



Analisis Sentimen Pelayanan Pelanggan Mini Market Alfamart Pada Media Sosial Twitter Dengan Naïve Bayes Classifier

* Awaludin Aziz¹, Agung Budi Susanto², Sudarno Wiharjo³

^{1,2,3} Program Studi S2 Teknik Informatika, Program Pascasarjana, Universitas Pamulang, Banten

Email: aziz.mcgr@gmail.com

ABSTRACT

Twitter is one of the social media that is currently popular, here the public is free to have opinions, write and comment anything. PT Sumber Alfaria Trijaya with its trademark Alfamart is a company engaged in the retail sector. Not infrequently consumers submit complaints, criticisms and suggestions through this social media. Community opinion can be used as evaluation material in improving services. In this study, sentiment analysis for Alfamart minimarket customer service was carried out based on data obtained from Twitter. This sentiment analysis aims to classify Alfamart's customer service tweets into positive, negative and neutral sentiments using the naive Bayes classifier algorithm. The data used is 2000 tweet data and then preprocessing is carried out so that 1691 tweets are clean data. Of the 1691 data analyzed, 1017 positive tweets, 297 negative tweets and 377 neutral tweets were obtained. Then the data will be divided into 80% training data and 20% test data. The results of the accuracy value are 70% with a Precision value of 70%, a Recall value of 70% and an F1-Score value of 66%.

Keywords: Alfamart; Text Classification; Twitter; Naïve Bayes Classifier.

ABSTRAK

Twitter adalah merupakan salah satu media sosial yang saat ini sedang populer, disini public bebas beropini, menulis serta berkomentar apapun. PT Sumber Alfaria Trijaya dengan merek dagangnya Alfamart adalah merupakan perusahaan yang bergerak dalam bidang retail. Tidak jarang konsumen menyampaikan keluhan, kritikan serta sarannya melalui sosial media ini. Opini masyarakat ini dapat dijadikan sebagai bahan evaluasi dalam meningkatkan pelayanan. Pada penelitian ini dilakukan analisis sentiment pelayanan pelanggan minimarket alfamart berdasarkan data yang diperoleh dari twitter. Analisis sentimen ini bertujuan untuk mengklasifikasikan tweet pelayanan pelanggan Alfamart ke dalam sentimen positif, negatif dan netral dengan menggunakan algoritma *naive bayes classifier*. Data yang digunakan sebanyak 2000 data tweet kemudian dilakukan *preprocessing* sehingga didapatkan data bersih 1691 tweet. Dari 1691 data yang di analisis, didapatkan 1017 tweet positif, 297 tweet negatif dan 377 tweet netral. Kemudian data akan di bagi menjadi 80% data latih dan 20% data uji. Hasil nilai akurasi sebesar 70% dengan nilai *Precision* sebesar 70 %, nilai *Recall* sebesar 70% dan nilai *F1-Score* sebesar 66%.

Kata kunci: Alfamart; Klasifikasi Teks; Twitter ; Naive Bayes Classifier.

1. PENDAHULUAN

Minimarket merupakan jenis bisnis retail yang sedang berkembang pesat, yang menjual barang kebutuhan sehari-hari dan menawarkan kemudahan karena lokasi atau letak yang bisa dijangkau dan dekat dengan konsumen [1]. Selain itu minimarket juga mengutamakan kepraktisan dan kecepatan yang didukung dari luas toko atau gerai yang

tidak terlalu besar sehingga konsumen berbelanja dengan waktu yang tidak terlalu lama. Keuntungan lain dengan belanja di Minimarket yaitu suasana yang nyaman dan aman dalam berbelanja, tidak sulit memilih barang-barang yang diperlukan, kualitas barang lebih terjamin bila dibandingkan belanja di pasar tradisional, harga pasti sehingga tidak perlu ditawar dan dapat berbelanja berbagai kebutuhan dalam satu tempat saja sehingga menghemat tenaga dan waktu [2].

