

Analisis Sentimen Mahasiswa Tentang Model Perkuliahan Hybrid Teaching pada Fakultas Ilmu Komputer UMI Menggunakan Machine Learning

Analysis of Student Sentiments About the Hybrid Teaching Lecture Model at the Faculty of Computer Science UMI Using Machine Learning

Nur Anisa Basri^{a,1,*}, Yulita Salim^{a,2*}, Herdianti Darwis^{a,3*}

^aTeknik Informatika, Universitas Muslim Indonesia, Makassar, Indonesia
¹nuranisa.basri@gmail.com; ²yulitasalim@umi.ac.id; ³herdianti.darwis@umi.ac.id;
*corresponding author

Informasi Artikel	ABSTRAK
Diserahkan : 12 Juli 2024 Diterima : 30 Juli 2024 Direvisi : 30 Juli 2024 Diterbitkan : 31 Juli 2024 Kata Kunci: Hybrid Teaching Sentimen Naïve Bayes KNN SVM Keywords: Hybrid Teaching Sentiment Naïve Bayes KNN SVM This is an open access article under the CC-BY-SA license. 	Model perkuliahan <i>hybrid teaching</i> yang dilaksanakan menuai kontroversi di kalangan mahasiswa. Banyak pendapat mahasiswa yang dikeluarkan terkait metode pembelajaran <i>hybrid teaching</i> di Fakultas Ilmu Komputer UMI. Penelitian ini bertujuan menganalisis sentimen mahasiswa terkait perkuliahan hybrid teaching dengan menggunakan algoritma <i>K-Nearest Neighbor</i> (KNN), <i>Naïve Bayes</i>, dan <i>Support Vector Machine</i> (SVM) menggunakan pelabelan NLTK, pengujian dengan <i>cross validation</i>, dan menggunakan <i>unigram tokenizing</i>. Teknik pelabelan NLTK yang digunakan pada penelitian ini menghasilkan tingkat keakuratan algoritma KNN dengan 67.74% dibandingkan dengan algoritma <i>Naïve Bayes</i> dan SVM yang memiliki nilai akurasi sebesar 100%. Sehingga algoritma <i>Naïve Bayes Classifier</i> dan SVM dapat digunakan dengan baik pada pengklasifikasian sentimen mahasiswa terhadap perkuliahan dengan metode pembelajaran <i>hybrid teaching</i> di Fakultas Ilmu Komputer UMI. ABSTRACT <i>The hybrid teaching lecture model implemented has generated controversy among students. Many student opinions were issued regarding the hybrid teaching learning method at the Faculty of Computer Science UMI. This research aims to analyze student sentiments related to hybrid teaching lectures using the K-Nearest Neighbor (KNN), Naïve Bayes, and Support Vector Machine (SVM) algorithms using NLTK labeling, testing with cross validation, and using unigram tokenizing. The NLTK labeling technique used in this study results in the accuracy of the KNN algorithm with 67.74% compared to the Naïve Bayes and SVM algorithms which have an accuracy value of 100%. So that the Naïve Bayes Classifier and SVM algorithms can be used well in classifying student sentiment towards lectures with hybrid teaching learning methods at the Faculty of Computer Science UMI.</i>

I. Pendahuluan

Peran Pendidikan dalam mengembangkan kualitas sumber daya manusia mampu mengubah pola pikir mahasiswa serta membawa perubahan dan kemajuan suatu bangsa[1]. Banyaknya institusi pendidikan tinggi termasuk Universitas Muslim Indonesia menerapkan ruang pembelajaran yang dilengkapi dengan teknologi, sehingga menimbulkan pertanyaan bagaimana lingkungan tersebut dapat dibentuk menjadi seperti yang diharapkan dan seefektif mungkin [2]. Maka dari itu, pengajaran *hybrid* atau *hybrid teaching* merupakan model pembelajaran efektif yang baru diterapkan di mana sebagian mahasiswa menghadiri perkuliahan secara langsung (*luring/offline*), sementara sebagian lainnya bergabung secara virtual dari rumah (*daring/online*). Pada proses pelaksanaan perkuliahan *hybrid*, *platform* siaran langsung memainkan peran perantara yang menghubungkan mahasiswa jarak jauh dengan dosen untuk hadir di tempat secara bersama-sama melalui ruang obrolan untuk berkomunikasi secara interaktif [3].

