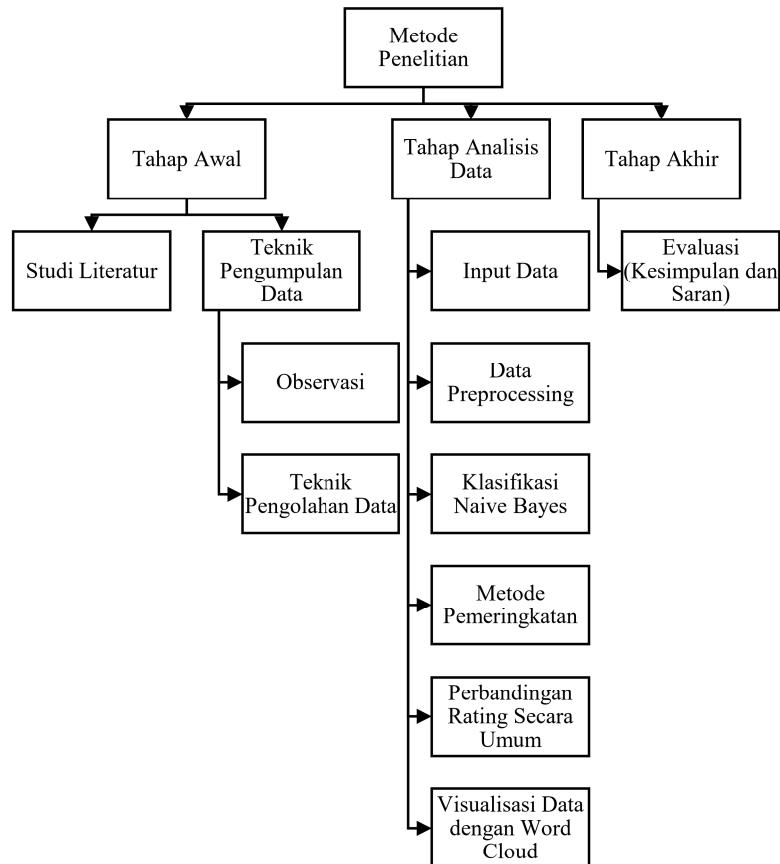


BAB III

METODOLOGI PENELITIAN

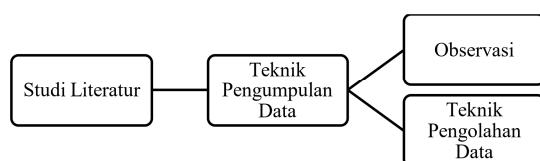
Dalam tahapan penelitian ini akan melalui tiga tahapan yaitu tahap awal penelitian, tahap analisis dan tahap akhir dapat dilihat pada Gambar 3.1



Gambar 3.1 Diagram metode penelitian

3.1 Tahap Awal Penelitian

Hal yang dilakukan pada penelitian ini:



Gambar 3.2 Bagan Tahapan Awal

Gambar diatas merupakan tahapan penelitian yang penulis lakukan dalam melakukan penelitian ini adapun rincian jelasnya adalah sebagai berikut:

3.1.1 Studi Literatur

Studi literatur ini dilakukan untuk mengetahui model penelitian yang terjadi pada bidang analisis sentimen, studi literatur ini dilakukan untuk mengulas jurnal-jurnal yang telah ada dan digunakan untuk menentukan studi kasus serta metode yang akan digunakan oleh penulis saat melakukan penelitian.

3.1.2 Teknik Pengumpulan Data

Pengumpulan data ini dilakukan untuk memperkuat alasan mengapa penelitian ini harus dilakukan, pengumpulan data diambil dari sumbernya langsung yaitu *website* resmi *Google Play Store* serta ulasan pengguna aplikasi Tokopedia, Shopee, dan Bukalapak. Berikut adalah tahapan pengumpulan data yang penulis lakukan dalam menyusun penelitian ini:

A. Observasi

Observasi ini dilakukan untuk mengamati objek yang akan diteliti. Pada penelitian ini dilakukan dengan mengamati ulasan pengguna situs belanja *online*, adanya banyak macam ulasan yang diberikan oleh pengguna pada aplikasi ini yang membuat penulis melakukan penelitian dalam presentase klasifikasi ulasan masyarakat yang ada di *Google Play Store*.

B. Teknik Pengolahan Data

Pengambilan data pada penelitian ini didapatkan dari ulasan para pengguna aplikasi Tokopedia, Shopee, dan Bukalapakdi *Google Play Store*, dengan menggunakan teknik *web scraping* (*add-on* dari *Software Chrome* yakni *Web Scraper*). Dataset yang terambil sebanyak masing-masing 120 ulasan dengan total menjadi 360 ulasan, jumlah data seimbang (*balanced*) setiap labelnya, karena dengan data yang tidak seimbang (*imbalanced*), klasifikasi yang dibangun memiliki kecenderungan untuk mengabaikan *minority*.

Perlu diperhatikan setelah melakukan teknik pengolahan data di atas sebelum melakukan langkah-langkah selanjutnya, ulasan yang telah didapatkan dari ketiga *online shop* harus memiliki kata yang baku. Terdapat beberapa kata sehari-hari yang tampil pada penelitian ini, oleh karenanya penulis merubah kata-kata sehari-hari tersebut kedalam kata baku sesuai dengan kamus yang digunakan yakni *Liu's word list*. Pada penelitian ini menggunakan metode perubahan secara manual karena kata yang ditemukan hanya sedikit. Berikut adalah kata yang di temukan dan diubah kedalam kata baku dalam Tabel 3.1

Tabel 3.1 Merubah kata menjadi baku.

Kata Sehari-hari	Kata Baku
Udah	Sudah
Dn	Dan
Drpd	Dari pada
Engga	Tidak
Nggk	Tidak
Gw	Saya
Bgt	Banget
Emang	Memang
Ga	Tidak
Tp	Tapi
Kdg	Kadang
Gak	Tidak

C. Variabel Penelitian

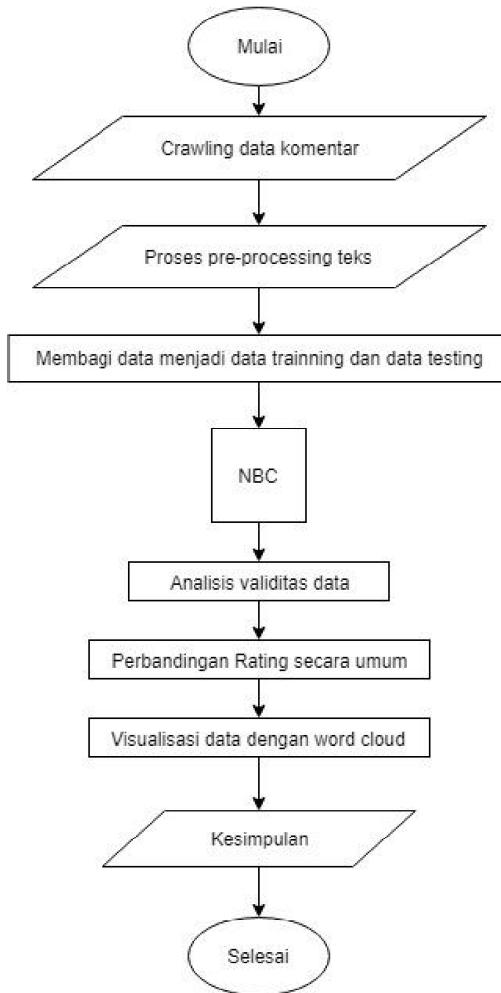
Data yang digunakan dalam penelitian ini dibagi menjadi data *training* sebanyak 90% dan data *testing* sebanyak 10% menggunakan *10-fold cross validation*. Variabel penelitian yang digunakan dalam penelitian ini diberikan dalam Tabel 3.2

Tabel 3.2 Variabel penelitian

Variabel	Keterangan	Skala Data
Y	Sentiment (Positif/Negatif) 0 = Sentiment Positif 1 = Sentimen Negatif	Nominal
X _k	Kata Kunci (Frekuensi Kemunculan Kata ke-k pada komentar ke-n)	Rasio

D. Diagram Alir

Langkah-langkah analisis ini diberikan dalam bentuk diagram alir pada Gambar 3.4



Gambar 3.4 Diagram alir penelitian

Gambar diatas menjelaskan alur dari kegiatan dalam penelitian ini mulai dari *crawling* data ulasan pada *Google Play Store* hingga membuat kesimpulan.

3.2 Tahap Analisis Data

Berikut merupakan bagan tahapan analisis data pada Gambar 3.6



Gambar 3.6 Teknik Analisis Data

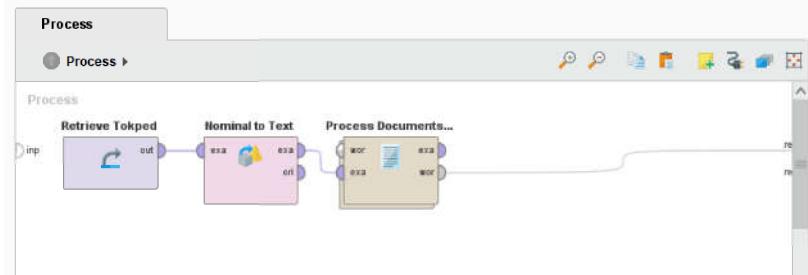
3.2.1 Input Data

1. Data didapatkan dari ulasan para pengguna aplikasi di website *Google Play Store* dengan Teknik *web scraping* (*add-on* dari *Software Chrome* yakni *Web Scraper*), kemudian tersimpan dalam bentuk *excel*. Pada penelitian ini file terbagi menjadi tiga, yakni Tokopedia, Shopee, dan Bukalapak.
2. Data yang telah tersimpan dalam bentuk *excel* kemudian diinputkan dalam *RapidMiner* untuk memudahkan proses analisis
3. Selanjutnya, sekumpulan data tersebut akan diinputkan pada tahap selanjutnya, yakni pada tahap data *preprocessing*.

