



# ANALISIS SENTIMEN PEMBELAJARAN DARING PADA MASA PANDEMI COVID-19 DI TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES

Siti Yana Nursyi'ah<sup>1)</sup>, Adhithia Erfina<sup>2)</sup>, dan Cecep Warman<sup>3)</sup>

<sup>1, 2, 3)</sup>Program Studi Sistem Informasi, Universitas Nusa Putra.

Jl. Raya Cibolang No. 21 Cibolang Kaler, Cisaat, Sukabumi, Jawa Barat 43152

e-mail: [siti.yana\\_si17@nusaputra.ac.id](mailto:siti.yana_si17@nusaputra.ac.id)<sup>1)</sup>, [Adhithia.erfina@nusaputra.ac.id](mailto:Adhithia.erfina@nusaputra.ac.id)<sup>2)</sup>,

[cecep.warman@nusaputra.ac.id](mailto:cecep.warman@nusaputra.ac.id)<sup>3)</sup>

\*Korespodensi: e-mail : [siti.yana\\_si17@nusaputra.ac.id](mailto:siti.yana_si17@nusaputra.ac.id)

## ABSTRAK

*Strategi melawan pandemi Covid-19 dengan pembatasan sosial memaksa semua institusi pendidikan menghentikan proses belajar dan menggantinya dengan pembelajaran daring. Kementerian Pendidikan dan Kebudayaan Republik Indonesia mempertegas kebijakan pembelajaran daring dalam masa pandemi dengan mengeluarkan kebijakan Belajar Dari Rumah (BDR). Kebijakan ini mewajibkan menggunakan jaringan internet melalui perantara smartphone, gadget, komputer, dan aplikasi sebagai pengganti tatap muka. Namun pembelajara daring yang dianggap sebagai strategi kemudian menjadi kontroversi karena perlu adaptasi. Perubahan mendadak dari pembelajaran tatap muka ke pembelajaran daring pada skala besar menyebabkan berbagai respons di masyarakat. Penelitian ini bertujuan untuk menganalisis opini publik terhadap pembelajaran daring pada masa pandemi Covid-19, penelitian ini diambil dari data Twitter dan diolah menggunakan algoritma Naïve Bayes.*

**Kata Kunci :** *Pandemi Covid-19, Pembelajaran Daring, Sentimen, Naïve Bayes.*

## ABSTRACT

*The strategy against the Covid-19 pandemic with social restrictions forced all educational institutions to stop the learning process and replace it with online learning. The Ministry of Education and Culture of the Republic of Indonesia has emphasized the online learning policy during the pandemic by issuing the Learning From Home (BDR) policy. This policy requires using the internet network through the intermediary of smartphones, gadgets, computers, and applications as a substitute for face-to-face. However, online learning, which was considered a strategy, became controversial because it needed adaptation. The sudden change from face-to-face learning to online learning on a large scale caused various responses in society. This study aims to analyze public opinion on online learning during the Covid-19 pandemic, this research was taken from Twitter data and processed using the Naïve Bayes algorithm.*

**Keywords:** *Covid-19 Pandemic, Online Learning, Sentiment, Naïve Bayes.*

## I. PENDAHULUAN

Kehadiran wabah pandemi Covid-19, merubah cara orang berpikir dan melakukan kegiatan sehari-hari. Covid-19 menimbulkan krisis kesehatan diseluruh dunia. Sekolah-sekolah pada seluruh jenjang dari tingkat dasar hingga perguruan tinggi di banyak negara ditutup untuk sementara waktu. Keadaan darurat ini, memaksa sebagian besar kegiatan termasuk diantaranya adalah pembelajaran dilakukan secara daring. Namun pembelajara daring yang dianggap sebagai strategi kemudian menjadi



kontroversi karena perlu adaptasi. Perubahan mendadak dari pembelajaran tatap muka ke pembelajaran daring pada skala besar menyebabkan berbagai respons di masyarakat.

Analisis sentimen merupakan teknik yang digunakan untuk melakukan evaluasi dan mengidentifikasi emosi dan opini baik positif maupun negatif. Data kumpulan opini yang akan dianalisis adalah data berupa teks yang dapat diambil dari kolom komentar yang ada di Twitter. Dalam Text mining suatu kelas lebih bersifat area subjek umum, maka diperlukan klasifikasi. Terdapat beberapa algoritma dalam klasifikasi salah satunya yaitu Naïve Bayes, merupakan sebuah metode klasifikasi yang mengacu pada teori Bayes yang digunakan untuk menghitung probabilitas ketidakpastian data. Maka dari itu diperlukan proses untuk memilih fitur yang relevan.

