

## PENENTUAN RESPON MASYARAKAT TERHADAP KEBIJAKAN PILKADA PADA MASA PANDEMI COVID-19 MENGGUNAKAN METODE ALGORITMA NAÏVE BAYES CLASSIFIER (NBC) DAN SUPPORT VECTOR MACHINE (SVM)

Arya Arief Budiman<sup>1</sup>, Mujiono Sadikin<sup>2</sup>, Arie Jaenul<sup>3</sup>

Program Studi Teknik Informatika. Fakultas Ilmu Komputer. Universitas Mercu Buana<sup>1,2</sup>  
Jurusan Teknik Elektro. Fakultas Teknik dan Ilmu Komputer. Jakarta Global University<sup>3</sup>  
[411517010015@student.mercubunana.ac.id](mailto:411517010015@student.mercubunana.ac.id)<sup>1</sup>, [Mujiono@mercubuana.ac.id](mailto:Mujiono@mercubuana.ac.id)<sup>2</sup>, [ariep@jgu.ac.id](mailto:ariep@jgu.ac.id)<sup>3\*</sup>

---

### ABSTRACT

---

**Purpose:** *The purpose of this study is to identify the sentiment analysis of the public response to the Indonesian Government's policy to carry out regional elections (pilkada) during the Covid-19 pandemic using the Naïve Bayes Classifier (NBC) algorithm and the Support Vector Machine (SVM) method.*

**Research methodology:** *The research method used in this study is to use quantitative research methods. The data used in this study were taken from public comments on a tweet in a Twitter post that is stored in the .csv format.*

**Results:** *The results obtained from this study are to compare 2 (two) algorithms, namely Naïve Bayes and SVM into 3 test scenarios. The test results show that the accuracy value obtained by SVM is much better than Naïve Bayes with the value in the first scenario Naïve Bayes 76%: 88% SVM, the second scenario Naïve Bayes 76%: 88% SVM, and the third scenario Naïve Bayes 78%: 90 % SVM.*

**Limitations:** *There are several limitations in this study, such as the data used only comes from the Twitter platform, the data used only focuses on Indonesian language posts, and only 2 (two) sentiment classification class labels are used, namely positive and negative.*

**Contribution:** *This research can be used as a reference by the General Election Commission (KPU) to determine election policies for the regions in the future during the Covid-19 period. This research is included in the disciplines of data mining and machine learning.*

**Keywords:** *Pilkada, Twitter, Naïve Bayes Classifier (NBC), Support Vector Machine (SVM)*

---

### ABSTRAK

---

**Tujuan:** Tujuan dari penelitian ini adalah untuk mengidentifikasi sentimen analisis respon masyarakat terhadap kebijakan Pemerintah Indonesia untuk melaksanakan pemilihan kepada daerah (pilkada) pada masa pandemic covid-19 dengan menggunakan metode algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM).

**Metodologi penelitian:** Metode penelitian yang digunakan dalam penelitian ini adalah dengan menggunakan metode penelitian kuantitatif. Data yang digunakan dalam penelitian ini diambil dari komentar-komentar masyarakat terhadap suatu *tweet* di postingan twitter yang disimpan dalam bentuk format .csv.

**Hasil:** Hasil yang didapatkan dari penelitian ini adalah dengan membandingkan 2 (dua) algoritma yaitu Naïve Bayes dan SVM kedalam 3 skenario pengujian. Hasil pengujian menunjukkan nilai akurasi yang didapatkan oleh SVM jauh lebih baik dibandingkan dengan Naïve Bayes dengan nilai pada skenario pertama Naïve Bayes 76%:88% SVM, skenario kedua Naïve Bayes 76%:88% SVM, dan Skenario ketiga Naïve Bayes 78%:90% SVM.

**Limitasi:** Terdapat beberapa limitasi dalam penelitian ini seperti data yang digunakan hanya bersumber dari *platform* twitter, data yang digunakan hanya berfokus pada postingan berbahasa Indonesia, dan label kelas klasifikasi sentiment yang digunakan hanya 2 (dua) yaitu positif dan negative.

