

BAB 4

HASIL PENELITIAN

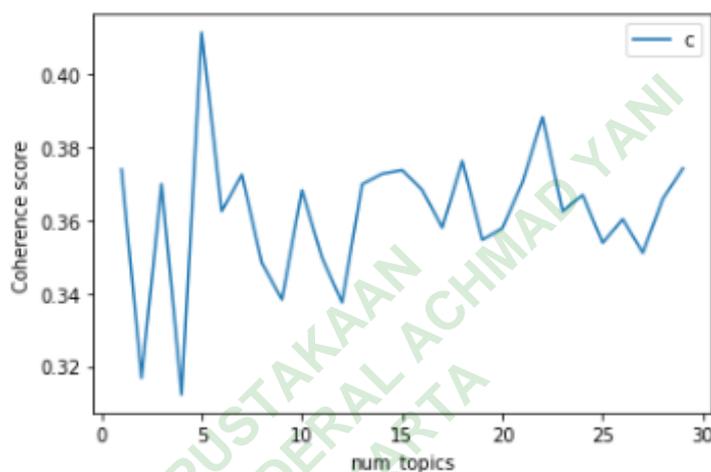
4.1 RINGKASAN HASIL PENELITIAN

Hasil penelitian dari penerapan dua metode yang dikombinasikan yaitu topic modelling dengan metode algoritma Latent Dirichlet Allocation (LDA) yang di kombinasikan dengan analisis sentiment dengan menggunakan metode Naïve Bayes Classifier (NBC). Untuk menganalisis topik serta mencari nilai sentiment dari setiap topik yang sudah didapatkan mengenai wabah virus corona dari media sosial Twitter berdasarkan data *tweet* yang di ambil dengan tiga kata kunci yaitu Vaksinasi, Omicron serta Delta dengan rentang dan jangka waktu pengambilan datanya sudah dituliskan dibagian latar belakang untuk vaksinasi data yang di ambil dari tanggal 3 januari 2021 sampai dengan tanggal 04 juli 2022 dengan jumlah data 5000 *tweet* sedangkan untuk kata kunci delta dari tanggal 11 oktober 2021 sampai dengan 04 juli 2022 dengan jumlah pengambilan datanya 5000 data *tweet* dan untuk kata kunci omicron dari tanggal 24 november 2021 hingga tanggal 04 juli 2022 dengan jumlah 5000 data *tweet* dari ketiga kata kunci tersebut keseluruhan datanya dilakukan penggabungan sehingga jumlah datanya menjadi 15.000 data *tweet* dan pada data tersebut dilakukan penghilangan *duplicate* data sehingga jumlah datanya menjadi 10.300 data *tweet* data tersebut nantinya akan digunakan untuk proses *training* dan *testing* pada metode Naïve Bayes Classifier dengan jumlah masing-masing data 1017 data untuk *training* yang sudah dilabeli secara manual dan 425 data untuk *testing* yang sudah dilabeli juga secara manual dan pada proses *training* dan *testing* NBC di dapatkan akurasi *training* 83.97% dan *testing* 89% dan didapatkan 5 topik ideal pada pembentukan *topic modelling* menggunakan metode LDA. Dan berikut ini pemaparan *topic modelling* dan *analisis sentiment* pada kasus COVID-19 berdasarkan pendapat masyarakat di media sosial Twitter menggunakan Bahasa pemrogramman Python dengan mengkombinasikan kedua metode LDA dan NBC, dengan bantuan beberapa *library* yang sudah ada pada Bahasa pemrogramman Python seperti *pandas*, *numpy*, *nlTK*, *emoji*, *sastrawi*,

sklearn, pickle, matplotlib, pyLDAvis, genism, corpora, PIL, os yang proses pengolahan datanya dilakukan di Jupyter Notebook.

4.2 PEMBEHASAN HASIL TOPIK COHERENCE

Untuk menemukan jumlah topik yang optimal, maka akan digunakan metode *topic coherence* untuk mengukur keterkaitan kata di dalam topik yang ada



Gambar 4.1 Grafik Diagram Pada Topic Coherence

Dari gambar grafik diagram diatas akan dipergunakan untuk menentukan jumlah topik dalam proses pemodelan topik dari data *tweet* yang sudah didapatkan dari Twitter. Dari jumlah topik yang sudah terlihat pada grafik nantinya akan dibandingkan berdasarkan nilai titik *coherence score* yang paling tinggi yang dapat dilihat pada Tabel 4.1

Tabel 4.1 Nilai Coherence

Num_topics	Chorence_score
1	0.3740726138978382
2	0.3168382990662958,
3	0.3700696962230816,
4	0.31227340476884086,
5	0.41162116489168776,
6	0.3626504483395103,

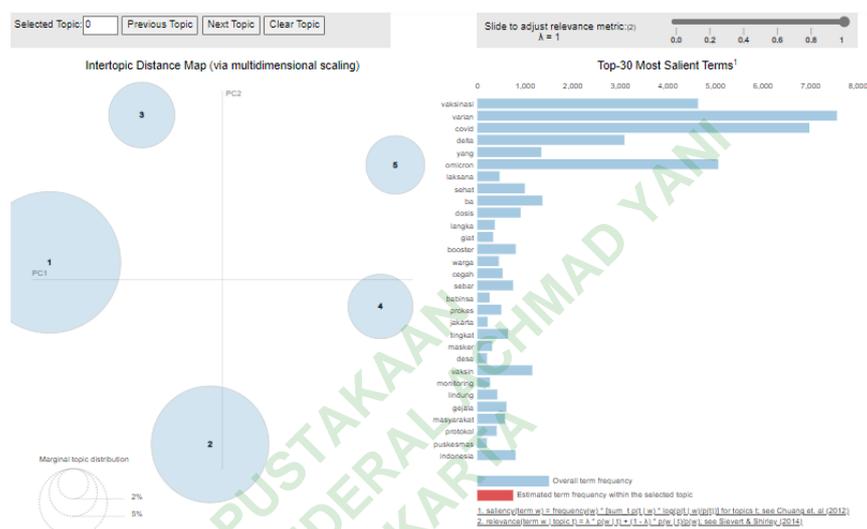
7	0.37260540935231495,
8	0.3484038103956858,
9	0.33837429577728995,
10	0.36836768037012446,
11	0.3498793107769193,
12	0.337561502485121,
13	0.36996245684430895,
14	0.37289011525593063,
15	0.37390073035631244,
16	0.368401395463423,
17	0.3580955346749033,
18	0.37637311888361225,
19	0.35472234736399766,
20	0.35798902809850613,
21	0.370494298750386,
22	0.38841836776802907,
23	0.36262813931207283,
24	0.367051832029148,
25	0.35390277482384774,
26	0.36042706573777195,
27	0.35117037758203073,
28	0.36597720867661904,
29	0.37435201036818994

Dari tabel di atas *coherence score* yang paling tinggi adalah pada nilai 0.41162116489168776 dengan jumlah topik ada 5.

4.3 HASIL TOPIK MODELLING

jumlah topik yang sudah didapatkan atau ditentukan dari metode *topic coherence* yaitu terdapat lima topik ideal lalu akan diproses lagi menggunakan metode LDA. Hasil *topic modelling* dengan jumlah lima topik yang sudah didapatkan akan dilakukan visualiasi ke dalam bentuk *Inter Distance Map* yang dapat dilihat pada Gambar 4.2.

Berdasarkan Gambar 4.2 yang sudah ditampilkan diatas pada jumlah topik lima hampir keseluruhan topik memiliki jarak sehingga jumlah topik tersebut dapat digunakan atau jumlah topik yang ideal untuk dilakukannya pembentukan *topic modelling* jadi data tersebut yang akan digunakan untuk proses selanjutnya.



