JATIKOM: Jurnal Aplikasi dan Teori Ilmu Komputer
ISSN: 2615-577X (Online)

Analisis Sentimen Tingkat Kepopuleran *E-commerce* Pada Pengguna Twitter Menggunakan Algoritma *Naive Bayes*

Acah Rahmawati[#1], Uus Muhammad Husni Tamyiz [#2], Imam Maruf Nugroho[#3]

[#]*Program Studi Teknik Informatika Sekolah Tinggi Teknologi Wastukancana Purwakarta*
Jl. Cikopak No.53, Sadang Purwakarta, Jawa Barat

[1]acah.rahmawati@stt-wastukancana.ac.id,[2]tamyiz_husni@wastukancana.ac.id,
[3]imam.ma@wastukancana.ac.id

Abstract*—The rapid growth of e-commerce in Indonesia has led to an increasing number of digital transactions. With numerous platforms available, consumers face challenges in choosing a reliable e-commerce service. Sentiment analysis based on social media, particularly Twitter, can be used to assess public perception of various e-commerce platforms. Twitter serves as a primary medium for users to express opinions openly, allowing interaction and the rapid dissemination of information. Through sentiment analysis, user opinions can be classified as positive or negative, providing insights into customer satisfaction with specific e-commerce services. This study employs the Naïve Bayes Classifier Algorithm to categorize user opinions on multiple e-commerce platforms in Indonesia. The model was tested using the Orange software to determine accuracy, precision, and recall in sentiment classification. The results indicate an overall accuracy of 78.4%, with a precision of 87% and recall of 77.2%. The accuracy scores for individual e-commerce platforms are as follows: Shopee (74.1%), Tokopedia (78%), Blibli (77.8%), Lazada (71.4%), and Bukalapak (76.5%). Based on the analysis, Tokopedia and Blibli emerged as the most popular e-commerce platforms according to user sentiment on Twitter. The findings suggest that sentiment analysis using the Naïve Bayes algorithm can serve as an effective evaluation tool for understanding public perception of e-commerce in Indonesia. This approach can assist companies in refining their marketing strategies and enhancing customer experience management. To improve sentiment classification accuracy, integrating additional machine learning methods such as Support Vector Machine (SVM) or Deep Learning could further optimize sentiment analysis outcomes.

Key Words*— *E-Commerce, Sentiment Analysis, Twitter, Naïve Bayes, Popularity

---

Abstrak*— Perkembangan e-commerce di Indonesia semakin pesat seiring dengan meningkatnya transaksi digital. Banyaknya platform yang tersedia menyebabkan masyarakat kesulitan dalam memilih e-commerce yang terpercaya. Analisis sentimen berbasis media sosial, khususnya Twitter, dapat digunakan untuk memahami opini publik terhadap berbagai platform e-commerce. Penelitian ini menerapkan Algoritma Naïve Bayes Classifier dalam proses klasifikasi sentimen pengguna terhadap beberapa e-commerce utama di Indonesia. Model yang dikembangkan diuji menggunakan perangkat lunak Orange untuk memperoleh nilai akurasi, presisi, dan recall dalam analisis sentimen. Hasil pengujian menunjukkan bahwa algoritma Naïve Bayes memiliki tingkat akurasi keseluruhan sebesar 78.4%, dengan presisi 87% dan recall 77.2%. Adapun akurasi masing-masing e-commerce adalah: Shopee (74.1%), Tokopedia (78%), Blibli (77.8%), Lazada (71.4%), dan Bukalapak (76.5%). Berdasarkan hasil analisis, Tokopedia dan Blibli memiliki tingkat kepopuleran tertinggi berdasarkan opini pengguna di Twitter. Kesimpulan dari penelitian ini menunjukkan bahwa analisis sentimen berbasis algoritma Naïve Bayes dapat digunakan sebagai alat evaluasi dalam mengidentifikasi persepsi publik terhadap e-commerce di Indonesia. Metode ini dapat membantu perusahaan dalam meningkatkan strategi pemasaran serta pengelolaan pengalaman pelanggan. Untuk meningkatkan efektivitas analisis sentimen, pendekatan lain seperti Support Vector Machine (SVM) atau Deep Learning dapat dikombinasikan dengan Naïve Bayes guna menghasilkan klasifikasi yang lebih akurat.