Adanya website kalam di Universitas Muslim Indonesia mampu memfasilitasi dosen maupun mahasiswa dalam pengajaran dan pembelajaran [4]. Unikasi secara interaktif [3]. Penugasan individu maupun kelompok yang dilakukan dengan model perkuliahan *hybrid teaching* pada Fakultas Ilmu Komputer UMI dijadikan

sebagai penunjang dalam mengelola tugas-tugas mandiri dan kelompok, mengambil keputusan yang tepat berdasarkan informasi dan data, serta mampu bertanggung-jawab atas tugas mandiri atau kelompok.

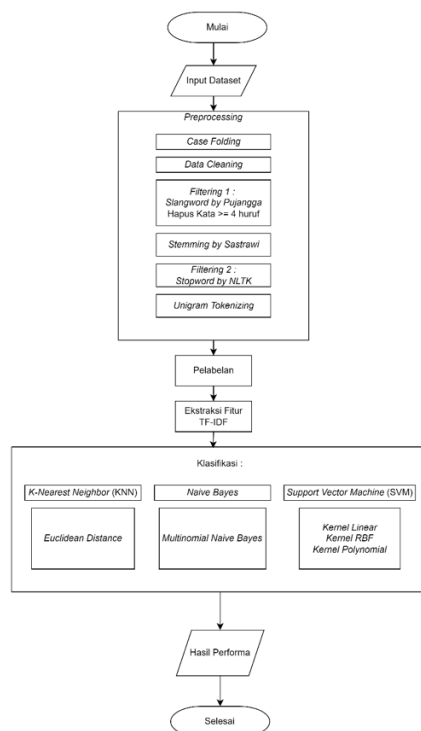
Masalah dalam analisis sentimen ini ialah keterbatasan data yang mencakup beragam sudut pandang mahasiswa. Jumlah sampel data yang terbatas atau tidak cukup representatif dapat menghambat kemampuan model untuk menghasilkan data yang akurat [5]. Selain itu, menentukan label sentimen (positif, negatif, atau netral) pada setiap teks mahasiswa melibatkan unsur subjektivitas. Interpretasi sentimen dapat bervariasi antara penilai sehingga berpotensi mempengaruhi validitas hasil analisis [6]. Untuk menyelesaikan masalah dalam analisis sentimen mahasiswa tentang model perkuliahan *hybrid teaching* Pada Fakultas Ilmu Komputer UMI menggunakan metode *Naïve Bayes*, metode lain ditambahkan seperti *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) [7]. Agar dapat mengetahui perbandingan ketiga algoritma dan memperluas sampel responden atau menggunakan teknik sampling yang lebih baik untuk mencakup berbagai latar belakang, jurusan, dan tingkat akademik mahasiswa serta memberi panduan yang jelas dan konsisten tentang apa yang dianggap sebagai sentimen positif, negatif, dan netral dapat membantu mengurangi ketidakpastian dalam *labelling* [8].

Adanya penggunaan TF-IDF mampu memberikan wawasan lebih dalam tentang kata-kata yang dikelompokkan pada sentimen mahasiswa terhadap model perkuliahan *hybrid teaching* [9]. Hal ini dapat membantu meningkatkan pemahaman terhadap persepsi dan respon mahasiswa serta dapat mengatasi masalah dalam analisis sentimen yang disebutkan dalam teks [10].

Adapun salah satu penelitian yang terkait yaitu *Support Vector Machine* untuk Analisis Sentimen Masyarakat Terhadap Penggunaan Antibiotik di Indonesia menerapkan algoritma SVM dengan *kernel linear*, RBF, dan *polynomial*, menggabungkan berbagai metode seperti pelabelan dengan *RoBERTa*, pelatihan dengan 5 *cross validation*, dan *tokenizing bigram* [11]. Dalam penelitian ini, berkaitan dalam membahas mengenai algoritma SVM dengan *kernel linear*, RBF, dan *polynomial*, menggabungkan berbagai metode seperti pelabelan dan melakukan pelatihan *cross validation*, meskipun berbeda dalam proses yang digunakan [12].

