

Analisis Sentimen Terhadap Event Big Sale 11.11 Shopee di Media Sosial Instagram Menggunakan Metode Naïve Bayes

Nurul Anizah^{a,1,*}, Yulita Salim^{a,2}, Lutfi Budi Ilmawan^{a,3}

^a Fakultas Ilmu Komputer, Universitas Muslim Indonesia, Jalan Urip Sumoharjo, Makassar, 90231, Indonesia

¹ nurulanizahh19@gmail.com; ² yulitasalim@umi.ac.id; ³ lutfibudi.ilmawani@umi.ac.id

*corresponding author

INFORMASI ARTIKEL	ABSTRAK
Diterima : 11 – 07 – 2022 Direvisi : 21 – 02 – 2023 Diterbitkan : 28 – 02 – 2023	Banyak E-commerce yang mempromosikan produk atau event pada Instagram, hal ini tentu memudahkan penjual dalam menjangkau pasar yang lebih luas. Termasuk acara yang diadakan oleh Shopee pada bulan November 2020 yaitu Event Big Sale 11.11 merupakan acara besar yang diadakan Shopee. Para pengunjung sudah bisa melihat produk yang penjual tawarkan melalui gambar tersebut. Oleh karena itu, diperlukan sebuah penelitian yang dapat menganalisa opini pengguna di Instagram dari suatu cara bagi sebuah produk Penelitian ini bertujuan untuk untuk mengklasifikasi opini pengguna terhadap Event Big Sale 11.11 Shopee yang banyak tersebut kedalam sentiment positif, negatif dan netral dengan menerapkan metode Naïve Bayes. Hasil dari penelitian dapat membantu masyarakat untuk mengambil keputusan dalam pembelian produk di Shopee. Metode Naive Bayes digunakan untuk melakukan klasifikasi opini masyarakat terhadap Event Big Sale 11.11. Karakteristik dari metode ini adalah tiap fitur atribut data dianggap independen, satu dan lainnya terpisah dan memiliki nilai sendiri. Metode ini dianggap memiliki performa yang handal dan kompetitif dalam proses pengklasifikasian karena asumsi independen atribut yang dimiliki sebuah data sangat sesuai dengan pengaplikasian di dunia nyata.
Kata Kunci: Analisis Sentimen Shopee Instagram Naïve Bayes	

This is an open access article under the [CC-BY-SA](#) license



I. Pendahuluan

Perkembangan teknologi memfasilitasi kegiatan masyarakat dalam bisnis jual beli dan memfasilitasi kegiatan lainnya. Teknologi yang berkembang dalam bisnis jual beli ini salah satunya adalah e-commerce. E-commerce adalah sarana transaksi komersial antara organisasi dan individu yang dilakukan secara digital. Perkembangan e-commerce di Indonesia bisa dikatakan sangat pesat. Apabila dibandingkan dengan Negara lain yang ada di sekitar. Seperti yang dapat kita lihat sendiri, e-commerce semakin menjamur seiring dengan berjalannya waktu. Tentunya hal ini juga menyesuaikan dengan permintaan pasar yang semakin mengandrungi kegiatan perbelanjaan online.

Pertumbuhan dalam penggunaan media sosial juga merupakan salah satu efek dari peningkatan jumlah pembelian barang dengan cara digital, tidak datang untuk melihat toko fisik secara langsung. Dengan pertumbuhan media sosial yang tinggi, media sosial juga sering digunakan dalam melakukan kegiatan bisnis seperti untuk membuat katalog dan menampilkan produk yang dijual, sehingga konsumen dapat melihat dan membeli produk tersebut [1].

Shopee adalah marketplace berbasis aplikasi mobile, yang sedang berkembang di Indonesia. Shopee menjadi situs wadah jual beli secara online yang telah melakukan perubahan untuk menarik minat pelanggan agar lebih banyak bertransaksi melalui situs tersebut. Dengan adanya Shopee, memudahkan penjualan untuk melakukan promosi dengan jangkauan pasar yang lebih luas, sedangkan bagi pelanggan menjadi lebih mudah untuk mengetahui sebuah produk karena pelanggan dapat memberikan opini atau pendapat pada sebuah produk.

Salah satu marketplace terbesar di Indonesia, Shopee juga sangat gencar melakukan promosi untuk menarik minat pelanggan, salah satunya adalah dengan mengadakan *Event Big Sale 11.11* Shopee. Event ini menawarkan banyak sekali keuntungan seperti gratis ongkir dan diskon sampai dengan 50%. Barang-barang yang promosi pada *Event Big Sale 11.11* Shopee juga beragam, mulai dari elektronik, pakaian, produk kecantikan dan lainnya.

