Implementation of Naïve Bayes and Decision Tree Algorithms for Sentiment Analysis on the Impact of Hydrometeorological Disasters [Implementasi Algoritma Naïve Bayes dan Decision Tree pada Sentimen Dampak Bencana Hidrometoeorologi]

Arry Wahyu Setiawan[1)], Suprianto[*,2)]

[1)]Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia
[2)] Program Studi Teknik Informat, Universitas Muhammadiyah Sidoarjo, Indonesia
*Email Penulis Korespondensi: suprianto@umsida.ac.id

Abstract. *Hydrometeorological disasters often trigger public responses on social media, so it's requiring analysis public opinion toward such disaster. This study compares the performance of Decision Tree and Naïve Bayes Algorithm to analyze performance on analyzing sentiments of respond related to hydrometeorological disaster. The dataset that used contains of 11890 comments. Comments labeled into Positive, Negative, and neutral then processed by preprocessing, TF-IDF weighting, and classification with 5-fold cross validation. Validation results Naïve Bayes had mean accuracy by 0.79 which higher than Decision Tree, which scored 0,65. Naïve Bayes also got 0.79 score on accuracy, precision, recall, and F1-Score compares to Decision tree that scored scored 0.63, 0.70, 0.63, and 0.64 respectively. Because the gap of the mean accuracy and test accuracy wasn't far off, concludes that Naïve Bayes have better generalization than Decision Tree. This result shows that Naïve Bayes is better than Decision Tree on classification of Hydrometeorological Disaster respond.*

Keywords *- Sentiment Analysis, Hydrometeorology Disaster, Naïve Bayes, Decision tree, TF-IDF, Social Media*

Abstrak. *Bencana hidrometeorologi memicu berbagai respons masyarakat di media sosial. maka diperlukan analisis untuk memahami opini publik terhadap bencana terkait. Tujuan penelitian ini adalah membandingkan kinerja algoritma Naïve Bayes dan Decision Tree dalam melakukan analisis sentimen terkait bencana Hidrometeorologi. Data digunakan berasal dari TikTok, Instagram Facebook yang terdiri dari 11.890 komentar dengan label kelas positif, negatif, dan netral yang kemudian diproses melalui tahapan pra-pemrosesan, pembobotan TF-IDF, dan klasifikasi dan validasi 5-fold cross validation. Hasil validasi menunjukkan Naïve Bayes memperoleh mean accuracy 0,79 lebih tinggi daripada Decision Tree yang mendapat skor 0,65. Pada data uji, Naïve Bayes memperoleh accuracy, precision, recall, dan F1-Score yang seragam 0,79 lebih unggul dari Decision Tree dengan accuracy 0,63, precision 0,70, recall 0,63, dan F1-score 0,64. Dikarenakan jarak mean acmecuracy dan test accuracy berdekatan disimpulkan Naïve Bayes mempunyai kemampuan generalisasi lebih baik daripada Decision Tree. Hasil ini menunjukkan Naïve Bayes unggul dalam mengklasifikasi sentimen terkait bencana Hidrometeorologi*

Kata Kunci *- Analisa Sentimen, Bencana Hidrometeorologi, Naibe Bayes, Decision*

I. PENDAHULUAN

Indonesia adalah salah satu negara yang memiliki penetrasi internet tertinggi di Asia Tenggara. Berdasarkan survey Asosiasi Penyelenggara Jasa Internet Indonesia periode 2023 hingga 2024, pengguna internet di indonesia telah mencapai 221 juta jiwa dari 278 juta penduduk [1]. Tingginya penggunaan internet berdampak pada terbentuknya pola menyampaian opini dan interaksi masyarakat di berbagai platform media sosial. Salah satu peristiwa yang memicu banyaknya respon public di sosial media per desember 2025 adalah bencana hidrometeorologi. Dengan demikian topik ini baik untuk dikaji lebih lanjut.

Media sosial tidak hanya menjadi sarana komunikasi, Media sosial dapat meberikan informasi respons masyarakat terhadap suatu peristiwa. Berberapa penelitian terdahulu menunjukan bahwa komentar media sosial terkait bencana memuat informasi kondisi terdampak dan respons masyarakat terhadap dampak bencana [2][3]. Informasi tersebut berguna untuk sebagai bahan evaluasi bagi pemerintah dan pemangku kepentingan untuk memahami persepsi masyarakat dalam pengambilan keputusan penanganan bencana. Namun tingginya volume komentar pada sosial media menyebabkan secara manual melelahkan dan tidak efisien sehingga diperlukan metode analisis sentimen berbasis algoritma klasifikasi untuk mengelola respon dengan efisien.

