

BAB III

METODOLOGI PENELITIAN

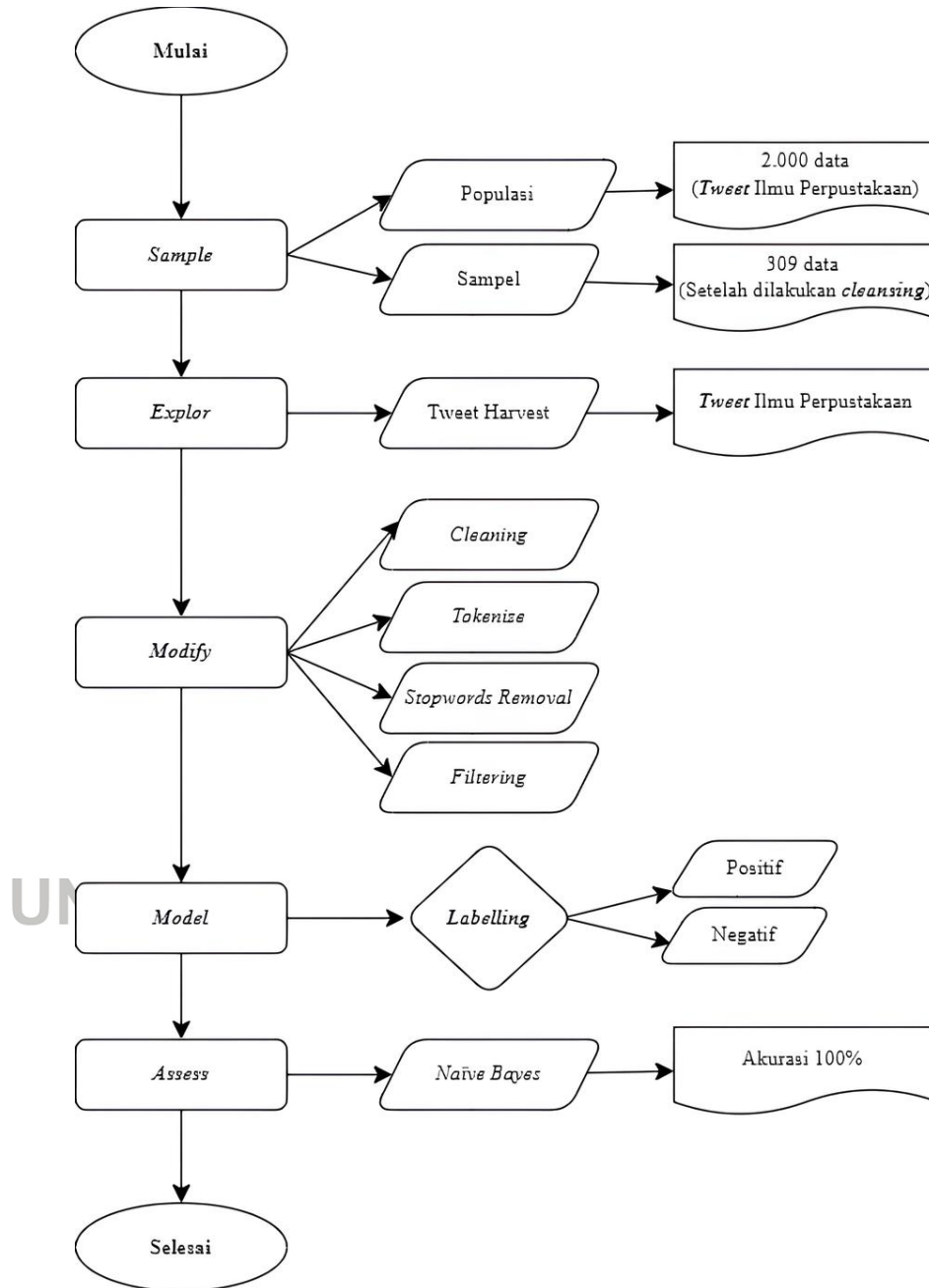
3.1 Jenis Penelitian

Metode kuantitatif dengan pendekatan deskriptif digunakan dalam penelitian ini. Menurut Sugiyono (2010) penelitian kuantitatif adalah Metode penelitian yang berlandaskan pada filsafat positivisme, digunakan untuk meneliti pada populasi atau sampel tertentu, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif, dengan tujuan untuk menguji hipotesis yang telah ditetapkan.

Metode ini digunakan Karena kemampuannya untuk mengumpulkan dan mengevaluasi data sentimen dengan cara yang tidak memihak dan metodis, strategi ini dipilih oleh para peneliti. Dengan menggunakan informasi yang diperoleh dari ulasan dan komentar di platform jejaring sosial X, teknik deskriptif diterapkan untuk mengkarakterisasi situasi saat ini. Perasaan positif dan negatif dari komentar pengguna platform jejaring sosial X akan diidentifikasi dan diukur melalui penelitian ini.

Penelitian ini menggunakan teori SEMMA yang dikembangkan pada tahun 1990-an. Langkah pertama, *Sample*, adalah mengambil sampel dari data ulasan yang diambil dari platform jejaring sosial media X. Tahap *Explore* ini penting untuk mendapatkan pemahaman awal tentang struktur data dan korelasi antar variabel.. *Modify*, terbagi atas *cleansing*, *tokenize*, *stopword*, dan *filtering*. *Model*, digunakan untuk memprediksi sentimen berdasarkan data yang dimodifikasi. Untuk memastikan bahwa model tersebut memberikan hasil yang akurat dan dapat dipercaya, tahap *Assess* dilakukan di tahap terakhir.

3.2 Desain Penelitian



Gambar 3. 1 Desain penelitian

Dari diagram diatas, Data diekstraksi dari populasi untuk digunakan sebagai sampel analitik pada tahap Sampel, yang merupakan awal dari prosedur ini. Data yang lebih besar yang akan menjadi sasaran analisis tambahan diwakili oleh sampel ini. Alih-alih memproses seluruh populasi, proses analisis dapat diselesaikan lebih cepat dengan menggunakan data sampel. Agar temuan analisis dapat menggambarkan keadaan secara akurat, prosedur pengambilan sampel harus dilakukan secara representatif. Data siap untuk tahap eksplorasi untuk mendapatkan informasi lebih lanjut setelah pengambilan sampel selesai.

Metodologi *Tweet Harvest*, yang mengumpulkan data dari media sosial, khususnya Twitter (sekarang X), digunakan untuk mengumpulkan data pada tahap Eksplorasi. Karena ada begitu banyak sudut pandang yang berbeda di media sosial, analisis sentimen sering kali menggunakan platform ini sebagai sumber data. Berbagai frasa pengguna yang mungkin menyampaikan perasaan positif, atau negatif dapat ditemukan dalam tweet yang dikumpulkan. Setelah itu, data ini diperiksa untuk mencari tren dan pola dalam opini publik. Setelah pengumpulan data selesai, data tersebut harus dimodifikasi agar dapat digunakan untuk analisis tambahan. Perubahan ini dilakukan dalam beberapa tahap untuk menjamin bahwa kualitas data tetap dalam kondisi terbaiknya.

Cleaning, Tokenize, Stopword, dan *Filtering* adalah empat langkah utama dari tahap Modifikasi. Tujuan dari pembersihan adalah untuk menghilangkan karakter asing, termasuk simbol, angka, dan tanda baca yang tidak perlu. Tokenisasi adalah proses membagi teks menjadi kata-kata atau token sehingga analisis tambahan dapat dilakukan. Tujuan dari penghilangan stopwords adalah untuk menghilangkan istilah yang sering digunakan yang tidak menambah apa pun pada analisis. Langkah terakhir dalam proses perubahan adalah penyaringan, yang mencari istilah yang benar-benar berkaitan dengan sikap yang diteliti. Data menjadi lebih bersih dan siap untuk diproses lebih lanjut dengan menyelesaikan tahap modifikasi ini.

Tahap Model digunakan untuk memberi label pada data setelah diubah. Proses pelabelan melibatkan pengklasifikasian data ke dalam klasifikasi sentimen, seperti positif dan negatif. Orang atau algoritma tertentu dapat melakukan pelabelan ini secara otomatis. Model pembelajaran mesin akan menggunakan data yang telah dilabeli sebagai data pelatihan. Model ini dapat mengidentifikasi pola dan ciri-ciri yang membedakan berbagai jenis sentimen melalui pelabelan. Setelah pelabelan selesai, model disiapkan untuk evaluasi dan pengujian kinerja.

Tujuan dari langkah Assess adalah untuk menilai model yang dibuat dengan algoritma Naive Bayes. Karena efisiensi dan kecepatannya, metode Naive Bayes berbasis probabilitas sering digunakan dalam kategorisasi teks. Dengan menggunakan data pelatihan, metode ini menentukan kemungkinan sebuah kata akan muncul dalam kategori emosi tertentu. Keakuratan dan kinerja model kemudian dinilai dengan membandingkan keluarannya dengan data uji. Dengan mengubah parameter atau menggunakan data yang lebih bersih, akurasi model dapat ditingkatkan jika masih buruk. Prosedur analisis sentimen dianggap selesai jika model telah dinilai dan hasilnya memuaskan.

Diagram alir ini menunjukkan bahwa analisis sentimen adalah prosedur metodis yang memerlukan sejumlah langkah penting. Setiap langkah sangat penting untuk menjamin keakuratan dan ketergantungan temuan analisis. Setiap tahap, mulai dari pengambilan sampel hingga evaluasi model, perlu dilakukan dengan cermat untuk mencegah bias pada temuan akhir. Bisnis, politik, dan penelitian sosial adalah beberapa industri yang sangat diuntungkan dari analisis sentimen. Organisasi dapat menggunakan data yang mereka miliki untuk membuat keputusan yang lebih baik dengan memahami tren suasana hati publik. Oleh karena itu, teknik seperti Naive Bayes sangat membantu ketika menganalisis data dalam jumlah besar dari media sosial dan sumber lainnya.

3.3 *Sample* (Populasi dan Sampel)

Mengumpulkan informasi tentang ulasan dan komentar terkait ilmu perpustakaan pada platform jejaring sosial X merupakan langkah pertama dalam desain penelitian ini. Popularitas aplikasi ini di kalangan pengguna yang sering membicarakan berbagai topik, termasuk ilmu perpustakaan, menjadi alasan pemilihan aplikasi ini.

Istilah populasi mengacu pada keseluruhan, totalitas, atau generalisasi dari unit, individu, objek, atau subjek yang memiliki kuantitas dan karakteristik tertentu yang perlu dipelajari. Ini dapat mencakup orang, benda, institusi, peristiwa, dan entitas lain yang dapat menghasilkan data untuk penelitian yang darinya kesimpulan dapat dibuat. Maka, populasi yang diambil dalam penelitian ini adalah semua tweet yang berhubungan dengan ilmu perpustakaan.

Agar dapat dianggap mewakili semua populasi yang diteliti, sampel haruslah merupakan bagian representatif dari populasi dengan atribut dan sifat yang sama yang menjadi ciri populasi. untuk membantu peneliti membuat generalisasi tentang populasi yang diwakilinya, prosedur pengambilan sampel sangat membantu. Maka sampel yang diambil untuk penelitian ini adalah 203 sentimen.

3.4 *Explore* (Objek dan Instrumen Penelitian)

Penelitian ini berfokus pada Ilmu Perpustakaan sebagai objek utama yang dikaji melalui analisis sentimen di platform jejaring sosial X. Ilmu Perpustakaan merupakan bidang yang berkaitan dengan pengelolaan informasi, sumber daya perpustakaan, serta layanan kepada pemustaka. Dalam era digital, banyak diskusi dan opini mengenai bidang ini yang berkembang di media sosial, termasuk di platform jejaring sosial X. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi dan menganalisis sentimen masyarakat terhadap ilmu perpustakaan dengan menggunakan teknik pemrosesan bahasa alami (Natural Language Processing/NLP) dan pembelajaran mesin (Machine Learning).

