



Plagiarism Checker X Originality Report

Similarity Found: 8%

Date: Monday, July 15, 2024

Statistics: 300 words Plagiarized / 3884 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

ANALISIS SENTIMEN KOMENTAR YOUTUBE TERHADAP FOOD VLOGGER DENGAN MENGGUNAKAN METODE NAÏVE BAYES 1Sulistia Maharani Harahap Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara Email : sulistiamaharani354@gmail.com 2Rakhmat Kurniawan Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara Email : rakhmat.Kr@uinsu.ac.id ABSTRACT YouTube is currently the media most frequently accessed by internet users in Indonesia with a percentage of 88%, then WhatsApp at 83%, Facebook at 81%, and Instagram at 80%. One of the themes that is often uploaded is about food reviews and people who upload them are called food vloggers.

Food vloggers influence millennial audiences, namely before and after food consumption. The purpose of food review videos is to provide reviews of food in terms of taste, price, quality, atmosphere, etc. Other people also have the freedom to comment on food review videos.

This research aims to analyze public sentiment towards food vloggers using data from YouTube social media. This research aims to analyze sentiment using the Naïve Bayes algorithm method for food vloggers used to analyze sentiment with an accuracy value of 44.8276%, precision of 16% or 0.16%, recall of 100% or 1.00%, and f1-score of 27% or 0.27%. This sentiment analysis can be a reference in understanding food vlogger preferences.

Keywords: sentiment, food vlogger, YouTube, naive Bayes, text preprocessing

PENDAHULUAN Perkembangan teknologi informasi membawa suatu perubahan dalam masyarakat. Lahirnya media sosial menjadikan pola perilaku masyarakat mengalami perubahan baik budaya, etika, dan norma yang ada. Penggunaan platform media sosial (medsos) mengalami peningkatan pesat saat ini [1].

Platform seperti Instagram, Facebook, Twitter, dan YouTube menyediakan beragam informasi kepada pengguna dalam bentuk gambar, video, dan teks. Di antara ketiga media sosial tersebut. Di Indonesia, YouTube memiliki peningkatan yang sangat tinggi. Banyak channel YouTube dengan berbagai kategori konten video yang di upload oleh user atau pengguna youtube [2].

YouTube saat ini merupakan media yang paling sering diakses oleh pengguna internet di Indonesia dengan persentase sebesar 88%, kemudian WhatsApp sebesar 83%, Facebook sebesar 81%, dan Instagram sebesar 80%. Salah satu tema yang banyak diunggah adalah mengenai ulasan makanan (food review) dan orang yang mengunggahnya disebut dengan food vlogger.

Food vlogger atau sebutan pada pembuat konten tentang makanan ini mampu memberikan pengaruh dan meningkatkan penonton untuk membeli produk baik makanan maupun minuman yang ditampilkan [3]. Kolom komentar yang disediakan oleh youtube dapat dijadikan sebuah informasi Analisis Media Sosial yang diklasifikasikan menjadi opini positif dan opini negatif.

Opini public terhadap tayangan di media sosial youtube dapat diketahui dengan analisis sentiment [4]. Model Natural Language Processing (NLP) telah mencapai kesuksesan fenomenal dalam tugas linguistik dan semantik seperti klasifikasi teks, terjemahan mesin, sistem dialog kognitif, pencarian informasi melalui Natural Language Understanding (NLU), dan Natural Language Generation (NLG) [5].

Pendekatan text mining dengan NLP (Natural Language Processing) dilakukan dalam menganalisa dan mencari makna dari setiap opini atau pendapat pengguna Youtube. Analisa sentimen memerlukan metode dalam mengklasifikasikan teks, salah satunya yakni metode Gaussian naïve bayes [7]. Algoritma Naïve Bayes merupakan suatu bentuk klasifikasi data dengan menggunakan metode probabilitas dan statistik.

Naïve bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasar pada penerapan aturan bayes dengan asumsi ketidaktergantungan yang kuat. Algoritma Naïve Bayes ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes yang merupakan memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya.

Analisis sentimen adalah kegiatan yang digunakan untuk menganalisis pendapat atau opini seseorang tentang suatu topik. Hasil dari penelitian ini adalah klasifikasi komentar kedalam sentiment positif, dan sentiment negatif. Hasil pengujian algoritma yang digunakan diperoleh nilai akurasi yang menunjukkan seberapa persen ketepatan algoritma Naïve Bayes Classifier untuk melakukan klasifikasi sentiment [7].

