

# ANALISIS SENTIMEN PROGRAM MERDEKA BELAJAR KAMPUS MERDEKA MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER

*SENTIMENT ANALYSIS OF MERDEKA BELAJAR KAMPUS MERDEKA USING  
NAÏVE BAYES CLASSIFIER ALGORITHM*

Priza Pandunata<sup>1\*</sup>, Saifur Rifqi Ali<sup>2</sup>, Yanuar Nurdiansyah<sup>3</sup>

E-mail: <sup>1</sup>priza@unej.ac.id, <sup>2</sup>saifurrifqi9@gmail.com, <sup>3</sup>yanuar\_pssi@unej.ac.id

<sup>1,2,3</sup> Program Studi Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Jember

## Abstrak

Program Merdeka Belajar – Kampus Merdeka (MBKM) adalah program yang dicanangkan oleh pihak Menteri Pendidikan dan Kebudayaan (Mendikbud) yang bertujuan untuk meningkatkan kualitas mahasiswa serta diharapkan dapat membantu mahasiswa dalam menguasai berbagai bidang keilmuan sehingga siap untuk memasuki dunia kerja. Program MBKM sendiri secara resmi digelar untuk pertama kali pada tahun 2020 silam, tentu seiring berjalannya waktu masyarakat telah merasakan dampak dari program MBKM ini sehingga tak jarang banyak menuai pro dan kontra. Topik penelitian analisis sentimen terhadap program MBKM ini bertujuan untuk menggali informasi sentimen yang terkandung dalam setiap komentar yang diperoleh dari penghimpunan data komentar masyarakat pada media sosial *twitter*, dari hasil penghimpunan data tersebut tentunya tujuan utama yang ingin dicapai adalah untuk melihat respons masyarakat serta berharap dapat menjadi bahan evaluasi dari pihak pemangku kepentingan khususnya pihak mendikbud. Algoritma *Naïve Bayes Classifier* difungsikan sebagai sebuah metode yang digunakan dalam proses pengklasifikasian *dataset* yang berjumlah 2000 data *twit*. Pengklasifikasian data dilakukan menggunakan tiga skema pengujian dengan perbandingan data *training* dan data *testing* sebesar 70%:30%, 80%:20% dan 90%:10%. Adapun nilai akurasi terbaik yang diperoleh dalam pengujian klasifikasi pada penelitian ini menunjukkan nilai sebesar 70% di mana nilai tersebut diperoleh dari pengujian menggunakan perbandingan data sebesar 90%:10%.

**Kata kunci:** MBKM, Analisis Sentimen, *Naïve Bayes Classifier*

## Abstract

*Merdeka Belajar – Kampus Merdeka (MBKM) Program is a program launched by the Minister of Education and Culture (Mendikbud) which aims to improve the quality of students and is expected to help students master various scientific fields so that they are ready to enter the world of work. The MBKM program itself was officially held for the first time in 2020, of course, over time the community has felt the impact of this MBKM program so that it is not uncommon to reap many pros and cons. The sentiment analysis research topic of the MBKM program aims to explore sentiment information contained in every comment obtained from collecting public comment data on twitter social media, from the results of collecting data, of course, the main goal to be achieved is to see the community's response and hope to be an evaluation material from stakeholders, especially the Minister of Education and Culture. The Naïve Bayes Classifier algorithm functions as a method used in the process of classifying datasets totaling 2000 comment data. Data classification was carried out using three test schemes with a comparison of training data and testing data of 70%:30%, 80%:20% and 90%:10%. The best accuracy value obtained in the classification test in this study showed a value of 70% where the value was obtained from the test using a data comparison of 90%:10%.*