Dengan adanya analisis sentimen dalam penelitian ini, mampu mengklasifikasi pandangan, perasaan, dan persepsi mahasiswa FIKOM UMI terkait model perkuliahan *hybrid teaching* dengan menggunakan pendekatan *machine learning* yang terdiri dari KNN, *Naïve Bayes*, dan SVM yang dapat menghasilkan prediksi sentimen yang memiliki akurasi relevan pada data mahasiswa FIKOM UMI.

II. Metode



Gambar 1. Tahapan Penelitian

Penelitian ini membandingkan tiga algoritma KNN, *Naïve Bayes* dan SVM dengan menggunakan *library* NLTK [13]. Melalui perbandingan algoritma dan pelabelan tersebut diharapkan dapat ditemukan akurasi yang

paling optimal dan efektif dalam melakukan klasifikasi [14]. Metode berisi tahapan atau prosedur dalam penelitian. Adapun formula yang digunakan untuk menyelesaikan permasalahan ketiga algoritma secara rinci dan melakukan perancangan sistem jika dibutuhkan.

Dalam Gambar 1 pada desain penelitian diawali dengan menginput dataset, kemudian melakukan *preprocessing* untuk menghilangkan beberapa permasalahan yang bisa mengganggu saat pemrosesan data. Lalu, melakukan ekstraksi fitur dengan menghitung frekuensi kemunculan *term* dalam dokumen. Setelah itu, klasifikasi dilakukan dengan mengkategorikan atau mengelompokkan data baru ke dalam kelas atau kategori yang sesuai dengan analisa yang telah dilakukan dalam pemodelan. Pada tahapan akhir, terdapat hasil performa yang dapat memberikan gambaran tentang seberapa baik model dalam melakukan klasifikasi atau prediksi.

III. Hasil dan Pembahasan

1) Pengambilan Data

Pada penelitian ini, dataset diambil dengan menggunakan kuesioner dari *Google Form* yang dibagikan kepada mahasiswa FIKOM UMI dan disimpan dalam bentuk format *excel*.

2) Preprocessing

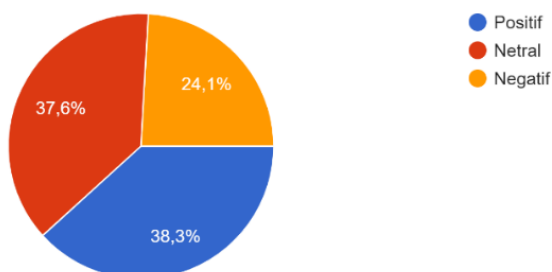
- Case Folding* dilakukan dengan mengubah huruf kapital dalam teks menjadi huruf kecil.
- Data *cleaning* yaitu pembersihan data yang dilakukan dengan menghilangkan tanda baca, karakter lain selain alfabet A-Z, dan menghapus data kosong serta duplikat.
- Filtering* dilakukan untuk menghapus kata-kata yang tidak diperlukan dan *stopward* dari pujangga sebagai referensi serta menghapus kata yang kurang dari 4 huruf.
- Stemming* yaitu memperoleh kata dasar dari kata-kata yang terdapat dalam teks. Pada proses stemming menggunakan library Sastrawi sebagai referensi.
- Unigram Tokenizing* dilakukan dengan membagi kalimat menjadi satu kata.

