

**ANALISIS SENTIMEN PEMBELAJARAN *ONLINE* SELAMA  
PANDEMI VIRUS CORONA (COVID-19) MENGGUNAKAN  
METODE *NAÏVE BAYES***

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**NIA HASTARI**

**17.11.0994**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**ANALISIS SENTIMEN PEMBELAJARAN *ONLINE* SELAMA  
PANDEMI VIRUS CORONA (COVID-19) MENGGUNAKAN  
METODE *NAÏVE BAYES***

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**NIA HASTARI**

**17.11.0994**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN PEMBELAJARAN *ONLINE* SELAMA PANDEMI  
VIRUS CORONA (COVID-19) MENGGUNAKAN METODE *NAÏVE*  
*BAYES***

yang disusun dan diajukan oleh

**Nia Hastari**

**17.11.0994**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 03 April 2023

**Dosen Pembimbing,**



**Yoga Pristyanto, S.Kom, M.Eng**

**NIK. 190302412**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN PEMBELAJARAN *ONLINE* SELAMA PANDEMI  
VIRUS CORONA (COVID-19) MENGGUNAKAN METODE *NAÏVE*  
*BAYES***

yang disusun dan diajukan oleh

**Nia Hastari**

**17.11.0994**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 03 April 2023

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Supriatin, M.Kom**  
**NIK. 190302239**

**Kusnawi, S.Kom., M.Eng.**  
**NIK. 190302112**

**Yoga Pristyanto, S.Kom., M.Eng.**  
**NIK. 190302412**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 09 Mei 2023

**DEKAN FAKULTAS ILMU KOMPUTER**



**Hanif Al Fatta, S.Kom., M.Kom.**  
**NIK. 190302096**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nia Hastari  
NIM : 17.11.0994

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Pembelajaran *Online* Selama Pandemi Virus Corona (COVID-19) Menggunakan Metode Naïve Bayes**

Dosen Pembimbing : Yoga Pristyanto, S.Kom, M.Eng.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 3 April 2023

Yang Menyatakan,



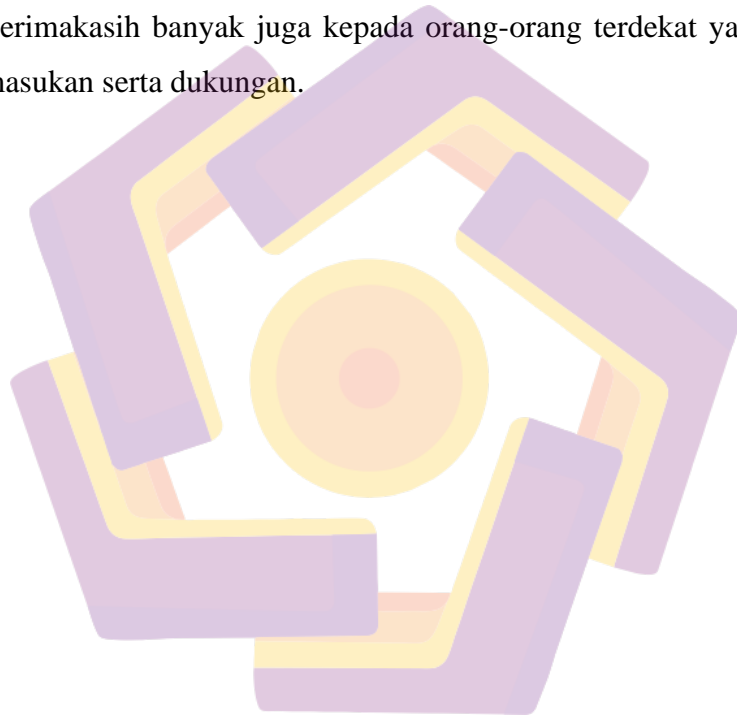
10000  
SERBUH RUPIAH  
METERAN  
TEMPEL  
09B20AJX943700746

Nia Hastari

## HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur, skripsi ini saya persembahkan.

1. Terimakasih untuk diri saya sendiri, yang tetap bertahan dan berjuang.
2. Terimakasih sebesar-besarnya kepada Bapak dan Mamak tercinta yang selalu sabar, menerima atas segala keterbatasan saya dan selalu mendukung dari segi moral, spiritual, psikis, dan materi yang tiada habis-habisnya untuk memastikan pencapaian yang ingin saya raih.
3. Terimakasih banyak juga kepada orang-orang terdekat yang memberikan masukan serta dukungan.