Kata Kunci— E-Commerce, Analisis Sentimen, Twitter, Naïve Bayes, Kepopuleran

I. PENDAHULUAN

Saat ini perdagangan menjadi salah satu daya tarik para pelaku usaha. Banyak pengusaha, baik perusahaan besar maupun ritel, beralih atau mengembangkan usaha ke arah digital yaitu berupa e-commerce. Tidak bisa dipungkiri lagi bahwa e-commerce dapat mempermudah transaksi antara penjual dan pembeli, selain dalam bentuk website ditambah dengan fasilitas e-commerce berupa aplikasi smartphone sehingga masyarakat bisa dengan mudah mengakses dimanapun dan kapanpun [1].

Masyarakat indonesia yang bertransaksi menggunakan e-commerce semakin bertambah, sehingga banyak bermunculan e-commerce baru yang membuat masyarakat bingung dalam memilih e-commerce yang dipercaya, Ada suatu pemberitaan kurang baik yang berakibat pandangan negatif terhadap satu e-commerce kejadian ini dari tweet dengan akun@CibiWiranegara, pesan dalam cuitan tersebut memuat mengenai pengungkapan pembicaraan di Whatsapp dengan

seorang kurir yang mengatakan terjadi keterlambatan pengiriman paket barang. untuk itu diperlukan sebuah penilaian yang dapat mengetahui tingkat kepopuleran e-commerce di indonesia berdasarkan analisis sentimen pada sosial media [2].

Analisis sentimen berguna untuk menentukan apakah opini atau komentar terhadap suatu permasalahan memiliki kecenderungan positif atau negatif dan dapat dijadikan sebagai acuan dalam meningkatkan suatu pelayanan, ataupun meningkatkan kualitas produk. Besarnya manfaat serta pengaruh dari melakukan analisis sentimen membuat penelitian dan aplikasi mengenai analisis sentimen berkembang pesat. Salah satu teknik pembelajaran dari text mining untuk analisis sentimen adalah Naïve Bayes Classifier. Metode Naïve Bayes ini salah satu metode terbaik dalam pengklasifikasian. Naïve Bayes menggunakan statistik sederhana berdasarkan teorema bayes yang mengasumsikan keberadaan atau ketiadaan dari suatu fitur tertentu dari suatu kelas yang tidak berhubungan dengan fitur lainnya [2].

Pada penelitian ini media sosial yang digunakan adalah twitter, sebab banyak masyarakat yang menggunakan twitter dalam menyampaikan opininya dari pada menyampaikan secara langsung. Twitter memungkinkan penggunanya saling berinteraksi satu sama lain sehingga memungkinkan untuk mendapatkan informasi yang dibutuhkan. Oleh karena itu, analisa sentimen merupakan salah satu solusi mengatasi masalah untuk mengelompokkan opini berupa opini positif atau negatif [2].

Naive Bayes merupakan klasifikasi paling sederhana dan paling umum digunakan. Naive Bayes menghitung probabilitas kelas berdasarkan distribusi kata-kata yang ada dalam dokumen. Naive Bayes memiliki beberapa keunggulan seperti sederhana, cepat dan akurasi yang tinggi. Banyak peneliti telah melakukan klasifikasi sentimen dengan menggunakan Naive Bayes [2].

Dalam penelitian ini dipilih metode Naïve Bayes karena dianggap sebagai metode yang berpotensi baik untuk melakukan klasifikasi data dari pada metode klasifikasi lainnya dalam hal akurasi dan komputasi. sentimen, kemudian membagi opini tersebut menjadi data sentimen positif dan negatif. Dan metode yang digunakan adalah metode klasifikasi yang dapat digunakan untuk mengklasifikasikan kumpulan data tweet [2].

Dari permasalahan diatas maka dari itu penulis menyusun laporan skripsi yang berjudul: ”Analisis Sentimen Tingkat Kepopuleran E-Commerce Pada Pengguna Twitter Menggunakan Algoritma Naive Bayes”.

