



SENTIMENT ANALYSIS OF PUBLIC OPINIONS ON THE EFFECTIVENESS OF ONLINE LEARNING USING NAÏVE BAYES ALGORITHM

Ratna Pinka Baptis Dakhi¹, Achmad Maezar Bayu Aji^{2*}

Sistem Informasi^{1,2}

Fakultas Teknologi Informasi^{1,2}

Universitas Nusa Mandiri^{1,2}

dakhi.ratnapinka@gmail.com¹, achmad.azb@nusamandiri.ac.id²

Received: March 30, 2022. **Revised:** May 10, 2022. **Accepted:** May 28, 2022. **Issue Period:** Vol.6 No.1 (2022), Pp. 273-279

Abstract: Covid-19 was declared a global pandemic by the World Health Organization (World Health Organization) on March 11, 2020. This has resulted in the government urging the public not to carry out activities outside the home which is an effort to break the chain of spreading Covid-19. Activities that were suddenly stopped had a negative impact on the community. Starting from the decline in economic growth to the lagging of students in studying. The Covid-19 pandemic has had an impact on various perspectives of human life today, especially in the field of education. So that the teaching and learning process continues. All students are required to study at home online or online. However, online learning raises pros and cons among the public, so research needs to be done to analyze public sentiment about online learning and find out the effectiveness of online learning during the Covid-19 pandemic by utilizing Twitter as a source of data to be researched. Based on the results of the analysis of public opinion sentiment about online learning using the Naïve Bayes Classifier algorithm with Rapidminer as the software used in data processing, the accuracy value is 60.00% with 65.67% precision and 53.30% recall

Keywords: *sentiment analysis, rapidminer, naive bayes*

I. PENDAHULUAN

Penyakit mematikan yang menelan banyak korban jiwa ini berawal mula pada akhir Desember 2019. Adanya sebuah kasus yaitu wabah penyakit yang tidak dapat dideteksi dengan gejala pernapasan akut terjadi di Wuhan, Provinsi Hubei, China. Komplikasi pernapasan akut ini, secara epidemiologi terhubung dengan Pasar Grosir Makanan Laut Huanan yang menjual hewan liar dan makanan laut. Hanya dalam kurun waktu tiga bulan merebaknya penyakit ini di seluruh dunia, Covid-19 dinyatakan sebagai pandemi global oleh Organisasi Kesehatan Dunia (World Health Organization) pada tanggal 11 Maret 2020.

Cara virus ini menyerang sistem pernapasan manusia adalah lewat udara yang sudah terkontaminasi virus melalui cipratan air liur ketika berbicara, bersin, batuk, hingga permukaan benda yang terkontaminasi dari satu manusia ke manusia yang lain. Oleh sebab itu, untuk menekan penyebaran virus Covid-19 ini, masyarakat yang berada di tempat umum diwajibkan menggunakan masker, menjaga jarak satu sama lain dengan jarak 1 meter, mencuci tangan, menyediakan hand sanitizer sampai harus bekerja dari rumah, beraktivitas di rumah, beribadah di rumah, hingga belajar di rumah. Intinya, masyarakat dihimbau dengan tegas untuk meniadakan perkumpulan dan menghindari kegiatan di luar rumah, sampai kondisi menjadi normal atau menurunnya kasus pandemi Covid-



DOI: 10.52362/jisicom.v6i1.822

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



19 ini. Akibatnya kantor ditutup, tempat ibadah ditutup, restoran atau rumah makan ditutup sampai sekolah pun harus ditutup. Kegiatan yang dihentikan secara mendadak ini berdampak buruk pada masyarakat.