Berdasarkan latar belakang diatas, maka penulis melakukan untuk mengambil judul "Analisis Sentimen Komentar YouTube Terhadap Food Vlogger Dengan Menggunakan Metode Naïve Bayes" yang bertujuan untuk menghasilkan klasifikasi dari komentar – komentar pengguna YouTube dengan kategori positif dan negative. Diharapkan sistem ini dapat menjadi bahan evaluasi para pemilik video meningkatkan kualitas dari saluran YouTube-nya.

METODE PENELITIAN Terkait penelitian yang akan dilakukan, penulis berencana untuk membuat sistem yang dapat melakukan analisis sentimen agar dapat mengetahui sentimen yang diberikan oleh para pengguna youtube terkait mengenai review – review makanan yang dilakukan oleh food vlogger, apakah sentimen tersebut bersifat positif atau negatif dengan menggunakan metode naïve bayes classifier. Adapun tahapan yang dilalui dalam penelitian ini meliputi : Gambar 1 Kerangka Berfikir 1. Identifikasi Masalah Prosedur penelitian yang digunakan dalam penelitian ini diawali dengan identifikasi suatu permasalahan atau permasalahan.

Pada tahap identifikasi masalah dirumuskan latar belakang masalah yang menjadi alasan mengapa penelitian ini dilakukan. 2. Studi Literatur Langkah selanjutnya adalah tinjauan pustaka, yaitu penelusuran jurnal ilmiah, tesis, dan disertasi untuk dijadikan referensi judul penelitian. 3. Pengumpulan Data Pengumpulan data merupakan tahapan yang paling penting dalam melakukan sebuah penelitian untuk mengumpulkan data yang diperlukan terkait penelitian yang akan dilakukan. 4. Disain Sistem Diagram alur proses yang menjadi sebagai acuan bagi peneliti sebelum sistem tersebut dirancang dikenal sebagai diagram alur proses.

Di dalam diagram alur proses ini mencakup tiga komponen utama yakni input, proses, dan output. Gambar 2 di bawah ini. Gambar 2 Desain Sistem 5. Pembobotan TF - IDF Dalam langkah ini, dilakukan perhitungan bobot untuk pada suatu kata (term) yang terdapat dalam suatu dataset yang berasal dari komentar pengguna YouTube. Proses TF-IDF dimulai dengan menghitung nilai term frequency (TF) dan document frequency (DF).

Selanjutnya, dihitung nilai invers document frequency (IDF) dan bobot akhir diperoleh dengan perkalian nilai TF dan IDF [7]. $TFIDF(t,d) = TF(t,d) \times IDF(t,d)$ (1) $TFIDF(t,d) =$

$1 + \frac{1}{N} + \frac{1}{N} + \dots + \frac{1}{N} = 1$(2) $P(x_i, y_j) = P(x_i)P(y_j)$??(3)

6. Algoritma Naïve Bayes Algoritma Naïve Bayes Classifier adalah metode yang menerapkan konsep probabilitas bersyarat.

Pada dasarnya algoritma NBC dapat **dinotasikan pada persamaan sebagai berikut.** ??

$P(y_j) = \sum_{i=1}^N P(x_i, y_j) = \sum_{i=1}^N P(x_i)P(y_j)$ (4) Untuk $P(x_i, y_j)$??

dan $P(x_i, y_j)$ dihitung pada saat pelatihan dimana persamaannya adalah sebagai berikut : $P(x_i, y_j) = \frac{1}{N} \sum_{i=1}^N P(x_i, y_j)$ (5) $P(x_i, y_j)$??

.....(6) 7. Evaluasi Hasil Hasil pengujian menggunakan metode

Naïve Bayes divisualisasikan dalam bentuk confusion matrix, yang dimana bentuk ini memberikan perbandingan hasil klasifikasi Naïve Bayes yang dibantu oleh metode pembobotan TF-IDF.

HASIL DAN PEMBAHASAN **Analisis Data Pada penelitian ini** penulis akan menganalisis data sentimen terhadap komentar youtube food vlogger dengan melakukan pengumpulan data. Proses pengumpulan data dilakukan dengan teknik **crawling data komentar youtube** food vlogger dengan menggunakan library python. Data yang dipakai **didapatkan dari hasil crawling** dan disimpan dan diproses dengan format.csv.

