



Implementasi Algoritma Decision Tree dan Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks

Tiyas Asih Qurnia Putri, Agung Triayudi*, Rima Tamara Aldisa

Fakultas Teknologi Komunikasi dan Informatika, Sistem Informasi, Universitas Nasional, Jakarta

Jl. Sawo Manila No.61, RW.7, Pejaten Bar., Ps. Minggu, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta, Indonesia

Email: ¹tiyasasihqurniaputri2019@student.unas.ac.id, ^{2,*}agungtriayudi@civitas.unas.ac.id, ³.rimatamaa@gmail.com

Email Penulis Korespondensi: agungtriayudi@civitas.unas.ac.id

Submitted: 17/01/2023; Accepted: 31/01/2023; Published: 31/01/2023

Abstrak-Indonesia masuk kedalam kategori negara dengan jumlah penduduk terbesar di dunia, keadaan tersebut menjadi peluang bisnis bagi pengusaha yang masuk kedalam pasar industry coffe shop. Peneliti memanfaatkan salah satu metode pengelompokan yakni klasifikasi data mining agar dapat membantu entitas bisnis untuk mengidentifikasi kelompok yang berbeda dalam database kepuasan pelanggan starbucks. Tujuan penelitian ini adalah dapat mengelompokkan kategori menjadi 3 kelas yaitu puas, cukup puas dan tidak puas menggunakan algoritma Decision tree & Naive Bayes. Sehingga bisa mengetahui opini masyarakat terhadap kepuasan pelanggan starbucks, pada penelitian ini memiliki tujuan untuk memperoleh hasil nilai accuracy, precision dan recall serta mencari tahu algoritma terbaik terhadap klasifikasi data mining kepuasan pelanggan starbucks. Pada penelitian ini menggunakan data uji yang diperoleh dari tweet dengan kata kunci "Starbucks" dari twitter. Hasil dari penelitian ini dimana dari proses klasifikasi sentimen terhadap kepuasan pelanggan starbucks diperoleh kategori netral, dapat dilihat dari ulasan menggunakan kata kunci "starbuck OR starbucks OR #starbucks" didapatkan hasil yakni komentar positif sebanyak 476 tweet dengan jumlah persentase sebesar 19,2%, komentar netral sebanyak 1743 tweet dengan jumlah persentase sebesar 70,3 % dan komentar negatif sebanyak 258 tweet dengan jumlah persentase sebesar 10,4%, sehingga dapat ditarik kesimpulan berdasarkan perhitungan polarity tersebut ulasan komentar terhadap starbuck memiliki kategori puas. Pada penelitian ini dapat ditarik kesimpulan bahwa kinerja algoritma Decision Tree lebih baik dari pada algoritma Naive Bayes, dapat dilihat dari penjelasan berikut. Algoritma Decision Tree pada hasil akurasi nilai sebesar 83%. Sedangkan algoritma Naive Bayes pada hasil akurasi nilai sebesar 74%.

Kata Kunci: Klasifikasi; Data Mining; Decision Tree; Naive Bayes; Sentimen; Twitter; Python.

Abstract-Indonesia is included in the category of countries with the largest population in the world, this situation is a business opportunity for entrepreneurs who enter the coffee shop industry market. Researchers utilize one of the grouping methods, namely data mining classification in order to help business entities to identify different groups in the Starbucks customer satisfaction database. The purpose of this research is to be able to group categories into 3 classes, namely satisfied, quite satisfied and dissatisfied using the Decision Tree & Naive Bayes algorithm. So that it can find out public opinion on Starbucks customer satisfaction, in this study the aim was to obtain accuracy, precision and recall values and find out the best algorithm for data mining classification of Starbucks customer satisfaction. In this study using test data obtained from tweets with the keyword "Starbucks" from Twitter. The results of this study where the sentiment classification process for Starbucks customer satisfaction obtained a neutral category, it can be seen from the reviews using the keywords "starbuck OR starbucks OR #starbucks" The results obtained were positive comments of 476 tweets with a percentage of 19.2%, neutral comments of 1743 tweets with a percentage of 70.3% and negative comments of 258 tweets with a percentage of 10.4%, so that conclusions can be drawn based on the polarity calculation, the comments on starbuck have a satisfied category. In this study, it can be concluded that the performance of the Decision Tree algorithm is better than the Naive Bayes algorithm, as can be seen from the following explanation. The Decision Tree algorithm results in an accuracy of 83%. Naive Bayes on value accuracy results by 74%.

Keywords: Classification; Data Mining; Decision Tree; Naive Bayes; Sentiment; Twitter; Python.

1. PENDAHULUAN

Tingginya masyarakat dalam mengkonsumsi kopi, menjadi peluang bisnis bagi pengusaha yang masuk kedalam pasar industry coffe shop [1]. Maka dari itu potensi berbisnis di Indonesia sendiri pun sangat tinggi, salah satunya pada industri coffe shop di Indonesia yaitu Starbucks merupakan contoh peluang bisnis yang berkembang pesat dan paling kompetitif di era perkembangan zaman saat ini. Dengan seiring perkembangan gaya hidup pada masyarakat mereka cenderung memiliki pikiran membeli dan minum kopi di coffe shop terlihat lebih keren dan elit dari kalangan anak muda maupun sampai orang tua sekalipun. Tercatat pada kuartal III – 2021 total ada 478 gerai resmi Starbuck di Indonesia [2]

Perkembangan tren teknologi yang semakin pesat ini terdapat sejumlah data yang tersedia pada industri informasi. Jika ingin data tersebut bermanfaat menjadi informasi yang terpercaya dan akurat maka harus dilakukan pengolahan data. Data mining merupakan industri yang berkembang pesat. Untuk mendapatkan informasi yang berharga dan dapat diandalkan setiap orang saat ini membutuhkan data untuk digunakan supaya dapat membantu dalam pembuatan keputusan bisnis yang bijaksana, prediksi hasil, dan banyak tugas lainnya. Terdapat beberapa macam model atau fungsi pada data mining yang terdiri dari association, classification, feature selection, dan clustering. Peneliti memanfaatkan salah satu metode pengelompokan yakni klasifikasi pada data mining agar dapat membantu entitas bisnis untuk mengidentifikasi kelompok yang berbeda dalam database kepuasan pelanggan starbucks. Factor yang dianalisis melalui klasifikasi dapat berdampak pada kepuasan pelanggan. Sebagian besar orang di Indonesia sendiri berusaha mencari bantuan orang lain pada saat ingin membuat suatu keputusan, tetapi



tidak selalu mudah untuk mengikuti saran siapapun. Dan itulah mengapa dengan bantuan data mining seseorang bisa percaya diri untuk membuat keputusan sendiri.