PT Sumber Alfaria trijaya adalah merupakan perusahaan ritel terbesar di Indonesia dengan nama brand Alfamart, alfamart menyediakan semua kebutuhan dan keperluan sehari-hari dengan harga yang terjangkau. Komitmen di dalam menjaga kualitas produk dan pelayanan yang terbaik serta tempat yang nyaman untuk berbelanja. Pembukaan gerai pertama Alfamart pada tahun 1999 di jalan beringin, Tangerang, Banten sebagai cikal bakal Alfamart. Alfamart terus tumbuh dan berkembang saat ini sudah memiliki 32 gudang untuk mendukung lebih dari 16.543 gerai, terbagi toko reguler 12.329 dan *franchise* 4.214 yang tersebar di seluruh wilayah tanah dengan melayani lebih dari 8 juta pelanggan setia setiap harinya [3].

Memberikan Kepuasan kepada pelanggan dengan berfokus pada produk dan pelayanan yang berkualitas unggul adalah merupakan salah satu dari misinya, untuk tetap bisa mempertahankan loyalitas pelanggan, maka faktor kepuasan dan pelayanan menjadi salah satu hal yang paling penting untuk dicermati. Kepuasan adalah perasaan senang atau kecewa seseorang yang berasal dari perbandingan antara kesannya terhadap kinerja (atau hasil) suatu produk dan harapan-harapannya. Hasilnya adalah kesetiaan pelanggan yang tinggi [4].

Dewasa ini kehidupan masyarakat mengalami banyak perubahan akibat kemajuan ilmu pengetahuan dan teknologi yang sangat pesat. Komunikasi dengan menggunakan media baru atau teknologi internet dapat menjangkau langsung dan cepat kepada semua pihak yang saat ini telah dinikmati secara luas. Teknologi internet adalah salah satu bukti yang bisa kita lihat sebagai pesatnya perkembangan teknologi di bidang informasi dan telekomunikasi. Perkembangan internet di Indonesia telah menunjukkan perkembangan yang signifikan. Menurut Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) jumlah pengguna internet sampai tahun 2020-Q2 mencapai 196,71 Juta orang [5]. dalam hal ini perusahaan akan membutuhkan feedback dari pelanggan

yang merupakan sumber informasi untuk meningkatkan kinerja pelayanan terhadap pelanggan. Feedback yang dibutuhkan dari pelanggan tidak hanya bersifat positif, melainkan juga bersifat negatif. Pelanggan seringkali memberikan keluhan-keluhan terhadap pelayanan yang telah diberikan atau memberikan usulan-usulan atau sekedar ingin mengetahui informasi yang ada saat ini. Salah satu cara untuk mendapatkan feedback dari pelanggan adalah dengan menggunakan media sosial. Berdasarkan data statistik dari sosmedtoday.com hingga Agustus 2014 tentang penggunaan media sosial di Indonesia, terdapat sekitar 70 juta masyarakat Indonesia yang aktif dalam media sosial. Hal ini bisa dibayangkan bahwa kecenderungan masyarakat cukup tinggi dalam mengakses media sosial yang ada saat ini.

Banyak cara yang disediakan oleh perusahaan untuk menampung partisipasi pelanggan baik dalam bentuk keluhan, informasi maupun saran dalam hal ini media sosial yang digunakan adalah twitter @alfamart dan @sahabatalfamart, data-data berupa keluhan, informasi atau saran yang ada pada media sosial Alfamart dapat dijadikan sebagai sumber data untuk memantau hasil feedback dari pelanggan [6]. Akan tetapi data-data yang disajikan dari hasil media sosial masih berbentuk terpecah belah sehingga butuh proses untuk dapat memantau hasil feedback dari pelanggan selain itu juga pemanfaatan data dari media sosial Twitter terkendala dengan sejumlah kesulitan teknis dalam proses pengumpulannya, pengolahannya, serta analisisnya.