3.2.2 Data Preprocessing

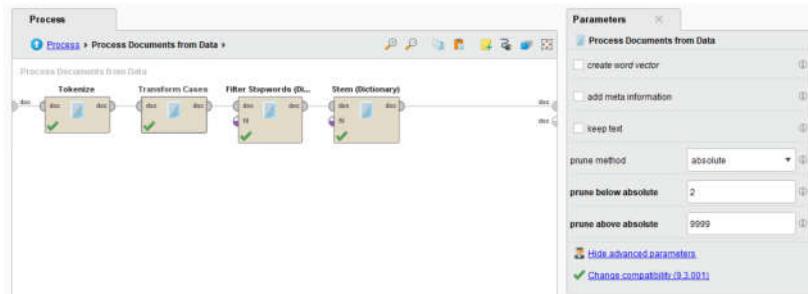
Berikutnya langkah yang harus dilakukan adalah :

1. Lakukan pemrosesan data yang telah diinputkan. Pilih *process documents*. Bisa dilihat seperti Gambar 3.8



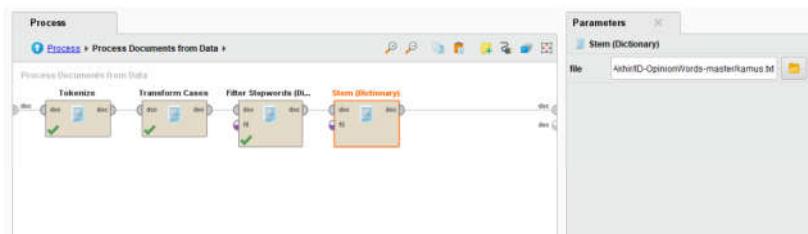
Gambar 3.8 Data preprocessing

2. Di dalam *process document* masukkan operator yang akan dijalankan. Pada tahap ini tahapan *text preprocessing* yang dilakukan adalah *case folding*, *tokenizing*, *filtering*, dan *stemming*. Bisa dilihat seperti Gambar 3.9



Gambar 3.9 Tahapan data preprocessing

3. Pada tahap *stemming* masukkan kamus yang digunakan, kamus yang digunakan pada penelitian ini adalah *Liu's Opinion word list*. Bisa dilihat seperti Gambar 3.10



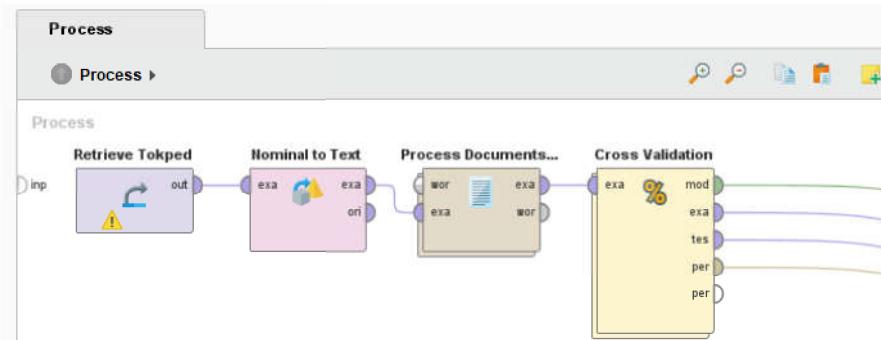
Gambar 3.10 Memasukkan kamus Liu's word list.

4. Hasil dari proses ini adalah suatu file berbentuk *excel* (.xls) dimana akan dibedakan setiap filenya sesuai dengan nama situs belanja dalam penelitian ini, yakni Tokopedia, Shopee, dan Bukalapak.
5. Masing-masing file akan diklasifikasikan pada proses selanjutnya, yakni proses klasifikasi *naive bayes*. Untuk mengetahui tingkat akurasi data.

3.2.3 Klasifikasi Naïve Bayes

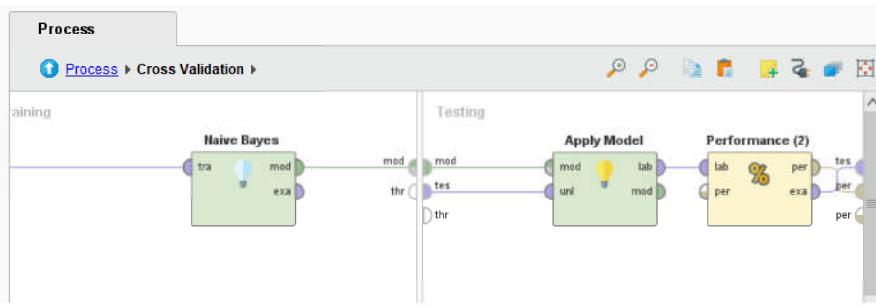
Langkah yang harus dilakukan pada tahap ini adalah :

1. Pilih *retrieve* atau bisa dengan *drag and drop* pada masing-masing data untuk membaca hasil dari *stemming* yang telah dilakukan pada tahap *preprocessing*.
2. Kemudian gunakan operator *Nominal to text* untuk mengubah bentuk nominal ke bentuk teks
3. Setelah itu pilih operator *Process Document from Data* untuk menyaring data seperti pada tahap *preprocessing*. Dan yang terakhir tambahkan operator *Cross Validation* untuk tahap validasi data. Bisa dilihat seperti Gambar 3.11



Gambar 3.11 Proses tahapan analisis

4. Di dalam operator *Cross Validation*, tambahkan operator *Naïve Bayes* pada kotak Training, kemudian pada kotak Testing diisi dengan operator *Apply Model* dan *Performance*
5. Setelah selesai memasukkan operator, tahap selanjutnya yaitu menghubungkan semua *output* dari masing-masing operator. Bisa dilihat seperti Gambar 3.12



Gambar 3.12 Proses training dan testing pada cross validation

6. Jalankan proses dan periksa hasilnya
7. Hasil dari proses pengujian ini yaitu berupa visualisasi data dalam bentuk grafik dan juga tabel yang kemudian akan diolah kembali pada proses visualisasi data *word cloud* untuk melihat kata yang paling banyak muncul. Data tersebut akan dijadikan sebagai bahan evaluasi tahap akhir

3.2.4 Perbandingan Rating Secara Umum

Pada langkah berikut ini akan ditentukan rating dari seluruh ulasan yang langsung diambil dari *Google Play Store*. Pada tahap ini akan dihitung seberapa besar reting yang didapatkan dari data yang telah diamati langsung dari *Google Play Store*. Berikut adalah tabel yang digunakan untuk menentukan rating secara umum :

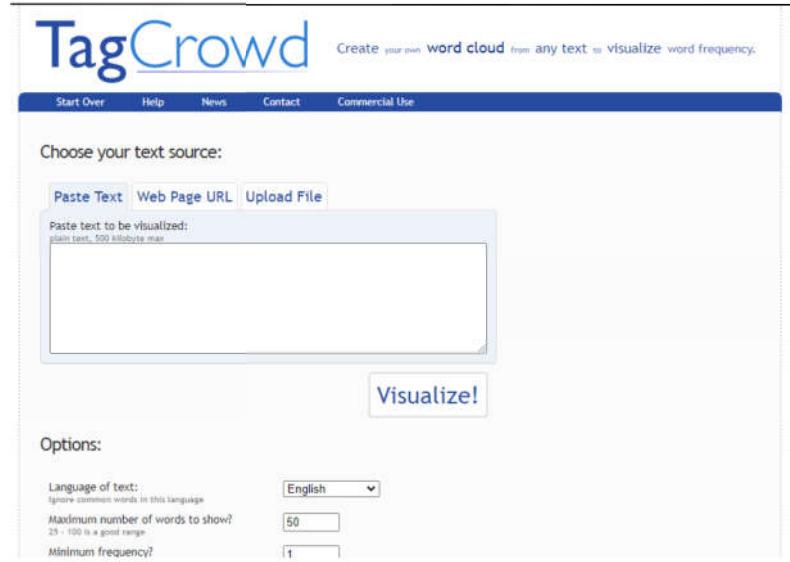
Tabel 3.3 Perbandingan rating secara umum

Rank	Situs	Jumlah Pengguna	Rating

3.2.5 Visualisasi Data dengan Word Cloud

Berikut adalah urutan langkah-langkah yang harus dilakukan dalam memvisualisasikan data dengan *wordcloud*. Dalam penelitian ini menggunakan *open source Tagcrowd*.

1. Pertama buka website *Tagcrowd* pada pencarian. Akan muncul seperti gambar 3.13.



Gambar 3.13 Halaman depan Tagword

2. Kemudian pilih upload file atau bisa memasukkan teks yang telah terbentuk pada langkah-langkah sebelumnya, yakni pada tahap *stemming*.
3. Setelah memilih file atau memasukkan teks → klik *Visualize*, untuk melihat hasil dari *wordcloud*.
4. Hasil akhir dari proses visualisasi *word cloud* ini akan dijadikan bahan acuan dalam penelitian untuk menentukan peringkat popularitas dari Tokopedia, Shopee, maupun Bukalapak sesuai dengan data dan pengujian yang telah dilakukan.
5. Kata-kata yang sering muncul dalam *word cloud* akan di evaluasi untuk dijadikan solusi bagi pemilik situs belanja maupun konsumen.