Berdasar penelitian-penelitian terdahulu, prosedur yang dikembangkan dalam klasifikasi domain teks khususnya konteks bencana secara garis besar menggunakan tahapan pengumpulan data, prapemrosesn teks, ekstrasi fitur dan klasifikasi. Berberapa penelitian terada pada penelitian Anugrah et al dan Prasetyo et al [4][5] yang menggunakan klasifikasi naïve bayes terhadap respon korban bencana. Sedangkan pada penlitian Rozi et al tahapan

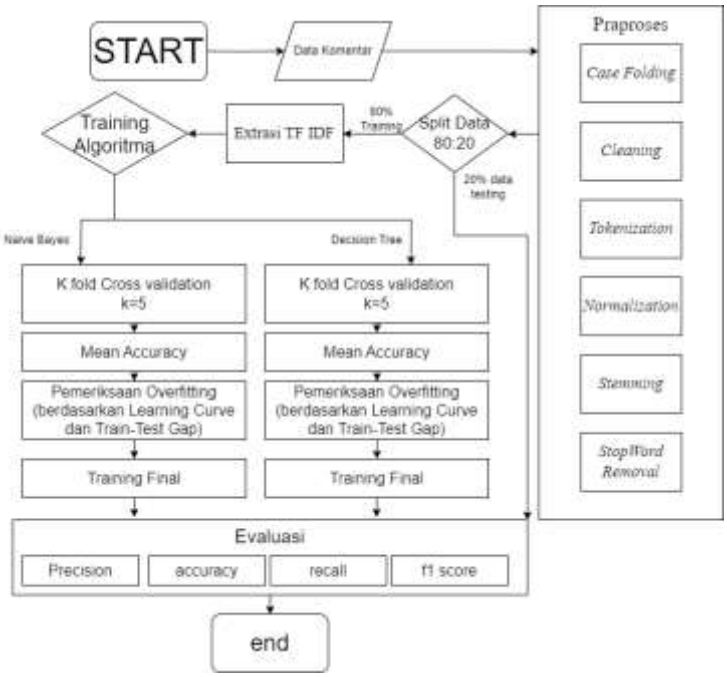
Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

dikembangkan menggunakan k means untuk mengelompokan data yang selanjutnya dikelompokan menggunakan Naïve Bayes [6]. Pada penelitian Das et al menunjukkan pendekatan TF-IDF memperkuat respentatif setiap kata pada klasifikasi domain teks [7]. Pendekatan tersebut diemplementtasi oleh Ioendri et al pada klasifikasi menggunakan Naïve Bayes serta diadopsi oleh Yuniar dan Kismiantini [8], Rizkina, Kustyana et al [9], Baehaqi serta Gerliandeva et al [10][11]pada penlitian analisis sentimen menggunakan Klasifikasi Naïve bayes. Pada konteks Decision Tree, algoritma tersebut digunakan oleh Maharani dan Fathoni [12], Pawar et al. [13], Akbar et al. [14], serta Trianto et al. [15] pada domain teks menghasilkan hasil akurasi yang cukup baik. Secara keseluruhan penelitian, menunjukan algoritma Naïve Bayes dan Decission Tree merupakan algoritma yang cukup efektif dalam klasifikasi teks dengan perbedaan pada pendekatan model klasifikasinya.

Berdasar Kajian penelitan terdahulu, terdapat keterbatasan pada prosedur analisis sentimen yang dikembangkan. Pada penelitian analisis sentimen yang mengambil konteks kebencanaan penelitian sebelum nya hanya terbatas di satu platform sosial media saja. Akibatnya, pemeriksaan respon kurang meluas ke berbagai kalangan platform karena setiap sosial media memiliki karakteristik pengguna dan gaya bahasa yang berbeda-beda selain itu pada penelitan terdahulu baik Naïve Bayes amupun Decision tree tidak membandingkan performa algoritma pada konteks multiplatform. Berdasarkan gap tersebut, kebaruan (state of art) penelitian ini mengembangkan metode klasifikasi teks konteks bencana dengan memperluuas sumber data lebih dari satu platform media sosial dengan membandingkan performa algoritma Naïve Bayes dan Decision Tree dalam klasifikasi sehingga dapat dihasilkan prosedur klasifikasi teks bencana yang lebih baik dari penelitian-penelitian sebelumnya.

II. METODE

Metodologi penelitian yang digunakan pada penelitian ini mengikuti alur proses sebagaimana ditunjukkan pada diagram alir Gambar 1:



Gambar 1. Diagram alir metode penelitian

A. Data Komentar

Data Komentar dikumpulkan melalui metode web scraping menggunakan python dengan library *playwright dan beatifulsoup*. Proses scraping mengambil komentar dari periode Oktober 2025 sampai dengan desember 2025. Komentar diambil dari postingan ber kata kunci bencana yang termasuk hidrometeorologi (contoh : banjir, longsor) dalam 3 platform sosial media, yaitu tiktok, instagram dan facebook. Lalu data akan dilabelkan dengan kriteria berikut:

a) Positif: mengandung dukungan, doa, harapan, atau ucapan terima kasih
b) Negatif: mengandung kritik, keluhan, kekhawatiran, atau kemarahan
c) Netral: berisi informasi faktual tanpa ekspresi emosional dan spam