**Keywords:** MBKM, Sentiment Analysis, *Naïve Bayes Classifier*

## 1. PENDAHULUAN

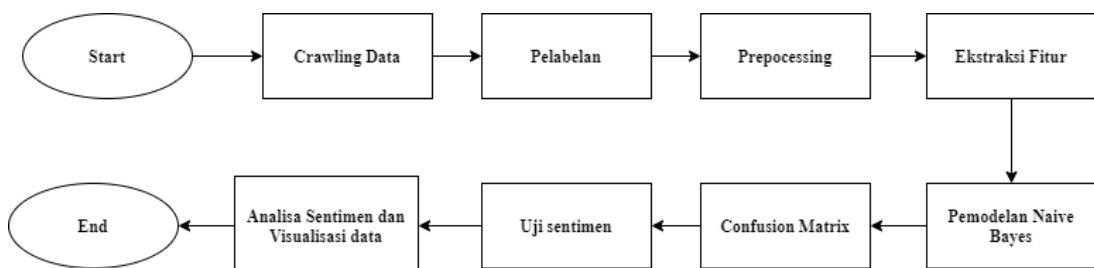
Dalam rangka mempersiapkan mahasiswa yang siap dalam menghadapi perubahan sosial budaya, dunia kerja, serta kemajuan teknologi tentu kompetensi yang dimiliki mahasiswa harus disiapkan untuk jauh lebih matang seiring dengan kebutuhan zaman[1]. Peningkatan kompetensi ini tentu tidak hanya berfokus pada dunia revolusi industri, namun harus terus diimbangi dengan sebuah perancangan masa depan yang dapat berubah sewaktu-waktu. Universitas diminta dapat menciptakan serangkaian kegiatan pembelajaran yang inovatif agar supaya mahasiswa dapat memperoleh capaian pembelajaran yang mencakup banyak aspek seperti aspek sikap, pengetahuan serta keterampilan secara optimal dan relevan[2].

Namun meskipun begitu tak jarang dalam prakteknya penyelenggaraan program MBKM ini masih banyak menuai kendala. Melalui penelitian yang dilakukan oleh Nensi Nofa Nofia[3] terungkap bahwa terdapat beberapa kendala dalam pelaksanaan kebijakan MBKM, di antaranya adalah penggunaan biaya yang cukup besar, perguruan tinggi kecil dengan keterbatasan sumber daya manusia, sarana prasarana dan lain sebagainya. Melihat hal tersebut, tentu tak sedikit dari peserta program MBKM yang turut menceritakan opininya selama berada di bawah naungan kerja MBKM. Berbagai macam bentuk opini yang disampaikan tentu dapat menarik keberagaman pandangan masyarakat terkait esensi program MBKM itu sendiri, umumnya opini yang disampaikan banyak tertuang melalui media sosial salah satunya ialah *twitter*[4].

Dari hasil olah data yang dihimpun dari komentar *twitter* tentu perlu untuk dilakukan sebuah analisis sentimen terhadap opini yang disampaikan. Analisis sentimen sendiri merupakan sebuah bentuk pemrosesan bahasa dan ekstraksi informasi yang bertujuan untuk mendapatkan perasaan penulis[5] yang diungkapkan melalui komentar dengan sentimen positif, netral atau negatif[6]. Pada penelitian terdahulu yang berjudul “Sentimen Analisis Pengguna *Twitter* Terhadap Kebijakan Merdeka Belajar Menggunakan *Algoritma Naïve Bayes*[7]” menjelaskan bahwa hasil dari uji performansi terhadap data dengan sentimen positif dan sentimen negatif, algoritma *Naïve Bayes Classifier* memperoleh nilai akurasi sebesar 80,55%. Dari hasil tersebut tentu penulis merasa bahwa algoritma *Naïve Bayes Classifier* cocok untuk digunakan sebagai metode uji sentimen, namun meski begitu pada penelitian tersebut penulis merasa bahwa perlu adanya pengujian terhadap data dengan sentimen netral guna hasil penelitian yang diperoleh dapat lebih akurat dan beragam[8].

## 2. METODOLOGI

Pada penelitian ini, terdapat tahapan penelitian yang dilakukan sebagaimana terlihat pada gambar 1.



Gambar 1. Tahapan Penelitian

### 2.1 Crawling Data

*Crawling* data adalah tahapan awal yang dilakukan dengan tujuan untuk menghimpun data dari unggahan *tweet* dengan memanfaatkan *twitter API*. Pada penelitian ini penulis memanfaatkan sebaran program MBKM sebagai kata kunci dalam pengumpulan data.

Adapun jumlah *dataset* yang digunakan pada penelitian ini dibatasi sebanyak 2000 data dan dihimpun sejak tanggal 1 Juli 2022 hingga 15 Agustus 2022.