Opini pelanggan yang diberikan biasanya bersifat positif, negatif atau netral. Opini yang diberikan oleh pelanggan inilah yang dapat digunakan sebagai bahan evaluasi oleh perusahaan untuk melakukan perbaikan dan juga dapat meningkatkan kualitas kerja yang baik ke depannya. Data-data opini pelanggan bisa kita olah dengan teknik text mining. Pengelompokan opini-opini bisa dilakukan dengan teknik text mining yaitu menggali informasi dari sebuah sumber data. Penggalian data biasanya menggunakan 2 metode yaitu klasifikasi dan Klustering. Klasifikasi digunakan untuk memprediksi suatu kelas untuk memetakan masing-masing data ke dalam target kelas dengan tujuan memprediksi secara akurat target kelasnya. Klasifikasi yang akan dilakukan dalam penelitian ini menggunakan metode naive bayes classifier. Algoritma Naïve Bayes Classifier digunakan pada analisis sentimen dikarenakan algoritma ini bertujuan sebagai metode klasifikasi ke dalam kategori positif, negatif dan netral.

Metode Naive Bayes adalah sebuah metode yang digunakan untuk menghitung probabilitas terjadinya suatu peristiwa berdasarkan pengaruh yang didapat dari hasil observasi.

Tujuan dari penelitian ini adalah untuk Menghasilkan klasifikasi opini pelanggan ke dalam 3 kelas yaitu positif, negatif dan netral menggunakan algoritma Naive Bayes, Mengetahui sentimen pelanggan di mini market Alfamart, dan Mengetahui tingkat akurasi hasil model klasifikasi komentar pelanggan.

2. METODE

Data yang digunakan adalah data sekunder. Data Sekunder adalah data yang diperoleh dan digunakan dalam bentuk teoritis yang didapat dari bahan-bahan kuliah, buku-buku, referensi yang relevan, serta dari hasil penjelelahan (*browsing*) di internet yang berkaitan dengan penelitian. Data yang akan diolah sebanyak 100 data sampling. Penelitian dilakukan dengan penelusuran literatur dari buku dan data yang bersumber dari alfmart maupun sumber-sumber buku atau jurnal lainnya. Observasi juga dilakukan pada Twitter dengan jumlah data yang digunakan sebanyak 1200 data tweet, yang diambil dari akun @alfamart menggunakan API Twitter.

3. HASIL DAN PEMBAHASAN

3.1. Pengambilan Data (*Crawling Data*)

Data yang terkumpul sebanyak 2000 tweet dikumpulkan dan disimpan kedalam file Csv (Command Separated Values) dengan format .csv. Tidak ada batasan dalam jumlah data yang diambil namun dengan 2000 data diharapkan dapat mewakili hasil opini masyarakat secara umum. Data yang tersimpan terdiri dari tanggal pembuatan tweet, id,user twitter yang memposting tweet.

no	tanggal	id	user	text
0	2022-05-30 23:18:50+00:00	1531414865713459200	alfamart	5. Keputusan yang dikeluarkan oleh tim Alfamart bersifat mutlak & tidak bisa diganggu gugat! Semakin BANYAK JAWAB, semakin BANYAK kesempatan untuk menang! Tidak Privat Akun#ApikasBelanjaOnlineAlfamart #KuisInteraktif #Giveaway #GercepTanpaBatas
1	2022-05-30 23:18:42+00:00	1531414829344641024	alfamart	1. FOLLOW Twitter @alfamart dan @alfagit_idn2 LOVE dan RETWEET Postingan ini! TULIS jawaban dan sertakan juga hashtag #KuisAlfamart4. Periode Kuis ini hingga 5 Juni 2022 pkl 21:00 WIB!ApikasBelanjaOnlineAlfamart #KuisInteraktif #Giveaway #GercepTanpaBatas
2	2022-05-30 23:17:49+00:00	1531414606761308162	alfamart	Pssst! #KuisAlfamart terbaru sudah hadir lho dipagi hari ini! 🤗 Yang mau menangkan hadiah A... voucher total senilai Rp 225.000, jangan sampai gak sudeh ya! Mumpung masih pekan gajian. Kita - kira masih inget gak nih pas gajian belinya apa aja?Hayoo 🤗🤗🤗 https://t.co/V7LWZHTMY7
3	2022-05-30 16:27:06+00:00	1531311249300258816	alfamart	Pssst! Yang belum bobo, coba nih di isi dulu Kali aja kesempatan ya kaaaan? 🤗 https://t.co/ra1QvgNYRth
4	2022-05-30 13:54:05+00:00	1531272738815746053	alfamart	@oohheni Slap Kak, untuk dapat lebih jelasnya dapat diteruskan melalui DM ya Kak :) ^-RF
...
1995	2021-07-16 02:14:53+00:00	1415857437874130946	Deppy_chan	@sahabatalfamart Saya dpt notr email untuk isi survey pontia tapi di aplikasi nya ga tersedia survei itu gimana? Instagram pontia juga udah hang. Bingung mau ngadu kmn
1996	2021-07-15 12:28:03+00:00	1415649359530582018	lilyaldr	@sahabatalfamart min ini gmn voucher diskon brnya galba dipake, padahal bin pernah dipake. Kalo di klik ada bacaan voucher sudah digunakan, saya cek di histori pesanan bin pernah pake voucher potongan brb. https://t.co/EW7GBITth
1997	2021-07-15 00:04:42+00:00	1415462290229370880	mahar_nuriman	@sahabatalfamart Sudah sy dm.
1998	2021-07-14 14:08:04+00:00	1415312142304112642	boykead	@sahabatalfamart @irpan17164995 @alfamart Alalmnya ada dibagian atas skrnnya min... Dibawah ada nomor HP nya juga... Telpn aja langsung, ngak usah nanyaz lagi
1999	2021-07-14 07:44:46+00:00	1415215680995085761	AntaAnwar_	@sahabatalfamart biaya admin top up link aja di alfamart berapa min?

