

Analisis Sentimen Opini Masyarakat Terhadap Stunting Pasca Debat Cawapres Pertama 2024 Dengan Algoritma *Bootstrap Aggregating Naïve Bayes*

Analysis of Public Opinion Sentiment on Stunting After the First Vice Presidential Debate in 2024 with Naïve Bayes Bootstrap Aggregating Algorithm

Fahrul Raditiar Yuliardi¹, Fatkhurokhman Fauzi², Tiani Wahyu Utami³

^{1,2,3}Universitas Muhammadiyah, Semarang

Corresponding author : fahrulry22@gmail.com

Abstrak

Stunting adalah kondisi gagal tumbuh pada anak-anak yang disebabkan oleh kekurangan gizi, kurangnya pengetahuan mengenai stunting, serta faktor ekonomi dan akses terhadap makanan bergizi. Debat calon wakil presiden pertama tahun 2024 menarik perhatian publik dan menjadi topik yang banyak dibicarakan di media sosial, termasuk X. Penelitian ini menggunakan metode Bootstrap Aggregating (Bagging) pada algoritma Naïve Bayes untuk menganalisis sentimen opini masyarakat yang diekspresikan melalui tweet terkait stunting. Metode ini dipilih karena kemampuannya dalam meningkatkan akurasi klasifikasi dengan mengatasi masalah ketidakseimbangan kelas dalam data sentimen. Hasil confusion matrix menunjukkan bahwa metode yang digunakan mampu mengklasifikasikan sentimen dengan akurasi sebesar 96%, precision 95%, recall 98%, dan F1-score 96%. Metode ini menunjukkan bahwa Bootstrap Aggregating Naïve Bayes efektif dalam mengklasifikasikan sentimen masyarakat mengenai stunting, terutama dalam konteks respons terhadap pernyataan dan diskusi yang muncul setelah debat cawapres. Mayoritas sentimen yang dianalisis adalah positif, hal ini menunjukkan dukungan dan harapan masyarakat untuk meningkatkan kesehatan dan kesejahteraan anak-anak. Penelitian ini memberikan wawasan tentang bagaimana debat cawapres 2024 dapat mempengaruhi persepsi masyarakat terhadap kebijakan pemerintah Indonesia.

Kata Kunci : Analisis Sentimen, Bootstrap Aggregating, Naive Bayes, Debat Calon Wakil Presiden, Stunting.

Abstract

Stunting is a condition of failure to grow in children caused by malnutrition, lack of knowledge about stunting, as well as economic factors and access to nutritious food. The debate for the first vice presidential candidate in 2024 has attracted public attention and has become a topic of much discussion on social media, including X. This study uses the Bootstrap Aggregating (Bagging) method on the Naïve Bayes algorithm to analyze the sentiment of public opinion expressed through stunting-related tweets. This method was chosen because of its ability to improve classification accuracy by overcoming the problem of class imbalance in sentiment data. The results of the confusion matrix show that the method used is able to classify sentiment with an accuracy of 96%, precision 95%, recall 98%, and F1-score 96%. This method shows that Bootstrap Aggregating Naïve Bayes is effective in classifying public sentiment regarding stunting, especially in the context of responding to statements and discussions that arise after the vice presidential debate. The majority of the sentiments analyzed were positive, which shows the support and expectations of the community to improve the health and welfare of children. This research provides insight into how the 2024 vice presidential debate can affect public perception of Indonesia government policies.

Keywords : Sentiment Analysis, Bootstrap Aggregating, Naive Bayes, Vice Presidential Candidate Debate, Stunting..

PENDAHULUAN

Pada masa balita, saat pertumbuhan dan perkembangan anak, proses pertumbuhan dan perkembangan sangat dipengaruhi. Salah satunya adalah stunting atau juga dikenal sebagai gangguan pertumbuhan pendek (Birman et al., 2023). Stunting merupakan suatu kondisi yang menyebabkan anak gagal tumbuh. Ada beberapa penyebab utama masalah stunting pada anak. Salah satunya adalah kurangnya gizi dan pengetahuan masyarakat mengenai stunting itu sendiri (Fitriani et al., 2022). Ketidakseimbangan asupan nutrisi selama seribu hari pertama kehidupan adalah penyebab stunting (Khoiriyah & Ismarwati, 2023). Faktor yang mempengaruhi terjadinya stunting antara lain terbatasnya akses terhadap makanan bergizi karena faktor genetik serta faktor ekonomi dan pengetahuan (Sairah et al., 2023). Stunting dapat berdampak pada risiko rendahnya tingkat kecerdasan, peningkatan kerentanan terhadap penyakit, dan risiko penurunan tingkat produktivitas di masa depan (Fitriani et al., 2022).

Dari 154 negara yang memiliki data tentang stunting, Indonesia menempati urutan tertinggi ke-27, dengan 80% balita yang mengalami stunting. Selain itu, Indonesia menempati peringkat ke-5 di Asia tertinggi dalam hal jumlah balita yang mengalami kasus stunting (Pascayantri et al., 2023). Pada tahun 2022, prevalensi stunting di Indonesia melebihi rekomendasi WHO yaitu 20%, yang sebesar 21,6% (Kementerian Kesehatan, 2023). Permasalahan stunting di Indonesia berkaitan dengan ketahanan pangan serta keadaan ekonomi dan sosial masyarakat. Dari sudut pandang ini, variabel yang diduga mempengaruhi stunting di Indonesia antara lain pengeluaran makanan, mencukupi kebutuhan gizi pada ibu hamil, ketidakcukupan pangan, tingkat pendidikan, dan fasilitas sanitasi yang memadai (Priatmadani et al., 2023).

