



## Analisis Sentimen terhadap Tugas Akhir Mahasiswa Menggunakan Naive Bayes

Rahmi Imanda<sup>1</sup>, Bagus Arfian Laksono<sup>2</sup>, Catur Meingar Subarkah<sup>3</sup>, Ikhwan Ghazali<sup>4</sup>, Yudo Novianto Putra<sup>5</sup>

<sup>12345</sup>Program Studi Teknik Informatika, Universitas Muhammadiyah Prof. DR. HAMKA (Uhamka)

Email: [rahmi.imanda@uhamka.ac.id](mailto:rahmi.imanda@uhamka.ac.id), [bagusarfnlaksono@gmail.com](mailto:bagusarfnlaksono@gmail.com), [diatmancaturms@gmail.com](mailto:diatmancaturms@gmail.com),  
[ikhwan06ghazali@gmail.com](mailto:ikhwan06ghazali@gmail.com), [yudhovianto@gmail.com](mailto:yudhovianto@gmail.com)

### Abstrak

Mahasiswa tingkat akhir sering mengalami kecemasan dalam menyelesaikan tugas akhir atau skripsi, yang dipicu oleh tekanan internal seperti perfeksionisme dan kurangnya kepercayaan diri, serta tekanan eksternal seperti tuntutan keluarga dan ketidakpastian masa depan. Penelitian ini bertujuan untuk menganalisis sentimen kecemasan mahasiswa terhadap tugas akhir menggunakan algoritma Naive Bayes. Metode yang digunakan mencakup pengumpulan data dari media sosial, preprocessing teks, labeling sentimen, visualisasi melalui word cloud, dan pemodelan klasifikasi. Model Naive Bayes yang dibangun mampu mencapai akurasi sebesar 73,95%, dengan performa lebih baik dalam mendeteksi sentimen positif. Hasil analisis menunjukkan bahwa mahasiswa dengan kecemasan cenderung menggunakan kata-kata bernada negatif, sedangkan dukungan sosial dan manajemen waktu menjadi faktor dominan dalam sentimen positif. Kesimpulannya, pendekatan ini efektif untuk memetakan kondisi psikologis mahasiswa dan dapat digunakan sebagai dasar pengembangan sistem dukungan berbasis teknologi dalam lingkungan akademik..

**Kata Kunci :** Kecemasan, Mahasiswa, Tugas Akhir, Analisis Sentimen, Naive Bayes

### PENDAHULUAN

Mahasiswa adalah seseorang yang sedang menempuh pendidikan pada tingkat perguruan tinggi, seperti universitas, institusi, sekolah tinggi, politeknik, atau akademi. Di akhir masa studinya, mahasiswa harus menyelesaikan tugas akhir yang merupakan salah satu syarat untuk memperoleh gelar sarjana. Namun, sekitar 17% mahasiswa mengalami kecemasan selama masa kuliah, khususnya saat mengerjakan skripsi. (Nur Faisyah et al., 2024)

Kecemasan dapat berupa respons fisiologis untuk mengantisipasi masalah yang mungkin muncul atau dapat menyebabkan gangguan jika berlebihan (Prabowo et al., 2021). Sampai saat ini, kecemasan terus menjadi penyakit masyarakat. Rasa gelisah dan cemas biasanya dianggap sebagai gejala gangguan mental atau penyakit jiwa, tetapi rasa cemas yang berlebihan juga dapat membahayakan bagian tubuh

Proses menyusun skripsi tidaklah lebih mudah seperti dengan menulis makalah untuk tugas kuliah setiap minggu. Menurut buku panduan penulisan tugas akhir atau skripsi, proses penyusunan tugas akhir/skripsi harus mengikuti tahapan yang terstruktur, diawali dari proses penyusunan proposal penelitian, pelaksanaan penelitian dan penulisan skripsi, pelaksanaan ujian skripsi, proses revisi, hingga akhirnya penyerahan skripsi. (Malfasari et al., 2018)

Setiap mahasiswa dapat memiliki respons yang berbeda terhadap tantangan dalam menyelesaikan skripsi. Menurut beberapa mahasiswa, tantangan tersebut dianggap sebagai beban berat yang sulit diatasi, sehingga menurunkan motivasi mahasiswa dalam menyelesaikan skripsi. Selain itu, faktor waktu juga berpengaruh dalam proses ini. Mahasiswa yang tidak dapat menyelesaikan tugas tepat waktu cenderung mengalami tekanan lebih besar dibandingkan mereka yang menyelesaikannya sesuai jadwal.

