

BAB III

HASIL DAN PEMBAHASAN

3.1. Hasil Pengumpulan Data

Dalam penelitian ini, data ulasan tentang hasil Aplikasi DANA yang diambil dari media Google PlayStore. Sebanyak 700 ulasan yang berhasil dikumpulkan dan disimpan dalam format CSV. Proses pengumpulan data dilakukan melalui metode *crawling* pada tanggal 27 Maret 2024. Hasil pengumpulan data pada Tabel 3.1.

Tabel 3.1 Hasil Pengumpulan Data

No.	User Nama	Waktu	Komentar
1.	Nitong99	2024-03-25 08:47:39	Buat APLIKASI DANA kebanyakan update, tapi masih lola lambat saat login maupun bertransaksi super lola, tolong lah update itu kan pembaruan, harusnya tingkatkan untuk kecepatan nya,, hapus fitur fitur yang memberatkan di Aplikasi Dana yang tidak Penting... Segera ditingkatkan!!!
2.	Suci Afriany	2024-03-24 03:52:21	Aplikasinya membantu, tapi cape setiap mau transaksi system error, dan nyoba buat upgrade baru bisa, ini masalahnya setiap mau transaksi yg error harus upgrade dulu. Masa iya upgrade trus trusan, tolong diperbaiki.
3.	M Hasan	2024-03-18 02:34:05	Lagi dan lagi, di saat pembayaran e-wallet ke shopeepay bermasalah lagi alesannya sistemnya bermasalah, terus waktu di tinggal sebentar malah "sesi kamu kadaluarsa, silahkan coba lagi" Saya coba ikutin cara developer di bawah ini ,sampai hapus data pun tidak bisa di gunakan tetap menyatakan sistem bermasalah, dari malam hingga pagi saya coba berulang kali kembali hasilnya tetap sama aja sistem bermasalah. Terakhir bisa minggu kemarin dan itu saya ulang" juga di jam 4 subuh

No.	User Nama	Waktu	Komentar
4.	DMS ALFARIZI	2024-03-15 03:32:58	Untuk fitur seperti transfer,nabung,inves,transaksi dll semua sudah oke menurutku, tapi tolong tambahkan fitur keamanan pas masuk ke apk dananya tolong buat fitur pin jadi pas masuk apk tidak langsung ke menu tampilan harus melalui pin untuk masuk terlebih dahulu supaya keamanannya makin oke, karna jika tidak nanti takutnya di hp kita otak atik sama org lain(hp dicuri) atau keluarga(anak) di apk dana kita
5.	Agung Nursetyawan	2024-03-18 14.08.24	Brengsek isi saldo malah gak bisa dipake gara ² minta upgrade premium. Coba upgrade malah gagal mulu padahal data sudah sesuai. Tolonglah kalo bikin aplikasi jangan kaya gini merugikan banyak orang. Permudah akses dan cara pemakaiannya gak perlu upgrade premium dulu baru bisa buat transaksi. Contoh noh aplikasi sebelah yang gak ribet harus upgrade premium segala.
Total			700

3.2. Hasil Labelling Data

Pelabelan data dilakukan oleh ahli bahasa (*expert*) dan model *lexicon based* untuk 700 data ulasan di Google Play Store. Dalam penelitian ini, ulasan dikategorikan menjadi tiga kelompok utama yaitu positif, negatif, dan netral. Dari data yang telah dilabeli *expert*, sebagian besar ulasan menunjukkan sentimen Negatif dan data yang diberi label *lexicon* sebagian besar ulasan menunjukkan Netral. Hasil pelabelan pada Tabel 3.2 dan 3.3.

Tabel 3.2 Hasil Labelling Data Expert

No.	Komentar	Sentimen
1.	Buat APLIKASI DANA kebanyakan update, tapi masih lola lambat saat login maupun bertransaksi super lola, tolong lah update itu kan pembaruan, harusnya tingkatkan untuk kecepatannya,, hapus fitur fitur yang memberatkan di Aplikasi Dana yang tidak Penting... Segera ditingkatkan!!	NEGATIF

No.	Komentar	Sentimen
2.	Aplikasinya membantu, tapi cape setiap mau transaksi system error, dan nyoba buat upgrade baru bisa, ini masalahnya setiap mau transaksi yg error harus upgrade dulu. Masa iya upgrade trus trusan, tolong diperbaiki.	NEGATIF
3.	Lagi dan lagi, di saat pembayaran e-wallet ke shopeepay bermasalah lagi alesannya sistemnya bermasalah, terus waktu di tinggal sebentar malah "sesi kamu kadaluarsa, silahkan coba lagi" Saya coba ikutin cara developer di bawah ini ,sampai hapus data pun tidak bisa di gunakan tetap menyatakan sistem bermasalah, dari malam hingga pagi saya coba berulang kali kembali hasilnya tetap sama aja sistem bermasalah. Terakhir bisa minggu kemarin dan itu saya ulang" juga di jam 4 subuh	NEGATIF
4.	Untuk fitur seperti transfer,nabung,inves,transaksi dll semua sudah oke menurutku, tapi tolong tambahkan fitur keamanan pas masuk ke apk dananya tolong buat fitur pin jadi pas masuk apk tidak langsung ke menu tampilan harus melalui pin untuk masuk terlebih dahulu supaya keamanannya makin oke, karna jika tidak nanti takutnya di hp kita otak atik sama org lain(hp dicuri) atau keluarga(anak) di apk dana kita	POSITIF
5.	Brengsek isi saldo malah gak bisa dipake gara ² minta upgrade premium. Coba upgrade malah gagal mulu padahal data sudah sesuai. Tolonglah kalo bikin aplikasi jangan kaya gini merugikan banyak orang. Permudah akses dan cara pemakaiannya gak perlu upgrade premium dulu baru bisa buat transaksi. Contoh noh aplikasi sebelah yang gak ribet harus upgrade premium segala.	NEGATIF
Total		700

Tabel 3.3 Hasil Labelling Data Lexicon

No.	Komentar	Sentimen
1.	Buat APLIKASI DANA kebanyakan update, tapi masih lola lambat saat login maupun bertransaksi super lola, tolong lah update itu kan pembaruan, harusnya tingkatkan untuk kecepatan nya,, hapus fitur fitur yang memberatkan di Aplikasi Dana yang tidak Penting... Segera ditingkatkan!!	Positif
2.	Aplikasinya membantu, tapi cape setiap mau transaksi system error, dan nyoba buat upgrade baru bisa, ini masalahnya setiap mau transaksi yg error harus upgrade dulu. Masa iya upgrade trus trusan, tolong diperbaiki.	Negatif
3.	Lagi dan lagi, di saat pembayaran e-wallet ke shopeepay bermasalah lagi alesannya sistemnya bermasalah, terus waktu di	Netral

