

**CESS**

**(Journal of Computer Engineering, System and Science)**

Available online: <https://jurnal.unimed.ac.id/2012/index.php/cess>

ISSN: 2502-714x (Print) | ISSN: 2502-7131 (Online)



## **Analisis Sentimen Publik di X Terhadap Rencana Kenaikan PPN 12% Menggunakan Bert**

### ***Analysis of Public Sentiment in X Towards The 12% PPN Increase Plan Using Bert***

**Ferdian Imawan<sup>1\*</sup>, Diqy Fakhrun Shiddieq<sup>2</sup>, Fikri Fahu Roji<sup>3</sup>**

<sup>1,2,3</sup>Program Studi Bisnis Digital, Fakultas Ekonomi, Universitas Garut

Jalan Raya Samarang No. 52a, Mekarwangi, Kec. Tarogong Kaler, Kabupaten Garut, Jawa Barat

Email: <sup>1</sup>[24025121015@fekon.uniga.ac.id](mailto:24025121015@fekon.uniga.ac.id), <sup>2</sup>[diqy@uniga.ac.id](mailto:diqy@uniga.ac.id), <sup>3</sup>[fikri@uniga.ac.id](mailto:fikri@uniga.ac.id)

*\*Corresponding Author*

#### **ABSTRAK**

Rencana kenaikan tarif Pajak Pertambahan Nilai (PPN) menjadi 12% telah menjadi salah satu isu kebijakan publik yang sedang diperbincangkan di masyarakat. Kebijakan ini memicu beragam tanggapan di media sosial X, yang mencerminkan adanya pro dan kontra terhadap rencana tersebut. Metode yang digunakan dalam penelitian ini adalah *pre-trained BERT Classification* yang digunakan untuk melakukan analisis sentimen, klasifikasi topik, serta memberikan akurasi tinggi. Tujuan penelitian ini untuk melakukan analisis sentimen dan memahami respon publik terhadap rencana kenaikan PPN 12%. Hasil analisis menunjukkan bahwa dominasi sentimen negatif sebesar 48,58%, yang mencerminkan kekhawatiran masyarakat terhadap potensi dampak kebijakan, diikuti oleh sentimen netral sebesar 42,39%, yang berfokus pada stabilitas ekonomi dan efek kebijakan secara umum. Sementara itu, sentimen positif sebesar 9,03%, merefleksikan optimisme terhadap manfaat kebijakan jangka panjang. Model BERT yang digunakan berhasil mencapai akurasi 83%, dengan nilai *precision*, *recall*, dan *F1-score* rata-rata sebesar 83%, 82%, dan 82%. Selain itu, visualisasi *word cloud* mendukung hasil analisis dengan menampilkan kata-kata dominan seperti "harga," "rakyat," dan "beban" pada sentimen negatif, serta "pajak" dan "daya beli" pada sentimen netral. Penelitian ini berkontribusi dalam menyediakan wawasan berbasis data untuk mendukung pemerintah dalam menyusun kebijakan mitigasi guna meminimalkan dampak negatif kebijakan.

**Kata Kunci:** Analisis sentimen; kebijakan publik; Pajak Pertambahan Nilai

#### **ABSTRACT**

The plan to increase the Value Added Tax (VAT) rate to 12% has become one of the most discussed public policy issues in society. This policy triggered a variety of responses on X social



media, reflecting the pros and cons of the plan. The method used in this research is pre-trained BERT Classification which is used to perform sentiment analysis, topic classification, and provide high accuracy. The purpose of this research is to conduct sentiment analysis and understand the public response to the plan to increase VAT by 12%. The results of the analysis show that the dominance of negative sentiment is 48.58%, which reflects the public's concern about the potential impact of the policy, followed by neutral sentiment of 42.39%, which focuses on economic stability and the effects of the policy in general. Meanwhile, positive sentiment amounted to 9.03%, reflecting optimism towards the long-term benefits of the policy. The BERT model used managed to achieve 83% accuracy, with average precision, recall, and F1-score values of 83%, 82%, and 82%. In addition, word cloud visualization supports the analysis results by displaying dominant words such as "price," "people," and "burden" in negative sentiments, as well as "tax" and "purchasing power" in neutral sentiments. This research contributes to providing data-driven insights to support the government in developing mitigation policies to minimize the negative impacts of the policy.

**Keywords:** *Sentiment analysis; public policy; Value Added Tax*

---

## 1. PENDAHULUAN

Pajak merupakan salah satu penyumbang terbesar bagi perekonomian Indonesia, dengan fungsi utama untuk menghimpun anggaran. Salah satu jenis pajak yang wajib dibayar oleh masyarakat adalah Pajak Pertambahan Nilai (PPN). PPN adalah salah satu penyumbang utama pendapatan negara, menjadi sorotan dalam beberapa tahun terakhir [1]. Pada tahun 2022, pemerintah menaikkan tarif PPN dari 10% menjadi 11% dengan tujuan untuk mencapai target penerimaan pajak dan mengatasi dampak ekonomi. Keputusan ini menimbulkan pro dan kontra di kalangan masyarakat. Di satu sisi, kenaikan tarif diharapkan dapat mendorong pemulihan ekonomi dan meningkatkan pendapatan negara, di sisi lain para pengusaha dan pedagang khawatir bahwa daya beli masyarakat akan menurun jika tarif PPN terus meningkat hingga 12% paling lambat pada tahun 2025 [2].

Rencana kenaikan tarif PPN menjadi 12% berpotensi menimbulkan masalah di kalangan masyarakat, karena dampak dari kenaikan tarif PPN ini tidak hanya dirasakan oleh sekelompok orang, tetapi berdampak pada seluruh masyarakat, baik sebagai pengusaha maupun sebagai konsumen akhir. Bagi pengusaha, selisih antara tarif PPN yang berlaku saat ini dan tarif sebelumnya menjadi beban yang cukup berat, terutama ketika transaksi dilakukan dalam jumlah besar. Selain itu, pengusaha juga dihadapkan pada tuntutan untuk menaikkan harga jual barang atau jasa, yang pada akhirnya akan memengaruhi konsumen [3]. Di sisi lain, harga produk atau jasa menjadi salah satu faktor kunci yang memengaruhi keputusan konsumen dalam melakukan pembelian [4]. Kenaikan tarif PPN memicu beragam respon dan opini di kalangan masyarakat, baik yang positif maupun negatif, mengingat PPN erat kaitannya dengan kegiatan konsumsi di negara berkembang seperti Indonesia. Masyarakat sering menggunakan media sosial sebagai platform untuk menyampaikan respon dan pemikiran mereka terhadap berbagai isu. Di Indonesia, X merupakan salah satu media sosial yang paling populer di kalangan pengguna, dan terkenal untuk menyebarkan informasi dan pendapat [5].

