

BAB IV

HASIL DAN PEMBAHASAN

4.1 Data Set

Data yang telah diperoleh saat proses *scraping* pada kolom komentar youtube Satu Persen dengan judul, “Guru di Indonesia Gak Dihargai! Ini Alasannya”, berjumlah 1020 komentar. Setelah itu dilakukan *cleaning* untuk menghilangkan item kosong, emoji, dan duplikat. Data yang telah dibersihkan berkurang menjadi 1014 komentar. *Labelling* dilakukan setelah data bersih, sehingga menghasilkan pembagian sentimen menjadi 3 bagian, yaitu positif, negatif, dan netral. Pada *labelling* keseluruhan data yang dipilih menjadi 1014 data, terbagi ke dalam 136 positif, 61 negatif, 817 netral.

Pada akun pembandingan dari youtube Berita Satu yang berjudul “Mengapa Guru di Indonesia Digaji Sangat Rendah?” mendapat hasil *scraping* sebanyak 948 komentar kemudian berkurang setelah proses *cleaning* menjadi 927 komentar, hasil *labelling* terbagi ke dalam 3 bagian, yaitu 134 positif, 17 negatif, 776 netral. Mengurangi jumlah data yang dianalisis membuat pengklasifikasian menjadi lebih efisien dan efektif pada metode pengklasifikasian *naïve bayes*. Dipilihnya data menggunakan salah satu metode pemilihan fitur yang sering digunakan, yaitu *document frequency*, metode ini telah dilakukan setelah proses *stemming* selesai (Bijaksana, 2020).

4.2 Preprocessing Data

Hasil preprocessing yang telah dilakukan melewati beberapa Langkah seperti *Cleaning*, *Labelling*, *Case folding*, *Tokenizing*, *Stopword removal*, *Stemming*, dan *Term frequency*.

	cleaned_text	stemmed_tokens
2	kasih pengemis tukang parkir	['kasihan', 'emis', 'tukang', 'parkir']
3	salah terpenting guru memiliki gairah pendidikan tulus membantu siswa belajar berkembang	['salah', 'penting', 'guru', 'miliki', 'gairah', 'didik', 'tulus', 'bantu', 'siswa', 'ajar', 'kembang']
4	pikir pekerjaan membutuhkan keahlian besar gaji diterima pekerja sebanding keterampilan mengatasi sulit pekerjaan bebaskan	['pikir', 'kerja', 'butuh', 'ahli', 'besar', 'gaji', 'terima', 'kerja', 'banding', 'terampil', 'atas', 'sulit', 'kerja', 'beban']
5	peningkatan penggunaan teknologi semestamerta tuntutan kualifikasi kompetensi guru dasarnya lapangan sekolah keterbelakangan teknologi guru memiliki akses penggunaan teknologi	['tingkat', 'guna', 'teknologi', 'semestamerta', 'tuntut', 'kualifikasi', 'kompetensi', 'guru', 'dasar', 'lapang', 'sekolah', 'belakang', 'teknologi', 'guru', 'miliki', 'akses', 'guna', 'teknologi']
	ngajar teman kelas siswa ngajar masyarakat menghormati berterimakasih ilmu ajarin relevan ajah masingmasing sulit ngajar anak tk angkat tangan	['ngajar', 'teman', 'kelas', 'siswa', 'ngajar', 'masyarakat', 'hormat', 'berterimakasih', 'ilmu', 'ajarin', 'relevan', 'ajah', 'masingmasing', 'sulit', 'ngajar', 'anak', 'tk', 'angkat', 'tangan']

Gambar 4.1 Hasil Preprocessing

4.3 Splitting Data

Splitting Data merupakan pembagian data dari sekumpulan data yang telah didapat kemudian dipisahkan menjadi 2 atau 3 subset, sehingga model data nantinya dapat dilatih dan diuji. Proses ini penting untuk membantu pembuatan model data sehingga mampu menghasilkan *output* akurat. Data akan dibagi menjadi 3 bagian untuk dibandingkan bagian manakah yang memiliki tingkat akurasi yang paling tinggi. Pembagian data akan dilakukan pada 2 akun, yaitu akun youtube Satu Persen dan akun pembandingnya youtube Berita Satu. Pada bagian pertama data akan dibagi menjadi 2 set, yaitu Data latih dan Data uji, pembagian dimulai dari 70% untuk Data latih dan 30% untuk Data uji.

Tabel 4.1 Pembagian Data Pertama Satu Porsen

3 Label	Data Latih	Data Uji
Positif	94	42
Negatif	45	16
Netral	570	247
Total	709	305

Data uji dan Data latih konten youtube Satu Porsen telah dipisahkan sesuai dengan perbandingan 70%:30% menghasilkan total pengujian data sebanyak 305 data.

Tabel 4.2 Pembagian Data Pertama Berita Satu

3 Label	Data Latih	Data Uji
Positif	90	44
Negatif	13	4
Netral	545	231
Total	648	279

Data uji dan Data latih konten youtube Berita Satu telah dipisahkan sesuai dengan perbandingan 70%:30% menghasilkan total pengujian data sebanyak 279 data. Lalu pembagian kedua, Data latih dan Data uji dibagi menjadi 80% dan 20%. Pembagian akan menghasilkan data sebagai berikut,

Tabel 4.3 Pembagian Data Kedua Satu Persen

3 Label	Data Latih	Data Uji
Positif	109	27
Negatif	49	12
Netral	653	164
Total	811	203

Hasil dari pembagian Data latih dan Data uji konten youtube Satu Persen menghasilkan total dari Data uji sebesar 203 data yang nantinya akan digunakan pada pengujian.