Gambar 4.2 Topic Modelling Berdasarkan 5 Topik

Berdasarkan dari dua gambar yang sudah ditampilkan diatas pada jumlah topik sepuluh sedikit topik yang memiliki jarak diantara satu sama lain, dan pada jumlah topik lima hampir keseluruhan topik memiliki jarak sehingga jumlah topik yang dapat digunakan atau jumlah topik yang ideal untuk dilakukannya pembentukan *topic modelling* jadi data yang akan digunakan adalah yang berjumlah lima topik dikarenakan data tersebut memiliki jarak diantara keseluruhannya.

4.4 HASIL WORDCLOUD PERTOPIK

Topik-topik yang sudah didapatkan dan ditentukan melalui metode *topic coherence* dari setiap kata-kata yang terdapat dalam topik tersebut akan divisualisasikan kedalam bentuk *wordcloud* fungsinya untuk menampilkan frekuensi dari setiap kata yang sering muncul yang dapat dilihat pada Gambar 4.3 sampai dengan Gambar 4.7 dengan menggunakan lima jumlah topik.



Gambar 4.3 Wordcloud Topik Ke 1

Pada segmen topik 1, dapat dilihat pada kata yang sering muncul adalah “Varian”, “Omicron”, “Delta” dan “Covid” dari empat kata ini merupakan kata yang sering muncul pada segmen topik 1. Maka dari itu dapat disimpulkan inti dari seluruh kata yang sering muncul adalah mengenai varian covid-19 seperti delta dan omicron conroh *tweet* terkait dengan *wordcloud* topik 1 dapat dilihat pada Tabel 4.2

Tabel 4.2 Contoh Tweet Yang Terkait Dengan Topik 1

No	Tweet
1	varian delta omnicron bandara picu kota wuhan kompak sebar seluruh kota dunia asli tukang tipu
2	sahabat yuk saksi episode varian delta anak jaga imun nutrisi minggu
3	pas varian delta gering pas varian omnicorn yo gering saiki gering teh mosok iyo isih keno deltacorn mah dadi rakepenak karo sing rung tahu keno
4	biasae mari kenek varian delta efek ngunu ikuh gaiso mangan akeh
5	kemenkumham ri kumham sulsel kemenkumham sulsel dokter tri agus pandemi selesai virus covid mudah mutasi timbul varian alfa delta omicron kemenkumham ri



Gambar 4.6 Wordcloud Topik Ke 4

Pada segmen topik 4 dapat dilihat bahwa kata yang sering muncul pada segmen topik 2 ini adalah “Covid”, “Vaksinasi”, “Jakarta” dan “Giat” serta “Kota” kata tersebut paling banyak muncul pada segmen topik 4 maka dapat disimpulkan segmen topik 4 bahwa polres mendampingi vaksinasi covid di kabupaten kota maupun desa contoh *tweet* pada segmen topik 4 dapat dilihat pada Tabel 4.5

Tabel 4.5 Tweet Yang Terkait Topik 4

No	Tweet
1	bupati fungsi booster tingkat sistem imun vaksinasi baharu kebal covid sebar covid varian omicron masif banding delta
2	fungsi booster tingkat sistem imun vaksin suntik baharu kebal covid sebar covid varian omicron masif banding delta
3	laku vaksinasi dosis lengkap disiplin protokol sehat cegah varian covid ikn cepat maju
4	cegah tular varian covid vaksinasi disiplin patuh prokes ikn cepat maju
5	vaksinasi serentak adl wujud upaya cegah sebar covid

4.5 HASIL ANALISIS PERTOPIK

Dari hasil penelitian, mengenai metode *topic modelling* untuk melakukan analisis topik dari setiap topiknya mengenai wabah COVID-19 yang pengambilan datanya sudah ditentukan dibagian latar belakang dan diperoleh topik yang dapat dilihat pada Tabel 4.7 dengan menggunakan lima jumlah topik.

Tabel 4.7 Kata Dalam Pertopik

No Topik	Kata-kata	Pembahasanya
1	'Delta', 'Varian', 'Omicron,' 'Covid' 'Gejala', 'Bahaya', 'Tahun','Bahaya' 'Haji', 'Mati', 'puncak', 'orang-orang' 'tular', 'tinggi', 'sakit'	Mengenai bahaya varian Covid delta dan omicron
2	'vaksinasi', 'monitoring', 'babinsa', 'laksanakan' 'gencar', 'polda', 'camat' 'polsek', 'hasil', 'puskesmas', 'jadwal', 'koramil', 'lurah' 'covid'	Babinsa gencar laksanakan monitoring vaksin covid
3	'covid', 'delta', 'langka', 'minyak goreng' 'masker' 'omicron' 'alfa', 'varian', 'tuntas', 'oksigen' 'omicron', 'tahun'. 'polri', 'ekonomi' 'langka'	Pada saat varian covid delta dan omicron terjadi kelangkaan minyak goreng, masker dan oksigen
4	'jakarta', 'kota', 'giat', 'damping', 'vaksinasi', 'kabupaten', 'juni', 'senin' 'polresta', 'desa', 'balai', 'sabtu'	Pelaksanaan vaksinasi di setiap kabupaten kota jakarta
5	'vaksinasi', 'booster, 'varian', 'omicron' 'covid', 'sebar', 'sehat', 'tingkat' 'masyarakat', 'dosis, 'cegah', 'indonesia', 'lindungi', 'ayo', 'waspada' 'perintah', 'virus'	Pelaksanaan vaksinasi tingkat booster di Indonesia lindungi masyarakat dari covid varian omicron

Berdasarkan dari jumlah lima topik terdapat kesamaan bahkan ada juga yang saling berhubungan yaitu :

1. pada topik 2 dan 4 membahas mengenai pelaksanaan vaksinasi COVID-

muncul adalah pelaksanaan cepat vaksin untuk mengatasi wabah covid varian omicron maupun delta contoh *tweet* teekait kata yang sering muncul dapat dilihat pada Tabel 4.8.

Tabel 4.8 Tweet Dari Keseluruhan Kata Yang Sering Muncul

No	Tweet
1	damping laksana cepat vaksinasi covid wilayah koramil margorejo sasar masyarakat umum lansia desa sukobubuk nakes puskesmas margorejosenin
2	terap protokol sehat putus sebar covid vaksinasi dosis tiga
3	teliti kembang rapid test identifikasi varian delta omicron
4	waktu omicron sih baca berita who bilang omicron varian patut waspada tidak ganas delta mutasi omicron lemah vaksinasi kalau naik angka sih
5	tahun covid varian delta sudah kaya mati ngelewatannya diputusin doi karena bosan saya tidak sayang

4.7 EVALUASI SERTA HASIL KLASIFIKASI

Pada model klasifikasi *training* dibutuhkan jumlah 1017 data dimana masing-masing data sudah dibagi menjadi positif berjumlah 518 data dan negatif berjumlah 499 data sehingga jumlah data keseluruhan menjadi 1017 data yang sudah dilabeli secara manual. Selanjutnya menguji k-fold cross validation dapat dilakukan angka perulangan yang sebanyak 10 kali agar mendapatkan angka yang sesuai. Dari setiap perhitungan tiap fold yang ada pada k-fold cross validation sudah terdapat nilai masing-masing dan hasil yang berbeda-beda pada tiap foldnya. Dan berikut hasil perhitungan 10 fold cross validation pada *accuracy* dan *f-measure* terdapat pada Tabel 4.9.

Tabel 4.9 K-fold Accuracy Dan F-measure

Fold	Accuracy	F-measure
Fold 1	84,3 %	84,3 %
Fold 2	85,3 %	85,33 %

Fold 3	80,4 %	80,34 %
Fold 4	87,3 %	87,25 %
Fold 5	82,4 %	82,34 %
Fold 6	83,8 %	83,78%
Fold 7	82,4 %	82,32 %
Fold 8	85,3 %	85,28 %
Fold 9	85,3 %	85,3 %
Fold 10	83,3 %	83,33 %

Dari hasil yang sudah didapatkan dari perhitungan cross validation yang sudah diperhitungkan sebanyak 10 kali maka menghasilkan nilai rata-rata akurasi yang baik yaitu 83,98 untuk *accuracy* dan untuk *f-measure* 83,95.