**2.2 Pelabelan**

Pelabelan data dilakukan guna mengidentifikasi komponen data yang sudah diambil dari proses *crawling* ke dalam sebuah kelas sentimen yang telah disiapkan. Dalam tahap ini pelabelan data dilakukan secara manual oleh Sarjana Jurusan Pendidikan Bahasa dan Sastra Indonesia selaku pihak ahli bahasa.

**2.3 Pre-processing**

*Pre-processing* adalah tahap di mana *dataset* akan diolah ke dalam bentuk yang lebih mudah untuk dipahami sistem[9]. Adapun rincian tahapan *preprocessing* pada penelitian kali ini antara lain *convert emoticon, remove punctuation, case folding, tokenizing, normalization, stopword removal dan stemming*.

**2.4 Ekstraksi Fitur**

Pada penelitian ini peneliti menggunakan dua model ekstraksi fitur antara lain adalah *Bag of Words Model* dan *Term Frequency-Inverse Dokumen Frequecy (TF-IDF)*. *Bag of Words* merupakan suatu cara sederhana dalam menghitung frekuensi kemunculan kata dalam dokumen[10]. Adapun TF-IDF merupakan sebuah proses mengubah *token (term)* menjadi nilai numerik yang dapat dibaca oleh komputer, tujuannya agar sistem dapat mengevaluasi seberapa pentingnya sebuah *term* pada suatu dokumen[11]. Adapun rumus perhitungan algoritma TF-IDF sebagaimana dituliskan pada persamaan (1) :

$$W_{t,d} = Tf_{t,d} \times idf_t \dots\dots\dots (1)$$

dimana  $Tf_{t,d}$  adalah bobot *Term Frequency* dan  $Idf_t$  adalah bobot *Inverse Document Frequency*.

Nilai IDF diperoleh dari persamaan (2) :

$$Idf_t = \log\left(\frac{N}{df(t)}\right) \dots\dots\dots (2)$$

dimana N merupakan total seluruh dokumen dan  $df(t)$  merupakan jumlah dokumen yang mengandung suatu term tertentu.

**2.4 Pemodelan Naïve Bayes Classifier**

Algoritma *Naïve Bayes Classifier* merupakan sebuah model yang difungsikan untuk keperluan klasifikasi data[12]. Umumnya tahapan yang biasa dilakukan dalam perhitungan *Naïve Bayes Classifier* antara lain seperti mencari nilai *prior probability*, lalu mencari nilai *conditional probability* dan diakhiri dengan mencari nilai *posterior probability*. Adapun persamaan dari perhitungan *prior probability* dirumuskan menggunakan persamaan (3) sebagai berikut:

$$P(v_j) = \frac{D_j}{D} \dots\dots\dots (3)$$

di mana D adalah jumlah dokumen dan  $D_j$  adalah jumlah dokumen yang mengandung suatu kelas tertentu.

Rumus perhitungan *conditional probability* sebagaimana persamaan (4) berikut :

$$P(X_i|V_j) = \frac{W_{i,j} + 1}{N + N_j} \dots\dots\dots (4)$$

di mana  $W_{i,j}$  adalah bobot term, N adalah jumlah *term* pada suatu dokumen dan  $N_j$  adalah total *term* pada dokumen.

Rumus perhitungan *posterior probability* sebagaimana persamaan (5) berikut :

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} p(x_1, x_2, x_3 \dots x_n)p(V_j) \dots \dots \dots (5)$$

di mana  $V_{MAP}$  adalah seluruh kategori yang akan diuji,  $V_j$  adalah kategori sentimen,  $P(X_i/V_j)$  adalah peluang  $X_i$  pada kategori  $V_j$  dan  $P(V_j)$  adalah peluang pada  $V_j$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil Penghimpunan *Dataset* dan Pelabelan