Oleh karena itu perlu dilakukan analisis sentimen yang bertujuan untuk menganalisis opini publik terhadap pembelajaran daring pada masa pandemi Covid-19 di Indonesia. Penelitian ini menggunakan kata kunci 'pembelajaran daring' pada tweet-tweet dalam bahasa Indonesia yang ditarik dari Twitter dan diolah menggunakan algoritma Naïve Bayes untuk menghasilkan analisis sentimen terhadap pembelajaran daring. Penelitian ini diharapkan dapat menjadi pembelajaran untuk tetap melaksanakan kegiatan belajar dari rumah pada masa pandemi Covid-19, dan dapat meningkatkan motivasi belajar dalam pelaksanaan pembelajaran daring.

## II. TINJAUAN PUSTAKA

### 2.1 Text Mining

*Text mining* adalah proses dimana sumber data biasanya didapatkan dari dokumen untuk mencari kata-kata yang dapat mewakili isi dokumen sehingga dapat dilakukan analisis [1]. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen, data yang digunakan dalam *text mining* adalah sekumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari *text mining* yaitu pengelompokan teks [2].

### 2.2 Analisis Sentimen

Analisis sentimen adalah suatu bidang yang berlangsung dalam penelitian berbasis teks. Analisis sentimen atau *opinion mining* adalah kajian tentang cara untuk memecahkan masalah dari opini masyarakat, sikap dan emosi suatu entitas tersebut dapat mewakili individu. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung beropini negatif atau positif [3].

### 2.3 Media Sosial Twitter

Pada dasarnya media sosial merupakan perkembangan mutakhir dari teknologi-teknologi web berbasis internet, yang memudahkan semua orang untuk dapat berkomunikasi, berpartisipasi, saling berbagi dan membentuk sebuah jaringan secara *online*, sehingga dapat menyebarluaskan konten mereka sendiri, diantara media sosial yang populer yaitu *Twitter* [4]. *Twitter* memberikan fasilitas bagi pengguna untuk dapat menuliskan pesan dalam *Twitter* hanya berisi 140 karakter. *Twitter* merupakan salah satu jejaring social yang paling mudah digunakan, karena hanya memerlukan waktu yang singkat tetapi informasi yang disampaikan dapat langsung menyebar secara luas [5].

### 2.4 Preprocessing

*Preprocessing* merupakan proses untuk mempersiapkan data mentah sebelum dilakukan proses lain. Pada umumnya, *preprocessing* data dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data menjadi bentuk yang lebih mudah yang diproses oleh sistem [6].