Salah satu media sosial yang banyak dimanfaatkan masyarakat untuk beropini adalah Instagram. Instagram sebagai salah satu media sosial populer dimana penggunaanya dapat mengekspresikan opini yang objektif tentang topik yang berbeda. Brand Development Lead Instagram APAC Paul Webster mengungkapkan, bahwa sejak diluncurkan pada tahun 2010 aplikasi Instagram telah memiliki 400 juta lebih pengguna aktif dari seluruh dunia [2]. Indonesia sendiri adalah salah satu negara dengan jumlah pengguna Instagram terbanyak. Banyaknya pengguna Instagram dapat dimanfaatkan untuk mendapatkan berbagai macam informasi, salah satunya adalah informasi tentang jual beli pada E-commerce. Banyak E-commerce yang mempromosikan produk atau event pada Instagram, hal ini tentu memudahkan mereka dalam menjangkau pasar yang lebih luas. Cara menggunakan media sosial Instagram ini cukup mudah karena pelaku bisnis tinggal membuka akun, mengupload gambar yang disertai dengan deskripsi produk. Dengan begitu komentar dan testimonial di Instagram dapat dimanfaatkan untuk menilai tingkat kepercayaan sebuah online shop di Instagram. Nilai kepercayaan tersebut didapatkan dengan melihat komentar positif atau negatif dari setiap gambar dan caption yang diposting sebuah online shop di Instagram. Penentuan komentar positif atau negatif ini merupakan salah satu bagian analisis sentiment [3]. Oleh karena itu, diperlukan sebuah penelitian yang dapat menganalisa opini pengguna di Instagram dari suatu cara bagi sebuah produk.

Opini pengguna di Instagram akan diklasifikasi menjadi tiga kelas, yaitu positif, negatif dan netral. Penentuan sentimen positif, negatif dan netral suatu opini dapat dikerjakan secara manual, namun dilihat dari jumlah pengguna yang banyak maka opini yang dihasilkan juga semakin banyak. Sehingga membutuhkan waktu dan usaha yang semakin banyak pula. Oleh karena itu, dibutuhkan sebuah analisa opini untuk mengklasifikasi opini pengguna terhadap *Event Big Sale 11.11* Shopee yang banyak tersebut kedalam sentiment positif, negatif dan netral dengan menerapkan metode Naïve Bayes. Berdasarkan latar belakang di atas, penulis bermaksud melakukan penelitian yang berjudul “Analisis Sentimen Terhadap *Event Big Sale 11.11* Shopee Di Media Sosial Instagram Menggunakan Metode Naïve Bayes”. Dalam penelitian ini, acara yang dianalisa adalah *Event Big Sale 11.11* yang diadakan oleh Shopee.

Metode Naive Bayes digunakan untuk melakukan klasifikasi opini masyarakat terhadap *Event Big Sale 11.11*. Karakteristik dari metode ini adalah tiap fitur atribut data dianggap independen, satu dan lainnya terpisah dan memiliki nilai sendiri. Metode ini dianggap memiliki performa yang handal dan kompetitif dalam proses pengklasifikasian karena asumsi independen atribut yang dimiliki sebuah data sangat sesuai dengan pengaplikasian di dunia nyata. Adapun keluaran dari penelitian berupa jurnal yang akan diterbitkan secara online sehingga dapat diketahui oleh masyarakat dan membantu keputusan dalam pembelian produk di Shopee.

II. Metode

A. Metode Analisis

1) Analisis Sentimen

Analisis Sentimen adalah salah satu teknik dalam mengekstrak informasi berupa pandangan (sentimen) seseorang terhadap suatu isu atau kejadian. Analisis Sentimen dapat digunakan untuk mengungkap opini publik terhadap suatu isu, kepuasan pelayanan, kebijakan, cyber bullying, memprediksi harga saham, dan analisis pesaing berdasarkan data tekstual [3]. Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat atau pendapat. Polaritas mempunyai arti apakah teks yang ada dalam dokumen, kalimat atau pendapat memiliki aspek positif atau negatif.

2) Instagram

Instagram adalah sebuah aplikasi yang digunakan untuk membagikan foto dan video. Makin populernya Instagram sebagai aplikasi yang digunakan untuk membagi foto mengakibatkan banyak pengguna yang terjun ke ranah bisnis seperti akun sosial bisnis yang turut mempromosikan produk-produknya lewat Instagram. Instagram sebagai media sosial yang juga banyak digunakan untuk media berbisnis, tentu akan menjadi wadah opini dari masyarakat terhadap event tersebut. Perkembangan Instagram sangatlah pesat dan dibuktikan dengan kepopuleran Instagram yang sudah mencapai sebanyak 150 juta pengguna. Ini merupakan pencapaian rekor yang fantastis. Trend Instagram merupakan sebuah cara marketing yang memakai produk sebagai sarana komunikasi.

3) Shopee

Shopee adalah sebuah aplikasi yang bergerak dibidang jual beli secara online dan dapat diakses secara mudah dengan menggunakan smartphone. Shopee hadir dalam bentuk aplikasi yang memudahkan penggunaanya dalam melakukan kegiatan berbelanja secara online tanpa harus ribet menggunakan perangkat komputer. Namun cukup menggunakan smartphone anda, Shopee akan menawarkan berbagai macam produk-produk fashion hingga produk untuk kebutuhan sehari-hari. Shopee gencar melakukan promosi untuk menarik minat pelanggan, salah satunya adalah dengan mengadakan *Event Big Sale 11.11* ini menawarkan banyak sekali keuntungan seperti gratis ongkir dan diskon sampai dengan 50%.