B. Prapemrosesan data

Prapemrosesan adalah proses dalam memperbaiki data agar ai lebih mudah dalam memahami suatu sentiment. Prapemrosesan ini terdiri dari

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

a) Case Folding : mengubah seluruh huruf menjadi huruf kecil
b) Cleaning : proses menghapus karakter yang tidak diperlukan
c) Tokenisasi :proses memecah kalimat menjadi kata-kata atau token
d) Normalisasi: mengubah kata tidak baku atau bahasa gaul menjadi bentuk baku
e) Stemming :proses mengubah kata berimbuhan menjadi kata dasar.
f) Stopword removal :proses menghapus kata-kata yang sering muncul tetapi memiliki sedikit pengaruh terhadap analisis

C. Pembagian Data 80:20

Pemisahan dataset menjadi 2 porsi, data latih dan data uji. Pembagian data menggunakan rasio 80:20, dimana 80% data duguakan untuk pelatihan model dan 20% sisanya untuk meguji kinerja daripada model. Pembagian tersebut menggunakan metode *stratified sampling.* Data latih digunakan untuk melakukan evaluasi pada model terhadap data yang belum pernah digunakan dalam pelatihan.

D. Pembobotan TF-IDF

Setelah preprocessing, data teks akan diubah menjadi bilangan numerik sehingga ai dapat melihat dan membaca tersebut. Metode yang dipakai peneliti adalah TF-IDF. TF-IDF akan memberikan bobot pada setiap kata dengan menyesuaikan frekuensi. Hasil perhitungan tf idf adalah matrix dari vector numerik yang akan mempresentasikan kepentingan kata dalam suatu dokumen yang ditentukan terhadap seluruh corpus. Vector ini lah yang akan dilakukan training pada algoritma naïve bayes. TF-IDF memiliki rumus sebagai berikut

$$\text{TF}(t, d)= \frac{ft,d}{\sum k\ fk,d} \quad \text{IDF}(t)= \log \frac{N}{nt} \tag{1}$$

TF merupakan nilai untuk menghitung frekuensi suatu kata dalam dokumen, sedangkan idf dugunakan untuk membobotkan berdasar tingkat pentingnya data pada seluruh dokumen dalam korpus. Parameter yang digunakan dama TF-IDF adalah menggunakan unigram, bigram, dan trigram sebagai fitur. Jumlah fitur yang digunakan adalah sebanyak 20.000 kata/frasa.

E Implementasi Algoritma *Naïve Bayes* dan *Decision Tree*

Setelah data diekstraksi menjadi data TF-IDF, dilakukan metode klasifikasi menggunakan algortima *Naive Bayes* dan *Decision Tree* untuk menghitung kemungkinan ada nya suatu data ke dalam kelas tertentu. Algoritma Naïve Bayes mnggunakan proses dimana mengklasifikasikan data berdasar nilai probabilistik menggunakan teorema Bayes. Rumus Naïve Bayes adalah sebagai berikut.

$$P(Ci \mid x)= \frac{P(x \mid Ci)P(Ci)}{P(x)} \tag{2}$$

Pada persamaan (2), P(Ci | x) merupakan probablitias posterior data x ke dalam kelas Ci, P(x | Ci) adalah probabilitas kemunculan fitur x pada kelas Ci. P(Ci) menyatakan probablitias prior kelas Ci, sedangkan P(x) berfungsi sebagai konstanta normalisasi

Disisi lain, Decision Tree (pohon keputusan) merupakan salah satu metode dalam klasifikasi dan prediksi yang berbentuk struktur pohon. Metode ini bekerja dengan cara membagi dataset ke dalam beberapa subset berdasarkan atribut tertentu melalui proses pemilihan fitur terbaik sebagai node (simpul). Struktur Decision Tree terdiri dari Root node (akar): atribut pertama yang digunakan untuk melakukan pemisahan data. Internal node (simpul): atribut lanjutan untuk membagi data. Leaf node (daun): hasil akhir berupa kelas atau keputusan Prinsip utama Decision Tree adalah melakukan pemilihan atribut terbaik untuk membagi data sehingga menghasilkan kelompok yang paling homogen. Untuk menentukan atribut terbaik, digunakan ukuran seperti Entropy dan Information Gain.