Pada penelitian ini *dataset* dihimpun dari hasil unggahan *tweet* dengan memanfaatkan *twitter API*. *Twitter API* sendiri merupakan suatu sarana yang dipersiapkan oleh pihak *twitter* guna memudahkan pengguna dalam membangun suatu perangkat lunak yang berintegrasi dengan *twitter*. Setelah *dataset* berhasil dihimpun langkah selanjutnya ialah pelabelan yang dilakukan oleh Sarjana Jurusan Pendidikan Bahasa dan Sastra Indonesia selaku pihak ahli bahasa. Dari hasil pelabelan *dataset* jumlah data yang teridentifikasi ke dalam sentimen positif berjumlah sebanyak 747 data, data yang teridentifikasi ke dalam sentimen netral sebanyak 699 data dan data yang teridentifikasi ke dalam sentimen negatif sebanyak 554 data dengan total keseluruhan *dataset* sebanyak 2000 data. Dari keseluruhan data tersebut, nantinya *dataset* akan dibagi ke dalam dua jenis data antara lain data *training* dan data *testing*. Adapun rasio dari perbandingan data *training* dan data *testing* ialah sebesar 70%:30%, 80%:20%, 90%:10%. Adapun hasil dari proses penghimpunan *dataset* beserta pelabelan oleh ahli bahasa sebagaimana terlihat pada tabel 1 berikut ini.

**Tabel 1. Hasil Pelabelan *Dataset***

Tweet	Label
Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka/PKL pendidikan dah 😊'	Positif
@alightlily @collegemenfess Mjb program studi di luar negeri selama 1 smt dri kampus merdeka'	Netral
MBKM,PANITIA,MAGANG Belum ada kejelasan semua :( "	Negatif

#### 3.2 Hasil *Pre-processing*

##### a. *Convert Emoji dan Convert Emoticon*

*Convert emoji dan convert emoticon* merupakan tahap mengubah bentuk *emoji* atau *emoticon* ke dalam bentuk *string* sesuai makna yang terkandung dalam bentuk *emoji* atau *emoticon* itu sendiri. Adapun hasil dari proses *convert emoji dan convert emoticon* dapat dilihat pada tabel 2.

**Tabel 2. Hasil *Convert Emoji dan Convert Emoticon***

Tweet	Hasil <i>Convert Emoji dan Convert Emoticon</i>
Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka/PKL pendidikan dah 😊'	Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka/PKL pendidikan dah <b>gembira</b> '

##### b. *Remove Punctuation*

*Remove Punctuation* adalah tahapan yang bertujuan untuk mebersihkan *noise* berupa *link URL, hastag, username, tanda baca dan angka*. Adapun hasil dari proses *remove punctuation* sebagaimana terlihat pada tabel 3.

**Tabel 3. Hasil *Remove Punctuation***

Tweet	Hasil <i>Remove Punctuation</i>
Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka/PKL pendidikan dah gembira	Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka PKL pendidikan dah gembira

c. *Case Folding*

*Case Folding* atau *Transform Cases* merupakan sebuah tahapan yang bertujuan untuk mengganti format huruf yang ada dalam *dataset* ke dalam format huruf kecil (*lower case*). Adapun hasil dari proses *case folding* dapat dilihat pada tabel 4.

**Tabel 4. Hasil Case Folding**

Tweet	Hasil Case Folding
Anjir cuy w harus nyiapin skill banget muat magang Kampus Merdeka PKL pendidikan dah gembira	anjir cuy w harus nyiapin skill banget muat magang kampus merdeka pkl pendidikan dah gembira

d. *Tokenizing*

*Tokenizing* adalah suatu proses di mana setiap kalimat yang ada pada *dataset* akan dipecah ke dalam bentuk token (*term*). Adapun hasil dari proses *tokenizing* dapat dilihat pada tabel 5.

**Tabel 5. Hasil Tokenizing**

Tweet	Hasil Tokenizing
anjir cuy w harus nyiapin skill banget muat magang kampus merdeka pkl pendidikan dah gembira	['anjir', 'cuy', 'w', 'harus', 'nyiapin', 'skill', 'banget', 'muat', 'magang', 'kampus', 'merdeka', 'pkl', 'pendidikan', 'dah', 'gembira']

e. *Normalization*

*Normalization* merupakan tahapan dimana sistem akan melakukan perbaikan kata yang bersifat tidak baku (*slang*) ke dalam bentuk kata baku sesuai kaidah dari Kamus Besar Bahasa Indonesia (KBBI). Pada tahap ini kamus perbaikan kata yang digunakan adalah kamus *colloquial-indonesian-lexicon* dan kamus tambahan yang disusun bersama dengan bimbingan dari pihak ahli bahasa. Adapun hasil dari proses *normalization* dapat dilihat pada tabel 6.

