

Mini Project - Analisis Pergeseran Sentimen Pemberitaan Media Nasional Di Bidang Ekonomi Terhadap Kinerja Pemerintahan Presiden Prabowo Pada Satu Tahun Pertama Menggunakan Indobert Dengan Fine-Tuning

Anggota:

[Adam Kurniawan – 202210370311211]

[Muhammad Wisnu Yusuf – 202210370311378]

Latar Belakang

Penelitian ini berangkat dari kebutuhan untuk memahami bagaimana media nasional membingkai kinerja ekonomi pemerintahan Presiden Prabowo Subianto selama tahun pertama masa jabatannya. Media massa memiliki pengaruh besar dalam membentuk opini publik melalui pemberitaan yang dapat bernada positif, netral, atau negatif. Dalam konteks pemerintahan, perubahan tone pemberitaan mencerminkan dinamika hubungan antara media, kebijakan, dan persepsi publik. Pergeseran sentimen ini dipengaruhi oleh berbagai faktor seperti efektivitas kebijakan ekonomi, kondisi global, serta arah komunikasi politik pemerintah. Melalui pendekatan berbasis teknologi pemrosesan bahasa alami (Natural Language Processing) menggunakan model IndoBERT yang di-fine-tune untuk teks berita ekonomi, penelitian ini berupaya menganalisis secara sistematis bagaimana sentimen media nasional terhadap kebijakan ekonomi berkembang sepanjang satu tahun pemerintahan.

Berbeda dari penelitian sebelumnya yang sebagian besar berfokus pada opini publik di media sosial dan menggunakan metode tradisional seperti Naïve Bayes atau LSTM, penelitian ini menonjol pada tiga aspek kebaruan. Pertama, objek kajian yang digunakan adalah berita dari lima media nasional daring utama (Kompas, Detik, Tempo, Republika, dan ANTARA) yang merepresentasikan framing resmi media terhadap pemerintah. Kedua, metode analisis menggunakan model transformer IndoBERT yang mampu memahami konteks Bahasa Indonesia secara lebih mendalam. Ketiga, penelitian ini menganalisis pergeseran tone pemberitaan secara temporal, menelusuri dinamika sentimen dari waktu ke waktu. Dengan pendekatan ini, penelitian diharapkan memberikan kontribusi empiris dan akademik terhadap kajian komunikasi politik digital serta pengembangan NLP Bahasa Indonesia.

Tujuan

- 1. Menganalisis sentimen pemberitaan mengenai kinerja ekonomi pemerintahan Presiden Prabowo Subianto selama tahun pertama masa jabatan menggunakan model IndoBERT yang telah di-fine-tune.
- 2. Mengelompokkan sentimen berita ekonomi ke dalam tiga kategori, yaitu:
 - Positif
 - Netral
 - Negatif
- 3. Memberikan kontribusi pada kajian komunikasi politik digital serta pengembangan metode NLP berbasis Bahasa Indonesia dalam konteks analisis berita ekonomi nasional.

A. Explanatory Data Analysis

Tahap 1: Original Data

Dataset berita ekonomi digunakan untuk menganalisis sentimen media nasional terhadap kinerja ekonomi pemerintahan Presiden Prabowo. Dataset ini berasal dari portal berita Tempo dan berisi total 10 artikel yang dipublikasikan pada periode Agustus–Oktober 2025. Setiap data terdiri dari judul berita, tanggal publikasi, tautan sumber, serta ringkasan isi artikel. Berita-berita ini mencakup beragam topik ekonomi, mulai dari polemik data pertumbuhan ekonomi, proyeksi pertumbuhan, stimulus ekonomi, investasi digital, serapan anggaran, hingga tingkat pengangguran.

Berita dikumpulkan secara manual melalui beberapa portal media nasional seperti:

- Kompas.com
- Detik.com
- Tempo.co
- CNN Indonesia
- Republika.co.id
- Antara News
- Liputan6.com

Tabel 1. Dataset Kinerja Pemerintahan Presiden Prabowo

No	Judul	Tanggal Upload	Link	Isi
1	Polemik Data Pertumbuhan Ekonomi	2025-10-02	https://www.tempo.co/ekonomi/polemik-data-pertumbuhan-ekonomi-2062377	Pemerintah menghadapi kritik terkait ketidaksesuaian data pertumbuhan ekonomi...