Mulai dari pemerosotan pertumbuhan ekonomi sampai tertinggalnya peserta didik dalam menuntut ilmu. Selain itu, upaya pencegahan juga dilakukan oleh para tenaga medis dan kesehatan yaitu dengan mengadakan kegiatan vaksinasi untuk semua masyarakat. Berdasarkan data yang dilansir dari website World Health Organization Di Indonesia sejak 3 Januari 2020 hingga 12 November 2021 kasus Covid-19 yang terkonfirmasi sebanyak 4.250.157 dengan kasus kematian mencapai 143.628, dilaporkan ke WHO. Sejak 9 November 2021 total dosis vaksin 207.306.431 sudah dikelola [1].

Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi juga menyalurkan dukungan kepada para pendidik, peserta didik, dosen dan mahasiswa supaya lebih semangat dan termotivasi selama belajar daring berupa kuota internet gratis. Bantuan paket kuota internet tahun 2021 periode September–November 2021 Kementerian Pendidikan, Kebudayaan, riset, dan teknologi dengan rincian bantuan antara lain sebagai berikut:

1. Peserta Didik Jenjang PAUD 7GB/Bulan
2. Peserta Didik Jenjang Pendidikan Dasar dan Menengah 10GB/Bulan
3. Pendidik Jenjang PAUD, Pendidikan Dasar dan Menengah 12GB/Bulan
4. Dosen dan Mahasiswa 15GB/Bulan dengan peraturan pembatasan akses di situs-situs yang telah diblokir oleh kementerian Komunikasi dan Informatika [2].

Berdasarkan munculnya pro dan kontra di kalangan masyarakat, maka penelitian ini dilakukan untuk menganalisis sentimen publik terhadap belajar daring ini. Analisis Sentimen (Opinion Mining) adalah prosedur dalam menginterpretasikan, menghasilkan dan menggarap data yang masih berbentuk tekstual secara cepat untuk memperoleh informasi yang terdapat dalam suatu kalimat opini [3]. Masyarakat era Industri 4.0, sudah menggunakan berbagai aplikasi media sosial untuk bersosialisasi secara daring satu sama lain. Contoh aplikasi yang sering digunakan adalah Facebook, Youtube, Instagram, WhatsApp, Twitter, Tik-Tok dan lain-lain. Oleh sebab itu, data yang akan diolah dalam penelitian ini diperoleh dari media sosial. Namun, media sosial yang dipakai adalah Twitter karena aplikasi media sosial Twitter telah menyediakan API (Application Programming Interface) yang berfungsi dalam pengambilan serta pengolahan data oleh para pengembang.

Menurut pendapat seorang Ilmuwan bernama Bing Liu mengungkapkan bahwa analisis sentimen atau penambahan opini merupakan ilmu yang mengkaji pendapat, sentimen, evaluasi, penilaian, sikap dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya [4]. Data yang akan diperoleh adalah hasil dari tweet para pengguna media sosial Twitter dengan tagar #belajar daring. Kemudian, data ini akan diolah dan diklasifikasi menggunakan algoritma Naïve Bayes untuk menganalisis sentimen publik melalui klasifikasi yaitu kelas positif dan kelas negatif. Dari hasil ini juga akan menentukan apakah belajar online efektif atau tidak selama pandemi Covid-19. Tujuan dari penelitian adalah untuk menganalisis sentimen publik tentang belajar daring serta mengetahui keefektifan belajar daring selama pandemi Covid-19 ini dengan memanfaatkan media sosial Twitter sebagai sumber data yang akan diteliti

II. METODE DAN MATERI

Metode yang digunakan oleh penulis dalam penelitian ini adalah metode kualitatif. Dan data yang digunakan dalam penelitian ini berupa kumpulan tweet mengenai opini masyarakat terkait belajar daring yang diterapkan oleh pemerintah sebagai salah satu upaya memutus mata rantai penyebaran wabah Covid-19. Data ini diambil dari media sosial Twitter pada bulan Desember 2021 berdasarkan kata kunci #belajar daring dan #belajar online menggunakan tools *Rapidminer* dan Twitter API (Application Programming Interface) sebanyak 436 tweet.