Langkah pertama dalam penelitian ini adalah mengumpulkan data dari aplikasi X menggunakan teknik *Tweet harvesting*. *Tweet harvesting* adalah metode otomatis yang memungkinkan pengunduhan data dalam jumlah besar berdasarkan kata kunci tertentu. Dalam penelitian ini, kata kunci yang digunakan adalah "Ilmu Perpustakaan" dengan batas maksimal 2.000 tweet dalam satu kali pengambilan. Data yang diperoleh akan mencakup berbagai opini, baik yang bersifat positif, negatif, yang akan dianalisis lebih lanjut untuk mendapatkan gambaran umum mengenai persepsi masyarakat terhadap bidang ini.

Setelah data dikumpulkan, langkah selanjutnya adalah merekamnya ke dalam format file CSV (Comma-Separated Values). Format ini dipilih karena kompatibel dengan berbagai alat analisis data dan memungkinkan pemrosesan yang lebih mudah dalam tahap selanjutnya. Data yang tersimpan dalam file CSV akan mencakup informasi seperti isi tweet, tanggal pengambilan, serta metadata lainnya yang relevan untuk analisis. Tahap ini penting untuk memastikan bahwa data yang digunakan dalam penelitian ini telah terorganisir dengan baik dan siap untuk diproses menggunakan metode analisis sentimen.

Setelah proses penyimpanan data selesai, analisis sentimen akan diterapkan pada dataset yang telah dikumpulkan. Analisis sentimen ini bertujuan untuk mengidentifikasi apakah suatu opini bersifat positif, negatif berdasarkan isi tweet yang telah dikumpulkan. Proses ini akan dilakukan dengan menggunakan alat analisis sentimen berbasis pembelajaran mesin, yang memungkinkan klasifikasi sentimen secara otomatis dengan tingkat akurasi yang tinggi. Hasil dari analisis ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai bagaimana ilmu perpustakaan dipersepsikan oleh masyarakat di media sosial, serta faktor-faktor yang mempengaruhi sentimen tersebut.

3.5 *Modify (Analisis Data)*

Pemrosesan informasi adalah tahap ketiga. Data yang belum diolah akan melalui beberapa tahap, yaitu:

- a. *Cleaning* : Menghilangkan elemen yang tidak penting
- b. *Tokenize* : Proses memecah teks menjadi unit-unit yang lebih kecil
- c. *Stopword Removal* : Mengeluarkan kata-kata yang sangat umum dan tidak penting untuk analisis teks
- d. *Filtering* : Teknik penyaringan teks untuk menjamin bahwa data hanya berisi elemen dan kata yang relevan.

3.6 *Model (Pelabelan Data)*

Tahap keempat dalam penelitian ini adalah analisis sentimen, yang bertujuan untuk mengklasifikasikan opini masyarakat mengenai ilmu perpustakaan menjadi kategori positif atau negatif. Analisis sentimen ini penting untuk memahami persepsi publik terhadap bidang ilmu perpustakaan, terutama di era digital yang semakin berkembang. Dengan teknik ini, peneliti dapat mengidentifikasi bagaimana ilmu perpustakaan dipandang oleh masyarakat, apakah memiliki citra positif atau justru masih mengalami stigma tertentu yang perlu diperbaiki.

Dalam proses analisis sentimen ini, data yang telah dikumpulkan akan diberi label secara manual dan otomatis. Pelabelan manual dilakukan dengan membaca dan memahami setiap opini untuk mengkategorikannya sebagai sentimen positif atau negatif. Sementara itu, pelabelan otomatis akan menggunakan teknik TF-IDF (*Term Frequency-Inverse Document Frequency*), sebuah metode statistik yang mengukur seberapa penting suatu kata dalam sebuah dokumen dibandingkan dengan keseluruhan dataset. Dengan teknik ini, sistem dapat secara otomatis menentukan pola kata yang sering muncul dalam opini positif maupun negatif, sehingga klasifikasi sentimen dapat dilakukan secara lebih efisien.

Penggunaan kombinasi pelabelan manual dan otomatis diharapkan dapat meningkatkan akurasi hasil analisis sentimen. Pelabelan manual memungkinkan adanya validasi terhadap data yang diperoleh, sementara pelabelan otomatis dengan TF-IDF membantu mempercepat proses klasifikasi dalam jumlah data yang besar. Hasil dari tahap ini akan memberikan gambaran yang lebih jelas mengenai bagaimana masyarakat memandang ilmu perpustakaan, serta faktor-faktor apa saja yang mempengaruhi sentimen tersebut. Temuan ini nantinya dapat digunakan sebagai dasar untuk mengembangkan strategi dalam meningkatkan citra dan pemahaman masyarakat tentang ilmu perpustakaan.

3.7 Assess (Uji Akurasi Data dan Implementasi)

Langkah kelima adalah uji keakuratan data menggunakan algoritma Naive Bayes. Berikut adalah persamaannya:

$$P(Y|X) = \frac{P(x|y)(Y)}{P(X)}$$

Persamaan di atas menunjukkan bahwa Y sebagai kelas spesifik, sedangkan X sebagai kelas yang belum diketahui. $(Y|X)$ merupakan probabilitas dari kelas berdasarkan hipotesa sebelumnya, sedangkan $P(X)$ merupakan probabilitas dari Y . $P(Y|X)$ adalah hasil perkalian antara *likelihood* dan *prior* dibagi *evidence*. *Likelihood* adalah probabilitas atribut data X pada kelas Y , *prior* adalah probabilitas kelas Y dari total data set, dan *evidence* adalah probabilitas atribut data X dari total data set.

Menafsirkan hasil analisis adalah langkah keenam. Untuk mendapatkan pemahaman tentang bagaimana masyarakat umum memandang ilmu perpustakaan, hasil analisis sentimen dan faktor-faktor yang mempengaruhinya akan dianalisis. Hasilnya akan dikontraskan dengan hasil penelitian sebelumnya untuk mengungkap tumpang tindih dan perbedaan.

Langkah terakhir adalah pengembangan rekomendasi. Berdasarkan hasil interpretasi, studi ini akan mengusulkan rekomendasi praktis untuk lembaga dan praktik-praktik terkait. Rekomendasi ini bertujuan untuk meningkatkan layanan dan strategi komunikasi agar lebih memenuhi kebutuhan dan harapan masyarakat.

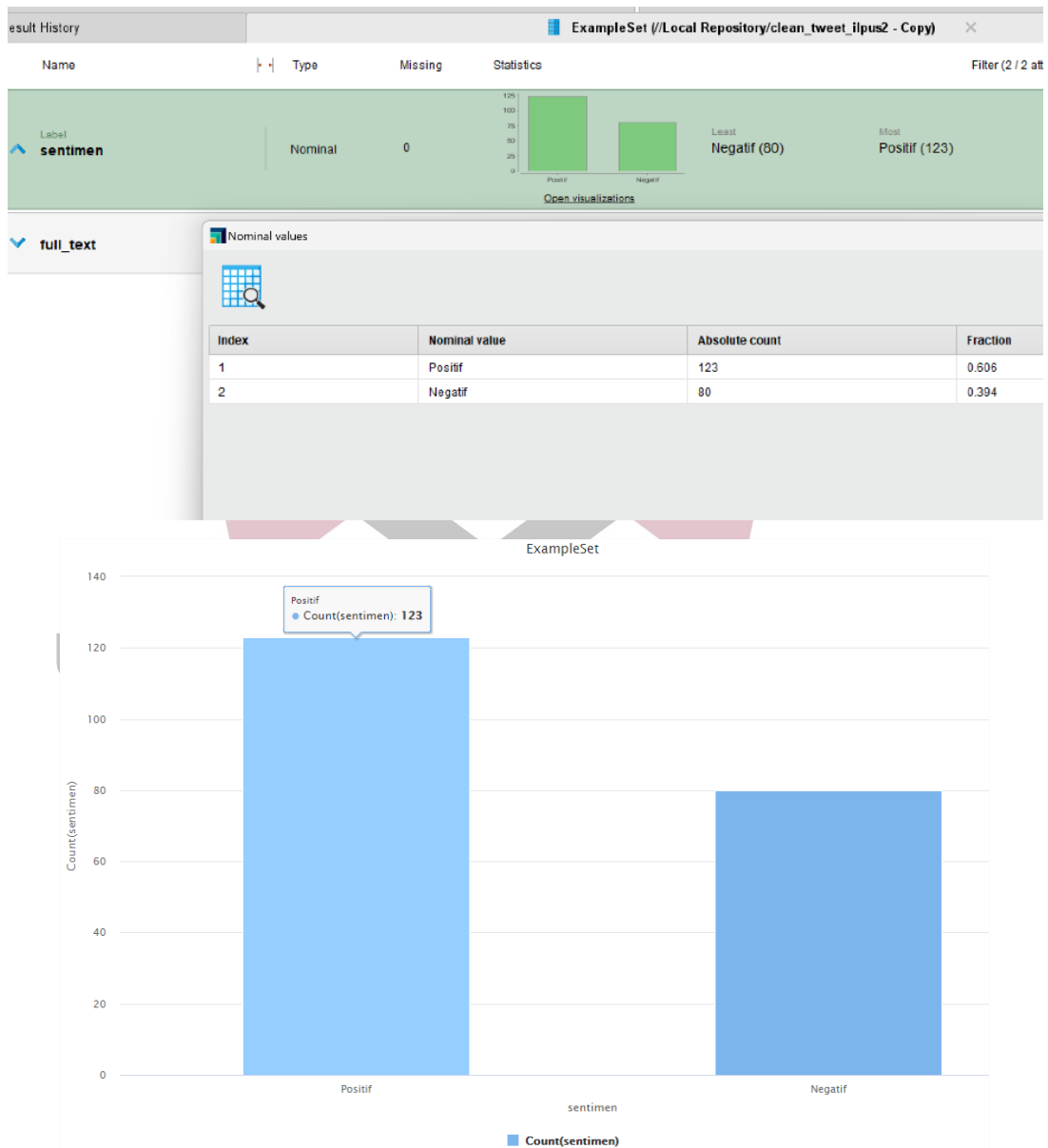
Tujuan dari metode penelitian ini adalah untuk memastikan bahwa data yang dikumpulkan dan dianalisis bersifat obyektif dan sistematis. Dengan pendekatan ini, diharapkan penelitian ini dapat memberikan kontribusi yang berharga bagi kemajuan literasi informasi di era digital.



BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Deskripsi Data



Gambar 4 1Diagram

Statistik data sentimen pada grafik pertama dipisahkan menjadi dua kelompok, yaitu positif dan negatif. Terdapat 203 total titik data untuk tipe data nominal dari label sentimen. Terdapat 123 titik data sentimen positif dan 80 titik sentimen negatif dari total ini. Grafik batang yang mewakili angka-angka ini juga ditampilkan, dengan batang yang mewakili emosi positif lebih besar daripada yang mewakili sentimen negatif. Perbandingan ini menunjukkan seberapa umum sentimen positif dalam data yang dianalisis.

Gambar pertama menyertakan tabel nilai nominal yang menunjukkan jumlah absolut dan proporsi dari setiap kategori sentimen di samping grafik batang. Dengan jumlah absolut 123, perasaan positif mencapai 50,6% dari keseluruhan data. Sebaliknya, ada 80 perasaan tidak menyenangkan, atau 39,4% dari keseluruhan. Informasi ini membuat distribusi sentimen dalam kumpulan data lebih mudah dipahami dan membantu dalam menentukan tren opini yang dianalisis. Kita dapat menentukan seberapa umum pikiran positif dalam diskusi dengan memeriksa distribusi ini.