Pada tahun 2021, pembahasan mengenai PPN sempat menjadi topik tren setelah disahkannya Undang-Undang Harmonisasi Peraturan Perpajakan (UU HPP), dan kini hingga tahun 2024 topik ini masih banyak diperbincangkan, terutama terkait dengan kenaikan tarif

PPN yang rencananya naik menjadi 12% yang telah diatur dalam UU HPP [6]. Setelah kenaikan tarif PPN tersebut diumumkan, masyarakat dan media berita menyampaikan beragam tanggapan pro dan kontra terhadap kebijakan tersebut di media sosial X. Oleh karena itu, penelitian ini dilakukan untuk mengamati tanggapan dari masyarakat yang menggunakan media sosial X untuk berkomentar mengenai isu kenaikan PPN 12%. Media sosial dikenal sebagai saluran partisipasi yang memungkinkan publik untuk berkontribusi secara aktif dalam berbagai hal, terutama dalam menanggapi dan mengkritisi berbagai kebijakan. Tulisan yang diposting di media sosial ini diharapkan mendapatkan respons dari pemerintah, meskipun sering kali maknanya bervariasi, mulai dari positif, negatif, dan netral. Keberagaman tanggapan ini mencerminkan kompleksitas komunikasi di media sosial dan menunjukkan pentingnya analisis sentimen untuk memahami respons masyarakat. Metode ini digunakan untuk memahami pendapat masyarakat terhadap berbagai subjek, termasuk tokoh, organisasi, layanan, isu, atau peristiwa [7]. Namun, proses analisis sentimen menghadapi tantangan, terutama dalam mengekstrak konteks dari tulisan, yang memerlukan pemahaman mendalam tentang makna dan nuansa teks tersebut.

Berbagai penelitian mengenai pendapat masyarakat terkait kebijakan publik telah dilakukan sebelumnya. Seperti penelitian [8], menggunakan metode netnografi menunjukkan adanya pro dan kontra terkait kenaikan PPN 11%. Penelitian oleh [9], menerapkan teknik analisis jaringan komunikasi deskriptif, menemukan bahwa arah opini yang terbentuk dalam jaringan tersebut adalah sentimen netral. Sebaliknya penelitian oleh [6], menyatakan bahwa sentimen masyarakat dominan negatif. Selain itu, penggunaan analisis sentimen dalam kebijakan publik, telah dilakukan sebelumnya seperti pada penelitian [10], memberikan rekomendasi kebijakan kepada pemerintah terkait pemilu 2024 menggunakan LSTM dan BERT, dalam penelitian tersebut peneliti mengungkapkan bahwa model BERT secara konsisten memberikan performa yang lebih baik dibandingkan dengan LSTM dan telah terbukti sangat efektif dalam memahami konteks yang kompleks dari data teks berbahasa alami. Berdasarkan beberapa hasil penelitian di atas, hasil yang ditampilkan masih berbeda-beda dan belum konsisten. Selain itu, pendekatan mendalam untuk memahami distribusi dan pola sentimen melalui visualisasi data serta analisis berbasis model BERT belum diterapkan secara spesifik dalam konteks kebijakan ini. Oleh karena itu, penelitian ini menarik untuk dilanjutkan.

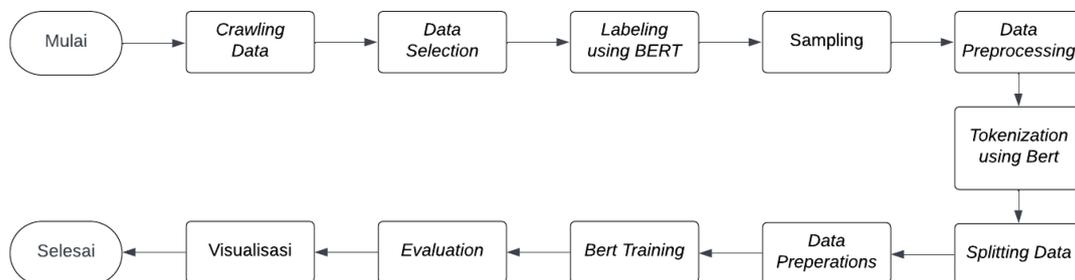
Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap rencana kenaikan PPN sebesar 12% dengan menggunakan model BERT. Penelitian ini diharapkan dapat memberikan kontribusi dalam menyediakan wawasan berbasis data yang mendukung pemerintah dalam perumusan kebijakan yang lebih responsif terhadap opini publik. Model BERT digunakan dalam penelitian ini karena mampu mempelajari representasi dua arah dari teks dengan mempertimbangkan konteks dari kiri dan kanan secara bersamaan di setiap lapisannya. Pendekatan ini telah terbukti efektif dalam memahami konteks bahasa alami pada berbagai penelitian *Natural Language Processing* (NLP) [11].

BERT merupakan model NLP berbasis *deep learning* yang dikembangkan oleh Google [12]. Dengan memanfaatkan encoder transformer, BERT mendukung berbagai tugas NLP, termasuk analisis sentiment. Implementasi BERT melibatkan dua fase utama yaitu *pre-training* untuk membangun representasi pemahaman bahasa secara umum, dan *fine-tuning* dengan data berlabel untuk menyelesaikan tugas-tugas spesifik [13]. Selama tahap *pre-training*, BERT menggunakan *Masked Language Model* untuk memprediksi token yang disembunyikan secara acak berdasarkan konteks sekitarnya, dengan memanfaatkan struktur transformer dua arah. BERT memiliki keunggulan dalam memahami konteks antar-token secara menyeluruh,

menangani teks panjang dengan efektif, serta menunjukkan performa optimal pada dataset kecil melalui proses *fine-tuning* [14].

