



Analisis Sentimen Polemik Kebijakan Pemerintah Makan Siang Gratis pada Twitter Menggunakan Metode *Neural Network Classification*

Dandy Tri Prasetyo¹, Deni Arifianto², Reni Umilasari^{3*}

^{1, 3} Program Studi Teknik Informatika, Universitas Muhammadiyah Jember, Indonesia

² Program Studi Manajemen Informatika, Universitas Muhammadiyah Jember, Indonesia

Email: wearesdy@gmail.com¹, deniarifianto@unmuhjember.ac.id², reni.umilasi@unmuhjember.ac.id³

*Penulis Korespondensi: reni.umilasi@unmuhjember.ac.id

Abstract. This study aims to investigate public sentiment toward the Indonesian government's free lunch program by analyzing discussions on the X social media platform. A total of 1,500 tweets were collected through a web scraping process using relevant keywords and hashtags. The research workflow consisted of text preprocessing, including data cleaning, case folding, tokenization, stop-word removal, and stemming. The processed text was then transformed into numerical features using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting method, followed by sentiment classification using the Neural Network Classification algorithm. Model performance was evaluated through K-Fold Cross Validation and a Confusion Matrix based on accuracy, precision, recall, and F1-score metrics. To address the issue of class imbalance, Random Over Sampling (ROS) was applied before the classification stage. The experimental results indicate that incorporating ROS improved the classification performance compared with the model trained on the original imbalanced dataset. Furthermore, the Neural Network Classification model effectively categorized public opinions into positive, negative, and neutral sentiments. The findings of this study are expected to provide valuable insights for policymakers in understanding public perceptions of the free lunch program and supporting future policy evaluation.

Keywords: Neural Networks; Random Over Sampling; Sentiment Analysis; Social Media X; TF-IDF.

Abstrak. Penelitian ini dilakukan untuk mengidentifikasi sentimen masyarakat terhadap kebijakan pemerintah mengenai program makan siang gratis melalui unggahan pada media sosial X. Data penelitian terdiri atas 1.500 cuitan yang diperoleh menggunakan teknik *scraping* berdasarkan kata kunci dan tagar yang relevan. Tahapan analisis diawali dengan prapemrosesan teks yang meliputi pembersihan data, *case folding*, *tokenisasi*, penghapusan *stopword*, dan *stemming*. Selanjutnya, fitur teks direpresentasikan menggunakan metode TF-IDF, kemudian dilakukan proses klasifikasi dengan algoritma *Neural Network Classification*. Kinerja model dievaluasi menggunakan K-Fold Cross Validation dan Confusion Matrix berdasarkan metrik akurasi, presisi, recall, dan F1-score. Untuk mengatasi ketidakseimbangan distribusi kelas, diterapkan metode *Random Over Sampling* (ROS) sebelum proses pelatihan model. Hasil penelitian menunjukkan bahwa penerapan ROS mampu meningkatkan performa klasifikasi dibandingkan model tanpa penyeimbangan data. Selain itu, *Neural Network Classification* mampu mengelompokkan sentimen publik ke dalam kategori positif, negatif, dan netral secara efektif. Temuan penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi pemerintah dalam mengevaluasi persepsi masyarakat terhadap pelaksanaan program makan siang gratis.

Kata Kunci: Analisis Sentimen; Media Sosial X; Neural Network; Random Over Sampling; TF-IDF.

1. LATAR BELAKANG

Program makan siang gratis mulai menjadi perhatian nasional sejak awal tahun 2024, sebagai bagian dari janji kampanye presiden terpilih. Program ini bertujuan meningkatkan gizi siswa, menurunkan angka stunting, serta mendukung proses pembelajaran di sekolah. Meskipun masih baru di Indonesia, kebijakan serupa telah berhasil diterapkan di berbagai negara, seperti Jepang dan Korea Selatan. Pemerintah berencana memulai uji coba program ini pada tahun ajaran 2025/2026 dengan prioritas di daerah yang memiliki tingkat kemiskinan dan stunting tinggi (Sitanggang et al., 2024). Namun, implementasi program ini tidak terlepas dari

tantangan, antara lain keterbatasan anggaran, kesiapan infrastruktur, serta pengawasan kualitas makanan.

Kemajuan pesat dalam teknologi informasi dan komunikasi telah memengaruhi metode penyampaian pandangan masyarakat mengenai persoalan sosial dan kebijakan publik. Platform media sosial, terutama X, merupakan salah satu sarana utama yang dimanfaatkan untuk mengekspresikan pandangan umum. Dengan sifatnya yang real-time serta kemudahan dalam berbagi informasi, X menjadi ruang diskusi sekaligus wadah masyarakat dalam memberikan komentar terhadap kebijakan pemerintah, termasuk kontroversi mengenai program makan siang gratis (Sujadi, 2022).