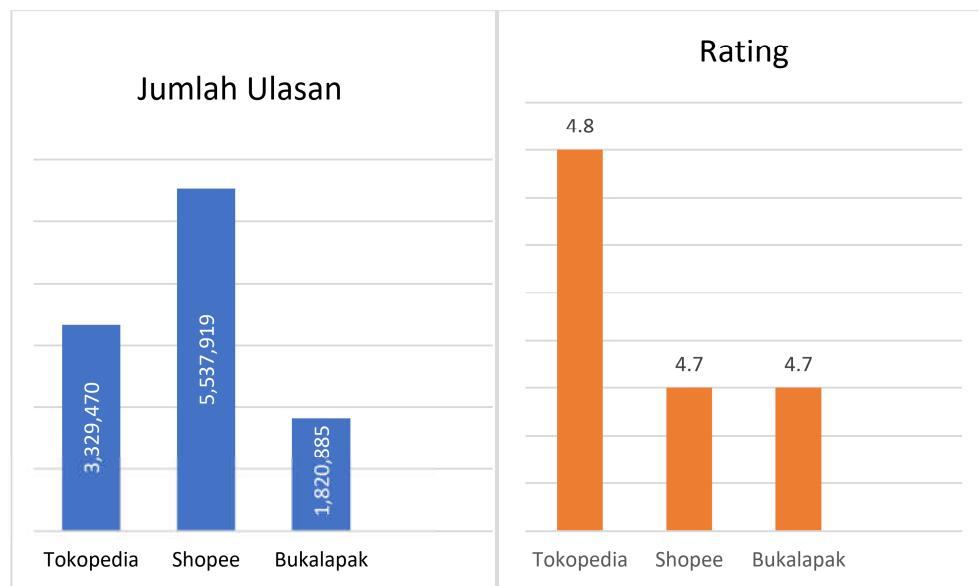
BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas hasil analisis berdasarkan pengolahan data pada platform *Google Play Store* yang memuat ulasan dari ketiga objek penelitian, yakni Tokopedia, Shopee, dan Bukalapak yang telah dilakukan. Metode yang digunakan dalam klasifikasi teks ulasan adalah *Naïve Bayes Classifier* (NBC).

4.1 Karakteristik Data dan Input Data

Google Play adalah layanan distribusi digital yang berfungsi sebagai toko media digital, yang menawarkan program musik, buku, film, dan televisi. Selain itu, *Google Play* juga menyediakan untuk menanggapi rating, pendapat, kritik, saran dan masalah *complain*. Secara manual, dilakukan pengumpulan data terhadap jumlah *rating* dari konsumen yang ditujukan kepada situs belanja Tokopedia, Shopee, dan Bukalapak. Gambar 4.1 menunjukkan jumlah rating dari konsumen pada bulan awal bulan Oktober tahun 2020.



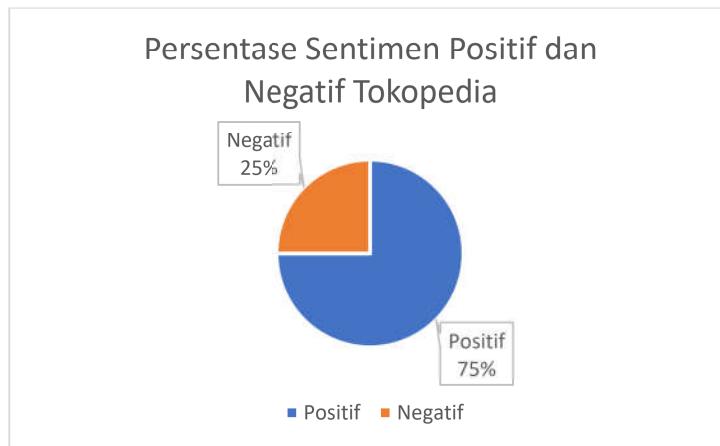
Gambar 4.1 menunjukkan jumlah ulasan dan rating di Google Play Store tahun 2020

Konsumen yang memberikan tanggapan, pendapat, kritik, saran dan masalah complain lebih banyak ditujukan pada akun Shopee dari pada akun Tokopedia dan Bukalapak. Dapat dilihat jumlah ulasan yang ditujukan pada ketiga akun pada tahun 2020, Tokopedia menjadi situs yang memiliki rating paling tinggi dengan 4,8 dari pada akun Shopee dan Bukalapak.

A. Situs Belanja online Tokopedia

Berikut merupakan prosentase ulasan positif maupun negatif dari *online shop* Tokopedia berdasarkan perhitungan manual. Perhitungan manual dilakukan oleh peneliti untuk melakukan perbandingan dengan hasil perhitungan menggunakan aplikasi *Rapidminer* untuk menghasilkan data yang lebih akurat. Perhitungan ini dilihat berdasarkan ulasan pengguna.

Apabila ulasan yang ditulis mengandung kata-kata positif dan rating yang bagus, maka ulasan tersebut masuk kedalam sentimen positif, begitu pula sebaliknya. Bisa dilihat seperti Gambar 4.2



Gambar 4.2 Persentase Sentimen Positif dan Negatif Tokopedia

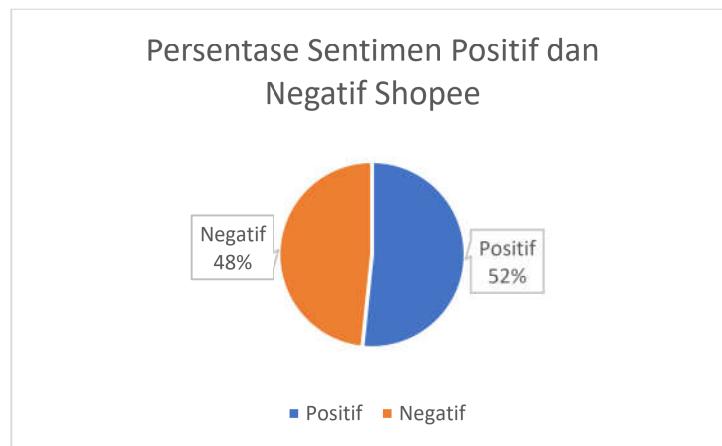
Berdasarkan Gambar 4.2 dapat diketahui bahwa data yang digunakan untuk klasifikasi terdiri dari 120 ulasan yang memuat sentimen negatif dan positif. Jumlah ulasan positif yang dihitung secara manual pada *online shop* Tokopedia adalah sebanyak 90 dan 30 ulasan negatif. Sentimen positif pada situs *online shop* Tokopedia memiliki persentase yang lebih banyak jika dibandingkan dengan persentase sentimen negatif. Persentase sentimen

negatif pada situs *online shop* Tokopedia secara keseluruhan sebesar 25%, sedangkan persentase sentimen positif adalah sebesar 75%.

B. Situs Belanja online Shopee

Berikut merupakan prosentase ulasan positif maupun negatif dari *online shop* Shopee berdasarkan perhitungan manual. Perhitungan manual dilakukan oleh peneliti untuk melakukan perbandingan dengan hasil perhitungan menggunakan aplikasi *Rapidminer* untuk menghasilkan data yang lebih akurat. Perhitungan ini dilihat berdasarkan ulasan pengguna.

Apabila ulasan yang ditulis mengandung kata-kata positif dan rating yang bagus, maka ulasan tersebut masuk kedalam sentimen positif, begitu pula sebaliknya. Bisa dilihat seperti Gambar 4.3



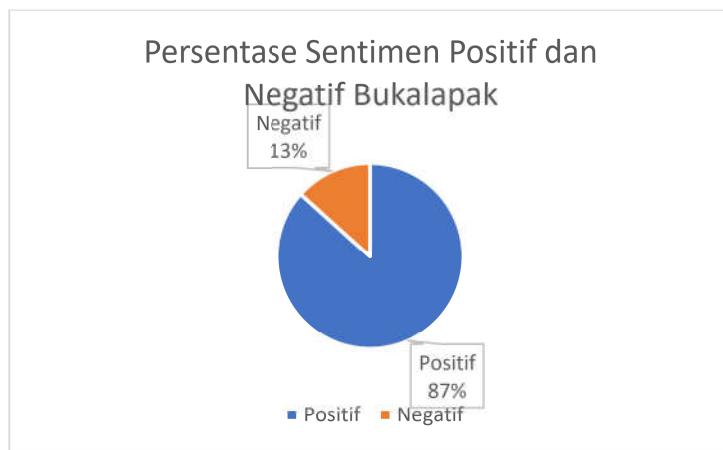
Gambar 4.3 Persentase Sentimen Positif dan Negatif Shopee

Berdasarkan Gambar 4.3 dapat diketahui bahwa data yang digunakan untuk klasifikasi terdiri dari 120 ulasan yang memuat sentimen negatif dan positif. Jumlah ulasan positif yang dihitung secara manual pada *online shop* Shopee adalah sebanyak 62 dan 58 ulasan negatif. Sentimen positif pada situs *online shop* Shopee memiliki persentase yang lebih banyak jika dibandingkan dengan persentase sentimen negatif. Persentase sentimen positif pada situs *online shop* Shopee secara keseluruhan sebesar 52%, sedangkan persentase sentimen negatif adalah sebesar 48%.

C. Situs Belanja online Bukalapak

Berikut merupakan prosentase ulasan positif maupun negatif dari *online shop* Bukalapak berdasarkan perhitungan manual. Perhitungan manual dilakukan oleh peneliti untuk melakukan perbandingan dengan hasil perhitungan menggunakan aplikasi *Rapidminer* untuk menghasilkan data yang lebih akurat. Perhitungan ini dilihat berdasarkan ulasan pengguna.