No.	Komentar	Sentimen
	tinggal sebentar malah "sesi kamu kadaluarsa, silahkan coba lagi" Saya coba ikutin cara developer di bawah ini ,sampai hapus data pun tidak bisa di gunakan tetap menyatakan sistem bermasalah, dari malam hingga pagi saya coba berulang kali kembali hasilnya tetap sama aja sistem bermasalah. Terakhir bisa minggu kemarin dan itu saya ulang" juga di jam 4 subuh	
4.	Untuk fitur seperti transfer,nabung,inves,transaksi dll semua sudah oke menurutku, tapi tolong tambahkan fitur keamanan pas masuk ke apk dananya tolong buat fitur pin jadi pas masuk apk tidak langsung ke menu tampilan harus melalui pin untuk masuk terlebih dahulu supaya keamanannya makin oke, karna jika tidak nanti takutnya di hp kita otak atik sama orng lain(hp dicuri) atau keluarga(anak) di apk dana kita	Netral
5.	Brengsek isi saldo malah gak bisa dipake gara ² minta upgrade premium. Coba upgrade malah gagal mulu padahal data sudah sesuai. Tolonglah kalo bikin aplikasi jangan kaya gini merugikan banyak orang. Permudah akses dan cara pemakaiannya gak perlu upgrade premium dulu baru bisa buat transaksi. Contoh noh aplikasi sebelah yang gak ribet harus upgrade premium segala.	Netral
Total		700

3.3. Hasil *Pre-Processing*

Sebelum data ulasan Aplikasi DANA digunakan, tahap *Pre-processing* dilakukan untuk mendapatkan data bersih. Tahapan yang dilakukan antara lain *Cleaning*, *Tokenizing*, *Case Folding*, *Stopword Removal*, dan *Stemming*. Berikut ini adalah sebelum dan setelah dilakukan *Pre-processing*.

3.3.1. *Case Folding*

Tahap pertama dalam preprocessing adalah *case folding*. *Case folding* adalah proses mengubah teks menjadi huruf kecil (*lowercase*) untuk menghilangkan variasi antara huruf besar dan huruf kecil dalam analisis teks. Hasil dari data *Case Folding* pada Tabel 3.4.

Tabel 3.4 Hasil Case Folding

No.	Case Folding
1.	buat aplikasi dana kebanyakan update, tapi masih lola lambat saat login maupun bertransaksi super lola, tolong lah update itu kan pembaruan, harusnya tingkatkan untuk kecepatannya,, hapus fitur fitur yang memberatkan di aplikasi dana yang tidak penting... segera ditingkatkan!!!
2.	aplikasinya membantu, tapi cape setiap mau transaksi system error, dan nyoba buat upgrade baru bisa, ini masalahnya setiap mau transaksi yg error harus upgrade dulu. masa iya upgrade trus trusan, tolong diperbaiki.
3.	lagi dan lagi, di saat pembayaran e-wallet ke shopeepay bermasalah lagi alesannya sistemnya bermasalah, terus waktu di tinggal sebentar malah ""sesi kamu kadaluarsa, silahkan coba lagi"" saya coba ikutin cara developer di bawah ini ,sampai hapus data pun tidak bisa di gunakan tetap menyatakan sistem bermasalah, dari malam hingga pagi saya coba berulang kali kembali hasilnya tetap sama aja sistem bermasalah. terakhir bisa minggu kemarin dan itu saya ulang" juga di jam 4 subuh
4.	untuk fitur seperti transfer,nabung,inves,transaksi dll semua sudah oke menurutku, tapi tolong tambahkan fitur keamanan pas masuk ke apk dananya tolong buat fitur pin jadi pas masuk apk tidak langsung ke menu tampilan harus melalui pin untuk masuk terlebih dahulu supaya keamanannya makin oke, karna jika tidak nanti takutnya di hp kita otak atik sama orang lain(hp dicuri) atau keluarga(anak) di apk dana kita
5.	brengsek isi saldo malah gak bisa dipake gara-gara minta upgrade premium. coba upgrade malah gagal mulu padahal data sudah sesuai. tolonglah kalo bikin aplikasi jangan kaya gini merugikan banyak orang. permudah akses dan cara pemakaiannya gak perlu upgrade premium dulu baru bisa buat transaksi. contoh noh aplikasi sebelah yang gak ribet harus upgrade premium segala
Total	700

3.3.2. *Cleaning*

Tahap berikutnya adalah menghapus karakter-karakter yang tidak diperlukan dari teks. Langkah ini bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan dan bisa mengganggu proses analisis. Ulasan dalam bahasa Indonesia seringkali mengandung simbol, kalimat atau kata yang tidak baku, angka, hashtag, dan tautan URL. Hasil dari data *Cleaning* ini pada Tabel 3.5.

Tabel 3.5 Hasil Cleaning

No.	Cleaning
1.	buat aplikasi dana kebanyakan update tapi masih lola lambat saat login maupun bertransaksi super lola tolong lah update itu kan pembaruan harusnya tingkatkan untuk kecepatan nya hapus fitur fitur yang memberatkan di aplikasi dana yang tidak penting segera ditingkatkan
2.	aplikasinya membantu tapi cape setiap mau transaksi system error dan nyoba buat upgrade baru bisa ini masalahnya setiap mau transaksi yg error harus upgrade dulu masa iya upgrade trus trusan tolong diperbaiki
3.	lagi dan lagi di saat pembayaran wallet ke shopeepay bermasalah lagi alesannya sistemnya bermasalah terus waktu di tinggal sebentar malah sesi kamu kadaluarsa silahkan coba lagi saya coba ikutin cara developer di bawah ini sampai hapus data pun tidak bisa di gunakan tetap menyatakan sistem bermasalah dari malam hingga pagi saya coba berulang kali kembali hasilnya tetap sama aja sistem bermasalah terakhir bisa minggu kemarin dan itu saya ulang juga di jam subuh
4.	untuk fitur seperti transfer nabung inves transaksi dll semua sudah oke menurutku tapi tolong tambahkan fitur keamanan pas masuk ke apk dananya tolong buat fitur pin jadi pas masuk apk tidak langsung ke menu tampilan harus melalui pin untuk masuk terlebih dahulu supaya keamanannya makin oke karna jika tidak nanti takutnya di hp kita otak atik sama orng lain hp dicuri atau keluarga anak di apk dana kita
5.	brengsek isi saldo malah gak bisa dipake gara minta upgrade premium coba upgrade malah gagal mulu padahal data sudah sesuai tolonglah kalo bikin aplikasi jangan kaya gini merugikan banyak orang permudah akses dan cara pemakaiannya gak perlu upgrade premium dulu baru bisa buat transaksi contoh noh aplikasi sebelah yang gak ribet harus upgrade premium segala
Total	700

3.3.3. *Tokenizing*

Tahap berikutnya dalam *preprocessing* adalah *Tokenizing*. Proses ini melibatkan pemecahan kalimat menjadi kata-kata individual atau token. Tujuannya adalah untuk mempersiapkan teks agar bisa dianalisis pada tingkat kata. Hasil dari *tokenizing* pada Tabel 3.6.