Tabel 1. Hasil *Preprocessing*

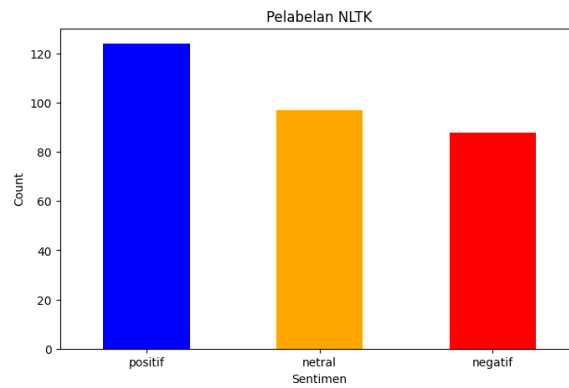
Data Mentah	<i>Case Folding</i>	Data Cleaning	<i>Filtering</i>	<i>Stemming</i>	<i>Tokenizing Unigram</i>
Pembelajaran metode ini cukup sulit Krn waktu pembelajaran kurang pasti, dan jika online secara zoom pelajaran dimengerti karena banyak kendala	2 pembelajaran metode ini cukup sulit Krn waktu pembelajaran kurang pasti, dan jika online secara zoom pelajaran dimengerti karena banyak kendala	2 pembelajaran metode ini cukup sulit Krn waktu pembelajaran kurang pasti, dan jika online secara zoom pelajaran dimengerti karena banyak kendala	metode pembelajaran metode ajar pembelajaran karena ajar jika ajar, jika, online, zoom, pelajaran online zoom ajar erti ajar, erti, kendala	metode cukup ajar, metode, cukup, sulit, karena ajar jika ajar, jika, online, zoom, pelajaran online zoom ajar erti ajar, erti, kendala	metode, cukup, sulit, karena ajar jika ajar, jika, online, zoom, pelajaran online zoom ajar erti ajar, erti, kendala

3) Pelabelan Sentimen

Pelabelan ini menggunakan dua cara yaitu pelabelan manual yang dilakukan oleh responden ketika mengisi kuesioner dan pelabelan kedua menggunakan referensi library NLTK. Dalam pelabelan manual dari 311 data menghasilkan 119 sentimen positif, 75 sentimen negative, dan 117 sentimen netral. Sedangkan, pelabelan menggunakan library NLTK yang dilakukan menghasilkan 122 sentimen positif, 89 sentimen negatif, dan 98 sentimen netral.



Gambar 2. Hasil Pelabelan Manual



Gambar 3. Hasil Pelabelan NLTK

4) Pembobotan Kata dengan TF-IDF

Pada tahapan pembobotan TF-IDF dilakukan dengan menghitung kemunculan term. Pembobotan kata diawali dengan menghitung setiap kata yang muncul pada sebuah dokumen (TF), kemudian menghitung jumlah kata yang dimuat dalam dokumen (DF). Setelah itu, menghitung nilai dari *Inverse Document Frequency* (IDF). Lalu nilai TF dikalikan dengan nilai IDF untuk menghasilkan nilai TF-IDF. Contoh perhitungan TF-IDF dapat dilihat pada Tabel 2:

Tabel 2. Pembobotan TF-IDF

Data Mentah	TF			DF	n/DF	IDF	W		
	D1	D2	D3				D1	D2	D3
Hybrid	1	1	1	3	1	0	0	0	0
Teaching	1	1	1	3	1	0	0	0	0
Efisien	1	0	0	1	3	0.47	0.47	0.00	0.00
Efektif	0	0	1	1	3	0.47	0.00	0.00	0.00
Ajar	1	0	0	1	3	0.47	0.47	0.00	0.00
Hemat	1	0	0	1	3	0.47	0.47	0.00	0.00
Bagus	0	1	0	1	3	0.47	0.00	0.47	0.00
Tenaga	1	0	0	1	3	0.47	0.47	0.00	0.00
Kampus	1	0	0	1	3	0.47	0.47	0.00	0.00
Sulit	0	1	0	1	3	0.47	0.00	0.47	0.00
Online	0	1	0	1	3	0.47	0.00	0.47	0.00
Offline	0	1	0	1	3	0.47	0.00	0.47	0.00
Siswa	0	1	0	1	3	0.47	0.00	0.47	0.00
Paham	0	1	0	1	3	0.47	0.00	0.47	0.00
Materi	0	1	0	1	3	0.47	0.00	0.47	0.00
Jangka	0	0	1	1	3	0.47	0.00	0.00	0.00
Kuliah	0	0	1	1	3	0.47	0.00	0.00	0.00

5) Evaluasi Hasil Klarifikasi

Performa dalam klasifikasi akan diuji melalui akurasi, presisi, *recall*, dan F1 score. Algoritma KNN memiliki akurasi yang rendah dibandingkan algoritma *Naïve Bayes* dan SVM.