II. LANDASAN TEORI

A. Analisis Sentimen

Analisis sentimen dalam Bahasa Indonesia adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif maupun sentimen negatif. Pendapat yang hampir senada dikemukakan oleh [3], di mana analisis sentimen digunakan untuk memahami komentar yang diciptakan oleh pengguna (internet) dan menjelaskan bagaimana sebuah produk maupun brand diterima oleh mereka.

Sedangkan menurut [3], analisis sentimen adalah proses yang digunakan untuk menentukan opini, emosi dan sikap yang dicerminkan melalui teks, dan biasanya diklasifikasikan menjadi opini negatif dan positif. Dari ketiga pendapat di atas, bisa diambil kesimpulan bahwa analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif[4].

B. *Text Mining*

Text mining adalah proses mengeksplorasi pola-pola yang ada dalam sumber data tekstual untuk menarik informasi dan pengetahuan dari sejumlah besar data. Tujuan dari text mining adalah untuk menganalisis koneksi dokumen menggunakan kata-kata yang mewakili konten untuk mendapatkan informasi yang berguna. Data yang digunakan sebagai input dalam proses text mining tidak terstruktur, sedangkan data mining menggunakan data terstruktur atau database sebagai input. Namun, penambangan teks menggunakan konsep dan teknik penambangan data. Hal ini karena teks diperlukan dalam format yang lebih terstruktur sebelum dapat dilakukan analisis lebih lanjut[5].

C. *Pre-Processing*

Tahap pre-processing merupakan tahapan dalam pengolahan data yang bertujuan untuk menghilangkan data-data yang tidak akan digunakan pada proses selanjutnya.

Pre-processing dilakukan untuk mempersiapkan dokumen teks yang masih belum terstruktur hingga siap untuk dilakukan analisis teks sesuai dengan metode yang digunakan dan data yang akan digunakan menjadi terstruktur.

Proses *Pre-processing*

masih	Di	Buka	kompetisi	bukalapak
Di	pertengahan	Maret	April	Ini
Ikuti	Kompetisi	Dan	Dapatkan	
	hadiahmu			
Untuk	Di	Bawa	Pulang	

Setelah Melakukan Proses *pre-processing*

masih	Buka	kompetisi	bukalapak
	pertengahan	Maret	April
Ikuti	kompetisi	Dapatkan	
Hadiahmu	Bawa	Pulang	

1.*Case Folding*

Case folding merupakan proses penyamaan huruf pada dokumen tweet dengan mengubah huruf kapital menjadi huruf kecil, supaya merenggangkan kata berlebih-lebihan dikarenakan perbedaan huruf dan akan menghasilkan dokumen tweet baru [6]. Case folding dilakukan karena data tweet yang di dapat terdapat banyak perbedaan bentuk huruf, tahapan ini dilakukan untuk merubah semua bentuk huruf pada kalimat menjadi huruf kecil (lowercase) atau dapat disebut juga penyeragaman bentuk huruf.

2. *Remove URL*

Seringnya muncul sebuah url dari data twitter membuat data tidak efektif dan tidak memiliki arti. Untuk itu perlu adanya penghapusan url tersebut Kemunculan web atau url ini disebabkan karena banyaknya user mempromosikan sebuah produk pada situs mereka supaya user yang lain langsung bisa masuk pada halaman web yang dimaksud [7].

3.Tokenization

Tokenisasi yaitu tahap pemotongan atau memisahkan per-kata atau bisa disebut pemenggalan kalimat menjadi kata (Artanti, 2018).

Menurut Nugroho (2018) [8] tahap tokenisasi ini semua kata yang ada dalam tiap dokumen dipisahkan dan dihilangkan tanda baca, serta dihilangkan jika terdapat simbol atau apapun yang bukan huruf.

4. *Filtering*

Filtering merupakan proses membuang kata yang tidak penting dari proses tokenisasi sebelumnya [9]. Filtering bertujuan untuk menghilangkan kata-kata yang sering muncul, namun tidak memiliki makna yang berarti stopword [10].