Cara mengatasi stunting memerlukan pendekatan yang komprehensif yang mempertimbangkan berbagai elemen, seperti pendidikan, kesehatan, dan ekonomi (Sairah et al., 2023). Oleh karena itu, upaya penanggulangan stunting juga harus melibatkan berbagai sektor, seperti sektor kesehatan, pendidikan, dan perekonomian, serta memerlukan partisipasi aktif Masyarakat setempat (Pratiwi, 2023). Akan tetapi, kendala yang sering terjadi dalam pelaksanaan masih ada, seperti sumber daya manusia yang terbatas, keterlambatan pencairan dana bantuan operasional kesehatan, dan kurangnya kampanye untuk memperkenalkan program stunting (Yuda et al., 2023). Peran tenaga kesehatan dokter, bidan dan perawat harus berperan sebagai komunikator dan motivator untuk membantu masyarakat melalui sosialisasi dan edukasi untuk meningkatkan pengetahuan tentang stunting (Purnaningsih et al., 2023). Meskipun makan siang gratis dapat membantu mengatasi stunting, keberhasilan program ini bergantung pada beberapa hal. Misalnya, makanan harus diolah dengan baik untuk memenuhi standar nutrisi, data harus dikumpulkan untuk menemukan keluarga yang membutuhkan bantuan, dan ada komunikasi yang efektif antara pemerintah, pemerintah daerah, dan komunitas. Isu makan siang gratis ini menjadi perdebatan sehingga dijadikan untuk topik pada perdebatan cawapres.

Debat Calon Wakil Presiden merupakan debat calon presiden yang diadakan di Indonesia dan sering menjadi perbincangan di masyarakat. Pada tahun 2024, perdebatan ini akan menjadi fokus utama pembahasan berita di media, baik media cetak, elektronik, maupun online. (Firdaus, 2019). Debat cawapres pertama tahun 2024 digelar pada Jumat, 22 Desember 2023 di Jakarta Convention Center (JCC). Diskusi berlangsung selama 150 menit dan dibagi menjadi enam bagian. Debat wakil presiden juga mempengaruhi pemilihan umum dengan mengungkapkan rekam jejak para kandidat dan posisi kebijakan

terhadap isu-isu penting. Setiap kandidat akan mempunyai kesempatan untuk menguraikan visi dan misi (Widayanti & Fridiyanti, 2024).

Dalam konteks debat cawapres pertama, kemampuan calon menjadi faktor penentu dapat mempengaruhi persepsi masyarakat adalah kemampuan para kandidat, seperti menyajikan berargumentasi, menyusun kata dan berkomunikasi. Calon wakil presiden nomer urut dua, mengangkat isu persoalan makan siang gratis untuk mencegah stunting menuju Generasi Indonesia Emas 2045. Ini adalah strategi politik yang dilakukan oleh semua kandidat. Pembahasan mengenai makan siang gratis memiliki berbagai reaksi dan opini dari masyarakat. Secara keseluruhan, evaluasi program makan siang gratis ini penting untuk memastikan bahwa intervensi yang dilakukan benar-benar efektif dalam mencegah stunting dan meningkatkan kualitas hidup anak-anak. Pemerintah biasanya bekerja sama dengan berbagai lembaga, seperti organisasi kesehatan, lembaga pendidikan, dan LSM, untuk melakukan evaluasi yang komprehensif. Jika dirancang dan diimplementasikan dengan baik, program makan siang gratis memiliki potensi untuk menjadi salah satu alat yang efektif dalam mengurangi stunting. Namun, kesuksesan program ini sangat tergantung pada bagaimana program tersebut diintegrasikan dengan inisiatif lain yang menangani berbagai faktor penyebab stunting. Komentar masyarakat beragam setelah masalah rencana makan siang gratis menjadi topik yang sensitif dan ramai diperdebatkan di media sosial, seperti X (Widayanti & Fridiyanti, 2024).

Analisis sentimen adalah proses menemukan teks yang diungkapkan oleh pengguna. Untuk tujuan penelitian ini, yaitu untuk mengklasifikasikan apakah teks di X bermakna negatif, positif atau bahkan netral (Anshor & Safuwani, 2023). Penelitian ini membahas bagaimana caranya melakukan analisis sentimen menggunakan opini masyarakat Indonesia tentang stunting di X. Oleh karena itu, penelitian ini adalah mengklasifikasikan opini masyarakat yang mengandung sentimen positif dan negatif terkait dengan masalah persoalan rencana makan siang gratis untuk mencegah stunting di Indonesia dengan menggunakan jejaring sosial X. X adalah platform media sosial yang memberikan ruang bagi pengguna untuk bertukar informasi, menyampaikan pendapat, dan berkomunikasi dengan saling melalui tweet, retweet, dan komentar serta dapat digunakan sebagai wadah untuk menyampaikan pendapat mengenai isu-isu politik (Rinaldo et al., 2023). Penggunaan media sosial tentunya menjadi layanan dan peluang untuk topik politik dan topik yang sedang dibahas (Prasetyo et al., 2023).