2000 rows x 5 columns

Gambar 1. Hasil Crawling Data

3.2. *Preprocessing*

Untuk mengerjakan proses *preprocessing* terdapat 4 tahapan proses untuk memperoleh hasil yang maksimal, sebagai berikut:

3.2.1. *Cleaning*

Keseluruhan dari proses *cleaning* dilakukan dengan menggunakan *regex* atau bisa juga disebut *regular expression*. *Regex* adalah konstruksi dalam suatu bahasa untuk mencocokkan teks berdasarkan pola tertentu, terutama untuk kasus-kasus kompleks.

3.2.2. *Case Folding*

Case Folding adalah proses merubah data tweet menjadi *lowercase*.

3.2.3. *Tokenizing*

Tokenizing pada penelitian ini adalah tahapan dalam memecah *string* atau input terhadap suatu teks yang telah melewati tahap sebelumnya yaitu *case folding* berdasarkan perkata yang menyusunnya dan menghilangkan *URL*, *@mention* dan *hashtag*.

3.2.4. *Stopword Removal (Filtering)*

Filtering merupakan tahapan *text preprocessing* dimana bertujuan untuk pembersihan *teks* dari *stopword*. *Stopword* adalah kata-kata yang tidak memiliki makna. Pada proses *filtering* untuk menghapus *stopword* digunakan *library NLTK*, dimana dalam *library* tersebut hanya terdapat 758 kata *stopword*. *Stopword* tambahan dapat ditambahkan secara manual perkata ke dalam *list stopwords* dengan menggunakan *fungsi .extend()*, penambahan *stopword* juga dapat dibuat terlebih dahulu kamus *stopword* dalam file *txt*. Contoh *stopword* dalam bahasa Indonesia adalah “yang”, “untuk”, “dan”, “pada”, “dengan”.

3.2.5. *Stemming*

Stemming adalah tahap mencari root (dasar) kata dari tiap kata hasil *filtering* dengan menghapus kata imbuhan di depan maupun di belakang kata. Tahap *stemming* dilakukan dengan menggunakan bantuan *library* pada bahasa pemrograman Python3 yang bernama *Sastrawi*.