**Tabel 6. Hasil Normalization**

Tweet	Hasil Normalization
['anjir', 'cuy', 'w', 'harus', 'nyiapin', 'skill', 'banget', 'muat', 'magang', 'kampus', 'merdeka', 'pkl', 'pendidikan', 'dah', 'gembira']	['anjir', 'cuy', 'gue', 'harus', 'menyiapkan', 'keahlian', 'banget', 'muat', 'magang', 'kampus', 'merdeka', 'pkl', 'pendidikan', 'deh', 'gembira']

f. *Stopword Removal*

*Stopword Removal* adalah tahap di mana sistem akan melakukan pemilahan terhadap kata yang bersifat umum lalu menghapusnya, namun meskipun begitu tahapan ini tidak sampai mengubah makna sentimen yang ada. Selain itu, pada tahap ini *detokenize* juga dilakukan dengan tujuan untuk menghapus bentuk *token* atau simbol yang diperoleh dari proses *tokenizing*. Adapun hasil dari proses *stopword removal* dapat dilihat pada tabel 7.

**Tabel 7. Hasil Stopword Removal**

Tweet	Hasil Stopword Removal
['anjir', 'cuy', 'gue', 'harus', 'menyiapkan', 'keahlian', 'banget', 'muat', 'magang', 'kampus', 'merdeka', 'pkl', 'pendidikan', 'deh', 'gembira']	anjir keahlian muat magang kampus merdeka pkl pendidikan gembira

g. *Stemming*

*Stemming* adalah proses pemetaan dan penguraian bentuk dari suatu kata imbuhan menjadi bentuk kata dasarnya. Adapun hasil dari proses *stemming* sebagaimana terlihat pada tabel 8.

**Tabel 8. Hasil Stemming**

Tweet	Hasil Stemming
anjir keahlian muat magang kampus	anjir ahli muat magang kampus merdeka
merdeka pkl pendidikan gembira	pkl didik gembira

### 3.3 Hasil Ekstraksi Fitur

Pada kesempatan ini penulis bermaksud untuk melakukan perhitungan ekstraksi fitur menggunakan algoritma TF-IDF dengan memanfaatkan beberapa sampel data training sebagai berikut yang dapat dilihat pada tabel 9.

**Tabel 9. Sampel Data Training**

No	Data Training	Label
1	ahli magang kampus merdeka pkl didik gembira	Positif
2	tukar program kampus merdeka info mana sertifikat	Positif
3	tidak perlu kampus merdeka	Negatif
4	dengar kabar korupsi proyek kampus merdeka kaget	Negatif
5	bantu iisma salah program kampus merdeka negeri ajar	Netral

Adapun proses perhitungan ekstraksi fitur menggunakan algoritma TF-IDF sebagaimana terlihat pada tabel 10 berikut.

**Tabel 10. Hasil Perhitungan TF-IDF**

NO	Term (t)	TF	TF	TF	TF	TF	DF	IDF
		D1	D2	D3	D4	D5		
1	ahli	1	0	0	0	0	1	0,69897
2	magang	1	0	0	0	0	1	0,69897
3	kampus	1	1	1	1	1	5	0
4	merdeka	1	1	1	1	1	5	0
5	pkl	1	0	0	0	0	1	0,69897
6	didik	1	0	0	0	0	1	0,69897
7	gembira	1	0	0	0	0	1	0,69897
8	tukar	0	1	0	0	0	1	0,69897
9	program	0	1	0	0	1	2	0,39794
10	info	0	1	0	0	0	1	0,69897
11	mana	0	1	0	0	0	1	0,69897
12	sertifikat	0	1	0	0	0	1	0,69897
13	tidak	0	0	1	0	0	1	0,69897
14	perlu	0	0	1	0	0	1	0,69897
15	dengar	0	0	0	1	0	1	0,69897
16	kabar	0	0	0	1	0	1	0,69897
17	korupsi	0	0	0	1	0	1	0,69897
18	proyek	0	0	0	1	0	1	0,69897
19	kaget	0	0	0	1	0	1	0,69897
20	bantu	0	0	0	0	1	1	0,69897
21	iisma	0	0	0	0	1	1	0,69897
22	salah	0	0	0	0	1	1	0,69897
23	negeri	0	0	0	0	1	1	0,69897
24	ajar	0	0	0	0	1	1	0,69897