2. Pengumpulan Data Pemrosesan crawling data komentar youtube food vlogger **menggunakan bahasa pemrograman python** pada aplikasi Google Colab. Gambar 3 Hasil Crawling Data 3. Preprocessing Data Tahap ini memiliki beberapa tahapan yaitu cleaning, tokenize, case folding, normalize, filtering.

Berikut adalah alur pada proses preprocessing : Tabel 1 Data Hasil Cleaning Hasil Komentar _Hasil Cleaning _Waduh...gratis lage...kebangeten...jadi senang bgt tuh...**Waduhgratis lagekebangetenjadi senang bgt tuh** _ _Abang ownernya sangat friendly _Abang ownernya sangat friendly _ _Mirip **bang evan hobby makan sihh** ownerx _Mirip bang evan hoby makan sih ownerx _ _Gimana nih kalau **Bang Evan Hobby Makan** Kulineran Kesini.. Ayo like yang banyak biar Bang Evan Kesini..

Waduhgratis lagekebangetenjadi senang bgt tuh _ _Gimana nih kalau Bang Evan Hoby Makan Kulineran Kesini Ayo like yang banyak biar Bang Evan Kesini _ _Ayo haha _Ayo haha _ _Case Folding : Pada tahap ini, seluruh huruf pada dataset diseragamkan **menjadi huruf kecil (Lowercase).** Tabel 2 Hasil Casefolding Hasil Casefolding _waduhgratis lagekebangetenjadi senang bgt tuh _ _abang ownernya sangat friendly _ _mirip bang evan hoby makan sih ownerx _ _gimana nih kalau bang evan hoby makan kulineran kesini ayo like yang banyak biar bang evan kesini _ _ayo haha _ 2. Tokenizing Pada tahap ini dilakukan untuk memisahkan dokumen dalam dataset menjadi kata per kata.

Tabel 3 Data Hasil Tokenizing Hasil Tokenizing _ [waduhgratis, lagekebangetenjadi,

seneng, bgt, tuh] _ [abang, ownernya, sangat, friendly] _ [mirip, bang, evan, hoby, makan, sih, ownerx] _ [gimana, nih, kalau, bang, evan, hoby, makan, kulineran, kesini, ayo, like, yang, banyak, biar, bang, evan, kesini] _ [ayo, haha] _ Stopword Removal : Tahap ini dilakukan menghapus kata dalam dokumen yang tidak memiliki arti seperti kata ganti, kata penghubung dan lainnya.

Tabel 4 Data Hasil Stopword Removal Hasil Stopword Removal _ [waduh, gratis, lage, kebangeten, jadi, seneng, bgt, tuh] _ [abang, ownernya, friendly] _ [bang, evan, hoby, makan, sih, ownerx] _ [gimana, nih, bang, evan, hoby, makan, kulineran, kesini, ayo, like, biar, bang, evan, kesini] _ [ayo, haha] _ Stemming Tahap ini dilakukan untuk menghapus seluruh imbuhan.

Tabel 5 Data Hasil Stemming Hasil Stemming _ Waduh gratis lage kebangeten jadi neng banget tuh _ abang ownernya friendly _ bang evan hoby makan sih ownerx _ gimana nih bang evan hoby makan kuliner kesini ayo like biar bang evan kesini _ ayo haha _ Normalization Pada tahap Normalization dilakukan untuk memperbaiki, menghapus kata yang tidak baku, singkatan serta penulisan kata yang tidak sesuai bahasa Indonesia yang baku.

Tabel 6 Data Hasil Normalization Hasil Normalization _ Waduh gratis lage kebangeten jadi neng banget tuh _ abang ownernya friendly _ bang evan hoby makan sih ownerx _ gimana nih bang evan hoby makan kuliner kesini ayo like biar bang evan kesini _ ayo haha _ 4. Labelisasi Dataset Pada tahap ini, data akan diproses secara otomatis yaitu melakukan perhitungan nilai skor sentimen dengan bahasa python.

Pada penelitian ini penulis menggunakan google collaboratory dengan menggunakan bahasa python. Gambar 4 Pelabelan Data 5. Pembobotan TF-IDF Tahap ini bertujuan untuk menghitung bobot tiap kata (term) yang terdapat didalam dokumen (N). Berikut contoh perhitungan dari proses TF-IDF dari dokumen yang diberikan sebelumnya disajikan dalam tabel 7.