## KATA PENGANTAR

*Assalamu'alaikum Warahmatullahi Wabarakatuh.*

*Alhamdulillahirabbil'Alamin.* Segala puji bagi Allah SWT yang telah memberikan rahmat, hidayah, dan ridha-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Pembelajaran Online Selama Pandemi Virus Corona (COVID-19) Menggunakan Metode Naïve Bayes”.

Skripsi ini disusun dan diajukan untuk memenuhi syarat perolehan gelar Sarjana Informatika(S.Kom) pada Fakultas Ilmu Komputer di Universitas AMIKOM Yogyakarta. Disamping itu, penulisan skripsi ini juga bertujuan untuk memberikan pengetahuan kepada pembaca.

Dalam menyelesaikan skripsi ini, banyak pihak yang telah memberikan bantuan. Atas segala bantuan yang telah diberikan, penulis ingin menyampaikan doa dan ucapan terimakasih yang tak terhingga kepada:

1. M. Suyanto, Prof, Dr, M.M. selaku rektor Universitas AMIKOM Yogyakarta
2. Hanif Al Fatta,S.Kom., M.Kom.selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta
3. Windha Mega Pradnya Duhita, M.Kom. selaku ketua Program Studi Informatika Universitas AMIKOM Yogyakarta
4. Andika Agus Slameto, M.Kom.. selaku dosen wali selama saya menempuh Pendidikan S1 di Universitas AMIKOM Yogyakarta
5. Yoga Pristyanto, S.Kom., M.Eng. selaku dosen pembimbing skripsi yang telah memberikan banyak pengarahan
6. Segenap civitas akademika jurusan Informatika, terutama seluruh dosen, terimakasih atas segenap ilmu dan bimbingannya.

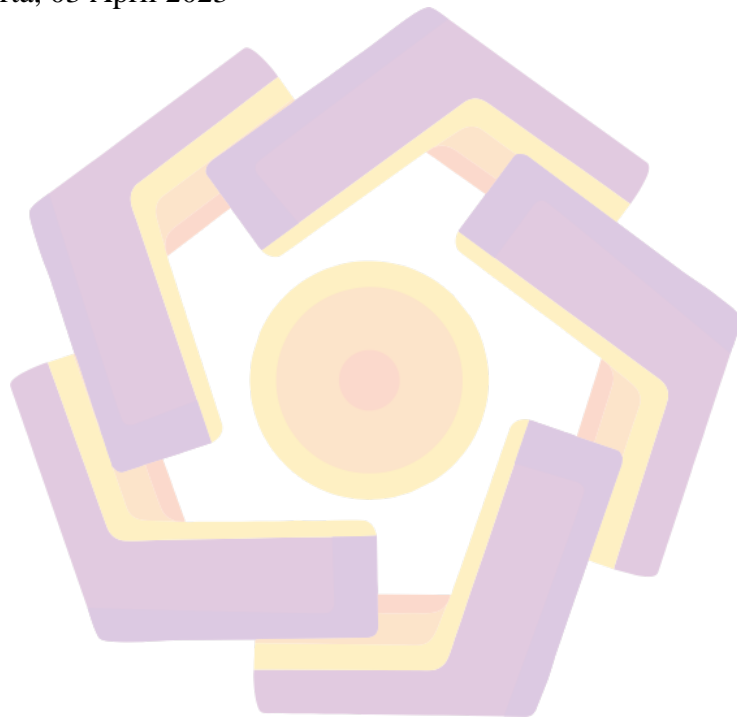
Atas segala kekurangan dan ketidak sempurnaan skripsi ini, penulis sangat mengharapkan masukan, kritik, dan saran yang bersifat membangun ke arah

perbaikan dan penyempurnaan skripsi ini, agar dalam penyusunan karya tulis selanjutnya dapat lebih baik.

Penulis berharap semoga skripsi ini dapat bermanfaat bagi seluruh pihak dan semoga amal baik yang telah diberikan kepada penulis mendapat balasan dari Allah SWT.