Tabel 3. Hasil *Preprocessing*

Algoritma	Cross Validation	Akurasi	Presisi	Recall	F1 Score
KNN Pelabelan Manual	5	80.95%	54.90%	66.67%	59.52%
KNN Pelabelan NLTK	5	67.74%	50%	66.67%	55.56%

Pengujian Tingkat akurasi menggunakan metode KNN melalui pelabelan NLTK menghasilkan nilai akurasi yaitu 67.74%. Akurasi dilakukan dengan jumlah *neighbor*(n) = 198 untuk mengetahui seberapa baik keberhasilan suatu model.

Tabel 4. Hasil Pengujian Pada Klasifikasi Naïve Bayes

Algoritma	Cross Validation	Akurasi	Presisi	Recall	F1 Score
Naïve Bayes Pelabelan Manual	5	100%	100%	100%	100%
Naïve Bayes Pelabelan NLTK	5	100%	100%	100%	100%

Pengujian ini menggunakan metode klasifikasi Naïve Bayes melalui pelabelan NLTK sehingga menghasilkan nilai akurasi yang sangat baik yaitu 100%.

Tabel 5. Hasil Pengujian Pada Klasifikasi SVM

Algoritma	Kernel	Cross Validation	Akurasi	Presisi	Recall	F1 Score
SVM Pelabelan Manual	Linear	5	100%	100%	100%	100%
	RBF					
	Poly					
SVM Pelabelan NLTK	Linear	5	100%	100%	100%	100%
	RBF					
	Poly					

Dalam pengujian metode SVM melalui pelabelan NLTK dengan 3 kernel yaitu *Linear*, RBF, dan *Polynomial* menghasilkan nilai akurasi yang sangat baik yaitu 100%.

Tabel 6. Tabel *Multiclass Confusion Matrix*

	Predicted Negative	Predicted Netral	Predicted Positive
True Negative	TN	FPNe	FNP
True Netral	FNeN	TNe	FNeP
True Positive	FPN	FNNe	TP

Keterangan:

TP : True Positive
 TNe : True Netral
 TN : True Negative
 FNeN : False Netral Negative
 FPN : False Positive Negative
 FNNe : False Negative Netral
 FNP : False Negative Positive
 FNeP : False Netral Positive

a) *K-Neares Neighbor*Tabel 7. Tabel *Multiclass Confusion Matrix*

	<i>Predicted Negative</i>	<i>Predicted Netral</i>	<i>Predicted Positive</i>
<i>True Negative</i>	0	0	20
<i>True Netral</i>	0	22	0
<i>True Positive</i>	0	0	20

Setelah melalui pengujian, terdapat hasil *true negative* yaitu 0 yang menunjukkan keberhasilan dalam mengklasifikasikan opini. Adapun nilai *true positive* sebanyak 20 yang menandakan opini positif berhasil diklasifikasikan dengan tepat dan terdapat 22 sebagai *true netral*.

$$\text{Akurasi} = \frac{0+20+22}{0+20+22+20+0+0+0+0} \times 100\%$$

$$\text{Akurasi} = \frac{42}{62} \times 100\%$$

$$\text{Akurasi} = 67.74\%$$

Maka didapatkan nilai akurasi sebesar 67.74% dengan perhitungan menggunakan rumus.

$$\text{Presisi}_{\text{negatif}} = \frac{0}{0+0+0} \times 100\% = \frac{0}{0} \times 100\% = \infty$$

$$\text{Presisi}_{\text{netral}} = \frac{22}{0+0+22} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Presisi}_{\text{positif}} = \frac{20}{0+20+20} \times 100\% = \frac{20}{40} \times 100\% = 50\%$$

$$\text{Average}_{\text{presisi}} = \frac{\infty+100\%+50\%}{3} = 50\%$$

Maka didapatkan nilai presisi sebesar 50% dengan perhitungan menggunakan rumus.

$$\text{Recall}_{\text{negatif}} = \frac{0}{0+0+20} \times 100\% = \frac{0}{20} \times 100\% = \infty$$

$$\text{Recall}_{\text{netral}} = \frac{22}{0+22+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Recall}_{\text{positif}} = \frac{20}{0+0+20} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Average}_{\text{recall}} = \frac{\infty+100\%+100\%}{3} = 66.67\%$$