Tabel 4.4 Pembagian Data Kedua Berita Satu

3 Label	Data Latih	Data Uji
Positif	103	31
Negatif	13	4
Netral	625	151
Total	741	186

Pembagian kedua dari konten youtube Berita Satu menghasilkan Data uji sebanyak 186 data untuk pengujian kelasnya. Pembagian ketiga dilakukan dengan membagi Data latih dan Data uji menjadi sebanyak 90% dan 10%. Pembagian telah disesuaikan dengan jumlah data yang ada sebagai berikut,

Tabel 4.5 Pembagian Data Ketiga Satu Persen

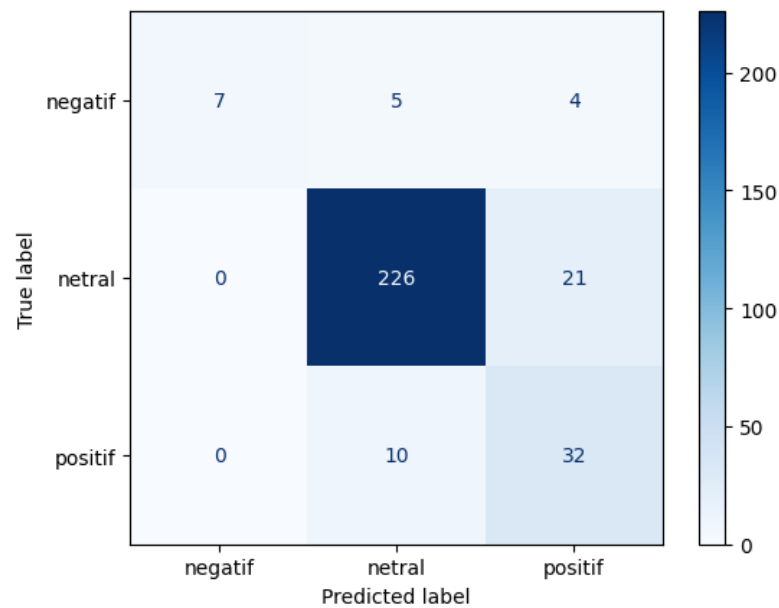
3 Label	Data Latih	Data Uji
Positif	124	10
Negatif	16	1
Netral	694	82
Total	834	93

4.4 Pengujian Klasifikasi Data

Pengujian data akan menggunakan *confusion matriks* sebagai bentuk evaluasi model sehingga akan menghasilkan nilai *accuracy*, *precision*, *recall*, dan *f1 Score*. Tiga pengujian akan dilakukan mulai dari pembagian 70% dengan 30%, 80% dengan 20%, dan 90% dengan 10%. Pengujian data dilakukan pada 2 akun, yaitu akun youtube Satu Persen dan akun pembandingnya Berita Satu.

4.4.1 Pengujian Kelas Pertama Youtube Satu Persen

Pada pengujian kelas pertama pembagian yang dilakukan untuk Data latih dan Data uji adalah 70% dan 30% dari data keseluruhan sebanyak 1014 data. Data uji yang didapat, yaitu 42 positif, 16 negatif, dan 247 netral.



Gambar 4.2 Confusion Matriks Pertama Satu Persen

Hasil yang didapat berupa pengukuran nilai *Accuracy*, *Precision*, *Recall*, dan *F1 Score*. *Confusion Matriks* telah menunjukkan hasil yang telah diperoleh dari data uji tersebut, bentuk tabel memperlihatkan bagian *true positive*, *true negative*, *true netral*, *false positive*, *false negative*, dan *false netral*. Berikut perhitungannya:

Classification Report:				
	precision	recall	f1-score	support
negatif	1.00	0.44	0.61	16
netral	0.94	0.91	0.93	247
positif	0.56	0.76	0.65	42
accuracy			0.87	305
macro avg	0.83	0.70	0.73	305
weighted avg	0.89	0.87	0.87	305

Gambar 4.3 Classification Matriks Pertama Satu Persen

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$Accuracy = \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{Total} \times 100\%$$

$$= \frac{7 + 225 + 32}{305} \times 100\% = \frac{264}{305} \times 100\% = 86.55\%$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Berikut merupakan hasil dari pengukuran seberapa tepatnya prediksi positif, nilai TP yang berarti *True Positive* dan FP berarti *False Positives* pada 3 sentimen, khusus untuk netral dinamakan *TP Netral* yang artinya *True netral* dan *FP Netral* artinya *False netral*:

Tabel 4.6 Hasil *Precision* Pertama Satu Persen

	Negatif	Netral	Positif
TP	7	226	32
FN	0+0	10+5	21+4
Precision	$\frac{7}{7+0} \times 100\%$	$\frac{226}{226+15} \times 100\%$	$\frac{32}{32+25} \times 100\%$
Hasil	100%	93.77%	56.14%

Dapat dilihat dari hasil model *precision* bagian negatif mendapat hanya 100% yang berarti nilai termasuk ke dalam kategori tinggi. Pada model bagian netral mendapatkan nilai 93.77% yang berarti masuk kategori tinggi,

kemudian bagian positif memiliki nilai 56.14% artinya termasuk ke dalam kategori sedang. Dari semua hasil rata-rata nilainya 83.29% menunjukkan banyak yang benar-benar positif berada di kategori tinggi.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut merupakan hasil dari pengukuran identifikasi dengan nilai TP yang berarti *True Positive* dan FN berarti *False Negative*, sebagai berikut:

Tabel 4.7 Hasil *Recall* Pertama Satu Persen

	Negatif	Netral	Positif
TP	7	246	32
FN	5+4	21+0	10+0
Recall	$\frac{7}{7+9} \times 100\%$	$\frac{225}{225+21} \times 100\%$	$\frac{32}{32+10} \times 100\%$
Hasil	43.75%	91.46%	76.19%

Penilaian kategori untuk *recall* pada bagian negatif mendapat nilai 43.75% artinya masuk kategori sedang, lalu bagian netral menghasilkan nilai 91.46% berarti hasil ini masuk ke kategori tinggi, sama dengan bagian negatif nilai dari bagian positif sebesar 76.19% tergolong tinggi. Identifikasi sentimen negatif memperlihatkan kurangnya prediksi namun secara rata-rata dari ketiga hasil mendapatkan nilai 70.47% yang berarti sedang atau cukup baik dalam mengukur keberhasilan deteksi positif.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2x \frac{presisi \times recall}{presisi + recall}$$

