

ANALISIS SENTIMEN OPINI MASYARAKAT TERKAIT PENGUNAAN THREAD TANPA IZIN OLEH NESSIE JUDGE

Wisah Sri Mulyani¹, Petrus Sokibi², Viar Dwi Kartika³, Kusnadi⁴, Ridho Taufiq Subagio⁵
Universitas Catur Insan Cendekia^{1,2,3}

Jl. Kesambi 202, Kota Cirebon, Jawa Barat Tlp : (0231) 220250

e-mail : wisah.mulyani.ti.19@cic.ac.id¹, petrus.sokibi@cic.ac.id², viardk@cic.ac.id³, kusnadi@cic.ac.id⁴,
ridho.taufiq@cic.ac.id⁵

Abstrak

YouTuber storytelling merupakan seorang content creator yang membuat konten pada platform YouTube dengan menjadi pencerita. Pada tanggal 4 dan 5 Februari 2023, ada dua user Twitter yang menulis tweet yang berisi bahwa pada konten Nessie Judge terdapat thread miliknya yang diambil dan dipakai untuk pembuatan konten, pihak Nessie Judge tidak izin terlebih dahulu untuk membawakan thread tersebut pada kontennya dan malah mengunggahnya pada platform YouTube. Nessie Judge termasuk sudah melanggar hak cipta. Masyarakat pun banyak yang ikut berkomentar. Tetapi banyaknya opini menjadikannya sulit dan membutuhkan banyak waktu untuk menganalisis satu per satu komentar tersebut. Maka dari itu dilakukan analisis sentimen untuk menganalisa dan mengklasifikasi pendapat dari masyarakat terhadap Nessie Judge terkait dengan penggunaan thread tanpa izin yang sudah dilakukannya dengan menggunakan metode Naïve Bayes Classifier. Analisis sentimen adalah bidang penelitian yang menganalisis pendapat, perasaan, evaluasi, sikap, dan perasaan orang tentang entitas yang diungkapkan dalam teks tertulis dan karakteristiknya. Hasil dari penelitian ini yaitu berupa presentase komentar positif, netral dan negatif yang didapatkan dari perhitungan data testing dengan menggunakan metode Naïve Bayes Classifier. Penelitian ini dapat menjadi acuan agar YouTuber lain lebih mempertimbangkan perihal izin dan tidak melakukan hal serupa karena akan merugikan orang lain dan dirinya sendiri.

Kata kunci: Analisis Sentimen, Naïve Bayes Classifier, YouTuber Storytelling, Nessie Judge, Twitter.

Abstract

YouTuber storytelling is a content creator who creates content on the YouTube platform by becoming a storyteller. On February 4 and 5 2023, two Twitter users wrote tweets stating that in Nessie Judge's content, there was a thread of their own that was taken and used for content creation, Nessie Judge did not get permission to bring this thread into her content and instead uploaded it on YouTube platforms. Nessie Judge includes copyright infringement. Many people also commented. But the sheer number of opinions makes it difficult and it takes a lot of time to analyze each. Therefore, sentiment analysis was carried out to analyze and classify the opinion of the public towards Nessie Judge regarding the use of threads without permission which she had done using the Naïve Bayes Classifier method. Sentiment analysis is a field of research that analyzes people's opinions, feelings, evaluations, attitudes, and feelings about entities expressed in written texts and their characteristics. The results of this study are the percentage of positive, neutral, and negative comments obtained from the calculation of testing data using the Naïve Bayes Classifier method. This research can be a reference so that other YouTubers are more considerate about permission and do not do the same thing because it will harm others and themselves.

Keywords: Sentiment Analysis, Naïve Bayes Classifier, YouTuber Storytelling, Nessie Judge, Twitter.

1. PENDAHULUAN

YouTuber *storytelling* merupakan seorang *content creator* yang membuat konten pada platform YouTube dengan menjadi pencerita atau pendongeng dari suatu cerita baik itu cerita *horror*, *true crime*, sejarah dan lain sebagainya. Beberapa YouTuber *storytelling* Indonesia yang cukup terkenal yaitu Nessie Judge, Nadia Omara, Hirotada Radifan dan Korea Reomit. Pada beberapa waktu lalu tepatnya pada tanggal 4 Februari 2023, nama Nessie Judge menjadi *trending* di Twitter karena seorang *user* menulis *tweet* yang berisi bahwa pada salah satu konten Nessie Judge terdapat *thread* miliknya yang dipakai untuk konten Nessie Judge dan menyebut bahwa konten yang dibawakan oleh Nessie itu merupakan konten daur ulang. Dan pada 5 Februari 2023, ada *user* lain yang juga mengatakan kalau *thread* miliknya juga diambil dan dipakai untuk pembuatan konten, pihak Nessie Judge tidak izin terlebih dahulu untuk membawakan *thread* tersebut pada kontennya dan malah mengunggahnya pada platform YouTube. Dengan adanya kejadian tersebut, Nessie Judge pun meminta maaf dan juga melakukan *take down* pada konten yang disebutkan. Tetapi masih banyak netizen yang juga menanggapi kalau Nessie tidak tulus dalam meminta maaf sehingga banyak opini terkait hal ini.