Representasi grafik batang yang lebih jelas dari distribusi sikap positif dan negatif dapat ditemukan pada gambar kedua. Grafik ini menunjukkan bagaimana jumlah data tentang sentimen positif jauh lebih banyak daripada sentimen negatif, dengan puncaknya di angka 123. Mayoritas komentar atau balasan yang terkumpul bersifat positif, seperti yang ditunjukkan oleh banyaknya batang sentimen positif. Namun, karena ada lebih sedikit pendapat yang tidak menguntungkan, bar untuk emosi negatif lebih pendek. Pembaca dapat lebih mudah memahami perbandingan jumlah perasaan secara sekilas dari representasi grafis ini.

Terdapat kecenderungan opini yang berlaku, seperti yang ditunjukkan oleh perbandingan grafik statistik positif dan negatif. Persentase sentimen positif yang lebih tinggi menunjukkan bahwa publik merespons atau mendukung topik atau masalah yang sedang dianalisis dengan lebih baik. Di sisi lain, meskipun jumlah sentimen negatif lebih sedikit, keberadaannya tetap signifikan bagi penelitian secara keseluruhan.

Ketika menyusun strategi atau membuat penilaian berdasarkan sentimen, data ini diperhitungkan. Data ini dapat dianalisis untuk meningkatkan layanan atau memenuhi permintaan audiens dengan lebih baik dengan analisis yang tepat.

Jika digabungkan, kedua gambar ini menawarkan penggambaran menyeluruh tentang distribusi dan persentase sentimen set data. Kita dapat dengan jelas melihat distribusi dan dominasi pemikiran positif dan negatif dengan menggabungkan data numerik dengan visualisasi grafis. Analisis sentimen menggunakan data ini untuk beberapa alasan, termasuk melacak opini publik, mengukur kebahagiaan konsumen, dan mengevaluasi layanan. Metode ini meningkatkan akurasi dan kemampuan interpretasi hasil analisis. Oleh karena itu, keputusan yang diambil berdasarkan informasi ini dapat lebih terfokus dan peka terhadap kebutuhan saat ini.

4.2 Temuan penelitian

Dalam penelitian ini, opini tentang ilmu perpustakaan menjadi sasaran analisis sentimen. Menurut penelitian tersebut, 123 sikap, atau sekitar 60,6% dari total data, menyatakan pandangan yang baik terhadap bidang ini. Hal ini menunjukkan bahwa sebagian besar pengguna atau responden memiliki opini positif terhadap ilmu perpustakaan, terutama dalam hal keuntungan, penggunaan, dan pengalamannya. Ketertarikan masyarakat terhadap ilmu perpustakaan dapat disimpulkan dari kesan positif ini. Kehadirannya dianggap sangat penting untuk memfasilitasi manajemen dan akses informasi. Kemajuan teknis yang digunakan dalam ilmu perpustakaan adalah elemen lain yang berkontribusi terhadap relevansi dan permintaan yang terus meningkat.

Di sisi lain, hampir 80 orang atau sekitar 39,5% menyatakan ketidaksetujuannya terhadap ilmu perpustakaan. Hal ini menunjukkan bahwa masih banyak orang yang memiliki pendapat negatif tentang bidang ini. Kurangnya antusiasme terhadap topik ini, kurangnya pengetahuan, dan pendapat tentang

kekurangan profesi perpustakaan adalah beberapa hal yang mungkin berkontribusi terhadap sikap yang kurang baik ini.

Dengan memetakan pendapat yang menghasilkan perasaan yang menguntungkan, kita dapat mengidentifikasi sejumlah elemen sikap yang mendominasi. Di antaranya adalah sikap bersyukur, di mana mereka yang disurvei percaya bahwa ilmu perpustakaan adalah disiplin ilmu yang bermanfaat. *“Ilmu perpustakaan sangat membantu dalam mengelola dan menemukan informasi yang dibutuhkan dengan cepat”* adalah contoh sudut pandang yang menggambarkan pola pikir ini. Optimisme adalah sikap lain yang lazim dalam kategori positif. Masa depan ilmu perpustakaan cukup menjanjikan, menurut banyak responden, terutama dengan penggunaan teknologi. *“Benerr pas aku liat kebanyakan yg dibutuhkan ilmu perpustakaan gak bmkng aja lembaga lain juga”* adalah contoh dari sudut pandang yang menggambarkan pola pikir ini. Selain itu, evaluasi yang baik juga menunjukkan mentalitas yang puas. Sistem ilmu perpustakaan saat ini dianggap oleh banyak orang sebagai sesuatu yang bermanfaat baik dalam konteks akademis maupun profesional. Misalnya, *“Ilmu Perpustakaan dia kata cuma nata buku hehhh kita belajar isi buku juga harus hafal satu satu keilmuannya juga”* adalah contoh sudut pandang yang menggambarkan pola pikir ini.

Sementara itu, sejumlah karakteristik sikap yang mendominasi juga terdapat pada kelompok pandangan negatif. Sikap skeptis adalah salah satunya, yang mengindikasikan kurangnya kepercayaan terhadap nilai ilmu perpustakaan di era digital. Sebagai contoh, *“Aku jurusan ilmu perpustakaan besok jadi buku ya”* adalah salah satu contoh dari pola pikir ini. Pendapat yang kurang baik juga termasuk sikap pesimis. Menurut beberapa responden, pengaruh jurusan ini di dunia kerja dan kesempatan kerja terbatas. Salah satu sudut pandang yang mencontohkan pola pikir ini adalah *“Lulusan ilmu perpustakaan pasti susah cari kerja. Lapangan pekerjaan terbatas, dan kebanyakan orang tidak tahu apa yang sebenarnya dilakukan pustakawan”* Hanya ada sedikit lowongan pekerjaan, dan kebanyakan orang tidak mengetahui tugas pustakawan yang sebenarnya. Pandangan negatif adalah sumber

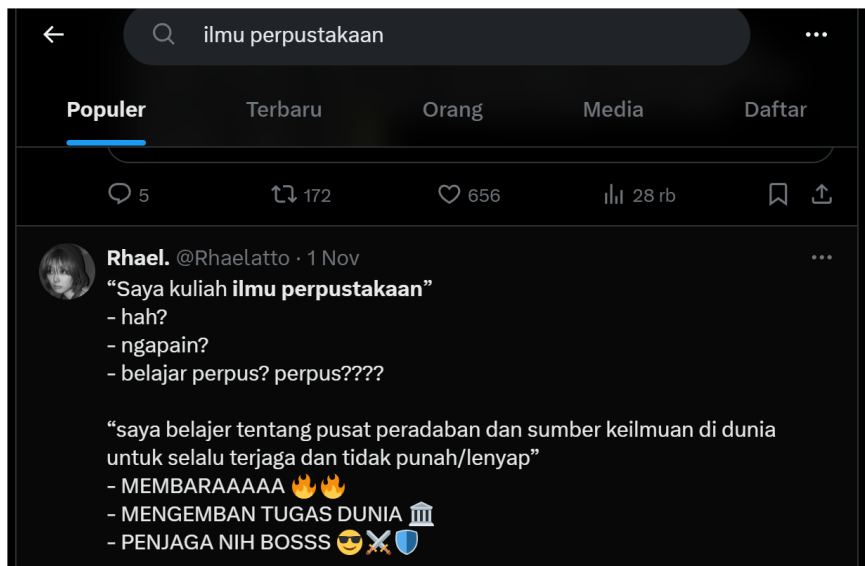
ketidakpuasan lainnya. Mungkin juga hanya sebuah lelucon sederhana, seolah-olah mereka selaras dengan perpustakaan dan buku. Pola pikir ini ditunjukkan dengan pernyataan, “*Anak ilmu perpustakaan kaya rak buku gt kah wkwk*”

Dengan tingkat akurasi 63,93%, model Naive Bayes digunakan dalam penelitian ini untuk menganalisis sentimen. Meskipun model ini memiliki tingkat ketergantungan yang tinggi dalam klasifikasi opini, akurasi prediksi masih dapat ditingkatkan. Secara keseluruhan, temuan penelitian ini menunjukkan bahwa, meskipun ada kritik dan kesulitan, publik masih memandang ilmu perpustakaan dengan baik. Memperoleh lebih banyak wawasan tentang opini publik dapat membantu dalam pembuatan rencana untuk membuat ilmu perpustakaan lebih menarik dan relevan di masa depan.

4.3 Interpretasi Hasil Penelitian

4.3.1 *Sample* (Populasi dan Sampel)

Tahap pertama dalam penelitian ini adalah pengambilan sampel, yang dilakukan melalui *tweet harvest* pada platform jejaring sosial X. Dalam analisis kuantitatif, ukuran sampel merupakan elemen penting yang menentukan validitas dan reliabilitas hasil penelitian. Untuk memastikan hasil analisis yang valid, sampel harus cukup besar untuk mewakili populasi yang diambil. Secara umum, ukuran sampel dipengaruhi oleh beberapa faktor seperti ukuran populasi, tingkat kepercayaan, *margin of error*, dan variasi dalam data. Di dunia digital, jumlah ulasan atau komentar pada platform jejaring sosial X bisa sangat bervariasi tergantung pada popularitas aplikasi dan intensitas diskusi tentang ilmu perpustakaan di platform tersebut. Oleh karena itu, peneliti perlu memperhitungkan volume data yang tersedia dan memastikan bahwa sampel yang diambil cukup representatif.



Gambar 4 2 Pencarian di X

Dalam konteks penelitian ini, karena kita berurusan dengan data dari aplikasi digital yang sering kali jumlahnya besar, metode sampling dapat mengandalkan teknik random sampling untuk menghindari bias. Sampel yang diambil adalah ulasan dan komentar dari pengguna yang menyebutkan kata kunci. Sampel ini diambil secara acak dari berbagai diskusi dan *review* untuk memastikan bahwa data yang diolah mencakup berbagai sudut pandang dan pengalaman pengguna yang beragam. Sampel ini mencakup berbagai kategori pengguna, mulai dari mahasiswa, dosen, hingga masyarakat umum.

Pemilihan sampel secara acak bertujuan untuk mendapatkan representasi yang lebih objektif mengenai sentimen masyarakat. Dalam analisis sentimen, ukuran sampel yang besar memberikan peluang lebih besar untuk menangkap variasi sentimen di berbagai kelompok pengguna. Selain itu, sampel ini diambil dari berbagai waktu, sehingga tren sentimen yang berubah-ubah juga dapat dianalisis secara lebih mendalam. Ini memungkinkan penelitian tidak hanya berfokus pada satu periode, tetapi juga pada perubahan persepsi masyarakat seiring waktu. Maka sampel yang diambil untuk penelitian ini adalah 203 sentimen.