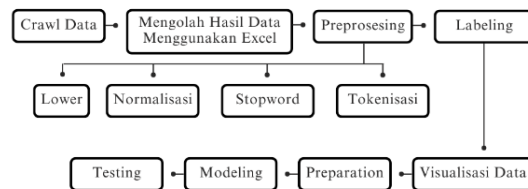
Dari penelitian (Prabowo et al., 2021) Berdasarkan hasil survei melalui kuesioner, disimpulkan bahwa beberapa faktor yang dapat memicu gangguan kecemasan meliputi harapan tinggi dari keluarga terhadap anaknya serta kebiasaan individu, seperti menuntut kesempurnaan dalam pekerjaan, ketergantungan pada sosok yang lebih kuat, kurangnya rasa percaya diri, dan kecenderungan menunda pekerjaan penting.

Menurut penelitian (Zakiyah, 2016), Mahasiswa yang sedang menjalani proses penyusunan skripsi kerap mengalami tingkat stres sedang sebesar 46%, ansietas pada tingkat sedang sebesar 52%, serta depresi tingkat sedang sebesar 58%. Kondisi ini dapat mempengaruhi seberapa cepat menyelesaikan skripsi mereka. Menurut Penelitian (Etika et al. 2016), banyak tantangan menghalangi mahasiswa untuk menyelesaikan skripsi mereka.

Tujuan penelitian secara umum ini adalah untuk menganalisis sentimen kecemasan mahasiswa menyusun tugas akhir. Tingkat kecemasan ini dianalisis melalui proses bimbingan, seminar proposal, penyusunan instrumen, dan tindakan peneliti dengan metode Naive Bayes

## METODE

Penelitian ini diambil pada Januari 2020- Februari 2025 diawali dengan proses crawl data, yaitu pengambilan data dari media sosial Twitter menggunakan bahasa pemrograman Python. Setelah data diperoleh, tahap berikutnya adalah pengolahan awal data menggunakan Excel, termasuk menghapus data yang bersifat duplikat. Selanjutnya dilakukan tahap preprocessing untuk membersihkan dan menyiapkan data. Proses ini mencakup beberapa langkah penting seperti lowercasing, normalisasi, stopwords removal, tokenisasi. Setelah data siap, tahap labeling dilakukan untuk mengelompokkan data ke dalam kategori sentimen (positif, negatif, atau netral). Data yang telah dilabeli kemudian melalui tahap visualisasi untuk menggambarkan distribusi sentimen yang ada. Selanjutnya, data memasuki proses preparation yang mencakup pembagian dataset menjadi data latih dan data uji. Model kemudian dibangun pada tahap modeling, dan dievaluasi pada tahap testing untuk mengukur akurasi dalam melakukan analisis sentimen.



**Gambar 1. Diagram Alir**

## HASIL DAN PEMBAHASAN

Hasil Pembahasan dan hasil analisis sentimen terhadap aplikasi X dilakukan dengan menerapkan algoritma Naive Bayes Classifier, di mana proses data dilakukan menggunakan platform Google Colab.

### 1. Crawl Data

Tahap awal dari penelitian dimulai dengan proses crawling data, yaitu pengambilan komentar publik dari media sosial X (sebelumnya dikenal sebagai Twitter), menggunakan kata kunci seperti “anxiety skripsi/tugas akhir”. Proses ini dilakukan dengan bantuan bahasa pemrograman Python dan menghasilkan dataset sebanyak 2.839 komentar. Data ini menjadi pondasi awal dari seluruh proses analisis. Dalam metode Naive Bayes, crawling data menjadi tahapan awal untuk menyediakan dataset yang akan digunakan dalam proses pelatihan (training) dan pengujian (testing) model. Tanpa data yang cukup dan relevan, algoritma Naive Bayes tidak dapat melakukan klasifikasi dengan baik.

```

filename = 'crawl.csv'
search_keyword = 'anxiety skripsi'
limit = 500
!npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST"
-l {limit} --token {twitter_auth_token}
  
```

Menjelaskan tentang crawl data yang menunjukkan untuk tweet yang mengandung kata “ anxiety skripsi/tugas akhir “

```

# Hasil Crawl Data

import pandas as pd

# Specify the path to your CSV file

file_path = f"tweets-data/{filename}"

# Read the CSV file into a pandas DataFrame

df = pd.read_csv(file_path, delimiter=",")

# Display the DataFrame

display(df)