Berdasarkan uraian di atas, Nessie termasuk telah melanggar hak cipta dan juga melakukan plagiarisme karena telah mengambil informasi tanpa izin dari pemiliknya dan mengkomersialkannya pada publik untuk kepentingan pribadi. Ketentuan tentang hak cipta ini tertuang dalam Undang-Undang Nomor 12 Tahun 1997 tentang Hak Cipta. Undang-Undang Nomor 12 Tahun 1997 khususnya pasal 11 ayat 1 menyatakan bahwa program komputer adalah ciptaan yang dilindungi dengan jangka waktu perlindungan selama 50 (lima puluh) tahun sejak pertama kali diumumkan, termasuk muatan *online* seperti teks, grafik, gambar, video dan audio pada situs web [1]. Maka dari itu, jelas bahwa Nessie Judge termasuk telah melanggar Undang-Undang tersebut karena kedapatan mengambil teks dan gambar tanpa izin yang terdapat dalam *thread* di Twitter.

Dengan begitu, pandangan masyarakat terkait YouTuber satu ini tentu menjadi beragam. Ada yang tetap mendukung, ada yang netral dan ada pula yang menentang. Terlebih pada masa sekarang ini opini ataupun sentimen dari masyarakat dapat bebas diungkapkan pada berbagai media. Terutama pada media sosial, salah satunya yaitu Twitter. Twitter adalah salah satu media sosial yang didirikan oleh Jack Dorsey, yang biasanya berfungsi untuk mengirimkan pesan yang dikenal dengan *tweet* [2]. Adanya kebebasan dalam berpendapat tersebut pun menghasilkan berbagai opini, baik itu opini yang terkesan positif atau mendukung, opini yang netral dan juga opini negatif atau menentang. Tetapi banyaknya opini menjadikannya sulit dan membutuhkan banyak waktu untuk menganalisis satu per satu komentar tersebut. Dengan begitu, adanya analisis sentimen sangat membantu. Analisis sentimen, juga dikenal sebagai penelitian opini publik, adalah bidang penelitian yang menganalisis pendapat, perasaan, evaluasi, sikap, dan perasaan orang tentang entitas yang diungkapkan dalam teks tertulis dan karakteristiknya. Entitas ini dapat berupa produk, layanan, organisasi, orang, peristiwa, masalah, atau topik [3]. Oleh karena itu, penulis tertarik untuk melakukan analisis sentimen terhadap masyarakat terkait masalah pelanggaran hak cipta atau pemakaian *thread* tanpa izin yang sudah dilakukan oleh Nessie Judge dengan menggunakan metode *Naïve Bayes Classifier*.

Berdasarkan permasalahan yang sudah diuraikan, maka tujuan dari penelitian yang akan dilakukan oleh penulis adalah membuat analisis sentimen yang mampu menganalisa tingkat positif, netral maupun negatif dari komentar masyarakat pada media sosial Twitter terhadap Nessie Judge terkait dengan penggunaan *thread* tanpa izin yang sudah dilakukannya. Pada penerapan metode *Naïve Bayes Classifier* ini akan melalui 3 tahapan yaitu perhitungan peluang *priors*, lalu perhitungan peluang *conditional*, serta pemilihan kategori.

1.1. Identifikasi Masalah

Berdasarkan uraian latar belakang masalah tersebut, penulis membuat identifikasi masalah sebagai berikut:

1. Opini masyarakat terhadap YouTuber *storytelling* Indonesia Nessie Judge terkait dengan penggunaan *thread* tanpa izin yang dilakukannya.
2. Dampak dari penggunaan *thread* tanpa izin yang dilakukan oleh Nessie Judge terhadap Nessie Judge-nya sendiri.
3. Cara mengimplementasikan metode *Naïve Bayes Classifier* pada analisis sentimen sehingga mendapat hasil positif, netral dan negatif.
4. Hasil dari penelitian ini berupa presentase dari komentar yang sudah diteliti.