## 2. METODE PENELITIAN

Penelitian ini akan dilakukan dengan melalui beberapa tahapan, tahapan penelitian ditunjukkan pada gambar 1.



Gambar 1. Metode penelitian

### 2.1 Crawling Data

Pada tahapan ini dilakukan proses pengumpulan data pada media sosial X yang dilakukan dengan metode crawling menggunakan bahasa pemrograman python dengan mengimport *library* pandas, dan juga node js, node JS digunakan untuk menjalankan tweet harvest yang nantinya digunakan di dalam *google collabs* [15]. Dataset yang dikumpulkan berfokus pada pengguna yang berasal dari Indonesia dan menggunakan bahasa Indonesia, dengan kata kunci pencarian "PPN 12%". Rentang waktu pengumpulan data ditetapkan dari 29 Oktober 2021, yaitu tanggal pengumuman rencana kenaikan tarif PPN 12%, hingga 20 September 2024, sebagai waktu pelaksanaan penelitian. Proses ini menghasilkan total 5.316 data tweet. Penelitian ini hanya menggunakan kolom *full text*, yang memuat teks lengkap dari setiap tweet, untuk menganalisis isi dan sentimen yang terkandung di dalamnya.

### 2.2 Data Selection

Pada tahapan ini, dilakukan proses seleksi karena tidak semua data diperlukan dalam proses *data mining* [16]. Tahapan ini dilakukan dengan cara pemilihan baris data yang dibutuhkan, pemilihan baris data mencakup pengecekan terhadap tweet yang relevan dengan kata kunci "PPN 12%". Tweet yang tidak relevan dengan pendapat masyarakat mengenai kenaikan PPN 12% akan dihapus, sehingga penting untuk menjaga *accuracy* hasil data mining.

### 2.3 Labelling using BERT

Pelabelan dalam penelitian ini dilakukan secara otomatis menggunakan *library* transformer dengan *model w11wo/indonesian-roberta-base-sentiment-classifier*, yang dipilih berdasarkan rekomendasi penelitian [17]. Karena model tersebut menunjukkan performa terbaik di antara model-model lain untuk klasifikasi sentimen dalam Bahasa Indonesia. Tujuan dari proses pelabelan ini adalah untuk mengelompokkan data ke dalam kelas-kelas sentimen yang akan digunakan dalam penelitian. Kelas sentimen yang akan digunakan terdiri dari tiga kategori, yaitu negatif, netral, dan positif. Selanjutnya, untuk memudahkan proses, kelas-kelas sentimen yang telah dilabeli akan dikonversi menjadi angka [18].

### 2.4 Sampling

Pada tahap ini, jumlah sampel dalam dataset yang sudah diberi label akan dihitung. Jika ditemukan ketidakseimbangan yang signifikan antara jumlah sampel, maka proses sampling

diperlukan untuk memastikan bahwa klasifikasi dapat dilakukan secara akurat tanpa bias terhadap sampel mayoritas [19]. Metode sampling yang digunakan adalah *Random Over Sampling*, data sampel minoritas akan diduplikasi untuk menyamakan jumlahnya dengan sampel mayoritas secara otomatis menggunakan bahasa pemrograman python.

## 2.5 Data Preprocessing

Tahap *preprocessing* dilakukan dengan menggunakan Python untuk memodifikasi *dataset* agar dapat diproses lebih lanjut. Proses ini bertujuan mempersiapkan data secara sistematis sehingga mendukung analisis sentimen. Tahapan *preprocessing* yaitu *case folding*, *data cleaning*, normalisasi, *stopwords removal*, dan *stemming* [20]. *Case folding* dilakukan untuk mengubah seluruh huruf menjadi huruf kecil. *Data cleaning* dilakukan untuk menghilangkan elemen yang tidak relevan, seperti URL, tag HTML, emoji, *username*, tanda baca (kecuali simbol seperti '%' yang relevan dengan penelitian). Normalisasi digunakan untuk menyempurnakan kata-kata tidak baku atau penyingkatan menjadi bentuk baku, termasuk konversi kata kunci seperti "12%" menjadi "duabelaspersen". *StopWords removal* digunakan untuk menghapus kata-kata dengan nilai informasi rendah. *Stemming* dilakukan untuk mengubah kata ke bentuk dasarnya dengan menghilangkan imbuhan.

## 2.6 Tokenization using BERT

Tahapan ini dilakukan menggunakan *BertTokenizer* dari *library Transformers*. Tokenisasi bertujuan untuk memecah teks menjadi unit-unit token yang lebih kecil sesuai dengan aturan yang diterima oleh model BERT. *BertTokenizer* akan memeriksa apakah setiap kata dalam sebuah kalimat terdapat dalam kosakata BERT. Jika suatu kata tidak ditemukan, kata tersebut akan dipecah menjadi sub kata yang lebih kecil sesuai dengan kosakata yang dikenali, dengan simbol awalan (##) untuk menunjukkan bagian kata yang tidak ditemukan dalam kosakata. Setelah proses tokenisasi, kalimat tersebut akan diubah menjadi *token IDs* yang merepresentasikan kata-kata dalam format yang dapat diterima oleh model. Untuk menyesuaikan format input yang diterima oleh model BERT, beberapa token khusus ditambahkan: token [CLS] di awal urutan untuk menandakan bahwa tugas klasifikasi sedang dilakukan, token [SEP] di akhir setiap urutan untuk memisahkan kalimat, dan token [PAD] digunakan untuk menyamakan panjang urutan dengan panjang maksimum yang telah ditentukan. Pada penelitian ini, panjang maksimum urutan token diatur menjadi 86 token, sesuai dengan panjang token maksimum pada *dataset* yang digunakan. Jika panjang kalimat kurang dari panjang maksimum tersebut, padding dilakukan dengan menambahkan token [PAD] hingga panjang urutan sesuai. Setelah proses padding, setiap token yang telah dilengkapi dengan padding kemudian diberikan mask dan diubah menjadi representasi numerik. Proses ini menghasilkan representasi numerik dari teks yang siap digunakan sebagai input untuk model BERT, memungkinkan model untuk memproses dan memahami teks dengan baik. Di tahap akhir, token tersebut diubah menjadi *attention mask*.