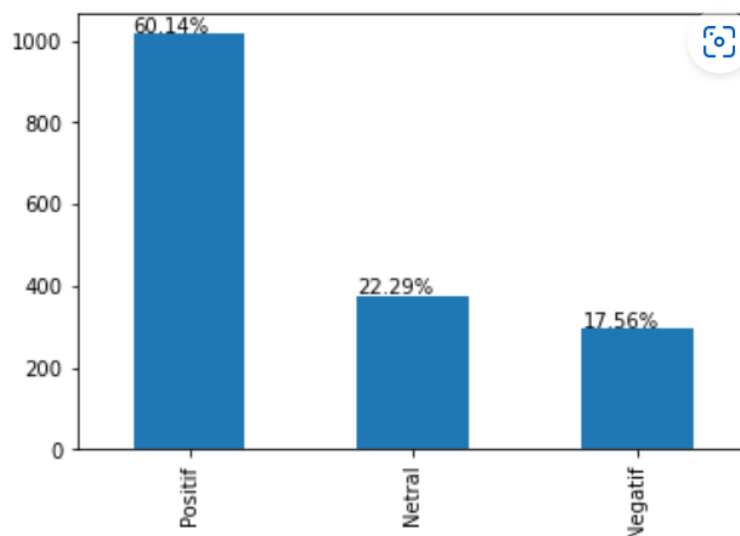
3.3. Pelabelan Data

Pelabelan data hasil *crawling* serta telah melewati tahapan *preprocessing* yang dilakukan dengan menggunakan *library python VaderSentiment* dengan cara melihat *polarity, subjectivity* yang dipunyai oleh teks tweet yang sudah dikumpulkan. *VaderSentiment* merupakan salah satu *library* yang disediakan oleh *Python* buat pemrosesan data dibidang *Natural Language Processing* yang bisa memberikan tag kata, ekstraksi kata, penerjemahan kata serta *sentiment analysis*.

Tabel 1. Hasil Pelabelan Data dengan *VaderSentiment*

Positif	Netral	Negatif	Jumlah
1017	377	297	1691
60,14%	22,29%	17,56%	100%

Didapatkan hasil akhir dari pelabelan dengan menggunakan *library VaderSentiment* sebanyak 1691 data tweet adalah 1017 tweet yang masuk dalam kelas positif, 377 tweet kelas netral dan 297 tweet kelas negatif. Persentase labeling data dengan *VaderSentiment* dapat dilihat pada Gambar 2



Gambar 2. Hasil persentase *labelling* data tweet dengan *VaderSentiment*

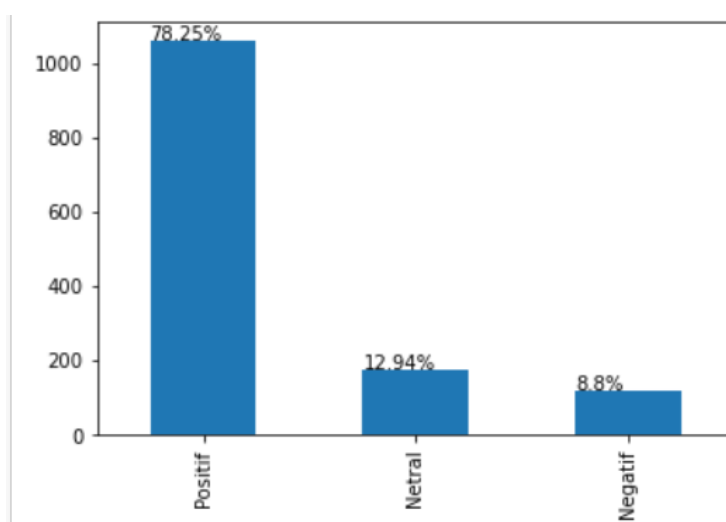
3.4. Ekstraksi Fitur (TF-IDF)

Data yang telah melewati *text preprocessing* masih berupa teks(kata), untuk analisis klasifikasi data wajib numerik ataupun angka. Data akan dikonversi terlebih dulu ke dalam bentuk angka dengan memakai pembobotan kata dengan TF- IDF.

3.5. Implementasi Klasifikasi Naïve Bayes

Setelah melewati *praprocessing* data dan *vectorizer* selanjutnya akan dibentuk model yang akan digunakan untuk melakukan klasifikasi pada data uji. Proses ini dijalankan dengan menggunakan *library* pada bahasa pemrograman *Python3* yang bernama *scikit-learn* untuk proses klasifikasi. Pada proses klasifikasi digunakan data uji sebanyak 20% dari data keseluruhan.