4.3.2 Explore (objek Penelitian)

Menemukan ulasan atau komentar yang relevan adalah tahap kedua dalam proses penelitian ini. Kata kunci “ilmu perpustakaan” akan digunakan untuk menemukan bahan-bahan ini. Untuk menghasilkan gambaran yang tepat tentang suasana hati masyarakat, pemilihan kata kunci ini mencoba untuk menjamin bahwa data yang dikumpulkan memang relevan dengan bidang perpustakaan. Dengan menggunakan kata kunci yang relevan, peneliti dapat mempersempit data yang tidak berhubungan dengan topik penelitian, sehingga proses analisis menjadi lebih terarah dan efektif. Selain itu, pemilihan kata kunci yang tepat akan membantu dalam menghilangkan data yang tidak perlu dari proses analisis.



```
File Edit View Insert Runtime Tools Help Cannot save changes
+ Code + Text Copy to Drive
-- Scrolling... (1) (2) (3) (4) (5) (6)
Your tweets saved to: /content/tweets-data/ilmu_perpustakaan.csv
Total tweets saved: 140
-- Scrolling... (1) (2) (3)
Your tweets saved to: /content/tweets-data/ilmu_perpustakaan.csv
Total tweets saved: 160
-- Scrolling... (1)
Your tweets saved to: /content/tweets-data/ilmu_perpustakaan.csv
Total tweets saved: 180
-- Scrolling... (1) (2)
Your tweets saved to: /content/tweets-data/ilmu_perpustakaan.csv
Total tweets saved: 200
-- Scrolling... (1)
Your tweets saved to: /content/tweets-data/ilmu_perpustakaan.csv
Total tweets saved: 202
[v2.6.1] No more tweets found, please check your search criteria and csv file result
-- Scrolling... (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) No more tweets found
Timeout reached 1 times, making sure again...
-- Scrolling... (1) [v2.6.1] No more tweets found, please check your search criteria and csv file result
```

Gambar 4 3 Tweet Harvest

Namun demikian, sejak platform berganti nama dari Twitter menjadi X, penghalang yang signifikan untuk pengambilan data dari platform jejaring sosial X adalah akses API yang terbatas. Sejak modifikasi ini, X telah meningkatkan mekanisme keamanan datanya, menghasilkan API yang lebih terbatas untuk akses data pengguna. Bagi para peneliti yang membutuhkan banyak data untuk analisis sentimen, hal ini

menjadi masalah. Pengambilan data secara otomatis menjadi tantangan dan jumlah data yang dapat diperoleh menjadi terbatas dengan tidak adanya akses API yang tidak terbatas. Jawaban dari para peneliti adalah dengan menggunakan sebuah program yang disebut Tweet Harvester 2.6.1, yang memungkinkan pengambilan data dengan cara yang berbeda.

Meskipun akses API dibatasi, *Tweet Harvester* 2.6.1 adalah alat yang dapat membantu pengumpulan data dari platform jejaring sosial X. Peneliti dapat menggunakan alat ini untuk mengambil data ulasan atau komentar dari platform jejaring sosial X dengan menggunakan kata kunci yang telah ditentukan sebelumnya. Peneliti dapat dengan mudah mengumpulkan data dalam jumlah yang besar dengan alat ini, meskipun tidak memiliki akses langsung ke API. Alat ini memungkinkan peneliti untuk mendapatkan data yang diperlukan untuk penelitian, meskipun faktanya tidak secepat menggunakan API. Selain itu, data yang dikumpulkan dapat disesuaikan berdasarkan parameter tertentu seperti bahasa dan subjek.

Data yang diambil dalam penelitian ini berjumlah sekitar 2.000 ulasan atau komentar yang dipilih berdasarkan jumlah opini dan bukan berdasarkan rentang waktu. Dengan berfokus pada jumlah pendapat, peneliti dapat memastikan bahwa sampel cukup besar untuk mewakili pandangan publik tentang ilmu perpustakaan di X.

	A
1	@sbmptnfess Ilmu perpustakaan
2	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan psikologi/ilmu perpustakaan di kampus diponegoro (aamiin allahumma aamiin)
3	@_SHISOYA_ Halooo kak aku boleh nanya2 tentang ilmu perpustakaan ga yaaa
4	@utbkac Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan ilmu komunikasi atau ilmu perpustakaan di kampus Unpad!! Aamiin.
5	@sbmptnfess bismillah ilmu perpustakaan universitas Indonesia aamiin. 13 juni 2024 jadi maba ui
6	@sbmptnfess bismillahirrahmanirrahim semoga aku lolos ke ilmu perpustakaan brawijaya 2024 AAAMIIN
7	Bidang Kajian Strategis HMPS Ilmu Perpustakaan Kabinet Panca Guna Bersama Menuju Ilpus yang Berdaya dan Berguna! #KastratPancaGu https://t.co/dlYKbNLF9D
8	[REKAP GUDANG ISU MEI] Hai Libs! Berikut rekap gudang isu selama bulan Mei. Kamu dapat bergabung dan berpartisipasi dalam menghin isu melalui https://t.co/KOgxrs7WSY Narahubung: Nikita Catriona (085774488297) https://t.co/EVfSp1m84
9	@sbmptnfess Aamiin... Moga kita bisa ikut okk UI bareng ya nderrr... Doain aku juga ya semoga aku lulus SNBT 2024 & amp
10	#NAME?
11	@Denald Sebagai anak lulusan ilmu perpustakaan sangat iri meliat ini

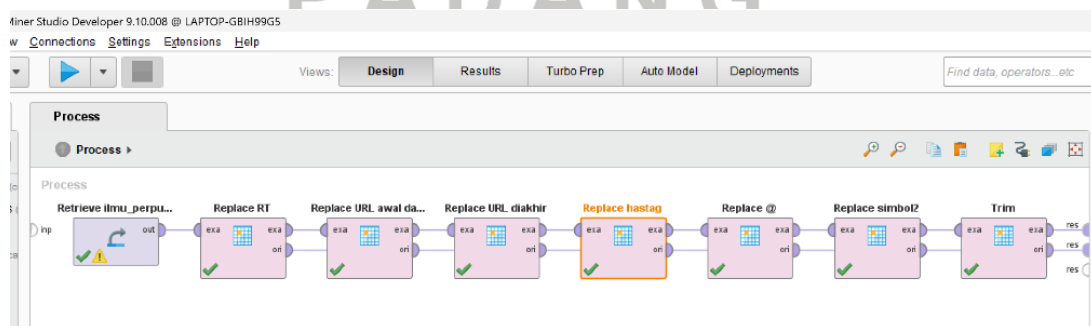
Gambar 4 4 Hasil data

Semua ulasan yang akan dianalisis menggunakan bahasa Indonesia, termasuk bahasa daerah, yang mencerminkan keberagaman linguistik di Indonesia. Keanekaragaman ini menunjukkan bahwa masyarakat menggunakan berbagai bentuk bahasa dalam mengekspresikan pendapat mereka, baik dalam konteks formal maupun informal. Oleh karena itu, penting bagi penelitian ini untuk mempertimbangkan aspek kebahasaan tersebut agar hasil analisis dapat mencerminkan realitas komunikasi yang sesungguhnya.

Penggunaan bahasa daerah dalam ulasan menambah kompleksitas dalam proses prapemrosesan data. Para peneliti harus memastikan bahwa setiap kata dan frasa yang digunakan dalam komentar dapat dipahami dengan benar agar tidak terjadi kesalahan dalam interpretasi sentimen. Selain itu, variasi dalam tata bahasa dan kosakata di berbagai daerah memerlukan pendekatan yang lebih cermat dalam membersihkan dan mengolah data sebelum dianalisis.

Dengan mempertimbangkan keberagaman bahasa yang ada, proses analisis sentimen dapat dilakukan dengan lebih akurat dan representatif. Pendekatan ini memungkinkan hasil penelitian lebih mencerminkan opini masyarakat secara luas tanpa mengabaikan nuansa budaya yang melekat dalam setiap bahasa. Dengan demikian, penelitian ini dapat memberikan wawasan yang lebih komprehensif terhadap sentimen publik dalam berbagai konteks linguistik di Indonesia.

4.3.3 *Modify (Analisis Data)*



Gambar 4 5 Langkah analisis data

Dalam proses analisis data, tahap *Modify* merupakan langkah krusial yang bertujuan untuk membersihkan dan menyusun data agar lebih siap digunakan dalam model analisis. Salah satu tahapan awal dalam proses ini adalah *cleaning* atau pembersihan data. *Cleaning* mencakup berbagai langkah seperti menghapus karakter khusus, tanda baca, angka yang tidak relevan, serta teks yang tidak memiliki makna signifikan. Selain itu, proses ini juga melibatkan normalisasi data, yaitu menyamakan format teks agar lebih konsisten, seperti mengubah semua huruf menjadi huruf kecil. Dengan melakukan *cleaning*, data yang digunakan menjadi lebih bersih dan mudah diolah dalam tahap selanjutnya.

Setelah data dibersihkan, langkah berikutnya adalah *tokenize* atau tokenisasi. Proses ini berfungsi untuk memecah teks menjadi unit-unit kecil yang disebut *token*, yang biasanya berupa kata atau frasa. Tokenisasi membantu dalam memahami struktur kalimat serta hubungan antara kata dalam suatu teks. Dalam konteks analisis sentimen atau pemrosesan bahasa alami (*Natural Language Processing*), tokenisasi sangat penting karena membantu model dalam mengenali pola dari kumpulan kata-kata dalam dataset. Dengan adanya tokenisasi, teks menjadi lebih terorganisir sehingga lebih mudah diproses oleh algoritma pembelajaran mesin.

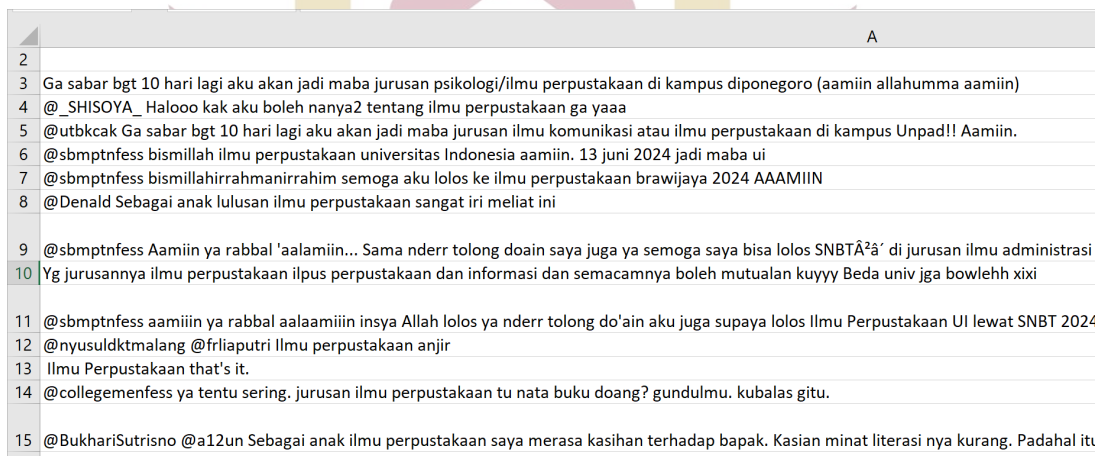
Tahapan berikutnya dalam *Modify* adalah *stop word removal*, yaitu proses menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis teks, seperti "dan", "atau", "yang", dan sebagainya. *Stop words* ini biasanya tidak memberikan informasi yang relevan dalam konteks analisis sentimen atau klasifikasi teks. Dengan menghilangkan *stop words*, jumlah fitur yang harus diproses oleh model menjadi lebih sedikit, sehingga dapat meningkatkan efisiensi dan akurasi dalam analisis data. Langkah ini juga membantu mengurangi kebisingan dalam data dan memastikan bahwa hanya kata-kata yang memiliki bobot informasi yang lebih tinggi yang digunakan dalam proses analisis.

Terakhir, proses *filtering* dilakukan untuk memilih dan mempertahankan hanya kata-kata yang benar-benar relevan dengan tujuan analisis. *Filtering* dapat mencakup berbagai teknik, seperti memilih kata-kata dengan frekuensi

tertentu atau menerapkan algoritma *stemming* dan *lemmatization* untuk mengubah kata-kata menjadi bentuk dasar mereka. Dengan *filtering*, data yang digunakan menjadi lebih fokus dan sesuai dengan kebutuhan analisis, sehingga dapat menghasilkan hasil yang lebih akurat dan dapat diandalkan. Secara keseluruhan, tahap *Modify* yang mencakup *cleaning*, *tokenize*, *stop word removal*, dan *filtering* merupakan langkah fundamental dalam analisis data yang bertujuan untuk meningkatkan kualitas dan efektivitas pemrosesan teks.