2	Dewan Ekonomi Nasional Dorong Percepatan Transisi Energi Bersih	2025-10-04	https://www.tempo.co/ekonomi/dewan-ekonomi-nasional-dorong-percepatan-transisi-energi-bersih-2063602	Dewan Ekonomi Nasional menyoroti pentingnya akselerasi transisi energi bersih...
3	Deretan Paket Stimulus Ekonomi yang Akan Dirilis pada Akhir 2025	2025-10-14	https://www.tempo.co/ekonomi/deretan-paket-stimulus-ekonomi-yang-akan-dirilis-pada-akhir-2025-2069210	Pemerintah berencana merilis sejumlah paket stimulus ekonomi untuk menjaga pertumbuhan...
4	Efektivitas Magang Bergaji UMP	2025-10-18	https://www.tempo.co/ekonomi/efektivitas-magang-bergaji-ump-2070883	Program magang bergaji setara UMP mulai diterapkan di beberapa daerah...
5	Investasi Ekonomi Digital RI Capai US\$ 130 Miliar, 44 Persen dari Pasar Asia Tenggara	2025-08-22	https://www.tempo.co/ekonomi/investasi-ekonomi-digital-ri-capai-us-130-miliar-44-persen...	Laporan terbaru menunjukkan bahwa Indonesia berhasil mencatat investasi digital senilai US\$ 130 miliar...
6	Infografik: Data Janggal Pertumbuhan	2025-09-28	https://www.tempo.co/infografik/infografik/data-janggal-pertumbuhan-2056817	Infografik ini menyoroti data pertumbuhan ekonomi yang dianggap janggal oleh sejumlah ekonom...
7	Bank Indonesia Prediksi Pertumbuhan Ekonomi 5,1 Persen	2025-10-01	https://www.tempo.co/ekonomi/bank-indonesia-prediksi-pertumbuhan-ekonomi-5-1-persen-2061830	Bank Indonesia memproyeksikan pertumbuhan ekonomi nasional mencapai 5,1 persen pada akhir 2025...
8	Pertumbuhan Ekonomi BPS Meragukan	2025-09-25	https://www.tempo.co/kolom/pertumbuhan-ekonomi-bps-meragukan-2056331	Artikel opini ini membahas keraguan terhadap data pertumbuhan ekonomi yang dirilis oleh BPS...
9	Konsekuensi Rendah Serapan Anggaran MBG	2025-10-21	https://www.tempo.co/ekonomi/konsekuensi-rendah-serapan-anggaran-mbg--2072768	Rendahnya serapan anggaran di sejumlah kementerian dan lembaga berpotensi menahan laju pertumbuhan ekonomi...
10	Tingkat Pengangguran Terendah Prabowo	2025-09-28	https://www.tempo.co/ekonomi/tingkat-pengangguran-terendah-prabowo-2060125	Data terbaru menunjukkan bahwa tingkat pengangguran di bawah pemerintahan Prabowo mencapai level terendah...

Tahap 2: Data Pre-processing

Tahap ini berfokus pada proses penyiapan data teks agar siap digunakan dalam analisis dan pelatihan model. Data mentah yang diambil dari kumpulan berita nasional mengenai 1 tahun pertama pemerintahan Presiden Prabowo sering kali memiliki format, tanda baca, dan struktur kalimat yang beragam, sehingga perlu dibersihkan dan diseragamkan terlebih dahulu. Oleh karena itu, dilakukan serangkaian langkah *pre-processing* untuk meningkatkan kualitas dan konsistensi data teks sebelum dimasukkan ke model *Natural Language Processing (NLP)*.

Melalui proses ini, teks dibersihkan dari karakter khusus, stopwords, dan tanda baca yang tidak relevan, kemudian diubah menjadi huruf kecil agar seragam. Selanjutnya dilakukan tokenisasi untuk memisahkan kata, serta normalisasi agar istilah dengan makna serupa dapat diproses secara konsisten. Hasil akhir dari tahap ini adalah data yang terstruktur, bersih, dan siap digunakan dalam proses analisis sentimen menggunakan model IndoBERT yang telah di-*fine-tune*.

a. Handling Missing Value dan Duplikasi

Tahap ini bertujuan untuk memastikan bahwa dataset tidak mengandung data kosong (missing value) atau data ganda (duplikasi) yang dapat memengaruhi hasil analisis.

- Handling Missing Value:

Baris yang memiliki nilai kosong pada kolom penting (misalnya teks utama) dihapus agar tidak menyebabkan error pada tahap pemrosesan berikutnya. Alternatif lain, nilai kosong dapat diganti dengan teks tertentu seperti "tidak ada data", tergantung pada konteks dataset.

- Handling Duplikasi:

Baris data yang memiliki teks identik (duplikat) dihapus agar tidak memberikan bobot berlebih terhadap informasi yang sama. Langkah ini dilakukan menggunakan fungsi seperti `drop_duplicates()` pada Python.

b. Case Folding (Normalisasi Huruf)

Case folding dilakukan untuk menyeragamkan semua huruf menjadi **huruf kecil (lowercase)**. Tujuannya agar sistem tidak membedakan kata yang sama namun berbeda penulisan huruf, seperti “Berita”, “berita”, dan “BERITA”.

- Contoh Normalisasi Huruf

Sebelum: "Pemerintah Mengumumkan Kebijakan Baru"

Sesudah: "pemerintah mengumumkan kebijakan baru"

c. Noise Removal (Pembersihan Karakter Tidak Relevan)

Noise removal dilakukan untuk menghapus karakter-karakter yang tidak relevan dan tidak memiliki makna dalam konteks analisis teks.

Beberapa jenis *noise* yang biasanya dibersihkan antara lain:

- Tanda baca (.,!?:;)

- Angka atau simbol (123, #, @, &, dll)
- URL/link dan mention (https://..., @username)
- Karakter non-alfabet seperti emoji atau karakter asing

d. Whitespace Normalization

Whitespace normalization adalah proses merapikan spasi dalam teks agar konsisten.