Penggunaan metode *Neural Network Classification* Temuan dari analisis ini diharapkan dapat berfungsi sebagai rekomendasi penting bagi pemerintah dalam menilai serta menyusun strategi kebijakan yang lebih efektif dan terfokus (Muzaki & Witanti, 2021). Oleh karena itu, studi ini tidak hanya memperkaya literatur terkait analisis sentimen, tetapi juga berkontribusi pada proses pengambilan keputusan dalam ranah kebijakan publik.

Lebih jauh, kebijakan makan siang gratis merupakan isu yang sensitif karena menyangkut penggunaan anggaran negara dan kesejahteraan masyarakat luas. Di satu sisi, kebijakan ini dipandang langkah positif untuk meningkatkan gizi dan mengurangi beban ekonomi keluarga. Namun di sisi lain, terdapat keraguan mengenai keberlanjutan program, potensi penyalahgunaan anggaran, serta distribusi yang tidak merata (Merinda Lestandy et al., 2021).

Kompleksitas opini publik inilah yang menjadi alasan pentingnya analisis sentimen dilakukan, agar pemerintah dapat memahami pandangan masyarakat secara komprehensif dan menjadikannya dasar dalam evaluasi kebijakan (Listyarini & Anggoro, 2021).

2. KAJIAN TEORITIS

Analisis Sentimen

Tujuannya adalah mengidentifikasi persepsi masyarakat terhadap suatu topik dengan menganalisis ekspresi yang terkandung dalam teks. Proses analisis sentimen meliputi penentuan dataset, prapemrosesan data, seleksi fitur, pelabelan data, klasifikasi, dan evaluasi hasil (Farid et al., 2022).

Text Preprocessing

Pemrosesan teks merupakan fase kritis dalam penanganan data tekstual, yang dirancang untuk menghapus data yang tidak esensial demi mempersiapkan data bagi algoritma klasifikasi.

Pembobotan Kata *TF-IDF*

Nilai *TF* menunjukkan frekuensi kemunculan kata di data, lalu *IDF* menunjukkan tingkat kelangkaan kata didalam dokumen. Tinggi nilai *TF-IDF*, pentingnya frasa ini terus meningkat sebagai fitur dalam proses analisis sentimen (Qaiser & Ali, 2018).

Neural Network Classification

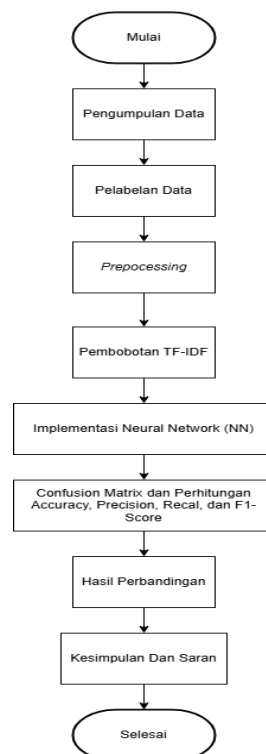
Neural Network Classification Model ini terdiri dari lapisan *input*, *hidden*, dan *output* yang dilatih menggunakan proses *forward propagation* dan *backpropagation* untuk mengoptimalkan bobot. Kemampuannya dalam mempelajari pola yang kompleks menjadikan metode ini efektif untuk klasifikasi teks pada analisis sentimen (Nisa & Candra, 2023).

K-Fold Cross Validation

Teknik ini menilai kinerja model yang mensyaratkan fragmentasi kumpulan data menjadi beberapa bagian yang dikenal sebagai lipatan. satu bagian ditentukan sebagai data uji sementara bagian-bagiannya dimanfaatkan sebagai data latih. Prosedur mengulang secara berurutan hingga seluruh data telah dialokasikan sebagai data validasi. Penerapan teknik ini berkontribusi pada penyediaan perkiraan kinerja model yang lebih tidak bias dan meminimalkan kemungkinan terjadinya *overfitting* (Xu et al., 2020).

3. METODE PENELITIAN

Langkah-langkah yang terlibat dalam fase penelitian ini adalah sebagai berikut:



Gambar 1. Tahapan Penelitian.