4) Pre-processing

Metode pengolahan data ini dilakukan dengan cara melakukan *preprocessing* yaitu tahapan awal dari text mining untuk mengubah data sesuai kebutuhan. Proses yang dilakukan pada tahap ini seperti *case folding*, *data cleaning*, *stop removal*, *stemming*, dan *tokenizing* [4].

a) Case Folding

Case folding adalah proses mengubah semua huruf dalam dokumen menjadi huruf kecil contohnya kalimat “Banyak Diskon ayo semua belanja” menjadi “banyak diskon ayo semua belanja”. Hanya huruf “a” sampai dengan “z” yang diterima.

b) Data Cleaning

Data cleaning merupakan proses pembersihan kata dengan menghilangkan semua tanda baca seperti koma (,), titik (.), at (@), tagar (#) dan tanda baca lainnya yang bertujuan untuk mengurangi noise pada proses ini.

c) Tokenizing

Tokenizing atau parsing adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya.

d) Stopword Removal

Stopword Removal proses untuk menghapus kata kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansi dengan teks. Untuk daftar stopwords atau stoplist yang digunakan bersumber dari banyak penelitian yang sudah ada sebelumnya [5].

e) Stemming

Stemming merupakan proses untuk mencari sistem (kata dasar) dari kata hasil stopwords removal (filtering). Pentingnya stemming dalam proses pembuatan sistem untuk menghilangkan imbuhan pada awalan dan akhiran [6].

5) Metode Naïve Bayes

Metode *Naïve Bayes* merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Metode naive bayes memiliki kelebihan, yaitu cepat dalam perhitungan, algoritma yang sederhana dan berakurasi tinggi. Metode *Naïve Bayes* yang hanya membutuhkan jumlah data pelatihan (*Data Training*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian [7].

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Data Training*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian [8]. *Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan [9].

Berikut adalah rumus dari metode klasifikasi Naïve Bayes untuk menghitung nilai probabilitas setiap kata.

a) Klasifikasi Naïve Bayes

Untuk melakukan klasifikasi Naïve Bayes menggunakan rumus berikut(1)

$$CNB = \arg \max_c P(c) \prod_{i=1}^n P(F_i|c) \quad (1)$$

Keterangan

$\arg \max$	: nilai <i>posterior probability</i> terbesar suatu kelas
$P(F_i c)$	: nilai <i>conditional probability</i>
$P(c)$	: nilai <i>prior probability</i>

b) Hitung Nilai Prior Probability (2)

$$P(c) = \frac{N_c}{N_{doc}} \quad (2)$$

Keterangan

$P(c)$	: nilai probabilitas data latih
N_c	: jumlah kelas positif, negatif dan netral
N_{doc}	: total data latih

c) Hitung Conditional Probability (3)

$$P(f_i|c) = \frac{\text{count}(f_i,c)}{\sum_{f \in V} \text{count}(f_i,c)} \quad (3)$$

Keterangan

$P(f_i c)$	: Nilai probabilitas tiap kata
$f_{i,c}$	: Jumlah kata tertentu yang terdapat pada seluruh dokumen pada kelas c
c	: Kelas (positif, negatif atau netral)
$F_{i,c}$	: Jumlah seluruh kata yang terdapat pada kelas c
V	: Jumlah kata berbeda pada semua dokumen

6) Metode Confusion Matrix

Metode pengujian *confusion matrix* merupakan salah satu metode yang digunakan untuk menguji akurasi, presisi, recall, dan *F1-score* dari sebuah metode klasifikasi [10]. Berikut adalah contoh tabel *confusion matrix* di bawah ini:

	Classified Positive	Classified Negative
Actual Positive	TP (True Positive)	FT (False Positive)
Actual Negative	FN (False Negative)	TN (True Negative)

Keterangan:

- TP : Jumlah data positif yang terdeteksi dengan benar
- FT : jumlah data negatif yang terdeteksi sebagai data positif
- FN : yaitu jumlah data positif namun terdeteksi sebagai data negatif
- TN : jumlah data negatif yang terdeteksi dengan benar

a) Akurasi

Akurasi adalah tingkat kedekatan nilai prediksi dengan nilai actual, untuk menghitung akurasi menggunakan rumus (4s).