Apabila ulasan yang ditulis mengandung kata-kata positif dan rating yang bagus, maka ulasan tersebut masuk kedalam sentimen positif, begitu pula sebaliknya. Bisa dilihat seperti Gambar 4.4



Gambar 4.4 Persentase Sentimen Positif dan Negatif Bukalapak

Berdasarkan Gambar 4.4 dapat diketahui bahwa data yang digunakan untuk klasifikasi terdiri dari 120 ulasan yang memuat sentimen negatif dan positif. Jumlah ulasan positif yang dihitung secara manual pada *online shop* Bukalapak adalah sebanyak 104 dan 16 ulasan negatif. Sentimen positif pada situs *online shop* Bukalapak memiliki persentase yang lebih banyak jika dibandingkan dengan persentase sentimen negatif. Persentase sentimen negatif pada situs *online shop* Bukalapak secara keseluruhan sebesar 13%, sedangkan persentase sentimen positif adalah sebesar 87%.

Berikut ditampilkan sentimen analisis studi kasus ini dari ketiga situs *Online Shop* berdasarkan perhitungan manual. Data dibawah ini adalah data yang akan dijadikan sebagai pembanding untuk megukur akurasi pada tahap berikutnya. Dapat dilihat pada Tabel 4.1

Tabel 4.1 Keseluruhan Sentimen Positif dan Negatif.

No.	Situs Belanja	Jumlah Sentimen		Persentase Sentimen	
		Positif	Negatif	Positif	Negatif
1	Tokopedia	90	30	75%	25%
2	Shopee	62	58	52%	48%
3	Bukalapak	104	16	87%	13%

4.2 Data Preprocessing

Sebelum dilakukan analisis, langkah selanjutnya yang akan dilakukan adalah pra proses pada data ulasan di *Google Play Store* yang memuat ulasan ketiga situs belanja online, yakni Tokopedia, Shopee, dan Bukalapak terdiri dari beberapa tahapan, yaitu *case folding*, *tokenizing*, *filtering*, dan *stemming*. Proses *case folding* adalah mengubah data teks menjadi *lowercase* dengan tujuan agar kata yang sama namun berbeda secara penulisan huruf kapital dan tidak, tidak dianggap kata yang berbeda. Contoh data sebelum dan sesudah melalui tahap *lowercase* ditampilkan pada Tabel 4.2

Tabel 4.2 Proses *case folding*

Sebelum proses <i>Case Folding</i>	Sesudah proses <i>Case Folding</i>
Aplikasi yang mempermudah semuanya di masa pandemi ini, saya puas sekali dengan aplikasi tokped yang bekerja sama dengan pra kerja dan mempermudah saya beli pelatihan untuk pra kerja,yuk download toko pedia sekarang juga.	aplikasi yang mempermudah semuanya di masa pandemi ini, saya puas sekali dengan aplikasi tokped yang bekerja sama dengan pra kerja dan mempermudah saya beli pelatihan untuk pra kerja,yuk download toko pedia sekarang juga.

Sebelum proses <i>Case Folding</i>	Sesudah proses <i>Case Folding</i>
Terpaksa saya kasih bintang 1, karena proses widraw dana dari hari jumat sampai sekarang belum cair hari senin juga proses widraw ulang belum cair juga sudah nunggu 1x28jam katanya akan cair 2x24 jam. Kurang sabar lagi perbaiki dulu sistem widrawal dana tokopedia.	terpaksa saya kasih bintang 1, karena proses widraw dana dari hari jumat sampai sekarang belum cair hari senin juga proses widraw ulang belum cair juga sudah nunggu 1x28jam katanya akan cair 2x24 jam. kurang sabar lagi perbaiki dulu sistem widrawal dana tokopedia.
Terimakasih toko pedia sudah memberikan sarana, skil academi untuk platihan prakerja yang bisa di pahami dan mengerti secara cepat,di sini saya mendapatkan pengalaman yang baru „, sukses selalu toko pedia,,,	terimakasih toko pedia sudah memberikan sarana, skil academi untuk platihan prakerja yang bisa di pahami dan mengerti secara cepat,di sini saya mendapatkan pengalaman yang baru „, sukses selalu toko pedia,,,

Setelah melakukan proses *case folding* pada data teks, langkah selanjutnya adalah proses *tokenizing* dimana proses memecah kalimat menjadi kata-kata. Contoh data sebelum dan sesudah melalui tahap *stopword* dan *tokenizing* ditampilkan pada Tabel 4.3

Tabel 4.3 Proses *Tokenizing*

Sebelum proses	Sesudah proses
aplikasi yang mempermudah semuanya di masa pandemi ini, saya puas sekali dengan aplikasi tokped yang bekerja sama dengan pra kerja dan mempermudah saya beli pelatihan untuk pra kerja,yuk download toko pedia sekarang juga.	aplikasi yang mempermudah semuanya di masa pandemi ini, saya puas sekali dengan aplikasi tokped yang bekerja sama dengan pra kerja dan mempermudah saya beli pelatihan untuk pra kerja,yuk download toko pedia sekarang juga.

Sebelum proses	Sesudah proses
terpaksa saya kasih bintang 1, karena proses widraw dana dari hari jumat sampai sekarang belum cair hari senin juga proses widraw ulang belum cair juga sudah nunggu 1x28jam katanya akan cair 2x24 jam. kurang sabar lagi perbaiki dulu sistem withdrawal dana tokopedia.	terpaksa saya kasih bintang 1, karena proses widraw dana dari hari jumat sampai sekarang belum cair hari senin juga proses widraw ulang belum cair juga sudah nunggu 1x28jam katanya akan cair 2x24 jam. kurang sabar lagi perbaiki dulu sistem withdrawal dana tokopedia.
terimakasih toko pedia sudah memberikan sarana, skil academi untuk platihan prakerja yang bisa dipahami dan mengerti secara cepat,di sini saya mendapatkan pengalaman yang baru,,, sukses selalu toko pedia,,,	terimakasih toko pedia sudah memberikan sarana, skil academi untuk platihan prakerja yang bisa di pahami dan mengerti secara cepat,di sini saya mendapatkan pengalaman yang baru,,, sukses selalu toko pedia,,,

Setelah melakukan proses *tokenizing* pada data teks, langkah selanjutnya adalah proses *filtering*. *Filtering* pada data ulasan merupakan proses membersihkan ulasan dari kata yang tidak diperlukan untuk mengurangi noise. Kata yang dihilangkan dalam ulasan adalah karakter HTML, *emoticons*, *hastag* (#), *username* (@username), URL, dan menghilangkan *stopword*. Contoh data sebelum dan sesudah melalui tahap *filtering* ditampilkan pada Tabel 4.4

Tabel 4.4 Proses *Filtering*.

Sebelum proses	Sesudah proses
aplikasi yang mempermudah semuanya di masa pandemi ini, saya puas sekali dengan aplikasi tokped yang bekerja sama dengan pra kerja dan mempermudah saya beli	aplikasi mempermudah semuanya masa pandemi saya puas sekali aplikasi tokped bekerja sama pra kerja mempermudah saya beli pelatihan untuk pra kerjayuk download toko pedia sekarang

Sebelum proses	Sesudah proses
pelatihan untuk pra kerja,yuk download toko pedia sekarang juga.	
terpaksa saya kasih bintang 1, karena proses widraw dana dari hari jumat sampai sekarang belum cair hari senin juga proses widraw ulang belum cair juga sudah nunggu 1x28jam katanya akan cair 2x24 jam. kurang sabar lagi perbaiki dulu sistem withdrawal dana tokopedia.	terpaksa saya kasih bintang karena proses widraw dana hari jumat sampai sekarang belum cair hari senin proses widraw ulang belum cair sudah nunggu jam katanya akan cair jam kurang sabar lagi perbaiki dulu sistem withdrawal dana tokopedia
terimakasih toko pedia sudah memberikan sarana, skil academi untuk platihan prakerja yang bisa di pahami dan mengerti secara cepat,di sini saya mendapatkan pengalaman yang baru „, sukses selalu toko pedia„,,	terimakasih toko pedia sudah memberikan sarana skil academi platihan prakerja bisa pahami mengerti secara cepat sini saya mendapatkan pengalaman baru sukses selalu toko pedia

Proses stemming dilakukan untuk mendapatkan kata dasar dengan cara menghilangkan awalan, akhiran, sisipan, dan *fixes* (kombinasi dari awalan dan akhiran). Pada penelitian ini kamus stemming yang digunakan adalah *Liu's word list*. Berikut adalah contoh data sebelum dan sesudah melalui tahap *stemming* ditampilkan pada Tabel 4.5

Tabel 4.5 Proses Stemming.