Pengumpulan Data

Twitter yang diperoleh melalui teknik *scraping* menggunakan hashtag #MakanSiangGratis dan #MakanBergiziGratis. Proses pengambilan data dilakukan pada periode 4 Juni 2024 hingga 25 Desember 2025 dan menghasilkan sebanyak 1.500 tweet. Selanjutnya data tersebut dilabeli secara manual oleh para ahli bahasa untuk menentukan sentimen yang sesungguhnya.

Preprocessing Data

Sebelum proses klasifikasi, seluruh data melalui tahap *text preprocessing* supaya data menjadi lebih bersih untuk dikelola. Tahapan yang dilakukan meliputi:

- Cleaning*, proses penghapusan URL, sebutan (mention), tagar (hashtag), emoji, angka, tanda baca, serta karakter lain yang tidak esensial.
- Case Folding*, transformasi seluruh abjad menjadi format huruf kecil.
- Tokenizing*, proses pemisahan sebuah kalimat menjadi kata individual.
- Stopword Removal*, eliminasi kata-kata lazim yang tidak berkontribusi signifikan terhadap makna dalam proses klasifikasi.
- Stemming*, proses reduksi setiap kata ke bentuk dasarnya.

Pembobotan TF-IDF

Untuk menentukan relevansi kata di dalam sebuah dokumen, relatif terhadap kumpulan seluruh dokumen yang ada. Prosesnya dimulai dengan menghitung *Term Frequency (TF)*, yaitu seberapa sering suatu kata muncul dalam dokumen spesifik. Selanjutnya, dihitung *Document Frequency (DF)* untuk mengidentifikasi berapa banyak dokumen yang mengandung kata tersebut. Terakhir, dihitung *Inverse Document Frequency (IDF)* menggunakan formula sebagai berikut:

$$IDF = \log\left(\frac{N}{df}\right) = \log(2)$$
$$\log_2 \approx 0.693$$

TF dan IDF selanjutnya dikorelasikan untuk memperoleh nilai bobot agregat TF-IDF:

$$TF-IDF = TF \times IDF$$

Analisis pembobotan mengungkapkan bahwa istilah-istilah umum seperti "makan", "gratis", dan "siang" memperoleh bobot yang rendah dikarenakan frekuensi kemunculannya yang tinggi di keseluruhan korpus dokumen. Sebaliknya, kata-kata yang lebih spesifik seperti "kritik" dan "anggaran" memiliki bobot yang lebih signifikan. Skor *TF-IDF* ini selanjutnya dimanfaatkan sebagai fitur masukan untuk model klasifikasi *Neural Network*.

Random Over Sampling

Sebelum proses pelatihan model dilakukan, penelitian ini menerapkan *Random Over Sampling (ROS)* Teknik ini bekerja dengan menambahkan data pada kelas minoritas hingga jumlah data setiap kelas menjadi lebih proporsional sehingga dapat mengoptimalkan kapasitas model dalam mengidentifikasi keseluruhan kategori sentimen.

Neural Network Classification

Tahap selanjutnya adalah vektorisasi data, yakni sebuah prosedur untuk mentransformasi kutipan menjadi bentuk kuantitatif guna keperluan pemrosesan oleh algoritma machine learning. Dalam studi ini diterapkan untuk menetapkan nilai bobot pada setiap entitas leksika sebuah naskah, sebanding dengan relevansinya terhadap keseluruhan kumpulan data.

Setelah proses *TF-IDF*, setiap dokumen direpresentasikan dalam bentuk vektor bobot kata, kemudian dilakukan perhitungan *weighted sum* menggunakan bobot dan bias pada neuron sebagai tahap forward propagation. Nilai *TF-IDF* dikalikan dengan bobot (0.10, 0.15, dan 0.20) serta ditambahkan bias untuk menghasilkan nilai $zh1$

Hasil perhitungan menunjukkan nilai z *hidden* untuk seluruh dokumen masih berada pada rentang positif, sehingga seluruh dokumen diprediksi ke kelas yang sama. Hal ini terjadi karena bobot yang digunakan masih bersifat awal (*initial weight*) dan belum melalui proses pelatihan, sehingga model belum mampu membedakan pola antar kelas secara optimal. Setelah proses *training*, bobot akan diperbarui sehingga menghasilkan prediksi yang lebih variatif dan akurat.

4. HASIL DAN PEMBAHASAN

Preprocessing Data

Dalam proses ini data dikumpulkan dengan teknik scraping melalui Google Colab dan platform X (Twitter) menggunakan API. Tujuannya proses ini adalah untuk mengumpulkan pendapat publik masalah program kebijakan pemerintah makan siang gratis. Pencarian ini difokuskan pada X dengan tagar #makansianggratis, dengan rentang waktu mulai dari 4 Juni 2024 hingga 25 Desember 2025. Proses scraping dilakukan secara otomatis agar data yang

diperoleh relevan dan sesuai dengan masalah yang ada. dari hasil scarping tersebut, berhasil dikumpulkan sebanyak 1.500 tweet berhasil dikumpulkan dan disimpan dalam format CSV untuk digunakan pada langkah berikutnya, yaitu preprocessing data.