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (4)$$

b) Presisi

Presisi adalah tingkat ketepatan atau ketelitian dalam pengklasifikasian untuk menghitung presisi menggunakan rumus (5)

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

c) Recall

Recall berfungsi untuk mengukur proporsi positif aktual yang benar diidentifikasi, untuk menghitung recall menggunakan rumus (6)

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

d) F1-Score

F1-score digunakan untuk mengetahui keseimbangan antara presisi dan *recall* yang didapat dari sistem yang akan dibangun untuk menghitung *f1-score* menggunakan rumus (7)

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \times 100\% \quad (7)$$

7) Simple Scrap

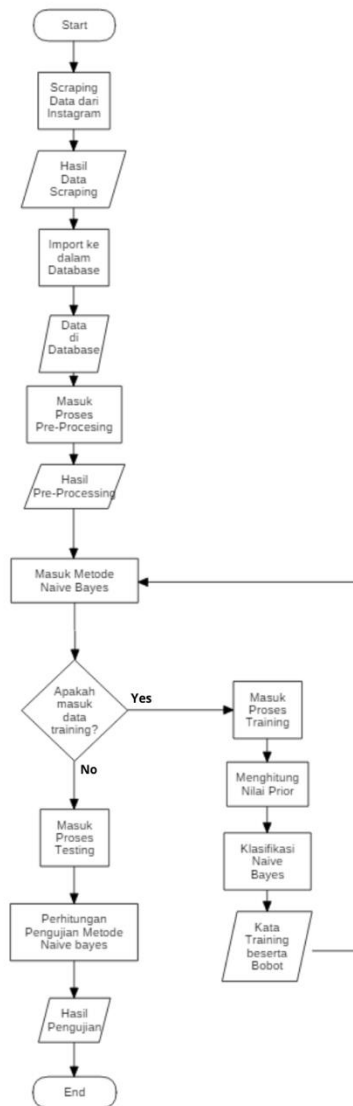
Simple scraper adalah sebuah *plugin google chrome* yang dapat digunakan untuk melakukan scraping pada sebuah halaman web. *Simple Scraper* tidak memerlukan pendaftaran karena dapat berjalan di browser tanpa menggunakan kode apa pun. Untuk melakukan proses scraping pada *simple scrap*, halaman yang akan diambil datanya diakses via *google chrome* kemudian mengaktifkan *plugin* dan memilih *section* yang akan diambil pada halaman web tersebut.

B. Teknik Analisis Data

Adapun teknik analisis data pada penelitian ini dikumpulkan melalui studi pustaka yang diperoleh dari buku-buku, literatur, referensi-referensi, sumber dari internet atau sumber lain yang diperlukan untuk merancang dan mengimplementasikan sistem yang berkaitan dengan penulisan skripsi yang dilakukan.

C. Flowchart Sistem

Flowchart Sistem Analisis Sentimen dan Klasifikasi Naïve Bayes



Gambar 1. Flowchart Sistem Analisis Sentimen dan Klasifikasi Naïve Bayes

Berikut penjelasan dari flowchart penelitian sebagai berikut:

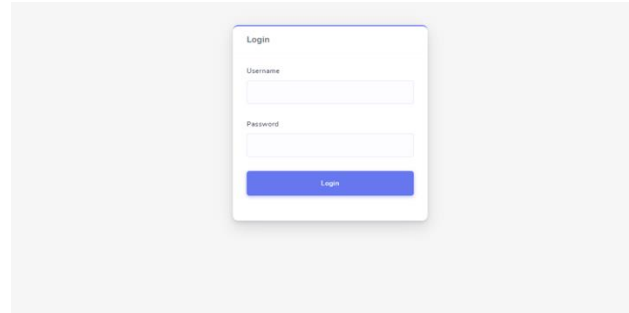
- 1) Memasukkan data komentar
Pada tahap ini, data yang telah diperoleh dari Simple Scraper dan sudah dilabeli dalam bentuk excel akan dimasukkan atau diimport ke dalam database. Data akan dipisah sesuai dengan kelas positif, negatif atau netral.
- 2) Melakukan proses pre-processing
Pada tahap ini, data pada database akan melalui proses-proses berikut:
 - a) Case folding yaitu mengubah semua huruf dalam dokumen menjadi huruf kecil.
 - b) Data cleaning yaitu proses pembersihan kata dengan menghilangkan semua tanda baca seperti koma (,), titik (.), at 9@), tagar (#) dan tanda baca lainnya.
 - c) Stop removal yaitu proses untuk menghapus kata-kata yang sering muncul dan bersifat umum, kurang menunjukkan relevansi dengan teks.
 - d) Stemming yaitu proses untuk mencari stem (kata dasar) dari kata hasil stopword removal (filtering), dan
 - e) Tokenizing yaitu tahap pemotongan string input berdasarkan tiap kata yang menyusunnya.
- 3) Melakukan klasifikasi Naïve Bayes
Menghitung probabilitas setiap kata terhadap semua dataset yang tersedia.
- 4) Melakukan pengujian
Melakukan pengujian *Confusion Matrix* yang digunakan untuk menguji akurasi, presisi, *recall* dan *F1-score* dari hasil klasifikasi.

