

ANALISIS SENTIMEN PENGGUNA ZOOM CLOUD MEETING MENGUNAKAN BIG DATA

Zidan Muharam¹, Iik Muhamad Malik Matin², Muhammad Taher Jufri³

¹ Fakultas Teknik Universitas Garut, Jl. Jati 42B, Garut, Jawa Barat, 44151, Indonesia

² Departemen Teknik Informatika dan Komputer Politeknik Negeri Jakarta, Jl. Prof. DR. G.A. Siwabessy, Kukusan, Kecamatan Beji, Kota Depok, Jawa Barat, Indonesia

³ Fakultas Komunikasi dan Informasi Universitas Garut, Jl. Raya Samarang Jl. Hampor No.52A, Langensari, Kec. Tarogong Kidul, Kabupaten Garut, Jawa Barat, 44151, Indonesia

Korespondensi: zidanmuharam321@gmail.com

ARTICLE HISTORY

Received: 19-12-2022

Revised: 29-12-2023

Accepted: 29-12-2023

Abstrak

Aplikasi *zoom cloud meetings* yang mulai *booming* digunakan sekarang ini karena adanya pandemi *covid-19*, sehingga membuat semua kegiatan dilakukan secara *virtual*. *Zoom cloud meetings* merupakan aplikasi yang memiliki berbagai fitur termasuk *video & audio conference*. Pada penelitian ini penulis menggunakan metode *Naïve Bayes* dalam menganalisa label sentimen positif dan negatif pada para pengguna aplikasi *zoom* di *Twitter*. Jumlah dataset setelah *preprocessing* sebanyak 1.000 data. Data hampir seimbang dengan label sentimen positif berbahasa indonesia sebanyak 583 dan sentimen positif berbahasa inggris sebanyak 194 dan label sentimen negatif berbahasa indonesia sebanyak 134 dan sentimen negatif berbahasa inggris sebanyak 89 data. Evaluasi model menggunakan *confusion matrix* diperoleh nilai akurasi dan nilai dari algoritma *Naïve Bayes* menggunakan data latih = 80% dan data uji = 20% menghasilkan nilai akurasi = 70%.

Kata kunci: Analisis Sentimen, *Covid-19*, *Naïve Bayes Classifier*, *twitter*, *Zoom Cloud Meetings*.

ZOOM CLOUD MEETING USER SENTIMENT ANALYSIS USING BIG DATA

Abstract

The *zoom cloud meetings* application, which is starting to boom, is being used now due to the *covid-19* pandemic, making all activities carried out virtually. *Zoom cloud meetings* is an application that has various features including *video & audio conferencing*. In this study, the authors used the *Naïve Bayes* method in analyzing positive and negative sentiment labels in the reviews of *zoom* application users on *Twitter*. The number of datasets after processing is 1,000 data. The data is almost balanced with a positive sentiment label of 583 and a positive sentiment in English with 194 and a negative sentiment label of 134 and a negative sentiment in English with 89 data. Evaluation of the model using the *confusion matrix* obtained accuracy values and values from the *Naïve*

Bayes algorithm using train data = 80% and test data = 20% resulting in accuracy values = 70%.

Key words: *Sentiment Analysis, Naïve Bayes Classifier, Covid-19, Zoom Cloud Meetings, twitter.*