Tabel 7 Hasil Akhir Preprocessing No Hasil Normalization _ 1 waduh gratis lage kebangeten jadi neng banget tuh _ 2 abang ownernya friendly _ 3 bang evan hoby makan sih ownerx _ 4 gimana nih bang evan hoby makan kuliner kesini ayo like biar bang evan kesini _ 5 ayo haha _ Tabel 8 Kata – Kata Pada Kalimat diatas waduh _ bang _ makan _ gratis _ evan _ kulnr _ lage _ hoby _ kesini _ kebangetan _ makan _ ayo _ jadi _ sih _ like _ neng _ ownernya _ biar _ banget _ gimana _ bang _ tuh _ nih _ evan _ abang _ bang _ kesini _ ownernya _ evan _ ayo _ frendly _ hoby _ haha _ Berdasarkan kata-kata seperti pada tabel 8 di atas, maka dihitunglah nilai term frequency.

D1 mewakili kalimat nomor 1, D2 mewakili kalimat nomor 2, D3 mewakili kalimat nomor 3, D4 mewakili kalimat nomor 4, dan D5 Tabel 9 Perhitungan Nilai TF dan DF Term _TF _DF _ _D1 _D2 _D3 _D4 _D5 _ _abang _0 _1 _0 _0 _0 _1 _ _ayo _0 _0 _0 _1 _1 _2 _ _banget _1 _0 _0 _0 _0 _1 _ _bang _0 _0 _1 _2 _0 _3 _ _biar _0 _0 _0 _1 _0 _1 _ _evan _0 _0 _1 _2 _0 _3 _ _frendly _0 _1 _0 _0 _0 _1 _ _gratis _1 _0 _0 _0 _0 _1 _ _gimana _0 _0 _1 _0 _1 _ _haha _0 _0 _0 _0 _1 _1 _ _hoby _0 _0 _1 _1 _0 _2 _ _jadi _1 _0 _0 _0 _0 _1 _ _kebangetan _1 _0 _0 _0 _0 _1 _ _kuliner _0 _0 _0 _1 _0 _1 _ _kesini _0 _0 _0 _2 _0 _2 _ _lage _1 _0 _0 _0 _0 _1 _ _like _0 _0 _0 _1 _0 _1 _ _makan _0 _0 _1 _1 _0 _2 _ _neng _1 _0 _0 _0 _1 _ _nih _0 _0 _0 _1 _0 _1 _ _ownerx _0 _1 _0 _0 _0 _1 _ _ownerx _0 _0 _1 _0 _0 _1 _ _sih _0 _0 _1 _0 _0 _1 _ _tuh _1 _0 _0 _0 _0 _1 _ _berikut disajikan tabel hasil perhitungan IDF melanjutkan dari tahap perhitungan sebelumnya.

Berikut disajikan contoh perhitungan untuk term pertama yaitu "abang" yang mana telah diketahui jumlah $d = 5$ dan $df = 1$????? = ????? ?? + 1 ????? + 1 + 1 ????? = ?????
 $5 + 1 + 1 + 1 + 1 = 2,098612$ Tabel 10 Perhitungan Nilai IDF Term _DF _Skor IDF _ _abang _
 1 _2,09861 _ayo _2 _1,69315 _banget _1 _2,09861 _bang _3 _1,40546 _biar _1
 _2,09861 _evan _3 _1,40546 _frendly _1 _2,09861 _gratis _1 _2,09861 _gimana _1
 _2,09861 _haha _1 _2,09861 _hoby _2 _1,69315 _jadi _1 _2,09861 _kebangetan _1
 _2,09861 _kuliner _1 _2,09861 _kesini _2 _1,69315 _lage _1 _2,09861 _like _1
 _2,09861 _Berikut disajikan perhitungan untuk term pertama yaitu "abang" yang mana telah diketahui nilai $tf = 1$ dan $idf = 2,098612$?? ???? = ???? ???? ?? ????? ?? ???? = 1 ??
 $2,098612 = 2,098612$ Tabel 11 Perhitungan Nilai TF-IDF (Wdt) Term _TF-IDF (Wdt) _ _
 _D1 _D2 _D3 _D4 _D5 _ _abang _0 _2,098612 _0 _0 _0 _ _ayo _0 _0 _0 _2,098612
 _2,098612 _banget _2,098612 _0 _0 _0 _0 _ _bang _0 _0 _2,098612 _2,81092 _0 _ _biar _0
 _0 _0 _2,098612 _0 _ _evan _0 _0 _2,098612 _2,81092 _0 _ _frendly _0 _2,098612 _0 _0 _0 _
 _gratis _2,098612 _0 _0 _0 _0 _ _gimana _0 _0 _0 _2,098612 _0 _ _haha _0 _0 _0 _0
 _2,098612 _hoby _0 _0 _2,098612 _2,098612 _0 _ _jadi _2,098612 _0 _0 _0 _0 _
 _kebangetan _2,098612 _0 _0 _0 _0 _ _kuliner _0 _0 _0 _2,098612 _0 _ _kesini _0 _0 _0 _2
 _0 _ _lage _2,098612 _0 _0 _0 _0 _ _like _0 _0 _0 _2,098612 _0 _ _makan _0 _0 _2,098612
 _2,098612 _0 _ _neng _2,098612 _0 _0 _0 _0 _ _nih _0 _0 _0 _2,098612 _0 _ _ownerx _0
 _2,098612 _0 _0 _0 _ownerx _0 _0 _2,098612 _0 _0 _sih _0 _0 _2,098612 _0 _0 _tuh
 _2,098612 _0 _0 _0 _0 _ _waduh _2,098612 _0 _0 _0 _0 _ _ Gambar 5 Pembobotan TF-IDF pada Sistem 6. Split Dataset Dalam penelitian ini akan diterapkan split data dengan perbandingan 8:2 yang artinya 80% dari total dataset akan digunakan sebagai data uji.