## 2.7 Splitting Data

Pada penelitian ini, proses *splitting* dilakukan menggunakan fungsi *train\_test\_split* dari *library sklearn*. *Dataset* dibagi menjadi tiga bagian utama yaitu *training*, *testing*, dan *validation*, yang masing-masing terdiri dari *input IDs*, *attention mask*, dan label. Pertama, *dataset* dibagi menjadi data *train* dan data *test* dengan rasio 90:10, kemudian, data *train* dibagi lagi menjadi data *train* dan data *validation* dengan rasio 85:15. Hasil akhirnya adalah distribusi 76,5% untuk data *training*, 13,5% untuk data *validation*, dan 10% untuk data *testing*.

Setiap bagian memiliki fungsi yang spesifik, data *training* digunakan untuk mengembangkan model, data *validation* untuk mengurangi *overfitting*, dan data *testing* untuk menguji serta mengevaluasi *accuracy* model.

## 2.8 Data Preparations (konversi ke bentuk tensor dan pemuatan *data loader*)

Pada tahap ini, *input IDs*, *attention masks*, dan label yang telah melalui proses data *splitting* dikonversi menjadi tensor menggunakan *library* PyTorch. Tensor, yang merupakan struktur data multidimensi, digunakan untuk merepresentasikan data secara numerik, memungkinkan kompatibilitas dengan *framework deep learning* seperti PyTorch. Proses konversi ini bertujuan untuk memastikan *dataset* dapat diproses dengan efisien oleh model dalam tugas analisis sentimen. Setelah *dataset* dikonversi menjadi tensor, *Random Sampler* diterapkan pada *dataset training* untuk memastikan pengambilan sampel secara acak dalam setiap iterasi. Sementara itu, untuk *dataset validation* dan *testing* digunakan *Sequential Sampler*, untuk menjaga urutan *dataset* tanpa pengacakan, sesuai dengan kebutuhan *evaluation* model. *Dataset* yang telah dipersiapkan kemudian dimuat ke dalam *Data loader* dengan *batch size* sebesar 32, mengikuti rekomendasi dari penelitian [21].

## 2.9 Bert Training

Penelitian ini menggunakan model *pre-trained BERT Classification*, model yang telah digunakan oleh peneliti [22]. Hasil penelitian tersebut menunjukkan bahwa model *pre-trained BERT* sangat efektif dalam implementasi analisis sentimen dan mampu menghasilkan tingkat *accuracy* yang tinggi. Sebelum *training* model, dilakukan tahapan *fine-tuning* untuk mempersiapkan model *pre-trained BERT*. *Fine-tuning* merupakan proses menyesuaikan model yang sudah dilakukan *training* agar lebih sesuai dengan *dataset* tertentu. Tahap ini memerlukan penentuan parameter utama, seperti *learning rate*, *batch size*, dan jumlah *epoch*, yang dapat disesuaikan untuk meningkatkan performa model [23]. Dalam penelitian ini, digunakan *optimizer AdamW* dan *scheduler get\_linear\_schedule\_with\_warmup* [24]. Penyesuaian dilakukan dengan menetapkan *batch size* sebesar 32, jumlah *epoch* sebanyak 10, dan *learning rate* sebesar  $1e-5$ . Pemilihan jumlah *epoch* mengacu pada penelitian [22], yang menunjukkan bahwa nilai ini dapat mencegah *underfitting*. Sementara itu, *learning rate* sebesar  $1e-5$  dipilih berdasarkan rekomendasi [25], yang melaporkan bahwa nilai tersebut menghasilkan *accuracy* lebih baik dibandingkan dengan *learning rate* yang lebih besar.

## 2.10 Evaluation

Pada tahap ini dilakukan *evaluation* yang bertujuan untuk menilai sejauh mana model dapat melakukan klasifikasi dengan baik pada berbagai kategori yang telah ditentukan. Dalam penelitian ini, salah satu metode evaluasi yang diterapkan adalah *confusion matrix*, sebuah *tools* yang umum digunakan untuk mengukur kinerja model klasifikasi dengan keluaran berupa dua kelas atau lebih. Metode ini memungkinkan untuk memahami performa model secara rinci melalui empat metrik utama *F1-score*, *precision*, *recall*, dan *accuracy*. Keempat metrik ini sering digunakan dalam analisis sentimen dan klasifikasi teks [26].

## 2.11 Visualisasi

Pada tahap ini, data divisualisasikan untuk mempermudah pembacaan dan pemahaman. Salah satu bentuk visualisasi yang digunakan adalah *word cloud*, yang berfungsi untuk menampilkan teks dalam bentuk grafis. Kata-kata yang sering muncul dalam data akan ditampilkan dengan ukuran yang lebih besar, sehingga memudahkan dalam mengidentifikasi pola atau informasi penting.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Data Selection

Pada tahap ini dataset yang telah dikumpulkan melalui proses pengecekan. Setelah proses pengecekan, jumlah tweet yang relevan dengan topik penelitian adalah 1517 tweet. Untuk menjaga *accuracy* hasil data mining, tweet yang tidak relevan akan dihapus dari dataset. Tabel 1. menyajikan contoh tweet beserta klasifikasi nya berdasarkan relevansi.

