



Pematangsiantar, 06 Agustus 2023

No : 007/SRT-LOA/KESATRIAVOL4NO2/VI/2023

Hal : Surat Penerimaan Naskah Publikasi Artikel Ilmiah

Kepada Yth:

Bapak/Ibu Penulis (Author)

Putri Rizki Aprilensia

Di Tempat

Handwritten signature and date: 20/8/2023

Assalamualaikum Wr. Wb

Salam Sejahtera

Terimakasih telah mengirimkan artikel ilmiah untuk diterbitkan pada KESATRIA:

Jurnal Penerapan Sistem Informasi (Komputer & Manajemen), ISSN Online : 2722-993X dengan Judul:

Analisis Sentimen Ulasan Pengguna Klik Indomaret Pada Google Play Store Menggunakan RapidMiner

Berdasarkan hasil *review*, artikel tersebut dinyatakan **DITERIMA** untuk dipublikasikan di Jurnal kami **Volume 4, Nomor 2, Oktober 2024**. Kami akan mengirimkan *softcopy* edisi tersebut pada akhir bulan penerbitan ke email penulis. Artikel tersedia secara online di <https://tunasbangsa.ac.id/pkm/index.php/kesatria>.

Berikut adalah beberapa hal penting yang kami ingin anda lakukan sehubungan dengan penerimaan paper tersebut:

- Mohon dilengkapi data *copyright transfer form* (CTF) yang tersedia di web jurnal.
- Konfirmasikan pengembalian *copyright transfer form* (CTF)

Anda dengan mengirim melalui email: agus.perdana@amiktunasbangsa.ac.id dan

lakukan konfirmasi melalui nomor whatsapp 089529766834 dengan subjek:

Proses pengiriman CTF sudah dilakukan.

Hormat kami,

Handwritten signature of Dr. Tutut Herawan
KESATRIA
Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)

Dr. Tutut Herawan
Editor In Chief KESATRIA



ANALISIS SENTIMEN ULASAN PENGGUNA KLIK INDOMARET PADA GOOGLEPLAY STORE MENGGUNAKAN RAPIDMINER

Kurniawan¹, Putri Rizki Aprilensia², Fatmasari³, Merrieayu Puspita Hannah⁴

[#]Sistem Informasi, Universitas Bina Darma

Palembang

kurniawan@binadarma.ac.id, aprilensiapr@gmail.com, fatmasari@binadarma.ac.id,
merrieayu@binadarma.ac.id

Klik Indomaret is an online shopping platform in Indonesia that offers a variety of products at competitive prices that allow users to shop for various product categories, including food, drinks, fashion, health and beauty products, home and others. Klik Indomaret also offers travel tickets and various virtual products. This platform provides a safe and comfortable online shopping experience with various payment methods, including bank transfers, cash on delivery, internet banking, electronic money, credit cards, online credit/debit cards, credit card installments, and payment points at Indomaret stores in throughout Indonesia. Based on data on the Google Play Store site, the Klik Indomaret application has been downloaded by approximately 5 million downloads with a rating of 4.5 and recorded 130 thousand user reviews of the Klik Indomaret application. The aim of this research is to conduct sentiment analysis on user reviews of the Klik Indomaret application using the Naive Bayes algorithm. Based on the research carried out, the interpretation of the results of naive bayes research with an accuracy rate of 49.97%, a precision rate of 49.98% and a recall rate of 99.81% and a specification of the value is 1.68% from 270 positive sentiments and 122 negative sentiments, and the positive reviews contain "easy words how the message and the goods are always there" while the negative reviews contain "severe symptoms when other applications are running smoothly". The results of the classification of the Naive Bayes method with the sum of 270 positive sentiments and 122 negative sentiments. Value accuracy 49.97%, margin of error 5.33%, positive prediction precision 49.98%, negative prediction precision 76.03%, positive data recall 99.81%, and negative data recall 76.03%

Keywords: Sentiment Analysis, Naive Bayes, Google Play Store

Klik Indomaret yaitu platform belanja online di Indonesia yang menawarkan berbagai macam produk dengan harga bersaing yang memungkinkan pengguna berbelanja berbagai kategori produk, termasuk makanan, minuman, fashion, produk kesehatan dan kecantikan, rumah dan lainnya. Klik Indomaret juga menawarkan tiket perjalanan, dan berbagai produk virtual. Platform ini memberikan pengalaman belanja online yang aman dan nyaman dengan berbagai metode pembayaran, antara lain transfer bank, cash on delivery, internet banking, uang elektronik, kartu kredit, kartu kredit/debit online, cicilan kartu kredit, dan titik pembayaran di toko Indomaret di seluruh Indonesia. Berdasarkan data pada situs Google Play Store aplikasi Klik Indomaret telah di download kurang lebih 5 juta unduhan dengan rating 4,5 dan tercatat 130 ribu ulasan komentar pengguna aplikasi Klik Indomaret. Tujuan penelitian ini adalah melakukan

analisis sentimen terhadap ulasan pengguna aplikasi Klik Indomaret dengan algoritma Naive Bayes. Berdasarkan penelitian yang telah dilakukan interpretasi hasil penelitian naive bayes dengan tingkat akurasi 49,97% , tingkat presisi 49,98% dan tingkat recall 99,81% dan spesifikasi nilainya 1,68% dari 270 sentimen positif dan 122 sentimen negatif ,dan ulasan positif berisi "perkataan gampang cara pesannya dan barangnya selalu ada" sedangkan ulasan negatif berisikan "lemot parah padahal aplikasi lain lancar aja". Hasil klasifikasi metode Naive Bayes dengan jumlah 270 sentimen positif dan 122 sentimen negatif. Nilai akurasi 49,97%, margin error 5,33%, presisi prediksi positif 49,98%, presisi prediksi negatif 76,03%, recall data positif 99,81%, dan recall data negatif 76,03%

Kata Kunci : Analisis Sentimen, Naive Bayes, Google Play Store

I. PENDAHULUAN

Indomaret juga menyediakan aplikasi yang disebut "Klik Indomaret yaitu platform belanja online di Indonesia yang menawarkan berbagai macam produk dengan harga bersaing. Platform ini tersedia sebagai aplikasi seluler di Google Play Store dan App Store, dan memungkinkan pengguna berbelanja berbagai kategori produk, termasuk makanan, minuman, fashion, produk kesehatan dan kecantikan, rumah dan lainnya. Klik Indomaret juga menawarkan tiket perjalanan, tiket perjalanan, dan berbagai produk virtual. Platform ini memberikan pengalaman belanja online yang aman dan nyaman dengan berbagai metode pembayaran, antara lain transfer bank, cash on delivery, internet banking, uang elektronik, kartu kredit, kartu kredit/debit online, cicilan kartu kredit, dan titik pembayaran di toko Indomaret di seluruh Indonesia. Klik Indomaret memiliki wilayah pengiriman luas yang mencakup banyak titik kota yang ada di Indonesia.

Berdasarkan data pada situs Google Play Store aplikasi klik indomaret telah di download kurang lebih 5 juta unduhan dengan rating 4,5 dan tercatat 130 ribu ulasan komentar pengguna aplikasi klik indomaret. Menurut Winarko (2015) Ada dua aspek review yang dipertimbangkan, yaitu rating nilai dan komentar lebih mendalam. Rating nilai menunjukkan suatu nilai angka, sedangkan komentar

lebih tekstual berfokus pada pendapat lebih mendalam. Oleh karena itu dibutuhkan analisis sentimen terhadap ulasan pengguna aplikasi klik indomaret dengan algoritma naive bayes. Analisis sentimen adalah proses mengolah data tekstual secara otomatis untuk mengetahui apakah opini yang terkandung dalam suatu kalimat adalah positif atau negatif dan Algoritma Naive Bayes Classifier dianggap cocok untuk digunakan dalam analisis sentimen karena bertujuan sebagai metode klasifikasi kedalam kategori positif dan negatif.

II. LANDASAN TEORI

Pada penelitian Permatasari et al (2021) mengatakan dalam analisis sentiment dapat digunakan beberapa teknik pembelajaran. Diantara beberapa metode pembelajaran tersebut, Naive Bayes Classifier dinilai sebagai metode yang memiliki kemampuan yang baik dalam melakukan klasifikasi data dibandingkan metode klasifikasi lainnya terutama dalam hal akurasi. Metode ini dapat digunakan untuk menganalisis data berbasis teks yang sangat besar. Analisis sentimen juga dapat melibatkan pembelajaran bahasa dengan menggunakan Part of Speech (POS) Tagging. Metode ini memberikan label (tag) kata-kata pada sebuah kalimat. Berikut gambar pengelompokan klasifikasi data dari sistem analisis sentimen



Gambar 2.1 Klasifikasi data system analisis *sentiment*

Teknik Analisis Sentimen *Mearching learning*

termasuk ke dalam Supervised Learning dalam metode Machine Learning. Sistem akan mengklasifikasi nilai sentimen berdasarkan hasil pembelajarannya pada data latih (training data). Metode ini juga tidak terlepas dari teknik Natural Language Processing (NLP) agar sistem dapat memahami bahasa alami manusia.

Dengan metode Machine Learning, output yang dihasilkan tidak tertutup pada positif, negatif, dan netral saja. Anda dapat membangun modelnya sesuai kondisi yang diinginkan. Output dapat dikelompokkan menjadi sentimen “sangat positif, positif, netral, negatif, dan sangat negatif”, “marah, senang, dan sedih”, atau yang lainnya.