Dari hasil perhitungan ekstraksi fitur di atas nilai TF dan IDF dipergunakan untuk keperluan mencari nilai TF-IDF itu sendiri, di mana cara yang dilakukan ialah mengalikan nilai keduanya. Setelah bobot TF-IDF ditemukan maka bobot pada masing-masing kelas akan dicari guna keperluan pada tahap selanjutnya yaitu penerapan model menggunakan algoritma *Naïve Bayes Classifier*.

### 3.4 Hasil Penerapan Model *Naïve Bayes Classifier*

Untuk melakukan proses penerapan model menggunakan algoritma *Naïve Bayes Classifier* penulis memanfaatkan sampel data *testing* sebagai berikut :

**Tabel 11. Sampel Data Testing**

No	Tweet	Label
1	uang saku telat cair	???

Setelah data testing disiapkan maka selanjutnya mencari nilai *prior probability* sebagaimana terlihat pada tabel berikut:

**Tabel 12. Hasil Perhitungan Prior Probability**

No	Kelas	Prior Probability $D(j) / D$
1	Positif	$2/5 = 0,4$
2	Netral	$1/5 = 0,2$
3	Negatif	$2/5 = 0,4$

Langkah selanjutnya ialah mencari nilai *conditional probability*, adapun hasil dari perhitungannya sebagaimana terlihat pada tabel berikut:

**Tabel 13. Hasil Perhitungan Conditional Probability**

$P(X_i V_j)$	Conditional Probability		
	Positif	Netral	Negatif
$P(uang)$	0,045945367	0,052717314	0,050077370
$P(saku)$	0,045945367	0,052717314	0,050077370
$P(telat)$	0,045945367	0,052717314	0,050077370
$P(cair)$	0,045945367	0,052717314	0,050077370

Langkah terakhir ialah mencari nilai *posterior probability*, adapun hasil dari perhitungannya sebagaimana terlihat pada tabel berikut:

**Tabel 14. Hasil Perhitungan Posterior Probability**

Kelas	Posterior Probability
Positif	$V_{map} = 0,4 \times 0,045945367 \times 0,045945367 \times 0,045945367 \times 0,045945367$ = 0,0000017825
Netral	$V_{map} = 0,2 \times 0,052717314 \times 0,052717314 \times 0,052717314 \times 0,052717314$ = 0,0000015447
Negatif	$V_{map} = 0,4 \times 0,050077370 \times 0,050077370 \times 0,050077370 \times 0,050077370$ = 0,0000025155

Melihat hasil penerapan algoritma *Naïve Bayes Classifier* di atas maka dapat ditarik kesimpulan bahwa melalui uji coba menggunakan sampel data *testing* “uang saku telat cair” data yang dimaksud termasuk ke dalam kelas sentimen negatif dikarenakan memperoleh nilai paling tinggi yaitu sebesar 0,0000025155.

### 3.5 Hasil Uji Performansi

Proses uji performansi menggunakan algoritma *confusion matrix* dilakukan dengan tujuan untuk mengetahui keberhasilan model dalam melakukan proses klasifikasi. Pada tahapan *confusion matrix* nilai yang akan diperoleh berupa nilai *recall*, *precision*, *f1-score* dan *accuracy*. Adapun hasil dari proses uji performansi algoritma *Naïve Bayes Classifier* sebagaimana terlihat pada tabel berikut:

Tabel 15. Hasil Uji Performansi

Keterangan	Skenario 1	Skenario 2	Skenario 3
Data Training	70%	80%	90%
Data Testing	30%	20%	10%
Recall	0,54	0,59	0,67
Presisi	0,61	0,66	0,72
F1-Score	0,53	0,60	0,68
Accuracy	56%	62%	70%

Dari paparan tabel di atas, pada tahapan uji performansi algoritma *Naïve Bayes Classifier* skenario 3 dengan perbandingan data *training* dan data *testing* sebesar 90%:10% memperoleh nilai akurasi paling tinggi yaitu sebesar 70%.