Berikut langkah-langkah untuk tahap *Modify* :

a. *Cleaning*



	A
2	
3	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan psikologi/ilmu perpustakaan di kampus diponegoro (aamiin allahumma aamiin)
4	@_SHISOYA_ Halooo kak aku boleh nanya2 tentang ilmu perpustakaan ga yaaa
5	@utbkck Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan ilmu komunikasi atau ilmu perpustakaan di kampus Unpad!! Aamiin.
6	@sbmptnfess bismillah ilmu perpustakaan universitas Indonesia aamiin. 13 juni 2024 jadi maba ui
7	@sbmptnfess bismillahirrahmanirrahim semoga aku lolos ke ilmu perpustakaan brawijaya 2024 AAAMIIN
8	@Denald Sebagai anak lulusan ilmu perpustakaan sangat iri meliat ini
9	@sbmptnfess Aamiin ya rabbal 'aalamiin... Sama nderr tolong doain saya juga ya semoga saya bisa lolos SNBT di jurusan ilmu administrasi
10	Yg jurusannya ilmu perpustakaan ilpus perpustakaan dan informasi dan semacamnya boleh mutualan kuyyy Beda univ jga bowlehh xixi
11	@sbmptnfess aamiin ya rabbal aalaamiin insya Allah lolos ya nderr tolong do'ain aku juga supaya lolos Ilmu Perpustakaan UI lewat SNBT 2024
12	@nyusuldktmalang @frliaputri Ilmu perpustakaan anjir
13	Ilmu Perpustakaan that's it.
14	@collegemenfess ya tentu sering. jurusan ilmu perpustakaan tu nata buku doang? gundulmu. kubalas gitu.
15	@BukhariSutrisno @a12un Sebagai anak ilmu perpustakaan saya merasa kasihan terhadap bapak. Kasian minat literasi nya kurang. Padahal itu

Gambar 4 6 Hasil setelah cleaning

Tahap awal dalam menyiapkan data teks mentah untuk analisis sentimen adalah pembersihan teks. Secara umum, data ulasan atau komentar platform jejaring sosial X terdiri dari item-item yang tidak penting seperti karakter khusus, angka, simbol, dan tanda baca yang berlebihan yang tidak memberikan nilai tambah untuk analisis.

Tujuan dari pembersihan ini adalah untuk menghilangkan komponen-komponen tersebut agar materi menjadi lebih jelas dan lebih fokus pada informasi penting. Karena model ini bekerja pada input yang lebih terstruktur, prosedur pembersihan tidak hanya meningkatkan efisiensi pemrosesan, tetapi juga berkontribusi pada peningkatan akurasi temuan analisis sentimen.

b. *Tokenize*

text	aaa	aaamiin	aalaamiin
sabar maba jurusan ilmu komunikasi ilmu perpustakaan kampus unpad aamiin	0	0	0
bismillah ilmu perpustakaan universitas indonesia aamiin juni maba	0	0	0
bismillahirrahmanirrahim semoga lolos ilmu perpustakaan brawijaya aaamiin	0	0.576	0
aamiin rabbal aalamiin nderr tolong doain semoga lolos snbtâ jurusan ilmu administrasi fiskaljurusan ilmu perpustakaan aamiin	0	0	0
jurusannya ilmu perpustakaan ilpus perpustakaan informasi semacamnya mutualan kuyy beda univ bowlehh xixi	0	0	0
aaamiin rabbal aalaamiin insya allah lolos nderr tolong lolos ilmu perpustakaan snbt aamiin allahumma aamiin	0	0	0.252
ilmu perpustakaan that	0	0	0
jurusan ilmu perpustakaan nata buku doang gundulmu kubalas gitu	0	0	0
anak ilmu perpustakaan kasihan kasian minat literasi korp surat beda kementrian	0	0	0
gatau kalo didunia luas jurusan ilmu perpustakaan	0	0	0

Gambar 4 7 Hasil kalimat tokenize

Setelah proses pembersihan data, langkah selanjutnya yang sangat penting adalah tokenisasi, yaitu proses memecah teks menjadi unit-unit yang lebih kecil yang disebut token. Token ini biasanya berupa kata atau frasa yang akan digunakan dalam analisis lebih lanjut. Tokenisasi membantu dalam menstrukturkan teks sehingga lebih mudah diproses oleh model, memungkinkan analisis yang lebih akurat dan sistematis.

Selain itu, tokenisasi berperan dalam memastikan bahwa kata-kata dalam teks dipahami dalam urutan yang tepat. Hal ini sangat penting untuk menghilangkan ambiguitas dalam bahasa, terutama dalam konteks analisis teks yang kompleks. Dengan adanya tokenisasi, model dapat menangkap makna sebenarnya dari sebuah kalimat dan menginterpretasikan hubungan antara kata-kata dengan lebih baik.

Studi yang melibatkan bahasa informal atau bahasa daerah juga sangat terbantu dengan adanya tokenisasi. Bahasa daerah sering kali memiliki pola atau struktur kalimat yang berbeda dari bahasa formal, sehingga tokenisasi membantu dalam menyesuaikan teks agar lebih mudah dianalisis oleh model. Dengan cara ini, tokenisasi tidak hanya meningkatkan pemahaman mesin terhadap teks tetapi juga memastikan bahwa analisis tetap relevan dengan berbagai bentuk bahasa yang digunakan.

c. *Stopword Removal*

...	@sbmptnfess Ilmu perpustakaan
1	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan psikologi ilmu perpustakaan di kampus diponegoro aamiin allahumma aamiin
2	Halooo kak aku boleh nanya2 tentang ilmu perpustakaan ga yaaa
3	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan ilmu komunikasi atau ilmu perpustakaan di kampus Unpad Aamiin
4	bismillah ilmu perpustakaan universitas Indonesia aamiin 13 juni 2024 jadi maba ui
5	bismillahirrahmanirrahim semoga aku lolos ke ilmu perpustakaan brawijaya 2024 AAAMIIN
6	Bidang Kajian Strategis HMPS Ilmu Perpustakaan Kabinet Panca Guna Bersama Menuju Ilpus yang Berdaya dan Berguna
7	[REKAP GUDANG ISU MEI] Hai Libs Berikut rekap gudang isu selama bulan Mei Kamu dapat bergabung dan berpartisipasi dalam menghimpun isuisu melalui Narahu
8	Aamiin Moga kita bisa ikut okk UI bareng ya nderrrr Doain aku juga ya semoga aku lulus SNBT 2024 amp
9	Sebagai anak lulusan ilmu perpustakaan sangat iri meliat ini
10	Ayo kita bersamasama berperan aktif dalam meningkatkan mutu dan kualitas pendidikan anak bangsa demi masa depan yang lebih gemilang

Gambar 4.8 Hasil penghilangan url

Pembersihan tidak hanya melibatkan penghapusan huruf dan simbol khusus, tetapi juga menghilangkan teks yang berisi tautan atau URL yang dapat dimasukkan ke dalam ulasan atau komentar. Menghapus tautan dari materi media sosial membantu menyederhanakan data karena tautan tersebut sering muncul dan tidak membantu dalam analisis sentimen. Dalam analisis sentimen, nama pengguna dan sebutan (seperti @nama_pengguna) juga sering kali tidak penting. Data teks menjadi lebih terkonsentrasi pada pemikiran dan pandangan yang disampaikan ketika mereka dihilangkan.

Emoji dan emotikon yang sering digunakan dalam tulisan biasa juga dihilangkan sebagai bagian dari proses pembersihan. Emoji, seperti yang ceria atau marah, dapat menyampaikan emosi, namun bagaimana mereka ditafsirkan bervariasi berdasarkan situasi. Para peneliti dapat berpikir untuk melestarikan dan mengklasifikasikan emoji jika penelitian ini membutuhkan pengenalan emoji sebagai komponen emosi. Penghapusan emoji, bagaimanapun, sering kali memfasilitasi pemrosesan dan analisis teks.

d. *Filtering*

R...	full_text
6	Yg jurusannya ilmu perpustakaan ilpus perpustakaan dan informasi dan semacamnya boleh mutualan kuyyy Beda univ jga bowlehh xixi
7	aamiin ya rabbal aalaamiin insya Allah lolos ya nderr tolong do'ain aku juga supaya lolos Ilmu Perpustakaan UI lewat SNBT 2024 ini aamiin allahumma aamiin
11	Sebagai anak ilmu perpustakaan saya merasa kasihan terhadap bapak Kasian minat literasi nya kurang Padahal itu di korp surat nya beda kementrian
15	Aamiin kamu pasti bisa nderr percaya sama Allah dan dirimu sendiri Doakan aku juga bisa lolos di pilihan 1 ilmu perpustakaan ui semangat sender dan kita semua
16	udah keliatan ini mah keterima nya bismillahirrahmanirrahim keterima snbt usu prodi ilmu perpustakaan
18	lolos di pil 1 universitas sumatera utara prodi ilmu perpustakaan
22	bismillahirrahmanirrahim tanggal 13 Juni 2024 dapat ucapan SELAMAT ANDA DINYATAKAN LOLOS dari SNPMB diterima di ke4 PTN yang sudah aku pilih UI Ilmu Perpustakaa
25	ilmu perpustakaan santai bgt dijamin bahagia healing everyday
28	Bapak saya dulu ilmu perpustakaan Karinya yg plg akhir di kompas gramedia grup jd redaksi
29	sebagai penyandang gelar S Ip alias ilmu perpustakaan dan hobi membaca novelwattpad aku bersyukur udh menemukan au au bagus di tahun ini terimakasih author2 keren
32	Pembelajaran ini ditujukan agar lulusan sarjana Ilmu Perpustakaan mampu mengoperasikan sistem informasi perpustakaan

Gambar 4 9 Data bersih

Setelah proses pembersihan data dilakukan, langkah selanjutnya adalah verifikasi manual oleh peneliti untuk memastikan bahwa prosedur pembersihan telah berhasil. Proses ini penting untuk menilai apakah data yang tersisa sudah cukup bersih dan dapat digunakan untuk analisis lebih lanjut. Tanpa verifikasi manual, ada kemungkinan bahwa data yang masih mengandung elemen tidak relevan dapat masuk ke dalam tahap analisis, yang dapat memengaruhi hasil penelitian secara keseluruhan.

Langkah verifikasi manual ini dilakukan untuk menghindari kesalahan yang mungkin terjadi selama pembersihan otomatis. Metode otomatis memang efisien dalam menghilangkan elemen yang tidak diperlukan, tetapi tidak selalu dapat membedakan antara informasi yang benar-benar penting dan yang harus dihapus. Oleh karena itu, keterlibatan manusia dalam mengecek hasil akhir sangat dibutuhkan guna memastikan kualitas data yang tinggi.

Selain itu, verifikasi manual juga berperan dalam meningkatkan efektivitas proses pembersihan data. Dalam beberapa kasus, algoritma pembersihan otomatis dapat secara tidak sengaja menghapus data yang masih memiliki nilai penting untuk penelitian. Dengan adanya pengecekan ulang secara manual, peneliti dapat memastikan bahwa tidak ada informasi berharga yang terbuang, sehingga analisis yang dilakukan tetap akurat dan relevan.

Pemeriksaan manual ini juga memungkinkan peneliti untuk menemukan data yang masih mengandung bagian yang tidak diinginkan atau noise. Noise dalam data dapat berupa kata-kata tidak relevan, simbol, atau format teks yang kurang sesuai untuk dianalisis. Dengan menghilangkan elemen-elemen ini secara manual, peneliti dapat memastikan bahwa data yang akan dianalisis benar-benar bersih dan dapat memberikan hasil yang lebih akurat.

Selain memastikan kebersihan data, verifikasi manual juga berfungsi untuk menjamin konsistensi data yang digunakan dalam analisis sentimen. Jika data yang digunakan tidak konsisten, hasil analisis dapat menjadi bias atau kurang dapat diandalkan. Oleh karena itu, langkah ini menjadi krusial dalam menghasilkan temuan yang valid dan dapat dipertanggungjawabkan dalam penelitian.