1.2. Batasan Masalah

Batasan permasalahan dalam penulisan ini dimaksudkan agar pembahasan dan penulisan laporan dapat dilakukan secara terarah dan mencapai sasaran. Batasan masalahnya antara lain:

1. Objek dari penelitian ini adalah Nessie Judge.
2. Pengambilan data melalui *Scraping* Twitter dengan kata kunci “Nessie Judge”.
3. Pengambilan data menggunakan Bahasa Indonesia dan Bahasa Inggris.
4. Data yang digunakan merupakan komentar dari pengguna Twitter yang membicarakan Nessie Judge terkait pengambilan *thread* tanpa izin untuk dijadikan konten YouTube per tanggal 4 Februari - 7 Februari 2023.
5. Analisis data dalam penelitian ini menggunakan analisis sentimen dengan metode *Naïve Bayes Classifier*.
6. Aplikasi yang digunakan untuk analisis data dalam penelitian ini adalah *software* Jupyter Notebook yang menggunakan Python versi 3.8.8 dan Microsoft Excel.
7. Dibuatkan *dashboard* menggunakan *library* dash Python dengan Jupyter Notebook sebagai *output*.

1.3. Tujuan

Adapun tujuan dari analisis sentimen pada penelitian ini sebagai berikut:

1. Memberikan informasi yang lebih akurat tentang opini masyarakat terhadap Nessie Judge terkait penggunaan *thread* tanpa izin yang dilakukannya.
2. Memberikan informasi mengenai dampak dari penggunaan *thread* tanpa izin yang sudah dilakukan oleh Nessie Judge terhadap Nessie Judge-nya sendiri.
3. Memberikan informasi mengenai pengklasifikasian data opini masyarakat terhadap Nessie Judge dengan metode *Naïve Bayes Classifier*.
4. Memberikan informasi mengenai tingkat akurasi dari metode *Naïve Bayes Classifier* dalam pengklasifikasian opini masyarakat terhadap Nessie Judge.
5. Memberikan informasi mengenai presentase komentar positif, netral dan negatif terkait permasalahan tersebut.

2. KAJIAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen, juga dikenal sebagai penelitian opini publik, adalah bidang penelitian yang menganalisis pendapat, perasaan, evaluasi, sikap, dan perasaan orang tentang entitas yang diungkapkan dalam teks tertulis dan karakteristiknya. Entitas ini dapat berupa produk, layanan, organisasi, orang, peristiwa, masalah, atau topik. Karena penelitian dan aplikasi analisis sentimen yang ada lebih fokus pada teks tertulis, ini menjadi area penelitian aktif dalam pemrosesan bahasa alami (NLP) [3]. Tipe analisis sentimen yang digunakan pada penelitian ini yaitu *fine-grained statement analysis* karena tipe analisis ini dapat mengelompokkan opini kedalam kategori positif, netral dan negatif. Metode klasifikasi yang digunakan pada penelitian ini yaitu *machine learning*, karena sifat-sifatnya dapat mengidentifikasi emosi (perspektif manusia) dalam teks.

2.2. Data Mining

Data mining atau penambangan data adalah proses menemukan korelasi, pola, dan tren baru yang signifikan dengan mencari melalui sejumlah besar data yang disimpan dalam arsip menggunakan teknik pengenalan pola dan teknik statistik dan matematika. Penambangan data adalah proses menemukan pola dan informasi yang menarik dari sejumlah besar data [4]. Beberapa jenis data mining yaitu clustering, klasifikasi, regresi, asosiasi, anomali detection, text mining, social network analysis, dan time series analysis.

2.3. Twitter

Twitter merupakan sebuah layanan *microblogging*. Pengguna Twitter mengikuti atau diikuti oleh pengguna lain. Seorang pengguna dapat mengikuti pengguna lain dan pengguna yang saat ini diikuti tidak

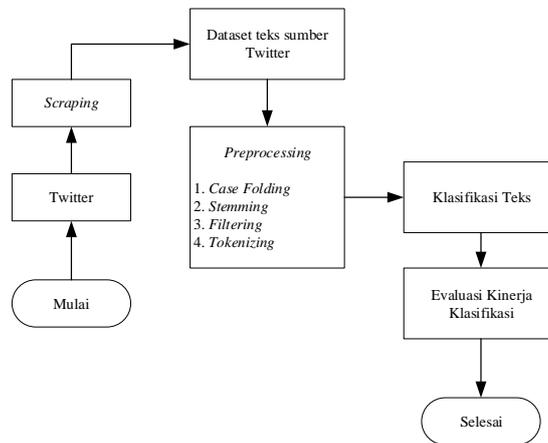
harus mengikuti kembali. Menjadi anggota Twitter berarti pengguna menerima semua pesan (disebut *tweet*) dari pengguna yang mengikuti mereka. Praktik umum membalas *tweet* telah berkembang menjadi budaya pemberian *tag* yang jelas: ‘RT’ atau *retweet* berarti balasan *tweet*, ‘@’ diikuti dengan tanda alamat atau nama pengguna tertentu, dan ‘#’ biasanya diindikasikan sebagai tagar diikuti dengan kata tertentu. Kosa kata *tag* ini didefinisikan dengan baik, dikombinasikan dengan batas ketat 140 karakter per pesan untuk kemudahan penggunaan dan singkatnya ekspresi. Mekanisme *retweet* memungkinkan pengguna untuk membagikan informasi pilihan mereka tanpa harus membeli informasi asli [2].