Menurut R. Apriani (2019) [11] *Filtering* merupakan pembuangan kata-kata tidak penting dari hasil token. Selain juga dilakukan penghapusan tanda baca dan *stopword*. *Stopword* diproses pada sebuah kalimat jika mengandung kata-kata yang sering keluar dan di anggap tidak penting seperti waktu, penghubung, dan lain sebagainya.

D. *Naïve Bayes*

Naive Bayes (Naive Bayes Classifier) adalah algoritma yang sangat efektif dalam permasalahan klasifikasi atau penggolongan. Algoritma ini bekerja berdasarkan probabilitas yang sudah ada untuk menentukan probabilitas yang akan datang. Meskipun dataset yang digunakan ada jutaan jumlahnya, *Naive Bayes* adalah pendekatan yang dianjurkan untuk digunakan. Untuk memahami lebih jauh tentang *Naive Bayes* maka perlu terlebih dahulu memahami dalil Bayes. Dalil Bayes dinamakan sesuai dengan nama penemunya Rev. Thomas Bayes. Algoritma ini bekerja berdasarkan probabilitas bersyarat, yaitu probabilitas yang menjelaskan tentang sesuatu akan terjadi berdasarkan peristiwa yang telah terjadi sebelumnya. *Bayes* merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema *Bayes* dengan asumsi independensi (ketidaktergantungan) yang kuat atau naïf. Metode *Bayesian classification* digunakan menganalisis dalam membantu tercapainya pengambilan keputusan terbaik suatu permasalahan dari sejumlah alternatif. Kaitan antara *Naive Bayes* dengan klasifikasi, korelasi hipotesis, dan bukti dengan klasifikasi adalah hipotesis dalam teorema *Bayes* merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadi masukan dalam model klasifikasi[6].

E. *Confusion Matriks*

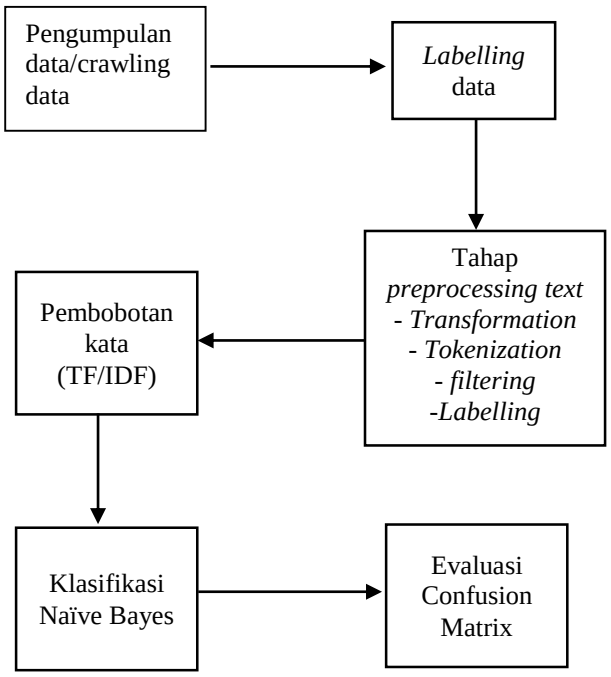
Pada data mining untuk mengukur atau ada beberapa cara untuk mengukur kinerja dari model yang dihasilkan salah satunya menggunakan *confusion matriks* (akurasi) *Confusion matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Presisi atau *confidence* adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. *Recall* atau *sensitivity* adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar[12].

III. METODE PENELITIAN

Dalam penelitian ini Twitter menggunakan lima E-Commerce dengan jumlah data Shopee yang diambil sebanyak 1000 data, Tokopedia yang diambil sebanyak 1000 data, Bukalapak yang diambil sebanyak 1000 data, Lazada yang diambil sebanyak 106 data, Blibli yang diambil sebanyak 1000 data.

A. Kerangka Penelitian

Rancangan penelitian yang akan dilakukan disajikan dalam bentuk kerangka penelitian yang meliputi studi pengumpulan data, tahap *preprocessing*, pembobotan TF-IDF, klasifikasi *naïve bayes* dan evaluasi (*confusion matrix*). Berikut gambar rancangannya pada gambar 3.1.