Berikut merupakan hasil dari *F1 Score* berdasarkan 3 sentimen, yaitu negatif, netral, positif:

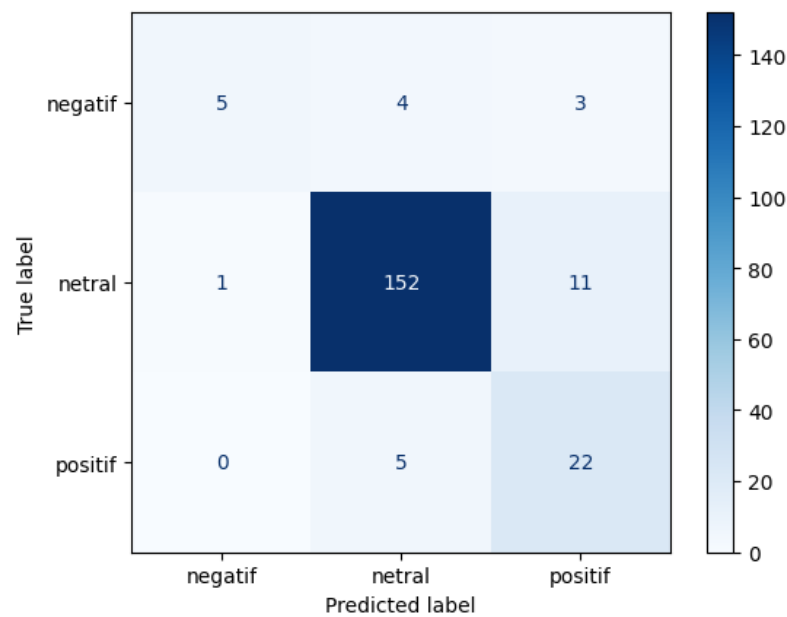
Tabel 4.8 Hasil *F1 Score* Pertama Satu Persen

	Negatif	Netral	Positif
<i>Precision</i>	1.00	0.94	0.56
<i>Recall</i>	0.44	0.91	0.76
F1 Score	$2x \frac{0.44}{1.44} x 100\%$	$2x \frac{0,8554}{1,85} x 100\%$	$2x \frac{0,4256}{1,32} x 100\%$
Hasil	61.11%	92.47%	64.48%

Dari hasil ketiga sentimen tersebut nilai bagian negatif mencapai 61.11% tergolong sedang, nilai bagian netral mencapai 92.47% berarti bagian ini tergolong tinggi lalu untuk nilai bagian positif 64.48% juga termasuk golongan tinggi. Melalui rata-ratanya dapat dilihat nilai mencapai 72.69% sehingga hasil nilai masuk ke dalam golongan tinggi atau baik, dalam menyeimbangkan kinerja model sehingga dapat menilai hasil yang unggul pada aspeknya.

4.4.2 Pengujian Kelas Kedua Youtube Satu Persen

Pada pengujian kelas kedua pembagian yang dilakukan untuk data latih dan data uji adalah sebesar 80% dan 20% dari data keseluruhan sebanyak 1019 data. Data uji yang didapat, yaitu 27 positif, 12 negatif, dan 164 netral. Hasilnya akan muncul pada *confusion matriks* dan pengukuran juga ditampilkan melalui *classification report* sebagai berikut ini.



Gambar 4.4 *Confusion Matriks* Kedua Satu Persen

Hasil *Confusion Matriks* menunjukkan bagian *true positive*, *true negative*, *false positive*, dan *false negative*. klasifikasi perhitungan menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1 Score* dari data uji yang telah dibagi menjadi 40%. Berikut merupakan perhitungan hasil kinerja model :

Classification Report:				
	precision	recall	f1-score	support
negatif	0.83	0.42	0.56	12
netral	0.94	0.93	0.94	164
positif	0.61	0.81	0.70	27
accuracy			0.88	203
macro avg	0.80	0.72	0.73	203
weighted avg	0.89	0.88	0.88	203

Gambar 4.5 Classification Matriks Kedua Satu Persen

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$\begin{aligned}
 Accuracy &= \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{\text{Total}} \times 100\% \\
 &= \frac{5 + 153 + 22}{203} \times 100\% = \frac{180}{203} \times 100\% = 88.66\%
 \end{aligned}$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Dibawah ini merupakan pengukuran tepat dari prediksi positifnya:

Tabel 4.9 Hasil *Precision* Kedua Satu Persen

	Negatif	Netral	Positif
TP	5	152	22
FN	5+0	5+4	11+3
<i>Precision</i>	$\frac{5}{5+1} \times 100\%$	$\frac{152}{152+9} \times 100\%$	$\frac{22}{22+14} \times 100\%$
Hasil	83.33%	94.40%	61.11%

Hasil dari tabel pengujian kelas kedua ini memperlihatkan bagian negatif sebesar 83.33% yang masuk kategori tinggi, bagian netral 94.40% berkategori tinggi, dan bagian positif 61.11% berkategori sedang. Rata-rata yang dihasilkan 79.61% yang berarti masuk kategori tinggi, dalam pengakurasion prediksi positif.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut adalah hasil identifikasi dari pengujian kelas kedua berdasarkan 3 sentimen yang ada:

Tabel 4.10 Hasil *Recall* Kedua Satu Persen

	Negatif	Netral	Positif
TP	5	152	22
FN	4+3	11+1	5+0
<i>Recall</i>	$\frac{5}{5+7} \times 100\%$	$\frac{152}{152+12} \times 100\%$	$\frac{22}{22+5} \times 100\%$
Hasil	41.66%	92.68%	81.48%

Pada tabel tertera hasil bagian negatif 41.66% yang masuk kategori sedang, bagian netral sebesar 92.68% masuk ke dalam kategori tinggi, dan bagian positif 81.48% masuk kategori tinggi. Berdasarkan semua hasil rata-rata yang diperoleh sebesar 71.94% dapat dikategorikan penilaian berada di tingkat tinggi untuk keberhasilannya mendeteksi sampel positif.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2 \times \frac{presisi \times recall}{presisi + recall}$$