2.1. Metode Pengumpulan Data

Dalam penelitian ini penulis menggunakan beberapa tahapan-tahapan yang diantaranya adalah:

1. Identifikasi Masalah

Pada awal tahapan dimulai dengan mengidentifikasi permasalahan yang berhubungan dengan belajar dalam jaringan (daring). Penulis menyimpulkan bahwa belajar daring menimbulkan pro dan kontra di kalangan





masyarakat, sehingga penelitian perlu dilakukan untuk menganalisis sentimen publik tentang belajar daring serta mengetahui keefektifan belajar daring selama pandemic Covid-19 ini dengan memanfaatkan media sosial Twitter sebagai sumber data yang akan diteliti.

2. Studi Pustaka

Pada tahap ini, penulis melakukan kegiatan pengumpulan informasi penting yang berasal dari berbagai sumber kepustakaan yang berkaitan dengan topik penelitian. Pada penelitian ini, penulis mendapatkan informasi dari jurnal, buku dan karya ilmiah.

3. *Crawling Data*

Data yang dikumpulkan oleh peneliti diperoleh dengan menggunakan *software Rapidminer*. Dalam *software Rapidminer* terdapat berbagai jenis operator salah satunya adalah Search operator yang dapat terhubung pada Twitter. Data yang dicrawling berupa pendapat masyarakat tentang belajar daring dengan yang diterapkan pada masa pandemi Covid-19 yang menimbulkan pro dan kontra di kalangan semua masyarakat. Untuk menentukan sentiment, data akan dilabeli menggunakan *software Rapidminer*. teks mining mempunyai definisi menggali informasi berbentuk teks yang diperoleh dari dokumen yang sumbernya berbeda-beda, dan bermaksud untuk mengumpulkan kosakata yang berhubungan dengan dokumen sehingga dapat dianalisis [5].

4. *Manual Labelling*

Pada tahap ini, data akan dilabel secara manual terhadap data yang akan diolah. Pelabelan ini akan dibagi ke dalam dua kelas yaitu kelas positif dan kelas negatif.

5. *Preprocessing*

Praproses data adalah bentuk penerapan dari text mining dalam mengolah suatu data teks yang tidak terstruktur menjadi terstruktur, menghasilkan data yang sama rata sehingga bisa diolah pada tahap seterusnya. Praproses data adalah bentuk penerapan dari text mining dalam mengolah suatu data teks yang tidak terstruktur menjadi terstruktur, menghasilkan data yang sama rata sehingga bisa diolah pada tahap seterusnya

6. *Feature Extraction*

Dalam analisis sentimen, data yang sudah diperoleh akan berbentuk kata. Namun, computer tidak dapat mengolah data selain data numerik sehingga dibutuhkan sebuah cara untuk mengekstrak kata menjadi numerik. Metode ini berguna dalam mempresentasikan data secara menyeluruh.

7. *Modelling*

Proses yang telah dilewati sebelumnya adalah proses dalam merubah dataset yang berupa teks menjadi document term matrix. Model yang telah dilatih dengan data training selanjutnya digunakan untuk pengklasifikasian data testing ke dalam dua kelas sentiment yaitu sentiment positif dan sentiment negatif. Dalam pengklasifikasian ini menggunakan metode Naïve Bayes Classifier diperoleh probabilitas yang berguna dalam penentuan apakah tweet masuk ke dalam kategori sentiment negatif atau sentiment positif. Gaya visualisasi yang sering digunakan adalah grafik dan histogram. Biasanya, hasil dari analisis sentimen ini sendiri berbeda-beda. Oleh karena itu, gaya visualisasi hasil berupa wordcloud, peta interaktif, dan gaya sparkline bisa efektif saat mempresentasikan hasil analisis [6].

8. *Model Evaluation*

Metode yang digunakan untuk evaluasi adalah confusion matrix. Confusion matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya.