Entropy digunakan untuk mengukur tingkat ketidakpastian atau ketidakhomogenan suatu data dengan memanfaatkan proporsi kelas ke I (pi). Entropy dapat dirumuskan sebagai berikut

$$\text{Entropy (S)} = -\sum_{i=1}^{n} pi \log_2 pi \tag{3}$$

Information Gain digunakan untuk menentukan atribut terbaik dalam pemisahan data dari entropy, dapat dirumuskan sebagai berikut:

$$\text{Gain (S, A)}=\text{Entropy(S)}- \sum_{\in \text{Values(A)}} \frac{|Sv|}{|S|} \text{Entropy (Sv)} \tag{4}$$

F. Validasi model dengan *K-fold Cross validation*

Validasi model dilakukan menggunakan metode *k-Fold Cross Validation* untuk memeroleh estimasi performa model dan mengungari bias akibat pembagian data. Pada metode ini dalah latih dibagi menjadi 5 bagian

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

(*fold*) dengan proporsi yang sama menggunakan *Stratified K-Fold*. Pada setiap iterasi satu *fold* digunakan sebagai data validasi. Sedangkan 4 *fold* lainya digunakan sebagai data latih. Proses tersebut diulang sebanyak 5 kali sehingga setiap *fold* mampu menjadi data training. Nilai k sama dengan 5 dipilih karena data testing memiliki besar 1/5 dari jumlah kesuluruhan data.

G. Akurasi rata-rata (*mean accuracy*)

Akurasi rata rata akan digunakan untuk mengukur performa model berdasar hasil validasi menggunakan *K-fold Cross Validation*. Nilai ini diperoleh dari rata-rata akurasi setiap fold. Sehingga didapatkan rumus:

$$\text{Mean Accuracy} = \frac{\sum_{i=1}^{k} Accuracy\ i}{k} \quad (5)$$

Pada Persamaan (5), *accuracy* merupakan nilai akurasi pada fold ke-i, sedangkan k adalah jumlah fold yang digunakan pada proses validasi. Pada penelitian ini digunakan k=5.

H. Pemeriksaan overfitting dengan learning curve

Pemeriksaan overfitting dilakukan menggunakan *learning curve* untuk melakukan evaluasi kemampuan generalisasi pada model. *Learning Curve* dapat menampilkan hubungan antara akurasi data latih dan akurasi data validasi apabila terdapat kurva saling menjauhi, maka model dapat disebutkan overfitting.

Pada penelitian ini, *learning curve* berisi data menggunakan hasil *K-Fold Cross Validation* pada data latih. Selain itu, pemeriksaan overfitting juga membandingkan nilai akurasi latih dan akurasi tes. Semakin kecil selisih suatu data tersebut, maka semakin baik kemampuan model dalam mengenali data yang belum pernah dipelajari.

I. Evaluasi model

Evaluasi Model dilakukan menggunakan data uji yang sudah dipisahkan pada tahap *split data*. Evaluasi model dilakukan untuk mengukur kemampuan model dalam mengklasifikasikan sentimen. Performa model diukut berdasarkan confusion matrix dengan kriteria sebagai *Accuracy*, Precision, *Recall, F1-Score. Accuracy* mencerminkan banyak nya komentar yang dipilih dengan benar oleh model. Dapat dinyatakan di persamaan 6:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision digunakan untuk mengukur ketepatan model dalam memprediksi suatu kelas. Dinyatakan di persamaan 7

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

Recall adalaht tingkat keberhasilan model dalam mengenali seluruh data yang diklasifikasikan terhadap seluruh data yang benar benar pada suatu kelas. Dapat dinyatakan di pernyataan 8

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

F1-Score digunakan dalam mengikut antara nilai precision dan recall, dapat dinyatakan di pernyataan 9

$$\text{F1 Score} = 2 \times \frac{\text{Precision(Recall)}}{\text{Precision+Recall}} \quad (9)$$

III. HASIL DAN PEMBAHASAN

A. Deskripsi data terkumpul

Dataset yang digunakan dalam penelitian ini diperoleh dari komentar netizen pada platform media sosial yang berkaitan dengan topik penelitian. Data dikumpulkan menggunakan teknik scraping. Setelah data terkumpul, data dilabeli menjadi 3 kelas, positif negatif dan netral. Pelabelan dilakukan berturut-turut netral, positif, negatif dengan 0, 1, dan 2. Data yang terkumpul dari platform sosial media terbatas dikarenakan penemuan pembatasan pengambilan pada setiap sosial media. data yang berhasil terkumpul terdapat sebanyak 11,890 komentar dari setiap aplikasi (tiktok, instagram, dan facebook). Dikarenakan pembatasan setiap platform berbeda maka didapatkan data komentar sebanyak 1638 dari facebook, 6120 dari tiktok, dan 4132 dari instagram. Setiap kelas positif, negatif, netral berturut turut mendapatkan jumlah 3230, 2884, dan 5776.