Setelah melakukan perhitungan cross validation yang sudah diperhitungkan sebanyak 10 kali maka dapat dilakukan penghitungan model dengan cross validation dengan penghitungan *confusion matrix* pada cross validation Untuk mengetes keakuratan aplikasi perlu dilakukan peroses penghitungan menggunakan *confusion matrix* untuk mengetahui nilai dari data actual serta prediksi dapat di lihat pada Tabel.

Tabel 4.10 Confusion Matrix Pada Cross Validatiion

Kelas Prediksi	Kelas Aktual	
	Positif	Negatif
Positif	87	13
Negatif	19	84

Hasil confusion matrix pada Tabel 4.10 mendapatkan hasil dengan TP = 87, TN = 84, FP = 13 dan FN = 19. Berikutnya setelah mendapatkan nilai *confusion matrix* serta nilai dari 10 kali k-fold cross validation maka dapat dilihat hasil dari nilai akurasi data *training* pada Tabel 4.11.

Tabel 4.11 Hasil Tata Training

Akurasi Data Training	Hasil
Accuracy	83.82%
F-measure	83.95%

Pada Tabel 4.11 dapat dilihat hasil dari akurasi data *training* untuk Accuracy mendapatkan nilai yang baik yaitu 83.82% dan untuk F-measure mendapatkan hasil yang baik juga yaitu 83.95%.

Dari hasil yang sudah didapatkan dari perhitungan cross validation yang sudah diperhitungkan sebanyak 10 kali maka menghasilkan nilai rata-rata akurasi yang baik yaitu 83,98 untuk *accuracy* dan untuk *f-measure* 83,95.

Pada tahap selanjutnya akan dilakukan proses data *testing* dengan jumlah data 425 yang sudah dilabeli secara manual dan untuk mengetahui nilai-nilai akurasi pada tahapan proses data *testing* maka dilakukan perhitungan *confusion matrix* untuk dapat membedakan nilai-nilai dari data *training* maupun data *testing*. Berikut hasil perhitungan *confusion matrix* pada data *testing* dapat dilihat pada Tabel 4.13.

Tabel 4.12 Confusion Matrix Pada Data Testing

Kelas Prediksi	Kelas Aktual	
	Positif	Negatif
Positif	211	26
Negatif	17	170

Maka akan dilakukan perhitungan data testing dengan data yang sudah diketahui TP = 211, TN = 170, FP = 26 dan FN = 17 dapat di hitung dengan

menggunakan rumus pada persamaan (3) untuk menghitung *accuracy* serta menghitung *f-measure* menggunakan rumus pada persamaan (6). Dan berikut menampilkan nilai dari *accuracy* dan nilai dari *f-measure* pada data *testing* dapat dilihat pada Tabel 4.13.

Tabel 4.13 Akurasi Pada Data Testing

Akurasi Data Testing	Hasil
Accuracy	89%
F-measure	90%

Dapat dilihat pada Tabel 4.13 nilai dari data *testing* seperti *accuracy* dan *f-measure* sudah dapat diketahui kedua nilai tersebut didapat dari jumlah data testingnya yang berjumlah 425 data yang sudah dilabeli secara manual menghasilkan nilai data *testing* sebesar 89% dan 90% sehingga dapat dilihat pada proses data *training* dan *testing* memiliki nilai yang baik dan signifikan.

4.8 HASIL KLASIFIKASI DATA KESELURUHAN

Setelah melakukan proses *training* dan *testing* dan sudah didapatkan akurasi yang baik maka selanjutnya akan dilakukan proses klasifikasi dari data keseluruhan dengan jumlah data 10.300 data dari data tersebut sudah di bagi untuk *training* sebesar 1017 data yang sudah terlabeli secara manual serta 425 data untuk proses *tasting* yang juga sudah terlabeli secara manual pada proses *tasting* dan *training* sebelumnya data yang digunakan belum terdapat *topic modelling* jadi data yang akan digunakan untuk proses klasifikasi keseluruhan yaitu data yang sudah di proses di metode Latent Dirichlet Allocation (LDA) sehingga data tersebut sudah terbagi topik-topiknya dan akan di proses dengan model prediksi pada metode Naïve Bayes Classifier (NBC) sehingga dapat dilihat hasilnya pada Tabel 4.14.

Tabel 4.14 Hasil Klasifikasi Keseluruhan

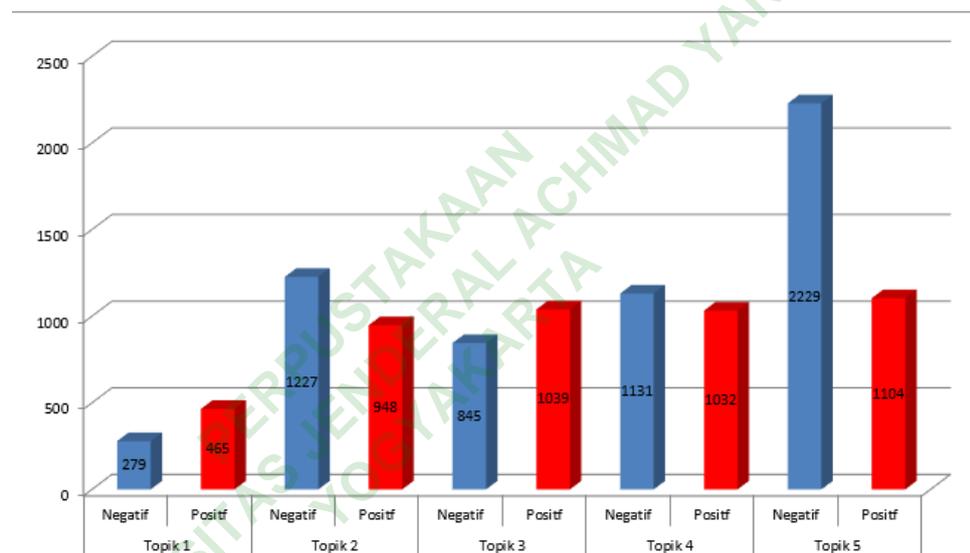
No	Tweet	Topic	Class	Result_Nbc
1	covid varian delta india booming wakanda terima tangan buka pesawat charteran dari india sekarang pmk hewan ternak gara-gara buka import sapi india	0	1	positif
2	kemenkes lapor temu subvarian omicron ba ba indonesia dampak sehat hebat varian delta masyarakat harap berhati-hati laksana protokol sehat ketat	1	0	negatif
3	anak covid tahun pas varian delta tidak yang bisa ngungkapin lihat bayi lemas tidak makan jadi kurus hrs maksa mnm macem obatvit	2	0	negatif
4	kepala sub bidang dukung sehat satgas covid brigjen tni pur alexander ginting varian delta sirkulasi indonesia hadap gelombang tiga akibat omicron	3	1	positif
5	data sebar varian sarscov provinsi milik satgas covid varian delta sirkulasi tengah masyarakat	4	1	positif

Pada Tabel 4.14 sudah dapat dilihat beberapa *tweet* beserta topic dan nilai sentimentnya, dari perhitungan topik modellingnya dimulai dari angka 0 sampai angka 4 sehingga jumlah *topic modelling* terdapat 5 topik yang ideal untuk digunakan. Hasil dari keseluruhan data dapat di lihat pada Tabel 4.15 dan dapat dilihat pada Gambar 4.9.