Berikut ini merupakan hasil dari penggabungan ukuran keseimbangan dari pengujian kelas kedua youtube Satu Persen:

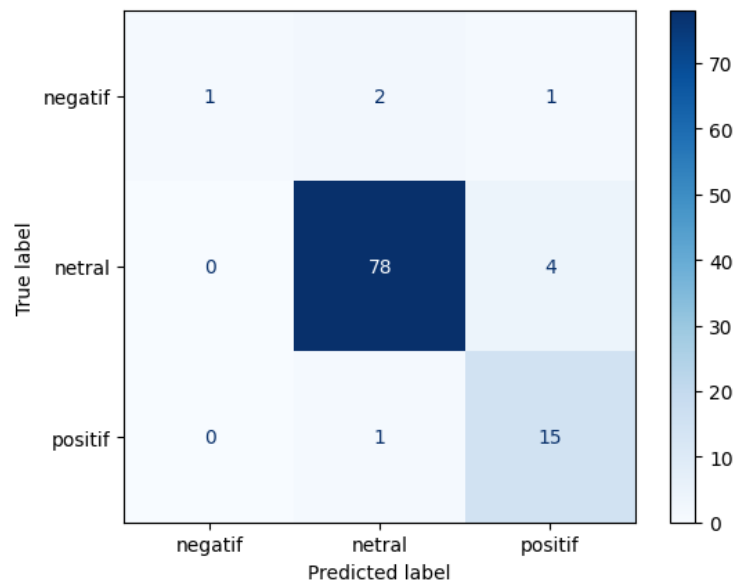
Tabel 4.11 Hasil *F1 Score* Kedua Satu Persen

	Negatif	Netral	Positif
<i>Precision</i>	0.83	0.94	0.61
<i>Recall</i>	0.42	0.93	0.94
<i>F1 Score</i>	$2x \frac{0.3486}{1.25} x 100\%$	$2x \frac{0.8742}{1,87} x 100\%$	$2x \frac{0.4941}{1.42} x 100\%$
Hasil	55.77%	93.49%	69.59%

Penggabungan dari bagian negatif sebesar 55.77% artinya masuk kategori sedang, bagian netral sebesar 93.49% masuk kategori tinggi, dan bagian positif 69.59% masuk kategori tinggi. Hasil bagian negatif masih kurang memenuhi kategori tinggi namun jika dilihat rata-rata semua nilainya menjadi sebesar 72.95% maka nilai dapat masuk ke dalam keseimbangan model kategori tinggi.

4.4.3 Pengujian Kelas Ketiga Youtube Satu Persen

Pada pengujian terakhir kelas ketiga pembagian data latih dan data uji sebanyak 90% dan 10%. data latih yang telah didapat yaitu sebanyak 14 positif, 6 negatif, 82 netral. Menghasilkan tampilan *confusion matriks* sebagai berikut



Gambar 4.6 Confusion Matriks Ketiga Satu Persen

Hasil *Confusion Matriks* menunjukkan bagian *true positive*, *true negative*, *false positive*, dan *false negative*. Lalu klasifikasi perhitungan menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1 Score* dari data uji yang telah dibagi menjadi 40%, membagi 3 sentimen, yaitu negatif, netral, dan positif. Berikut hasil perhitungan rinci dari *classification report* yang didapat dari pengujian ketiga:

	precision	recall	f1-score	support
negatif	1.00	0.25	0.40	4
netral	0.96	0.95	0.96	82
positif	0.75	0.94	0.83	16
accuracy			0.92	102
macro avg	0.90	0.71	0.73	102
weighted avg	0.93	0.92	0.92	102

Gambar 4.7 Classification Report Ketiga Satu Persen

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$Accuracy = \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{Total} \times 100\%$$

$$= \frac{1 + 78 + 15}{102} \times 100\% = \frac{95}{102} \times 100\% = 93.13\%$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan cara membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Dibawah ini merupakan pengukuran tepat dari prediksi positifnya berdasarkan 3 sentimen yang ada:

Tabel 4.12 Hasil *Precision* Ketiga Satu Persen

	Negatif	Netral	Positif
TP	1	78	15
FN	1+0	1+2	4+1
<i>Precision</i>	$\frac{1}{1+0} \times 100\%$	$\frac{78}{78+3} \times 100\%$	$\frac{15}{15+5} \times 100\%$
Hasil	100%	96.29%	75%

Pengukuran tepat prediksi positif bagian negatif menghasilkan 100% yang termasuk kategori tinggi, bagian netral menghasilkan nilai 96.29% sehingga masuk kategori tinggi, lalu nilai bagian positif dengan 75% akan masuk kategori tinggi. Dari hasil pengukuran tersebut setiap bagian masuk ke dalam kategori yang berbeda namun setelah mendapat nilai rata-rata sebesar 90.43% Tingkat *precision* berubah menjadi tinggi yang artinya pengukuran baik dalam memprediksi positif secara tepat.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut adalah hasil identifikasi dari pengujian kelas ketiga berdasarkan 3 sentimen yang ada:

Tabel 4.13 Hasil *Recall* Ketiga Satu Persen

	Negatif	Netral	Positif
TP	1	78	15
FN	2+1	4+0	1+0
<i>Recall</i>	$\frac{1}{1+3} \times 100\%$	$\frac{78}{78+4} \times 100\%$	$\frac{15}{15+1} \times 100\%$
Hasil	25%	95.12%	93.75%

Hasil yang diperoleh ada 25% untuk bagian negatif yang masuk kategori rendah, sebesar 95.12% untuk bagian netral masuk kategori tinggi, sebesar 93.75% bagian positif masuk kategori tinggi. Hasil *recall* berkategori tinggi hanya berdasarkan 2 sentimen yang ada. Rata-rata keseluruhan telah didapat nilai 71.29% yang bisa dikategorikan tinggi, hasil identifikasi model berarti berhasil menemukan besar sampel positif yang sebenarnya.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2 \times \frac{presisi \times recall}{presisi + recall}$$