Tabel 3.6 Hasil Tokenizing

No.	Tokenizing
1.	['buat', 'aplikasi', 'dana', 'kebanyakan', 'update', 'tapi', 'masih', 'lola', 'lambat', 'saat', 'login', 'maupun', 'bertransaksi', 'super', 'lola', 'tolong', 'lah', 'update', 'itu', 'kan', 'pembaruan', 'harusnya', 'tingkatkan', 'untuk', 'kecepatan', 'nya', 'hapus', 'fitur', 'fitur', 'yang', 'memberatkan', 'di', 'aplikasi', 'dana', 'yang', 'tidak', 'penting', 'segera', 'ditingkatkan']
2.	['aplikasinya', 'membantu', 'tapi', 'cape', 'setiap', 'mau', 'transaksi', 'system', 'error', 'dan', 'nyoba', 'buat', 'upgrade', 'baru', 'bisa', 'ini', 'masalahnya', 'setiap', 'mau', 'transaksi', 'yg', 'error', 'harus', 'upgrade', 'dulu', 'masa', 'iya', 'upgrade', 'trus', 'trusan', 'tolong', 'diperbaiki']
3.	['lagi', 'dan', 'lagi', 'di', 'saat', 'pembayaran', 'wallet', 'ke', 'shopeepay', 'bermasalah', 'lagi', 'alesannya', 'sistemnya', 'bermasalah', 'terus', 'waktu', 'di', 'tinggal', 'sebentar', 'malah', 'sesi', 'kamu', 'kadaluarsa', 'silahkan', 'coba', 'lagi', 'saya', 'coba', 'ikutin', 'cara', 'developer', 'di', 'bawah', 'ini', 'sampai', 'hapus', 'data', 'pun', 'tidak', 'bisa', 'di', 'gunakan', 'tetap', 'menyatakan', 'sistem', 'bermasalah', 'dari', 'malam', 'hingga', 'pagi', 'saya', 'coba', 'berulang', 'kali', 'kembali', 'hasilnya', 'tetap', 'sama', 'aja', 'sistem', 'bermasalah', 'terakhir', 'bisa', 'minggu', 'kemarin', 'dan', 'itu', 'saya', 'ulang', 'juga', 'di', 'jam', 'subuh']
4.	['untuk', 'fitur', 'seperti', 'transfer', 'nabung', 'inves', 'transaksi', 'dll', 'semua', 'sudah', 'oke', 'menurutku', 'tapi', 'tolong', 'tambahkan', 'fitur', 'keamanan', 'pas', 'masuk', 'ke', 'apk', 'dananya', 'tolong', 'buatkan', 'fitur', 'pin', 'jadi', 'pas', 'masuk', 'apk', 'tidak', 'langsung', 'ke', 'menu', 'tampilan', 'harus', 'melalui', 'pin', 'untuk', 'masuk', 'terlebih', 'dahulu', 'supaya', 'keamanannya', 'makin', 'oke', 'karna', 'jika', 'tidak', 'nanti', 'takutnya', 'di', 'hp', 'kita', 'otak', 'atik', 'sama', 'orng', 'lain', 'hp', 'dicuri', 'atau', 'keluarga', 'anak', 'di', 'apk', 'dana', 'kita']
5.	['brengek', 'isi', 'saldo', 'malah', 'gak', 'bisa', 'dipake', 'gara', 'minta', 'upgrade', 'premium', 'coba', 'upgrade', 'malah', 'gagal', 'mulu', 'padahal', 'data', 'sudah', 'sesuai', 'tolonglah', 'kalo', 'bikin', 'aplikasi', 'jangan', 'kaya', 'gini', 'merugikan', 'banyak', 'orang', 'permudah', 'akses', 'dan', 'cara', 'pemakaiannya', 'gak', 'perlu', 'upgrade', 'premium', 'dulu', 'baru', 'bisa', 'buat', 'transaksi', 'contoh', 'noh', 'aplikasi', 'sebelah', 'yang', 'gak', 'ribet', 'harus', 'upgrade', 'premium', 'segala']
Total	700

3.3.4. Stopword Removal

Tahap *Stopword Removal* digunakan untuk menghapus kata-kata yang tidak memiliki makna penting dalam teks, sehingga hanya kata-kata yang relevan dan

signifikan yang tersisa untuk dianalisis. Hasil dari data *Stopword Removal* pada Tabel 3.7

Tabel 3.7 Hasil Stopword Removal

No.	Stopword Removal
1.	['aplikasi', 'dana', 'kebanyakan', 'update', 'lola', 'lambat', 'login', 'bertransaksi', 'super', 'lola', 'tolong', 'update', 'pembaruan', 'tingkatkan', 'kecepatan', 'hapus', 'fitur', 'fitur', 'memberatkan', 'aplikasi', 'dana', 'ditingkatkan']
2.	['aplikasinya', 'membantu', 'cape', 'transaksi', 'system', 'error', 'nyoba', 'upgrade', 'transaksi', 'error', 'upgrade', 'iya', 'upgrade', 'trus', 'trusan', 'tolong', 'diperbaiki']
3.	['pembayaran', 'wallet', 'shopeepay', 'bermasalah', 'alesannya', 'sistemnya', 'bermasalah', 'tinggal', 'sebentar', 'sesi', 'kadaluarsa', 'silahkan', 'coba', 'coba', 'ikuti', 'developer', 'hapus', 'data', 'sistem', 'bermasalah', 'malam', 'pagi', 'coba', 'berulang', 'kali', 'hasilnya', 'sistem', 'bermasalah', 'minggu', 'kemarin', 'ulang', 'jam', 'subuh']
4.	['fitur', 'transfer', 'nabung', 'inves', 'transaksi', 'oke', 'menurutku', 'tolong', 'tambahkan', 'fitur', 'keamanan', 'pas', 'dananya', 'tolong', 'buatkan', 'fitur', 'pin', 'pas', 'langsung', 'menu', 'tampilan', 'pin', 'keamanannya', 'oke', 'karna', 'takutnya', 'hp', 'otak', 'atik', 'orng', 'hp', 'dicuri', 'keluarga', 'anak', 'dana']
5.	['brengek', 'isi', 'saldo', 'dipake', 'gara', 'upgrade', 'premium', 'coba', 'upgrade', 'gagal', 'mulu', 'data', 'sesuai', 'tolonglah', 'aplikasi', 'kaya', 'gini', 'merugikan', 'orang', 'permudah', 'akses', 'pemakaiannya', 'upgrade', 'premium', 'transaksi', 'contoh', 'noh', 'aplikasi', 'sebelah', 'ribet', 'upgrade', 'premium']
Total	700