Data mining adalah proses pengolahan dan pengujian data untuk menemukan hubungan, informasi, dan kemungkinan yang relevan dan berharga. Salah satu teknik data mining adalah klasifikasi. Klasifikasi adalah salah satu jenis analisis data yang menghasilkan model yang mendeskripsikan kelas data. Ada berbagai pendekatan untuk klasifikasi data yang dapat memberikan hasil yang akurat dan berkualitas tinggi, seperti *Support Vector Machine (SVM)*, *K-Nearest Neighbors (KNN)*, *Naïve Bayes*, dan *Decision Tree*. *Naïve bayes* adalah algoritma klasifikasi yang menggunakan prinsip probabilitas *Bayes*. Dalam metode ini, setiap fitur atribut diasumsikan terkait secara independen dengan suatu kelas, dan probabilitas kelas dapat dihitung dengan menggabungkan probabilitas setiap fitur (Saputri et al., 2021). Keunggulan algoritma ini adalah cepat, efisien, mudah diimplementasikan, dan memungkinkan pengembangan model yang memiliki akurasi dan daya prediksi tinggi serta dapat menangani data dengan banyak fitur (Khalim et al., 2023). Kelemahan *naïve bayes* terletak pada kelas yang tidak seimbang sehingga menimbulkan masalah atau tantangan. Ini karena mesin learning akan menghasilkan prediksi yang akurat terhadap kelas data latih yang besar, tetapi kurang akurat terhadap kelas data latih yang sedikit (Nursimpati & Saifudin, 2019). Metode

klasifikasi ini menghasilkan pendugaan parameter yang tidak konsisten, di mana perubahan kecil pada data dapat menyebabkan perubahan signifikan pada model. Oleh karena itu, pendekatan *bootstrap*, seperti *Bootstrap Aggregating* (Bagging), diperlukan untuk menciptakan pendugaan parameter yang lebih konsisten. Bagging berguna untuk meningkatkan akurasi prediksi dari berbagai penduga atau algoritma, seperti regresi logistik atau pohon klasifikasi.

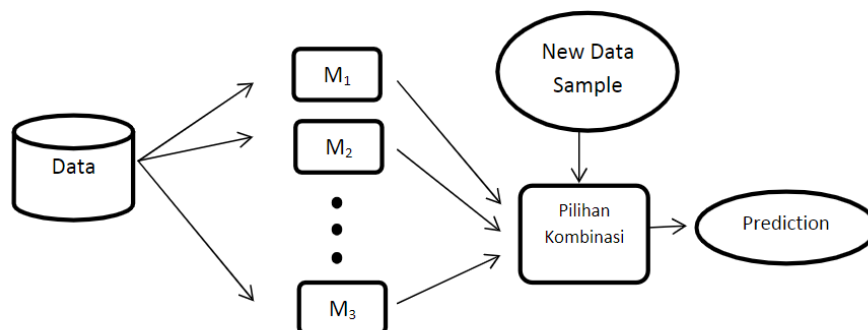
Bootstrap Aggregating (*bagging*) merupakan metode yang menggunakan prinsip metode *ensemble*, yaitu metode yang menggunakan beberapa model secara kombinasi. Pada klasifikasi bagging, suatu model dibentuk dari beberapa Kumpulan data baru yang diperoleh melalui random sampling (*bootstrap*). Hasil akhir klasifikasi ditentukan melalui voting pada hasil klasifikasi masing-masing model. Bagging dirancang untuk meningkatkan stabilitas dan meningkatkan akurasi klasifikasi yang dihasilkan dengan metode tunggal. Bagging pertama kali diperkenalkan oleh Breiman pada tahun 1994 (A. Nugroho & Religia, 2021). Berdasarkan latar belakang di atas, penelitian ini menggunakan algoritma *Bootstrap Aggregating Naive Bayes* untuk mengklasifikasikan opini positif dan negatif. Hasil metode ini dievaluasi menggunakan *Confusion Matrix* untuk menghitung *accuracy*, *precision*, *recall* dan *f1-score*, hal ini akan menjadi sebuah kesimpulan dari hasil evaluasi tersebut.

METODE

Klasifikasi adalah proses menggabungkan data atau objek menjadi kelompok-kelompok sesuai dengan karakteristik yang diperhatikan. Pengambilan keputusan, pengelolaan data, dan pengujian sistem dapat dibantu oleh klasifikasi. Metode klasifikasi biasanya menggunakan berbagai algoritma, seperti *k-nearest neighbors* (KNN), *support vector machine* (SVM), *Naive Bayes*, dan *decision tree*. Setiap algoritma memiliki kelebihan dan kekurangan, jadi penting untuk memilih algoritma yang tepat untuk mencapai akurasi yang tinggi.

A. *Bootstrap Aggregating* (*Bagging*)

Metode kombinasi model adalah teknik yang melibatkan penggabungan beberapa model, seperti $M_1, M_2, \dots, M_k$, untuk membentuk model gabungan yang lebih baik, yang disebut M^* . Metode ini bertujuan untuk meningkatkan performa model gabungan. Secara umum, terdapat dua jenis metode kombinasi model, yaitu *Bagging* dan *Boosting*. Keduanya dapat digunakan untuk keperluan klasifikasi dan prediksi (Han & Kamber, 2006). Ilustrasi dari metode gabungan ini dapat dilihat pada gambar 1.1.