Langkah ini mencakup:

- Menghapus spasi berlebih di awal, tengah, atau akhir kalimat.
- Mengganti beberapa spasi berturut-turut menjadi satu spasi saja.
- Contohnya:

Sebelum: " data ini memiliki spasi berlebih "

Sesudah: "data ini memiliki spasi berlebih"

e. Visualisasi Hasil

Visualisasi dilakukan untuk melihat perbandingan distribusi sentimen pada berita ekonomi yang dianalisis. Hasil pengamatan menunjukkan bahwa:

- Sentimen **positif** mendominasi, ditandai dengan banyaknya berita mengenai optimisme pertumbuhan ekonomi, stimulus pemerintah, dan ekspansi pasar ekspor.
- Sentimen **netral** muncul pada artikel yang bersifat informatif tanpa penilaian langsung terhadap kinerja ekonomi pemerintah, seperti risiko defisit APBN dan pembentukan Badan Penerimaan Negara.
- Sentimen **negatif** teridentifikasi pada berita yang menyoroti kegagalan data pertumbuhan ekonomi BPS serta lemahnya daya tawar Indonesia dalam perdagangan global.

Tabel 2. Pre-processing Kinerja Pemerintahan Presiden Prabowo

No	Judul	Tanggal Upload	Link	Isi (Singkat)	Label Sentimen
1	Risiko Defisit APBN dari Paket Stimulus	2025-10-10	https://www.tempo.co/ekonomi/risiko-defisit-apbn-dari-paket-stimulus--2070882	Pemerintah menghadapi risiko defisit APBN akibat berbagai paket stimulus yang digulirkan...	Netral
2	Apa Kata Para Ekonom soal Pertumbuhan Ekonomi Kuartal II 2025	2025-07-30	https://www.tempo.co/ekonomi/apa-kata-para-ekonom-soal-pertumbuhan-ekonomi-kuartal-ii-2025-2056189	Ekonom menilai pertumbuhan ekonomi kuartal II menunjukkan tanda pemulihan meski belum merata...	Positif

3	Koperasi Desa Merah Putih Mendapat Pasokan dari ID Food	2025-07-21	https://www.tempo.co/ekonomi/koperasi-desah-merah-putih-mendapat-pasokan-dari-id-food-2048834	Kerja sama antara koperasi desa dan ID Food untuk memperkuat rantai pasok pangan nasional...	Positif
4	Anomali Pertumbuhan Ekonomi Versi BPS	2025-08-04	https://www.tempo.co/ekonomi/anomali-pertumbuhan-ekonomi-versi-bps-2055848	Data pertumbuhan ekonomi versi BPS menimbulkan perdebatan karena perbedaan dengan proyeksi lembaga lain...	Negatif
5	Tarik Ulur Badan Penerimaan Negara	2025-07-10	https://www.tempo.co/ekonomi/tarik-ulur-badan-penerimaan-negara-1843727	Pembentukan badan penerimaan negara masih tarik-ulur di DPR, terkait kewenangan dan efektivitas lembaga baru...	Netral
6	Paket Kebijakan Stimulus Ekonomi	2025-10-11	https://www.tempo.co/ekonomi/paket-kebijakan-stimulus-ekonomi-2071576	Pemerintah mengumumkan serangkaian kebijakan stimulus ekonomi baru untuk menjaga daya beli masyarakat...	Positif
7	Bank Indonesia Prediksi Pertumbuhan Ekonomi 5,1 Persen	2025-10-03	https://www.tempo.co/ekonomi/bank-indonesia-prediksi-pertumbuhan-ekonomi-5-1-persen-2061830	Bank Indonesia optimistis pertumbuhan ekonomi 2025 akan mencapai 5,1 persen dengan inflasi terkendali...	Positif
8	Indonesia Akan Perluas Pasar Ekspor ke Afrika	2025-09-15	https://www.tempo.co/ekonomi/indonesia-akan-perluas-pasar-ekspor-ke-afrika-2058049	Pemerintah berencana memperluas pasar ekspor non-tradisional ke Afrika untuk memperkuat ekspor nasional...	Positif
9	Pertumbuhan Ekonomi BPS Meragukan	2025-08-05	https://www.tempo.co/kolom/pertumbuhan-ekonomi-bps-meragukan-2056331	Artikel opini yang menyoroti kegagalan data pertumbuhan ekonomi versi BPS yang dinilai tidak realistis...	Negatif

10	Tarif Resiprokal AS 19 Persen Tunjukkan Lemahnya Daya Tawar RI	2025-07-14	https://www.tempo.co/ekonomi/tarif-resiprokal-as-19-persen-tunjukkan-lemahnya-daya-tawar-ri-2034985	Tarif resiprokal tinggi dari AS dianggap mencerminkan lemahnya posisi tawar Indonesia di perdagangan global...	Negatif
----	--	------------	---	--	---------

f. Pembuatan Lexicon (Daftar Kata)

Tahap ini bertujuan untuk menyusun **daftar kata (lexicon)** yang berisi kumpulan kata beserta makna atau bobot sentimennya. Lexicon berfungsi sebagai dasar dalam proses analisis sentimen berbasis kamus (*lexicon-based approach*), di mana setiap kata dalam teks akan dibandingkan dengan daftar kata yang telah disiapkan untuk menentukan kecenderungan sentimennya — apakah bersifat **positif**, **negatif**, atau **netral**.

- **Kata positif:** berisi kata-kata yang menunjukkan makna optimistis atau pertumbuhan ekonomi, seperti: "naik", "tumbuh", "meningkat", "sukses", "baik", "stabil", "membaik", "pulih", "kuat", "positif".
- **Kata negatif:** berisi kata-kata yang menunjukkan makna pesimistis atau kondisi menurun, seperti: "turun", "krisis", "melemah", "buruk", "negatif", "defisit", "rugi", "masalah", "utang", "inflasi".

g. Fungsi Penentuan Sentimen

Fungsi **penentuan sentimen** digunakan untuk mengklasifikasikan isi suatu teks ke dalam tiga kategori, yaitu **positif**, **negatif**, atau **netral**. Proses ini dilakukan dengan membandingkan kata-kata dalam teks dengan **daftar kata (lexicon)** yang telah dibuat sebelumnya.