3.3.5. Stemming

Tahap *stemming* adalah proses mengubah kata berimbuhan menjadi akar katanya. Pada tahap ini, akar kata dari setiap kata dicari menggunakan modul Sastrawi pada Python. Hasil dari *stemming* pada Tabel 3.8.

Tabel 3.8 Hasil Stemming

No	Stemming
1.	aplikasi dana banyak update lola lambat login transaksi super lola tolong update baru tingkat cepat hapus fitur fitur berat aplikasi dana tingkat

No.	Stemming
2.	<p>aplikasi bantu cape transaksi system error nyoba upgrade transaksi error</p> <p>upgrade iya upgrade trus trusan tolong baik</p> <p>bayar wallet shopeepay masalah alesannya sistem masalah tinggal sebentar sesi</p>
3.	<p>kadaluarsa silah coba coba ikutin developer hapus data sistem masalah malam</p> <p>pagi coba ulang kali hasil sistem masalah minggu kemarin ulang jam subuh</p> <p>fitur transfer nabung inves transaksi oke turut tolong tambah fitur aman pas</p>
4.	<p>dana tolong buat fitur pin pas langsung menu tampil pin aman oke karna takut</p> <p>hp otak atik orng hp curi keluarga anak dana</p> <p>brengsek isi saldo dipake gara upgrade premium coba upgrade gagal mulu data</p>
5.	<p>sesuai tolong aplikasi kaya gin rugi orang mudah akses pakai upgrade premium</p> <p>transaksi contoh noh aplikasi belah ribet upgrade premium</p>
Total	700

3.4. Hasil Pembobotan Kata (TF-IDF)

Pembobotan kata dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengidentifikasi kata-kata penting dalam ulasan aplikasi DANA di Google PlayStore. Dataset yang digunakan terdiri dari 700 baris dengan 2214 kolom, di mana setiap kolom mewakili kata unik dalam korpus data. Hasil yang ditampilkan dalam Gambar 3.1 TF-IDF menunjukkan bahwa sebagian besar kata memiliki bobot rendah dengan nilai 0, yang mengindikasikan bahwa kata-kata tersebut jarang muncul dalam dokumen.

	aaasssuuuuu	aaman	abai	abal	abdetan	abis	abistu	acc	account	\
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
..
695	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
696	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
697	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
698	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
699	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

	activate	...	yha	ynya	yotube	you	youtube	yt	ytub	yuk	zonk	\
0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
..
695	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
696	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
697	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
698	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
699	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

	Sentiment
0	Netral
1	Netral
2	Negatif
3	Positif
4	Negatif
..	...
695	Negatif
696	Negatif
697	Netral
698	Negatif
699	Positif

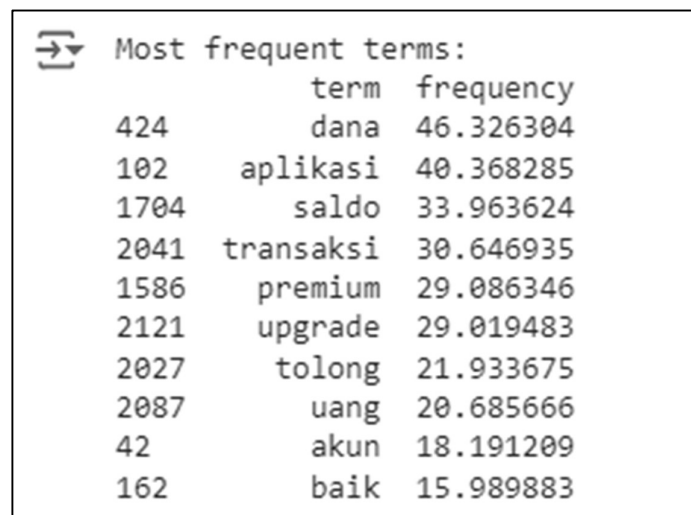
[700 rows x 2217 columns]

Gambar 3.1 Pembobotan Kata (TF-IDF)

Berdasarkan hasil skoring keseluruhan kata dengan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), dapat ditemukan kata-kata yang paling sering muncul dan memiliki bobot signifikan dalam korpus data. Gambar pada 3.2 menunjukkan bahwa kata "dana" memiliki *frekuensi* tertinggi dengan skor 46.326304, diikuti oleh kata "aplikasi" dengan skor 40.368285, dan "saldo" dengan skor 33.963624. Kata-kata lain yang juga memiliki bobot tinggi termasuk "transaksi" 30.646935, "premium" 29.086346, "upgrade" 29.019483, "tolong" 21.933675, "uang" 20.685666, "akun" 18.191209, dan "baik" 15.989883.

Frekuensi tinggi dari kata-kata ini menunjukkan topik yang dominan dan relevan dalam Google PlayStore ulasan pengguna aplikasi DANA. Kata "dana" sering muncul dalam dataset, menunjukkan bahwa topik pembicaraan banyak berkaitan dengan uang

atau sumber daya keuangan, kata "aplikasi" juga sering muncul, menunjukkan bahwa topik pembicaraan sering membahas tentang perangkat lunak atau aplikasi DANA, mungkin terkait dengan keuangan atau layanan teknologi, kemunculan kata "saldo" menunjukkan bahwa banyak diskusi berkaitan dengan saldo keuangan, kata "transaksi" menunjukkan bahwa ada banyak pembicaraan tentang aktivitas transaksi keuangan, seperti pembayaran, transfer uang, atau pembelian, kata "premium" dan "upgrade" dapat menunjukkan bahwa ada pembahasan tentang layanan atau produk premium yang mungkin melibatkan biaya tambahan atau fitur khusus, kata "tolong" mungkin muncul dalam konteks permintaan bantuan atau dukungan, menunjukkan bahwa pengguna sering meminta bantuan terkait layanan atau aplikasi, kemunculan kata "uang" menunjukkan bahwa banyak diskusi yang secara langsung berkaitan dengan uang, seperti pengelolaan uang atau masalah keuangan, kata "akun" sering muncul dalam konteks pembicaraan tentang akun pengguna, seperti pembuatan, pengelolaan, atau masalah dengan akun, kata "baik" bisa muncul dalam konteks evaluasi atau penilaian, seperti memberikan penilaian positif terhadap layanan atau aplikasi.