III. Hasil dan Pembahasan

Pada Hasil penelitian ini akan membahas hasil yang didapatkan selama penelitian berlangsung. Data yang digunakan untuk penelitian ini adalah dari data komentar di Instagram Shopee dari Event Big Sale 11.11. Data komentar yang digunakan dengan total 450 data dengan 225 Data Training dan 225 Data Testing. Berikut dibawah ini penulis sajikan tampilan dari implementasi halaman.

A. Implementasi Tampilan Halaman

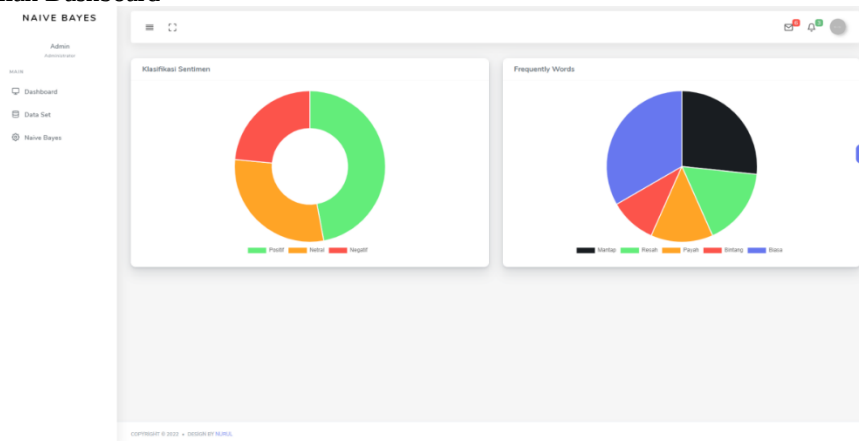
1) Tampilan Login



Gambar 1. Tampilan Login

Pada gambar 1. Halaman login merupakan halaman dimana untuk portal hak akses dari akun admin agar dapat masuk ke halaman admin sebagai pengelola website.

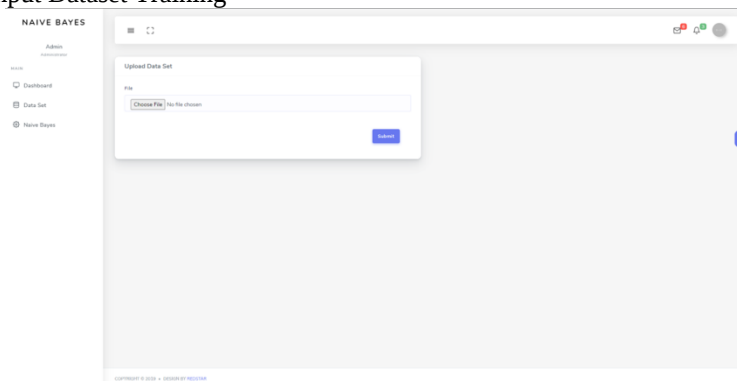
2) Tampilan Dashboard



Gambar 2. Tampilan Dashboard

Pada gambar 2, halaman dashboard dijadikan portal utama dari sistem klasifikasi naïve bayes.

3) Tampilan Input Dataset Training



Gambar 3. Tampilan Input Dataset Training

Pada gambar 3. Halaman *input dataset training* merupakan tempat untuk *export excel data training* agar data komentar dapat masuk kedalam database.

4) Tampilan List Dataset

NAIVE BAYES

Admin Administrator

MAIN

- Dashboard
- Data Set
- Naive Bayes

List Data Set

Show: 10 entries

Search

Nomor	Komentar	Kelas
1	Kalau aku sich suka smua min apa lg banyak promonya trsmn smng dech	Positif
2	Semua brand,min,acal ngasih promo gedg2an...	Positif
3	Tertarik gara" harga normal nya Trun dratis bukan yg di diskon nya Mantuauul Abis, kya iphone Se yg harga nya jutaan jadi 111k	Positif
4	shopee selalu di hati	Positif
5	Klu sy mau gabung jd reseller bisa GK ya	Netral
6	Min, kapan yg SelfieShopeeSKZ di unumumn???	Netral
7	pengen beli Adidas.cita ukuran 44 - auto skip	Netral
8	JANGAN YG MENANG BOT JUGA	Negatif
9	Gimana nih shopee.uth begadang malah ad org yg pale bot	Negatif
10	PANTESAN GW SELALU KALAH CEPET PMS FLASH SALE AJG LAH SISTERHAU KEK TN BSA DIBOL PAKE BOT. BUAT YG GAK TAU NI LGA KESEBAR VIDEO ORNG2 YG MAKE SITEM BOT BUAT DAPETN BARANG FLASH SALE BHAR CPT. ALIAS NGE HACK PAKE PC.ITU SIA SIA BEGADANG MLU	Negatif

Showing 1 to 10 of 10 entries

Previous 1 Next

Copyright © 2023 - DESIGN BY REDOTAR

Gambar 4. Tampilan List Dataset

Pada gambar 4. Halaman Data *Training* merupakan halaman untuk melihat list data *training* yang sudah ada di database.