Kepuasan pelanggan starbucks merupakan rancangan studi penelitian dalam menganalisis kepuasan para pelanggan yang diutarakan dari pengguna social media twitter. Maka dari itu, perusahaan mengharuskan untuk dilakukannya evaluasi supaya dapat mengetahui aspek apa saja yang dapat mempengaruhi kepuasan para pelanggan, sehingga kesetiaan para pelanggan akan tercipta jika pelanggan merasakan kepuasan terhadap manfaat layanan yang diberikan dari perusahaan [3]. Kepuasan pelanggan mencakup ketidaksamaan terkait pentingnya tingkat layanan dengan hasil yang dapat dirasakan [4]

Berikut beberapa penelitian sebelumnya terkait dengan yang diteliti oleh penulis. Pertama, penelitian sebelumnya yang dilakukan oleh Nurhakim tentang analisis sentimen terhadap ulasan kepuasan pelanggan pada marketplace tokopedia di jejaring sosial twitter menggunakan algoritma naive bayes. Penelitian ini bertujuan sebagai tolak ukur terkait kepuasan para pelanggan dengan dilakukannya penelitian terhadap ulasan opini pelanggan tentang Tokopedia pada jejaring media sosial Twitter. Penelitian ini mendapatkan hasil pada algoritma Multinomial Naive Bayes didapat nilai akurasi sebesar 79,02%, precision sebesar 80,30%, recall sebesar 77,94% dan specificity sebesar 80,15% [5]. kedua, penelitian sebelumnya yang dilakukan oleh Saurina tentang analisis sentimen ulasan pelanggan batik ecoprint menggunakan naive bayes classifier dan KNN. Penelitian ini memiliki tujuan dimana upaya mengetahui ulasan pelanggan terhadap batik ecoprint dan menjadi tolak ukur dalam penilaian dalam kepuasan terhadap pelanggan menjadi 2 kategori kelas yakni positif dan negatif. Hasil dari penelitian ini dimana dari proses Analisa sentiment terhadap batik ecoprint diperoleh kategori netral, dengan memperoleh hasil dimana pada algoritma Naive Bayes nilai klasifikasi presesi sebesar 85%, recall 88% dan f1-score 88%. Sedangkan algoritma KNN nilai klasifikasi presesi sebesar 76%, recall 76% dan f1-score 74% [6]. Ketiga, penelitian sebelumnya yang dilakukan oleh Sepri tentang penerapan algoritma Naive Bayes untuk analisis kepuasan penggunaan aplikasi bank. Penelitian ini memiliki tujuan dimana dalam melakukan analisa sentimen terhadap pengguna aplikasi mobile bank muamalat pada media playstore, agar mendapatkan tingkat akurasi yang baik dan tepat menggunakan algoritma Naive Bayes. Sehingga peneliti mengharapkan supaya hasil tersebut menjadi tolak ukur untuk penilaian bank muamalat sebagai penentuan kebijakan berbasis komputerisasi. Hasil dari penelitian ini dapat menyimpulkan dimana menggunakan algoritma Naive Bayes dapat dilihat dari nilai akurasi sebesar 87%, berdasarkan data training yakni 200 ulasan komentar negatif dan 200 ulasan komentar positif, sedangkan data testing yakni 70 ulasan komentar negatif dan 70 ulasan komentar positif [7]. Keempat, penelitian yang dilakukan oleh Petiwi tentang analisis sentimen gofood berdasarkan twitter menggunakan metode Naive Bayes dan support vector machine. Penelitian ini memiliki tujuan dimana dari kumpulan data ulasan opini masyarakat tersebut akan dilakukan pengelompokan menjadi 3 kelas kategori yakni, positif, negatif dan netral. Hasil dari penelitian ini dimana pada algoritma Naive Bayes didapat nilai akurasi sebesar 74,6% dan 91,5% melalui pengujian antarmuka dan algoritma Support Vector Machine didapat nilai akurasi sebesar 83% dan 98,5% [8]. Kelima, penelitian yang dilakukan oleh Jacarria Pangestu & Kodar tentang Implementasi Multinomial Naive Bayes Untuk Klasifikasi Sentimen Terhadap Pelayanan Perusahaan Otobus Menggunakan Data Facebook. Penelitian ini bertujuan dimana peneliti ingin melakukan analisis terkait informasi yang didapat terhadap perusahaan otobus supaya menjadi acuan dalam mengukur kualitas pelayanan di masa akan datang. Hasil dari penelitian ini dimana mendapatkan persentase sentimen negatif sebesar 23% , sentimen netral sebesar 53%, dan sentimen Positif sebesar 25% serta pengujian data pada algoritma Multinomial Naive Bayes didapat nilai akurasi sebesar 95%, precision sebesar 95%, dan recall sebesar 95% [9]. Dan yang terakhir, penelitian yang dilakukan oleh Syafii tentang komparasi algoritma Naive Bayes dan algoritma C4.5 dalam klasifikasi pelanggan produk indihome. Penelitian ini memiliki tujuan yakni dimana untuk memperoleh pola informasi terhadap klasifikasi data pelanggan dengan algoritma C4.5 dan Naive Bayes untuk memperoleh tingkat akurasi terbaik guna untuk dijadikan sumber dalam pengambilan keputusan oleh PT.Telkom Indonesia. Hasil dari penelitian ini dimana pengujian data pada algoritma Naive Bayes didapat nilai akurasi sebesar 99.81% dan pengujian data pada algoritma C4.5 didapat nilai akurasi 94.63% [10].

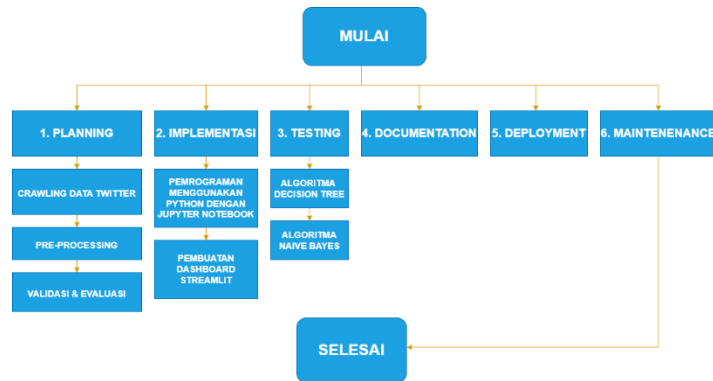
Berdasarkan penelitian sebelumnya diatas, didapatkan suatu permasalahan tentang “Penerapan Klasifikasi data mining Untuk Kepuasan Pelanggan Starbucks Menggunakan Algoritma Decision tree & Naive Bayes”. Tujuan penelitian ini adalah dapat mengelompokan kategori menjadi 3 kelas yaitu puas, cukup puas adan tidak puas menggunakan algoritma Decision tree & Naive Bayes, sehingga bisa mengetahui opini masyarakat terhadap kepuasan pelanggan starbuck serta pada penelitian ini memiliki tujuan untuk memperoleh hasil nilai accuracy, precision dan recall serta mencari tahu algoritma terbaik terhadap klasifikasi data mining kepuasan pelanggan starbucks. Pada penelitian ini menggunakan data uji yang diperoleh dari tweet dengan kata kunci "Starbucks" dari twitter.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada perencanaan sebuah sistem terhadap penelitian ini menggunakan Metode Agile Software Development [11].Metode Agile Software Development merupakan salah satu metode untuk mengembangkan software pada

tahapan-tahapan yang berulang (iterasi) [12]. Gambar berikut ini merupakan tahapan dari penelitian yang akan dilakukan :



Gambar 1. Tahapan penelitian

Adapun penjelasan gambar 1 diatas, Tahapan pertama dimulai dari planning yang terdiri dari Crawling Data dari twitter, Pre-processing, serta Validasi dan Evaluasi. Selanjutnya, Implementation dimana menggunakan bahasa pemrograman Python dan framework streamlite. Kemudian, Testing yang terdiri dari pengujian data menggunakan algoritma Decision tree & Naive Bayes. Setelah itu, Documentation merupakan catatan modul dan baris-baris perintah pada program sehingga menghasilkan nilai input dan output. Lalu, Deployment yang dilakukan pada saat perancangan sistem menggunakan jupyter notebook dengan bahasa pemrograman python, serta tampilan design berupa dashboard menggunakan framework streamlit. Yang terakhir, Maintenance dimana sistem perhitungan yang telah berhasil diimplementasikan terhadap algoritma tersebut secara berkala akan dilakukan update perkembangan arus yang terdapat pada twitter terhadap ulasan komentar starbuck, serta penambahan dataset baru dan lebih banyak jumlahnya.