Berikut ini merupakan hasil dari penggabungan ukuran keseimbangan:

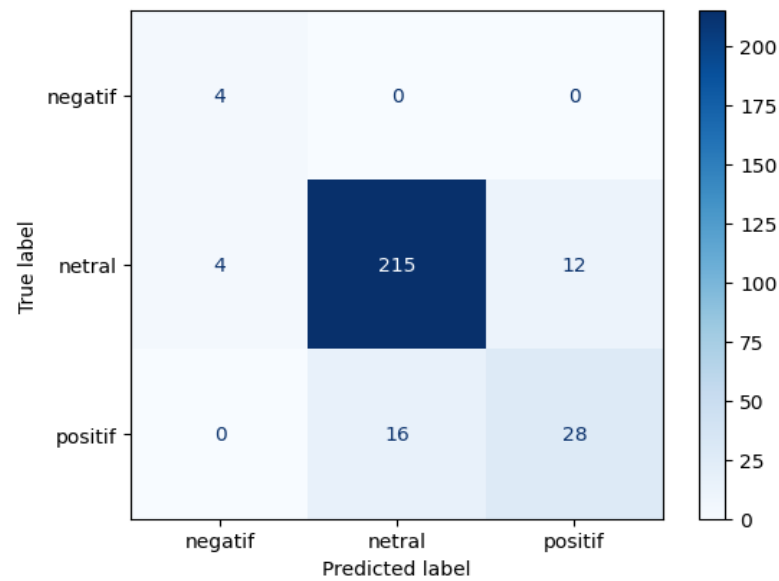
Tabel 4.14 Hasil *F1 Score* Ketiga Satu Persen

	Negatif	Netral	Positif
<i>Precision</i>	1.00	0.96	0.75
<i>Recall</i>	0.25	0.95	0.94
<i>F1 Score</i>	$2x \frac{0.25}{1.25} x 100\%$	$2x \frac{0.912}{1.91} x 100\%$	$2x \frac{0.705}{1.69} x 100\%$
Hasil	40%	95.49%	83.43%

Nilai bagian negatif menghasilkan nilai 40% termasuk kategori tinggi, nilai bagian netral menghasilkan 95.49% termasuk kategori tinggi, dan nilai bagian positif yang menghasilkan 83.43% termasuk ke dalam kategori tinggi. Berarti semua nilai memiliki tingkat keseimbangan yang baik. Pada rata-rata setelah dihitung mendapatkan 72.97%, hasil ini masuk kategori tinggi untuk keseimbangan kinerja model.

4.4.4 Pengujian Kelas Pertama Youtube Berita Satu

Pada pengujian kelas pertama pembagian yang dilakukan untuk Data latih dan Data uji adalah 70% dan 30% dari data keseluruhan sebanyak 927 data. Data uji yang didapat, yaitu 44 positif, 4 negatif, dan 231 netral.



Gambar 4.9 Confusion Matriks Pertama Berita Satu

Confusion Matriks menunjukkan hasil yang telah diperoleh dari data uji tersebut, bentuk tabel memperlihatkan bagian *true positive*, *true negative*, *true netral*, *false positive*, *false negative*, dan *false netral*. Berikut perhitungannya yang ditampilkan melalui *classification report*:

Classification Report:				
	precision	recall	f1-score	support
negatif	0.50	1.00	0.67	4
netral	0.93	0.93	0.93	231
positif	0.70	0.64	0.67	44
accuracy			0.89	279
macro avg	0.71	0.86	0.75	279
weighted avg	0.89	0.89	0.89	279

Gambar 4.8 Classification Report Pertama Berita Satu

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$Accuracy = \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{Total} \times 100\%$$

$$= \frac{4 + 215 + 28}{279} \times 100\% = \frac{247}{279} \times 100\% = 88.53\%$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Berikut merupakan hasil dari pengukuran seberapa tepatnya prediksi positif, nilai TP yang berarti *True Positive* dan FP berarti *False Positives* pada 3 sentimen, khusus untuk netral dinamakan *TP Netral* yang artinya *True netral* dan *FP Netral* artinya *False netral*:

Tabel 4.15 Precision Pertama Berita Satu

	Negatif	Netral	Positif
TP	4	215	28
FN	4+0	16+0	12+0
Precision	$\frac{4}{4 + 4} \times 100\%$	$\frac{215}{215 + 16} \times 100\%$	$\frac{28}{28 + 12} \times 100\%$
Hasil	50%	93.07%	70%

Dapat dilihat dari hasil model *precision* bagian negatif mendapat hanya 50% yang berarti nilai termasuk ke dalam kategori sedang. Pada model bagian netral mendapatkan nilai 93.07% yang berarti masuk kategori tinggi, kemudian bagian positif memiliki nilai 70% artinya termasuk ke dalam

kategori tinggi. Dari semua hasil rata-rata nilainya 71.02% menunjukkan banyak yang benar-benar positif berada di kategori tinggi.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut merupakan hasil dari pengukuran identifikasi dengan nilai TP yang berarti *True Positive* dan FN berarti *False Negative*, sebagai berikut:

Tabel 4.16 Recall Pertama Berita Satu

	Negatif	Netral	Positif
TP	4	215	28
FN	4+0	16+0	16+0
Recall	$\frac{4}{4 + 4} \times 100\%$	$\frac{215}{215 + 16} \times 100\%$	$\frac{28}{28 + 16} \times 100\%$
Hasil	100%	93.07%	63.63%

Penilaian kategori untuk *recall* pada bagian negatif mendapat nilai 100% artinya masuk kategori tinggi, lalu bagian netral menghasilkan nilai 93.07% berarti hasil ini masuk ke kategori tinggi, sama dengan bagian negatif nilai dari bagian positif sebesar 63.63% tergolong tinggi. Identifikasi sentimen negatif memperlihatkan kurangnya prediksi namun secara rata-rata dari ketiga hasil mendapatkan nilai 85.56% yang berarti tinggi atau sangat baik dalam mengukur keberhasilan deteksi positif.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2x \frac{presisi \times recall}{presisi + recall}$$