5) Tampilan Input Dataset Testing

NAIVE BAYES

Admin Administrator

MAIN

- Dashboard
- Data Set
- Naive Bayes

Upload Data Set Testing

File

Choose File No file chosen

Submit

Copyright © 2023 - DESIGN BY REDOTAR

Gambar 5. Halaman Input Dataset Training

Pada gambar 5. Halaman input *Dataset Testing* merupakan tempat untuk *export excel* data training agar data komentar dapat masuk kedalam database.

6) Tampilan List Dataset Testing

NAIVE BAYES

Admin Administrator

MAIN

- Dashboard
- Data Set
- Naive Bayes

List Data Set Testing

Show: 10 entries

Search

Nomor	Komentar	Kelas
1	Kalau aku sich suka smua min apa lg banyak promonya trsmn smng dech	Positif
2	Semua brand,min,acal ngasih promo gedg2an...	Positif
3	Tertarik gara" harga normal nya Trun dratis bukan yg di diskon nya Mantuauul Abis, kya iphone Se yg harga nya jutaan jadi 111k	Positif
4	shopee selalu di hati	Positif
5	Klu sy mau gabung jd reseller bisa GK ya	Netral
6	Min, kapan yg SelfieShopeeSKZ di unumumn???	Netral
7	pengen beli Adidas.cita ukuran 44 - auto skip	Netral
8	JANGAN YG MENANG BOT JUGA	Negatif
9	Gimana nih shopee.uth begadang malah ad org yg pale bot	Negatif
10	PANTESAN GW SELALU KALAH CEPET PMS FLASH SALE AJG LAH SISTERHAU KEK TN BSA DIBOL PAKE BOT. BUAT YG GAK TAU NI LGA KESEBAR VIDEO ORNG2 YG MAKE SITEM BOT BUAT DAPETN BARANG FLASH SALE BHAR CPT. ALIAS NGE HACK PAKE PC.ITU SIA SIA BEGADANG MLU	Negatif

Showing 1 to 10 of 10 entries

Previous 1 Next

Copyright © 2023 - DESIGN BY REDOTAR

Gambar 6. Tampilan List Dataset Testing

Pada gambar 6. Halaman Data *Testing* merupakan halaman untuk melihat list data *testing* yang sudah ada di database.

7) Tampilan Pengujian Metode Naïve Bayes

The screenshot shows a web application titled 'NAIVE BAYES' with a sidebar menu containing 'Admin', 'Dashboard', 'Data Set', and 'Naive Bayes'. The main content area is titled 'List Data Set Testing' and displays a table with 10 rows of data. Each row contains a 'Nomor' (Number), a 'Komentar' (Comment), and three columns for sentiment classification: 'ARGMAX Positif', 'ARGMAX Netral', and 'ARGMAX Negatif'. The 'Hasil' (Result) column shows the predicted sentiment for each comment. The comments are related to a Shopee sale event.

Nomor	Komentar	ARGMAX Positif	ARGMAX Netral	ARGMAX Negatif	Hasil
1	Kalau aku scth sctk semua min apa lg brpki promosi kemnang shang dech	0.625	0.082	0.094	Positif
2	Semua brand min jual ngasih promo gnde2an...	0.625	0.082	0.094	Positif
3	Telarang gara" harga normal nya Trun dradits bukan yg di diskon nya Martasauli Ahihi, kya sphone Se yg harga nya jutaan jdi 111k	0.625	0.082	0.094	Positif
4	shopee selalu di hat	0.625	0.082	0.094	Positif
5	Klu ny mau gabung jd reseller bisa GK ya	0.082	0.625	0.094	Netral
6	Mis, kapan yg SeftaShopeeGZ di umumkan???	0.082	0.625	0.094	Netral
7	pengen beli Adidas jute ukuran 44 x auto ship	0.082	0.094	0.625	Netral
8	JANGANQ YG MENANG BOT JUGA	0.625	0.082	0.094	Negatif
9	Gimana nih shopee jd begabung masin ad yng yg pale bot	0.082	0.094	0.625	Negatif
10	PANTEGAN GW SELALU KALAH CEPET RAG PLASH SALE AUG LAH SISTERANU KEX TN BSA DIBOROL PAKE BOT BUAT YG GAK TRU NI UDA KESEBAR VIDEO ORNG2 YG NAKIL SOTEN BOT BUAT DAPETIN BARRING PLASH SALE BARI CPT ALIAS NGK HACK PAKE PC??? SA DA BEGADING NGU	0.082	0.094	0.625	Negatif

Gambar 7. Tampilan Pengujian Metode Naïve Bayes

Pada gambar 7. Halaman pengujian merupakan halaman proses klasifikasi *naive bayes* dimana setiap kata training dan testing menjadi satu dan dapat dilihat hasil dari klasifikasinya.