1	conversation	created_at	favorite_coun	full_text	id_str
2	2,00398E+18	Wed Dec 24 23:58:53 +0000 2025	0	MBG anak sehat Indonesia hebat #FaktaIndonesia #MakanBergizi	2,00398E+18
3	2,00398E+18	Wed Dec 24 23:58:23 +0000 2025	0	Pentingnya program MBG untuk masa depan #FaktaIndonesia #M	2,00398E+18
4	2,00398E+18	Wed Dec 24 23:58:09 +0000 2025	0	MBG adalah langkah nyata membangun generasi sehat #MakanBe	2,00398E+18
5	2,00398E+18	Wed Dec 24 23:57:58 +0000 2025	0	MBG buat anak lebih fokus belajar #MakanBergiziGratis #Generas	2,00398E+18
6	2,00398E+18	Wed Dec 24 23:57:49 +0000 2025	0	Dukung Program Makan Bergizi Gratis #MakanBergiziGratis #Gene	2,00398E+18
7	2,00398E+18	Wed Dec 24 23:57:38 +0000 2025	0	Perut terisi belajar pun berisi! #MakanBergiziGratis #GenerasiSel	2,00398E+18
8	2,00398E+18	Wed Dec 24 23:57:29 +0000 2025	0	MBG Pondasi masa depan cerah #MakanBergiziGratis #GenerasiSe	2,00398E+18
9	2,00398E+18	Wed Dec 24 23:57:19 +0000 2025	0	Bersama dukung program MBG #MakanBergiziGratis #GenerasiSe	2,00398E+18
10	2,00398E+18	Wed Dec 24 23:57:05 +0000 2025	0	MBG anak sehat Indonesia hebat #FaktaIndonesia #MakanBergizi	2,00398E+18
11	2,00397E+18	Wed Dec 24 23:24:15 +0000 2025	0	Dukung Bersama Program Makan Bergizi Gratis #ManfaatMakanBe	2,00397E+18
12	2,00397E+18	Wed Dec 24 23:24:08 +0000 2025	0	Dukung Bersama Program Makan Bergizi Gratis #ManfaatMakanBe	2,00397E+18
13	2,00385E+18	Wed Dec 24 15:18:50 +0000 2025	0	Pemprov Kalteng terus mematangkan Program Makan Bergizi Gra	2,00385E+18
14	2,00385E+18	Wed Dec 24 15:18:43 +0000 2025	0	MBG penuh kalori dan protein untuk menyiapkan generasi emas	2,00385E+18
15	2,00384E+18	Wed Dec 24 15:01:24 +0000 2025	0	DPR Usul Alihkan Dana MBG untuk Pemulihan Bencana Sumatera	2,00384E+18
16	2,00378E+18	Wed Dec 24 11:08:14 +0000 2025	1	paket kemasan tahan lama sehingga siswa tetap memperoleh giz	2,00378E+18
17	2,00378E+18	Wed Dec 24 10:38:42 +0000 2025	0	MBG 3B (Bumil Busui Balita non-PAUD) menyasar fase krusial 1.000 HPK investa	2,00378E+18
18	2,00376E+18	Wed Dec 24 09:16:13 +0000 2025	0	Makan Bergizi Gratis adalah langkah strategis pemerintah dalam i	2,00376E+18
19	2,00376E+18	Wed Dec 24 09:16:12 +0000 2025	1	Kebijakan Makan Bergizi Gratis menunjukkan keseriusan pemerir	2,00376E+18
20	2,00376E+18	Wed Dec 24 09:16:11 +0000 2025	1	Program Makan Bergizi Gratis menjadi fondasi penting dalam me	2,00376E+18

Gambar 2. Hasil Scripping Data.

Pelabelan Data

Sebanyak 1500 komentar dari X yang berkaitan dengan MBG telah melalui proses anotasi data secara manual. Validasi terhadap anotasi ini dilakukan oleh Ibu Dr. Eka Nova Ali Vardani, M.Pd., seorang pakar tata bahasa dari Universitas Muhammadiyah Jember, guna menjamin ketepatan dan keseragaman hasil pelabelan.

Tabel 1. Hasil Pelabelan Data.