2.4. Algoritma Naïve Bayes Classifier

Metode *Naïve Bayes Classifier* adalah metode dalam data mining untuk mengklasifikasikan data menggunakan perhitungan probabilitas. Klasifikasi dibagi menjadi dua fase yang terdiri dari pembelajaran dan pengklasifikasian. Fase pembelajaran membaca data yang diketahui kelasnya, sedangkan fase pengklasifikasian membentuk data yang akan diuji. *Naïve bayes* diklasifikasikan dan diprediksi untuk nilai-nilai dimasa depan berdasarkan nilai sebelumnya [5].

3. METODE PENELITIAN

Metode penelitian dalam penelitian menggunakan algoritma *Naïve Bayes Classifier* untuk menganalisa sentimen pada teks opini masyarakat terkait penggunaan *thread* tanpa izin oleh YouTuber *Storytelling* Indonesia Nessie Judge dengan alur sebagai berikut:



Gambar 1. Flowchart Alur Metode Pengolahan dan Analisis Data

4. HASIL DAN PEMBAHASAN

4.1. Contoh Perhitungan Algoritma Naïve Bayes Classifier

a) Proses Training

Sebuah dokumen *training* sudah di *preprocessing* dan diklasifikasi secara manual sehingga menghasilkan data sebagai berikut:

Tabel 1. Contoh Kasus Data Training

<i>Tweet</i>	<i>Text</i>	Kelas
<i>Tweet 1</i>	hate fan lihat youtube kak ajar etika izin susah	0
<i>Tweet 2</i>	maaf tulus ajar salah salah perkara hate love manner manner tingkat ya	-1
<i>Tweet 3</i>	mau si <i>thread</i> dapat opini jelek bagus ya wajar izin dulu iya misal dapat opini bagus izin <i>thread</i> protes orang tim etika izin kocak	0
<i>Tweet 4</i>	masalah source kak tolong banget akun gede tonton okay cantum <i>source</i> bahas izin all hide main problem put <i>source desc</i> basi kak lihat hindar	1
<i>Tweet 5</i>	imho etika point izin ranah <i>socmed</i> bebas orang <i>post</i> ulang daur ulang <i>share save</i> konsumsi pribadi baik izin unggah unggah begitu maksud	0
<i>Tweet 6</i>	mbak izin biar enak dua belah pihak nyaman ambil minimal izin hasil uang cerita orang orang izin manner mbak	-1

Tweet 7 makin tinggi pohon makin kencang angin tiup cocok nessie pure masalah kok 1
 komersil tulis orang lain izin kemarin *tweet* creator daur ulang

Dari data tabel 1 dibuat sebuah model probabilitas prior dengan mengacu pada persamaan:

$$P(c) = \frac{Nc}{N}$$

Kelas positif muncul 2 kali dalam dokumen *training*. Total dokumen adalah 7 dokumen. Maka dari itu, jika dimasukkan kedalam persamaan:

$$P(c) = \frac{2}{7}$$

Begitu pula untuk kelas negatif dan netral. Setiap dokumen yang muncul dibagi dengan total seluruh dokumen.

Tabel 2. Hasil Perhitungan Probabilitas Prior Untuk Setiap Kelas

Atribut Kelas	P(class)
Positif (1)	2/7
Negatif (-1)	2/7
Netral (0)	3/7

Model probabilitas kondisional, berikut contoh perhitungan probabilitas kondisional untuk *term* “izin” pada kelas positif:

$$P(izin|positif) = \frac{3}{45}$$

Setiap *term* yang ditemukan pun dihitung probabilitas kemunculan untuk setiap kelasnya kemudian dibagi dengan total *term* yang ditemukan. Pada persamaan diatas, 3 merupakan jumlah *term* “izin” pada kelas positif, dan 45 merupakan jumlah total *term* atau kata pada kelas positif. Contoh *term* yang ditemukan terdapat pada tabel 3.