Sebagai contoh, kali ini kita akan membangun model Analisis Sentimen untuk menganalisis ulasan produk pada situs google playstore. Tujuan dari pembangunan model ini yaitu untuk mengklasifikasi ulasan-ulasan dari suatu produk ke dalam kategori positif, negatif, atau netral.

2.5 Metode Data Mining

Metode dalam data mining adalah knowledge in database (KDD), . Setiap metode mempunyai kelebihan-masing yang dapat diterapkan sesuai dengan analisis masalah dan hasil yang diharapkan (saputra,2022).

Tahapan data selection

Web Scraping

Web scraping adalah proses mengumpulkan dan memproses data dari website secara otomatis. Proses ini biasanya dilakukan oleh program atau bot yang disebut crawler, yang akan mengunjungi situs web dan mengumpulkan data dari halaman web tersebut. Setelah data diumpulkan, crawler akan mengolah dan mengumpulkan data tersebut ke dalam suatu database atau format lain untuk digunakan secara lanjutan. *Web scraping* digunakan dalam berbagai aplikasi, seperti web indexing, web mining, data mining, monitoring harga online, pengumpulan data produk, mengumpulkan data weather, mengidentifikasi perubahan website, penelitian, dan mengawasi online presence dan reputasi.

Web scraping dapat dilakukan dengan menggunakan bahasa pemrograman seperti Python, yang memiliki beberapa paket yang dapat digunakan untuk web scraping. Proses web scraping melibatkan beberapa langkah, mulai dari mengakses URL situs web, membaca dan memproses *HTML*, mengumpulkan data, dan menyimpan data tersebut dalam format yang diinginkan.

Tahap Preprocessing

Text processing

Data set harus melalui tahap text prprocessing terlebih dahulu karena data set tidak bisa digunakan tanpa melalui tahap pengelolaan data. *Text preprocessing* adalah suatu proses pengelolaan data set yang tidak bersih seperti kesalahan sistem saat pencatatan sehingga terjadinya data duplikat . data yang belum diolah atau data tidak bersih kategorinya seperti format data yang tidak beraturan , adanya data kosong , tipe data yang berbeda-beda, adanya atribut yang tidak penting, dan lain sebagainya .semakin bersih pra proses yang dilakukan , maka kemungkinan besar hasil data tersbeut semakin akurat (irwansyah saputra, 2023). Terdapat tahapan text preprocessing, sebagai berikut (rahman isnain et al., 2021):

A. cleaning

cleaning merupakan tahapan untuk membersihkan data pada set dan meyeleksi kata yang tidak diperlukan , tidak memiliki arti ,atau arti yang memengaruhi sentiment seperti html, link, mention dan hastag.

B. Tokenize

Tokenize adalah pemecahan data set ke dalam bentuk token atau potongan kata agar mempermudah tahapan selanjutnya. Contohnya

kalimat 'syaa ingin makan' dipenggal menjadi [;saya', 'ingin', 'makan'] (Baihaqi et al., 2020).

C. Transform cases

Tahapan merubah kalimat data teks menjadi teks yang seragam .tahapan ini selalu ada dalam proses text preprocessing karena data yang ada tidak selalu terstruktur dalam penggunaan hurufnya. Dengan adanya tahapan ini dapat berperan dalam penyamarataan penggunaan huruf kapital. Sebagai contoh "Data" dan "data" akan terbaca sebagai dua kata yang berbeda , sehingga melalui proses ini sistem dapat membaca secara efektif.

D. Stopward removal

Tahapan ini untuk menghapus kata yang sering muncul tapi tidak memiliki arti penting dan maknanya tidak berpengaruh pada sistem, seperti "oh", "di", "pada", dan sebagainya (taufiqurrahman et al.,2021).

E. Filter

Filter adalah tahapan untuk menghapus kata-kata yang terlalu pendek dan terlalu panjang dengan minimal 3 huruf dan maksimal 25 huruf (hajati et al.,2021).

Tahapan data minning

Naïve Bayes

Naïve Bayes adalah bagian dari teori keputusan Bayesian. Disebut Naive karena rumusnya membuat beberapa Naive Asumsi. Kemampuan pemrosesan teks Python yang membagi dokumen menjadi vektor digunakan. Ini dapat digunakan untuk mengklasifikasikan teks. Pengklasifikasian dapat dimasukkan ke dalam bentuk yang dapat dibaca manusia. Ini adalah metode klasifikasi yang populer selain kondisional independensi, overfitting, dan metode Bayesian (Kaviani, 2017).

Naïve Bayes adalah salah satu algoritma klasifikasi data. Algoritma Naive Bayes adalah klasifikasi probabilistik sederhana. Algoritma ini menghitung satu set probabilitas dengan menghitung frekuensi dan kombinasi nilai dalam kumpulan data tertentu. Probabilitas Fitur tertentu dalam data muncul sebagai anggota dalam urutan probabilitas dan diperoleh dengan menghitung frekuensi setiap nilai fitur di kelas dari pelatihan kumpulan data. Kumpulan data pelatihan adalah subset yang digunakan untuk melatih algoritma klasifikasi.

Evaluasi

Confusion Matrix

Confusion matrix merupakan suatu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem untuk diukur keakuratannya. Confusion matrix juga merupakan salah satu cara dalam melakukan visualisasi terhadap hasil pembelajaran sistem, visualisasi yang

ditampilkan memuat dua kategori atau lebih (Rahman, dkk., 2017).

Tabel di bawah merupakan contoh hasil confusion matrix prediksi dua kelas.

Nilai prediksi	Nilai actual	
	Positive	Negative
Positive	True positive	True Negative
Negative	False Positive	False Negative

- Tabel diatas merupakan tabel confusion matrix dengan keterangan sebagai berikut ;
 - 1) *TP (True positive)* = jumlah data nilai aktual negatif dan nilai prediksi kelas positif.
 - 2) *TN (True Negative)* = jumlah data nilai aktual negatif dan nilai prediksi negatif
 - 3) *FP (False Positive)* = jumlah data nilai aktual positif dan nilai prediksi negatif
 - 4) *FN (False Negative)* = jumlah data nilai aktual negatif dan nilai prediksi positif

Akurasi adalah nilai rasio data tweet yang sudah terdeteksi dalam pengujian .nilai akurasi dapat menunjukkan kedekatan antara nilai prediksi sistem dan prediksi manusia (Forest & bayes, 2021) ,

III. METODOLOGI PENELITIAN

Waktu dan Tempat Penelitian

Tabel 3.1 Waktu Pelaksanaan Penelitian

No.	Tahapan	Maret 2024	April 2024	Mei 2024	Juni 2024	Juli 2024
1	Data selection					
2	Preprocessing					
3	Data mining					
4	evaluasi					
5	Pembuatan laporan					

Keterangan: Dari table diatas, peneliti akan menjelaskan bahwa peneliti melakukan penelitian selama lima bulan pada tahun dua ribu dua puluh empat yang telah diuraikan pada table diatas.

1. Data selection : proses pengambilan data pada platform google play data yang diambil setelah itu data akan diklasifikasikan kedalam kelas negatif dan positif.
2. Preprocessing : data yang sudah ada

kemudian diolah pada tahapan preprocessing yang bertujuan untuk mengubah data mentah menjadi data terstruktur.

3. Data mining : proses mengklasifikasikan data dalam bentuk kelas negatif dan positif .

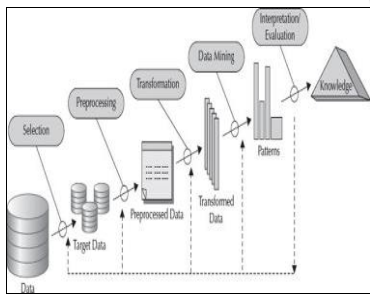
4. Evaluasi : tahapan ini merupakan melakukan perhitungan confusion matrix yang meliputi *accuracy*, *presisi* dan *recall*.

Populasi dan sampel penelitian

Penelitian ini menggunakan populasi dan sample berupa ulasan pada media sosial google play store dengan kata kunci klik indomaret.

Metodologi Penelitian

Metode penelitian ini memiliki suatu rancangan bagaimana alur sistem penelitian ini akan dirancang. Dalam penelitian ini peneliti menggunakan metode KDD (knowledge discovery in database). terdapat lima tahapan dalam kdd yaitu ,data selection ,preprocessing transformation, data mining dan *evaluation* (sabarmathi abd chnnaiyan ,2017). Tahapan KDD dalam kegiatan penelitian yang dilakukan terlihat seperti gambar 3.1



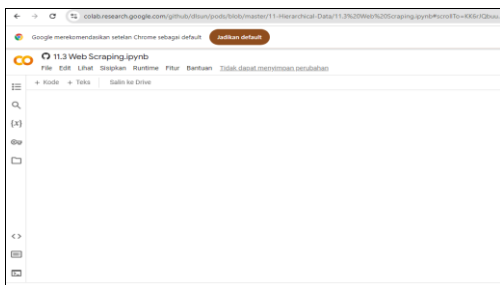
Gambar 3.2 Tahapan Penelitian KDD

Data Selection

Pada tahap ini dilakukan pengolahan data yang sudah diambil menggunakan teknik scraping pada web tools google scraping. data digunakan dalam penelitian ini adalah data ulasan yang ada di google play store. data yang dikumpulkan merupakan data teks yang diambil menggunakan teknik scraping menggunakan web scraping.

1. Scraping Data

Pada awal untuk melakukan proses pengambilan data dari google play store ,peneliti harus membuat codingan dari tools *web googlescraping.ipynb*.