No	Komentar	Sentimen	Validasi
1	Viral video unboxing isi paket makan siang gratis di sma bogor yang isinya hanya semangka kangkung susu tanpa nasi seusai viral siswa yang merekam video buat klarifikasi berisi permintaan maaf	Netral	✓
2	Semua pihak kemungkinan sudah mempertimbangkan berbagai aspek sebelum mengambil langkah.	Netral	✓
3	Anak-anak jadi lebih fokus dan berenergi.	Positif	✓
4	Siswa dan siswi kalbar memberikan apresiasi positif terhadap pelaksanaan makan bergizi gratis	Positif	✓
7	Menu di indonesia terlalu beragam... perlu standar menu nasional agar tidak memicu tantangan dari variasi rasa dan budaya.	Negatif	✓
8	Topik ini memang cukup luas dan dapat dibahas dari banyak sisi.	Netral	✓
9	Program ini sangat berarti bagi masyarakat kecil.	Positif	✓
10	Semua pihak tentu telah mempertimbangkan berbagai hal sebelumnya.	Netral	✓
11	Makan bergizi gratis makan bergizi gratis	Netral	✓
12	Perluasan makan bergizi gratis menghadirkan gizi terpenuhi perilaku makan anak semakin baik imt lebih seimbang serta dukungan ekonomi lewat ribuan sppg di seluruh indonesia makanbergizigratis indonesiamaju	Positif	✓
1500		Negatif	✓

Preprocessing Text

Proses ini dilakukan dengan scraping menggunakan *script google colab* berikut hasil *preprocessing text*:

Cleaning

Tujuan dari proses pembersihan ini adalah untuk mengeliminasi karakter yang tidak dibutuhkan, seperti tanda baca, angka, alamat web, sebutan tagar, dan ikon emoji.

=== HASIL CLEANING ===

	cleaned_text	cleaned
0	viral video unboxing isi paket makan siang gra...	viral video unboxing isi paket makan siang gra...
1	beraninya sama bocil nih partai	beraninya sama bocil nih partai
2	dokter gizi soroti menu makan siang viral di s...	dokter gizi soroti menu makan siang viral di s...
3	program makan bergizi gratis merupakan solusi ...	program makan bergizi gratis merupakan solusi ...
4	viral video siswa smkn kota bogor yang memperl...	viral video siswa smkn kota bogor yang memperl...
5	wajar salah ini manusia bukan nabi bos	wajar salah ini manusia bukan nabi bos
6	viral di media sosial video siswa smk di kota ...	viral di media sosial video siswa smk di kota ...
7	negara ini butuh anggaran besar untuk ikn dan ...	negara ini butuh anggaran besar untuk ikn dan ...
8	bhabinkamtibmas desa mekarwangi briпка rusiwan...	bhabinkamtibmas desa mekarwangi briпка rusiwan...
9	bhabinkamtibmas desa mekarwangi briпка rusiwan...	bhabinkamtibmas desa mekarwangi briпка rusiwan...

Gambar 3. Hasil Data Cleaning.

Case Folding

Prosedur *case folding* bertugas mengonversi semua huruf kapital suatu teks menjadi huruf kecil, yang bertujuan untuk menyeragamkan format dan memelihara keseragaman data.

=== HASIL CASE FOLDING ===

	cleaned	case_folded
0	viral video unboxing isi paket makan siang gra...	viral video unboxing isi paket makan siang gra...
1	beraninya sama bocil nih partai	beraninya sama bocil nih partai
2	dokter gizi soroti menu makan siang viral di s...	dokter gizi soroti menu makan siang viral di s...
3	program makan bergizi gratis merupakan solusi ...	program makan bergizi gratis merupakan solusi ...
4	viral video siswa smkn kota bogor yang memperl...	viral video siswa smkn kota bogor yang memperl...
5	wajar salah ini manusia bukan nabi bos	wajar salah ini manusia bukan nabi bos
6	viral di media sosial video siswa smk di kota ...	viral di media sosial video siswa smk di kota ...
7	negara ini butuh anggaran besar untuk ikn dan ...	negara ini butuh anggaran besar untuk ikn dan ...
8	bhabinkamtibmas desa mekarwangi briпка rusiwan...	bhabinkamtibmas desa mekarwangi briпка rusiwan...
9	bhabinkamtibmas desa mekarwangi briпка rusiwan...	bhabinkamtibmas desa mekarwangi briпка rusiwan...

Gambar 4. Hasil Case Folding Data.

Tokenizing

Tahapan ini dibuat untuk memecah kalimat atau teks menjadi bagian-bagian yang lebih kecil, yang disebut sebagai token, agar proses analisis menjadi lebih efektif dan tertata.