Tabel 4.15 Hasil Klasifikasi Keseluruhan

No	Topic	Class	Count
0	0	1	465
1	0	0	279

2	1	0	1227
3	1	1	948
4	2	1	1039
5	2	0	845
6	3	0	1131
7	3	1	1032
8	4	0	2229
9	4	1	1104



Gambar 4.9 Hasil Diagram Dari Keseluruhan Data

Pada Tabel 4.15 dan Gambar 4.9 dapat dilihat topik dan masing-masing jumlah sentimentnya pada topik 1 jumlah data positif berjumlah 465 data dan negatif berjumlah 279 data untuk topik 2 data positif berjumlah 948 dan negatif berjumlah 1227 dan topik 3 data positif berjumlah 1039 dan negatif berjumlah 845 untuk topik 4 jumlah data positif 1131 dan untuk negatif berjumlah 1032 serta topik 5 data positif berjumlah 1104 dan negatif berjumlah 2229 dari keseluruhan jumlah data yang sudah terlihat dan terbagi jika di jumlahkan akan menghasilkan jumlah data yang sesuai yaitu 10.300 data yang sudah terdapat hasil sentimen pertopik. Hasil sentimen pertopik dapat dilihat pada Tabel 4.16.

Tabel 4.16 Hasil Tweet Positif Pada Topik 1

No	Tweet	Topic	Class	Result_Nbc
1	covid varian delta india booming wakanda terima tangan buka pesawat charteran dari india sekarang pmk hewan ternak gara-gara buka import sapi india	0	1	positif
2	tahu mbah mbah buyut tahun covid varian delta usia tahun untung kumpul	0	1	positif
3	jare omicron ora gejala anosmia ib iki ora mambu mosok jek varian delta	0	1	positif
4	semenjak lord rangga bungkam terus covid susul varian delta perang dunia stand with lord rangga	0	1	positif
5	gara sentimen varian delta tidak berani datang	0	1	positif

Tabel 4.17 Hasil Tweet Negatif Pada Topik 1

No	Tweet	Topic	Class	Result_Nbc
1	siap skenario tangan delta gubernur jawa tengah ganjar pranowo rumah sakit jawa tengah siaga lonjak covid varian omicron	0	0	negatif
2	varian delta ganas haritu mati laku keluarga kakak kakak ipar sepupu belah suami left me traumatized until now	0	0	negatif
3	keno delta nang twitter lgsg sikat jare ngerti seko ndi keno delta soale tes tes sing nang masyarakat ora iso bukti kuwi varian opo	0	0	negatif
4	jalan vaksinasi ganjar lapor presiden jokowi kait siap daerah	0	0	negatif

	hadap tingkat aktif isolasi pusat skenario tanggulang varian delta			
5	dikuarantin sebab varian delta sebab radang paru mari hati awam elak sesak sentiasa sanitize duduk meja kerusi sentuh rongga badan selagi tgn sanitize	0	0	negatif

Tabel 4.18 Hasil Tweet Positif Pada Topik 2

No	Tweet	Topic	Class	Result_Nbc
1	rakyat suruh ambil vaksin untuk lawan virus yer sekarang dosis galak untuk lawan varian omicron kan habis tuu yang ve ribu juta vekes	1	1	positif
2	antigen atau apain simpul varian delta uji lab cepet gitu jenis virus	1	1	positif
3	diskusi angkat tema omicron trend dunia ancam indonesia mana delta varian narasumber baik so lewat talkshownya live youtube sobat holopis	1	1	positif
4	pasar malem bebas pdhal belah ktr lurah kondangan dangdut bebas warkop bebas tangsel tembus papar hati tangsel dominan varian delta take care sana putar	1	1	positif
5	baik ekonomi indonesia tumbuh positif kuartal ii kuartal iv tahun alami koreksi kuartal iii muncul varian delta lengkap	1	1	positif

Tabel 4.19 Hasil Tweet Negatif Pada Topik 2

No	Tweet	Topic	Class	Result_Nbc
1	bagaimana akurat virus varian alat yang pakai test stock versi delta omicron	1	0	negatif
2	negeri eropa dampak delta tidak parah drpada varian vaksin works banget vaksin mrna	1	0	negatif
3	contoh cuba sekat datang varian delta amp omicron varian selinap dalam komuniti bukti dasar kesihatan awam ekang virus	1	0	negatif
4	persis protokol covid alfa psbb janji vaksin temu pandemi selesai delta vaksinasippkm omicron omicron boosterppkm varian muter pandemi	1	0	negatif
5	maaf vaksin spesifik varian delta rebak indonesia pakai vaksin keluar negeri strainnya beda nyambung alias	1	0	negatif

Tabel 4.20 Hasil Tweet Positif Pada Topik 3

No	Tweet	Topic	Class	Result_Nbc
1	gws sebenarnya varian omicron tidak parah delta	2	1	positif
2	kemenkesri darimana tahu yang positif varian delta omicron dan lain-lain alat tes sekarang udah dasar varian	2	1	positif
3	sudah lolos varian delta delta dan lain-lain eh pas omicron tidak lolos	2	1	positif
4	dr koko vaksin buat ringan gejala bagaimana yang bilang kayak digebukin sakit bayang bagaimana remuk digebukin sert	2	1	positif

	kalu sekarang omicron vaksinya yang ringan omicronya			
5	delta tidak timbul parah hy mudah sebar sudah bilang teman dari mangga karena webinar bulan oktober ketua dari litbangkes yang bilang gt jadi bulan juni juli	2	1	positif

Tabel 4.21 Hasil Tweet Negatif Pada Topik 3

No	Tweet	Topic	Class	Result_Nbc
1	benar dari endridok antigen tentuin varian klu memang brarti prof musniumar tidak antigen saja varian delta yang jahat tapi tidak tes pake antigen	2	0	negatif
2	lahir corona varian devok sampai skrng varian omnikron sampai nanti delta kawin omnikron ngelahirin varian omnibuslaw tidak kenal nama wfh	2	0	negatif
3	pas varian delta kemarin beliau ngetweet percovidan pake bahasa lokal saking tinggi angka covid id kemarin beliau tahu covid id tweet lokal biar gampang paham	2	0	negatif
4	denger varian gabung delta sama gitu tapi yang nama virus memang gmpng ubah ubah kayak mood	2	0	negatif
5	bu dokter paru bahaya varian delta	2	0	negatif

Tabel 4.22 Hasil Tweet Positif Pada Topik 4

No	Tweet	Topic	Class	Result_Nbc
1	sekretaris daerah diy masyarakat waspada aktivitas disiplin protokol sehat wajib laksana salah langkah menagkal tular omicron kaca alam covid varian delta	3	1	positif
2	presiden tinjau vaksinasi lansia anakanak kabupaten bintang harap vaksinasi tiga vaksinasi booster pagar sebar covid varian delta omicron presiden	3	1	positif
3	vaksin lindung masyarakat serang varian delta omicron cegah laju omicron	3	1	positif
4	dinas sehat menghimbau segenap masyarakat waspada varian omicron yang muncul daerah masyarakat waspada disiplin prokes pakai masker dlm aktivitas selamat	3	1	positif
5	terap protokol sehat putus sebar covid varian omicron ba ba vaksinasi dosis tiga	3	1	positif

Tabel 4.23 Hasil Tweet Negatif Pada Topik 4

No	Tweet	Topic	Class	Result_Nbc
1	plt wali kota cimahi ngatiyana imbau masyarakat waspada covid cimahi nyata positif covid varian delta	3	0	negatif
2	Masyarakat perintah bersamasama antisipasi lonjak varian delta ulang	3	0	negatif
3	tinggi gelombang varian delta indonesia tingkat covid sebab laku jalan negeri ppls presiden jokowi imbau masyarakat jalan negeri	3	0	negatif

4	menteri koordinator maritim investasi luhut panjaitan warga negara indonesia pergi negeri tahan omicron indonesia tembus angka lampau varian delta	3	0	negatif
5	omicron varian warga panik ikn satu indonesia	3	0	negatif