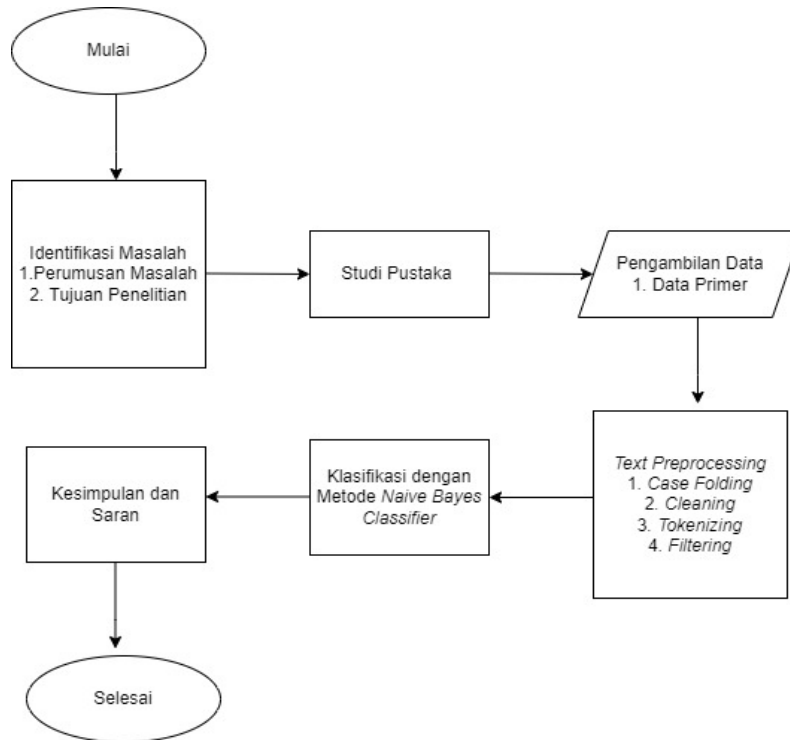
1. Pendahuluan

Hampir setiap negara di dunia, termasuk Indonesia, saat ini sedang menghadapi kesulitan akibat pandemi *Covid-19*. *Covid-19* atau Penyakit *Coronavirus* 2019 adalah infeksi menular yang disebabkan oleh virus *corona* baru. Penyakit ini menyebar begitu cepat sehingga seluruh dunia waspada. Oleh karena itu, banyak negara telah menerapkan karantina wilayah dan pembatasan seluruh aktivitas di negaranya masing-masing untuk menahan penyebaran virus *Covid-19* [1]. Jarak Sosial, atau Jarak Sosial, adalah serangkaian tindakan pengendalian infeksi nonfarmasi yang dirancang untuk menghentikan atau menunda penyebaran infeksi ini. Oleh karena itu, diharapkan semua orang bekerja di rumah untuk menghentikan penyebaran virus *Covid-19*. Agar tetap aktif di rumah, Anda memerlukan rapat virtual untuk berkomunikasi dengan tim dan karyawan lain. Konferensi virtual sekarang tersebar luas [2].

Penggunaan aplikasi *cloud conferencing Zoom* yang akhir-akhir ini mulai *booming* akibat merebaknya *virus corona* yang melanda hampir di setiap negara, termasuk Indonesia. Salah satu langkah pemerintah untuk menekan penyebaran *virus corona* adalah dengan bekerja dari rumah, dimana para pekerja didorong untuk bekerja dari rumah atau *work from home* (WFH), begitupun para pelajar dan mahasiswa agar belajar dari rumah atau *study from home* (SPH) [3]. Penelitian ini berfokus pada penentuan *mood* label. Analisis sentimen secara otomatis memahami, mengekstrak, dan memproses data tekstual untuk memperoleh informasi emosional yang terkandung dalam pernyataan opini, atau pendapat pribadi, perilaku, dan emosi (menjelaskan orang, peristiwa, topik) tentang suatu entitas [4][5][6][7]. Proses komputasi yang memberikan informasi berharga. Untuk orang-orang yang berasal dari kumpulan data tidak terstruktur. Besarnya dampak dan manfaat dari analisis sentimen berarti penelitian dan aplikasi berbasis analisis sentimen berkembang sangat pesat dan banyak perusahaan yang fokus pada layanan analisis sentimen.

2. Metode

Analisis sentimen pada aplikasi *zoom cloud meeting* dilakukan dengan tahapan pengumpulan data, *preprocessing*, klasifikasi, dan evaluasi. Tahapan analisis sentimen ditunjukkan pada gambar 1. *Preprocessing* teks pada penelitian ini terdiri dari *case folding*, *cleaning*, *tokenizing* dan *filtering*. Kemudian teks diklasifikasi menggunakan Metode *Naïve Bayes Classifier*. Untuk mengetahui performa model dilakukan analisis pada *confusion matrix*.



Gambar 1. Flowchart Penelitian

2.1 Pengumpulan Data

Pengambilan data dilakukan dengan cara *web scraping* data karena lebih mudah dan lebih efisien dibandingkan dengan cara manual. Data yang dikumpulkan berupa tweet berbahasa Indonesia dan bahasa Inggris yang dilakukan pada bulan Juli 2022 selama satu bulan. Data yang diperoleh dari twitter mengenai ulasan penggunaan aplikasi *Zoom Cloud Meetings*. Dari hasil pengambilan data yang dilakukan diperoleh 1.000 data. Data yang diperoleh kemudian dibagi menjadi 80% untuk data training dan 20% untuk data uji.