=== HASIL TOKENIZING ===

	case_folded	tokenized
0	viral video unboxing isi paket makan siang gra...	[viral, video, unboxing, isi, paket, makan, si...
1	beraninya sama bocil nih partai	[beraninya, sama, bocil, nih, partai]
2	dokter gizi soroti menu makan siang viral di s...	[dokter, gizi, soroti, menu, makan, siang, vir...
3	program makan bergizi gratis merupakan solusi ...	[program, makan, bergizi, gratis, merupakan, s...
4	viral video siswa smkn kota bogor yang memperl...	[viral, video, siswa, smkn, kota, bogor, yang, ...
5	wajar salah ini manusia bukan nabi bos	[wajar, salah, ini, manusia, bukan, nabi, bos]
6	viral di media sosial video siswa smk di kota ...	[viral, di, media, sosial, video, siswa, smk, ...
7	negara ini butuh anggaran besar untuk ikn dan ...	[negara, ini, butuh, anggaran, besar, untuk, i...
8	bhabinkamtibmas desa mekarwangi briпка rusiwan...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...
9	bhabinkamtibmas desa mekarwangi briпка rusiwan...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...

Gambar 5. Hasil Tokenizing Data.

Stopword Removal

Langkah berikutnya penghapusan kosakata yang bersifat generik dan tidak menambah nilai signifikan pada evaluasi, termasuk entitas leksikal seperti "dan", "di", dan "yang".

```
... === HASIL STOPWORD REMOVAL ===
```

	tokenized	stopwords_removed
0	[viral, video, unboxing, isi, paket, makan, si...	[viral, video, unboxing, isi, paket, makan, si...
1	[beraninya, sama, bocil, nih, partai]	[beraninya, bocil, nih, partai]
2	[dokter, gizi, soroti, menu, makan, siang, vir...	[dokter, gizi, soroti, menu, makan, siang, vir...
3	[program, makan, bergizi, gratis, merupakan, s...	[program, makan, bergizi, gratis, solusi, penu...
4	[viral, video, siswa, smkn, kota, bogor, yang,...	[viral, video, siswa, smkn, kota, bogor, menu,...
5	[wajar, salah, ini, manusia, bukan, nabi, bos]	[wajar, salah, manusia, nabi, bos]
6	[viral, di, media, sosial, video, siswa, smk, ...]	[viral, media, sosial, video, siswa, smk, kota...
7	[negara, ini, butuh, anggaran, besar, untuk, i...	[negara, butuh, anggaran, ikn, makan, siang, g...
8	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...
9	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...

Gambar 6. Hasil Stopword Removal Data.

Stemming

Tahap akhir, yang dikenal sebagai stemming, bertujuan untuk mengonversi kata-kata turunan ke bentuk dasarnya guna menyelaraskan berbagai variasi leksikal yang mengandung arti serupa.

```
=== HASIL STEMMING ===
```

	stopwords_removed	stemmed
0	[viral, video, unboxing, isi, paket, makan, si...	[viral, video, unboxing, isi, paket, makan, si...
1	[beraninya, bocil, nih, partai]	[berani, bocil, nih, partai]
2	[dokter, gizi, soroti, menu, makan, siang, vir...	[dokter, gizi, sorot, menu, makan, siang, vira...
3	[program, makan, bergizi, gratis, solusi, penu...	[program, makan, gizi, gratis, solusi, turun, ...]
4	[viral, video, siswa, smkn, kota, bogor, menu,...	[viral, video, siswa, smkn, kota, bogor, menu,...
5	[wajar, salah, manusia, nabi, bos]	[wajar, salah, manusia, nabi, bos]
6	[viral, media, sosial, video, siswa, smk, kota...	[viral, media, sosial, video, siswa, smk, kota...
7	[negara, butuh, anggaran, ikn, makan, siang, g...	[negara, butuh, anggar, ikn, makan, siang, gra...
8	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...
9	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...	[bhabinkamtibmas, desa, mekarwangi, briпка, ru...

Gambar 7. Hasil Stemming Data.

Pembobotan TF-IDF

Term Frequency (TF)

Menghitung seberapa sering sebuah kata muncul dalam sebuah dokumen. Peningkatan nilai TF mengindikasikan frekuensi kemunculan kata yang lebih tinggi.

```
=== TF Raw (D1-D5, ..., D1499) ===
```

	a	aaaaaaa	aaahhh	abai	abal
D1	0	0	0	0	0
D2	0	0	0	0	0
D3	0	0	0	0	0
D4	0	0	0	0	0
D5	0	0	0	0	0
...					
	a	aaaaaaa	aaahhh	abai	abal
D1499	0	0	0	0	0

Gambar 8. Hasil TF.