```

Merupakan kode yang digunakan untuk mendownload hasil crawl data yang berupa file csv

tweet_id	tweet_text	sentiment	user_id	user_name
1	Tue Dec 31 09:49:30 kalau terus bercampur dengan kecemasan kapan draf s negatif	https://x.com/cakartuning		
2	Sat Dec 14 16:30:40 hari ini rasanya hari terproduktif aing di tahun 2024 sis negatif	https://x.com/lonlonwinlo		
3	Thu Nov 28 21:43:05 Aku dan vitaminku against the world. Sebelum berangkat positif	https://x.com/onondodam		
4	Sat Nov 09 08:20:19 faktor anime terhadap kecemasan pada mahasiswa ting netral	https://x.com/zeezearel		
5	Sat Oct 24 08:16:45 Dijamin Ampuh! 5 Cara Mengatasi Kecemasan Saat Me positif	https://x.com/akiakuratco		
6	Thu Oct 23 17:57:46 kembali saya ingatkan mengenai judul yang belum sem netral			
7	Thu Sep 19 06:54:58 info joki skripsi dengan judul hubungan komunikasi ter netral	https://x.com/PaiParaPermat30509		
8	Tue Sep 17 10:10:31 @desy_nw Selamat sore. Perkenalkan saya Kanza mahi netral	https://x.com/yoyouwannao		
9	10 Sat Sep 14 08:08:05 + Aku dan vitaminku against the world. Sebelum berangkat positif			
10	2828 Wed Mar 23 14:28:06 @jagamrks golek bantuan profesional rider mental lwo positif	https://x.com/yoyoshayoshi_		
2829	Wed Mar 23 01:23:31 enjoying my life ~ assalamualaikum revisian & me netral	https://x.com/lavlavienroc		
2830	Sun Mar 20 11:45:51 @jusussie219 makasih rujee kamu juga ya take care of positif	https://x.com/del delightful_tata		
2831	Thu Mar 17 16:51:00 Kalau skripsi adalah bentuk praktiknya Melit berart positif	https://x.com/beebeelehan		
2832	Mon Mar 14 12:07:27 preparing my mental health for bimbingan skripsi evok positif	https://x.com/aenaeveeth		
2833	Thu Mar 10 14:26:07 isi objek skripsi kuu. jadi bacanya aku yang wslr inggris positif	https://x.com/papacarcial		
2834	Thu Mar 10 13:52:17 isi otak gue banyak bgt sih jujur ya. Mari kita tulis saman netral	https://x.com/titi_titikomati		
2835	Wed Mar 09 14:26:21 Skripsi ini sangat menguji kesabaran ketabahan dan me positif	https://x.com/balbabysweety		
2836	Sun Mar 06 14:21:20 @shyreen Alhamdulillah baik jg kaa. stay safe stay heal positif	https://x.com/pok_jokialitask		
2837	Sun Feb 27 14:50:06 #NKDh Bereskan mental health mu segera Menyelesaika positif	https://x.com/d3-d3dmar1		

Gambar 2. Hasil Crawl Data

merupakan data menunjukkan tampilan output dari hasil membaca dan menampilkan file CSV hasil crawling Twitter

## 2. Labeling

Labeling komentar berdasarkan sentimen, yaitu positif, negatif, atau netral. Proses labeling ini dilakukan secara manual dan dibantu model dasar untuk mempercepat klasifikasi. Hasilnya, sebanyak 1.273 komentar dikategorikan sebagai positif, dan 952 komentar sebagai negative

twit	label
Ya Allah bismillah atas kecemasan ketakutan dan kebimbangan ku atas tugas akhir ini semoga Engkau memudahkan jalanku ya Allah. I just want something good tolong bantu hamba ya Allah. Mudahkan dan lancarkan skripsi journey hamba ini...aamiin	Positive
Tiap hari kecemasan gara² mikirin skripsi Tapi giliran buka revisian rasanya	Negative

<p>udah kaya ga ada harapan</p> <p>sorry kalo salty tapi lu kalo bunuh diri tuh skripsi lu ga kelar orang tua lu sedih lu bakal bikin bayak orang rugi karna mereka sedih kehilangan lo mereka harus meluangkan waktu buat ngelayat. ga ada untungnya malah nyusahin.</p> <p>Tolol amat ama skripsi aja mau bunuh diri. Lebih baik DO aja. Nnt semester depan lanjut lagi</p>	<p>Netral</p> <p>Negative</p>
---	-------------------------------

**Tabel 2. Labeling**

### 3. Preprocessing

Preprocessing bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan, menyamakan format penulisan, dan mempersiapkan data agar lebih mudah dianalisis.