Tabel 3. Contoh Term Pada Dokumen

Term	Positif	Netral	Negatif
izin	3/45	6/47	1/30
maaf	0/45	0/47	1/30
thread	0/45	2/47	0/30
ajar	0/45	1/47	1/30
etika	1/45	2/47	0/30
manner	0/45	1/47	2/30
tim	0/45	1/47	0/30
tonton	0/45	0/47	1/30
minimal	0/45	1/47	0/30
orang	2/45	3/47	0/30

Dalam proses perhitungan, kemungkinan *term* tidak terdapat dalam suatu kelas sangat besar. Apabila *term* tersebut tidak terdapat pada kelas, maka akan bernilai 0. Hal ini dapat menyebabkan seluruh perhitungan menjadi tidak akurat. Oleh karena itu, digunakan teknik *smoothing*. Teknik *smoothing* yang digunakan adalah *laplace smoothing*. *Laplace smoothing* merupakan teknik penambahan angka 1 untuk setiap *term* yang ditemukan agar angka 0 dapat dihilangkan dan hasil perhitungan menjadi lebih akurat. Hasil yang didapat dari proses ini menjadi model untuk melakukan klasifikasi. Berikut adalah contoh perhitungan probabilitas untuk *term* “izin” pada kelas positif:

$$P(\text{izin}|\text{positif}) = \frac{3 + 1}{45 + 122} = 0,02395$$

Pada persamaan diatas angka 3 merupakan jumlah *term* “izin” yang ditemukan. Angka 45 merupakan total *term* pada kelas positif sedangkan 122 adalah total *term* pada seluruh dokumen. Total *term* tersebut didapat dari penjumlahan *term* pada tabel 3.

Tabel 4. Model Perhitungan Klasifikasi

Term	Positif	Netral	Negatif
izin	0,02395209	0,04142011	0,01315789
maaf	0,00598802	0,00591715	0,01315789
thread	0,00598802	0,01775147	0,00657894
ajar	0,00598802	0,01183431	0,01315789
etika	0,01197605	0,01775147	0,00657894
manner	0,00598802	0,01183431	0,01973684
tim	0,00598802	0,01183431	0,00657894
tonton	0,00598802	0,00591715	0,01315789
minimal	0,00598802	0,01183431	0,00657894
orang	0,01796407	0,02366863	0,00657894

Pada tabel 4 merupakan hasil dari perhitungan setiap *term* yang sudah ditambahkan 1. Probabilitas tersebut selanjutnya dicari nilai yang paling besar dan dijadikan sebagai kelas untuk *term* tersebut. Misal *term* “izin” pada tabel 4 kelas positif menunjukkan 0,023 sedangkan pada kelas netral 0,041 dan 0,013 pada kelas negatif. Dengan hasil tersebut didapatkan *term* “izin” pada kelas netral karena nilai netral lebih besar.

b) Proses Testing

Alur proses *testing* kurang lebih sama dengan alur *training*. Hanya saja setelah proses selesai, data *testing* akan dihitung nilai probabilitas akhir. Data terdapat pada tabel 5. Data ini diambil dari tabel 1.

Tabel 5. Data Testing

Tweet	Text	Kelas
Tweet 8	gila nes kecewa rasa <i>subscribe</i> sama tonton video sumpah coba hubungi akun minimal minta maaf karena izin jawab kayak begini	?
Tweet 9	tim riset tim kerja pakai otak minimal pikir minimal dm <i>thread</i> izin begitu cuit konten sekali materi <i>request subscribe</i>	?
Tweet 10	ajar nessie pakai <i>thread</i> twitter konten baik izin hindar drama kayak begini <i>still enjoy watch</i> video terima kasih sudah teman hari beres rumah	?

Perhitungan probabilitas pada bagian $\prod_{k=1}^n P(x_k|C_i)$, dihitung terlebih dahulu agar memudahkan kalkulasi selanjutnya. Hasilnya terdapat pada tabel 7. Untuk sebuah *term* yang kemunculannya lebih dari satu kali, nilai kemunculan tersebut menjadi pangkat dari nilai probabilitas kondisionalnya pada tabel 5 dengan *term frequency* masing-masing. Kemudian jumlahkan nilainya untuk masing-masing kelas.

Tabel 6. Perhitungan Term Dari Data Testing

Term	Positif	Netral	Negatif
izin	1/23	0	2/37
maaf	0/23	0	1/37
thread	1/23	0	1/37
ajar	1/23	0	0/37
tim	0/23	0	2/37
tonton	0/23	0	1/37
minimal	0/23	0	3/37
subscribe	0/23	0	2/37
video	1/23	0	1/37

Tabel 7. Hasil Nilai Probabilitas Kondisional

Term	TF	Positif	Netral	Negatif
izin	3	0,02409638	0	0,03092783
maaf	1	0,01204819	0	0,02061855
thread	2	0,02409638	0	0,02061855
ajar	1	0,02409638	0	0,01030927
tim	2	0,01204819	0	0,03092783
tonton	1	0,01204819	0	0,02061855
minimal	3	0,01204819	0	0,04123711
subscribe	2	0,01204819	0	0,03092783
video	2	0,02409638	0	0,02061855

Misalnya term “izin” memiliki kemunculan sebanyak 3 kali. Kemudian setiap probabilitas dari setiap kelas akan dipangkatkan dengan 3. Untuk nilai kemunculan yang sama agar proses kalkulasi tidak membebani, maka proses tersebut dilakukan satu kali. Hal ini mempercepat proses pencarian probabilitas. Pada tabel 7 disebutkan hasil dari pemangkatan seluruh term yang ditemukan.