Tabel 1. Contoh Tweet Saat Dilakukan Pengecekan

Tanggal	Tweet	Keterangan
29-10-2021	@keabadian_ Hai Kak. Dalam hal batas akhir pelaporan SPT Masa PPN bertepatan dengan hari libur (Sabtu Minggu libur nasional pemilu cuti bersama nasional)	Tidak Relevan
.....	.....	.....
14-09-2021	PPN mau jadi 12% udah puncak komedi sih. Bangkok cuma 7% <a href="https://t.co/eoKv76ALLT">https://t.co/eoKv76ALLT</a>	Relevan

#### 3.2. Labelling using BERT

Pada Proses labelling dilakukan dengan menambahkan kolom baru bernama sentimen untuk menyimpan hasil klasifikasi otomatis yang dikelompokkan ke dalam tiga kategori, yaitu positif, negatif, dan netral. Setelah proses pelabelan, kategori sentimen tersebut dikonversi ke dalam bentuk numerik, di mana sentimen negatif direpresentasikan dengan nilai 0, netral dengan nilai 1, dan positif dengan nilai 2. Distribusi hasil labelling menunjukkan bahwa 48,58% data termasuk dalam kategori sentimen negatif (737 data), 42,39% dalam kategori netral (643 data), dan 9,03% dalam kategori positif (137 data).

#### 3.3. Sampling

*Dataset* yang telah melalui proses pelabelan menunjukkan ketidakseimbangan distribusi pada label positif, netral, dan negatif, di mana jumlah data dengan label negatif lebih dominan dibandingkan dengan label lainnya. Ketidakseimbangan ini berpotensi memengaruhi kinerja model, sehingga diperlukan penyesuaian melalui proses *random oversampling* untuk menyeimbangkan jumlah data pada setiap label. Sebelum dilakukan *random oversampling*, dataset terdiri atas 137 data dengan label positif, 643 data dengan label netral, dan 737 data dengan label negatif, dengan total keseluruhan 1.517 data. Setelah proses *random oversampling*, jumlah data untuk masing-masing label menjadi sama, yaitu 737 data, sehingga total dataset meningkat menjadi 2.211 data. Dengan demikian, proses *random oversampling* berhasil menyeimbangkan jumlah data pada setiap label, yang diharapkan dapat meningkatkan performa model dalam melakukan analisis sentimen.

#### 3.4. Preprocessing

Pada tahap ini, dataset akan disesuaikan dengan lima tahapan proses *preprocessing*.

Tabel 2. Proses *Preprocessing*

Proses	Tweet
<i>Dataset</i>	@Hendardo_rdondo @Abu_waras Lah terus pajak kenapa dinaikin? Kan rakyat yang bayar jadinya. PPN tahun ini jd 11% naik dari 10%. Nanti naik lg jd 12%

Proses	Tweet
<i>Case Folding</i>	@hendardo_rdondo @abu_waras lah terus pajak kenapa dinaikin? kan rakyat yang bayar jadinya. ppn tahun ini jd 11% naik dari 10%. nanti naik lg jd 12%.
<i>Data Cleaning</i>	lah terus pajak kenapa dinaikin kan rakyat yang bayar jadinya ppn tahun ini jd 11% naik dari 10% nanti naik lg jd 12%
<i>Normalisasi</i>	lah terus pajak kenapa dinaikkan kan rakyat yang bayar jadinya pajak pertambahan nilai tahun ini jadi sebelas persen naik dari sepuluh persen nanti naik lagi jadi duabelaspersen
<i>Stopwords Removal</i>	pajak dinaikkan rakyat bayar pajak pertambahan nilai sebelas persen sepuluh persen duabelaspersen
<i>Stemming</i>	pajak naik rakyat bayar pajak tambah nilai sebelas persen puluh persen duabelaspersen

### 3.5. Splitting Data

Pada tahap ini, dataset terbagi menjadi tiga kelas, yaitu data *training*, *validation*, dan *testing*. Data *training* memiliki dimensi *input* sebesar (1690, 81) dengan jumlah label sebanyak 1690 dan *mask* sebesar (1690, 81), yang menunjukkan bahwa data ini digunakan untuk melatih model dengan jumlah sampel terbesar. Selanjutnya, data *validation* yang berfungsi untuk mengukur kinerja model selama *training* tanpa mempengaruhi bobot model memiliki dimensi input sebesar (299, 81), jumlah label sebanyak 299, dan *mask* sebesar (299, 81). Terakhir, data *testing* yang digunakan untuk mengevaluasi performa akhir model terhadap data yang belum pernah dilihat memiliki dimensi input sebesar (222, 81), jumlah label sebanyak 222, dan *mask* sebesar (222, 81). Proporsi pembagian data ini menunjukkan keseimbangan yang memadai untuk memastikan model dapat dilakukan *training*, *validation*, dan *testing* secara optimal, serta menghindari potensi *overfitting* atau *underfitting*.

### 3.6. Bert Training

Selama proses pelatihan, metrik yang diamati meliputi *accuracy* dan data *loss* pada data *train* dan *validation*. Gambar 2. menunjukkan hasil pelatihan model per *epoch*.

Epoch	Train Accuracy	Train Loss	Val Accuracy	Val Loss
1	0.4083	1.0881	0.5251	1.0411
2	0.5811	0.9607	0.6522	0.8570
3	0.6876	0.7749	0.7625	0.6361
4	0.7657	0.6096	0.7592	0.5753
5	0.8142	0.4909	0.8060	0.4850
6	0.8456	0.4211	0.7960	0.4803
7	0.8716	0.3624	0.8294	0.4670
8	0.8876	0.3181	0.8194	0.4635
9	0.9059	0.2921	0.8294	0.4490
10	0.9083	0.2787	0.8261	0.4492

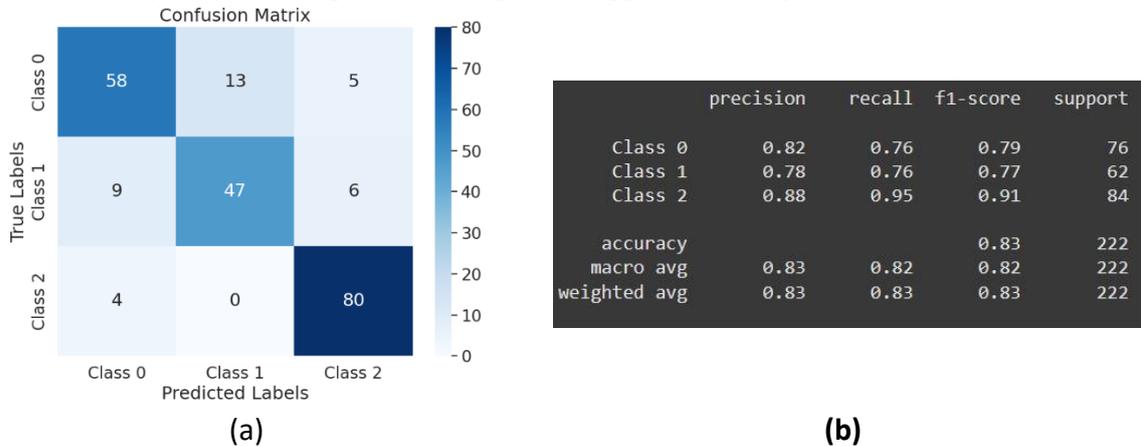
Gambar 2. *Accuracy* dan *loss training* model per *epoch*