Maka didapatkan nilai *recall* sebesar 66.67% dengan perhitungan menggunakan rumus.

$$\text{F1 Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\%$$

$$\text{F1 Score}_{\text{negatif}} = 2 \times \frac{0 \times 0}{0+0} \times 100\% = \frac{0}{0} \times 100\% = \infty$$

$$\text{F1 Score}_{\text{netral}} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$\text{F1 Score}_{\text{positif}} = 2 \times \frac{0.5 \times 1}{0.5+1} \times 100\% = \frac{1}{1.5} \times 100\% = 66.67\%$$

$$\text{Average}_{\text{F1 Score}} = \frac{\infty+100\%+66.67\%}{3} = 55.56\%$$

Maka didapatkan nilai *F1 Score* sebesar 55.56% dengan perhitungan menggunakan rumus.

b) *Naïve Bayes*Tabel 8. Tabel *Multiclass Confusion Matrix*

	<i>Predicted Negative</i>	<i>Predicted Netral</i>	<i>Predicted Positive</i>
<i>True Negative</i>	20	0	0
<i>True Netral</i>	0	22	0
<i>True Positive</i>	0	0	20

Setelah melalui pengujian, terdapat hasil *true negative* sebanyak 20 yang menunjukkan keberhasilan dalam mengklasifikasikan 20 opini secara benar. Adapun nilai *true positive* yaitu 20 yang menandakan 20 opini negatif berhasil diklasifikasikan dengan tepat dan terdapat 22 sebagai *true netral*.

$$\text{Akurasi} = \frac{20+22+20}{20+22+20+0+0+0+0+0} \times 100\%$$

$$\text{Akurasi} = \frac{62}{62} \times 100\%$$

$$\text{Akurasi} = 100\%$$

Maka didapatkan nilai akurasi sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{Presisi}_{\text{negatif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Presisi}_{\text{netral}} = \frac{22}{22+0+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Presisi}_{\text{positif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Average}_{\text{presisi}} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai presisi sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{Recall}_{\text{negatif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Recall}_{\text{netral}} = \frac{22}{0+22+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Recall}_{\text{positif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Average}_{\text{recall}} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai recall sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{F1 Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\%$$

$$\text{F1 Score}_{\text{negatif}} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$\text{F1 Score}_{\text{netral}} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$\text{F1 Score}_{\text{positif}} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$\text{Average}_{\text{F1 Score}} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai F1 Score sebesar 100% dengan perhitungan menggunakan rumus.

c) SVM

Tabel 9. Tabel *Multiclass Confusion Matrix*

	Predicted Negative	Predicted Netral	Predicted Positive
True Negative	20	0	0
True Netral	0	22	0
True Positive	0	0	20

Setelah melalui pengujian dengan kernel *linear*, RBF dan *polynomial*, terdapat hasil *true negative* sebanyak 20 yang menunjukkan keberhasilan dalam mengklasifikasikan 20 opini secara benar. Adapun nilai *true positive* yaitu 20 yang menandakan 20 opini negatif berhasil diklasifikasikan dengan tepat dan terdapat 22 sebagai *true netral*.

$$\text{Akurasi} = \frac{20+22+20}{20+22+20+0+0+0+0+0+0} \times 100\%$$

$$\text{Akurasi} = \frac{62}{62} \times 100\%$$

$$\text{Akurasi} = 100\%$$

Maka dapatkan nilai akurasi sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{Presisi}_{\text{negatif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Presisi}_{\text{netral}} = \frac{22}{22+0+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Presisi}_{\text{positif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Average}_{\text{presisi}} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai presisi sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{Recall}_{\text{negatif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Recall}_{\text{netral}} = \frac{22}{0+22+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Recall}_{\text{positif}} = \frac{20}{20+0+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Average}_{\text{recall}} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai recall sebesar 100% dengan perhitungan menggunakan rumus.

$$\text{F1 Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\%$$

$$F1\ Score_{negatif} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$F1\ Score_{netral} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