Dengan melakukan verifikasi manual, peneliti dapat meningkatkan kualitas data yang akan digunakan dalam analisis sentimen. Data yang bersih, relevan, dan konsisten akan menghasilkan model analisis yang lebih akurat serta memberikan wawasan yang lebih mendalam mengenai sentimen masyarakat terhadap ilmu perpustakaan. Oleh sebab itu, meskipun metode otomatis sangat membantu dalam proses pembersihan data, keterlibatan manusia tetap menjadi faktor penting dalam memastikan keakuratan dan keandalan hasil penelitian.

4.3.4 Model (Pelabelan Data)

Data dilabel menjadi 2 jenis yaitu positif dan negatif. Berikut adalah ciri-ciri ketentuan sentiment positif dan negatif.

a. Positif

Sentimen positif mengacu pada opini, komentar, atau ekspresi yang menunjukkan pandangan baik terhadap suatu topik. Beberapa ciri-ciri sentimen positif antara lain:

- a. Penggunaan kata-kata dengan konotasi positif
- b. Adanya ekspresi kebahagiaan atau kepuasan

- c. Kalimat yang mencerminkan harapan baik atau dukungan
- d. Pujian terhadap suatu produk, layanan, atau individu
- e. Minimnya kata negatif atau kritik

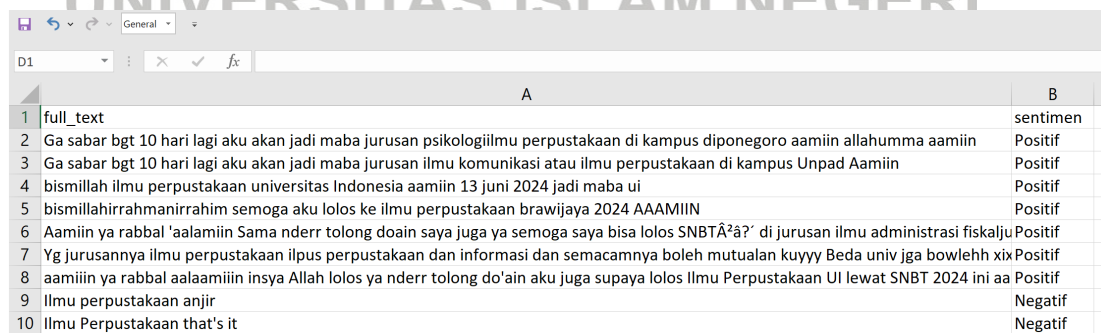
b. Negatif

Sentimen negatif mencerminkan opini atau ekspresi ketidakpuasan terhadap suatu hal. Berikut adalah ciri-cirinya:

- a. Penggunaan kata-kata bernada negative
- b. Ekspresi kekecewaan, marah, atau frustrasi
- c. Kalimat yang mencerminkan ketidakpercayaan atau ketidakpuasan
- d. Kritik terhadap produk, layanan, atau individu
- e. Adanya kata-kata penolakan atau kekecewaan berulang

Dalam menentukan label positif dan negatif, terbagi menjadi 2 cara yaitu manual dan otomatis. Berikut adalah penjelasannya.

1. Manual



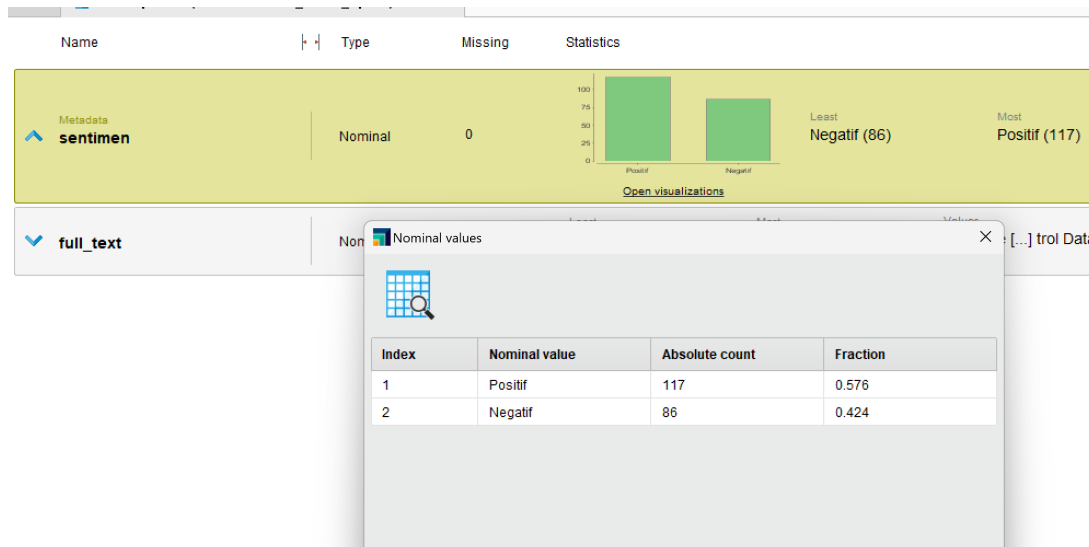
	A	B
1	full_text	sentimen
2	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan psikologi ilmu perpustakaan di kampus diponegoro aamiin allahumma aamiin	Positif
3	Ga sabar bgt 10 hari lagi aku akan jadi maba jurusan ilmu komunikasi atau ilmu perpustakaan di kampus Unpad Aamiin	Positif
4	bismillah ilmu perpustakaan universitas Indonesia aamiin 13 juni 2024 jadi maba ui	Positif
5	bismillahirrahmanirrahim semoga aku lolos ke ilmu perpustakaan brawijaya 2024 AAAMIIN	Positif
6	Aamiin ya rabbal 'aalamiin Sama nderr tolong doain saya juga ya semoga saya bisa lolos SNBT di jurusan ilmu administrasi fiskal	Positif
7	Yg jurusan ilmu perpustakaan ilpus perpustakaan dan informasi dan semacamnya boleh mutualan kuyyy Beda univ jga bowlehh xix	Positif
8	aamiin ya rabbal aalaamiin insya Allah lolos ya nderr tolong do'ain aku juga supaya lolos Ilmu Perpustakaan UI lewat SNBT 2024 ini aa	Positif
9	Ilmu perpustakaan anjir	Negatif
10	Ilmu Perpustakaan that's it	Negatif

Gambar 4 10 Pelabelan manual

Pelabelan manual memiliki manfaat utama dalam menghasilkan akurasi yang tinggi karena manusia mampu mengenali makna halus dan emosi dalam teks, seperti ironi atau sarkasme, yang sering kali sulit dipahami oleh komputer. Dengan

kemampuan ini, pelabelan manual dapat memberikan hasil interpretasi yang lebih tepat dibandingkan dengan metode otomatis. Selain itu, manusia dapat mempertimbangkan konteks secara lebih mendalam, memastikan bahwa setiap teks dikategorikan dengan benar sesuai dengan sentimen yang terkandung di dalamnya. Keakuratan ini menjadikan pelabelan manual sebagai pendekatan yang andal dalam analisis sentimen, meskipun tetap memerlukan keterampilan linguistik yang baik. Dengan adanya pemahaman kontekstual yang lebih dalam, metode ini dapat mengurangi kesalahan klasifikasi yang sering muncul dalam analisis berbasis mesin. Oleh karena itu, pelabelan manual sering digunakan dalam penelitian yang membutuhkan ketelitian tinggi dalam mengkategorikan sentimen teks.

Meskipun lebih akurat, pelabelan manual membutuhkan konsistensi yang tinggi, perhatian terhadap detail, dan pemahaman yang mendalam tentang bahasa serta konteksnya. Setiap teks harus dianalisis dengan cermat agar hasil yang diperoleh tidak bias dan tetap objektif sesuai dengan tujuan penelitian. Proses ini memerlukan waktu yang lebih lama dibandingkan dengan metode otomatis, terutama jika jumlah data yang dianalisis sangat besar. Namun, kelemahan tersebut dapat diimbangi dengan hasil yang lebih berkualitas, yang pada akhirnya meningkatkan performa model analisis sentimen. Data yang telah diberi label dengan baik akan membantu model dalam mengenali pola sentimen dengan lebih baik saat diterapkan pada data baru. Oleh karena itu, meskipun lebih memakan waktu, pelabelan manual tetap menjadi metode yang berharga dalam pengembangan model analisis sentimen yang presisi dan akurat.

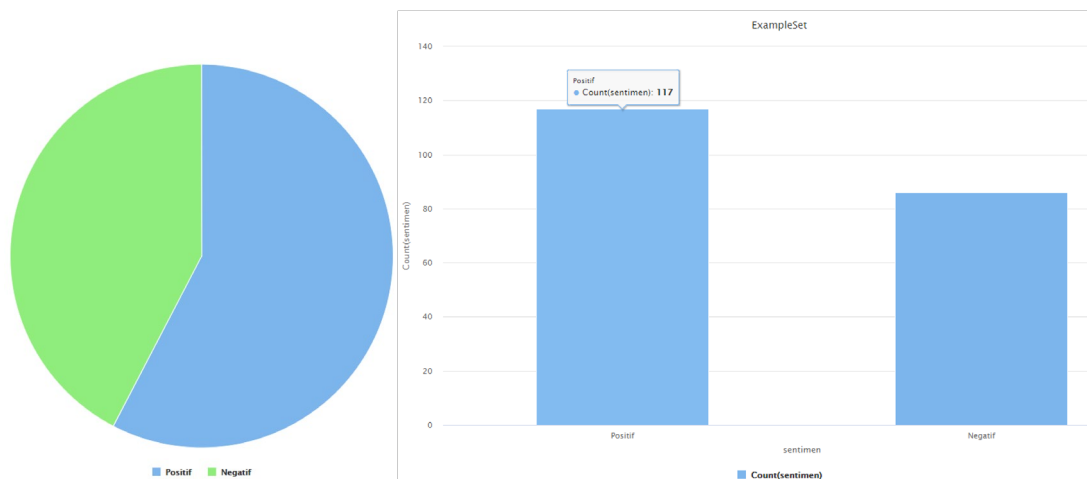


Gambar 4 11 Hasil pelabelan manual

Berdasarkan hasil pelabelan manual, dari seluruh data yang dianalisis, sebanyak 117 data diklasifikasikan sebagai memiliki sikap positif. Hal ini menunjukkan bahwa mayoritas ulasan atau komentar yang dikumpulkan mengandung pandangan yang mendukung, optimis, atau menunjukkan apresiasi terhadap topik yang dibahas.

Namun, di sisi lain, terdapat 86 data yang dikategorikan sebagai memiliki emosi negatif. Data ini mencerminkan adanya ketidakpuasan, kritik, atau opini yang kurang mendukung terhadap subjek yang diteliti. Keberadaan sentimen negatif ini menunjukkan bahwa masih ada aspek yang perlu diperhatikan atau diperbaiki agar persepsi masyarakat lebih baik di masa depan.

Perbandingan antara sentimen positif dan negatif ini memberikan gambaran mengenai persepsi umum yang berkembang dalam data yang dianalisis. Dengan memahami proporsi sentimen yang ada, penelitian ini dapat memberikan wawasan yang lebih mendalam mengenai faktor-faktor yang mempengaruhi opini publik serta menjadi dasar untuk rekomendasi perbaikan atau pengembangan lebih lanjut.