Berdasarkan gambar 2. hasil *training* model, terjadi peningkatan yang signifikan pada nilai *Train Accuracy* dan *Validation Accuracy* seiring bertambahnya jumlah *epoch*. Pada epoch pertama, nilai *Train Accuracy* tercatat sebesar 40,83% dengan *Train Loss* sebesar 1,0881, sementara *Validation Accuracy* mencapai 52,51% dengan *Validation Loss* sebesar 1,0411. Hingga *epoch* ke-10, *Train Accuracy* meningkat menjadi 90,83% dengan penurunan *Train Loss* hingga 0,2787. Sementara itu, *Validation Accuracy* mencapai 82,61% dengan *Validation Loss* yang turun menjadi 0,4492. Penurunan nilai *loss* yang konsisten dan peningkatan *accuracy*

menunjukkan bahwa model telah berhasil mempelajari pola data secara efektif selama proses *training*. Nilai *Train Accuracy* menunjukkan performa yang tinggi, *Validation Accuracy* cenderung stabil di sekitar 82% mulai dari *epoch* ke-7, dengan *Validation Loss* yang tetap rendah dan konsisten. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data *validation* tanpa mengalami gejala *overfitting* yang signifikan.

### 3.7. Evaluation

Setelah dilakukan pelatihan, model akan dievaluasi untuk mengukur kinerja model yang telah dilakukan proses *training* model dengan menggunakan *confusion matrix*.



Gambar 3. Hasil *evaluation* model (a) *confusion matrix* (b) performa model

Berdasarkan gambar 3. Menunjukkan hasil *confusion matrix* dan performa model. Nilai *confusion matrix* Untuk Class 0 (negatif), terdapat 76 sampel, di mana model mengklasifikasikan 58 sampel dengan benar (*True Negative*). Namun, 13 sampel salah diklasifikasikan sebagai Class 1 (*False Neutral*) dan 5 sampel salah diklasifikasikan sebagai Class 2 (*False Positive*). Selanjutnya, pada Class 1 (netral) dengan total 62 sampel, model berhasil mengklasifikasikan 47 sampel dengan benar (*True Neutral*). Namun, terdapat 9 sampel yang salah diklasifikasikan sebagai Class 0 (*False Negative*) dan 6 sampel sebagai Class 2 (*False Positive*). Untuk Class 2 (positif), dari total 84 sampel, model mengklasifikasikan 80 sampel dengan benar (*True Positive*). Namun, terdapat 4 sampel yang salah diklasifikasikan sebagai Class 0 (*False Negative*) dan tidak ada sampel yang salah diklasifikasikan sebagai Class 1. Hasil dari performa model memiliki kemampuan yang baik dalam mengklasifikasikan sentimen ke dalam tiga kategori, yaitu kelas 0, kelas 1, dan kelas 2. Precision model mencapai 82% untuk kelas 0, 78% untuk kelas 1, dan 88% untuk kelas 2, menunjukkan tingkat ketepatan prediksi yang cukup tinggi. Recall model menunjukkan performa sebesar 76% untuk kelas 0, 76% untuk kelas 1, dan 95% untuk kelas 2, yang mengindikasikan kemampuan model dalam mengidentifikasi mayoritas sampel dengan benar pada masing-masing kategori. Nilai *F1-score* yang dicapai adalah 79% untuk kelas 0, 77% untuk kelas 1, dan 91% untuk kelas 2, yang menggambarkan keseimbangan antara *precision* dan *recall*. *Accuracy* keseluruhan model tercatat sebesar 83%, dengan nilai *macro average* untuk *precision*, *recall*, dan *F1-score* masing-masing sebesar 83%, 82%, dan 82%. Selain itu, nilai *weighted average* untuk *precision*, *recall*, dan *F1-score* berada pada kisaran yang sama, yaitu 83%, yang menunjukkan konsistensi performa model pada semua kategori. Distribusi data yang ditunjukkan oleh nilai *support* masing-masing kelas (76 untuk kelas 0, 62 untuk kelas 1, dan 84 untuk kelas 2) tidak terlalu

jauh berbeda, sehingga membantu model dalam menangkap pola dengan lebih baik tanpa bias yang signifikan terhadap kategori tertentu.

### 3.8. Visualisasi

Pada tahapan penelitian ini wordcloud digunakan untuk menampilkan kata-kata dominan berdasarkan kategori sentimen (positif, netral, negatif). Proses ini mencakup pengolahan teks, pemilihan kata-kata penting, dan penyajian visual. Selain membantu menilai kinerja model sentimen, wordcloud memberikan gambaran yang lebih mendalam mengenai kata-kata serta topik utama yang diungkapkan oleh pengguna X, sehingga menjadi alat yang berguna untuk menganalisis sentimen teks secara visual.