**3.6 Hasil Uji Sentimen**

Uji sentimen dilakukan untuk melihat hasil dari pengujian sistem terkait prediksi sentimen yang terkandung dalam sebuah unggahan *twit*. Adapun data yang akan diujikan dalam tahapan uji sentimen merupakan data yang diperoleh dari hasil cuitan pengguna di media sosial *twitter* namun di luar *dataset* yang sudah ada. Adapun hasil dari uji sentimen yang dimaksud sebagaimana terlihat pada tabel 16 berikut.

Tabel 16. Hasil Uji Sentimen

Data Uji Sentimen	Klasifikasi Sentimen		
	Skenario 1	Skenario 2	Skenario 3
kerja tugas magang mudah asal jalan ikhlas	Positif	Positif	Positif
kaget dana program mbkm ngaret	Negatif	Negatif	Negatif
takut ikut mbkm belum siap	Netral	Netral	Netral

Dari paparan tabel di atas, diketahui hasil dari tahap pengujian sentimen pada masing-masing skenario memperoleh prediksi sentimen yang sama persis. Meskipun perbandingan data *training* dan data *testingnya* berbeda.

**3.6 Hasil Analisa Sentimen dan Visualisasi Data**

Pada tahap pencarian hasil analisa sentimen dan visualisasi data penulis memanfaatkan *library wordcloud* sebagai cara dalam mencari kata yang sering bermunculan pada sebuah dokumen dan disajikan ke dalam bentuk visual[13]. Adapun bentuk visualisasi data pada setiap kelas sentimen sebagaimana gambar berikut :



Gambar 2. Hasil Visualisasi Data Sentimen Positif





Gambar 3. Hasil Visualisasi Data Sentimen Netral



Gambar 4. Hasil Visualisasi Data Sentimen Negatif

Melihat hasil visualisasi data pada masing-masing kelas sentimen di atas, terlihat sentimen positif menunjukkan kata yang cukup dominan yaitu kata “kuliah” dan “giat”. Sedangkan pada sentimen netral kata yang mendominasi hasil dari visualisasi data tersebut ialah kata “semester”. Lalu pada hasil visualisasi data pada kelas sentimen negatif terlihat bahwa kata yang muncul paling besar ialah kata “cair” lalu kemudian muncul kata “uang saku”, hal ini seolah menggambarkan bahwa pada kelas sentimen negatif, banyak peserta program MBKM yang mengeluh ataupun tidak senang terkait mekanis pencairan uang saku MBKM.

#### 4. KESIMPULAN

Berdasarkan hasil analisa sentimen dan visualisasi data yang dilakukan pada setiap kelas sentimen diperoleh kesimpulan bahwa pengguna *twitter* banyak memberikan opini positif terkait program Merdeka Belajar Kampus Merdeka (MBKM) dikarenakan senang terhadap suasana perkuliahan di era Kampus Merdeka serta merasa semakin giat dalam mengikuti prosesi pembelajaran. Sedangkan sentimen netral banyak tergambar karena ekspresi pengguna merasa lumrah terhadap syarat semester yang perlu dipenuhi oleh mahasiswa ketika mendaftar program MBKM. Sedangkan visualisasi data pada sentimen negatif cukup mengejutkan di mana pengguna *twitter* ramai memperbincangkan terkait kendala yang terjadi selama pagelaran program MBKM yaitu perihal pencairan tunjangan gaji atau uang saku. Melihat hasil pengujian model yang telah dilakukan sebanyak tiga kali dengan tingkat rasio perbandingan data *training* dan data *testing* sebesar 70%:30%, 80%:20% dan 90%:10% nilai akurasi terbaik yang diperoleh ialah sebesar 70% di mana nilai tersebut didapat dari rasio perbandingan data sebesar 90%:10%. Sedangkan pada pengujian dengan rasio perbandingan data sebesar 80%:20% memperoleh nilai akurasi sebesar 62% dan pada pengujian dengan rasio perbandingan data sebesar 70%:30% nilai akurasi yang didapat hanya sebesar 56%.