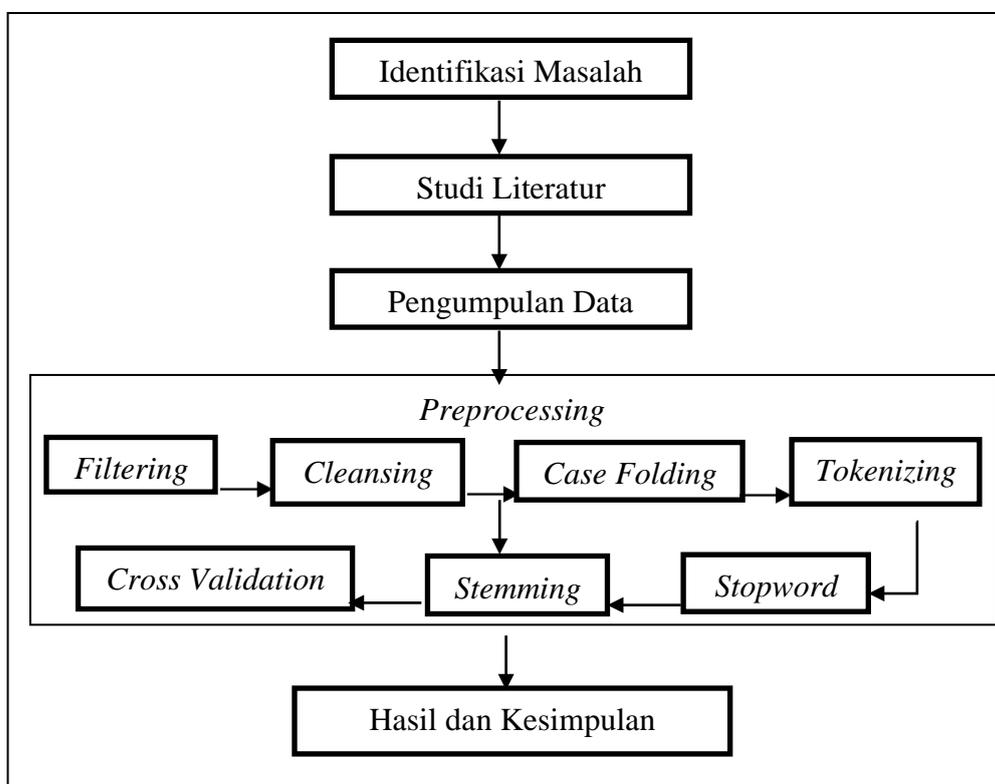
## 2.5 Algoritma *Naïve Bayes*

Algoritma *Naïve Bayes* merupakan metode algoritma yang dapat melakukan klasifikasi data pada kelas tertentu *Naïve Bayes* menempuh dua tahap pelatihan dilakukan proses analisis terhadap sampel dokumen berupa pemilihan *vocabulary*, yaitu kata yang mungkin muncul dalam koleksi dokumen sampel yang sedapat mungkin dapat menjadi representasi dokumen [7].

### III. METODE PENELITIAN

Metode penelitian merupakan suatu proses berpikir untuk menentukan masalah, mengumpulkan data baik melalui studi literatur maupun melalui studi lapangan, melakukan pengolahan data hingga memberikan kesimpulan dari permasalahan yang diteliti. Dengan metodologi maka siklus pemecahan masalah pada penelitian dapat dilaksanakan secara baik dan terstruktur.

Metode penelitian yang akan digunakan pada penelitian ini sebagaimana ditunjukkan pada Gambar 1 dibawah ini :

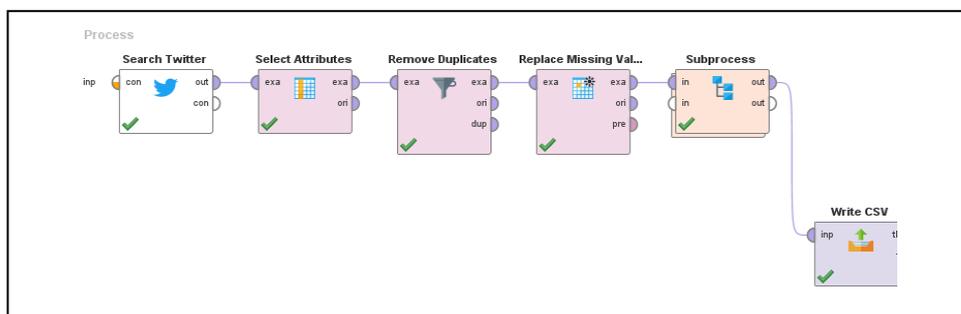


**Gambar 1.** Metodologi Penelitian

Dalam penelitian ini melalui beberapa tahapan yang dilakukan yaitu identifikasi masalah, pencarian literatur, pengumpulan data dengan cara menggunakan *crawling* dengan kata kunci “Belajar Daring”, aplikasi yang digunakan yaitu menggunakan *Rapid Miner 9.9.2*. Setelah melakukan *crawling* data, kemudian dilakukan proses *preprocessing* dengan tahapan *filtering*, *cleansing*, *case folding*, *tokenizing*, *stopword*, *stemming*, dan *cross validation* menggunakan algoritma *Naïve Bayes* dengan bentuk keluaran sistem adalah nilai akurasi. Proses selanjutnya yaitu penggunaan seleksi fitur Chi Square, terakhir yaitu hasil dan kesimpulan dari penelitian yang telah dilakukan.

## IV. HASIL DAN PEMBAHASAN

### 4.1 Pengumpulan Data



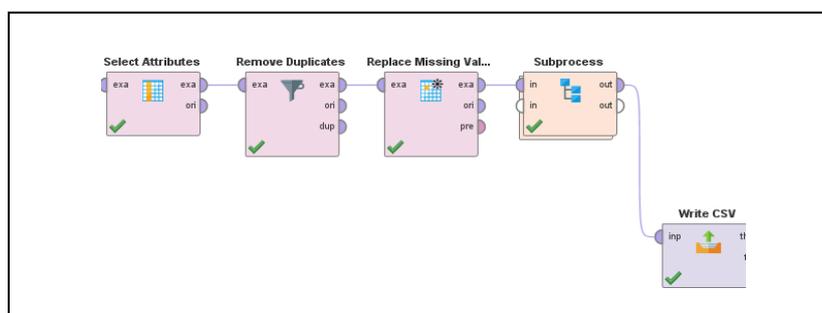
**Gambar 2.** *Crawling Data*

Data yang digunakan pada penelitian ini diambil dari kumpulan *tweets* bahasa Indonesia dengan kata kunci “belajar daring”, dari tahap pengumpulan data didapatkan hasil crawling data sebanyak 2000 data.