Gambar 3.3 Halaman awal tools scraping data

Setelah masuk kedalam tools peneliti harus melakukan codingan untuk melakukan crawling data ulasan. Dibawah ini contoh codingan pada saat crawling data menggunakan *googlescraping.ipynb*

```

from google_play_scraper import app

import pandas as pd

import numpy as np

[5]: scrape_jumlah_ulasan yang diinginkan
from google_play_scraper import sort, review

result, continuation_token = review(
    "Indomaret:KlikIndomaret",
    lang="id", # Misal kita mau scrape data ulasan aplikasi shopee yang berada di google play store
    country="id", # Misal setting bahasa nya menjadi bahasa Indonesia
    sort=Sort.NEWEST, # # # kemudian kita gunakan sort_relevance untuk mendapatkan ulasan yang paling relevan
    count=50, # # # kemudian di filter_score kita gunakan None untuk mengambil semua score atau rating bintang 1 sampai 5
    filter_score=0 # # # kemudian di filter_score kita gunakan None untuk mengambil semua score atau rating bintang 1 sampai 5

df_husa = pd.DataFrame(np.array(result), columns=['review'])

df_husa = df_husa[df_husa['review'].str.contains('review', na=False)]
    
```

Gambar 3.4 Halaman coding tools

Hasil proses crawling data menggunakan tools *scraping.ipynb* maka mendapatkan data dengan jumlah 1000 ulasan. Untuk hasil data contoh pada excel dapat dilihat seperti gambar 3.5 dibawah ini

A	B	C	D	E	F
1	Ulasan	score	konten		
2	Riska Diah	5	2024-08-16 2:20 Memainkan dan kompres file JavaScript, CSS, dan HTML, untuk mempermudah ukurannya. Kompres gambar dan aset media lainnya tanpa mengura		
3	Agung Budi Win	1	2024-08-11 12:5 Kenapa ya ketika belanja di klik indomaret scan barcode barang belanjanya lama banget. Bahkan sering gagal. jadi lebih cepet kalo keyboard di pancaran dari paca scan ba		
4	DHARA PUSPA	1	2024-08-07 17:2 Tolong dong diperbaiki lagi apk nya. Dipencet ga bisa. Hapus belanjanya lama banget. Untuk itu2 yang ga bisa pergi belanja keluar sebenarnya sangat terbantu dgn adar		
5	Indah nur hafidha	1	2024-07-29 9:40 Untuk pesanan yg dipick up sendiri lemot seketik. untuk koin ke umum GG masalah maini kadang juga. Tapi masa ini prosesnya update lama. contoh pesan jam sepuluh dianti		
6	Ema Farida	1	2024-07-29 18:22 Belanja gampang, lemot lagi, banyak lag. Nyoba order stoki dan kayak gak mau lg. Mendang aplikasi kompetennya. Koneksi yg lebih बेवस्था baru dan salah satunya ada ham pro		
7	Arial Rudianto	1	2024-07-12 14:2 Kalaupun 1 jam sampai, ini udah sebulan kali order. nyampe nya 2 jam lebih. membeli Saya sdkt. Ternyata 3 jam lebih pesanya saya belum sampai juga. Parah banget. kalau belan		
8	Hamdan Sanjaya	1	2024-07-31 1:50 Masih bingung cara belanjanya. app lemot banget. Terus juga seharusnya tampilan produk yang dipilih untuk ambil offline. Saya di Jakarta, saya mau beli produk untuk orang y		
9	Quave Rayna	1	2024-08-11 6:30 Udah Zlatan lebih pake aplikasi ini. Sangat membantu terlihat lagi sama ini saya dapat kurir yg baik walaupun sudah banyak pengantaran kurir. Mungkin harus di tingkatkan lu		
10	Ria Indah Lestari	3	2024-08-14 11:30 Saya jeman dari yang, dan belum dikun hingga sekarang sudah pake 15.00 aplikasi sering jekel. pengiriman juga sering sama. alng pesannya saya bahkan belum diantar, dan li		
11	Dipa Vu	1	2024-07-31 23:33 Setelah update malah jadi error tiba tiba bisa dibuka. "kami mendeteksi aktivitas yg menyalahi ketentuan berlaku". Aplikasinya yg error malah nyalahi hp konsumen. hp nya sama p		
12	Mochamad Rii	1	2024-08-09 1:26 Niat bikin apk kaga sih fitur daftar aja lenyapnya minta ampun, belum lagi rethesh nya padahal hku realme 13 tapi masuk apk ini lag nya minta ampun apa masalah jaringan? Di		
13	Corneilus Praseli	1	2024-08-08 6:45 Aplikasi marketplace itu dibuat untuk mempermudah transaksi, bukan jadi mempersulit. Sy yg awalnya lebih suka berbelanja di indomaret, akhirnya lebih memilih alfabet sebag		
14	ahk horobito	3	2024-07-22 20:20 kenapa sekiranya aplikasi lemot banget kadang bug jka... discore2 ga langsung muncul. kadang bolak bolak keluar aplikasi. ditunggu sekiranya kalau barang kosong idan		
15	Nid indedy	1	2024-07-31 15:15 Kalau jelek semot lebih gampang banget. Tapi lemot n jadi ga bisa nungguin pesanan mau order. Untuk indomaret cibe diperbaiki lagi supaya pelayanan lebih nye		
16	Apa Nurhasanah	5	2024-06-18 4:48 Mohon diperbaiki lagi mengenai display barang yg ada pada "promo minggu ini". Kanarin sering sekal barang yg sama bisa muncul 2,3x ke. Udah mah nge bug untuk muncul i		
17	adi dolly	4	2024-07-15 16:14 bingung du ya... pesanan dari saya cuma di tambahkan fitur batasan pesanan mungkin dalam waktu 5-10 menit waktu untuk orang berfikir atau mungkin ada yang bermasala		
18	Anissa Rifqah A	1	2024-07-15 19:23 Pengiriman sering pake grab yg sering nanya lagi. Malahnya dimana. padahal udah dibuat bilanginnya benar benar. pengirimannya super super lama. ga on time sama sekali. info		
19	Zaki Yusari	1	2024-08-01 14:45 Sangat ter nya bebelang bag an app klik indomaret ini. Berarti instali app terus daftar dulu. mengali nomor hp dan kata sandi terus daftar. Di klik web sebandar dan app bisa digi		
20	denisa aha	3	2024-08-11 5:55 Parah lenyapnya. udah banyak yg ngirfu kalo aplikasinya lemot tapi gada perbaikan. apa thnya gak seanehngi bka sebandar atau komed2 diini? gak pernah kebaca sama Manag		
21	Safya Rapsa	1	2024-08-16 9:50 padahal udah banyak yang kasih ulasan terkait aplikasi yang super lemot, tapi tidak ada perubahan apapun. aplikasi malah saja lemot, yang tadinya mau belanja jadi ga mood i		
22	Shelan	5	2024-07-05 2:56 Aplikasinya bagus. kalau bisa sih metode pembayarannya yang selengkap itu bisa juga melakukan pembayaran lewat shoppe paylater biar enak belanjanya di klik indomaret		
23	achmad sidiq	1	2024-08-01 10:30) padahal lebih sering bisa di indomaret tapi ini gampang banget. itu aplikasi ini hudoan lebih mahal. kasusnya sama persis. di aplikasi klik komany tapi pan ka tidak ba		
24	Neneng Harjani	5	2024-07-01 23:22 Tolong di perbaiki kualitas koneksi nya. Setiap ya bisa aplikasi ini sryah! hp saya langsung hang dan loading nya lama banget padahal aplikasi yang lain lemot. 🙄 barang suda		
25	Cristi Tanya W	1	2024-07-30 8:09 Lemot banget dan sangat buruk klo dibanding aplikasi retail sebandar. Sistem callback cocok cash lg ga jalan. Dapin buat belanja langsung ke indomaret ternyata ga bisa dipake		
26	Nanasa Rukhaha	4	2024-07-05 7:43 ini aplikasi sangat membantu. sekali. bukannya fitur sama sekali tidak ada aplikasi ini alhamdulillah bawar. Tapi kenapa sampai susah uninstal? dan kenapa bisa ditambah blok di huda. 🙄 ada kelas		

Gambar 3.5 Hasil *scraping* data

2. Pelabelan Data

Hasil dari proses scraping data dalam file excel pada gambar 3,5 diatas yang nantinya akan dilakukan proses labeling data untuk menentukan klasifikasi pendapat atau pandangan dari hasil ulasan yang telah di *scraping* tadi. Pada proses labelling ini dibedakan menjadi 3 kelas yaitu class positif, class negatif dan class netral ,contoh dari labelling data seperti tabel 3.1 dibawah ini.

Tabel 3.1 Data Hasil Labelling

Class Positif	Class Netral	Class Negatif
Sangat membantu disaat engga bisa keluar lg gk enak badan	tolong diperbaiki, nglag dan lemot parah	Sudah berkali kali di update ga ada perbaikan tetap nge lag, ga malu kalian sama alfamart? Kasihan citra indomaret jadi buruk gegara kalian!..
Swangat membantu khususnya saya pribadi , yg kerus parah waktu . Susah untuk belanja langsung ke tokonya . Saya sangat suka Aplikasi ini sangat bagus tingkatan terus ya	Aplikasinya sangat lamaa saat melakukan pesanan dan mencari produk	Sudah berkali kali di update ga ada perbaikan tetap nge lag, ga malu kalian sama alfamart? Kasihan citra indomaret jadi buruk gegara kalian!..