2.2 Confusion Matrix

Penggunaan Confusion Matrix merupakan salah satu metode untuk mengukur kinerja pada suatu metode pengelompokkan. Cara kerja Confusion Matrix dilakukan dengan komparasi dataset yang sudah dilakukan pengujian menjadi empat kategori. Berikut rumus untuk melakukan pengukuran nilai akurasi untuk klasifikasi text :

$$\text{Recall} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

2.3 Algoritma Decision Tree

Decision tree merupakan salah satu algoritma yang dapat digunakan untuk memprediksi dari data yang sebelumnya sudah tersedia [13]. Berikut langkah-langkah dalam membuat pohon keputusan :

- a. Memilih atribut untuk dijadikan sebagai akar yang berdasarkan dari nilai gain tertinggi dari atribut yang sudah ada. Berikut rumus persamaan untuk menghitung nilai gain :

$$\text{Gain} (S,A) = \text{Entropy}(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy} (S_i) \tag{4}$$

Keterangan :

A : Properti

N : Jumlah partisi atribut A

S : Sekumpulan kasus

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S.

Berikut merupakan rumus persamaan untuk menghitung nilai entropy :

$$\text{Entropy} (s) = \sum_{i=1}^n - p_i * \text{Log}_2 p_i \tag{5}$$

Keterangan:

S : Sekumpulan kasus

A : Properti

N : Total Jumlah

pi : Rasio dari Si terhadap S.

- b. Membuat cabang pada tiap-tiap nilai.
- c. Membagi setiap kasus pada cabang
- d. Membagi kasus pada cabang
- e. Mengulangi setiap tahapan supaya mendapatkan kelas yang sama dari setiap cabang maupun setiap kasus [14]

2.4 Algoritma Naive Bayes

Algoritma Naive Bayes merupakan salah satu algoritma yang sering digunakan pada penelitian dalam upaya mengatasi klasifikasi pada teks dengan menggunakan perhitungan probabilitas [15]. Naive Bayes adalah salah satu algoritma yang banyak digunakan pada penelitian dengan jumlah data yang banyak dan data yang kosong dengan menggunakan perhitungan probabilitas, serta statistic dalam prediksi suatu data terkait dari pengujian yang telah dilakukan sebelumnya istilah tersebut biasa dikenal dengan missing value [16].Berikut adalah rumus persamaan perhitungan algoritma Naive Bayes :

$$P(P|H) = \frac{P(H) \times P(H)}{P(Y)} \tag{6}$$

Keterangan:

H : Hipotesis Y untuk kelas tertentu

Y : data kelas yang tidak diketahui

P (H| Y) : Probabilitas hipotetis H (probabilitas posterior) berdasarkan kondisi y

P(H) : probabilitas hipotesis H (probabilitas sebelumnya)

P (Y | H) : Probabilitas hipotesis Y mengingat kondisi hipotesis

P (Y) : Probabilitas Y

2.5 Anaconda Navigator & Jupyter Notebook.

Anaconda navigator merupakan tampilan Graphical User Interface (GUI) pada dekstop dimana salah satu dari output anaconda sebagai upaya pengguna dalam membuat aplikasi dan mengelola paket, lingkungan dan saluran Conda tanpa menggunakan perintah pada baris[17]. Jupyter notebook merupakan lingkungan yang cukup interaktif dalam menjalankan baris perintah berupa kode pada browser dan termasuk tools yang berguna dalam membuat sebuah pekerjaan yang berhubungan dengan analisis imuwan data [18].

3. HASIL DAN PEMBAHASAN

Penelitian ini dilakukan dengan cara membandingkan ulasan setiap pelanggan Starbucks dengan menggunakan kata kunci "starbuck", "starbucks", dan "#starbucks". Tujuan dari penelitian ini adalah mendapatkan nilai accuracy, precision, dan recall serta mencari tahu algoritma terbaik terhadap klasifikasi data mining kepuasan pelanggan Starbucks. Penelitian ini dilakukan dalam beberapa tahap, tahap pertama yaitu melakukan crawling data yang bertujuan mendapatkan data dari website/link yang diinginkan sesuai dengan kata kunci yang diinput. Tahap selanjutnya adalah pre-processing agar data-data yang didapatkan bersih dan dapat dengan mudah dilakukan pengklasifikasian. Tahap terakhir yaitu melakukan pengklasifikasian dan analisis data menggunakan algoritma decision tree dan Naïve Bayes ke dalam tiga kategori yaitu pelanggan yang puas, cukup puas, dan tidak puas.

3.1 Crawling Data

Pada tahap pertama, dilakukan pembuatan program dengan menggunakan bahasa pemrograman Python. Dimana program tersebut dapat digunakan untuk mengakses API dari Twitter yang menghasilkan keluaran berupa format JSON dan menyimpan data tersebut dalam format CSV. Selanjutnya, dilakukan proses crawling data yang bersumber dari cuitan pelanggan Starbucks di Twitter dengan kata kunci "starbuck", "starbucks", dan "#starbucks". Data-data yang diambil berupa data pesan, jumlah favorit, jumlah retweet, tanggal postingan, username, dan jumlah followers-nya. Data-data yang diperoleh dari proses crawling data adalah sebagai berikut.

Unnamed: 0	Message	Favorite Count	Retweet Count	Created At	Username	Followers
0	@sianakah @tenyarifes Starbucks ga bikin kembu...	0	0	2022-12-05 18:41:43+00:00	Wilyan Inan Puspasari	150
1	Kengawan minggu lepas aku gi warehouse sale s...	0	0	2022-12-05 16:37:49+00:00	shamira	148
2	@BMULYAD1 @dinaeight @Lenny_diary lah kan rih...	0	0	2022-12-05 16:34:53+00:00	Naren.	216
3	@discounffees Besok ote ke mod sekalan tukar ...	0	0	2022-12-05 16:29:46+00:00	akbarfith @j's	224
4	@RahN_ @Lenny_diary @tom_hutomo Mari korekai l...	0	0	2022-12-05 16:28:01+00:00	Mulyadi Setian	66
...
2473	Pikunnya ystuhan 1Kaca mata kelinggatan di el...	0	0	2022-11-28 15:08:41+00:00	Ayu Suerianti	464
2474	Starbuck. Nda pedai order 1 hal, hrpa lg 1 hal...	0	0	2022-11-28 14:54:17+00:00	dhilho✓	110
2475	@saifuzamir Aah ada hahaha baru lalu semalam...	0	0	2022-11-28 14:53:39+00:00	esso too	11933
2476	@ngaliamless kalengann starbuck aja bin nemu, yg...	0	0	2022-11-28 14:50:52+00:00	nad	132
2477	nyesel baru nyobain tuku skrg, brr kta org2 tu...	0	0	2022-11-28 14:36:44+00:00	aku suka ka Abel	21

Gambar 2. Crawling

Dari Gambar 2 di atas, dapat di lihat bahwa dari proses crawling data didapatkan 2.478 data yang sesuai dengan kata kunci yang telah diberikan. Akan tetapi data ini masih perlu diolah agar didapatkan data yang lebih akurat sehingga siap untuk digunakan untuk tahap klasifikasi data mining.