Tabel 4.24 Hasil Tweet Positif Pada Topik 5

No	Tweet	Topic	Class	Result_Nbc
1	periksa masyarakat vaksin terkena varian gejala dapat rata-rata otg gejala ringan vaksin lindung pada masyarakat serang varian delta omicron cegah laju omicron	4	1	positif
2	waspada covid varian sintas omicron stay safe jokowi	4	1	positif
3	pandemi covid ojo kesusu lepas masker disiplin prokes dn waspast slurr naik positivity rate dn kasus covid akibat varian	4	1	positif
4	jokowi waspada varian omicron rangkai bhayangkara	4	1	positif
5	pemda diy ingat potensi tular sub varian omicron	4	1	positif

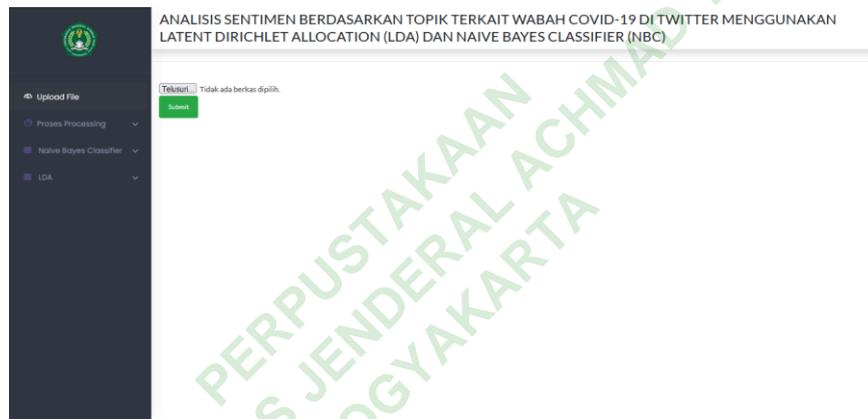
Tabel 4.25 Hasil Tweet Negatif Pada Topik 5

No	Tweet	Topic	Class	Result_Nbc
1	persen varian lapor indonesia varian omicron	4	0	negatif
2	indonesia dominasi varian omicron	4	0	negatif
3	menkes budi gunadi sadikin ungkap ciri-ciri varian omicron subvarian omicron dominasi wilayah	4	0	negatif

4	epidemiolog ugm lonjak covid garagara klaster libur lebaran varian omicron	4	0	negatif
5	hatihati varian omicron doa misi jokowi	4	0	negatif

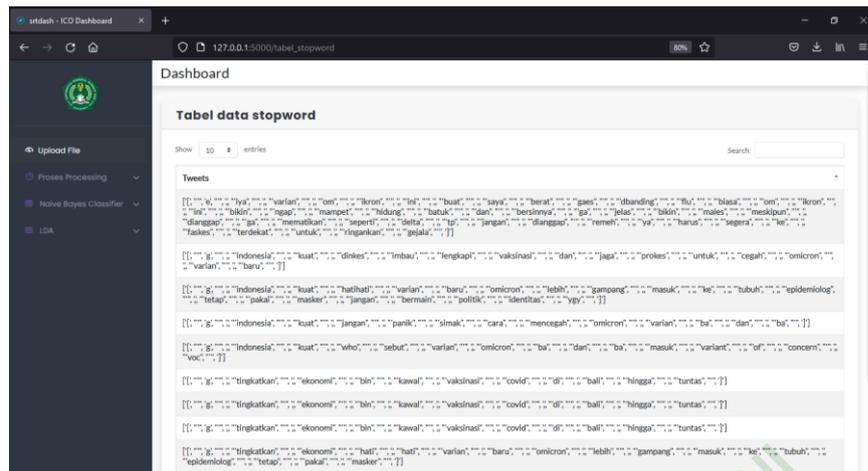
4.9 WEB DASHBOARD

web dashboard ini dibuat dengan menggunakan Python Flask yang mana pada *web dashboard* ini nantinya akan menampilkan hasil-hasil data yang sudah diolah di Jupyter Notebook berupa data file Exel tampilan *web dashboard* dapat dilihat pada Gambar 4.10 sampai 4.21.



Gambar 4.10 Tampilan Awal Dashbord

Pada Gambar 4.10 dapat dilihat tampilan awal pada *web dashboard* yang mana awal untuk menjalankan *website* tersebut dengan cara mengupload data yang sudah didapatkan dari proses *web scriping* lalu data akan diproses dengan *proprocessing* dan akan menampilkan hasil pada Gambar 4.11



The screenshot shows a web dashboard titled 'Tabel data stopwords'. The main content area displays a table of stopwords for tweets. The table has a search bar and a 'Show 10 entries' indicator. The stopwords listed are:

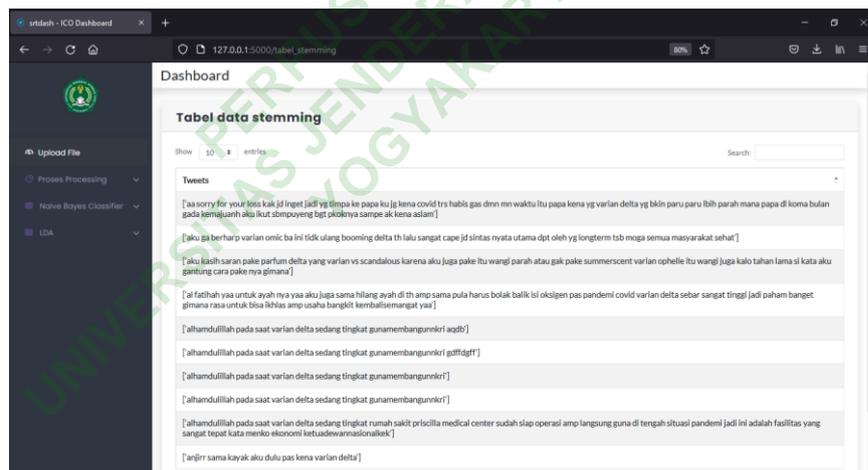
```

[[{"k": "ya", "v": "verlan", "d": "demi", "l": "liron", "b": "bi", "s": "buat", "s": "saya", "t": "beras", "g": "gasi", "d": "dhanding", "l": "lu", "s": "biasa", "o": "oni", "l": "liron", "t": "bi", "b": "bikin", "s": "ngap", "m": "manpet", "h": "hidung", "b": "batuk", "d": "dan", "b": "bersinyal", "s": "sa", "j": "jelas", "b": "bikin", "m": "males", "m": "meskipun", "d": "dianggap", "s": "ga", "m": "memastikan", "s": "seperti", "l": "delta", "t": "tr", "j": "jangan", "d": "dianggap", "r": "remeh", "l": "ya", "h": "harus", "s": "segera", "k": "ke", "t": "tugas", "t": "terdekat", "t": "untuk", "t": "riagankan", "t": "pejaka", "t": "}}]]
[[{"s": "s", "i": "Indonesia", "k": "kuat", "d": "dinkes", "l": "limbau", "t": "lengkap", "t": "vaksinasi", "d": "dan", "t": "jaga", "t": "proses", "t": "untuk", "t": "cegah", "t": "omicon", "t": "varian", "t": "baru", "t": "}}]]
[[{"s": "s", "i": "Indonesia", "k": "kuat", "t": "hati-hati", "t": "varian", "t": "baru", "t": "omicon", "t": "lebih", "t": "gampang", "t": "masuk", "t": "ke", "t": "tubuh", "t": "epidemiolog", "t": "tetap", "t": "pakai", "t": "masker", "t": "jangan", "t": "bermain", "t": "politik", "t": "identitas", "t": "ygy", "t": "}}]]
[[{"s": "s", "i": "Indonesia", "k": "kuat", "t": "jangan", "t": "panik", "t": "simak", "t": "cara", "t": "mencegah", "t": "omicon", "t": "varian", "t": "ba", "t": "dan", "t": "ba", "t": "}}]]
[[{"s": "s", "i": "Indonesia", "k": "kuat", "t": "who", "t": "sebut", "t": "varian", "t": "omicon", "t": "ba", "t": "dan", "t": "ba", "t": "masuk", "t": "varian", "t": "of", "t": "concern", "t": "ygy", "t": "}}]]
[[{"s": "s", "i": "tingkatkan", "t": "ekonomi", "t": "bin", "t": "kawal", "t": "vaksinasi", "t": "covid", "t": "di", "t": "ball", "t": "hingga", "t": "tuntas", "t": "}}]]
[[{"s": "s", "i": "tingkatkan", "t": "ekonomi", "t": "bin", "t": "kawal", "t": "vaksinasi", "t": "covid", "t": "di", "t": "ball", "t": "hingga", "t": "tuntas", "t": "}}]]
[[{"s": "s", "i": "tingkatkan", "t": "ekonomi", "t": "bin", "t": "kawal", "t": "vaksinasi", "t": "covid", "t": "di", "t": "ball", "t": "hingga", "t": "tuntas", "t": "}}]]
[[{"s": "s", "i": "tingkatkan", "t": "ekonomi", "t": "hati", "t": "hati", "t": "varian", "t": "baru", "t": "omicon", "t": "lebih", "t": "gampang", "t": "masuk", "t": "ke", "t": "tubuh", "t": "epidemiolog", "t": "tetap", "t": "pakai", "t": "masker", "t": "}}]]