Gambar 1.1 Langkah *Bagging*

Gambar 1.1 menunjukkan proses kerja dari metode kombinasi model, yang bertujuan untuk meningkatkan akurasi model. Metode ini mencakup dua teknik utama, yaitu *Bagging* dan *Boosting*, yang menghasilkan model klasifikasi atau prediksi seperti

$M_1, M_2, \dots, M_k$. Strategi yang digunakan untuk menggabungkan prediksi dari model-model ini adalah dengan menggunakan strategi *voting* untuk menentukan hasil akhir pada objek yang belum diketahui kategorinya. Metode *Bagging* dikembangkan oleh Breiman (1996) dan didasarkan pada konsep *bootstrap* dan *aggregating*, yang menggabungkan manfaat dari kedua teori tersebut. *Bootstrap* sendiri diterapkan berdasarkan teori *sampling* acak dengan pengembalian (Tibshirani dan Efron, 1993). Oleh karena itu, *bootstrap* $X^b = (X_1^b, X_2^b, \dots, X_n^b)$ dilakukan dengan menerapkan pengambilan sampel acak dengan pengembalian dari data pelatihan $= (X_1, X_2, \dots, X_n)$. Metode ini dianggap efektif untuk meminimalkan kesalahan dalam proses klasifikasi. Sebagai hasilnya, model yang dibentuk menggunakan teknik klasifikasi (*classifier*) kemungkinan besar akan memiliki kinerja yang lebih baik. Proses penggabungan beberapa *classifier* ini disebut *aggregating*.

Terkadang, *classifier* gabungan menghasilkan performa yang lebih baik daripada hanya menggunakan satu *classifier* saja. Hal ini disebabkan oleh penggabungan kelebihan dari beberapa *classifier* pada tahap akhir. Oleh karena itu, *bagging* bermanfaat untuk membantu membangun *classifier* yang lebih baik menggunakan sampel data pelatihan.

Berikut ini adalah langkah-langkah dalam menerapkan teknik *Bagging*:

1. Lakukan replikasi *bootstrap* X^b sebanyak m dari sejumlah n data pelatihan. Ulangi langkah ini untuk $b = 1, 2, \dots, B$, dimana m adalah jumlah data yang diambil dari data pelatihan, n adalah ukuran sampel dari data pelatihan, dan B adalah jumlah replikasi *bootstrap* yang dilakukan.
2. Gunakan *simple majority vote* untuk memilih label yang paling sering muncul dari hasil penilaian sebagai aturan pengambilan keputusan akhir.

B. Naïve Bayes

Naive Bayes adalah metode klasifikasi yang menggunakan teorema *Bayes* untuk menghitung probabilitas kelas dari suatu dokumen berdasarkan kemunculan kata-kata dalam dokumen tersebut. Metode ini sangat sederhana dan cepat karena menganggap bahwa fitur-fitur dokumen tidak bergantung satu sama lain. Namun, asumsi ini tidak selalu benar, dan dalam praktiknya *Naive Bayes* sering bersaing dengan klasifikasi yang lebih canggih. (Chen et al., 2021). Untuk memilih fitur dalam *preprocessing* teks, beberapa metode dan algoritma termasuk *ensemble feature selection*, *Chi-Square*, dan *TF-IDF*. Selain itu, ada juga teknik *preprocessing* teks yang dapat membantu pemilihan fitur, seperti *lemmatization*, *stemming*, dan *removal stopwords*. (Lewis, 1998). Analisis sentimen, klasifikasi dokumen, dan klasifikasi teks lainnya adalah beberapa contoh bagaimana klasifikasi *naive bayes* dapat digunakan. (McCallum & Nigam, 1998). Perhatikan rumus *Naive Bayes* sebagai berikut (Ling et al., 2014) :

$$p(v|a) = \frac{P(v) P(a|v)}{P(a)} \quad (2.4)$$

Keterangan:

v = target label (kelas)

a = kata dengan kelas yang belum diketahui

Probabilitas $P(v|a)$ menggambarkan kemungkinan terjadinya kejadian v jika diketahui bahwa kejadian a telah terjadi. $P(v)$ yang disebut *prior*, adalah probabilitas dasar dari terjadinya nilai target label tertentu tanpa mempertimbangkan nilai dari fitur (variabel prediktor). Sementara itu, $P(a|v)$ yang dikenal sebagai *likelihood*, adalah

probabilitas terjadinya nilai fitur tertentu ketika nilai target labelnya telah diketahui. Terakhir $P(a)$, yang disebut sebagai *evidence*, adalah probabilitas terjadinya sekumpulan nilai fitur (variabel prediktor).