Yogyakarta, 03 April 2023

Penulis





## DAFTAR ISI

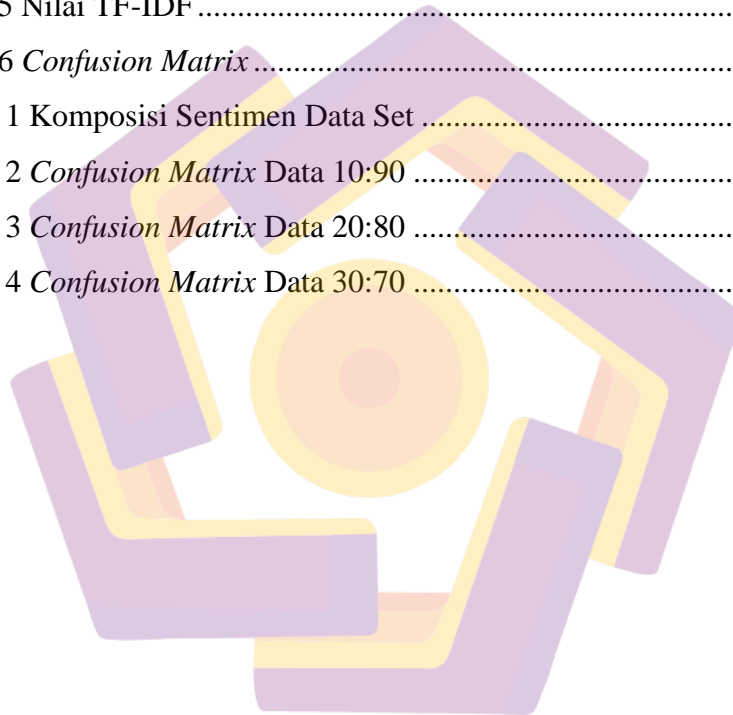
HALAMAN JUDUL .....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN .....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI .....	iv
HALAMAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
DAFTAR ISTILAH .....	xiii
INTISARI .....	xiv
<i>ABSTRACT</i> .....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah .....	2
1.4 Maksud dan Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Metode Penelitian .....	3
1.6.1 Metode Pengumpulan Data.....	4
1.6.1.1 Metode Kepustakaan.....	4
1.6.1.2 Pengumpulan Data .....	4
1.6.2 Analisa Kebutuhan.....	4
1.6.2.1 Metode Analisis Data.....	4
1.7 Sistematika Penulisan .....	5
BAB II TINJAUAN PUSTAKA .....	7
2.1 Kajian Pustaka .....	7
2.2 Landasan Teori .....	13
2.1.1 <i>Data Mining</i> .....	13

2.1.2	<i>Text Mining</i> .....	13
2.1.3	<i>Text Preprocessing</i> .....	14
2.1.4	Analisis Sentimen .....	14
2.1.5	Instagram.....	15
2.1.6	<i>Naïve Bayes Classifier</i> .....	16
2.1.7	Pembobotan kata .....	17
2.1.8	<i>Python</i> .....	18
2.1.9	<i>Confusion Matrix</i> .....	18
BAB III METODE PENELITIAN .....		20
3.1	Objek Penelitian.....	20
3.2	Alur Penelitian .....	20
3.1.1	Pengumpulan Dataset.....	21
3.1.2	Pelabelan Dataset .....	21
3.1.3	Preprocessing Data.....	21
3.1.3.1	<i>Cleansing</i> .....	22
3.1.3.2	<i>Case Folding</i> .....	22
3.1.3.3	<i>Tokenizing</i> .....	22
3.1.3.4	<i>Convert Slang Word</i> .....	23
3.1.3.5	<i>Stopword Removal</i> .....	23
3.1.3.6	<i>Stemming</i> .....	23
3.1.4	Seleksi Data .....	24
3.1.5	Pembobotan TF-IDF dan <i>Naïve Bayes Classifier</i> .....	25
3.1.5.1	<i>Term Frequency</i> .....	26
3.1.5.2	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> .....	30
3.1.6	Evaluasi.....	31
3.3	Alat dan Bahan.....	33
3.3.1	Data Penelitian .....	33
3.3.2	Alat Penelitian.....	33
BAB IV HASIL DAN PEMBAHASAN .....		34
4.1	Pengumpulan Dataset .....	34
4.2	<i>Preprocessing Data</i> .....	35

4.2.1	<i>Cleansing</i> .....	35
4.2.2	<i>Case Folding</i> .....	36
4.2.3	<i>Tokenization</i> .....	36
4.2.4	<i>Convert Slang Word</i> .....	36
4.2.5	Stopword Removal.....	37
4.2.6	<i>Stemming</i> .....	37
4.3	Seleksi Data .....	38
4.3.1	Perbandingan 10:90 .....	38
4.3.2	Perbandingan 20:80 .....	39
4.3.3	Perbandingan 30:70 .....	40
4.4	Analisis dan Evaluasi.....	40
4.4.1	Analisis .....	40
4.4.1.1	TF-IDF .....	40
4.4.1.2	<i>Naïve Bayes Classifier</i> .....	41
4.4.1.3	Pengujian dengan <i>Confusion Matrix</i> .....	42
4.4.1.3.1	Pengujian 1.....	42
4.4.1.3.2	Pengujian 2.....	43
4.4.1.3.3	Pengujian 3.....	44
BAB V	PENUTUP .....	43
5.1	Kesimpulan .....	43
5.2	Saran .....	43
DAFTAR	PUSTAKA .....	45

## DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian .....	9
Tabel 2. 2 <i>Confusion Matrix</i> .....	18
Tabel 3. 1 Contoh Hasil <i>Preprocessing</i> .....	24
Tabel 3. 2 Contoh Koleksi Data Training .....	25
Tabel 3. 3 <i>Term Frequency</i> .....	26
Tabel 3. 4 Nilai <i>Term Frequency</i> .....	28
Tabel 3. 5 Nilai TF-IDF .....	31
Tabel 3. 6 <i>Confusion Matrix</i> .....	32
Tabel 4. 1 Komposisi Sentimen Data Set .....	35
Tabel 4. 2 <i>Confusion Matrix</i> Data 10:90 .....	42
Tabel 4. 3 <i>Confusion Matrix</i> Data 20:80 .....	43
Tabel 4. 4 <i>Confusion Matrix</i> Data 30:70 .....	44



## DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian .....	20
Gambar 3. 2 Alur Preprocessing .....	22
Gambar 4. 1 Kumpulan Data Set .....	34
Gambar 4. 2 Perintah <i>Cleansing</i> .....	35
Gambar 4. 3 Perintah <i>Case Folding</i> .....	36
Gambar 4. 4 Perintah <i>Tokenization</i> .....	36
Gambar 4. 5 Perintah <i>Convert Slang Word</i> .....	36
Gambar 4. 6 Perintah <i>Stopword Removal</i> .....	37
Gambar 4. 7 Perintah <i>Stemming</i> .....	37
Gambar 4. 8 Sampel <i>Data Preprocessing</i> .....	38
Gambar 4. 9 Perintah <i>Split Data</i> 10:90 .....	38
Gambar 4. 10 Diagram Pembagian Data Testing 10% .....	39
Gambar 4. 11 Diagram Pembagian Data Training 90% .....	39
Gambar 4. 12 Perintah split data 20:80.....	39
Gambar 4. 13 Diagram Pembagian Data Training 80% .....	39
Gambar 4. 14 Diagram Pembagian Data Testing 20% .....	39
Gambar 4. 15 Perintah split data 30:70.....	40
Gambar 4. 16 Diagram pembagian data training 70% .....	40
Gambar 4. 17 Diagram pembagian data testing 30% .....	40
Gambar 4. 18 Perintah TF-IDF .....	41
Gambar 4. 19 Hasil Pembobotan TF-IDF .....	41
Gambar 4. 20 Perintah <i>Naïve Bayes Classifier</i> .....	41
Gambar 4. 21 Hasil Implementasi Model Klasifikasi Pada Data Testing .....	45

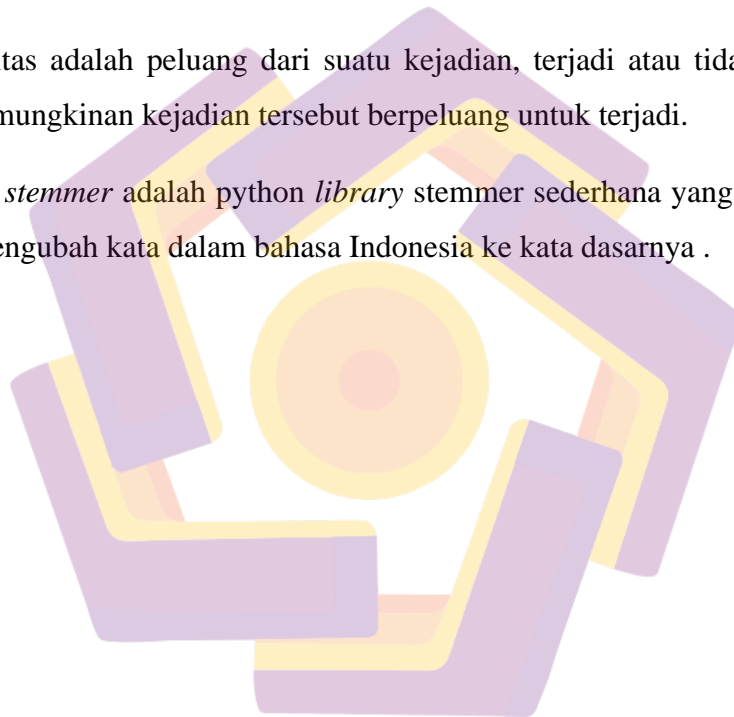
## DAFTAR ISTILAH

*Library* yaitu sekumpulan kode yang memiliki fungsi-fungsi tertentu yang ditulis dalam bahasa pemrograman, dan dapat dipanggil kedalam program lain.