Gambar 4. Visualisasi sentimen masyarakat (a) wordcloud negatif (b) wordcloud netral



Gambar 5. Visualisasi sentimen masyarakat wordcloud positif

Hasil visualisasi menunjukkan bahwa kata-kata dominan dalam sentimen negatif, seperti “duabelaspersen,” “rakyat,” “harga,” dan “beban,” mencerminkan kekhawatiran masyarakat terhadap kenaikan biaya hidup akibat kebijakan ini. Kata-kata seperti “subsidi” dan “rugi” mengindikasikan keresahan mengenai kurangnya dukungan finansial bagi kelompok rentan. Hal ini menunjukkan perlunya kebijakan mitigasi yang dapat langsung membantu masyarakat berpenghasilan rendah. Sebaliknya, dalam sentimen positif, kata-kata seperti “indonesiamaju” dan “sejahtera” mencerminkan dukungan dari sebagian masyarakat terhadap kebijakan ini, dengan harapan akan membawa manfaat jangka panjang bagi perekonomian. Sementara itu, kata-kata dalam sentimen netral, seperti “pajak” dan “daya beli,” menunjukkan bahwa masyarakat juga fokus pada dampak kebijakan ini terhadap stabilitas ekonomi secara umum. Hasil analisis ini menegaskan dominasi sentimen negatif,

yang mencerminkan keresahan masyarakat terhadap dampak kebijakan kenaikan PPN, terutama pada kenaikan biaya hidup, termasuk kebutuhan pokok dan bahan bakar. Untuk meredakan keresahan ini, beberapa langkah dapat dipertimbangkan oleh pemerintah:

### **3.8.1. Meningkatkan transparansi komunikasi**

Kata-kata seperti “rakyat” dan “harga” yang muncul secara dominan dalam sentimen negatif menunjukkan pentingnya komunikasi terbuka. Pemerintah perlu menjelaskan tujuan kenaikan PPN secara lebih rinci, termasuk manfaat jangka panjangnya bagi perekonomian dan rencana alokasi penerimaan pajak.

### **3.8.2. Memberikan subsidi tepat sasaran**

Dominasi kata-kata seperti “subsidi” dan “rugi” menunjukkan keresahan masyarakat berpenghasilan rendah. Pemerintah dapat memperluas program bantuan seperti subsidi bahan pokok atau Bantuan Langsung Tunai (BLT) untuk mengurangi dampak kebijakan ini.

### **3.8.3. Mendukung sektor Usaha Kecil dan Menengah (UKM)**

Meskipun tidak secara eksplisit tercermin dalam hasil visualisasi, kebijakan ini berpotensi memengaruhi daya saing UKM yang menjadi tulang punggung perekonomian. Pemerintah dapat memberikan insentif pajak sementara atau bantuan finansial untuk menjaga keberlanjutan UKM selama masa transisi.

Hasil yang diperoleh dalam penelitian ini menunjukkan bahwa pelatihan model BERT berhasil mencapai peningkatan *Train Accuracy* yang signifikan, dari 40,83% pada epoch pertama menjadi 90,83% pada epoch ke-10. Sementara itu, *Validation Accuracy* relatif stabil di sekitar 82% setelah 10 epoch, menunjukkan performa model yang cukup baik pada data *validation*. *Evaluation* menggunakan *confusion matrix* menghasilkan *accuracy* keseluruhan sebesar 83%, dengan rata-rata nilai *precision*, *recall*, dan *F1-score* masing-masing sebesar 83%, 82%, dan 82%. Analisis sentimen melalui visualisasi *wordcloud* menunjukkan dominasi sentimen negatif yang merepresentasikan keresahan masyarakat terhadap kebijakan tersebut. Sentimen netral mencerminkan pandangan yang berfokus pada stabilitas ekonomi, sedangkan sentimen positif menunjukkan dukungan terhadap kebijakan dalam jangka panjang. Namun, terdapat beberapa kelemahan pada metode yang digunakan. Pertama, ketidakseimbangan data sebelum dilakukan *random oversampling* yang dapat memengaruhi distribusi kelas dan potensi bias pada model. Kedua, penggunaan pelabelan otomatis berpotensi menghasilkan kesalahan pada label yang memengaruhi *train accuracy*. Ketiga, keterbatasan model dalam melakukan generalisasi pada data baru perlu diperhatikan, terutama jika data memiliki pola yang berbeda dari *data train*. Terakhir, perbedaan signifikan antara nilai *Train Accuracy* dan *Validation Accuracy* dapat menjadi indikasi *overfitting*. Untuk mengatasi hal ini, penelitian selanjutnya disarankan untuk mengatur ulang *hyperparameter*, seperti *learning rate* dan jumlah *epoch*, serta mempertimbangkan penggunaan teknik regularisasi untuk meningkatkan generalisasi model.

## **4. KESIMPULAN**

Hasil penelitian menunjukkan dominasi sentimen negatif terhadap rencana kenaikan PPN 12%, yang mencerminkan kekhawatiran masyarakat terhadap kenaikan biaya hidup dan penurunan daya beli. Visualisasi *wordcloud* memperkuat temuan ini dengan kata-kata dominan seperti “harga,” “rakyat,” dan “beban.” Sentimen positif, meskipun lebih kecil, tercermin melalui kata-kata seperti “indonesiamaju” dan “sejahtera,” yang menunjukkan

optimisme terhadap manfaat jangka panjang kebijakan. Model BERT terbukti efektif dalam analisis sentimen berbasis teks media sosial dengan *accuracy* 83% serta nilai *precision*, *recall*, dan *F1-score* rata-rata masing-masing 83%, 82%, dan 82%. Temuan ini memberikan pemahaman mendalam mengenai respons masyarakat, yang dapat menjadi pertimbangan penting bagi pembuat kebijakan. Penelitian ini diharapkan menjadi landasan untuk kajian lanjutan dalam memahami opini publik terkait kebijakan ekonomi.