## 5. DAFTAR RUJUKAN

- [1] Dirjen Dikti Kemendikbud, “Buku Panduan Pelayanan Merdeka Belajar dan Kampus Merdeka,” *Merdeka Belajar-Kampus Merdeka*, pp. 1–33, 2020, [Online]. Available: <http://dikti.kemdikbud.go.id/wp-content/uploads/2020/04/Buku-Panduan-Merdeka-Belajar-Kampus-Merdeka-2020>
- [2] Y. Yuherman, W. Nugroho, and D. Sunarsi, “Dampak Kebijakan MBKM Pada Kesiapan Sumber Daya Manusia dan Fasilitas Fakultas Hukum Usahid Jakarta,” *Moral. J. Ilmu Huk.*, vol. 7, no. 2, p. 222, 2021, doi: 10.52947/morality.v7i2.235.
- [3] N. N. Nofia, “Analisis Tantangan Implementasi Kebijakan ‘Merdeka Belajar Kampus Merdeka’ Pada Perguruan Tinggi Islam Negeri Di Indonesia,” *PRODU Prokurasi Edukasi J. Manaj. Pendidik. Islam*, vol. 1, no. 2, pp. 61–72, 2020, doi: 10.15548/p-prokurasi.v1i2.3328.
- [4] F. A. Girnanfa and A. Susilo, “Studi Dramaturgi Pengelolaan Kesan Melalui Twitter Sebagai Sarana Eksistensi Diri Mahasiswa di Jakarta,” *J. New Media Commun.*, vol. 1, no. 1, pp. 58–73, 2022, doi: 10.55985/jnmc.v1i1.2.
- [5] S. Mukherjee, “Sentiment Analysis,” *ML.NET Reveal.*, pp. 113–127, 2021, doi: 10.1007/978-1-4842-6543-7\_7.
- [6] J. Gautam, M. Atrey, N. Malsa, A. Balyan, R. N. Shaw, and A. Ghosh, “Twitter Data Sentiment Analysis Using Naive Bayes Classifier and Generation of Heat Map for Analyzing Intensity Geographically,” *Adv. Intell. Syst. Comput.*, vol. 1319, no. July, pp. 129–139, 2021, doi: 10.1007/978-981-33-6919-1\_10.
- [7] H. D. Prasetyo, T. Pramiyati, and I. N. Isnainiyah, “Sentimen Analisis Pengguna Twitter Terhadap Kebijakan Merdeka Belajar Menggunakan Algoritma Naive Bayes,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, no. April, pp. 559–568, 2021.
- [8] Y. Nurdiansyah, F. Rahman, and P. Pandunata, “Analisis Sentimen Opini Publik Terhadap Undang-Undang Cipta Kerja pada Twitter Menggunakan Metode Naive Bayes Classifier,” *Pros. Semin. Nas. Sains Teknol. dan Inov. Indones.*, vol. 3, no. November, pp. 201–212, 2021, doi: 10.54706/senastindo.v3.2021.158.
- [9] S. Kumar, A. K. Kar, and P. V. Ilavarasan, “Applications of text mining in services management: A systematic literature review,” *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 1, p. 100008, 2021, doi: 10.1016/j.jjime.2021.100008.
- [10] N. Bayat, E. Rastegari, and Q. Li, “Human Gait Recognition Using Bag of Words Feature-Representation Method,” *Hum. Factors Wearable Technol.*, vol. 29, 2022, doi: 10.54941/ahfe1001481.
- [11] H. Zakiyuddin and Apriani, “Penerapan Algoritma Cosine Similarity Dan Pembobotan TF-IDF Pada System Penerimaan Mahasiswa Baru Universitas Bumigoram Mataram,” *J. Bumigora Inf. Technol.*, vol. 3, no. 1, pp. 19–27, 2021.
- [12] D. F. Zhafira, B. Rahayudi, and I. Indriati, “Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube,” *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
- [13] S. Trihandaru, H. A. Parhusip, B. Susanto, and C. F. R. Putri, “Word Cloud of UKSW Lecturer Research Competence Based on Google Scholar Data,” *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 7, no. 2, pp. 52–59, 2021, doi: 10.23917/khif.v7i2.13123.