Pelayanan baik...pesanan selalu lengkap...barang yg dikirim tidak ada yg exp	Belanja diklik indomaret dan langsung ke toko beda harganya, tdk sperti toko sebelah	Sudah berkali kali di update ga ada perbaikan tetap nge lag, ga malu kalian sama alfamart? Kasihan citra indomaret jadi buruk gegara kalian!..
Adanya belanja online di indomaret memudahkan smuanya. Free ongkir pula. Trimakasih indomaret semoga makin sukses dan banyak promonya	Pesan di klik indomaret loading bangett. Mohon untuk dapat ditingkatkan kembali agar app nya tidak loading	Tolong dong diperbaiki lagi apk nya. Dipencet2 ga bisa. Hapus belanjaan lemotnya ga karuan. Untuk ibu2 yang ga bs pergi belanja keluar sebetulnya sangat terbantu dgn adanya klik. Tapi semakin lama kok makin menurun kualitas nya aplikasi ini
pakai ini sejak jaman pandemi dan selalu memuaskan. promonya banyak dan sesuai sama toko. pengiriman oke. kemaren ada barang gak diproses tapi setelah komplain semua beres dan dikirim esok harinya. kalo ada brg habis. uang dikembalikan full pertahankan memudahkan ibu rumah tangga banget	Mudah dgunakan, banyak diskon n tebus murah yg menarik. Sayangnya loadingnya luumbaatttt. G ush salahin hp saya ya, aplikasi merah lancar jaya soalnya. Tp saya ttp lbh suka klik indomart meskipun klo pk aplikasi merah 15 menit barangnya sampai. Soalnya promonya lbh sesuai sm yg saya perlu	Aplikasi tambah lama tambah berat, hp udah bagus masih aja lelet, kebanyakan fitur jadi berat, buka klik food aja lemot parah.

Dalam kasus ini didalam class positif mendapatkan bintang lebih dari 3 yang artinya kata-kata tersebut tidak mengandung unsur hatespeech atau ujaran kebencian, sedangkan didalam class negatif mendapatkan bintang kurang dari 3 yang menyatakan bahwa komentar tersebut mengandung unsur hatespeech atau ujaran kebencian dan didalam class netral mendapatkan bintang 3 saja. Setiap data akan melalui proses preprocessing dengan melakukan perubahan bentuk data yang belum terstruktur menjadi terstruktur sesuai dengan kebutuhan seperti mengatasi kata yang berulang kata baku, kata asing, serta karakter yang tidak memiliki arti.

Preprocessing

Pada tahap ini dilakukan analisis data yang sudah diambil, melalui beberapa tahapan, berikut tahapan yang akan dilalui.

1. Tahapan proses preprocessing

Preprocessing adalah tahapan proses untuk membersihkan data dari kata-kata atau tweet yang tidak diperlukan serta kata-kata yang tidak memiliki makna. Proses ini dilakukan dengan isi data dari proses pengambilan data atau scraping data google playstore, adapun proses beberapa langkah dari proses preprocessing memiliki urutan sebagai berikut:

A. Cleaning

cleaning adalah proses penghapusan simbol, tanda baca, huruf kapital dan bilangan yang sering muncul pada tweet pengguna klik indomaret sehingga data tersebut menjadi data yang tidak efektif dan tidak

memiliki arti seperti : (#@\$%^&*<,./?). proses ini dijalankan menggunakan program, sehingga cleaning ini berjalan secara otomatis sebelum menyimpan hasil decode dalam bentuk file excel, peneliti membuat hasil data set dalam bentuk excel dikarenakan bahasa pemrograman python memiliki library yang menyediakan layanan baca dan menulis suatu data bertipe file csv atau excel, contoh penerapan proses cleaning dapat dilihat dari seperti tabel 3.2 di bawah ini.

Ulasan Sebelum Cleaning	Ulasan Sesudah Cleaning
Sangat membantu disaat engga bisa keluar lg gk enak badan	sangat membantu disaat engga bisa keluar lg gk enak badan
Lemoott buwanget aplikasinya, tidak dpt menampilkan halaman!	lemoott buwanget aplikasinya tidak dpt menampilkan halaman
Cukup membantu buat beli barang anak di pesantren	cukup membantu buat beli barang anak di pesantren
Sangat memudahkan juga harga lebih hemat daripada kudu ke toko langsung.	sangat memudahkan juga harga lebih hemat daripada kudu ke toko langsung
Belum ada pengalaman pake buat belanja	belum ada pengalaman pake buat belanja

B. Remove stopwords

Remove stopwords adalah proses penghapusan kata-kata yang kurang bermakna atau kata yang tidak memiliki arti seperti kata dan atau kamu, saya contohnya proses penerapan pada tahap stopwords dapat dilihat pada tabel 3.3 dibawah.

Tabel 3.3 Contoh Data Hasil Stopword

Ulasan Sebelum Remove Stopword	Ulasan Sesudah Remove Stopword
Sangat membantu disaat engga bisa keluar lg gk enak badan	sangat membantu disaat engga keluar gak enak badan
Lemoott buwanget aplikasinya, tidak dpt menampilkan halaman!	lemoott buwanget aplikasinya dpt menampilkan halaman
Cukup membantu buat beli barang anak di pesantren	cukup membantu beli barang anak pesantren
Sangat memudahkan juga harga lebih hemat daripada kudu ke toko langsung.	sangat memudahkan harga hemat ke toko langsung
Belum ada pengalaman pake	belum pengalaman pake

buat belanja	belanja
--------------	---------

C. Tokenization

Tokenization adalah proses untuk memecahkan kalimat untuk menjadi beberapa bagian yang dinamakan token, sebuah token dapat dianggap menjadi satu bentuk sebuah kata, frasa, atau suatu elemen yang berarti. Contohnya proses pada tahap tokenization dapat dilihat pada tabel 3.4 dibawah.

Tabel 3.4 Contoh Data Hasil Tokenization

Ulasan Sebelum Tokenization	Ulasan Sesudah Tokenization
Sangat membantu disaat engga bisa keluar lg gk enak badan	Sangat, membantu, disaat, engga, bisa, keluar, lg, gk, enak, badan,
Lemoott buwanget aplikasinya, tidak dpt menampilkan halaman!	Lemoott, buwanget, aplikasinya, tidak dpt, menampilkan halaman
Cukup membantu buat beli barang anak di pesantren	cukup, membantu, buat, beli, barang, anak, di pesantren
Sangat memudahkan juga harga lebih hemat daripada kudu ke toko langsung.	sangat, memudahkan, juga, harga, lebih hemat, daripada, kudu, ke toko, langsung
Belum ada pengalaman pake buat belanja	Belum, ada, pengalaman, pake, buat, belanja

D. filtering

filtering adalah Filter adalah tahapan untuk menghapus kata-kata yang terlalu pendek dan terlalu panjang dengan minimal 4 huruf dan maksimal 25 huruf. dapat dilihat pada tabel 3.5 dibawah.

Tabel 3.5 Contoh Data Hasil Stemming

Ulasan Sebelum filtering	Ulasan Sesudah filtering
Sangat, membantu, disaat, engga, bisa, keluar, lg, gk, enak, badan,	Sangat bantu saat engga bisa luar enak badan,
Lemoott, buwanget, aplikasinya, tidak dpt, menampilkan, halaman	Lemoott buwanget aplikasinya tidak nampilkan halaman
cukup, membantu, buat, beli, barang, anak, di pesantren	cukup bantu buat beli barang anak pesantren
sangat, memudahkan, juga, harga, lebih hemat, dari pada, kudu, ke toko, langsung	Sangat mudahkan juga harga lebih hemat dari pada kudu toko

	langsung
Belum, ada, pengalaman, pake, buat, belanja	Belum ngalaman, pake buat belanja

Transformation

Dalam kasus ini setelah semua ulasan dikumpulkan dari hasil crawling data tersebut telah melewati proses preprocessing, proses selanjutnya adalah membuat fitur yang berguna untuk mempermudah proses pengklasifikasian data tersebut, biasanya proses ini dibuat dengan proses ekstraksi fitur.

1. Klasifikasi

Dalam kasus ini peneliti menggunakan metode naive bayes classifier sebagai metode pengklasifikasian data ulasan untuk mendapatkan sentimen analisis untuk melakukan proses pengklasifikasian ini diperlukan hasil dari data yang sudah diolah dari proses sebelumnya yaitu hasil dari proses preprocessing dan hasil dari pembobotan kata setekah dat berhasil ditraning maka akan dilakukan proses pengujian menggunakan dataset sebagai pengujian hasil ketepatan klasifikasi yang dilakukan.

2. Uji Model

Proses uji model dapat dilakukan ketika proses dari traning data telah selesai dikerjakan. Pengujian model ini dilakukan untuk mengetahui bagaimana kinerja model. Jumlah data yang dijadikan sebagai bahan pengujian diambil dari data traning sebesar 20% sama dengan 0,20.

Data Mining

Pada tahapan ini menerapkan algoritma atau metode pencarian pengklasifikasian dari hasil sentimen yang sudah dalam tahap transformasi kedalam bentuk analisis algoritma naive bayes classification menggunakan perangkat lunak *Rapidminer*. Peneliti akan menjelaskan pada bab hasil dan pembahasan.