B. Hasil pra-pemrosesan data

Preprocessing data dilakukan dengan 6 fase, Case Folding, cleaning, Tokenisasi, Normalisasi, Stemming, Dan Stopword Removal. Sebagai contoh maka diambil satu data komentar sampel: “Dan pemerintah dengan entengnya bilang kondisi sudah membaik :(”

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Tabel 1. Sampel Preprocessing Data

Tahapan	Komentar
Komentar Asli	Dan pemerintah dengan entengnya bilang kondisi sudah membaik :(
Case Folding	dan pemerintah dengan entengnya bilang kondisi sudah membaik :(
Cleaning	dan pemerintah dengan entengnya bilang kondisi sudah membaik
Tokenisasi	['dan', 'pemerintah', 'dengan', 'entengnya', 'bilang', 'kondisi', 'sudah', 'membaik']
Normalisasi	['dan', 'pemerintah', 'dengan', 'entengnya', 'berkata', 'kondisi', 'sudah', 'membaik']
Stemming	['dan', 'pemerintah', 'dengan', 'enteng', 'bilang', 'kondisi', 'sudah', 'baik']
Stopword Removal	['pemerintah', 'enteng', 'bilang', 'kondisi', 'baik']

Data pada dataset sebelum melalui proses Preprocessing memiliki 111,955 kata. Kata yang sering muncul adalah Pemerintah. Rata-rata jumlah kata per komentar adalah 9 kata dengan komentar terpendek hanya memiliki 1 kata dan komentar terpanjang memiliki 628 kata. Setelah preprocessing, kata berkurang menjadi 90,987 kata.

C. Split Data

Dari 11.890 komentar, data akan dibagi menjadi 2 bagian dengan rasio 80:20. Sebanyak 80% digunakan dalam data latih. Jumlah data latih adalah 9.512 data. 20% dari sisa data latih digunakan dalam data training sebanyak 2.378 data. Pembagian dilakukan menggunakan metode stratified sampling sehingga proporsi setiap kelas sentimen tetap terjaga pada kedua kelompok data.

D. Hasil Pembobotan TF-IDF

Setelah proses pembagian data, data latih kemudian dibobotkann menggunakan metode TF-IDF. TF-IDF menghasilkan representasi numerik dari setiap dokumen berdasar korpus. Hasil ini kemudian akan digunakan dalam proses klasifikasi Naïve Bayes dan Decision Tree. Sampel dari hasil tf idf adalah sebagai berikut

Tabel 2. Lima sampel data yang telah dibobotkan TF-IDF

Fitur	Bobot TF-iDF
Sedih	0.0829
perintah	0,0206
tanya	0,0116
lihat	0,0106
doa	0,0106

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

Fitur	Bobot TF-iDF
Sedih	0.0829

E. Perbandingan Parameter Model

Pada algoritma Naïve bayes pada penelitian ini menggunakan parameter alpha sebesar setengah satuan sebagai nilai laplace smoothing. Parameter ini digunakan untuk mengurangi munculnya probabilitas nol pada proses klasifikasi sehingga model dapat menghasilkan hasil prediksi dengan stabil.

Pada Algoritme Decision Tree digunakan parameter entropy sebagai dasar pemilihan atribut terbaik pada setiap proses pemisahan node. Untuk kedalaman pohon sendiri dibatasi dengan max_depth 20 satuan untuk mengurangi overkompleksitas model dan meminimalisir overfitting. Parameter min_sample_split 5 satuan digunakan untuk menentukan satuan minimal sampel yang digunakan sebelum suatu node dapat bercabang. Setelahnya class_weight dengan himpunan {0:1. 1:1. 2:1} digunakan untuk memberikan bobot yang sama pada seluruh kelas sentimen.

F. Hasil validasi K-Fold Cross Validation

K-fold Cross Validation dilakukan untuk evaluasi kestabilan modle pada data latih. Proses validasi dilakukan dengan membagi data latih menjadi 5 bagian fold dari 9.512 data latih komentar menggunakan Stratified K-fold dengan tujuan menjaga proporsi setiap kelas sentimen. Hasil akurasi pada setiap fold adalah seperti pada tabel 3

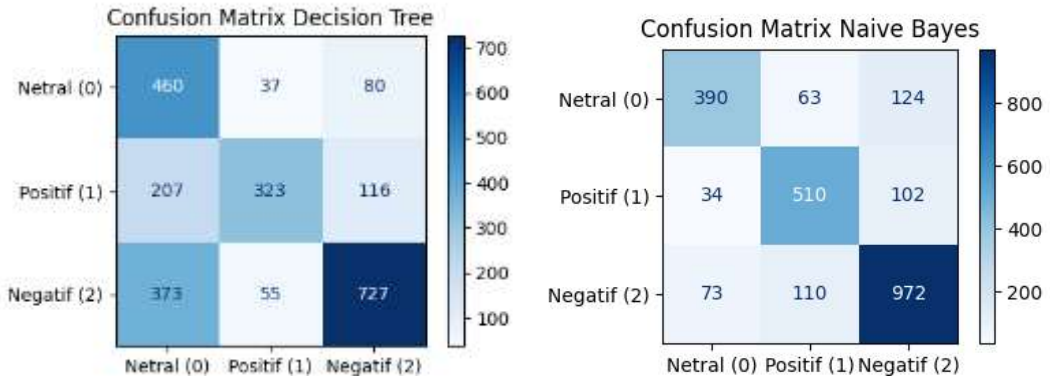
Tabel 3 Akurasi Setiap fold pada K-fold cross validation