3.2 Pre-processing

Data-data yang dihasilkan dari proses crawling data selanjutnya akan diolah kembali dengan menggunakan metode pre-processing. Metode pre-processing bertujuan untuk melakukan eliminasi data yang tidak sesuai dengan parameter yang telah ditentukan. Berikut ini tahapan pre-processing yang terdiri dari :

3.2.1 Case Folding

Case folding adalah proses mengubah semua huruf menjadi huruf kecil. Hal ini disebabkan karena data yang kita miliki tidak selalu konsisten dan terstruktur dalam menggunakan huruf kapital. Berikut ini adalah data yang dihasilkan dari proses case folding :

Unnamed: 0	Message	Favorite Count	Retweet Count	Created At	Username	Followers	Tokenization	Stop_Removal
0	starbucks ga bikin kembung untungnya tapi b...	0	0	2022-12-05 16:41:43+00:00	Wayan Intan Puspasari	150	[starbucks, ga, bikin, kembung, untungnya, L...	[starbucks, ga, bikin, kembung, untungnya, b...
1	kengawen minggu lepas aku gi warehouse sale s...	0	0	2022-12-05 16:37:45+00:00	shamira	148	[kengawen, minggu, lepas, aku, gi, warehouse, sale...	[kengawen, minggu, lepas, gi, warehouse, sale...
2	lah kan mas mulyadi nih yang berprasangka bu...	0	0	2022-12-05 16:34:53+00:00	Naren	216	[lah, kan, mas, mulyadi, nih, yang, berprasa...	[mas, mulyadi, nih, berprasangka, rusak, bl...
3	besok oke ke mod sekalian tukar starbucks rew...	0	0	2022-12-05 16:29:49+00:00	akbar69.ath	224	[besok, oke, ke, mod, sekalian, tukar, starbu...	[besok, oke, mod, tukar, starbucks, rewarr...
4	mari koreksi li hati apakah hati akan cond...	0	0	2022-12-05 16:28:01+00:00	Mulyadi Selian	66	[mari, koreksi, li, hati, apakah, hati, aka...	[mari, koreksi, li, hati, hati, ootlong, me...
5	lah ngapain dicatat ni mah emang mas mulya...	0	0	2022-12-05 16:27:20+00:00	Naren	216	[lah, ngapain, dicatat, ni, mah, emang, mas, mulya...	[ngapain, dicatat, mah, emang, mas, mulyadi...
6	tolong dicatat masuk akal ala pacinta kopi ...	0	0	2022-12-05 16:18:02+00:00	Mulyadi Selian	66	[tolong, dicatat, masuk, akal, ala, pacinta...	[tolong, dicatat, masuk, akal, ala, pacinta...
7	duh orang dibalang gue berangkat argo lawu be...	0	0	2022-12-05 16:14:13+00:00	caramel-machiato	33	[duh, orang, dibalang, gue, berangkat, argo...	[duh, orang, dibalang, gue, berangkat, argo...
8	let us see aku udah minum seminggau nih so...	0	0	2022-12-05 16:10:26+00:00	WTS BORN PINK CAT1 CAT3 D2	16	[let, us, see, aku, udah, minum, seminggau...	[let, us, see, udah, minum, seminggau...
9	mana buktinya kalau kamu banyak duit kalau ...	0	0	2022-12-05 16:08:37+00:00	Mulyadi Selian	66	[mana, buktinya, kalau, kamu, banyak, duit, ka...	[buktinya, duit, modal, foto, minum, kopi, s...

Gambar 3. Case Folding

Pada Gambar 3 dapat dilihat bahwa telah dilakukan perubahan pada semua kata yang terdapat pada teks kalimat dari huruf "a" sampai "z" menjadi huruf kecil.

3.2.2 Cleaning

Cleaning adalah proses menyiapkan data dengan cara menghapus atau memodifikasi data yang salah, tidak relevan, tidak akurat, duplikat, maupun yang tidak terformat. Setelah melewati tahap ini barulah data siap untuk diolah. Berikut ini adalah data yang dihasilkan dari proses cleaning:

Unnamed: 0	Message	Favorite Count	Retweet Count	Created At	Username	Followers
0	starbucks ga bikin kembung untungnya tapi b...	0	0	2022-12-05 16:41:43+00:00	Wayan Intan Puspasari	150
1	kengawen minggu lepas aku gi warehouse sale s...	0	0	2022-12-05 16:37:45+00:00	shamira	148
2	lah kan mas mulyadi nih yang berprasangka bu...	0	0	2022-12-05 16:34:53+00:00	Naren	216
3	besok oke ke mod sekalian tukar starbucks rew...	0	0	2022-12-05 16:29:49+00:00	akbar69.ath	224
4	mari koreksi li hati apakah hati akan cond...	0	0	2022-12-05 16:28:01+00:00	Mulyadi Selian	66
5	lah ngapain dicatat ni mah emang mas mulya...	0	0	2022-12-05 16:27:20+00:00	Naren	216
6	tolong dicatat masuk akal ala pacinta kopi ...	0	0	2022-12-05 16:18:02+00:00	Mulyadi Selian	66
7	duh orang dibalang gue berangkat argo lawu be...	0	0	2022-12-05 16:14:13+00:00	caramel-machiato	33
8	let us see aku udah minum seminggau nih so...	0	0	2022-12-05 16:10:26+00:00	WTS BORN PINK CAT1 CAT3 D2	16
9	mana buktinya kalau kamu banyak duit kalau ...	0	0	2022-12-05 16:08:37+00:00	Mulyadi Selian	66

Gambar 4. Cleaning

Pada Gambar 4 dilakukan penghapusan kata pada teks yang tidak memiliki arti, seperti kata penghubung, tanda baca, simbol atau karakter, angka, emoji dan URL.

3.2.3 Tokenizing

Tokenizing adalah metode untuk melakukan pemisahan kata dalam suatu kalimat dengan tujuan untuk proses analisis teks lebih lanjut. Berikut ini adalah data yang dihasilkan dari proses tokenizing:

Unnamed: 0	Message	Favorite Count	Retweet Count	Created At	Username	Followers	Tokenization
0	starbucks ga bikin kembung untungnya tapi b...	0	0	2022-12-05 16:41:43+00:00	Wayan Intan Puspasari	150	[starbucks, ga, bikin, kembung, untungnya, t...
1	kengawen minggu lepas aku gi warehouse sale s...	0	0	2022-12-05 16:37:45+00:00	shamira	148	[kengawen, minggu, lepas, aku, gi, warehouse...
2	lah kan mas mulyadi nih yang berprasangka bu...	0	0	2022-12-05 16:34:53+00:00	Naren	216	[lah, kan, mas, mulyadi, nih, yang, berprasa...
3	besok oke ke mod sekalian tukar starbucks rew...	0	0	2022-12-05 16:29:49+00:00	akbar69.ath	224	[besok, oke, ke, mod, sekalian, tukar, starb...
4	mari koreksi li hati apakah hati akan cond...	0	0	2022-12-05 16:28:01+00:00	Mulyadi Selian	66	[mari, koreksi, li, hati, apakah, hati, aka...
5	lah ngapain dicatat ni mah emang mas mulya...	0	0	2022-12-05 16:27:20+00:00	Naren	216	[lah, ngapain, dicatat, ni, mah, emang, mas...
6	tolong dicatat masuk akal ala pacinta kopi ...	0	0	2022-12-05 16:18:02+00:00	Mulyadi Selian	66	[tolong, dicatat, masuk, akal, ala, pacinta...
7	duh orang dibalang gue berangkat argo lawu be...	0	0	2022-12-05 16:14:13+00:00	caramel-machiato	33	[duh, orang, dibalang, gue, berangkat, argo...
8	let us see aku udah minum seminggau nih so...	0	0	2022-12-05 16:10:26+00:00	WTS BORN PINK CAT1 CAT3 D2	16	[let, us, see, aku, udah, minum, seminggau...
9	mana buktinya kalau kamu banyak duit kalau ...	0	0	2022-12-05 16:08:37+00:00	Mulyadi Selian	66	[mana, buktinya, kalau, kamu, banyak, duit...