Berikut merupakan data hasil penelitian sampel teks untuk preprocessing dengan mengekstrak kata dan kalimat tertentu. Dimulai melalui data asli dari hasil proses understanding, kemudian dilanjutkan dengan proses lower, normalisasi, stopword, tokenisasi

- LOWER

```
data['text_cleaning'] = data['text_cleaning'].str.lower()
```

```
data.head()
```

Tahap pertama adalah lower, yaitu tahap dengan mengubah seluruh huruf menjadi kecil pada sebuah teks. Dengan mengubah seluruh teks menjadi huruf kecil, model dapat lebih konsisten dalam mengenali pola dan menghindari kesalahan interpretasi akibat perbedaan kapitalisasi.

- NORMALISASI

```
norm = {"klo" : "kalau", "enggak" : "tidak", "ngga" : "tidak", "ga" : "tidak",
"mager" : "malas", "thn" : "tahun"}, def normalisasi (str_text):
for i in norm:
str_text = str_text.replace(i, norm[i])
return str_text
```

```
data['text_cleaning'] = data['text_cleaning'].apply(lambda x: normalisasi(x))
```

Selanjutnya, dilakukan normalisasi, yaitu mengganti kata-kata tidak baku, singkatan, atau slang dengan kata baku yang lebih umum digunakan. Jika kata-kata ini tidak dinormalisasi, model dapat kesulitan memahami konteksnya dan menganggapnya sebagai kata yang berbeda

#### • STOPWORD

```
import Sastrawi

from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory, StopWordRemover, ArrayDictionary

more_stop_word = []

stop_words = StopWordRemoverFactory().get_stop_words()

new_array = ArrayDictionary(stop_words)

stop_words_remover = StopWordRemover(new_array)

def atopword(str_text):

str_text = stop_words_remover_new.remover(str_text)

return str_text
```

```
data['text_cleaning'] = data['text_cleaning'].apply(lambda x: atopword(x))
```

Stopword, yaitu tahap menghapus suatu kata yang tidak memberikan informasi dalam analisis. Kata-kata seperti "dan", "ke", "di", "yang", dan "dengan" sering kali tidak memberikan kontribusi berarti dalam pemahaman konteks suatu teks, sehingga dihapus untuk meningkatkan efisiensi analisis. Dengan mengurangi jumlah kata yang tidak relevan, model dapat lebih memusatkan perhatian pada kata-kata yang memiliki bobot lebih besar dalam menentukan makna suatu kalimat.

#### • TOKENISASI

```
tokenized = data['text_cleaning'].apply(lambda x:x.split())
```

tokenized

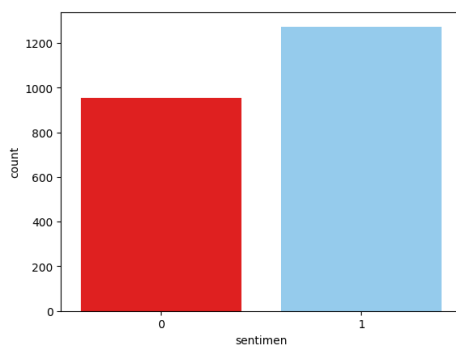
Tokenisasi, yaitu membagi teks menjadi bagian yang kecil atau disebut token. NLP, tokenisasi dapat dilakukan dalam berbagai tingkat, seperti per kata atau per kalimat. Dengan contoh, kalimat "aku sangat senang" akan diubah menjadi ["aku", "sangat", "senang"].66

#### 4. Visualisasi kata

Visualisasi word cloud digunakan untuk menganalisis data teks. Word cloud menampilkan semua kata dalam teks, di mana kata dengan frekuensi tertinggi akan ditampilkan dengan ukuran font lebih besar dalam visualisasi.

Berdasarkan hasil word cloud, kata "skripsi" sering muncul dalam komentar. Sementara itu, dalam pembahasan terkait "anxiety skripsi," kata-kata yang sering digunakan meliputi "skripsi," "tapi," "anxiety," "mental," dan lainnya. Hal ini menunjukkan bahwa masyarakat Indonesia banyak membahas skripsi dengan kata-kata kunci tersebut. Maka, dilakukan analisis lebih lanjut untuk





**Gambar 5. Presentase World Cloud**

Data yang digunakan berjumlah 2839, dimana hasil dengan sentimen positif berjumlah dengan kode (1) sebanyak 1273 dan sentimen negatif dengan kode (2) berjumlah sebanyak 952.