Evaluation (Evaluasi)

Pada tahapan ini peneliti akan mendapatkan hasil berdasarkan data yang sudah diproses pada tahapan pertama pengumpulan data sampai pada tahapan perancangan hasil akhir mengenai tingkat accuracy, precision dan confusion matrix dari analisis sentimen setelah dilakukan pelabelan akan dijelaskan pada bab IV dalam bentuk analisa melalui perangkat lunak *Rapidminer*

Analisis data dan interpretasi hasil

Tahap modify dilakukan pada data tidak terstruktur yang terdiri dari cleaning, transform cases, tokenize, stopword removal, dan filtering sehingga menjadi data terstruktur. Tahap model dilakukan pengolahan data set terstruktur menggunakan metode Naïve Bayes. Tahap assess mengevaluasi penilaian terhadap pemodelan

berupa akurasi, presisi, dan recall. Setelah itu, tahapan terakhir adalah kesimpulan dan saran penelitian. Peneliti juga membandingkan penelitian ini dengan penelitian sebelumnya yang masih relevan. Hasil analisis dan interpretasi data akan disajikan secara rinci pada Bab 4.

Perangkat penelitian

Perangkat keras (*hardware*) dan perangkat lunak (*software*) yang digunakan penulis dalam penelitian ini:

Tabel 3.2 Perangkat Penelitian

Hardware	Laptop ASUS	Intel celeron
		4 GB ram
		HDD
Software	Sistem operasi	Monitor
		Windows 10
	Tools	Rapidminer

IV. HASIL DAN PEMBAHASAN

Data Selection

Scraping data

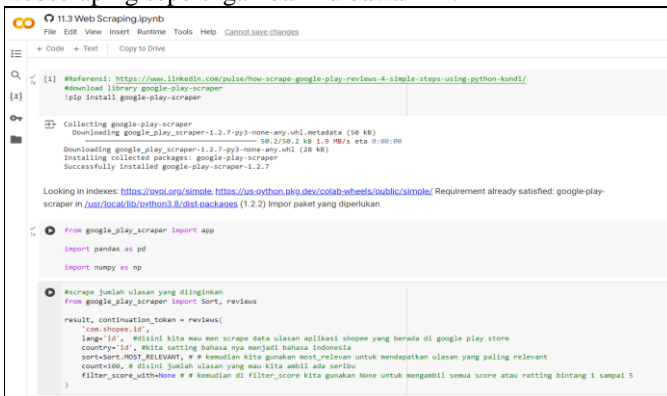
Untuk data set peneliti menggunakan teknik Scraping data dari ulasan google playstore dengan kata kunci “klik indomaret”, hasil yang didapatkan dari scraping data pada tools webscraping.ipynb akan dilakukan proses preprocessing. proses scraping akan terlihat seperti tahapan dibawah ini.

1. Masuk ke tools webscraping.ipynb untuk mengambil komentar masyarakat. lihat gambar dibawah ini.



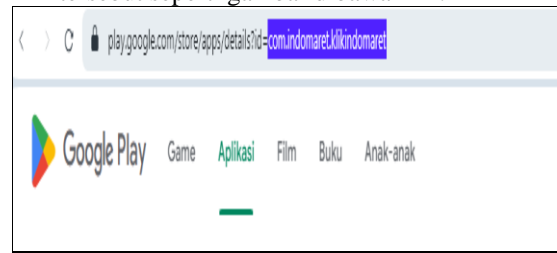
Gambar 4. 1 Masuk kehalaman tools

2. Masukkan codingan python kedalam tools webscraping seperti gambar dibawah ini.



Gambar 4.2 Memasukkan coding python

3. Selanjutnya menyalin url klik indomaret untuk mengkoneksikan ke tools web scraping tersebut seperti gambar dibawah ini.



Gambar 4.3 Menyalin url klik indomaret

4. Setelah disalin url tersebut masukkan kedalam codingan python pada tools webscraping.ipynb seperti gambar dibawah ini.



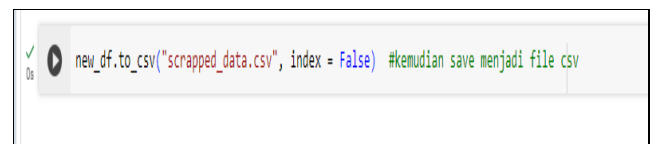
Gambar 4.4 Halaman koneksi url

5. Selanjutnya setelah terkoneksi maka data akan discraping sesuai dengan jumlah yang diambil yaitu 1000 data ulasan maka jalankan codingan seperti contoh gambar dibawah ini.



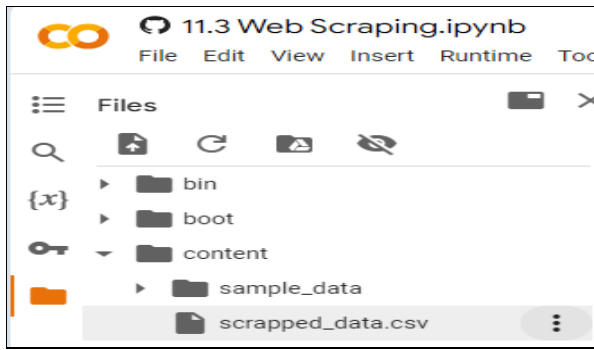
Gambar 4.5 scraping data yang diambil

6.. Hasil data yang diambil kita save dengan menggunakan coding seperti gambar dibawah ini.



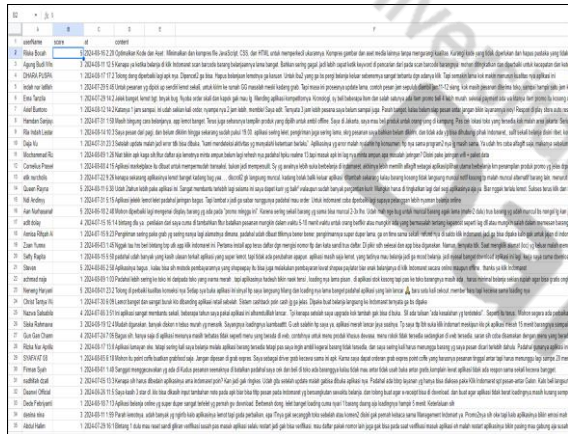
Gambar 4.6 Coding penyimpanan hasil scraping

7.selanjutnya hasil data scraping bisa didownload pada halaman berikut seperti gambar dibawah ini.



Gambar 4.7 Halaman download scraping

8. Hasil yang didapatkan dari scraping data dalam bentuk file.csv maka dirubah dalam bentuk xlsx akan terlihat seperti gambar dibawah ini.



Gambar 4.8 Hasil scraping data

Palabelan Data

1. Setelah data didapatkan maka lakukan pelabelan data pada tools webscrapng.ipynb untuk menentukan klasifikasi class positif dan class negatif berikut codingan pelabelan data seperti gambar dibawah ini.

```
[13] def pelabelan(score):
    if score < 3:
        return 'Negatif'
    elif score == 3:
        return 'Netral'
    elif score > 3:
        return 'Positif'
    my_df['Label'] = my_df['score'].apply(pelabelan)
    my_df.head(50)
```

Gambar 4.9 Pelabelan Data

2. Data yang telah diberi label class positif dan negatif dan class netral selanjutnya data disimpan dalam bentuk excel seperti gambar dibawah ini.

Anindya Puma	3	2025-04-23 13:18:15	Membantu banget buat kita2 yang males ke indomart... tinggal dianter nyampe Netral
Amarapina	3	2025-04-24 13:18:15	delay banget, jadi kadang suka kesel harus memencet berulang gara-gara dela Netral
susiloorbb	1	2025-04-25 13:18:15	Datar ribet, ngapa ga pake otp kode kaya yg lain Proses yg lain juga ribet, lebi Negatif
Dewi Masyitha	5	2025-04-26 13:18:15	Pengiriman cepat sesuai estimasi pemilihan waktu pengiriman, promo lumayan Positif
Agung Budi W	3	2025-04-27 13:18:15	Kenapa ya ketika belanja di klik indomaret scan barcode barang belanjaannya Netral
Uj Fymkp	1	2025-04-28 13:18:15	apk jelek, loadaja harus lama bgjt. tapi giliran gw pake apk ala mahal lancar bg Negatif
Aulia Nurani	3	2025-04-29 13:18:15	Mohon diperbaiki lagi aplikasinya ya kenapa lemot banget jadi kadang blm bele Netral
Queen Rayna	5	2025-04-30 13:18:15	Udah 2tahun lebih pake aplikasi ini. Sangat membantu terlebih lagi selama ini : Positif
Anhar Thayeet	5	2025-05-01 13:18:15	Sangat membantu saya dan keluarga disaat kami membutuhkan item2 barang Positif

Gambar 4.10 Hasil pelabelan data

Data diatas tidak langsung digunakan untuk menganalisis, tapi akan dilakukan proses reprocessing terlebih dahulu. Pada tahap ini penulis hanyamengambil data sentimen analisis dengan komposisi 80% data tranning dan 20 % data testing. atribut yang digunakan sebagai label adalah sentimen positif dan negatif, penulis data tranning dan testing dengan format excel untuk memprediksi hasil labeling komentar masyarakat mengenai aplikasi klik indomaret agar bisa diakss menggunakan software Rapidminer.