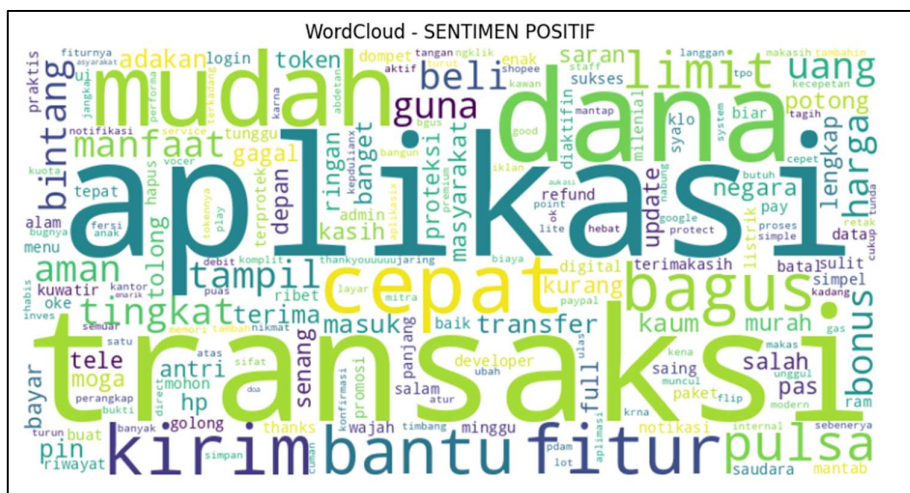


	term	frequency
424	dana	46.326304
102	aplikasi	40.368285
1704	saldo	33.963624
2041	transaksi	30.646935
1586	premium	29.086346
2121	upgrade	29.019483
2027	tolong	21.933675
2087	uang	20.685666
42	akun	18.191209
162	baik	15.989883

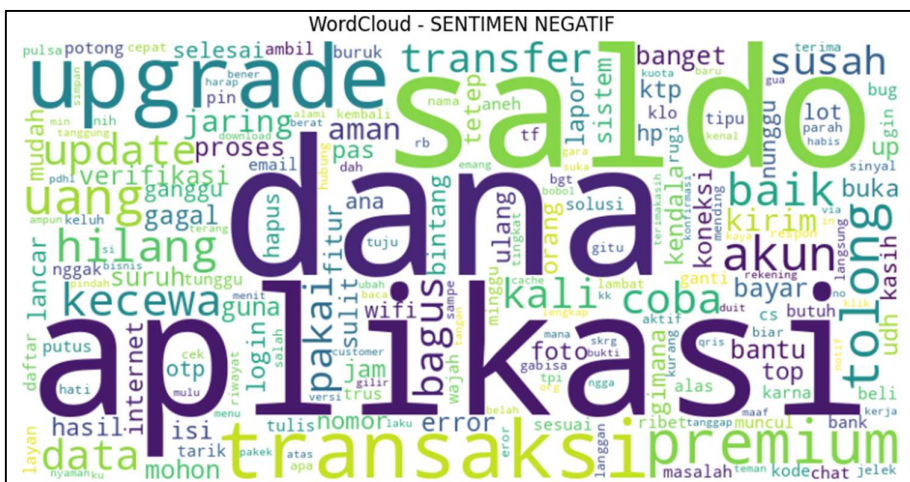
Gambar 3.2 Skoring TF

3.5. Hasil Visualiasi

Kata-kata yang sering muncul dalam teks ulasan divisualisasikan menggunakan *wordcloud*. Dalam pembuatan *wordcloud*, peneliti menggunakan library *wordcloud* dan hasil ekstraksi fitur dengan TF-IDF. Visualisasi kata yang sering muncul dalam ulasan positif pada Gambar 3.3, visualisasi kata yang sering muncul dalam ulasan negatif pada Gambar 3.4, dan visualisasi kata yang sering muncul dalam ulasan netral pada Gambar 3.5.



Gambar 3.3 WordCloud Positif



Gambar 3.4 WordCloud Negatif

Tabel 3.9 Pembagian Data KNN

Rasio	Akurasi
90:10	95%
80:20	95%
70:30	96%

Tabel 3.10 Pembagian Data Naive Bayes

Rasio	Akurasi
90:10	95%
80:20	95%
70:30	96%

Setelah dilakukan pengujian dengan 3 model rasio yaitu 90:10, 80:20, dan 70:30 didapatkan akurasi untuk masing masing rasio sebesar 95% untuk pengujian menggunakan 90:10, 95% untuk pengujian dengan rasio 80:20, dan 96% untuk pengujian dengan rasio 70:30. Dengan berbagai model rasio pengujian tersebut, menunjukkan bahwa model KNN dan Naive Bayes yang dilatih dengan teknik teks pre-processing yang benar serta pembobotan TF-IDF memberikan performa yang baik.

Hasil evaluasi menunjukkan bahwa model KNN dan Naive Bayes dengan rasio pengujian 70:30 mencapai akurasi sebesar 0.9619, Ini berarti model berhasil mengklasifikasikan sekitar 96.19% data dengan benar. Pada rasio pengujian 80:20, model menunjukkan performa dengan akurasi sebesar 0.9571, mengindikasikan bahwa sekitar 95.71% data diklasifikasikan dengan benar. Dan pada rasio pengujian 90:10, akurasi model menunjukkan akurasi yang sama 0.9571, yang berarti model mampu mengklasifikasikan sekitar 95.71% data dengan benar.

Perbedaan akurasi ini menunjukkan bahwa performa model sedikit bervariasi tergantung pada rasio pembagian data latih dan data uji. Pada rasio 70:30, model menunjukkan performa yang lebih baik dibandingkan dengan rasio 90:10 dan 80:20. Hal ini dapat disebabkan oleh berbagai faktor, seperti jumlah data uji yang lebih besar pada rasio 70:30, yang memungkinkan model untuk mempelajari lebih banyak pola dari data.