Algoritma	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Naïve Bayes	0.80	0.78	0.78	0.77	0,78
Decision Tree	0,64	0,64	0.63	0.64	0,67

Maka bisa dari tabel 3 didapat mean accuracy dari masing-masing algoritma Naïve Bayes dan Decision Tree Berturut-turut adalah 0,78 dan 0,63

G. Hasil Pengujian Model pada Data Tes

Pada tahap ini model diuji dengan data uji yang sudah dipisah sebelumnya. Data uji berjumlah 2378 yang merupakan 20% dari total keseluruhan set data. Hasil dari klasifikasi model setiap algoritma kepada data uji disajikan pada gambar 2.



Gambar 2 Confusion Matrix (a) Decision Tree (b) Naïve Bayes

Maka dapat dihitung Accuracy, Precision, Recall, dan F1-ScoreI dari masing-masing model. Nilai-nilai tersebut adalah sebagai berikut

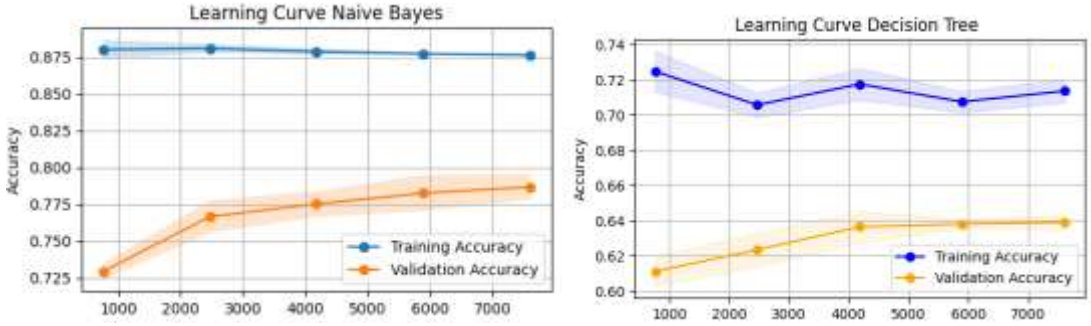
Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Tabel 4 Hasil Evaluasi Model terhadap data uji

Algortima	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Naïve Bayes	0.79	0.79	0.79	0.79
Decision Tree	0,63	0,70	0.63	0.64

H. Pemeriksaan Overfitting

Pemeriksaan overfitting dilakukan dengan membandingkan akurasi model pada data latih, akurasi rata-rata hasil crossvalidation dan akurasi pada data uji. Kedua algortima menunjukkan nilai training accuracy yang lebih tinggi dibandingkan testing acccuracy. Terdapat selisih performa antara data latih dan uji sebesar 0,0875 untuk Naïve Bayes, sedangkan Decision Tree memiliki selisih sebesar 0,0776 Selain itu, Learning Curve menunjukkan akurasi latih Naïve Bayes berada pada skala 0,87 sampai dengan 0,88. Kurva validasi meningkat secara bertahap dari 0,73 menuju 0,79 seiring penambahan data. Disisi lain Decision Tree mendapatkan kurva akurasi latih sebesar 0,71 sampai dengan 0,72, dan kurva akurasi validasi pada skala 0,61 sampai dengan 0,64. Kestabilan model divisualisasikan melalui learning curve seperti pada Gambar 3.



Gambar 3 Learning Curves (a) Naïve Bayes (b) Decision Tree

I. Pembahasan

Berdasar Hasil pengujian, algoritma Naïve Bayes mampu menunjukkan performa yang lebih baik daripada Decision Tree pada tahapan evaluasi. Hal ini dapat dilihat dari nilai mean accuracy hasil K-Fold Cross Validation sebesar 0,7869 lebih tinggi daripada Decision Tree yang memeroleh nilai sebesar 0,6493. Hal ini juga terjadi saat pengujian menggunakan data uji, Naïve Bayes memperoleh akurasi sebesar 0,79 sedangkan Decision Tree memperoleh nilai sebesar 0,64. Hasl tersebut menunjukkan bahwa Naïve Bayes lebih akurat dalam melakukan klasifikasi sentimen berbasis data teks dibandingkan Decision Tree pada dataset komentar dampak bencana Hidrometeorologi.

Berdasar confusion matrix, Naïve Bayes mengalami kesalahan klasifikasi pada sentimen terutama kelas netral. 124 data diprediksi sebagai negatif dan 63 data diprediksi sebagai positif. Sehingga menghasilkan nilai recall netral sebesar 0,68. Nilai recall ini lebih rendah dibanding kelas lainnya. Sedangkan recall tertinggi terdapat pada data kelas negatif sebesar 0,84. Besar recall berbanding lurus dengan jumlah data yang tersedia. Semakin banyak data semakin baik juga recall nya.