Gambar 6 Perbandingan Data Latih dan Data Uji 7. Klasifikasi Naive Bayes Setelah tahapan pembobotan kata selesai, maka akan dilakukan tahap pengklasifikasian terhadap data uji dengan metode naive bayes. Tabel 12 Contoh Data Latih Dokumen
 _Data Opini _Label _ _D1 _Parah lu koh tidak bayar kasihan orang usaha _Negatif _ _D2
 _masakan enak goyang nasi padang _Positif _ _D3 _sampayang enak minum udah mulai

panas_Positif __D4_bayarlah bang abis makan_Negatif __D5_pantesan makin gemuk makan gratis terus_Negatif __1.Data latih sentiment positif Tabel 13 Data Latih Positif #
 _Term_Frekuensi (?? ??) __1_masakan_1 __2_enak_2 __3_goyang_1 __4_nasi_1 __5_padang_1 __6_sampayang_1 __7_minum_1 __8_udah_1 __9_mulai_1 __10_panas_1 __Pada tabel diatas, dapat diketahui bahwa : Jumlah frekuensi keseluruhan sentiment positif (n) = 11 Jumlah kosakata keseluruhan sentiment positif = 10 2.Data latih sentiment negatif Tabel 14 Data Latih Sentimen Negatif # _Term _Frekuensi (?? ??)
 __1_parah_1 __2_lu_1 __3_koh_1 __4_tidak_1 __5_bayar_1 __6_kasih_1 __7_orang_1 __8_usaha_1 __9_bayarlah_1 __10_bang_1 __11_abis_1 __12_makan_2 __13_pantesan_1 __14_makin_1 __15_gemuk_1 __16_gratis_1 __17_terus_1 __Pada tabel diatas, dapat diketahui bahwa : Jumlah frekuensi keseluruhan sentiment negatif (n) = 18 Jumlah kosakata keseluruhan sentiment negatif = 17 Data Uji Tabel 15 Contoh Data Uji Data Uji _Label __gila murah banget cinta nasi padang surga banget __
 __Tahap dalam proses perhitungan klasifikasi kelas pada daa uji dimulai dari perhitungan nilai prior probability, conditiona probability, dan posterior probability.