Gambar 3

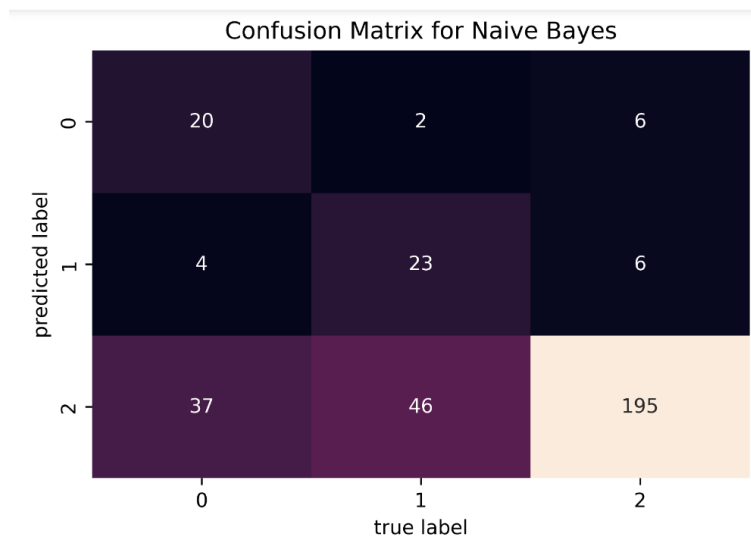


Gambar 3. Data training Naïve Bayes

Dari Gambar 3 data diatas dapat di ketahui untuk data untuk Training sebesar 1352 terbagi menjadi 78.25% data positif, 12.94% negative dan 8.8 % Netral

3.6. Pengujian Model Klasifikasi Naïve Bayes

Untuk mengetahui performa dari Algoritma *Naive Bayes*, dapat dilakukan pengujian terhadap model yang telah dibuat. Hasil klasifikasi akan divisualisasi dalam bentuk *confusion matrix*. *Confusion matrix* merupakan salah satu metode yang dapat digunakan untuk pengukuran kinerja suatu metode klasifikasi. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Berikut adalah pengujian model klasifikasi dengan menggunakan *library python sklearn.metric* yang didalamnya memiliki *confusion_matrix* dan divisualisasi dalam dengan menggunakan *seaborn* yang merupakan pustaka visualisasi dengan sumber terbuka dibangun diatas pustaka *matplotlib*.



Gambar 4. Code python dan Visualisasi Confusion Matrix

Gambar 4 menunjukkan hasil visualisasi Confusion Matrix. Setelah diketahui *Confusion Matrix* dari model yang dibuat selanjutnya dilakukan perhitungan nilai akurasi dari model klasifikasi *Naive Bayes* yang telah dibuat menggunakan *sklearn.metrics* dengan mengimport *accuracy score* yang disediakan oleh *library scikit.learn*.

3.7. Evaluasi Model

Setelah melakukan pengujian model langkah selanjutnya yang dilakukan adalah evaluasi model. Evaluasi model bertujuan untuk menghasilkan *confusion matrix* dengan ukuran 3 x 3. *Confusion Matrix* memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh model klasifikasi dengan hasil klasifikasi sebenarnya. Hasil evaluasi model dengan *confusion matrix* dilihat pada Tabel 2 Hasil Confusion Matrix.

Tabel 2. Hasil Confusion Matrix

Actual Class	Predict Class		
	Positif	Netral	Negatif
Positif	20	2	6
Netral	4	23	6
Negatif	37	46	195

3.8. Hasil dan Visualisasi Klasifikasi Sentimen

3.8.1. Sentimen Positif

Data ulasan positif adalah hasil pelabelan yang masuk pada kelas positif menggunakan analisis sentimen berbasis Vader. Ulasan positif tersebut di

identifikasi berdasarkan banyaknya frekuensi kata dalam ulasan. berikut adalah hasil visualisasi ulasan positif dari hasil ekstraksi informasi yang didapatkan dari ulasan-ulasan yang ditulis oleh pengunjung. diperoleh informasi bahwa pada kelas sentimen positif kata yang paling sering muncul adalah kata-kata *alfamart, pay, follow, twitter, retweet, team, absolute, happy*.