Berikut merupakan hasil dari *F1 Score* berdasarkan 3 sentimen, yaitu negatif, netral, positif:

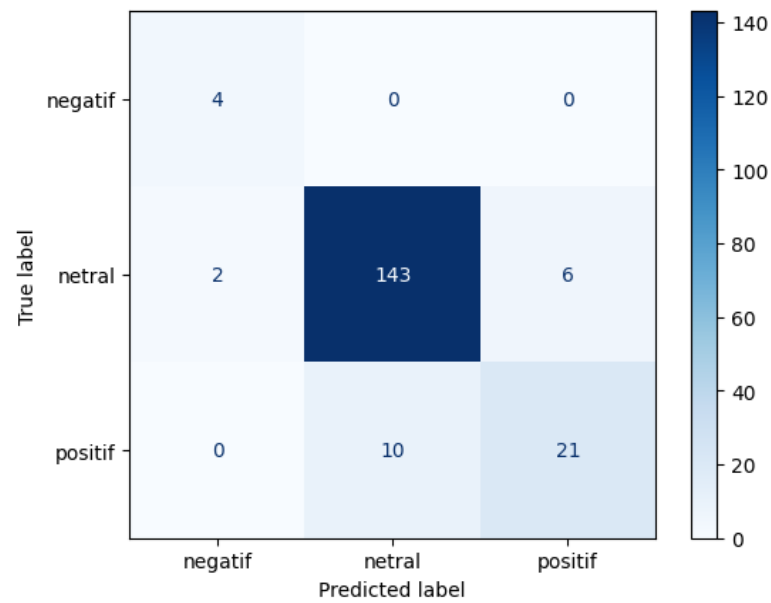
Tabel 4.17 *F1 Score* Pertama Berita Satu

	Negatif	Netral	Positif
<i>Precision</i>	0.50	0.98	0.70
<i>Recall</i>	1.00	0.93	0.64
<i>F1 Score</i>	$2x \frac{0.50}{1.50} \times 100\%$	$2x \frac{0,9114}{1,91} \times 100\%$	$2x \frac{0,448}{1,34} \times 100\%$
Hasil	66.66%	95.43%	66.86%

Nilai bagian negatif mencapai 66.66% tergolong tinggi, nilai bagian netral mencapai 95.43% berarti bagian ini tergolong tinggi lalu untuk nilai bagian positif 66.86% juga termasuk golongan tinggi. Melalui rata-ratanya dapat dilihat nilai mencapai 76.31% sehingga hasil nilai masuk ke dalam golongan tinggi atau baik, dalam menyeimbangkan kinerja model sehingga dapat menilai hasil yang unggul pada aspeknya.

4.4.5 Pengujian Kelas Kedua Youtube Berita Satu

Pengujian dengan perbandingan 80%: 20% dari youtube Berita Satu dengan jumlah data ujinya 186 data, yaitu 31 positif, 4 negatif, dan 151 netral.



Gambar 4.10 Confusion Matriks Kedua Berita Satu

Confusion Matriks menunjukkan hasil yang telah diperoleh dari data uji tersebut, bentuk tabel memperlihatkan bagian *true positive*, *true negative*, *true netral*, *false positive*, *false negative*, dan *false netral*. Berikut perhitungannya yang ditampilkan melalui *classification report*:

Classification Report:				
	precision	recall	f1-score	support
negatif	0.67	1.00	0.80	4
netral	0.93	0.95	0.94	151
positif	0.78	0.68	0.72	31
accuracy			0.90	186
macro avg	0.79	0.87	0.82	186
weighted avg	0.90	0.90	0.90	186

Gambar 4.11 Classification Matriks Kedua Berita Satu

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$Accuracy = \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{Total} \times 100\%$$

$$= \frac{4 + 143 + 21}{186} \times 100\% = \frac{247}{279} \times 100\% = 90.32\%$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Berikut merupakan hasil dari pengukurannya yang akan ditampilkan pada tabel dibawah ini:

Tabel 4.18 *Precision* Kedua Berita Satu

	Negatif	Netral	Positif
TP	4	143	21
FN	2+0	10+0	6+0
<i>Precision</i>	$\frac{4}{4 + 2} \times 100\%$	$\frac{143}{143 + 10} \times 100\%$	$\frac{21}{21 + 6} \times 100\%$
Hasil	66.66%	93.46%	77.77%

Dapat dilihat dari hasil model *precision* bagian negatif mendapat hanya 66.66% yang berarti nilai termasuk ke dalam kategori tinggi. Pada model bagian netral mendapatkan nilai 93.46% yang berarti masuk kategori tinggi, kemudian bagian positif memiliki nilai 77.77% artinya termasuk ke dalam kategori tinggi. Dari semua hasil rata-rata nilainya 79.30% menunjukkan banyak yang benar-benar positif berada di kategori tinggi.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut merupakan hasil dari pengukuran identifikasi dengan nilai TP yang berarti *True Positive* dan FN berarti *False Negative*, sebagai berikut:

Tabel 4.19 Recall Kedua Berita Satu

	Negatif	Netral	Positif
TP	4	143	21
FN	4+0	6+2	10+0
<i>Recall</i>	$\frac{4}{4 + 4} \times 100\%$	$\frac{143}{143 + 8} \times 100\%$	$\frac{21}{21 + 10} \times 100\%$
Hasil	100%	94.70%	67.74%

Recall pada bagian negatif mendapat nilai 100% artinya masuk kategori tinggi, lalu bagian netral menghasilkan nilai 94.70% berarti hasil ini masuk ke kategori tinggi, sama dengan bagian negatif nilai dari bagian positif sebesar 67.74% tergolong kategori tinggi. Rata-rata dari ketiga hasil mendapatkan nilai 87.48% yang berarti tinggi atau sangat baik dalam mengukur keberhasilan deteksi positif.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2x \frac{presisi \times recall}{presisi + recall}$$

Berikut ini merupakan hasil dari penggabungan ukuran keseimbangan:

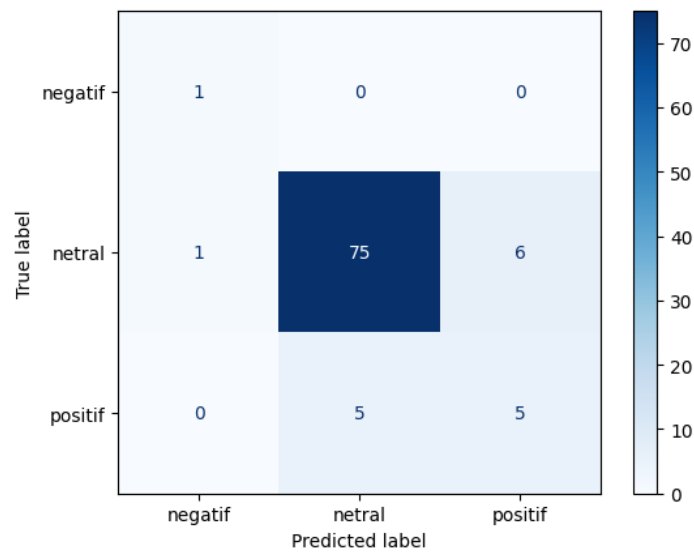
Tabel 4.20 Hasil *F1 Score* Ketiga Satu Persen

	Negatif	Netral	Positif
<i>Precision</i>	1.00	0.93	0.70
<i>Recall</i>	1.00	0.95	0.68
<i>F1 Score</i>	$2x \frac{1}{2.00} x 100\%$	$2x \frac{0.8835}{1.88} x 100\%$	$2x \frac{0.476}{1.38} x 100\%$
Hasil	100%	93.98%	68.98%

Nilai bagian negatif menghasilkan nilai 100% termasuk kategori tinggi, nilai bagian netral menghasilkan 93.98% termasuk kategori tinggi, dan nilai bagian positif yang menghasilkan 68.98% termasuk ke dalam kategori tinggi. Berarti semua nilai memiliki tingkat keseimbangan yang baik. Pada rata-rata setelah dihitung mendapatkan 87.65%, hasil ini masuk kategori tinggi untuk keseimbangan kinerja model.

4.4.6 Pengujian Kelas Ketiga Youtube Berita Satu

Pengujian kelas ketiga dengan perbandingan 90%: 10% dari youtube Berita Satu dengan keseluruhan datanya 93 data terbagi jumlahnya dalam 3 label, yaitu 10 positif, 1 negatif, dan 82 netral.



Gambar 4.12 Confusion Matriks Ketiga Berita Satu

Confusion Matriks menunjukkan hasil yang telah diperoleh dari data uji tersebut, bentuk tabel memperlihatkan bagian *true positive*, *true negative*, *true netral*, *false positive*, *false negative*, dan *false netral*. Berikut perhitungannya yang ditampilkan melalui *classification report*:

Classification Report:				
	precision	recall	f1-score	support
negatif	0.50	1.00	0.67	1
netral	0.94	0.91	0.93	82
positif	0.45	0.50	0.48	10
accuracy			0.87	93
macro avg	0.63	0.80	0.69	93
weighted avg	0.88	0.87	0.87	93

Gambar 4.13 Classification Report Ketiga Berita Satu

a. *Accuracy*

Memprediksi seberapa benar klasifikasi model dengan membagi prediksi benar dengan total data.

$$Accuracy = \frac{TP_{Negatif} + TP_{Netral} + TP_{positif}}{Total} \times 100\%$$

$$= \frac{1 + 75 + 5}{93} \times 100\% = \frac{247}{279} \times 100\% = 87.09\%$$

b. *Precision*

Pengukuran seberapa tepat prediksi positifnya, dengan membandingkan prediksi benar positif dengan semua prediksi positif.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

Berikut merupakan hasil dari pengukurannya yang akan ditampilkan pada tabel dibawah ini:

Tabel 4.21 *Precision* Kedua Berita Satu

	Negatif	Netral	Positif
TP	1	75	5
FN	1+0	5+0	6+0
<i>Precision</i>	$\frac{1}{1+0} \times 100\%$	$\frac{75}{75+5} \times 100\%$	$\frac{5}{5+6} \times 100\%$
Hasil	50%	93.75%	45.45%

Dapat dilihat dari hasil model *precision* bagian negatif mendapat hanya 50% yang berarti nilai termasuk ke dalam kategori sedang. Pada model bagian netral mendapatkan nilai 93.75% yang berarti masuk kategori tinggi, kemudian bagian positif memiliki nilai 45.45% artinya termasuk ke dalam kategori sedang. Dari semua hasil rata-rata nilainya 63.06% menunjukkan banyak yang benar-benar positif berada di kategori tinggi.

c. *Recall*

Mengidentifikasi semua data positif yang sebenarnya dengan memprediksi benar positifnya lalu bandingkan dengan semua data sebenarnya positif.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Berikut merupakan hasil dari pengukuran identifikasi dengan nilai TP yang berarti *True Positive* dan FN berarti *False Negative*, sebagai berikut:

Tabel 4.22 Recall Kedua Berita Satu

	Negatif	Netral	Positif
TP	1	75	5
FN	0+0	6+1	5+0
<i>Recall</i>	$\frac{1}{1+0} \times 100\%$	$\frac{75}{75+7} \times 100\%$	$\frac{5}{5+5} \times 100\%$
Hasil	100%	91.46%	50%

Recall pada bagian negatif mendapat nilai 100% artinya masuk kategori tinggi, lalu bagian netral menghasilkan nilai 91.46% berarti hasil ini masuk ke kategori tinggi, sama dengan bagian negatif nilai dari bagian positif sebesar 50% tergolong kategori tinggi. Rata-rata dari ketiga hasil mendapatkan nilai 80.48% yang berarti tinggi atau sangat baik dalam mengukur keberhasilan deteksi positif.

d. *F1 Score*

Penggabungan dari *precision* dan *recall* bertujuan untuk mengukur suatu keseimbangan kinerja model.