## DAFTAR PUSTAKA

- [1] R. E. Yani, E. Simandalahi, and A. R. Nasution, "Pengaruh PPN (Pajak Pertambahan Nilai) terhadap Pendapatan Nasional," *Eksis J. Ilm. Ekon. dan Bisnis*, vol. 15, no. 1, p. 30, 2024, doi: 10.33087/eksis.v15i1.424.
- [2] M. Christanti Kwan, B. Sarjono, P. Studi Manajemen Pemasaran, P. Studi Perpajakan, and P. Ubaya, "Dampak Kenaikan Tarif Pajak Pertambahan Nilai Pada Perilaku Konsumen Di Indonesia," vol. 02, no. 03, pp. 338–348, 2024.
- [3] P. A. Cakranegara, W. Kurniadi, F. Sampe, J. Pangemanan, and M. Yusuf, "The impact of goods product pricing strategies on consumer purchasing power: a review of the literature," *J. Ekon.*, vol. 11, no. 03, p. 2022, 2022, [Online]. Available: <http://ejournal.seaninstitute.or.id/index.php/Ekonomi>
- [4] R. Mustika Sari and Prihartono, "Pengaruh Harga Dan Kualitas Produk Terhadap Keputusan Pembelian (Survey Pelanggan Produk Sprei Rise)," *J. Ilm. MEA (Manajemen, Ekon. dan Akuntansi)*, vol. 5, no. 3, pp. 1171–1184, 2021.
- [5] Y. Cahyana, "Analisis Sentiment Pembelajaran Tatap Muka Terbatas (PTMT) Selama Pandemi Covid-19 Menggunakan Algoritma Naïve Bayes," *Petir*, vol. 16, no. 2, pp. 200–215, 2023, doi: 10.33322/petir.v16i2.1964.
- [6] J. Akuntansi, D. A. N. Keuangan, N. Fauziah, M. Alkautsar, Y. Suryaman, and F. F. Roji, "Pelabelan VADER Dalam Menganalisis Persepsi Masyarakat Terhadap Kenaikan Tarif PPN di Indonesia," vol. 12, no. 2, pp. 228–238, 2024.
- [7] A. C. T. Angel, V. H. Pranatawijaya, and W. Widiatry, "Analisis Sentimen dan Emosi dari Ulasan Google Maps Untuk Layanan Rumah Sakit di Palangka Raya Menggunakan Machine Learning," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 35–49, 2024, doi: 10.24002/konstelasi.v4i1.8924.
- [8] M. Faisol and N. Norsain, "Netnografi: Perspektif Netizen Terhadap Kenaikan Tarif PPN 11%," *J. Akad. Akunt.*, vol. 6, no. 2, pp. 167–182, 2023, doi: 10.22219/jaa.v6i2.24536.
- [9] M. F. Hadi and C. Suratnoaji, "Opini Digital Isu Kenaikan PPN 11%," *Da'watuna J. Commun. Islam. Broadcast.*, vol. 4, no. 2, pp. 415–430, 2023, doi: 10.47467/dawatuna.v4i2.3821.
- [10] T. Pemilu, M. Khadapi, and V. M. Pakpahan, "Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi," vol. 6, pp. 130–137, 2024.
- [11] R. Merdiansah, S. Siska, and A. Ali Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 221–228, 2024, doi: 10.55338/jikoms.v7i1.2895.
- [12] E. Septiana and C. Caroline, "Optimalisasi Evaluasi Pelaksanaan Pelatihan Melalui Analisis Sentimen Otomatis Dengan Model Text Classification," pp. 141–154, 2024.
- [13] S. Rahayu, N. S. Harahap, and S. Agustian, "Application of Langchain Technology to the Fiqh Question Answering System of Four Madhhab Penerapan Teknologi LangChain

- pada Question Answering System Fikih Empat Madzhab,” vol. 4, no. July, pp. 974–983, 2024.
- [14] N. Husin, “Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN),” *J. Esensi Infokom J. Esensi Sist. Inf. dan Sist. Komput.*, vol. 7, no. 1, pp. 75–84, 2023, doi: 10.55886/infokom.v7i1.608.
- [15] F. A. Artanto, “Analisis Sentimen Opini Publik terhadap Fenomena Bunuh Diri Mahasiswa di Twitter Menggunakan Algoritma Naive Bayes,” *J. Sains Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 70–77, 2024, doi: 10.54259/satesi.v4i1.2908.
- [16] N. F. Fattah, “Penerapan Data Mining Untuk Klasifikasi Kualitas Air Dengan Algoritma Support Vector Machine Pada Dinas Lingkungan Hidup Dan Pertanahan Provinsi Sumsel,” *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 11, no. 2, pp. 145–158, 2024, doi: 10.30656/prosisko.v11i2.8285.
- [17] M. Roberta, “Prediksi Sentimen Pada Teks Media Sosial Corporate University,” pp. 302–316.
- [18] E. D. Ratnasari, D. A. Rudira, and A. Surya Buana, “Klasifikasi Penyakit Daun Sawi Hijau Dengan Metode Cnn,” *Pros. Semin. Nas. Teknol. Dan Sains*, vol. 3, pp. 338–393, 2024.
- [19] S. S. Utomo, T. A. Cahyanto, and B. H. Prakoso, “Penggunaan Algoritma Random Over Sampling Untuk Mengatasi Masalah Imbalance Data Pada Klasifikasi Gizi Balita,” pp. 1–9, 2020.
- [20] U. Khairani, V. Mutiawani, and H. Ahmadian, “Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 887–894, 2024, doi: 10.25126/jtiik.1148315.
- [21] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, no. Mlm, pp. 4171–4186, 2019.
- [22] R. Mas, R. W. Panca, K. Atmaja<sup>1</sup>, and W. Yustanti<sup>2</sup>, “Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers),” *Jeisbi*, vol. 02, no. 03, p. 2021, 2021.
- [23] A. S. Yazid and E. Winarko, “Fine-Tuning BERT untuk Menangani Ambiguitas Pada POS Tagging Bahasa Indonesia,” *J. Linguist. Komputasional*, vol. 6, no. 2, pp. 57–64, 2023, doi: 10.26418/jlk.v6i2.148.
- [24] J. Lee and K. Han, “Multimodal Interactions Using Pretrained Unimodal Models for SIMMC 2.0,” 2021, [Online]. Available: <http://arxiv.org/abs/2112.05328>
- [25] R. Yunanto, E. P. Wibowo, and R. Rianto, “a Bert Model To Detect Provocative Hoax,” *J. Eng. Sci. Technol.*, vol. 18, no. 5, pp. 2281–2297, 2023.
- [26] M. Haris, A. Suharso, E. H. Nurkifli, P. S. Informatika, U. S. Karawang, and T. Timur, “Analisis Sentimen Pada Game *E-football* Di Google Play Store Menggunakan Algoritma Indobert,” vol. 8, no. 6, pp. 12108–12121, 2024.