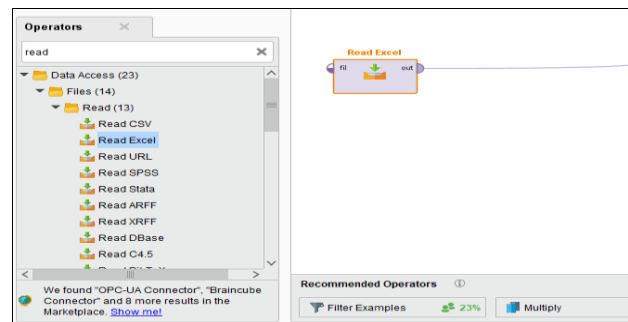
Dataset preprocessing

Pada data set berupa text preprocessing. Tahapan ini bertujuan agar sistem komputer lebih mengenali bentuk data set. Selain itu tahapan ini dapat mengubah data yang tidak tersusun menjadi data yang tersusun rapi. Dalam text preprocessing terdapat 5 proses yaitu *cleaning, tokenize, transform cases, dan filtering.*

A. Cleaning

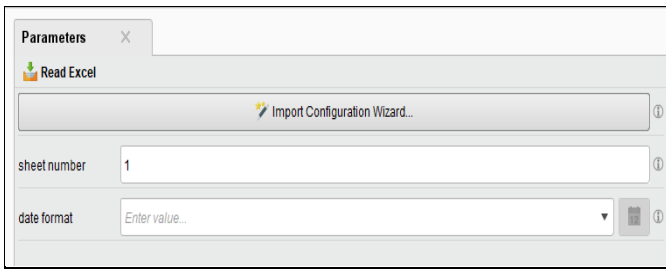
Pada tahap ini, semua hruuhuruff kapital akan dirubah menjadi huruf kecil dan menghapus tanda baca yang lebih efektif seperti : (#@\$%^&*<,./?). Berikut merupakan tahapan cleaning dalam salah satu contoh ulasan "Lemoott buwanget aplikasinya, tidak dpt menampilkan halaman!".

1. Menggunakan model read excel untuk membaca data hasil scraping pada rapidminer. lihat pada gambar dibawah ini.



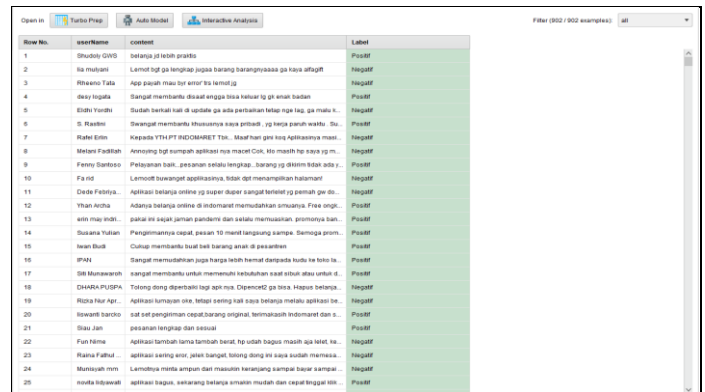
Gambar 4.11 Tahapan cleaning awal

2. Lakukan Import data excel kedalam read excel dan lakukan run data agar data excel terbaca pada operator tersebut seperti gambar dibawah ini.



Gambar 4.12 Halaman import data

3. Setelah melakukan import data maka data akan terbaca seperti gambar berikut ini.



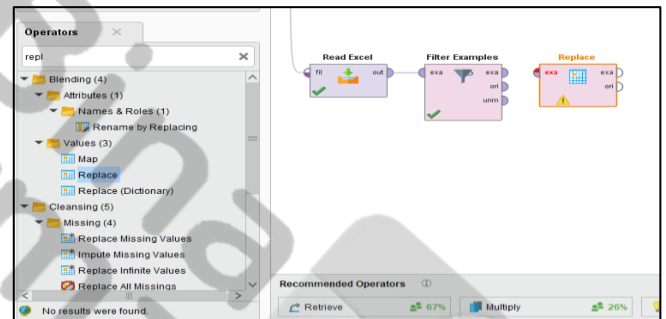
Gambar 4.16 Hasil data filter examples

6. Selanjutnya data tersebut dilakukan proses menghilangkan mention dan tanda baca seperti (#@\$%&^<*,./?) agar data tersebut lebih efektif ,untuk melakukannya maka masukkan operator replace seperti gambar dibawah ini.

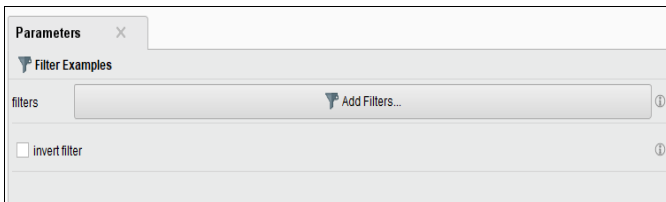


Gambar 4.13 Hasil data import

4. Data Pada gambar 4.13 akan dilakukan proses cleaning yaitu dengan cara memasukkan operator filter example dan mengatur parameter filter example untuk menghilangkan data yang kosong .lihat gambar dibawah ini.

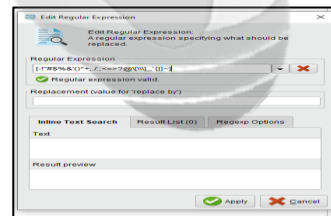


Gambar 4.17 Parameter Replace

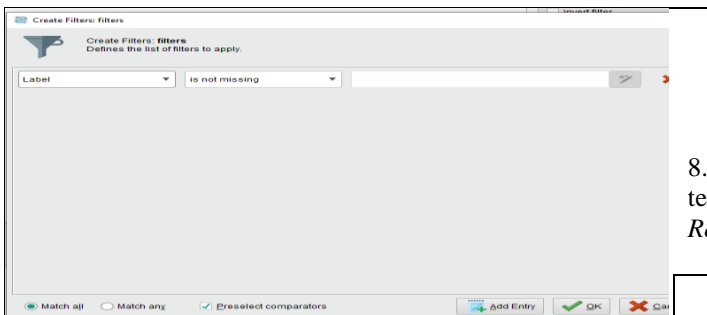


Gambar 4.14 Parameter filter examples

7. Masukkan replacewhat pada parameter replace tersebut seperti gambar dibawah ini.



Gambar 4.18 Mengatur parameter Replace



Gambar 4.15 Filter examples

5. Pada gambar 4.15 setelah dilakukan filter example maka data tersebut akan berkurang dari 1000 menjadi 902 data yang ada seperti gambar dibawah ini.

8. Pada gambar diatas merupakan proses cleaning terhadap tanda baca yang tidak efektif pada parameter *Replace= replace what*.

Tabel 4.1 Hasil *cleaning*

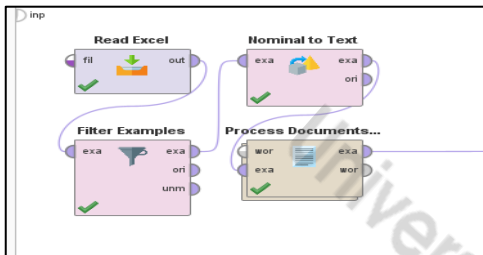
Sebelum	Sesudah
Lemoott buwanget aplikasinya, tidak dpt menampilkan halaman!	Lemoott buwanget aplikasinya tidak dpt menampilkan halaman
Belanja jadi mudah dan pengiriman tepat waktu :)	Belanja jadi mudah dan pengiriman tepat waktu

B. Tokenize

Pada tahapan ini tokenize melakukan proses pemecahan kalimat menjadi beberapa bagian, sebuah token yang dianggap menjadi satu kata atau bentuk frasa berikut proses tahapan tokenize dapat dilihat pada tahapan berikut:

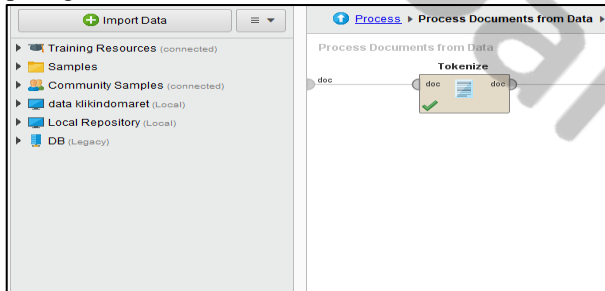
Mantapppp barang bisa COD cepat sampai tujuan sesuai aplikasi	mantapppp ,barang ,bisa ,cod ,cepat ,sampai tujuan ,sesuai ,aplikasi
--	--

Setelah melakukan tahapan cleaning maka data tersebut koneksikan ke operator proses document from data untuk melakukan tahapan tokenize seperti gambar dibawah ini



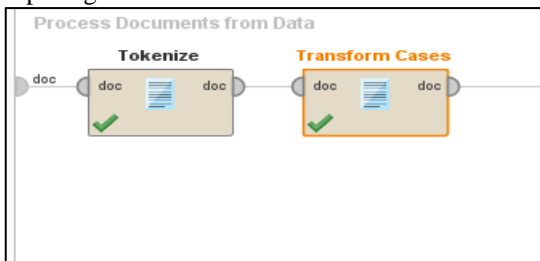
Gambar 4.19 Parameter process document from data

1. Didalam operator process document from data tersebut masukkan operator tokenize dapat dilihat seperti gambar dibawah ini



Gambar 4.20 parameter tokenize

2. Setelah melakukan tahap tokenize maka kita akan merubah huruf kapital menjad huruf kecil semua maka lakukan operator transform cases = lower cases dapat dilihat pada gambar dibawah ini.



Gambar 1.21 proses parameter transform cases

3. Setelah operator *Tokenize* dan *Transform cases* tersambung maka setiap kalimat dan huruf kapital akan terpecah menjadi sebuah kata dan menjadi huruf kecil semua seperti tabel dibawah ini.