Sebelum proses	Sesudah proses
aplikasi mempermudah semuanya masa pandemi saya puas sekali aplikasi tokped bekerja sama pra kerja mempermudah saya beli pelatihan	‘aplikasi’ ‘mempermudah’ ‘semuanya’ ‘masa’ ‘pandemi’ ‘saya’ ‘puas’ ‘sekali’ ‘aplikasi’ ‘tokped’ ‘bekerja’ ‘sama’ ‘pra’ ‘kerja’ ‘mempermudah’ ‘saya’ ‘beli’

Sebelum proses	Sesudah proses
untuk pra kerjayuk download toko pedia sekarang	'pelatihan' 'untuk' 'pra' 'kerjayuk' 'download' 'toko' 'pedia' 'sekarang'
terpaksa saya kasih bintang karena proses widraw dana hari jumat sampai sekarang belum cair hari senin proses widraw ulang belum cair sudah nunggu jam katanya akan cair jam kurang sabar lagi perbaiki dulu sistem withdrawal dana tokopedia	Terpaksa' 'saya' 'kasih' 'bintang' 'karena' 'proses' 'widraw' 'dana' 'hari' 'jumat' 'sampai' 'sekarang' 'belum' 'cair' 'hari' 'senin' 'proses' 'widraw' 'ulang' 'belum' 'cair' 'sudah' 'nunggu' 'jam' 'katanya' 'akan' 'cair' 'jam' 'kurang' 'sabar' 'lagi' 'perbaiki' 'dulu' 'sistem' 'withdrawal' 'dana' 'tokopedia'
terimakasih toko pedia sudah memberikan sarana skil academi platihan prakerja bisa pahami mengerti secara cepat sini saya mendapatkan pengalaman baru sukses selalu toko pedia	'Terimakasih' 'toko' 'pedia' 'sudah' 'memberikan' 'sarana' 'skil' 'academi' 'platihan' 'prakerja' 'bisa' 'pahami' 'mengerti' 'secara' 'cepat' 'sini' 'saya' 'mendapatkan' 'pengalaman' 'baru' 'sukses' 'selalu' 'toko' 'pedia'

Hasil dari proses *stemming* dipakai sebagai kata kunci dari data ulasan yang akan dianalisis dengan menggunakan metode NBC (*Naive Bayes Classification*). Tabel 4.6, Tabel 4.7 dan Tabel 4.8 menunjukkan perhitungan frekuensi tiap kata kunci pada situs belanja *online*.

Tabel 4.6 Frekuensi data setelah proses *stemming* Tokopedia.

Word	Total Occurences
tokopedia	219
pelatihan	141
prakerja	100
skill	79
academy	64

Word	Total Occurences
mudah	59
ulasan	59
terimakasih	58
kasih	48
semoga	47
terima	42
membantu	39
sukses	39
mengikuti	36
bermanfaat	33

Berdasarkan Tabel 4.6 menunjukkan perhitungan frekuensi data setelah proses *stemming* pada *online shop* Tokopedia terdapat beberapa kata kunci. Pada ulasan pertama didapatkan bahwa kata “tokopedia” disebutkan pada sebanyak 219 kali, disusul dengan kata “pelatihan” dengan 141 kali, kata “prakerja” sebanyak 100 kali, dan seterusnya.

Tabel 4.7 Frekuensi data setelah proses stemming Shopee.

Word	Total Occurences
shopee	253
ongkir	90
barang	79
ulasan	72
belanja	61
gratis	57
tolong	53
akun	39
pas	39
bayar	38
aplikasi	36
produk	35
bagus	32
banget	32
kecewa	31

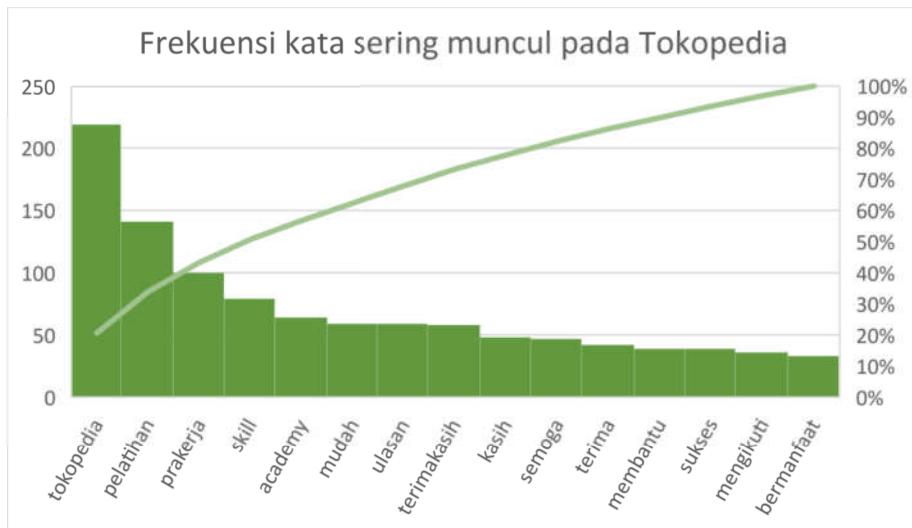
Berdasarkan Tabel 4.7 menunjukkan perhitungan frekuensi data setelah proses *stemming* pada *online shop* Shopee terdapat beberapa kata kunci. Pada ulasan pertama didapatkan bahwa kata “shopee” disebutkan pada sebanyak 253 kali, disusul dengan kata “ongkir” dengan 90 kali, kata “barang” sebanyak 79 kali, dan seterusnya.

Tabel 4.8 Frekuensi data setelah proses *stemming* Bukalapak.

Word	Total Occurrences
pelatihan	275
bukalapak	218
prakerja	175
mudah	118
mengikuti	97
materi	94
sukses	85
membeli	84
skill	78
ulasan	77
academy	73
membantu	72
peserta	70
kasih	60
terimakasih	56

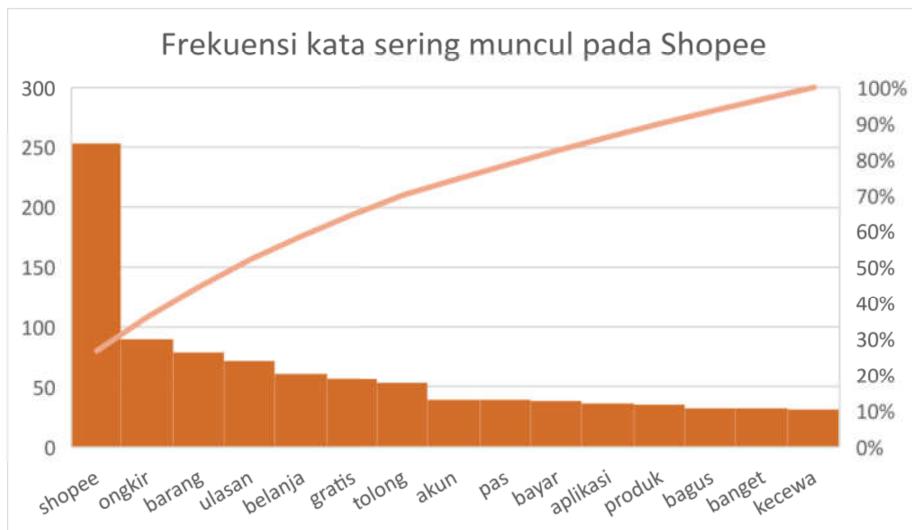
Berdasarkan Tabel 4.8 menunjukkan perhitungan frekuensi data setelah proses *stemming* pada *online shop* Bukalapak terdapat beberapa kata kunci. Pada ulasan pertama didapatkan bahwa kata “pelatihan” disebutkan pada sebanyak 275 kali, disusul dengan kata “bukalapak” dengan 218 kali, kata “prakerja” sebanyak 175 kali, dan seterusnya.

Frekuensi setiap kata dalam ulasan Tokopedia, Shopee, dan Bukalapak ditunjukkan pada Gambar 4.8, Gambar 4.9 dan Gambar 4.10.



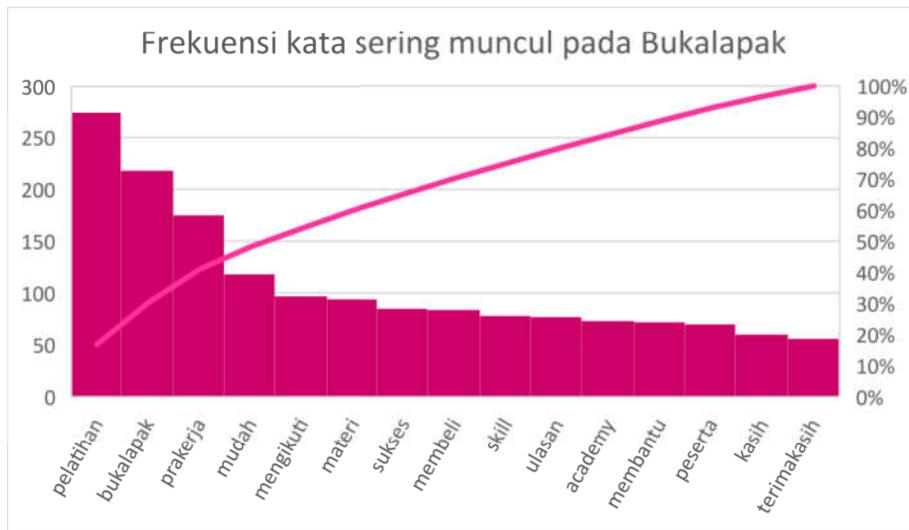
Gambar 4.8 Frekuensi data ulasan Tokopedia

Berdasarkan Gambar 4.8 dapat diketahui frekuensi dari 15 kata tertinggi situs belanja *online* Tokopedia. Kata “pelatihan” merupakan kata yang paling banyak dibicarakan oleh konsumen, sedangkan kata “prakerja” merupakan kata yang paling banyak kedua dibicarakan oleh konsumen. Hal tersebut dikarenakan bahwa Tokopedia kemungkinan pada saat ini banyak menjual jasa pelatihan daripada barang. Kata “skill” merupakan kata berikutnya yang paling banyak dibicarakan.