Gambar 5. Tokenizing

Pada Gambar 5 menunjukkan telah dilakukan pemisahan teks pada kalimat menjadi 1 kata sesuai dengan penulisan spasi pada teks.

3.2.4 Stopword Removal

Stopword removal adalah proses menghilangkan kata penghubung. Hal ini bertujuan untuk mengurangi jumlah kata dalam sebuah dokumen yang nantinya akan mempengaruhi kecepatan dan performa. Berikut ini adalah data yang dihasilkan dari proses Stopword removal:

Unmmed: 0	Message	Favorite Count	Retweet Count	Created At	Username	Followers	Tokenization	Stop_Removal
0	starbucks ga bikin kembang untungnya lagi lu...	0	0	2022-12-05 16:41:43+00:00	Wayan Intan Puspasari	150	[starbucks, ga, bikin, kembang, untungnya, l...	[starbucks, ga, bikin, kembang, untungnya, b...
1	kengkawan minggu lepas aku gi warehouse sale...	0	0	2022-12-05 16:37:45+00:00	shamira	148	[kengkawan, minggu, lepas, aku, gi, warehouse, sale...	[kengkawan, minggu, lepas, gi, warehouse, sale...
2	lah kan mas mulyadi nih yang berprasangka bu...	0	0	2022-12-05 16:34:53+00:00	Naren.	216	[lah, kan, mas, mulyadi, nih, yang, berprasangka, bu...	[mas, mulyadi, nih, berprasangka, buruk, bl...
3	besok otw ke mod sekalian tukar starbucks new.	0	0	2022-12-05 16:29:46+00:00	akbar69.eth	224	[besok, otw, ke, mod, sekalian, tukar, starbucks, new.	[besok, otw, mod, tukar, starbucks, rewards...
4	mari koreksi isi hati apakah hati akan cond...	0	0	2022-12-05 16:28:01+00:00	Mulyadi Selian	66	[mari, koreksi, isi, hati, apakah, hati, akan, cond...	[mari, koreksi, isi, hati, apakah, hati, aka...
5	lah ngapain dicatet ini mah emang mas mulya...	0	0	2022-12-05 16:27:20+00:00	Naren.	216	[lah, ngapain, dicatet, ini, mah, emang, mas, mulya...	[ngapain, dicatet, mah, emang, mas, mulyadi...
6	tolong dicatet masuk akal ala pecinta kopi...	0	0	2022-12-05 16:19:02+00:00	Mulyadi Selian	66	[tolong, dicatet, masuk, akal, ala, pecinta, kopi...	[tolong, dicatet, masuk, akal, ala, peccinta...
7	duh orang dibilang gue berangkat argo lawu be...	0	0	2022-12-05 16:14:13+00:00	caramel-machiato	33	[duh, orang, dibilang, gue, berangkat, argo, lawu, be...	[duh, orang, dibilang, gue, berangkat, argo...
8	let us see aku udah minum semingguan sih so...	0	0	2022-12-05 16:10:25+00:00	WTS BORN PINK CAT1 CAT3 D2	16	[let, us, see, aku, udah, minum, semingguan, sih, so...	[let, us, see, udah, minum, semingguan, sih...
9	mana buktinya kalau kamu banyak duit kalau...	0	0	2022-12-05 16:08:37+00:00	Mulyadi Selian	66	[mana, buktinya, kalau, kamu, banyak, duit, kalau...	[buktinya, duit, modal, foto, minum, kopi, b...

Gambar 6. Stopword Removal

Pada Gambar 6 merupakan tahap dalam memilih kata yang terdapat pada teks yang tidak memiliki arti seperti kata keterangan, kata sambung dan lainnya, yang tidak dibutuhkan pada saat pemodelan data.

3.2.5 Stemming

Stemming adalah menghilangkan imbuhan pada suatu kata. Berikut ini adalah data yang dihasilkan dari proses stemming:

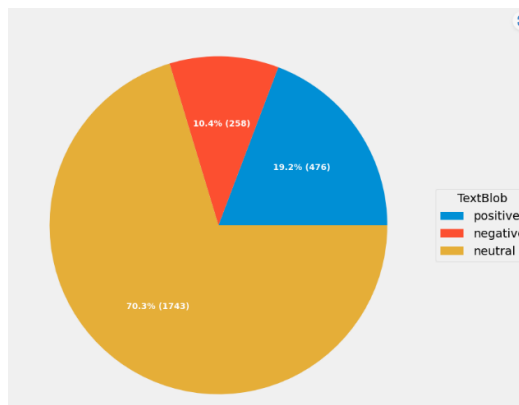
Message	Created At	Username	Tokenized	Stop_Rem	Tweet_Stemmed
starbucks ga bikin kembang untungnya lagi lu...	2022-12-05 16:41:43+00:00	Wayan Intan Puspasari	["starbu", "lu", "starbu", "starbucks", "ga", "bikin", "kembang", "untung", "badan", "seger", "fresh", "sih", "wkwk", "yaa", "americano", "coffee", "sho"]	["starbu", "lu", "starbu", "starbucks", "ga", "bikin", "kembang", "untung", "badan", "seger", "fresh", "sih", "wkwk", "yaa", "americano", "coffee", "sho"]	["starbu", "lu", "starbu", "starbucks", "ga", "bikin", "kembang", "untung", "badan", "seger", "fresh", "sih", "wkwk", "yaa", "americano", "coffee", "sho"]
kengkawan minggu lepas aku gi warehouse sale...	2022-12-05 16:37:45+00:00	shamira	["kengkaw", "kengkawan", "minggu", "lepas", "gi", "warehouse", "sale", "starbucks", "nak", "tumbler", "dm", "ok", "esok", "update", "yg", "si"]	["kengkaw", "kengkawan", "minggu", "lepas", "gi", "warehouse", "sale", "starbucks", "nak", "tumbler", "dm", "ok", "esok", "update", "yg", "si"]	["kengkaw", "kengkawan", "minggu", "lepas", "gi", "warehouse", "sale", "starbucks", "nak", "tumbler", "dm", "ok", "esok", "update", "yg", "si"]
lah kan mas mulyadi nih yang berprasangka bu...	2022-12-05 16:34:53+00:00	Naren.	["lah", "lu", "mas", "mulyadi", "nih", "prasan", "buruk", "bilang", "kalo", "orang", "starbucks", "karna", "lgensi", "sadar"]	["lah", "lu", "mas", "mulyadi", "nih", "prasan", "buruk", "bilang", "kalo", "orang", "starbucks", "karna", "lgensi", "sadar"]	["lah", "lu", "mas", "mulyadi", "nih", "prasan", "buruk", "bilang", "kalo", "orang", "starbucks", "karna", "lgensi", "sadar"]
besok otw ke mod sekalian tukar starbucks new.	2022-12-05 16:29:46+00:00	akbar69.eth	["besok", "otw", "mod", "tukar", "starbucks", "rewards", "birthday", "cake"]	["besok", "otw", "mod", "tukar", "starbucks", "rewards", "birthday", "cake"]	["besok", "otw", "mod", "tukar", "starbucks", "rewards", "birthday", "cake"]
mari koreksi isi hati apakah hati akan cond...	2022-12-05 16:28:01+00:00	Mulyadi Selian	["mari", "koreksi", "isi", "hati", "condong", "tuman", "nongkrong", "starbucks", "mantau", "pilih", "hati", "nongkrong", ""]	["mari", "koreksi", "isi", "hati", "condong", "tuman", "nongkrong", "starbucks", "mantau", "pilih", "hati", "nongkrong", ""]	["mari", "koreksi", "isi", "hati", "condong", "tuman", "nongkrong", "starbucks", "mantau", "pilih", "hati", "nongkrong", ""]
lah ngapain dicatet ini mah emang mas mulya...	2022-12-05 16:27:20+00:00	Naren.	["lah", "ni", "ngapain", "dicatet", "mah", "emang", "mas", "mulyadi", "aja", "negative", "thinking", "orang", "minum", "kopi", "starbucks"]	["lah", "ni", "ngapain", "dicatet", "mah", "emang", "mas", "mulyadi", "aja", "negative", "thinking", "orang", "minum", "kopi", "starbucks"]	["lah", "ni", "ngapain", "dicatet", "mah", "emang", "mas", "mulyadi", "aja", "negative", "thinking", "orang", "minum", "kopi", "starbucks"]
tolong dicatet masuk akal ala pecinta kopi...	2022-12-05 16:19:02+00:00	Mulyadi Selian	["tolong", "tolong", "tolong", "catat", "masuk", "akal", "ala", "cinta", "kopi", "starbucks", "bentuk", "masuk", "akal", "paksa", "paksa", "masuk", "ak"]	["tolong", "tolong", "tolong", "catat", "masuk", "akal", "ala", "cinta", "kopi", "starbucks", "bentuk", "masuk", "akal", "paksa", "paksa", "masuk", "ak"]	["tolong", "tolong", "tolong", "catat", "masuk", "akal", "ala", "cinta", "kopi", "starbucks", "bentuk", "masuk", "akal", "paksa", "paksa", "masuk", "ak"]
duh orang dibilang gue berangkat argo lawu be...	2022-12-05 16:14:13+00:00	caramel-machiato	["duh", "duh", "duh", "orang", "bilang", "gue", "berangkat", "argo", "lawu", "besuk", "jam", "napa", "dah", "nangkring", "starbucks", "dah"]	["duh", "duh", "duh", "orang", "bilang", "gue", "berangkat", "argo", "lawu", "besuk", "jam", "napa", "dah", "nangkring", "starbucks", "dah"]	["duh", "duh", "duh", "orang", "bilang", "gue", "berangkat", "argo", "lawu", "besuk", "jam", "napa", "dah", "nangkring", "starbucks", "dah"]
let us see aku udah minum semingguan sih so...	2022-12-05 16:10:25+00:00	WTS BORN PINK CAT1 CAT3 D2	["let", "us", "see", "aku", "udah", "minum", "minggu", "sih", "so", "Tara", "Suka", "Starbucks", "kopien", "nyoba", "animos", "the", "litter", ""]	["let", "us", "see", "aku", "udah", "minum", "minggu", "sih", "so", "Tara", "Suka", "Starbucks", "kopien", "nyoba", "animos", "the", "litter", ""]	["let", "us", "see", "aku", "udah", "minum", "minggu", "sih", "so", "Tara", "Suka", "Starbucks", "kopien", "nyoba", "animos", "the", "litter", ""]
mana buktinya kalau kamu banyak duit kalau...	2022-12-05 16:08:37+00:00	Mulyadi Selian	["mana", "buktinya", "bukti", "duit", "modal", "foto", "minum", "kopi", "starbucks", "sany", "face", "sneezing", "face", ""]	["mana", "buktinya", "bukti", "duit", "modal", "foto", "minum", "kopi", "starbucks", "sany", "face", "sneezing", "face", ""]	["mana", "buktinya", "bukti", "duit", "modal", "foto", "minum", "kopi", "starbucks", "sany", "face", "sneezing", "face", ""]
mana video sana yg ngomong bahasa jepari	2022-12-05 16:08:37+00:00	Gen	["mana", "video", "sana", "yg", "ngomong", "bahasa", "jepang", "nyebut", "starbucks"]	["mana", "video", "sana", "yg", "ngomong", "bahasa", "jepang", "nyebut", "starbucks"]	["mana", "video", "sana", "yg", "ngomong", "bahasa", "jepang", "nyebut", "starbucks"]
jam terpantau ada di starbucks bareng	2022-12-05 16:08:37+00:00	qj	["jam", "ter", "jam", "ter", "jam", "pantau", "starbucks", "bareng", "curhat", "sedih", "to", "sedih", "banget", "ya", "lidup", "gue", "ehh", "jam", "pantau", "haj"]	["jam", "ter", "jam", "ter", "jam", "pantau", "starbucks", "bareng", "curhat", "sedih", "to", "sedih", "banget", "ya", "lidup", "gue", "ehh", "jam", "pantau", "haj"]	["jam", "ter", "jam", "ter", "jam", "pantau", "starbucks", "bareng", "curhat", "sedih", "to", "sedih", "banget", "ya", "lidup", "gue", "ehh", "jam", "pantau", "haj"]
tadi aku burn out teruk aku buat bukan lepi	2022-12-05 16:08:37+00:00	RCZMN	["tadi", "aku", "burn", "out", "teruk", "aku", "buat", "bukan", "lepi", "starbucks", "hang", "and", "bell", "starbucks", "grrrrrr", "face", "with", "Big", "eyes", "lepas", "ts"]	["tadi", "aku", "burn", "out", "teruk", "aku", "buat", "bukan", "lepi", "starbucks", "hang", "and", "bell", "starbucks", "grrrrrr", "face", "with", "Big", "eyes", "lepas", "ts"]	["tadi", "aku", "burn", "out", "teruk", "aku", "buat", "bukan", "lepi", "starbucks", "hang", "and", "bell", "starbucks", "grrrrrr", "face", "with", "Big", "eyes", "lepas", "ts"]
kemarin masker pake yg varian coklat kanj	2022-12-05 16:08:37+00:00	ichel	["kemarin", "kemarin", "kemarin", "masker", "pake", "yg", "varian", "coklat", "kayak", "chocolatos", "yg", "coffee", "kayak", "bi", "starbucks", "face", "vr"]	["kemarin", "kemarin", "kemarin", "masker", "pake", "yg", "varian", "coklat", "kayak", "chocolatos", "yg", "coffee", "kayak", "bi", "starbucks", "face", "vr"]	["kemarin", "kemarin", "kemarin", "masker", "pake", "yg", "varian", "coklat", "kayak", "chocolatos", "yg", "coffee", "kayak", "bi", "starbucks", "face", "vr"]
ke kenangan janji jiwa juga boleh starback	2022-12-05 16:08:37+00:00	ohhh..hll	["ke", "se", "kenang", "janji", "jiwa", "starback"]	["ke", "se", "kenang", "janji", "jiwa", "starback"]	["ke", "se", "kenang", "janji", "jiwa", "starback"]
ini ada banyak buaya yang kirim martabak	2022-12-05 16:08:37+00:00	Michael.	["ini", "ada", "banyak", "buaya", "yang", "kirim", "martabak", "amp", "starbucks", "rumah", "a", "w", "sabbie", "anonim", ""]	["ini", "ada", "banyak", "buaya", "yang", "kirim", "martabak", "amp", "starbucks", "rumah", "a", "w", "sabbie", "anonim", ""]	["ini", "ada", "banyak", "buaya", "yang", "kirim", "martabak", "amp", "starbucks", "rumah", "a", "w", "sabbie", "anonim", ""]
td searching klou di starbucks brewed coffee	2022-12-05 16:08:37+00:00	salsssss	["td", "se", "td", "se", "td", "searching", "klou", "starbucks", "brewed", "coffee", "blonde", "roast", "yg", "grande", "kafeinnya", "sekitr", "an", "mg"]	["td", "se", "td", "se", "td", "searching", "klou", "starbucks", "brewed", "coffee", "blonde", "roast", "yg", "grande", "kafeinnya", "sekitr", "an", "mg"]	["td", "se", "td", "se", "td", "searching", "klou", "starbucks", "brewed", "coffee", "blonde", "roast", "yg", "grande", "kafeinnya", "sekitr", "an", "mg"]
terus karyawan starbucks gimana nasibnya	2022-12-05 16:08:37+00:00	Naren.	["terus", "karyaw", "karyawan", "starbucks", "gimana", "nasib", "mas", "kalo", "ngga", "gur", "ngopi", "starbucks", ""]	["terus", "karyaw", "karyawan", "starbucks", "gimana", "nasib", "mas", "kalo", "ngga", "gur", "ngopi", "starbucks", ""]	["terus", "karyaw", "karyawan", "starbucks", "gimana", "nasib", "mas", "kalo", "ngga", "gur", "ngopi", "starbucks", ""]