**Kontribusi:** Penelitian ini dapat dijadikan sebagai rujukan oleh Komisi Pemilihan Umum (KPU) untuk menentukan kebijakan pemilihan kepada daerah di masa mendatang pada masa covid-19. Penelitian ini masuk ke dalam area disiplin ilmu data mining dan juga *machine learning*.

**Kata kunci:** Pilkada, Twitter, Naïve Bayes Classifier (NBC), Support Vector Machine (SVM)

---

## PENDAHULUAN

Sejak awal tahun 2020 sampai saat ini tidak ada satu orang pun yang dapat mengabaikan adanya pandemi Covid-19. Bukan hanya negara yang disibukkan oleh fenomena yang muncul tiba-tiba, tapi organisasi, bahkan individu. Semula banyak pihak yang menyangkal, menganggap apa yang terjadi di Wuhan (kota pertama kali terjadinya Covid-19) hanya bersifat lokal, dan dengan sendirinya menghilang seperti yang terjadi pada pengalaman-pengalaman sebelumnya.

Fenomena Covid-19 memicu timbulnya berbagai analisis tentang dampak yang terjadi serta respon yang perlu disiapkan, agar dunia lebih siap di masa yang akan datang untuk menghadapi guncangan serupa di masa depan. Berbagai upaya telah dilakukan untuk mengatasinya, dari inisiatif-inisiatif yang bersifat domestik, sampai kerjasama bilateral, namun belum menunjukkan hasil yang berarti (Irawan, 2020).

Kebijakan Pemilihan Kepala daerah (Pilkada) 2020 diadakan selama Pandemi COVID-19. Ada 270 kabupaten yang memilih Pemimpin Daerah. Fase Pilkada dimulai dari akhir Agustus hingga Desember 2020 (Azanella, 2020).

PILKADA serentak membuka pintu potensi penyebaran virus covid-19 yang lebih besar di masyarakat. Karena itu perlu ditunda hingga pandemi mereda, dan Pilkada bisa dilakukan secara bersamaan. Namun, disisi lain, "PILKADA" harus dilakukan agar mendapatkan pemimpin daerah dapat menangani pandemi secara paling efektif (Azanella, 2020).

Analisis Sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif. Salah satu contoh penggunaan analisis sentimen dalam dunia nyata adalah identifikasi kecenderungan pasar dan opini pasar terhadap suatu objek barang. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang

pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentiment (Liu, 2010).

Dari kebijakan pilkada yang dilakukan oleh pemerintah bahwa akan diadakannya pilkada serentak tahun 2020 pada masa pandemi covid-19, adanya respon dari masyarakat terhadap kebijakan ini. Untuk itu respon dari masyarakat yang dituangkan melalui postingan twitter baik sebelum dilaksanakannya pilkada maupun setelah diadakannya pilkada dijadikan sebagai penelitian untuk menentukan kebijakan yang akan datang.