Gambar. 1. Kerangka Penelitian

B. Pengumpulan *Data*

Proses pengumpulan data dalam penelitian ini mengenai *e-commerce* yang di peroleh dari twitter dengan *query* "Shopee", "Tokopedia", "Bukalapak", "Lazada", dan Blibli dengan metode crawling data. Data tweet yang diambil merupakan tweet dari akun pengguna twitter

C. *Crawling Data*

Data diambil dengan cara *crawling* data pada media sosial twitter dengan memanfaatkan search *Application Program Interface* (API) yang disediakan oleh *twitter*. Crawling data

dilakukan pada tanggal 10 Maret 2022 Kemudian setelah mendapatkan akses API *key* maka dapat dicari data yang diinginkan melalui fasilitas kolom pencarian di *twitter* yang berhubungan dengan penelitian ini. Dalam penelitian, data dibagi menjadi 3 kategori komentar yaitu komentar positif negatif, dan netral yang diberikan oleh pengguna *twitter* dalam bahasa Indonesia. Gambar 3.2 berikut ini merupakan alur dari proses *crawling* data menggunakan *Tools Orange*

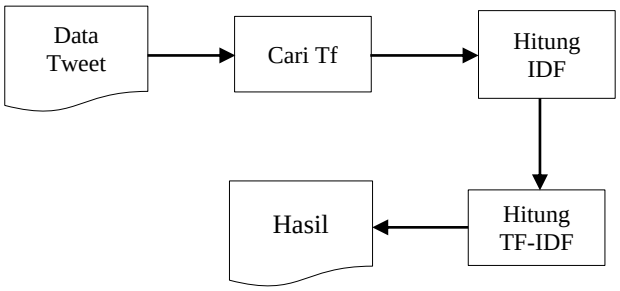
Gambar. 1. Alur Crawling Data

D. Pelabelan

Pada tahap ini, data yang telah diproses sebelumnya akan diproses pelabelan data. Pelabelan data dilakukan dalam 2 (dua) tahap, tahap pertama dilengkapi dengan alat orange untuk melihat hasil emosional penggunaan emosi multibahasa Indonesia. Untuk tahap kedua, pelabelan data dilakukan secara manual, menggunakan rumus Ms. Excel IF (A2<0;" negatif"; IF(A2>0;" positif";" netral" tahap kedua dilakukan karena emosi pada tahap pertama hanya dalam bentuk digital dan tidak dapat dijadikan label untuk proses selanjutnya. Dengan tag, data tweet akan mudah dikelompokkan.

E. Pembobotan Kata

Setelah dilakukannya preprocessing dokumen (transformation, tokenization, filtering), selanjutnya dilakukan proses term weighting yang nanti nya akan diberikan bobot atau nilai dimana bobot tersebut mengindikasikan pentingnya sebuah term terhadap dokumen. Penghitungan bobot tiap term dicari pada setiap dokumen bertujuan untuk dapat mengetahui ketersediaan dan kemiripan suatu term di dalam dokumen. Semakin banyak kata tersebut muncul pada koleksi dokumen, semakin tinggi nilai atau bobot term tersebut. Setelah tahapan pemberian bobot selesai barulah dilanjutkan ke proses klasifikasi. Dalam pembobotan kata, metode yang digunakan dalam melakukan pembobotan adalah metode TF-IDF.



Gambar. 3. Proses Pembobotan Kata (TF-IDF)

IV. HASIL DAN PEMBAHASAN

A. *Hasil Term Frequency*

Hasil Term Frequency Jika kata tersebut tidak muncul maka diberi nilai 0, dan jika terdapat maka diberi nilai 1. Hasil Perhitungan TF. Metode TF-IDF merupakan suatu cara untuk memberikan bobot hubungan suatu kata (*term*) terhadap dokumen. Metode ini menggabungkan dua konsep untuk perhitungan bobot, yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen tertentu dan inverse frekuensi dokumen yang mengandung kata tersebut. Frekuensi kemunculan kata di dalam dokumen yang diberikan menunjukkan seberapa penting kata itu di dalam dokumen tersebut.