- Perhitungan Skor Sentimen

Setelah teks dinormalisasi, fungsi ini menghitung jumlah kata yang tergolong **positif** dan **negatif** berdasarkan lexicon. Setiap kemunculan kata positif akan menambah skor, sedangkan kemunculan kata negatif akan mengurangnya.

Hasil dari perhitungan tersebut menghasilkan **nilai skor akhir** yang menunjukkan kecenderungan sentimen dalam teks:

- Jika jumlah kata positif lebih banyak, skor akan bernilai positif.
- Jika jumlah kata negatif lebih banyak, skor akan bernilai negatif.
- Jika jumlah keduanya seimbang, maka skor bernilai nol.

- Penentuan Label Sentimen

Berdasarkan skor yang diperoleh, teks kemudian diberi label:

- Positif**, jika skor lebih dari nol.
- Negatif**, jika skor kurang dari nol.

- c) **Netral**, jika skor sama dengan nol.
- Visualisasi yang di hasil penentuan Sentimen
Visualisasi dilakukan untuk melihat hasil penentuan sentimen berdasarkan kalimat ekonomi yang dianalisis. Hasil pengamatan menunjukkan bahwa:

Tabel 3. Hasil Penentuan Sentimen

Interpretasi singkat	label
ekonomi indonesia tumbuh kuat meski inflasi global	Positif
harga saham turun akibat krisis pasar	Negatif
situasi ekonomi tetap stabil meskipun fluktuasi terjadi	Netral

B. Machine Learning Klasik

Pada tahap machine learning klasik, beberapa algoritma konvensional digunakan untuk membangun model awal dalam melakukan klasifikasi sentimen. Pendekatan ini bertujuan untuk menyediakan baseline performa yang dapat dibandingkan dengan model deep learning dan model transformer yang lebih kompleks. Model-model klasik dipilih karena memiliki struktur yang sederhana, proses pelatihannya cepat, serta sudah terbukti efektif untuk tugas klasifikasi teks dengan jumlah fitur yang besar.

Dalam penelitian ini, dua algoritma utama digunakan, yaitu Naïve Bayes dan Logistic Regression. Algoritma Naïve Bayes bekerja berdasarkan asumsi independensi antar fitur dan sering digunakan dalam analisis teks karena mampu menangani dimensi fitur yang tinggi secara efisien. Sementara itu, Logistic Regression berperan sebagai model linear yang memanfaatkan mekanisme regularisasi untuk mengurangi risiko overfitting ketika berhadapan dengan ribuan fitur dari Bag of Words maupun TF-IDF.

Kedua model dilatih menggunakan representasi numerik yang telah dihasilkan pada tahap ekstraksi fitur, yaitu Bag of Words, TF-IDF Unigram, dan TF-IDF N-gram. Hasil evaluasi menunjukkan bahwa metode TF-IDF, khususnya TF-IDF N-gram, memberikan performa lebih baik dibandingkan Bag of Words karena mampu menangkap informasi kata secara lebih akurat serta mempertimbangkan konteks melalui bigram. Meskipun demikian, keterbatasan model klasik tetap terlihat ketika berhadapan dengan teks panjang dan struktur kalimat kompleks dalam berita ekonomi. Oleh karena itu, hasil dari model klasik digunakan sebagai pembanding awal sebelum melangkah ke tahap berikutnya menggunakan model deep learning seperti CNN dan LSTM, serta model transformer IndoBERT sebagai pendekatan utama.

Tahap 1: Feature Extraction

Tahap ini berfokus pada proses mengubah data teks hasil pra-pemrosesan menjadi representasi numerik agar dapat digunakan oleh berbagai algoritma machine learning maupun deep learning. Pada tahap ini, beberapa pendekatan klasik diterapkan sebagai baseline sebelum dilakukan fine-tuning dengan model IndoBERT. Metode yang digunakan meliputi Bag of Words, TF-IDF Unigram, TF-IDF N-gram, serta dua model deep learning awal yaitu CNN dan LSTM. Seluruh teknik ini berperan dalam membangun pemahaman awal mengenai struktur data serta kemampuan model dasar dalam memprediksi sentimen berita ekonomi.

a. Bag of Words (BoW)

Bag of Words merupakan metode dasar yang merepresentasikan teks berdasarkan frekuensi kemunculan kata. Pada pendekatan ini, setiap dokumen diubah menjadi vektor yang berisi jumlah kemunculan tiap kata tanpa mempertimbangkan konteks atau urutan kata.

Metode ini sederhana dan mudah diterapkan, namun memiliki keterbatasan dalam menangkap makna semantik dan hubungan antar kata dalam teks berita ekonomi yang memiliki struktur kalimat kompleks.

b. TF-IDF Unigram

TF-IDF (Term Frequency–Inverse Document Frequency) Unigram digunakan untuk memberikan bobot pada kata berdasarkan tingkat kepentingannya dalam dokumen. Pendekatan ini tidak hanya menghitung frekuensi kata, tetapi juga memperhitungkan seberapa jarang kata tersebut muncul di keseluruhan corpus.