Dalam algoritma *naïve bayes classifier*, setiap dokumen direpresentasikan sebagai kumpulan atribut " $a_1, a_2, \dots, a_n$ " dimana a_1 adalah kata pertama, a_2 adalah kata kedua dan seterusnya. Sementara itu, v yaitu himpunan kategori *tweet*. Pada tahap klasifikasi, algoritma ini akan mencari probabilitas tertinggi untuk setiap kategori yang uji menggunakan *naïve bayes classifier*, yang dilambangkan V_{MAP} dan dirumuskan sesuai dengan persamaan berikut:

$$V_{MAP} = \underset{v_j \in v}{\operatorname{argmax}} \prod_{a=1}^n (P(a_i|v_j)P(v_j)) \quad (2.5)$$

Keterangan:

- a = Kata yang belum terklasifikasi
- v = Kategori kelas dalam data training
- $P(v_j)$ = Probabilitas kelas j
- $(P(a_i|v_j))$ = Probabilitas kata a_i kelas v_j

Metode *Naive Bayes* terdiri dari dua tahap proses klasifikasi teks, yaitu proses pelatihan dan proses klasifikasi. Klasifikasi masalah mendapat kesimpulan dari analisis sentimen yang diuji berupa opini positif atau negatif. Dalam buku pengantar Pengambilan Informasi karya (Manning dkk, 2009), tahapan pelatihannya adalah:

1. Menghitung *prior* setiap kategori $P(v_j)$

$$P(v_j) = \frac{n_j}{n_{doc}} \quad (2.6)$$

Dimana:

- n_j = Jumlah *tweet* dalam data *training* dengan kelas j
- n_{doc} = Jumlah total *tweet* dalam data *training*

2. Menghitung probabilitas (*likelihood*) masing-masing kata (a_i) untuk setiap kategori ($P(a_i|v_j)$).

$$(P(a_i|v_j)) = \frac{n_i}{\text{kosa kata}} \quad (2.7)$$

Dimana:

- n_i = Jumlah kemunculan kata a_i dalam *tweet* yang berkategori v_j
- kosa kata = Banyaknya kata dalam data *training*

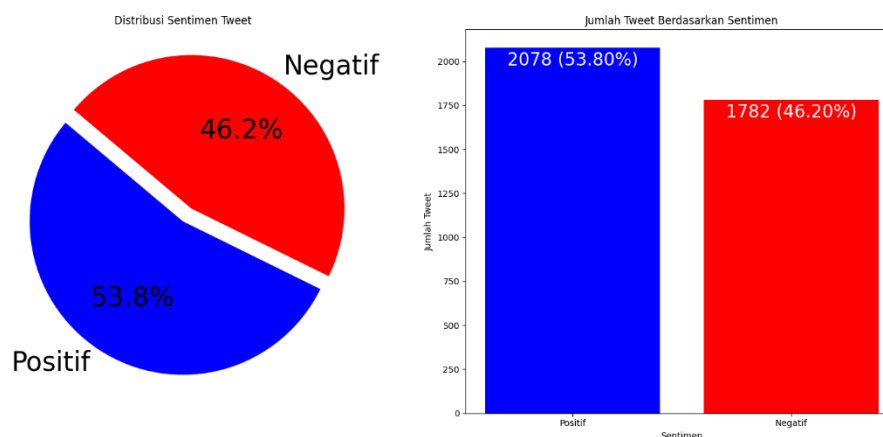
Permasalahan pada pelatihan *likelihood* adalah jika kata yang dicari tidak ditemukan pada dokumen maka nilai hasilnya adalah 0. Untuk mengatasi masalah ini dilakukan pemulusan persamaan sebelumnya (2.7) sehingga persamaan:

$$P(a_i|v_j) = \frac{n_i + 1}{n + \text{kosa kata}}$$

HASIL DAN PEMBAHASAN

A. Distribusi Data

Dalam penelitian ini, setelah proses *scrapping* data dari media sosial X sebesar 5.009 *tweet*. Tetapi setelah *preprocessing* untuk menghilangkan duplikasi dan *tweet* berbahasa asing yang tidak membahas *stunting* jumlahnya berkurang menjadi 3.860 data. Data dibagi menjadi 80% *training* dan 20% *testing*. Data dikategorikan menjadi sentimen positif dan negatif seperti pada gambar dibawah ini:



Gambar 1.2 Distribusi Data *Tweet*

Berdasarkan gambar 1.2 dapat dilihat bahwa mayoritas opini masyarakat mempunyai komentar positif terhadap *stunting*. Hal tersebut terbukti berdasarkan jumlah total 3.860 *tweet*. *Tweet* “Positif” sebanyak 2.078 (53.80%) dan *tweet* “Negatif” sebanyak 1.782 (46.20%). Persentase *tweet* positif lebih tinggi dibandingkan dengan *tweet* negatif. *Tweet* diklasifikasikan sentimen positif apabila mengandung pernyataan positif seperti pujian, dukungan motivasi, atau yang bersifat positif lainnya tentang makan siang gratis untuk mencegah *stunting*. Untuk *tweet* diklasifikasikan sentimen negatif apabila mengandung pernyataan negatif seperti ketidakpuasan, ketidaknyamanan, kritik atau yang bersifat negatif lainnya tentang makan siang gratis untuk mencegah *stunting*.