Gambar 4.9 Frekuensi data ulasan Shopee

Berdasarkan Gambar 4.9 dapat diketahui frekuensi dari 15 kata tertinggi situs belanja *online* Shopee. Kata “ongkir” merupakan kata yang paling banyak dibicarakan oleh konsumen, sedangkan kata “barang” merupakan kata yang paling banyak kedua dibicarakan oleh konsumen. Hal tersebut dikarenakan bahwa Shopee adalah *market place* yang banyak menjual barang daripada jasa. Kata “ulasan” merupakan kata berikutnya yang paling banyak dibicarakan.



Gambar 4.10 Frekuensi data ulasan Bukalapak

Berdasarkan Gambar 4.10 dapat diketahui frekuensi dari 15 kata tertinggi situs belanja *online* Bukalapak. Kata “pelatihan” merupakan kata yang paling banyak dibicarakan oleh konsumen, sedangkan kata “prakerja” merupakan kata yang paling banyak kedua dibicarakan oleh konsumen. Hal tersebut dikarenakan bahwa Tokopedia kemungkinan pada saat ini banyak menjual jasa pelatihan daripada barang. Kata “mudah” merupakan kata berikutnya yang paling banyak dibicarakan.

4.3 Klasifikasi Naïve Bayes (NBC)

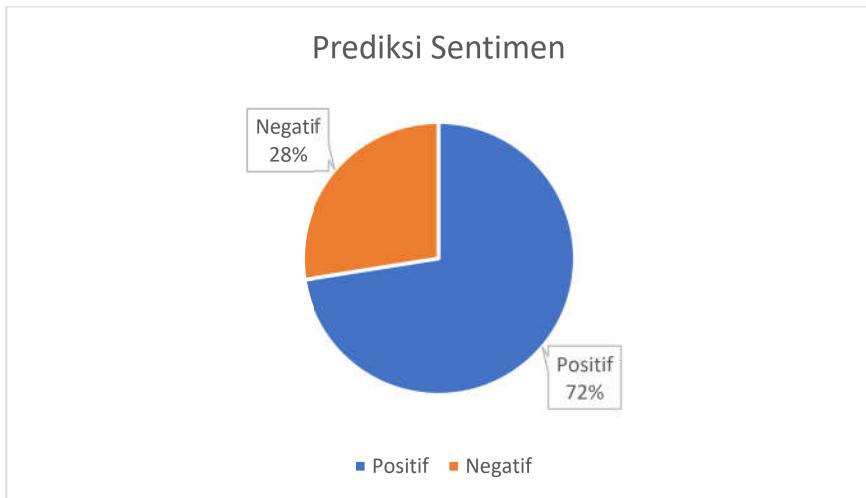
Naïve Bayes Classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema *Bayes*. Ciri

utama dari *Naïve Bayes Classifier* ini adalah asumsi yg sangat kuat (*naif*) akan independensi dari masing-masing kondisi/kejadian.

A. Klasifikasi *Naïve Bayes* pada *Online Shop Tokopedia*

Index	Nominal value	Absolute count	Fraction
1	Pos	87	0.725
2	Neg	33	0.275

Gambar 4.11 Prediksi sentiment *RapidMiner* pada Tokopedia



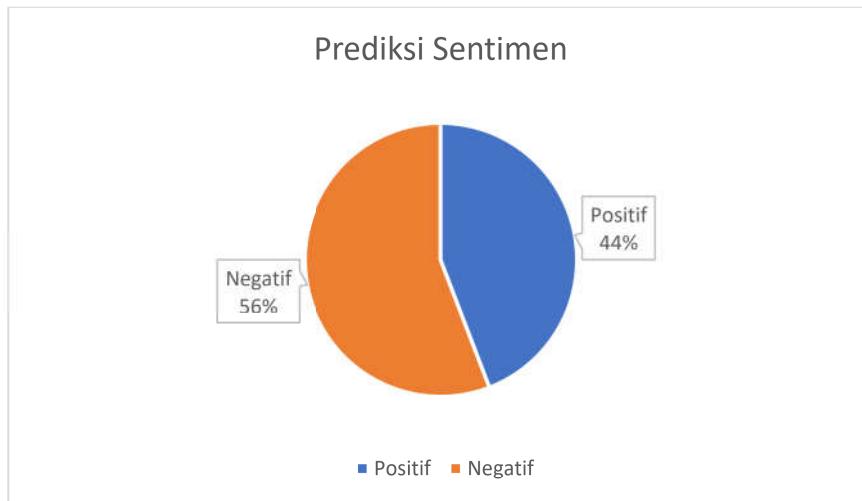
Gambar 4.12 Presentase nilai prediksi sentiment pada Tokopedia

Berdasarkan Gambar 4.11 dan 4.12 dapat diketahui presdiksi sentimen klasifikasi dengan metode *Naïve Bayes Classifier* (NBC) pada *online Shop Tokopedia*. Jika nilai pecahan sentimen negatif lebih tinggi jika dibandingkan dengan probabilitas sentimen positif, maka akan memberikan hasil keputusan negatif begitupula sebaliknya. Pada data ulasan Tokopedia, didapatkan prediksi nilai sentimen positif sebesar 87 dan prediksi nilai negative sebesar 33, Prediksi sentimen positif pada situs *online shop* Tokopedia memiliki persentase yang lebih bayak jika dibandingkan dengan persentase prediksi sentimen negatif. Persentase prediksi sentimen negatif pada situs *online shop* Tokopedia secara keseluruhan sebesar 28%, sedangkan persentase prediksi sentimen positif adalah sebesar 72%.

B. Klasifikasi *Naïve Bayes* pada *Online Shop* Shopee

Index	Nominal value	Absolute count	Fraction
1	Neg	67	0.558
2	Pos	53	0.442

Gambar 4.13 Prediksi sentimen *RapidMiner* pada Shopee



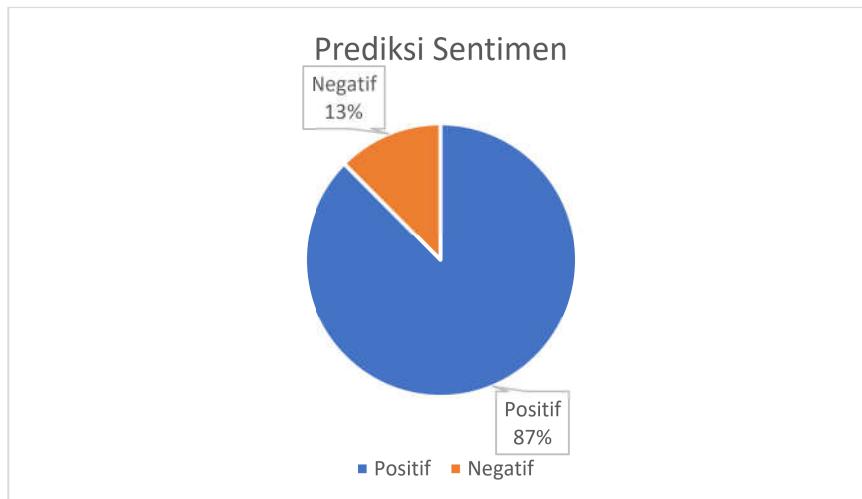
Gambar 4.14 Presentase nilai prediksi sentimen pada Shopee

Berdasarkan Gambar 4.13 dan 4.14 dapat diketahui presdiksi sentimen klasifikasi dengan metode *Naïve Bayes Classifier* (NBC) pada *online Shop* Shopee. Jika nilai pecahan sentimen negatif lebih tinggi jika dibandingkan dengan probabilitas sentimen positif, maka akan memberikan hasil keputusan negatif begitupula sebaliknya. Pada data ulasan Shopee, didapatkan prediksi nilai sentimen positif sebesar 53 dan prediksi nilai negative sebesar 67, Prediksi sentimen negatif pada situs *online shop* Shopee memiliki persentase yang lebih bayak jika dibandingkan dengan persentase prediksi sentimen positif. Persentase prediksi sentimen negatif pada situs *online shop* Shopee secara keseluruhan sebesar 56%, sedangkan persentase prediksi sentimen positif adalah sebesar 44%.

C. Klasifikasi *Naïve Bayes* pada *Online Shop* Bukalapak

Index	Nominal value	Absolute count	Fraction
1	Pos	105	0.875
2	Neg	15	0.125

Gambar 4.15 Prediksi sentiment *RapidMiner* pada Bukalapak



Gambar 4.16 Presentasc nilai prediksi sentiment pada Bukalapak

Berdasarkan Gambar 4.15 dan 4.16 dapat diketahui presdiksi sentimen klasifikasi dengan metode *Naïve Bayes Classifier* (NBC) pada *online Shop* Bukalapak. Jika nilai pecahan sentimen negatif lebih tinggi jika dibandingkan dengan probabilitas sentimen positif, maka akan memberikan hasil keputusan negatif begitupula sebaliknya. Pada data ulasan Bukalapak, didapatkan prediksi nilai sentimen positif sebesar 105 dan prediksi nilai negative sebesar 15, Prediksi sentimen positif pada situs *online shop* Bukalapak memiliki persentase yang lebih bayak jika dibandingkan dengan persentase prediksi sentimen negatif. Persentase prediksi sentimen negatif pada situs *online shop* Bukalapak secara keseluruhan sebesar 13%, sedangkan persentase prediksi sentimen positif adalah sebesar 87%.

Tabel 4.9 Hasil Perhitungan NBC Pada situs belanja.

Situs Belanja	Perhitungan <i>Naïve Bayes</i>	
	Positif	Negatif
Tokopedia	72%	28%
Shopee	44%	56%
Bukalapak	87%	13%

4.4 Uji Validitas Data dan Akurasi *Naïve Bayes*

Pengujian pada penelitian ini menggunakan *tools RapidMiner* untuk menghasilkan suatu output klasifikasi dan keakurasi data. *10-fold cross validation* digunakan dalam penelitian ini untuk menguji keakuratan data.