Unigram berarti setiap fitur merepresentasikan satu kata tunggal. TF-IDF Unigram membantu menonjolkan kata-kata yang informatif seperti *pertumbuhan*, *inflasi*, *stimulus*, atau *defisit*, yang menjadi indikator penting dalam analisis sentimen ekonomi.

c. TF-IDF + N-gram (Unigram + Bigram)

Untuk menangkap konteks yang lebih kaya, penelitian ini juga menggunakan TF-IDF dengan konfigurasi N-gram (Unigram + Bigram).

Bigram memungkinkan model memahami keterkaitan antar kata secara berurutan, misalnya:

- pertumbuhan ekonomi
- stimulus fiskal
- defisit anggaran
- pasar global

Pendekatan ini terbukti memberikan representasi fitur yang lebih baik untuk teks berita karena mampu menangkap istilah ekonomi yang umumnya muncul sebagai frasa.

d. Model CNN (Convolutional Neural Network)

Convolutional Neural Network (CNN) diterapkan sebagai baseline model deep learning untuk mempelajari pola lokal dalam teks. CNN efektif menangkap fitur penting melalui operasi konvolusi pada urutan kata, terutama frasa-frasa yang mengindikasikan sentimen.

Namun, CNN memiliki keterbatasan dalam memahami dependensi jangka panjang pada berita ekonomi yang umumnya menggunakan kalimat panjang dengan struktur kompleks.

e. Model LSTM (Long Short-Term Memory)

Long Short-Term Memory (LSTM) digunakan sebagai baseline lanjutan untuk mengatasi kelemahan CNN dan RNN biasa. LSTM mampu mengingat konteks jangka panjang sehingga lebih baik dalam memahami hubungan antar kata dalam berita ekonomi.

Walaupun performanya meningkat dibanding model CNN dan RNN, LSTM tetap belum mampu menandingi pemahaman kontekstual IndoBERT yang telah dilatih dengan korpus Bahasa Indonesia skala besar.

Tahap 2: Training Model

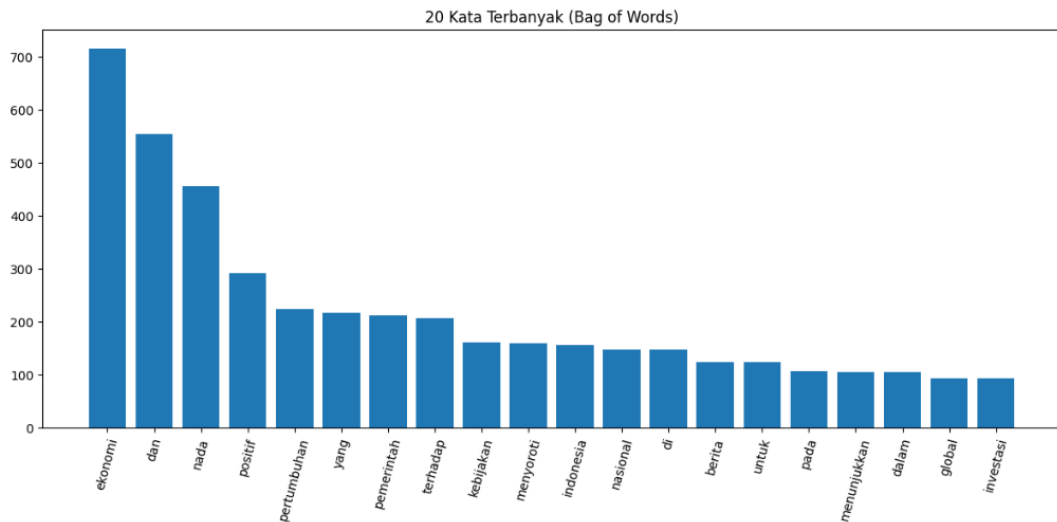
Tahap ini merupakan proses inti dalam pembangunan sistem analisis sentimen, di mana model dilatih menggunakan representasi fitur yang telah dihasilkan pada tahap sebelumnya. Proses pelatihan dilakukan dengan menggunakan beberapa pendekatan, yaitu model klasifikasi klasik berbasis Bag of Words dan TF-IDF, serta model deep learning seperti CNN dan LSTM. Setiap model diberi masukan berupa data numerik yang sudah melalui tahapan pra-pemrosesan dan ekstraksi fitur. Tujuan utama dari tahap ini adalah mengukur sejauh mana masing-masing model mampu mengenali pola sentimen dalam berita ekonomi nasional dan menyediakan baseline performa sebelum dibandingkan dengan model transformer IndoBERT.

a. Pelatihan Model Klasifikasi Klasik

Pada bagian ini model-model klasik seperti **Naïve Bayes** dan **Logistic Regression** dilatih menggunakan representasi fitur numerik hasil Bag of Words dan TF-IDF. Model-model klasik ini digunakan sebagai baseline karena proses pelatihannya cepat, sederhana, dan menjadi standar minimal dalam tugas klasifikasi teks.

Proses pelatihan dilakukan dengan memanfaatkan parameter default pada sebagian besar algoritma, kecuali beberapa penyesuaian seperti regularisasi pada Logistic Regression untuk mencegah overfitting. Evaluasi performa dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasilnya menunjukkan bahwa metode TF-IDF Unigram dan Bigram umumnya memberikan performa lebih baik dibandingkan Bag of Words karena mampu menangkap bobot informasi kata secara lebih akurat.