B. Evaluasi Model

Evaluasi model dilakukan dengan melihat performa metode *bootstrap aggregating naïve bayes* dalam melakukan prediksi terhadap data *testing*. Evaluasi model berupa *accuracy*, *precision*, *recall*, dan *f-1 score*. Dalam mengklasifikasikan sentimen *stunting* diperlukan pertimbangan dalam memilih algoritma dengan kegagalan klasifikasi terkecil. Adapun hasil analisis metode *bootstrap aggregating naïve bayes* dalam menentukan peforma sebagai berikut:

Tabel 1.1 Evaluasi Model *Bagging Naive Bayes*

Replikasi	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F-1 Score</i>
5	92%	91%	95%	93%
10	95%	94%	96%	95%
25	96%	94%	98%	96%
50	96%	95%	98%	96%
100	96%	94%	98%	96%
150	96%	95%	98%	96%

Berdasarkan tabel 1.1 jumlah replikasi dari 5 hingga 150, kinerja model klasifikasi terus membaik di semua metrik evaluasi, termasuk *accuracy*, *precision*, *recall*, dan *f1-score*. Dengan lebih banyak replikasi, model menjadi lebih stabil dan mampu mengidentifikasi kelas positif dengan lebih akurat dan konsisten. Pada replikasi 50 sudah mencapai titik konvergen maka model sudah cukup dengan akurasi mencapai 96% keseimbangan yang baik antara *presisi* dan *recall*, menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membuat prediksi yang benar dan menangkap semua kasus positif. Secara keseluruhan, peningkatan jumlah replikasi berkontribusi positif terhadap kualitas dan keandalan hasil model. Berikut adalah *confusion matrix* dari replikasi konvergen yaitu 50 kali.

Tabel 1.2 *Confusion Matrix Bagging Naive Bayes*

Data Test	Hasil Prediksi	
	Negatif	Positif
Negatif	335	23
Positif	9	405

Dari tabel 1.2 dapat diketahui bahwa data masuk dalam *True Negatif* atau yang diprediksi tepat negatif sebanyak 335 data dan *False Negatif* atau yang diprediksi positif sebenarnya negatif sebanyak 23 data. *True Positif* atau yang diprediksi tepat positif sebanyak 405 data dan *False Positif* atau yang diprediksi negatif sebenarnya positif sebanyak 9 data dengan jumlah 772 data *testing*. Setelah mendapatkan *confussion matrix* selanjutnya adalah menentukan kinerja algoritma klasifikasi sebagai berikut:

$$accuracy = \frac{TN + TP}{TN + FP + FN + TP} = \frac{335 + 405}{335 + 23 + 9 + 405} = \frac{740}{772} = 0,96 = 96\%$$

$$precision = \frac{TP}{TP + FP} = \frac{405}{405 + 23} = \frac{405}{428} = 0,95 = 95\%$$

$$recall = \frac{TP}{TP + FN} = \frac{405}{405 + 9} = \frac{405}{414} = 0,98 = 98\%$$

$$f1 - score = 2 \times \frac{precision \times recall}{precision + recall} = 2 \times \frac{0,95 \times 0,98}{0,95 + 0,98} = 0,96 = 96\%$$

Dari keempat hasil kinerja metode *bootstrap aggregating naïve bayes* tersebut didapatkan bahwa 96% model dapat mengklasifikasikan sentimen positif maupun negatif. Dari keseluruhan sentimen, yang berkategori positif 95% diidentifikasi dengan benar. Untuk *recall* atau pengukuran proporsi kategori positif yang diidentifikasi dengan benar sebanyak 98%.

Tabel 1.3 Performa Metode Bagging Naive Bayes

Replikasi	Accuracy	Precision	Recall	F1-Score
50	96%	95%	98%	96%

Setelah dilakukan analisis metode *bootstrap aggregating naïve bayes*, didapatkan hasil yang terbaik dalam mengklasifikasikan sentimen terkait *stunting*. Metode tersebut berhasil mengklasifikasikan sejumlah 772 data *testing* dengan replikasi terbesar yaitu di replikasi 50 kali mempunyai *accuracy* sebesar 96%, *precesion* sebesar 95%, *recall* sebesar 98% dan *f1-score* sebesar 96%.

KESIMPULAN

1. Analisis sentimen menunjukkan bahwa opini masyarakat mengenai *stunting* pasca debat cawapres pertama di X dapat diklasifikasikan dengan baik. Kebanyakan sentimen adalah positif sebesar 56.80% dan negatif sebesar 46.20%, sentimen positif yang berisi seperti pujian, dukungan motivasi, atau yang bersifat positif lainnya tentang makan siang gratis sebagai langkah

untuk mencegah *stunting*. Dari grafik batang sentimen positif maupun negatif memberikan penegasan kuantitatif tentang seberapa sering kata-kata tertentu muncul dalam data dengan sentimen keduanya, kata "*stunting*" jelas menjadi fokus utama, menunjukkan bahwa diskusi terkait penanganan *stunting* sangat dominan dalam sentimen positif dan negatif.

2. Hasil evaluasi model *bagging naïve bayes* dengan *confusion matrix* menggunakan replikasi yang konvergen sebesar 50 kali menunjukkan *accuracy* sebesar 96%, dengan *precision* sebesar 95%, *recall* sebesar 98%, dan *f1-score* sebesar 96%. Hal ini menunjukkan bahwa metode ini mampu meningkatkan prediksi dalam data dan memberikan hasil yang baik dalam klasifikasi sentimen.