### 5. Data Preparation

Tahap dalam proses data yang dilakukan sebelum data digunakan untuk analisis lebih lanjut atau pelatihan model machine learning. Tahap ini memiliki tujuan untuk mengetahui data berada dalam keadaan optimal, bebas dari kesalahan, serta siap untuk digunakan dalam model prediktif atau analitik. Dalam tahap ini, berbagai teknik diterapkan, seperti pembersihan data, penanganan nilai yang hilang, transformasi variabel, normalisasi, serta pembagian data set menjadi data latih dan data uji. Dengan melakukan Data preparation secara menyeluruh, kualitas dataset dapat ditingkatkan sehingga model machine learning mampu menghasilkan prediksi yang lebih akurat dengan memiliki tahapan Yaitu membagi data menjadi data pelatihan dan data pengujian, menggunakan CountVectorizer untuk mengubah teks menjadi fitur numerik, melakukan oversampling menggunakan SMOTE pada data latihan, contoh data hasil resampling

- Memisahkan Data Latih menjadi Data Uji

```
X=data_clean['text_cleaning']
y = data_clean['sentimen']
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.15,
random_state=42)
```

Pengkelompokan dataset mengubah menjadi data latih dan data uji memiliki tujuan untuk memastikan kemampuan model dalam melakukan generalisasi secara optimal. Dengan pendekatan ini, model dapat dievaluasi menggunakan data yang tidak digunakan selama proses pelatihan, sehingga performanya terhadap data baru dapat diukur secara lebih objektif. Selain itu, pemisahan ini juga berperan dalam mencegah overfitting, yaitu kondisi ketika model terlalu terpaku pada data latih dan kesulitan dalam mengolah data yang belum ditemukan.

- CountVektorizer Teks Menjadi Numerik

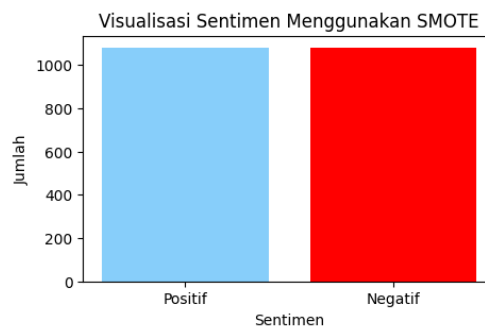
```
vectorizer = CountVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)
```

Model machine learning tidak dapat langsung memahami teks sebagai input, karena model bekerja dengan data dalam bentuk numerik. Oleh sebab itu, perlu dilakukan proses transformasi teks menjadi vektor numerik menggunakan metode seperti CountVectorizer. Teknik ini mengubah kata-kata dalam teks menjadi representasi berbasis angka yang dapat digunakan dalam perhitungan matematis oleh model. Dengan pendekatan ini, informasi dalam teks tetap dipertahankan dalam format yang dapat diolah oleh algoritma machine learning

- Oversampling Menggunakan SMOTE

```
smote = SMOTE(random_state=42)
x_train_resampled, y_train_resampled = smote.fit_resample(x_train_vec,
y_train)
```

Tahap ini merupakan langkah dalam tahap pra pemrosesan data untuk keperluan klasifikasi, yang bertujuan untuk mencegah model menjadi bias terhadap kelas tertentu akibat ketidakseimbangan data.



**Gambar 6. Visualisasi Sentimen Menggunakan SMOTE**

Jumlah data dengan sentimen positif jauh lebih sedikit dibandingkan sentimen negatif, model cenderung memiliki bias terhadap kelas yang lebih dominan atau lebih banyak muncul. Untuk mengatasi masalah ini, digunakan teknik SMOTE (Synthetic Minority Over-sampling Technique) yang secara sintesis menambah jumlah sampel pada kelas minoritas. Dengan metode ini, distribusi kelas dalam dataset menjadi lebih seimbang, sehingga model dapat belajar dengan lebih baik tanpa bias terhadap kelas mayoritas.

- Hasil Data Resampling

```
sentimen_counts = y_train_resampled.value_counts()

plt.figure(figsize=(5, 3))

plt.bar(sentimen_counts.index, sentimen_counts.values, color=['red',
'lightskyblue'])

plt.xlabel('Sentimen')

plt.ylabel('Jumlah')

plt.title('Visualisasi Sentimen Menggunakan SMOTE')

plt.xticks(sentimen_counts.index, ['Negatif', 'Positif'])
```

```
plt.show()
```

Setelah melakukan proses oversampling dengan SMOTE, selanjutnya yaitu evaluasi distribusi data antar kelas merata. Cara yang dapat digunakan adalah dengan melakukan visualisasi data hasil resampling. Dengan menampilkan jumlah sampel dari masing-masing kelas dalam bentuk grafik batang atau diagram lainnya, untuk memastikan bahwa data yang digunakan untuk pelatihan telah seimbang, sehingga model memiliki peluang yang sama dalam mempelajari pola dari setiap kelas.

