nilai *F-measure*, nilai *precision*, dan nilai *recall* masing-masing dihitung menggunakan persamaan (5), persamaan (6), dan persamaan (7).

IV. HASIL DAN PEMBAHASAN

A. Paparan Data

Data *tweet* yang diambil menggunakan *keyword* “merdeka belajar”, “kampus merdeka”, “merdeka belajar kampus merdeka”, “mbkm”, dan “program mbkm” antara bulan Agustus 2020 sampai dengan Februari 2021. Data yang diperoleh sebanyak 4149 *tweet*. Data tersebut merupakan *raw data* termasuk di dalamnya *tweet* berita, *tweet* pengumuman, *tweet* pertanyaan, dan *tweet* mengandung *link*. Gambar 3 merupakan contoh data *tweet* yang diperoleh pada tahap pengumpulan data:

	Text	Timestamp	ID	Username	Retweets	Likes
3915	[cm] ak mw tny dong, especially kakak2 pgsd. k...	2021-11-04 22:47:50	1456287079302713346	@collegemenfess	0.0	1.0
2796	dari kampus merdeka kita belajar, mending sarj...	2021-10-26 11:38:34	1452857160988172290	@DimasCr96	0.0	2.0
2034	Kampus apa rentenir. Udah kek punya utang aja ...	2021-11-23 21:20:05	1463150364929118209	@imfinegodthanks	0.0	0.0
944	Gimana ya sbnrnya dilema udah pada pengen offl...	2021-11-09 16:32:31	1458004565022756864	@bendiatmaja	2.0	3.0
3572	#InfoGenWaniPooll Sebuah petisi meminta Mendik...	2021-11-18 19:30:00	1461310720461574149	@gen1031fmsby	0.0	0.0

Gambar 3. Contoh data *tweet*

B. Data Preprocessing

Data preprocessing terhadap data *tweet* meliputi *cleaning* dan *labeling*. Proses *cleaning* dengan menghapus *tweet* berita, *tweet* pengumuman, *tweet* pertanyaan, dan *tweet* mengandung *link* menyisakan data siap pakai sebanyak 1431 *tweet*. *Labeling* atau pemberian *label* dilakukan secara manual oleh peneliti. Diantara data tersebut, 711 *tweet* merupakan *tweet* dengan label *positive* dan 720 *tweet* merupakan *tweet* dengan label *negative*. Gambar 4 merupakan hasil data *preprocessing* yang siap digunakan untuk kebutuhan analisis sentimen:

	Text	Label
47	anjir halo kampus gw.lngw udah smt 7 kampus me...	neg
70	bingung mau seneng apa sedih gara-gara lolos p...	pos
113	gara-gara webinar kemarin tiap ada orang bilan...	pos
429	Tanpa kampus merdeka sebenarnya stat tuh udah ...	neg
48	Anjirrr ada adik tingkat merdeka belajar dari...	neg

Gambar 4. Data *tweet* hasil *preprocessing*

C. Analisis Sentimen

Pada tahap analisis sentimen, terdapat beberapa *library python* yang digunakan. Gambar 5 menunjukkan *library python* yang digunakan pada penelitian ini:

```

1 import pandas as pd
2 import numpy as np
3 from sklearn.feature_extraction.text import TfidfVectorizer
4 import time
5 from sklearn import svm
    
```

Gambar 5. *Library Python*

Library *pandas* digunakan untuk mengambil data train yang berasal dari file berekstensi *.csv*. Library *numpy.random* digunakan untuk mengacak *data test* yang akan digunakan. Library *sklearn.feature_extraction.text.TfidfVectorizer* digunakan untuk tahap *feature extraction*. Library *time* digunakan untuk menghitung lama waktu tahap *training* dan lama waktu tahap *testing*. Library *sklearn.svm.SVC* digunakan untuk tahap *classification* menggunakan *Support Vector Machine*.

Gambar 6 menunjukkan potongan kode yang menggunakan *library pandas* (*pd*) untuk membaca data untuk selanjutnya dijadikan sebagai *data train* dan *data test*. Pengacakan data test dilakukan oleh *library numpy* (*np*).

```
1 df = pd.read_csv('fix.csv')
2 df['split'] = np.random.randn(df.shape[0], 1)
3 msk = np.random.rand(len(df)) <= 0.15
4 train = df
5 test = df[msk]
```

Gambar 6. Potongan kode *read data*

```
1 vectorizer = TfidfVectorizer(min_df = 2,
2                             max_df = 0.8,
3                             sublinear_tf = True,
4                             use_idf = True)
5
6 train_vectors = vectorizer.fit_transform(train['Text'])
7 test_vectors = vectorizer.transform(test['Text'])
```