Tabel 8. Hasil Perkalian Nilai Probabilitas Kondisional Dengan TF-nya

Term	Positif	Netral	Negatif
izin	0,00001399121	0	0,00002958341
maaf	0,01204819	0	0,02061855
thread	0,00058063552	0	0,0004251246
ajar	0,02409638	0	0,01030927
tim	0,00014515888	0	0,00095653066
tonton	0,01204819	0	0,02061855
minimal	0,0000017489	0	0,00007012367
subscribe	0,00014515888	0	0,00095653066
video	0,00058063552	0	0,0004251246
Total	0,04966009	0	0,05440939

Dari tabel diatas, didapatkan nilai probabilitas setiap term. Untuk mendapatkan kelas yang diinginkan dari dokumen adalah dengan mengalikan probabilitas dokumen dengan probabilitas kelas yang didapat dari hasil perhitungan. Kemudian diambil nilai paling besar diantara kelas tersebut. Sebagai contoh tweet 8. Didapatkan nilai probabilitas dari tweet 8 terhadap seluruh kelas dengan cara mengalikan nilai probabilitas prior dengan total nilai probabilitas kondisional untuk masing-masing kelas.

Probabilities dari tweet 8 terhadap kelas positif:

$$P(\text{positif}|\text{tweet 8}) = \frac{2}{7} * 0,04966009 = 0,01418859714$$

Probabilities dari tweet 8 terhadap kelas negatif:

$$P(\text{negatif}|\text{tweet 8}) = \frac{2}{7} * 0,05440939 = 0,01554554$$

Dari hasil perhitungan probabilitas diatas, diketahui bahwa probabilitas pada tweet 8 terhadap kelas negatif menghasilkan nilai yang lebih tinggi, sehingga tweet 8 dapat diklasifikasikan kedalam kelas negatif.

c) Perhitungan Performance Measure

Pada penelitian ini, dilakukan performance measure yang bertujuan untuk mengevaluasi kinerja model yang digunakan untuk mengklasifikasikan emosi atau opini dalam teks. Dalam hal ini, model yang dimaksud yaitu metode Naïve Bayes Classifier yang digunakan pada penelitian ini. Perhitungan performance measure ini dilakukan terhadap 10 dataset yang tertera pada tabel 1 dan tabel 5. Berikut cara pencarian True Positive, True Negative, False Positive dan False Negative atau nilai confusion matrix dari 10 dataset:

Tabel 9. Pencarian Nilai Confusion Matrix

Kelas Berdasarkan Model	Tweet	Positif	Netral	Negatif
Netral	Tweet 1	0,029	0,048	0,032
Negatif	Tweet 2	0,038	0,064	0,042
Netral	Tweet 3	0,067	0,112	0,074
Positif	Tweet 4	0,038	0,064	0,042
Netral	Tweet 5	0,067	0,112	0,074
Negatif	Tweet 6	0,019	0,032	0,021
Positif	Tweet 7	0,019	0,032	0,021
Negatif	Tweet 8	0,020	0,000	0,022
Negatif	Tweet 9	0,030	0,000	0,033
Positif	Tweet 10	0,015	0,000	0,016

Dari tabel 9. didapatkan hasil pencarian *True Positive*, *True Negative*, *False Positive* dan *False Negative* atau nilai *confusion matrix* sebagai berikut:

Tabel 10. Hasil Pencarian Nilai Confusion Matrix

	Positif	Netral	Negatif
Positif	0	2	1
Netral	0	3	0
Negatif	0	2	2

Hasil pencarian ini didapatkan berdasarkan nominal terbesar dari nilai perhitungan pada tabel 9. Berikut rumus untuk perhitungan *performance measure*:

$$\text{Akurasi} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Presisi} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$F\text{-measure} = 2 * (\text{Presisi} * \text{Recall}) / (\text{Presisi} + \text{Recall})$$

Di mana:

TP: *True Positive*, yaitu data positif yang diklasifikasikan sebagai positif oleh model.

TN: *True Negative*, yaitu data negatif yang diklasifikasikan sebagai negatif oleh model.

FP: *False Positive*, yaitu data negatif yang diklasifikasikan sebagai positif oleh model.