## 6. TF IDF

TF-IDF (Term Frequency-Inverse Document Frequency) merupakan metode yang digunakan untuk menentukan tingkat kepentingan suatu kata dalam sebuah dokumen berdasarkan frekuensi kemunculannya dan seberapa jarang kata tersebut muncul di seluruh kumpulan dokumen. Metode ini menghitung nilai 0 (TF) dan Inverse Document Frequency (IDF) berdasarkan kumpulan dokumen yang ada dalam suatu korpus.

Kata	TF	IDF	TF IDF
Skripsi	2501	1.2779	3195.9559
Anxiety	589	2.6094	1536.9589
Depresi	644	2.5681	1653.8259
Mahasiswa	129	4.1980	541.5380

**Tabel 3. TF-IDF**

Kata “Skripsi” memiliki frekuensi paling tinggi yaitu 2501 kali, namun IDF-nya hanya 1.2779 karena kemungkinan kata ini sering muncul di berbagai dokumen, sehingga nilainya tidak terlalu langka. Namun, hasil TF-IDF-nya tetap tinggi (3195.9559) karena TF-nya sangat besar.

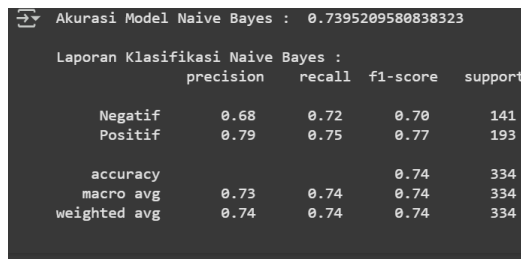
Kata “Mahasiswa” memiliki nilai TF paling kecil (129), namun IDF-nya tinggi (4.1980) karena kemungkinan kata ini tidak muncul di banyak dokumen lain, sehingga masih memberikan bobot yang signifikan pada TF-IDF (541.5380).

Sedangkan “Anxiety” dan “Depresi” memiliki nilai TF menengah dan IDF yang tinggi, menghasilkan TF-IDF yang cukup besar (masing-masing 1536.9589 dan 1653.8259). Ini menunjukkan bahwa meskipun tidak terlalu sering muncul, kata-kata tersebut cukup penting dan membedakan dokumen dibandingkan kata-kata lain.

## 7. Modeling

Modeling dalam Machine Learning adalah tahap inti dalam pengembangan model kecerdasan buatan yang bertujuan untuk membuat prediksi atau klasifikasi berdasarkan pola dalam data. Proses ini melibatkan pemilihan algoritma, pelatihan model dengan dataset, pengujian kinerja model, serta optimasi agar model dapat memberikan hasil yang akurat.

```
print("Akurasi Model Naive Bayes : ", accuracy)
print("\nLaporan Klasifikasi Naive Bayes : \n", classification_rep)
```



	precision	recall	f1-score	support
Negatif	0.68	0.72	0.70	141
Positif	0.79	0.75	0.77	193
accuracy			0.74	334
macro avg	0.73	0.74	0.74	334
weighted avg	0.74	0.74	0.74	334

**Gambar 7. Akurasi Model Naive Bayes**

Model yang dibangun menggunakan algoritma Naive Bayes. Model ini dibangun dengan memanfaatkan data latih, lalu dilakukan pengujian menggunakan data uji. Berdasarkan hasil evaluasi, model mampu meraih akurasi sebesar 73,95%, dengan nilai precision pada kelas positif (79%) lebih tinggi dibandingkan pada kelas negatif (68%). Selain itu, nilai recall untuk sentimen positif (75%) juga menunjukkan kemampuan model dalam mendeteksi komentar yang bernuansa positif. Nilai F1-score yang dicapai yaitu 77% untuk kelas positif dan 70% untuk negatif, mencerminkan performa yang seimbang, meskipun terdapat kecenderungan model lebih unggul pada sentimen positif.

## 8. Testing

```
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB

# Data training sederhana
train_texts = ["saya sangat senang", "saya bahagia", "saya sangat sedih", "saya kecewa"]
train_labels = [1, 1, 0, 0] # 1 = Positif, 0 = Negatif

# Inisialisasi vectorizer dan model Naive Bayes
vectorizer = CountVectorizer()
X_train = vectorizer.fit_transform(train_texts)

# Melatih model
naive_bayes = MultinomialNB()
naive_bayes.fit(X_train, train_labels)

# Input testing
new_text = input("\nMasukkan Teks Baru: ")
new_text_vec = vectorizer.transform([new_text])
prediction_sentimen = naive_bayes.predict(new_text_vec)

if prediction_sentimen[0] == 1:
    sentimen_label = "positif"
elif prediction_sentimen[0] == 0:
```

```
sentimen_label = "negatif"
```

Proses testing dilakukan dengan menggunakan data teks baru untuk menguji apakah model dapat mengklasifikasikan sentimen teks tersebut sebagai positif atau negatif. Tahapan ini memberikan simulasi penerapan model dengan memberikan sebuah kalimat, lalu sistem secara otomatis memberikan label sentimen berdasarkan pelatihan.