2.2 Preprocessing

Dilakukannya *text preprocessing* bertujuan untuk membersihkan data dan mengubah data menjadi lebih terstruktur. Tahapan *text preprocessing* menggunakan *tool python 3.10.5*. Empat tahapan yang dilakukan pada *text preprocessing* yaitu:

a. Case Folding

Case folding dilakukan untuk mengubah bentuk huruf pada data menjadi huruf lowercase.

b. Cleaning

Cleaning digunakan untuk menghilangkan bagian yang tidak penting pada data seperti tanda baca, angka, emoji, dan lain-lain.

c. Tokenizing

Tokenizing merupakan proses pemecahan teks kalimat menjadi potongan kata agar kalimat yang terbentuk menjadi lebih bermakna.

d. Filtering

Tahap *filtering* adalah tahap mengambil kata-kata penting dari hasil proses tokenizing. Penelitian ini menggunakan stopwords berbahasa Inggris berdasarkan modul NLTK.

2.3 Klasifikasi

Klasifikasi sentimen pada makalah ini dilakukan dengan menggunakan algoritma *Naïve bayes Classifier*. Algoritma yang digunakan untuk melakukan klasifikasi ulasan pengguna yang berbentuk kelas positif maupun negatif.

2.4 Evaluasi

Evaluasi merupakan tahapan untuk mengukur performa model pada makalah ini diukur menggunakan metode *confusion matrix*. Hasil klasifikasi biner pada suatu dataset dapat direpresentasikan dengan matriks 2 x 2 yang disebut *confusion matrix*. Berikut merupakan matriks tabel pada *confusion matrix* yang ditunjukkan pada Tabel 1.

Tabel 1. Matrik *confusion*

Prediksi	Aktual	
	Positif	Negatif
Positif	<i>True Positive</i> (TP)	<i>False Negative</i> (FN)
Negatif	<i>False Positive</i> (FP)	<i>True Negative</i> (TN)

Dimana:

- True Positive (TP) : memprediksi positif dan itu benar
- True Negative (TN) : memprediksi negatif dan itu benar
- False Positive (FP) : memprediksi positif dan itu salah
- False Negative (FN) : memprediksi negatif dan itu salah

Dari Tabel 1. dapat diperoleh rumus untuk menghitung *accuracy*, *precision*, *recall* dan *F-1 score*.

- a. *Accuracy* yaitu menggambarkan akurasi model dalam mengklasifikasikan dengan benar.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- b. *Precision* menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model.

$$\text{presisi} = \frac{TP}{TP+FP} \quad (2)$$

- c. *Recall* atau sensitivity: menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

- d. *F1-Score* menggambarkan perbandingan rata-rata *precision* dan *recall* yang dibobotkan

$$F - \text{Score} = 2 \times \frac{\text{Precision} \times \text{TPR}}{\text{Precision} + \text{TPR}} \quad (4)$$

3. Hasil dan Pembahasan

3.1 Pengujian

Setelah perancangan telah selesai, maka pada tahapan ini akan membahas tentang hasil pengujian yang di teliti. Pengujian analisis sentimen atau *opinion mining* adalah deklarasi sikap – sikap terhadap objek atau orang. Analisis sentimen dilakukan untuk mengidentifikasi pendapat atau kecenderungan seseorang terhadap suatu masalah atau objek. Tidak masalah jika orang tersebut cenderung memiliki pandangan atau pendapat *negatif* atau *positif*.

3.1.1 Pengambilan Data Menggunakan Api Twitter

Proses pengambilan data atau *crawling* data *Twitter* ini menggunakan API key *Twitter*, dan proses pengambilan data *Twitter* menggunakan bahasa pemrograman *Python*. Untuk penelitian sebelumnya proses pengambilan data atau *crawling* data melalui *google play* dan untuk pengkodingannya melalui *jupyter* sedangkan disini peneliti untuk pengkodingannya melalui *google colab.library* yang dibutuhkan dimasukan selanjutnya masukkan *api_key*, *api_key_secret*, *access_token*, *access_token_secret* yang telah didapatkan dari *twitter*. Proses pengambilan atau *crawling* data dari API *twitter* telah dilakukan dan menjalankan semua kode diatas maka akan didapat *file excel* dengan ekstensi *file csv* yang belum diolah dari hasil *crawling* data. Seperti pada gambar 4. dibawah.

3.1.2. Hasil Sentimen dengan Angka

Apabila Sentimen positif memiliki angka 0.0 sedangkan untuk sentimen negatif memiliki angka diatas 0.0 seperti 0.35

3.1.3. Hasil Sentimen Dengan Kata

Apabila ada satu akun yang memposting disebut sentimen positif, sedangkan untuk sentimen negatif ada yang memposting lebih dari satu akun.

3.1.4. Hasil Pengolahan Berdasarkan Tahapan *Preprocessing*

1) *Cleaning*

Mengubah semua huruf pada data menjadi huruf *lowercase*.