3.8.2. Sentimen Negatif

Data ulasan negatif adalah hasil pelabelan yang masuk pada kelas negatif menggunakan analisis sentimen berbasis Vader. Ulasan negatif tersebut diidentifikasi berdasarkan banyaknya frekuensi kata dalam ulasan. berikut adalah hasil visualisasi ulasan negatif dari hasil ekstraksi informasi yang didapatkan dari ulasan-ulasan yang ditulis oleh pengunjung, diperoleh informasi bahwa pada kelas sentimen negatif kata yang paling sering muncul adalah kata-kata *complaining, Pay, promo, stupid, shopeepay, shalma, voucher, meaning*.

3.8.3. Sentimen Netral

Data ulasan Netral adalah hasil pelabelan yang masuk pada kelas netral menggunakan analisis sentimen berbasis Vader. Ulasan netral tersebut diidentifikasi berdasarkan banyaknya frekuensi kata dalam ulasan. berikut adalah hasil visualisasi ulasan positif dari hasil ekstraksi informasi yang didapatkan dari ulasan-ulasan yang ditulis oleh pengunjung. diperoleh informasi bahwa pada kelas sentimen netral kata yang paling sering muncul adalah kata-kata *info, chek, continue, buy, sahabat, shalma*.

4. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan, dapat disimpulkan bahwa hasil crawling data dari twitter sebanyak 2000 data tweet kemudian di proses melalui proses preprocessing dihasilkan data bersih sebanyak 1691 data tweet, Pada proses klasifikasi digunakan data test sebanyak 20% dan data training 80% dari data keseluruhan, dari 339 tweet data test memakai Algoritma Naïve Bayes dengan seleksi fitur Term- Frequency, didapatkan jumlah sentimen positif 83.19% yang lebih dominan dibanding sentimen negatif 9.44% ataupun netral 7.37 % , untuk data training dari 1352, didapatkan jumlah sentimen positif 78.25% yang lebih dominan dibanding sentimen negatif 8.8 % ataupun

netral 12.94 %. Sedangkan nilai akurasi pada Algoritma Naïve Bayes dengan seleksi fitur Term-Frequency 70.20 %.

5. DAFTAR PUSTAKA

- [1] E. Elisa, “Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 2, pp. 472–478, Jun. 2018, doi: 10.29207/resti.v2i2.280.
- [2] F. Wibowo, A. U. Khasanah, and F. I. F. S. Putra, “Analisis Dampak Kehadiran Pasar Modern terhadap Kinerja Pemasaran Pasar Tradisional Berbasis Perspektif Pedagang dan Konsumen di Kabupaten Wonogiri,” *Benefit J. Manaj. dan Bisnis*, vol. 7, no. 1, pp. 53–65, 2022, doi: 10.23917/benefit.v7i1.16057.
- [3] S. Y. B. Ginting, “Keberadaan Mini Market Alfamart Dan Indomaret Kaitannya Dengan Tingkat Penghasilan Pedagang Tradisional Di Wilayah Kecamatan Talang Kelapa Kabupaten Banyuasin Sumatera Selatan,” *J. Swarnabhumi*, vol. 3, no. 1, pp. 67–76, 2018.
- [4] S. P. Lestari and M. Samaniatun, “Pengaruh Pelayanan dan Kepuasan Pelanggan Terhadap Loyalitas Pelanggan Pada Perusahaan Ritel Alfamart DR Wahidin H398 Semarang,” *J. Egaliter*, vol. 4, no. 6, 2020.
- [5] APJII, “Gudang Data,” <https://apjii.or.id/gudang-data/hasil-opm-2022>, 2022. .
- [6] N. Amaliah, “Kualitas Karyawan Untuk Meningkatkan Pelayanan Prima Terhadap Pelanggan Komplain Dialfamart Transkalimantan Km. 8 Handil Bakti Kabupaten Barito Kuala,” Universitas Islam Kalimantan, 2020.