*Python* adalah bahasa pemrograman interpretatif multiguna dengan kode yang mudah diimplementasikan dapat digunakan di berbagai platform yang berfokus pada keterbacaan kode dan merupakan bahasa populer yang berkaitan dengan *Data Science*.

Probabilitas adalah peluang dari suatu kejadian, terjadi atau tidak dan seberapa besar kemungkinan kejadian tersebut berpeluang untuk terjadi.

Sastrawi *stemmer* adalah python *library* *stemmer* sederhana yang memungkinkan untuk mengubah kata dalam bahasa Indonesia ke kata dasarnya .



## INTISARI

Virus Corona yang menyerang saluran pernapasan manusia hingga menyebabkan kematian menyebar ke berbagai dunia termasuk Indonesia. Penyebaran virus membuat pemerintah mengeluarkan kebijakan salah satunya kebijakan untuk belajar dari rumah yang memunculkan banyak pendapat oleh masyarakat mengenai kebijakan tersebut, salah satunya pendapat yang dikeluarkan pada media sosial Instagram melalui kolom komentar pada beberapa postingan akun @nadiemmakarim selaku Menteri Pendidikan.

Pendapat yang dikeluarkan masyarakat dapat menjadi peluang pemanfaatan data untuk menghasilkan suatu pengetahuan dengan mengklasifikasikan komentar-komentar menggunakan teknik analisis sentiment dengan melalui tahapan *preprocessing data* berupa *cleansing*, *case folding*, *tokenizing*, *convert slang word*, *stopword removal* dan *stemming*. Kemudian dilakukan pembobotan pada kata dengan *term frequency – invers document frequency* (TF-IDF) untuk dilanjutkan proses klasifikasi ke dalam kelas positif, netral dan negatif menggunakan metode *Naïve Bayes Classifier*.

Pengujian menggunakan *naïve bayes* dilakukan dengan memanfaatkan permodelan data uji dan data latih yang berbeda yaitu dengan perbandingan 10:90, 20:80 dan 70:30 dengan jumlah data total sebanyak 2875 data. Dari hasil penelitian, algoritma *Naïve Bayes* dinilai cukup baik dalam melakukan analisis sentimen pembelajaran online selama pandemi covid-19. Di dapatkan hasil akurasi tertinggi yaitu sebesar 70,49% dengan permodelan 10% data testing dan 90% data training. Hal ini menunjukkan semakin besar jumlah data latih yang digunakan semakin baik tingkat akurasi yang dihasilkan.

**Kata kunci:** sentimen, *naïvebayes*, Instagram, belajar, online.

## **ABSTRACT**

*The Corona virus which attacks the human respiratory tract to cause death has spread to various parts of the world, including Indonesia. The spread of the virus made the government issue policies, one of which was a policy to learn from home, which gave rise to many opinions by the public regarding this policy, one of which was the opinion issued on Instagram social media through the comments column on several account posts @nadiemmakarim as the Minister of Education.*

*Opinions issued by the public can be an opportunity to use data to produce knowledge by classifying comments using sentiment analysis techniques by going through the stages of data preprocessing in the form of cleansing, case folding, tokenizing, converting slang words, stopword removal and stemming. Then the words are weighted with term frequency – inverse document frequency (TF-IDF) to continue the classification process into positive, neutral and negative classes using the Naïve Bayes Classifier method.*

*Testing using naïve Bayes was carried out by utilizing different modeling of test data and training data, namely with a ratio of 10:90, 20:80 and 70:30 with a total of 2875 data. From the results of the study, the Naïve Bayes algorithm is considered quite good at analyzing online learning sentiments during the co-19 pandemic. The highest accuracy results were obtained, namely 70.49% with 10% testing data modeling and 90% training data. This shows that the greater the amount of training data used, the better the level of accuracy produced.*

**Keyword:** *sentiment, naïvebayes, Instagram, learn, online.*