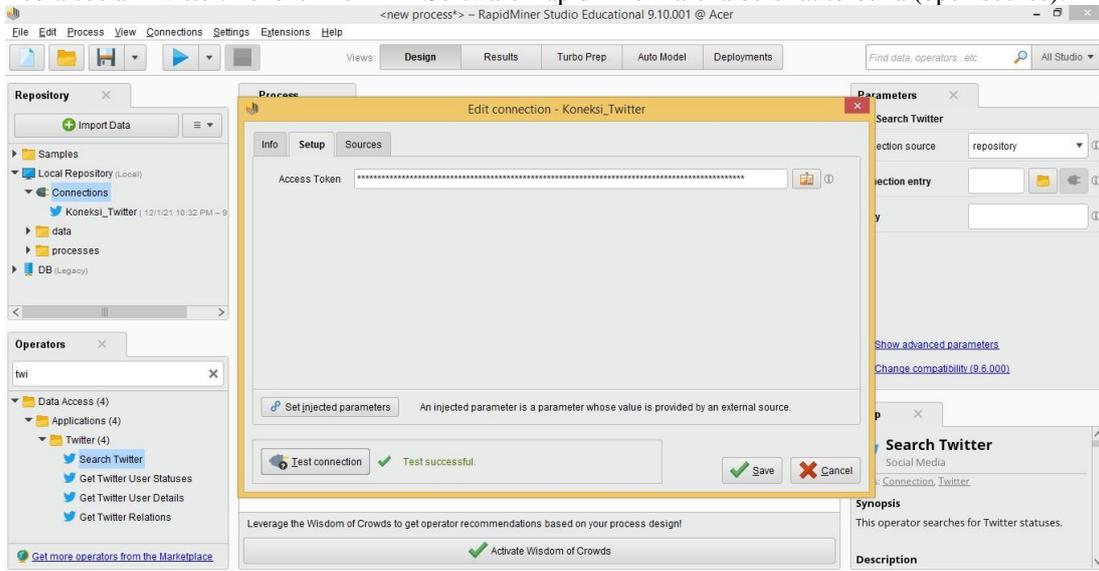
III. PEMBAHASAN DAN HASIL

3.1. *Crawling Data*

Tahap penting pada saat melakukan penelitian adalah adanya dataset yang tersedia untuk diolah. Pada tahap ini, data yang akan diolah diambil dari media sosial Twitter dengan menggunakan *software Rapidminer*. Tahap



selanjutnya adalah menginstal software Rapidminer yang akan digunakan sebagai media pada saat crawling data dari media sosial Twitter. Peneliti memilih Software Rapidminer karena bersifat terbuka (open source).



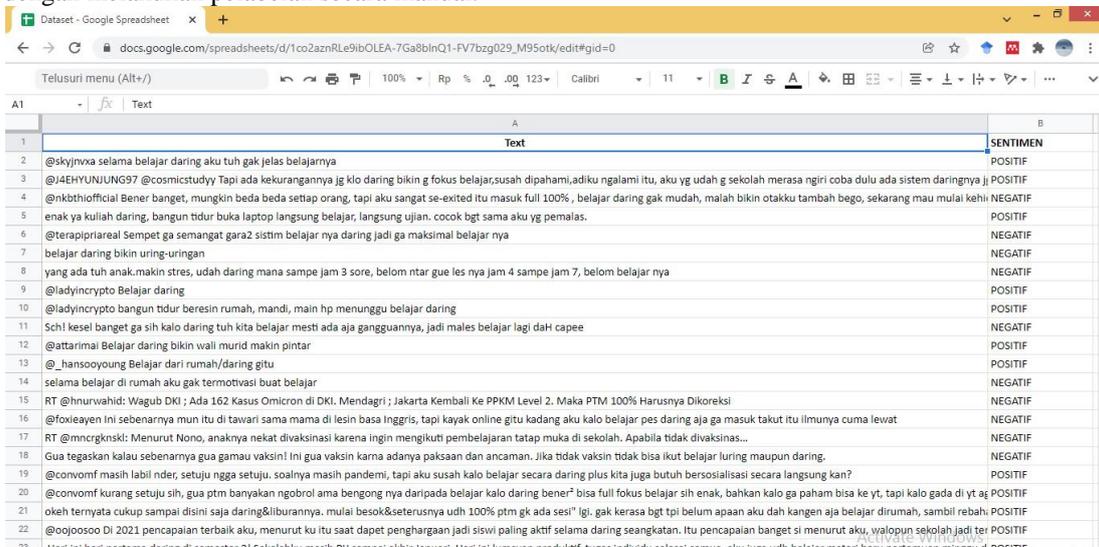
Sumber: Penelitian (2022)