Sebelum pengujian, beberapa kalimat pelatihan yang merupakan kalimat positif dan negatif digunakan, seperti “saya sangat senang” (positif) dan “saya kecewa” (negatif). Teks-teks tersebut diubah menjadi bentuk numerik menggunakan teknik CountVectorizer, setelah pelatihan selesai, model diuji dengan menerima masukan berupa kalimat baru. Kalimat tersebut diproses menggunakan CountVectorizer yang sama, lalu dianalisis menggunakan model Naive Bayes. Hasil analisis berupa prediksi apakah sentimen dari kalimat itu adalah “positif” atau “negatif”.

## KESIMPULAN

Hasil penelitian menunjukkan bahwa analisis sentimen mahasiswa terhadap tugas akhir yang digunakan untuk memahami tingkat kecemasan. Dengan menggunakan algoritma Naive Bayes, model berhasil mencapai akurasi 73,95% dan menunjukkan efektivitas yang lebih tinggi dalam mengenali sentimen positif. Hasil penelitian mengungkap bahwa kecemasan mahasiswa dipengaruhi oleh tekanan internal dan eksternal, sementara dukungan sosial dan strategi manajemen waktu berperan penting dalam membentuk sentimen positif. Penelitian ini memberikan kontribusi dalam pengembangan sistem deteksi psikologis berbasis teks dan mendorong institusi pendidikan untuk menyediakan pendekatan pendampingan yang lebih suportif

Secara keseluruhan, penelitian ini memberikan pengetahuan dalam memahami kecemasan mahasiswa dalam lingkungan akademik, khususnya dalam menyelesaikan tugas akhir. Dengan menerapkan metode analisis yang sesuai serta mengidentifikasi faktor-faktor penyebab kecemasan, penelitian ini tidak hanya mengungkap permasalahan yang dihadapi mahasiswa, tetapi juga menyajikan solusi yang dapat diterapkan untuk meningkatkan kesejahteraan mahasiswa.

## DAFTAR PUSTAKA

- Agustina Dewi, P., Aulia, R., Khairi, il, & Ula, M. (n.d.). SENASTIKA Universitas Malikussaleh DETEKSI RISIKO DEPRESI DAN KECEMASAN MAHASISWA MENGGUNAKAN ALGORITMA NAIVE BAYES.
- Akhnaf, A. F., Putri, R. P., Vaca, A., Hidayat, N. P., Az-Zahra, R. I., & Rusdi, A. (2022). SELF AWARENESS DAN KECEMASAN PADA MAHASISWA TINGKAT AKHIR. *Jurnal Muara Ilmu Sosial, Humaniora, Dan Seni*, 6(1), 107. <https://doi.org/10.24912/jmishumsen.v6i1.13201.2022>
- Anjarsari, T., Ratna, I., Astutik, I., Informatika, ), Sains, F., & Teknologi, D. (n.d.). DETEKSI DINI GANGGUAN KECEMASAN MENGGUNAKAN METODE NAÏVE BAYES.
- Etika, N., & Hasibuan, W. F. (2016). Deskripsi Masalah Mahasiswa Yang Sedang Menyelesaikan Skripsi. In Available online at [www.journal.unrika.ac.id](http://www.journal.unrika.ac.id) *Jurnal KOPASTA* *Jurnal KOPASTA* (Vol. 3, Issue 1). [www.journal.unrika.ac.id](http://www.journal.unrika.ac.id)
- Fardiana Risa, D., Pradana, F., & Abdurrachman Bachtiar, F. (n.d.). IMPLEMENTASI METODE NAÏVE BAYES UNTUK MENDETEKSI STRES SISWA BERDASARKAN TWEET PADA SISTEM MONITORING STRES. <https://doi.org/10.25126/jtiik.202184372>
- Kurnia Indriyanti Purnama Sari, S. M. R. W. (2023). KECEMASAN AKADEMIK MAHASISWA KEBIDANAN; LITERATURE REVIEW. *JURNAL PENGEMBANGAN ILMU DAN PRAKTIK KESEHATAN*, 2, 1–10.