Inverse Document Frequency (IDF)

Kemunculan sebuah kata dalam sebuah dokumen diukur. Semakin Banyak kemunculannya semakin tinggi pula nilai *TF* yang dihasilkan.

```

=== DF & IDF (D1-D5, ..., D1499) ===
      DF      IDF
a         3  2.698680
aaaaaaa   1  3.175802
aaahhh    1  3.175802
abai       1  3.175802
abal       1  3.175802
...
      DF      IDF
zulkifli   1  3.175802

```

Gambar 9. Hasil IDF.

Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF mengonversi teks diubah menjadi representasi numerik yang mencerminkan signifikansi leksikal. Nilai ini, yang dihasilkan dari perkalian TF dan IDF, selanjutnya berfungsi sebagai masukan untuk model klasifikasi.

```

=== TF-IDF Non-zero (D1-D5, ..., D1499) ===
                                TFIDF_nonzero
0  [0.20267, 0.182315, 0.138078, 0.273472, 0.0946...
1  [0.718693, 0.79395, 0.486338, 0.67467]
2  [0.288709, 0.034349, 0.217968, 0.131219, 0.030...
3  [0.042446, 0.036952, 0.041982, 0.045158, 0.193...
4  [0.145669, 0.131039, 0.097035, 0.160859, 0.110...
...
                                TFIDF_nonzero
1498 [0.159843, 0.191651, 0.195764, 0.383303, 0.128...

```

Gambar 10. Hasil TF-IDF.

Evaluasi Model Neural Network Dan Neural Network dan Random Over Sampling

Hasil evaluasi dan perbandingan kinerja model *Neural Network Classification* yang dilatih menggunakan data teknik balancing, yaitu *Neural Network-Random Over Sampling* serta tanpa teknik *balencing*. Penilaian dilakukan dengan menerapkan teknik *K-Fold Cross Validation* yang mencakup nilai K sebesar 2, 3, 4, 5, 6, dan 10.

Tabel 2. Hasil Neural Neural dan Neural Network + ROS.

<i>Label</i>	<i>F1-Score NN Tanpa Balencing</i>	<i>F1-Score NN+ROS</i>	<i>Recall NN Tanpa Balencing</i>	<i>Recall NN+ROS</i>	<i>Precision NN Tanpa Balencing</i>	<i>Precision NN+ROS</i>
Negatif (-1)	81%	84%	82%	83%	89%	88%
Netral (0)	64%	66%	59%	64%	70%	69%
Positif (1)	88%	90%	94%	94%	83%	86%
<i>Accuracy</i>	<i>NN tanpa Balencing</i>			73.40%		
	<i>NN + ROS</i>			73.67%		

Penerapan metode *Random Over Sampling (ROS)* menunjukkan peningkatan performa model dibandingkan tanpa *balencing*. Hal ini terlihat dari kenaikan nilai *F1-score* pada semua kelas, yaitu kelas negatif 81% jadi 84%, netral 64% jadi 66%, dan positif 88% jadi 90%, serta peningkatan *recall* terutama pada kelas netral (59% jadi 64%). *Precision* cenderung stabil dengan sedikit peningkatan pada kelas positif. Selain itu, *accuracy* juga mengalami kenaikan dari 73.40% menjadi 73.67%. Secara keseluruhan, *balencing* dengan *ROS* mampu membuat model lebih seimbang dalam memprediksi setiap kelas dan mengurangi bias terhadap kelas mayoritas.

5. KESIMPULAN DAN SARAN

Penelitian ini mengindikasikan bahwa teknik Klasifikasi Jaringan Saraf Tiruan menunjukkan keunggulan yang memadai dalam mengategorikan sentimen publik, dengan akurasi meningkat dari 73.40% menjadi 73.67% setelah penerapan *Random Over Sampling (ROS)*. Hasil klasifikasi menunjukkan bahwa kelas positif memiliki angka terbagus dari *F1-score* mencapai 90%, sementara kelas negatif dan netral juga mengalami peningkatan setelah *balencing*. Penerapan *ROS* terbukti mampu mengurangi ketidakseimbangan data, meningkatkan kinerja model terutama pada kelas minoritas, serta membuat model lebih seimbang dan tidak bias terhadap kelas mayoritas. Secara keseluruhan, kombinasi *NN+ROS* efektif untuk analisis sentimen pada data yang tidak seimbang.