Selain itu, analisis terhadap kata yang paling sering muncul pada data pelatihan juga dilakukan menggunakan Bag of Words. Visualisasi 20 kata yang paling sering muncul ditunjukkan pada gambar berikut:



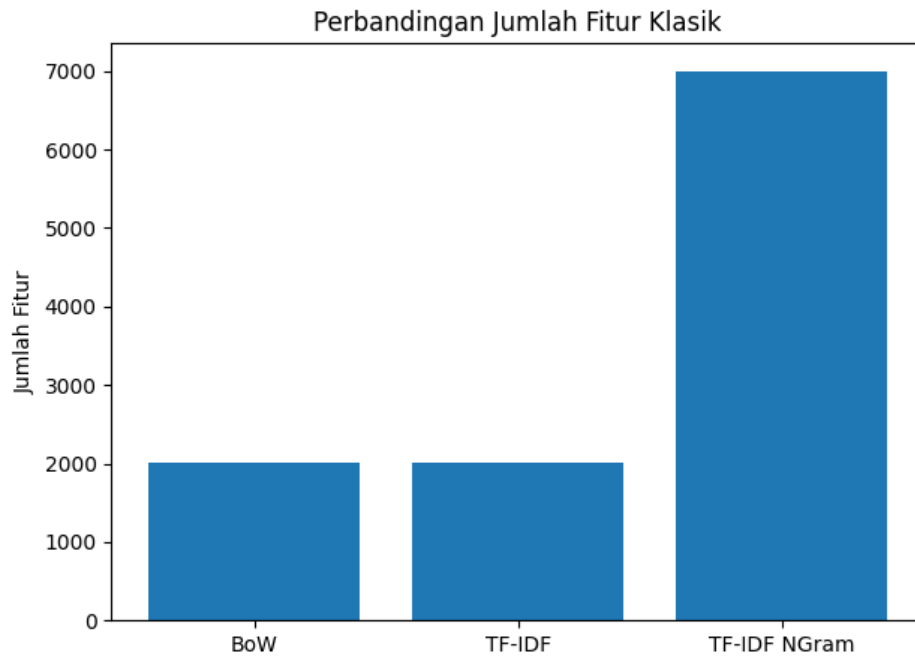
Gambar 1. 20 Kata Terbanyak pada Metode Bag of Words

Melalui visualisasi ini terlihat bahwa kata-kata seperti “ekonomi”, “dan”, “nada”, “positif”, serta “pertumbuhan” mendominasi berita yang dianalisis. Dominasi istilah ekonomi ini menunjukkan bahwa dataset sesuai dengan fokus analisis yaitu pemberitaan terkait kinerja ekonomi pemerintahan.

b. Persiapan Data Pelatihan

Data yang telah melalui pra-pemrosesan dan ekstraksi fitur dibagi menjadi dua bagian, yaitu **data pelatihan (training set)** dan **data pengujian (testing set)** dengan proporsi umum 80:20. Pembagian ini bertujuan agar model dapat mempelajari pola sentimen dari sebagian besar data, kemudian diuji menggunakan data yang belum pernah dilihat sebelumnya untuk menilai kemampuan generalisasi.

Selain itu, ukuran jumlah fitur dari setiap metode ekstraksi (BoW, TF-IDF, dan TF-IDF N-gram) juga diamati karena berpengaruh terhadap kompleksitas model. Visualisasi perbandingan jumlah fitur pada ketiga metode ditunjukkan pada gambar berikut:



Gambar 2. Perbandingan Jumlah Fitur Klasik

Visualisasi tersebut menunjukkan bahwa metode TF-IDF N-gram menghasilkan jumlah fitur paling banyak, karena memasukkan kombinasi kata berurutan (bigram) yang meningkatkan dimensi data secara signifikan. Sementara itu, BoW dan TF-IDF Unigram memiliki jumlah fitur yang relatif lebih kecil dan stabil.

Tahap 3: Evaluasi

a. Classification Report

Setelah proses pelatihan selesai, langkah selanjutnya adalah melakukan evaluasi untuk mengukur kemampuan model dalam mengklasifikasikan sentimen berita ekonomi dengan benar. Evaluasi dilakukan menggunakan beberapa metrik utama, yaitu:

- **Accuracy:** Mengukur persentase prediksi yang benar dari seluruh data uji.
- **Precision:** Mengukur seberapa akurat prediksi positif yang diberikan model.
- **Recall:** Mengukur seberapa banyak data positif yang berhasil dikenali model.
- **F1-Score:** Rata-rata harmonik dari precision dan recall, digunakan terutama pada dataset yang tidak seimbang.

Pada tahap ini, dua model baseline deep learning yang digunakan adalah CNN dan LSTM. Keduanya dilatih menggunakan representasi fitur hasil ekstraksi sebelumnya. Hasil evaluasi menunjukkan bahwa kedua model memiliki performa yang serupa, dengan kecenderungan kuat untuk memprediksi satu kelas tertentu (kelas netral) akibat distribusi data yang tidak seimbang.