```

Gambar 4.13 Table Data Stopwords

Pada Gambar 4.13 merupakan proses *stopwords* proses *stopwords* sendiri merupakan proses filtering kata-kata apa saja yang tidak perlu digunakan untuk proses pengolahan data selanjutnya. Proses selanjutnya dapat dilihat pada Gambar 4.14.



The screenshot shows a web dashboard titled 'Tabel data stemming'. The main content area displays a table of stemmed tweets. The tweets listed are:

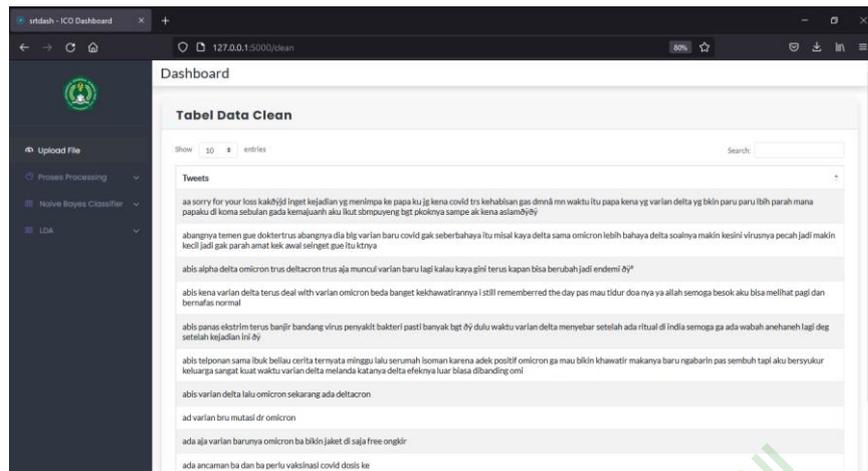
```

[sa sorry for your loss kak jd inget jadi yg tanpa ke papa ku jg kena covid trs habis gas dmn mn waktu itu papa kena yg varian delta yg bikin paru paru parah mana papa di koma bulan pada kemajuanh aku ikut sbmpuyeng bgt pliknya sampe ak kena asiam]
[aku ga berharap varian omic ba ini tsdk ulang booming delta th lalu sangat cape jd sintas nyata utama oplet oleh yg longterm tob moga semua masyarakat sehat]
[aku baah saran pake parfum delta yang varian vs scandalous karena aku juga pake itu wangj parah atau gak pake summerscent varian ophelle itu wangj juga kalo tahan lama si kata aku gantung cara pake nya gimana]
[al fatihah yaa untuk ayah nya yea aku juga sama hilang ayah di th amp sama zula harus bolak balik bi oksigen pas pandemi covid varian delta sebar sangat tinggi jdi paham banget gimana rasa untuk bisa btihas amp usaha bangkit kembalimengant yaa]
[alhamdulillah pada saat varian delta sedang tingkat gunamembangunnkri agdb]
[alhamdulillah pada saat varian delta sedang tingkat gunamembangunnkri gdfdfdf]
[alhamdulillah pada saat varian delta sedang tingkat gunamembangunnkri]
[alhamdulillah pada saat varian delta sedang tingkat gunamembangunnkri]
[alhamdulillah pada saat varian delta sedang tingkat rumah sakit pricilla medical center sudah siap operasi amp langsung guna di tengah situasi pandemi jadi ini adalah fasilitas yang sangat tepat kata menteri ekonomi ketuadevansionalak]
[anjirr sama kayak aku duhu pas kena varian delta]

```

Gambar 4.14 Tabel Data Stemming

Pada Gambar 4.14 dapat dilihat merupakan proses penampilan table *stemming* yang mana proses *stemming* sendiri merupakan proses perubahan kata menjadi kata dasar jadi semua kalimat yang ada pada table *dataset* akan dirubah menjadi kata dasar pada proses *stemming*. Proses selanjutnya dapat dilihat pada Gambar 4.15.

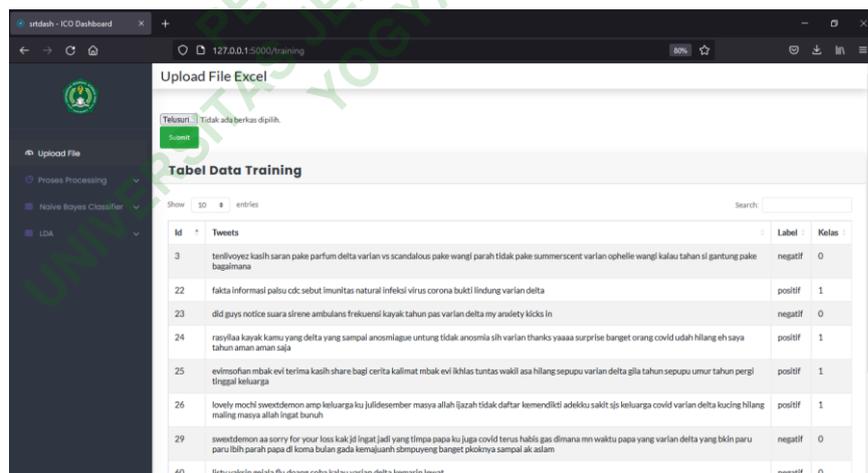


The screenshot shows a web dashboard titled 'Tabel Data Clean'. It features a sidebar with navigation options: 'Upload File', 'Proses Processing', 'Naive Bayes Classifier', and 'LDA'. The main content area displays a list of tweets under the heading 'Tabel Data Clean'. The tweets are as follows:

- aa sorry for your loss kak/jd inget kejadian yg meminima ke papa ku jg terna covid trs kehabisan gas dirnml mn waktu itu papa kena yg varian delta yg bkin paru paru lbh parah mana papaku di koma sebulan gada kemajuaahn aku likut sbmpuyeng bgt pkoknya sampe ak kena aslam/bly
- abangnya temen gue dokter trus abangnya dia big varian baru covid gak sebahaya itu misal kaya delta sama omicron lebih bahaya delta soalnya makin kesini virusnya pecah jadi makin kecil jadi gak parah amat bsk awal serngt gue itu ktrnya
- abis alpha delta omicron trus delta trus trus aja muncul varian baru lagi kalau kaya gini terus kapan bisa berubah jadi endemi dy*
- abis kena varian delta terus deal with varian omicron beda banget kekhawatirannya I still remembered the day pas mau tidur doa nya ya allah semoga besok aku bisa melihat pagi dan bernafas normal
- abis panas ekstrim terus banjir bandang virus penyakit bakteri pasti banyak bgt dy dulu waktu varian delta menyebar setelah ada ritual di india semoga ga ada wabah anehaneh lagi deg setelah kejadian ini dy
- abis telepon sama lbk bellau cerita temesta minggu lalu serumah isoman karena adek positif omicron ga mau bkin khawatir makanya baru ngabarin pas sembuh tapi aku bersyukur keluarga sangat kuat waktu varian delta melanda katanya delta efeknya luar biasa dibanding omi
- abis varian delta lalu omicron sekarang ada delta tron
- ad varian brn mutasi dr omicron
- ada aja varian barunya omicron ba bkin jaket di saja free ongkir
- ada ancaman ba dan ba perlu vaksinasi covid dosis ke

Gambar 4.15 Tabel Data Clean

Pada Gambar 4.15 merupakan proses akhir dari *preprocessing* data yang mana data sudah bersih dan sudah dapat digunakan untuk proses pengolahan data selanjutnya. Dan untuk selanjutnya akan masuk ke proses Naïve Bayes Classifier yang mana pada proses ini ada beberapa tampilan gambar yang akan ditampilkan dari *training*, *tasting* dan klasifikasi data keseluruhan yang sudah didapatkan dari proses LDA. Proses data *training* dapat dilihat pada Gambar 4.16.