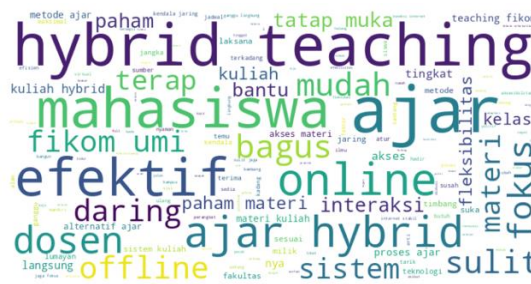
$$F1\ Score_{positif} = 2 \times \frac{1 \times 1}{1+1} \times 100\% = \frac{2}{2} \times 100\% = 100\%$$

$$Average_{F1\ Score} = \frac{100\%+100\%+100\%}{3} = 100\%$$

Maka didapatkan nilai F1 Score sebesar 100% dengan perhitungan menggunakan rumus.

6) Wordcloud

Pada Gambar 3 kata dengan frekuensi terbesar adalah ‘*hybrid*, *teaching*, *mahasiswa*, *ajar*, *bagus*, *fokus*, *efektif*, *offline*, *mahasiswa*, *online*, *sistem*’. Dalam Gambar 3 menggunakan pelabelan NLTK yang dapat memberikan sebuah informasi bahwa sistem pembelajaran *hybrid teaching* yang efektif, bagus bagi mahasiswa meskipun perkuliahan dilakukan secara online dan offline. Hal ini menjadi perhatian utama terkait opini mahasiswa dalam data teks dimiliki



Gambar 4. Wordcloud Pelabelan Manual

Pada Gambar 4 kata dengan frekuensi terbesar adalah ‘*ajar*, *bagus*, *fokus*, *efektif*, *offline*, *mahasiswa*, *online*, *sistem*, *hybrid*, *teaching*’. Dalam Gambar 4 menggunakan pelabelan NLTK yang dapat memberikan sebuah informasi bahwa sistem pembelajaran *hybrid teaching* sangat bagus dan baik bagi mahasiswa meskipun ada yang sulit memahami atau kurang fokus terkait materi perkuliahan dilakukan secara *online* dan *offline*. Hal ini menjadi perhatian utama terkait opini mahasiswa dalam data teks dimiliki



Gambar 5. Wordcloud Pelabelan NLTK

IV. Kesimpulan dan saran

Dari penelitian yang telah dilakukan, maka dapat disimpulkan bahwa sentimen mahasiswa FIKOM UMI terkait perkuliahan *hybrid teaching* menggunakan *Naive Bayes* yang ditambahkan dengan metode KNN dan SVM diidentifikasi dengan adanya penggunaan berbagai metode sehingga mampu mengetahui tingkat akurasi, presisi, *recall*, dan *f1 score* dalam pengujian dataset. Adapun performa dalam berbagai metode melalui skenario yang digunakan yaitu *stopword* dari pujangga sebagai referensi *library* dari *filtering* sehingga menunjukkan performa dengan akurasi pada algoritma KNN menggunakan *library* pelabelan NLTK sebesar 67.74%, presisi 50%, *recall* 66.67%, *f1 score* 55.56% sedangkan pada pelabelan manual, akurasi mencapai 80.95%, presisi 54.90%, *recall* 66.67%, dan *f1 score* 59.52%. Adapun tingkat akurasi, presisi, *recall*, dan *f1 score* tertinggi yang terdapat pada algoritma *Naive Bayes* dan SVM kernel *linear*, RBF, serta *polynomial* sebesar 100%. Maka dari itu, saran untuk peneliti selanjutnya di masa yang akan datang ialah dengan memperluas dataset penelitian dan ruang lingkup yang dapat menampung semua mahasiswa dari berbagai Universitas dan menggunakan beragam pelabelan seperti *Roberta*, *Vader*, dan *Textlob* untuk mengetahui tingkat akurasi dalam melakukan pengujian suatu model yang dapat dikembangkan dalam penelitian ini.