A. *10-Fold Cross Validation* pada *Online Shop Tokopedia*

Dalam penelitian ini, seluruh tahapan uji coba algoritma klasifikasi pada dataset Tokopedia yang telah diolah menggunakan *tools RapidMiner*, pada tahap uji coba menggunakan klasifikasi *Naive Bayes* dengan metode *Cross Validation* serta nilai *K-Fold* sebagai parameter, pada tahap ini nilai *K-Fold* yang digunakan ialah K=10 dengan hasil akurasi sebagai berikut:

accuracy: 92.50% +/- 7.30% (micro average: 92.50%)			
	true Pos	true Neg	class precision
pred. Pos	84	3	96.55%
pred. Neg	6	27	81.82%
class recall	93.33%	90.00%	

Gambar 4.17 Hasil pengujian berdasarkan algoritma *Naive Bayes* dengan *K-Fold* pada Tokopedia

Pada gambar 4.17 dapat diketahui bahwa pada tahap pengujian menggunakan algoritma *Naive Bayes* dengan nilai *K-Fold Cross Validation* 10 menghasilkan akurasi dengan jumlah 92,50% serta nilai *precision* nya dari masing-masing *prediction true* dan *prediction false* bernilai 96,55% dan 81,82% adapun *recall* dari masing-masing *prediction* bernilai 90,00% dan 93,33%.

B. 10-Fold Cross Validation pada Online Shop Shopee

Dalam penelitian ini, seluruh tahapan uji coba algoritma klasifikasi pada dataset Shopee yang telah diolah menggunakan *tools RapidMiner*, pada tahap uji coba menggunakan klasifikasi *Naive Bayes* dengan metode *Cross Validation* serta nilai *K-Fold* sebagai parameter, pada tahap ini nilai *K-Fold* yang digunakan ialah K=10 dengan hasil akurasi sebagai berikut:

accuracy: 67.50% +/- 10.72% (micro average: 67.50%)

	true Neg	true Pos	class precision
pred. Neg	46	24	65.71%
pred. Pos	15	35	70.00%
class recall	75.41%	59.32%	

Gambar 4.18 Hasil pengujian berdasarkan algoritma *Naive Bayes* dengan *K-Fold* pada Shopee

Pada gambar 4.18 dapat diketahui bahwa pada tahap pengujian menggunakan algoritma *Naive Bayes* dengan nilai *K-Fold Cross Validation* 10 menghasilkan akurasi dengan jumlah 67,50% serta nilai *precision* nya dari masing-masing *prediction true* dan *prediction false* bernilai 65,71% dan 70,00% adapun *recall* dari masing-masing *prediction* bernilai 59,32% dan 75,41%.

C. 10-Fold Cross Validation pada Online Shop Bukalapak

Dalam penelitian ini, seluruh tahapan uji coba algoritma klasifikasi pada dataset Bukalapak yang telah diolah menggunakan *tools RapidMiner*, pada tahap uji coba menggunakan klasifikasi *Naive Bayes* dengan metode *Cross Validation* serta nilai *K-Fold* sebagai parameter, pada tahap ini nilai *K-Fold* yang digunakan ialah K=10 dengan hasil akurasi sebagai berikut:

accuracy: 95.83% +/- 5.89% (micro average: 95.83%)			
	true Pos	true Neg	class precision
pred. Pos	102	3	97.14%
pred. Neg	2	13	86.67%
class recall	98.08%	81.25%	

Gambar 4.19 Hasil pengujian berdasarkan algoritma *Naive Bayes* dengan *K-Fold* pada Bukalapak

Pada gambar 4.19 dapat diketahui bahwa pada tahap pengujian menggunakan algoritma *Naive Bayes* dengan nilai *K-Fold Cross Validation* 10 menghasilkan akurasi dengan jumlah 95,83% serta nilai *precision* nya dari masing-masing *prediction true* dan *prediction false* bernilai 97,14% dan 86,14% adapun *recall* dari masing-masing *prediction* bernilai 81,25% dan 98,08%.

4.5 Rating Secara Umum

A. Berdasarkan Jumlah Pengguna

Berikut adalah paparan rating dari seluruh ulasan yang langsung diambil dari *Google Play Store*. Pada tahap ini akan dihitung seberapa besar jumlah pengguna yang didapatkan dari data yang telah diamati langsung dari *Google Play Store*:

Tabel 4.10 Jumlah pengguna pada setiap situs

Rank	Situs	Jumlah Pengguna
1	Shopee	5,537,919
2	Tokopedia	3,329,470
3	Bukalapak	1,820,885

Tabel 4.10 menjelaskan bahwa secara jumlah pengguna Shopee menempati peringkat 1 untuk *Online Shop* pada tahun 2020. Kemudian disusul dengan Tokopedia dan Bukalapak.

B. Berdasarkan Rating Aplikasi

Berikut adalah paparan rating dari seluruh ulasan yang langsung diambil dari *Google Play Store*. Pada tahap ini akan dihitung seberapa besar Rating yang didapatkan dari data yang telah diamati langsung dari *Google Play Store*:

Tabel 4.11 Total Rating aplikasi pada setiap situs

Rank	Situs	Rating
1	Tokopedia	4,8
2	Shopee	4,7
3	Bukalapak	4,7

Tabel 4.11 menjelaskan bahwa secara Rating aplikasi Tokopedia menempati peringkat 1 untuk *Online Shop* pada tahun 2020. Kemudian disusul dengan Shopee dan Bukalapak.

C. Berdasarkan Tingkat akurasi *Naïve Bayes*

Berikut adalah paparan tingkat akurasi *Naïve Bayes* dari seluruh ulasan yang langsung diambil dari *Google Play Store* dan telah diolah dengan *RapidMiner*. Pada tahap ini akan dihitung seberapa besar tingkat akurasi yang didapatkan dari data yang telah diambil secara *Crawling* dari *Google Play Store* dan diolah menggunakan *RapidMiner*:

Tabel 4.12 Tingkat akurasi *Naïve Bayes* pada setiap situs

Rank	Situs	Tingkat Akurasi
1	Bukalapak	95,83%
2	Tokopedia	92,50%
3	Shopee	67,50%

Tabel 4.12 menjelaskan bahwa secara tingkat akurasi Bukalapak menempati peringkat 1 untuk *Online Shop* pada tahun 2020. Kemudian disusul dengan Tokopedia dan Shopee.

Berdasarkan pemeringkatan yang dilakukan pada tahap ini, didapatkan bahwa setiap *online shop* memiliki keunggulan masing-masing menurut data dari *Google Play Store* dan metode perhitungan *naive bayes*. Berikut tampilan pemeringkatan, dapat dilihat pada tabel 4.12

Tabel 4.13 Pemeringkatan Keseluruhan.

Kategori	Tokopedia	Shopee	Bukalapak
Jumlah Pengguna	2	1	3
Rating Aplikasi	1	2	3
Tingkat Akurasi	2	3	1

Berdasarkan tabel 4.13, didapatkan bahwa *online shop* Tokopedia memimpin sebagai *online shop* yang paling disarankan diantara Shopee dan Bukalapak menurut data dari *Google Play Store* dan metode perhitungan *Naive Bayes*.

4.6 Visualisasi Data dengan *WordCloud*

Pendapat dari konsumen belanja *online* di Tokopedia, Shopee dan Bukalapak dapat divisualisasikan dalam bentuk *Wordcloud*. *Wordcloud* merupakan representasi grafis dari sebuah dokumen yang dilakukan dengan plotting kata-kata yang sering muncul pada sebuah dokumen pada ruang dua dimensi. Frekuensi dari kata yang sering muncul ditunjukkan melalui ukuran huruf kata tersebut. Semakin besar ukuran kata menunjukkan semakin besar frekuensi kata tersebut muncul dalam dokumen. Berikut adalah tampilan visualisasi data dengan *Wordcloud* berdasarkan hasil uji yang telah dilakukan.

A. Tokopedia



Gambar 4.20 Visualisasi *Wordcloud* Tokopedia.

Berdasarkan Gambar 4.20 dapat diketahui bahwa tiga kata yang sering muncul pada akun Tokopedia adalah tokopedia, pelatihan, dan prakerja. Peningkatan pelayanan dan respon positif kepada pengguna terhadap pelatihan prakerja dari *online shop* Tokopedia perlu dipertahankan dan diberi apresiasi. Selain respon positif dari pengguna, respon negatif juga perlu diketahui oleh Tokopedia untuk meningkatkan pelayanan konsumen.

B. Shopee



Gambar 4.21 Visualisasi *Wordcloud* Shopee

Berdasarkan Gambar 4.21 dapat diketahui bahwa tiga kata yang sering muncul pada akun Shopee adalah shopee, ongkir, dan barang. Peningkatan pelayanan dan respon positif kepada pengguna terhadap ongkos kirim suatu barang dari *online shop* Shopee perlu dipertahankan dan diberi apresiasi. Selain respon positif dari pengguna, respon negatif juga perlu diketahui oleh Shopee untuk meningkatkan kepuasan dalam berbelanja.

C. Bukalapak



Gambar 4.22 Visualisasi *Wordcloud* Bukalapak

Berdasarkan Gambar 4.22 dapat diketahui bahwa tiga kata yang sering muncul pada akun Bukalapak adalah pelatihan, bukalapak dan prakerja. Peningkatan pelayanan dan respon positif kepada pengguna terhadap pelatihan prakerja dari *online shop* Bukalapak perlu dipertahankan dan diberi apresiasi. Selain respon positif dari pengguna, respon negatif juga perlu diketahui oleh Bukalapak untuk meningkatkan pelayanan konsumen.