Perhitungan nilai prior probability $P(?? ??) = \frac{|????????|}{|????????| + |????????|} = \frac{1}{1+1} = 0,5$ P (Positive) = 1 2 = 0,5 P (Negative) = 1 2 = 0,5 Perhitungan nilai probability setiap term Data Latih Probabilitas setiap term dapat dicari dengan menggunakan rumus : $P(?? ?? | ?? ??) = \frac{P(?? ??) \times P(?? ?? | ?? ??)}{P(?? ??) + 1 \times P(?? ?? | ?? ??)}$ Probabilitas data latih sentiment positif
 $P(\text{masakan}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{enak}|\text{positif}) = \frac{2 + 1}{11 + 10} = 0,1428$
 $P(\text{goyang}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{nasi}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$
 $P(\text{padang}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{sampayang}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{minum}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{udah}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{mulai}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$ $P(\text{panas}|\text{positif}) = \frac{1 + 1}{11 + 10} = 0,0952$
 Probabilitas data latih sentiment negative $P(\text{parah}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{lu}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{koh}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{tidak}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{bayar}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{kasihan}|\text{negaif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{orang}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{usaha}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{bayarlah}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{bang}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{abis}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{makan}|\text{negatif}) = \frac{2 + 1}{18 + 17} = 0,0857$ $P(\text{pantesan}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{makin}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{gemuk}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 $P(\text{gratis}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$ $P(\text{terus}|\text{negatif}) = \frac{1 + 1}{18 + 17} = 0,0571$
 Dengan menghitung nilai probabilitas dari semua kategori sentimen maka terdapat perubahan nilai probabilitas dari setiap term pada data latih.

Sehingga perubahan probabilitas sebagai berikut : Tabel 16 Perubahan Probabilitas Term Pada Sentimen Positif # _Term _Frekuensi (?? ??) _Probabilitas $P(?? ?? | ?? ??)$ _
 _1_masakan_1 $\frac{(1+1)}{(11+27)} = 0,052$ __2_enak_2 $\frac{(2+1)}{(11+27)} = 0,079$ __3

$\frac{1}{11+27} = 0,052$ $\frac{4}{11+27} = 0,052$ $\frac{5}{11+27} = 0,052$ $\frac{1}{11+27} = 0,052$ $\frac{6}{11+27} = 0,052$ $\frac{7}{11+27} = 0,052$ $\frac{8}{11+27} = 0,052$ $\frac{9}{11+27} = 0,052$ $\frac{10}{11+27} = 0,052$

Tabel 17 Perubahan Probabilitas Term Pada Sentimen Negatif

_Term _Frekuensi (?? ??) _Probabilitas ??(?? ?? | ?? ??)

$\frac{1}{18+27} = 0,044$ $\frac{2}{18+27} = 0,044$ $\frac{3}{18+27} = 0,044$ $\frac{4}{18+27} = 0,044$ $\frac{5}{18+27} = 0,044$ $\frac{6}{18+27} = 0,044$ $\frac{7}{18+27} = 0,044$ $\frac{8}{18+27} = 0,044$ $\frac{9}{18+27} = 0,044$ $\frac{10}{18+27} = 0,044$ $\frac{11}{18+27} = 0,044$ $\frac{12}{18+27} = 0,044$ $\frac{2}{2+18} = 0,066$ $\frac{13}{18+27} = 0,044$ $\frac{14}{18+27} = 0,044$ $\frac{15}{18+27} = 0,044$ $\frac{16}{18+27} = 0,044$ $\frac{17}{18+27} = 0,044$

2. Data Uji Dalam tahap klasifikasi data uji, term yang terdaftar pada tabel 15 frekuensi data uji akan diklasifikasi dicari nilai probabilitasnya dengan membandingkan dan mencocokkan dengan term yang ada pada sentimen positif dan sentimen negatif data latih. a.

Kelas sentiment positif $P(\text{gila}|\text{positif}) = \frac{0}{11+27} = 0,026$ $P(\text{murah}|\text{positif}) = \frac{0}{11+27} = 0,026$ $P(\text{banget}|\text{positif}) = \frac{0}{11+27} = 0,026$ $P(\text{cinta}|\text{positif}) = \frac{0}{11+27} = 0,026$ $P(\text{nasi}|\text{positif}) = \frac{1}{11+27} = 0,052$ $P(\text{padang}|\text{positif}) = \frac{1}{11+27} = 0,052$ $P(\text{surga}|\text{positif}) = \frac{0}{11+27} = 0,026$ Berdasarkan nilai probabilitas data uji pada sentiment positif diatas, nilai $P(\text{gila}|\text{positif}) \times P(\text{murah}|\text{positif}) \times P(\text{banget}|\text{positif}) \times P(\text{cinta}|\text{positif}) \times P(\text{nasi}|\text{positif}) \times P(\text{padang}|\text{positif}) \times P(\text{surga}|\text{positif}) = 1,6063620352E-12$