B. Pengujian Confusion Matrix

Setelah sistem desain dan dianalisis secara rinci, maka akan menuju tahap pengujian sistem. Pengujian program untuk metode Naïve Bayes ini menggunakan Confussion Matrix. Diawali dengan melakukan pembobotan kata dengan TF/IDF kemudian pembentukan model untuk menentukan peluang positif, netral, dan negatif yang digunakan untuk perbandingan data uji dengan *Naive Bayes*. Setelah itu membandingkan setiap kata dari data uji dengan data latih kemudian diambil peluang positif, netral, dan negatif dari setiap kata. Lalu dihitung setiap peluang yang mendapatkan nilai tertinggi. Berikut dibawah ini merupakan hasil dari klasifikasi komentar tersebut. Dibawah ini merupakan tabel dari Multiclass *Confussion Matrix*.

	Predicted Positif	Predicted Netral	Predicted Negatif
Actual Positif	TP	FNeP	FNP
Actual Netral	FNNE	TNe	FPNe
Actual Negatif	FPN	FNeN	TN

Keterangan :

- TP : True Positif
- TNe : True Netral
- TN : True Negatif
- FNeN : False Netral Negatif
- FPN : False Positif Negatif
- FNNe : False Negatif Netral
- FPNe : False Positif Netral
- FNP : False Negatif Positif
- FNeP : False Netral Positif

Kemudian dari data yang sudah masuk klasifikasi *Naive Bayes*, maka data di akumulasi berdasarkan tabel diatas sebagai berikut:

	Predicted Positif	Predicted Netral	Predicted Negatif
Actual Positif	4	1	0
Actual Netral	0	2	3
Actual Negatif	0	3	3

Pada tabel 18. Setelah melakukan pengujian didapatkan nilai true negatif dengan nilai 3 yang berarti berhasil mengklasifikasikan 3 komentar negatif dengan benar, true netral dengan nilai 2 berarti terdapat 2 komentar yang diklasifikasikan sebagai komentar netral.

Berikut dibawah ini merupakan rumus untuk menghitung akurasi, presisi, *recall*, dan *f1-score*, untuk perhitungan akurasi menggunakan rumus (4)

$$Akurasi = \frac{TP+TNe+TN}{TP+TNe+TN+FNeN+FPN+FNNe+FPNe+FNP+FNeP} \times 100\% \quad (4)$$

$$Akurasi = \frac{4 + 2 + 3}{4 + 2 + 3 + 3 + 0 + 0 + 3 + 0 + 1} \times 100\%$$

$$= \frac{9}{16} \times 100\% \\ = 56\%$$

Maka didapatkan nilai akurasi sebesar 56%, untuk perhitungan presisi menggunakan rumus (5).

$$\begin{aligned} \text{Presisi} &= \frac{TP}{TP+FN+FP} \times 100\% \\ \text{Presisi} &= \frac{4}{4+1+0} \times 100\% \\ &= \frac{4}{5} \times 100\% \\ &= 80\% \end{aligned} \quad (5)$$

Maka didapatkan nilai presisi sebesar 80%. Untuk perhitungan recall menggunakan rumus (6).

$$\begin{aligned} \text{Recall} &= \frac{TP}{TP+FN+FP} \times 100\% \\ \text{Recall} &= \frac{4}{4+0+0} \times 100\% \\ &= \frac{4}{4} \times 100\% \\ &= 100\% \end{aligned} \quad (6)$$

Maka didapatkan nilai *recall* sebesar 100%. Untuk perhitungan *f1-score* menggunakan rumus (7)

$$\begin{aligned} \text{F1-Score} &= 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \times 100\% \\ \text{F1-Score} &= 2 \times \frac{80\% \times 100\%}{80\% + 100\%} \times 100\% \\ \text{F1-Score} &= 2 \times \frac{0,8 \times 1}{0,8 + 1} \times 100\% \\ &= 89\% \end{aligned} \quad (7)$$

Maka didapatkan nilai *f1-score* sebesar 89%.

C. Pembahasan

Berikut di bawah ini merupakan hasil dari nilai presisi, recall, dan *f1-score* dari pengujian Confusion Matrix.

	Predicted Positif	Predicted Netral	Predicted Negatif
Presisi	50%	40%	80%
Recall	50%	33%	100%
F1-score	50%	36%	89%

Hasil dari akurasi untuk presisi prediksi Negatif sebesar 50%, untuk prediksi Netral sebesar 40% dan Positif sebesar 80%. Hasil dari akurasi untuk *Recall* prediksi Negatif sebesar 50%, untuk prediksi Netral sebesar 33% dan Positif sebesar 100%. Hasil dari akurasi untuk *f1-Score* prediksi Negatif sebesar 50%, untuk prediksi Netral sebesar 36% dan Positif sebesar 89%. Berdasarkan hasil *testing* yang sudah diterapkan pada aplikasi, bahwa algoritma *Naive Bayes* cukup berhasil memprediksi kategori sentimen yang benar karena hasil uji akurasi *Naive Bayes* mendapatkan akurasi yaitu 56% yang artinya *Naive Bayes* mengklasifikasikan data teks dengan sangat baik.