Sedangkan pada rasio 90:10 dan 80:20, jumlah data uji yang lebih sedikit mungkin tidak cukup untuk menangkap variasi yang ada dalam data, sehingga menyebabkan penurunan akurasi. Secara keseluruhan, hasil ini menunjukkan pemilihan rasio yang tepat dalam pembagian data latih dan data uji untuk mendapatkan performa model yang optimal. Pemilihan rasio yang baik dapat membantu model mempelajari pola dengan lebih baik dan menghasilkan prediksi yang lebih akurat (Kurniawan & Anubhakti, 2023) Dari ketiga rasio tersebut peneliti mengambil rasio yang terbaik yaitu 70:30.

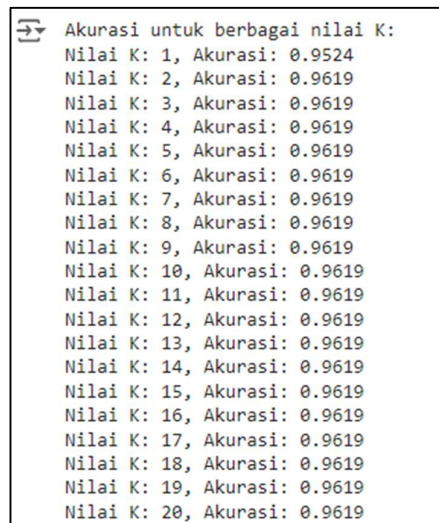
3.6.1. Metode K-Nearest Neighbor (KNN)

Pada metode KNN data yang digunakan yaitu sebanyak 700 data. Data tersebut dibagi menjadi dua kelompok yaitu data *training* (data latih) dan data *testing* (data uji) dengan menggunakan rasio 70:30 yang mendapatkan performa lebih baik. Sebanyak 490 data digunakan sebagai data *training* dan 210 data sebagai data *testing*.

Setelah menentukan rasio terbaik, dilakukan pengujian untuk mencari akurasi optimal pada berbagai nilai K dalam algoritma K-Nearest Neighbors (KNN). Jarak K dihitung menggunakan Euclidean Distance. Beberapa nilai K digunakan untuk membandingkan hasil akurasi dari masing-masing nilai. Uji ini dilakukan baik saat validasi maupun evaluasi akhir, untuk menentukan nilai K yang paling optimal di antara semua nilai yang diuji. Dengan membandingkan akurasi dari setiap nilai K, kita dapat menemukan nilai K yang memberikan kinerja terbaik untuk model tersebut (Herman et al., 2020)

Evaluasi akurasi model K-Nearest Neighbor (KNN) dengan berbagai nilai K, K1 menunjukkan akurasi 0.9524 dan K2 hingga K20 menunjukkan bahwa setiap nilai K menghasilkan akurasi yang sama, yaitu 0.9619. Hal ini menunjukkan bahwa model KNN memiliki performa yang sangat baik dan stabil pada dataset yang digunakan. Perubahan nilai K tidak mempengaruhi akurasi secara signifikan, karena

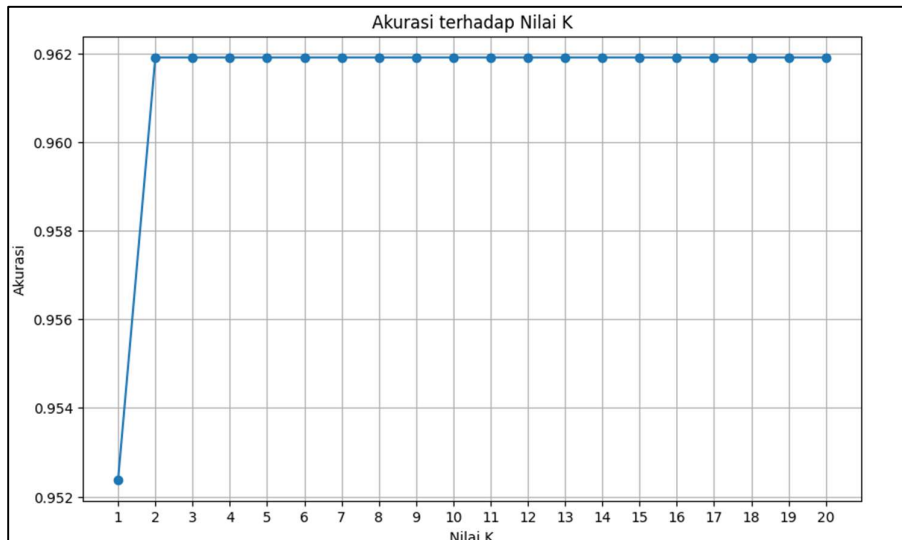
ketidakseimbangan kelas, di mana salah satu kelas memiliki jumlah sampel lebih banyak dibandingkan dengan kelas lainnya. Ketidakseimbangan ini bisa membuat model lebih mudah memprediksi kelas mayoritas dengan benar, sehingga meningkatkan akurasi keseluruhan tanpa benar-benar mencerminkan kemampuan model dalam memprediksi kelas minoritas dengan akurat (Nur Ariyanti & Cahya Wihandika, 2019) Hasilnya pada Gambar 3.6.



```
➔ Akurasi untuk berbagai nilai K:  
Nilai K: 1, Akurasi: 0.9524  
Nilai K: 2, Akurasi: 0.9619  
Nilai K: 3, Akurasi: 0.9619  
Nilai K: 4, Akurasi: 0.9619  
Nilai K: 5, Akurasi: 0.9619  
Nilai K: 6, Akurasi: 0.9619  
Nilai K: 7, Akurasi: 0.9619  
Nilai K: 8, Akurasi: 0.9619  
Nilai K: 9, Akurasi: 0.9619  
Nilai K: 10, Akurasi: 0.9619  
Nilai K: 11, Akurasi: 0.9619  
Nilai K: 12, Akurasi: 0.9619  
Nilai K: 13, Akurasi: 0.9619  
Nilai K: 14, Akurasi: 0.9619  
Nilai K: 15, Akurasi: 0.9619  
Nilai K: 16, Akurasi: 0.9619  
Nilai K: 17, Akurasi: 0.9619  
Nilai K: 18, Akurasi: 0.9619  
Nilai K: 19, Akurasi: 0.9619  
Nilai K: 20, Akurasi: 0.9619
```

Gambar 3.6 Hasil Akurasi Nilai K

Grafik yang menunjukkan hubungan antara nilai K dan akurasi dalam model K-Nearest Neighbor (KNN) menunjukkan bahwa sumbu horizontal mewakili nilai K, yang berkisar dari 1 hingga 20, sementara sumbu vertikal mewakili akurasi model. Dari grafik tersebut, terlihat K1 mendapatkan akurasi 0.952 dan K2 hingga K20 mendapatkan akurasi yang sama 0.962 untuk setiap nilai K yang diuji. Hasil ini pada Gambar 3.7.