2) *Tokenizing*

Tokenizing merupakan proses pemecahan teks kalimat menjadi potongan kata-kata guna membuat kalimat yang terbentuk menjadi lebih bermakna.

3) *Filtering*

Tahap *filtering* adalah tahap mengambil kata-kata penting dari hasil proses tokenizing. Penelitian ini menggunakan *stopword* berbahasa Inggris berdasarkan modul NLTK.

4) *Case Folding*

Mengubah semua huruf pada data menjadi huruf *lowercase*.

3.1.5. Hasil Klasifikasi Menggunakan Metode *Naïve Bayes Classifier*

a) Hasil Data *Training* dan Data *Testing*

Dalam pengujian, aturan umum data training dan testing harus dipisahkan terlebih dahulu dengan kata lain tidak boleh menggunakan data yang sama untuk training dan testing. Pada tahap ini peneliti membagi dataset menjadi data training dan testing.

b) Hasil Nilai Akurasi

Nilai akurasi yang didapatkan dari proses pengujian model sebesar 70% dengan proses perhitungan berdasarkan jumlah nilai dari *confusion matrix* dan dibagi dengan jumlah seluruh data.

4. Kesimpulan

Kesimpulan dari penelitian yang dilakukan adalah berdasarkan data *zoom cloud meeting* di *twitter* pada bulan Juli 2022, didapatkan data sentimen positif berbahasa indonesia sebanyak 583 data dan sentimen positif berbahasa inggris 194 data, sedangkan untuk data sentimen negatif sebanyak 134 berbahasa indonesia dan sentimen negatif berbahasa inggris sebanyak 89 data. Berdasarkan tahapan *preprocessing* data yang didapat sebanyak 1000 data yang dapat diolah. Berdasarkan klasifikasi yang telah dilakukan dengan metode *naïve bayes classifier*, yang dimana dengan menggunakan data latih dan data uji. Diketahui bahwa dengan data latih dan data uji persentasenya sebesar 80 : 20 menghasilkan tingkat akurasi sebesar 70%.

Ucapan Terima Kasih

Penelitian ini dapat dilaksanakan dengan baik berkat bantuan dari berbagai pihak, untuk itu penulis mengucapkan terimakasih kepada Prodi Teknik Elektro, Fakultas Teknik, Universitas Garut, dan pihak lain yang telah memberikan kerjasama yang baik dalam penelitian ini.

Daftar Pustaka

- [1] Ratino, N. Hafidz, S. Anggraeni, and W. Gata, "Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naïve Bayes," *J. JUPITER*, vol. 12, no. 2, pp. 1–11, 2020.
- [2] M. Rezki, D. N. Kholifah, M. Faisal, P. Priyono, and R. Suryadithia, "Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes," *J. Infortech*, vol. 2, no. 2, pp. 264–270, 2020, doi: 10.31294/infortech.v2i2.9286.
- [3] Okto Vierten Masrel, "Bekerja dari Rumah (Work From Home) Dari Sudut Pandang Unit Kepatuhan Internal," *Kanwil DJKN Jawa Barat*, 2020. <https://www.djkn.kemenkeu.go.id/artikel/baca/13014/Bekerja-dari-Rumah-Work-From-Home-Dari-Sudut-Pandang-Unit-Kepatuhan-Internal.html> (accessed Feb. 28, 2022).
- [4] A. Rohanah, B. A. Dermawan, and ..., "Klasifikasi Ulasan Pengguna Zoom Cloud Meetings Menggunakan Metode Information Gain dan Naïve Bayes Classifier," *J.*

-
- Inform.*, vol. 6, no. 2, pp. 348–357, 2021, [Online]. Available: <http://openjournal.unpam.ac.id/index.php/informatika/article/view/10728>
- [5] A. Fauzi, A. N. Rais, M. F. Akbar, and W. Gata, “Analisis Sentimen Berinternet Pada Media Sosial AMIK BSI Tegal Dengan Menggunakan Algoritma Naive Bayes,” *Semin. Nas. Teknol. Inf. Univ. Ibn Khaldun Bogor*, pp. 46–54, 2018.
- [6] G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” *Integer J.*, vol. 2, no. 1, pp. 32–41, 2017, [Online]. Available: <https://t.co/jrvaMsgBdH>
- [7] M. W. Pertiwi, “Analisis Sentimen Opini Publik Mengenai Sarana dan Transportasi Mudik Tahun 2019 Pada Twitter Menggunakan Algoritma Naive Bayes, Neural Network, K-NN dan SVM,” *Inti Nusa Mandiri*, vol. 14, no. 1, pp. 27–32, 2019.