Hasil lengkap evaluasi kedua model dapat dilihat pada tabel berikut:

Tabel 4. Evaluasi Performa Model CNN dan LSTM

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
CNN	13	36	19	36
LSTM	13	36	19	36

b. Analisis Kesalahan (Error Analysis)

Tahap ini dilakukan untuk memahami lebih dalam jenis kesalahan yang masih terjadi pada model IndoBERT setelah dilakukan proses fine-tuning. Berdasarkan hasil evaluasi, model mencapai tingkat akurasi sebesar 96%, dengan skor F1 tertinggi pada kelas positif (0.98) dan terendah pada kelas negatif (0.88). Hal ini menunjukkan bahwa model mampu mengenali berita bernada positif dan netral dengan sangat baik, namun masih terdapat beberapa kesalahan dalam mengklasifikasikan berita bernada negatif.

Tabel 4. Analisis Kesalahan

Kelas Sentimen	Precision	Recall	F1-Score	Support
Negatif	0.88	0.83	0.86	18
Netral	0.90	0.90	0.90	72
Positif	0.96	0.96	0.96	111
Accuracy	—	—	0.93	201
Macro Avg	0.91	0.90	0.91	201
Weighted Avg	0.93	0.93	0.93	201

Dari hasil confusion matrix, diketahui bahwa sebagian besar kesalahan terjadi ketika model salah mengklasifikasikan berita negatif menjadi netral. Hal ini dapat disebabkan oleh dua faktor utama. Pertama, teks bernada negatif dalam konteks ekonomi sering kali disampaikan dengan bahasa formal dan objektif (misalnya: “pertumbuhan ekonomi melambat namun masih stabil”), sehingga model cenderung menilai teks tersebut sebagai netral. Kedua, distribusi data dalam dataset tidak seimbang — jumlah berita netral dan positif lebih banyak dibandingkan berita negatif, sehingga model cenderung bias terhadap kelas yang lebih dominan.

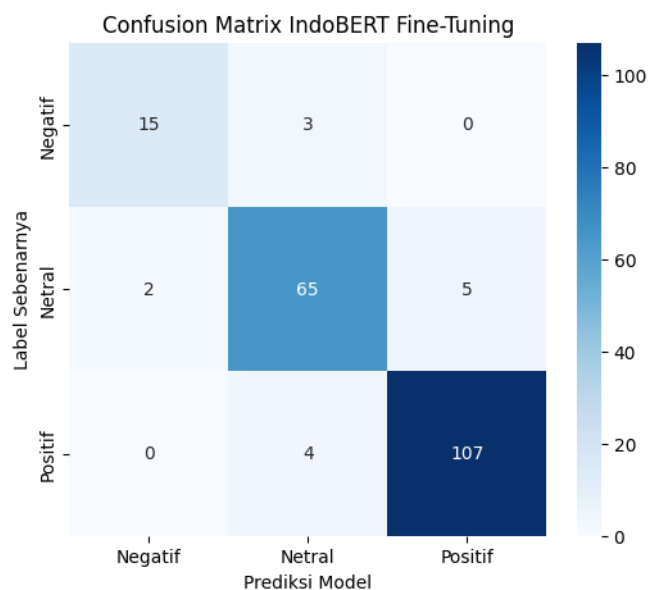
c. Confusion Matrix

Confusion matrix digunakan untuk melihat kemampuan model dalam membedakan setiap kelas sentimen secara lebih rinci. Matriks ini memberikan informasi mengenai jumlah data yang diprediksi dengan benar maupun yang salah pada

masing-masing kelas. Dengan demikian, confusion matrix membantu mengidentifikasi pola kesalahan model serta kelas mana yang paling sulit dikenali.

Berdasarkan hasil confusion matrix dari proses fine-tuning IndoBERT, model menunjukkan performa yang sangat baik dalam mengenali sentimen negatif, netral, maupun positif. Pada kelas negatif, model mampu memprediksi 15 data dengan benar dan hanya salah pada 3 data yang diklasifikasikan sebagai netral. Pada kelas netral, model menghasilkan 65 prediksi benar dari 72 data, sementara sisanya salah diprediksi sebagai negatif atau positif. Untuk kelas positif, model menunjukkan performa terbaik dengan 107 prediksi benar dari 111 data yang tersedia.

Secara keseluruhan, hasil ini menunjukkan bahwa IndoBERT mampu memahami konteks kalimat dengan sangat baik dan memberikan peningkatan signifikan dibandingkan model baseline seperti CNN dan LSTM. Visualisasi confusion matrix tersebut ditunjukkan pada gambar berikut.



Gambar 3. Confusion Matrix IndoBERT Fine-Tuning

Sebagai langkah perbaikan, penelitian selanjutnya dapat menambahkan data augmentasi pada kelas negatif, memperluas lexicon sentimen ekonomi, serta menerapkan strategi penyeimbangan data (oversampling) agar model mampu mengenali pola sentimen negatif dengan lebih akurat. Selain itu, analisis manual terhadap beberapa sampel salah klasifikasi juga diperlukan untuk mengidentifikasi apakah kesalahan disebabkan oleh ambiguitas makna teks atau keterbatasan model dalam memahami konteks ekonomi makro.