TABEL 1. HASIL PERHITUNGAN TF

term	D 1	D 2	D 3	D 4	D 5	D 6	D 7	D 8	D 9	D 10
Shope e	1	0	0	0	0	0	0	0	0	1
promo	0	2	0	0	1	0	2	1	0	0
hadiah	0	0	1	0	0	0	0	0	0	0
bodoh	0	0	0	1	0	0	0	0	0	0
hemat	0	0	0	0	1	0	0	0	0	0
murah	0	1	0	0	0	0	1	0	0	0
Lama	0	0	0	0	0	0	0	0	0	1

B. *Hasil Inverse Document Frequency*

Kemudian total dari seluruh dokumen(I) dibagi oleh DF menghasilkan nilai I/DF. Untuk mendapatkan nilai IDF, lakukan operasi logaritma terhadap nilai I/DF.

TABEL 2. HASIL DATA PERHITUNGAN IDF

DF	D/DF	LOG	1+IDF
2	5	0.69	1.69
6	2	0.22	1.22
1	10	1	2
1	10	1	2
1	10	1	2
2	5	0.69	1.69
1	10	1	2
Nilai Bobot Setiap Dokumen			

C. *Hasil Term Frequency- Inverse Document Frequency*

Untuk mendapatkan nilai TF-IDF (Term Frequency-Inverse Document Frequency) kalikan bobot tiap kata dari seluruh dokumen(TF) dengan nilai dari IDF

Hasil Perhitungan TF-IDF dapat dilihat pada tabel berikut ini:

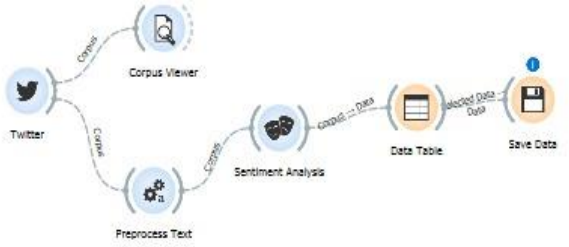
TABEL 3. HASIL PERHITUNGAN TF-IDF

W									
D1	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
2	0	0	0	0	0	0	0	0	1,6
0	2	0	0	1	0	2	1	0	0
0	0	2	0	0	0	0	0	0	0
0	0	0	2	0	0	0	0	0	0

W									
D1	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
0	0	0	0	2	0	0	0	0	0
0	1	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	2

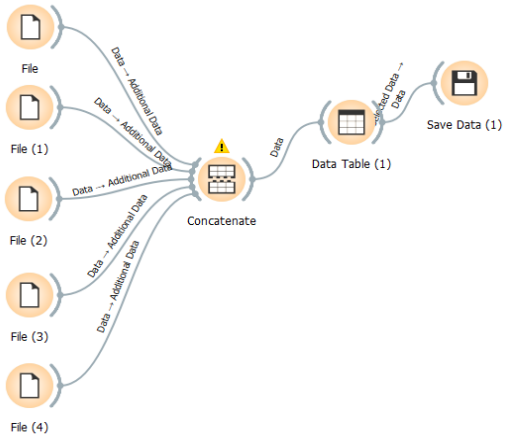
D. *Hasil Klasifikasi Naive Bayes*

Pada hasil klasifikasi ini penulis akan menganalisa dan menghitung hasil dari *performance* algoritma *Naïve Bayes* dengan menghitung akurasi, *precision*, dan *recall* menggunakan *Tools Orange*. Proses Crawling Data sebagai berikut:



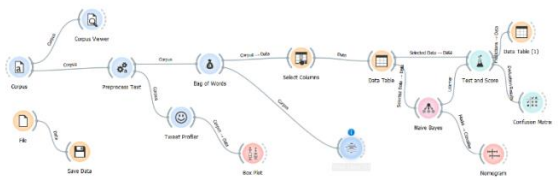
Gambar .4. Proses Crawling Data

Setelah proses *crawling* data selanjutnya Pada gambar dibawah proses penggabungan data dilakukan di orange. Proses penggabungan data sebagai berikut