Gambar 7. Potongan kode *feature extraction*

Gambar 7 menunjukkan potongan kode yang menggunakan *library TfidfVectorizer* untuk merubah *data train* dan *data test* dalam bentuk *matrix tf-idf*. Bentuk transformasi disimpan dalam variabel *train_vectors* untuk *data train* dan variabel *test_vectors* untuk *data test*. Library *TfidfVectorizer* menggunakan parameter. Parameter *min_df* menunjukkan nilai ambang bawah (*cut-off*) *df*. Parameter *max_df* menunjukkan nilai ambang batas (*threshold*) *df*. Parameter *sublinear_tf* menunjukkan bahwa *tf scling* diaplikasikan secara *sublinear*. Parameter *use_idf* menunjukkan bahwa akan dilakukan pembobotan ulang terhadap *idf*.

```
1 classifier_linear = svm.SVC(kernel='linear')
2 t0 = time.time()
3 classifier_linear.fit(train_vectors, train['Label'])
4 t1 = time.time()
5 prediction_linear = classifier_linear.predict(test_vectors)
6 t2 = time.time()
7 time_linear_train = t1-t0
8 time_linear_predict = t2-t1
```

Gambar 8. Potongan kode *classification* menggunakan SVM

Gambar 8 menunjukkan potongan kode yang menggunakan *library SVM* untuk melakukan tahap *classification*. Pada baris pertama ditunjukkan bahwa *classification* menggunakan SVM menerapkan *linear kernel* untuk proses *classification*. Proses *classification* dilakukan pada *train_vectors* (hasil *feature extraction*) terhadap label pada masing-masing dataset. Proses *prediction* dilakukan berdasarkan hasil proses *classification* terhadap *test_vectors* (hasil *feature extraction*). Variabel *time_linear_train* digunakan untuk menghitung lama waktu tahap *training*. Variabel *time_linear_predict* digunakan untuk menghitung lama waktu tahap *testing*. Selanjutnya hasil perhitungan waktu ditampilkan bersama dengan *classification report*.

```
1 review = "alhamdulillah akhirnya bisa ikut mbkm, bakal dapet duit jajan nih"
2 review_vector = vectorizer.transform([review])
3 print(classifier_linear.predict(review_vector))

['pos']

1 review = "mbkm apa gak jelas banget sih. ruwit, telat bayar pula"
2 review_vector = vectorizer.transform([review])
3 print(classifier_linear.predict(review_vector))

['neg']
```

Gambar 9. Potongan kode prediksi sentimen

Gambar 9 menunjukkan potongan kode bagaimana proses prediksi terhadap suatu sentimen. Kalimat yang disimpan pada variabel *review* akan ditransformasikan menggunakan *library* *TfidfVectorizer* untuk selanjutnya disimpan sebagai variabel *review_vector*. Variabel *review_vector* yang selanjutnya dilakukan prediksi menggunakan *library* *SVM*. Hasil prediksi sentimen berupa label *pos* untuk sentimen positif dan label *neg* untuk sentimen negatif.

D. Pengujian

Pada tahap pengujian *library python* yang digunakan adalah *library* *sklearn.metrics.classification_report*. *Library* tersebut digunakan untuk menghitung nilai *precision*, *recall*, dan *F-measure* untuk masing-masing kelas (kelas positif maupun kelas negatif). Gambar 10 menunjukkan potongan kode tahap pengujian:

```
1 from sklearn.metrics import classification_report
2
3 print("Training time: %fs; Prediction time: %fs" % (time_linear_train, time_linear_predict))
4 report = classification_report(test['Label'], prediction_linear, output_dict=True)
5 print('pos:', report['pos'])
6 print('neg:', report['neg'])
```

Training time: 0.132716s; Prediction time: 0.005528s
pos: {'precision': 0.9487179487179487, 'recall': 0.9487179487179487, 'f1-score': 0.9487179487179487, 'support': 39}
neg: {'precision': 0.95, 'recall': 0.95, 'f1-score': 0.9500000000000001, 'support': 40}

Gambar 10. Potongan kode classification report

Selain menunjukkan hasil pengujian terhadap kelas positif dan kelas negatif, Gambar 10 juga menunjukkan waktu yang diperlukan untuk proses training (*training time*) dan waktu yang diperlukan untuk proses *prediction* (*prediction time*). Tabel 1 menunjukkan lama waktu yang diperlukan untuk proses *training* dan proses *prediction*:

Tabel 1. Waktu proses

Proses	Training	Prediction
Time (s)	0.132716	0.005528

Secara lebih ringkas, Tabel 2 menunjukkan nilai *precision*, *recall*, dan *F-measure* untuk masing-masing kelas:

Tabel 2. Hasil pengujian

Uji	Kelas Positif	Kelas Negatif
Precision	0.948	0.95
Recall	0.948	0.95

Uji	Kelas Positif	Kelas Negatif
<i>F-measure</i>	0.948	0.95

Nilai *precision* yang tinggi menunjukkan kecilnya nilai *False Positive* (terprediksi *positive*, sebenarnya *negative*). Nilai *recall* yang tinggi menunjukkan kecilnya nilai *False Negative* (terprediksi *negative*, sebenarnya *positive*). Nilai *F-measure* yang baik mengindikasikan bahwa model klasifikasi sentimen mempunyai *precision* dan *recall* yang baik.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil pembahasan, dapat disimpulkan bahwa analisis sentimen menggunakan Support Vector Machine dapat diterapkan untuk memprediksi arah sentimen seseorang terhadap Program Merdeka Belajar – Kampus Merdeka, baik sentimen positif maupun sentimen negatif. Nilai *precision* yang tinggi menunjukkan kecilnya nilai *False Positive* (terprediksi *positive*, sebenarnya *negative*). Nilai *recall* yang tinggi menunjukkan kecilnya nilai *False Negative* (terprediksi *negative*, sebenarnya *positive*). Nilai *F-measure* yang baik mengindikasikan bahwa model klasifikasi sentimen mempunyai *precision* dan *recall* yang baik. Secara umum dapat disimpulkan bahwa nilai akurasi untuk masing-masing kelas dilihat dari nilai *F-measure*, kelas positif sebesar 94.8% dan kelas negatif sebesar 95%.

Pengembangan penelitian yang mungkin dikerjakan adalah dengan melakukan analisis sentimen terhadap kebijakan publik yang lain untuk selanjutnya dimanfaatkan sebagai bahan penilaian dan evaluasi. Penggunaan teknik klasifikasi lain masih sangat mungkin untuk digunakan. Guna memperoleh nilai akurasi yang lebih baik, penelitian selanjutnya dapat menggunakan lebih banyak *dataset*.

DAFTAR PUSTAKA

- [1] Prodjo, W.A. (2020, Januari 25). 4 Alasan Nadiem Makarim Mengeluarkan kebijakan kampus Merdeka. *Kompas Online*. [Online]. Tersedia: <https://edukasi.kompas.com/read/2020/01/25/20283891/4-alasan-nadiem-makarim-mengeluarkan-kebijakan-kampus-merdeka>.
- [2] Irawan, P. (2017). Analisis Opini Publik tentang Kualitas Pelayanan Publik Pemerintah Kota Palembang dalam Rubrik “Lapor Mang Sripo” pada Surat Kabar Sriwijaya Post. [Online]. Tersedia: <http://eprints.radenfatah.ac.id/869/1/PURNAMA%20IRAWAN%2012530064%20JURNALIS%20TIK.pdf>.
- [3] Zaenudin, A. (2018, Maret 21). Bagaimana Twitter Mempengaruhi Opini Publik dan Preferensi Politik. *Tirto-id*. [Online]. Tersedia: <https://tirto.id/bagaimana-twitter-mempengaruhi-opini-publik-dan-preferensi-politik-cGre>.
- [4] Novantirani, A., Sabariah, M.K., Effendy, V., “Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Dalam Kota dengan Metode Support Vector Machine,” dalam e-Proceeding of Engineering : Vol.2, No.1 April 2015, hal. 1177.
- [5] Dharmapatni, P.M.N., & Merawati, N.L.P. (September, 2020). Penerapan Algoritma Support Vector Machine Dalam Sentimen Analisis Terkait Kenaikan Tarif BPJS Kesehatan. *Jurnal BITE*. [Online]. Vol 2 No 2. Tersedia: <https://journal.universitاسbumigora.ac.id/index.php/bite/article/view/904>.
- [6] Santosa, B. & Umam, A. (2018). *Data Mining dan Big Data Analytics*. Bantul, Yogyakarta: Penebar Media Pustaka.
- [7] DIKTI. (2020). *Buku Panduan Merdeka Belajar – Kampus Merdeka*. Jakarta: Direktorat Jendral Pendidikan Tinggi Kemdikbud RI.
- [8] Aditya, D.B. (2021, Oktober 30). Tantangan dan Harapan FIB dalam Mendukung Merdeka Belajar Kampus Merdeka. *Unair News*. [Online]. Tersedia: <http://news.unair.ac.id/2021/10/30/tantangan-dan-harapan-fib-dalam-mendukung-merdeka-belajar-kampus-merdeka/>

- [9] MonkeyLearn. (n.d.) Understanding TF-IDF: A Simple Introduction. [Online]. Tersedia: <https://monkeylearn.com/blog/what-is-tf-idf>.
- [10] MonkeyLearn. (n.d.) Sentiment Analysis: the Definitive Guide. [Online]. Tersedia: <https://monkeylearn.com/sentiment-analysis>.