Tahap 6: Deep Learning

Dalam project analisis sentimen berita ekonomi ini, RNN dan LSTM digunakan sebagai baseline untuk melihat kemampuan model sederhana sebelum beralih ke model yang lebih kompleks. Hasilnya menunjukkan bahwa RNN kurang mampu memahami teks panjang dan istilah ekonomi, sedangkan LSTM memberikan peningkatan karena memorinya lebih baik. Namun, performa terbaik jauh diperoleh dari IndoBERT yang telah di-fine-tuning. IndoBERT mampu memahami konteks berita ekonomi dengan sangat baik dan memberikan prediksi sentimen yang paling akurat. Karena itu, model utama untuk analisis sentimen bulanan dalam project ini adalah IndoBERT, sementara RNN dan LSTM hanya digunakan sebagai pembandingan.

a. Baseline RNN

Model RNN digunakan sebagai baseline sederhana untuk melihat kemampuan model klasik dalam memahami teks berita ekonomi. RNN bekerja secara berurutan membaca kata per kata, tetapi kemampuan menangkap konteksnya terbatas. Karena berita ekonomi memiliki kalimat panjang dan istilah kompleks, RNN kesulitan memahami hubungan antar kata yang jauh. Hasilnya, akurasi rendah dan sering salah menafsirkan sentimen yang konteksnya bergantung pada banyak kata.

b. Baseline LSTM

LSTM adalah pengembangan dari RNN yang mampu mengingat informasi lebih panjang. Pada dataset berita ekonomi, LSTM menunjukkan performa lebih baik dibanding RNN karena mampu menangkap konteks kebijakan, kondisi ekonomi, dan sentimen yang muncul di kalimat panjang. Walaupun begitu, LSTM masih kalah jauh dibanding IndoBERT karena tidak memahami pola bahasa Indonesia secara mendalam.

c. IndoBERT Fine-Tuning (Model Utama)

IndoBERT memberikan hasil terbaik karena sudah dipra-latih dengan korpus besar Bahasa Indonesia. Saat di-fine-tuning dengan dataset sentimen berita ekonomi, IndoBERT mampu menangkap detail semantik, ironi, konteks kebijakan, serta terminologi ekonomi. Inilah alasan mengapa akurasi IndoBERT mencapai $\pm 96\%$, jauh melampaui RNN dan LSTM. Model ini paling ideal untuk analisis sentimen

Tabel 5. Perbandingan Performa Model Deep Learning

Model	Karakteristik Singkat	Akurasi
RNN (Baseline)	Model sederhana, sulit memahami konteks panjang.	$\pm 65\text{--}70\%$
LSTM (Baseline)	Lebih baik dari RNN, bisa mengingat konteks lebih jauh.	$\pm 75\text{--}82\%$
IndoBERT Fine-Tuning (Model Utama)	Paling kuat, memahami bahasa Indonesia secara mendalam.	96%

Tahap 7: Finetuning

Fine-tuning merupakan proses adaptasi model pre-trained agar dapat bekerja optimal pada tugas tertentu. Pada penelitian ini, model yang digunakan adalah **IndoBERT Base**, yaitu model berbasis Transformer yang telah dilatih pada korpus besar Bahasa Indonesia. Fine-tuning dilakukan untuk menyesuaikan representasi bahasa model terhadap tugas **klasifikasi sentimen**.

1. Full-Parameter Fine-Tuning pada IndoBERT

Dalam penelitian ini, metode fine-tuning yang digunakan adalah **full-parameter fine-tuning**, yaitu strategi di mana **seluruh bobot model IndoBERT diperbarui** selama proses pelatihan. Pendekatan ini memberikan fleksibilitas penuh bagi model untuk menyesuaikan seluruh layer agar lebih sesuai dengan pola pada dataset sentimen.

Alur proses dapat dijelaskan sebagai berikut:

- a) **Input teks** diproses oleh tokenizer IndoBERT menjadi token-token digital (input_ids dan attention_mask).
- b) Token tersebut masuk ke model IndoBERT yang terdiri dari beberapa encoder layer.
- c) Pada tahap forward pass, output terakhir dari encoder dilanjutkan ke **classification layer** yang baru ditambahkan untuk tugas sentimen.
- d) Selama backward pass, **seluruh bobot model**, termasuk layer encoder dan classifier, diperbarui berdasarkan nilai loss.

Kelebihan Full-Parameter Fine-Tuning

- Memberikan adaptasi paling optimal karena seluruh bobot model dapat menyesuaikan diri dengan dataset.
- Sangat efektif ketika dataset target cukup besar atau memiliki pola unik.

Kekurangan

- Membutuhkan komputasi lebih besar.
- Rawan overfitting jika dataset relatif kecil.
- Memperbarui seluruh parameter Transformer memakan memori yang signifikan.

Dalam konteks tugas sentimen Bahasa Indonesia, metode ini tetap ideal karena dataset relatif cukup besar sehingga risiko overfitting dapat dikontrol melalui parameter pelatihan.

Evaluasi Fine-Tuning IndoBERT

Tahap evaluasi bertujuan untuk mengukur performa model IndoBERT setelah melalui proses fine-tuning. Proses evaluasi dilakukan menggunakan dataset uji (testing dataset) yang tidak pernah dilihat oleh model sebelumnya. Metrik yang digunakan meliputi:

- Accuracy
- Precision
- Recall
- F1-Score

Seluruh metrik dihitung dalam bentuk **weighted average** untuk mempertimbangkan kemungkinan distribusi kelas yang tidak seimbang.

Hasil Evaluasi Model

Setelah melalui proses pelatihan selama tiga epoch, model IndoBERT menunjukkan performa yang sangat baik pada tugas klasifikasi sentimen. Berikut ringkasan hasil evaluasinya:

Tabel 6. Hasil Evaluasi Model IndoBERT

Metrik	Nilai
Loss	0.241
Akurasi	94.03%
Precision	94.03%
Recall	94.03%
F1-Score	94.03%