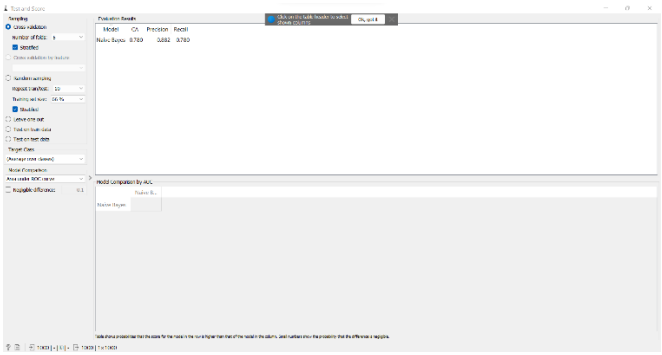
Gambar.5. Proses Penggabungan Data

Pada gambar 6, setelah roses penggabungan data selanjutnya proses pengujian data, proses pengujian data dilakukan diaplikasi orange dan pengujian data sebagai berikut:



Gambar.6. Proses Pengujian Data

Setelah proses penggabungan dan proses pengujian selanjutnya Hasil dari Klasifikasi Naïve bayes pada Semua e-commerce dengan akurasi, presisi dan recall sebagai berikut:



Gambar.7. Hasil Klasifikasi Semua E-Commerce

E. *Confusion Matrix* Semua *E-Commerce*

Pada data mining untuk mengukur atau ada beberapa cara untuk mengukur kinerja dari model yang dihasilkan salah satunya menggunakan *confusion matriks* (akurasi) *Confusion matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Presisi atau *confidence* adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. *Recall* atau *sensitivity* adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar.

TABEL 4. CONFUSION MATRIX SEMUA E-COMMERCE

		Prediksi		
		Negatif	Netral	Positif
Aktual	Negatif	572	239	4
	Netral	65	1691	42
	Positif	120	562	810

1. Akurasi

Merupakan rasio prediksi benar (positif, netral dan negatif) dengan keseluruhan data. Berikut ini rumus Akurasi.

$$\text{Akurasi} = \frac{\text{TP} + \text{TNet} + \text{TN}}{\text{Jumlah } Data} = \frac{810 + 1691 + 572}{4105}$$

$$= 0{,}7485992692$$

2. Presisi

Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Berikut ini rumus Presisi

$$Precission = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{810}{810 + 120}$$

$$= 0{,}8709677419$$

3. Recall

Merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Berikut ini rumus Recall

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{810}{810 + 239} = 0{,}7721639657$$

F. Hasil Evaluasi

Evaluasi dipakai sebagai parameter dari kebenaran dalam penelitian ini, dengan akurasi perhitungan dari tabel confusion matrix (matriks klasifikasi). Untuk menghitung akurasi, precision, recall. Berikut ini dapat dilihat pada tabel 5.

TABEL 5. HASIL EVALUASI

	CA	***Precision***	***Recall***
Shopee	0.741	0.612	0.978
Tokopedia	0.78	0.531	0.957
Blibli	0.778	0.654	0.883
Lazada	0.714	1	0.533
Bukalapak	0.765	1	0.856