Daftar Pustaka

- [1] J. R. Gultom, D. Sundara, and M. D. Fatwara, "Pembelajaran Hybrid Learning Model Sebagai Strategi Optimalisasi Sistem Pembelajaran Di Era Pandemi Covid-19 Pada Perguruan Tinggi Di Jakarta," *Mediastima*, vol. 28, no. 1, pp. 11–22, 2022, doi: 10.55122/mediastima.v28i1.385.
- [2] L. Belluano, P. Lestari, P. Purnawansyah, and ..., "Development of academic information system using WebAssembly technology," 2021, *pdfs.semanticscholar.org*. [Online]. Available: <https://pdfs.semanticscholar.org/15f2/7a0d455e4690bf49d32a2de2f4d3bb7b7637.pdf>
- [3] E. I. Nurlaili, A. Arif, and F. Rahmawati, "Persepsi Mahasiswa Terhadap Pembelajaran Hybrid di Mata Kuliah Matematika Ekonomi," *JPEKA J. Pendidik.* ..., 2021, [Online]. Available: <https://journal.unesa.ac.id/index.php/jpeka/article/view/17031>
- [4] P. Ganovia, S. Sherly, and H. Herman, "Efektivitas hybrid learning dalam proses pembelajaran untuk siswa kelas XI SMA Kalam Kudus Pematangsiantar," *J. Pendidik. Tambusai*, 2022, [Online]. Available: <https://jptam.org/index.php/jptam/article/view/3141>
- [5] H. Ramanizar, A. Fajri, R. Binsar Sinaga, H. Mubarak, A. D. Pangestu, and D. S. Prasvita, "Analisis Sentimen Pengguna Twitter terhadap Konflik antara Palestina dan Israel Menggunakan Metode Naïve Bayesian Classification dan Support Vector Machine," *Semin. Nas. Mhs. Ilmu Komput. dan Apl. Jakarta-Indonesia*, no. September, pp. 166–175, 2021.
- [6] I. As'ad, M. A. Asis, H. M. Pakka, R. Mursalin, and Y. M. Noor, "K-Nearest Neighbors Analysis for Public Sentiment towards Implementation of Booster Vaccines in Indonesia," *Ilk. J. Ilm.*, vol. 15, no. 2, pp. 365–372, 2023, [Online]. Available: <http://103.133.36.110/index.php/ILKOM/article/view/1561>
- [7] A. R. Manga, A. N. Handayani, H. W. Herwanto, and ..., "Analysis of the Ensemble Method Classifier's Performance on Handwritten Arabic Characters Dataset," 2023, *researchgate.net*. doi: 10.33096/ilkom.v15i1.1357.186-192.
- [8] F. H. Amin, R. Jefri, and S. Luhriyani, "Metode Pembelajaran Bahasa Inggris Terapan Berbasis Proyek dengan Sistem Hybrid," *Semin. Nas. Has.* ..., 2022, [Online]. Available: <http://eprints.unm.ac.id/30268/>
- [9] R. Mukarramah, D. Atmajaya, and L. B. Ilmawan, "Performance comparison of support vector machine (SVM) with linear kernel and polynomial kernel for multiclass sentiment analysis on twitter," 2021, *academia.edu*. [Online]. Available: <https://www.academia.edu/download/84003190/pdf.pdf>
- [10] A. Yudhana, H. Herman, S. Suwanti, and ..., "Evaluating The Application of Library Information System Technology using The PIECES Method in Remote Areas," *Ilk. J.* ..., 2023, [Online]. Available: <https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/1539>
- [11] N. F. Mustamin, A. Buang, F. Aziz, and N. H. Nur, "Ensemble Techniques Based Risk Classification for Maternal Health During Pregnancy," *Ilk. J. Ilm.*, 2024, [Online]. Available: <https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/2005>
- [12] H. Darwis, N. Wanaspati, and S. Anraeni, "Support Vector Machine untuk Analisis Sentimen Masyarakat Terhadap Penggunaan Antibiotik di Indonesia," *Indones. J. Comput.* ..., 2023, [Online]. Available: <http://3.8.6.95/ijcs/index.php/ijcs/article/view/3320>
- [13] A. S. Laswi, U. Ulvah, and D. Dasril, "The Application of Weighted Ranking Method Using Combination of ROC and CPI to Select Eligible Family for Keluarga Harapan Program Aids," *Ilk. J. Ilm.*, 2023, [Online]. Available: <https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/1614>
- [14] M. P. Munthe, A. S. R. Ansori, and R. R. Septiawan, "Analisis Sentimen Komentar Pada Saluran Youtube Food Vlogger Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes," 2021.