FN: *False Negative*, yaitu data positif yang diklasifikasikan sebagai negatif oleh model.

Perhitungan *performance measure* untuk kelas positif:

$$\text{TP} = 0$$

$$\text{TN} = 2 + 2 = 4$$

$$\text{FP} = 0 + 0 = 0$$

$$\text{FN} = 0 + 2 = 2$$

$$\text{Akurasi} = (0 + 4) / (0 + 4 + 0 + 2) = 0,67$$

$$\text{Presisi} = 0 / (0 + 0) = 0$$

$$\text{Recall} = 0 / (0 + 2) = 0$$

$$F\text{-measure} = 2 * (0 * 0) / (0 + 0) = 0$$

Perhitungan *performance measure* untuk kelas negatif:

$$\text{TP} = 0$$

$$\text{TN} = 0 + 3 = 3$$

$$\text{FP} = 0 + 0 = 0$$

$$\text{FN} = 1 + 2 = 3$$

$$\text{Akurasi} = (0 + 3) / (0 + 3 + 0 + 3) = 0,5$$

$$\text{Presisi} = 0 / (0 + 0) = 0$$

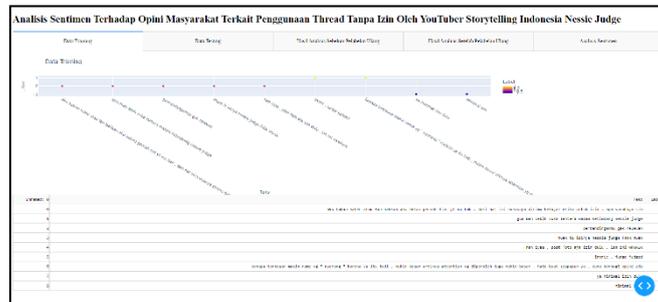
$$\text{Recall} = 0 / (0 + 3) = 0$$

$$F\text{-measure} = 2 * (0 * 0) / (0 + 0) = 0$$

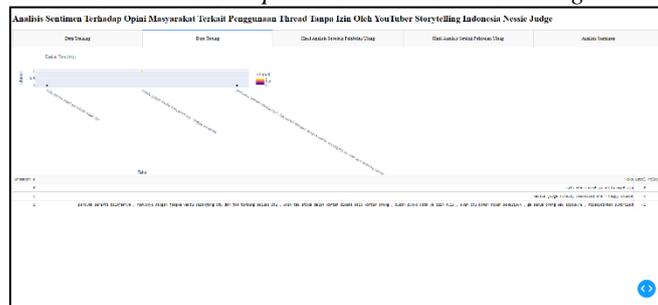
Dari perhitungan *performance measure* tersebut, didapatkan hasil akurasi sebesar 0,67 atau 67% untuk kelas positif dan 0,5 atau 50% untuk kelas negatif dari perhitungan analisis sentimen pada 10 data sampel.

4.2. Hasil Program

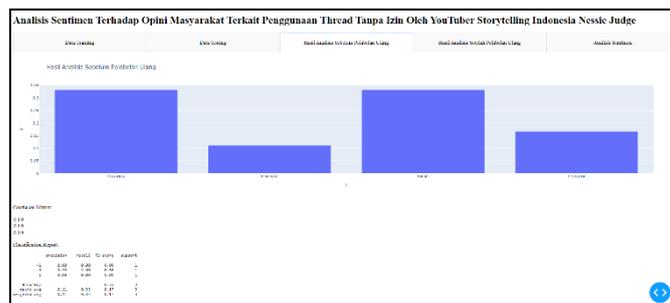
Tampilan *dashboard* contoh perhitungan analisis sentimen Naïve Bayes Classifier menggunakan *library* dash Python.



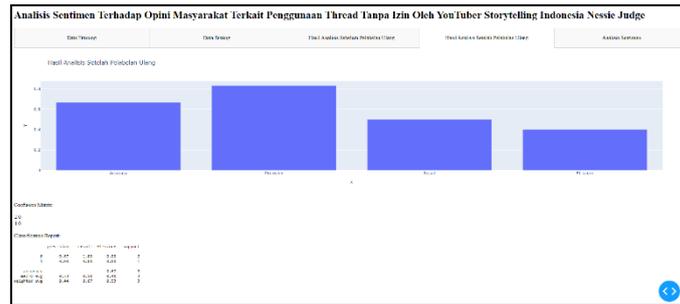
Gambar 2. Tampilan Halaman Data Training