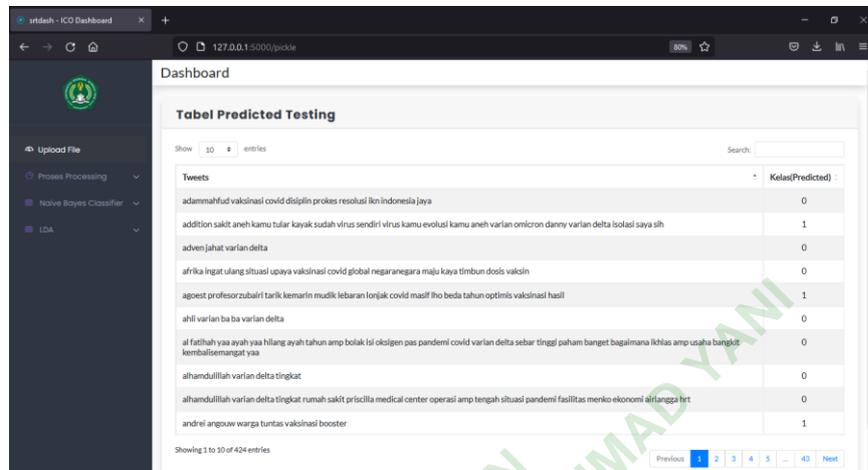
The screenshot shows a web dashboard titled 'Tabel Data Training'. It features a sidebar with navigation options: 'Upload File', 'Proses Processing', 'Naive Bayes Classifier', and 'LDA'. The main content area displays a table of tweets under the heading 'Tabel Data Training'. The table has columns for 'Id', 'Tweets', 'Label', and 'Kelas'. The data is as follows:

Id	Tweets	Label	Kelas
3	tenivoyez kasih saran pake parfum delta varian vs scandalous pake wangi parah tidak pake summerscent varian ophelle wangi kalau tahan snt gantung pake bagaimana	negatif	0
22	fakta informasi palu cd: sebut imunitas natural infeksi virus corona bukti lindung varian delta	positif	1
23	did guys notice suara sirene ambulans frekuensi kayak tahun pas varian delta my anxiety kicks in	negatif	0
24	rasylla kayak kamu yang delta yang sampai anosmiague untung tidak anosmia sih varian thanks yaaaaa surprise banget orang covid udah hilang eh saya tahun aman saja	positif	1
25	evimofan mbak evi terima kasih share bagi cerita kalimat mbak evi ikhlas tuntas wakil asa hilang sepupu varian delta gila tahun sepupu umur tahun pergi tinggal keluarga	positif	1
26	lovely mochi sweetdemon amp keluarga ku JulideseMBER masya allah ijazah tidak daftar kemendikti adekku sakit sjs keluarga covid varian delta kucing hilang maling masya allah ingat buruh	positif	1
29	sweetdemon aa sorry for your loss kak jd inget jadi yang timpa papa ku juga covid terus habis gas dimana mn waktu papa kena varian delta yang bkin paru paru lbh parah papa di koma bulan gada kemajuaahn sbmpuyeng banget pkoknya sampe ak aslam	negatif	0
60	listy vaksin gejala flu doang coba kalau varian delta kemarin lewat	negatif	0

Gambar 4.16 Tabel Data Training

Pada Gambar 4.16 proses upload data dilakukan kembali pada upload data ini yang diupload merupakan data yang khusus data training yang mana data tersebut sudah terlabeli secara manual dengan jumlah data 1017 data yang sudah

dilabeli secara manual. Pada proses selanjutnya dapat dilihat pada Gambar 4.17 yaitu penampilan pemodelan pickle.



Dashboard

Tabel Predicted Testing

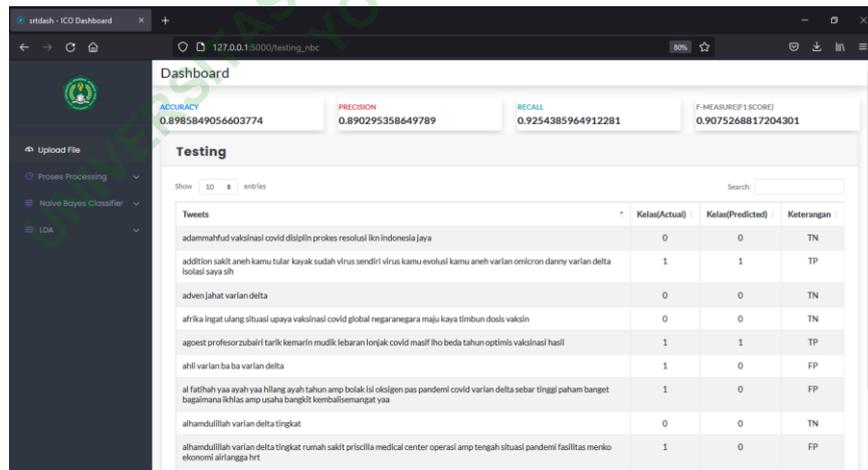
Show 10 entries

Tweets	Kelas(Predicted)
adammahfud vaksinasi covid disiplin proses resolusi lkn indonesia jaya	0
addition sakit aneh kamu tular kayak sudah virus sendiri virus kamu evolusi kamu aneh varian omicron danny varian delta isolasi saya sih	1
adven jahat varian delta	0
afrika ingat ulang situasi upaya vaksinasi covid global negaranegara maju kaya timbun dosis vaksin	0
agoest profesor zubair tarik kemarin mudik lebaran lonjak covid masif lho beda tahun optimis vaksinasi hasil	1
ahli varian ba ba varian delta	0
al fatihah yaa ayah yaa hilang ayah tahun amp bolak isi oksigen pas pandemi covid varian delta sebar tinggi paham banget bagaimana ikhlas amp usaha bangkit kembalisesangat yaa	0
alhamdulillah varian delta tingkat	0
alhamdulillah varian delta tingkat rumah sakit pricilla medical center operasi amp tengah situasi pandemi fasilitas merko ekonomi alitangga lrt	0
andrei angouw warga tuntas vaksinasi booster	1

Showing 1 to 10 of 424 entries

Gambar 4.17 Pemodelan Pickle

Pada Gambar 4.17 merupakan proses pembentukan model pickle yang mana nantinya model tersebut akan digunakan untuk membuka model untuk proses klasifikasi data keseluruhan. Proses selanjutnya adalah hasil dari data *testing* dapat dilihat pada Gambar 4.18.