- Machmud, A., Wibisono, B., & Suryani, N. (2025). Analisis Sentimen Cyberbullying Pada Komentar X Menggunakan Metode Naïve Bayesid 4 (\*) Corresponding Author. In *Computer Science (CO-SCIENCE)* (Vol. 5, Issue 1). <http://jurnal.bsi.ac.id/index.php/co-science>
- Malfasari, E., Devita, Y., Erlin, F., Studi Ilmu Keperawatan Stikes Payung Negeri Pekanbaru Jl Tamtama No, P., & Baru Pekanbaru Riau, L. (2018). FAKTOR-FAKTOR YANG MEMPENGARUHI KECEMASAN MAHASISWA DALAM MENYELESAIKAN TUGAS AKHIR DI STIKES PAYUNG NEGERI PEKANBARU. In *Jurnal Ners Indonesia* (Vol. 8, Issue 2).
- Marjan, F., Sano, A., & Ifdil, I. (2018). Tingkat kecemasan mahasiswa bimbingan dan konseling dalam menyusun skripsi. *JPGI (Jurnal Penelitian Guru Indonesia)*, 3(2), 84. <https://doi.org/10.29210/02247jpgi0005>
- Nur Faisyah, S., Iqbal, M., Alifia, N., Rosa, P., Wafi Rizulloh, M., Yunita Haryanti, D., Studi, P. S., Keperawatan, I., Studi DIII Keperawatan, P., Ilmu Kesehatan, F., Muhammadiyah Jember, U., & Korespondensi, P. (2024). Tingkat Kecemasan Mahasiswa Keperawatan dalam Menyelesaikan Tugas Akhir di Universitas Muhammadiyah Jember. *The Indonesian Journal of Health Science*, 15(2). <https://doi.org/10.32528/tijhs.v15i2.1604>
- Prabowo, P. S., Piter, J., & Sihombing, T. (2021). Gambaran Gangguan Kecemasan pada Mahasiswa Fakultas Kedokteran Universitas “X” Angkatan 2007.
- Rahayu, K., Fitria, V., Septhya, D., Rahmaddeni, R., & Efrizoni, L. (2023). Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan pada Pengguna Twitter Berbasis Machine Learning. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 108–114. <https://doi.org/10.57152/malcom.v3i2.780>
- Ridha Dwiki Putri, D., Reza Fahlevi, M., Sadikin, M., Utami, R., Rizki Fajar Utomo, M., & Studi Rekayasa Perangkat Lunak, P. (n.d.). Prediksi Tingkat Depresi Remaja Menggunakan Metode Naïve Bayes Classifier: Analisis Faktor Psikologis Dan Lingkungan (Vol. 5, Issue 4).
- Septiningsih, D., Ratnasari, F., & Tangerang, S. Y. (2021). PENGARUH TEKNIK EMOTIONAL FREEDOM (EFT) TERHADAP PENURUNAN TINGKAT KECEMASAN PADA PASIEN Effect Of Emotional Freedom Technique (Eft) On The Reduction Of Anxiety Levels In Patients. *Nusantara Hasana Journal*, 1(5), Page.
- Shafa Azizah, R., Kamayani, M., & Kunci, K. (2023). Analisis Sentimen Terhadap Kesehatan Mental Selama Pandemi Covid-19 Berdasarkan Algoritma Naïve Bayes dan Deep Learning. *Jurnal ICT : Information Communication & Technology*, 23(1), 38–43. <https://ejournal.ikmi.ac.id/index.php/jict-ikmi>
- Sri Rahayu, A. S. M. (2023). EFETIVITAS AKUPRESUR DAN TERAPI MUSIK TERHADAP KECEMASAN MAHASISWA TINGKAT AKHIR. *JURNAL BONEO CENDEKIA*, 7, 1–10.
- Wakhyudin, H., Dwi, A., & Putri, S. (2020). ANALISIS KECEMASAN MAHASISWA DALAM MENYELESAIKAN SKRIPSI.
- Zakiah. (2016). GAMBARAN KARAKTERISTIK, TINGKAT STRES, ANSIETAS, DAN DEPRESI PADA MAHASISISWA KEPERAWATAN YANG SEDANG MENGERJAKAN SKRIPSI. 2, 1–6.
- Zulfahmi, A., & Andriany, D. (2021). Kematangan vokasional dengan kecemasan dalam menghadapi dunia kerja pada mahasiswa tingkat akhir. *Cognicia*, 9(2), 64–75. <https://doi.org/10.22219/cognicia.v9i2.15728>