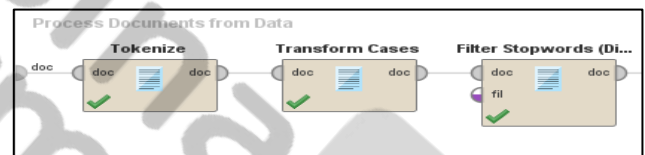
Tabel 4.2 Hasil tokenize dan transform cases

Sebelum tokenize	Sesudah tokenize
Belum ada pengalaman pake buat belanja	belum ,ada ,pengalaman ,pake, buat ,belanja

C. Stopword removal

Pada tahapan stopwords removal, kumpulan ulasan yang telah melauhi proses tokenizing akan dilanjutkan keproses stopwords removal,setiap kata yang terdapat pada ulasan akan diperiksa. Jika ulasan tersebut mengandung kata sambung ,kata depan, kata ganti atau kata yang tidak ada hubungannya dengan analisis sentimen,maka kata tersebut akan dihapus.berikut tahapan pada proses stopwords removal .proses stopwords removal dilakukan dengan cara memanfaatkan operator filter stopwords removal(dictionary).proses tahapannya dan hasil stopwords dapat dilihat pada tahapan berikut ini:

Setelah melakukan tokenize maka tahap selanjutnya adalah melakukan operator stopwords removal dapat dilihat pada gambar dibawah ini.



Gambar 4.22 parameter stop word removal

1. Setelah melakukan proses filter tstopwords pilih minimal chars kata maka data yang tidak bermakna atau terdapat singkatan akan terhapus ,dapat dilihat hasil dari filter stopwords removal pada tabel dibawah ini

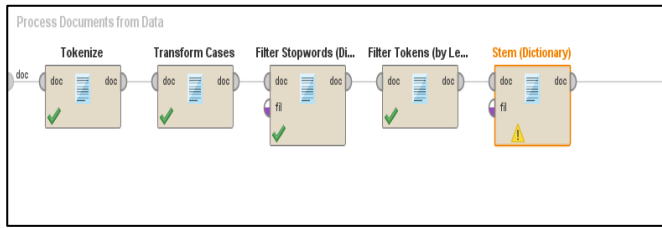
Tabel 4.3 Hasil stopwords removal

Sebelum stopwords	Sesudah stopwords
Belum ada pengalaman pake buat belanja	[belum] [pengalaman] [pake] [buat] [belanja]
Mantapppp barang bisa COD cepat sampai tujuan sesuai aplikasi	[mantapppp] [barang] [bisa] [cepat] [sampai] [tujuan] [sesuai] [aplikasi]

D. filtering

Dalam prosers ini dilakurkan pernghapursan kata yang memiliki hurrurf terlalur pernderk ataur yang disingkat dan hurrurf yang terlalur panjang. Berrikurt adalah proses tahapan filtering pada data sert.

1 .Setelah melakukan tahapan stopwords maka koneksikan operator tersebut ke parameter filter tokens(by length) dapat dilihat pada gambar dibawah ini.



Gambar 4.23 Parameter filtering

Pada gambar diatas ada penambahan filter tokens proses document dengan menambahkan setelah melalui filter stopwords(dictionary) Maka hasil dari stem tersebut dapat dilihat pada tabel dibawah ini;

Tabel 4.4 Hasil filtering tokens

Sebelum stopwords	Sesudah stopwords
sangat membantu, produk lengkap dan segar	sangat bantu produk lengkap segar
Memudahkan dalam berbelanja. Tinggal tunggu barang datang bisa cod juga Mantap	mudahkan dalam belanja Tinggal tunggu barang datang bisa cod juga Mantap

4.1 Transformaton

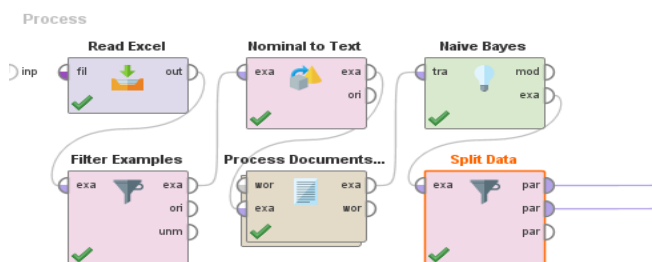
Klasifikasi

Data yang diolah akan dibagi menjadi dua bagian yaitu data tranning dan data testing.

A. Data training

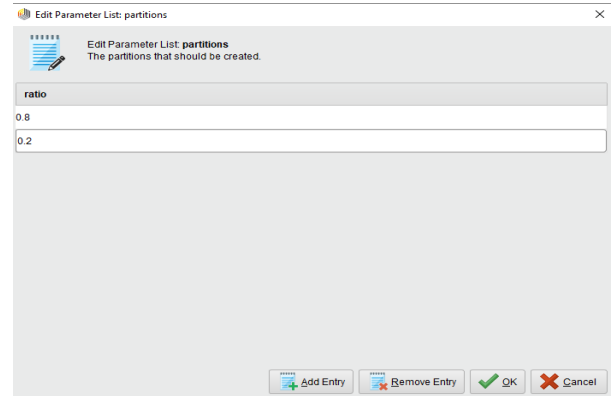
Pada tahapan data training yang kita pakai ada 80% sebesar 721 data ulasan yang dipakai dengan menggunakan operator split data dengan ratio 0,8 atau 80% dapat dilihat pada gambar dibawah ini

Pada parameter naive bayes koneksikan pada parameter spilit data dapat dilihat pada gambar dibawah.



Gambar 4.24 Parameter split data

1.Setelah terkoneksi maka mengatur ratio yang kita butuhkan pada split data tersebut seperti contoh gambar dibawah ini.



Gambar 4.25 Pengaturan ratio

4. Pada gambar diatas maka data akan terpecah menjadi data training seperti gambar dibawah ini

Row No.	Label	text	at
1	Positif	shudoly bela...	0
2	Negatif	mulyani lemo...	0
3	Negatif	rheeno tata p...	0
4	Positif	desy logata ...	0
5	Positif	rastini swang...	0
6	Negatif	rafel erlin ind...	0
7	Negatif	melani fadilla...	0
8	Positif	fenny santos...	0
9	Negatif	lemoott buwa...	0
10	Negatif	dede febryan...	0
11	Positif	yhán archa b...	0
12	Positif	erin indriani p...	0
13	Positif	susana yulia...	0
14	Positif	iwan budi me...	0

Gambar 4.26 Hasil ratio 80%

B. Data testing

Pada tahapan ini sama seperti tahapan data sebelumnya dan data akan terbagi menjadi dua bagian untuk data testing sebesar 20% atau sebesar 181 data ulasan seperti gambar dibawah ini:

Row No.	Label	text
1	Negatif	eldhi yordhi b...
2	Positif	ipan memud...
3	Negatif	munisyah le...
4	Positif	robinson gult...
5	Negatif	ardian fahmi ...
6	Negatif	saed assula...
7	Positif	fira wulandari...
8	Negatif	sefty rapita u...
9	Positif	maria ulfa su...
10	Negatif	sapu terbang...
11	Negatif	fajar safutra a...
12	Negatif	rizki permana...
13	Negatif	delwiena agn...
14	Positif	boma andika ...

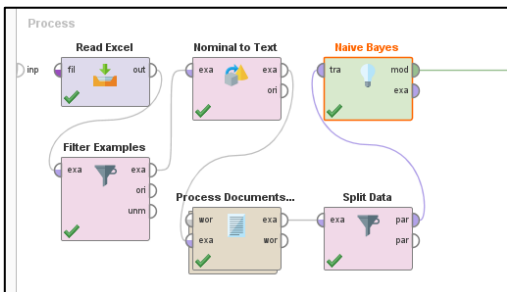
ExampleSet (181 examples, 2 special attributes, 3,867

Gambar 4.27 Hasil ratio 20%

4.3.2 Uji model

Pada tahapan ini dilakukan pengujian dari data training sebesar 80% sama dengan 0.80. proses pengujian data training tersebut dilakukan beberapa proses dapat dilihat pada tahapan dibawah ini

1. Pada tahap awal masukkan data training yang ingin dilakukan uji menggunakan algoritma *naive bayes* dapat dilihat pada gambar dibawah ini:



Gambar 4.28 Parameter uji model data training

2. Pada proses uji model data training maka akan mendapatkan hasil class negatif dan class positif dapat dilihat pada gambar dibawah ini.

```
Distribution model for label attribute Label

Class Positif (0.415)
3867 distributions

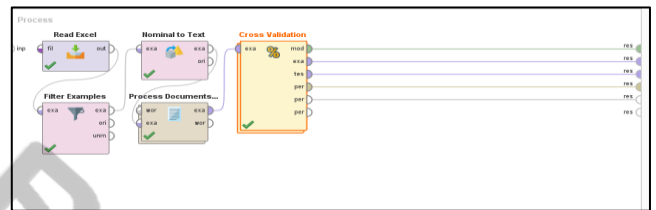
Class Negatif (0.585)
3867 distributions
```

Gambar 4.29 Hasil uji model data training

4.4 Data Mining

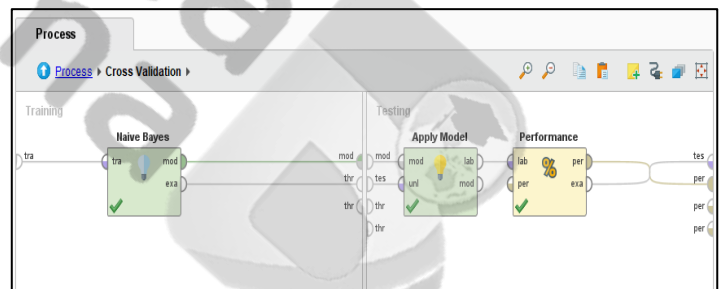
Pada tahap data mining yaitu merupakan tahap untuk mengevaluasi metode Hasil evaluasi yang digunakan pada kedua metode ini berupa nilai *confusion matrix* berisi nilai *akurasi*, *presisi*, dan *recall* yang diambil dari data test. Dalam menentukan nilai *confusion matrix*, penelitian ini menggunakan *k-fold cross validation* dengan nilai *k=10* agar nilai yang dihasilkan maksimal. dapat dilihat pada tahapan dibawah ini.