Dashboard

ACCURACY: 0.8985849056603774 PRECISION: 0.890295358649789 RECALL: 0.9254385964912281 F-MEASURE(F1 SCORE): 0.9075268817204301

Testing

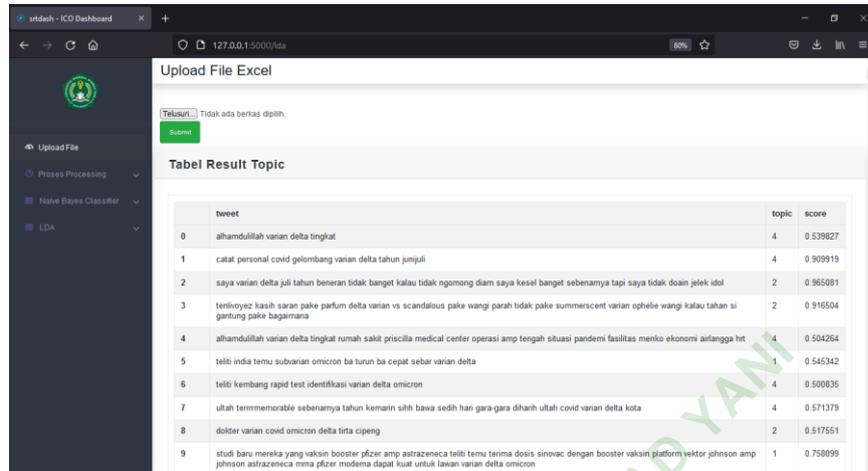
Show 10 entries

Tweets	Kelas(Actual)	Kelas(Predicted)	Keterangan
adammahfud vaksinasi covid disiplin proses resolusi lkn indonesia jaya	0	0	TN
addition sakit aneh kamu tular kayak sudah virus sendiri virus kamu evolusi kamu aneh varian omicron danny varian delta isolasi saya sih	1	1	TP
adven jahat varian delta	0	0	TN
afrika ingat ulang situasi upaya vaksinasi covid global negaranegara maju kaya timbun dosis vaksin	0	0	TN
agoest profesor zubair tarik kemarin mudik lebaran lonjak covid masif lho beda tahun optimis vaksinasi hasil	1	1	TP
ahli varian ba ba varian delta	1	0	FP
al fatihah yaa ayah yaa hilang ayah tahun amp bolak isi oksigen pas pandemi covid varian delta sebar tinggi paham banget bagaimana ikhlas amp usaha bangkit kembalisesangat yaa	1	0	FP
alhamdulillah varian delta tingkat	0	0	TN
alhamdulillah varian delta tingkat rumah sakit pricilla medical center operasi amp tengah situasi pandemi fasilitas merko ekonomi alitangga lrt	1	0	FP

Gambar 4.18 Hasil Testing

Pada proses penampilan hasil testing dapat dilihat jumlah perhitungan nilainya yang sesuai dengan proses yang ada pada Jupyter Notebooknya,

selanjutnya masuk ketahapan LDA pada tahapan LDA ada beberapa tampilan yang akan di perlihatkan dapat dilihat pada Gambar 4.19 dan 4.20.



tweet	topic	score
0 alhamdulillah varian delta tingkat	4	0.539827
1 catat personal covid gelombang varian delta tahun junjuli	4	0.909919
2 saya varian delta juli tahun beneran tidak banget kalau tidak ngomong diam saya kesel banget sebenarnya tapi saya tidak doain jelek idol	2	0.965081
3 tenivoyez kasih saran pake parfum delta varian vs scandalous pake wangi parah tidak pake summerscent varian ophelie wangi kalau tahan si gantung pake bagaimana	2	0.916504
4 alhamdulillah varian delta tingkat rumah sakit priscilla medical center operasi amp tengah situasi pandemi fasilitas menko ekonomi antanga ht	4	0.504264
5 teli india temu subvarian omicron ba turun ba cepat sebar varian delta	4	0.545342
6 teli kembang rapid test identifikasi varian delta omicron	4	0.500835
7 ulah temmemorable sebenarnya tahun kemarin sih bawa sedih hari gara-gara diarah ulah covid varian delta kuta	4	0.571379
8 dokter varian covid omicron delta birta cipeng	2	0.517551
9 studi baru mereka yang vaksin booster pfizer amp astrazeneca teli temu terima dosis sinovac dengan booster vaksin platform vektor johnson amp johnson astrazeneca mina pfizer modera dapat kuat untuk lawan varian delta omicron	1	0.750999

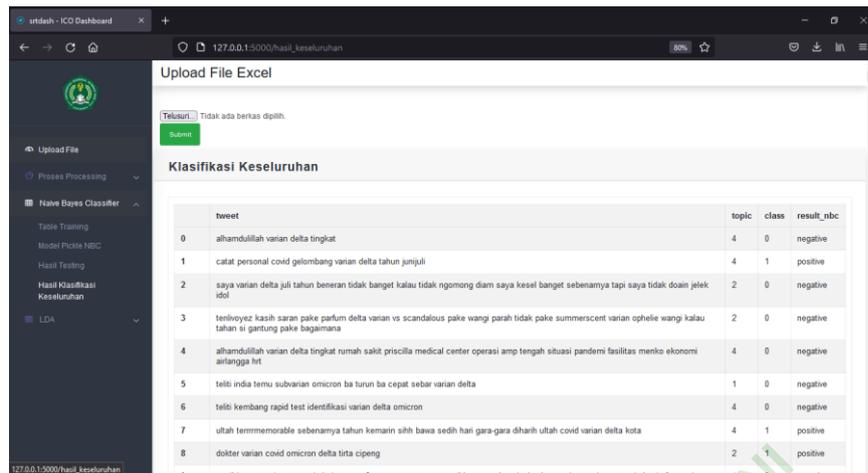
Gambar 4.19 Tabel Result Topik

Pada Gambar 4.19 merupakan penampilan data yang sudah didapatkan dari proses terakhir di metode LDA yang mana data tersebut sudah memiliki *topic* dan nilai *score* masing-masing di setiap topiknya. Selanjutnya pada Gambar 4.20 dapat dilihat *wordcloud* dari setiap topiknya.



Gambar 4.20 Wordcloud Pertopik

Gambar 4.20 dapat dilihat hasil dari *wordcloud* pertopiknya dari topik pertama sampai topik ke lima dapat terlihat kasus apa saja yang paling banyak dibahas disetiap topikya. Dan selanjutnya pada Gambar 4.21 adalah hasil dari keseluruhan data yang akan terlihat label dan topiknya.



tweet	topic	class	result_nbc
0 alhamdulillah varian delta tingkat	4	0	negative
1 catat personal covid gelombang varian delta tahun junjuli	4	1	positive
2 saya varian delta juli tahun beheran tidak banget kalau tidak ngomong diam saya kesel banget sebenarnya tapi saya tidak doain jelek idol	2	0	negative
3 tenivoyez kasih saran pake parfum delta varian vs scandalous pake wangi parah tidak pake summerscent varian ophelle wangi kalau tahan si gantung pake bagaimana	2	0	negative
4 alhamdulillah varian delta tingkat rumah sakit priscilla medical center operasi amp tengah situasi pandemi fasilitas menko ekonomi ariangga lit	4	0	negative
5 telis india temu subvarian omicron ba turun ba cepat sebar varian delta	1	0	negative
6 telis kembang rapid test identifikasi varian delta omicron	4	0	negative
7 ulah temmemorable sebenarnya tahun kemarin sihh bawa sedih hari gara-gara diarah ulah covid varian delta kota	4	1	positive
8 dokter varian covid omicron delta birta cipeng	2	1	positive

Gambar 4.21 Hasil Keseluruhan Data

Pada Gambar 4.21 proses akan di kembalikan ke metode Naïve Bayes Classifier sebelum-sebelumnya dilakukan proses *training* dan *testing* terlebih dahulu pada metode NBC dan menghasilkan nilai akurasi yang baik setelah mendapatkan nilai akurasi yang baik maka dapat dilanjutkan dengan memasukkan data yang sudah didapatkan di metode LDA yang sudah terdapat topiknya kedalam proses NBC sehingga menghasilkan nilai sentimen pertopik yang dapat dilihat pada Gambar 4.21 diatas.