## **TINJAUAN PUSTAKA DAN PENGEMBANGAN HIPOTESIS**

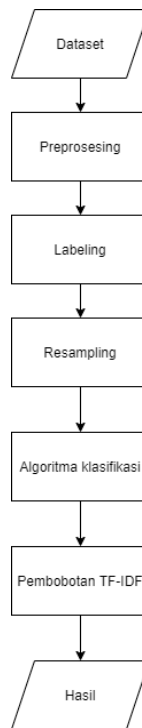
Beberapa penelitian terkait diantaranya adalah penelitian yang dilakukan oleh (Buntoro, 2017) dengan judul analisis sentimen calon Gubernur DKI Jakarta 2017 di Twitter dengan menggunakan metode Lexicon Based, Support Vector Machine, dan Naïve Bayes dengan hasil rata-rata akurasi yang didapatkan untuk metode naïve bayes adalah sebesar 74,60% dan untuk Support Vector Machine sebesar 75,62%. Berikutnya penelitian yang dilakukan oleh (Rofiqoh et al., 2017), penelitian ini menganalisis sentiment terhadap tingkat kepuasan pengguna penyedia layanan telekomunikasi seluler pada Twitter dengan metode Support Vector Machine dan Lexicon Based Features menghasilkan nilai accuracy sebesar 79%, precision sebesar 65%, recall sebesar 97%, dan f-measure sebesar 78%. Selanjutnya analisis sentimen tentang opini film pada dokumen twitter berbahasa Indonesia menggunakan Naïve Bayes dengan perbaikan kata tidak baku dilakukan oleh (Antinasari et al., 2017), berdasarkan hasil pengujian yang telah dilakukan didapatkan nilai accuracy, precision, recall, dan f-measure sebesar 98.33%, 96.77%, 100%, dan 98.36%. Dan (Buntoro, 2016), melakukan penelitian dengan menggunakan metode klasifikasi Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) dengan preprocessing data menggunakan tokenisasi, cleansing dan filtering. Data yang digunakan adalah tweet dalam bahasa Indonesia dengan tagar HateSpeech (#HateSpeech), dengan jumlah dataset sebanyak 522 tweet yang didistribusikan secara merata menjadi dua sentimen HateSpeech dan GoodSpeech. Hasil akurasi tertinggi didapatkan saat menggunakan metode klasifikasi Support Vector Machine (SVM) dengan tokenisasi unigram, stopword list Bahasa Indonesia dan emoticons, dengan nilai rata-rata akurasi mencapai 66,6%, nilai presisi 67,1%, nilai recall 66,7% nilai TP rate 66,7% dan nilai TN rate 75,8%. Metode Support Vector Machine digunakan oleh (Pravina et al., 2019) untuk menganalisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter dengan hasil accuracy sebesar 40%, precision 40%, 100% recall, dan f-measure sebesar 57,14%.

Dari beberapa penelitian diatas metode algoritma Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) dapat digunakan untuk menganalisis sentimen yang dilakukan oleh masyarakat dengan menggunakan media sosial salah satunya adalah dengan menggunakan twitter. Untuk itu metode tersebut digunakan dalam penelitian ini untuk mengidentifikasi sentimen analisis respon masyarakat terhadap kebijakan Pemerintah Indonesia untuk melaksanakan pemilihan kepada daerah (pilkada) pada masa pandemic covid-19.

## **METODE PENELITIAN**

Pada penelitian ini digunakan metode penelitian kuantitatif, dimana jenis penelitian ini muncul dalam bentuk angka atau mendeskripsikan fenomena atau fakta penelitian, atau mengumpulkan data secara mendalam dan lengkap untuk menjelaskan fenomena yang terjadi di masyarakat. penelitian ini juga mengadopsi bentuk eksplorasi yaitu data

yang diperoleh dari penelitian atau eksplorasi lapangan untuk memperoleh informasi lebih lanjut, penelitian ini meliputi studi tentang perilaku, persepsi, motivasi dan perilaku yang digambarkan dalam kata-kata dan bahasa dalam konteks yang natural. Dalam penelitian ini kelengkapan dan kedalaman data yang diteliti sangat penting. Penelitian ini pengumpulan datanya melalui pengambilan komentar-komentar yang diberikan masyarakat terhadap suatu tweet di postingan twitter yang dapat disimpan dalam bentuk format .csv. Diagram alir penelitian seperti pada gambar 1 dibawah ini.