DAFTAR REFERENSI

- Dwijayanti, A. (2024). Policy Spillover: Analisis Jaringan Dampak Kebijakan Makan Siang Bergizi Gratis terhadap Sektor Pertanian. *Jurnal Administrasi Publik*, 20(2), 281–308. <https://doi.org/10.52316/jap.v20i2.435>
- Farid, M., #1, N., Ferdiana Kusuma, S., Ngagel, J., & Selatan, J. (2022). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen pada Media Sosial Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Berbasis Deep Learning.
- Fauziyyah, A. K. (2020). Analisis Sentimen Pandemi Covid19 Pada Streaming Twitter Dengan Text Mining Python. *Jurnal Ilmiah SINUS*, 18(2), 31. <https://doi.org/10.30646/sinus.v18i2.491>
- Iswantoro, D., & Handayani UN, D. (2022). Klasifikasi Penyakit Tanaman Jagung Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 900. <https://doi.org/10.33087/jiubj.v22i2.2065>
- Kearney, M. (2019). rtweet: Collecting and analyzing Twitter data. *Journal of Open Source Software*, 4(42), 1829. <https://doi.org/10.21105/joss.01829>
- Listyarini, S. N., & Anggoro, D. A. (2021). Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network (CNN). *Jurnal Pendidikan Dan Teknologi Indonesia*, 1(7), 261–268. <https://doi.org/10.52436/1.jpti.60>

- Machine Learning and Computer Science, 4(1), 78–84. <https://doi.org/10.57152/malcom.v4i1.1018>
- Merinda Lestandy, Abdurrahim Abdurrahim, & Lailis Syafa'ah. (2021). Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent *Neural Network* dan Naïve Bayes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4), 802–808. <https://doi.org/10.29207/resti.v5i4.3308>
- Muzaki, A., & Witanti, A. (2021). Sentiment Analysis Of The Community In The Twitter To The 2020 Election In Pandemic Covid-19 By Method Naive Bayes Classifier. *Jurnal Teknik Informatika (Jutif)*, 2(2), 101–107. <https://doi.org/10.20884/1.jutif.2021.2.2.51>
- Nisa, C., & Candra, F. (2023). Klasifikasi Jenis Rempah-Rempah Menggunakan Algoritma Convolutional Neural Network. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 78–84. <https://doi.org/10.57152/malcom.v4i1.1018>
- Noor, I. M., & Turan, M. (2020). Sentiment Analysis on New Currency in Kenya using Twitter Dataset. *IJID (International Journal on Informatics for Development)*, 8(2), 81. <https://doi.org/10.14421/ijid.2019.08206>
- Nooraeni, R., Safiruddin, A. B., Afifah, A. F., Agung, K. D., & Rosyad, N. N. (2020). Analisis Sentimen Publik terhadap Sistem Zonasi Sekolah Menggunakan Data Twitter dengan Metode Naïve Bayes Classification. *Faktor Exacta*, 12(4), 315. <https://doi.org/10.30998/faktorexacta.v12i4.5205>
- Program, T., Siang, M., Tundo, G., & Rachmawati, D. N. (2024). Implementasi Algoritma Naive Bayes untuk Analisis Sentimen. In *Jurnal Indonesia : Manajemen Informatika dan Komunikasi (JIMIK)* (Vol. 5, Issue 3). <https://journal.stmiki.ac.id>
- Qaiser, S., & Ali, R. (2018). Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *International Journal of Computer Applications*, 181(1), 25–29. <https://doi.org/10.5120/ijca2018917395>
- Safuruddin, A., Hermawan, A., & Wibowo, A. P. (2020). Implementation Of BackPropagation Nueral Network In Sentiment Analisys On Twitter To Public Figures. *Compiler*, 9(2), 101. <https://doi.org/10.28989/compiler.v9i2.834>
- Saiful Anwar, M. (2019). Prosiding Konferensi Ilmiah Mahasiswa Unissula (Kimu) 2 Sistem Pencarian E-Journal Menggunakan Metode Stopword Removal Dan Steming Berbasis Android.
- Sitanggang, A., Umaidah, Y., Umaidah, Y., Adam, R. I., & Adam, R. I. (2024). Analysis Sentiment Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naive Bayes. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3). <https://doi.org/10.23960/jitet.v12i3.4902>
- Sujadi, H. (2022). Analisis Sentimen Pengguna Media Sosial Twitter Terhadap Wabah Covid-19 Dengan Metode Naive Bayes Classifier Dan Support Vektor Machine. *Infotech Journal*, 8(1), 22–27. <https://doi.org/10.31949/infotech.v8i1.1883>