$$F1\ Score = 2x \frac{presisi \times recall}{presisi + recall}$$

Berikut ini merupakan hasil dari penggabungan ukuran keseimbangan:

Tabel 4.23 Hasil *F1 Score* Ketiga Satu Persen

	Negatif	Netral	Positif
<i>Precision</i>	0.50	0.94	0.45
<i>Recall</i>	1.00	0.91	0.50
<i>F1 Score</i>	$2x \frac{0.1}{1.50} x 100\%$	$2x \frac{0.8554}{1.85} x 100\%$	$2x \frac{0.225}{0.95} x 100\%$
Hasil	100%	92.47%	47.36%

Nilai bagian negatif menghasilkan nilai 100% termasuk kategori tinggi, nilai bagian netral menghasilkan 92.47% termasuk kategori tinggi, dan nilai bagian positif yang menghasilkan 47.36% termasuk ke dalam kategori sedang. Rata-rata dari ketiga nilai setelah dihitung mendapatkan 79.94%, hasil ini masuk kategori tinggi untuk keseimbangan kinerja model.

4.5 Hasil Akurasi Pengujian

Hasil akurasi pengujian dari ketiga pembagian data latih dan data uji telah diperoleh berdasarkan *classification report*, data menunjukkan hasil yang lebih besar cenderung pada sentimen netral, berikut ini adalah tabel hasil dari ketiga perbandingan tersebut.

Tabel 4.24 Keseluruhan Hasil Pengujian Satu Persen

Pembagian	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
70% : 30%	86.55%	71.02%	70.47%	72.69%
80% : 20%	88.66%	79.30%	71.94%	72.95%
90% : 10%	93.13%	90.43%	71.29%	72.97%

Terlihat dari hasil tabel youtube Satu Persen diatas bahwa hasil akurasi tertinggi ada pada pembagian ketiga sebesar 90%:10% dengan data yang mengalami pengujian sebanyak 14 positif, 6 negatif, 82 netral. Hasil akurasi metode *naïve bayes* tertinggi ialah 93.13%.

Tabel 4.25 Keseluruhan Hasil Pengujian Berita Satu

Pembagian	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
70% : 30%	88.53%	83.29%	85.56%	76.31%
80% : 20%	90.32%	79.61%	87.48%	87.65%
90% : 10%	87.09%	63.06%	80.48%	79.94%

Berdasarkan hasil tabel tersebut, dari 3 pengujian youtube Berita Satu sebagai akun pembandingan dari youtube Satu Persen diperoleh hasil *accuracy* tertinggi ada pada pembagian kedua sebesar 80%:20% dengan data sebanyak, 31 positif, 4 negatif, dan 151 netral. Hasil akurasi metode *naïve bayes* tertinggi ialah 90.32%. Nilai akurasi ini masuk ke dalam rentang nilai klasifikasi performa yang sangat baik berdasarkan rentang tabel performa dibawah ini.

Tabel 4.26 Batasan Nilai Akurasi

Batasan Nilai	Klasifikasi
90%-100%	Sangat baik
80%-90%	Baik
70%-80%	Cukup
60%-70%	Buruk
$\leq 60\%$	Sangat Buruk

Tabel rentang nilai diatas memperlihatkan bahwa metode klasifikasi *naïve bayes* berada direntang nilai yang sangat baik artinya metode mampu untuk mendapatkan jumlah ketepatan pada kelas pengujian ketiga(Sang, 2021).

Tabel 4.27 Kategori Penilaian

Batasan Nilai	Kategori penilaian
67%-100%	Tertinggi
33%-66%	Sedang
$\leq 33\%$	Rendah

Tabel diatas merupakan batasan nilai yang dibuat menjadi 3 kategori, yaitu tertinggi, sedang, dan rendah. Kategori ini dibuat untuk mengetahui suatu nilai masuk ke dalam penilaian yang mana untuk *precision*, *recall*, dan *f1 score*. Dari hasil ketiga perbandingan youtube Satu Persen pengujian 90%:10% yang unggul khusus pada nilai *precision* dan *f1 score* sebesar 90.43% untuk *precision* dan 93.15% untuk *f1 score*, lalu perbandingan 80% : 20% unggul hanya pada nilai *recall* sebesar 71.94% yang masuk kategori tertinggi dalam tabel penilaian.

Pada akun pembanding youtube Berita Satu diketahui pengujian 80%:20% unggul tidak hanya pada hasil *accuracy* tetapi juga pada nilai *recall* dan *f1 score* sebesar 87.48% untuk *recall* dan 87.65% untuk *f1 score*, lalu perbandingan 70% :

a. Topik 1

['guru', 'gaji', 'rendah', 'ulang', 'pikir', 'sistem', 'butuh', 'kompetensi', 'kerja', 'naik']

Pertimbangan diambil dari kumpulan komentar relevan:

Gaji guru, terutama guru honorer, seringkali berada di bawah standar kelayakan hidup, yang mengakibatkan kesulitan finansial bagi mereka. Perlu ada kebijakan untuk meningkatkan gaji guru agar setara dengan profesi lain yang memiliki fungsi sosial penting.

b. Topik 2

['guru', 'sekolah', 'didik', 'ajar', 'swasta', 'emang', 'bahas', 'banyak', 'udah', 'kaya']

Pertimbangan diambil dari kumpulan komentar relevan:

Sistem pendidikan saat ini membutuhkan perbaikan untuk memastikan bahwa guru memiliki kompetensi yang tinggi. Bisa dicapai melalui pelatihan berkelanjutan dan insentif yang mendorong peningkatan kompetensi profesional.

c. Topik 3

['guru', 'gak', 'honorer', 'pns', 'sistem', 'kayak', 'org', 'kalo', 'gitu', 'sd']

Pertimbangan diambil dari kumpulan komentar relevan:

Birokrasi yang berbelit sering menjadi penghalang bagi guru untuk mendapatkan hak-hak mereka. Perlu adanya kebijakan yang lebih terbuka dan transparan dalam menyediakan kesempatan karier yang lebih baik bagi guru, termasuk pengangkatan guru honorer menjadi PNS.