Gambar 4. Diagram Alir Penelitian

#### a) Pengumpulan Data

Metode pengumpulan data yang digunakan untuk penelitian adalah dengan mengambil data tweet dari Twitter, hal ini dapat dilakukan dengan mengintegrasikan Twitter API dengan Rapid Miner, pengumpulan data di mulai sebelum pilkada dan sesudah pilkada, PILKADA 2020 terlaksana pada tanggal 9 Desember 2020, Rapidminer didasarkan pada query istilah objek dalam aplikasi pengolahan data mining yang terhubung ke Twitter API. dari Kicauan yang didapat berisi "upaya pemerintah melaksanakan kebijakan pilkada". Untuk tweet tentang PILKADA, kumpulan data kueri mencakup hasil tweet dari (#) dan tweet biasa. Tweet ini dipecah atau diubah menjadi beberapa kumpulan data, seperti kumpulan data #PilkadaBencana Serentak, Pilkada2020 (karena covid-19) dan kebijakan pemilihan kepala daerah pada saat covid - 19. Dan tabel 1 dibawah ini merupakan jumlah data tweets terkait dengan pilkada 2020.

Tabel 5.Data Twitter

Nama Data	Data Tweets
Pilkada	39077

### b) Pra Pemrosesan Data (*Preprocessing*)

Pada tahap preprocessing, dilakukan 4 langkah sebagai berikut.

1. Data Cleansing membersihkan dari Hastag, Username, URL, ReTweet, angka dan emoticon pada tweets.
2. Case folding mengubah huruf besar menjadi huruf kecil untuk di proses.
3. Tokenizing membagi teks dan memeriksa setiap kata individual.
4. Stopwords removal penghapusan kata sambung.
5. Stemming membentuk kata – kata dasar setelah dari proses Tokenizing.
6. Remove Duplicate digunakan untuk menghapus data yang sama.

Dan gambar 2 berikut menunjukkan data sebelum dan sesudah preprocessing.

	Created-At	Text	Id
0	2021-02-08 06:59:56	@geloraco Ditiadakan pilkada 2022 dan 2023, ad...	1358566194211483651
1	2021-02-08 06:59:09	RT @Indonesianews15: Bukan Risma, Tapi Gibran ...	1358565997423169536
2	2021-02-08 06:57:58	RT @tribunnews: Peneliti LIPI Usulkan Pilkada ...	1358565701762441217
3	2021-02-08 06:56:17	Jika Skenario Buruk Terjadi, Demokrat Tak Bisa...	1358565275314999296
4	2021-02-08 06:56:04	RT @SartonoHutomo: Komisioner KPU Beberkan Beb...	1358565219946008576

	Tweet
2816	
6941	aaaasssliii gak pilkada
18164	aaah rambut topi baju orang udah gak drama ind...
29997	aamiiinnn yaallaaah enek gejala ketok batuk pi...
26699	aang dorong awas milu sidoarjo riset pilkada c...
...	...
36451	zona merah pilkada tuh kayak
23907	zona merah psbb neng kerja neng lihat orang
24644	zona merah psbb neng kerja neng lihat orang se...
32567	zoom meeting monitoring pam pilkada t a serah ...
193	zuhro milu pilkada satu

(a)

(b)

Gambar 5.(a) Sebelum Preprocessing dan (b) Setelah Preprocessing

### c) Pelabelan Kelas Data

Tahap selanjutnya dari pengumpulan data adalah labeling data. Dimana pada tahap labeling ini akan dilakukannya secara manual. Untuk hasil dari labeling berupa sentimen positif dan sentiment negative. Lalu dilakukannya perhitungan menggunakan algoritma Support Naïve Bayes Classifier (NBC), dan Support Vector Machine. Pertama yaitu Coarse-grained sentiment analysis kita mencoba melakukan proses analysis pada level Dokumen. Singkatnya adalah kita mencoba mengklasifikasikan orientasi sebuah dokumen secara keseluruhan. Orientasi ini ada 3 jenis: Positif, Netral, Negatif. Akan tetapi, ada juga yang menjadikan nilai orientasi ini bersifat kontinu / tidak diskrit. Kedua adalah Fined-grained sentiment analysis kategori kedua ini yang sedang Naik Daun sekarang. Maksudnya adalah para researcher sebagian besar fokus pada jenis ini. Obyek yang ingin diklasifikasi bukan berada pada level dokumen melainkan sebuah kalimat pada suatu dokumen. Dan tabel 2 dibawah ini menunjukkan nilai klasifikasi.