Gambar 7. Stemming

Pada Gambar 7 dapat dilihat bahwa data yang dihasilkan dari proses step removal berupa kata akan dihilangkan imbuhan (Azhar et al., 2022).

3.2.6 Labelling

Pada proses ini dimana data yang dihasilkan dari proses Stopword Removal akan dilakukan perhitungan berdasarkan polarity dari ulasan opini yang didapatkan, sehingga memperoleh 3 kategori pada klasifikasi yakni, label negatif dimana (< 0) dan label yang bernilai netral (nilai = 0), sedangkan untuk label positif dimana (nilai > 1).

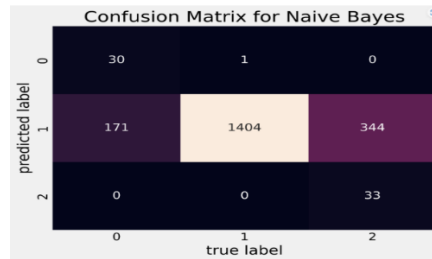


Gambar 8. Labelling

Pie chart pada Gambar 8 menjelaskan dimana klasifikasi ulasan opini masyarakat terhadap starbucks menggunakan jejaring media sosial twitter, dapat dilihat label yang bernilai positif sebesar 19,2%, label yang bernilai negatif sebesar 10,4% dan label yang bernilai netral sebesar 70,3%.

3.3 Validasi dan Evaluasi

3.3.1 Analisis Menggunakan Naive Bayes



Gambar 9. Confusion Matrix Naive Bayes

Pada gambar 9 menjelaskan dimana Confusion Matrix tersebut berbentuk matriks dengan ukuran 3 x 3 yang mengkategorikan masing-masing kelas klasifikasi menjadi Neutral untuk angka 1, positif untuk angka 2 dan negatif untuk angka 0.

Tabel 1. Hasil Confusion Matrix

Confusion Matrix		Predict Class (Kelas Prediksi)		
		Negatif	Neutral	Positif
Actual Class (kelas sebenarnya)	Negatif	30	1	0
	Neutral	171	1404	344
	Positif	0	0	33

Pada Tabel 1 menjelaskan hasil Confusion Matrix dari masing-masing kelas yang terdapat pada kondisi kelas sebenarnya (actual class), dimana dapat memprediksi model dengan benar dan dapat dijelaskan sebagai berikut, nilai TP sebanyak 30 data untuk kelas negatif, nilai TP sebanyak 1404 data untuk kelas neutral dan nilai TP sebanyak 33 data untuk kelas positif. Dalam melakukan evaluasi pada matriks untuk mendapatkan perhitungan nilai akurasi dengan algoritma Naive Bayes menggunakan bahasa pemrograman python yang terdapat pada jupyter notebok dan dapat dilihat pada gambar sebagai berikut:

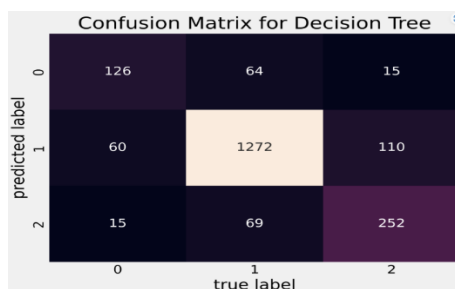
	precision	recall	f1-score	support
0	0.97	0.15	0.26	201
1	0.73	1.00	0.84	1405
2	1.00	0.09	0.16	377
accuracy			0.74	1983
macro avg	0.90	0.41	0.42	1983
weighted avg	0.81	0.74	0.66	1983

Gambar 10. Hasil Accuracy Naive Bayes

Pada Gambar 10 merupakan hasil perhitungan dari matriks menggunakan bahasa pemrograman python yang terdapat pada jupyter notebok yakni 0,74 atau 74%. Berikut terdapat penjelasan perhitungan terhadap nilai akurasi secara manual dari perhitungan matriks :

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{True Positif} + \text{True Negatif}}{\text{Total data matrix}} \times 100 \% \\
 &= \frac{1467}{1983} \times 100 \% = 74\%.
 \end{aligned}$$

3.3.2 Analisis Menggunakan Decision Tree



Gambar 11. Confusion Matrix Decision Tree

Pada gambar 11 menjelaskan dimana Confusion Matrix tersebut berbentuk matriks dengan ukuran 3 x 3 yang mengkategorikan masing-masing kelas klasifikasi menjadi Neutral untuk angka 1, positif untuk angka 2 dan negative untuk angka 0.

Tabel 2. Hasil Confusion Matrix

Confusion Matrix		Predict Class (Kelas Prediksi)		
		Negatif	Neutral	Positif
Actual Class (Kelas Sebenarnya)	Negatif	126	64	15
	Neutral	60	1272	110
	Positif	15	69	252

Pada Tabel 2 menjelaskan hasil Confusion Matrix dari masing-masing kelas yang terdapat pada kondisi kelas sebenarnya (actual class), dimana dapat memprediksi model dengan benar dan dapat dijelaskan sebagai berikut, nilai TP sebanyak 126 data untuk kelas negatif, nilai TP sebanyak 1272 data untuk kelas neutral dan nilai TP sebanyak 252 data untuk kelas positif. Dalam melakukan evaluasi pada matriks untuk mendapatkan perhitungan nilai akurasi dengan algoritma Decision Tree menggunakan bahasa pemrograman python yang terdapat pada jupyter notebok dan dapat dilihat pada gambar sebagai berikut:

	precision	recall	f1-score	support
0	0.61	0.63	0.62	201
1	0.88	0.91	0.89	1405
2	0.75	0.67	0.71	377
accuracy			0.83	1983
macro avg	0.75	0.73	0.74	1983
weighted avg	0.83	0.83	0.83	1983