BAB V

PENUTUP

Berdasarkan hasil analisis yang telah diperoleh, kesimpulan dan saran yang didapatkan dari penelitian ini adalah sebagai berikut.

5.1 Kesimpulan

Berdasarkan hasil analisis yang telah diperoleh, kesimpulan yang didapatkan dari penelitian ini adalah sebagai berikut.

1. Konsumen yang memberikan tanggapan, pendapat, kritik, saran dan masalah *complain* lebih banyak ditujukan pada online shop Shopee dari pada Tokopedia dan Bukalapak. Sementara itu, tingkat akurasi data pada *online shop* Shopee kurang begitu baik, sehingga memungkinkan untuk menggunakan uji yang lain seperti *Random Forest* ataupun *Support Vector Machine*.
2. Berdasarkan visualisasi *wordcloud*, penulis mengambil 3 sentimen positif teratas yang banyak dibicarakan oleh konsumen pada *online shop* :
 - a) Tokopedia : pelatihan, prakerja, skill
 - b) Shopee : ongkir, barang, ulasan
 - c) Bukalapak : pelatihan, prakerja, mudah
3. *Online Shop* yang paling popular menurut penelitian ini adalah Bukalapak dengan rata-rata penghitungan manual, dengan nilai sentimen positif 87%, disusul dengan Tokopedia dengan nilai sentimen positif 75%, kemudian di posisi paling akhir adalah Shopee dengan nilai sentimen positif 52%
4. Dengan uji validitas metode NBC menggunakan *10-fold cross validation* didapatkan tingkat akurasi untuk Tokopedia 92,50%, Shopee 67,50% dan Bukalapak 95,83%. Artinya, metode NBC tidak sepenuhnya memiliki tingkat akurasi tinggi dalam klasifikasi sentimen analisis.
5. Pada penelitian ini Tokopedia merupakan *online shop* yang paling popular dibandingkan dengan Bukalapak dan Shopee. Berdasarkan perhitungan pemeringkatan keseluruhan didapatkan untuk online shop

Tokopedia mendapatkan ranking 1, disusul dengan Shopee dan Bukalapak.

6. Berdasarkan pemeringkatan yang dilakukan, didapatkan bahwa *online shop* Tokopedia memimpin sebagai *online shop* paling populer diantara Shopee dan Bukalapak menurut data dari ulasan *Google Play Store* dan metode perhitungan *Naive Bayes*.

5.2 Saran

Berdasarkan hasil analisis yang telah dijelaskan, maka saran yang dapat diberikan pada penelitian yang akan datang adalah sebagai berikut.

1. Pada penelitian klasifikasi teks, membutuhkan data sentimen yang cukup besar untuk mengurangi resiko dari imbalanced data. Pengklasifikasian sentimen awal sebaiknya dilakukan dengan banyak orang sehingga dapat mengurangi obyektifitas pada hasil sentimen.
2. Pada proses *stemming*, perlu merubah kata tidak baku yang digunakan untuk bahasa sehari-hari dikarenakan pada data ulasan banyak menggunakan bahasa sehari-hari. Apabila ditemukan sedikit kata tidak baku maka dapat merubah data secara manual saja, jika sebaliknya maka harus menggunakan tools khusus.
3. Penulis menyarankan untuk tidak menggunakan 1 metode pada penelitian yang akan datang dikarenakan metode NBC belum tentu memiliki tingkat akurasi seperti yang diharapkan dari pada metode lainnya.

DAFTAR PUSTAKA

- Aprilla Dennis. (2013). Belajar Data Mining dengan RapidMiner. *Innovation and Knowledge Management in Business Globalization: Theory & Practice, Vols 1 and 2*, 5(4), 1–5. Retrieved from http://esjournals.org/journaloftechnology/archive/vol1no6/vol1no6_6.pdf%5Cnhttp://www.airccse.org/journal/nsa/5413nsa02.pdf
- Bentley, W. &. (2007). *System Analysis & Design Methods Seventh Edition*. 317.
- Cnnindonesia.com. (2019). Daftar Toko Online yang Paling Menguasai Pasar RI Selama 2019. Retrieved August 25, 2020, from <https://www.cnnindonesia.com/teknologi/20200124205259-206-468540/daftar-toko-online-yang-paling-menguasai-pasar-ri-selama-2019>
- Fatmawati, V., & Imron, A. (2017). Intuisi 9 (1) (2017). *Intuisi Jurnal Psikologi Ilmiah*, 9(1), 26–38.
- Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71. <https://doi.org/10.26623/transformatika.v18i1.2317>
- Gamallo, P., & Garcia, M. (2015). *Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets*. <https://doi.org/10.3115/v1/s14-2026>
- 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)*, 5(2), 293. <https://doi.org/10.24114/cess.v5i2.18186>
- Imam, A., & Fajriab, H. (2015). Implementasi Text Mining pada Mesin Pencarian Twitter untuk Menganalisis Topik - Topik Terkait “KPK dan Jokowi.” *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika UMS 2015*, 570–581.
- Inews.id. (2020). 10 E-Commerce Terpopuler di Indonesia. Retrieved August 25, 2020, from <https://www.inews.id/finance/bisnis/10-e-commerce-terpopuler-di-indonesia-tokopedia-terdepan-shopee-geser-bukalapak>
- Jyoti, & Rao, S. (2016). A survey on sentiment analysis and opinion mining. In *ACM International Conference Proceeding Series* (Vol. 12-13-Augu). <https://doi.org/10.1145/2979779.2979832>
- Liu, B. (2012). Sentiment analysis and opinion mining. In *Synthesis Lectures on Human Language Technologies* (Vol. 5). <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Makinudin, & Hadiyanto. (2006). *Analisis Sosial*. Bandung: Yayasan Akatiga.

- McLeod, P. (2008). *Sistem Informasi Manajemen*.
- Mittal, A., Nagar, A., Gupta, K., & Nahar, R. (2015). Comparative Study of Various Frequent Pattern Mining Algorithms. *Ijarcce*, 4(4), 550–553. <https://doi.org/10.17148/ijarcce.2015.44127>
- Mukhlash, I., Yuanda, D., & Iqbal, M. (2018). Mining fuzzy time interval periodic patterns in smart home data. *International Journal of Electrical and Computer Engineering*, 8(5), 3374–3385. <https://doi.org/10.11591/ijece.v8i5.pp3374-3385>
- Murnawan, M. (2017). Pemanfaatan Analisis Sentimen Untuk Pemeringkatan Popularitas Tujuan Wisata. In *Jurnal Penelitian Pos dan Informatika* (Vol. 7). <https://doi.org/10.17933/jppi.2017.070203>
- Nurwahyuni, S., Bina, U., & Informatika, S. (2019). Online KRL access menggunakan metode Naive Bayes. *Swabumi*, 7(1), 31–38.
- Republikpos.com. (2016). Profil perusahaan tokopedia dan kisah pendirinya. Retrieved August 25, 2020, from <http://republikpos.com/2016/10/profilperusahaan-tokopedia-dan-kisah-pendirinya>
- Ruhyan, N. (2019). Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil / Genap Pada Twitter Dengan Metode Klasifikasi Naive Bayes. *Jurnal IKRA-ITH Informatika*, 3(1), 94–99.
- Silviosilviando. (2016). Analisis web e-commerce bukalapak. Retrieved August 25, 2020, from <https://silviosilviando.wordpress.com>.
- Surohman, S., Aji, S., Rousyati, R., & Wati, F. F. (2020). Analisa Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes Classifier Dan KNearest Neighbor. *EVOLUSI : Jurnal Sains Dan Manajemen*, 8(1), 93–105. <https://doi.org/10.31294/evolusi.v8i1.7535>
- Syadid, F. (2019). Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf- Idf) Dan Metode Multi Layer Perceptron (Mlp) Neural Network. *Skripsi Universitas Islam Negeri Syarif Hidayatullah Jakarta*, 1–89.
- Tokopedia.com. (2020). Tentang Tokopedia. Retrieved August 25, 2020, from <https://www.tokopedia.com/about?lang=id>
- Triawati, C. (2009). Metode Pembobotan Statistical Concept Based untuk Klastering dan Kategorisasi Dokumen Berbahasa Indonesia. In *Universitas*

Telkom. Institut Teknologi Telkom Bandung.

- Wardani, F. K., Hananto, V. R., & Nurcahyawati, V. (2019). Analisis Sentimen Untuk Pemerikatan Popularitas Situs Belanja Online Di Indonesia Menggunakan Metode Naive Bayes (Studi Kasus Data Sekunder). *Jsika*, 08(01), 1–9.
- Wardhana, S. R. (2017). Analisis Sentimen Pada Opini Pengguna Aplikasi Mobile Untuk Evaluasi Faktor Kebergunaan. *Tesis*.
- Wartaekonomi.co.id. (2020). Tiga Tantangan yang Dihadapi E-Commerce Shopee. Retrieved August 25, 2020, from <https://www.wartaekonomi.co.id/read134357/tiga-tantangan-yang-dihadapi-ecommerce-shopee.html>
- Wong, J. (2010). *Internet Marketing for Beginners*. Jakarta: PT Elex Media Komputindo.
- Xhemali, D., J. Hinde, C., & G. Stone, R. (2009). Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages. *International Journal of Computer Science*, 4(1), 16–23. Retrieved from <http://cogprints.org/6708/>