Tabel 6. Nilai Klasifikasi

	Positif	Negatif
Pilkada	8308	3990

#### d) Resampling

Menerapkan model resampling yang terpilih sehingga di sini terjadi transformasi dataset yang tidak seimbang menjadi dataset yang seimbang untuk evaluasi ulang tetapi hanya dengan 3 model dengan akurasi tinggi (Brilianti, 2020).

Karena terdapat sentiment yang tidak seimbang pada tweet pilkada, dilakukan resampling dengan metode random yang berguna untuk menyeimbangkan data. Cara yang digunakan ialah dengan memangkas data yang paling banyak, yaitu downsampling. Berikut adalah nilai klasifikasi setelah dilakukan downsampling yang terlihat pada tabel 3.

Tabel 7. Nilai Resampling

	Positif	Negatif
Pilkada	3990	3990

#### e) Pembobotan TF-IDF

Pembobotan kata adalah proses pemberian bobot untuk setiap kata yang terdapat dalam sebuah dokumen. Dalam pencarian informasi peringkat berdasarkan frekuensi kata, salah satu metode yang paling populer adalah metode TF-IDF (Term Frequency - Inversed Document Frequency) (Gunawan et al., 2018).

#### f) Klasifikasi

Setelah processing data, labeling, dan resampling tahap selanjutnya adalah klasifikasi analisis sentimen. Tahap ini adalah tahap untuk memberikan pelatihan dan mengimplementasikan berbagai algoritma machine learning. Merupakan tahap pemilihan teknik penambangan dengan menentukan algoritma yang akan digunakan. Penelitian ini menggunakan tools yang digunakan untuk melakukan pemodelan sesuai dengan teknik yang telah ditentukan, tools tersebut adalah Python. Tahap klasifikasi sentimen merupakan proses pengujian klasifikasi data tweet dengan menggunakan Algoritma klasifikasi yang digunakan yaitu, algoritma sebagai modelnya. Hasil pengetesan setiap model adalah untuk mengklasifikasikan tweet.

Naive Bayes merupakan metode pengklasifikasian yang sering digunakan dalam sentimen analisis karena sederhana dan mudah dalam melakukan pengklasifikasian dokumen positif dan tweet negatif untuk mendapatkan nilai akurasi terbaik pada setiap algoritma (Kurniawan & Susanto, 2019).

Konsep klasifikasi dengan Support Vector dapat diartikan Machine adalah mencari hyperplane terbaik yang berfungsi sebagai pemisah dua kelas data. Support Vector Machine mampu bekerja pada dataset yang berdimensi tinggi dengan menggunakan kernel trik (Rofiqoh et al., 2017).

## HASIL DAN PEMBAHASAN

Pada penelitian ini penulis melakukan eksperimen pada studi kasus pilkada 2020 pada saat pandemic ini dengan menggunakan 2 algoritma, Support Vector Machine dan Naive Bayes. Terdapat 3 skenario yang menjadi pembagian presentase data. Pembagian presentasi data merupakan pemisahan data training dan testing berdasarkan presentasinya, contoh 90% : 10% berarti 90% merupakan data training dan 10%

merupakan data testing. Pada penelitian ini pembagian presentase dibagi menjadi 3 skenario eksperimen yaitu eksperimen pertama menggunakan pembagian data 90% : 10%, eksperimen kedua menggunakan pembagian data 80% : 20% dan eksperimen yang ketiga menggunakan pembagian data 70% : 30%.