V. KESIMPULAN

Berdasarkan kesimpulan yang didapatkan dari penelitian ini bahwa analisis sentimen tingkat kepopuleran *E-Commerce* pada pengguna Twitter menggunakan lima *E-Commerce* dengan jumlah data Shopee yang diambil sebanyak 1000 data, Tokopedia yang diambil sebanyak 1000 data. Dengan mengklasifikasikan sentimen ke dalam 3 kategori yaitu positif, netral, dan negatif. Penggunaan Algoritma Naïve Bayes Classifier sangat efektif untuk digunakan sebagai proses klasifikasi Twitter yang dibutuhkan dalam sistem analisis sentiment ini. Model yang dihasilkan dalam diuji pada orange untuk mendapatkan nilai akurasi, *precision* dan *recall*. Dari algoritma Naïve Bayes sehingga didapat pengujian dengan nilai akurasi 0.784, *precision* 0.870, *recall* dan 0.772, untuk keseluruhan data pada *E-Commerce*. Sedangkan setiap *E-Commerce* nya bernilai untuk Shopee akurasi 0.741, *precision* 0.612, *recall* 0.978. Untuk Tokopedia akurasi 0.78, *precision 0*.531, *recall* 0.957. Untuk Blibli akurasi 0.778, *precision* 0.654, *recall* 0.883. Untuk Lazada akurasi 0.714, *precision* 1, *recall* 0.533. Untuk Bukalapak akurasi 0.765, *precision* 1, *recall* 0.856. Berdasarkan hasil analisis tersebut membuktikan bahwa yang lebih populer ada pada *E-Commerce* Tokopedia, dan Blibli dari 5 *e-commerce.*

REFERENSI

[1] A. S. Widagdo, B. S. W.A, and A. Nasiri, "Analisis Tingkat Kepopuleran E-Commerce Di Indonesia Berdasarkan Sentimen Sosial Media Menggunakan Metode Naïve Bayes," *J. Inf. J. Penelit. dan Pengabdi. Masy.*, vol. 6, no. 1, pp. 1–5, 2020, doi: 10.46808/informa.v6i1.159.

[2] F. Afshoh, "Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Jual Rokok Pada Media Sosial Twitter," *Inform. Progr. Stud. Komunikasi, Fak. Inform. D A N Surakarta, Univ. Muhammadiyah*, pp. 1–17, 2017.

[3] N. & Yi, "Analisis Sentimen," 2003.

[4] H. Nurrun Muchammad Shiddieqy, S. Paulus Insap, and W. Wing Wahyu, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. March, pp. 57–64, 2016.

[5] D. T. Seimbang and T. Mining, "Infotek : Jurnal Informatika dan Teknologi Analisa Sentimen Kicauan Twitter Tokopedia Dengan Optimalisasi Data Tidak Seimbang Menggunakan Algoritma SMOTE Studi Sistem Informasi , Universitas Bina Sarana Informatika Pendahuluan dapat berupa sebuah pujian ," vol. 5, no. 1, pp. 108–116, 2022.

[6] M. Yasid, "Analisis Sentimen Maskapai Citilink Pada Twitter dengan Metode Naive Bayes," *J. Ilm. Inform.*, vol. 7, no. 2, pp. 82–86, 2019.

[7] S. Mujilahwati, "PRE-PROCESSING TEXT MINING PADA DATA TWITTER," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 2089–9815, 2016.

[8] A. Nugroho, "Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstrasi Fitur N-Gram," *J-SAKTI (Jurnal Sains Komput. dan Inform.*, vol. 2, no. 2, p. 200, 2018, doi: 10.30645/j-sakti.v2i2.83.

[9] M. W. A. Putra, Susanti, Erlin, and Herwin, "Analisis Sentimen Dompet Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier," *IT J. Res. Dev.*, vol. 5, no. 1, pp. 72–86, 2020, doi: 10.25299/itjrd.2020.vol5(1).5159.

[10] B. W, I. Riski, K. Dwi, R. Nooraeni, T. Siahaan, and Y. Dhea, "Analisis Text Mining dari Cuitan Twitter Mengenai Infrastruktur di Indonesia dengan Metode Klasifikasi Naïve Bayes," *Eig. Math. J.*, vol. 1, no. 2, pp. 92–101, 2019, doi: 10.29303/emj.v1i2.36.

[11] R. Apriani *et al*., "ANALISIS SENTIMEN DENGAN NAÏVE BAYES TERHADAP KOMENTAR APLIKASI TOKOPEDIA," *J. Rekayasa Teknol. Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019, [Online]. Available: https://rekayasa.nusaputra.ac.id/article/view/86

[12] T. Rosandy, "Perbandingan Metode Naive Bayes Classifier dengan Metode Decision Tree Untuk Menganalisa Kelancaran Pembiayaan," *J. TIM Darmajaya*, vol. 02, no. 01, pp. 52–62, 2016.