Gambar 3. Tampilan Halaman Data Testing



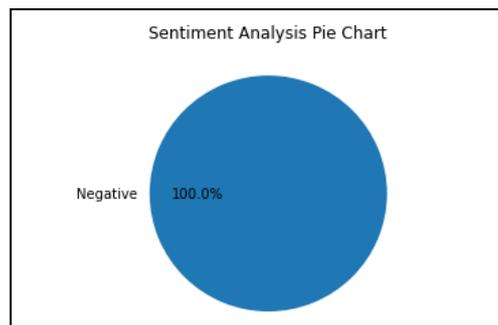
Gambar 4. Tampilan Halaman Hasil Analisis Sebelum Pelabelan Ulang



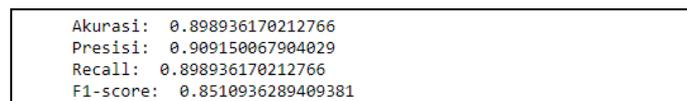
Gambar 5. Tampilan Halaman Hasil Analisis Setelah Pelabelan Ulang



Gambar 6. Tampilan Halaman Hasil Analisis Sentimen



Gambar 7. Tampilan Visualisasi Data Hasil Perhitungan Seluruh Data



Gambar 8. Tampilan Hasil Perhitungan Performance Measure

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan masalah dan pembahasan analisis yang telah diuraikan di atas, maka dapat diambil kesimpulan sebagai berikut:

1. Analisis sentimen ini telah berhasil dalam menganalisis opini masyarakat terkait penggunaan *thread* tanpa izin yang dilakukan oleh salah satu YouTuber *Storytelling* Indonesia Nessie Judge. Dengan hasil sebesar 100% di kelas negatif dan 0% di kelas positif dan netral. Yang berarti bahwa masyarakat sangat menentang hal yang telah dilakukan oleh Nessie Judge.
2. Dampak positif dari kejadian tersebut terhadap Nessie Judge yaitu adanya kenaikan *subscriber* dan nama Nessie Judge menjadi *trending* menjadikan dia lebih dikenal dan mendapatkan penghasilan tambahan. Sedangkan dampak negatifnya yaitu *viewers* konten Nessie yang lebih rendah dari

sebelum kejadian, pandangan orang terhadap Nessie Judge menjadi buruk, dan video yang berkaitan harus di *take down*.

3. Berdasarkan hasil implementasi terhadap analisis yang telah dilakukan, dapat disimpulkan bahwa analisis ini dapat diterapkan untuk membantu menentukan opini atau komentar dari masyarakat di media sosial Twitter terkait penggunaan *thread* tanpa izin oleh salah satu YouTuber *storytelling* Nessie Judge secara cepat dan akurat.
4. Hasil pengujian *performance measure* atau uji akurasi metode *Naïve Bayes Classifier* pun cukup besar, yakni dengan akurasi sebesar 90%, presisi 91%, *recall* 90%, dan *f1-score* 85%.
5. Hasil penelitian ini dapat menjadi acuan untuk YouTuber atau *content creator* lain agar lebih memperhatikan terkait izin dan tidak melakukan hal serupa karena dapat merugikan pihak lain dan diri sendiri.

4.2 Saran

Adapun saran yang dapat penulis berikan untuk pengembangan analisis sentimen ini yaitu:

1. Dapat ditampilkan hasil perhitungan untuk keseluruhan data yang sudah di analisis sebelumnya pada *dashboard* sehingga hasil analisis dapat tervisualisasi.
2. Dapat dijadikan acuan untuk melakukan analisis di media sosial lain terkait suatu topik. Seperti untuk menganalisa *rating* atau komentar pada suatu aplikasi. Sehingga analisis sentimen dapat dilakukan untuk mengetahui pro atau kontra dan *user experience* yang dapat digunakan untuk meningkatkan kualitas suatu produk.

DAFTAR PUSTAKA

- [1] A. Surniandari, "UUIITE Dalam Melindungi Hak Cipta Sebagai Hak Atas Kekayaan Intelektual (HKI) Dari Cybercrime," *Cakrawala: Jurnal Humaniora Bina Sarana Informatika*, vol. 16, no. 1, 2016.
- [2] T. Krisdiyanto, "Analisis sentimen opini masyarakat Indonesia terhadap kebijakan PPKM pada media sosial Twitter menggunakan Naïve bayes classifiers," *Jurnal CoreIT*, pp. 32–37, 2021.
- [3] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge university press, 2020.
- [4] D. T. Larose and C. D. Larose, *Discovering knowledge in data: an introduction to data mining*, vol. 4. John Wiley & Sons, 2014.
- [5] T. E. Putri, R. T. Subagio, and P. Sobiki, "Classification System Of Toddler Nutrition Status using Naïve Bayes Classifier Based on Z-Score Value and Anthropometry Index," in *Journal of Physics: Conference Series*, IOP Publishing, 2020, p. 012005.