Gambar 3.7 Grafik Akurasi Nilai K

Hasil evaluasi akurasi dari model K-Nearest Neighbor (KNN) menunjukkan bahwa K1 mencapai akurasi sebesar 0.952, sedangkan K2 hingga K20 menghasilkan akurasi yang sama, yaitu 0.962. Hal ini menunjukkan bahwa model KNN menunjukkan performa yang sangat baik dan stabil pada dataset yang digunakan, di mana perubahan nilai K tidak berpengaruh secara signifikan terhadap akurasi. Stabilitas ini mengindikasikan bahwa model KNN efektif dalam memprediksi label data uji dengan tingkat ketepatan yang konsisten. Meskipun terdapat perbedaan pada K1, nilai akurasi K2 hingga K20 tetap konsisten pada 0.962. Berdasarkan hasil ini, peneliti memilih nilai K=5 sebagai parameter final untuk model KNN.

Untuk mengevaluasi kinerja suatu model klasifikasi yang memprediksi sentimen dari teks. Evaluasi ini melibatkan perhitungan nilai-nilai seperti akurasi (*Accuracy*). Kemudian, dilakukan perhitungan persentase dari masing-masing kategori tersebut terhadap total jumlah data. *Akurasi* model dihitung berdasarkan prediksi yang dihasilkan pada data uji pada Gambar 3.8 dan 3.9.



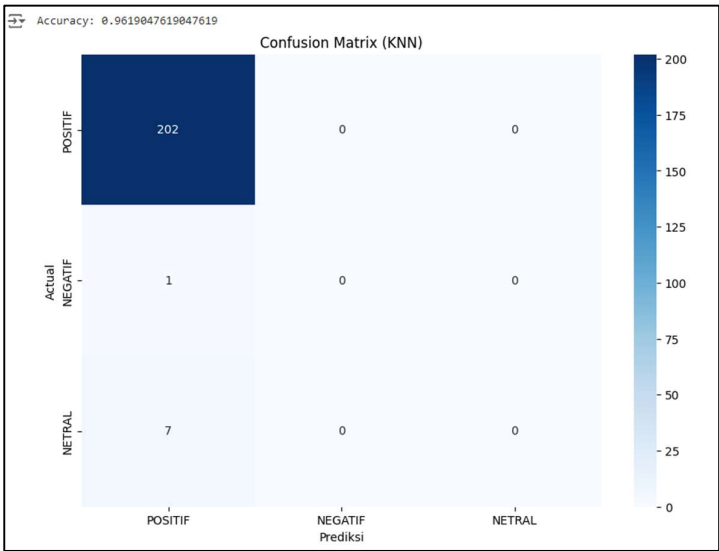
Gambar 3.8 Hasil Klasifikasi KNN Expert



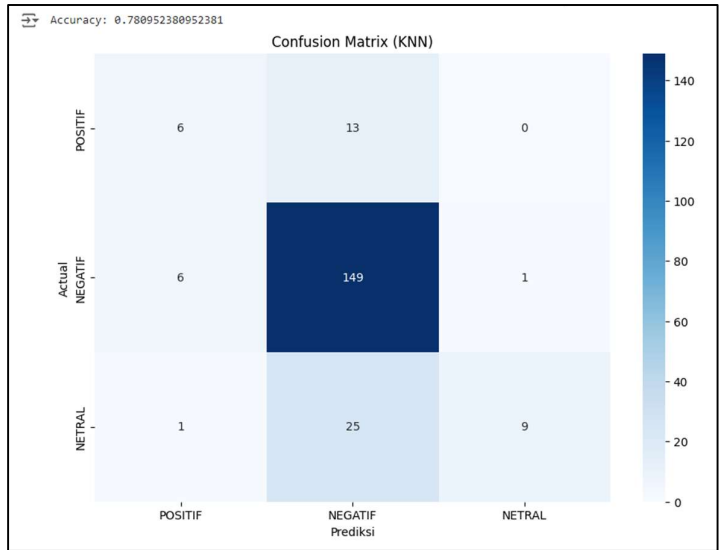
Gambar 3.9 Hasil Klasifikasi KNN Lexicon

Akurasi dari model diatas adalah akurasi yang digunakan oleh peneliti menggunakan data yang diberi label oleh (*expert*) sebesar 0.9619 atau 96%. Sebelumnya peneliti juga melakukan proses sampai tahap akurasi menggunakan data yang diberi label model *lexicon* dan mendapatkan akurasi sebesar 0.7809 atau 78%.

Untuk tampilan *Confusion Matrix* pada Gambar 3.10 dan 3.11.



Gambar 3.10 Hasil Confusion Matriks KNN Expert



Gambar 3.11 Hasil Confusion Matriks KNN Lexicon

Adapun perhitungan manual berdasarkan hasil Gambar 3.10 dan 3.11, pada Tabel 3.11 dan 3.12.

Tabel 3.11 Hasil Data Uji KNN Expert

Kategori	Prediksi Positif	Prediksi Negatif	Prediksi Netral
Aktual Positif	202	0	0
Aktual Negatif	1	0	0
Aktual Netral	7	0	0

Akurasi adalah rasio dari jumlah prediksi yang benar dan membaginya terhadap total jumlah prediksi.

Berikut adalah hasil perhitungan manual akurasi pada Persamaan (3.1).

$$Akurasi = \frac{202+0+0}{202+0+0+1+0+0+7+0+0} = \frac{202}{210} = 0.9619\% \quad (3.1)$$

Tabel 3.12 Hasil Data Uji KNN Lexicon

Kategori	Prediksi Positif	Prediksi Negatif	Prediksi Netral
Aktual Positif	6	13	0
Aktual Negatif	6	149	1
Aktual Netral	1	25	9

Berikut adalah hasil perhitungan manual akurasi pada Persamaan (3.2).

$$Akurasi = \frac{6+149+9}{6+13+0+6+149+1+1} = \frac{164}{210} = 0.7809\% \quad (3.2)$$

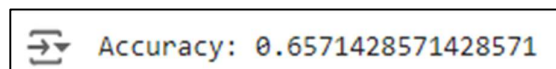
Hasil pengujian model menggunakan labelling *expert* menunjukkan akurasi sebesar 0.9619 atau 96% pada data uji, Berarti dari 210 data yang diuji model mampu memprediksi dengan benar 96 kasus. Dan hasil pengujian model menggunakan lexicon mendapatkan akurasi sebesar 0.7809 atau 78% pada data uji, Berati dari 210 data yang diuji model mampu memprediksi dengan benar 78 kasus. Akurasi ini dihitung berdasarkan jumlah prediksi yang benar dibagi dengan total jumlah kasus.