Kelas sentiment negative $P(\text{gila}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{murah}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{banget}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{cinta}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{nasi}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{padang}|\text{negatif}) = \frac{0}{18+27} = 0,022$ $P(\text{surga}|\text{negatif}) = \frac{0}{18+27} = 0,022$ Berdasarkan nilai probabilitas data uji pada sentiment positif diatas, nilai untuk sentiment positif adalah sebagai berikut : $P(\text{gila}|\text{positif}) \times P(\text{murah}|\text{positif}) \times P(\text{banget}|\text{positif}) \times P(\text{cinta}|\text{positif}) \times P(\text{nasi}|\text{positif}) \times P(\text{padang}|\text{positif}) \times P(\text{surga}|\text{positif}) = 1,247178944E-13$

Evaluasi Hasil Hasil pengujian menggunakan metode Naïve Bayes divisualisasikan dalam bentuk confusion matrix, yang dimana bentuk ini memberikan perbandingan hasil klasifikasi Naïve Bayes yang dibantu oleh metode pembobotan TF-IDF.

Tabel 11 Confusion Matrix Hasil Klasifikasi ij _Kelas Prediksi (j) __ _Positif _Negatif _
 _Kelas Aktual (i) _Positif _Negatif
 Perhitungan accuraci: $\frac{33+0+16+10}{33+0+16+10} \times 100\% = 44,8276$ Perhitungan Precision $\frac{33}{33+0+16} \times 100\% = 16\%$ Perhitungan Recal $\frac{33}{33+0} \times 100\% = 100\%$

Perhitungan $F1\text{score} = \frac{2 \times \frac{16}{16+100}}{16+100} \times 100\% = 27\%$ Gambar 7
 Classification Report KESIMPULAN Berdasarkan hasil penelitian analisis sentimen komentar youtube food vlogger menggunakan data hasil crawling dengan kata kunci 'food vlogger' diperoleh data sebanyak 1000 komentar kemudian data tersebut melalui tahap preprocessing yang dimana tahapan didalamnya meliputi cleaning, case folding, tokenizing, stopword removal, stemming, normaization, agar data yang diperoleh dapat digunakan pada tahap pelabelan dengan metode TF-IDF. Selanjutnya data yang telah dilabeli akan diklasifikasi dengan memakai metode Naïve Bayes.

Dari hasil confusion matrix diperoleh performa klasifikasi dengan metode Naïve Bayes yang diperkuat dengan pembobotan TF-IDF yang menghasilkan persentase akurasi sebesar 44.8276 % dari total 1000 komentar, dengan persentase sebesar 58% untuk untuk precision, 16% untuk nilai recall 100% untuk nilai f1-score 27%. Daftar Pustaka [1] Jannah, F. R., Megiananta Aprilistya, A., & Khadijah, S. (2023).

Analisis Framing Pemberitaan Perseteruan Farida Nurhan Dan Food Vlogger Codeblu Di Kompas.Com Dan Viva.Co.Id. Prosiding Seminar Nasional, 1040–1048. [2] Calvin Jonathan, Theresia Herlina Rochadiani, & Thamrin Sofian. (2023). Analisis Sentimen Komentar Video Youtube Flat Earth Theory Dengan Menggunakan Metode Unsupervised Dan Supervised Learning.

Decode: Jurnal Pendidikan Teknologi Informasi, 3(2), 378–387. <https://doi.org/10.51454/decode.v3i2.210> [3] Afifah, N. U. R. (2022). PURCHASE INTENSITY ANALYSIS OF MUSLIM MILLENNIAL CONSUMERS WITH YOUTUBE FOOD CONCENT. 7(1), 57 – 72 [4] Chely Aulia Misrun, Haerani, E., Fikry, M., & Budianita, E. (2023). Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier.

Jurnal CoSciTech (Computer Science and Information Technology), 4(1), 207–215. <https://doi.org/10.37859/coscitech.v4i1.4790> [5] Riset, J., Hasibuan, M.S. and Serdano, A. (2022) 'Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan Support Vector Machine dan Naive Bayes Policy Sentiment Analysis Face-to-face Learning Using Supports Vector and Naive Bayes Engines', 6(2), pp. 199–204. [6] Mujahidin, S., Prasetio, B., & Utomo, M. C. C. (2022).