Akurasi pada tiap masing-masing scenario bisa berbeda, karena model yang telah terbentuk akan dievaluasi dengan menggunakan konsep confusion matrix. Adapun rumus yang digunakan dalam menentukan akurasi ialah:

$$\text{Akurasi} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) * 100$$

- TP = True Positif
- TN = True Negatif
- FP = False Positif
- FN = False Negatif

Hasil dari kinerja model dengan menggunakan pembagian data atau percentage split 90% : 10% bahwa untuk studi kasus pilkada memiliki nilai akurasi 88% untuk algoritma SVM seperti terlihat pada gambar 3 dan 76% untuk algoritma Naïve Bayes seperti terlihat pada gambar 4.

```
0.8847117794486216
[[1071 117]
 [ 159 1047]]
      precision    recall  f1-score   support

     0       0.87     0.90     0.89     1188
     1       0.90     0.87     0.88     1206

 accuracy                   0.88     2394
 macro avg       0.89     0.88     0.88     2394
 weighted avg    0.89     0.88     0.88     2394
```

Gambar 6. Hasil Kinerja SVM Percentage split 90:10

```
0.758563074352548
[[1075  92]
 [ 486 741]]
      precision    recall  f1-score   support

     0       0.69     0.92     0.79     1167
     1       0.89     0.60     0.72     1227

 accuracy                   0.76     2394
 macro avg       0.79     0.76     0.75     2394
 weighted avg    0.79     0.76     0.75     2394
```

Gambar 7. Hasil Kinerja Naïve Bayes Percentage split 90:10

Hasil dari kinerja model dengan menggunakan pembagian data atau percentage split 80% : 20% bahwa untuk studi kasus pilkada memiliki nilai akurasi 88% untuk algoritma SVM ditunjukkan pada gambar 5 dan 76% untuk algoritma Naïve Bayes ditunjukkan pada gambar 6.

```
0.8834586466165414
[[716 90]
 [ 96 694]]
precision recall f1-score support
0 0.88 0.89 0.89 806
1 0.89 0.88 0.88 790
accuracy 0.88 1596
macro avg 0.88 0.88 0.88 1596
weighted avg 0.88 0.88 0.88 1596
```

Gambar 8. Hasil Kinerja SVM Percentage split 80:20

```
0.7568922305764411
[[705 57]
 [331 503]]
precision recall f1-score support
0 0.68 0.93 0.78 762
1 0.90 0.60 0.72 834
accuracy 0.76 1596
macro avg 0.79 0.76 0.75 1596
weighted avg 0.79 0.76 0.75 1596
```

Gambar 9. Hasil Kinerja Naïve Bayes Percentage split 80:20

Hasil dari kinerja model dengan menggunakan pembagian data atau percentage split 70% : 30% bahwa untuk studi kasus pilkada memiliki nilai akurasi 90% untuk algoritma SVM seperti yang terlihat pada gambar 7 dan 78% untuk algoritma Naïve Bayes seperti yang terlihat pada gambar 8 dibawah ini.

```
0.9035087719298246
[[361 39]
 [ 38 360]]
precision recall f1-score support
0 0.90 0.90 0.90 400
1 0.90 0.90 0.90 398
accuracy 0.90 798
macro avg 0.90 0.90 0.90 798
weighted avg 0.90 0.90 0.90 798
```

Gambar 10. Hasil Kinerja SVM Percentage split 70:30

```
0.7781954887218046
[[375 22]
 [155 246]]
precision recall f1-score support
0 0.71 0.94 0.81 397
1 0.92 0.61 0.74 401
accuracy 0.78 798
macro avg 0.81 0.78 0.77 798
weighted avg 0.81 0.78 0.77 798
```

Gambar 11. Hasil Kinerja Naïve Bayes Percentage split 70:30

Dari ketiga scenario eksperimen yaitu eksperimen pertama menggunakan pembagian data 90% : 10%, eksperimen kedua menggunakan pembagian data 80% : 20% dan eksperimen yang ketiga menggunakan pembagian data 70% : 30% mendapatkan hasil perbandingan antara metode algoritma naïve bayes dan SPM. Pada tabel 4 dibawah ini merupakan hasil perbandingan dari kedua algoritma tersebut.