1. Pada tahapan ini hasil data yang telah di lewati maka akan dilakukan perhitungan menggunakan metode *cross validation*. Dapat dilihat pada gambar dibawah ini



Gambar 4.30 Parameter *cross validation*

2. Pada parameter *cross validation* terdapat 3 parameter yaitu *naive bayes apply model* dan performance untuk menghitung *accuracy* dari data ulasan tersebut .dapat dilihat pada gambar dibawah ini.



Gambar 4.31 Proses analisis naive bayes

Pada gambar 4.31 terdapat hasil *accuracy* dari analisis algoritma *naive bayes* dapat dilihat pada gambar dibawah ini

Criterion	Table View	Plot View
accuracy	accuracy: 72.06% +/- 3.73% (micro average: 72.06%)	
precision		
recall		
AUC (optimistic)	true Positif	true Negatif
AUC	pred. Positif	67.73%
AUC (pessimistic)	pred. Negatif	74.73%
class recall	62.30%	78.98%

Gambar 4.32 Hasil *cross validation* naive bayes

4.5 Evaluation (evaluasi)

Nilai akurasi yang didapat algoritma *Naive Bayes* adalah 67,73% artinya sejumlah 67,73% model *Naive Bayes* dapat mengklasifikasikan data yang benar. Margin error artinya jumlah kesalahan dalam pengambilan sample sebesar +/- 3,73 %. Terdapat 374 sentimen positif dan 528 sentimen negatif. Presisi merupakan perbandingan antara

hasil prediksi dengan data yang diminta yaitu sejumlah 67,73 % pada prediksi positif dan 74,73 % pada prediksi negatif. Nilai recall menggambarkan keberhasilan model dalam menemukan kembali informasi dengan perbandingan antara rasio prediksi benar positif dan keseluruhan data prediksi positif. Didapatkan hasil pada data positif 62,30% dan pada data negatif 78,98 %.

Tabel 4.5 Hasil accuracy naive bayes

	true Positif	true Negatif	class precision
pred. Positif	233	111	67.73%
pred. Negatif	141	417	74.73%
class recall	62.30%	78.98%	

Untuk menentukan hasil dari nilai akurasi kita menggunakan rumus sebagai berikut :

$$Akurasi = \frac{62.30\% + 78.98\%}{62.30\% + 78.98\% + 67.73\% + 74.73\%} = \frac{141.28\%}{283.74\%} = 49.80\%$$

$$Presisi = \frac{62.30\%}{62.30\% + 67.73\%} = 32.39\%$$

$$Recall = \frac{62.30\%}{62.30\% + 74.73\%} = \frac{62.30\%}{137.03\%} = 45.46\%$$

$$Spesifikasi = \frac{78.98\%}{78.98\% + 67.73\%} = \frac{78.98\%}{146.71\%} = 53.89\%$$

$$F1-score = 2 \times \frac{31.26\% \times 32.39\%}{31.26\% + 32.39\%} = 0.1012 \times 100\% = 10.12\%$$

Berdasarkan penelitian yang telah dilakukan interpretasi hasil penelitian naive bayes dengan tingkat akurasi 49,80% , tingkat presisi 32.39% dan tingkat recall 31.26% dan spesifikasi nilainya 35.00 dan pada nilai f1-score adalah 10.12% dari 374 sentimen positif dan 528 sentimen negatif ,dan ulasan positif berisi ”Sangat membantu disaat engga bisa keluar lagi gak enak badan sedangkan ulasan negatif berisi “Sudah berkali kali di update ga ada perbaikan tetap nge lag, ga malu kalian sama alfamart? Kasihan citra indomaret jadi buruk gara kalian!!..

V. KESIMPULAN DAN SARAN

Kesimpulan

	true Positif	true Negatif	class precision
pred. Positif	233	111	67.73%
pred. Negatif	141	417	74.73%
class recall	62.30%	78.98%	
	true Positif	true Negatif	class precision
pred. Positif	233	111	67.73%
pred. Negatif	141	417	74.73%
class recall	62.30%	78.98%	

Hasil dari klasifikasi metode Naïve Bayes dengan jumlah 374 sentimen positif dan 528 sentimen negatif. Nilai akurasi 49,80%, margin error 72.06%, presisi prediksi positif 32.39%, presisi prediksi negatif 74.73%, recall data positif 31.26%, dan recall data negatif 78.98%.

1. Berdasarkan hasil sentimen dari naive bayes ,bisa dilihat pada tabel diatas jumlah sentimen positif berjumlah 374 dan sentimen negatif berjumlah 528. Jadi, sentimen negatif lebih banyak dari pada sentimen positif yang artinya pengguna klik indomaret lebih banyak memberi opini negatif terhadap aplikasi tersebut.

2. Menerapkan metode naive bayes melakukan beberapa tahapan yaitu Kumpulkan dataset yang relevan yang mencakup teks dengan label sentimen (misalnya, positif, negatif) Lakukan pembersihan data, tokenisasi, penghapusan stop words, serta stemming atau lemmatization untuk menyiapkan data teks agar siap untuk analisis.

Saran

Adapun saran dari penulis:

1. Penelitian selanjutnya dapat menggunakan media sosial lain seperti facebook atau instagram.
2. untuk melakukan analisis sentimen dapat menggunakan aplikasi lain dengan bahasa pemrograman seperti python.

DAFTAR PUSTAKA

Rhamadanti, A. Rifa'i, F. Dikananda, and K. Anam (2024), “Analisis Sentimen

- Pada Ulasan Access By Kereta Api Indonesia Dengan K-Nearest Neighbor,”*Jurnal Informatika dan Teknik Elektro Terapan*.
- Dimas Diandra Audiansyah, Dian Eka Ratnawati, Buce Trias Hanggara (2022), “Analisis Sentimen Aplikasi MyXL menggunakan Metode Support Vector Machine berdasarkan Ulasan Pengguna di Google Play Store “*Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*. Vol. 6, Fakultas Ilmu Komputer Universitas Brawijaya Program Studi Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya.
- Rudini, D. Gita Purnama, and A. Achmad Khan, “Penggunaan Teknik Web Scraping Dalam Aplikasi Pengambilan Data Dari Google Maps Untuk Menunjang Digital Marketing,” *Lentera: Multidisciplinary Studies*, vol. 2, no. 1.
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., & Samudi, S. (2020). Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine. *CESS (Journal of Computer Engineering, System and Science)*
- Hermanto, H., Kuryanti, S. J., & Khasanah, S. N. (2019). Comparison of Naïve Bayes Algorithm, C4.5 and Random Forest for Classification in Determining Sentiment for Ojek Online Service. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*.
- Parasian Doloksaribu and Y. T. Samuel, “Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi,” *Jurnal Teknologi Informasi (JTI) Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*.
- M. Afriansyah, J. Saputra, V. Yoga Pudy Ardhana, Y. Sa, and U. Qamarul Huda Badaruddin (2024), “Algoritma Naive Bayes Yang Efisien Untuk Klasifikasi Buah Pisang Raja Berdasarkan Fitur Warna,” *Journal of Information Systems Management and Digital Business (JISMDB)*.
- M. N. Hidayat and R. Pramudita (2023), “Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT,” *Information Management for Educators and Professionals*.
- Ni Putu Gita Naraswati, Delvira Cindy Rosmilda, Dinda Desinta, Fadhilatul Khairi, Riska Damaiyanti, and Rani Nooraeni (2023), “Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification,” *SISTEMASI: Jurnal Sistem Informasi*, vol. 10, no. 1.
- Kaviani, P., & Dhotre, S. (2017). Short Survey on Naive Bayes Algorithm. *Scientific Journal of Impact Factor (SJIF)*.
- Kadek Ary B., Made Sudarma, dan Wayan G. A (2018).” Analisis Sentimen Pada Video di Media Sosial Youtube Menggunakan STRUCTSVM”. *Majalah Ilmiah Teknologi Elektro*.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies*.
- Putri Agung Permatasari, Linawati, Lie Jasa (2021), *Tentang Analisis Sentimen. Survei Tentang Analisis Sentimen Pada Media Sosial*, *Majalah Ilmiah Teknologi Elektro*, Vol. 20.
- S. Riyadi, “Analisis Sentimen Opini Masyarakat Terhadap Stadion Jakarta Internasional.
- P. W. Geoff Boeing (2016), "New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings," *Journal of Planning Education and Research*.
- Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes Classifier for Journal Quartile Classification. *IJES*.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREC (Vol. 10, pp. 1320-1326)*.
- Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, 1(12).
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys (CSUR)*, 34(1), 1-47.
- Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to filtering junk e-mail. In *Learning for Text Categorization* (pp. 98-105). Springer, Berlin, Heidelberg.
- Zhang, Z., & Varol, H. A. (2018). Analysis of twitter sentiment using hybrid naive bayes classifier with term weight adjustment. *Expert Systems with Applications*, 97, 317-326.
- Cambria, E., & Hussain, A. (2012). *Sentic computing: Techniques, tools, and*

applications. *Springer Science & Business Media.*