### 4.2 Preprocessing

*Preprocessing* merupakan proses untuk mempersiapkan data mentah sebelum dilakukan proses lain. Pada umumnya, *preprocessing* data dilakukan dengan cara mengeliminasi data yang tidak sesuai atau mengubah data menjadi bentuk yang lebih mudah yang diproses oleh sistem. Berikut tahapan *preprocessing* adalah sebagai berikut :

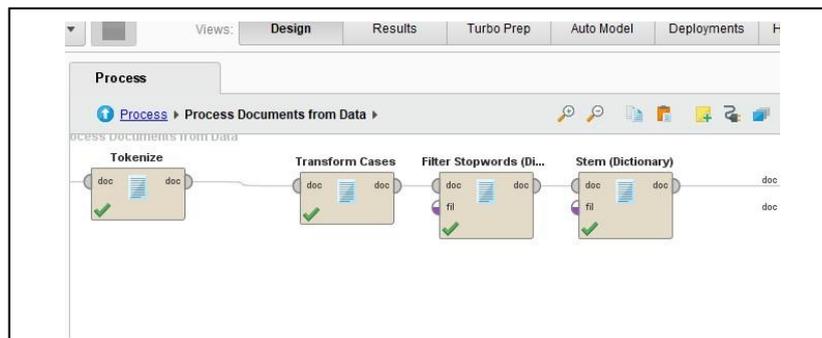
#### 4.2.1 Filtering dan Cleansing



**Gambar 3.** *Proses filtering dan cleansing*

Pada tahapan ini dilakukan proses *filtering* dan *cleansing* yaitu menggunakan *select atribut* dan *remove duplicate* untuk mengambil kata-kata penting, membersihkan dokumen dari kata yang tidak diperlukan, dan membersihkan kata yang *double* atau duplikat. Hasil dari data *filtering* dan *cleansing* menghasilkan data sebanyak 700 data.

Sebelum proses selanjutnya dataset akan meliputi tahap pelabelan manual terhadap data opini menggunakan Microsoft Excell. Dataset dilabelkan kedalam dua kelas sentimen yakni kelas positif dan negatif.



**Gambar 3.** Tahap *Preprocessing*

#### 4.2.2 Case Folding

Proses selanjutnya yaitu proses case folding dengan menggunakan *transform cases* untuk mengubah teks menjadi huruf kecil.

#### 4.2.3 Tokenizing

Pada proses ini yaitu untuk menghapus semua karakter non alphabet, misalnya simbol, spasi, angka menggunakan tokenize.

#### 4.2.4 Stemming

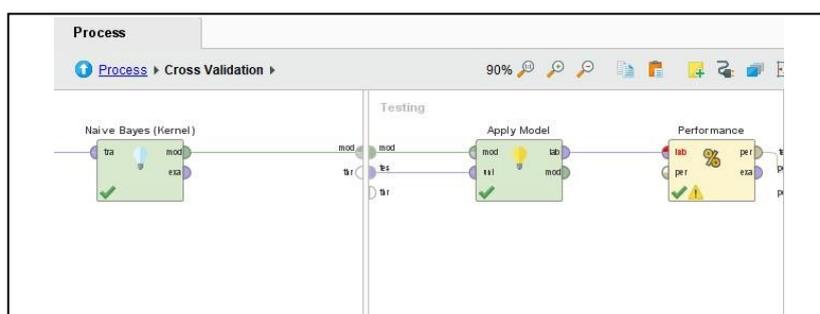
Proses dilanjutkan dengan menghilangkan imbuhan pada masing-masing kata sehingga menjadi kata dasar, dan dalam tahap ini bertujuan untuk membersihkan suatu kata dari pengejaan yang kurang tepat.

#### 4.2.5 Stopword

Stopword diproses pada sebuah kalimat yang mengandung kata-kata yang sering keluar dan dianggap tidak penting seperti waktu atau penghubung. Untuk itu perlu penghapusan.

#### 4.2.6 Cross Validation

Proses *cross validation* menggunakan algoritma *Naïve Bayes* yang meliputi proses *training* dan *testing* (*apply model dan performance*). Setelah itu dilakukan proses pemodelan dengan menggunakan algoritma *Naïve Bayes* untuk mendapatkan nilai *accuracy*.