IV. Kesimpulan dan saran

Kesimpulan pada penelitian ini yaitu sistem Analisis Sentimen pada event Big Sale 11.11 Shopee dengan menggunakan klasifikasi *Naive Bayes* mampu menganalisa sentimen komentar pada Instagram dari Akun Shopee berdasarkan kategorinya yaitu Positif, Netral, dan Negatif. Proses klasifikasi dimulai dari pre-processing yang dimana terdiri dari Case Folding, Data Cleaning, Stopword Removal, Stemming, dan Tokenizing. Setelah melakukan pre-processing, maka data akan masuk ke training metode *Naive Bayes* dengan cara klasifikasi, menentukan nilai prior probability, dan menentukan nilai probabilitas setiap kata. Setelah itu maka masuk testing atau pengujian. Kedua ialah berdasarkan pengujian dari perbandingan nilai dari aplikasi dengan penghitungan manual, 257 dari 450 data memiliki kecocokan dan pengujian menggunakan Metode *Naive Bayes* ini menunjukkan 56% akurat. Hal tersebut menjadi nilai kecocokan yang cukup tinggi dikarenakan pada akun Shopee terutama pada postingan Instagram tentang event Big Sale 11.11 Shopee lebih banyak menggunakan kalimat netral. Yang terakhir ialah sistem ini dapat dijadikan rekomendasi kepada perusahaan terkait tentang opini masyarakat yang diklasifikasi menggunakan *Naive Bayes* sehingga menjadi tolak ukur bagi akun Shopee yang diharapkan membantu perkembangan perusahaan.

Berdasarkan penelitian yang telah dilakukan, beberapa saran dapat acuan untuk pengembangan penelitian kedepannya diantaranya ialah data yang digunakan untuk data training lebih banyak lagi agar tingkat akurasi lebih tinggi. Selain itu data training yang diberikan dari kalimat Positif, Netral, dan Negatif disamakan agar nilai prior yang ditampilkan tidak terjadi perbedaan yang signifikan. Nilai prior yang berbeda

signifikan akan mempengaruhi hasil dari pengujian tersebut, seperti terjadi perbedaan hasil klasifikasi. Kedua ialah pengambilan data saat ini masih menggunakan aplikasi pihak ketiga yang kemudian data tersebut akan terunduh dalam bentuk file excel. Untuk peneliti selanjutnya dapat menggunakan pengambilan data dalam aplikasi sendiri, sehingga peneliti tidak perlu unduh dan unggah file excel lagi. Yang terakhir yaitu lebih informatif kepada masyarakat akan adanya hasil klasifikasi dari komentar pada akun Shopee, karena hal tersebut mengacu pada perkembangan dari event dan perusahaan akan mengetahui event tersebut berjalan dengan baik atau tidak. lagi.

Daftar Pustaka

- [1] I. N. Kabiru and P. K. Sari, "Analisa Konten Media Sosial E-commerce Pada Instagram Menggunakan Metode Sentiment Analysis Dan Lda-based Topic Modeling (studi Kasus: Shopee Indonesia)," *eProceedings Manag.*, vol. 6, no. 1, pp. 12–19, 2019.
- [2] W. Prihatiningsih, "Motif Penggunaan Media Sosial Instagram Di Kalangan Remaja," *Communication*, vol. 8, no. 1, p. 51, 2017, doi: 10.36080/comm.v8i1.651.
- [3] Y. Puspitarani, "Sentimen Analysis terhadap Nilai Kepercayaan sebuah Online Shop di Instagram," *JITTER*, vol. II, no. 1, pp. 76–81, 2015.
- [4] W. A. Luqyana, I. Cholissodin, and R. S. Perdana, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4704–4713, 2018.
- [5] G. B. Prasetyo, H. E. Dien, D. Afif, and R. Prasetyo, "Instagram Dan Facebook Menggunakan Metode Naive Bayes Dan Support Vector Machine," *Semin. Inform. Apl. Polinema*, pp. 336–340, 2020.
- [6] A. Fauzi, M. F. Akbar, and Y. F. A. Asmawan, "Sentimen Analisis Berinternet Pada Media Sosial dengan Menggunakan Algoritma Bayes," *J. Inform.*, vol. 6, no. 1, pp. 77–83, 2019, doi: 10.31311/ji.v6i1.5437.
- [7] H. F. Putro, R. T. Vlandari, and W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *J. Teknol. Inf. dan Komun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [8] E. Manalu, F. A. Sianturi, and M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan Pada CV. Papadan Mama Pastries," *J. Mantik Penusa*, vol. 1, no. 2, pp. 16–21, 2017.
- [9] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creat. Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2015.
- [10] M. F. Rahman, D. Alamsah, M. I. Darmawidjadja, and I. Nurma, "Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN)," *J. Inform.*, vol. 11, no. 1, p. 36, 2017, doi: 10.26555/jifo.v11i1.a5452.