Gambar 12. Hasil Accuracy Decision Tree

Pada Gambar 12 merupakan hasil perhitungan dari matriks menggunakan bahasa pemrograman python yang terdapat pada jupyter notebok yakni 0,83 atau 83%. Berikut terdapat penjelasan perhitungan terhadap nilai akurasi secara manual dari perhitungan matriks :

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{True Positif} + \text{True Negatif}}{\text{Total data matrix}} \times 100 \% \\
 &= \frac{1650}{1983} \times 100 \% = 83\%.
 \end{aligned}$$

4. KESIMPULAN

Hasil dari penelitian ini dimana dari proses klasifikasi sentimen terhadap kepuasan pelanggan starbucks diperoleh kategori netral, dapat dilihat dari ulasan menggunakan kata kunci “starbuck OR starbucks OR #starbucks” didapatkan hasil yakni komentar positif sebanyak 476 tweet dengan jumlah persentase sebesar 19,2%, komentar neutral sebanyak 1743 tweet dengan jumlah persentase sebesar 70,3 % dan komentar negatif sebanyak 258 tweet dengan jumlah persentase sebesar 10,4%, sehingga dapat ditarik kesimpulan berdasarkan perhitungan polarity tersebut ulasan komentar terhadap starbuck memiliki kategori puas. Pada penelitian ini dapat ditarik kesimpulan bahwa kinerja algoritma Decision Tree lebih baik dari pada algoritma Naive Bayes, dapat dilihat dari penjelasan berikut. Algoritma Decision Tree pada hasil akurasi nilai sebesar 83%. Sedangkan algoritma Naive Bayes pada hasil akurasi nilai sebesar 74%.

REFERENCES

- [1] A. S. D. Herlambang, E. Komara, and A. Sulisty Herlambang, “Pengaruh Kualitas Produk, Kualitas Pelayanan, Dan Kualitas Promosi Terhadap Kepuasan Pelanggan (Studi kasus pada Starbucks Coffee Reserve Plaza Senayan),” *Jurnal Ekonomi, Manajemen dan Perbankan (Journal of Economics, Management and Banking)*, vol. 7, no. 2, 2021.
- [2] J. P. Putra, T. Janji, and R. Sitinjak, “Pengaruh kualitas produk dan harga terhadap kepuasan pelanggan Kopi Starbucks di Summarecon Mall Kelapa Gading 3,” *Jurnal Ilmiah Akuntansi dan Keuangan*, vol. 4, no. 8, p. 2022, 2022, [Online]. Available: <https://journal.ikopin.ac.id/index.php/fairvalue>
- [3] N. Bayes Yunitasari, H. Siti Hopipah, and R. Mayasari, “Optimasi Backward Elimination untuk Klasifikasi Kepuasan Pelanggan Menggunakan Algoritme k-Nearest Neighbor (k-NN) dan,” *Technomedia Journal (TMJ)*, vol. 6, no. 1, 2021, doi: 10.33050/tmj.v6i1.
- [4] A. K. Febrian, Y. H. Chrisnanto, D. Pupita, N. Sabrina, and J. Achmad Yani, “SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Studi Komparasi Metode Klasifikasi K-Nearest Neighbor dan Naive Bayes dalam Mengidentifikasi Kepuasan Pelanggan Terhadap Produk,” *Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika*, p. 333, 2022, doi: 10.31284/p.snestik.2022.2717.
- [5] M. A. Nurhakim, Y. Widiastiw, and N. Chamidah, “Analisis Sentimen Terhadap Ulasan Kepuasan Pelanggan Pada Marketplace Tokopedia Di Jejaring Sosial Twitter Menggunakan Algoritma Naive Bayes,” 2022.



- [6] N. Saurina, T. Rahayuningsih, L. Retnawati, F. Teknik, U. Wijaya, and K. Surabaya, “Analisis Sentimen Ulasan Pelanggan Batik Ecoprint Menggunakan Naïve Bayes Dan KNN Classifier,” vol. 9, no. 2, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [7] D. Sepri, P. Algoritma, N. Bayes, U. Analisis, K. Penggunaan, and A. Bank, “media cetak,” *Journal of Computer System and Informatics (JoSYC)*, vol. 2, no. 1, pp. 135–139, 2020.
- [8] M. I. Petiwi, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 1, p. 542, Jan. 2022, doi: 10.30865/mib.v6i1.3530.
- [9] D. Jacarria Pangestu and A. Kodar, “Implementasi Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Pelayanan Perusahaan Otobus Menggunakan Data Facebook (Studi Kasus: Grup Facebook Murni Jaya Lovers),” vol. 7, no. 3, 2022.
- [10] A. Syafii, G. Dwilestari, and A. Ajiz, “KOMPARASI ALGORITMA NAÏVE BAYES DAN ALGORITMA C4.5 DALAM KLASIFIKASI PELANGGAN PRODUK INDIHOME,” *JURSIMA Jurnal Sistem Informasi dan Manajemen*, vol. 10, no. 2, 2022, [Online]. Available: <https://ejournal.stmikgici.ac.id/>
- [11] W. I. Rahayu, A. Anindita, and M. N. Fauzan, “PENENTUAN VALIDASI DATA PEMILIH DAN KLASIFIKASI HASIL PEMILU DPRD KAB.BONE UNTUK MEMPREDIKSI PARTAI PEMENANG MENGGUNAKAN METODE NAIVE BAYES Program Studi D4 Teknik Informatika 123 Politeknik Pos Indonesia 123,” 2022.
- [12] B. Hendrik and B. R. Suteja, “Identifikasi Risiko Program Maintenance dalam Pengelolaan Proyek Berbasis Agile Menggunakan Pohon Klasifikasi,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3545.
- [13] A. Tangkelayuk and E. Mailoa, “Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes Dan Decision Tree,” vol. 9, no. 2, pp. 1109–1119, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [14] F. Fatmawati and N. Narti, “Perbandingan Algoritma C4.5 dan Naïve Bayes Dalam Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Daring,” *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 4, no. 1, pp. 1–12, May 2022, doi: 10.35746/jtim.v4i1.196.
- [15] M. A. Djamaludin, A. Triayudi, and E. Mardiani, “Analisis Sentimen Tweet KRI Nanggala 402 di Twitter menggunakan Metode Naïve Bayes Classifier,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 6, no. 2, p. 2022, 2022, doi: 10.35870/jti.
- [16] A. Rozaqi, A. Triayudi, and R. T. Aldisa, “Analisis Sentimen Vaksinasi Booster Berdasarkan Twitter Menggunakan Algoritma Naïve Bayes dan K-NN,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 1, p. 184, Oct. 2022, doi: 10.30865/json.v4i1.4907.
- [17] N. A. Susanti and M. Walid, “KLASIFIKASI DATA TWEET UJARAN KEBENCIAN DI MEDIA SOSIAL MENGGUNAKAN NAIVE BAYES CLASSIFIER,” 2022. [Online]. Available: www.kaggle.com
- [18] A. Jurnal et al., “Perbandingan Metode K-Nearest Neighbors dan Naïve Bayes Classifier Pada Klasifikasi Status Gizi Balita di Puskesmas Muara Jawa Kota Samarinda,” *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 1, 2022, doi: 10.30872/atasi.v1i1.25.
- [19] K. Grąbczewski, *Meta-Learning in Decision Tree Induction*, Warsaw: Springer International Publishing, 2014.
- [20] F. T. Hristea, *The Naïve Bayes Model*, London: Springer Cham Heidelberg New York Dordrecht London, 2013.