Gambar 1. Proses Membuat Connection Search Twitter Pada Rapidminer

Pada gambar ini, penulis melakukan koneksi pada media sosial Twitter dengan menggunakan operator Search Twitter. Setelah Rapidminer terkoneksi dengan API (Application Programming Interface) maka crawling data sudah bisa dilakukan.

3.2. Pelabelan Data

Setelah data diambil, langkah selanjutnya adalah menentukan sentimen positif dan negatif terhadap teks dengan melakukan pelabelan secara manual.



Sumber: Penelitian (2022)

Gambar 2. Dataset Yang Dilabeli

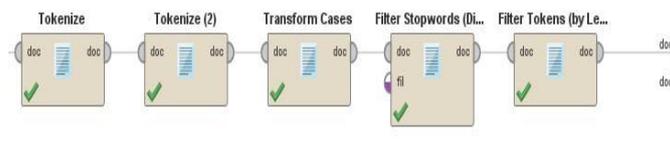


DOI: 10.52362/jisicom.v6i1.822

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

3.3. Preprocessing Text

Data yang sudah dilabeli secara manual selanjutnya akan melewati tahap penting untuk meningkatkan ketepatan klasifikasi. Proses ini akan menghilangkan kata-kata yang dianggap tidak penting seperti simbol-simbol, username, link URL dan tanda baca yang dapat menghambat dalam proses selanjutnya.



Sumber: Penelitian (2022)

Gambar 3. Tampilan Preprocessing Pada Rapidminer

3.4. Extraction Feature

Setelah melewati preprocessing data, dataset yang berupa kata akan diekstrak ke dalam bentuk numerik untuk mempermudah computer dalam melakukan proses pengolahan. Metode yang dibutuhkan dalam ekstraksi data teks Term Frequency Inverse Document Frequency (TF-IDF).

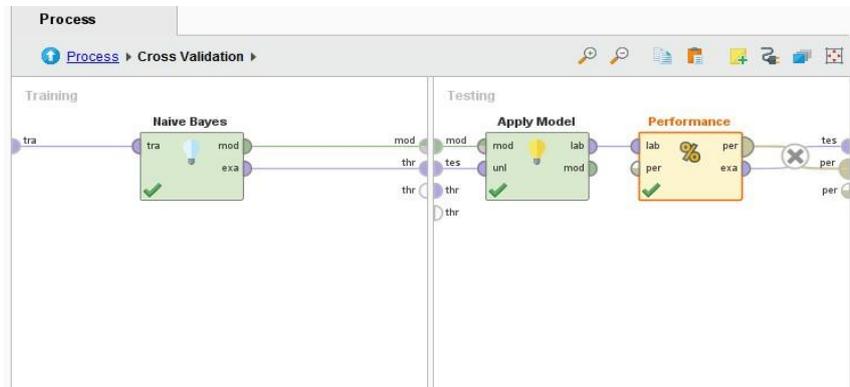
besok	besoknya	best	betah	betapa	bete	bgitu	bgst	biar	biarpun	bias
0	0	0	0	0	0	0	0	0.312	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.861	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0.228	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Sumber: Penelitian (2022)

Gambar 4. Hasil Term Matrix Pada Rapidminer

3.5 Modelling

Dalam pengklasifikasian ini menggunakan metode Naïve Bayes Classifier diperoleh probabilitas yang berguna dalam penentuan apakah tweet masuk ke dalam kategori sentimen negatif atau sentiment positif.