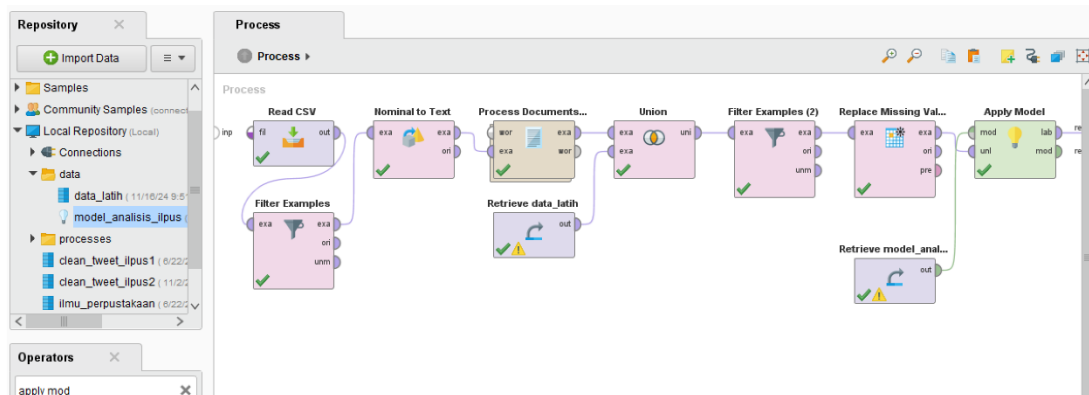


Gambar 4 12 Diagram

Meskipun jumlah sentimen negatif lebih rendah dibandingkan sentimen positif, analisis terhadap data ini tetap penting untuk memahami faktor-faktor yang mempengaruhi persepsi kurang baik terhadap ilmu perpustakaan. Dengan mengidentifikasi aspek-aspek yang menjadi keluhan atau kritik, langkah-langkah perbaikan dapat dirancang untuk meningkatkan citra dan relevansi ilmu perpustakaan di masyarakat.

Melalui analisis yang mendalam, dapat ditemukan pola atau tren dalam sentimen negatif yang dapat dijadikan dasar untuk inovasi dan pengembangan kebijakan di bidang perpustakaan. Dengan demikian, pemahaman yang lebih baik terhadap kritik yang ada dapat membantu menciptakan strategi yang lebih efektif dalam meningkatkan apresiasi dan pemahaman masyarakat terhadap ilmu perpustakaan.

2. Otomatis



Gambar 4 13 Proses pelabelan otomatis

Pelabelan data teks berdasarkan sentimen dengan menggunakan model atau algoritme pembelajaran mesin dikenal sebagai pelabelan otomatis. Metode ini memungkinkan sistem untuk mengkategorikan data secara lebih cepat tanpa memerlukan intervensi manusia secara langsung. Dengan pendekatan ini, proses analisis sentimen dapat dilakukan secara efisien, terutama ketika berhadapan dengan jumlah data yang sangat besar.

Dalam konteks analisis sentimen untuk ilmu perpustakaan di platform jejaring sosial X, pelabelan otomatis menjadi solusi yang lebih efektif dibandingkan dengan pelabelan manual. Penggunaan metode ini dapat menghemat waktu dan tenaga, terutama dalam penelitian dengan skala data yang luas. Selain itu, dengan algoritme yang telah dilatih sebelumnya, proses klasifikasi dapat dilakukan secara konsisten tanpa terpengaruh oleh subjektivitas manusia.

Untuk menentukan sentimen suatu data secara otomatis, algoritme seperti Naïve Bayes dapat diterapkan pada dataset yang telah diberi label secara manual sebelumnya. Dengan memanfaatkan model yang telah terlatih, sistem dapat memprediksi apakah suatu teks memiliki sentimen positif atau negatif berdasarkan pola yang telah dipelajari. Pendekatan ini memungkinkan analisis sentimen dilakukan secara lebih sistematis dan akurat, meskipun tetap perlu dilakukan evaluasi berkala untuk memastikan kinerjanya tetap optimal.

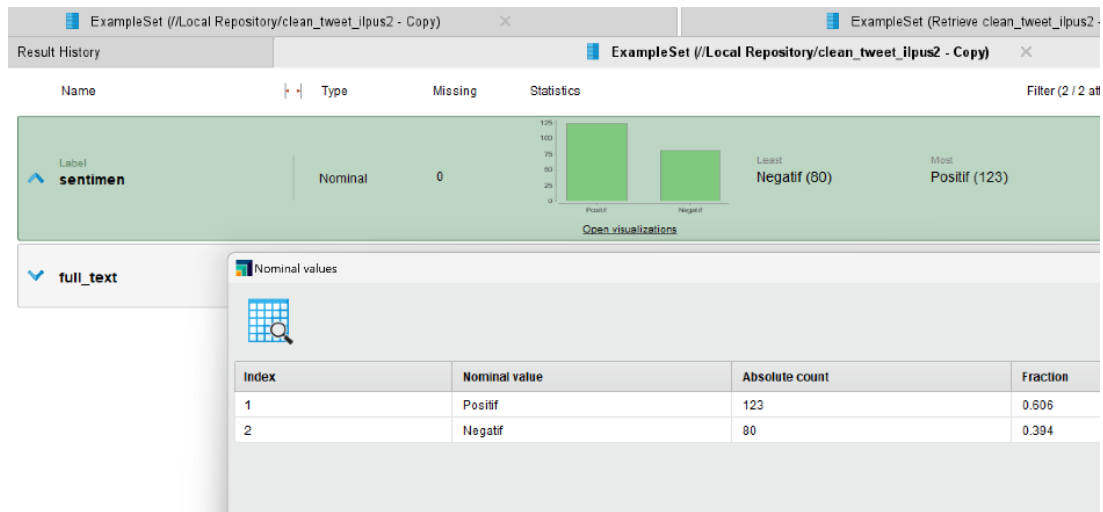
R...	..	prediction(...	..	..	text
1	.	Positif	.	.	sabar maba jurusan ilmu komunikasi ilmu perpustakaan kampus unpad aamiin
2	.	Positif	.	.	bismillah ilmu perpustakaan universitas indonesia aamiin juni maba
3	.	Positif	.	.	bismillahirrahmanirrahim semoga lolos ilmu perpustakaan brawijaya aaamiin
4	.	Positif	.	.	aamiin rabbal aalamiin nderr tolong doain semoga lolos snbtá jurusan ilmu administrasi fiskal jurusan ilmu perpustakaan aamiin
5	.	Positif	.	.	jurusannya ilmu perpustakaan ilpus perpustakaan informasi semacamnya mutualan kuyyy beda univ bowlehh xixi
6	.	Positif	.	.	aamiin rabbal aalaamiin insya allah lolos nderr tolong lolos ilmu perpustakaan snbt aamiin allahumma aamiin
7	.	Negatif	.	.	ilmu perpustakaan that

Gambar 4 14 Pelabelan otomatis

Gambar di atas menampilkan hasil analisis sentimen terhadap berbagai teks yang berkaitan dengan ilmu perpustakaan. Kolom pertama menunjukkan nomor urut data, kolom kedua berisi hasil prediksi sentimen yang dikategorikan sebagai "Positif" atau "Negatif," sedangkan kolom terakhir berisi teks asli dari komentar atau ulasan yang dianalisis. Mayoritas data yang ditampilkan dalam gambar ini memiliki sentimen positif, yang mengindikasikan bahwa banyak pengguna memiliki pandangan yang baik terhadap ilmu perpustakaan, baik dalam konteks akademik maupun sosial.

Selain itu, terdapat pula komentar yang mengaitkan ilmu perpustakaan dengan berbagai disiplin ilmu lainnya, seperti administrasi fiskal dan informasi. Hal ini mencerminkan bahwa ilmu perpustakaan memiliki keterkaitan dengan berbagai bidang dan dapat menjadi pilihan studi yang menjanjikan. Beberapa komentar juga menunjukkan adanya pemahaman bahwa ilmu perpustakaan memiliki cakupan yang luas, baik dari aspek akademis maupun profesional.

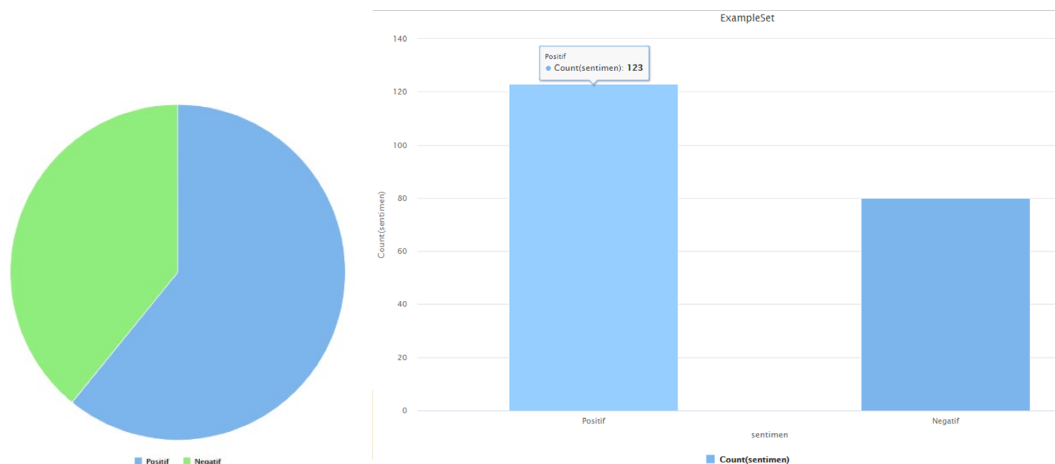
Namun, terdapat beberapa data yang diklasifikasikan sebagai sentimen negatif. Sentimen negatif ini mencerminkan adanya pandangan atau pengalaman kurang baik terhadap ilmu perpustakaan, baik dari segi persepsi umum maupun pengalaman pribadi pengguna. Keberadaan sentimen negatif ini menjadi perhatian bagi akademisi dan praktisi di bidang ilmu perpustakaan untuk terus meningkatkan pemahaman masyarakat serta memperbaiki aspek-aspek yang mungkin masih kurang optimal dalam bidang ini.



Gambar 4 15 Hasil sentimen pelabelan otomatis

Hasil analisis sentimen dengan pelabelan otomatis menunjukkan bahwa dari keseluruhan data yang diproses, sebanyak 123 data dikategorikan sebagai sentimen positif. Hal ini mengindikasikan bahwa mayoritas pengguna memiliki pandangan yang baik terhadap ilmu perpustakaan, baik dari segi manfaat, relevansi, maupun pengalaman mereka dalam mengakses layanan perpustakaan. Tingginya jumlah sentimen positif ini menunjukkan bahwa ilmu perpustakaan masih memiliki peran penting dan mendapat apresiasi dari masyarakat.

Di sisi lain, terdapat 80 data yang termasuk dalam kategori sentimen negatif. Sentimen ini menunjukkan bahwa masih ada kelompok pengguna yang memiliki pandangan kurang baik terhadap ilmu perpustakaan, yang mungkin disebabkan oleh keterbatasan akses, kurangnya pemahaman, atau pengalaman yang kurang memuaskan. Meskipun jumlahnya lebih sedikit dibandingkan sentimen positif, adanya opini negatif ini tetap menjadi aspek penting yang perlu diperhatikan dalam upaya meningkatkan kualitas dan persepsi terhadap ilmu perpustakaan.



Gambar 4.16 Diagram

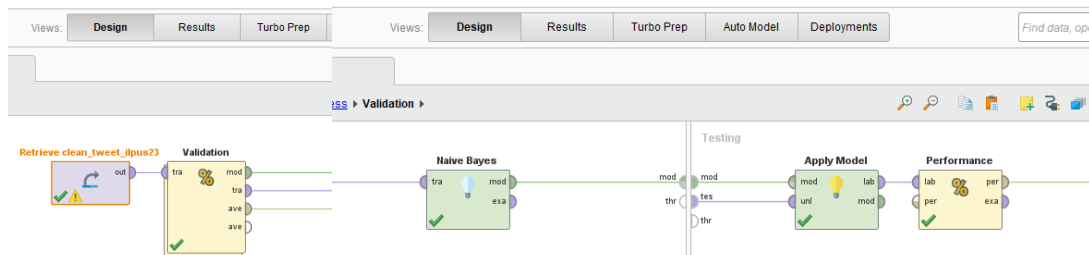
Pelabelan otomatis dalam analisis sentimen memberikan kemudahan bagi peneliti dalam mengklasifikasikan data dalam jumlah besar secara efisien. Dengan menggunakan algoritme pembelajaran mesin, proses ini dapat dilakukan dengan lebih cepat dibandingkan dengan pelabelan manual. Keunggulan utama dari metode ini adalah kemampuannya untuk menangani dataset yang luas tanpa memerlukan terlalu banyak campur tangan manusia, sehingga mempercepat proses analisis dan pengambilan keputusan.