**Gambar 4.** Tahap *Cross Validation Process-Naïve Bayes*



The screenshot shows a software interface with a sidebar on the left containing a 'Criterion' list with 'accuracy' selected. The main area has 'Table View' selected and displays the following information:

accuracy: 70.25% +/- 3.81% (micro average: 70.25%)

	true negatif	true positif	class precision
pred. negatif	140	24	85.37%
pred. positif	95	141	59.75%
class recall	59.57%	85.45%	

**Gambar 4.** Nilai *accuracy* algoritma *Naïve Bayes* pada sentimen pembelajaran daring pada masa pandemi Covid-19

Nilai *accuracy* yang didapatkan algoritma *Naïve Bayes* adalah 70.25%, dengan nilai negatif 85.37% sedangkan positif 50.75%.

Berdasarkan analisis dan pengujian yang telah dilakukan, maka mendapatkan hasil sebagai berikut :

1. Data yang didapatkan dari hasil *crawling* data berupa data opini sebanyak 2000 data yang dikumpulkan.
2. Implementasi algoritma *Naïve Bayes* pada penelitian ini menghasilkan nilai *accuracy* 70.25%, dengan nilai negatif 83.37% sedangkan positif 50.75%.

## V. KESIMPULAN

Penelitian ini menerapkan algoritma *Naïve Bayes* untuk klasifikasi sentimen. Analisis sentimen menggunakan data Twitter dengan kata kunci “belajar daring” dalam bahasa Indonesia. Penelitian ini menunjukkan 50.75% sentimen positif, 83.3% sentimen negatif. Persepsi negatif dihasilkan karena ketidakpuasan masyarakat terhadap pembelajaran daring. Beberapa Twit menunjukkan kekecewaan dengan kata ‘stress’, ‘malas’, ‘bosan’. Berdasarkan hasil ini pembelajaran daring belum maksimal diterapkan di Indonesia pada masa pandemi Covid-19. Penelitian lebih lanjut diharapkan menerapkan algoritma yang berbeda untuk mendapatkan hasil yang lebih akurat dalam analisis opini.

## DAFTAR PUSTAKA

- [1] S. Anisah, A. S. Honggowibowo, and A. Pujiastuti, “Klasifikasi Teks Menggunakan Chi Square Feature Selection Untuk Menentukan Komik Berdasarkan Periode, Materi Dan Fisik dengan Algoritma Naive Bayes,” *Compiler*, vol. 5, no. 2, pp. 59–66, 2016, doi: 10.28989/compiler.v5i2.171.
- [2] A. Nisa, E. Darwiyanto, and I. Asror, “Analisis Sentimen Menggunakan Naive Bayes Classifier dengan Chi-Square Feature Selection Terhadap Penyedia Layanan Telekomunikasi,” *e-Proceeding Eng.*, vol. 6, no. 2, pp. 8650–8659, 2019.
- [3] T. Rosandy, “PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DENGAN METODE DECISION TREE (C4.5) UNTUK MENGANALISA KELANCARAN PEMBIAYAAN (Study Kasus: KSPPS / BMT AL-FADHILA,” *J. Teknol. Inf. Magister Darmajaya*, vol. 2, no. 01, pp. 52–62, 2016.
- [4] E. F. U. Latifah, “Perbandingan Kinerja Machine Learning Berbasis Algoritma Support Vector Machine dan Naive Bayes (Studi Kasus: Data Tanggapan Mengenai Traveloka Melalui Media Sosial Twitter),” 2018, [Online]. Available: <https://dspace.uin.ac.id/handle/123456789/7996>.



- [5] T. W. U. Imelya Susianti, Sri Soerya Ningsih, M. Al Haris, “Analisis Sentimen Pada Twitter Terkait New Normal Dengan Metode Naïve Bayes Classifier,” *Pros. Semin. Edusainstech FMIPA UNIMUS*, pp. 354–363, 2020, [Online]. Available: <https://prosiding.unimus.ac.id/index.php/edusaintek/article/view/576/578>.
- [6] J. LING, I. P. E. N. KENCANA, and T. B. OKA, “Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square,” *E-Jurnal Mat.*, vol. 3, no. 3, p. 92, 2014, doi: 10.24843/mtk.2014.v03.i03.p070.
- [7] I. Listiowarni and E. R. Setyaningsih, “Feature Selection Chi-Square dan K-NN pada Pengkategorian Soal Ujian Berdasarkan Cognitive Domain Taksonomi Bloom,” *J. Komput. Terap.*, vol. 4, no. 1, pp. 21–30, 2018.