Tabel 8. Perbandingan Nilai Akurasi

Skenario	Akurasi (%)	
	Naïve Bayes	Support Vector Machine
Pertama (90%:10%)	76	88
Kedua (80%:20%)	76	88
Ketiga (70%:30%)	78	90
Rata-rata	76.67	88.67

## KESIMPULAN

Kesimpulan yang didapatkan dari penelitian ini adalah sebagai berikut:

1. Penelitian ini dapat mengidentifikasi sentiment analisis respon masyarakat terhadap kebijakan pemerintah Indonesia untuk melaksanakan pilkada pada masa pandemic covid-19 dengan data yang didapatkan bersumber dari postingan twitter.
2. Metode algoritma Naïve Bayes dan Support Vector Machine (SVM) dapat digunakan untuk mengklasifikasikan sentiment analisis.
3. Nilai akurasi SVM lebih baik dari Naïve Bayes dengan rata-rata nilai akurasi yang didapatkan pada naïve bayes sebesar 76.67% dan SVM sebesar 88.67%.

## LIMITASI DAN STUDI LANJUTAN

Terdapat beberapa limitasi dalam penelitian ini seperti data yang digunakan hanya bersumber dari platform twitter, data yang digunakan hanya berfokus pada postingan berbahasa Indonesia, dan label kelas klasifikasi sentiment yang digunakan hanya 2 (dua) yaitu positif dan negative. Studi lanjutan yang dapat dilakukan pada penelitian ini adalah dapat menggunakan berbagai *platform* sosial media, lalu dapat menggunakan berbagai algoritma *machine learning* lainnya, serta dapat menambah label kelas klasifikasi bukan hanya 2 (dua) positif dan negatif.

## UCAPAN TERIMA KASIH

Ucapan terima kasih kepada dosen pembimbing penelitian yaitu Dr. Mujiono Sadikin, M.T., CISA, CGEIT dan Arieep Jaenul, S.Pd., M.Sc.Eng

## REFERENSI

- Antinasari, P., Perdana, R. S., & Fauzi, M. A. (2017). *Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku*. 1(12), 1733–1741.
- Azanella, L. A. (2020). *Alasan Pro dan Kontra Pilkada Serentak di Tengah Pandemi Covid-19*. Kompas.Com. <https://www.kompas.com/tren/read/2020/09/24/072900565/alasan-pro-dan-kontra-pilkada-serentak-di-tengah-pandemi-covid-19?page=all>
- Brilianti, S. P. (2020). *Peningkatan Performa Analisis Sentimen Dengan Resampling dan Hyperparameter pada Ulasan Aplikasi BNI Mobile*. September 2018, 140–153. <https://doi.org/10.30864/eksplora.v9i2.333>
- Buntoro, G. A. (2016). ANALISIS SENTIMEN HATESPEECH PADA TWITTER DENGAN METODE NAÏVE BAYES CLASSIFIER DAN SUPPORT VECTOR

- MACHINE. *Jurnal Dinamika Informatika*, 5(September).
- Buntoro, G. A. (2017). *Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter*. 2(1), 32–41.
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). *Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes*. 4(2), 113–118.
- Irawan, A. J. H. (2020). Fenomena Covid-19: Dampak Globalisasi Dan Revitalisasi Multilateralisme. *Jurnal Ilmiah Hubungan Internasional*, 0(0), 47–52. <https://doi.org/10.26593/jihi.v0i0.3877.47-52>
- Kurniawan, I., & Susanto, A. (2019). Implementasi Metode K-Means dan Naive Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019. *Eksplora Informatika*, 9(1), 1–10. <https://doi.org/10.30864/eksplora.v9i1.237>
- Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing, Second Edition*, 627–666.
- Pravina, A. M., Cholissodin, I., & Adikara, P. P. (2019). Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3), 2789–2797. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4793>
- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). *Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Features*. 1(12), 1725–1732.