Meskipun efisien, hasil yang diperoleh dari pelabelan otomatis masih perlu divalidasi dengan pendekatan manual guna memastikan akurasi yang lebih tinggi. Setiap algoritme memiliki batasan dalam memahami konteks dan nuansa bahasa yang kompleks, terutama dalam analisis sentimen yang melibatkan beragam ekspresi dan gaya bahasa. Oleh karena itu, diperlukan evaluasi lebih lanjut untuk menyesuaikan hasil klasifikasi dengan makna yang sebenarnya dari setiap ulasan atau komentar yang dianalisis.

Perbedaan interpretasi dalam penggunaan istilah, struktur kalimat, dan konteks dapat menyebabkan model menghasilkan klasifikasi yang kurang tepat. Beberapa kata atau frasa mungkin memiliki arti yang berbeda tergantung pada situasi penggunaannya, yang dapat mempengaruhi hasil analisis sentimen. Oleh karena itu,

kombinasi antara pendekatan otomatis dan manual menjadi solusi yang ideal untuk meningkatkan kualitas serta keakuratan hasil analisis sentimen dalam penelitian ini.

4.3.5 Assess (Uji Akurasi Data)



Gambar 4 17 Tahap akurasi

Data yang telah melalui prosedur pembersihan dan pelabelan dimasukkan ke dalam model analisis sentimen untuk memulai tahap uji akurasi penelitian ini. Algoritma Naive Bayes, salah satu teknik klasifikasi berbasis probabilitas, kemudian digunakan untuk melakukan prosedur validasi. Berdasarkan atribut yang ada, algoritma ini menentukan probabilitas bahwa data akan masuk ke dalam kategori tertentu. Dalam hal analisis sentimen, Naive Bayes akan menggunakan kata-kata yang ada di dalam teks untuk memutuskan apakah teks tersebut memiliki kemungkinan besar untuk memiliki sikap positif atau negatif.

r 9.10.008 @ LAPTOP-GBIH99G5
 Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments Find data, operators...etc

ExampleSet (//Local Repository/clean_tweet_ilpus23) ExampleSet (//Local Repository/clean_tweet_ilpus23)

PerformanceVector (Performance) ExampleSet (Retrieve clean_tweet_ilpus23)

Table View Plot View

accuracy: 63.93%

	true Positif	true Negatif	class precision
pred. Positif	27	14	65.85%
pred. Negatif	8	12	60.00%
class recall	77.14%	46.15%	

Gambar 4 18 Hasil akurasi

Gambar yang ditampilkan menunjukkan hasil evaluasi performa model klasifikasi sentimen menggunakan algoritma Naive Bayes dalam perangkat lunak *RapidMiner*. Dari hasil pengujian, model memiliki tingkat akurasi sebesar 63,93%, yang berarti bahwa model dapat mengklasifikasikan data dengan benar dalam 63,93% dari total sampel yang diuji. Akurasi ini mencerminkan sejauh mana model mampu mengenali pola sentimen dalam data yang diberikan.

- a. *True Positif* (TP) = 27
- b. *True Negatif* (TN) = 12
- c. *False Positif* (FP) = 14
- d. *False Negatif* (FN) = 8

Berdasarkan confusion matrix yang ditampilkan, model berhasil mengklasifikasikan 27 data dengan benar sebagai sentimen positif (*true positive*) dan 12 data dengan benar sebagai sentimen negatif (*true negative*). Namun, terdapat 14 data yang seharusnya negatif tetapi diklasifikasikan sebagai positif (*false positive*), serta 8 data yang seharusnya positif tetapi diklasifikasikan sebagai negatif (*false negative*). Kesalahan klasifikasi ini menunjukkan bahwa model masih mengalami beberapa keterbatasan dalam membedakan sentimen tertentu.

4.4 Implikasi Penelitian

Beberapa dampak dari hasil penelitian ini, sebagai berikut.

4.4.1 Pemahaman Persepsi Masyarakat terhadap Ilmu Perpustakaan

Hasil penelitian ini memberikan wawasan mengenai bagaimana masyarakat memandang ilmu perpustakaan, di mana mayoritas sentimen yang dianalisis cenderung positif. Hal ini menunjukkan bahwa ilmu perpustakaan masih memiliki reputasi yang baik dan dianggap memiliki peran penting dalam dunia pendidikan serta literasi informasi. Masyarakat yang memberikan sentimen positif umumnya melihat ilmu perpustakaan sebagai bidang yang relevan dan mendukung perkembangan intelektual.

Namun, meskipun mayoritas sentimen bersifat positif, masih terdapat kelompok masyarakat yang memiliki persepsi negatif terhadap ilmu perpustakaan. Beberapa di antaranya merasa bahwa bidang ini kurang menarik atau memiliki keterbatasan dalam hal prospek kerja. Sentimen negatif ini menunjukkan bahwa masih ada aspek dalam ilmu perpustakaan yang perlu disosialisasikan dengan lebih baik agar pemahaman masyarakat terhadap bidang ini lebih luas dan akurat.

Dengan adanya pemahaman lebih mendalam mengenai persepsi masyarakat, institusi akademik dan perpustakaan dapat mengambil langkah-langkah strategis untuk meningkatkan daya tarik ilmu perpustakaan. Mempromosikan peran pustakawan dalam era digital dan menyoroti pentingnya literasi informasi dapat menjadi cara efektif untuk memperkuat citra positif ilmu perpustakaan di mata masyarakat.

4.4.2 Strategi Peningkatan Citra Ilmu Perpustakaan

Analisis sentimen yang dilakukan dalam penelitian ini memberikan data berharga bagi institusi pendidikan dan perpustakaan dalam menentukan strategi peningkatan citra ilmu perpustakaan. Dengan memahami pola sentimen yang berkembang di masyarakat, institusi dapat menyusun pendekatan yang lebih efektif dalam mempromosikan ilmu perpustakaan sebagai bidang yang modern, dinamis, dan relevan dengan perkembangan zaman.

Jika masih ditemukan sentimen negatif, maka langkah-langkah perbaikan perlu segera diterapkan untuk mengubah persepsi masyarakat. Salah satu cara yang dapat dilakukan adalah dengan meningkatkan inovasi dalam layanan perpustakaan, seperti digitalisasi koleksi, penyediaan akses informasi berbasis teknologi, serta kolaborasi dengan berbagai disiplin ilmu lainnya. Langkah-langkah ini dapat membantu memperluas cakupan manfaat ilmu perpustakaan dan meningkatkan daya tariknya.

Selain itu, penting bagi institusi pendidikan dan pustakawan untuk lebih aktif dalam menyampaikan informasi tentang peran strategis ilmu perpustakaan di era digital. Kampanye edukasi yang menarik dan berbasis media sosial dapat membantu membentuk pemahaman yang lebih baik di kalangan masyarakat, sehingga citra ilmu perpustakaan semakin positif dan dihargai secara luas.

4.4.3 Meningkatkan Peran Media Sosial dalam Promosi Ilmu Perpustakaan

Hasil analisis sentimen menunjukkan bahwa media sosial memiliki pengaruh besar dalam membentuk opini publik terhadap ilmu perpustakaan. Di era digital saat ini, masyarakat cenderung mencari informasi melalui platform media sosial, sehingga institusi pendidikan dan perpustakaan harus lebih aktif dalam memanfaatkan kanal ini untuk membangun komunikasi yang positif dan efektif dengan masyarakat.

Melalui media sosial, institusi dapat membagikan informasi mengenai manfaat ilmu perpustakaan, perkembangan teknologi dalam bidang ini, serta berbagai inovasi yang telah diterapkan. Selain itu, interaksi langsung dengan pengguna melalui media sosial juga memungkinkan adanya diskusi dan umpan balik yang dapat membantu meningkatkan layanan serta mengubah persepsi negatif yang mungkin masih ada.

Untuk memaksimalkan potensi media sosial, perpustakaan dan institusi pendidikan perlu menciptakan konten yang menarik, edukatif, dan mudah dipahami. Video pendek, infografis, serta diskusi interaktif dapat menjadi cara efektif untuk meningkatkan kesadaran masyarakat tentang pentingnya ilmu perpustakaan. Dengan strategi ini, citra ilmu perpustakaan dapat semakin positif dan lebih banyak orang yang tertarik untuk mempelajarinya lebih lanjut.

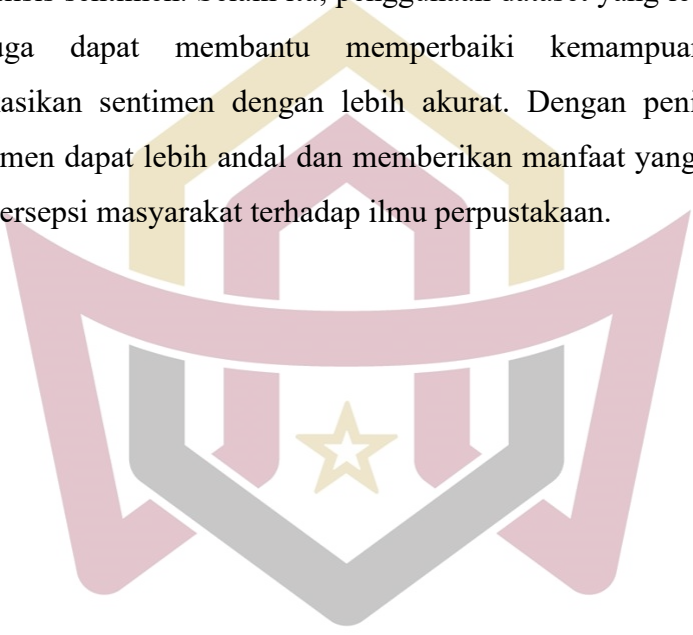
4.4.4 Evaluasi Model Analisis Sentimen

Penggunaan algoritma Naïve Bayes dalam penelitian ini menghasilkan akurasi sebesar 63,93%, yang menunjukkan bahwa model ini cukup baik dalam mengklasifikasikan sentimen masyarakat terhadap ilmu perpustakaan. Dengan tingkat akurasi tersebut, model dapat menjadi alat bantu yang berguna dalam memahami pola opini publik dan memberikan wawasan yang lebih mendalam mengenai persepsi masyarakat.

Namun, hasil ini juga mengungkapkan bahwa model masih mengalami keterbatasan dalam mengidentifikasi sentimen negatif dengan benar. Hal ini dapat disebabkan oleh berbagai faktor, seperti kompleksitas bahasa yang digunakan dalam

ulasan, penggunaan bahasa informal atau slang, serta keterbatasan dalam data latih yang digunakan. Oleh karena itu, diperlukan evaluasi lebih lanjut untuk meningkatkan keakuratan model dalam mengenali sentimen negatif dengan lebih baik.

Untuk penelitian selanjutnya, disarankan untuk menggabungkan metode lain seperti deep learning atau Support Vector Machine (SVM) guna meningkatkan performa analisis sentimen. Selain itu, penggunaan dataset yang lebih besar dan lebih beragam juga dapat membantu memperbaiki kemampuan model dalam mengklasifikasikan sentimen dengan lebih akurat. Dengan peningkatan ini, hasil analisis sentimen dapat lebih andal dan memberikan manfaat yang lebih besar dalam memahami persepsi masyarakat terhadap ilmu perpustakaan.



UNIVERSITAS ISLAM NEGERI
IMAM BONJOL
PADANG