3.6.2. Metode Naive Bayes

Pada metode Naive Bayes data yang digunakan yaitu sebanyak 700 data. Data tersebut dibagi menjadi dua kelompok yaitu data *training* (data latih) dan data *testing* (data uji) dengan menggunakan rasio 70:30 yang mendapatkan performa lebih baik. Sebanyak 490 data digunakan sebagai data *training* dan 210 data sebagai data *testing*. Akurasi model dihitung berdasarkan prediksi yang dihasilkan pada data uji pada Gambar 3.12 dan 3.13.



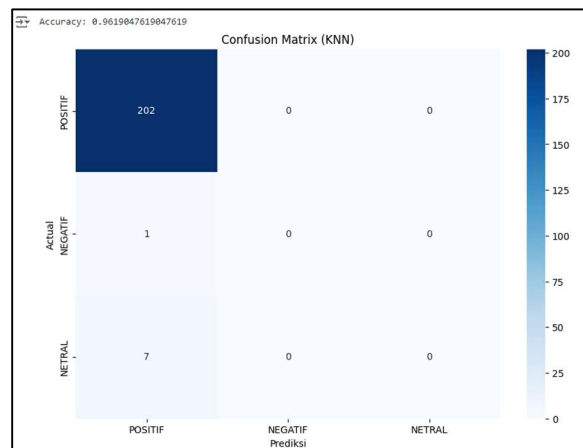
Gambar 3.12 Hasil Klasifikasi Naive Bayes Expert



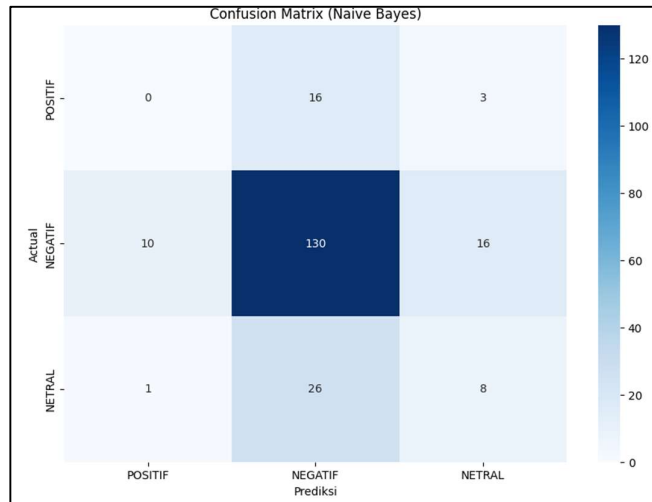
Gambar 3.13 Hasil Klasifikasi Naive Bayes Lexicon

Akurasi dari model diatas adalah akurasi yang digunakan oleh peneliti menggunakan data yang diberi label oleh (*expert*) sebesar 0.9619 atau 96%. Dan sebelumnya peneliti juga melakukan proses sampai tahap akurasi menggunakan data yang diberi label model *lexicon* dan mendapatkan akurasi sebesar 0.6571 atau 65%.

Untuk tampilan Confusion Matrix pada Gambar 3.14 dan 3.15.



Gambar 3.14 Hasil Confusion Matriks Naive Bayes Expert



Gambar 3.15 Hasil Confusion Matriks Naive Bayes Lexicon

Adapun perhitungan manual berdasarkan hasil Gambar 3.14 dan 3.15, pada Tabel 3.13 dan 3.14.

Tabel 3.13 Hasil Data Uji Naive Bayes Expert

Kategori	Prediksi Positif	Prediksi Negatif	Prediksi Netral
Aktual Positif	202	0	0
Aktual Negatif	1	0	0
Aktual Netral	7	0	0

Akurasi adalah rasio dari jumlah prediksi yang benar dan membaginya terhadap total jumlah prediksi.

Berikut adalah hasil perhitungan manual *akurasi* pada Persamaan (3.3).

$$Akurasi = \frac{202+0}{202+0+0+1+0+0+7+0} = \frac{202}{210} = 0.9619\% \quad (3.3)$$

Tabel 3.14 Hasil Data Uji Naive Bayes Lexicon

Kategori	Prediksi Positif	Prediksi Negatif	Prediksi Netral
Aktual Positif	0	16	3
Aktual Negatif	10	130	16
Aktual Netral	1	26	8

Berikut adalah hasil perhitungan manual *akurasi* pada Persamaan (3.4).

$$Akurasi = \frac{0+130}{0+16+3+10+130+16+1+26+} = \frac{138}{210} = 0.6571\% \quad (3.4)$$

Hasil pengujian model menggunakan labelling *expert* menunjukkan *akurasi* sebesar 0.9619 atau 96% pada data uji, Berarti dari 210 data yang diuji model mampu memprediksi dengan benar 96 kasus. Dan hasil pengujian model menggunakan *lexicon* mendapatkan akurasi sebesar 0.6571 atau 65% pada data uji, Berati dari 210 data yang diuji model mampu memprediksi dengan benar 65 kasus. *Akurasi* ini dihitung berdasarkan jumlah prediksi yang benar dibagi dengan total jumlah kasus.

Dari kedua metode diatas terdapat perbedaan akurasi jika seorang peneliti menggunakan data yang sama tetapi menggunakan labelling yang berbeda. jika peneliti menggunakan data yang diberi labelling model *lexicon*, pada metode KNN mendapatkan akurasi sebesar 78% dan metode Naive Bayes mendapatkan akurasi sebesar 65%. Dan dapat dilihat terjadi perbedaan jika dibandingkan dengan menggunakan data yang diberi label oleh (*expert*) menggunakan metode KNN dan Naive Bayes mendapatkan akurasi yang sama sebesar 96%.

Perbedaan *akurasi* tersebut disebabkan karena terjadinya perubahan kelas sentimen yang diberi label oleh seorang ahli bahasa (*expert*) dan label menggunakan model *lexicon*. Terdapat penambahan pada kelas positif yang diberi label *lexicon*, dari 23 data kelas positif menjadi 104 data positif. Selanjutnya penambahan pada kelas netral dari 5 data kelas negatif menjadi 520 data netral. Dan menurunnya kelas negatif yang diberi label *expert*, dari 672 data kelas negatif menjadi 76 data netral. Hal inilah yang membuat terjadi perbedaan akurasi dari kedua metode tersebut.