Implementasi Analisis Sentimen Masyarakat Mengenai Kenaikan Harga BBM Pada Komentar Youtube Dengan Metode Gaussian naïve bayes. Voteteknika (Vocational Teknik Elektronika Dan Informatika), 10(3), 17. <https://doi.org/10.24036/voteteknika.v10i3.118299> [7] Lubis, A.H. and Harahap, Y.F. (2023) 'Analisis Sentimen Masyarakat Terhadap Resesi Ekonomi Global 2023

INTERNET SOURCES:

<1% -

<https://jurnal.univpgri-palembang.ac.id/index.php/redoks/article/download/14160/8040>

<1% - <https://djournals.com/klik/article/download/1246/755>

<1% -

<https://www.statista.com/statistics/1399512/indonesia-share-of-youtube-users-by-age-group/>

<1% -

https://www.academia.edu/91908971/Do_Millennials_Believe_in_Food_Vlogger_Reviews_A_Study_of_Food_Vlogs_as_a_Source_of_Information

1% -

<https://www.djkn.kemenkeu.go.id/kpknl-semarang/baca-artikel/14366/Pengaruh-Positif-dan-Negatif-Media-Sosial-Terhadap-Masyarakat>

<1% -

<https://www.kompas.id/baca/riset/2020/10/27/berebut-pelanggan-lewat-konten-di-media-sosial/>

<1% - <https://id.techinasia.com/fakta-perkembangan-youtube-di-indonesia>

1% -

<https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/download/16903/16620>

1% - <https://journal.ipm2kpe.or.id/index.php/JKS/article/download/4630/2955>

1% - <http://repository.uin-suska.ac.id/72414/1/CHELY%20AULIA%20MISRUN.pdf>

<1% -

<https://www.unite.ai/id/nlp-bangkit-dengan-model-transformator-analisis-komprehensif-t5-bert-dan-gpt/>

<1% - <https://ieeexplore.ieee.org/document/9965301>

1% -

<https://media.neliti.com/media/publications/124041-ID-prediksi-kehadiran-menggunakan-metode-kl.pdf>

<1% -

<https://medium.com/@abbinizarm/research-sentimental-analysis-on-youtube-comments-435ac9c484d6>

<1% -

<http://www.repository.nusaputra.ac.id/id/eprint/93/1/RITA%20APRIANI%20Si20.pdf>

<1% -

<https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/16897>
<1% - http://repository.upi.edu/18374/4/T_BIND_1103372_Chapter3.pdf
<1% - <https://deepublishstore.com/blog/latar-belakang-penelitian/>
<1% -
<https://news.detik.com/berita/d-4850130/metode-pengumpulan-data-kuantitatif-dan-kualitatif>
<1% - <https://dltsierra.medium.com/algorithm-tf-idf-633e17d10a80>
<1% - <https://jurnal.unidha.ac.id/index.php/jteksis/article/download/1263/787>
<1% - https://kc.umh.ac.id/16924/7/BAB_III.pdf
<1% -
<https://blog.kazee.id/mengenal-data-crawling-pengertian-manfaat-dan-cara-kerjanya>
<1% - <https://www.youtube.com/watch?v=sm60ogJSPXU>
<1% - <https://site.amalsholeh.com/hobby-makan-dan-amalsholeh-com/>
<1% -
<https://esairina.medium.com/pre-processing-data-teks-menggunakan-python-64e18f7fbd6f>
<1% -
<https://apriiiana.blogspot.com/2017/09/penerapan-tokenisasi-stopword-removal.html>
<1% - <https://www.ruangguru.com/blog/contoh-kata-baku-dan-tidak-baku>
<1% - <http://jurnal.unmuhjember.ac.id/index.php/ELKOM/article/download/18424/4751>
<1% -
<https://dibimbing.id/blog/detail/memahami-teknik-validasi-silang-k-fold-dalam-machine-learning>
<1% - <http://eprints.umg.ac.id/808/3/7.%20BAB%20II.pdf>
<1% - <https://proceeding.unesa.ac.id/index.php/sniis/article/download/881/365/3193>
<1% -
https://www.researchgate.net/publication/375565359_Analisis_Sentimen_Komentar_Video_Youtube_Flat_Earth_Theory_Dengan_Menggunakan_Metode_Unsupervised_Dan_Supervised_Learning
<1% - <https://sinta.kemdikbud.go.id/journals/google/9619>
<1% - <https://www.mendeley.com/catalogue/c9799a43-9f74-32c4-98cb-7bc35bbb0fb2/>
<1% - <https://jurnalnasional.ump.ac.id/index.php/JRST/article/download/15145/5421>
<1% - <https://journal.unnes.ac.id/nju/index.php/JIPK/article/view/36674>