Sedangkan pada Decision Tree, pola kesalahan klasifikasi berbeda daripada Naïve Bayes. Model Decision Tree memperoleh recall tertinggi pada kelas netral. Decision tree juga memiliki kelemahan dalam mengenali kelas positif. Decision tree hanya mampu mengenali 50% data kelas positif. Perbedaan ini dipicu karena Decision Tree memiliki proses yang berbeda dari naïve bayes, Naïve bayes mengambil semua probabilitas semua fitur secara bersamaan sedangkan Decision Tree hanya memilih sebagian fitur dalam pembentukan percabangan.

Pada pemeriksaan overfitting dengan training accuracy, mean accuracy, serta hasil K-Fold Cross validation serta didukung oleh learning curve, Naïve Bayes memperlihatkan overfitting ringan dengan indikasi gap sebesar 0,0875. Sementara Decision Tree memperoleh gap sebesar 0,0776. Meski gap berdekatan, pola learning curve kada keduanya berbeda. Naïve Bayes, kurva training accuracy berada pada rentang 0,87 hingga 0,88 dan cenderung stabil, sedangkan kurva validation accuracy terus meningkat dari 0,73 menuju 0,79 seiring bertambahnya data latih, sehingga kedua kurva semakin mendekat satu sama lain sehingga menggambarkan tidak terdapat overfitting berat. Pola ini menunjukkan bahwa model akan membaik seiring bertambah nya data. Sedangkan pada Decision Tree kurva akurasi latih hanya berada pada rentang 0,71 dan 0,72 dan kurva akurasi validasi berada pada rentang 0,61 sampai 0,64 dengan jarak diantara 2 kurva relatif konstan sepanjang penambahan data. Rendahnya nilai-nilai tersebut menunjukkan bahwa Decision Tree belum ammpu mempelajari pola dengan baik.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

IV. SIMPULAN

Hasil pengujian menunjukkan bahwa algoritma *Naïve Bayes* secara konsisten menunjukkan performa yang lebih baik daripada *Decision Tree*. Pada prose *K-Fold Cross Validation*, Naïve Bayes mampu memperoleh *Mean Accuracy* sebesar 0,79 dibandingkan *Decision Tree* yang hanya memperoleh 0,65. Pada evaluasi data uji, Algoritma Naïve Bayes juga lebih unggul dalam *accuracy, precision, recall,* dan *F1-Score* yang seragam sebesar 0,79. Disisi lain *Decision Tree* memperoleh *accuracy, precision, recall,* dan *F1-Score* berturut turut sebesar 0,63, 0,70, 0,63 dan 0,64.

Analisis *confusion matrix* menunjukkan bahwa *Naïve Bayes* mengalami kesalahan klasifikasi terbanyak pada kelas netral dengan recall sebesar 0,68. Sedangkan performa terbaik pada kelas terbanyak yaitu kelas negatif. Hal ini menunjukkan kemampuan Naïve Bayes berbanding lurus dengan banyak nya data. Sedangkan *confusion matrix* milik *Decision tree* menunjukkan kekuatan pada kelas netral dengan recall tertinggi 0,80 dan terendah pada kelas positif hanya 0,50. Perbedaan pola ini menunjukkan Naïve Bayes unggul dalam mengklasifikasikan data daripada *Decision Tree.*

Pada Pemeriksaan overfitting melalui train accuracy, cross validation accuracy dan test accuracy memperkuat bahwa naïve bayes lebih unggul daripada *Decision Tree.* Naïve Bayes menunjukkan train-test gap sebesar 0,0875 dengan kurva learning curve yang saling mendekat seiring bertambahnya data. Menunjukkan overfitting ringan seiring bertambahnya data latih sehingga mengindikasikan kemampuan generalisasi yang tetap baik. Disisi lain Decision Tree memperoleh train-test gap sebesar 0,0776, namun dengan train accuracy yang rendah yakni 0,71. menunjukkan bahwa model belum mempelajari pola data secara maksimal bahkan pada data latihnya sendiri.

Penelitian ini memiliki berberapa keterbatasan yang perlu diperhatikan. Diantaranya pengumpulan data yang dipengaruhi oleh rate limit pada masing-msing platform media sosial sehingga jumlah komentar yang diperoleh dari Tiktok, Instagram, dan Facebook tidak sepenuhnya seimbang. Distribusi data pada setiap kelas sentimen cukup signifikan yang berpotensi memengaruhi kemampuan model sehingga cenderung mengklasifikasikan kepada data mayoritas. TF-IDF juga membuat keterbatasan sehingga model hanya terbatas pada frekuensi dan belum mampu menangkap konteks kalimat seperti sentimen berbentuk sarkasme. Keterbatasan tersebut terlihat ketika adanya kesalahan klasifikasi pada komentar yang mempunyai kemiripan kosa kata antar kelas sentimen.