DAFTAR PUSTAKA

- Annisa, Z., & Ulama, B. S. S. (2023). *Analisis Sentimen Data Ulasan Pengguna Aplikasi "Pedulilindungi" Pada Google Play Store Menggunakan Metode Naïve Bayes Classifier Model Multinomial*. *Jurnal Sains Dan Seni Its*.
- Anshor, A. H., & Safuwani, A. (2023). *Analisis Sentimen Opini Warganet X Terhadap Tes Screening Genose Pendeteksi Virus Covid-19 Menggunakan Metode Naïve Bayes Berbasis Particle Swarm Optimization*. *Jurnal Informatika Teknologi Dan Sains*.
- Ardiansyah, R. (2019). *Analisis Sentimen Calon Presiden Dan Wakil Presiden Periode 2019-2024 Pasca Debat Pilpres Di X*. *Scientico: Computer Science And Informatics Journal*.
- Birman, Y., Efriza, & Rosmaini. (2023). *Edukasi Pentingnya Pencegahan Stunting Di Puskesmas Gunung Kota Padang Panjang*. *Jurnal Pengabdian Masyarakat Kesehatan (Jurabdikes)*.
- Chen, H., Hu, S., Hua, R., & Zhao, X. (2021). *Improved Naive Bayes Classification Algorithm For Traffic Risk Management*. *Eurasip Journal On Advances In Signal Processing*, 2021, 1–12.
- Fahrudin, T. M., Sari, A. R. F., Lisanthoni, A., & Lestari, A. A. D. (2022). *Analisis Speech-To-Text Pada Video Mengandung Kata Kasar Dan Ujaran Kebencian Dalam Ceramah Agama Islam Menggunakan Interpretasi Audiens Dan Visualisasi Word Cloud*. *Skanika*.
- Firdaus, R. (2019). *Konstruksi Realitas Sosial Dalam Berita: Analisis Framing Model Murray Edelman Dalam Berita Debat Cawapres 2019 Di Tribunnews.Com Periode 17-19 Maret 2019*.
- Fitriani, Barangkau, Hasan, M., Ruslang, Hardianti, E., Khaeria, Oktavia, R., & Selpiana. (2022). *Cegah Stunting Itu Penting!* *Jurnal Pengabdian Kepada Masyarakat (Jurdikmas) Sosiosaintifik*.
- Fu, R., & Li, X. (2021). *Malicious Attacks On The Web And Crawling Of Information Data By Python Technology*. *Security And Privacy*, 4.
- Gao, X., Yuan, F., & Fan, J. (2022). *Python Based Crawling Program For Collecting Data From Websites*. *Other Conferences*.
- Harahap, P. N., & Sulindawaty, S. (2020). *Implementasi Data Mining Dalam Memprediksi Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus Pt.Arma Anugerah Abadi Cabang Sei Rampah)*. *Matics*, 11(2), 46.
- Irvandi, Irawan, B., & Nurdiawan, O. (2023). *Naive Bayes Dan Wordcloud Untuk Analisis Sentimen Wisata Halal Pulau Lombok*. *Infotech Journal*.

- Kementrian Kesehatan. (2023, January 25). *Prevalensi Stunting Di Indonesia Turun Ke 21,6% Dari 24,4%*.
- Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes. *Jati (Jurnal Mahasiswa Teknik Informatika)*.
- Khoiriyah, H., & Ismarwati, I. (2023). Faktor Kejadian Stunting Pada Balita : Systematic Review. *Jurnal Ilmu Kesehatan Masyarakat*.
- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). *Analisis Sentimen Ulasan Aplikasi Dana Dengan Metode Random Forest* (Vol. 6, Issue 9).
- Lette, A. R., & Jumetan, M. A. (2023). Penyuluhan Tentang Bahaya Asap Rokok Serta Kaitannya Dengan Penyebab Stunting Pada Kaum Bapak Jemaat Overa Fatululat, Amfoang Tengah, Nusa Tenggara Timur. *Jurnal Abdi Masyarakat Indonesia*.
- Lewis, D. D. (1998). Naive (Bayes) At Forty: The Independence Assumption In Information Retrieval. *European Conference On Machine Learning*.
- Ling, J., Putu, I., Kencana, E. N., & Oka, T. B. (2014). *Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square*. 3(3), 92–99.
- Liu, B. (2012). *Sentiment Analysis And Opinion Mining*.
- Lu, Z. (2023). Research On Python Crawling Algorithm In Model Data Visualization. *2023 World Conference On Communication & Computing (Wconf)*, 1–6.
- Mccallum, A., & Nigam, K. (1998). A Comparison Of Event Models For Naive Bayes Text Classification. *Aaai Conference On Artificial Intelligence*.
- Natasuwarna Jurusan Sistem Informasi, A. P., & Pontianak, S. (2020). *Seleksi Fitur Support Vector Machine Pada Analisis Sentimen Keberlanjutan Pembelajaran Daring Support Vector Machine Feature Selection On Online Learning Sustainability Sentiment Analysis* (Vol. 19, Issue 4).
- Nugroho, A., & Religia, Y. (2021). Analisis Optimasi Algoritma Klasifikasi Naive Bayes Menggunakan Genetic Algorithm Dan Bagging. *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 504–510.
- Nugroho, M. R., Sasongko, R. N., & Kristiawan, M. (2021). Faktor-Faktor Yang Mempengaruhi Kejadian Stunting Pada Anak Usia Dini Di Indonesia. *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*.
- Nur Rais, A., Hasan, N., & Faizal Amir, R. (N.D.). *Komparasi Adaboost Dan Bagging Dengan Naïve Bayes Pada Dataset Bank Direct Marketing*. 9(1), 2021.
- Nursimpati, Y. L., & Saifudin, A. (2019). *Penerapan Teknik Bagging Berbasis Naïve Bayes Untuk Seleksi Penerimaan Mahasiswa*. 4(2), 2622–4615.
- Oktavia, D., & Ramadhan, Y. R. (2023). Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial X Menggunakan Algoritma Support Vector Machine (Svm). *Media Online*, 4(1), 407–417.
- Panggalo, S. (2022). Kajian Deskriptif Tentang Stilistika Dan Pragmatik. *Jiip - Jurnal Ilmiah Ilmu Pendidikan*.
- Parlika, R., Pradika, S. I., Hakim, A. M., & Kholilulrachmann, M. (2020). Analisis Sentimen X Terhadap Bitcoin Dan Cryptocurrency Berbasis Python Textblob. *Jurnal Ilmiah Teknologi Informasi Dan Robotika*.
- Pascayantri, A., Armadany, F. I., Jabbar, A., Wahyuni, W., Malik, F., Fitrawan, L. O. M., Munasari, D., Afriana, D., Ringgu, H., Annisaa, N. A., Nurfenti, N., Newulasindo, S. M., & Holidin, W. (2023). Edukasi Pencegahan Stunting Sejak Dini Mulai Dari Ibu Hamil Dan Anak Di Puskesmas Abeli Kota Kendari. *Ejoin : Jurnal Pengabdian Masyarakat*.