Sumber: Penelitian (2022)

Gambar 5. Pemodelan Naive Bayes Pada Rapidminer

3.6 Evaluation Model

Setelah pemodelan data dilakukan, maka tahapan terakhir dari proses data adalah evaluasi model. Terdapat beberapa kriteria evaluasi model dalam masalah klasifikasi diantaranya akurasi, presision, recall dan F1-Score. Untuk mengetahui kriteria-kriteria evaluasi model pada masalah klasifikasi ini diperlukan sebuah confusion matrix. Confusion Matrix atau disebut juga matriks eror adalah sebuah table spesifik yang memberikan gambaran performa dari suatu algoritma atau model yang digunakan.

accuracy: 62.77% +/- 5.41% (micro average: 62.78%)

	true POSITIF	true NEGATIF	class precision
pred. POSITIF	121	63	65.76%
pred. NEGATIF	106	164	60.74%
class recall	53.30%	72.25%	

Sumber: Penelitian (2022)

Gambar 6. Hasil Confusion Matrix

Berdasarkan hasil pengujian klasifikasi Naïve Bayes menggunakan aplikasi Rapidminer dihasilkan confusion matrix yang dapat diketahui pada Gambar 6, langkah selanjutnya adalah melakukan perhitungan ketepatan klasifikasi.

$$Akurasi = \frac{164 + 121}{164 + 121 + 63 + 106} = \frac{285}{454} = 62,77\%$$

$$Precision = \frac{121}{121 + 63} = \frac{121}{184} = 65,67\%$$

$$Recall = \frac{121}{121 + 164} = \frac{121}{285} = 53,30\%$$

Sumber: Penelitian (2022)

Gambar 6. Hasil Pengujian Klasifikasi Naïve Bayes

IV. KESIMPULAN



DOI: 10.52362/jisicom.v6i1.822

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



Berdasarkan hasil penelitian analisis sentimen opini publik tentang belajar online menggunakan algoritma *Naïve Bayes Classifier* dengan *Rapidminer* sebagai perangkat lunak yang dimanfaatkan dalam pengolahan data diperoleh nilai akurasi 60.00% dengan *precision* 65,67% dan *recall* 53,30%. Pada penelitian ini, tidak dilakukan perbandingan keakuratan algoritma sehingga disarankan untuk penelitian selanjutnya dilakukan perbandingan keakuratan dengan algoritma klasifikasi lainnya seperti Algoritma Support Vector Machine untuk mendapatkan hasil yang lebih akurat dan tepat.

REFERENASI

- [1] “Indonesia: WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data.” <https://covid19.who.int/region/searo/country/id> (accessed Nov. 14, 2021).
- [2] “Bantuan Kuota Data Internet 2021 - Kemendikbud.” <https://kuota-belajar.kemdikbud.go.id/> (accessed Nov. 14, 2021).
- [3] G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” *INTEGGER J. Inf. Technol.*, vol. 1, no. 1, pp. 32–41, 2017.
- [4] Y. Lin, X. Wang, and A. Zhou, “Opinion spam detection,” *Opin. Anal. Online Rev.*, no. May, pp. 79–94, 2016, doi: 10.1142/9789813100459_0007. (accessed Dec. 07, 2021).
- [5] “Mengenal peran sentiment analysis beserta cara kerjanya.”, <https://www.ekrut.com/media/sentiment-analysis-adalah> (accessed Dec. 07,2021).
- [6] “Apa itu Text Mining? – School of Information Systems.” <https://sis.binus.ac.id/2021/04/23/apa-itu-text-mining/> (accessed Dec. 07,2021).