Berdasarkan Keterbatasan tersebut, penelitian senaljutnya disarankan menggunakan dataset yang lebih besar dan seimbang. Lebih lagi dapat menerapkan suatu metode respresesntasi teks yang mampu memahami konteks semantik seperti *word embedding*. Selain itu, Peneliti selanjutnya dapat melakukan explorasi penggunaan algoritma lain atau melakukan perubahan parameter sehingga optimal dalam klasifikasi dokumen.

V. UCAPAN TERIMA KASIH

Alhamdullillah saya telah dapat menyelesaikan jurnal ini.Saya ucapkan terima kasih atas rekan-rekan yang membantu saya. Terima kasih juga saya ucapkan kepada ibu Yunianita yang telah membantu pembentukan program saya pada validasi dan metode pengumpulan dataset. Saya juga ucapkan terima kasih kepada orang-orang yang kemungkinan tidak saya ingat dan sebutkan Namanya.

VI. REFERENSI

[1] A. Khairani and A. Guspa, “Pengaruh Self Control terhadap Toxic Online Disinhibition Effect pada Generasi Z Pengguna Aplikasi X,” vol. 3, no. 1, pp. 243–253, 2025.

[2] D. Erokhin and N. Komendantova, “International Journal of Disaster Risk Reduction Social media data for disaster risk management and research,” *Int. J. Disaster Risk Reduct.*, vol. 114, no. June, p. 104980, 2024, doi: 10.1016/j.ijdrr.2024.104980.

[3] M. Karimiziarani and H. Moradkhani, “International Journal of Disaster Risk Reduction Social response and Disaster management : Insights from twitter data Assimilation on Hurricane Ian,” *Int. J. Disaster Risk Reduct.*, vol. 95, no. July, p. 103865, 2023, doi: 10.1016/j.ijdrr.2023.103865.

[4] S. M. Anugerah, R. Wijaya, and M. A. Bijaksana, “Sentimen Analysis Social Media for Disaster using Naïve Bayes and IndoBERT,” vol. 11, no. 1, pp. 51–58, 2024.

[5] V. R. Prasetyo, G. Erlangga, D. A. Prima, U. Surabaya, and P. Korespondensi, “Analisis sentimen untuk identifikasi bantuan korban bencana alam berdasarkan data di Twitter menggunakan metode k-means dan Naïve Bayes,” vol. 10, no. 5, pp. 1055–1062, 2023, doi: 10.25126/jtiik.2023107077.

[6] I. F. Rozi, A. T. Firdausi, and K. Islamiyah, “Analisis sentimen pada Twitter mengenai pascabencana menggunakan metode Naïve Bayes dengan fitur n-gram,” pp. 33–39.

[7] M. Das, S. Kamalanathan, and P. Alphonse, “A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset,” vol. 5571, pp. 0–2, 2021.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

[8] P. Yuniar, “Analisis Sentimen Ulasan pada Gojek Menggunakan Metode Naïve Bayes,” vol. 23, no. 2, pp. 164–175, 2023.
[9] N. Q. Rizkina and F. N. Hasan, “Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes,” vol. 4, no. 4, pp. 1136–1144, 2023, doi: 10.47065/josh.v4i4.3803.
[10] A. Gerliandeva, Y. H. Chrisnanto, and H. Ashaury, “Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naïve Bayes dan Ekstraksi Fitur TF-IDF serta N-grams Optimization of Sentiment Classification on Online Comments using Multinomial Naïve Bayes and TF-IDF Feature Extraction and N-grams,” no. X, pp. 259–272, 2024, doi: 10.56873/jpkm.v9i2.5585.
[11] J. Khatib and S. Dalam, “Indonesian Journal of Computer Science,” vol. 13, no. 1, pp. 1051–1063, 2024.
[12] J. Ilmiah, T. Informasi, and U. Sriwijaya, “Analisis Sentimen Pengguna Terhadap Faktor Penggunaan PayPal Menggunakan Metode Decision Tree,” vol. 18, no. 1, pp. 71–83, 2024.
[13] S. Pawar, K. Pandith, and M. G. Rao, “Text Document Categorization Using Random Forest and c4 . 5 Decision Tree Classifier,” vol. 7, no. 2, pp. 211–220, 2023.
[14] F. M. Akbar, R. Hermansyah, S. Lusa, D. I. Sensuse, and N. Safitri, “Analisis sentimen untuk evaluasi reputasi merek motor XYZ berkaitan dengan isu rangka motor di Twitter menggunakan machine learning,” vol. 11, no. 3, pp. 647–654, 2024, doi: 10.25126/jtiik.938663.
[15] G. A. Trianto *et al.*, “KLASIFIKASI OPINI TERHADAP RESESI INDONESIA 2023 PADA TWITTER MENGGUNAKAN ALGORITMA DECESION,” pp. 1–9, 2023.

Conflict of Interest Statement:
The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..