- Prasetyo, S. D., Hilabi, S. S., & Nurapriani, F. (2023). *Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes Dan Knn*. *Jurnal Komtekinfo*.
- Pratiwi, D. A. (2023). *Implementasi Kebijakan Penanggulangan Stunting Di Kota Batam*. *Jurnal El-Riyasah*.
- Priatmadani, P., Anjarweni, H. U., Putri, S. M., Pramana, A. S., Palupi, R. L., & Budiasih, B. (2023). *Determinan Prevalensi Balita Stunting Di Indonesia Tahun 2021*. *Seminar Nasional Official Statistics*.
- Purnaningsih, N., Lu'lu' Raniah, D., Sriyanto, D. F., Azzahra, F. F., Pribadi, B. T., Tisania, A., Ayuka, I. R., Cahyani, Z., Komunikasi, D. S., Masyarakat, P., & Manusia, E. (2023). *Upaya Pencegahan Dan Penanggulangan Stunting Di Desa Muncanglarang, Kabupaten Tegal (Stunting Prevention And Countermeasures In Muncanglarang, Tegal)* (Vol. 5, Issue 1).
- Putri, Y. A. N., L.Ginting, G., & Panjaitan, M. W. L. (2023). *Optimasi Penerimaan Siswa Baru Dengan Penerapan Algoritma Text Mining Dan Tf-Idf*. *Journal Of Computing And Informatics Research*.
- Ridwan, A. (2020). *Penerapan Teknik Bagging Pada Algoritma Naive Bayes Dan Algoritma C4.5 Untuk Mengatasi Ketidakseimbangan Kelas*. In *Jurnal Bisnis Digitasl Dan Sistem Informasi* (Vol. 1).
- Rinaldo, R., Sari, A. P., & Fardiana, E. (2023). *Digital Opinion #Puanadalahharapan Di Media Sosial X Menggunakan Social Network Analysis*. *Jurnal Ilmiah Multidisiplin*.
- Sairah, S., Nurcahyani, M., & Chandra, A. (2023a). *Analisis Penyebab Kejadian Stunting Pada Anak Usia Dini*. *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*.
- Saputri, I. I., Purnawansyah, P., & Herman, H. (2021). *Implementasi Metode Naive Bayes Pada Pengenalan Tulisan Tangan Lontara*. *Buletin Sistem Informasi Dan Teknologi Islam*.
- Sartono, B. (2015). *Tinjauan Terhadap Keunggulan Pohon Klasifikasi Ensemble Untuk Memperbaiki Kemampuan Prediksi Pohon Klasifikasi Tunggal* (Vol. 9, Issue 2).
- Septiani, D., & Isabela, I. (2022). *Sintesia: Jurnal Sistem Dan Teknologi Informasi Indonesia Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks*.
- Vegita, Y., Prianto, C., & Pane, S. F. (2023). *Komparasi Model Klasifikasi Naïve Bayes Dan C4.5 Pada Data Prestasi Kerja Pns*. *Jurnal Informatika Upgris*.
- Widayanti, C., & Fridiyanti, Y. N. (2024). *Analisis Pengaruh Debat Calon Presiden 2024 Pertama Terhadap Elektabilitas Calon Presiden Perspektif Pandangan Masyarakat*. *Journal Of Social And Economics Research*.
- Yuda, A. K., Septina, Z., Maharani, A., & Nurdiantami, Y. (2023). *Tinjauan Literatur: Perkembangan Program Penanggulangan Stunting Di Indonesia*. *Jurnal Epidemiologi Kesehatan Indonesia*.